



**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

**“Detección de posibles sitios de unión a plaguicidas en las proteínas
TBP y TRF1 de mosquitos transmisores de enfermedades”**

TESIS

**PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS**

Presenta

Edgar Fernández Ureta

Director de Tesis

Dra. Carmen Nina Pastor Colón



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Instituto de
Investigación en
Ciencias
Básicas y
Aplicadas

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

Coordinación de Programas Educativos

Posgrado en Ciencias



**DR. JEAN MICHEL GRÉVY MACQUART
COORDINADOR DEL POSGRADO EN CIENCIAS
PRESENTE**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada **Detección de posibles sitios de unión a plaguicidas en las proteínas TBP y TRF1 de mosquitos transmisores de enfermedades**, que presenta el alumno **Edgar Fernández Ureta (10024365)** para obtener el título de **Maestro en Ciencias**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Carlos Daniel Amero Tello CIQ-UAEM	APROBADO	
Dr. Rodrigo Said Razo Hernández CIDC-UAEM	APROBADO	
Dr. Enrique Rudiño Piñera IBT-UNAM	APROBADO	
Dra. Lina Andrea Rivillas Acevedo CIDC-UAEM	APROBADO	
Dra. Carmen Nina Pastor Colón CIDC-UAEM	APROBADO	



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

CARLOS DANIEL AMERO TELLO | Fecha:2020-11-21 18:43:45 | Firmante

jLrjBtgZAIP+bvzbzGsahIUOjFhh/xxCUJliQVnzmJfCofA8wTHNamsrPGqKc3H7B0tP4k3mDVbxXyAudeQyBW6E5I8AF8ST6PYtLTaVZzcEamNb3CHU5/EejGocCYVtSte+1IZIXC6q8l+u0X5uVn35oh9yArAOj9kWPkLQUJLfm3xN1RChwz8iCzTy58uG7z5IF8S6Qbn92P1qTAA1UzN1Ofal1FkXn6gkhi3n7t+IX8/4VZT1NII+yEo1w33m+DFdfHrKIZvzYg89o7axERGKGcH6+coqW8m8rti1f/OREQDM+IDL+va4w1IJ94mzezYntNKI44gZrB//T2RbUxw==

RODRIGO SAID RAZO HERNANDEZ | Fecha:2020-11-21 19:25:03 | Firmante

ZVQB6/QvOXdbllspg/Odnbilquxs+WS4L2DssY1BL3jTmEgtAJ/qh2WmF5tyl4IKGShb1ncJuRRNi6EvCY8J21P8K4Qngbc8MsDoTiiWTaXVQksLbgdZWUkWGxmgpPxaL3o9OhpQ4ldVTwai044VcY4A0ByxNixtWW37k0tYGTACabtNjGavZBk+y+cfrOXJTZaAJPVGYrWxjjiZZkq7hh6xc9rjNkkb0uwyHuobTk/ILXSWIFnxQRGA6xJsarQUI8dqxf+oUO49EQDL0Qt15+UTccRia1YECaKM8+GEPBII75LI29qWTY14Hr05QgVdGsn/FopxVFIktsmd2mrsOg==

ENRIQUE RUDIÑO PIÑERA | Fecha:2020-11-23 09:16:03 | Firmante

UImj7KGhi6dLjSwRksXk5wsObzyenYxCASD0mkcfqejdZD/+NjZtiqB17t3DGDzFSI/3hpvuuzVzwXck01AXzww63Y57ug1sEhBWIeLMfa83JaxbXibcuCZVqSQLI9suHQEo31/8r42LKZhwDZzbmfkH61vDxD+KxryDF/6Qnuwii/6wvN8d46qGu35I/0tzQoyalsIH6KuSzbT+CLPCaAH2QvTX4FYQfQIZJsRkHNWf9UPGZIAK9QFCokVrCnp0BmJFMEHLvG6ZWdfM4bMkCOCT5gLO2ctvenoFLBqgJBrDsRODHbJ7pblwNu1QeuV7UTubbMuHPxf7Pd6y1cYA==

LINA ANDREA RIVILLAS ACEVEDO | Fecha:2020-11-23 09:19:00 | Firmante

r6PvR6vB/V4qmiWx/kOimfx0m6EwNlw/RBtkJxuNS2uEIKgNYHXObV9dMx135+DbCb+40RF+G71+R3IKDQu/fftnQqBMkyJJS8qQXSSErhT857ErLK1qoNqQ8Md7wPZJq3l+spXV3HFE3gfmtRoDJZfrc6ar/0xd5G1uVemmmDrlcOvTfTtPye/ghjpp7PwIOAi7/CCXMFFVekz6/KgLiLOQu6hWxREuCHgMDTX4Y5R53yKvM4zqhcCi7fbGPB1s75WEdSuz19i6Ew9L93DPcOxDhgsSzkNo3C4LFN0W50j6lYatdDscK7Ne0eHYcYUzy7vsD05zOVZRUwFwMAGW20Q==

CARMEN NINA PASTOR COLON | Fecha:2020-11-23 09:36:56 | Firmante

NHpgccGmNni1wv/0kwnuLOb/b6YWJOOwcCjC5t+NqhWDdtZChC4diHbM3FF3ZDd/udtktDdE31EZcmI3nab/GTAZpskSf8DNTBUua0HMKZ8RgcDtZD7bNjaWAqfwOpU4u/DEk8yGVziPIHc8C9yzMbS19tJjyL+H8QUTb08BkUnSKWOvtXLXWTT9WPtSI7ITWYmgjH+D5cSgmgrvaLeazQYebLYaeBkDwGesxcvFouFLyBnLQ18PPNXjHVAN/6d6FvF8MOg0i7TbrQcR48cgjIKCvpklyfrAKTubKuCaxnKH10Ds/mpv8kM4DnuikUFHoIG5T5Jcr7pQSaqd8WXYeA==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



cpA4oe

<https://efirma.uaem.mx/noRepudio/FLZKllaBFkE4rkHYV0ILQDfe8CufoRis>



Jurado Evaluador

Dr. Carlos Daniel Amero Tello

Centro de Investigaciones Químicas, UAEM

Dr. Rodrigo Razo Hernández

Centro de Investigación en Dinámica Celular, UAEM

Dr. Enrique Rudiño Piñera

Instituto de Biotecnología, UNAM

Dra. Lina Andrea Rivillas Acevedo

Centro de Investigación en Dinámica Celular, UAEM

Dra. Carmen Nina Pastor Colón

Centro de Investigación en Dinámica Celular, UAEM

Agradecimientos

A la doctora Nina Pastor, ha sido un regalo extraordinario haberla tenido como tutora. De los seres más extraordinarios con los que he tenido la dicha de compartir.

A mis profesores, los doctores Carlos Amero, Lina Rivillas, Rodrigo Razo y César Millán.

Al Laboratorio de Dinámica de Proteínas.

A CONACyT por la beca de maestría (becario 956525) y el Proyecto de infraestructura CONACyT INFR-2014-02-231509 con el que se adquirió equipo de supercómputo en el Laboratorio de Dinámica de Proteínas.

A LANCAD por tiempo de supercómputo en Yoltla y en Xiuhcóatl.

A mis padres, por su amor y apoyo incondicional.

A mis amigos queridos por toda su ayuda y buenas vibras.

Índice

Jurado Evaluador	0
Agradecimientos	1
Resumen	4
Abstract.....	5
Lista de Siglas y Abreviaturas	6
Lista de Figuras	7
Lista de Tablas.....	8
1. Introducción	9
1.1 Artrópodos vectores	9
1.2 Estrategias de control y erradicación	10
1.3 Insecticidas utilizados a gran escala e incremento de la resistencia	10
1.4 Blancos moleculares de acción de pesticidas en uso, y efectos secundarios	11
1.5 La transcripción como posible blanco de pesticidas.....	12
1.6 Interacciones moleculares y su modelado.....	14
1.7 Las interacciones proteína-proteína y las dificultades para inhibirlas	16
2. Fundamento teórico.....	17
Justificación	17
Hipótesis	17
Objetivos.....	17
3 Metodología.....	18
3.1 Búsqueda de secuencias completas de TBP y TRF1	19
3.2 Modelado por homología de candidatos	19
3.3 Análisis de diferencias en la superficie de TBP y TRF1 respecto a TBP de humano	19
3.4 Simulaciones por dinámica molecular	20
3.5 Localización de cavidades y selección de estructuras para acoplamiento molecular.	21
3.6 Selección de compuestos útiles en ZINC15	21
3.7 Cribado molecular.....	22
3.8 Análisis de la selectividad con respecto a la TBP de humanos.	22
3.9 Análisis toxicológico de los compuestos propuestos.....	23
4. Resultados.....	24

4.1 Búsqueda de secuencias de TBPs y TRF1	24
4.2 Generación de Modelos	24
4.3 Análisis de diferencias en la superficie de TBP y TRF1 respecto a TBP de humano	25
4.4 Simulaciones por dinámica molecular	26
4.5 Localización de cavidades y selección de estructuras para acoplamiento molecular.	29
6.6 Mejores candidatos de acuerdo a energías y selectividad del acoplamiento molecular	33
4.7 Análisis de toxicidad de los compuestos propuestos	38
5. Discusión	39
5.1 Selectividad.....	39
5.2 Inhibiendo ARN Pol III	40
5.3 Toxicidad de los compuestos propuestos.....	42
5.4. Metodología empleada.....	42
Conclusiones.....	43
Perspectivas	43
Referencias	44

Resumen

Los mosquitos constituyen uno de los principales vectores de enfermedades a nivel mundial. La estrategia más socorrida para el control de mosquitos es el uso de químicos sintéticos que atacan su sistema nervioso. La mayoría de los mosquitos blanco son resistentes a los insecticidas en uso, además de ser estos tóxicos para mamíferos y el ambiente. Por lo tanto, se requieren nuevos blancos y compuestos que cumplan con esta función. La transcripción es el primer paso en la expresión del genoma. En este proceso participan factores de transcripción pertenecientes a la familia de las TBP, como la propia TBP y TRF1. TBP se encuentra en todas las células eucariotas y es esencial para que ocurra transcripción en todas las etapas de la vida. TRF1 es específico de insectos y se expresa en el sistema nervioso embrionario y en células germinales masculinas. Por tanto, inhibir la función de estas proteínas en mosquitos podría ser letal para ellos. Ambos factores de transcripción divergen en secuencia respecto a la TBP de humanos, haciendo posible la selectividad. Encontramos compuestos menos tóxicos y agresivos contra la salud humana y el ambiente, que se unen selectivamente a los factores de transcripción de mosquitos respecto a la TBP de humanos. El sitio de unión de los compuestos propuestos coincide con el sitio de unión de Brf1, factor imprescindible para la formación del complejo de pre-iniciación de la transcripción de la ARN Pol III. ARN Pol III sintetiza ARN de transferencia, ribosomales, U6 snARN, RNasa P ARN, RNasa MRP ARN, 7SL ARN, Vault ARNs, y transcribe elementos de ADN repetidos intercalados cortos que codifican ARN en todas las células eucariotas; por lo tanto, su función es vital.

Abstract

Mosquitoes are one of the main vectors of disease worldwide. The most common strategy for mosquito control is the use of synthetic chemicals that attack their nervous system. Most of the target mosquitoes are becoming resistant to the insecticides in use; in addition, these insecticides are toxic to mammals and damage the environment. Therefore, new targets and compounds are required to fulfill this function. Transcription is the first step in genome expression. Transcription factors belonging to the TBP family participate in this process, such as TBP itself and TRF1. TBP is found in all eukaryotic cells and is essential for transcription to occur at all stages of life. TRF1 is insect-specific and is expressed in the embryonic nervous system and in male germ cells. Therefore, inhibiting the function of these proteins in mosquitoes could be lethal for them. Both transcription factors diverge in sequence from human TBP, making selectivity possible. We found less toxic and aggressive compounds against human health and the environment, which selectively bind to mosquito transcription factors compared to human TBP. The binding site of the proposed compounds coincides with the binding site of Brf1, an essential factor for the formation of the pre-initiation complex of RNA Pol III. RNA Pol III synthesizes transfer and ribosomal RNA, U6 snRNA, RNase P RNA, RNase MRP RNA, 7SL RNA, Vault RNAs, and transcribes short interspersed repeated DNA elements encoded RNAs in all eukaryotic cells; therefore, its function is vital.

Lista de Siglas y Abreviaturas

2D	Dos dimensiones
Å	Ångström (0.1 nm)
ADN	Ácido desoxirribonucleico
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
LogP	Logaritmo del coeficiente de Partición Octanol/Agua
NAMD	<i>Not (just) Another Molecular Dynamics</i>
NCBI	<i>National Center for Biotechnology Information</i>
ns	Nanosegundos
PDB	<i>Protein Data Bank</i>
Phyre2	P rotein H omology/analog Y R ecognition E ngine V 2.0
RMSD	<i>Root-Mean-Square Deviation</i> (raíz de la desviación cuadrática media)
TBP	<i>TATA-Binding Protein</i>
TRF1	<i>TBP-Related Factor 1</i>
VMD	<i>Visual Molecular Dynamics</i>

Lista de Figuras

Figura 1: Diagrama del procedimiento computacional empleado para construir las colecciones de estructuras a partir de las dinámicas moleculares -----	19
Figura 2: Diagrama de flujo resumen de la metodología empleada -----	24
Figura 3: Porcentaje de identidad con respecto a humanos de las TBPs de <i>Aedes aegypti</i> y <i>Aedes albopictus</i> -----	27
Figura 4: Diferencias en superficie y porcentaje de identidad con respecto a humanos de las TRF1s de <i>Aedes aegypti</i> y <i>Aedes albopictus</i> , y las TBPs de <i>Culex quinquefasciatus</i> , <i>Anopheles darlingi</i> y <i>Anopheles sinensis</i> -----	27
Figura 5: RMSD de las simulaciones de todos los modelos de los factores de transcripción -----	28
Figura 6: Superposición de los tres centros de grupos más poblados de las simulaciones -----	29
Figura 7: Cavidades detectadas por Metapocket en la superficie de los cinco factores de transcripción analizados -----	30
Figura 8: A: Cavidades detectadas por el servidor Metapocket en las superficies de los centros de grupo más poblados de las especies <i>Aedes aegypti</i> y <i>Aedes albopictus</i> . B: Superposición de las TRF1 de ambas especies de <i>Aedes</i> y análisis de conservación -----	31
Figura 9: A: Cavidades detectadas por el servidor Metapocket en las superficies de los centros de grupo más poblados de las especies <i>Anopheles darlingi</i> y <i>Anopheles sinensis</i> . B: Superposición de las TBP de <i>Anopheles darlingi</i> y <i>Anopheles sinensis</i> y su análisis de conservación -----	32
Figura 10: Cavidades detectadas por el servidor Metapocket en las superficies de los centros de grupo más poblados de la especie <i>Culex quinquefasciatus</i> -----	32
Figura 11: Cálculo de volumen realizado por POVME a lo largo de las simulaciones de los factores de transcripción en la cavidad que corresponde con parte del sitio de interacción de Brf1 -----	33
Figura 12: Cavidad donde se realizó el cribado molecular -----	34
Figura 13: Mejores candidatos por selectividad encontrados en <i>Aedes aegypti</i> (TRF1) -----	35
Figura 14: Mejores candidatos por selectividad encontrados en <i>Anopheles darlingi</i> (TBP) -----	36
Figura 15: Mejores candidatos por selectividad encontrados en <i>Culex quinquefasciatus</i> (TBP) ---	37
Figura 16: Predicción de toxicidad realizada con el servidor ADMETSAR -----	38
Figura 17: Alineamiento de secuencia de los factores de transcripción analizados -----	39
Figura 18: Superposición del sitio de unión de la cavidad analizada en mosquitos vectores junto al ligando acoplado y el complejo de unión de TBP y Brf1 -----	40
Figura 19: Productos de la transcripción mediada por ARN Pol III -----	41

Lista de Tablas

Tabla 1: Secuencias de factores de transcripción pertenecientes a la familia de las TBP encontradas en el NCBI. Porcentaje de identidad con respecto a la TBP de humanos -----	25
Tabla 2: Parámetros de evaluación de los modelos generados por homología -----	26
Tabla 3: Grupos más poblados de las simulaciones realizadas -----	29

1. Introducción

1.1 Artrópodos vectores

Los artrópodos constituyen el filo más numeroso y diverso del reino animal; describe a animales que poseen esqueleto y apéndices articulados, e incluye a insectos, arácnidos, crustáceos y miriápodos (1). Los artrópodos juegan un importante papel como plagas y vectores de enfermedades de plantas y animales. Arácnidos de la familia ixodidea son transmisores de la Enfermedad de Lyme, tifus, babesiosis, ehrliquiosis y anaplasmosis (2–6). La mayoría de los virus de plantas son transmitidos por insectos hemípteros (7). Insectos como cucarachas y moscas son transmisores de bacterias y protozoos causantes de enfermedades (8–10). Dípteros como la mosca tsé-tsé pueden transmitir nemátodos patógenos (11). El 17 % de todas las enfermedades infecciosas son transmitidas por artrópodos vectores de los cuales la mayoría son insectos que se alimentan de sangre (12), por ejemplo, insectos tritominos son transmisores del protozoo *Trypanosoma cruzi* causante del Mal de Chagas (13–15), y los artrópodos culícidos (mosquitos) constituyen unos de los principales vectores de enfermedades a nivel mundial.

Los trabajadores de salud pública se enfrentan a desafíos a causa de la carga de enfermedades transmitidas por mosquitos, a pesar de los grandes esfuerzos realizados en programas de control (16). El panorama del riesgo de enfermedades transmitidas por mosquitos ha cambiado dramáticamente en las últimas décadas, debido a la aparición y el resurgimiento de los ciclos de transmisión urbana impulsados por los invasores *Aedes aegypti* y *Aedes albopictus* (17). Ambas especies son vectores competentes de cinco importantes arbovirus (chikungunya, dengue, fiebre del Valle de Rift, fiebre amarilla y Zika) y ambos están extensamente distribuidos, haciendo a estas especies un factor importante de diseminación de estas enfermedades a nivel mundial (18). Un estimado de cuatro millones de personas viven en áreas de riesgo de contraer dengue (16). El vector de malaria *Anopheles darlingi* tiene una amplia distribución geográfica en América Central y Sudamérica, extendiéndose desde el sur de México hasta el norte de Argentina y desde el este de la cordillera de los Andes hasta la costa del océano Atlántico. *A. darlingi* se considera uno de los vectores más eficientes de malaria en regiones tropicales. Los estudios en la biología y comportamiento de este mosquito han sido relativamente pocos, especialmente teniendo en cuenta su importante rol en la transmisión de malaria. El mosquito *Anopheles sinensis* es el principal vector de malaria en países del sureste asiático, incluyendo China (19–22).

Culex quinquefasciatus (el mosquito doméstico del sur) es un importante mosquito vector de virus como el virus del Nilo Occidental y el virus de la encefalitis de San Luis, así como de nemátodos que causan filariasis linfática. *C. quinquefasciatus* es una especie dentro del complejo de especies *Culex pipiens* y se puede encontrar en climas tropicales y templados de todo el mundo. La capacidad de *C. quinquefasciatus* para ingerir sangre de aves, ganado y humanos contribuye a su capacidad de vectorizar patógenos entre especies (23).

1.2 Estrategias de control y erradicación

Para enfrentar el problema que representan estos vectores se han ideado varias estrategias, por ejemplo, la producción de mosquitos del género *Anopheles* transformados genéticamente que evitan que *Plasmodium*, el protozoo causante de la malaria, se desarrolle dentro de ellos y así impedir su transmisión y diseminación (24). El control de poblaciones de mosquitos del complejo *Culex pipiens* y *Aedes albopictus* se ha intentado a través de la infección de poblaciones de estas especies con *Wolbachia*, una bacteria endosimbiótica que induce muerte embrionaria (25). Sin embargo, el control de vectores no ha logrado prevenir las epidemias recientes ni evitar la expansión de enfermedades arbovirales importantes como el dengue (26). La principal estrategia para el control de brotes de arbovirus es el uso de químicos sintéticos de acción rápida que causan la muerte a especímenes adultos del género *Aedes* mediante atomización espacial (26–28). La mayoría de los insecticidas recomendados para estas especies son de la clase de los piretroides (27). Respecto al manejo específico del vector de arbovirus *Aedes aegypti*, se ha propuesto e implementado el control larval como estrategia primaria, incluyendo la aplicación de larvicidas químicos y toxinas bacterianas. Se han empleado agentes biológicos contra larvas que incluyen copépodos depredadores, peces, y larvas de *Toxorhynchites*. Es probable que el mayor obstáculo para el control larval de *A. aegypti* sea la imposibilidad de detectar, acceder y eliminar o tratar los focos de cría domiciliarios, lo cual constituye un reto y una tarea costosa que usualmente conduce a un bajo control; adicionalmente, su eficiencia reducida en algunas ocasiones limita su adopción generalizada (26).

