

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

Análisis *in silico* de compuestos derivados del cumato y floretina para suprimir la expresión
génica en *Drosophila melanogaster* modificadas genéticamente

Juan Nicolás Fuentes Pinto

Grupo de Investigación en Bioquímica y Microbiología (GIBIM)

Trabajo de Grado para Optar al Título en Química

Directora

Dra. Stelia Carolina Méndez Sánchez (GIBIM)

Codirector

Dr. Luis Carlos Vesga Gamboa (GIBIM)

Universidad Industrial de Santander

Facultad de Ciencias

Escuela de Química

Bucaramanga

2023

Dedicatoria

A mi nona Aurora, por enseñarme que la ciencia es más que una pasión.

A mis padres, por enseñarme el valor de la perseverancia y dedicación para cumplir mis metas.

A Happy, Tororo y Kitty por ser mis compañeros gatunos en mis mañanas, tardes y noches de escritura.

A la depresión, quien me enseñó que es nuestro compañero más fiel.

Agradecimientos

Agradezco a Dios por darme la paciencia y perseverancia para poder realizar este logro tan importante en mi vida.

A mis padres, por darme su constante apoyo, amor y la oportunidad de demostrarles que soy mejor persona gracias al esfuerzo de ellos.

A mis hermanos, por sacarme sonrisa, quitarme el estrés y acompañarme durante todo el proceso.

A Shirleidy García por acompañarme en la instancia final de mi proyecto y ser paciente para cumplir nuestros objetivos.

A la profesora Stelia por confiar en mi y darme la oportunidad de desarrollar este proyecto a pesar de todas las dificultades que se presentaron en su realización.

A Luis Carlos por ser ese amigo y guía en este proyecto, de igual manera, espero que se convierta en un excelente doctor para poder seguir trabajando con él.

A mi compañera María Lucía Otero, quien fue mi ayudante, compañera de travesuras y problemas durante la realización del proyecto.

*A la universidad Industrial de Santander, Vicerrectoría de Investigación y Extensión VIE por la financiación del proyecto: “Análisis *in silico* y experimental de derivados de Floretina, Cumato y Vanilato para suprimir expresión génica en moscas *Drosophilla melanogaster* modificadas genéticamente” bajo el código:2678.*

Muchas gracias a todos, este logro no es solo mío.

Contenido

Introducción.....	14
1. Marco de referencia y antecedentes.	16
1.1. Características estructurales de las proteínas CymR y TtgR.....	16
1.2. Reconocimiento molecular de las proteínas CymR y TtgR.....	17
1.3. Usos de métodos computacionales para el reconocimiento molecular.....	19
1.3.1. Modelado de homología.....	19
1.3.2. Docking molecular.....	21
1.3.3. Dinámica Molecular.....	25
1.3.4. Campos de Fuerza.....	27
1.4. Usos de sistemas binarios para la generación de machos estériles.	29
1.5. Uso de proteínas represoras reguladas por una molécula pequeña.....	30
2. Objetivos.	32
2.1. Objetivo general.	32
2.2. Objetivos específicos.....	32
3. Metodología.....	32
3.1 Fase preliminar.	33
3.1.1. Modelo de Homología.....	34
3.1.2. Preparación de proteína y ligandos.	35
3.1.3. Screening virtual de derivados del cumato y Floretina.....	36
3.1.4. Simulación de Dinámica Molecular.....	37
4. Resultados y Discusión.....	37
4.1. Proteína CymR.....	37

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

4.1.1.	Resultados de Modelo de Homología	38
4.1.2.	Validación del modelo de Homología CymR	39
4.1.3.	Identificación de sitios drogables por SiteMap para la proteína CymR.....	40
4.1.4.	Docking molecular para CymR.....	43
4.1.5.	Dinámica Molecular CymR	50
4.1.6.	Análisis de interacciones proteína-ligando CymR	54
4.2.	Proteína TtgR	58
4.2.1.	Identificación de Target para TtgR	58
4.2.2.	Docking molecular para TtgR.....	60
4.2.3.	Docking molecular para TtgR en el sitio de Alta afinidad.....	61
4.2.4.	Resultados Dinámica Molecular para TtgR	65
4.2.5.	Análisis de interacciones proteína-ligando TtgR	69
5.	Conclusiones.	74
6.	Recomendaciones.....	75
	Divulgación de Resultados.....	75
	Referencias Bibliográficas.....	76
	Apéndices	92

Lista de Tablas

Tabla 1. Puntuaciones de acoplamiento y energía de afinidad de los 10 compuestos mejores calificados junto con sus propiedades ADME.	46
Tabla 2. Puntuaciones de acoplamiento y energía libre de unión de los compuestos mejores calificados en el sitio de unión de alta afinidad de proteína TtgR junto con sus propiedades ADME.	62

Lista de Ecuaciones

Ecuación 1. Función de puntuación basada en campos de fuerza de AutoDock.	22
Ecuación 2. Funciones de puntuación empírica de GlideScore	23
Ecuación 3. Funciones de puntuación basadas en el conocimiento PMF	24
Ecuación 4. Segunda ley de Newton para cálculos en dinámica molecular.....	26
Ecuación 5. Segunda ley de Newton con la aceleración para cálculos en dinámica molecular.....	26
Ecuación 6. Función energía potencial.	26
Ecuación 7. Energía potencial total.....	28
Ecuación 8. Modelo matemático para obtener SiteScore en maestro Schrödinger.....	41
Ecuación 9. Modelo matemático para obtener DrugScore en maestro Schrödinger.....	41

Lista de Figuras

Figura 1. Esquema metodológico para el desarrollo del proyecto de investigación	33
Figura 2. Superposición modelo de Homología con template 2HYT.....	38
Figura 3. Gráfico de Ramachandran del modelo de homología. CymR	40
Figura 4. Sitios de unión obtenidos con la herramienta de SiteMap de Maestro Schrödinger.....	42
Figura 5. Visualización del inhibidor cumato para la proteína CymR.....	44
Figura 6. Visualización del inhibidor cumato para la proteína CymR.....	48
Figura 7. Valores de RMSD para Complejos CymR-Ligandos.....	51
Figura 8. Valores de RMSD para los derivados de Cumato.	52
Figura 9. Cambios posicionales del ácido 3-bromo-2-metoxibenzoico durante la réplica 1.	53
Figura 10. Interacciones del Complejo CymR-Ligando a lo largo de la simulación.	55
Figura 11. Estructura cristalina de la proteína TtgR con código PDB: 2UXI.....	59
Figura 12. Visualización del inhibidor floretina para la proteína TtgR	60
Figura 13. Interacciones Predichas en el sitio de alta afinidad por el Docking Molecular para TtgR.	64
Figura 14. Valores de RMSD para TtgR.....	66
Figura 15. Valores de RMSD para la floretina y compuesto en estudio.....	67
Figura 16. Cambios posicionales de la floretina durante la réplica 1.....	68
Figura 17. Interacciones del Complejo TtgR-Ligando a lo largo de la simulación.	70

Lista de Apéndices

	pág.
Apéndice 1 Base de datos de compuestos para CymR.....	92
Apéndice 2. Base de datos de compuestos para TtgR.....	102
Apéndice 3. RMSD de las proteínas TtgR y CymR sin ligando.	103
Apéndice 4. RMSD de la proteína CymR sin ligando y con Cumato.	103
Apéndice 5. RMSD de la proteína CymR sin el ligando y con Ácido 4-amino-3-hidroxibenzoico	103
Apéndice 6. RMSD de la proteína CymR sin el ligando y con Ácido 3-bromo-2-metoxibenzoico	103
Apéndice 7. RMSD de la proteína TtgR sin el ligando y Floretina	103
Apéndice 8. RMSD de la proteína TtgR sin el ligando y Quercetina	103
Apéndice 9. RMSD de la proteína TtgR sin el ligando y Hesperetina.....	103
Apéndice 10. RMSD de la proteína TtgR sin el ligando y Hesperetina.....	103

Resumen

Título: Análisis *in silico* de compuestos derivados del cumato y floretina para suprimir la expresión génica en *Drosophila melanogaster* modificadas genéticamente.

Autor: Juan Nicolás Fuentes Pinto, Luis Carlos Vesga, Stelia Carolina Méndez Sánchez.

Palabras Clave: *Drosophilla melanogaster*; supresión génica; docking molecular; dinámica molecular.

