



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

BÚSQUEDA *IN SILICO* DE MOLÉCULAS ORGÁNICAS PEQUEÑAS  
MODULADORAS DE LA ACTIVIDAD DE LA GTPASA EFL1

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
MIRYAM SAMANTHA MALDONADO LÓPEZ

TUTOR PRINCIPAL  
Dra. Nuria Victoria Sánchez Puig  
[Instituto de Química, UNAM](#)

MIEMBROS DEL COMITÉ TUTOR  
Dra. Karina Martínez Mayorga  
[Instituto de Química, UNAM](#)

Dr. Rogelio Rodríguez Sotres  
[Facultad de Química, UNAM](#)

Ciudad Universitaria, Cd.Mx., Febrero 2020



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## **Agradecimientos**

Al financiamiento otorgado en el proyecto CONACyT 283909 “*Escrutinio de moléculas orgánicas pequeñas contra la GTPasa EFL1 como blanco en el Síndrome Shwachman-Diamond*” para la realización de este trabajo.

A la DGTIC por el tiempo de cómputo otorgado por en la plataforma de supercómputo “Miztli” a través del proyecto LANCAD-UNAM-DGTIC-215 (2018).

A la Dra. Nuria Sánchez Puig por permitirme formar parte de su grupo de investigación y por su asesoría constante. A los miembros del comité tutor por sus valiosas aportaciones y facilitarme sus recursos.

A mi familia que constantemente me ha impulsado. A Rogelio Pedraza quien siempre ha apoyado este sueño y a todas las maravillosas personas que tuve la fortuna de conocer durante la maestría, gracias amigos.

## Índice

1	Antecedentes.....	2
1.1	Los ribosomas .....	2
1.1.1	Biogénesis ribosomal .....	2
1.1.2	Síndrome de Shwachman-Diamond.....	3
1.1.3	La GTPasa EFL1 (Elongation Factor like-1, EFL1).....	5
1.2	Métodos computacionales .....	11
1.2.1	Cribado virtual (Virtual screening) .....	11
1.2.2	Acoplamiento molecular (Docking).....	12
1.2.3	Dinámica molecular Newtoniana (MD) .....	13
1.2.4	Predicción de las afinidades de unión a ligando: Métodos de energía libre ....	16
2	Hipótesis.....	19
3	Objetivos.....	20
3.1	Objetivo general .....	20
3.2	Objetivos particulares .....	20
4	Metodología.....	21
4.1	Metodología computacional .....	21
4.1.1	Búsqueda por similitud molecular con sordarina .....	21
4.1.2	Acoplamiento molecular .....	21
4.1.3	Simulaciones de dinámica molecular .....	23
4.1.4	Predicción de energía de unión a ligando.....	24
4.2	Metodología experimental.....	24
4.2.1	Expresión y purificación de ScEfl1 y HsEFL1 .....	24
4.2.2	Desplazamiento térmico monitoreado por fluorescencia (Thermofluor) .....	25
5	Resultados y Discusión.....	26
5.1	Cribado virtual basado en la similitud molecular con sordarina .....	26
5.2	Acoplamiento molecular.....	27
5.3	Simulaciones de dinámica molecular y predicción de la energía de unión de ligando .....	45

5.3.1	Complejo EF-2•sordarina .....	45
5.3.2	Complejo HsEFL1-ligando17.....	51
5.3.3	Complejo ScEfl1-ligando 03 .....	56
5.4	Resultados experimentales.....	61
5.4.1	Purificación de proteínas ScEfl1 y HsEFL1 .....	61
5.4.2	Desplazamiento térmico monitoreado por fluorescencia (Thermofluor) .....	61
6	Conclusiones .....	65
7	Perspectivas .....	66
8	Bibliografía .....	67
9	Anexo .....	70
9.1	Principales sitios de unión a ligando predichos por la herramienta SiteMap en HsEFL1 y ScEfl1 en conformación traslocada.....	70
9.2	Clasificación de sitios de unión a ligando.....	72
9.3	Interacciones proteína-ligando obtenidos para el acoplamiento de los ligandos	72
9.4	Puntajes de interacción calculados para el acoplamiento molecular .....	73
9.5	Ligandos evaluados mediante dinámica molecular .....	74
9.6	RMSD del esqueleto proteico y de los átomos pesados de los ligandos en complejo.....	75
9.7	Compuestos con energía de unión favorable con HsEFL1 y ScEfl1 mediante la metodología LIE .....	79
9.8	Monitoreo de la purificación de HsEFL1 y ScEfl1 .....	85

## Abreviaturas

DMSO	Dimetil sulfóxido
EF-2	Del inglés Elongation factor 2
EFL1	Del inglés Elongation factor like-1
eIF6	Del inglés Eukaryotic translation initiation factor 6
GDP	Guanosin difosfato
GppNH <sub>2</sub>	guanosina-5'-β-amino-difosfato, análogo no hidrolizable de guanosin trifosfato
GTP	Guanosin trifosfato
HsEFL1	Elongation factor-like 1 de <i>H. sapiens</i>
LIE	del inglés linear interaction energy
MCS	Máxima subestructura común
MM-GBSA	Del inglés Molecular mechanics energies combined generalized Born and surface area continuum solvation
MOE	Molecular Operating Environment
PMSF	Fluoruro de fenilmetilsulfonilo
RMSD	Del inglés Root mean square deviation
RNA	Ácido ribonucleico
rRNA	RNA ribosomal
ScEfl1	Elongation factor-like 1 de <i>S. cerevisiae</i>
SDBS	Proteína alterada en el Síndrome Shwachman-Bodian-Diamond
Sdo1	Ortólogo de SBDS en <i>S. cerevisiae</i>
SDS	Síndrome de Shwachman-Diamond / Síndrome Shwachman-Bodian-Diamond
SDS-PAGE	Electroforesis desnaturizante de proteínas en gel de poliacrilamida

SD-URA	Synthetic Drop-out sin uracilo
snoRNPs	Ribonucleoproteínas pequeñas nucleolares
Tif6	Factor de antiasociación de <i>S. cerevisiae</i>
T <sub>m</sub>	Temperatura de fusión

## Resumen

La biogénesis ribosomal es un proceso complejo en eucariontes que resulta en la maduración de las subunidades ribosomales 60S y 40S para producir ribosomas capaces de llevar a cabo la traducción. Uno de los últimos pasos conocidos en la maduración de la subunidad 60S es la remoción del factor de anti-asociación Tif6/eIF6, el cual evita el acoplamiento prematuro entre las subunidades ribosomales. Factor de Elongación de Tipo 1 (EFL1) es una GTPasa que en conjunto con SBDS, su intercambiador de nucleótidos de guanina, remueve al factor Tif6/eIF6 de la superficie de la subunidad pre-60S. Los defectos en la función de estas proteínas derivan en una ribosomopatía llamada síndrome de Shwachman-Diamond (SDS), un trastorno autosómico recesivo caracterizado por insuficiencia pancreática exocrina, disfunción de la médula ósea, anomalías esqueléticas y un riesgo incrementado de padecer leucemia. Aproximadamente el 90% de los pacientes con SDS tienen mutaciones bialélicas en el gen *SBDS*, y en menor medida mutaciones en *EFL1*. Mutaciones en ambas proteínas previenen la liberación de eIF6 y su posterior reciclaje al núcleo. Su liberación depende de un cambio conformacional en EFL1 similar al que sufre el factor de elongación 2 (EF-2) durante la translocación en la síntesis de proteínas. La toxina sordarina es capaz de estabilizar el estado translocado en EF-2 mediante la unión en una región interdominio. Por lo que se prevé que existan moléculas pequeñas que se unan al sitio equivalente en EFL1 y sean capaces de estabilizar la conformación translocada requerida para liberar a eIF6. Con la finalidad de encontrar este tipo de moléculas en el presente trabajo se realizó un escrutinio virtual de moléculas orgánicas pequeñas de bibliotecas de compuestos disponibles comercialmente con potencial modulador de la GTPasa EFL1. Para ello se realizó primeramente un cribado por similitud con sordarina, seguido de estudios de acoplamiento molecular o *docking* con los compuestos seleccionados y simulación de dinámica molecular. Se seleccionaron compuestos para los que se predijo una energía de interacción favorable utilizando ambos métodos, y que fueron capaces de estabilizar la conformación translocada según estudios de *simulated annealing*. Los resultados obtenidos sugieren que existen ligandos con puntajes favorables de energía de unión capaces de estabilizar la conformación translocada de EFL1, identificado así, moléculas candidatas que podrían interactuar con la GTPasa EFL1 tanto de levadura como con su ortólogo de humano. Se realizaron ensayos de desplazamiento de la temperatura de desnaturalización de EFL1 por fluorescencia como prueba de concepto para establecer un método de escrutinio de los ligandos capaces de unirse a EFL1. Estos ensayos mostraron un incremento en la temperatura de desnaturalización de EFL1 como resultado de la unión a GDP, por lo que esta técnica se podrá usar para evaluar la interacción de los compuestos seleccionados.



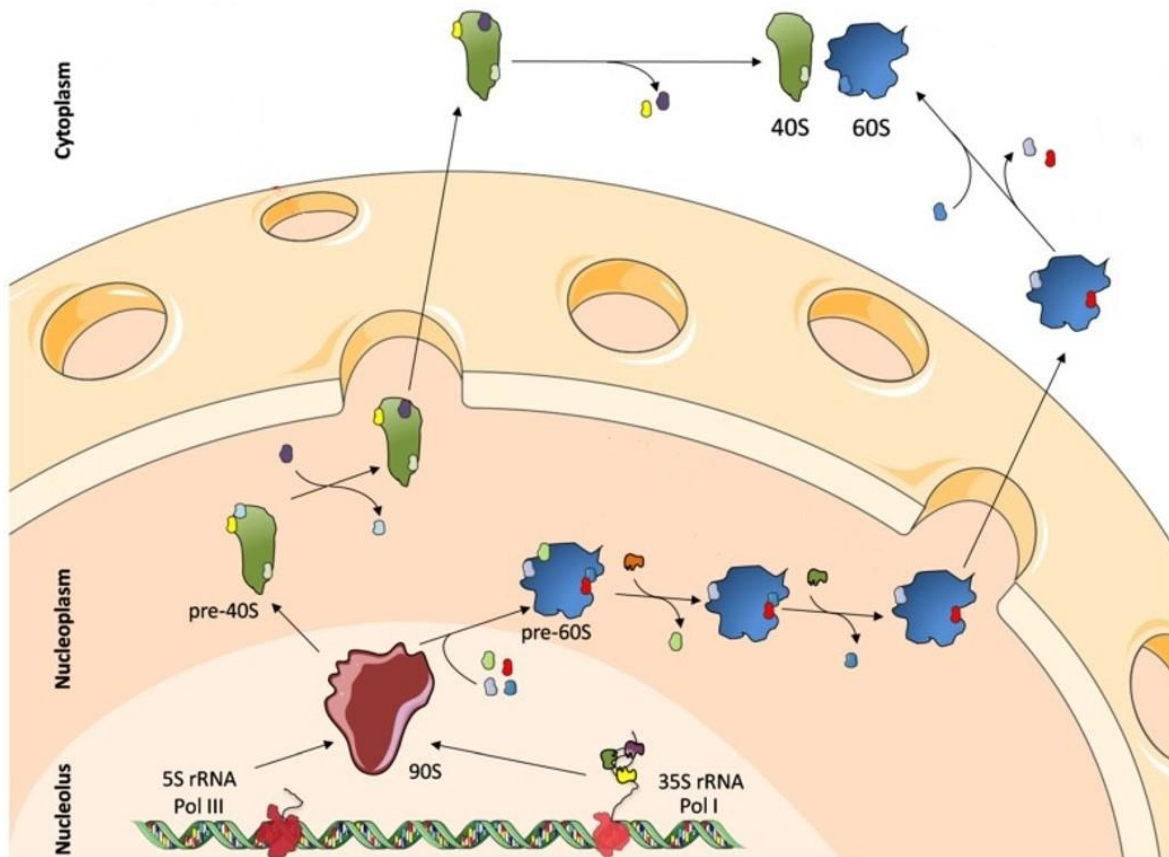
# 1 Antecedentes

## 1.1 Los ribosomas

Los ribosomas son una maquinaria molecular muy dinámica de la célula y son esenciales para la vida ya que producen todas las proteínas requeridas para el crecimiento celular. Su biogénesis requiere el adecuado plegamiento de sus RNAs y la incorporación de numerosas proteínas estructurales. Debido a su función biológica, el correcto ensamblaje de este complejo ribonucleoproteico es crítico para la adecuada expresión génica. Los ribosomas en eucariontes están compuestos por una subunidad mayor 60S y una subunidad menor 40S. La subunidad 60S está constituida por tres rRNA (5.8S, 5S y 25S) y 46 proteínas, mientras que la subunidad 40S contiene un rRNA (18S) y 33 proteínas <sup>1</sup>.

### 1.1.1 *Biogénesis ribosomal*

La biogénesis ribosomal es el proceso, que en eucariontes comienza en el nucléolo y termina en el citoplasma, a través del cual se sintetizan los ribosomas (Figura 1). En este complejo proceso participan las tres RNA polimerasas, más de 200 proteínas accesorias y cerca de 75 ribonucleoproteínas pequeñas nucleolares (snoRNPs). Se inicia con la transcripción de los pre-rRNA 5S y 35S en el nucléolo, los cuales sufren modificaciones post-transcripcionales y se asocian con proteínas ribosomales y factores de transactivación para formar la partícula 90S. Cuando esta partícula es escindida libera las partículas pre-40S y pre-60S que siguen vías de maduración independientes. La subunidad pre-40S se exporta rápidamente al citoplasma tras unas cuantas modificaciones, mientras que la maduración de la pre-60S es mucho más compleja pues sufre varios cambios en su composición en el nucleoplasma. En el citoplasma los principales eventos de maduración incluyen la liberación de factores del túnel de salida, el ensamblaje del tallo ribosomal y el desacoplamiento de factores de anti-asociación <sup>2-7</sup>. Los defectos en los componentes estructurales del ribosoma o en aquellos involucrados en su biogénesis, derivan en el desarrollo de patologías conocidas como ribosomopatías, enfermedades que, si bien tienen en común la disfunción ribosomal, difieren significativamente en el mecanismo y los síntomas clínicos que se presentan. El síndrome de Shwachman-Diamond pertenece a este heterogéneo grupo de afecciones <sup>8,9</sup>.



**Figura 1.** Biogénesis ribosomal y procesamiento del pre-rRNA en levadura. Representación esquemática de la maduración del ribosoma que inicia con la transcripción de los pre-rRNAs 35S por la RNA Pol I y del 5S por la RNA Pol III. Estos rRNAs se asocian con proteínas ribosomales y accesorias formando la partícula 90S. Luego de la escisión la partícula 90S se separa en las subunidades pre-60S (azul) y pre-40S (verde), que siguen rutas independientes de maduración y exportación. La maduración de la pre-60S involucra un mayor número de pasos en el nucleoplasma. Finalmente la maduración en el citoplasma resulta en las subunidades ribosomales 60S y 40S las cuales se ensamblan para formar el ribosoma 80S traduccionalmente competente. (Modificado de Olombrada M. *et al.* "Fungal Ribotoxins: A Review of Potential Biotechnological Applications" *Toxins* vol. 9,2 71. 2017)

### 1.1.2 Síndrome de Shwachman-Diamond

El síndrome de Shwachman-Diamond (SDS) fue reportado por primera vez en 1964 por Schwachman *et al.* en un grupo de 5 niños que presentaban neutropenia y disfunción pancreática. Esto pacientes habían estado siendo tratados en una clínica de fibrosis quística en la Universidad de Harvard, sin embargo esta condición ya había descartado en por ausencia de dificultades respiratorias y concentraciones normales de electrolitos en sudor <sup>10</sup>. SDS es una enfermedad autosómica recesiva, caracterizada por insuficiencia pancreática exocrina, hematopoyesis no efectiva y un riesgo incrementado de leucemia.

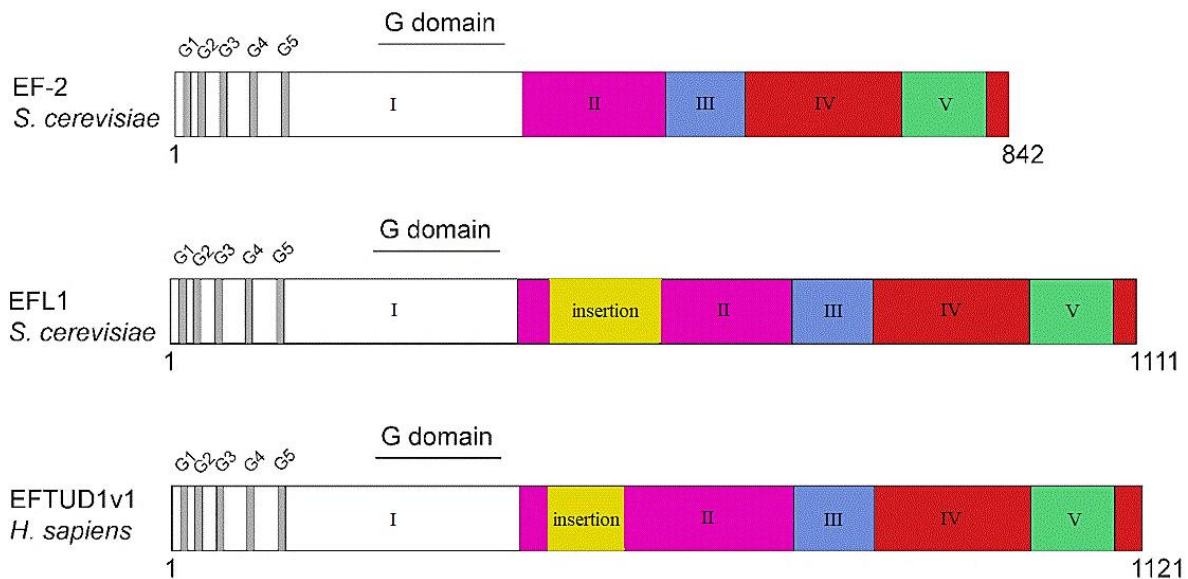
Esta enfermedad junto con la fibrosis quística es la causa más común de insuficiencia pancreática en niños. Los pacientes frecuentemente presentan esteatorrea y escaso crecimiento derivado de la insuficiencia pancreática. Además, presentan infecciones severas como resultado de la neutropenia y otros signos que incluyen anemia, trombocitopenia, baja estatura y anormalidades esqueléticas <sup>11</sup>. Los pacientes con SDS también poseen un mayor riesgo de desarrollar mielodisplasia y transformación maligna. Las medidas de apoyo para los pacientes con SDS incluyen enzimas pancreáticas, antibióticos, transfusiones y factor estimulante de colonia de granulocito. La única terapia definitiva es el trasplante de médula ósea <sup>12</sup>. Aproximadamente el 90% de los pacientes que presentan los criterios clínicos para el diagnóstico de SDS tienen mutaciones en el gen *SBDS*. Se ha estimado que la frecuencia portadora para esta mutación es de alrededor 1 en 60,000 individuos <sup>13</sup>. Este gen está altamente conservado y presenta cinco exones que comprenden 7.9 kb localizados en 7q11 en la región centromérica del cromosoma <sup>13,14</sup>. El gen *SBDS* codifica una proteína altamente conservada de 250 aminoácidos y existe un pseudogen adyacente *SBDSP* con el que comparte 97% de homología pero contiene deleciones y cambios de nucleótidos que evitan la formación de una proteína funcional. Cerca del 75% de los pacientes con SDS presentan mutaciones en *SBDS* que resultan de un evento de conversión génica con el pseudogen <sup>14</sup>. Las mutaciones que más comúnmente se presentan son la 183-184TA→CT, que genera un codón de paro en el aminoácido 62 (K62X), y la 258+2T→C que interrumpe el sitio donador del empalme en el intrón 2. Este cambio en el sitio de empalme produce un corrimiento en el marco de lectura que codifica para una proteína truncada en el aminoácido 84 (C84fsX3) <sup>14</sup>. Estas mutaciones pueden clasificarse con base en el fenotipo de la proteína *SBDS* como: mutaciones que truncan a la proteína, mutaciones que causan inestabilidad en la proteína provocando su desnaturalización y mutaciones que afectan epítopes de la superficie de la proteína. El mRNA y la proteína *SBDS* se expresan ampliamente en los tejidos humanos <sup>15,16</sup>. La completa ausencia de expresión de *SBDS* es letal en modelos murinos <sup>17</sup>. A pesar de que la mutación que origina la proteína trunca *SBDS* K62X es común en los pacientes con SDS, no se han identificado pacientes homocigotos para esta mutación, lo que sugiere que probablemente la pérdida total de la expresión de *SBDS* es letal en humanos.

Los estudios con el ortólogo de *SBDS* en levadura sugieren que *SBDS* puede actuar como factor intercambiador de nucleótidos de guanina para la GTPasa citoplasmática denominada Factor de Elongación de Tipo 1 (EFL1), involucrada también en la biogénesis de la subunidad ribosomal mayor. *SBDS* disminuye a la mitad la constante de Michaelis de

EFL1 por GTP a través de un aumento en la constante de disociación por GDP<sup>18,19</sup>. A su vez EFL1 promueve la disociación del factor de anti-asociación eIF6 (Tif6 ortólogo en levadura) dependiente de GTP de la subunidad ribosomal 60S *in vitro*<sup>20</sup>. Levaduras con mutaciones en Sdo1, el ortólogo de SBDS, son incapaces de reciclar a Tif6 al núcleo y presentan muy lento crecimiento. Estudios genéticos demostraron la supresión del fenotipo de lento crecimiento por mutaciones en Tif6<sup>2</sup>. La proteína Tif6/eIF6 se asocia con la subunidad ribosomal 60S y evita la asociación prematura de la subunidad ribosomal 40S<sup>21,22</sup>. Las mutaciones en Tif6 que suprimen el fenotipo de lento crecimiento están localizadas en una región que reduce la unión de Tif6 con la subunidad 60S; estas mutaciones también suprimen defectos en la biogénesis ribosomal que resultan de mutaciones en EFL1 y en Sdo1<sup>20</sup>. Lo anterior sugiere que la mutación supresora promueve la liberación de Tif6 y su reciclaje al núcleo debilitando la afinidad de la interacción de Tif6 con las subunidades pre-60S. Estos datos sugieren un modelo en que Sdo1 y EFL1 interaccionan con la subunidad pre-60S y facilitan la remoción de Tif6 que permite la asociación de las subunidades ribosomales 60S y 40S<sup>2</sup>. La liberación de Tif6 depende de un cambio conformacional en EFL1 la cual realiza la hidrólisis de GTP y pasa a ocupar un sitio en común con Tif6 en la subunidad ribosomal 60S<sup>23</sup>.

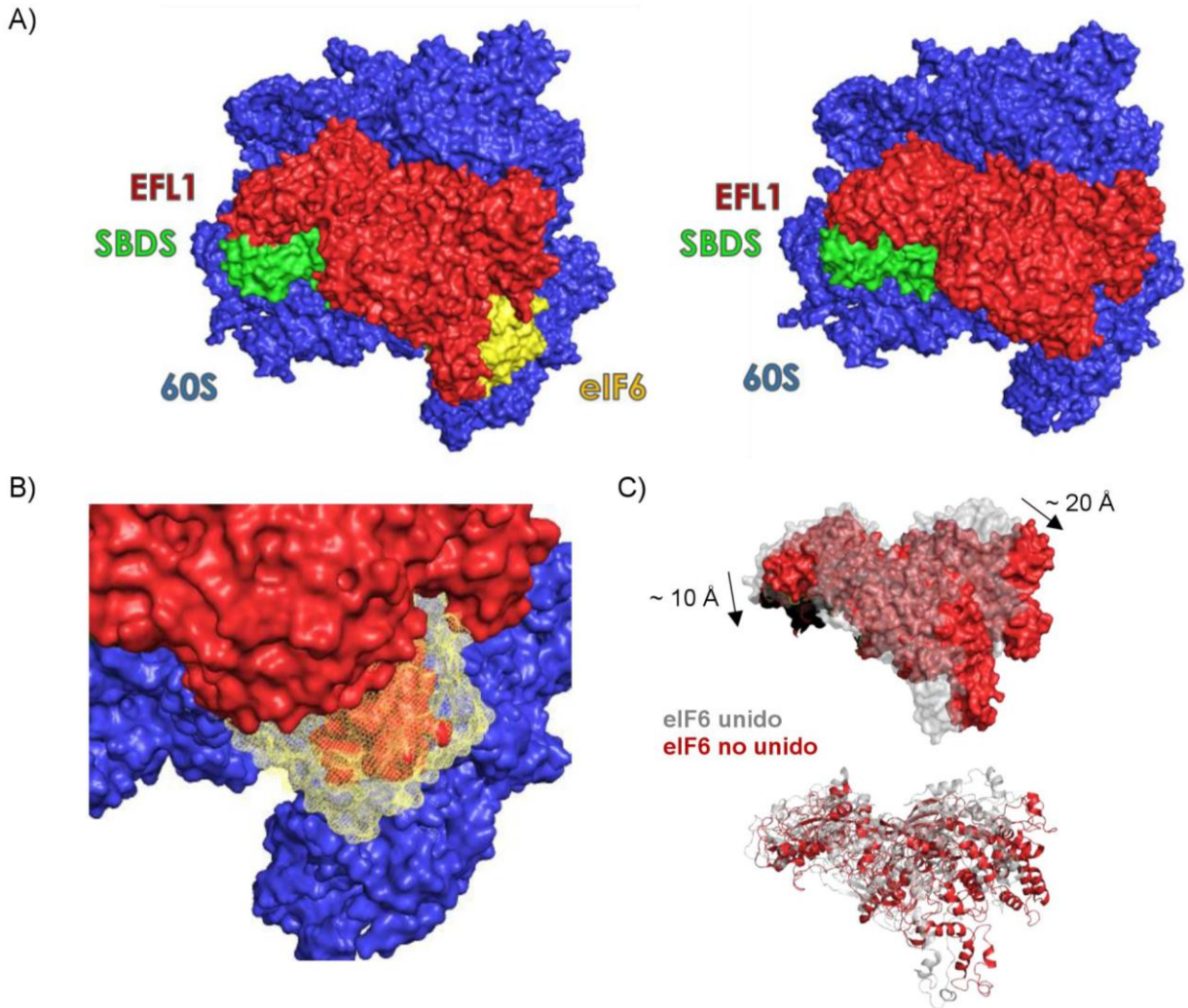
### **1.1.3 La GTPasa EFL1 (Elongation Factor like-1, EFL1)**

La proteína EFL1 es una GTPasa homóloga al factor de elongación 2 (EF-2) con 39% de identidad de secuencia, EF-2 dirige la translocación del tRNA desde el sitio A al sitio P durante la síntesis proteica. De forma análoga, EFL1 remueve a eIF6 de la superficie de la 60S a través de un cambio conformacional similar al que efectúa EF-2 durante la elongación. Modelos estructurales proponen que EFL1 presenta la estructura básica de los factores de elongación EF-G/EF-2, es decir, cinco dominios siendo el dominio de unión a nucleótidos de guanina el dominio I y los restantes enumerados del II al V (Figura 2). El dominio I de EFL1 contiene un subdominio G, conservado en todas las GTPasas, compuesto por cinco motivos (G1-G5) encargados de reconocer e hidrolizar al GTP. Los motivos G5 y G4 reconocen específicamente la base de guanina y el motivo G1 interacciona con los fosfatos  $\alpha$  y  $\beta$  de la molécula de GTP, mientras que los motivos G3 y G2 (switch I y switch II) están implicados en el mecanismo de hidrólisis de GTP<sup>22,23</sup>. Los otros cuatro dominios de EFL1 son muy similares a los descritos para EF-2, a excepción de una inserción desestructurada de tamaño variable (160 residuos en la EFL1 de levadura) en el dominio II a través de la cual interactúa con SBDS<sup>24</sup>.



**Figura 2.** Distribución de los dominios estructurales del factor de elongación 2 y las proteínas EFL1 de humano y levadura.

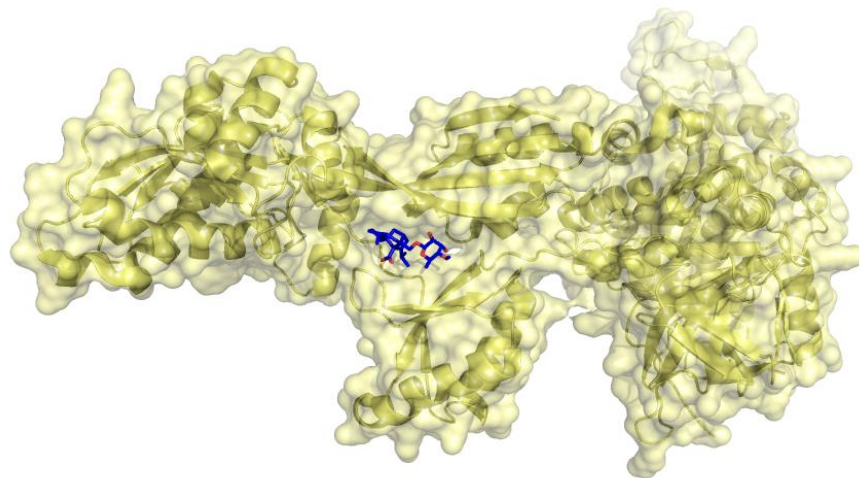
Recientemente, en el grupo de trabajo se describió la presencia de mutaciones en los residuos M882 y R1095 en el dominio IV de EFL1 en pacientes que presentan cuadros clínicos similares a aquellos observados en el Síndrome Shwachman-Diamond <sup>26</sup>. Al igual que lo observado en alteraciones en SBDS, las mutantes de EFL1 no pueden reciclar a Tif6 a núcleo por lo que esta se encuentra relocalizada en el citoplasma <sup>2</sup>. El alineamiento de las secuencias de EFL1 mostró que el residuo R1095 está contenido dentro de una firma consenso que se conserva entre todos los miembros EFL1 sugiriendo una función esencial para él. El residuo equivalente a M882 en el parólogo EF-2 de levadura es I676. Con base en la estructura cristalina de EF-2 de levadura, el residuo I676 contacta directamente con R823 que corresponde a R1086 y R1095 en los ortólogos EFL1 de levadura (ScEfl1) y humano (HsEFL1), respectivamente. Suponiendo que el plegamiento entre las proteínas de la familia EFL1 y EF-2 se conserva, esto sugiere que los residuos mutados en los pacientes interactúan entre sí en la proteína silvestre y que las mutaciones pueden interrumpir esta interacción. Además, se ha demostrado que el C-terminal de Efl1 es esencial para su función de tal manera que deleciones que se extienden hasta el residuo R1086 resultan en mutaciones extragénicas en Tif6 <sup>25</sup>.



**Figura 3.** Un cambio conformacional en EFL1 remueve al factor de anti-asociación eIF6 de la superficie de la pre-60S. A) Mapas de superficie de la subunidad pre-60S unida a EFL1 y SBDS previa a la liberación de eIF6 (izquierda) y después de su liberación (derecha). B) El volumen previamente ocupado por eIF6 resaltado en amarillo en el mapa de superficie ahora es ocupado por EFL1 tras su cambio conformacional. C) Superposición de los modelos atómicos de EFL1 en presencia (gris) o ausencia (rojo) de eIF6 <sup>23</sup>.