Ha existido un esfuerzo internacional para controlar o erradicar la malaria que ha dado como resultado nuevas herramientas para controlar esta enfermedad y sus vectores. Estrategias actuales de manejo integrado de vectores que incluyen el uso de mosquiteros tratados con insecticida y pulverización residual interior, pueden no ser suficientes para erradicar la transmisión de malaria en áreas endémicas (19). El control de *Anopheles sinensis* y *Anopheles darlingi* se basa fundamentalmente en el uso de mosquiteros tratados con insecticida y pulverización residual interior. Para el control de los mosquitos del complejo *Culex pipiens* se emplean insecticidas organofosforados, carbamatos y piretroides (29,30).

1.3 Insecticidas utilizados a gran escala e incremento de la resistencia

Los insecticidas juegan un papel principal en el control de la mayoría de vectores de enfermedades como mosquitos, moscas, pulgas, piojos, y chinches. Estos químicos han sido extensivamente usados desde 1940 para el control de mosquitos (22,31). Históricamente cuatro principales categorías de insecticidas han sido aplicados: organoclorados, organofosforados, carbamatos y piretroides (31). En 1955 la Organización Mundial de la Salud (OMS) propuso la erradicación global de los vectores transmisores de malaria con el uso doméstico de DDT (dicloro difenil tricloroetano) atomizado; sin embargo, la euforia de los insecticidas terminó pronto y en 1976, la OMS cambió oficialmente el discurso de erradicación por control, cuya causa fue la aparición de resistencia al DDT en un amplio rango de mosquitos vectores (32). La resistencia a los insecticidas se define como la capacidad heredada de una población para sobrevivir a una dosis de insecticida, administrada

bajo las mismas condiciones, que normalmente habría resultado letal para individuos de otra población susceptible de la misma especie (33). En 1975 La OMS informó que 256 millones de personas vivían en áreas donde la resistencia al DDT estaba socavando los esfuerzos de control de la malaria. Los problemas de resistencia continuaron con el cambio a nuevos insecticidas como los organofosforados, carbamatos y piretroides (32).

Los neonicotinoides son una familia de insecticidas que actúan en el sistema nervioso central de los insectos y, con menor toxicidad, en vertebrados (aves y mamíferos). Están ahora registrados y aprobados para su uso en cientos de cultivos en más de 120 países por lo que actualmente se encuentran entre los insecticidas más usados a nivel mundial. Sin embargo, el uso de ciertos productos químicos de esta familia está siendo restringido en países debido a una posible conexión con el desorden del colapso de colonias apícolas, o muerte de las abejas. Los neonicotinoides se consideran neurotóxicos porque actúan como agonistas contra los receptores nicotínicos de acetilcolina. En insectos, los neonicotinoides causan parálisis que lleva a la muerte, frecuentemente en pocas horas. Debido a que los neonicotinoides bloquean una ruta neuronal específica que es más abundante en insectos que en mamíferos, estos insecticidas son por tanto selectivos frente a insectos en comparación a mamíferos. Son utilizados extensivamente en campo en cientos de cultivos y están siendo empleados para tratar ectoparásitos en animales domésticos, pero no son utilizados como insecticidas de mosquitos (34).

La resistencia a los insecticidas es un problema en expansión y de larga duración al que nos enfrentamos en el control de artrópodos plaga. El manejo efectivo de la resistencia a los insecticidas es esencial para preservar la utilidad de los insecticidas actuales y futuros. El control químico es sin lugar a dudas la principal, en ocasiones la única, herramienta con que se cuenta. El uso intensivo y extensivo de insecticidas ha traído consigo una presión de selección y la consecuente resistencia, la cual es creciente. Se ha reportado resistencia a carbamatos, piretroides y organofosforados en *C. quinquefasciatus* y *Anopheles gambiae* (29,30). Se han realizado estudios en *Aedes aegypti* que han arrojado como resultados la presencia de resistencia a organofosforados y piretroides en poblaciones del Caribe, islas Martinica, Guyana Francesa, Brasil, Europa y Tailandia (35–40).

1.4 Blancos moleculares de acción de pesticidas en uso, y efectos secundarios

Todos los insecticidas químicos actuales son neurotóxicos, ya que actúan afectando el sistema nervioso de los organismos blanco. El sistema nervioso de insectos está altamente desarrollado y no se diferencia en demasía del sistema nervioso de mamíferos. El sistema nervioso periférico, aunque menos complejo, también presenta llamativas similitudes, por lo que, la mayoría de los insecticidas no son selectivos entre especies con respecto a los blancos moleculares de insectos y mamíferos, incluyendo humanos, que son altamente sensibles a sus efectos tóxicos (41).

Además, se han encontrado otros efectos adversos de estos pesticidas convencionales para la salud humana. Estudios epidemiológicos reportan que la exposición a pesticidas organoclorados incrementa el riesgo de padecer cáncer. Un estudio realizado con dos

organoclorados, clordecona y lindano, arrojó que pueden actuar como disruptores endocrinos mediante la activación del receptor de estrógeno α ($ER\alpha$), aumentando la proliferación de células endoteliales, la migración y la actividad de MMP2 (metalproteinasa2), relacionada con la angiogénesis. Los organoclorados aumentaron la producción de óxido nítrico mediante una mejora de la actividad de eNOS (sintetasa endotelial de óxido nítrico) sin modificación del estrés oxidativo. Se ha proporcionado evidencia de que las dos toxinas aumentaron la neovascularización *in vivo*. Lo más interesante es que todos los procesos anteriores se evitaron parcial o completamente después del silenciamiento de $ER\alpha$. En conjunto, estos datos destacan que los organoclorados modulan los procesos angiogénicos celulares mediante la activación de $ER\alpha$. Este estudio refuerza aún más las evidencias de los efectos nocivos de estos pesticidas en la carcinogénesis, particularmente en la modulación de la angiogénesis, un paso crítico en el desarrollo del tumor, a través de $ER\alpha$ (42). Los efectos toxicológicos de las mezclas de pesticidas en dosis bajas en la salud humana son en gran medida desconocidos, aunque existen crecientes preocupaciones sobre su seguridad. La toxicidad acumulativa de organofosforados y organoclorados produce efectos estrogénicos y el desarrollo de polineuropatía retardada inducida por organofosforados (43). Los insecticidas carbamatos y organofosforados tienen como mecanismo de acción la inhibición de la enzima colinesterasa (44–47). Esta enzima se encuentra extracelularmente (48). Todos los piretroides y DDT interactúan con el canal de sodio, al mantenerlo abierto durante más tiempo, aumentan la probabilidad de desarrollar potenciales de acción, creando así una condición de hiperexcitabilidad, cuyo principal signo clínico son los temblores. La mayoría de los organoclorados (excepto la clordecona), así como ciertos piretroides (tipo II), bloquean los canales de cloruro del receptor GABA-A e interactúan con canales de cloruro activados por glicina, inhibiéndolos (41,49).

Por lo tanto, estos insecticidas no tienen que atravesar membranas celulares para ejecutar su efecto. Teniendo en cuenta que la mayoría de los insecticidas usados interactúan con moléculas, dígame receptores o enzimas, que también se encuentran presentes en otros organismos como mamíferos y por tanto ejercen efectos adversos en la salud de poblaciones humanas, en este estudio se propone la búsqueda de selectividad molecular, con lo cual, los insecticidas propuestos no afectarían la salud humana.

1.5 La transcripción como posible blanco de pesticidas

La transcripción es el primer paso en la expresión del genoma. Es ejecutada por ARN polimerasas, enzimas que usan una cadena de ADN como molde para sintetizar ARN, y existen tres formas de ARN polimerasas en eucariotas, nombradas ARN polimerasa I (Pol I), Pol II y Pol III (50). Los factores de transcripción (TFs) son proteínas que se unen al ADN al reconocer secuencias específicas localizadas en estructuras reguladoras como promotores o potenciadores. Esta unión controla procesos como el reclutamiento de ARN polimerasas, la metilación del ADN, y modificaciones químicas de nucleosomas y su desplazamiento. El resultado es la activación o represión de la expresión génica. Por lo tanto, los TFs tienen un papel crucial en la interpretación de la información genómica y son actores centrales en las redes de regulación genética (51).

La proteína de unión a cajas TATA (TBP), es un factor de transcripción requerido por las tres ARN polimerasas para llevar a cabo el proceso de iniciación de la transcripción. TBP tiene 180 residuos conservados filogenéticamente en la región C-terminal, que contiene dos repeticiones directas que flanquean un segmento altamente básico conocido como la repetición básica (52–54). El extremo N-terminal es desordenado y altamente divergente en secuencia y longitud (55,56). TBP tiene como función la formación del complejo de pre-iniciación, el cual es requerido para la iniciación de la transcripción, por lo que establece interacciones a lo largo de su superficie con otros factores de transcripción como TFIIA, TFIIB, MOT1, Brf1, TAF1, NC2, y las proteínas Spt3 y Spt8 del complejo SAGA (57–60). Hay estructuras cristalográficas disponibles de algunos de estos complejos con TBP (57–60). TBP presenta múltiples interacciones proteína-proteína y, por tanto, una extensa superficie molecular que puede ser motivo de estudio y que brinda la opción de utilizar más de un compuesto de forma simultánea en la creación del insecticida (61). La tasa de mutación de TBP es extremadamente baja por lo que la adquisición de resistencia a insecticidas que ataquen a este blanco debe ser un proceso lento (61,62).

Anteriormente, la TBP era considerada como el factor de transcripción universal y el único requerido por las tres ARN polimerasas para el reconocimiento del promotor y el inicio de la transcripción en eucariotas (63–65). En *Drosophila melanogaster* fue descrito un gen que codifica para una proteína estrechamente relacionada a TBP; esta proteína fue nombrada TRF1 (TBP-related-factor), y es específica de insectos (53). Se ha demostrado que TRF1 puede desempeñar el rol de la TBP en la transcripción de ciertos conjuntos de genes no regulados por la TBP. Hibridación *in situ* en embriones de *Drosophila melanogaster* mostró que el nivel más alto de transcripción de TRF1 fue detectado en células neuronales y germinales (53,63,66). Se ha confirmado a través de estudios bioquímicos que TRF1, al igual que TBP, puede interactuar específicamente con TFIIA, TFIIB, unirse al promotor de la caja TATA y dirigir la transcripción mediada por la Pol II *in vivo*. De forma similar a TBP, TRF1 es miembro de un gran complejo multiprotéico. Mutaciones en los genes que codifican TRF1 resultaron en infertilidad masculina y anomalías en el sistema nervioso. Inhibición de TRF1, pero no de TBP, en extractos nucleares de embriones de *Drosophila* condujo a una inhibición de la transcripción mediada por la ARN polimerasa III, lo cual confirma el rol de TRF1 en ese proceso (63,66). Aunque la estructura de TBP está conservada, han sido descritas diferencias en secuencia en las TBP de parásitos eucariotas, lo cual trae consigo diferencias estructurales y de superficie (61).

En este estudio se propone hacer un análisis similar en los factores de transcripción de los tres géneros de mosquitos mencionados. Teniendo en cuenta que impidiendo que ocurra cualquiera de estas interacciones proteína-proteína no ocurre la formación del complejo pre-iniciación y con ello no puede ocurrir transcripción, en este estudio se proponen como blancos moleculares a TBP y TRF1. Ya TBP ha sido propuesta como blanco terapéutico para el tratamiento de enfermedades parasitarias (61).

1.6 Interacciones moleculares y su modelado

Las interacciones entre biomoléculas son fundamentales para todo proceso biológico. Haciendo uso de esas interacciones, los organismos vivos mantienen complejas redes regulatorias y metabólicas. El trabajo experimental y las simulaciones computacionales son las principales herramientas científicas para entender estos procesos y encontrar nuevas moléculas que puedan ser utilizadas como sustancias bioactivas para modificarlos y controlarlos. El análisis de interacciones moleculares individuales es un paso fundamental para entender los procesos vitales. El estudio de estas interacciones requiere un conocimiento exhaustivo de la estructura tridimensional de las moléculas que se van a considerar. Desde la década de los 90, el conocimiento de estructuras moleculares se ha venido incrementando aceleradamente y han sido desarrollados muchos algoritmos que abordan el acoplamiento molecular (docking), así como los espacios conformacionales y las dificultades para modelar la energía (67).

Los sitios de unión en las proteínas son generalmente cavidades donde ocurren las interacciones moleculares. La característica que debe tener un blanco molecular es que posea un sitio de unión que permita la unión de un compuesto (ligando), y que esto ocurra con alta afinidad y especificidad (54,68). La capacidad de predecir si compuestos pequeños se pueden unir con alta afinidad y especificidad a una proteína en particular, basados únicamente en su estructura 3D, ha sido una meta que han tenido los biólogos estructurales y científicos computacionales hace mucho tiempo. La promesa es que una predicción precisa de la accesibilidad farmacológica de las proteínas puede capitalizar las enormes inversiones ya realizadas en iniciativas de genómica estructural mediante la identificación de blancos farmacológicos y el uso de esta información en campañas de identificación y validación de blancos o dianas (69). Para el desarrollo y búsqueda de nuevos fármacos la principal estrategia empleada es la búsqueda y selección de compuestos químicos evaluados contra un blanco molecular de forma experimental, lo cual es un proceso lento y costoso. El cribado virtual (virtual screening) proporciona una alternativa mucho más rápida y barata de hacer esto. La idea es evaluar computacionalmente, mediante acoplamiento molecular, grandes bibliotecas de compuestos químicos (potenciales fármacos) en blancos de estructuras conocidas, determinar cuáles son los compuestos que se unen con mejor afinidad, y posteriormente evaluar experimentalmente solamente estos compuestos. Esto representa un ahorro considerable de tiempo y recursos. El cribado virtual enfrenta varios desafíos, como el muestreo de las diferentes conformaciones de moléculas flexibles y el cálculo de la energía de unión absoluta en un entorno acuoso. Aún con las limitaciones actuales, nuevos compuestos han sido descubiertos usando estas herramientas computacionales puesto que éstas brindan acceso a un inmenso número de posibles ligandos, la mayoría de los cuales pueden ser fácilmente adquiridos y probados. Además, estructuras de ligandos conocidos en complejos con sus receptores han sido correctamente predichos computacionalmente utilizando las estructuras de forma independiente (70).

El campo del acoplamiento molecular ha tomado fuerza en las tres últimas décadas debido a que los estudios computacionales de biología molecular estructural y de diseño de fármacos

basado en estructura han demostrado ser herramientas sumamente importantes y necesarias para llevar a cabo estudios moleculares. Esto ha ocurrido gracias al dramático crecimiento en disponibilidad, poder de las computadoras y la creciente facilidad de acceso a bases de datos de pequeñas moléculas y proteínas. El acoplamiento molecular tiene una amplia variedad de usos y aplicaciones en el descubrimiento de fármacos. Estas aplicaciones incluyen la búsqueda y optimización de compuestos como potenciales fármacos mediante cribado virtual; estudios de correlación estructura-actividad de ligandos; estudios en los cuales, con auxilio de la cristalografía de rayos X, se obtiene información y predicciones de efectos de compuestos en mutagénesis; estudios de sustratos e inhibidores enzimáticos; y estudios de mecanismos químicos y bibliotecas combinatorias de diseño(71). El objetivo del software automatizado de acoplamiento molecular es predecir la unión no covalente de dos moléculas, utilizando sus estructuras para comprender y predecir reconocimiento molecular, tanto estructuralmente, encontrando modos de unión probables, como energéticamente, prediciendo su afinidad.

El acoplamiento molecular generalmente se realiza entre una molécula pequeña y una macromolécula blanco. Esto a menudo se conoce como acoplamiento de ligando-proteína, pero existe un creciente interés en el acoplamiento proteína-proteína (71,72). En este estudio, se llevará a cabo el acoplamiento ligando-proteína.

Durante las últimas dos décadas, una gran variedad de herramientas y programas de acoplamiento molecular, como AutoDock (73), AutoDock Vina (72), LeDock (74), rDock (75), UCSF DOCK (76), LigandFit (77), Glide (78), GOLD (79), MOE Dock (80), y Surflex-Dock (81), entre muchos otros, se han desarrollado para ambos usos comerciales y académicos. Para un programa de acoplamiento molecular, los dos componentes más críticos son el algoritmo de muestreo y la función de puntaje, que determinan su poder y efectividad. (82). Han existido muchos estudios cuyo objetivo es comparar los diferentes programas de acoplamiento y sus respectivas funciones de muestreo y puntaje (83). Diferentes estrategias de muestreo y funciones de puntaje son empleadas por diferentes programas de acoplamiento, por tanto, es importante tener en cuenta la evaluación y comparación del rendimiento de estos programas, puesto que los resultados de la evaluación pueden revelar las ventajas y limitaciones de cada programa de acoplamiento, y esto, ayudar a los usuarios a tomar decisiones razonables entre diferentes programas de acoplamiento molecular.

En un estudio realizado en 2016 que comparaba diez de los programas de acoplamiento molecular más utilizados (AutoDock, AutoDock Vina, LeDock, rDock, UCSF DOCK, LigandFit, Glide, GOLD, MOE Dock y Surflex-Dock), se encontró que AutoDock Vina tuvo el mejor poder de puntaje y una buena función de búsqueda, significativamente mejor que Autodock (82). Los softwares de la familia Autodock son ampliamente utilizados en las investigaciones que emplean acoplamiento proteína-ligando y son notables por estar entre los pocos programas de acoplamiento que están disponibles gratuitamente para académicos y uso industrial. Autodock Vina se ejecuta más rápido, reproduce y clasifica con mejor precisión las poses de complejos proteína-ligando conocidos, en comparación con Autodock4, por lo tanto, se recomienda a los investigadores emplearlo de preferencia a Autodock4 para realizar cribado virtual (84). AutoDock Vina es uno de los programas de

acoplamiento molecular más utilizados. Ha demostrado estar a la altura de programas como Glide y Gold. Es de acceso libre, y permite asignar las cadenas laterales de los aminoácidos de interés como flexibles, lo cual lo hace una herramienta eficiente para realizar cribado virtual (72,85).

Los estudios encaminados al descubrimiento de compuestos novedosos en el campo de la medicina, la industria, y demás ramas económicas y de investigación que estriban en compuestos químicos para su desarrollo, dependen de la calidad de las bases de datos empleadas para el perfeccionamiento e implemento de las herramientas computacionales. Por lo tanto, se han desarrollado gran cantidad de bases de datos o bibliotecas de compuestos como Binding Moad (86), PDB bind (87), Ligand-Protein DataBase (88), Relibase (89), BindingDB (90)[95], PDBLIG (91), MSDsite (92), eF-Site (93), PDB-Ligand (94), Super-Ligands (95), Protein Ligand Database (PLD) (96), Het-PDB Navi (97), sc-PDB (98), PDBSite (99), Ligand-Depot (86), AffinDB (100), LIGAND (101), y KiBank (102). ZINC15 es una base de datos de acceso público y es además un conjunto de herramientas, desarrollado para permitir el acceso inmediato a compuestos para el cribado virtual, que se ha vuelto ampliamente utilizada para este fin, descubrimiento de ligandos, cribado de farmacóforo, evaluación comparativa, y desarrollo de campos de fuerza. Tiene más de 120 millones de compuestos “similares a fármacos “, que pueden ser fácilmente adquiridos; es una base de datos diseñada con el objetivo de reducir la brecha entre la quimioinformática y la biología, por lo que implementa herramientas dirigidas a investigadores que no son especialistas en bioinformática. Es una base de datos de reconocido prestigio y calidad (103).

1.7 Las interacciones proteína-proteína y las dificultades para inhibirlas

Los factores de transcripción son considerados blancos de difícil acceso molecular, a causa de que tienen grandes superficies de interacción proteína-proteína, lo que se traduce en la carencia de cavidades profundas. Sin embargo, estudios similares se han llevado a cabo con éxito en familias de proteínas como la familia de factores de transcripción Myc (104), que son oncoproteínas que controlan varias funciones celulares relacionadas con la proliferación celular eficiente, crecimiento y metabolismo y regeneración de tejidos. Muchos de los blancos considerados de difícil acceso farmacológico han sido considerados eventualmente blancos satisfactorios, ejemplo de esto es la familia de proteínas BCL-2, la cual fue considerada en un principio un blanco de difícil acceso farmacológico (104–106). Hasta el momento no se han desarrollado insecticidas que inhiban interacciones de TBP con otros factores de transcripción, aunque sí existe un estudio similar en TBP de parásitos eucariotas con resultados promisorios, en el que se determinó que TBP es un blanco potencial (61).