Descripción: En la actualidad, los insectos son una problemática que se asocia a la pérdida de producción agrícola y transmisión de enfermedades. En consecuencia, los métodos genéticos de control de plagas han adquirido creciente importancia, donde los sistemas de expresión genética para el control poblacional permiten emplear proteínas represoras como estrategia para la inducción de letalidad sexo específica en insectos. Por tanto, en el laboratorio del profesor Omar Akbari (Universidad de California, San Diego UCSD) se generaron moscas *Drosophilla melanogaster* modificadas genéticamente. Estas moscas presentan sistemas binarios conformados por: una proteína represora (CymR y TtgR, que son reguladas por cumato y floretina, respectivamente) y la proteína VP16 como factor de transcripción que pueden inducir la expresión de proteínas fluorescentes como la proteína GFP y luciferasa *in vivo*. En cultivos *In vitro*, se ha demostrado que estos sistemas reprimen la transcripción (en presencia de la molécula represora), sin embargo, en modelos *in vivo* no se obtienen los mismos resultados. En nuestro laboratorio surgió como hipótesis que una de las posibles razones por las cuales no se evidencia la inhibición de la expresión *in vivo*, se debe a la baja biodisponibilidad de las moléculas represoras en el músculo del vuelo de los insectos. Por consiguiente, se identificaron compuestos con posible capacidad inhibidora para las proteínas CymR y TtgR, y se evaluó su estabilidad en el sitio de

unión propuesto mediante ensayos de docking molecular y dinámica molecular. A partir de ello, se determinó que el ácido 4-amino-3-hidroxibenzoico presenta interacciones tipo puente de hidrógeno con los aminoácidos encargados de la estabilidad del complejo proteína-ligando, los cuales son: Arg126, Asn136 y Arg170 con una frecuencia de 100%, 50% y 30% respectivamente, logrando una estabilidad en el sitio de unión de la proteína CymR durante 200ns. Asimismo, la quercetina presenta interacciones tipo puente de hidrógeno con los aminoácidos encargados de la inhibición de la proteína TtgR Asn110, Arg130 y Aps172 con una frecuencia de 10%, 20% y 40%, respectivamente, permaneciendo en el sitio de unión de la proteína TtgR a lo largo de 200ns. Por consiguiente, los resultados sugieren que el ácido 4-amino-3-hidroxibenzoico y la quercetina son compuestos promisorios para inhibir la expresión génica de la proteína GFP en moscas transgénicas que expresan las proteínas CymR y TtgR, como sistemas represores.

* Trabajo de Grado ** Facultad de Ciencias, Escuela de Química. Director: Dr. Stelia Carolina Méndez Sánchez. Codirector: Dr. Luis Carlos Vesga

Abstract

Title: In silico analysis of cumate and phloretin derived compounds to suppress gene expression in genetically modified *Drosophila melanogaster*.

Author(s): Juan Nicolás Fuentes Pinto, Luis Carlos Vesga, Stelia Carolina Méndez Sánchez

Key Words: *Drosophilla melanogaster*; gene suppression; molecular docking; molecular dynamics.

Description: Insects are currently a problem associated with loss of agricultural production and disease transmission. Consequently, genetic methods of pest control have become increasingly important, where gene expression systems for population control allow the use of repressor proteins as a strategy for the induction of sex-specific lethality in insects. Therefore, genetically modified *Drosophilla melanogaster* flies were generated in the laboratory of Professor Omar Akbari (University of California, San Diego UCSD). These flies have binary systems consisting of a repressor protein (CymR and TtgR, which are regulated by coumate and floretin, respectively) and the VP16 protein as a transcription factor that can induce the expression of fluorescent proteins such as GFP and luciferase *in vivo*. *In vitro*, these systems have been shown to repress transcription (in the presence of the repressor molecule), but the same results are not obtained in *in vivo* models. In our laboratory, it was hypothesised that one of the possible reasons for the lack of evidence of inhibition of expression *in vivo* is due to the low bioavailability of the repressor molecules in insect flight muscle. Therefore, compounds with potential inhibitory capacity for the CymR and TtgR proteins were identified, and their stability at the proposed binding site was assessed by molecular docking and molecular dynamics assays. From this, it was determined that 4-amino-3-hydroxybenzoic acid exhibits hydrogen bridge-type interactions with the amino acids responsible

for the stability of the protein-ligand complex, which are: Arg126, Asn136 and Arg170 with a frequency of 100%, 50% and 30% respectively, achieving stability at the CymR protein binding site for 200ns. Likewise, quercetin exhibits hydrogen bridge-like interactions with the TtgR inhibitory amino acids Asn110, Arg130 and Aps172 with a frequency of 10%, 20% and 40%, respectively, remaining in the TtgR protein binding site for 200ns. Therefore, the results suggest that 4-amino-3-hydroxybenzoic acid and quercetin are promising compounds for inhibiting GFP gene expression in transgenic flies expressing CymR and TtgR proteins as repressor systems.

* Bachelor's degree Work ** Faculty of Science. School of Chemistry. Director: Dr. Stelia Carolina Méndez Sánchez. Codirector: Dr. Luis Carlos Vesga

Introducción.

El control preciso de la expresión génica es uno de los principales objetivos para la biotecnología, puesto que, presenta un amplio campo de aplicaciones biológicas como la terapia génica mediante la reprogramación celular en la ingeniería de tejidos (Fussenegger et al., 1998; Thomas et al., 2000) o la construcción de sensores para circuitos génicos sintéticos (Deans et al., 2007; Kramer & Fussenegger, 2005). Aunque existen diferentes métodos para este control, es de vital prioridad encontrar métodos que puedan presentar una toxicidad casi nula, es por ello, que el uso de sistemas de expresión génica es un método alternativo donde se hace uso de una proteína represora controlada por un compuesto químico (Bertram & Hillen, 2008; Venken et al., 2011). Ahora bien, los sistemas de expresión genética regulados son una poderosa herramienta que contribuye a generar métodos de control poblacional (Stebbins & Yin, 2001; You et al., 2004), sin embargo, la mayoría de estos sistemas dependen de un represor transcripcional, que se une estrechamente a operadores en la región promotora de genes diana en ausencia de un inductor, impidiendo así la transcripción. (Kaczmarczyk et al., 2013).

En la actualidad, se ha reportado casos satisfactorios de represión de la expresión génica controlada por tetraciclina en moscas. Estos logros se han obtenido mediante el uso de un potenciador ocular específico diseñado para impulsar la expresión del transgén TetO-LacZ. La expresión de LacZ se pudo observar seis horas después de que las larvas fueran sometidas a un medio carente de tetraciclina. Asimismo, también se demostró la capacidad de suprimir la expresión del transgén TetO-LacZ en los embriones alimentando a las madres con altas concentraciones de tetraciclina (Bello et al., 1998; McGuire et al., 2004; Thomas et al., 2000).

Por otro lado, en el ámbito de la ingeniería genética, el represor de vanilato en ratones ha propiciado el desarrollo de un sistema novedoso. Este sistema ha sido minuciosamente concebido

para la regulación de transgenes heterólogos en mamíferos y se caracteriza por su capacidad exclusiva de respuesta a la presencia de la molécula de ácido vanílico. Este enfoque ofrece múltiples configuraciones posibles, permitiendo la variación en el número de operadores, la utilización de diversos dominios de transactivación de mamíferos y la modulación de las distancias entre el sitio del operador específico y el promotor mínimo. De esta manera, proporciona una herramienta sumamente versátil con un extenso espectro de aplicaciones potenciales (Gitzinger et al., 2012).

En el laboratorio Akbari de la Universidad de California, San Diego UCSD se generaron moscas *Drosophila melanogaster* modificadas genéticamente con sistemas binarios, compuestos por una proteína represora y una región de la proteína VP16 (Gamez et al., 2021), que hace las veces de factor de transcripción (Goodrich et al., 1993). Estos sistemas represores se denominan CymR-VP16 y TtgR-VP16, que a su vez, son regulados por el cumato (Eaton, 1997) y la floretina (Alguel et al., 2007a), respectivamente. En cultivos *In vitro*, se demostró que este sistema funciona y logra reprimir la transcripción, sin embargo, al pasar a un modelo *In vivo* no se obtienen los mismos resultados (Gamez et al., 2021). Por consiguiente, se identificaron compuestos con posible capacidad inhibidora para las proteínas CymR y TtgR, y se evaluó su estabilidad en el sitio de unión propuesto mediante ensayos de docking molecular y dinámica molecular.

Para llevar a cabo este proyecto, se presenta como primera instancia una descripción sobre las estructuras de las proteínas y algunos conceptos teóricos sobre métodos computacionales que envuelven el estudio. De igual manera, se presenta una descripción de la metodología, comenzando con la selección de los derivados de cumato y floretina, a los cuales se les realizó un análisis *In silico*, por medio de docking molecular y dinámica molecular.

1. Marco de referencia y antecedentes.

1.1. Características estructurales de las proteínas CymR y TtgR.

La proteína represora CymR pertenece a la familia de factores de transcripción Rrf2. Esta familia, de manera general, posee 3 características fundamentales: a) el regulador de biogénesis del grupo Fe-S IscR (Schwartz et al., 2001); b) el represor de detección de óxido nítrico (NO) NsrR, que tiene el papel de controlar la expresión de los sistemas de desintoxicación de NO (Isabella et al., 2009) y c) el regulador global sensible al hierro RirA en *Rhizobiaceae* (Johnston et al., 2007). Estos 3 grupos controlan o coordinan un grupo Fe-S que es el precursor en la transducción de señales (Isabella et al., 2009; Schwartz et al., 2001). Adicionalmente, la proteína CymR posee un mecanismo de control diferente a los de esta familia, puesto que, forma un complejo regulador con la proteína CysK. En este complejo, CymR es la proteína de unión al ADN, y CysK aumenta la estabilidad en la interacción de ADN-CymR. Asimismo, la proteína CysK a través de su sustrato O-acetilserina (OAS) (Albanesi Daniela et al., 2005; Hullo Marie-Françoise et al., 2007; Tanous et al., 2008), actúa como sensor para la reserva de cisteína en la célula, de tal manera, que permite la transferencia de información al represor CymR, por tanto, en presencia de cisteína, se forma el complejo CymR-CysK y se produce la represión transcripcional.