EFL1 sufre un cambio conformacional dependiente de la unión e hidrólisis GTP en el ribosoma que preceden la disociación de eIF6 (Figura 3) <sup>2</sup>. La similitud estructural entre EFL1 y EF-2 sugiere que, el cambio conformacional que sufre EFL1 es análogo al que efectúa EF-2 durante la elongación en la síntesis proteica <sup>26</sup>. Se cree que la conformación adoptada por EFL1 requerida para desacoplar a eIF6 es similar al estado traslocado que genera la toxina sordarina en EF-2 (Figura 4). La sordarina es un compuesto sintetizado por *Sordaria arenosa* que inhibe la síntesis de proteínas, puesto que atrapa a EF-2 en una

conformación traslocada. La sordarina (Figura 5) forma puentes de hidrógeno con los residuos Gln490, Ala562 y Phe798 de EF-2. Entre estos residuos únicamente Phe798 presenta alta conservación en la familia de las proteínas EF-2, de igual forma el residuo equivalente en ScEfl1 y HsEFL1 es Phe. Sin embargo, la interacción se da entre el grupo amino del esqueleto de Phe798 y el grupo alcohol de sordarina. El residuo Gln490 interactúa mediante el grupo amida de su cadena lateral con otro de los grupos alcohol de sordarina. Este residuo se encuentra en una posición en la que en el resto de las EF2 existe Arg, posiblemente este residuo da especificidad de unión de ScEfl1 a sordarina. De acuerdo con los alineamientos, en las proteínas ScEfl1 y HsEFL1, al igual que para los miembros de la familia de EF2, en esta posición existe una Arg (Figura 6). Además, existen extensas interacciones de van der Waals entre sordarina y cadenas laterales aromáticas o alifáticas de EF-2 (Figura 7 A). Muchos de los residuos en contacto con sordarina están localizados en conectores interdominio, lo cual explica cómo la unión de sordarina puede inducir un gran cambio conformacional en EF-2 <sup>28</sup>. También se ha sugerido que el cambio conformacional en EFL1 induce una rotación de dominio en SBDS que conduce a un rearrreglo estructural del ribosoma, desestabilizando la interacción de eIF6 con la subunidad pre-60S y provocando la disociación de eIF6 <sup>22</sup>. La imposibilidad de acoplar la hidrólisis de GTP por la EFL1 con la liberación de Tif6 de la subunidad 60S se ha descrito como la principal característica molecular de SDS <sup>29</sup>.



**Figura 4.** Estructura cristalográfica de EF-2 de *Saccharomyces cerevisiae* (amarillo) en la conformación traslocada unida a sordarina (azul) proveniente del PDB 1N0U <sup>27</sup>.

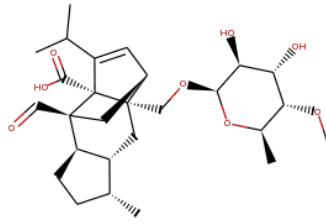


Figura 5. Estructura química de la sordarina.

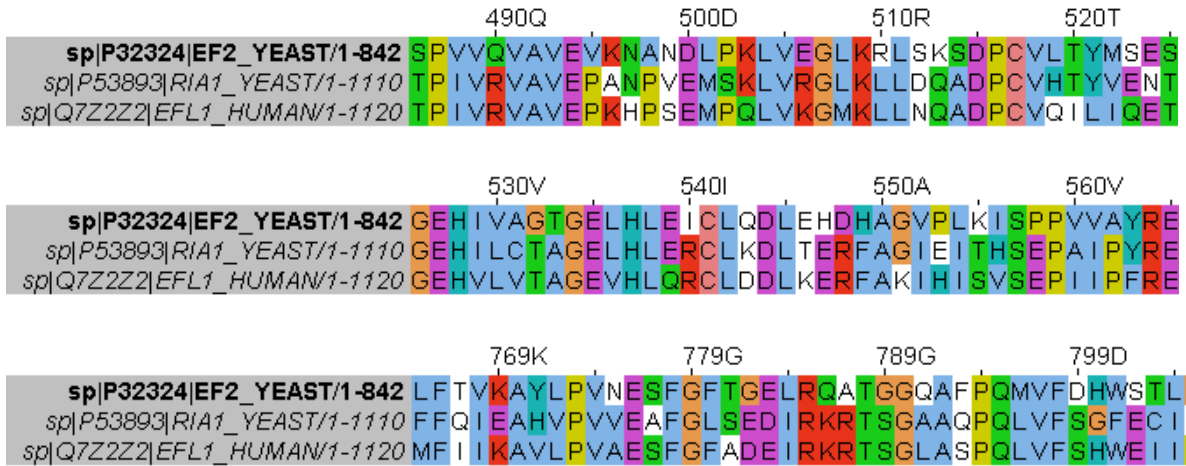
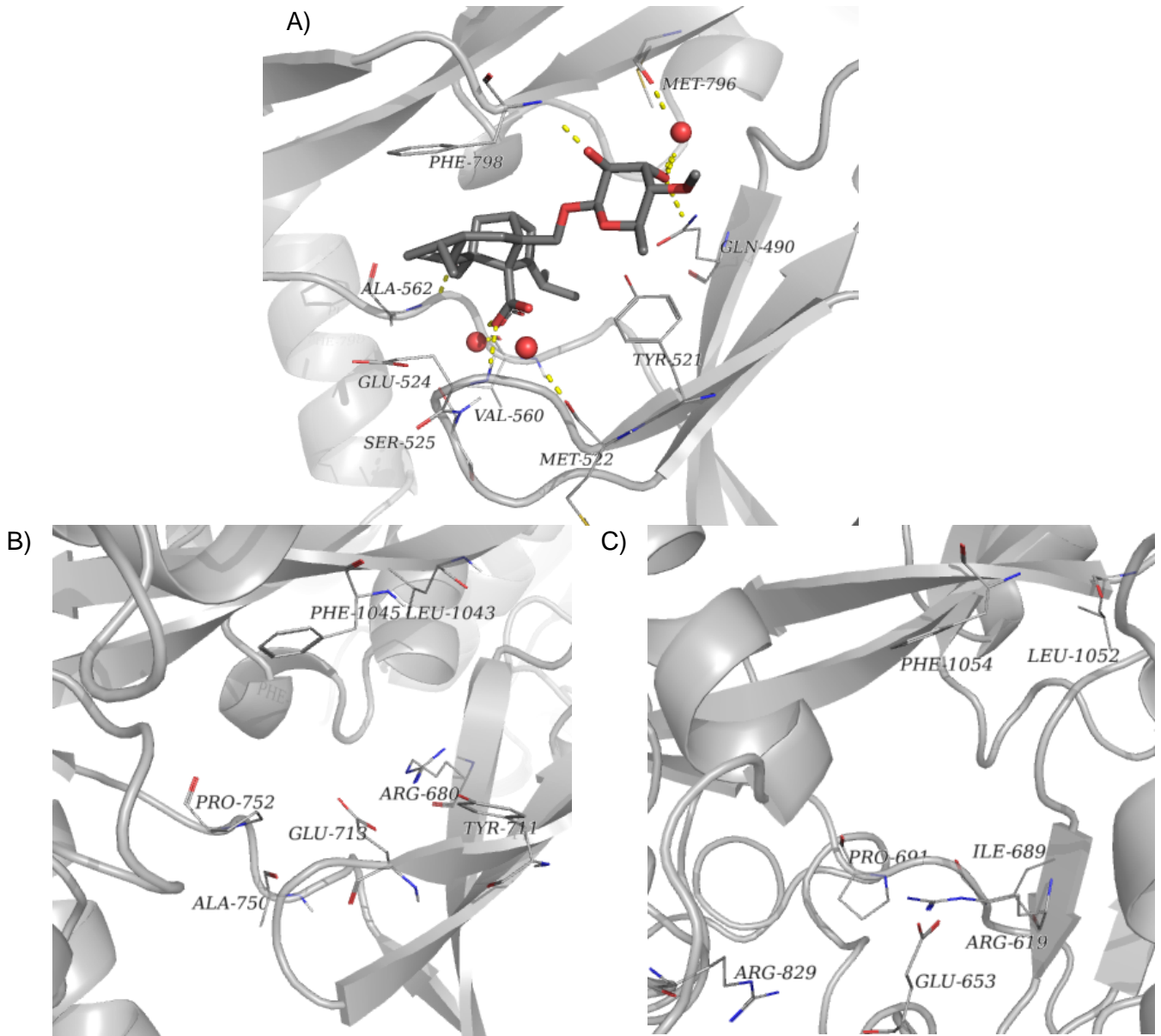


Figura 6. Alineamiento de secuencia primaria del bolsillo de unión de las proteínas EF-2, ScEfl1 y HsEFL1. Se muestran los residuos que forman el sitio de unión en cada una de ellas.





**Figura 7.** Las proteínas EFL1 presentan un bolsillo de unión similar al que existe en EF-2. A) Sitio de unión a sordarina en la región entre los dominios III, IV y V de EF-2. Las interacciones polares se muestran en amarillo, además se muestran las cadenas laterales de los residuos involucrados en la unión con el ligando. B) Región interdominio equivalente al sitio de unión de sordarina de EF-2 en su homólogo ScEfl1 y C) HsEFL1. Se representan los residuos equivalentes a los mostrados en A de acuerdo con los alineamientos.

Ya que la proteína EFL1 es homóloga a EF-2 (39% de identidad de secuencia), consideramos que podrían existir moléculas orgánicas pequeñas que se unan a EFL1 en la región entre los dominios III al V (Figura 7 B y C), como la sordarina a EF-2, y que emulen el cambio conformacional necesario para liberar a Tif6 / eIF6 de la subunidad 60S. Debido

a lo anterior proponemos a EFL1 como blanco terapéutico de SDS, para el cual se pueden encontrar moléculas orgánicas pequeñas que se unan a ella y rescatar su función.

## 1.2 Métodos computacionales

### 1.2.1 Cribado virtual (*Virtual screening*)

El cribado virtual es un filtrado computacional (*in silico*) de moléculas para seleccionar candidatos o *hits* computacionales que después pueden ser evaluados experimentalmente<sup>29</sup>. Existen dos enfoques generalmente aceptados para el cribado virtual: métodos basados en ligando y métodos basados en la estructura del blanco (acoplamiento o “*docking*”). El cribado virtual basado en ligando (LBVS-por sus siglas en inglés) se enfoca en búsquedas por similitud molecular con ligandos conocidos como referencia. Se fundamenta en el principio de que compuestos similares a los ligandos conocidos probablemente muestren actividad biológica similar<sup>30</sup>. Dicha similitud puede basarse en la comparación de descriptores moleculares, esto es el resultado final de un procedimiento lógico y matemático que transforma la información química codificada dentro de una representación simbólica de una molécula en un número útil. Los descriptores, según su enfoque, pueden clasificarse en tres grupos:

- i. Una dimensión (1D). Están basados en propiedades que pueden derivarse de la fórmula estructural de un compuesto.
- ii. Dos dimensiones (2D). Usan descriptores derivados del patrón de unión covalente de una molécula.
- iii. Tres dimensiones (3D). Se basan en coordenadas atómicas y por lo tanto dependen de la conformación. Debido a su naturaleza flexible las moléculas orgánicas pueden adoptar diferentes estados en el espacio 3D, denominados conformeros que comparten la misma topología, pero difieren en su disposición geométrica<sup>31</sup>.

Mientras que el cribado virtual basado en estructura (SBVS-por sus siglas en inglés) aplica diferentes técnicas de modelaje para imitar la interacción de unión de un ligando a un blanco biomolecular. Tradicionalmente el SBVS se conforma de varios pasos incluyendo la preparación del receptor y la biblioteca de compuestos, el acoplamiento de cada compuesto en el sitio de unión en el receptor, el post-procesamiento y la clasificación de los resultados para bioensayos por una función de puntaje<sup>32</sup>. Por lo tanto, la mayor diferencia entre LBVS y SBVS es que el último requiere de información estructural del blanco, si dicha información no existe, es posible obtener esta información con modelos por homología<sup>34</sup>.

### **1.2.2 Acoplamiento molecular (Docking)**

Los estudios de acoplamiento molecular son empleados para predecir modos de unión de ligandos a su receptor usando métodos computacionales, dichos estudios están constituidos por dos componentes clave:

- i. Búsqueda de los posibles modos de unión en el sitio de interés de la macromolécula, usando algoritmos de muestreo.
- ii. Puntaje calculado para cada modo de unión de la molécula, utilizando funciones de puntaje <sup>35</sup>.

Algunos algoritmos de muestreo corresponden a: algoritmos de correspondencia basados en la forma del ligando en un sitio activo en la proteína usando características de su forma e información química <sup>36</sup>, algoritmos de construcción incremental basados en la fragmentación del ligando por ruptura de sus enlaces rotables o algoritmos que acoplan un primer fragmento en el sitio activo de la proteína para luego añadir los fragmentos restantes incrementalmente<sup>36,37</sup>. Los métodos estocásticos incluyen métodos de Monte Carlo, que generan poses del ligando con base en rotaciones de enlaces <sup>38</sup> y los algoritmos genéticos en los que los grados de libertad del ligando se codifican en cadenas binarias llamadas genes <sup>39,40</sup>. La dinámica molecular (MD) se basa en el progreso en el movimiento del sistema en pequeñas etapas, en el contexto del acoplamiento las simulaciones de MD representan tanto la flexibilidad del ligando como de la proteína de forma más efectivamente que otros algoritmos, además son eficaces en optimizaciones locales <sup>42,43</sup>.

Por otro lado, el propósito de las funciones de puntaje es distinguir las poses correctas de las incorrectas en un tiempo computacional razonable, estas se clasifican en funciones basadas en campos de fuerza, empíricas y basadas en conocimiento <sup>43</sup>. Las funciones basadas en campos de fuerzas evalúan la energía de unión calculando la suma de las interacciones no enlazadas (electrostática y van der Waals). En las funciones empíricas la energía de unión se descompone en varios componentes energéticos (como puentes de hidrógeno, interacciones iónicas, efectos hidrofóbicos y entropía de unión) <sup>44-46</sup>, funciones de puntaje basadas en el conocimiento utilizan el análisis estadístico de estructuras cristalinas de complejos ligando-proteína para obtener las frecuencias y/o distancias de contacto interatómicas entre el ligando y la proteína, construidas para favorecer contactos permitidos y penalizar interacciones repulsivas <sup>48,49</sup>.

## Metodologías de acoplamiento

Las metodologías de acoplamiento pueden clasificarse de acuerdo con la flexibilidad de la proteína y el ligando:

- i. Acoplamiento entre ligando y receptor rígidos: En este caso el espacio de búsqueda es muy limitado, pues se consideran sólo tres grados de libertad de traslación y tres de rotación. La flexibilidad del ligando podría ser abordada mediante el uso de un conjunto precalculado de sus conformaciones o permitiendo un grado de superposición átomo-átomo<sup>34</sup>.
- ii. Acoplamiento entre ligando flexible y receptor rígido: Para los sistemas cuyo comportamiento sigue el paradigma de ajuste inducido es muy importante considerar las flexibilidades tanto del ligando como del receptor. Un enfoque que considera un equilibrio entre precisión y tiempo computacional es tratar el ligando como flexible mientras que el receptor se mantiene rígido <sup>35</sup>.
- iii. Acoplamiento entre ligando y receptor flexibles: Existen varios métodos para implementar la flexibilidad del receptor, como puede ser mediante el uso de bibliotecas de rotámeros, permitiendo la flexibilidad de las cadenas laterales del receptor o empleando varias conformaciones de la proteína <sup>34</sup>.

### 1.2.3 *Dinámica molecular Newtoniana (MD)*

Las simulaciones de dinámica molecular monitorean los procesos dependientes del tiempo en sistemas moleculares para estudiar sus propiedades estructurales, dinámicas y termodinámicas mediante la resolución numérica de una ecuación de movimiento, que es la formulación de las reglas que rigen el movimiento ejecutado por la molécula. Es decir, la MD proporciona información sobre la dependencia del tiempo y la magnitud de las fluctuaciones en ambas posiciones y velocidades. El comportamiento temporal de las moléculas se describe mejor por la ecuación de movimiento mecánica cuántica, como lo es la ecuación de Schrödinger dependiente de tiempo. Sin embargo, debido a que esta ecuación es extremadamente difícil de resolver para sistemas grandes, una descripción más simple de mecánica clásica se emplea a menudo para aproximar el movimiento efectuado por los átomos pesados de la molécula.

En la forma más simplificada, la ecuación de movimiento de Newton establece que:

$$F_i = m_i a_i = m_i \ddot{r}_i \quad (1)$$

donde  $F_i$  es la fuerza actuando sobre la partícula  $i$ ,  $m_i$  es la masa de la partícula  $i$ ,  $a_i$  es su aceleración y  $\ddot{r}_i$  es la segunda derivada de la partícula en posición  $r$  con respecto al tiempo. La fuerza  $F_i$  está determinada por el gradiente de la función de la energía potencial,  $U(r)$ , la cual es una función de todas las coordenadas atómicas  $r$ .

$$F_i = -\nabla_i U(r) \quad (2)$$

La ecuación (2) es una ecuación diferencial de segundo orden. Una formulación más general de la ecuación de movimiento de Newton está dada en términos del sistema Hamiltoniano,  $H$  en ecuación (3).

$$H = H(r, p) = K(p) + U(r) = \sum_i \frac{p_i^2}{2m_i} + U(r) \quad (3)$$

Donde  $K(p)$  es la energía cinética y  $p_i$  es el momentum de la partícula  $i$ .

Entonces la ecuación de movimiento de Newton en términos del Hamiltoniano se escribe como un par de ecuaciones diferenciales de primer orden acopladas:

$$\dot{r}_k = \frac{\partial H(r, p)}{\partial p_k}; \quad \dot{p}_k = -\frac{\partial H(r, p)}{\partial r_k} \quad (4)$$

Si se sustituye la definición de  $H$  de (3) en (4) se obtiene nuevamente (1). Estas dos ecuaciones diferenciales de primer orden en (4) son más fáciles de resolver que una ecuación diferencial de segundo orden.

### Algoritmos computacionales

Para resolver la ecuación de movimiento de Newton se requiere de un procedimiento de integración. Un método estándar para resolver ecuaciones diferenciales ordinarias es el enfoque de la diferencia finita. En dicho enfoque se obtienen las coordenadas moleculares y las velocidades a tiempo  $t + \Delta t$  a partir de las coordenadas y velocidades al tiempo previo  $t$ . La elección del intervalo de tiempo  $\Delta t$  depende de las propiedades del sistema simulado, este intervalo debe ser significativamente más pequeño que las características de tiempo del movimiento estudiado.

Para entender el método de diferencia finita, un buen punto de partida es la expansión de Taylor sobre el tiempo  $t$  de la posición al tiempo  $t + \Delta t$ ,

$$r(t + \Delta t) = r(t) + \dot{r}(t)\Delta t + \frac{1}{2}\ddot{r}(t)\Delta t^2 + \dots \quad (5)$$

Que puede ser escrito como:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \dots \quad (6)$$

Donde  $v(t)$  es la velocidad del vector y  $a(t)$  es la aceleración. Es conveniente reescribir la ecuación en una forma discreta y reemplazar  $a$  según la ecuación (1). Usando  $r_n$  para indicar la posición en el paso  $n$  (al tiempo  $t$ ) y  $r_{n+1}$  para indicar la posición en el siguiente paso,  $n + 1$  (al tiempo  $t + \Delta t$ ). Tal que la ecuación (6) se puede expresar como:

$$r_{n+1} = r_n + v_n\Delta t + \frac{1}{2}\left(\frac{F_n}{m}\right)\Delta t^2 + O(\Delta t^3) \quad (7)$$

Donde  $O(\Delta t^n)$  está en términos del orden de  $\Delta t^n$  o menores. Con esta información la velocidad  $v_{n+1}$  al tiempo  $n + 1$  puede estimarse como:

$$v_{n+1} = (r_{n+1} - r_n)/2 \quad (8)$$

Juntas, las ecuaciones (7) y (8) forman el algoritmo de integración. Dada la posición  $r_n$ , la velocidad  $v_n$  y la fuerza  $F_n$  en el paso  $n$ , estas ecuaciones permiten obtener la posición  $r_{n+1}$ , y la velocidad  $v_{n+1}$  en el paso  $n + 1$ .

El algoritmo de integración más usado en el estudio de biomoléculas es el integrador de Verlet. Este integrador está basado en dos expansiones de Taylor, una expansión hacia adelante y otra hacia atrás. La suma de las dos expresiones da origen a la siguiente ecuación:

$$r_{n+1} = 2r_n - r_{n-1} + \frac{F_n}{m}\Delta t^2 + O(\Delta t^4) \quad (9)$$

Además posee una expresión complementaria para propagar las velocidades:

$$v_n = \frac{r_{n+1} - r_{n-1}}{2\Delta t} + O(\Delta t^2) \quad (10)$$

El algoritmo representado en las ecuaciones (9) y (10) es un método numérico estable para la resolución de la ecuación de movimiento de Newton para sistemas desde fluidos simples hasta biopolímeros. Las ventajas de emplear el algoritmo de Verlet son que la posición de integración es bastante acertada, se requiere una sola evaluación de la fuerza por ciclo de integración lo que es menos costoso computacionalmente. Sin embargo la propagación de la velocidad está sujeta a grandes errores, además  $v_n$  sólo puede calcularse si ya se conoce  $r_{n+1}$ . Para mejorar la ecuación de velocidad se han realizado modificaciones al esquema básico de Verlet. Una de estas modificaciones es el algoritmo de integración *leap-*

*frog*, el cual es empleado en las simulaciones de este trabajo, en dicho algoritmo las velocidades son evaluadas en el punto medio de la evaluación de la posición y viceversa. Puede escribirse como:

$$r_{n+1} = r_n + v_{n+1/2}\Delta t \quad (11)$$

$$v_{n+1/2} = v_{n-1/2} + \frac{F_n}{m}\Delta t \quad (12)$$

donde  $v_{n\pm 1/2}$  representa la velocidad en el tiempo de medio paso. El algoritmo de *leap-frog* involucra tres pasos:

1. Emplea la posición actual  $r_n$  para calcular la fuerza actual  $F_n$ .
2. Usa la fuerza actual  $F_n$  y la velocidad anterior de medio paso  $v_{n-1/2}$  para calcular la siguiente velocidad de medio paso  $v_{n+1/2}$ .
3. Finalmente usa la posición actual  $r_n$  y la siguiente velocidad de medio paso  $v_{n+1/2}$  (desde el paso 2) para calcular la posición en el siguiente paso,  $r_{n+1}$ .

La velocidad actual  $v_n$ , que es necesaria para calcular la energía cinética, se puede calcular como:

$$v_n = (v_{n+1/2} + v_{n-1/2})/2 \quad (13)$$

Este algoritmo mejora la ecuación de velocidad, además la evaluación directa de velocidades da un manejo útil para controlar la temperatura de la simulación <sup>50</sup>.

#### **1.2.4 Predicción de las afinidades de unión a ligando: Métodos de energía libre**

Los métodos de energía libre usan muestreo conformacional para generar promedios termodinámicos, en contraste con los métodos de acoplamiento que se enfocan en una conformación unida. El método de predicción de afinidad de unión denominado *linear interaction energy* (LIE) ofrece un buen equilibrio entre velocidad computacional y precisión. Este enfoque estima la energía libre de unión a través de la diferencia en la energía de interacción del ligando con sus alrededores en los estados unido al receptor y libre. Se basa en una representación adecuada de las interacciones ligando-receptor mediante los términos incluidos en un campo de fuerza de mecánica molecular, y considera tanto la solvatación como los efectos entrópicos. Para ello necesita un método de muestreo como la MD para generar conjuntos de configuraciones y obtener promedios termodinámicos a partir de estos <sup>51</sup>.

El método LIE<sup>52</sup> implica ejecutar dos simulaciones de MD con un campo de fuerza empírico: uno para el ligando en solución y el otro para el ligando en el sitio de unión de la proteína. Los cuadros guardados de las simulaciones representan ensamblajes de conformaciones de Boltzmann y se utilizan para calcular las energías de interacción electrostáticas promediadas por Boltzmann ( $U_{elec}$ ) y van der Waals ( $U_{vdw}$ ) del ligando con su entorno en los estados unido y libre. La energía libre de unión se estima entonces como:

$$\Delta G_{unión} = \alpha(\langle U_{vdw} \rangle_{unido} - \langle U_{vdw} \rangle_{libre}) + \beta(\langle U_{elec} \rangle_{unido} - \langle U_{elec} \rangle_{libre}) + \gamma$$

donde los corchetes angulares indican los promedios de Boltzmann. Los cambios en la energía interna del solvente y la proteína se consideran explicados por los factores  $\alpha$  y  $\beta$ , con base en su energía de reorganización en respuesta al ligando, mientras que  $\gamma$  es un término constante que también se puede implementar como dependiente del área de superficie<sup>52</sup>. En la derivación inicial del método LIE, el factor no polar se estimó utilizando un enfoque empírico puro, a través de una calibración en un conjunto de 18 complejos proteína-ligando. El valor obtenido de  $\alpha = 0.18$  ha reproducido con éxito desde entonces las energías libres de unión en una amplia variedad de sistemas de proteína-ligando, incluyendo ligandos neutros pequeños de P450CAM<sup>53</sup>, compuestos cargados tales como bloqueadores de canales de potasio<sup>54,55</sup> e incluso compuestos flexibles grandes tales como los peptidomiméticos en las proteasas aspárticas<sup>56</sup>. Åqvist y Hansson determinaron un primer conjunto de valores refinados para el factor de contribución polar,  $\beta$ , en función de la naturaleza química del ligando con base en los cálculos de perturbación de la energía libre (FEP) realizados para diferentes entidades químicas. Los valores corresponden a las desviaciones de la teoría de respuesta lineal, que están directamente relacionadas con la capacidad del ligando para participar en la red de enlaces de hidrógeno del disolvente acuoso. De acuerdo con lo anterior, los ligandos con carga neta poseen un valor de  $\beta=0.5$ , mientras que los ligandos no cargados tienen un valor asociado de  $\beta=0.43$ ,  $\beta=0.37$  o  $\beta=0.33$  dependiendo de si poseen cero, uno o más de un grupo hidroxilo<sup>56</sup>. Para la mayoría de los compuestos considerados normalmente en un proceso de diseño de fármacos, los factores  $\beta$  estimados son cercanos a los valores "clásicos" y son suficientes para la mayoría de los cálculos de unión a ligandos.

En este trabajo realizamos una búsqueda virtual de moléculas orgánicas pequeñas que se unen a la GTPasa EFL1 y emulen el cambio conformacional necesario para liberar a Tif6.



Usando como blanco de búsqueda un modelo de la proteína EFL1 humana y de levadura generados por homología con la estructura cristalográfica de EF-2 unida a sordarina.

## **2 Hipótesis**

Si el cambio conformacional que sufre EFL1 para liberar a eIF6 es similar al estado traslocado que genera la toxina sordarina al unirse a su homólogo EF-2, entonces una molécula orgánica pequeña podría estabilizar la conformación traslocada de EFL1 al acoplarse al sitio equivalente de unión de la sordarina.

## 3 Objetivos

### 3.1 Objetivo general

Identificar moléculas orgánicas pequeñas con potencial para estabilizar la conformación traslocada de EFL1 a través de su unión al sitio equivalente en que se une la toxina sordarina a su homólogo EF-2.

### 3.2 Objetivos particulares

- Realizar el cribado virtual para obtener moléculas similares a la sordarina.
- Elaborar el acoplamiento molecular de ligandos en la región entre los dominios III-V en la estructura de la EFL1 de *Saccharomyces cerevisiae* y de *Homo sapiens*.
- Realizar la predicción de la energía de unión al ligando empleando dinámica molecular como método de muestreo.
- Analizar la interacción de los ligandos con las proteínas EFL1 mediante dinámica molecular.
- Evaluar la estabilidad conformacional del esqueleto de las proteínas EFL1 durante la trayectoria de dinámica molecular.
- Realizar ensayos de DSF para evaluar la unión de EFL1 con ligandos conocidos de manera cualitativa.

## 4 Metodología

### 4.1 Metodología computacional

#### 4.1.1 Búsqueda por similitud molecular con sordarina

La biblioteca de compuestos para la búsqueda se seleccionó de la base de datos para escrutinio virtual (*virtual screening*) de ZINC (<http://zinc15.docking.org/>) con peso molecular entre 425 y 580 Da, y logP de 1 a 3. Se evaluó la similitud bidimensional de los compuestos en la biblioteca con sordarina (ZINC000044430825) empleando el software PowerMV versión 0.61 (NISS). Dicha búsqueda se realizó con base en el índice de similitud de Tanimoto según los descriptores binarios de pares de átomos, pares de fragmentos, firmas o huellas digitales de farmacóforo y el índice de similitud de Euclidean según el descriptor continuo del número de carga. PowerMV calcula un valor de distancia entre 0.00 y 1.00 de acuerdo con la similitud existente entre cada molécula de la biblioteca de compuestos y la molécula de referencia, siendo 0.00 el valor que representa la mayor similitud, esto es el vecino más cercano. Se solicitaron los 500 vecinos más cercanos para cada descriptor y se obtuvo el valor promedio para el índice de similitud. Se seleccionaron las moléculas con un valor calculado para el descriptor menor al valor promedio, es decir los más cercanos a 0.00, para evaluar su similitud tridimensional con sordarina. Posteriormente se realizó el alineamiento 3D por tipo de átomos y mediante subestructura común (MCS) de cada una de estas moléculas con la sordarina empleando MarvinSketch de ChemAxon, en el cual se asigna un valor entre 0 y 1 al alineamiento, siendo 1 el mejor alineamiento. Finalmente se obtuvo el valor promedio y la desviación estándar (DS) para los dos tipos de alineamiento 3D para conservar únicamente las moléculas con valores superiores al promedio sumando la desviación estándar.

#### 4.1.2 Acoplamiento molecular

Las moléculas seleccionadas tras el alineamiento 3D fueron preparadas con Ligprep para realizar el acoplamiento molecular (Schrödinger, LLC, New York, NY, 2017). Se corrigieron errores estereoquímicos y las conformaciones de anillos, también se generaron los estados de ionización y tautómeros para valores de pH de  $7 \pm 2$  usando Epik3.9 (Schrödinger, LLC, New York, NY, 2017).

La estructura de EF-2 de *Saccharomyces cerevisiae* en el estado traslocado en complejo con sordarina fue obtenida de la estructura cristalográfica (PDB ID: 1N0U) así como en estado no traslocado (PDB ID: 1N0V) presentes en *Protein Data Bank* ([www.rcsb.org/](http://www.rcsb.org/))<sup>51</sup>. La estructura tridimensional de ScEfl1 fue obtenida por Alfonso Méndez

Godoy mediante modelado por homología con MODELLER usando como molde la estructura cristalográfica de EF-2 (39% de identidad de secuencia), mientras que para la estructura tridimensional de HsEFL1 se empleó el modelo atómico generado por homología con EF-2 y EF-G acoplado en los mapas de densidad de estudios de criomicroscopía electrónica cryo-EM, refinado empleando molecular dynamic flexible fitting (MDFF) por Weis, *et al.* (PDB ID: 5ANC). La corrección de errores estereoquímicos y la adición de átomos faltantes en las estructuras proteicas se efectuó empleando el paquete MOE (*Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2018) y la herramienta *Protein preparation wizard* (Schrödinger, LLC, New York, NY, 2017). Se asignaron estados correctos de protonación a residuos de histidina y se agregaron los átomos de hidrógeno. Finalmente se efectuó una minimización energética empleando potenciales optimizados para simulaciones líquidas (OPLS3, Schrödinger, Inc., New York, NY, 2013) como campo de fuerza. Además se utilizó la herramienta SiteMap (Schrödinger, LLC, New York, NY, 2017) para conocer los probables sitios de unión en HsEFL1 en la conformación traslocada y no traslocada. Los cálculos efectuados por SiteMap consisten en tres etapas. En la primera etapa, los sitios puntuales relevantes son seleccionados basados en sus propiedades geométricas y energéticas, los puntos se agrupan para definir sitios. Luego propiedades hidrofóbicas, hidrofílicas y otras propiedades clave se computan en una malla de puntos y se preparan los contornos de los mapas. Finalmente se calculan propiedades del sitio como el tamaño, el grado de reclusión, hidrofobicidad, e hidrofiliidad entre otras

57.