2. Fundamento teórico

Justificación

En los tres géneros de mosquitos vectores que se estudian en este proyecto existe resistencia a insecticidas; además los insecticidas actuales causan daño al ambiente y la salud humana, por lo que se hace necesario la búsqueda de nuevos compuestos y blancos moleculares que efectúen esta función. En los factores de transcripción TBP y TRF1 existen diferencias en secuencia respecto a la TBP de humanos, lo que conlleva a diferencias estructurales que pueden proporcionar sitios blanco novedosos que presenten selectividad con respecto a humanos, y por tanto evitar efectos adversos a la salud humana. Inhibiendo TBP se detiene la transcripción en todas las etapas de vida de estos vectores e inhibiendo a TRF1 se afecta su desarrollo en etapas larvales.

Hipótesis

Existen sitios de unión en las superficies de las TBPs y TRFs de artrópodos vectores de enfermedades que pueden ser propuestos como sitios novedosos de unión a compuestos selectivos que impidan la interacción con otras proteínas del complejo pre-iniciación.

Objetivos

Objetivo general

-Identificar *in silico* compuestos con probable actividad pesticida, que se unan selectivamente a los factores transcripcionales TBP y TRF de mosquitos con respecto a la TBP de humano.

Objetivos específicos

-Seleccionar un conjunto de TBPs y TRFs de vectores artrópodos, y modelar las estructuras.

-Identificar las diferencias en secuencia respecto a la TBPs de humanos y mapearlas sobre la superficie molecular.

-Obtener ensamblajes de estructuras representativas de TBPs y TRFs de vectores y localizar posibles sitios de unión para moléculas pequeñas sobre sus superficies.

-Realizar cribado virtual partiendo de bibliotecas de compuestos químicos aptos para usarse como plaguicidas sobre las superficies de TBPs y TRFs de vectores, y compararlas con la TBP de humano, identificando los compuestos con mejor selectividad y afinidad hacia las TBPs y TRFs de vectores.

3 Metodología

En este acápite se describe la estrategia computacional utilizada para llevar a cabo la investigación, comenzando por la búsqueda de secuencias completas de factores de transcripción de mosquitos vectores, hasta el análisis toxicológico realizado a los compuestos propuestos como potenciales insecticidas.

La figura 1 muestra un diagrama resumen de los procedimientos computacionales realizados a partir de las dinámicas moleculares obtenidas, con el objetivo de obtener las colecciones de estructuras donde se lleva a cabo el cribado virtual.

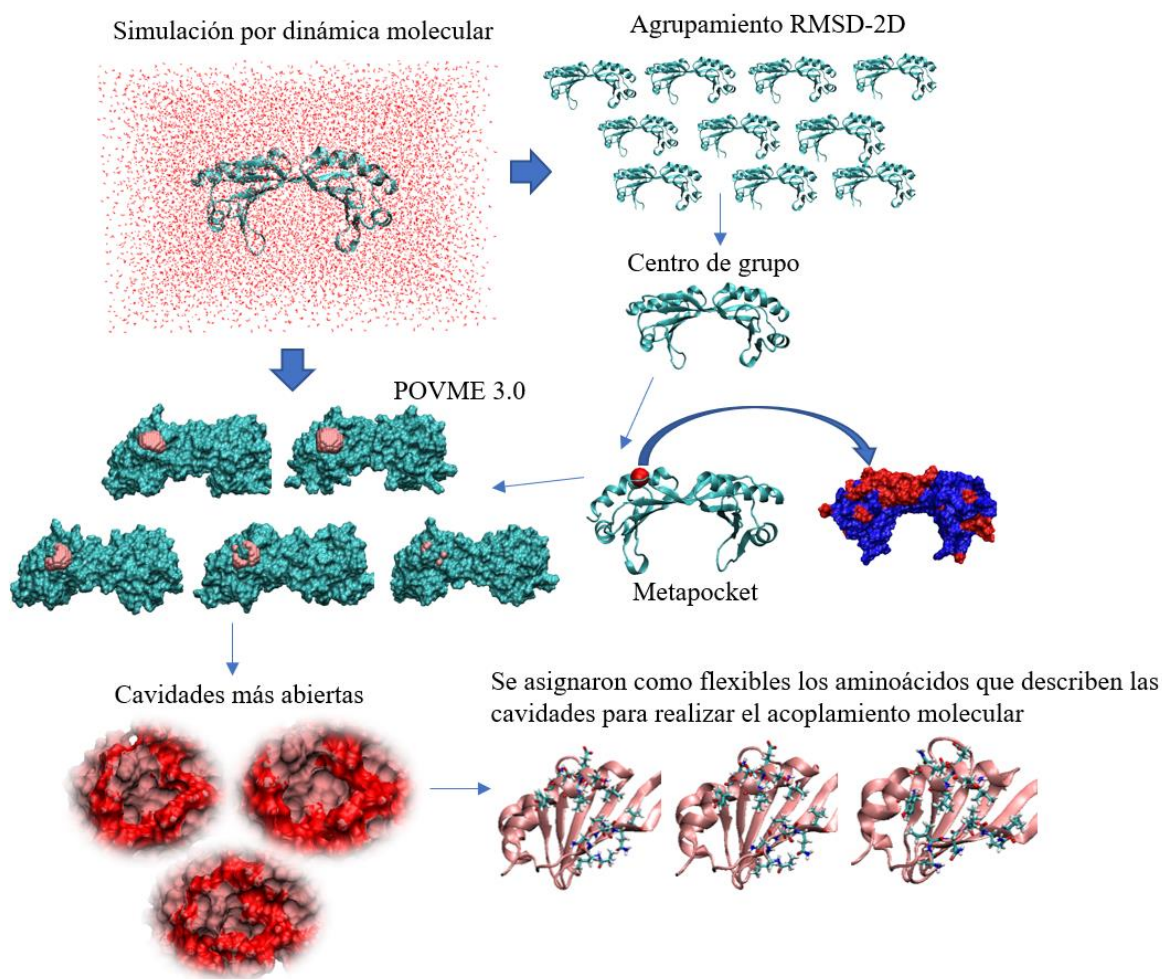


Figura 1: Diagrama del procedimiento computacional empleado para construir las colecciones de estructuras a partir de las dinámicas moleculares para llevar a cabo el cribado virtual.

3.1 Búsqueda de secuencias completas de TBP y TRF1

La búsqueda de las secuencias en formato FASTA de los factores de transcripción se realizó en el NCBI. Se buscaron secuencias de TBP de artrópodos vectores que poseyeran los 180 aminoácidos de unión a ADN. Para extraer la secuencia de 180 aminoácidos de interés para este estudio, se tomó como referencia la secuencia de la estructura cristalográfica de la TBP de humanos 1NVP, y se realizó un alineamiento de ésta, con las secuencias en formato FASTA de los factores de transcripción extraídos del NCBI, con el servidor Clustal Omega (107).

3.2 Modelado por homología de candidatos

Teniendo en cuenta que la gran conservación en secuencia que presentan las TBP es un elemento muy favorable para el modelado por homología, se generaron tres modelos con tres servidores diferentes por cada factor de transcripción. Los modelos fueron construidos empleado tres predictores de estructuras: I-TASSER (108), MODELLER (109) y Phyre2 (110). Inicialmente los factores de transcripción propuestos se modelaron con Phyre2, que genera los modelos de manera rápida; a estos modelos se le realizó el análisis de conservación en secuencia y su mapeo sobre toda la superficie de la proteína. Una vez seleccionados los factores de transcripción con mayor divergencia para trabajar en este proyecto, fueron modeladas sus estructuras con I-Tasser y Modeller.

3.3 Análisis de diferencias en la superficie de TBP y TRF1 respecto a TBP de humano

Teniendo los modelos de los factores de transcripción, se realizó una alineación y superposición estructural para lo cual se utilizó el programa VMD (111). Esto se realizó para determinar diferencias en las superficies de los factores de transcripción de las cinco especies de mosquitos vectores con respecto a la TBP de humanos. Posteriormente se estudiaron estas zonas de divergencia para determinar si en ellas se encontraban o formaban cavidades potencialmente selectivas. El análisis de identidad y mapeo de divergencia sobre la superficie de los factores de transcripción provee información valiosa acerca de la distribución de estas diferencias; si se encuentran dispersas en pequeños parches, se dificulta encontrar cavidades que posean diferencias significativas y sean, por tanto, selectivas; cuando las diferencias convergen en un área, aumenta la probabilidad de encontrar una cavidad selectiva. Por estas razones se excluyeron los factores de transcripción que presentaran un porcentaje de identidad con la TBP de humanos mayor de 80 % y que no presentaran la zona divergente agrupada en un parche o zona susceptible de ser analizada, y por tanto, fuente de una cavidad divergente.

3.4 Simulaciones por dinámica molecular

Las simulaciones por dinámica molecular fueron realizadas para generar los ensambles de estructuras que iban a ser utilizadas en el acoplamiento molecular. La idea es que los ensambles creados describan a las cavidades a lo largo de la simulación y por tanto representen la dinámica de la superficie de la proteína. Por ese motivo fueron seleccionadas cavidades que se abrieran varias veces a lo largo de la simulación, no solamente en un momento determinado.

Las simulaciones moleculares también proporcionan información acerca de los modelos generados; los factores de transcripción modelados son proteínas ordenadas y por tanto su estructura no debe poseer una variación significativa con respecto a la estructura inicial de la simulación. Esta estructura inicial es el modelo generado, minimizado y con hidrógenos, por tanto, se espera que la dinámica molecular converja en un momento dado y se mantenga estable por el resto de la simulación. Para medir esto se calcularon los RMSD respecto a la estructura inicial en función del tiempo, y además, se compararon todas las estructuras de cada simulación por pares. Posteriormente se realizó un agrupamiento estructural con corte de 2 Å a nivel de cadena principal de todas las simulaciones, y se superpusieron los centros de grupo.

Para realizar las simulaciones por dinámica molecular las estructuras fueron optimizadas previamente. Se le añadieron hidrógenos empleando el servidor CHARMM_GUI (112) y se emplearon 50 pasos de minimización utilizando el método de *Steepest descent* sobre los hidrógenos manteniendo rígido el resto de la estructura. Posteriormente se realizó una minimización también de 50 pasos sobre todos los átomos del sistema empleando el esquema *Conjugate gradient*. Estas minimizaciones se realizaron con el potencial CHARMM36 (113) y el programa charmm38b1 (114).

Las simulaciones por dinámica molecular fueron realizadas con el programa NAMD2.11 (115) con un tiempo de integración de 2 fs y una frecuencia de guardado de 500 pasos. Se utilizó el potencial CHARMM36. Las simulaciones se realizaron en solvente explícito e incluyen aproximadamente 36000 átomos. Se empleó una caja ortorrómbica de dimensión de 92x64x60 Å³ construida con el servidor CHARMM-GUI con una concentración de 0.15M de NaCl. Las simulaciones se realizaron en un ensamble NPT. Fueron simulados 101 ns para cada estructura a una temperatura de 323 K, esto con el objetivo de lograr mayor movimiento de la proteína y por tanto un mayor muestreo conformacional. El primer ns se tomó como tiempo de equilibración para las trayectorias.

Posteriormente se efectuó un agrupamiento para lo cual se calculó un RMSD en dos dimensiones (RMSD-2D) sobre los átomos de la cadena principal, con un corte de 2.0 Å. Esto se efectuó con el objetivo de ver la diversidad estructural de los ensambles generados en la dinámica molecular. Para todas las estructuras se obtuvo un grupo que representaba a más del 75% de las estructuras generadas por la dinámica, con los tres primeros grupos más poblados se describe más de un 95% de las estructuras obtenidas de la simulación, lo cual

sugiere una buena convergencia de las dinámicas. Para llevar a cabo este análisis se utilizó el programa SIMULOID (116).

3.5 Localización de cavidades y selección de estructuras para acoplamiento molecular

El siguiente objetivo fue localizar las cavidades formadas a lo largo de la simulación en cada factor de transcripción, y que se situaran en zonas divergentes en secuencia, para proponerlas como sitios potenciales de unión a compuestos. Para ello se tomaron las estructuras de los centros de grupos más poblados calculadas previamente con el programa SIMULOID y se enviaron al servidor Metapocket (117), que detecta cavidades a lo largo de toda la superficie de la proteína. Posteriormente se seleccionaron las cavidades que se encontraban en zona de divergencia con respecto a la TBP de humano. Las coordenadas del centro de las cavidades detectadas por Metapocket que fueron seleccionadas como potencialmente selectivas, posteriormente fueron empleadas por el programa POVME 3.0 (118) para calcular su volumen a lo largo de la simulación. Estos cálculos se efectuaron con espaciamiento de 1.0 Å, una distancia de corte de 1.09 Å y esferas de inclusión de 5 Å. Este cálculo se realizó sobre las 3000 estructuras derivadas de las simulaciones de cada factor de transcripción. Por cada cavidad seleccionada se extrajeron las cinco estructuras con mayores volúmenes de acuerdo al cálculo realizado con POVME, y que se encontraran distribuidas a lo largo de la simulación para no sesgar el muestreo. Con las estructuras seleccionadas se construyó un ensamble de cinco estructuras por cada cavidad.

3.6 Selección de compuestos útiles en ZINC15

Los ligandos fueron obtenidos de la base de datos ZINC15 en formato pdbqt, porque es el formato empleado por Autodock Vina, y contiene información sobre los radios y cargas atómicas tanto de la proteína como del ligando, además de la lista de los enlaces sobre los que se permite libre giro.

Los factores de transcripción de interés en este estudio se encuentran dentro del núcleo de las células. La membrana nuclear presenta poros de gran tamaño (119), por tanto, fueron seleccionados compuestos que fueran capaces de atravesar las membranas celulares, puesto que, un vez que los compuestos de interés se encuentren dentro de citoplasma, existe una alta probabilidad de que lleguen al interior del núcleo. Mälkiä, A. y colaboradores plantean que las principales propiedades de las moléculas relacionadas a la permeabilidad de membranas biológicas son la lipofilidad, capacidad de formación de puentes de hidrógenos, la carga y el tamaño (120). Por esos motivos, la selección de los compuestos se hizo tomando en cuenta las siguientes propiedades: hasta 500 Da de peso molecular, LogP mayor que 0 y menor que 5, con presencia de hasta 5 donadores y 10 aceptores de enlaces de hidrógeno. Teniendo en cuenta que los resultados de los programas de acoplamiento molecular comienzan a perder

precisión en presencia de ligandos con carga y con más de 10 enlaces rotables, fueron seleccionados ligandos neutros y con hasta siete enlaces rotables. Finalmente fueron seleccionados 6200 ligandos.

3.7 Cribado molecular

El acoplamiento molecular se realizó utilizando el programa Autodock Vina. Para realizar el cribado molecular se construyó una colección de estructuras empleando la estrategia que se muestra en la Figura 2.

Las colecciones construidas de cinco estructuras por cada género de mosquitos fueron utilizadas para hacer el cribado molecular empleando los 6200 ligandos obtenidos de la base de datos ZINC 15. El acoplamiento molecular se efectuó de forma dirigida sobre las zonas de interés, asignando como flexibles los aminoácidos que describen las cavidades (excluyendo glicina, alanina y prolina) y los ligandos. La caja empleada para el acoplamiento molecular se construyó con un espaciamiento de 1 Å; las dimensiones se asignaron con la ayuda del plugin de Vina en Pymol 1.7.7.2 (121) para su uso en el acoplamiento molecular. El acoplamiento molecular flexible se realizó con una exhaustividad de 24 y generando cinco poses para cada compuesto. La exhaustividad controla cuantas veces se repiten los cálculos, y es el parámetro con el cual Vina controla el número de evaluaciones durante cada optimización local, partiendo de conformaciones aleatorias. La exhaustividad que trae el programa preconcebidamente es de 8. En este trabajo se aumentó a 24 para mejorar el muestreo.

3.8 Análisis de la selectividad con respecto a la TBP de humanos.

Se tomaron las estructuras cristalográficas de la TBP de humanos 1NVP, 1NGM y 1C9B depositadas en el PDB y se les realizó dinámica molecular. Las dinámicas moleculares se realizaron siguiendo fielmente el protocolo descrito en el acápite 3.4.

La conformación de las colecciones de estructuras para el cribado molecular en humano se realizó utilizando exactamente el mismo protocolo que para generar las colecciones en los factores de transcripción de mosquitos, procedimiento descrito en el acápite 3.5. Para la TBP de humanos fueron conformados ensamblajes análogos a todas las cavidades de vectores analizadas, donde fueron evaluados los mismos compuestos.

3.9 Análisis toxicológico de los compuestos propuestos

Los compuestos propuestos fueron analizados empleando el predictor de toxicidad admetSAR (122). Fueron analizados parámetros de toxicidad para mamíferos, en este caso, de carcinogenicidad y de mutagénesis de AMES. El ensayo de AMES está diseñado para detectar un amplio rango de compuestos que pueden causar daños genéticos y producir mutaciones mediante el análisis de revertantes en *Salmonella* (123). Fueron analizados estos parámetros por la importancia de conocer si los compuestos que se están proponiendo pueden ser potencialmente cancerígenos o mutagénicos. Además, fueron analizados parámetros de predicción de corrosión e irritación ocular por la enorme importancia que tiene conocer si los compuestos propuestos pueden ser potencialmente dañinos o corrosivos para los ojos, puesto que la mayoría de los insecticidas son empleados mediante atomización y pueden entrar en contacto con ellos. El servidor admetSAR contiene bases de datos de disímiles moléculas que presentan toxicidad para diferentes sistemas de órganos y funciones ambientales. Se suministraron a dicho servidor las estructuras químicas en formato SMILE de los compuestos propuestos como potenciales insecticidas para que predijera si eran potencialmente tóxicos. Este servidor compara las estructuras proporcionadas con su base de datos propia, y, empleando algoritmos de aprendizaje automático (*machine learning*), que incluyen bosque aleatorio (*random forest*, RF), máquina de vectores de soporte (*support vector machine*, SVM) y vecinos k-cercanos (*k-nearest neighbors*, kNN) predice por similitud si las moléculas proporcionadas son potencialmente tóxicas (122). Fueron analizados también parámetros de biodegradabilidad y de toxicidad en abejas, peces y crustáceos porque estos parámetros describen potenciales daños al ambiente.

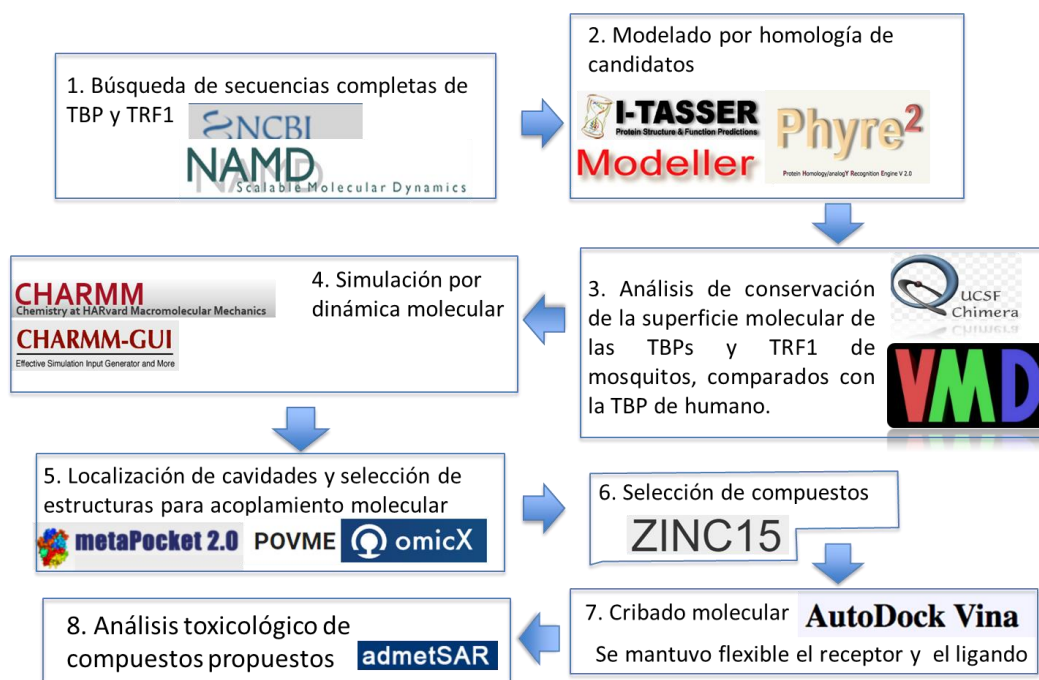


Figura 2: Diagrama de flujo resumen de la metodología empleada.