Por otra parte, la estructura monomérica de la proteína TtgR perteneciente a la familia TetR consta de 9 alfa hélices que forman dos dominios (Alguel et al., 2007a). El dominio de unión al ADN está formado por alpha1-alpha3 (residuos del 1-53) mientras que el dominio de unión al ligando está formado por alpha4-alpha9 (residuos 54-210), con un ángulo de aproximadamente 80° entre los dos dominios (Alguel et al., 2007b). La principal característica de la estructura de TtgR es que hay una curvatura de 65° hacia dentro en el centro de alpha4, característica que no posee ningún otro miembro de la familia (Alguel et al., 2007b). Dentro de esta curvatura destacan

aminoácidos como lo son His70, que da origen a un espacio significativo que permite la entrada del ligando al sitio de unión (Alguel et al., 2007b). Asimismo, la superficie molecular del dímero de TtgR presenta una hendidura alrededor de una curva, ideal para la unión del ligando, esta hendidura o bolsillo es de forma triangular y presenta, en gran parte, características hidrofóbicas (Alguel et al., 2007a). Por último, todos los miembros de la familia TetR se unen al ADN como un homodímero, presentando distancias de 42 Å, conservando un dominio de unión relativamente diverso (Alguel et al., 2007a).

1.2. Reconocimiento molecular de las proteínas CymR y TtgR.

El reconocimiento molecular se refiere al proceso en el que las macromoléculas biológicas interactúan entre sí o con varias moléculas pequeñas a través de interacciones no covalentes para formar un complejo específico (Janin, 1995). Las proteínas son una clase muy importante de macromoléculas porque desempeñan una variedad de funciones en la célula, incluidas funciones estructurales, mecánicas, bioquímicas y de señalización celular (Williams & Daviter, 2016). Estas funciones se realizan a través de su interacción física directa con otras moléculas que reciben el nombre de ligando (Williams & Daviter, 2016).

Ahora bien, estudios sobre reguladores transcripcionales han demostrado que algunas proteínas represoras actúan uniéndose directamente a una amplia gama de compuestos tóxicos generando supresión génica (Jennings et al., 2013). TtgR es un represor de unión a múltiples fármacos que controla negativamente la transcripción del operón *ttgABC*, así como su propia expresión. Se ha demostrado que la proteína TtgR se acopla a compuestos que pertenecen a diferentes clases funcionales, como pueden ser antibióticos, flavonoides y disolventes orgánicos. Para los antibióticos, se ha reportado la unión de la proteína TtgR con cloranfenicol y tetraciclina, donde los anillos cíclicos de estas moléculas estaban rodeados por los residuos hidrofóbicos

Leu92, Leu93, Val96, Ile141 y Phe168. Asimismo, los grupos hidroxilo de la tetraciclina interactúan con los átomos de nitrógenos de los aminoácidos Asp172 y Asn110 (Alguel et al., 2007b). Adicionalmente, para esta proteína también se ha reportado la unión con flavonoides como lo son la quercetina, naringenina y floretina.

La quercetina y la naringenina, tienen una estructura similar conformada por un anillo de cromenona y un anillo de hidroxifenilo, estos dos compuestos se unen en lugares similares que los antibióticos, específicamente, los anillos de cromenona están rodeados por Leu92, Leu93, Ile141 y Phe168, mientras que el otro anillo, está rodeado por Leu66, Val 96, Val171 e Ile175. Además, estas dos moléculas pueden formar puentes de hidrógeno con el aminoácido Asn110 e His114. Sin embargo, de los dos flavonoides la quercetina presenta mayor afinidad con el sitio de unión, puesto que, este presenta más grupos hidroxilos que la naringenina, por tanto, la densidad de electrones está más definida (Alguel et al., 2007b).

Del mismo modo, la floretina al poseer un anillo de hidroxifenilo, se encuentra en la parte inferior del bolsillo de unión y está profundamente conectada mediante puentes de hidrógeno con los residuos His114, Arg130, Arg176 y Asn110, además, el anillo cíclico es rodeado por los aminoácidos hidrofóbicos Leu66, Leu1123, Val134 y Phe168. Esta molécula es el inhibidor por excelencia de esta proteína, puesto que, es la única que presenta interacción con el aminoácido Arg176, quien hace la función de bloqueo del ADN (Alguel et al., 2007a).

Por otra parte, para la proteína CymR solo se ha reportado en la literatura un inhibidor, el cumato, sin embargo, es una proteína que carece de estudio, dado que, aun no se conoce muy bien los aminoácidos que permiten la unión del ADN, ni tampoco se conocen los aminoácidos que pueden conformar su sitio de unión. Por tanto, para este proyecto, se realizó un estudio de sitemap

en el software Maestro Schrödinger donde se utilizan cálculos mecánico cuánticos para encontrar posibles sitios de unión con alta afinidad y drogabilidad.

1.3. Usos de métodos computacionales para el reconocimiento molecular.

Los métodos informáticos en la química computacional son una herramienta poderosa que permite estudiar los complejos inhibidor-receptor y las propiedades ADME (Di & Kerns, 2015). Los métodos básicamente se basan en comparar compuestos mediante su estructura, pero para ello, es necesario conocer la estructura de la proteína que se quiere inhibir, puesto que, hace posible el uso de las herramientas como docking molecular (Leelananda & Lindert, 2016; Sliwoski et al., 2014). Además, las simulaciones de dinámica molecular se pueden utilizar cuando se conoce la estructura del target para brindar información de como los ligandos se unen a las proteínas, sobre las vías de interacción y la flexibilidad de la interacción (Hanson et al., 2015; Y. Wang et al., 2010).

1.3.1. Modelado de homología.

Como se mencionaba en la sección anterior, los métodos computacionales se basan en comparar compuestos mediante su estructura, para ello, es necesario conocer la estructura de la proteína objetivo. Ahora bien, es importante reconocer la importancia del funcionamiento de la proteína, sin embargo, en las bases de datos no se encuentran reportadas todas las estructuras cristalinas de las proteínas identificadas en un organismo, es por ello, que el modelado de homología es útil, puesto que, se emplea cuando la proteína de estudio se puede relacionar con una secuencia conocida y una estructura desconocida con el fin de generar un modelo tridimensional (Robinson et al., 2014).

Para obtener un modelo de homología óptimo, en primera instancia, se debe identificar homólogos específicos de la proteína de estudio y determinar la línea evolutiva, esto se hace

mediante alineaciones de secuencias por pares en serie con la ayuda de técnicas de búsqueda en base de datos como BLAST (John & Sali, 2003). El siguiente paso, es realizar un alineamiento de los homólogos con la proteína base, con el fin de identificar regiones comunes entre las dos estructuras, de esta manera, el homólogo que presente mayor similitud se toma como plantilla y nuevamente se hace un alineamiento, sin embargo, se utilizan algoritmos más refinados como pueden ser PRALINE (Simossis & Heringa, 2005) y T-Coffe (Notredame et al., 2000). PRALINE y T-Coffe son algoritmos de alineación progresiva global que evalúan constantemente cada paso de la alineación con el fin de identificar que secuencia o pares de bloques a alinear, es decir, permiten conocer la secuencia óptima sin requerir un precálculo o alineación previa(Notredame et al., 2000) . La principal diferencia entre los dos algoritmos es que T-Coffe permite alineación de múltiples secuencias heterogéneas a través de una biblioteca de alineaciones por pares (Notredame et al., 2000), mientras que PRALINE ofrece la opción de permitir la incorporación de estructura secundaria o información transmembrana para optimar la alienación (Simossis & Heringa, 2005), sin embargo, los dos algoritmos ofrecen un porcentaje de similitud mayor al 30% entre la secuencia diana y la secuencia de la plantilla (Sanjeevi et al., 2022).

Una vez se identifique la plantilla a utilizar, el siguiente paso es construir una estructura tridimensional para la proteína de estudio en la que se toman las coordenadas de los residuos aminoacídicos de la proteína plantilla. De esta manera, se consolida la cadena principal de la proteína, sin embargo, se deben determinar las posiciones de los bucles y átomos de la cadena lateral, puesto que, esto permite evaluar las interacciones proteína-ligando en los sitios activos y las interacciones proteína-proteína en la interfaz de contacto (John & Sali, 2003).

Por último, se optimiza y se refina todo el modelo según criterios energéticos como lo es la minimización de energía para cada átomo del modelo, para finalmente evaluar la calidad del

modelo obtenido para que las características estructurales cumplan con las reglas fisicoquímicas (Muhammed & Aki-Yalcin, 2019). Esta evaluación se realiza mediante el diagrama de Ramachandran que permite la visualización de los ángulos de torsión de la cadena polipeptídica con el fin de identificar si los ángulos son estereoquímicamente factibles (Hollingsworth & Karplus, 2010). Así, con el modelo de homología, se procede a evaluar las interacciones químicas entre la proteína y moléculas pequeñas, mediante el acoplamiento molecular.

1.3.2. Docking molecular.

El docking molecular es una herramienta clave para el diseño de fármacos asistido por computadora. El objetivo del acoplamiento proteína-ligando es predecir los modos de unión predominantes de un ligando con una proteína de estructura tridimensional conocida (Leelananda & Lindert, 2016). El docking molecular consta de un proceso de dos pasos: el primer paso es acoplar el ligando en el sitio activo de la proteína, obteniendo diferentes conformaciones del complejo proteína-ligando, y, el segundo paso es clasificar estas conformaciones mediante una función de puntuación (Liu & Wang, 2015; Rapp et al., 2009).