El acoplamiento se efectuó empleando Glide standard-precision (SP) y extra-precision (XP), el muestreo del ligando se manejó como flexible y se solicitaron al menos 5 poses por ligando. La afinidad del ligando por el sitio de unión se predijo empleando Glide score (Gscore) que implementa el campo de fuerza OPLS3. Posteriormente se calculó la energía libre de unión de los complejos acoplados mediante un enfoque de mecánica molecular usando Prime MM-GBSA (*Molecular Mechanics/Generalized Born Surface Area*). Además, se utilizó MOE empleando *Triangle Matcher* como método de acoplamiento. La función de puntaje fue *London dG* donde las 10 mejores salidas fueron refinadas mediante campo de fuerza clasificadas por estimación de la energía libre de unión de MM/GBVI (*Molecular Mechanics/Generalized Born volume integration*). Finalmente se utilizó el software PyRx 0.9.5 para Windows con el módulo de AutoDock Vina 2.0 incluido en él. El archivo SDF que contenía los ligandos, así como los PDB de las estructuras proteicas

fueron convertidos al formato pdbqt requerido por AutoDock Vina para efectuar el acoplamiento. Posteriormente se efectuó el acoplamiento de los ligandos en el sitio de unión a sordarina en EF-2 de *S. cerevisiae*, el homólogo estructural de EFL1, para comparar la afinidad de la proteína por los ligandos que se unirían a EFL1.

#### **4.1.3 Simulaciones de dinámica molecular**

Todos los cálculos de MD se efectuaron empleando el paquete Gromacs 5.0.4<sup>58</sup>. La topología de la proteína fue generada empleando los parámetros del campo de fuerza GROMOS 54A7. Adicionalmente se generaron las topologías y archivos de coordenadas de los ligandos empleando el servidor ATB<sup>60</sup> versión 2.2. Los complejos proteína-ligando, obtenidos del acoplamiento molecular fueron incluidos en el centro de una caja cúbica, luego el sistema fue solvatado con moléculas de agua SPE, finalmente se añadieron iones Na<sup>+</sup> y Cl<sup>-</sup> necesarios para que el sistema alcanzara la neutralidad. Los sistemas neutralizados se sometieron a una minimización energética empleando el algoritmo de *steepest descent*. La equilibración del sistema se realizó comenzando con una simulación de ensamble NVT de 200 ps con intervalos de integración de 2 fs, a 310 K (37°C) usando el termostato de Berendsen para mantener la temperatura cada 0.1 ps. Seguido por una simulación NPT, de 200 ps con paso de integración de 2 fs, a una temperatura de 310 K controlada por el termostato de Berendsen con tiempo constante de 0.1 ps y 1 Bar de presión usando el baróstato de Parrinello-Rahman isotrópico cada 2 ps y una compresibilidad de  $4 \times 10^{-5} \text{ Bar}^{-1}$ . La producción de la MD del complejo proteína-ligando se efectuó por un tiempo de 10 ns empleando el algoritmo de integración *leap-frog*, se mantuvieron las distancias de los enlaces covalentes con el algoritmo *linear constraint solver* (LINCS), se determinaron dos grupos de acoplamiento de la temperatura (proteína-ligando y solvente), las interacciones no enlazantes fueron terminadas a un radio de 14 Å empleando el esquema de Verlet, las interacciones electrostáticas de larga distancia fueron calculadas con el método Particle Mesh Ewald (PME). Para las simulaciones bajo el esquema de *simulated annealing* la producción se realizó con ciclos de calentamiento y enfriamiento del sistema; con incremento hasta 400 K (127°C) en 50 ps y se mantuvo dicha temperatura por 250 ps, el enfriamiento se realizó durante 1000 ps hasta 310 K manteniendo la temperatura constante por 1000 ps. Esta estrategia se emplea para suministrar al sistema la energía requerida para sobrepasar mínimos locales en la simulación.

#### **4.1.4 Predicción de energía de unión a ligando**

Para cada compuesto, el  $\Delta G$  de unión se calculó usando energías promedio de interacción proteína-ligando obtenidas a partir de las simulaciones de MD de los complejos, con valores de energía de interacción almacenados cada 0.5 ps. Se empleó el valor de  $\alpha = 0.18$  y  $\beta$  de acuerdo con la naturaleza química del ligando,  $\beta=0.5$  para aquellos con carga neta,  $\beta=0.43$ ,  $\beta=0.37$  o  $\beta=0.33$  para compuestos sin carga según cantidad de grupos hidroxilo. Mientras que las energías de interacción coulombicas ( $U_{elec}$ ) y de Lennard-Jones ( $U_{vdw}$ ) promedio para el ligando libre se obtuvieron a partir de simulaciones adicionales del ligando solvatado usando los mismos parámetros que para los complejos proteína-ligando.

## **4.2 Metodología experimental**

### **4.2.1 Expresión y purificación de ScEfl1 y HsEFL1**

La expresión y purificación de EFL1 se efectuó como se describe en Gijsbers *et al.*<sup>61</sup> Inocular *S. cerevisiae* BCY123 albergando el plásmido pRS426-ScEfl1TevHis o pRS426-hsEFL1TevHis en 20 mL de medio SD-URA suplementado con 2% de glucosa y Dropout-Ura e incubar toda la noche a 30°C y 200rpm. El preinoculo se empleó para inocular 1.5 L del mismo medio hasta una  $DO_{600nm}$  de 0.05 y se incubó sin agitación por 19h. Posteriormente se coloca el cultivo en agitación de 200rpm a 30°C hasta alcanzar una  $DO_{600nm}$  de 2.0. La expresión de la proteína se realizó con una concentración final de 2.8% de galactosa y se incubó a 30°C y 200 rpm por 16 a 18h. Las células se cosecharon por centrifugación a 4500rpm durante 6min (rotor JA-14 Beckman-Coulter). El paquete celular fue almacenado a -20°C hasta la purificación de la proteína.

Para purificar la proteína EFL1 se resuspendió la pastilla celular en 30 mL de buffer de lisis suplementado con 1 mM PMSF en etanol y 1 mM benzimidina (50 mM Tris-HCl pH 8, 1 M NaCl, 40 mM imidazol, 5 mM  $MgCl_2$ , 10% glicerol). La lisis de las células en suspensión se efectuó por rompimiento por fricción en el contenedor del BeadBeater con perlas de vidrio previamente humedecidas con el buffer de lisis. El rompimiento se realiza en 5 ciclos de 2 min prendido, con 15 min de reposo entre cada ciclo. El lisado celular se centrifugó a 15500 rpm por 55min en rotor JA-17 (Beckman-Coulter). El sobrenadante se decantó y filtró a través de una membrana de 0.45  $\mu m$  (Millipore). La etiqueta de histidinas adicionada en el extremo C-terminal de ambas proteínas permitió la purificación de la proteína por cromatografía de afinidad a iones  $Ni^{2+}$ . Todas las soluciones y el agua utilizadas en la purificación se filtraron a través de una membrana de 0.45  $\mu m$ . La columna se equilibró con 3 volúmenes de cama de agua desionizada, filtrada y sonicada seguido de 3 volúmenes de cama del buffer de lisis. El sobrenadante de la lisis celular introdujo a la columna y lavó 3

volúmenes de cama de buffer de lisis. Posteriormente se lavó la columna con 3 volúmenes de buffer de alta sal (50 mM Tris-HCl pH 8, 2 M NaCl) y se efectuó un segundo lavado con 3 volúmenes de cama de buffer de lisis. Finalmente la proteína se eluyó con 3 volúmenes de buffer de elución (50 mM Tris-HCl pH 8, 300 mM NaCl, 250 mM imidazol, 5 mM MgCl<sub>2</sub>). Para guardar la columna, esta se lavó con 3 volúmenes de cama de agua desionizada, filtrada y sonicada. Seguido de un lavado con 3 volúmenes de cama de etanol al 20% v/v y se almacenó. Subsecuentemente, las fracciones obtenidas de la elución de la cromatografía de afinidad a iones Ni<sup>2+</sup> se purificaron por cromatografía de exclusión molecular en una columna Superdex 200 HiLoad 16/60. La columna se equilibró con 150 mL de buffer para exclusión molecular (50 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM MgCl<sub>2</sub>). Se inyectó 1 mL de con un flujo de 0.7 mL/min. Las fracciones con la proteína se juntaron, concentraron y se guardaron a -80°C hasta su posterior uso.

#### **4.2.2 Desplazamiento térmico monitoreado por fluorescencia (Thermofluor)**

El proceso de desnaturalización expone las regiones hidrofóbicas de la proteína, lo que resulta en un incremento en la fluorescencia de fluoróforos cuya fluorescencia se incrementa al interactuar con regiones hidrofóbicas de la proteína. Los ensayos de desplazamiento térmico monitoreado por fluorescencia fueron realizados en un sistema de detección en tiempo real para PCR (Step One Plus Applied Biosystem), empleando una placa de 96 pozos. Para monitorear la desnaturalización de la proteína, se empleó Sypro orange como colorante fluorescente a una concentración final de 5X. La concentración de la proteína fue de 4 μM en amortiguador 50 mM HEPES pH 8, 150 mM NaCl, 2 mM MgCl<sub>2</sub> y 3% de DMSO. Se empleó una rampa de calentamiento de 25 a 90°C con incremento de 1°C/min. La intensidad de la fluorescencia se midió a una longitud de onda de excitación de 490 nm y a 530 nm para la emisión <sup>61</sup>.



## 5 Resultados y Discusión

A continuación se presentan los resultados obtenidos de la búsqueda por similitud estructural con la sordarina. Así como del acoplamiento molecular de los compuestos con mayor similitud 3D con la sordarina a las proteínas ScEfl1 y HsEFL1 en su conformación traslocada, obtenidos mediante las metodologías descritas anteriormente.

### 5.1 Cribado virtual basado en la similitud molecular con sordarina

Una estrategia para seleccionar compuestos con actividades similares es emplear el conocimiento contenido en bases de datos de compuestos. Se describe numéricamente al compuesto de interés y se analiza su similitud con los compuestos anotados en una base de datos. Existen varias formas de caracterizar numéricamente a un compuesto, típicamente existen dos estilos de descriptores; cadenas de bits de ceros y unos y vectores de números continuos. En los descriptores de cadenas de bits, cada bit nota la presencia o ausencia de un fragmento molecular en particular. Mientras que en un descriptor de vector de número continuos los elementos del vector notan los valores de propiedades moleculares. Para medir la similitud en PowerMV 0.61, entre una molécula blanco y cada molécula candidata por descriptores de cadenas de bits se usa la distancia de Tanimoto. Para medir la similitud entre moléculas en vectores continuos se usa la distancia de Euclidian. Los descriptores basados en pares de átomos comparan el tipo de átomos entre la molécula candidata y la molécula blanco; en el estilo Carhart cada átomo además se define basado en sus conexiones con átomos distintos a hidrógeno y el número de electrones pi. El descriptor basado en fragmentos toma en cuenta los principios bioisostéricos, por ejemplo un grupo disulfuro (-S-) se usa frecuentemente para reemplazar un grupo ester (-O-), por lo que se clasifican estos grupos como del mismo tipo. El descriptor continuo número de carga ponderado incluye electronegatividad, lipofilidad atómica y XLogP<sup>62</sup>. Los valores promedio de similitud obtenidos para los 500 vecinos más cercanos a sordarina de acuerdo con cada descriptor molecular en PowerMV versión 0.61 de los compuestos presentes en la biblioteca se recopilaron en la Tabla 1.

**Tabla 1.** Evaluación de la similitud 2D con sordarina de los compuestos en la biblioteca ZINC15 según los descriptores moleculares empleados por PowerMV 0.61.

<b>Descriptor</b>	<b>Promedio</b>
Pares de átomos	0.181 <sup>T</sup>
Pares de átomos de Carhart	0.498 <sup>T</sup>
Pares de fragmentos	0.690 <sup>T</sup>
Huellas digitales de farmacóforo	0.586 <sup>T</sup>
Número de carga ponderado	0.662 <sup>E</sup>

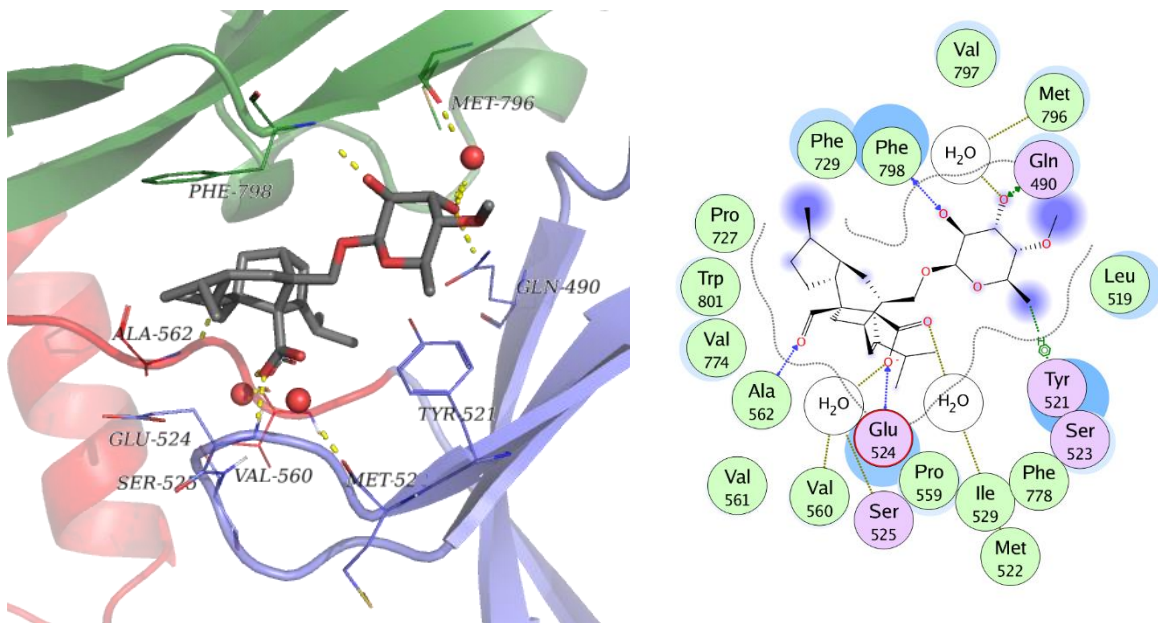
<sup>T</sup> Similitud de Tanimoto; <sup>E</sup> Similitud de Euclidian

Se obtuvieron 812 moléculas con valores de similitud menores al promedio para cada descriptor, compuestos denominados como vecinos más cercanos, donde 0.00 es el valor que representa la mayor similitud. En seguida se realizó el alineamiento 3D con la sordarina por tipo de átomos y mediante subestructura común (MCS) de estas 812 moléculas empleando MarvinSketch de ChemAxon. El valor promedio del alineamiento por tipo de átomos fue de 0.332, mientras que para el alineamiento por subestructura común fue de 0.339. Cerca de 70 moléculas poseían valores superiores al promedio para ambos tipos de alineamiento 3D. Estos compuestos fueron seleccionados para efectuar su acoplamiento en las proteínas ScEfl1 y HsEFL1 luego de corroborar su disponibilidad comercial.

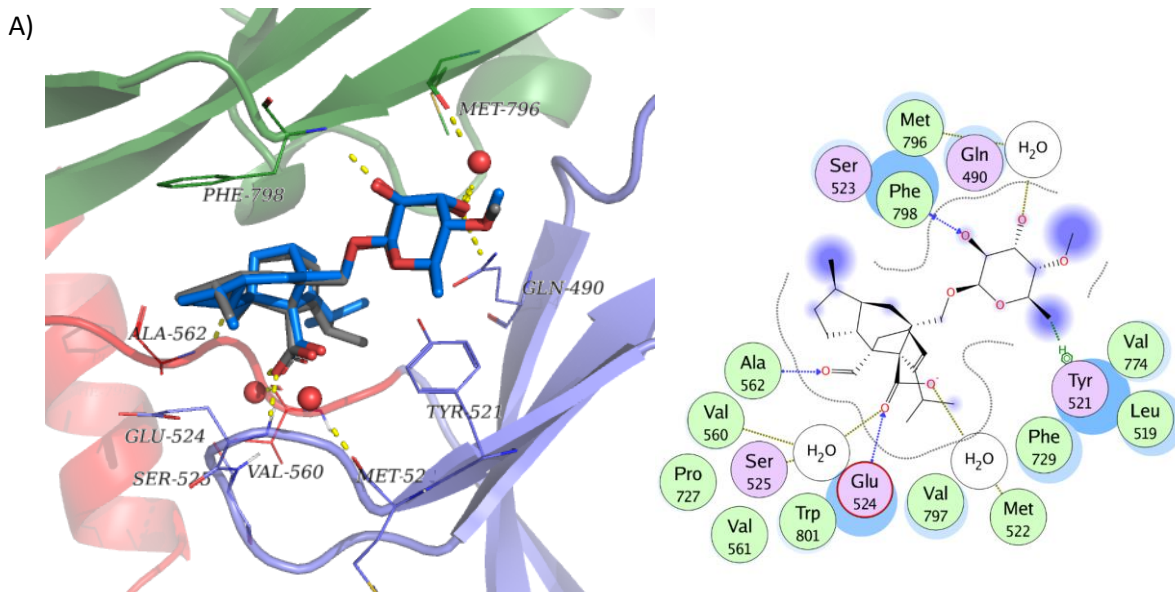
## 5.2 Acoplamiento molecular

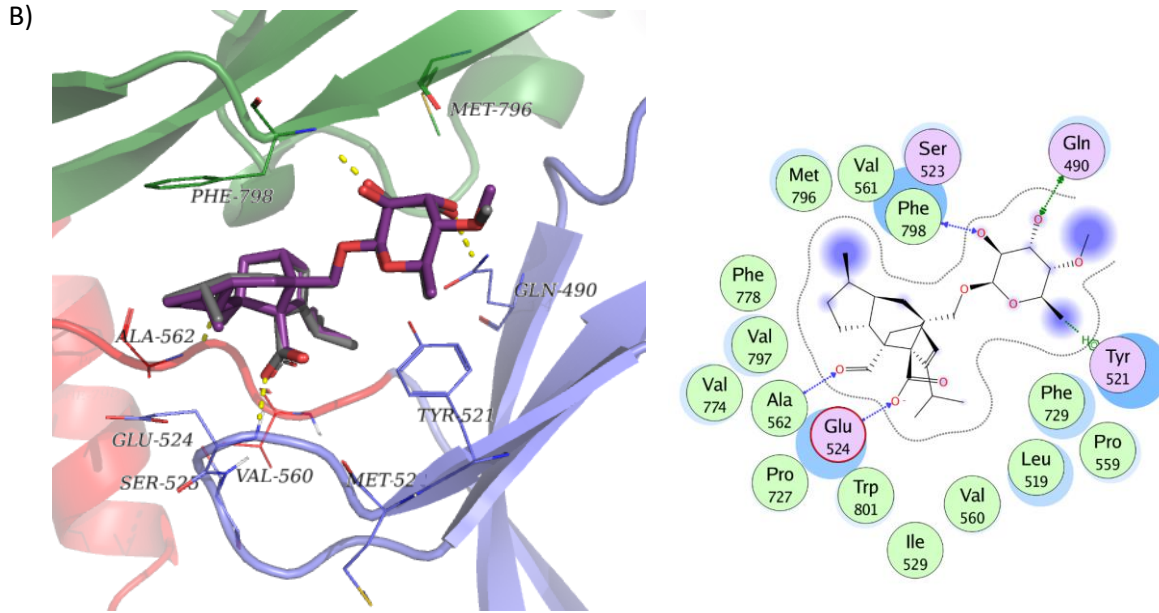
En primer lugar se realizó el acoplamiento de la sordarina en la estructura de EF-2, tanto en presencia como ausencia de moléculas de agua en el sitio de unión. Esto para evaluar si la presencia de agua era requerida para una predicción acertada de la pose del ligando como se observa en la estructura cristalográfica. Al comparar la pose de la sordarina con mayor afinidad predicha por MOE del acoplamiento en presencia de moléculas de agua, contra la conformación de sordarina co-cristalizada en EF-2 (Figura 8) se obtuvo una desviación cuadrática media (RMSD) de 0.4 Å y un puntaje de -39.99 kcal/mol. Mientras que para el acoplamiento en EF-2 excluyendo las moléculas de agua el RMSD calculado fue de 0.5 Å y el puntaje de -29.00 kcal/mol. Dado que la pose predicha del ligando posee un RMSD menor a 2 Å, se puede inferir que la presencia de moléculas de agua no es un factor determinante para la predicción de la pose de unión del ligando en este sistema, ya que la conformación es prácticamente la misma que la de la sordarina co-cristalizada (Figura 9). En cambio, la presencia de las moléculas de agua sí favorece el puntaje de

energía de interacción con EF-2. Ya que como se muestra en la Figura 9 A las moléculas de agua en el sitio de unión se encuentran mediando interacciones con residuos que no contactan directamente con sordarina como la Met522, Val560 y Met796.



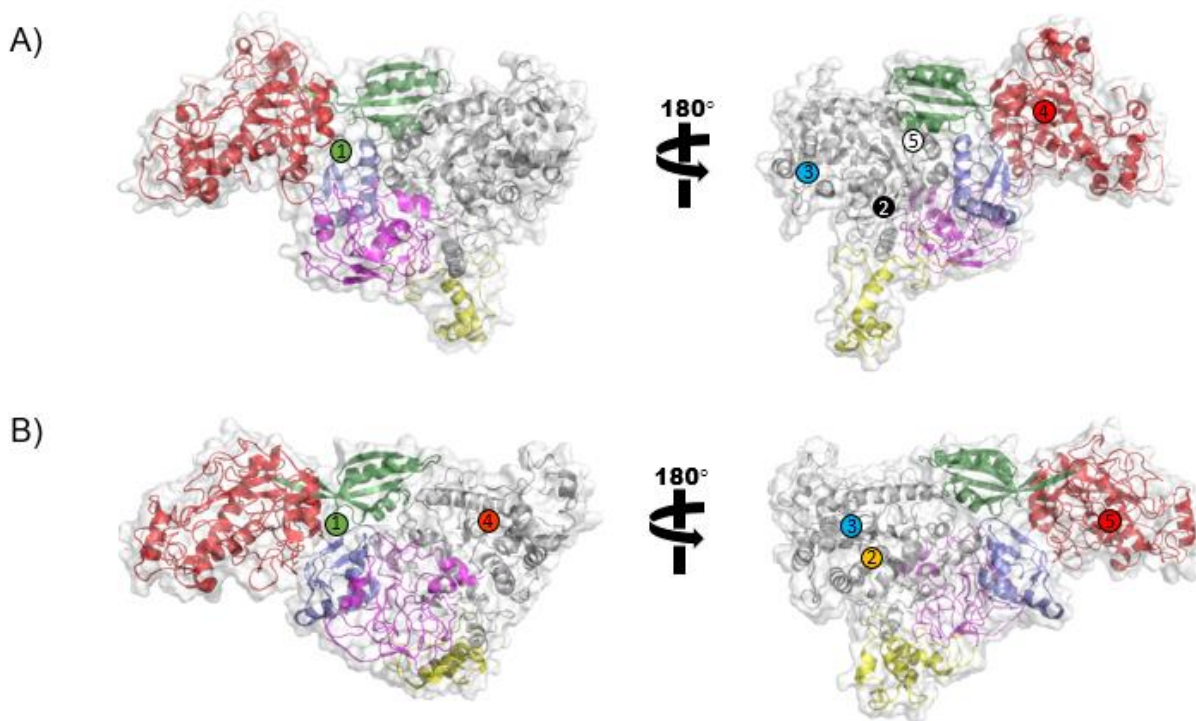
**Figura 8.** Sordarina co-cristalizada con EF-2 y su diagrama de interacción en el sitio de unión entre los dominios III (azul), IV (rojo) y V (verde). Se muestran las moléculas de agua mediante las cuales interactúan los residuos en EF-2 (Ver Anexo 9.3 para conocer el tipo de interacción proteína-ligando).





**Figura 9.** Pose de unión de sordarina predicha por MOE y diagrama de interacción en EF-2 en presencia (A) o ausencia (B) de moléculas de agua en el sitio de unión entre los dominios III (azul), IV (rojo) y V (verde). La conformación del ligando cristalizado se muestra en gris como referencia (Ver Anexo 9.3 para conocer el tipo de interacción proteína-ligando).

Primeramente, se realizó la predicción de los principales bolsillos de unión de moléculas orgánicas pequeñas en la conformación no traslocada y traslocada de la proteína EFL1 humana. Los cinco mejores sitios probables de unión predichos por la herramienta SiteMap (Schrödinger, LLC, New York, NY, 2017) se muestran en la Figura 10 y los residuos que los conforman se enlistan en las Tabla 5 y 6 del Anexo 9.1. Tras la preparación de las moléculas con Ligprep (Schrödinger, LLC, New York, NY, 2017) se realizó el acoplamiento molecular sobre las estructuras tridimensionales de ScEfl1 y HsEFL1. Además cabe mencionar que el acoplamiento se realizó sin moléculas de agua en el sitio de unión. Ya que como anteriormente se señaló para el sistema sordarina•EF-2, estas no son necesarias para predecir de forma adecuada la conformación de unión del ligando, pero sí favorecen el valor puntaje de interacción calculado.

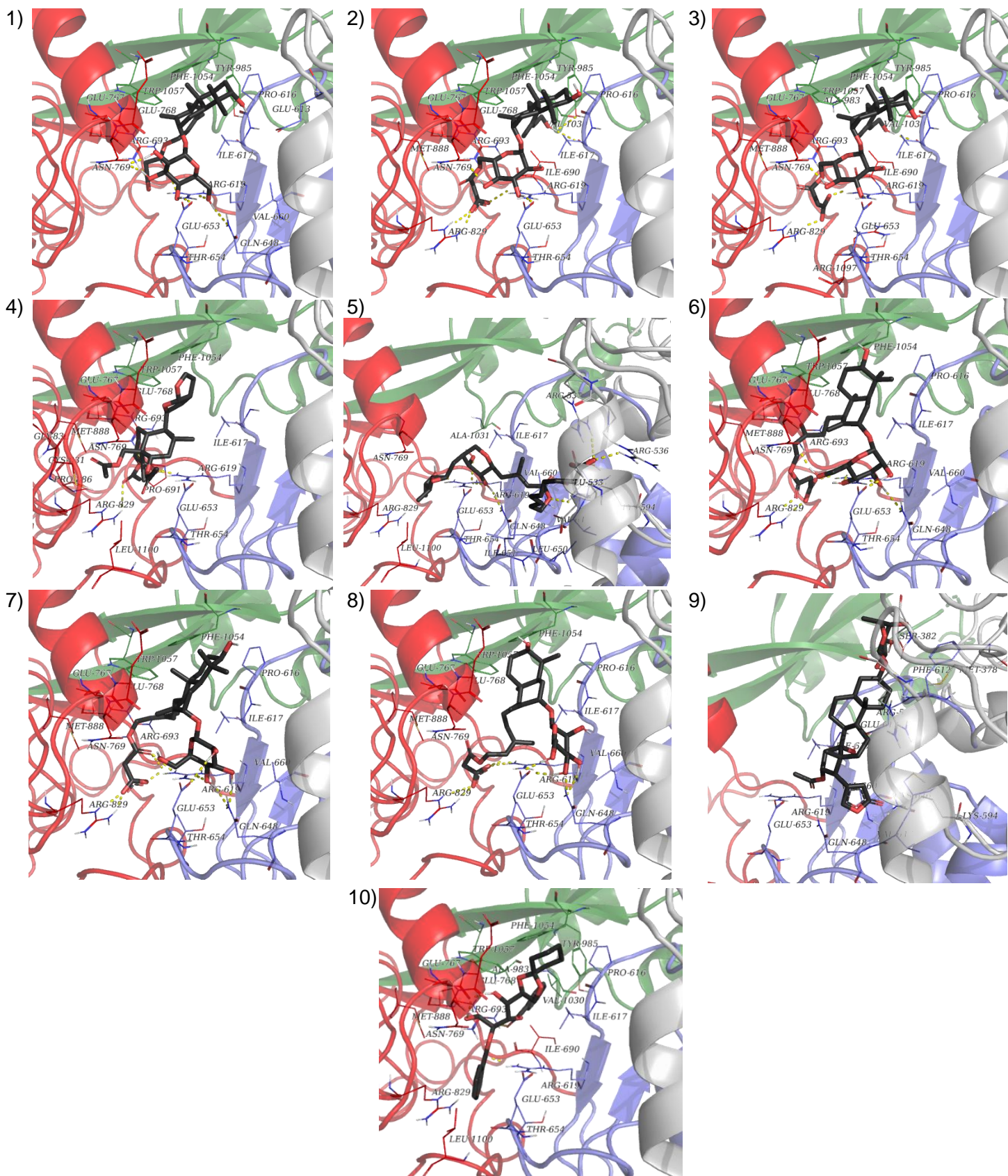


**Figura 10.** Principales sitios de unión a ligando predichos por la herramienta SiteMap en HsEFL1 en la conformación A) no traslocada (PDB ID 5ANB: K) y B) traslocada (PDB ID 5ANC: K) <sup>22</sup>. El sitio equivalente de unión a sordarina (1) se indica en verde. Dominio I (gris), II e inserción (magenta-amarillo) III (azul), IV (rojo) y V (verde).

SiteMap puede emplearse para clasificar la “drogabilidad” de las proteínas, como medida de su capacidad para unirse fuertemente a moléculas pequeñas absorbidas pasivamente. La hidrofobicidad representa el factor más importante en el Dscore que establece el programa SiteScore debido a que sitios no drogables o difícilmente drogables típicamente son mucho más hidrofílicos y menos hidrofóbicos que los sitios drogables. El valor de *Dscore* calculado para el sitio equivalente de unión a sordarina (sitio 1) que se muestra en verde en la Tabla 5 y 6 en HsEFL1 del Anexo 9.1, sugiere que presenta características de un sitio “drogable” (ver Tabla 7 en Anexo 9.2). El sitio número 3, resaltado en cyan en las tablas mencionadas, corresponde al sitio de unión a nucleótidos de guanina, lo cual explica el alto valor de *Dscore* calculado, por su capacidad predicha para unirse a moléculas pequeñas de forma pasiva como se señaló previamente.

Los 10 ligandos con mejor puntaje obtenido del acoplamiento flexible en el sitio equivalente de unión a sordarina empleando Glide XP sobre las proteínas EFL1, se resumen en la Tabla 8 del Anexo 9.4. Además se muestra el valor calculado para la energía libre de unión con Prime MM-GBSA para los complejos proteína-ligando obtenidos del acoplamiento. Cabe mencionar que el puntaje obtenido para los ligandos que se muestran en la Tabla 8 fue mucho menor en los sitios 2 al 5 comparados con los que se observaron en el sitio 1. Por lo cual existiría mayor especificidad de unión de los ligandos hacia el sitio 1 ubicado entre los dominios III al V. Así mismo se muestra la pose del ligando en el sitio de unión de la proteína para los 10 compuestos con energía de interacción predicha más favorable. Se muestran las cadenas laterales de los residuos involucrados en la unión con el ligando así como las interacciones por puente de hidrógeno (Figura 11 y Figura 12).

Así mismo se realizó el acoplamiento con MOE en la región entre los dominios III, IV y V se obtuvieron los ligandos enlistados en la Tabla 9 el Anexo 9.4 como los que poseen interacción más favorable con las proteínas HsEFL1 y ScEfl1. La pose del ligando en el sitio de unión de la proteína se ilustra en la Figura 13 y Figura 14.