4. Resultados

4.1 Búsqueda de secuencias de TBPs y TRF1

Secuencias encontradas en el NCBI

Factor de Transcripción	Número de Referencia	Porcentaje de identidad respecto a la TBP de humano
<i>Aedes aegypti</i> (TBP)	XP_001662432.1	87.22 %
<i>Aedes albopictus</i> (TBP)	XP_019552045.1	87.78 %
<i>Aedes aegypti</i> (TRF1)	XP_001654561.1	61.67 %
<i>Aedes albopictus</i> (TRF1)	XP_029709094.1	62.78 %
<i>Culex quinquefasciatus</i> (TBP)	XP_001849783.1	63.33 %
<i>Anopheles darligi</i> (TBP)	ETN59952.1	62,78 %
<i>Anopheles sinensis</i> (TBP)	KFB39418.1	61, 67 %

Tabla 1: Secuencias de factores de transcripción pertenecientes a la familia de las TBP encontradas en el NCBI. Porcentaje de identidad con respecto a la TBP de humanos.

4.2 Generación de Modelos

Los modelos generados fueron estables estructuralmente puesto que en las simulaciones por dinámica molecular convergieron y se estabilizaron, como se espera ocurra en modelos de proteínas ordenadas. Además, los parámetros de valoración de los modelos, generados por cada predictor de estructuras, resultaron todos satisfactorios.

Factor de Transcripción	I-TASSER		MODELLER	Phyre2
	C-Score	TM-Score	GA341	Conf. promedio
<i>Aedes aegypti</i> (TRF1)	1.60	0.94±0.05	1.00	0.98
<i>Aedes albopictus</i> (TRF1)	1.61	0.94±0.05	1.00	0.98
<i>Culex quinquefasciatus</i> (TBP)	1.62	0.94±0.05	1.00	0.98
<i>Anopheles darligi</i> (TBP)	1.61	0.94±0.05	1.00	0.98
<i>Anopheles sinensis</i> (TBP)	1.61	0.94±0.05	1.00	0.98

Tabla 2: Parámetros de evaluación de los Modelos Generados

4.3 Análisis de diferencias en la superficie de TBP y TRF1 respecto a TBP de humano

El análisis de diferencias en secuencia, y el agrupamiento de estas diferencias en las superficies de los factores de transcripción, se utilizó como base para la selección de los factores de transcripción definitivos a estudiar. En la figura 3 se observan estas diferencias mapeadas en las superficies de las TBP de ambas especies de mosquitos del género *Aedes*. En las TBPs de ambas especies existe aproximadamente un 90% de identidad en secuencia con respecto a la TBP de humano, lo cual representa una alta coincidencia. Además, las diferencias se encuentran esparcidas en pequeños parches a lo largo de la superficie, lo que tampoco favorece las probabilidades de encontrar una cavidad selectiva. Por estas razones, las TBP de *Aedes aegypti* y *Aedes albopictus* fueron excluidas de este estudio, puesto que encontrar diferencias estructurales en sus superficies es improbable, y por tanto la presencia de cavidades divergentes con respecto a la TBP de humanos que puedan ser fuente de selectividad. Entonces, de los mosquitos del género *Aedes*, fueron seleccionados los TRF1, que presentan una divergencia significativa y por tanto fueron considerados potenciales blancos de acceso de fármacos selectivos.

En los restantes factores de transcripción analizados, dígame los TRF1 de *Aedes* y las TBPs de *Anopheles* y *Culex*, como se refleja en la figura 4, se encontró una divergencia con respecto a la TBP de humanos de aproximadamente cuarenta por ciento. Además, gran parte de la superficie divergente se agrupó en una región extensa apta de ser analizada y, por tanto, fuente potencial de cavidades selectivas.

Finalmente, los factores de transcripción considerados para este estudio fueron: Las proteínas de unión a caja TATA (TBPs) de: *Anopheles darlingi* GenBank: ETN59952.1, *Anopheles sinensis* GenBank: KFB39418.1, *Culex quinquefasciatus* NCBI Reference Sequence: XP_001849783.1; y los TRF1 de: *Aedes aegypti* NCBI Reference Sequence: XP_001654561.1 y *Aedes albopictus* NCBI Reference Sequence: XP_029709094.1.

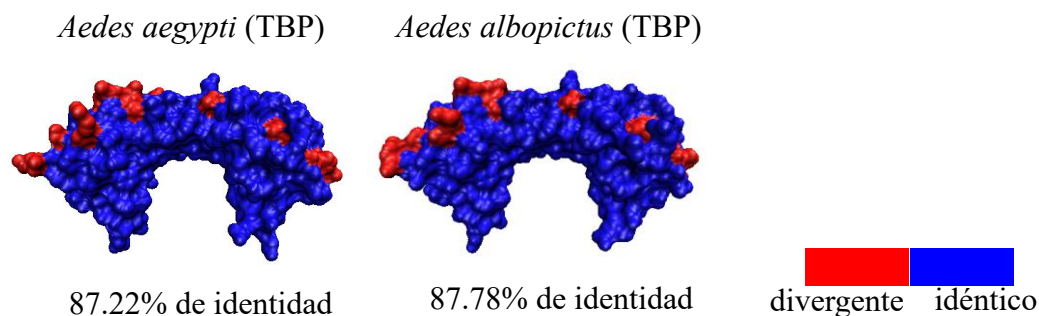


Figura 3: Porcentaje de identidad en secuencia con respecto a la TBP de humanos de las TBPs de *Aedes aegypti* y *Aedes albopictus*. Mapeo sobre la superficie molecular de estas diferencias en secuencia: el color rojo representa los aminoácidos diferentes con respecto a la TBP de humanos y el color azul representa los que son idénticos.

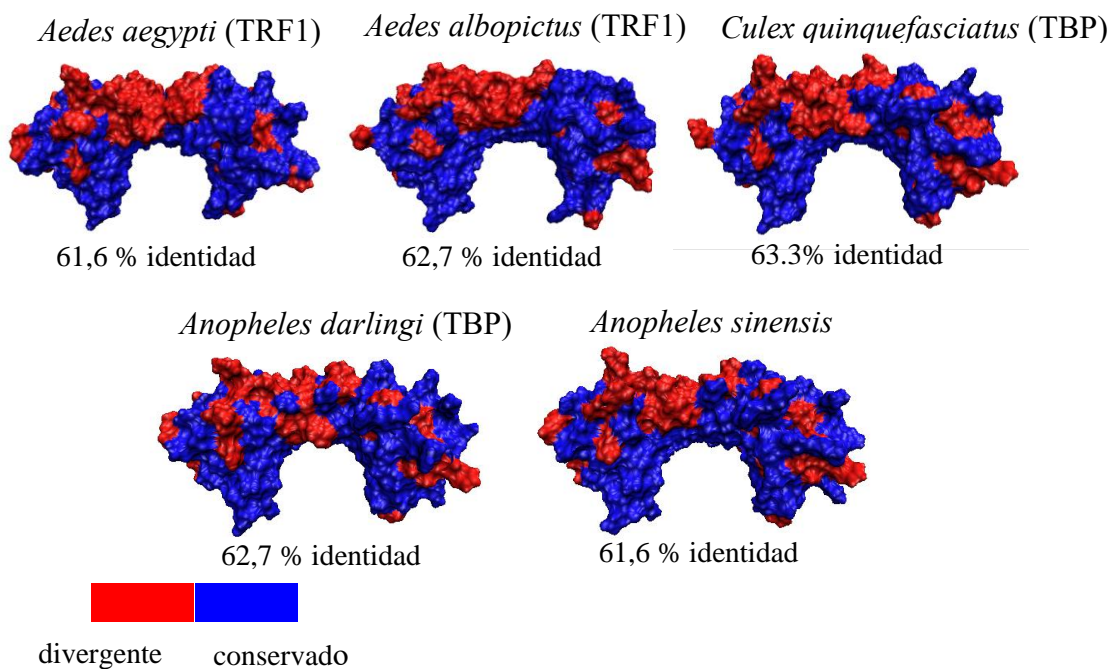


Figura 4: Diferencias en superficie y porcentaje de identidad con respecto a humanos de las TRF1s de *Aedes aegypti* y *Aedes albopictus*, y las TBPs de *Culex quinquefasciatus*, *Anopheles darlingi* y *Anopheles sinensis*. Mapeo sobre la superficie molecular de estas diferencias en secuencia: el color rojo representa los aminoácidos diferentes con respecto a la TBP de humanos y el color azul representa los que son idénticos.

4.4 Simulaciones por dinámica molecular

Las simulaciones obtenidas convergieron y se estabilizaron; esto habla de persistencia y calidad de los modelos generados. Se obtuvo en todos los casos un grupo más poblado que describió a más de un 75 % de las estructuras obtenidas en la simulación. Los tres primeros grupos más poblados describieron a más de un 95 % del total de estructuras generadas en la dinámica. La tabla 3 muestra los principales grupos de las simulaciones de todos los factores de transcripción con un corte de 2 Å.

En la figura 5 se muestran los RMSD respecto a la estructura inicial a lo largo del tiempo de todas las simulaciones. En cada gráfica se observa en color azul el RMSD de todas las estructuras generadas por dinámica molecular a partir de los modelos creados por el servidor Phyre2, en color verde las estructuras generadas a partir de los modelos creados por I-TASSER y en color rojo las estructuras generadas a partir de los modelos proporcionados por MODELLER. En todos los casos el RMSD obtenido es inferior a 4 Å, lo cual es indicador de movimientos moderados propios de proteínas con estructuras ordenadas. La figura 6

muestra una superposición de los tres primeros centros de grupos más poblados. Esta superposición ilustra que las diferencias más significativas se encuentran en los estribos de las proteínas y en los extremos amino y carboxilo terminales. En estas zonas de la proteína es donde tienen lugar los mayores movimientos y, por tanto, se infiere sean estos fragmentos los responsables de los mayores RMSD.

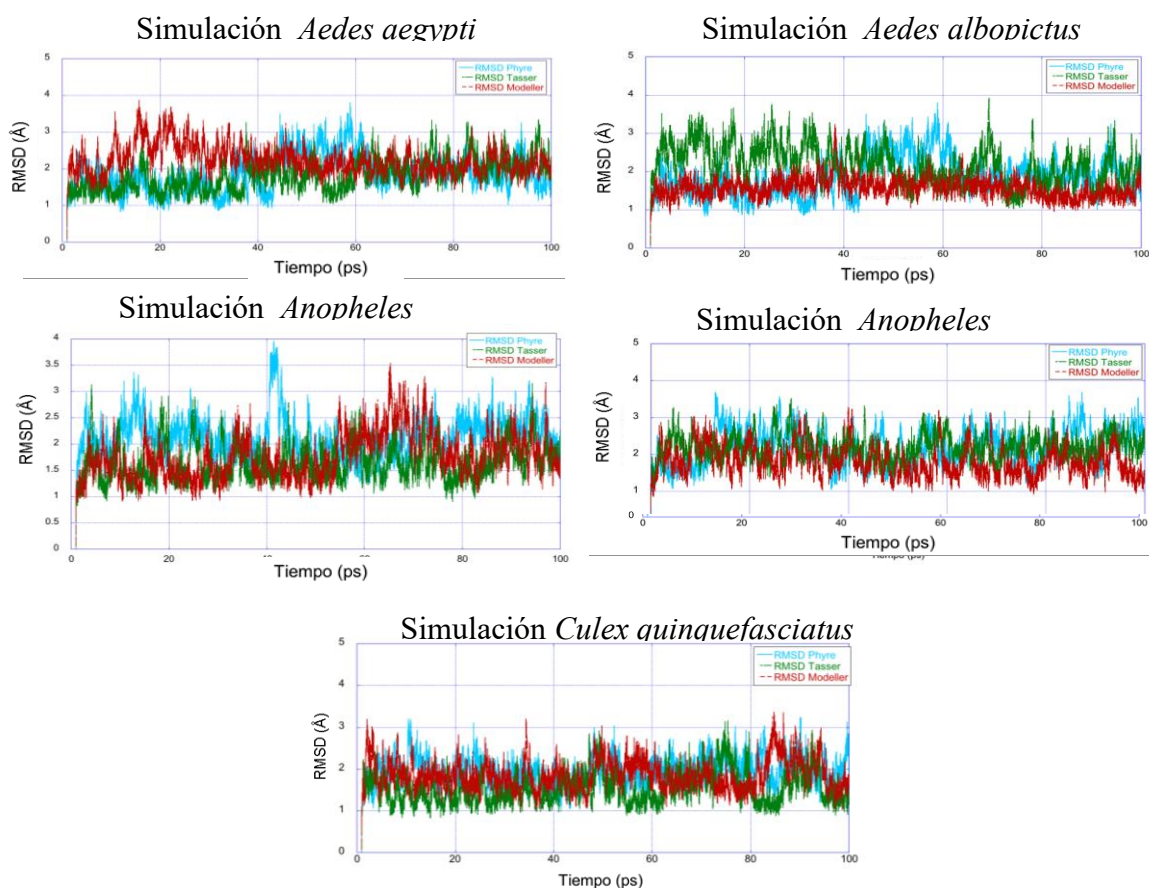


Figura 5: RMSD de las simulaciones de todos los modelos de los factores de transcripción. En azul se muestran los RMSD de las simulaciones realizadas a partir de los modelos generados por el servidor Phyre2, en color verde de I_TASSER y en color rojo de MODELLER.

	<i>Aedes aegypti</i> (TRF1)	<i>Aedes albopictus</i> (TRF1)	<i>Anopheles darlingi</i> (TBP)	<i>Anopheles sinesis</i> (TBP)	<i>Culex quinquefasciantus</i> (TBP)
Primer grupo más poblado	<u>2534</u> (84.4%)	<u>2776</u> (92.5%)	<u>2529</u> (84.3%)	<u>2365</u> (78.8%)	<u>2671</u> (89.0%)
Segundo grupo más poblado	<u>267</u> (8.9%)	<u>96</u> (3.2%)	<u>148</u> (4.9%)	<u>341</u> (11.36%)	<u>165</u> (5.5%)
Tercer grupo más poblado	<u>103</u> (3.4%)	<u>71</u> (2.3%)	<u>143</u> (4.7%)	<u>182</u> (6.0%)	<u>111</u> (3.7%)

Tabla 3: Grupos más poblados de las simulaciones realizadas. Cantidad de estructuras representadas en cada grupo y porcentaje que representa cada grupo de la totalidad de estructuras.

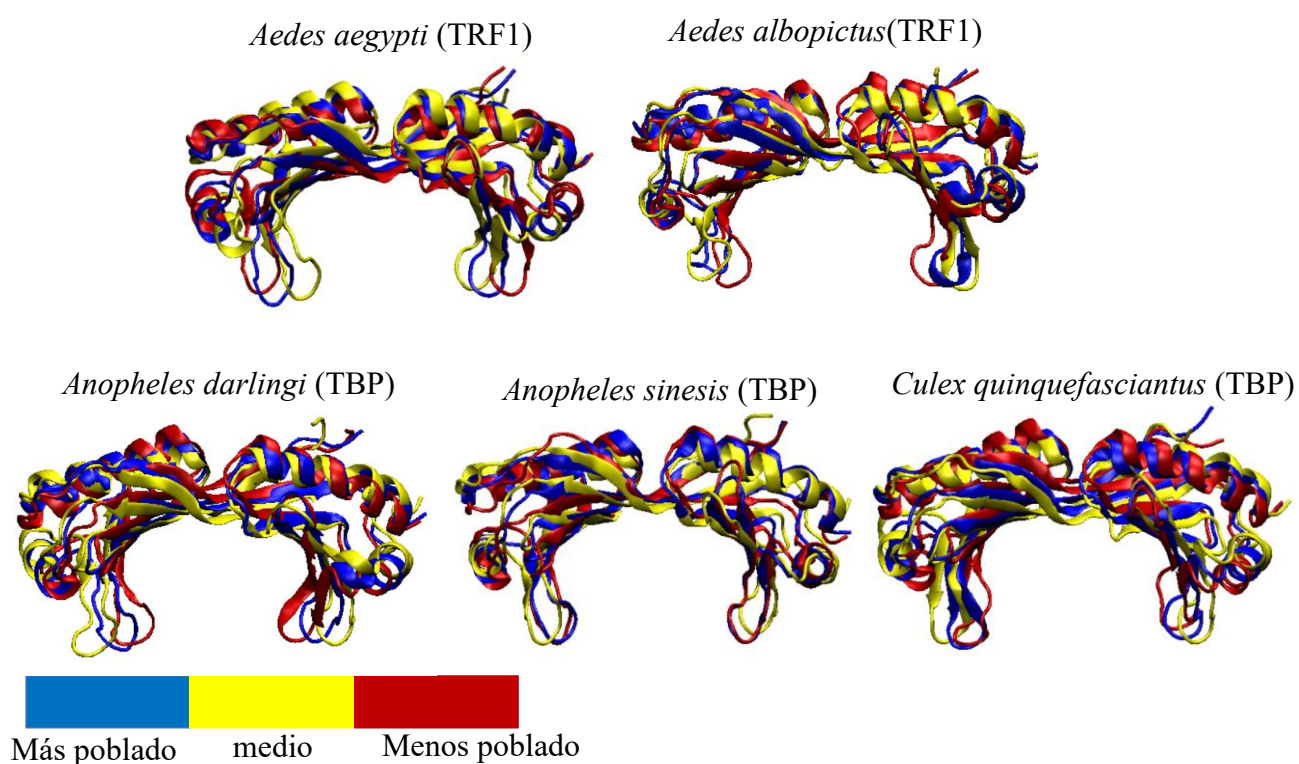


Figura 6: Superposición de los centros de los tres grupos más poblados de las dinámicas moleculares de las especies estudiadas. En azul el centro del grupo más poblado, en amarillo en centro del grupo medianamente poblado y en rojo el centro del grupo con menos población de los tres.

4.5 Localización de cavidades y selección de estructuras para acoplamiento molecular

Analizados los centros de grupo de todas las dinámicas con el servidor Metapocket, se seleccionaron las cavidades que coincidieran con parches de divergencia en las superficies de los factores de transcripción. La figura 7 muestra la totalidad de las cavidades detectadas por Metapocket en la superficie de los centros de grupo más poblados de los factores de transcripción analizados. Se encuentra señalada una cavidad que se encontró en zona de considerable divergencia y que coincide con parte del sitio de unión de Brf1, por lo cual se seleccionó como una cavidad prometedora a ser analizada.

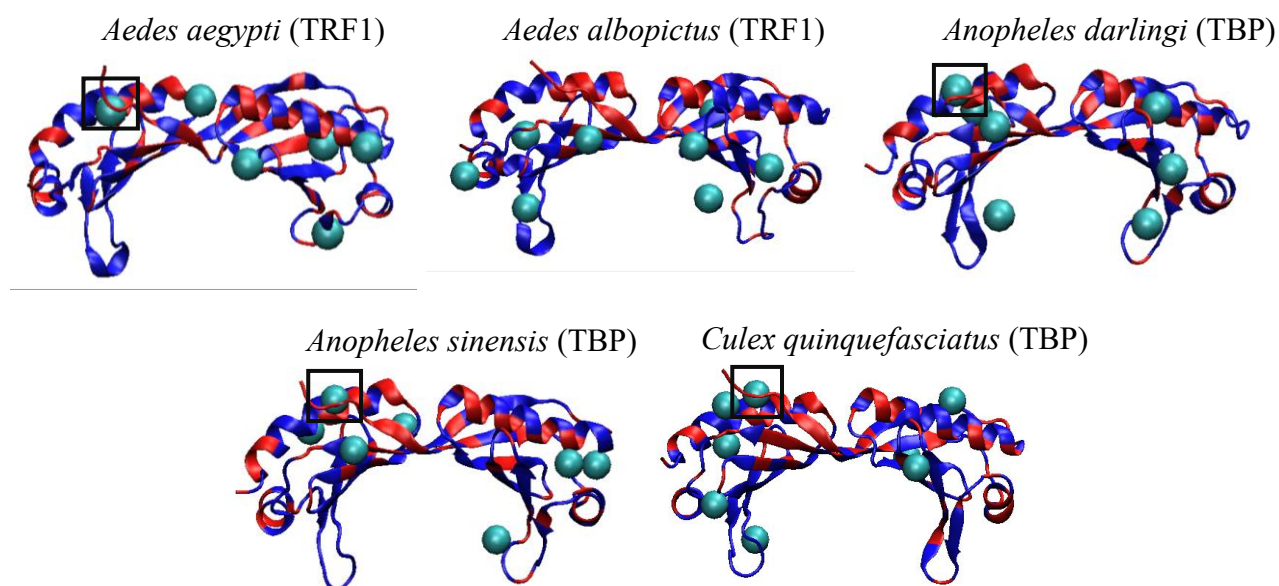


Figura 7: Cavidades detectadas por Metapocket en la superficie de los cinco factores de transcripción analizados (esferas color cyan). Se encuentra señalada en cada estructura una potencial cavidad selectiva, que coincide con parte del sitio de unión de Brf1. En rojo se muestran los residuos divergentes con respecto a TBP de humanos, y en azul, los residuos idénticos.