1.3.2.1. Funciones de puntuación.

Las funciones de puntuación son algoritmos matemáticos de muestreo que permite la visualización de diferentes modos de unión que pueden existir entre dos moléculas químicas, teniendo como parámetros los seis grados de libertad de traslación, rotación y de conformación (Meng et al., 2011). El principal objetivo de las funciones de puntuación es seleccionar los complejos proteína-ligando correctos. Las funciones de puntuación se pueden dividir en tres: Funciones de puntuación basadas en campos de fuerza (Fischer et al., 2008; N. Huang et al., 2006), funciones de puntuación empíricas (Artemenko, 2008; Martin & Schomburg, 2008) y funciones de puntuación basadas en conocimiento (Gohlke et al., 2000; S.-Y. Huang & Zou, 2006).

1.3.2.1.1. Función de puntuación basada en campo de fuerza.

Se desarrollan utilizando la mecánica molecular clásica, donde las interacciones electrostáticas y las interacciones de van der Waals contribuyen a la energía de interacción del complejo ligando-proteína (Leelananda & Lindert, 2016). El software AutoDock (Huey et al., 2012) utiliza este tipo de funciones de puntuación que es representada por la **Ecuación 1**.

Ecuación 1. Función de puntuación basada en campos de fuerza de AutoDock.

Tomado de (Meng et al., 2011).

$$V = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i,j} \left(\frac{q_i q_j}{e(r_{ij}) r_{ij}} \right) + W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\sigma^2)}$$

Como se muestra en la **Ecuación 1**, para dos átomos *i*, *j*, la energía atómica por pares se evalúa mediante la suma de van der Waals, puentes de hidrógeno, energía de Coulomb y la desolvatación, siendo *W* factores ponderados para calibrar la energía libre empírica (Meng et al., 2011). Esta función de puntuación presenta un problema que consiste en un efecto de desolvatación (Friesner et al., 2004a), es decir, los grupos cargados son favorecidos por ambientes acuosos, pero los grupos no polares permanecen en ambientes no acuosos, este efecto favorece al ligando que presenten una alta carga, sin embargo, esto no es lo apropiado porque una función de puntuación debe evaluar todos los efectos antes y después de la formación del complejo, para este caso en específico, la función debe evaluar el efecto de solvente antes de la formación y como este solvente interactúa con el complejo (Sethi et al., 2019). Con lo mencionado anteriormente es necesario implementar otras funciones de puntuación que permitan calcular valores energéticos de manera independiente con el fin de optimizar el proceso de selección del ligando.

1.3.2.1.2 Función de puntuación empírica.

Se obtienen utilizando datos de estructuras determinadas experimentalmente y ajustando esta información a parámetros establecidos, de tal manera, que la energía libre se calcula como la suma ponderada de los términos que no están correlacionados (Leelananda & Lindert, 2016). Un ejemplo de función de puntuación empírica es la **Ecuación 2**, utilizada por GlideScore que cuenta con bastantes variables de energía independientes con el fin de optimizar la selección de compuestos con fuerte afinidad de aquellos que con baja capacidad de unión.

Ecuación 2. Funciones de puntuación empírica de GlideScore.

Tomado de Friesner et al., 2004a.

$$\Delta G_{binding} = \Delta G_0 \Delta G_{bond} \sum_{il} g_1(\Delta r) g_2(\Delta \alpha) + \Delta G_{metal} \sum_{aM} f(r_{aM}) \\ + \Delta G_{lipo} \sum_{iL} f(r_{iL}) + \Delta G_{rot} H_{rot}$$

Como se puede observar en la **Ecuación 2**, ΔG es la energía libre estimada de unión, ΔG_0 es la constante de regresión y los términos ΔG_{bond} , ΔG_{metal} , ΔG_{lipo} , ΔG_{rot} , son coeficientes de regresión para cada termino de energía libre correspondiente. Ahora bien, el término de puentes de hidrógeno se calcula para todas las posiciones complementarias de puentes de hidrógeno entre los átomos del ligando y receptor. El término Δr es la desviación de la longitud del puente hidrógeno H...O/N con respecto a 1,85 Å y $\Delta \alpha$ es la desviación del ángulo del puente de hidrógeno N/O-H...O/N con respecto a 180° (Friesner et al., 2004a). El término metálico se calcula para todos los átomos aceptores y aceptores/donores en el ligando y cualquier átomo metálico del receptor. El término lipofílico se calcula para todos los átomos lipofílicos del ligando y del receptor. El último término identifica los enlaces rotatorios congelados. Los enlaces se consideran congelados si los átomos a ambos lados del enlace rotatorio están en contacto con el receptor (Eldridge et al., 1997). Adicionalmente, Glide utiliza dos formas de GlideScore: GlideScore SP y GlideScore XP,

ambas utilizan términos similares pero están formuladas con objetivos diferentes, la primera es una función más sencilla que se utiliza para ligandos que tienen alta probabilidad de unión a la proteína, mientras que GlideScore XP es una función más compleja que impone restricciones a las posiciones que violan los principios fisicoquímicos establecidos, por ejemplo, analiza grupos cargados y fuertemente polares que pueden estar expuestos en el disolvente (Friesner et al., 2004b).

Por último, las funciones de puntuación empírica poseen una metodología similar a las funciones de puntuación basadas en conocimiento, puesto que, ambas parten de análisis experimentales, sin embargo, se diferencian en que las empíricas solo analizan la estructura cristalina de la proteína mientras que las funciones basadas en conocimiento analizan el complejo proteína-ligando.

1.3.2.1.3 Función de puntuación basado en conocimiento.

Las funciones basadas en conocimiento son potenciales estadísticos y se derivan de información que ha sido determinada experimentalmente sobre el complejo proteína-ligando (Leelananda & Lindert, 2016). El análisis estadístico se centra en las estructuras cristalinas de complejos para obtener las frecuencias de contacto interatómico entre la proteína y el ligando donde se toma como base cuanto más fuerte sea una interacción, mayor será la frecuencia de su aparición. La función tiene en cuenta los contactos favorables y las interacciones repulsivas entre cada átomo del ligando, colocando a la proteína dentro de una esfera o grid con un límite específico (Sethi et al., 2019). Un ejemplo claro de función de puntuación basada en conocimiento es la **Ecuación 3**, utilizada por el software Potential of Mean Force (PMF).

Ecuación 3. Funciones de puntuación basadas en el conocimiento PMF.

Tomado de (Meng et al., 2011)

$$PM_{score} = \sum_{r < r_{cut-off}^{ij}} k_B A_{ij}(r) \quad A_{ij}(r) = -k_B T \ln \left[f_{Vol-Corr}^{ij}(r) \frac{\rho_{seg}^{ij}(r)}{\rho_{bulk}^{ij}} \right]$$

En la ecuación 3, k_B es la constante de Boltzmann; T es la temperatura absoluta; r la distancia del par de átomos $f_{Vol-Corr}^{ij}(r)$ es el factor de corrección del volumen del ligando; $\frac{\rho_{seg}^{ij}(r)}{\rho_{bulk}^{ij}}$ hace referencia a la función de distribución radial de un átomo de proteína y un átomo del ligando (Meng et al., 2011). La función de puntuación basada en conocimiento tiende a ser simples computacionalmente debido a que se tiene una información previa sobre el complejo, sin embargo, cuando se quiere estudiar interacciones con halógenos y metales del complejo se presenta un gran problema (Sethi et al., 2019).

Estos cálculos de predicción de unión entre ligando y proteína que son obtenidos mediante docking molecular son de baja confianza, es por ello, que se complementan con otros métodos computacionales como dinámica molecular que permiten la estimación del efecto de moléculas de agua y calcula con mayor precisión la energía de afinidad (Santos et al., 2019)

1.3.3. Dinámica Molecular.

La dinámica molecular es una técnica en rápida evolución y crecimiento. La metodología de esta técnica permite calcular las propiedades dinámicas y termodinámicas de equilibrio de un sistema a temperatura finita (Iftimie et al., 2005). Asimismo, las simulaciones de dinámica molecular también se pueden utilizar para estudiar interacciones proteína-ligando, pero de una manera más detallada. De esta forma, las simulaciones de dinámica molecular emplean ecuaciones que se resuelven en pequeñas variaciones de tiempo para calcular las fuerzas que actúan sobre cada partícula de un sistema determinando las nuevas posiciones y velocidades en función del

tiempo (Konteaty et al., 2011). Esta variación del movimiento se realiza bajo un campo de fuerza que se aplica a cada uno de los átomos y se describe mediante la **Ecuación 4**.

Ecuación 4. Segunda ley de Newton para cálculos en dinámica molecular.

Tomado de Tuckerman & Martyna, 2000.

$$F_i = m_i \ddot{r}_i$$

En esta ecuación, m_i representa las masas de N partículas, r_i la posición de las partículas en función del tiempo y F_i la fuerza de cada una de las partículas en función de cada una de las posiciones (Tuckerman & Martyna, 2000). Ahora bien, al involucrar la aceleración a partir de las fuerzas que actúan sobre un átomo i, y se calculan las fuerzas a partir del campo de fuerza ejercido sobre el átomo (**Ecuación 5**).

Ecuación 5. Segunda ley de Newton con la aceleración para cálculos en dinámica molecular.

$$a_i(t) = \frac{d^2 r_i(t)}{dt^2} = - \frac{dV(r(t))}{m_i dr_i(t)}$$

Donde $dV(r(t))$ representa la función de energía potencial ejercida por el campo de fuerza sobre cada átomo de la proteína (De Vivo et al., 2016), asimismo, el modelo matemático que surge para esta energía potencial está representado por la **Ecuación 6**.