**Figura 11.** Mejores 10 acoplamientos en HsEFL1 obtenidos con Glide XP en la región entre los dominios III (azul), IV (rojo) y V (verde). Las interacciones polares se muestran en amarillo.

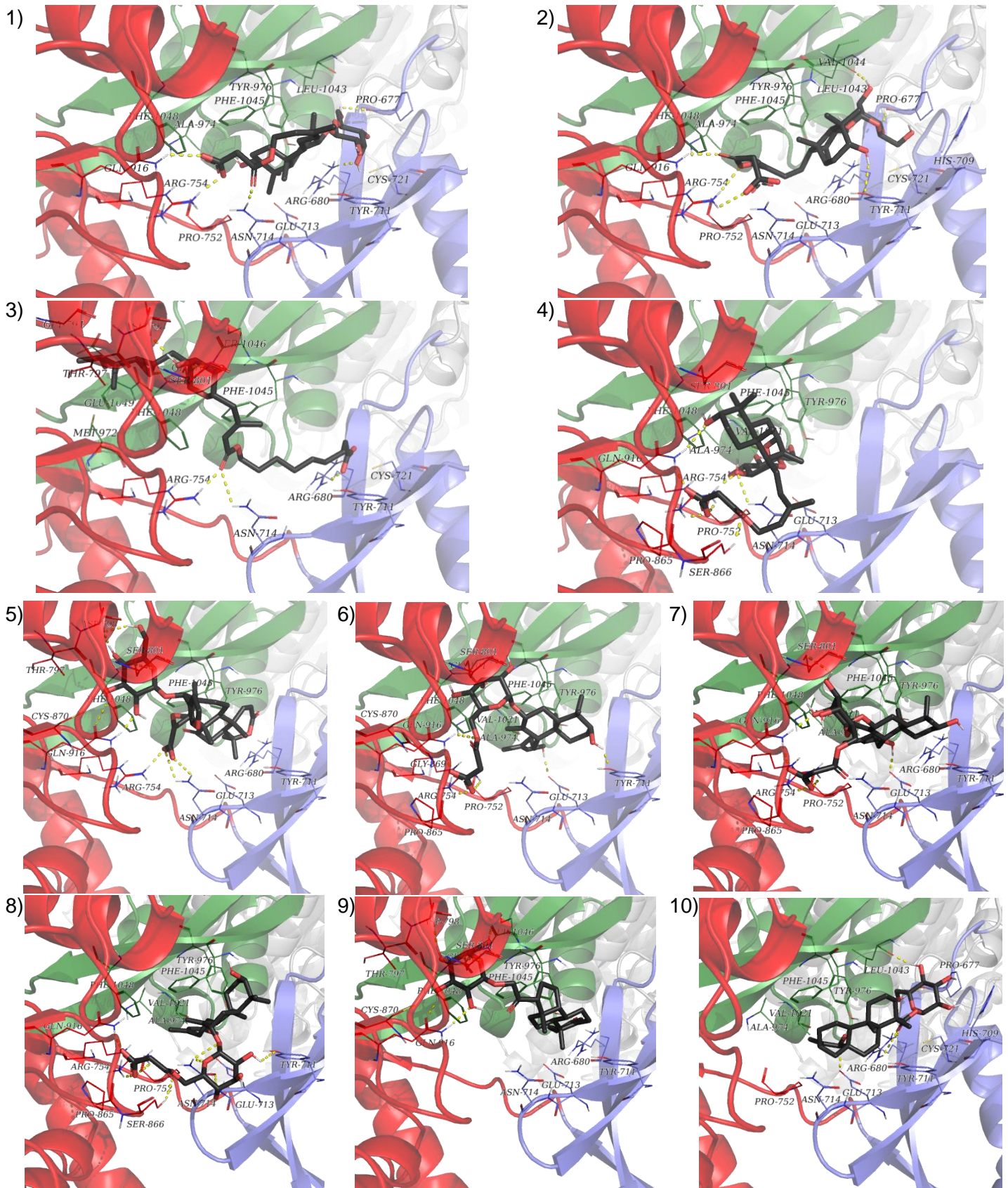
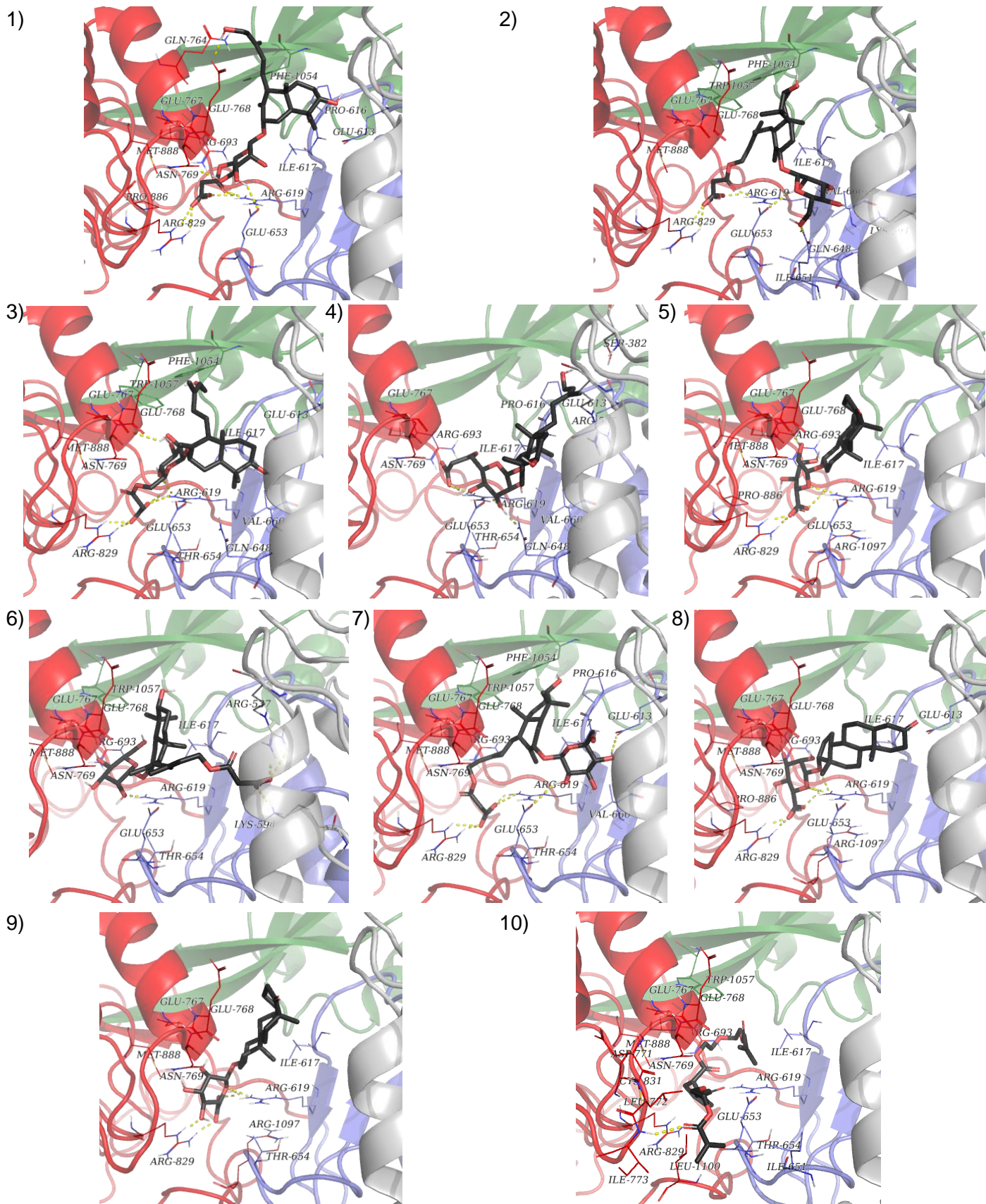
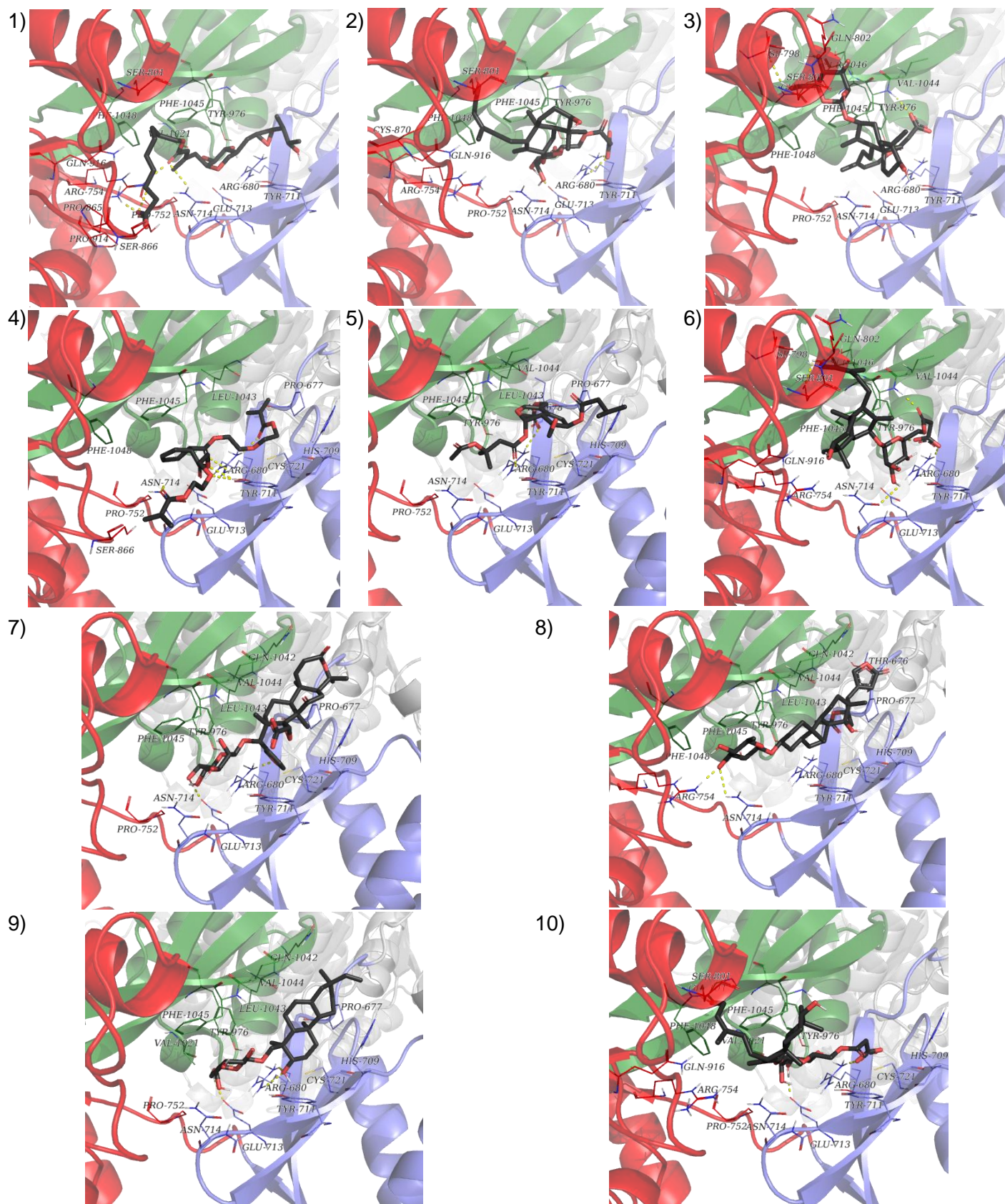


Figura 12. Mejores 10 acoplamientos en ScEfl1 obtenidos con Glide XP en la región entre los dominios III (azul), IV (rojo) y V (verde). Las interacciones polares se muestran en amarillo.





**Figura 13.** Se muestran los 10 mejores acoplamientos en HsEFL1 obtenidos con MOE en la región entre los dominios III (azul), IV (rojo) y V (verde).



**Figura 14.** Se muestran los 10 mejores acoplamientos en ScEfl1 obtenidos con MOE en la región entre los dominios III (azul), IV (rojo) y V (verde).

En los acoplamientos obtenidos para la proteína HsEFL1 se observan interacciones polares importantes de los ligandos con los residuos Arg619 y Phe1054 empleando Glide (Figura 11) y MOE (Figura 13). Así como con los residuos Arg680 y Phe1045 en los acoplamientos en ScEfl1 efectuados con Glide (Figura 12) y con MOE (Figura 14). Estos residuos son los equivalentes a Gln490 y Phe798 presentes en su homólogo estructural EF-2, los cuales se encuentran formando puentes de hidrógeno con la sordarina en la estructura cristalizada 1N0U. Finalmente en la Tabla 10 del Anexo 9.4 se presentan los 10 ligandos para los cuales se calculó el mejor puntaje de acoplamiento empleando AutoDock Vina en PyRx hacia las proteínas HsEFL1 (Figura 15) y ScEfl1 (Figura 16). Dado que el algoritmo de búsqueda conformacional, así como la función de puntaje son distintas para cada uno de los programas usados en el acoplamiento, no se predicen las mismas moléculas entre las 10 mejores según su puntaje asignado. Sin embargo existen compuestos coincidentes en los tres programas de acoplamiento molecular usado como ZINC000085946246 y ZINC000257470512 para los acoplamientos en ScEfl1, y ZINC000253534724, ZINC000077292071 y ZINC000036052121 para los acoplamientos en HsEFL1.

Recordemos que el puntaje de interacción predicho por MOE para sordarina en EF-2 excluyendo las moléculas de agua, puesto que no parecen influir sobre la conformación de la pose de unión predicha, fue de -29.00 kcal/mol. En el caso de los compuestos con mejor puntaje predicho empleando MOE que se enlistan en la

Tabla 9, el compuesto con mejor afinidad calculada en HsEFL1 fue ZINC000253534724 (-27.22 kcal/mol) y en ScEfl1 fue ZINC000077312672 (-28.95 kcal/mol), ambos compuestos contienen un anillo de tetrahidropirano al igual que la sordarina.

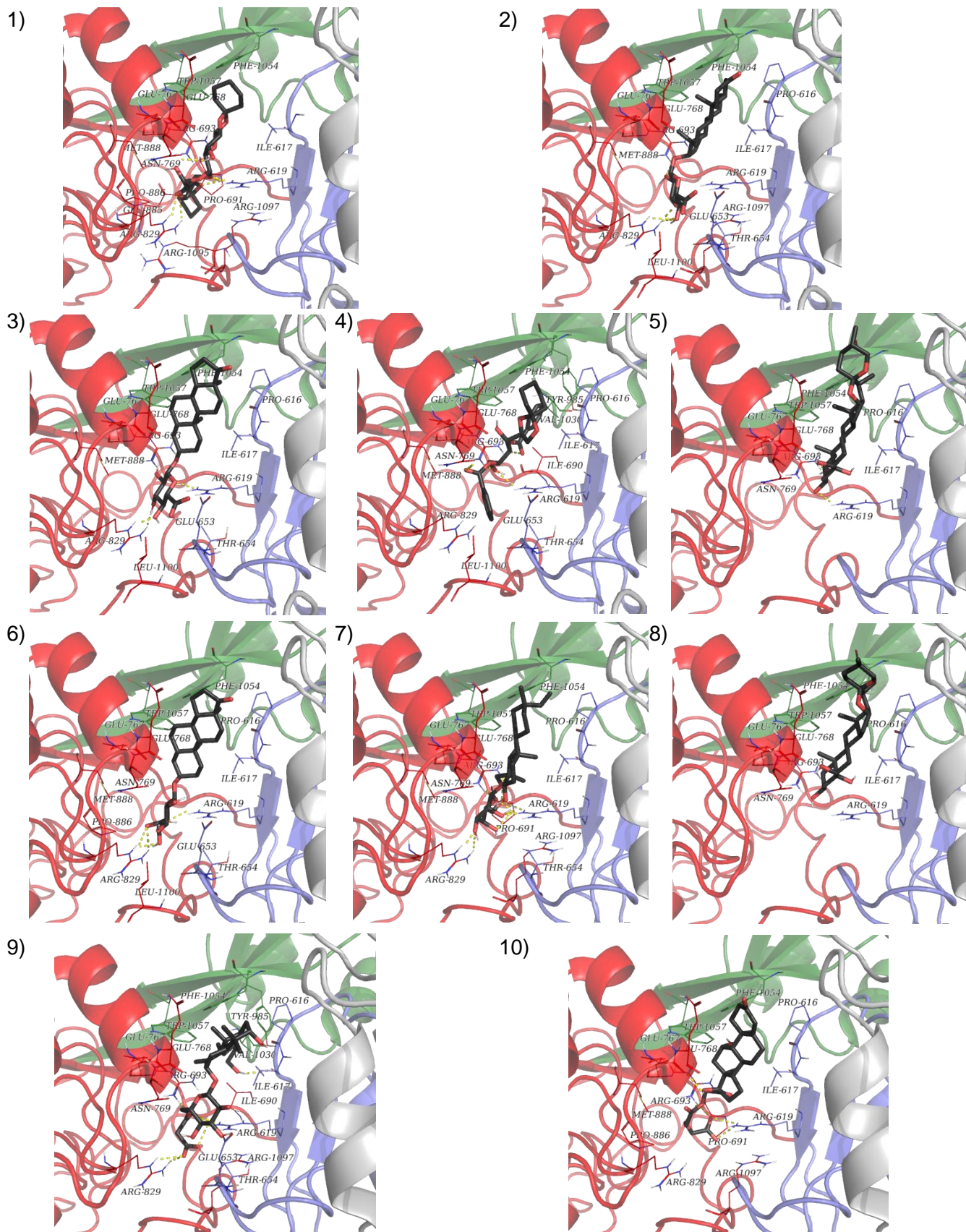


Figura 15. Se muestran los 10 mejores acoplamientos en HsEFL1 obtenidos con AutoDock Vina en PyRx en la región entre los dominios III (azul), IV (rojo) y V (verde).

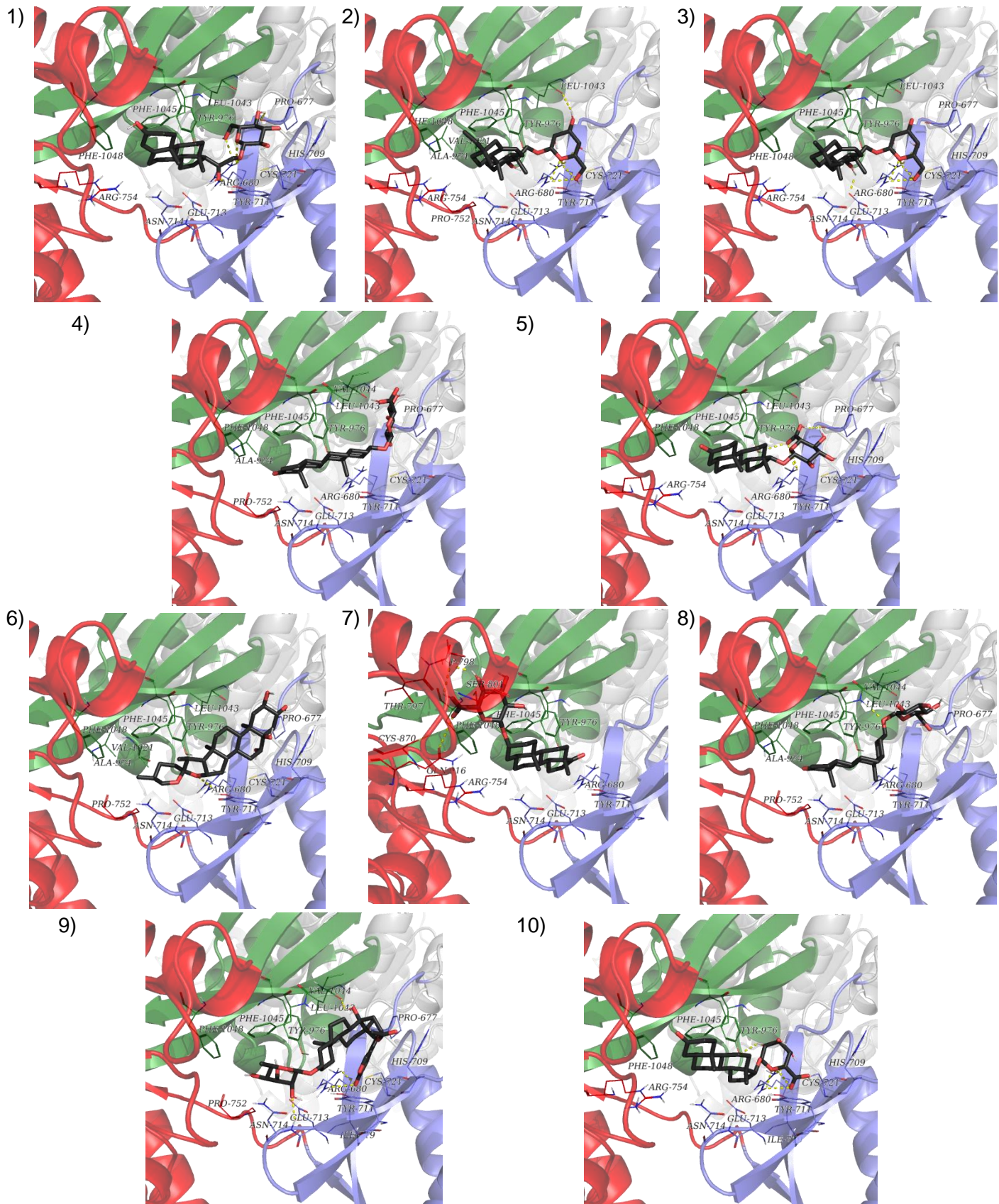
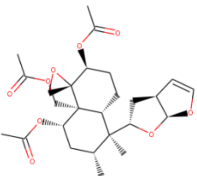
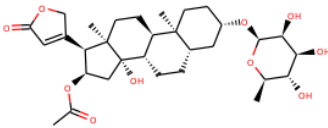
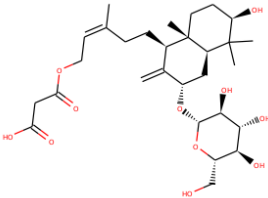
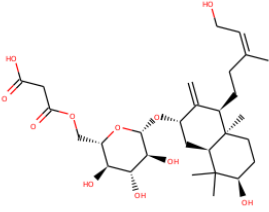
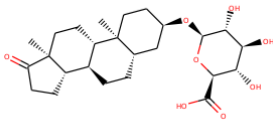
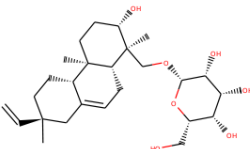
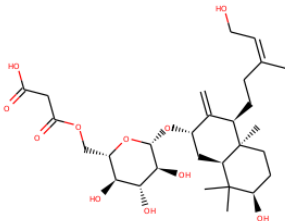
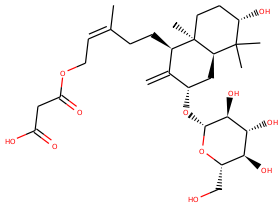
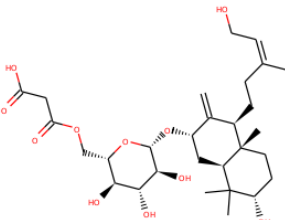
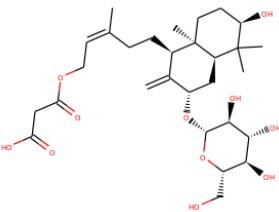
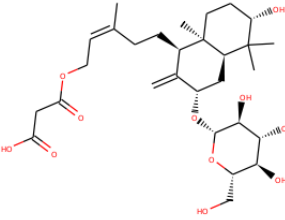
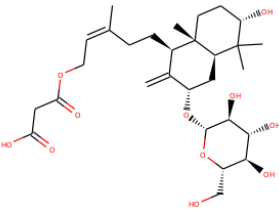
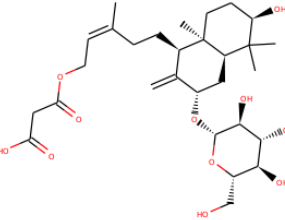
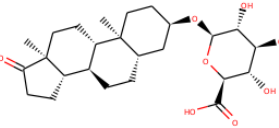


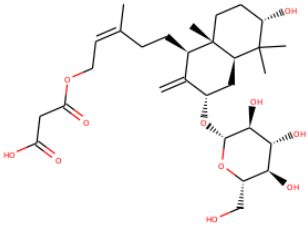
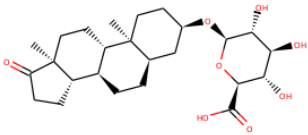
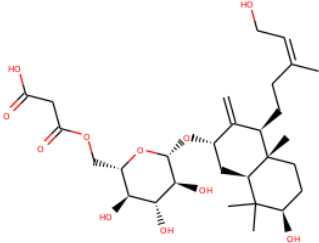
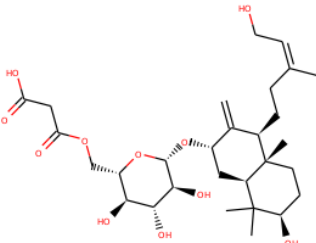
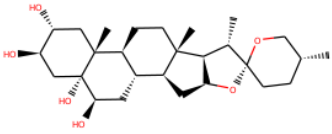
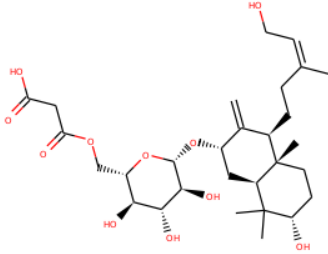
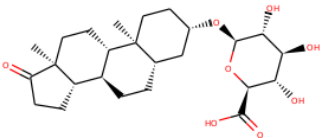
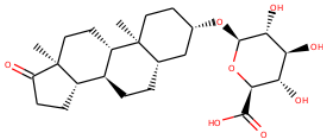
Figura 16. Se muestran los 10 mejores acoplamientos en ScEfl1 obtenidos con AutoDock Vina en PyRx en la región entre los dominios III (azul), IV (rojo) y V (verde).

Además se realizó el acoplamiento en EF-2 (PDB ID: 1N0U) de los ligandos con mayor similitud 3D con sordarina pero seleccionados para la proteína EFL1. En la Tabla 2 se presenta el puntaje calculado para el acoplamiento en EF-2 de *S. cerevisiae* de los 20 ligandos que obtuvieron los mejores puntajes de acoplamiento tanto en ScEfl1 como en HsEFL1.

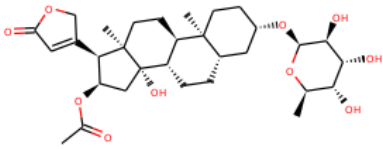
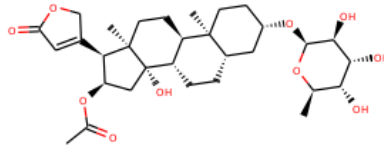
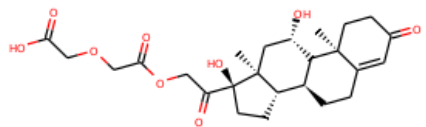
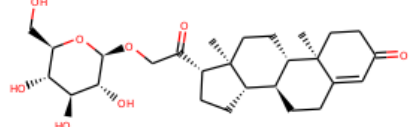
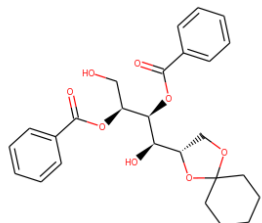
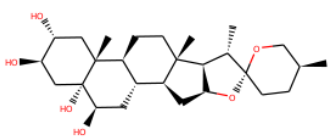
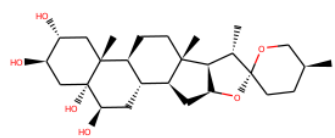
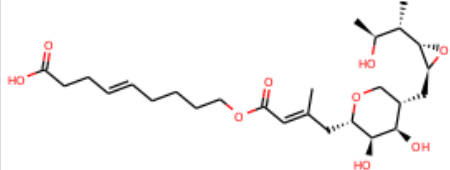
**Tabla 2.** Puntajes calculados para el acoplamiento molecular de los ligandos en EF-2 de *S. cerevisiae* empleando Glide.