Nótese que no en todas las estructuras en las cuales se realizó este análisis fueron detectadas las mismas cavidades. En la figura 8A se muestran más detalladamente estas diferencias. En *Aedes aegypti* Metapocket detectó la cavidad de interés en la zona divergente situada en la región de unión al sitio Brf1, sin embargo, esta cavidad no fue detectada en *Aedes albopictus*.

En la figura 8B se muestra un análisis de conservación de la superficie en estas dos especies, lo cual sugiere que los resultados, tanto de análisis de cavidades como posteriormente de acoplamiento molecular, pueden ser extrapolables dentro del género *Aedes*.

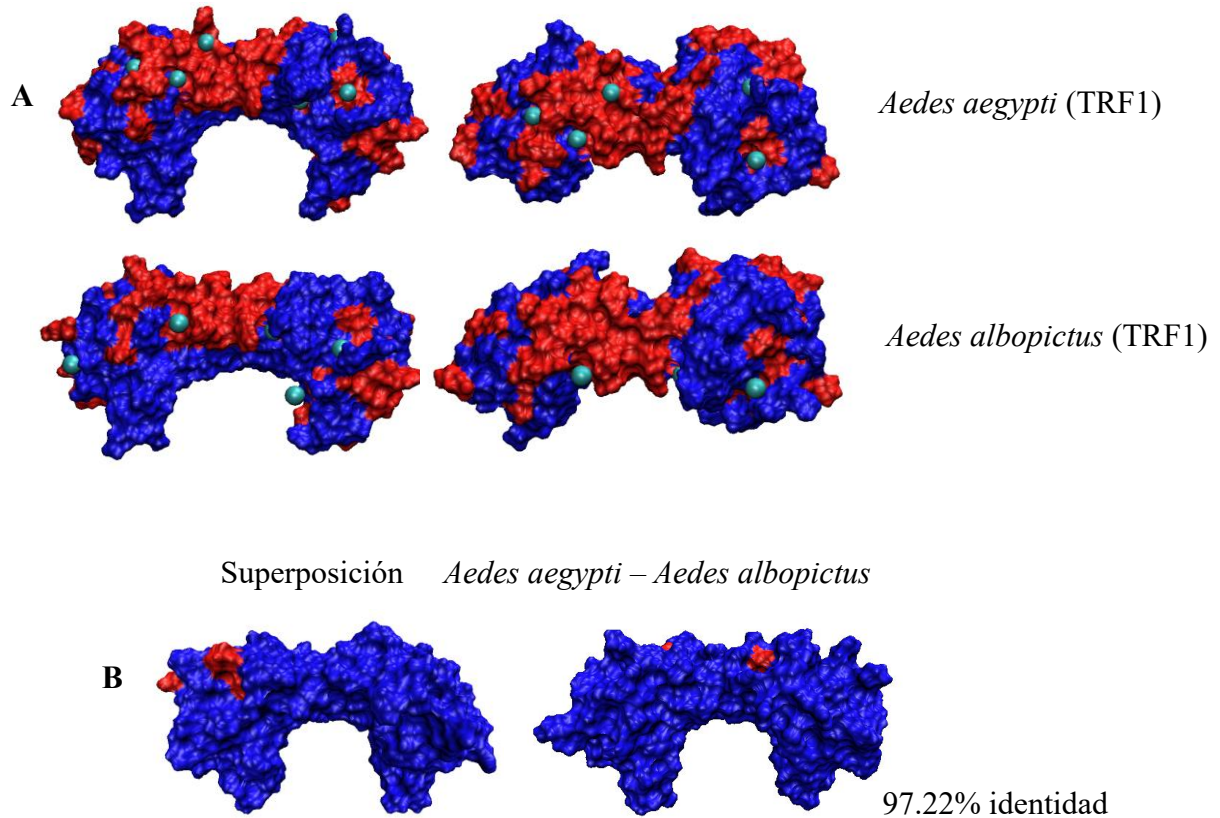


Figura 8: A: Cavidades detectadas por el servidor Metapocket en las superficies de los centros de grupo más poblados de las especies *Aedes aegypti* y *Aedes albopictus*; se marca también la divergencia (rojo) e identidad (azul) con respecto a la TBP de humano. B: Superposición de las TRF1 de ambas especies de *Aedes* y análisis de conservación, la superficie azul muestra la zona de identidad y la superficie roja la zona de divergencia entre ambas especies.

De forma similar ocurrió con las TBP de las especies de *Anopheles*. En la especie *Anopheles darlingi* fueron detectadas cavidades que no se detectaron en *Anopheles sinensis* y viceversa. Ambas especies tienen alto porcentaje de identidad por lo que, de igual manera, se infiere que los resultados encontrados pueden ser extrapolables dentro de las dos estructuras analizadas del género *Anopheles*.

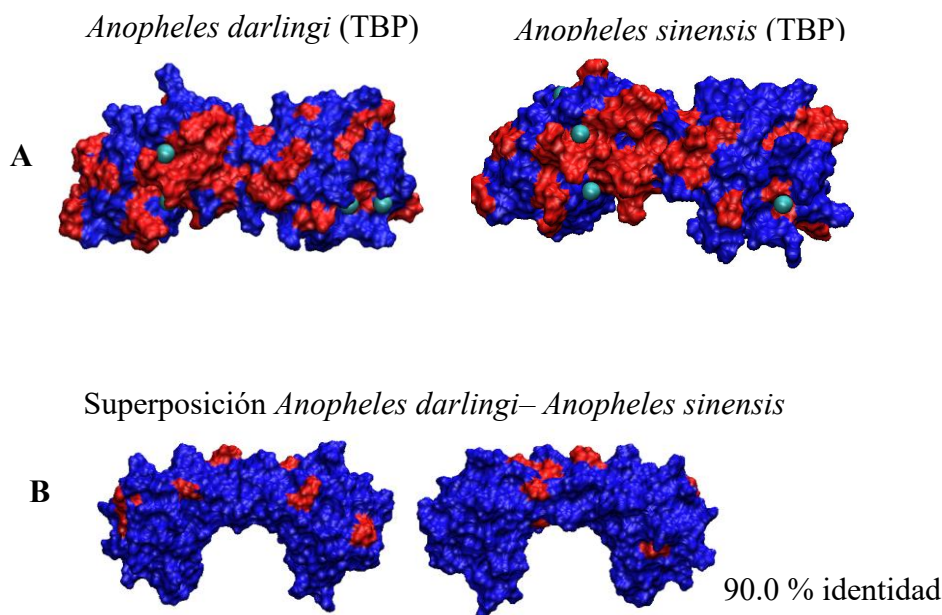


Figura 9: A: Cavidades detectadas por el servidor Metapocket en las superficies de los centros de grupo más poblados de las especies *Anopheles darlingi* y *Anopheles sinensis*. Se señala también la divergencia (rojo) e identidad (azul) con respecto a la TBP de humano. B: Superposición de las TBP de *Anopheles darlingi* y *Anopheles sinensis* y su análisis de conservación, la superficie azul muestra la zona de identidad y la superficie roja la zona de divergencia entre ambas especies.

En el caso de *Culex quinquefasciatus*, sólo se cuenta con esa única especie por género. La figura 10 muestra las cavidades detectadas por Metapocket en los dos centros de los grupos más poblados. Ocurrió de forma similar: fueron detectadas cavidades distintas en ambas estructuras.

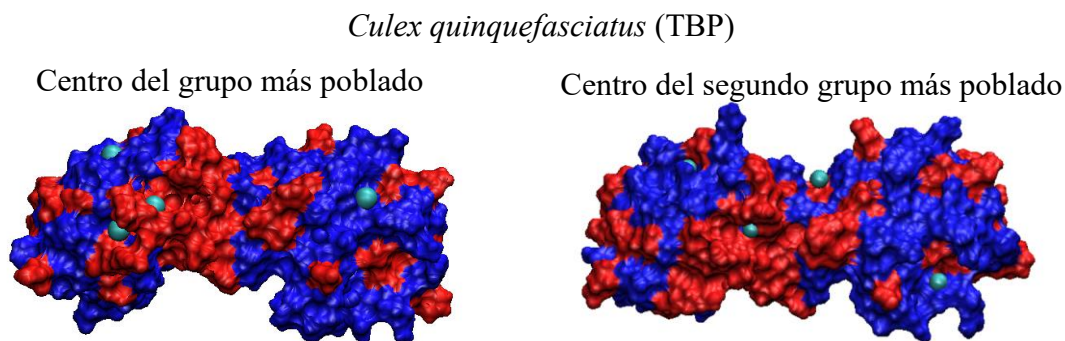


Figura 10: Cavidades detectadas por el servidor Metapocket en las superficies de los centros de grupo más poblados de la especie *Culex quinquefasciatus*. De la misma forma, se muestra en color rojo zona de divergencia y en color azul los aminoácidos idénticos respecto a TBP de humano.

Para todos los factores de transcripción fueron utilizadas estas coordenadas proveídas por Metapocket para el posterior análisis de las cavidades de interés con el programa POVME 3.0 (118). Se obtuvo el volumen de las cavidades a lo largo de la simulación y se extrajeron las estructuras que presentaran las cavidades más abiertas y que se encontraran distribuidas a lo largo de la simulación. En la figura 11 se muestra este cálculo de volumen realizado por POVME en la cavidad que coincide con parte del sitio de unión de Brf1 (cavidad señalada en la figura 7).

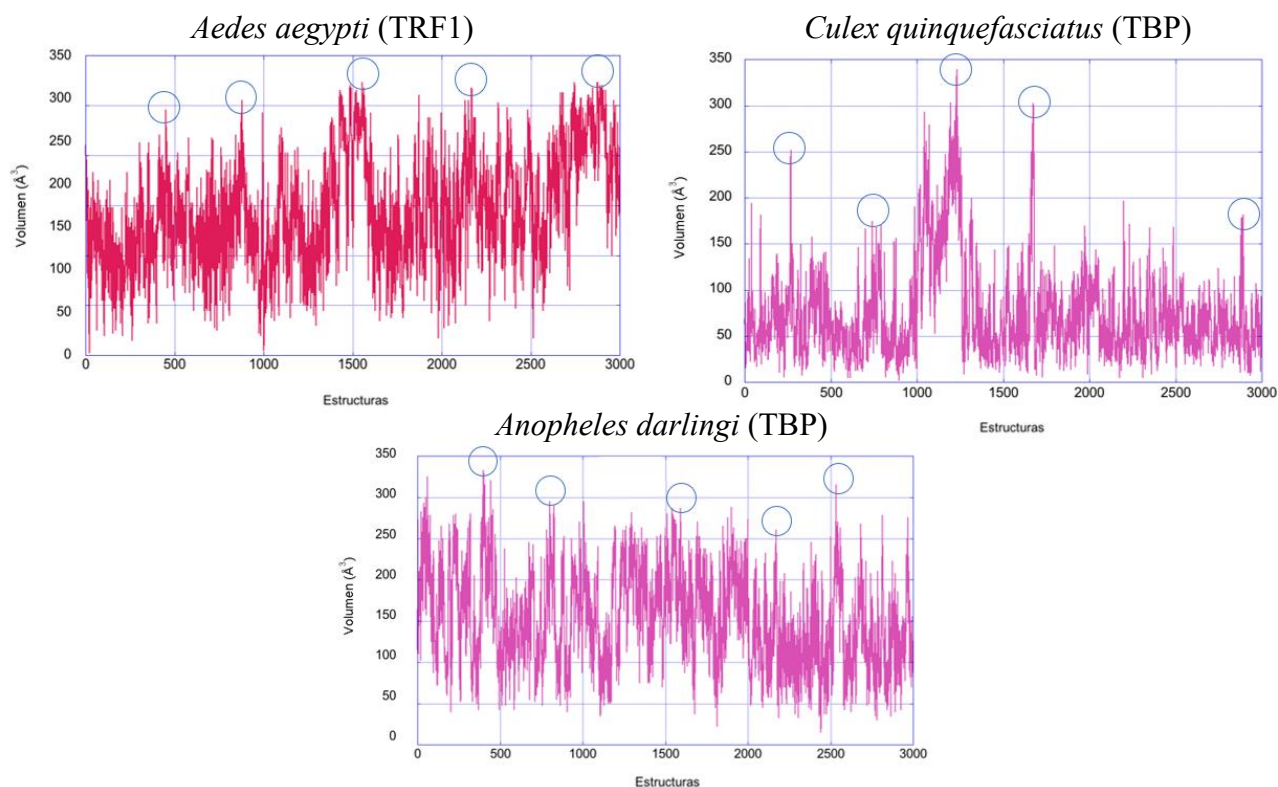


Figura 11: Cálculo de volumen realizado por POVME a lo largo de las simulaciones de los factores de transcripción en la cavidad que corresponde con parte del sitio de interacción de Brf1. Se encuentran señaladas en círculos las estructuras que presentaron mayor volumen en dicha cavidad, y fueron seleccionadas para conformar las colecciones.

Como vimos anteriormente, en los tres géneros de mosquitos fue detectada la cavidad que se muestra en la figura 12. Esta cavidad coincide con la mayor zona de divergencia encontrada en los factores de transcripción analizados y con parte de la zona de unión de Brf1, por lo que se infirió sería una cavidad prometedora. Ahí se llevó a cabo el cribado molecular de los 6200 ligandos obtenidos de la base de datos ZINC 15. Esta misma zona fue analizada en la TBP de humanos para determinar selectividad.

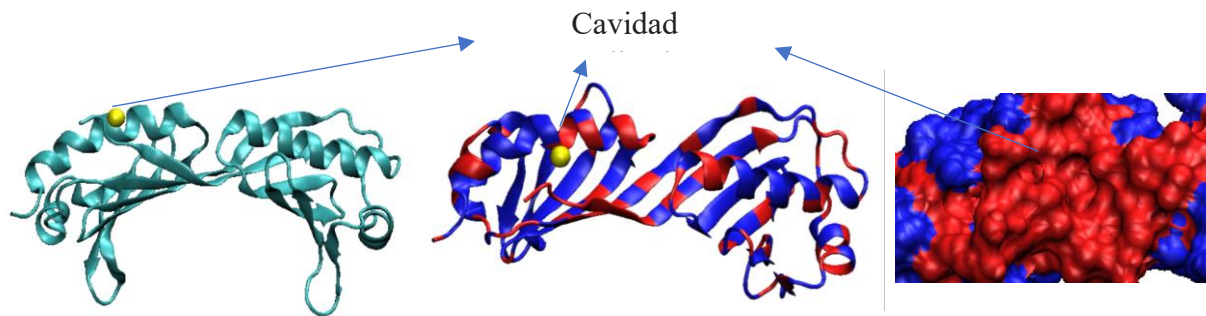
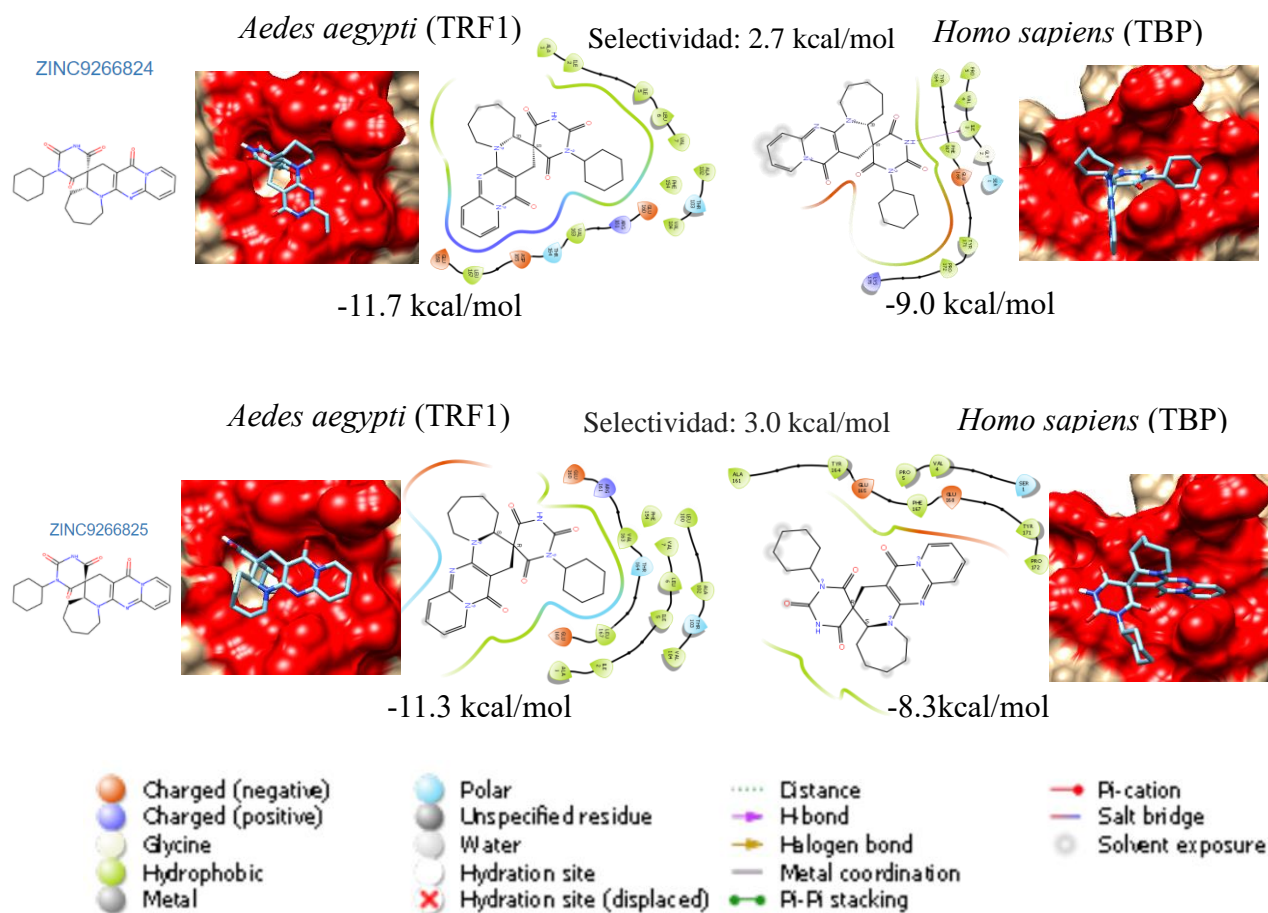


Figura 12: Cavidad donde se realizó en cribado molecular. La esfera de color amarillo representa el centro de la cavidad que se muestra. El color azul muestra zona de identidad y el color rojo zona de divergencia con respecto a la TBP de humanos.

6.6 Mejores candidatos de acuerdo a energías y selectividad del acoplamiento molecular

Se encontraron compuestos que se unieron preferencialmente a los factores de transcripción de los mosquitos vectores con respecto a la TBP de humanos. En las figuras 13, 14 y 15 se muestran los mejores compuestos candidatos en función de las mejores energías de unión y selectividades.




 Aminoácidos asignados como flexibles

Figura 13: Mejores candidatos por selectividad encontrados en *Aedes aegypti* (TRF1). En la figura se muestran las poses de los mejores candidatos encontrados, y en rojo se señalan los aminoácidos que fueron considerados flexibles. Se muestran los mapas de interacciones para explicar los resultados de la selectividad. En rojo se muestran los aminoácidos cargados negativamente, en azul se muestran los aminoácidos cargados positivamente, las glicinas en color blanco hueso, en verde los aminoácidos hidrofóbicos, en gris se muestran metales, en azul los aminoácidos polares, el agua se muestra en gris claro, y sus interacciones en color blanco. Las interacciones no covalentes se muestran de la siguiente forma: los enlaces de hidrógenos con líneas color violeta, enlaces con halógenos son representados con líneas color marrón, la coordinación con metales en líneas color gris claro, las interacciones Pi-Pi en líneas color verde, las interacciones Pi-cation con líneas color rojo, los puentes salinos con líneas moteadas. La exposición al solvente se representa con nubes de color gris.