Ecuación 6. Función energía potencial.

$$= \sum_i^{bond} \frac{k_{li}}{2} (r_i - l_{0,i})^2 + \sum_i^{angles} \frac{k_{ai}}{2} (\alpha_i - \alpha_{0,i})^2 + \sum_i^{torsion} \left\{ \sum_k^M \frac{V_{ik}}{2} [1 + \cos(\alpha_{ik} * \theta_{ik} - \theta_{0,ik})] \right\} \\ + \sum_{i,j}^{pairs} \left[\left(\frac{r_{0,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0,ij}}{r_{ij}} \right)^6 \right] + \sum_{i,j}^{pairs} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$

Como se puede observar en la **Ecuación 6**, los primeros términos representan todas las interacciones intramoleculares de los átomos de manera descriptiva, es decir, tomas las variaciones

en la energía potencial como una función del estiramiento, la flexión y la torsión de los enlaces de los átomos involucrados, asimismo, representan sumas sobre longitudes de enlace, ángulos y ángulos diédricos, respectivamente (De Vivo et al., 2016).

Los dos últimos términos, representan las interacciones electrostáticas y de Van der Waals entre átomos y se denotan como “no enlazados”, estas contribuciones actúan sobre cada átomo de la proteína de manera que el sistema aun no este cubierto por la parte enlazada (De Vivo et al., 2016). Por consiguiente, la energía se expresa como función potencia inversa de la distancia entre átomos r_{ij} . Ahora bien, las fuerzas no ligadas surgen debido a las interacciones de Van der Waals, modeladas utilizando el potencial Lennard-jones 6-12 e interacciones cargadas que son modeladas por la ley de Coulomb (Durrant & McCammon, 2011).

Para realizar estudios de dinámica molecular, es necesario el uso de campos de fuerza empíricos que actúen sobre las moléculas de interés, estos campos son utilizados por los respectivos softwares para estimar la energía y la fuerza asociada entre una proteína y un ligando (Yu & MacKerell, 2017).

1.3.4. Campos de Fuerza.

En una simulación de dinámica molecular los átomos no se encuentran aislados, por el contrario, están sujetos a fuerzas que pueden ser interatómicas o intermoleculares, estas fuerzas se pueden denominar campos de fuerzas. Los campos de fuerza son un conjunto de ecuaciones matemáticas que calculan la energía potencial de un sistema en función de las coordenadas atómicas, de esta manera, se pueden generar comportamientos del sistema, por ejemplo, flexiones de ángulos, ángulos impropios, estiramiento de enlaces, entre otros (Damm et al., 1997).

Los campos de fuerza más utilizados en la literatura son AMBER, CHARMM; GROMOS Y OPLS-AA (Bowers et al., 2006; Brooks et al., 2009; Case et al., 2005; Páll et al., 2015). Estos

cuatro campos de fuerzas expresan la energía potencial de un sistema como la suma de interacciones enlazantes y no enlazantes mediante la **Ecuación 7**.

Ecuación 7. Energía potencial total

Tomado de (Y. Wang et al., 2010)

$$E_{Vtotal} = \sum E_{enlazantes} + \sum E_{coulomb} + \sum E_{LJ}$$

Las interacciones enlazantes dependen de la estructura del enlace covalente en las moléculas, por tanto, se puede utilizar variables que expresen la longitud y ángulo del enlace, ahora bien, las interacciones no enlazantes son expresadas por interacciones electrostáticas e interacciones de van der Waals y potencial de Lennard-Jones (Heffelfinger & Swol, 1994).

Por otra parte, los cálculos de acoplamiento molecular y dinámica son expresados en términos de energía libre de unión, por tanto, existen métodos para determinar este valor, de esta manera, los métodos más utilizados son *Molecular Mechanics Poisson-Boltzmann Surface area* (MM/PBSA) y *Molecular Mechanics Generalized-Born Surface area* (MM/GBSA(Hou et al., 2011; E. Wang et al., 2019)), ambos métodos describen la energía de unión como la diferencia entre la energía de unión del complejo y la energía de unión del ligando y receptor, por tanto, estos métodos tienen presente los cambios energéticos que se pueden presentar antes y después de la conformación del complejo, pero, también tiene presente los cambios energéticos del receptor y ligando de manera independiente (Hou et al., 2011; E. Wang et al., 2019). En el presente estudio se utilizó métodos computacionales como docking molecular y dinámica molecular para encontrar posibles inhibidores a los sistemas binarios CymR-VP16 y TtgR-VP16.

1.4. Usos de sistemas binarios para la generación de machos estériles.

La principal técnica que se ha propuesto para el control de insectos de interés médico y agrícola es la liberación de machos estériles SIT (Sterile Insect Technique), la cual ha sido estudiada desde los años 1940 donde se ha utilizado la irradiación como principal fuente de esterilidad (Hendrichs & Robinson, 2009). Sin embargo, este proceso es costoso y con una alta dificultad, puesto que, se pueden generar el mismo número de machos y hembras, por lo cual, es necesario separar manualmente desde el estadio de pupas las moscas para obtener machos estériles que contribuyen a la disminución poblacional (Thomas et al., 2000). Es por ello, que se ha implementado el estudio de diferentes técnicas que permitan la regulación de la expresión génica y que puedan servir como un sistema de control poblacional.

No obstante, es importante tener presente que a largo plazo se espera no eliminar la especie que se quiere controlar, puesto que, no se conoce el impacto ambiental que se pueda generar, es por ello, que se pretende crear un sistema controlable que pueda ser usado únicamente cuando la población del insecto de interés este elevada (Phuc et al., 2007), o cuando los insectos se alimenten de humanos o de cultivos de interés agrícola promoviendo la transmisión de enfermedades y plagas.

Es por ello, que se están desarrollando métodos genéticos basados en transgenes para insectos de importancia agrícola y sanitaria. Por ejemplo, se desarrolló un sistema de letalidad específico para hembras en *Bombyx mori* (Tan Anjiang et al., 2013). Este sistema exitoso tiene una función autorreguladora, es decir, que actúa como transactivador e inductor, por lo cual, se induce la letalidad de las hembras durante las etapas embrionarias y larvarias tempranas debido a la acumulación masiva de la tTAV (proteína transactivadora de la tetraciclina) en ausencia de tetraciclina en la dieta (Tan Anjiang et al., 2013). Asimismo, en moscas de la fruta se ha

desarrollado un sistema que aprovecha un transactivador de transcripción reprimible por tetraciclina (TRTT) (Schneider, 2000). Este sistema consiste en crear moscas de la fruta que expresen el gen que codifica TRTT de modo que la expresión se limite a las hembras, por lo cual, permite una fácil selección de los machos, puesto que, en ausencia de tetraciclina, todas las crías hembras mueren, mientras que los machos no se ven afectados por el tratamiento con tetraciclina porque nunca expresan el transactivador (Schneider, 2000).

De esta manera, las proteínas represoras son una herramienta alternativa, la cual permite la unión al ADN o ARN para inhibir la expresión de uno o más genes. Asimismo, muchas de estas proteínas pueden ser reguladas con metabolitos o compuestos que pueden activarlas o inhibirlas (Kaczmarczyk et al., 2013).

1.5. Uso de proteínas represoras reguladas por una molécula pequeña.

En la literatura se han reportado proteínas represoras reguladas con compuestos químicos en células en cultivo demostrando alta efectividad, algunos de estos estudios se describen a continuación. Fussenegger y colaboradores (2000) describieron el represor Pip (pristinamicin-induced protein) (Fussenegger et al., 2000). De esta manera, se diseñó mediante ingeniería genética un sistema de expresión inducible por antibióticos de streptogramina (pristinamicina, virginamicina y sinercid) obteniendo una alta eficacia en la activación y represión de genes clonados (Fussenegger et al., 2000). Adicionalmente, existen antibióticos de streptogramina, como la mezcla quinupristina/dalfopristina (Bergeron & Montay, 1997), que presenta mayor biodisponibilidad y mejor farmacocinética en animales, por lo cual, este sistema posee una significativa ventaja en la regulación, puesto que, puede ser inhibido por diferentes derivados antibióticos de streptogramina (Fussenegger et al., 2000).

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

Del mismo modo, Gossen y Bujard (1992) fusionaron el receptor de tetraciclina con el activador de VP16 y fue adaptado para células HeLa (cáncer de cuello uterino) acoplado a la expresión de luciferasa, con este sistema ellos observaron una eficiente regulación de la expresión de luciferasa usando concentraciones muy bajas, cercanas a 1 µg/mL de tetraciclina (Gossen & Bujard, 1992), por lo cual, este sistema es el más antiguo con eficiente capacidad de regulación *in vitro* (Gossen & Bujard, 1992).

Asimismo, Mullick y colaboradores (2006) usaron un represor bacteriano CymR para regular la expresión génica, por lo cual, diseñaron una proteína binaria compuesta del represor CymR unido al factor de transcripción VP16, muy parecido al sistema que será usado en este proyecto. Este sistema tiene la capacidad de inducir la expresión génica, sin embargo, en presencia de cumato se anula la unión al ADN, por lo cual, se suprime la expresión génica (Mullick et al., 2006).

Por último, Gamez y colaboradores (2021) evaluaron varios sistemas transactivacionales binarios derivados de bacterias en *Drosophila melanogaster*, incluyendo el operón p-CymR de *Pseudomonas putida*, el operón PipR de *streptomyces coelicolor*, el operón TtgR de *Pseudomonas putida* y el operón VanR de *Caulobacter crescentus*. Este trabajo proporcionó la primera caracterización de estos sistemas en un modelo animal *in vivo* (Gamez et al., 2021).