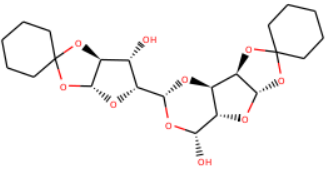
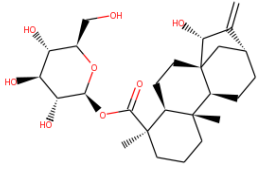
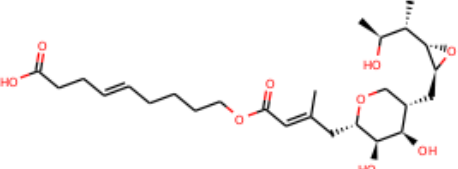
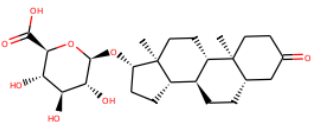
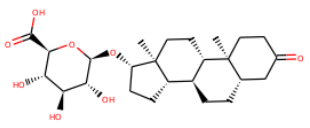
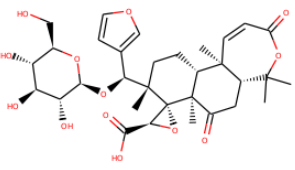
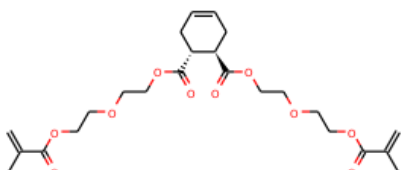
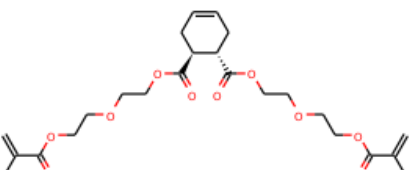
No.	Ligandos con mejor puntaje en HsEFL1		Ligandos con mejor puntaje en ScEfl1	
	Ligando	XP Glide score en EF-2 (en HsEFL1) kcal/mol	Ligando	XP Glide score en EF-2 (en ScEFL1) kcal/mol
1	 ZINC000085340910	-7.49 (-8.94)	 ZINC000257470514	-7.58 (-7.42)
2	 ZINC000253532104	-7.30 (-8.54)	 ZINC000253534723	-7.27 (-8.13)
3	 ZINC000008234384	-7.30 (-7.16)	 ZINC000085946246	-7.27 (-7.16)

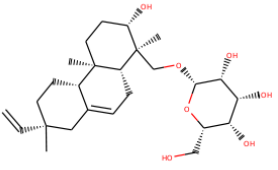
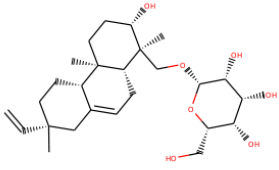
No.	Ligandos con mejor puntaje en HsEFL1		Ligandos con mejor puntaje en ScEfl1	
	Ligando	XP Glide score en EF-2 (en HsEFL1) kcal/mol	Ligando	XP Glide score en EF-2 (en ScEFL1) kcal/mol
4	 <p>ZINC000253534723</p>	-7.27 (-6.68)	 <p>ZINC000253532111</p>	-7.24 (-9.80)
5	 <p>ZINC000253534724</p>	-7.24 (-9.41)	 <p>ZINC000253532106</p>	-7.23 (-9.32)
6	 <p>ZINC000253532111</p>	-7.24 (-8.56)	 <p>ZINC000253532108</p>	-7.23 (-8.72)
7	 <p>ZINC000253532106</p>	-7.23 (-8.49)	 <p>ZINC000008234384</p>	-7.23 (-6.94)

Ligandos con mejor puntaje en HsEFL1		Ligandos con mejor puntaje en ScEfl1		
No.	Ligando	XP Glide score en EF-2 (en HsEFL1) kcal/mol	XP Glide score en EF-2 (en ScEFL1) kcal/mol	
8	 <p>ZINC000253532108</p>	-7.23 (-10.21)	 <p>ZINC000008234385</p>	-7.23 (-6.94)
9	 <p>ZINC000253534722</p>	-7.14 (-9.42)	 <p>ZINC000253534722</p>	-7.14 (-8.37)
10	 <p>ZINC000038140890</p>	-6.74 (-5.59)	 <p>ZINC000253534724</p>	-6.45 (-8.44)
11	 <p>ZINC000077292071</p>	-6.42 (-7.16)	 <p>ZINC000077292071</p>	-6.42 (-6.94)



No.	Ligandos con mejor puntaje en HsEFL1		Ligandos con mejor puntaje en ScEfl1	
	Ligando	XP Glide score en EF-2 (en HsEFL1) kcal/mol	Ligando	XP Glide score en EF-2 (en ScEFL1) kcal/mol
12	 <p>ZINC000257470512</p>	-6.21 (-7.40)	 <p>ZINC000257470512</p>	-6.21 (-6.36)
13	 <p>ZINC000004692041</p>	-6.05 (-6.42)	 <p>ZINC000008214538</p>	-6.04 (-8.08)
14	 <p>ZINC000036052121</p>	-6.02 (-7.31)	 <p>ZINC000038140888</p>	-6.02 (-6.11)
15	 <p>ZINC000038140888</p>	-6.02 (-5.71)	 <p>ZINC000077312672</p>	-5.84 (-9.03)

	Ligandos con mejor puntaje en HsEFL1		Ligandos con mejor puntaje en ScEfl1	
No.	Ligando	XP Glide score en EF-2 (en HsEFL1) kcal/mol	Ligando	XP Glide score en EF-2 (en ScEFL1) kcal/mol
16	 <p>ZINC000096102533</p>	-5.94 (-5.74)	 <p>ZINC000096023887</p>	-5.84 (-8.06)
17	 <p>ZINC000077312672</p>	-5.84 (-8.94)	 <p>ZINC000040164488</p>	-5.71 (-6.92)
18	 <p>ZINC000040164488</p>	-5.71 (-6.79)	 <p>ZINC000255240337</p>	-5.40 (-8.69)
19	 <p>ZINC000033604952</p>	-5.26 (-5.30)	 <p>ZINC000033604951</p>	-5.26 (-5.33)

No.	Ligandos con mejor puntaje en HsEFL1		Ligandos con mejor puntaje en ScEfl1	
	Ligando	XP Glide score en EF-2 (en HsEFL1) kcal/mol	Ligando	XP Glide score en EF-2 (en ScEFL1) kcal/mol
20	 ZINC000085946247	-5.18 (-6.53)	 ZINC000085946247	-5.18 (-7.16)

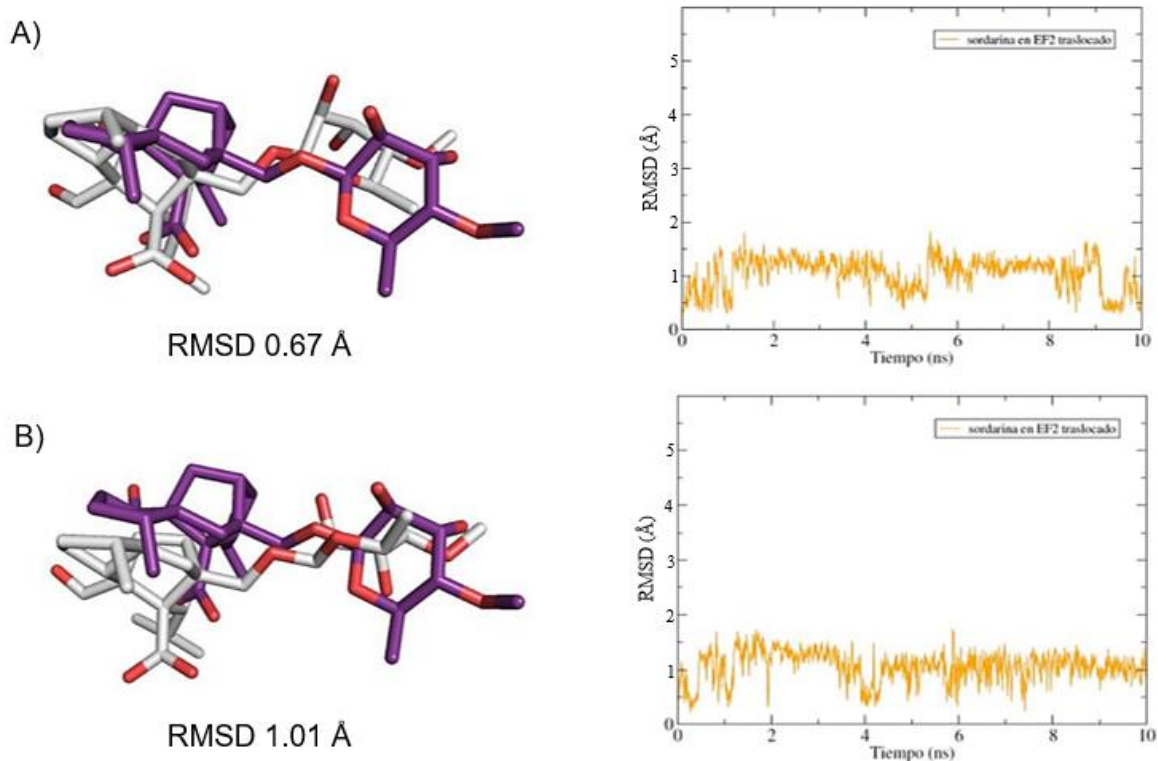
En general los ligandos parecen poseer energías de unión menos favorables hacia EF-2 que hacia EFL1, sugiriendo que las moléculas se unirían preferiblemente a ésta última. Los compuestos para los cuales se predecía una mejor interacción con EF-2 fueron descartados. Posteriormente se inspeccionaron los modos de unión predichos por Glide, MOE y Autodock vina de las moléculas con puntajes de XP Glide menores a -5.0 kcal/mol. Los compuestos para los que se observó mayor similitud entre las poses obtenidas de los tres distintos acoplamientos fueron seleccionados para efectuar un análisis más detallado empleando simulaciones de dinámica molecular (Tabla 11 Anexo 9.5).

### 5.3 Simulaciones de dinámica molecular y predicción de la energía de unión de ligando

Tras realizar la producción de las simulaciones de acuerdo con lo descrito en la sección 4.1.3, se analizaron las trayectorias producidas, graficando el RMSD (del inglés *Root Mean Square Deviation*) para cada sistema en función del tiempo. La conformación final de los ligandos en cada simulación fue comparada con la conformación de inicio predicha por el acoplamiento.

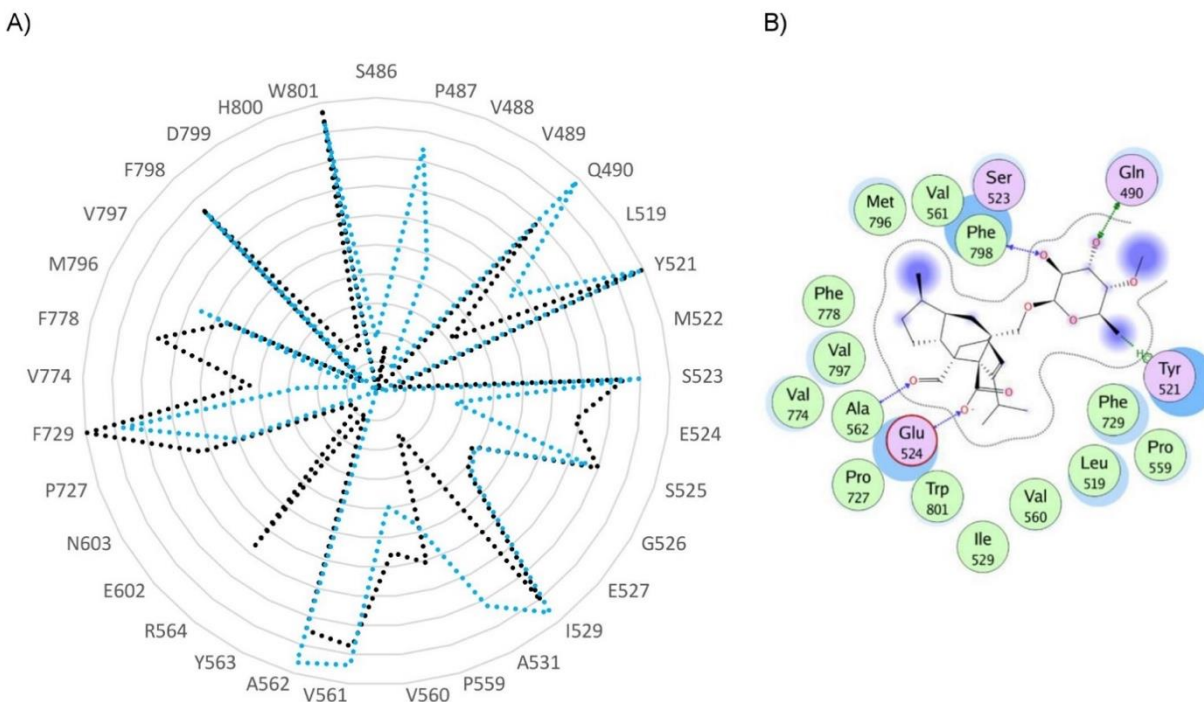
#### 5.3.1 Complejo EF-2•sordarina

Con el objetivo de elucidar la estabilidad conformacional de la sordarina y EF-2 se realizaron las simulaciones empleando una metodología denominada *simulated annealing*. Esta técnica consiste en ciclos de calentamiento y enfriamiento del sistema para evaluar a qué conformación converge el sistema después de la desnaturalización y subsecuente re-naturalización del sistema.



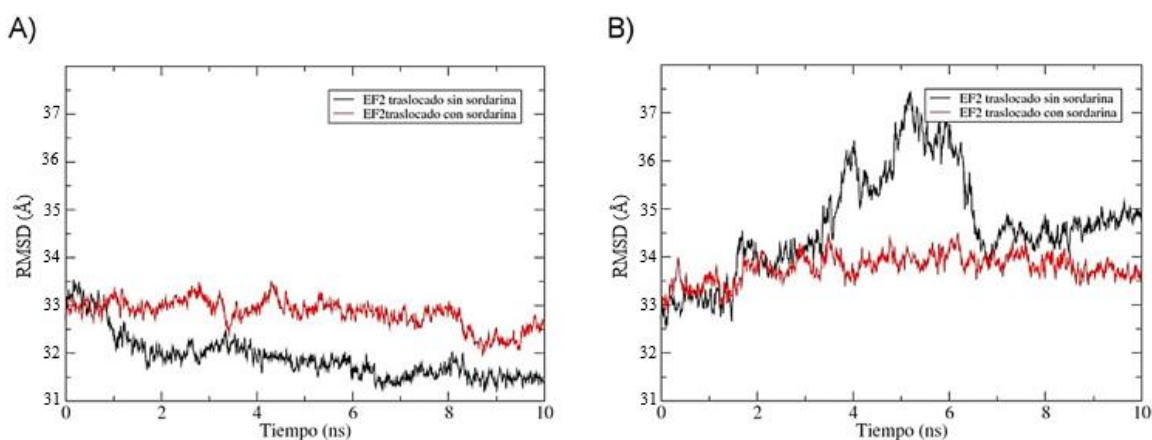
**Figura 17.** Comparación de la conformación final (blanco) contra la inicial (morado) de sordarina obtenida del acoplamiento entre los dominios III al V en la conformación traslocada de EF-2 a 310 K (A) y empleando *simulated annealing* (B) con sus gráficas de RMSD de los átomos pesados de la sordarina.

Como se ilustra en la Figura 17 la variación en el RMSD de los átomos pesados de la sordarina no es mayor a 2 Å, es decir, la conformación de la sordarina al inicio y final de la simulación para ambos casos es prácticamente la misma. Esto indica que el ligando se encuentra en una conformación estable en el sitio de unión. Para corroborar lo anterior, además se evaluaron las frecuencias de contacto mayores a 5.0 %, con los residuos de EF-2 que se encontraban a 4 Å de la sordarina durante la simulación entera (Figura 18). Tanto en la simulación a 310 K como con *simulated annealing* los contactos con la Y521, F798 y W801 se mantienen constantes; así mismo otros residuos con alta frecuencia de contacto son la Q490, I529, A562 y F729. Entre estos residuos con importante probabilidad de contacto de EF-2 se encuentran Q490, A562 y F798 que forman puentes de hidrógeno con el ligando de acuerdo con la estructura cristalográfica del complejo EF-2•sordarina (PDB ID: 1N0U) <sup>28</sup>. La frecuencia en el contacto con sordarina para residuos como R564 y F778 disminuye bajo el esquema con *simulated annealing* si se compara con la frecuencia de contacto en la simulación a 310 K. Esto es de esperarse, debido al incremento de la energía cinética durante el *simulated annealing*, de modo que la distancia podría aumentar para las interacciones débiles.



**Figura 18.** A) Frecuencia de contacto de residuos en EF-2 que se encuentran a 4 Å de sordarina durante la simulación a 310 K (negro) y en *simulated annealing* (azul) donde cada círculo representa un incremento de 10% de probabilidad de contacto. B) Diagrama de interacción de la pose de unión inicial de sordarina con EF-2.

La transición de la conformación del estado traslocado al no traslocado fue evaluada a través del cambio en el RMSD del esqueleto de la proteína (átomos N, C $\alpha$ , C, O). Se comparó cada cuadro de la simulación con la conformación en el estado no traslocado de EF-2 (PDB ID: 1N0V) cuando la proteína se encontraba sin ligando, o bien en complejo con sordarina (Figura 19). En el panel B es posible notar que la conformación del esqueleto de EF-2 en complejo con sordarina (línea roja) guarda una diferencia constante en el RMSD con la conformación no traslocada de la misma proteína que es contra la cual se compara, la superposición de las estructuras se ilustra en la Figura 20 paneles B y D cuando la sordarina se encuentra presente. Es decir la conformación no cambia de aquella en el estado traslocado (inicial) y se encuentra estable pues el RMSD no presenta fluctuaciones mayores a 1 Å. A diferencia de esto, el esqueleto de EF-2 muestra cambios más significativos en el RMSD en la simulación en la que se removieron las coordenadas de sordarina, lo cual indica una menor estabilidad en su conformación durante la simulación.

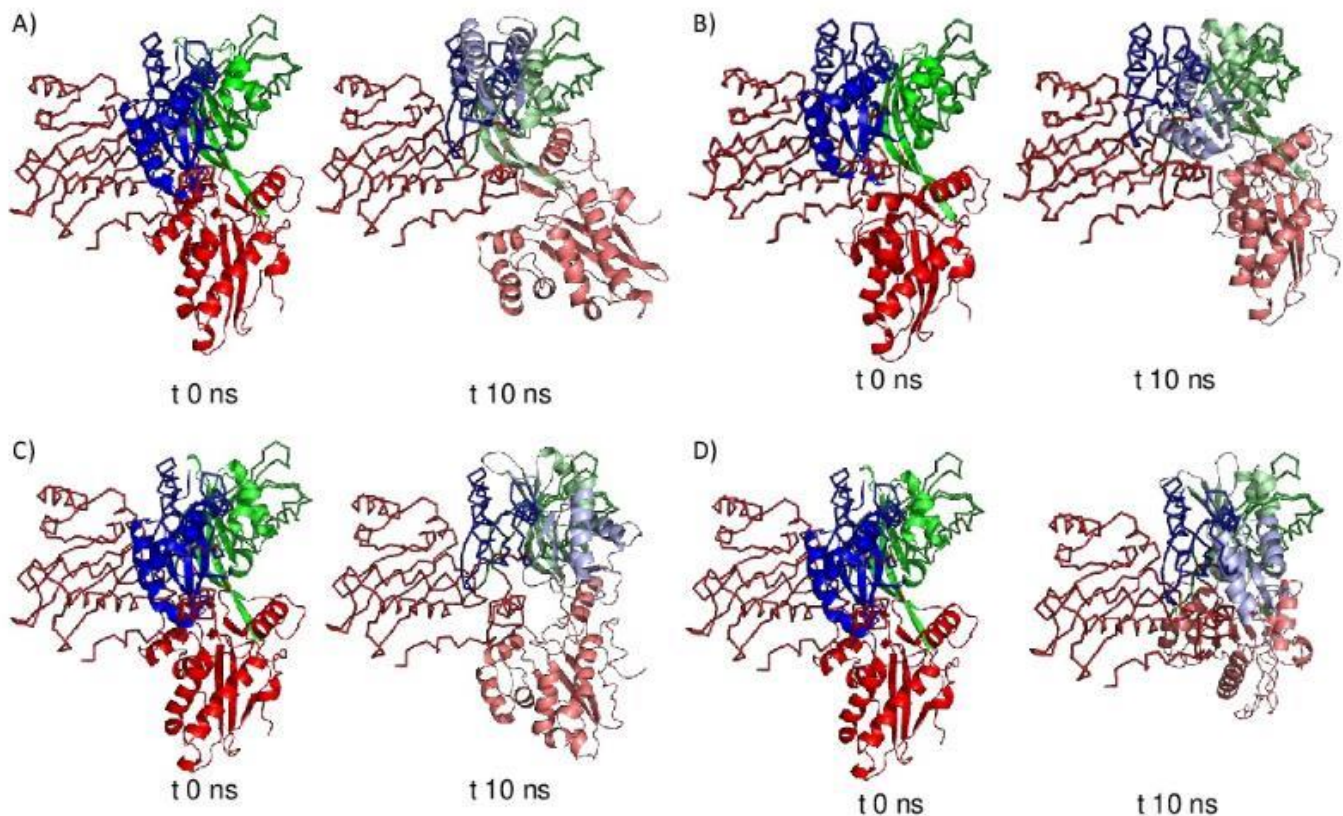


**Figura 19.** Transición de la conformación del esqueleto de EF-2 calculado para los átomos N, C $\alpha$ , C, O desde el estado traslocado hacia el no traslocado a 310 K (A) y empleando *simulated annealing* (B) en ausencia (negro) o presencia de sordarina (rojo).

La Figura 20 muestra la superposición de la estructura de EF-2 al inicio y al final de la simulación fijando los dominios I y II. Se puede notar en la Figura 20 A que si bien, tras 10 ns de simulación en ausencia de sordarina la conformación de los dominios IV y V apenas ha comenzado a cambiar, la correspondiente conformación en el dominio III en azul claro se ha modificado y ha adquirido una conformación que se asemeja a la de dicho dominio en la estructura no traslocada de EF-2 (RMSD 8.14 Å). Al comparar la conformación del dominio III al inicio de la simulación con la del estado no traslocado el RMSD era de 14.32

Å. Mientras que al comparar la conformación del dominio III después de 10 ns de simulación del complejo EF-2•sordarina con su conformación en la estructura no traslocada el RMSD es 9.42 Å. Esto indica que en presencia de sordarina la conformación de dicho dominio varía menos que cuando no está presente, dado que el RMSD mantiene un valor más cercano al de inicio cuando se compara con la conformación no traslocada de EF-2.

Una vez efectuada la simulación empleando el programa de *simulated annealing* (Figura 20 C y D) el RMSD del dominio III en EF-2•sordarina comparado con su conformación en EF-2 no traslocado fue de 19.42 Å y para EF-2 sin ligando fue de 21.77 Å. Esta diferencia en el RMSD se correlaciona con el comportamiento observado en el gráfico de la Figura 19 B, donde existe una gran fluctuación en el RMSD del esqueleto de EF-2 cuando esta se encuentra sin ligando (línea negra), revelando una menor estabilidad conformacional. Por lo cual la sordarina presenta un rol estabilizador en la conformación de la proteína.



**Figura 20.** Superposición de EF-2 alineando los dominios I y II (no visibles), se representa como cartoon los dominios III en azul, IV en rojo y V en verde al tiempo 0 (tonos oscuros) y 10 ns (tonos claros) con la conformación no traslocada en representación de listón. En ausencia (A,

C) o presencia (B, D) de sordarina durante la simulación a 310 K (A y B) y bajo el esquema de *simulated annealing* (C y D) desde la conformación traslocada de EF-2.

Posteriormente se evaluó la estabilidad conformacional de EFL1 en complejo con los ligandos listados en la Tabla 11 mediante MD. Los gráficos de RMSD en la Figura 32 y Figura 33 del Anexo 9.6 se generaron mediante la comparación de las conformaciones a lo largo de la trayectoria respecto a las de inicio, es decir la conformación traslocada, del esqueleto de la proteína HsEFL1 y ScEfl1, respectivamente. En estos gráficos se puede apreciar que a lo largo de los 10 ns de simulación, en general no existen fluctuaciones muy marcadas en el esqueleto de la proteína de los complejos analizados, por lo que parece ser que la presencia de los ligandos no altera la estabilidad de la estructura proteica. Los complejos dejan de presentar incremento significativo en el RMSD alrededor de los 4 ns, con excepción del complejo HsEFL1-09 (Figura 32 B) que muestra una fluctuación luego de los 5 ns y el complejo ScEfl-16 (Figura 33 D) que manifiesta incremento hasta los 8 ns de simulación, lo que indicaría menor estabilidad en la conformación de la proteína.

Para investigar la estabilidad de unión de los ligandos en la región interdominio de las proteínas, se calculó el RMSD de la posición de cada ligando en el complejo con las proteínas HsEFL1 (Figura 34 Anexo 9.6) y ScEfl1 (Figura 36 Anexo 9.6). El ligando 16 mostró las mayores fluctuaciones en el RMSD tanto para el complejo con HsEFL1 como con ScEfl1, lo que implicaría una baja interacción de este compuesto en la región equivalente de unión a sordarina en dichas proteínas. A continuación se realizó una extensión de la simulación para el complejo HsEFL1-16 y ScEfl1-16 hasta 100 ns para evaluar si el ligando era capaz de estabilizarse en el sitio de unión tras un mayor tiempo de simulación (Figura 35 y Figura 37 Anexo 9.6). Este ligando no llega a obtener una conformación estable en ScEfl1, pero en complejo con HsEFL1 sí logra mantener una conformación que no varía más de 2 Å a partir de los 60 ns de simulación.

Una vez realizadas las simulaciones de los ligandos en complejo con HsEFL1 y ScEfl1, para realizar la estimación de las energías de unión se aplicó la metodología LIE empleando el factor  $\beta$  correspondiente de acuerdo la naturaleza química del ligando determinados por Hansson *et al.*<sup>54</sup>. Siendo 0.37 para compuestos neutros con un solo grupo hidroxilo, 0.33 si posee dos o más grupos hidroxilo y 0.5 para compuestos cargados. Además, se usaron las energías de interacción coulómbicas y de Lennard-Jones promedio para el ligando libre



obtenidas de simulaciones del ligando en agua. Las energías de unión predichas para los ligandos de los complejos con HsEFL1 y ScEfl1 analizados mediante simulación molecular se recopilan en la Tabla 3 y Tabla 4.

La energía de unión más negativa a HsEFL1 corresponde al ligando número 17, mientras que la menos favorable corresponde al compuesto 16, el cual además mostró una de las mayores fluctuaciones en el RMSD (Figura 34 D Anexo 9.6). Lo que sugiere que no habrá interacción propicia entre dicho compuesto y la región ubicada entre los dominios III, IV y V de HsEFL1. Este resultado difiere de los obtenidos en los acoplamientos, ya que en ellos el compuesto número 16 obtuvo puntajes favorables (Glide= -9.43 kcal/mol, MOE= -22.95 kcal/mol, AutoDock Vina= -6.5 kcal/mol). Para los complejos con la proteína ScEfl1, la energía de unión más favorable corresponde al compuesto número 3 y la más desfavorable al compuesto 9, el cual también obtenía puntajes negativos en el acoplamiento molecular (Glide= -6.917 kcal/mol, MOE= -18.341 kcal/mol, AutoDock Vina= -9.40 kcal/mol). Aplicando el cálculo de energía de unión mediante el enfoque LIE es posible disminuir la cantidad de ligandos falsos positivos que en muchas ocasiones el acoplamiento molecular no es capaz de discriminar.

**Tabla 3.** Energía libre de unión a HsEFL1 obtenida para cada ligando empleando el método LIE.

No.	Compuesto	LIE $\Delta G_{unión}$ kJ/mol (DS)
1	ZINC000096023887	-17.550 (6.779)
2	ZINC000257470512	-22.074 (7.902)
3	ZINC000015070573	-20.019 (6.800)
4	ZINC000077291909	-5.978 (5.173)
5	ZINC000004692041	26.482 (11.186)
6	ZINC000253532104	16.589 (15.699)
7	ZINC000036052121	-7.384 (5.115)
8	ZINC000077312672	1.591 (14.724)
9	ZINC000040164488	26.541 (12.811)
10	ZINC000008214538	-16.587 (6.504)
11	ZINC000253534724	6.833 (14.680)
12	ZINC000085946247	-30.460 (7.539)
13	ZINC000257459537	-13.453 (5.836)
14	ZINC000038139805	-8.225 (4.617)
15	ZINC000035465157	-13.330 (5.521)
16	ZINC000253534722	31.643 (12.993)
17	ZINC000003295242	-46.721 (13.193)

18	ZINC000253532106	3.708 (12.170)
19	ZINC000253532108	-11.348 (9.901)
20	ZINC000040163012	-9.134 (5.309)

**Tabla 4.** Energía libre de unión a ScEfl1 obtenida para cada ligando empleando el método LIE.

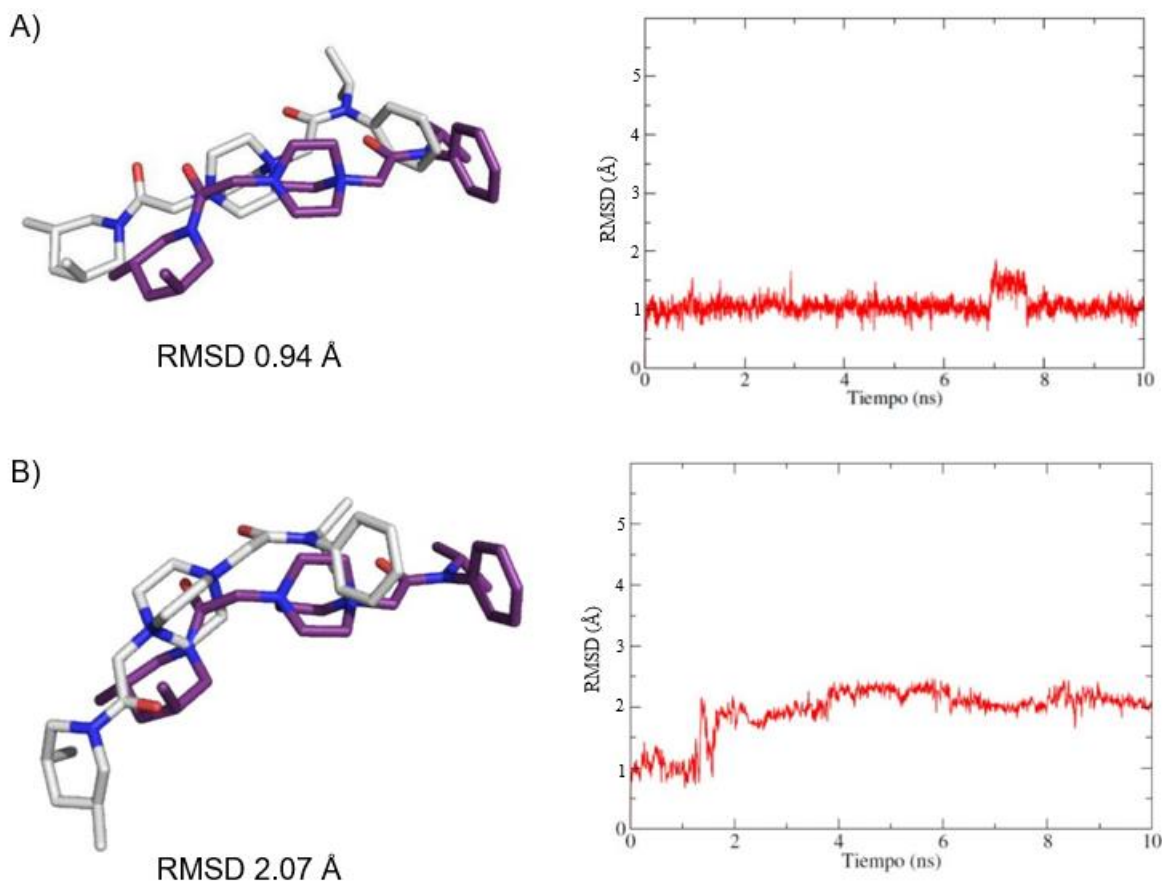
No.	Compuesto	LIE $\Delta G_{unión}$ kJ/mol (DS)
1	ZINC000096023887	-5.360 (5.720)
2	ZINC000257470512	-2.371 (5.482)
3	ZINC000015070573	-16.088 (11.365)
4	ZINC000077291909	-1.682 (4.898)
5	ZINC000004692041	21.054 (11.262)
6	ZINC000253532104	12.546 (13.211)
7	ZINC000036052121	-8.685 (5.144)
8	ZINC000077312672	21.024 (14.162)
9	ZINC000040164488	24.933 (13.342)
10	ZINC000008214538	-9.252 (6.138)
11	ZINC000253534724	6.389 (5.421)
12	ZINC000085946247	-2.140 (6.461)
13	ZINC000257459537	-1.777 (4.779)
14	ZINC000038139805	-15.246 (5.268)
15	ZINC000038140888	-3.139 (4.594)
16	ZINC000033604951	2.116 (5.610)
17	ZINC000247930162	-3.039 (5.343)
18	ZINC000096102535	-13.884 (4.646)
19	ZINC000252515685	-2.544 (6.286)
20	ZINC000254045588	3.049 (5.017)

Los complejos de los ligandos con los  $\Delta G_{unión}$  más negativos de acuerdo con la metodología LIE fueron seleccionados para evaluar la estabilidad conformacional del esqueleto de la proteína en el complejo, con simulaciones a 310 K y bajo *simulated annealing*.