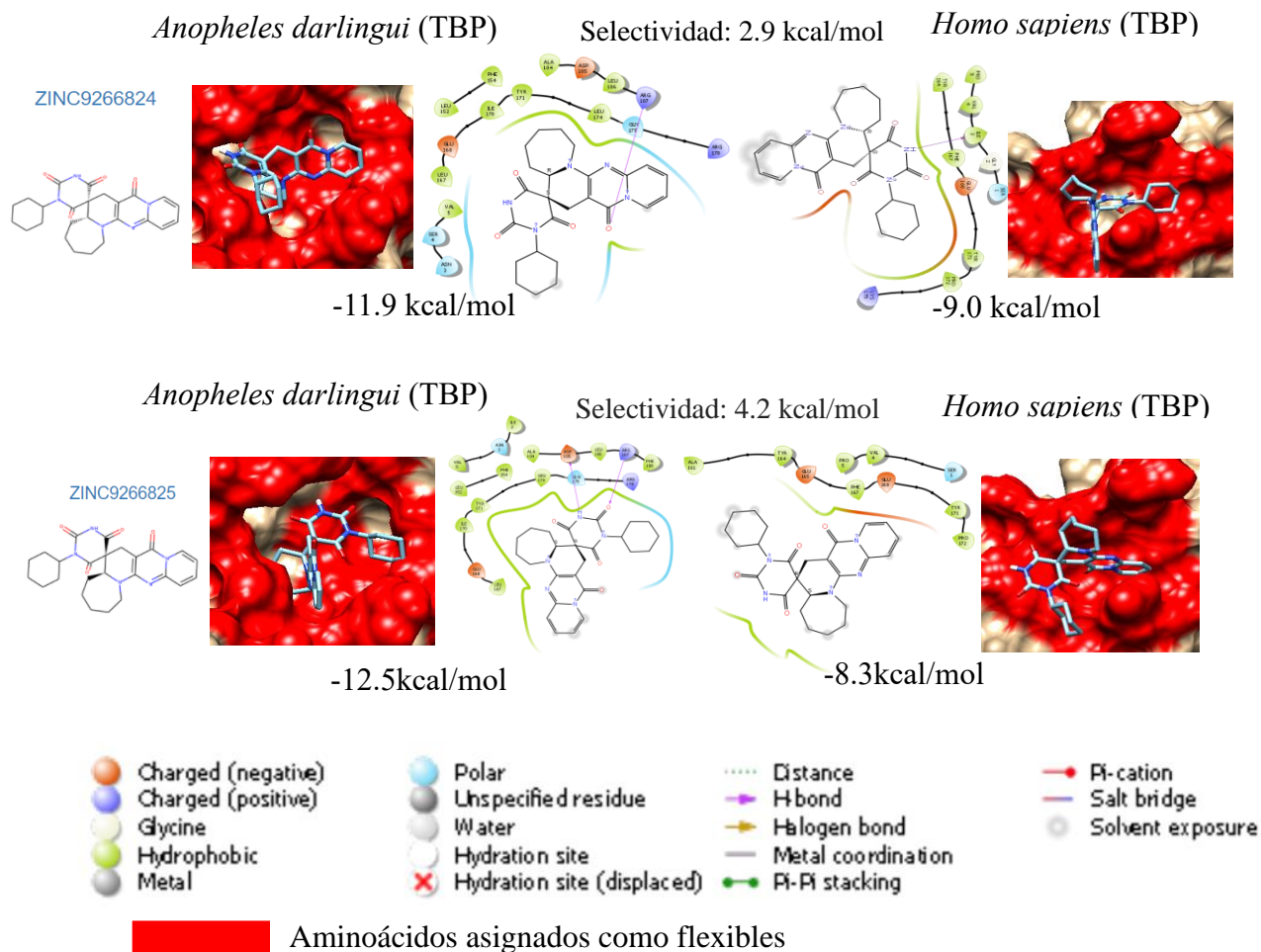


Figura 14: Mejores candidatos por selectividad encontrados en *Anopheles darlingi* (TBP). En la figura se muestran las poses de los mejores candidatos encontrados, y en rojo se señalan los aminoácidos que fueron considerados flexibles. Se muestran los mapas de interacciones para explicar los resultados de la selectividad. En rojo se muestran los aminoácidos cargados negativamente, en azul se muestran los aminoácidos cargados positivamente, las glicinas en color blanco hueso, en verde los aminoácidos hidrofóbicos, en gris se muestran metales, en azul los aminoácidos polares, el agua se muestra en gris claro, y sus interacciones en color blanco. Las interacciones no covalentes se muestran de la siguiente forma: los enlaces de hidrógenos con líneas color violeta, enlaces con halógenos son representados con líneas color marrón, la coordinación con metales en líneas color gris claro, las interacciones Pi-Pi en líneas color verde, las interacciones Pi-cación con líneas color rojo, los puentes salinos con líneas moteadas. La exposición al solvente se representa con nubes de color gris.

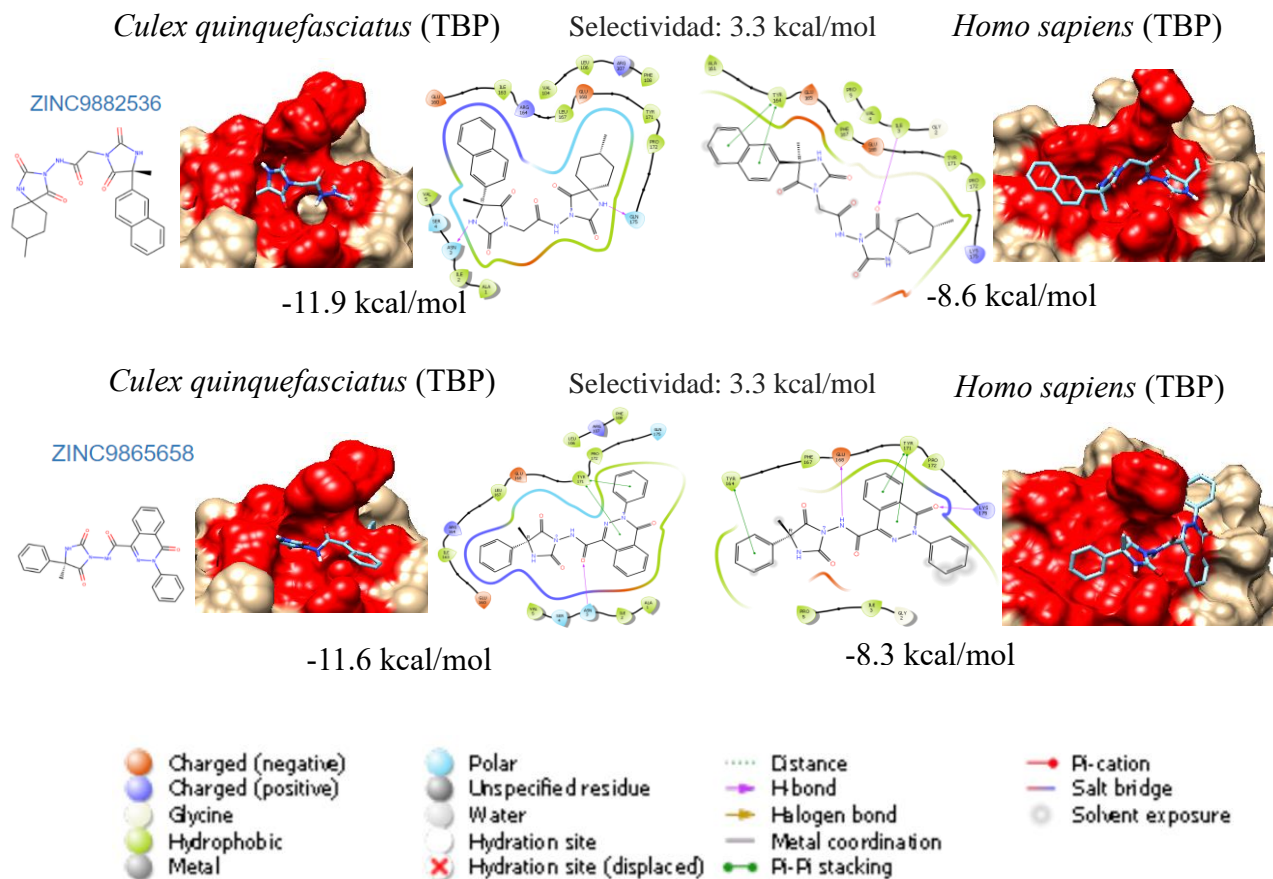
 Aminoácidos asignados como flexibles

Figura 15: Mejores candidatos por selectividad encontrados en *Culex quinquefasciatus* (TBP). En la figura se muestran las poses de los mejores candidatos encontrados, y en rojo se señalan los aminoácidos que fueron considerados flexibles. Se muestran los mapas de interacciones para explicar los resultados de la selectividad. En rojo se muestran los aminoácidos cargados negativamente, en azul se muestran los aminoácidos cargados positivamente, las glicinas en color blanco hueso, en verde los aminoácidos hidrofóbicos, en gris los metales, en azul los aminoácidos polares, el agua se muestra en gris claro, y sus interacciones en color blanco. Las interacciones no covalentes se muestran de la siguiente forma: los enlaces de hidrógenos con líneas color violeta, enlaces con halógenos son representados con líneas color marrón, la coordinación con metales en líneas color gris claro, las interacciones Pi-Pi en líneas color verde, las interacciones Pi-cación con líneas color rojo, los puentes salinos con líneas moteadas. La exposición al solvente se representa con nubes de color gris.

En la figura 13 se muestran resultados de los mejores candidatos obtenidos mediante acoplamiento molecular en la TRF1 de *Aedes aegypti*. En este caso la selectividad viene dada por las interacciones hidrofóbicas. Los ligandos quedaron más cubiertos en la cavidad analizada en *Aedes* que en la misma cavidad en la TBP de humanos, donde los ligandos quedaron con mayor superficie expuesta al solvente, y, por lo tanto, disminuyó la energía de interacción. Esto se mantiene para el resto de los resultados obtenidos: los compuestos evaluados, en todos los casos, quedaron más cubiertos en las cavidades de los factores de transcripción de mosquitos que en las mismas cavidades en la TBP de humanos.

En la figura 14 se muestran resultados de los mejores candidatos obtenidos mediante acoplamiento molecular en la TBP de *Anopheles darlingi*. En este caso se observa la presencia de un puente de hidrógeno entre la arginina 107 y uno de los grupos carboxilos del ligando. Esta arginina está presente en los factores de transcripción de los tres géneros de mosquitos en la posición 107 y se encuentra mutada en la TBP de humanos por una lisina, el cual, aunque también es un aminoácido cargado positivamente, presenta menos donadores de puente de hidrógeno que la arginina. Este puente de hidrógeno de la arginina se mantuvo con ambos ligandos (son isómeros), y en el caso del ligando ZINC9266825 se añade otro puente de hidrógeno del aspartato 105. Este aspartato se mantiene en esta posición también en la TBP de humanos, sin embargo, en esta TBP no realizó ninguna interacción significativa con el ligando.

En la figura 15, donde se muestran resultados de los mejores candidatos obtenidos mediante acoplamiento molecular en la TBP *Culex quinquefasciatus*, se puede observar que la asparagina 3 forma un puente de hidrógeno con uno de los grupos carboxilos del compuesto acoplado. Esta asparagina está presente en los factores de transcripción de las tres especies de mosquitos analizados, sin embargo, en la TBP de humanos se encuentra mutada a una isoleucina. De la misma manera ocurre con la glutamina 175 que se encuentra mutada en humanos por una lisina. La tirosina 171 forma interacciones pi-pi con dos de los ciclos del compuesto acoplado tanto en la TBP de *Culex* como en la TBP de humanos.

En el caso del compuesto ZINC9865658, de esta misma figura, en la TBP de humanos ocurrió una interacción pi-pi y un puente de hidrógeno más que en la TBP de *Culex*, sin embargo, en la TBP de humanos el compuesto quedó con mayor exposición al solvente que en la cavidad presente en la TBP de *Culex*, donde quedó completamente cubierto, obteniéndose una selectividad de 3.3 kcal/mol.

4.7 Análisis de toxicidad de los compuestos propuestos

Utilizando el predictor de toxicidad admetSAR (122), se encontró que los compuestos propuestos en esta investigación como potenciales insecticidas novedosos son menos agresivos contra la salud humana y el ambiente que insecticidas ampliamente utilizados en la actualidad, como la permetrina y el carbamil. La figura 15 muestra estos resultados; los parámetros que se señalan en color verde indican que los compuestos analizados no son dañinos o perjudiciales con respecto a ese factor; los señalados en color rosa, indican que el compuesto es perjudicial.

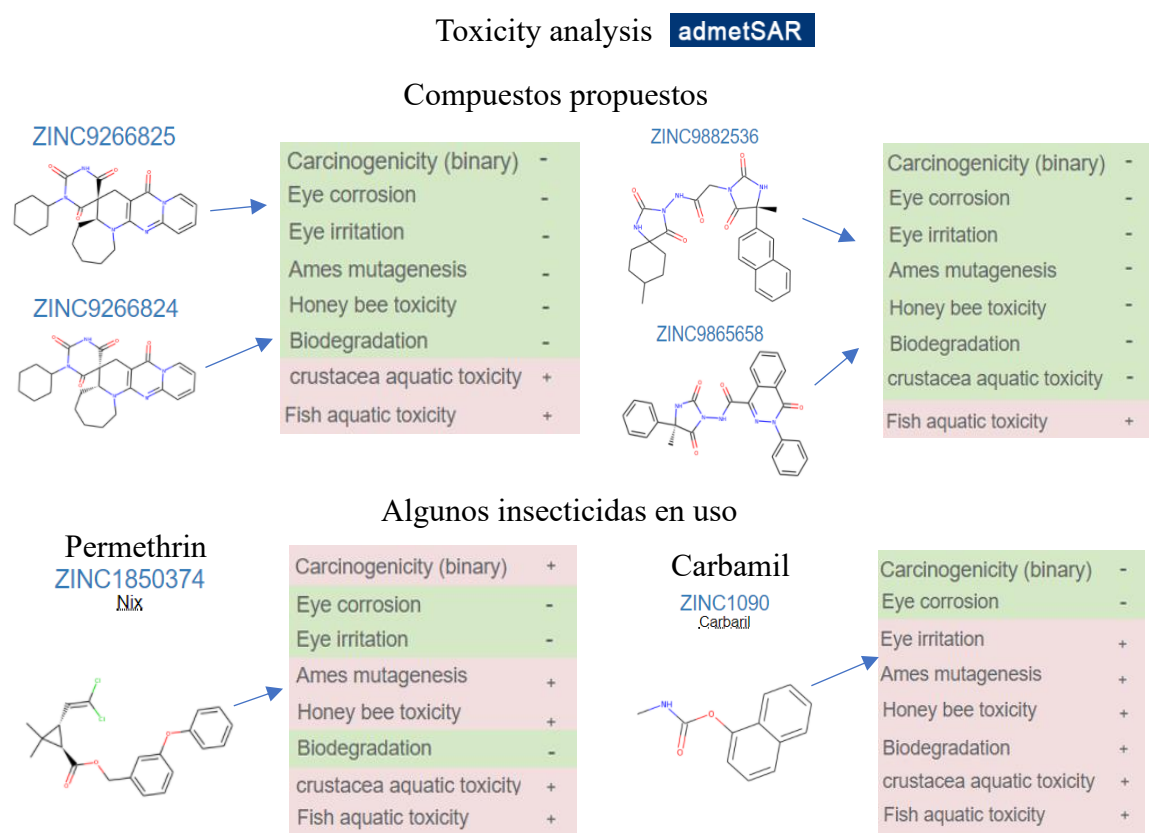


Figura 16: Predicción de toxicidad realizada con el servidor ADMETSAR. Se señala en color verde inocuidad y en color rojo toxicidad.

Los compuestos ZINC9266824 y ZINC9266825, encontrados como mejores candidatos en los géneros *Aedes* y *Anopheles*, se encuentran a la venta comercializados por las compañías Otava, Pharmeks, IBScreen, Vitas-M, Princeton BioMolecular Research, Molport SC Economical, eMolecules, Mcule y Lab Network, con precios que oscilan entre 25 a 35 USD 1 mg ó 40 a 56 UDS 5mg. Los compuestos ZINC9882536 y ZINC9865658, encontrados como mejores candidatos en el género *Culex*, se encuentran comercializados por las compañías Molport, eMolecules y Mcule, con precios que oscilan en 55 y 72 USD 1 mg ó 59 y 77 USD 5mg.

La mejor energía y selectividad se encontró en el género *Anopheles*, sin embargo, los mejores candidatos en este género son los mismos encontrados en el género *Aedes*, por lo tanto, estos compuestos pueden ser propuestos como insecticidas para ambas especies.

5.2 Inhibiendo a ARN Pol III

La cavidad analizada coincide con parte del sitio de unión de Brf1 (124). En la figura 17 se muestra la coincidencia de la cavidad y los compuestos analizados con parte del sitio de unión de Brf1.

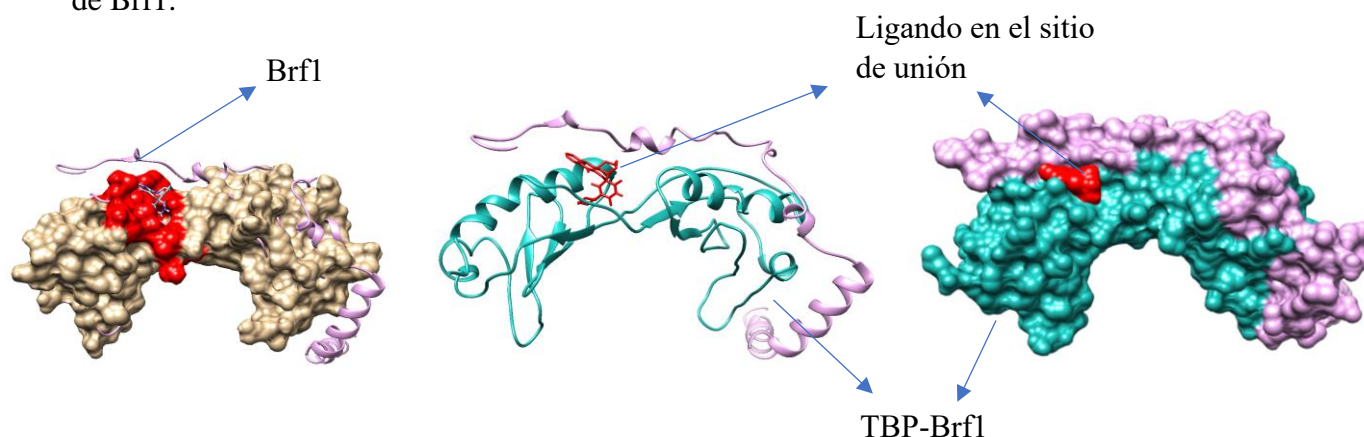


Figura 18: Superposición del sitio de unión de la cavidad analizada en mosquitos vectores junto al ligando acoplado y el complejo de unión de TBP y Brf1. Las TBP se muestran en color marrón y verde-azul; en rojo se muestran los aminoácidos que describen la cavidad analizada, que fueron considerados como flexibles. Brf1 se muestra en color rosa, formando un complejo junto a TBP. El compuesto en el sitio de unión es señalado también en color rojo y representa a todos los compuestos propuestos; se puede observar cómo coinciden con parte del sitio de unión de Brf1.