Aunque se demostró una sólida transactivación espacial específica del tejido de la expresión del gen reportero de luciferasa para cada sistema, surgió una problemática en la supresión efectiva de la expresión génica en los músculos de vuelo. Esta dificultad plantea dos hipótesis: en primer lugar, los compuestos administrados (cumato, virginiamicina, floretina y ácido vanílico, respectivamente) no llegaron adecuadamente a los músculos de vuelo debido a la complejidad de la ruta que debían seguir a través del cuerpo de las moscas, enfrentando obstáculos

en su paso por el intestino medio, el transporte a través de la hemolinfa y llegar a los músculos de vuelo. En segundo lugar, se considera que las concentraciones de los compuestos utilizados podrían no haber sido suficientes para lograr una supresión completa de la expresión génica en los músculos de vuelo. Por lo tanto, este estudio resalta la necesidad de investigaciones adicionales para comprender de una mejor manera las razones subyacentes por las cuales estos compuestos no funcionaron según lo esperado y para explorar posibles estrategias destinadas a mejorar la supresión de la expresión génica en tejidos específicos de *Drosophila melanogaster* (Gamez et al., 2021).

Por todo lo mencionado anteriormente, este trabajo de investigación cumplió los siguientes objetivos:

2. Objetivos.

2.1. Objetivo general.

Identificar compuestos derivados del cumato y floretina que puedan suprimir la expresión génica *in vivo* en moscas *Drosophila melanogaster* modificadas genéticamente para expresar proteínas represoras CymR y TtgR y proteínas GFP.

2.2. Objetivos específicos.

-Determinar compuestos derivados del cumato y floretina mediante un análisis *in silico* que puedan interaccionar con las proteínas binarias CymR-VP16 y TtgR-VP16.

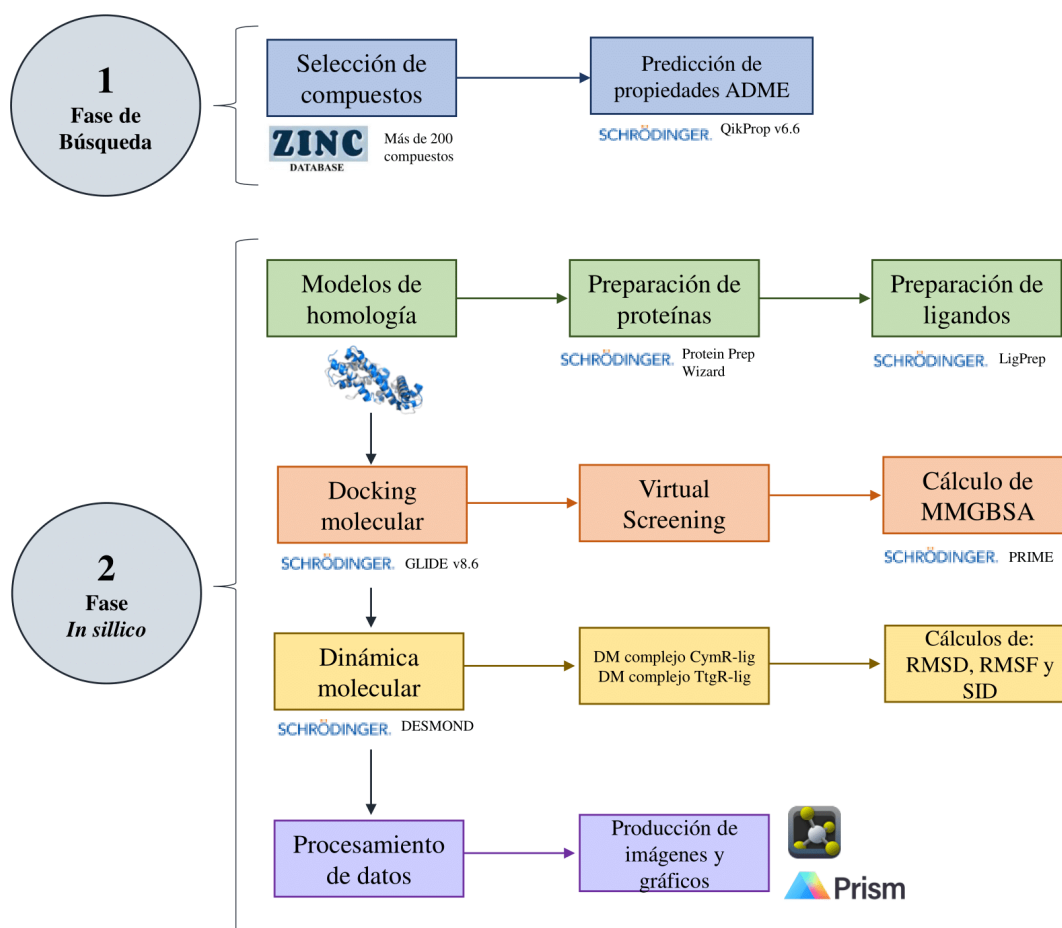
- Validar la estabilidad del complejo proteína-ligando mediante un estudio de dinámica molecular.

3. Metodología.

Este trabajo de grado se llevó a cabo siguiendo el esquema metodológico descrito en la **Figura 1**. Como primera instancia, se creó una base de datos de aproximadamente 210 compuestos

derivados comerciales de cumato y floretina. Posteriormente se hizo todo el estudio computacional de las dos proteínas y sus respectivos ligandos donde se identificaron los mejores ligandos según las interacciones aminoacídicas y puntaje de acoplamiento. Asimismo, se realizaron simulaciones de dinámica molecular para estudiar la estabilidad de la interacción para el complejo.

Figura 1. Esquema metodológico para el desarrollo del proyecto de investigación



3.1 Fase preliminar.

En esta primera fase se realizó una búsqueda bibliográfica sobre las dos proteínas CymR y TtgR para conocer su actividad biológica, sitios activos, residuos aminoacídicos que participan en el sitio de unión de las proteínas y estructura cristalina en general. Asimismo, se hizo una búsqueda

en base de datos con posibles derivados comerciales del cumato y la floretina. De esta manera, los compuestos se descargaron de las bases de datos Zinccdatabase, DrugBank y PubChem en formato SDF y, posteriormente, se importaron en el software Maestro Schrödinger.

Posteriormente, se realizó un análisis con la herramienta Qikprop v6.6 presente en el paquete Drug Discovery de Schrödinger para determinar los parámetros fisicoquímicos absorción gastrointestinal, coeficiente de reparto (LogP) y coeficiente de solubilidad en agua (LogS) de cada compuesto y su efecto genotóxico (Ioakimidis et al., 2008). Los compuestos que presentaron una alta capacidad de permear membranas (conforme a los parámetros de Lipinski) se analizaron mediante docking molecular con las proteínas represoras como se describe a continuación.

3.2 Fase *in silico*.

3.1.1. Modelo de Homología.

Dado que, la proteína CymR no tiene reportado en la literatura una estructura cristalina, es necesario realizar un modelo de homología, es decir, realizar una predicción de la estructura tridimensional de la proteína utilizando como template la secuencia y estructura conocida de una proteína que presente una alta similitud con la proteína CymR.

En este orden de ideas, la estructura primaria de la proteína CymR-VP16 se obtuvo a partir del plásmido reportado bajo el código p9O7H1 en addgene, el cual fue diseñado en la Universidad de California San Diego (UCSD) (Gamez et al., 2021). Ahora bien, para la selección del template modelo se usó criterios de similitud con secuencias reportadas en la base de datos Protein Data Bank, mediante el software BLAST Homology Search (John & Sali, 2003). Mediante la herramienta Prime Schrödinger-Relased 2021-1 con la opción Prime-STA se hizo un alineamiento de secuencias para obtener una estructura terciara semejante a la del template modelo (Jacobson et al., 2004). El modelo de homología se creó bajo el criterio de minimización de energía con un

modelo de solvatación VSGB y campo de fuerza OPLS3e (Li et al., 2011; Roos et al., 2019). Por último, se realizó un procedimiento de refinamiento y optimización a la estructura proteica obtenida con el asistente de proteínas Protein Preparation Wizard (Madhavi Sastry et al., 2013) que se ubica en la configuración de la herramienta Refine Loops del Software Maestro Schrödinger que tiene como finalidad solucionar problemas estructurales.

3.1.2. Preparación de proteína y ligandos.

El modelo de homología presentado en la sección anterior ha sido refinado, eliminando así la necesidad de realizar preparaciones adicionales, puesto que, se encuentra en un estado óptimo para futuros estudios. En contraste, la proteína TtgR requiere un proceso de preparación con el fin de alcanzar estados con energía minimizada y una mayor estabilidad, lo que contribuirá a reducir posibles errores computacionales en las investigaciones subsiguientes.

En este contexto, la proteína TtgR posee una estructura cristalina optima de código PDB 2UXI con una resolución de 2,50 Å reportada en la base de datos Protein Data Bank (Alguel et al., 2007a). De esta manera, se procedió a preparar la proteína TtgR y los derivados de cumato y floretina con el fin de obtener estados con energía minimizada y estables para reducir los errores computacionales posteriores. Además, esta preparación en los ligandos permite corregir posibles errores en las estructuras de Lewis.