### 5.3.2 Complejo HsEFL1-ligando17

Dado que el ligando 17 obtuvo el puntaje de energía de interacción más negativo con HsEFL1, se investigó la estabilidad que presentaba la conformación tanto de ligando como del esqueleto proteico en el complejo HsEFL1-ligando 17. El RMSD del ligando fue muy estable a lo largo de la simulación a 310 K, su conformación sólo varió 0.94 Å respecto a la

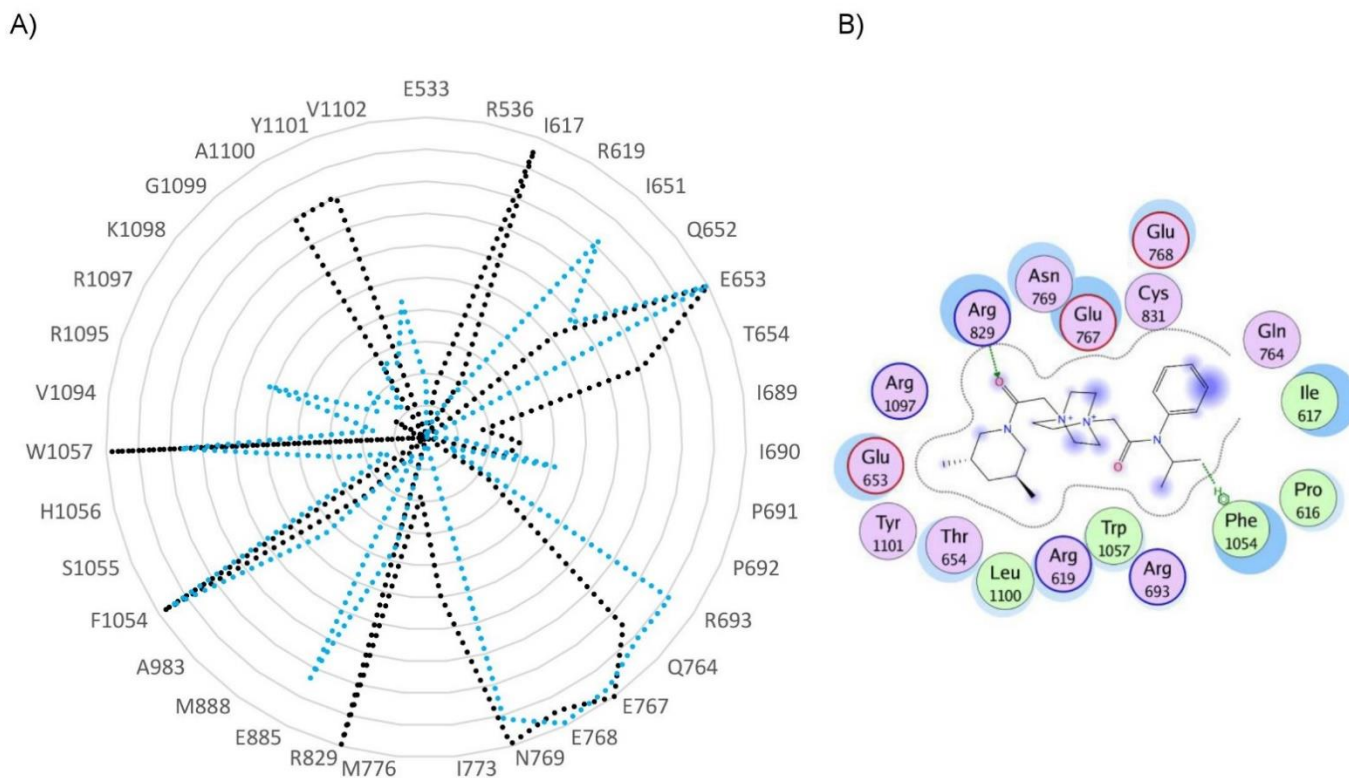
conformación predicha por el acoplamiento (Figura 21). En cambio bajo *simulated annealing* la conformación del ligando se modifica 2.07 Å respecto a la conformación de inicio. A pesar de ello, la conformación del ligando se estabiliza después de los 4 ns manteniéndose constante hasta el final de la simulación.



**Figura 21.** Comparación de la conformación final (blanco) contra la inicial (morado) del ligando 17 en complejo obtenida del acoplamiento entre los dominios III al V en la conformación traslocada de HsEFL1 a 310 K (A) y empleando *simulated annealing*. (B) Gráficas de RMSD de los átomos pesados del ligando.

Posteriormente se evaluaron las frecuencias de contacto mayores a 5.0 %, con los residuos de HsEFL1 que se encontraban a 4 Å del ligando 17 durante el total del tiempo de simulación (Figura 22). Tanto en la simulación a 310 K como en *simulated annealing* los contactos con la E653, E767, y F1054 se mantienen constantes; así mismo otros residuos con alta frecuencia de contacto son Q764, E768, N769 y W1057. La diferencia en la frecuencia de contacto con los residuos entre la simulación a 310 K y la realizada bajo

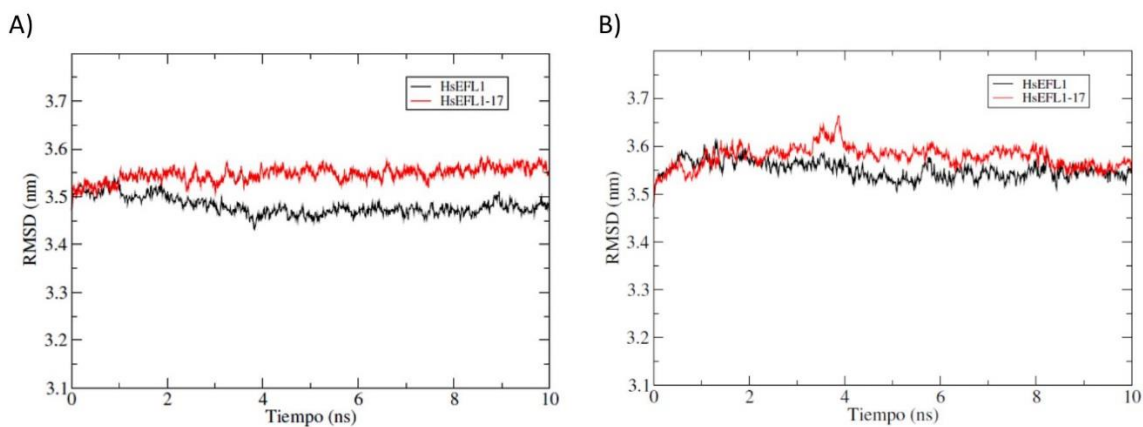
*simulated annealing* se debe a que el ligando 17 adopta una nueva conformación dentro del sitio de unión en esta última simulación (Figura 21 B).



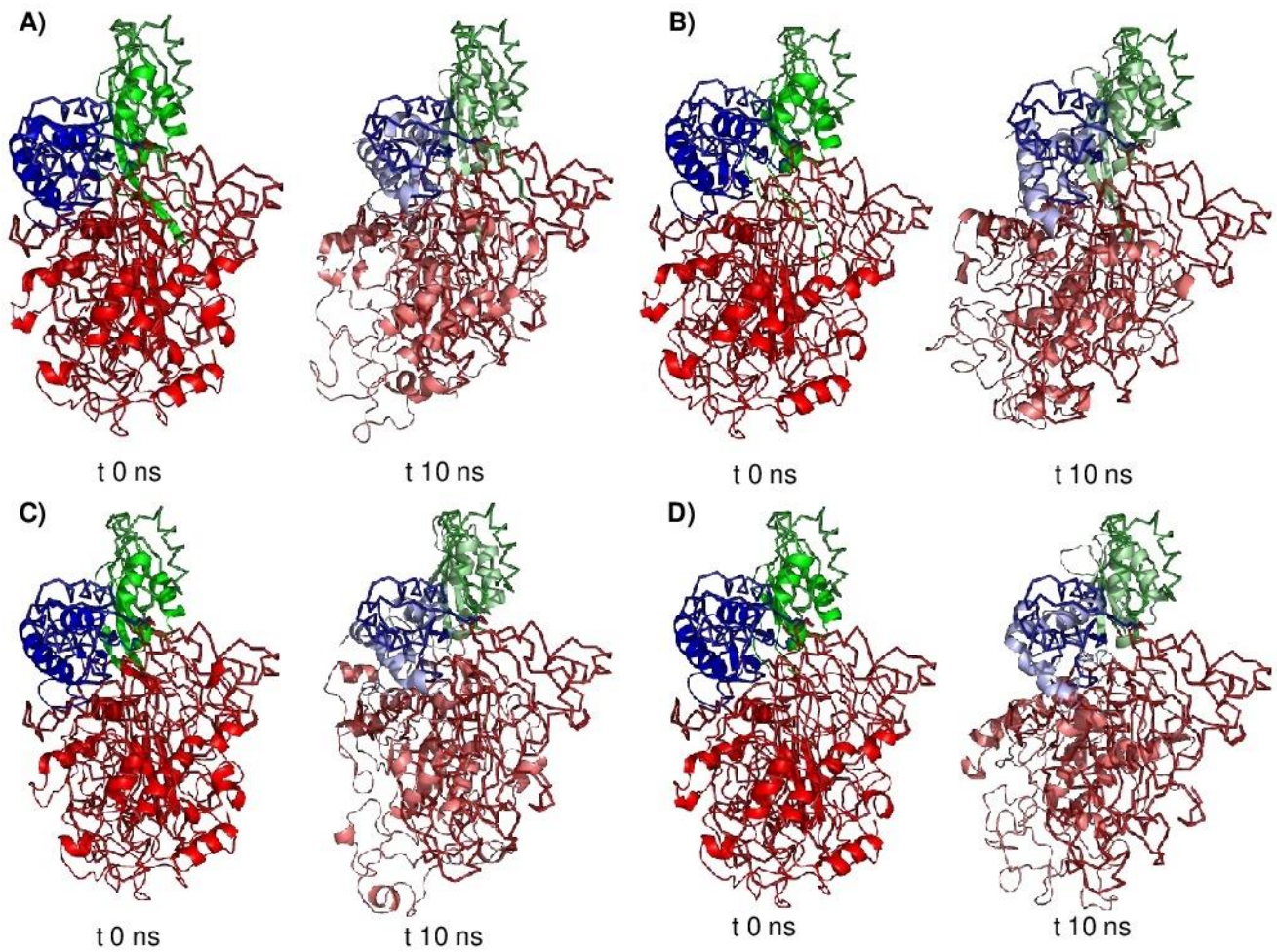
**Figura 22.** A) Frecuencia de contacto de residuos en HsEFL1 que se encontraban a 4 Å de sordarina durante la simulación a 310 K (negro), *simulated annealing* (azul) donde cada círculo representa un incremento de 10% de probabilidad de contacto. B) Diagrama de interacción de la pose de unión inicial para el ligando 17 con HsEFL1.

La transición de HsEFL1 desde su conformación traslocada hacia la conformación no traslocada no es evidente aún a los 10 ns de simulación a 310 K o utilizando *simulated annealing* (Figura 23). Una vez realizada la superposición de los dominios I y II de HsEFL1 en la conformación inicial y al final de la simulación a 310 K el RMSD de los dominios III al V al inicio de la simulación, comparados con su conformación en la estructura no traslocada de HsEFL1 (representación de listón) fue de 10.99 Å. Luego al comparar la conformación los dominios III al V con la estructura no traslocada de HsEFL1, el RMSD fue de 17.41 Å. Mientras que al comparar la conformación al finalizar la simulación del complejo HsEFL1-17 a 310 K con su conformación en la estructura no traslocada el RMSD fue 20.09 Å. Utilizando *simulated annealing* en ausencia de ligando el RMSD del esqueleto de la proteína calculado para la conformación final de los dominios III al V comparados con la del estado

no traslocado resultó de 17.63 Å. Mientras que al comparar la conformación del dominio III después de 10ns de simulación del complejo con el ligando 17 con su conformación en la estructura no traslocada el RMSD es 20.84 Å (Figura 24). Dado que las diferencias en el RMSD al final de las simulaciones tomando como referencia la estructura no traslocada de HsEFL1 son casi idénticas en ambas simulaciones, parece que el *simulated annealing* no acelera la evolución en la conformación de este sistema. Aunque en el caso del complejo con el ligando el valor del RMSD es mayor, lo que indica que en el complejo los dominios III a V poseen una conformación menos similar a la que existe en la estructura no traslocada de HsEFL1.



**Figura 23.** Transición de la conformación del esqueleto de HsEFL1 calculado para los átomos N, C $\alpha$ , C, O desde el estado traslocado hacia el no traslocado a 310 K (A) y empleando *simulated annealing* (B) en ausencia (negro) o presencia de ligando (rojo).

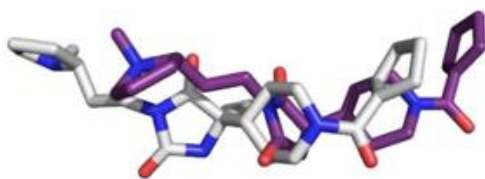


**Figura 24.** Superposición de HsEFL1 alineando los dominios I y II (no visibles), se representa como cartoon los dominios III en azul, IV en rojo y V en verde a los tiempos 0 (tonos oscuros) y 10 ns (tonos claros) con la conformación no traslocada en representación de listón. En ausencia (A, C) o presencia (B, D) del ligando 17 durante la simulación a 310 K (A y B) y bajo el esquema de *simulated annealing* (C y D) desde la conformación traslocada.

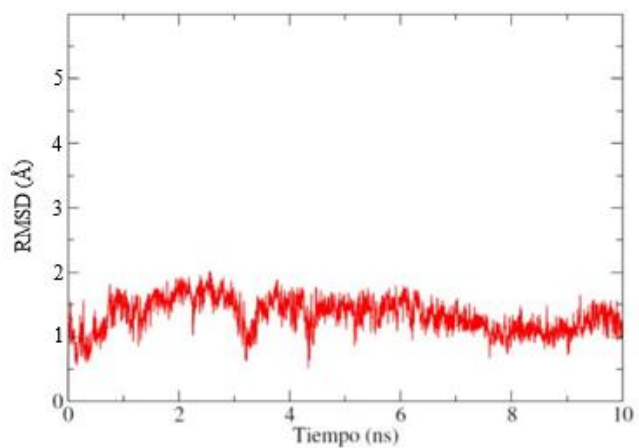
### 5.3.3 Complejo ScEfl1-ligando 03

Puesto que el ligando 3 obtuvo el puntaje de energía de interacción más favorable con ScEfl1 empleando la metodología LIE, se indagó la estabilidad conformacional tanto del ligando como del esqueleto proteico en el complejo ScEfl1-ligando 03. De acuerdo con el gráfico de RMSD del ligando en el complejo ScEfl1-ligando 03 la diferencia de la conformación del este comparado con su pose inicial no es significativa, ya que fue menor a 2 Å durante la simulación a 310 K. Hacia el final de la simulación su conformación se modificó a 1.16 Å respecto a su conformación de inicio predicha por el acoplamiento. Así mismo durante la simulación empleando *simulated annealing* el cambio en la conformación del ligando no superó los 2 Å respecto a la pose inicial (Figura 25). Así mismo, la frecuencia de contacto con los residuos de ScEfl1 que se encontraban a 4 Å del ligando, cuyo valor superaba el 5.0 % a lo largo de la simulación se aprecian en la (Figura 26). Tanto en la simulación a 310 K como usando *simulated annealing* los contactos con la P677, I678, Y711 y E713 se mantienen constantes; así mismo otros residuos con alta frecuencia de contacto son I719, C721, Q1042, L1043 y F1045. La diferencia en la frecuencia de contacto de los residuos en ScEfl1 con el ligando 3 entre la simulación a 310 K y con *simulated annealing* son menores que las observadas para HsEFL1 con el ligando 17. Pues como se ha mencionado el ligando 3 no sufre fluctuaciones mayores a 2 Å en su RMSD hacia el final de las simulaciones, manteniendo aproximadamente los mismos contactos entre ambas simulaciones durante la producción de estas.

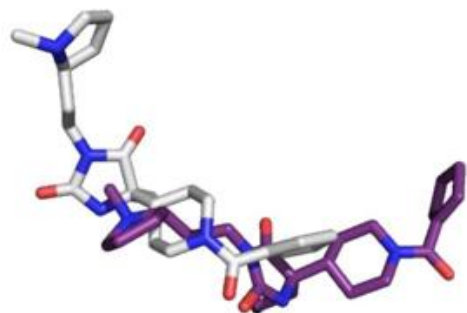
A)



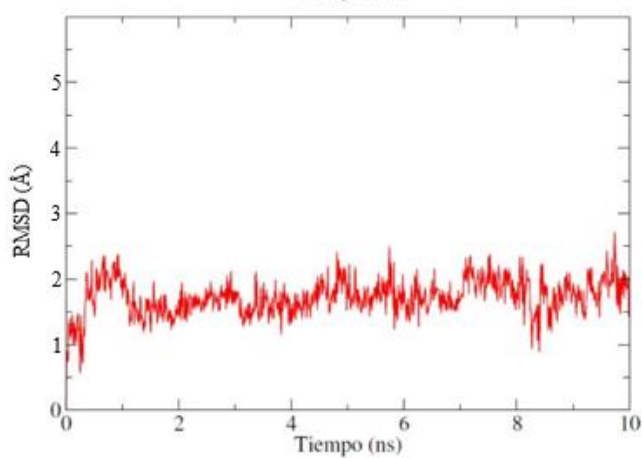
RMSD 1.16 Å



B)



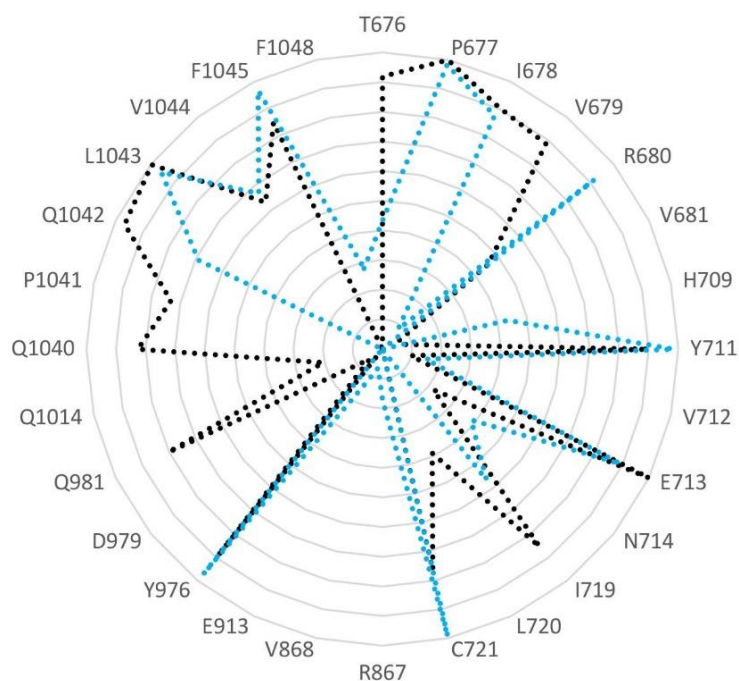
RMSD 1.76 Å



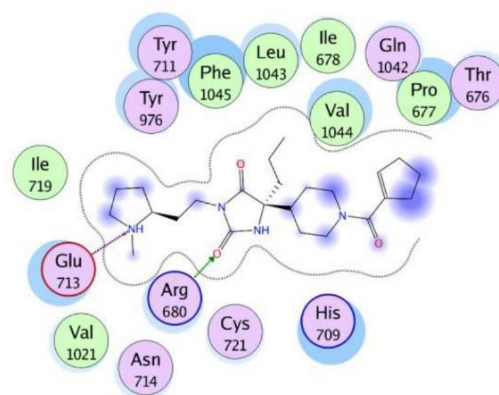
**Figura 25.** Comparación de la conformación final (blanco) contra la inicial (morado) del ligando 17 en complejo obtenida del acoplamiento entre los dominios III al V en la conformación traslocada de ScEFL1 a 310 K (A) y empleando *simulated annealing*. (B) Gráficas de RMSD de los átomos pesados del ligando.



A)

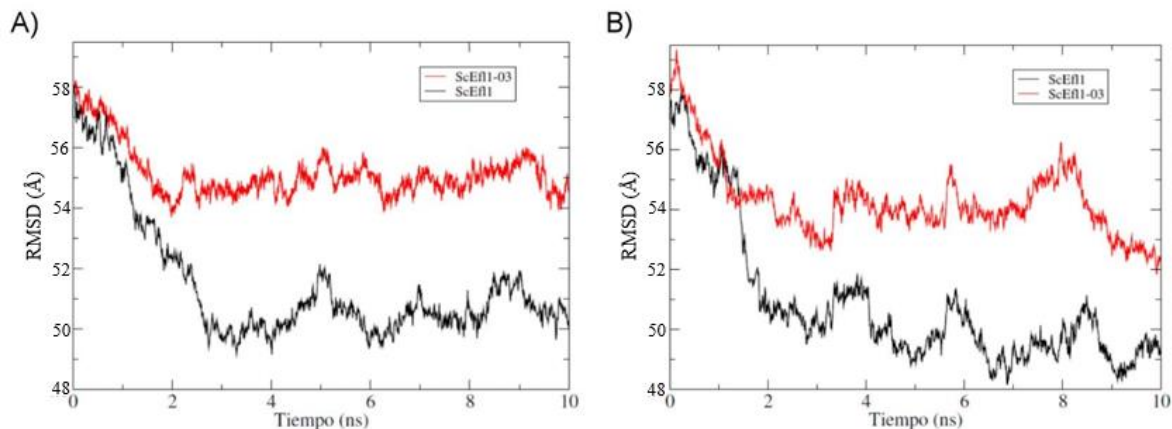


B)



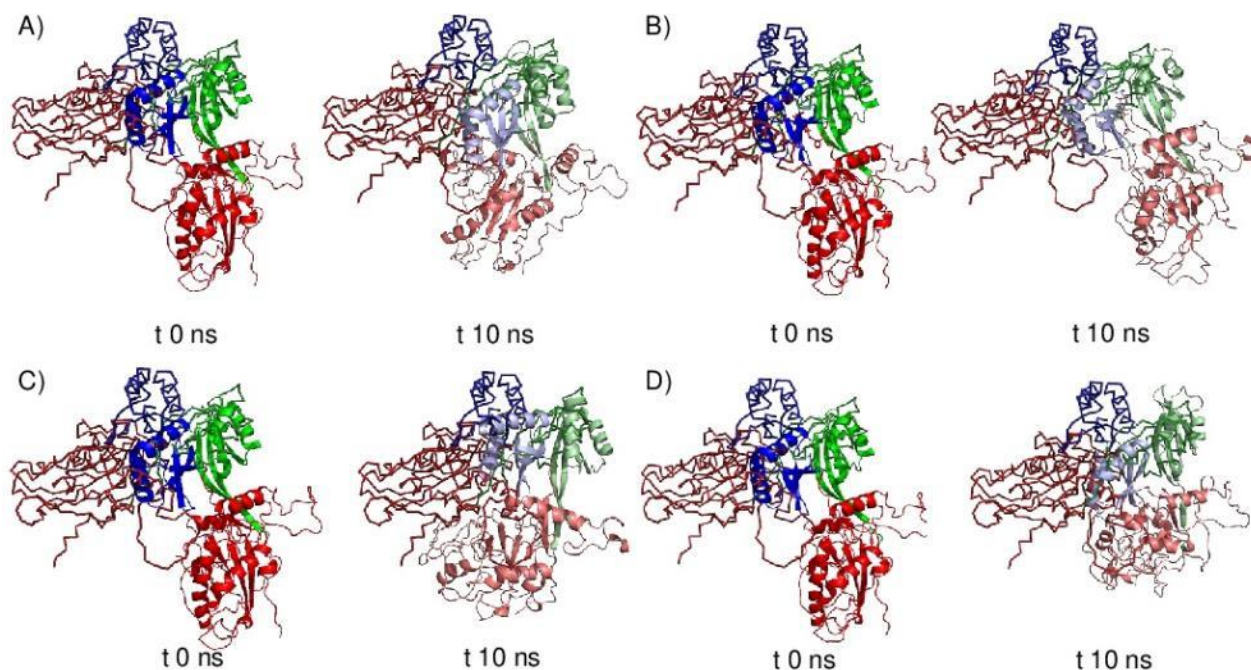
**Figura 26.** A) Frecuencia de contacto de residuos en ScEfl1 que se encontraban a 4 Å del ligando 17 durante la simulación a 310 K (negro), *simulated annealing* (azul) donde cada círculo representa un incremento de 10% de probabilidad de contacto. B) Diagrama de interacción de la pose de unión inicial para el ligando 3 con ScEfl1.

La transición en la conformación de ScEfl1 es muy notoria durante los primeros 3 ns de simulación particularmente en el esqueleto de la proteína en ausencia del ligando 3 representado por la línea negra en el gráfico de la Figura 27. Además el cambio es más pronunciado en las simulaciones bajo *simulated annealing* debido a los ciclos de calentamiento y enfriamiento que sufre el sistema. El esqueleto de ScEfl1 en complejo con el ligando 3 mantiene una conformación más alejada de la que corresponde al estado no traslocado, denotando una mayor estabilidad conformacional que cuando las coordenadas del ligando han sido removidas.



**Figura 27.** Transición de la conformación del esqueleto de ScEf11 calculado para los átomos N, C $\alpha$ , C, O desde el estado traslocado hacia el no traslocado a 310 K y empleando *simulated annealing* (B) en ausencia (negro) o presencia de ligando (rojo).

La superposición estructural de los dominios I y II de ScEf11 de la conformación al inicio y final de la simulación a 310 K en ausencia del ligando 3 se muestran en la Figura 28. Se calculó que el RMSD de los dominios III al V al final de la simulación a 310 K comparados con la conformación en la estructura no traslocada de ScEf11 (representación de listón) es de 38.74 Å. Cuando al principio de la simulación el RMSD de dichos dominios respecto a su conformación en el estado no traslocado fue de 43.26 Å (Figura 28 A). Mientras que al comparar la conformación al finalizar la simulación del complejo ScEf11-03 a 310 K con su conformación en la estructura no traslocada el RMSD fue 51.23 Å (Figura 28 B). Es decir, la diferencia en el RMSD del esqueleto de ScEf11 tomando como referencia la estructura no traslocada es menor cuando no existe el ligando, por lo que la conformación de los dominios III al V es más similar a la conformación que poseen en la estructura no traslocada.



**Figura 28.** Superposición de ScEfl1 alineando los dominios I y II (no visibles), se representa como cartoon los dominios III en azul, IV en rojo y V en verde a los tiempos 0 (tonos oscuros) y 10 ns (tonos claros) con la conformación no traslocada en representación de listón. En ausencia (A, C) o presencia (B, D) del ligando 3 durante la simulación a 310 K (A y B) y bajo el esquema de *simulated annealing* (C y D) desde la conformación traslocada de ScEfl1.

Se puede notar que en muchos casos la información obtenida del *docking* no corresponde con la observada en las simulaciones de MD, pues compuestos cuyo puntaje de interacción parecía favorable en el *docking* mostraron baja estabilidad en el sitio de unión de la proteína, así como una energía de unión desfavorable calculada mediante la metodología LIE. En el caso del acoplamiento los cálculos son sólo aproximados, pues normalmente se hacen en el vacío, además se omiten los efectos entrópicos y de solvatación. Además mediante el *docking* se intenta reproducir potenciales químicos que determinan la conformación preferida de unión al receptor asumiendo que este es rígido. Es importante emplear herramientas como MD ya que permite la inclusión explícita de disolvente en el sistema, así como la exploración del espacio conformacional del receptor. Sin embargo aún empleando este enfoque existen problemas adicionales inherentes a las limitaciones del método, como la inexacta reproducción de los efectos electrostáticos, la incertidumbre de la constante

dieléctrica del medio que podría apantallar la intensidad de la interacción ligando-proteína, la convergencia al mínimo global, etc. Además una ventaja de emplear MD es que se pueden obtener valores promediados estadísticamente de la trayectoria de MD para cualquier propiedad del sistema que sea medible directamente en cada punto de la misma.

## **5.4 Resultados experimentales**

En esta sección se presentan los resultados experimentales de los ensayos de desplazamiento térmico monitoreado por fluorescencia. Realizados como prueba en la búsqueda de una técnica que identificara la interacción con moléculas orgánicas pequeñas empleando ligandos conocidos de las GTPasas EFL1.

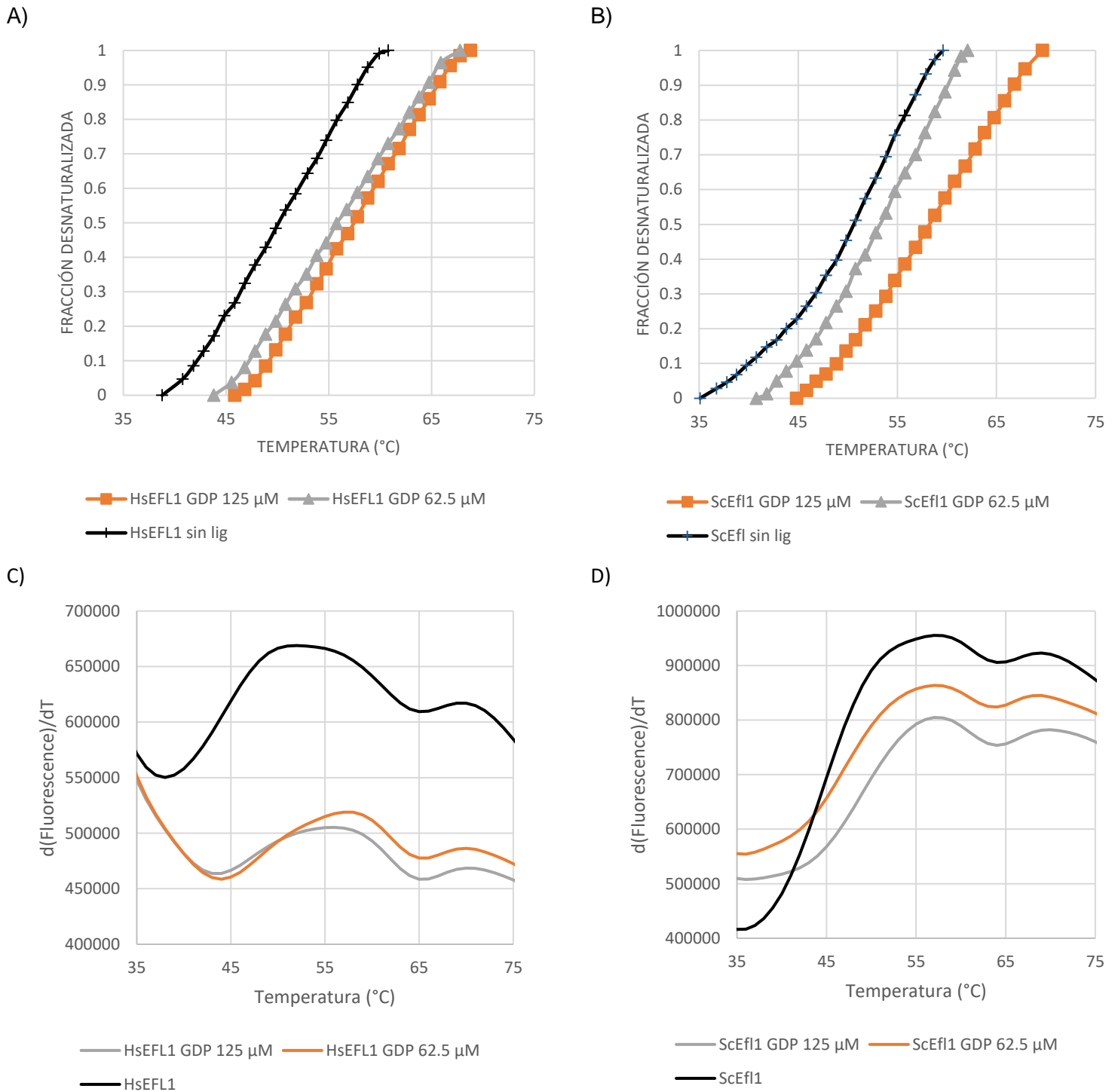
### **5.4.1 Purificación de proteínas ScEfl1 y HsEFL1**

Para evaluar la interacción de los ligandos es necesario aislar las proteínas ScEfl1 y HsEFL1, esto se logra mediante la sobreexpresión de proteína recombinante. A continuación, se muestran los resultados obtenidos para la purificación de las proteínas ScEfl1 y HsEFL1, siguiendo la metodología descrita en la sección 4.2.1. Las fracciones colectadas durante la cromatografía de afinidad a Ni<sup>2+</sup> se analizaron mediante SDS-PAGE (Figura 38 y Figura 39 Anexo 9.8). En estos geles es posible observar que el carril de la elución presenta una banda correspondiente a una proteína con peso molecular aproximado de 135 kDa, que es cercano al de la proteína ScEfl1 de 127.3 kDa y HsEFL1 de 127.5 kDa. Una vez realizada la electroforesis de las fracciones provenientes de la cromatografía de afinidad, la fracción eluída se concentró hasta un volumen final de 1 mL y se efectuó cromatografía de exclusión molecular que se analizó mediante electroforesis SDS-PAGE. Finalmente, las fracciones que contenían la proteína correspondiente ScEfl1 (Figura 40 Anexo 9.8) y HsEFL1 (Figura 41 Anexo 9.8), se concentraron hasta un volumen final de 1 mL, que fue almacenado en alícuotas de 100 µL a -70°C.