Brf1, Bdp1 y TBP son componentes esenciales del complejo TFIIB (125). TFIIB es requerido para que ocurra la transcripción mediada por la ARN-Polimerasa III (125). ARN-Pol III transcribe múltiples genes: todos los genes de ARNt conocidos en las células eucariontes; los ARNr 5S, que forman parte de las subunidades mayores de los ribosomas; los ARNsn U6, esenciales para que ocurra corte y empalme (splicing) en los pre-ARNm; ARN P RNasa, que son componentes de la RNasa P, una endorribonucleasa que procesa los extremos 5' de los pre-ARNt; ARN MPR RNasa, que son componentes de la ARN RNasa mitocondrial (MRP), enzima que originalmente fue identificada como una endonucleasa formada por ARN y proteínas que procesaba cebadores (primers) de la replicación del ADN en las mitocondrias, sin embargo la mayoría de las MPR se encuentran en el núcleo (un 99%), donde participan en el procesamiento de los pre-ARNm; ARN 7SL, que constituyen el andamio de señalización del reconocimiento de partículas; ARN Vault, que son componentes de las ribonucleoproteínas Vault (RNPs), enormes partículas citoplasmáticas presentes en muchas células eucariontes desde protozoos hasta humanos, involucradas en procesos de

resistencia a fármacos en tumores humanos y se cree estén también implicadas en ensamblaje y/o transporte macromolecular (126). Teniendo en cuenta lo anterior, se infiere que inhibir a la Pol III en los mosquitos puede ser letal para ellos. El factor de transcripción Brf1 interacciona con los aminoácidos ALA 1, ILE 2, SER3, MET160, ARG161, GLU162, THR164, GLU 165, GLU168, TYR171 de las TBP. Todos estos aminoácidos interaccionan con los compuestos propuestos, por lo tanto, se infiere que, una vez unidos estos compuestos a los factores de transcripción en esa cavidad, evitarían la interacción de estos aminoácidos con el factor de transcripción Brf1, imprescindible para la acción de la PolIII. Teniendo en cuenta lo anterior, se infiere que inhibir a la Pol III en los mosquitos puede ser letal para ellos.

En la figura 19 se muestran los transcritos producto de la acción de la ARN Pol III.

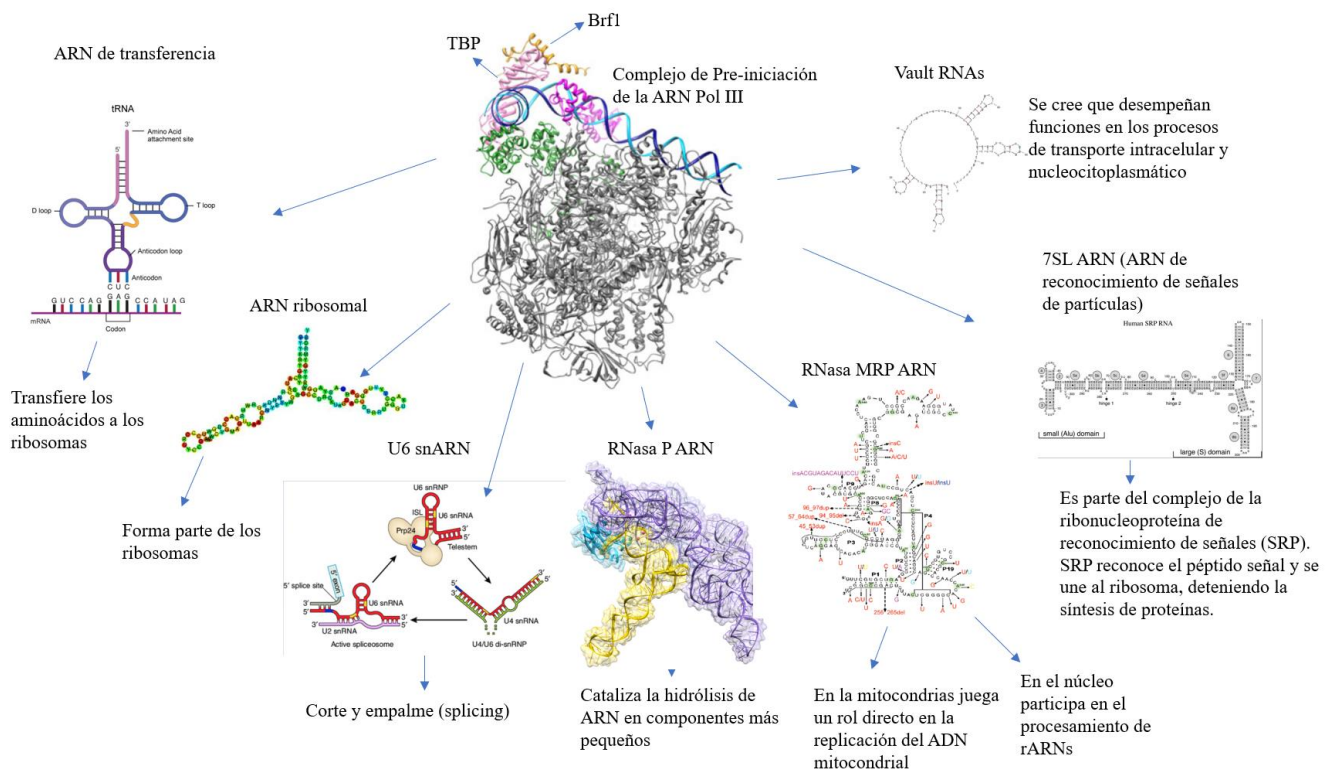


Figura 19: Productos de la transcripción mediada por ARN Pol III.

5.3 Toxicidad de los compuestos propuestos.

Teniendo en cuenta el daño a la salud humana y el ambiente que ocasionan los compuestos insecticidas empleados en la actualidad, se hace necesario proponer compuestos que, además de tener la funcionalidad de eliminar a estos insectos, sean cada vez menos agresivos contra la salud humana y el ambiente. Por esos motivos, además de evaluar con el predictor de toxicidad admetSAR los compuestos propuestos en este trabajo, se realizó también una evaluación a moléculas ampliamente empleadas como insecticidas, para tener un punto de comparación de ambas generados por un mismo servidor.

La figura 15 muestra que los compuestos propuestos en este trabajo, se predice, sean menos agresivos contra la salud humana que compuestos extensamente utilizados como la permetrina y el carbamil. Ninguno de los compuestos propuestos en este proyecto se predice sean mutagénicos, cancerígenos ni irritantes oculares. Esto contrasta con los insecticidas en uso también evaluados en el servidor admetSAR, el cual arrojó que la permetrina se predice sea cancerígena y mutagénica; y el carbamil fue descrito como mutagénico e irritante ocular.

Con respecto al análisis de toxicidad ambiental, los compuestos propuestos fueron referidos por el servidor admetSAR como tóxicos hacia peces y crustáceos, y los insecticidas en uso evaluados fueron descritos como tóxicos hacia peces, crustáceos y abejas. Aunque lo ideal es desarrollar un compuesto con función insecticida totalmente inocuo hacia la salud humana y el ambiente, es notorio advertir que los compuestos que acá se proponen, son, según el predictor de toxicidad empleado, menos agresivos contra la salud humana y el ambiente, que compuestos actualmente en uso.

5.4. Metodología empleada

El objetivo de esta investigación es proponer posibles cavidades como blanco de insecticidas novedosos. Para esto, se realizó un mapeo de identidad de secuencia sobre la superficie y posteriormente se buscaron cavidades que coincidieran en zona de divergencia. Además, dichas cavidades deben coincidir con la zona de unión de algún factor de transcripción esencial para la función de alguna de las polimerasas, de forma tal, que al impedir o inhibir su función sea inhibida alguna de las polimerasas. Posteriormente había que probar que dichas cavidades eran selectivas. Esta es la razón por la cual la metodología empleada se centró en la búsqueda de estas cavidades de manera local y no se realizó acoplamiento molecular ciego a lo largo de la superficie. Consideramos que esto es una perspectiva importante a llevar a cabo posteriormente.

Conclusiones

El presente estudio propone a los factores de transcripción pertenecientes a la familia de las TBP de mosquitos vectores de enfermedades, como un blanco potencial para la unión de insecticidas, atractivo para ser evaluado experimentalmente con los compuestos que en este trabajo se proponen. El análisis de diferencias de superficie llevado a cabo señala que la cavidad aquí propuesta es una zona atractiva de estudio puesto que diverge en secuencia con respecto a la TBP de humanos y coincide con parte de la superficie de unión a Brf1.

La mejor selectividad se encontró en la TBP de la especie *Anopheles darlingui*, y en esta especie se encontraron los mismos compuestos que se unieron a la TRF1 de la especie *Aedes aegypti*, por lo que estos compuestos pueden ser propuestos como insecticidas para ambas especies. Además, esta coincidencia puede proporcionar información acerca del tipo de compuestos que se unen a dicho sitio, puesto que ambos compuestos son isómeros, por tanto, comparten el mismo armazón. Esto abre posibilidades de búsquedas de moléculas con estructuras similares para mejorar la afinidad y selectividad.

Los compuestos sugeridos en este estudio como potenciales insecticidas se predice sean menos agresivos contra la salud humana y el ambiente, esto es otro factor por el cual se hace interesante su estudio. Los compuestos que se proponen en este trabajo pueden ser adquiridos a través de varios proveedores. Por lo tanto, los estudios experimentales son factibles de llevar a cabo.

Perspectivas

1. Evaluar nuevas bibliotecas de ligandos por cribado molecular para detectar otros compuestos que puedan ser propuestos como potenciales insecticidas.
2. Utilizar otros programas de acoplamiento molecular con otras funciones de puntaje para evaluar los resultados obtenidos con Autodock Vina.
3. A medida que nuevas secuencias de factores de transcripción de mosquitos vectores sean depositadas en el NCBI, aumentar el número de estos a analizar.
4. Extender el estudio a factores transcripcionales de otros artrópodos perjudiciales a la salud humana, animal y de importancia agroeconómica.
5. Realizar acoplamiento molecular ciego en los factores de transcripción analizados, puesto que no existe referencia estructural de complejos proteína – ligando en TBP.

Referencias

1. ZHANG Z-Q. Phylum Athropoda. In: Zhang, Z.-Q. (Ed.) Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness (Addenda 2013). Zootaxa [Internet]. 2013 Aug 30;3703(1):17. Available from: doi: 10.11646/zootaxa.3703.1.6
2. David T. Dennis, Nekomoto TS, John C. Victor, Paul WS, Piesman J. Reported Distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the United States. J Med Entomol [Internet]. 1998;35(5):629–38. Available from: doi:10.1093/jmedent/35.5.629
3. Spielman A, Clifford C, Piesman J, Corwin M. Human Babesiosis on Nantucket Island , USA : Description of the Vector, *Ixodes* (IXODES) DAMANINI, N. Sp. 1979;(3):218–34. Available from: doi:10.1093/jmedent/15.3.218
4. Sprong H, Wielinga PR, Fonville M, Reusken C, Brandenburg AH, Borgsteede F, et al. *Ixodes ricinus* ticks are reservoir hosts for *Rickettsia helvetica* and potentially carry flea-borne *Rickettsia* species. Parasites and Vectors [Internet]. 2009;2(1):1–7. Available from: doi:10.1186/1756-3305-2-41
5. Richter PJ, Kimsey RB, Madigan JE, Barlough JE, Dumler JS, Brooks DL. *Ixodes pacificus* (Acari: Ixodidae) as a Vector of *Ehrlichia equi* (Rickettsiales: Ehrlichieae). ELSEVIER [Internet]. 1996;33(1):1–5. Available from: doi:10.1093/jmedent/33.1.1
6. Ogden NH, Lindsay LR, Hanincová K, Barker IK, Bigras-Poulin M, Charron DF, et al. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. Appl Environ Microbiol [Internet]. 2008;74(6):1780–90. Available from: doi: 10.1128/AEM.01982-07
7. Hogenhout SA, Ammar E-D, Whitfield AE, Redinbaugh MG. Insect Vector Interactions with Persistently Transmitted Viruses. Annu Rev Phytopathol [Internet]. 2008;46(1):327–59. Available from: doi:10.1146/annurev.phyto.022508.092135
8. Zurek L, Schal C. Evaluation of the German cockroach (*Blattella germanica*) as a vector for verotoxigenic *Escherichia coli* F18 in confined swine production. Vet Microbiol [Internet]. 2004;101(4):263–7. Available from: doi:10.1016/j.vetmic.2004.04.011
9. Fotedar R, Banerjee U, Shriniwas. Vector potential of the German cockroach in dissemination of *Pseudomonas aeruginosa*. J Hosp Infect [Internet]. 1993;23(1):55–9. Available from: doi:10.1016/0195-6701(93)90131-I
10. Katholi CR, Toé L, Merriweather A, Unnasch TR. Determining the prevalence of *onchocerca volvulus* infection in vector populations by polymerase chain reaction screening of pools of black flies. J Infect Dis [Internet]. 1995;172(5):1414–7. Available from: doi: 10.1093/infdis/172.5.1414

11. Poinar Jr. GO, Mondet B, Gouteux JP, Laveissiere C. *Hexameris glossinae* sp.nov. (Nematoda: Mermithidae), a parasite of tse-tse flies in West Africa. *Can J Zool* [Internet]. 1981;59(5):858–61. Available from: doi: 10.1139/z81-122
12. Shaw WR, Attardo GM, Aksoy S, Catteruccia F. A comparative analysis of reproductive biology of insect vectors of human disease. *Curr Opin Insect Sci* [Internet]. 2015;10:142–8. Available from: doi:10.1016/j.cois.2015.05.001
13. Noireau F, Diosque P, Jansen AM. *Trypanosoma cruzi*: Adaptation to its vectors and its hosts. *Vet Res* [Internet]. 2009;40(2). Available from: doi:10.1051/vetres/2009009
14. Kollien AH, Schaub GA. The development of *Trypanosoma cruzi* in triatominae. *Parasitol Today* [Internet]. 2000;16(9):381–7. Available from: doi:10.1016/S0169-4758(00)01724-5
15. Lee BY, Bacon KM, Connor DL, Willig AM, Bailey RR. The potential economic value of a *Trypanosoma cruzi* (chagas disease) vaccine in Latin America. *PLoS Negl Trop Dis* [Internet]. 2010;4(12):1–8. Available from: doi:10.1371/journal.pntd.0000916
16. Brady OJ, Gething PW, Bhatt S, Messina JP, Brownstein JS, Hoen AG, et al. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. *PLoS Negl Trop Dis* [Internet]. 2012;6(8). Available from: doi: 10.1371/journal.pntd.0001760
17. Weaver SC, Reisen WK. Present and future arboviral threats. *Antiviral Res* [Internet]. 2010;85(2):328–45. Available from: doi: 10.1016/j.antiviral.2009.10.008
18. Yen C, Hyde TB, Costa AJ, Fernandez K, Tam JS, Hugonnet S, et al. The development of global vaccine stockpiles. *Lancet Infect Dis* [Internet]. 2015;15(3):340–7. Available from: doi: 10.1016/S1473-3099(14)70999-5
19. Hiwat H, Bretas G. Ecology of *Anopheles darlingi* Root with respect to vector importance: A review. *Parasites and Vectors* [Internet]. 2011;4(1):1–13. Available from: doi: 10.1186/1756-3305-4-177
20. Jung J, Jung Y, Min GS, Kim W. Analysis of the population genetic structure of the malaria vector *Anopheles sinensis* in South Korea based on mitochondrial sequences. *Am J Trop Med Hyg* [Internet]. 2007;77(2):310–5. Available from: doi:10.4269/ajtmh.2007.77.310
21. Ma Y, Yang M, Fan Y, Wu J, Ma Y, Xu J. Population structure of the malaria vector *anopheles sinensis* (Diptera: Culicidae) in China: Two gene pools inferred by microsatellites. *PLoS One* [Internet]. 2011;6(7). Available from: doi: 10.1371/journal.pone.0022219
22. Zhong D, Chang X, Zhou G, He Z, Fu F, Yan Z, et al. Relationship between Knockdown Resistance, Metabolic Detoxification and Organismal Resistance to Pyrethroids in *Anopheles sinensis*. *PLoS One* [Internet]. 2013;8(2). Available from: doi: 10.1371/journal.pone.0055475
23. Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B, et al.

- Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science* (80-) [Internet]. 2010;330(6000):86–8. Available from: doi: 10.1126/science.1191864
24. Rafikov M, Bevilacqua L, Wyse APP. Optimal control strategy of malaria vector using genetically modified mosquitoes. *J Theor Biol* [Internet]. 2009;258(3):418–25. Available from: doi: 10.1016/j.jtbi.2008.08.006
 25. Atyame CM, Cattel J, Lebon C, Flores O, Dehecq JS, Weill M, et al. Wolbachia-based population control strategy targeting *Culex quinquefasciatus* mosquitoes proves efficient under semi-field conditions. *PLoS One* [Internet]. 2015;10(3):1–15. Available from: doi: 10.1371/journal.pone.0119288
 26. Achee NL, Grieco JP, Vatandoost H, Seixas G, Pinto J, Ching-Ng L, et al. Alternative strategies for mosquito-borne arbovirus control. *PLoS Negl Trop Dis*. 2019;13(1):1–22.
 27. Achee NL, Gould F, Perkins TA, Reiner RC, Morrison AC, Ritchie SA, et al. A Critical Assessment of Vector Control for Dengue Prevention. *PLoS Negl Trop Dis* [Internet]. 2015;9(5):1–20. Available from: doi: 10.1371/journal.pntd.0003655
 28. Corbel V, Achee NL, Chandre F, Coulibaly MB, Dusfour I, Fonseca DM, et al. Tracking Insecticide Resistance in Mosquito Vectors of Arboviruses: The Worldwide Insecticide resistance Network (WIN). *PLoS Negl Trop Dis* [Internet]. 2016;10(12):10–3. Available from: doi: 10.1371/journal.pntd.0005054
 29. Corbel V, N’Guessan R, Brengues C, Chandre F, Djogbenou L, Martin T, et al. Multiple insecticide resistance mechanisms in *Anopheles gambiae* and *Culex quinquefasciatus* from Benin, West Africa. *Acta Trop* [Internet]. 2007;101(3):207–16. Available from: doi: 10.1016/j.actatropica.2007.01.005
 30. Raymond M, Berticat C, Weill M, Pasteur N, Chevillon C. Insecticide resistance in the mosquito *Culex pipiens*: What have we learned about adaptation? *Genetica* [Internet]. 2001;112–113:287–96. Available from: doi: 10.1023/A:1013300108134
 31. Kamgang B, Marcombe S, Chandre F, Nchoutpouen E, Nwane P, Etang J, et al. Insecticide susceptibility of *Aedes aegypti* and *Aedes albopictus* in Central Africa. *Parasites and Vectors* [Internet]. 2011;4(1):1–8. Available from: doi: 10.1186/1756-3305-4-79
 32. Hemingway J, Ranson H. Insecticide Resistance In Insect Vectors of Human Disease. 2000;371–91. Available from: doi:10.1146/annurev.ento.45.1.371
 33. Oppenoorth FJ. Biochemistry of insecticide resistance. *Pestic Biochem Physiol* [Internet]. 1984;22(2):187–93. Available from: doi: 10.1016/0048-3575(84)90088-9
 34. Wang X, Anadón A, Wu Q, Qiao F, Ares I, Martínez-Larrañaga M-R, et al. Mechanism of Neonicotinoid Toxicity: Impact on Oxidative Stress and Metabolism. *Annu Rev Pharmacol Toxicol* [Internet]. 2018;58(1):471–507. Available from: doi: 10.1146/annurev-pharmtox-010617-052429
 35. Rawlins SC. Spatial distribution of insecticide resistance in caribbean populations of