La estructura cristalina para la proteína TtgR-VP16 se preparó con la herramienta Protein Preparation Wizard (PrepWiz) con la opción Protein Preparation Tools, donde se agregó todos los átomos de hidrógeno polares, se optimizó los puentes de hidrógeno y se obtuvo todos los estados de protonación a un pH fisiológico (7.0 ± 0.2) (Madhavi Sastry et al., 2013; Martin & Schomburg, 2008), además, con la herramienta Prime de Schrödinger-Released 2021-1, se pudo identificar

cadena laterales y residuos faltantes que se pudo corregir mediante la minimización del sistema con un RMSD debajo del umbral y un campo de fuerza OPLS3e (Roos et al., 2019).

Por otra parte, los candidatos a inhibidores se prepararon con la herramienta LigPrep de Schrödinger que permite obtener las configuraciones de quiralidad y estereoisomería con la más baja energía (Shelley et al., 2007), además, se obtienen estados de protonación y tautómeros basados en el valor de pKa teórico calculado y otros ajustes predeterminados del programa.

3.1.3. Screening virtual de derivados del cumato y Floretina.

En primer lugar, se hizo una identificación y evaluación de posibles sitios drogables en la estructura cristalina, para ellos, se utilizó la herramienta de SiteMap de Schrödinger. El algoritmo de SiteMap permite localizar sitios de unión con un alto grado de confianza y puede predecir la drogabilidad de estos sitios, además, permite la optimización de los sitios al proporcionar información de las interacciones ligando-receptor con el fin de mejorar el acoplamiento del ligando (Halgren, 2009).

Seguidamente, por parámetros de DrugScore y SiteScore que se obtuvieron del algoritmo de SiteMap, se seleccionó el mejor sitio de unión para los ensayos de docking molecular. Para el estudio del acoplamiento de los ligandos preparados se realizó usando Glide (Glide v8.6), (Friesner et al., 2004a) con la opción de acoplamiento flexible de ajuste inducido y sin ajuste inducido con precisión estándar (SP) y extra (XP) . Los ligandos se acoplaron dentro de un grid de 30Å y 22 Å (lo suficientemente grande para abarcar toda la proteína), generando al menos 20 poses por ligando (Sherman et al., 2006). Posteriormente, el refinamiento de los complejos de proteína-ligando se realizó mediante la herramienta Prime con opciones estándar (Jacobson et al., 2004). En este paso, se minimizó todas las cadenas laterales dentro de los 6 Å de cada pose de ligando acoplado. Por

último, las poses de acoplamiento se seleccionaron mediante inspección visual y puntuación de docking score en función de las interacciones de los aminoácidos relevantes.

3.1.4. Simulación de Dinámica Molecular.

Las poses de acoplamiento seleccionadas se validaron por medio de una simulación de dinámica molecular (MD), con el fin de estudiar la estabilidad del ligando y sus interacciones dentro del sitio de unión. Para ello, las simulaciones de MD se realizaron con la herramienta Desmond de Maestro Schrödinger (Bowers et al., 2006), utilizando el campo de fuerza OPLS4 que permite un mejor rendimiento en la predicción de afinidades de unión proteína-ligando. Asimismo, mediante la herramienta System Builder, los complejos proteína-ligando, se introdujeron en dos cajas ortorrómbicas con distancias de $a=1.0 \text{ \AA}$, $b=1.0 \text{ \AA}$ y $c=1.5 \text{ \AA}$ (CymR) y $a=1.0 \text{ \AA}$, $b=1.0 \text{ \AA}$ y $c=1.8 \text{ \AA}$ (TtgR) (Lanig et al., 2006) desde los bordes de la caja hasta cualquier átomo de las proteínas, se llenó con agua tipo TIP3P (Jorgensen et al., 1983) y se neutralizó con la adición de iones Na^+ y Cl^- . Luego, todos los sistemas fueron equilibrados mediante simulaciones cortas bajo el conjunto NPT durante 5ns a temperatura constante de 300K y 1 atm de presión, utilizando el termostato y baróstato Berendsen (Harger & Ren, 2019; Lemak & Balabaev, 1994; Vesga et al., 2021). Posteriormente, se hizo una producción para cada sistema de 200ns para obtener un total de 10.000 frames y se realizó al menos tres réplicas independientes. Por último, la estabilidad del ligando dentro del sitio de unión y los cambios conformacionales de la proteína se evaluaron mediante el seguimiento del Root-Mean-Square-Deviation (RMSD) y las interacciones proteína-ligando que se calcularon con la herramienta Simulation Interaction Diagram (SID).

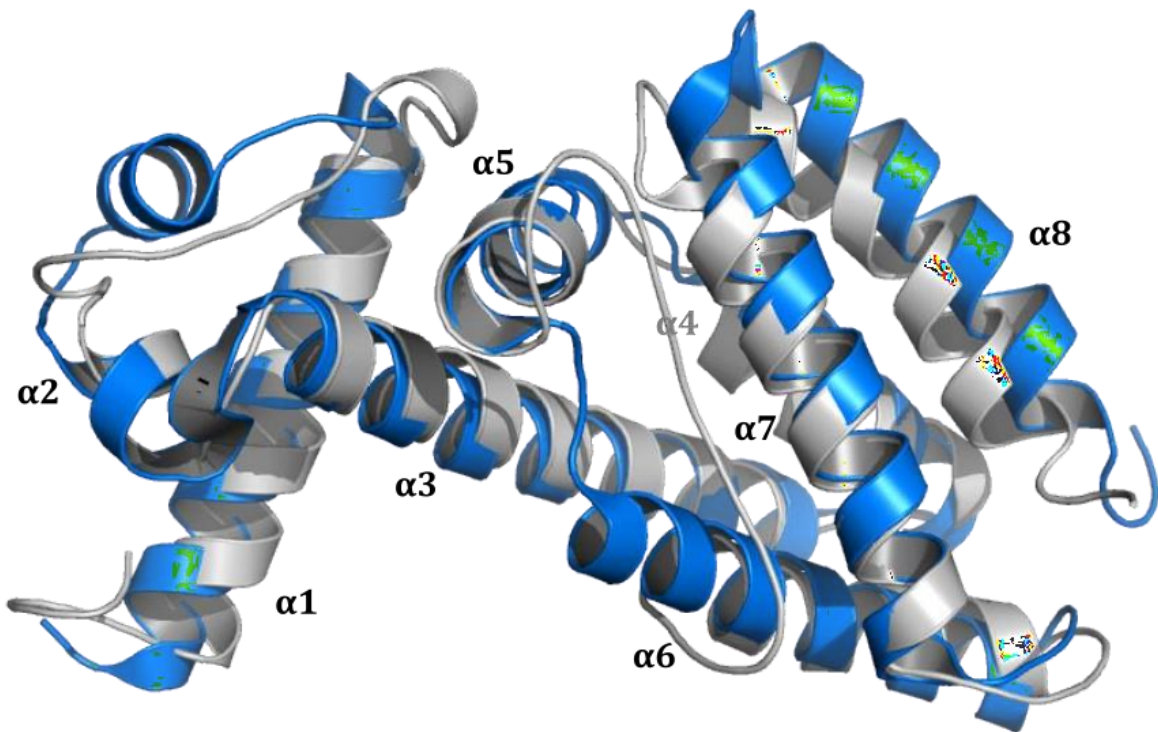
4. Resultados y Discusión

4.1. Proteína CymR

4.1.1. Resultados de Modelo de Homología.

El modelo de homología para la proteína CymR (**Figura 2**) se diseñó utilizando como target la secuencia del plásmido p9O7H1, que se encontraba en formato de lenguaje de nucleótidos. Para llevar a cabo el modelamiento, se realizó la respectiva transducción del plásmido con la herramienta Blastx del sitio web NCBI BLAST (National Center for Biotechnology Information) obteniendo la estructura primaria de la proteína CymR. Luego, se utilizó como template la estructura cristalina bajo el código PDB: 2HYT (1.64Å) (Joint Center for Structural Genomics, 2006), donde se obtuvo una identidad del 40% y un score de 110, con su respectivo refinamiento y minimización del sistema.

Figura 2. Superposición modelo de Homología con template 2HYT.



Nota Elaboración propia (2023). Superposición del modelo de homología obtenido para la proteína CymR (color blanco) y la proteína template de código PDB 2HYT (color azul).*

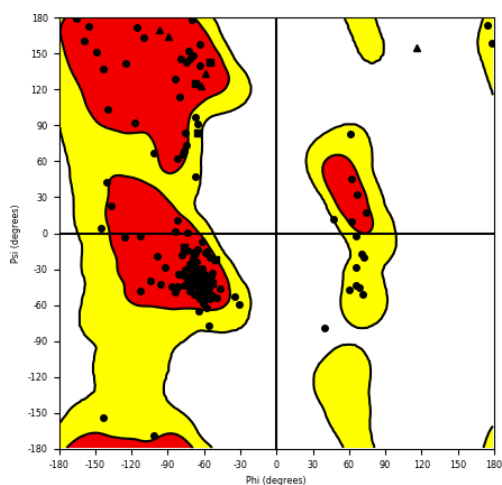
La proteína template (**Figura 2**) es un regulador transcripcional de la familia TetR que contiene dos dominios relevantes en su estructura. El primero es un dominio de unión al ADN en el extremo N-terminal de tipo hélice-giro-hélice (HTH) y el segundo es un dominio central regulador involucrado en la dimerización y la unión del ligando (Ramos Juan L. et al., 2005). Los dominios de unión al ADN están constituidos por hélices $\alpha 1$, $\alpha 2$ y $\alpha 3$, mientras que la hélice $\alpha 4$ es un puente entre los dos dominios. Por otra parte, el dominio central este compuesto por las $\alpha 5$, $\alpha 6$, $\alpha 7$, $\alpha 8$, $\alpha 9$ y $\alpha 10$, causantes de la dimerización y del acoplamiento de la tetraciclina mediante un catión divalente (Ramos Juan L. et al., 2005).