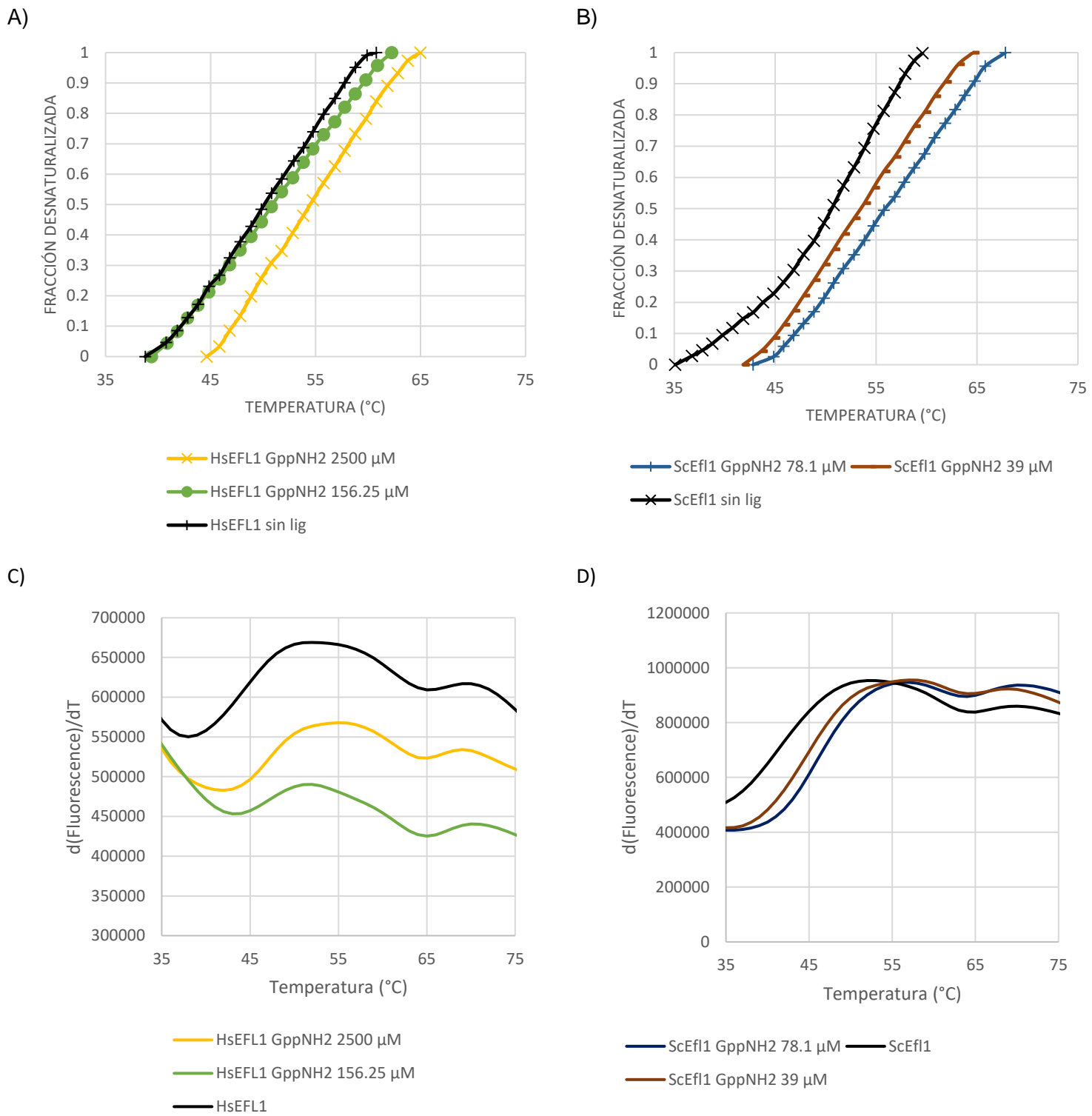
### **5.4.2 Desplazamiento térmico monitoreado por fluorescencia (Thermofluor)**

Los ensayos de desplazamiento térmico monitoreado por fluorescencia se realizaron en un sistema de detección en tiempo real para PCR según lo descrito en la sección 4.2.2. Se obtuvieron las curvas de fusión de las proteínas HsEFL1-TevHis y ScEfl1-TevHis para el ligando conocido guanosina difosfato (GDP) (Figura 29) y su análogo no hidrolizable guanosina-5'-β-amino-difosfato (GppNH<sub>2</sub>) (Figura 30) empleando concentraciones máximas saturantes de los ligandos ( $K_d$  ScEfl1-GDP ≈ 10 µM,  $K_d$  ScEfl1-GppNH<sub>2</sub> ≈ 677 µM). Cuando la fracción

desnaturalizada de la proteína es 0.5 representa el valor de la temperatura de fusión ( $T_m$ ) para el ensayo.



**Figura 29.** Curvas de fusión de la proteína HsEFL1-TevHis (A, C) y ScEfl1-TevHis (B, D) empleando distintas concentraciones de GDP.



**Figura 30.** Curvas de fusión de la proteína HsEFL1-TevHis (A, C) y ScEfl1-TevHis (B, D) empleando distintas concentraciones del análogo no hidrolizable GppNH<sub>2</sub>.

Como se muestra en las tablas los valores de  $T_m$  obtenidos para ScEfl1 y HsEFL1 parecen coincidir con la  $T_m$  de la primera transición en la desnaturalización térmica de ScEfl1 obtenida mediante *Differential Scanning Calorimetry* (DSC) en trabajos previos de nuestro grupo de investigación cuyo valor corresponde a 54°C <sup>19</sup>. La unión con otras moléculas tiende a estabilizar la proteína, por lo que el desplazamiento de la  $T_m$  hacia valores superiores sugiere que existe interacción entre el compuesto analizado y proteína. Lo cual se observa con los compuestos GDP y GppNH<sub>2</sub> en las GTPasas EFL1. La técnica de thermofluor parece ser una herramienta útil para evaluar la unión de los ligandos seleccionados a través del escrutinio virtual con las GTPasas EFL1.

## 6 Conclusiones

- La ausencia de moléculas de agua en el sitio de unión a sordarina en los dominios III, IV y V de EF-2 no influyen sobre la pose de unión de sordarina predicha por acoplamiento molecular pero sí sobre el valor del puntaje de la interacción. Esto se debe a que existen interacciones con residuos más lejanos que están mediadas por el agua.
- Los ligandos con  $\Delta G_{unión}$  más favorable predicho mediante la metodología LIE presentan poca fluctuación en el valor de su RMSD a lo largo de la simulación de dinámica molecular.
- El ligando 3 estabiliza la conformación traslocada de ScEfl1 impidiendo la transición hacia su conformación no traslocada.
- El ligando 17 estabiliza la conformación traslocada de HsEFL1 e impide su transición hacia la conformación no traslocada.
- Los residuos más importantes para la interacción entre ligando y proteína mantienen contacto constante en la simulación a 310 K y con *simulated annealing*. Esto sugiere que se favorecerá la formación del complejo a través de dichos residuos.
- El *docking* puede ser empleado como un método relativamente rápido de cribado virtual, sin embargo mediante el uso de MD en este trabajo fue posible identificar falsos positivos obtenidos por dicho enfoque. Se pueden mencionar el ligando 16 en HsEFL1 y el ligando 9 en ScEfl1 cuyos puntajes de interacción obtenidos en *docking* resultaron favorables. Pero una vez calculado mediante la metodología LIE usando MD, que considera adicionalmente la flexibilidad tanto de las proteínas como los ligandos, así como los efectos entrópicos y de solvatación, se encontró que no se predice una interacción favorable.



## 7 Perspectivas

- Evaluar la estabilidad conformacional del esqueleto proteico en simulaciones de MD empleando *simulated annealing* para el resto de los complejos de HsEFL1 con los ligandos para los cuales se obtuvieron  $\Delta G_{unión}$  negativas de acuerdo con la metodología LIE.
- Evaluar la estabilidad conformacional del esqueleto proteico en simulaciones de MD empleando *simulated annealing* para el resto de los complejos de ScEfl1 con los ligandos para los cuales se obtuvieron  $\Delta G_{unión}$  negativas de acuerdo con la metodología LIE.
- Evaluar la estabilidad de los ligandos en el sitio de unión entre los dominios III al V, para los cuales se obtuvieron  $\Delta G_{unión}$  negativas de acuerdo con la metodología LIE, en simulaciones de MD empleando *simulated annealing* en complejo con HsEFL1.
- Evaluar la estabilidad de los ligandos en el sitio de unión entre los dominios III al V, para los cuales se obtuvieron  $\Delta G_{unión}$  negativas de acuerdo con la metodología LIE, en simulaciones de MD empleando *simulated annealing* en complejo con ScEfl1.
- Realizar ensayos thermofluor con los ligandos con valores de  $\Delta G_{unión}$  favorables que posean estabilidad en el sitio de unión y se prediga que estabilizan la conformación traslocada de HsEFL1.
- Realizar ensayos thermofluor con los ligandos con valores de  $\Delta G_{unión}$  favorables que posean estabilidad en el sitio de unión y se prediga que estabilizan la conformación traslocada de ScEfl1.

## 8 Bibliografía

1. Jenner, L. *et al.* Crystal structure of the 80S yeast ribosome. *Curr. Opin. Struct. Biol.* **22**, 759–767 (2012).
2. Menne, T. F. *et al.* The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nat. Genet.* **39**, 486–495 (2007).
3. Woolford, J. L. & Baserga, S. J. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* **195**, 643–681 (2013).
4. Kemmler, S., Occhipinti, L., Veisu, M. & Panse, V. G. Yvh1 is required for a late maturation step in the 60S biogenesis pathway. *J. Cell Biol.* **186**, 863–880 (2009).
5. Lo, K.-Y., Li, Z., Wang, F., Marcotte, E. M. & Johnson, A. W. Ribosome stalk assembly requires the dual-specificity phosphatase Yvh1 for the exchange of Mrt4 with P0. *J. Cell Biol.* **186**, 849–862 (2009).
6. Panse, V. G. & Johnson, A. W. Maturation of eukaryotic ribosomes: acquisition of functionality. *Trends Biochem. Sci.* **35**, 260–266 (2010).
7. Gerhardy, S., Menet, A. M., Peña, C., Petkowski, J. J. & Panse, V. G. Assembly and nuclear export of pre-ribosomal particles in budding yeast. *Chromosoma* **123**, 327–344 (2014).
8. Nakhoul, H. *et al.* Ribosomopathies: Mechanisms of Disease. *Clin. Med. Insights Blood Disord.* **7**, 7 (2014).
9. Narla, A. & Ebert, B. L. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**, 3196–3205 (2010).
10. Shwachman, H., Diamond, L. K., Oski, F. A. & Khaw, K. T. THE SYNDROME OF PANCREATIC INSUFFICIENCY AND BONE MARROW DYSFUNCTION. *J. Pediatr.* **65**, 645–663 (1964).
11. Ganapathi, K. A. & Shimamura, A. Ribosomal dysfunction and inherited marrow failure. *Br. J. Haematol.* **141**, 376–387 (2008).
12. Burroughs, L., Woolfrey, A. & Shimamura, A. Shwachman Diamond Syndrome – a review of the clinical presentation, molecular pathogenesis, diagnosis, and treatment. *Hematol. Oncol. Clin. North Am.* **23**, 233–248 (2009).
13. Goobie, S. *et al.* Shwachman-Diamond syndrome with exocrine pancreatic dysfunction and bone marrow failure maps to the centromeric region of chromosome 7. *Am. J. Hum. Genet.* **68**, 1048–1054 (2001).
14. Boocock, G. R. B. *et al.* Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat. Genet.* **33**, 97–101 (2003).
15. Woloszynek, J. R. *et al.* Mutations of the SBDS gene are present in most patients with Shwachman-Diamond syndrome. *Blood* **104**, 3588–3590 (2004).
16. Austin, K. M., Leary, R. J. & Shimamura, A. The Shwachman-Diamond SBDS protein localizes to the nucleolus. *Blood* **106**, 1253–1258 (2005).
17. Zhang, S., Shi, M., Hui, C.-C. & Rommens, J. M. Loss of the mouse ortholog of the shwachman-diamond syndrome gene (*Sbds*) results in early embryonic lethality. *Mol. Cell. Biol.* **26**, 6656–6663 (2006).
18. García-Márquez, A., Gijsbers, A., de la Mora, E. & Sánchez-Puig, N. Defective Guanine Nucleotide Exchange in the Elongation Factor-like 1 (EFL1) GTPase by Mutations in the Shwachman-Diamond Syndrome Protein. *J. Biol. Chem.* **290**, 17669–17678 (2015).
19. Gijsbers, A., García-Márquez, A., Luviano, A. & Sánchez-Puig, N. Guanine nucleotide exchange in the ribosomal GTPase EFL1 is modulated by the protein mutated in the Shwachman-Diamond syndrome. *Biochem. Biophys. Res. Commun.* **437**, 349–354 (2013).
20. Senger, B. *et al.* The nucle(ol)ar Tif6p and Efl1p are required for a late cytoplasmic step of ribosome synthesis. *Mol. Cell* **8**, 1363–1373 (2001).

21. Ceci, M. *et al.* Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579–584 (2003).
22. Weis, F. *et al.* Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nat. Struct. Mol. Biol.* **22**, 914–919 (2015).
23. Wittinghofer, A. & Vetter, I. R. Structure-function relationships of the G domain, a canonical switch motif. *Annu. Rev. Biochem.* **80**, 943–971 (2011).
24. Asano, N. *et al.* Direct interaction between EFL1 and SBDS is mediated by an intrinsically disordered insertion domain. *Biochem. Biophys. Res. Commun.* **443**, 1251–1256 (2014).
25. Stepensky, P. *et al.* Mutations in EFL1, an SBDS partner, are associated with infantile pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in a Shwachman-Diamond like syndrome. *J. Med. Genet.* (2017). doi:10.1136/jmedgenet-2016-104366
26. Spahn, C. M. T. *et al.* Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J.* **23**, 1008–1019 (2004).
27. Jørgensen, R. *et al.* Two crystal structures demonstrate large conformational changes in the eukaryotic ribosomal translocase. *Nat. Struct. Biol.* **10**, 379–385 (2003).
28. Finch, A. J. *et al.* Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes Dev.* **25**, 917–929 (2011).
29. Scior, T. *et al.* Recognizing pitfalls in virtual screening: a critical review. *J. Chem. Inf. Model.* **52**, 867–881 (2012).
30. Klopmand, G. Concepts and applications of molecular similarity, by Mark A. Johnson and Gerald M. Maggiora, eds., John Wiley & Sons, New York, 1990, 393 pp. Price: \$65.00. *J. Comput. Chem.* **13**, 539–540 (2004).
31. Xue, L. & Bajorath, J. Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. *Comb. Chem. High Throughput Screen.* **3**, 363–372 (2000).
32. Cheng, T., Li, Q., Zhou, Z., Wang, Y. & Bryant, S. H. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *AAPS J.* **14**, 133–141 (2012).
33. Cavasotto, C. N. Homology models in docking and high-throughput docking. *Curr. Top. Med. Chem.* **11**, 1528–1534 (2011).
34. Meng, X.-Y., Zhang, H.-X., Mezei, M. & Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **7**, 146–157 (2011).
35. Brint, A. T. & Willett, P. Algorithms for the identification of three-dimensional maximal common substructures. *J. Chem. Inf. Comput. Sci.* **27**, 152–158 (1987).
36. DesJarlais, R. L., Sheridan, R. P., Dixon, J. S., Kuntz, I. D. & Venkataraghavan, R. Docking flexible ligands to macromolecular receptors by molecular shape. *J. Med. Chem.* **29**, 2149–2153 (1986).
37. Leach, A. R. & Kuntz, I. D. Conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comput. Chem.* **13**, 730–748 (1992).
38. Hart, T. N. & Read, R. J. A multiple-start Monte Carlo docking method. *Proteins* **13**, 206–222 (1992).
39. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727–748 (1997).
40. Morris, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).
41. Weiner, S. J. *et al.* A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).
42. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).

43. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **3**, 935–949 (2004).
44. Böhm, H. J. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput. Aided Mol. Des.* **12**, 309–323 (1998).
45. Head, R. D. *et al.* VALIDATE: A New Method for the Receptor-Based Prediction of Binding Affinities of Novel Ligands. *J. Am. Chem. Soc.* **118**, 3959–3969 (1996).
46. Jain, A. N. Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J. Comput. Aided Mol. Des.* **10**, 427–440 (1996).
47. Ishchenko, A. V. & Shakhnovich, E. I. SMoG2001: an improved knowledge-based scoring function for protein-ligand interactions. *J. Med. Chem.* **45**, 2770–2780 (2002).
48. Feher, M., Deretey, E. & Roy, S. BHB: a simple knowledge-based scoring function to improve the efficiency of database screening. *J. Chem. Inf. Comput. Sci.* **43**, 1316–1327 (2003).
49. Becker, O. M. *Computational biochemistry and biophysics*. (New York: M. Dekker, c2001, 2001).
50. Linear interaction energy: method and applications in drug design. - PubMed - NCBI. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/22183545>. (Accessed: 26th November 2017)
51. Aqvist, J., Medina, C. & Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **7**, 385–391 (1994).
52. Almlöf, M., Brandsdal, B. O. & Åqvist, J. Binding affinity prediction with different force fields: Examination of the linear interaction energy method. *J. Comput. Chem.* **25**, 1242–1254 (2004).
53. Luzhkov, V. B. & Aqvist, J. Mechanisms of tetraethylammonium ion block in the KcsA potassium channel. *FEBS Lett.* **495**, 191–196 (2001).
54. Osterberg, F. & Aqvist, J. Exploring blocker binding to a homology model of the open hERG K<sup>+</sup> channel using docking and molecular dynamics methods. *FEBS Lett.* **579**, 2939–2944 (2005).
55. Bjelic, S. *et al.* Computational inhibitor design against malaria plasmepsins. *Cell. Mol. Life Sci.* **64**, 2285–2305 (2007).
56. Åqvist, J. & Hansson, T. On the Validity of Electrostatic Linear Response in Polar Solvents. *J. Phys. Chem.* **100**, 9512–9521 (1996).
57. Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **49**, 377–389 (2009).
58. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinform. Oxf. Engl.* **29**, 845–854 (2013).
59. Malde, A. K. *et al.* An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **7**, 4026–4037 (2011).
60. Gijsbers, A., Nishigaki, T. & Sánchez-Puig, N. Fluorescence Anisotropy as a Tool to Study Protein-protein Interactions. *JoVE J. Vis. Exp.* e54640–e54640 (2016). doi:10.3791/54640
61. Lo, M.-C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* **332**, 153–159 (2004).
62. Liu, K., Feng, J. & Young, S. S. PowerMV: A Software Environment for Molecular Viewing, Descriptor Generation, Data Analysis and Hit Evaluation. *J. Chem. Inf. Model.* **45**, 515–522 (2005).

## 9 Anexo

### 9.1 Principales sitios de unión a ligando predichos por la herramienta SiteMap en HsEFL1 y ScEfl1 en conformación traslocada

**Tabla 5.** Principales sitios de unión a ligando predichos por SiteMap en la conformación no traslocada de HsEFL1.

Sitio	Site score <sup>a</sup>	Tamaño	Dscore <sup>b</sup>	Residuos				
				Dominio I	II	III	IV	V
1	1.028	228	1.04	97, 99, 100, 103, 129, 130, 132, 133	420, 589, 590, 593, 594, 595, 610, 616, 617, 619,	612, 613, 614, 615, 638, 639, 641, 642, 646, 647, 648,649, 650, 651, 653, 660, 662	691, 692, 693, 764, 767, 769, 832, 828, 829, 831, 833, 834, 878, 885, 886, 888, 1052, 1057	1053, 1054, 1055, 1056,
2	1.046	121	0.963	33, 46, 47, 50, 51, 55, 56, 57, 58, 59, 60, 61, 64, 73, 75, 92	418, 425, 426, 427, 428, 441, 442, 488, 489, 494, 495, 498, 567, 568, 569, 570			
3	1.039	77	1.065	241, 243, 247, 248, 251, 255, 296, 297, 299, 300, 301, 304, 318, 319, 321, 322, 325, 348, 349				
4	0.981	103	1.019		621, 623		742, 686, 739, 882, 883, 884, 967, 970, 971, 974, 1065, 1066, 1088, 1092, 1093, 1095, 1096, 1097, 1098, 1099, 1102, 1102, 1103, 1104, 1106, 1108, 1110, 1111, 1112, 1113, 1118, 1120	
5	0.852	56	0.83	27, 28, 29, 30, 67, 68, 96, 97, 126				1035, 1038, 1041, 1042, 1045,

**Tabla 6.** Principales sitios de unión a ligando predichos por SiteMap en la conformación traslocada de HsEFL1.

Sitio	Site score <sup>a</sup>	Tamaño	Dscore <sup>b</sup>	Residuos				
				Dominio I	II	III	IV	V
1	1.078	138	0.985			616, 617,619, 648, 653,654,	690, 691, 692, 693, 694, 764, 767, 768, 769, 771, 772, 773, 829, 831,832, 885, 886, 888, ,1030, 1054, 1057, 1097, 1098,1099	983, 985, ,1100,1101, 1095
2	1.242	101	1.28	22, 24, 34, 35, 38, 39, 42, 79, 86, 88, 114, 142, 144, 231, 233,235, 236, 241,242,243 , 246, 247, 250, 351,352, 353,354,355 ,356, 359,				
3	0.984	56	1.025	149, 152,153, 238, 239, 241, 301, 304, 305, 308, 338, 341				
4	0.835	53	0.803	189, 191,192,193 ,194,195, 268, 214, 215, 216, 218, ,225, 226, 227, 265, 269,				
5	0.809	60	0.779				729, 730, 732, 879, 883, 884, 908, 909, 966, 967, 970, 971, 1069, 1070, 1107,1108, 1104, 1105, 1106, 1120	

<sup>a</sup> El *SiteScore* promedio para 157 sitios submicromolares investigados es 1.0. Así, una puntuación de más de 1 sugiere un sitio de unión a ligando promisorio particular. La literatura sugiere que un *SiteScore* de 0.80 distingue con precisión entre los sitios de unión a fármacos y los que no poseen unión a fármacos

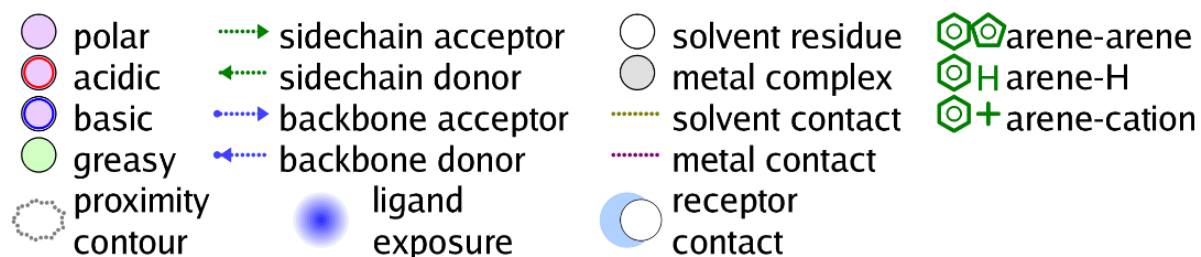
<sup>b</sup> *Druggability score (Dscore)*, para 22 de 67 sitios subnanomolares estudiados, los sitios con valores *Dscore* inferiores a 0.83 son clasificados como "no drogable", mientras que aquellos con valores *Dscore* entre 0.83 y 0.98 como "difícil" y aquellos que tienen mayores valores *Dscore* como "drogable".<sup>57</sup>

## 9.2 Clasificación de sitios de unión a ligando

**Tabla 7.** Características de Sitios "no drogables", difíciles, y "drogables" definidas para SiteMap en Halgren, T. A.<sup>57</sup>

Categoría	Características
No drogable	(a) fuertemente hidrófilo; relativamente pequeño en tamaño; poco o ningún carácter hidrófobo; o (b) requiere unión covalente; o (c) muy pequeño o muy poco profundo
Difícil	Suficientemente hidrófilo para requerir la administración como profármaco; menos hidrófobo que un sitio drogable típico
Drogable	Tamaño, reclusión e hidrofobicidad razonables con hidrofobicidad poco excepcional

## 9.3 Interacciones proteína-ligando obtenidos para el acoplamiento de los ligandos



**Figura 31.** Leyenda de interacciones proteína-ligando presentes en el diagrama en dos dimensiones de acuerdo con MOE.

## 9.4 Puntajes de interacción calculados para el acoplamiento molecular

**Tabla 8.** Puntajes y energías de unión calculadas empleando el programa Glide para el acoplamiento molecular de los ligandos con mayor similitud 3D con sordarina en HsEFL1 y ScEfl1 en la conformación traslocada.

No.	Acoplamiento en HsEFL1			Acoplamiento en ScEfl1		
	Ligando	XP Glide score (kcal/mol)	MMGBSA dG Bind (kcal/mol)	Ligando	XP Glide score (kcal/mol)	MMGBSA dG Bind (kcal/mol)
1	ZINC000253532108	-10.21	-29.45	ZINC000253532111	-9.80	-29.41
2	ZINC000253534722	-9.43	-32.22	ZINC000253532106	-9.32	-43.81
3	ZINC000253534724	-9.41	-28.61	ZINC000077312672	-9.03	-42.57
4	ZINC000085340910	-8.94	-34.63	ZINC000253532108	-8.72	-26.18
5	ZINC000077312672	-8.94	-36.74	ZINC000255240337	-8.69	-33.70
6	ZINC000253532111	-8.56	-31.39	ZINC000253534724	-8.44	-36.73
7	ZINC000253532104	-8.54	-43.58	ZINC000253534722	-8.37	-32.71
8	ZINC000253532106	-8.49	-30.13	ZINC000253534723	-8.13	-41.01
9	ZINC000257470512	-7.40	-34.75	ZINC000008214538	-8.08	-59.55
10	ZINC000036052121	-7.31	-50.00	ZINC000096023887	-8.06	-53.72

**Tabla 9.** Puntajes calculados empleando MOE para el acoplamiento molecular de los ligandos con mayor similitud 3D con sordarina en HsEFL1 y ScEfl1.

No.	Acoplamiento en HsEFL1		Acoplamiento en ScEfl1	
	Ligando	S* (kcal/mol)	Ligando	S* (kcal/mol)
1	ZINC000253534724	-27.22	ZINC000077312672	-28.95
2	ZINC000253532111	-24.71	ZINC000253534723	-26.79
3	ZINC000253534723	-23.95	ZINC000253532108	-26.28
4	ZINC000253534722	-22.95	ZINC000033604951	-25.47
5	ZINC000077292071	-22.70	ZINC000040163012	-25.42
6	ZINC000253532104	-22.51	ZINC000253534724	-25.37
7	ZINC000253532108	-22.25	ZINC000255240337	-25.19
8	ZINC000040164488	-22.23	ZINC000257470514	-24.88
9	ZINC000008234385	-21.39	ZINC000085946246	-24.37
10	ZINC000033604952	-20.40	ZINC000253534722	-24.31

\*S: Energía libre de unión ligando-proteína



**Tabla 10.** Puntajes calculados empleando AutoDock Vina para el acoplamiento molecular de los ligandos con mayor similitud 3D con sordarina en HsEFL1 y ScEfl1.