- Aedes aegypti* and its significance. *Rev Panam Salud Publica/Pan Am J Public Heal* [Internet]. 1998;4(4):243–51. Available from: doi: 10.1590/s1020-49891998001000004
36. Montella IR, Martins AJ, Viana-Medeiros PF, Lima JBP, Braga IA, Valle D. Insecticide resistance mechanisms of Brazilian *Aedes aegypti* populations from 2001 to 2004. *Am J Trop Med Hyg* [Internet]. 2007;77(3):467–77. Available from: doi: 10.4269/ajtmh.2007.77.467
 37. Marcombe S, Darriet F, Tolosa M, Agnew P, Duchon S, Etienne M, et al. Pyrethroid resistance reduces the efficacy of space sprays for Dengue control on the island of Martinique (Caribbean). *PLoS Negl Trop Dis* [Internet]. 2011;5(6):1–9. Available from: doi: 10.1371/journal.pntd.0001202
 38. Dusfour I, Thalmensy V, Gaborit P, Issaly J, Carinci R, Girod R. Multiple insecticide resistance in *aedes aegypti* (Diptera: Culicidae) populations compromises the effectiveness of dengue vector control in French Guiana. *Mem Inst Oswaldo Cruz* [Internet]. 2011;106(3):346–52. Available from: doi: 10.1590/S0074-02762011000300015
 39. Baldacchino F, Caputo B, Chandre F, Drago A, della Torre A, Montarsi F, et al. Control methods against invasive *Aedes* mosquitoes in Europe: A review. *Pest Manag Sci* [Internet]. 2015;71(11):1471–85. Available from: doi: 10.1002/ps.4044
 40. Plernsub S, Saingamsook J, Yanola J, Lumjuan N, Tippawangkosol P, Walton C, et al. Temporal frequency of knockdown resistance mutations, F1534C and V1016G, in *Aedes aegypti* in Chiang Mai city, Thailand and the impact of the mutations on the efficiency of thermal fogging spray with pyrethroids. *Acta Trop* [Internet]. 2016;162:125–32. Available from: doi:10.1016/j.actatropica.2016.06.019
 41. Costa LG. The neurotoxicity of organochlorine and pyrethroid pesticides [Internet]. 1st ed. Vol. 131, *Handbook of Clinical Neurology*. Elsevier B.V.; 2015. 135–148 p. Available from: doi:10.1016/B978-0-444-62627-1.00009-3
 42. Clere N, Lauret E, Malthiery Y, Andriantsitohaina R, Faure S. Estrogen receptor alpha as a key target of organochlorines to promote angiogenesis. *Angiogenesis* [Internet]. 2012;15(4):745–60. Available from: doi: 10.1007/s10456-012-9288-7
 43. Hernández AF, Parrón T, Tsatsakis AM, Requena M, Alarcón R, López-Guarnido O. Toxic effects of pesticide mixtures at a molecular level: Their relevance to human health. *Toxicology* [Internet]. 2013;307:136–45. Available from: doi: 10.1016/j.tox.2012.06.009
 44. Bajgar J, Fusek J, Kuca K, Bartosova L, Jun D. Treatment of Organophosphate Intoxication Using Cholinesterase Reactivators: Facts and Fiction. *Mini-Reviews Med Chem* [Internet]. 2007;7(5):461–6. Available from: doi: 10.2174/138955707780619581
 45. Nouira S, Abroug F, Elatrous S, Boujdaria R, Bouchoucha S. Prognostic value of serum cholinesterase in organophosphate poisoning. *Chest* [Internet]. 1994;106(6):1811–4. Available from: doi:10.1378/chest.106.6.1811

46. Kolbezen MJ, Metcalf RL, Fukuto TR. Insecticide Structure and Activity: Insecticidal Activity of Carbamate Cholinesterase Inhibitors. *J Agric Food Chem* [Internet]. 1954;2(17):864–70. Available from: doi: 10.1021/jf60037a003
47. Nunes GS, Skládal P, Yamanaka H, Barceló D. Determination of carbamate residues in crop samples by cholinesterase- based biosensors and chromatographic techniques. *Anal Chim Acta* [Internet]. 1998;362(1):59–68. Available from: doi: 10.1016/S0003-2670(97)00547-3
48. López-Durán RM, Valencia-Quintana R, Sánchez-Alarcón J, Pérez-Aguilar B, Salinas-Arreortua N, Serrano H, et al. La estructura y función de las colinesterasas: Blanco de los plaguicidas. *Rev Int Contam Ambient* [Internet]. 2018;34:69–80. Available from: doi: 10.20937/RICA.2018.34.ESP02.06
49. Vale C, Fonfría E, Bujons J, Messeguer A, Rodríguez-Farré E, Suñol C. The organochlorine pesticides γ -hexachlorocyclohexane (Lindane), α -endosulfan and dieldrin differentially interact with GABAA and glycine-gated chloride channels in primary cultures of cerebellar granule cells. *Neuroscience* [Internet]. 2003;117(2):397–403. Available from: doi: 10.1016/S0306-4522(02)00875-8
50. Cramer P. Eukaryotic Transcription Turns 50. *Cell* [Internet]. 2019;179(4):808–12. Available from: doi:10.1016/j.cell.2019.09.018
51. de Mendoza A, Sebé-Pedrós A. Origin and evolution of eukaryotic transcription factors. *Curr Opin Genet Dev* [Internet]. 2019;58–59:25–32. Available from: doi: 10.1016/j.gde.2019.07.010
52. Nikolov DB, Chen H, Halay ED, Hoffmann A, Roeder RG, Burley SK. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A* [Internet]. 1996;93(10):4862–7. Available from: doi: 10.1073/pnas.93.10.4862
53. Akhtar W, Veenstra GJC. TBP-related factors: A paradigm of diversity in transcription initiation. *Cell Biosci* [Internet]. 2011;1(1):23. Available from: doi: 10.1186/2045-3701-1-23
54. Nisius B, Sha F, Gohlke H. Structure-based computational analysis of protein binding sites for function and druggability prediction. *J Biotechnol* [Internet]. 2012;159(3):123–34. Available from: doi:10.1016/j.jbiotec.2011.12.005
55. Adachi N, Senda M, Natsume R, Senda T, Horikoshi M. Crystal structure of *Methanococcus jannaschii* TATA box-binding protein. *Genes to Cells* [Internet]. 2008;13(11):1127–40. Available from: doi: 10.1111/j.1365-2443.2008.01233.x
56. Adachi N, Natsume R, Senda M, Muto S, Senda T, Horikoshi M. Purification, crystallization and preliminary X-ray diffraction analysis of *Methanococcus jannaschii* TATA box-binding protein (TBP). *Acta Crystallogr Sect D Biol Crystallogr* [Internet]. 2004;60(12 II):2328–31. Available from: doi: 10.1107/S0907444904023960
57. Sainsbury S, Bernecky C, Cramer P. Structural basis of transcription initiation by

- RNA polymerase II. *Nat Rev Mol Cell Biol* [Internet]. 2015;16(3):129–43. Available from: doi:10.1038/nrm3952
58. Pugh BF. Control of gene expression through regulation of the TATA-binding protein. *Gene* [Internet]. 2000;255(1):1–14. Available from: doi: 10.1016/S0378-1119(00)00288-2
 59. Anandapadamanaban M, Andresen C, Helander S, Ohyama Y, Siponen MI, Lundström P, et al. High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nat Struct Mol Biol* [Internet]. 2013;20(8):1008–14. Available from: doi:10.1038/nsmb.2611
 60. Papai G, Frechard A, Kolesnikova O, Crucifix C, Schultz P, Ben-Shem A. Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature* [Internet]. 2020;577(7792):711–6. Available from: doi:10.1038/s41586-020-1944-2
 61. Santiago Á, Razo-Hernández RS, Pastor N. The TATA-binding Protein DNA-binding domain of eukaryotic parasites is a potentially druggable target. *Chem Biol Drug Des* [Internet]. 2020;95(1):130–49. Available from: doi: 10.1111/cbdd.13630
 62. Hoshiyama D, Kuma KI, Miyata T. Extremely reduced evolutionary rate of TATA-box binding protein in higher vertebrates and its evolutionary implications. *Gene* [Internet]. 2001;280(1–2):169–73. Available from: doi: 10.1016/S0378-1119(01)00766-1
 63. Kopytova D V., Krasnov AN. The family of TRF (TBP-like factors) proteins. *Russ J Genet* [Internet]. 2007;43(3):241–5. Available from: doi:10.1134/S1022795407030039
 64. Crowley TE, Hoey T, Liu JK, Jan YN, Jan LY, Tjian R. A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila*. *Nature* [Internet]. 1993;361(6412):557–61. Available from: doi: 10.1038/361557a0
 65. Hansen SK, Takada S, Jacobson RH, Lis JT, Tjian R. Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* [Internet]. 1997;91(1):71–83. Available from: doi: 10.1016/S0092-8674(01)80010-6
 66. Davidson I. The genetics of TBP and TBP-related factors. *Trends Biochem Sci* [Internet]. 2003;28(7):391–8. Available from: doi: 10.1016/S0968-0004(03)00117-8
 67. Lengauer T, Rarey M. Computational methods for biomolecular docking. *Curr Opin Struct Biol* [Internet]. 1996;6(3):402–6. Available from: doi: 10.1016/S0959-440X(96)80061-3
 68. Pérot S, Sperandio O, Miteva MA, Camproux AC, Villoutreix BO. Druggable pockets and binding site centric chemical space: A paradigm shift in drug discovery. *Drug Discov Today* [Internet]. 2010;15(15–16):656–67. Available from: doi: 10.1016/j.drudis.2010.05.015
 69. Hajduk PJ, Huth JR, Tse C. Predicting protein druggability. *Drug Discov Today* [Internet]. 2005;10(23–24):1675–82. Available from: doi: 10.1016/S1359-6446(05)03624-X

70. Shoichet BK. Virtual screening of chemical libraries. *Nature* [Internet]. 2004;432(7019):862–5. Available from: doi: 10.1038/nature03197
71. Garret M, Lim-Wilby M. Molecular Docking. *Encycl Ref Genomics Proteomics Mol Med* [Internet]. 2008;443:1149–53. Available from: doi: 10.1007/3-540-29623-9_3820
72. Trott O, Olson A. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem* [Internet]. 2009;32:174–82. Available from: doi: 10.1002/jcc.21334
73. Morris G, Huey R, Lindstrom W. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem* [Internet]. 2009;32:174–82. Available from: doi: 10.1002/jcc.21256
74. Zhao H, Caflisch A. Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics. *Bioorganic Med Chem Lett* [Internet]. 2013;23(20):5721–6. Available from: doi: 10.1016/j.bmcl.2013.08.009
75. Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N, Garmendia-Doval AB, Juhos S, Schmidtke P, et al. rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Comput Biol* [Internet]. 2014;10(4):1–8. Available from: doi: 10.1371/journal.pcbi.1003571
76. Allen WJ, Balias TE, Mukherjee S, Brozell SR, Moustakas DT, Lang PT, et al. DOCK 6: Impact of new features and current docking performance. *J Comput Chem* [Internet]. 2015;36(15):1132–56. Available from: doi: 10.1002/jcc.23905
77. Venkatachalam CM, Jiang X, Oldfield T, Waldman M. LigandFit: A novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graph Model* [Internet]. 2003;21(4):289–307. Available from: doi: 10.1016/S1093-3263(02)00164-X
78. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem* [Internet]. 2004;47(7):1739–49. Available from: doi: 10.1021/jm0306430
79. Jones G, Willett P, Glen RC, Leach AR, Taylor. Development and Validation of a Genetic Algorithm for Flexible Docking. *J Mol Biol* [Internet]. 1997;267(3):727–48. Available from: doi: 10.1006/jmbi.1996.0897
80. Corbeil CR, Williams CI, Labute P. Variability in docking success rates due to dataset preparation. *J Comput Aided Mol Des* [Internet]. 2012;26(6):775–86. Available from: doi: 10.1007/s10822-012-9570-1
81. Jain AN. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem* [Internet]. 2003;46(4):499–511. Available from: doi: 10.1021/jm020406h
82. Wang Z, Sun H, Yao X, Li D, Xu L, Li Y, et al. 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* [Internet]. 2016;18(18):12964–75. Available from: doi: 10.1039/c6cp01555g
83. Cole JC, Murray CW, Nissink JWM, Taylor RD, Taylor R. Comparing protein-ligand docking programs is difficult. *Proteins Struct Funct Genet* [Internet]. 2005;60(3):325–32. Available from: doi: 10.1002/prot.20497
 84. Chang MW, Ayeni C, Breuer S, Torbett BE. Virtual screening for HIV protease inhibitors: A comparison of AutoDock 4 and Vina. *PLoS One* [Internet]. 2010;5(8):1–9. Available from: doi: 10.1371/journal.pone.0011955
 85. Sandeep G, Nagasree KP, Hanisha M, Kumar MMK. AUDocker LE: A GUI for virtual screening with AUTODOCK Vina. *BMC Res Notes* [Internet]. 2011;4(Figure 3):3–6. Available from: doi: 10.1186/1756-0500-4-445
 86. Benson ML, Smith RD, Khazanov NA, Dimcheff B, Beaver J, Dresslar P, et al. Binding MOAD, a high-quality protein-ligand database. *Nucleic Acids Res* [Internet]. 2008;36(SUPPL. 1):674–8. Available from: doi: 10.1093/nar/gkm911
 87. Wang R, Fang X, Lu Y, Yang CY, Wang S. The PDBbind database: Methodologies and updates. *J Med Chem* [Internet]. 2005;48(12):4111–9. Available from: doi: 10.1021/jm048957q
 88. Roche O, Kiyama R, Brooks CL. Ligand-protein database: Linking protein-ligand complex structures to binding data. *J Med Chem* [Internet]. 2001;44(22):3592–8. Available from: doi: 10.1021/jm000467k
 89. Hendlich M, Bergner A, Günther J, Klebe G. Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. *J Mol Biol* [Internet]. 2003;326(2):607–20. Available from: doi: 10.1016/S0022-2836(02)01408-0
 90. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res* [Internet]. 2007;35(SUPPL. 1):198–201. Available from: doi: 10.1093/nar/gkl999
 91. Chalk AJ, Worth CL, Overington JP, Chan AWE. PDBLIG: Classification of small molecular protein binding in the protein data bank. *J Med Chem* [Internet]. 2004;47(15):3807–16. Available from: doi: 10.1021/jm040804f
 92. Golovin A, Dimitropoulos D, Oldfield T, Rachedi A, Henrick K. MSDsite: A database search and retrieval system for the analysis and viewing of bound ligands and active sites. *Proteins Struct Funct Genet* [Internet]. 2005;58(1):190–9. Available from: doi: 10.1002/prot.20288
 93. Kinoshita K, Furui J, Nakamura H. Identification of protein functions from a molecular surface database, eF-site. *J Struct Funct Genomics* [Internet]. 2002;2(1):9–22. Available from: doi: 10.1023/A:1011318527094
 94. Shin JM, Cho DH. PDB-Ligand: A ligand database based on PDB for the automated

- and customized classification of ligand-binding structures. *Nucleic Acids Res* [Internet]. 2005;33(DATABASE ISS.):238–41. Available from: doi: 10.1093/nar/gki059
95. Michalsky E, Dunkel M, Goede A, Preissner R. SuperLigands - A database of ligand structures derived from the Protein Data Bank. *BMC Bioinformatics* [Internet]. 2005;6:1–6. Available from: doi: 10.1186/1471-2105-6-122
 96. Puvanendrapillai D, Mitchell JBO. L/D Protein Ligand Database (PLD): Additional understanding of the nature and specificity of protein-ligand complexes. *Bioinformatics* [Internet]. 2003;19(14):1856–7. Available from: doi: 10.1093/bioinformatics/btg243
 97. Yamaguchi A, Iida K, Matsui N, Tomoda S, Yura K, Go M. Het-PDB Navi.: A Database for Protein-Small Molecule Interactions. *J Biochem* [Internet]. 2004;135(1):79–84. Available from: doi: 10.1093/jb/mvh009
 98. Kellenberger E, Muller P, Schalon C, Bret G, Foata N, Rognan D. sc-PDB: An annotated database of druggable binding sites from the Protein Data Bank. *J Chem Inf Model* [Internet]. 2006;46(2):717–27. Available from: doi: 10.1021/ci050372x
 99. Ivanisenko VA, Pintus SS, Grigorovich DA, Kolchanov NA. PDBSite: A database of the 3D structure of protein functional sites. *Nucleic Acids Res* [Internet]. 2005;33(DATABASE ISS.):183–7. Available from: doi: 10.1093/nar/gki105
 100. Block P. AffinDB: a freely accessible database of affinities for protein-ligand complexes from the PDB. *Nucleic Acids Res* [Internet]. 2006;34(90001):D522–6. Available from: doi: 10.1093/nar/gkj039
 101. Goto S. LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res* [Internet]. 2002;30(1):402–4. Available from: doi: 10.1093/nar/30.1.402
 102. Zhang J, Aizawa M, Amari S, Iwasawa Y, Nakano T, Nakata K. Development of KiBank, a database supporting structure-based drug design. *Comput Biol Chem* [Internet]. 2004;28(5–6):401–7. Available from: doi: 10.1016/j.compbiolchem.2004.09.003
 103. Sterling T, Irwin JJ. ZINC 15 - Ligand Discovery for Everyone. *J Chem Inf Model* [Internet]. 2015;55(11):2324–37. Available from: doi: 10.1021/acs.jcim.5b00559
 104. Whitfield JR, Beaulieu ME, Soucek L. Strategies to inhibit Myc and their clinical applicability. *Front Cell Dev Biol* [Internet]. 2017;5(FEB):1–13. Available from: doi: 10.3389/fcell.2017.00010
 105. Dang C V., Reddy EP, Shokat KM, Soucek L. Drugging the “undruggable” cancer targets. *Nat Rev Cancer* [Internet]. 2017;17(8):502–8. Available from: doi: 10.1038/nrc.2017.36
 106. McCormick F. KRAS as a therapeutic target. *Clin Cancer Res* [Internet]. 2015;21(8):1797–801. Available from: doi: 10.1158/1078-0432.CCR-14-2662

107. Sievers F, Higgins DG. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. *Methods Mol Biol* [Internet]. 2013;1079:263–71. Available from: doi: 10.1007/978-1-62703-646-7
108. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: Protein structure and function prediction. *Nat Methods* [Internet]. 2014;12(1):7–8. Available from: doi: 10.1038/nmeth.3213
109. Fiser A, Šali A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. *Methods Enzymol* [Internet]. 2003;374:461–91. Available from: doi: 10.1016/S0076-6879(03)74020-8
110. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* [Internet]. 2015;10(6):845–58. Available from: doi: 10.1038/nprot.2015-053
111. Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *ELSEVIER* [Internet]. 2017;122(October 1995):694–701. Available from: doi: 10.1016/j.carbon.2017.07.012
112. Sunhwan J, Taehoon K, Vindy A. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM SUNHWAN. *J Comput Chem* [Internet]. 2007;32:174–82. Available from: doi: 10.1002/jcc.20945
113. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 Dihedral Angles. *J Chem Theory Comput* [Internet]. 2012;8(9):3257–73. Available from: doi: 10.1021/ct300400x
114. Brooks, B R, Brooks CL, Mackerell J, Nilson L, Petrella R. CHARMM: The Biomolecular Simulation Program. *J Comput Chem* [Internet]. 2012;32:174–82. Available from: doi: 10.1002/jcc.21287
115. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem* [Internet]. 2005;26(16):1781–802. Available from: doi: 10.1002/jcc.20289
116. Mezei M. Simulaid: A Simulation Facilitator and Analysis Program. *J Comput Chem* [Internet]. 2010;32:174–82. Available from: doi: 10.1002/jcc.21551
117. Huang B. Metapocket: A meta approach to improve protein ligand binding site prediction. *Omi A J Integr Biol* [Internet]. 2009;13(4):325–30. Available from: doi: 10.1089/omi.2009.0045
118. Durrant JD, De Oliveira CAF, McCammon JA. POVME: An algorithm for measuring binding-pocket volumes. *J Mol Graph Model* [Internet]. 2011;29(5):773–6. Available from: doi: 10.1016/j.jmgm.2010.10.007
119. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, et al. The molecular architecture of the nuclear pore complex. 2007;450(November).
120. Mälkiä A, Murtomäki L, Urtti A, Kontturi K. Drug permeation in biomembranes: In

- vitro and in silico prediction and influence of physicochemical properties. *Eur J Pharm Sci*. 2004;23(1):13–47.
121. Yuan S, Chan HCS, Hu Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip Rev Comput Mol Sci* [Internet]. 2017;7(2):1–10. Available from: doi: 10.1002/wcms.1298
 122. Yang H, Lou C, Sun L, Li J, Cai Y, Wang Z, et al. AdmetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* [Internet]. 2019;35(6):1067–9. Available from: doi: 10.1093/bioinformatics/bty707
 123. Mortelmans K, Zeiger E. The Ames Salmonella/microsome mutagenicity assay Kristien. *EISEVIER* [Internet]. 2000;455(1–2):29–60. Available from: doi 10.1016/S0027-5107(00)00064-6
 124. Nadph TEA, Juo ZS, Kassavetis GA, Wang J, Geiduschek EP. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* [Internet]. 2003;422(April):1–6. Available from: doi: 10.1038/nature01534
 125. Ishiguro A, Kassavetis GA, Geiduschek EP. Essential Roles of Bdp1, a Subunit of RNA Polymerase III Initiation Factor TFIIB, in Transcription and tRNA Processing. *Mol Cell Biol* [Internet]. 2002;22(10):3264–75. Available from: doi 10.1128/mcb.22.10.3264-3275.2002
 126. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. The expanding RNA polymerase III transcriptome. *Trends Genet* [Internet]. 2007;23(12):614–22. Available from: doi: 10.1016/j.tig.2007.09.001