No.	Acoplamiento en HsEFL1		Acoplamiento en ScEfl1	
	Ligando	Afinidad de unión (kcal/mol)	Ligando	Afinidad de unión (kcal/mol)
1	ZINC000096102533	-9.20	ZINC000008214538	-10.00
2	ZINC000040164488	-9.00	ZINC000085946246	-10.00
3	ZINC000077292071	-9.00	ZINC000085946247	-9.80
4	ZINC000036052121	-8.80	ZINC000077292071	-9.60
5	ZINC000038140888	-8.70	ZINC000040164488	-9.40
6	ZINC000008234384	-8.50	ZINC000038140888	-9.40
7	ZINC000085946247	-8.30	ZINC000008234384	-9.40
8	ZINC000038140890	-8.20	ZINC000008234385	-9.20
9	ZINC000253534724	-8.10	ZINC000257470512	-9.10
10	ZINC000004692041	-8.00	ZINC000095099842	-9.10

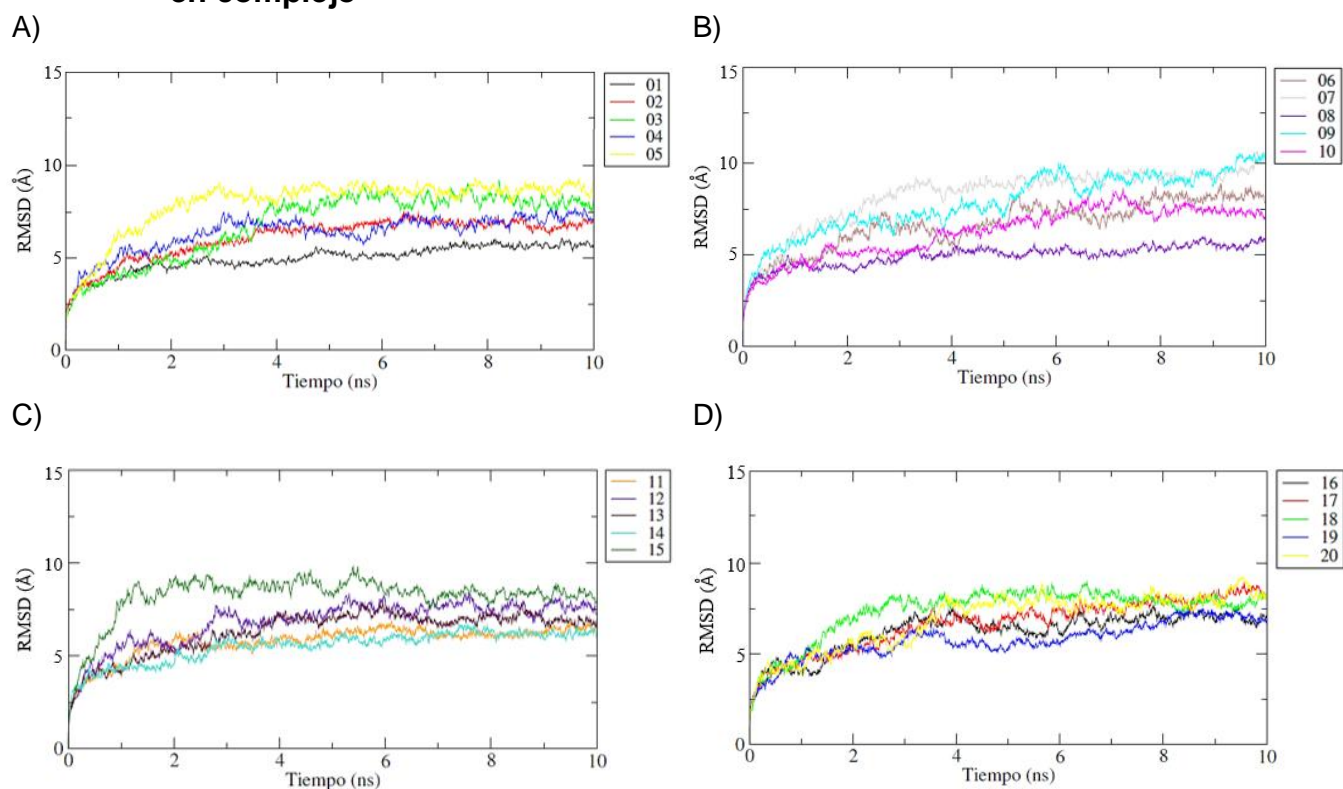
## 9.5 Ligandos evaluados mediante dinámica molecular

**Tabla 11.** Ligandos presentes en complejo con las proteínas EFL1 evaluados mediante dinámica molecular.

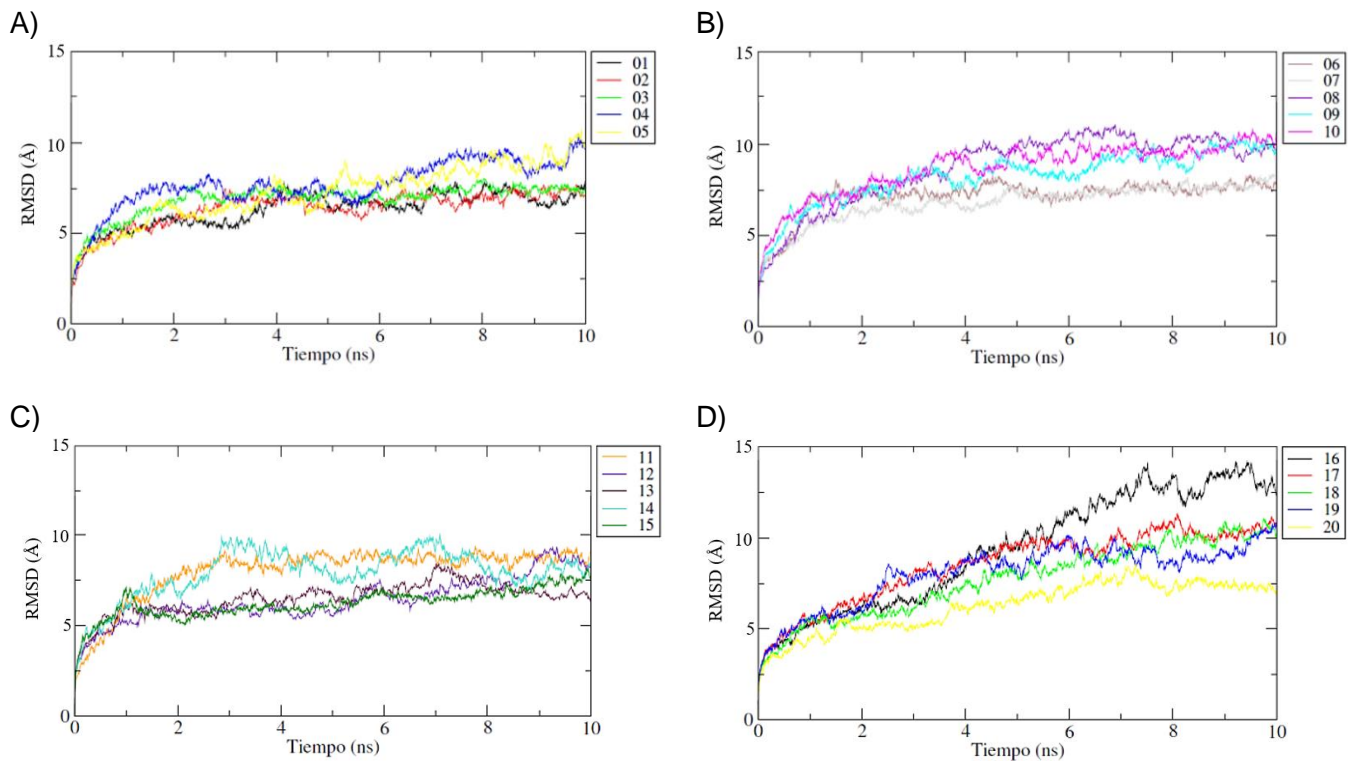
No.	Ligando en complejo con HsEFL1	Ligando en complejo con ScEfl1
1	ZINC000096023887	ZINC000096023887
2	ZINC000257470512	ZINC000257470512
3	ZINC000015070573	ZINC000015070573
4	ZINC000077291909	ZINC000077291909
5	ZINC000004692041	ZINC000004692041
6	ZINC000253532104	ZINC000253532104
7	ZINC000036052121	ZINC000036052121
8	ZINC000077312672	ZINC000077312672
9	ZINC000040164488	ZINC000040164488
10	ZINC000008214538	ZINC000008214538
11	ZINC000253534724	ZINC000253534724
12	ZINC000085946247	ZINC000085946247
13	ZINC000257459537	ZINC000257459537

14	ZINC000038139805	ZINC000038139805
15	ZINC000035465157	ZINC000038140888
16	ZINC000253534722	ZINC000033604951
17	ZINC000003295242	ZINC000247930162
18	ZINC000253532106	ZINC000096102535
19	ZINC000253532108	ZINC000252515685
20	ZINC000040163012	ZINC000254045588

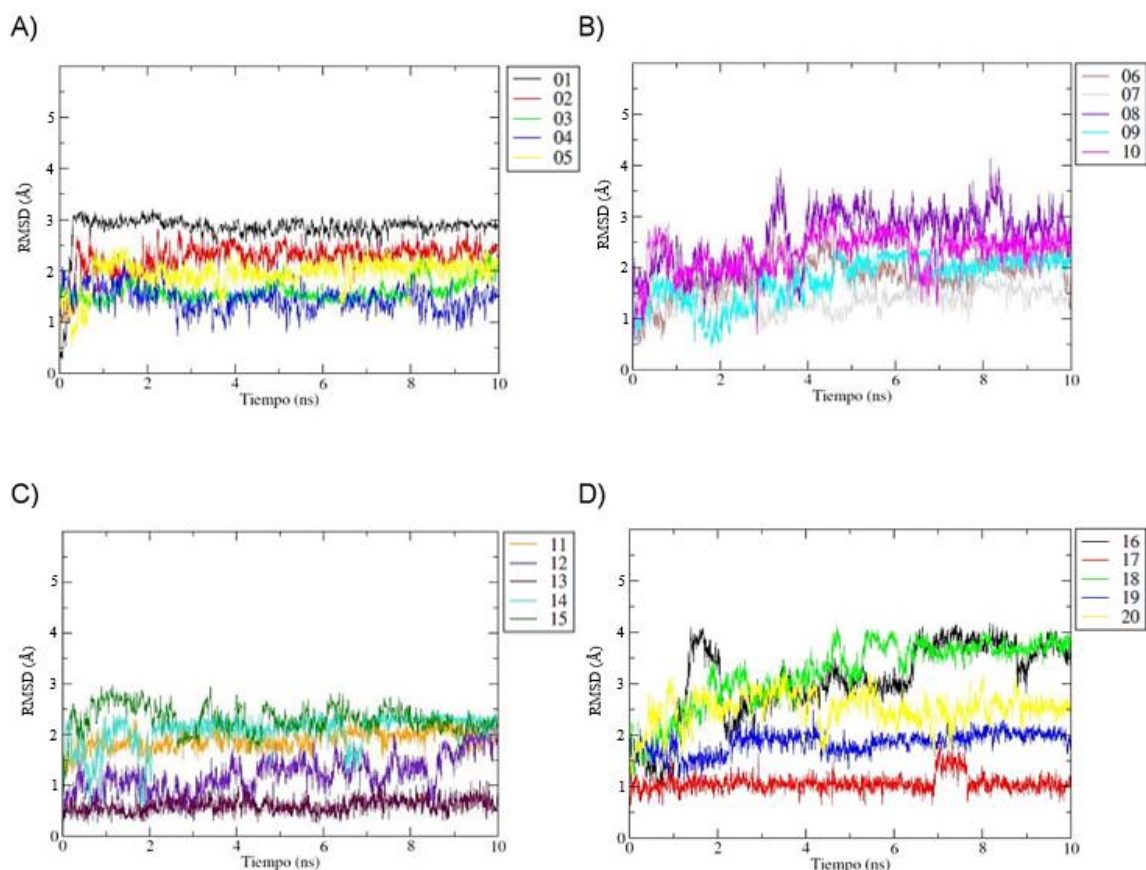
## 9.6 RMSD del esqueleto proteico y de los átomos pesados de los ligandos en complejo



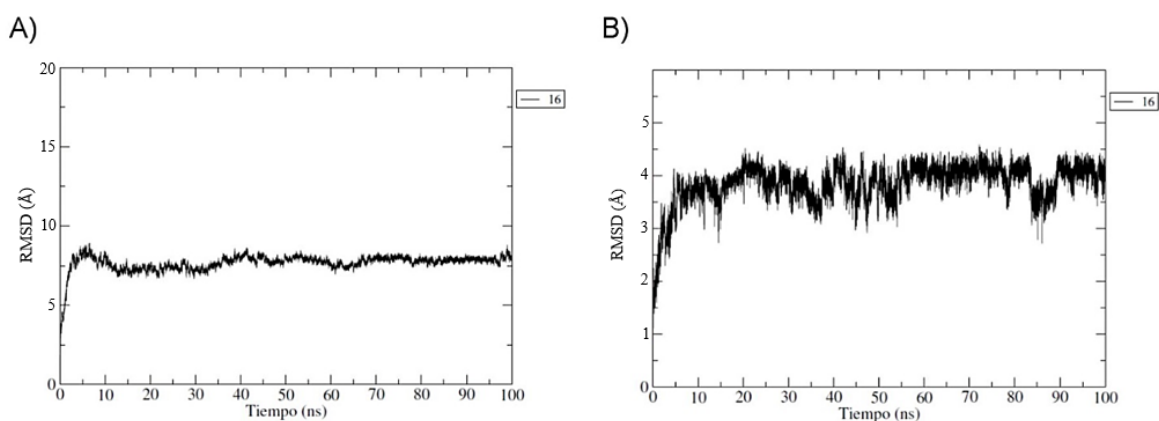
**Figura 32.** No existen fluctuaciones marcadas en el esqueleto de la proteína de los complejos analizados, la presencia de los ligandos no altera la estabilidad de la estructura proteica. RMSD del esqueleto de la proteína HsEFL1 calculado para los átomos N, C $\alpha$ , C, O en complejo con cada uno de los ligandos analizados. A) ligandos 1 al 5, B) ligandos 6 al 10, C) ligandos 11 al 15 y D) ligandos 16 al 20 de acuerdo con la nomenclatura de la Tabla 11.



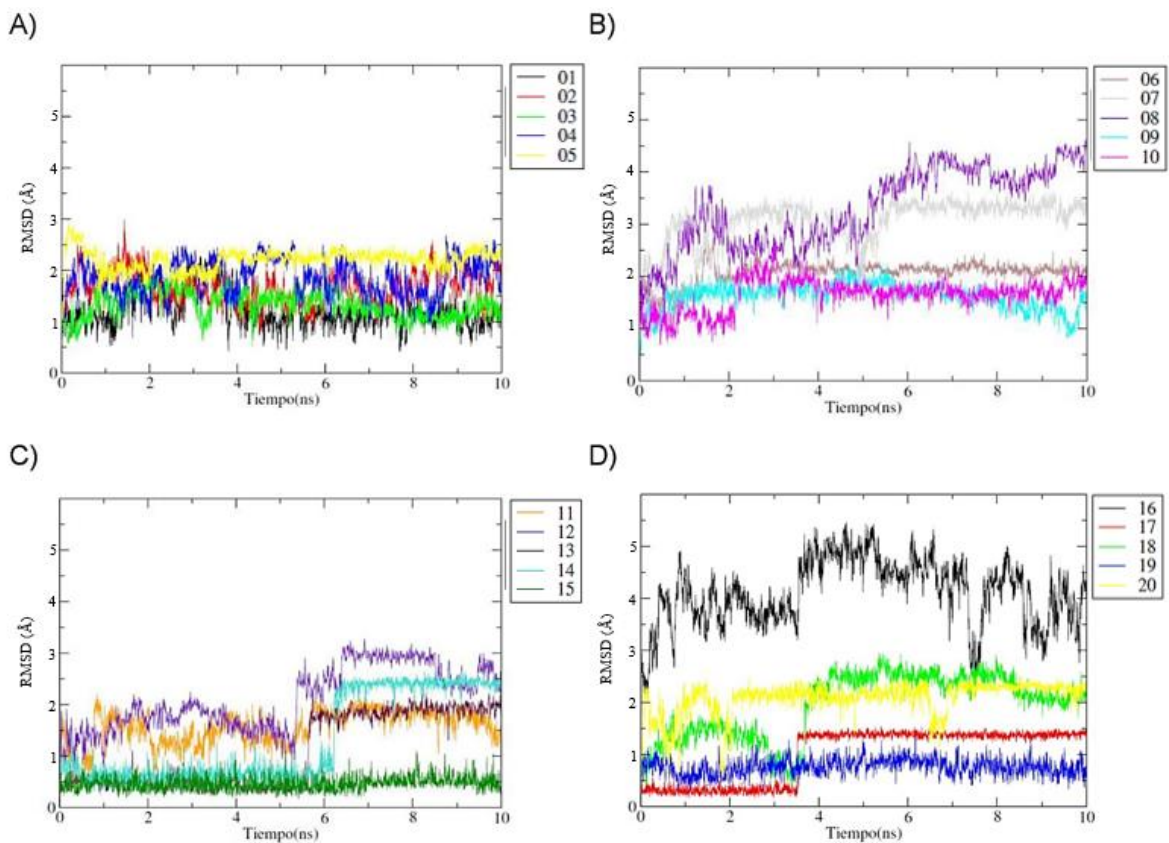
**Figura 33.** No existen fluctuaciones marcadas en el esqueleto de la proteína de los complejos analizados, la presencia de los ligandos no altera la estabilidad de la estructura proteica. RMSD del esqueleto de la proteína ScEf11 calculado para los átomos N, C $\alpha$ , C, O en complejo con cada uno de los ligandos analizados. A) ligandos 1 al 5, B) ligandos 6 al 10, C) ligandos 11 al 15 y D) ligandos 16 al 20 de acuerdo con la nomenclatura de la Tabla 11.



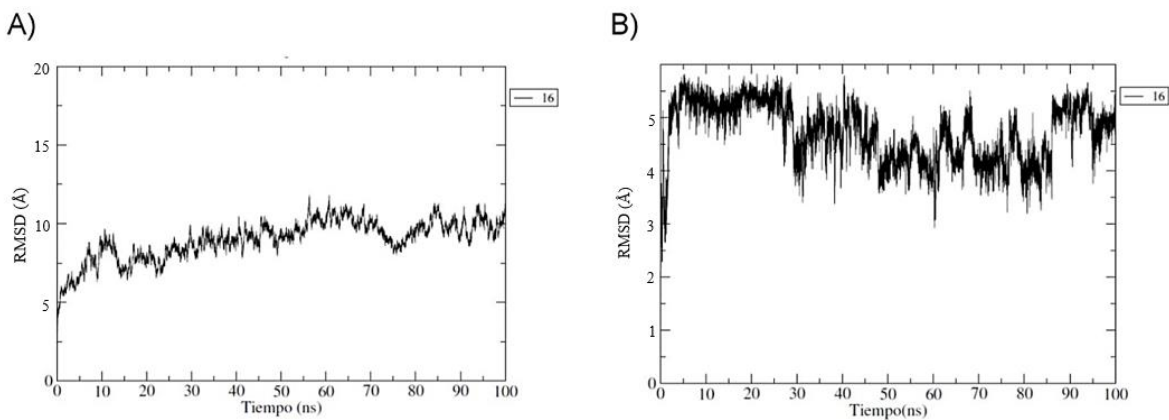
**Figura 34.** Estabilidad en el sitio de unión de los ligandos. RMSD del ligando calculado en complejo con la proteína HsEFL1. A) ligandos 1 a 5, B) ligandos 6 al 10, C) ligandos 11 al 15 y D) ligandos 16 al 20 de acuerdo con la nomenclatura de la Tabla 11.



**Figura 35.** A) RMSD del esqueleto de la proteína HsEFL1 calculado para los átomos N, Ca, C, O. B) RMSD del ligando en el sitio de unión en el complejo HsEFL1-ligando 16.



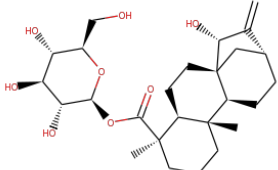
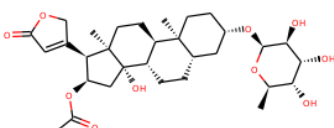
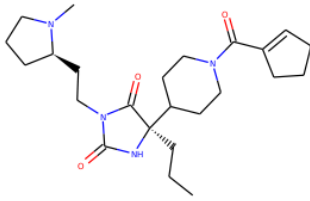
**Figura 36.** Estabilidad en el sitio de unión de los ligandos. RMSD del ligando calculado en complejo con la proteína ScEfl1. A) ligandos 1 al 5, B) ligandos 6 al 10, C) ligandos 11 al 15 y D) ligandos 16 al 20 de acuerdo con la nomenclatura de la Tabla 11.

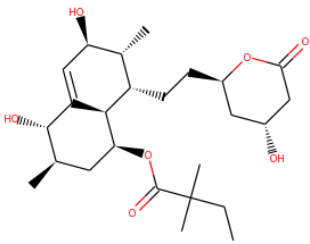
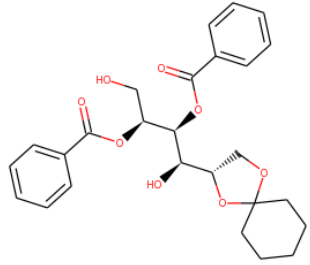
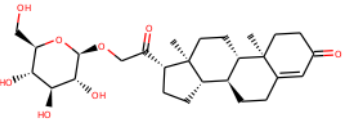
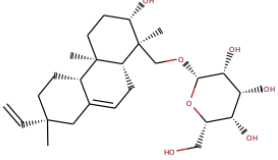
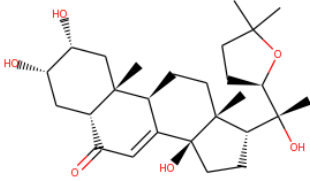


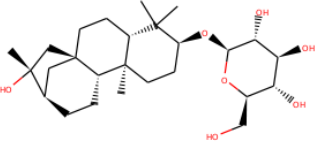
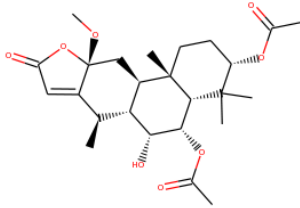
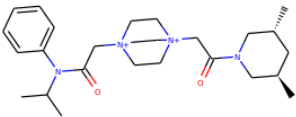
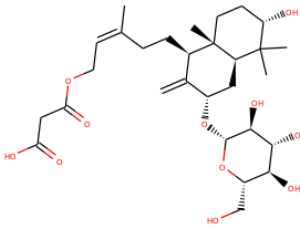
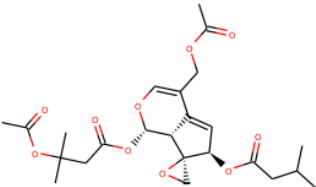
**Figura 37.** A) RMSD del esqueleto de la proteína ScEfl1 calculado para los átomos N, Ca, C, O. B) RMSD del ligando en el sitio de unión en el complejo ScEfl1-ligando 16.

## 9.7 Compuestos con energía de unión favorable con HsEFL1 y ScEfl1 mediante la metodología LIE

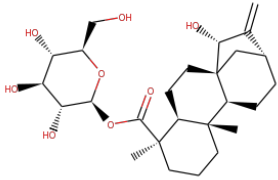
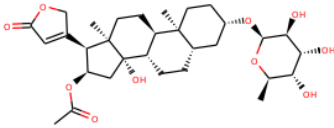
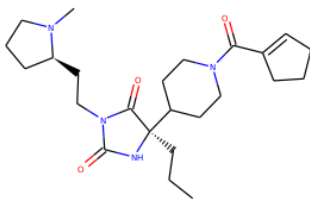
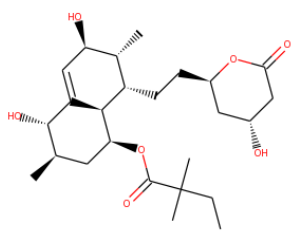
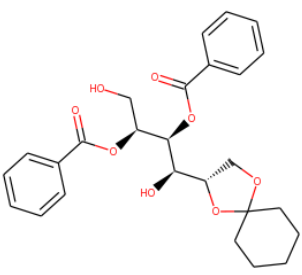
Tabla 12. Compuestos con energía libre de unión favorables calculadas con la metodología LIE en complejo con HsEFL1 y ScEfl1.

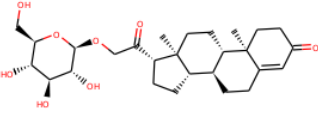
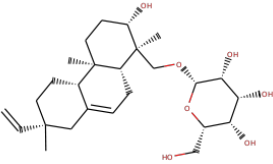
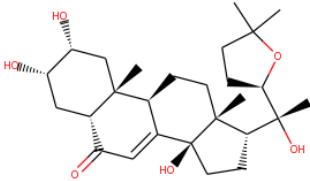
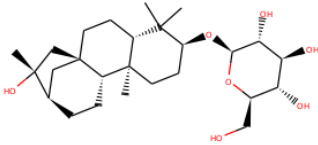
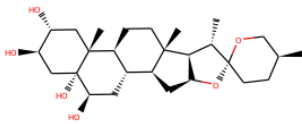
Proteína en complejo	Ligando No.	Compuesto	Estructura	LIE $\Delta G_{unión}$ kJ/mol (DS)
HsEFL1	1	ZINC000096023887		-17.550 (6.779)
	2	ZINC000257470512		-22.074 (7.902)
	3	ZINC000015070573		-20.019 (6.800)

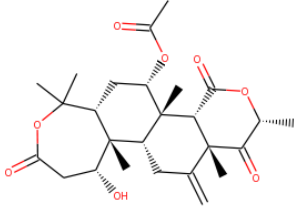
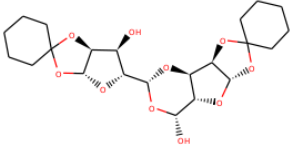
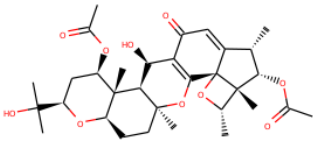
4	ZINC000077291909		-5.978 (5.173)
7	ZINC000036052121		-7.384 (5.115)
10	ZINC000008214538		-16.587 (6.504)
12	ZINC000085946247		-30.460 (7.539)
13	ZINC000257459537		-13.453 (5.836)

14	ZINC000038139805		-8.225 (4.617)
15	ZINC000035465157		-13.330 (5.521)
17	ZINC000003295242		-46.721 (13.193)
19	ZINC000253532108		-11.348 (9.901)
20	ZINC000040163012		-9.134 (5.309)

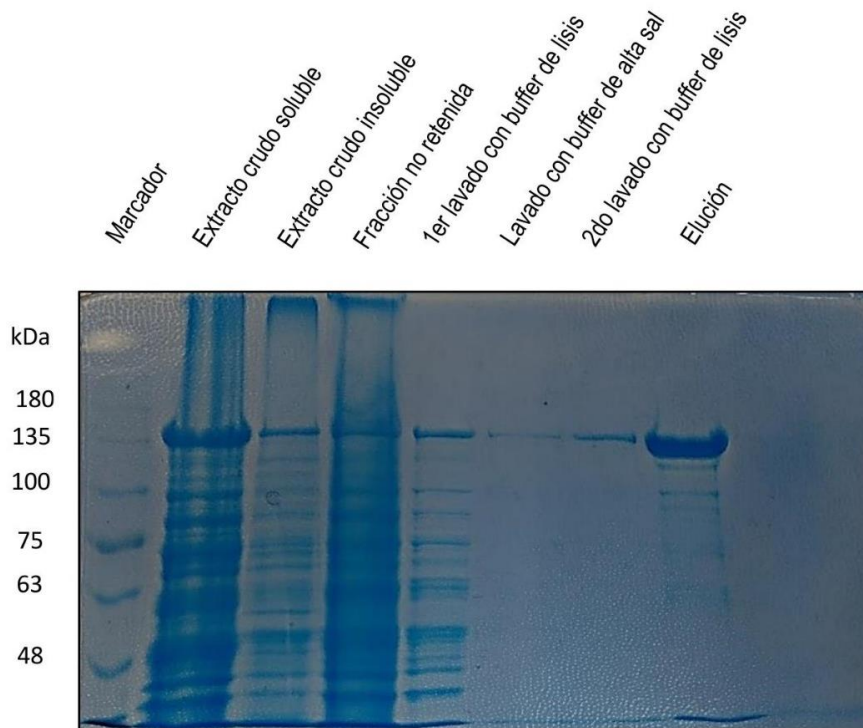


ScEfl1	1	ZINC000096023887		-5.360 (5.720)
	2	ZINC000257470512		-2.371 (5.482)
	3	ZINC000015070573		-16.088 (11.365)
	4	ZINC000077291909		-1.682 (4.898)
	7	ZINC000036052121		-8.685 (5.144)

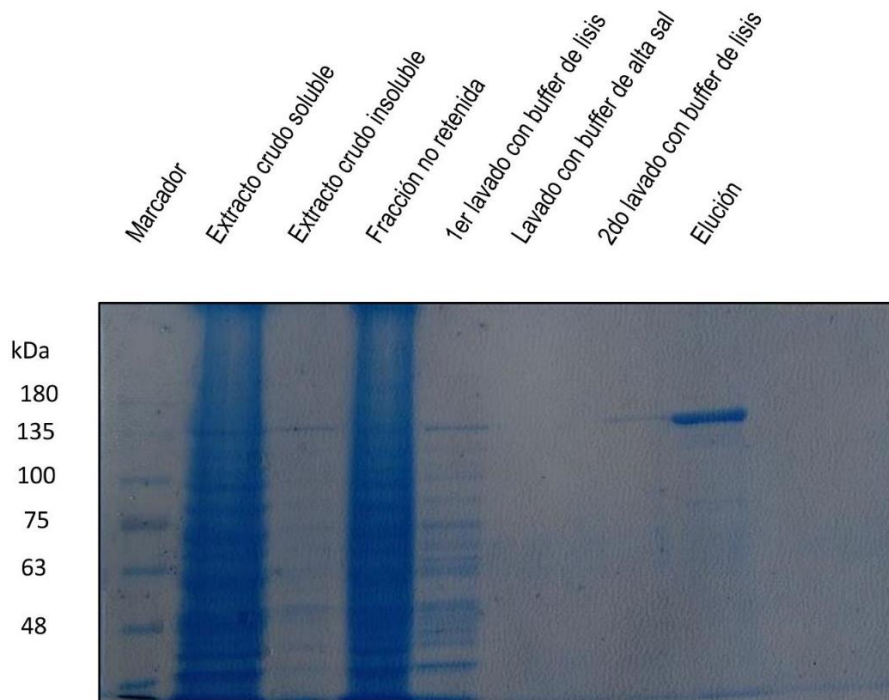
10	ZINC000008214538		-9.252 (6.138)
12	ZINC000085946247		-2.140 (6.461)
13	ZINC000257459537		-1.777 (4.779)
14	ZINC000038139805		-15.246 (5.268)
15	ZINC000038140888		-3.139 (4.594)

17	ZINC000247930162		-3.039 (5.343)
18	ZINC000096102535		-13.884 (4.646)
19	ZINC000252515685		-2.544 (6.286)

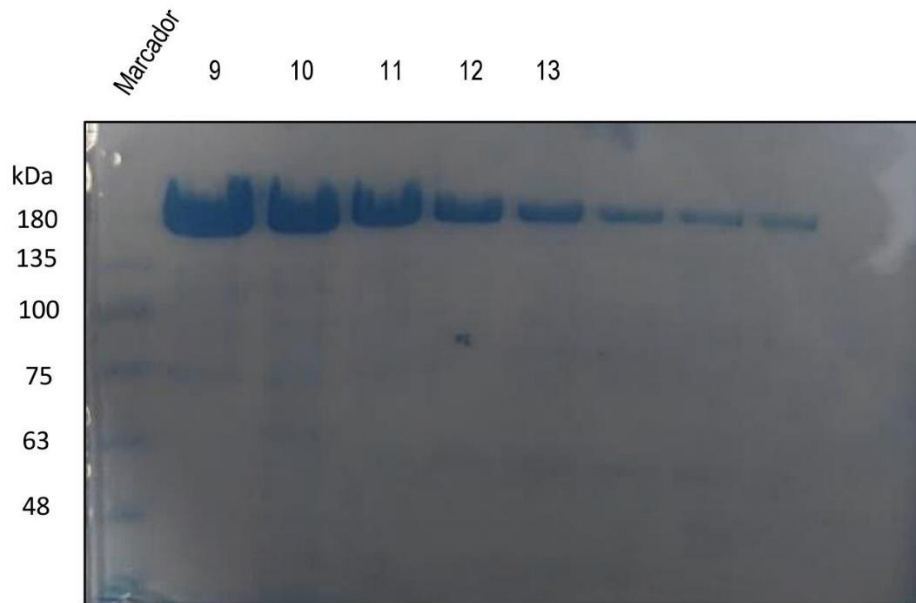
## 9.8 Monitoreo de la purificación de HsEFL1 y ScEfl1



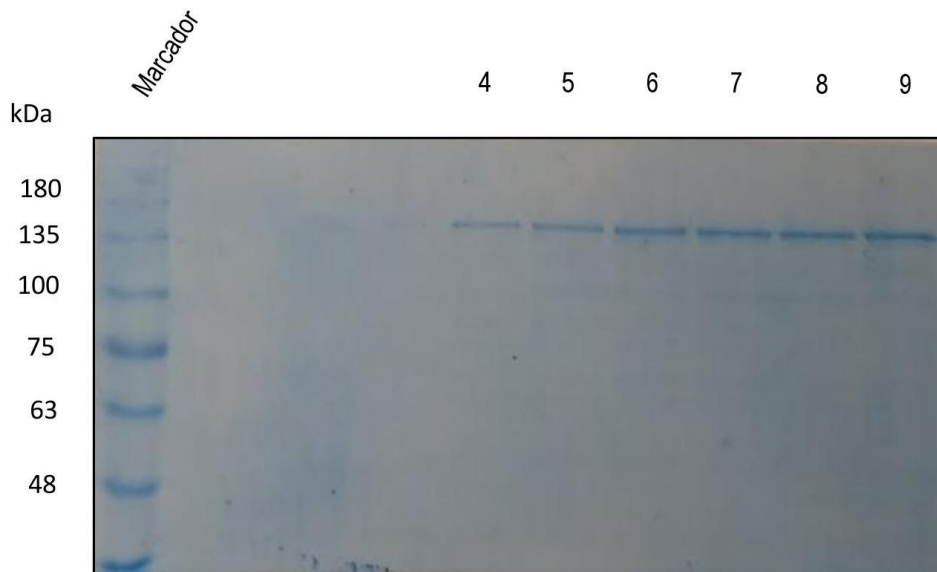
**Figura 38.** Electroforesis SDS-PAGE en gel de poliacrilamida al 7.5% de las fracciones obtenidas mediante cromatografía de afinidad a  $\text{Ni}^{2+}$  de la proteína ScEfl1.



**Figura 39.** Electroforesis SDS-PAGE en gel de poliacrilamida al 7.5% de las fracciones obtenidas mediante cromatografía de afinidad a  $\text{Ni}^{2+}$  de la proteína HsEFL1.



**Figura 40.** Electroforesis SDS-PAGE en gel de poliacrilamida al 7.5% de las fracciones 9 a la 13 recolectadas de la cromatografía de exclusión molecular de la proteína.



**Figura 41.** Electroforesis SDS-PAGE en gel de poliacrilamida al 7.5% de las fracciones 4 a la 9 recolectadas de la cromatografía de exclusión molecular de la proteína HsEFL1.