4.1.2. Validación del modelo de Homología CymR.

La proteína modelada consta de 257 aminoácidos con una desviación cuadrática media (RMSD) de 2.640 Å (Rajagopalan et al., 2013), este valor se calcula tomando solo los residuos que se han alineado con éxito, por tanto, un valor cercano a 2.3 Å, indica que la estructura tiene regiones flexibles y óptimas para sus posteriores estudios (Kufareva & Abagyan, 2012). Asimismo, se obtuvo un puntaje de 0.283 en el alineamiento, donde es importante resaltar que una puntuación superior a 0,7-0,8 indica que hay una falla en el cálculo estructural y por ende no hay suficiente similitud estructural (Schrödinger Release, 2021).

Ahora bien, según el análisis la gráfica de Ramachandran correspondiente al modelo (Figura 3), se pudo observar que únicamente 2 residuos de aminoácidos (representando un 0,77% del total) se encuentran fuera de la región permitida en el diagrama. Esta observación indica que la conformación tridimensional del modelo es confiable y óptima para llevar a cabo este modelo previsto.

Figura 3. Gráfico de Ramachandran del modelo de homología.



Nota Elaboración propia (2023). Los puntos negros son la representación de cada aminoácido de la estructura primaria de la proteína CymR. Asimismo, las superficies rojas y amarillas representan las regiones permitidas para obtener un modelo de homología óptimo.*

4.1.3. Identificación de sitios drogables por SiteMap para la proteína CymR.

Como se mencionó anteriormente, la proteína CymR carece de estudio estructural, es por ello, que se realizó un estudio estructural con la herramienta SiteMap de Maestro Schrödinger con el fin de encontrar los posibles sitios de unión del ligando.

SiteMap es una técnica computacional para identificar posibles sitios de unión y predecir su farmacabilidad en aplicaciones de descubrimiento de nuevos fármacos, caracterizar sitios de unión y evaluar de manera rigurosa los posibles ligandos en aplicaciones de optimización de

nuevos fármacos (Santiago et al., 2011). SiteMap identifica correctamente el sitio de unión en más del 86% de los casos donde la estabilidad del ligando mejora la calidad de puntuación del sitio (Halgren, 2009). De esta manera, SiteMap proporciona datos cuantitativos y gráficos que permite estudiar la estructura cristalina de una proteína con el fin de aumentar la estabilidad del ligando o las propiedades físicas de la proteína (Santiago et al., 2011). El cálculo de SiteMap comienza con una búsqueda que determina una o más regiones que pueden estar cerca de la superficie de la proteína, los cuales se denominan sitios. En la segunda etapa, se generan unos mapas del sitio donde se tienen en cuenta las características hidrofílicas e hidrofóbicas del sitio. Los mapas hidrófilos se dividen a su vez en regiones donadoras,ceptoras y de unión a metales. Por último, una etapa de evaluación, que concluye el cálculo de dos parámetros SiteScore y DrugScore (Halgren, 2009). SiteScore se basa en una suma ponderada del número de sitio (n) con un límite de 100, puntuación del recinto (e) y la puntuación hidrófila (p) que posee un límite de 1,0. Por otra parte, DrugScore contiene términos que promueven la unión de los ligandos, es por ello, que utiliza las mismas propiedades de SiteScore, pero con diferentes coeficientes (Santiago et al., 2011) y condiciones, por ejemplo, la puntuación hidrofílica no tiene límite. La **Ecuaciones 8** y la **Ecuación 9** representan el cálculo para SiteScore y DrugScore respectivamente.

Ecuación 8. Modelo matemático para obtener SiteScore en maestro Schrödinger

$$SiteScore = 0,0733sqrt(n) + 0,6688e - 0,20p$$

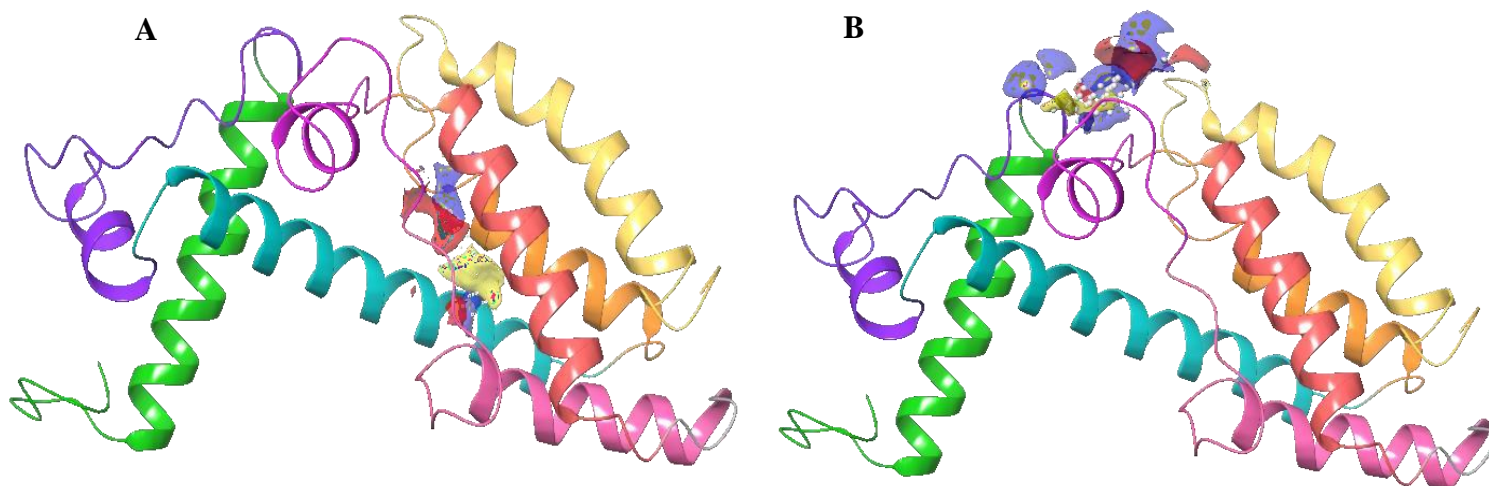
Ecuación 9. Modelo matemático para obtener DrugScore en maestro Schrödinger

$$DrugScore = 0,094qrt(n) + 0,60e - 0,324p$$

En este sentido, se usaron estos modelos matemáticos para identificar los posibles sitios de unión de la proteína CymR y su capacidad farmacológica. El estudio permitió identificar dos candidatos potenciales como posibles sitios de unión para la proteína CymR (**Figura 4**). Se pudo

observar que el primer candidato (**Figura 4A**) presenta un SiteScore (0.948) y DrugScore (0.826) abarcando las alfas hélices $\alpha 3$ (color cian), $\alpha 4$ (color naranja), $\alpha 6$ (color rosado), $\alpha 7$ (rojo) y $\alpha 8$ (amarillo). En esta configuración se evidencia que el sitio activo exhibe una superficie propicia a interacciones hidrofóbicas (amarilla), una superficie que actúa como donante de hidrógeno (azul) y una superficie capaz de aceptar hidrógeno (roja), por consiguiente, estas interacciones entre la proteína y ligando contribuyen a mejorar la estabilidad del complejo, como se ha documentado en la literatura (Meng et al., 2011). Además, es importante destacar que se ha demostrado en la literatura que sitios de unión con un SiteScore cercano a 1.0 pueden considerarse como sitios activos en una proteína (Halgren, 2009). No obstante, la presencia de un ligando con un tamaño considerable puede tener dificultades durante el proceso de acoplamiento molecular.

Figura 4. Sitios de unión obtenidos con la herramienta de SiteMap de Maestro Schrödinger.



Nota Elaboración propia (2023). Posibles sitios de unión; A: Primer sitio de unión, B: segundo sitio de unión; ambos representados de color blanco para la proteína de CymR, cada conjunto de aminoácidos está distribuidos en sus alfas hélices: $\alpha 1$ color verde, $\alpha 2$ color morado, $\alpha 3$ color ciano, $\alpha 4$ color naranja, $\alpha 5$ color magenta, $\alpha 6$ color rosado, $\alpha 7$ color rojo, $\alpha 8$ color amarillo.*

Por otra parte, el segundo candidato (**Figura 4B**) presenta un SiteScore (0.667) y DrugScore (0.667), y su entorno se limita a la presencia de la hélice α_6 (color magenta). En consecuencia, la interacción entre proteína y ligando puede ser afectada, dado que, este sitio activo presenta superficies propicias para la aceptación de hidrógeno. Aunque, un sitio con una alta hidrofiliidad puede implicar un mejor rendimiento, no se recomienda en estructuras con un RMSD mayor a 1,485 Å. Sin embargo, en casos en los que las estructuras superen los 2 Å de RMSD, resulta conveniente identificar sitios de unión que sean principalmente hidrofóbicos, pero que también presenten interacciones hidrofílicas (Chung et al., 2009). Por lo tanto, se optó por seleccionar el primer sitio activo en función de los valores de los parámetros de SiteScore y DrugScore para llevar a cabo estudios de docking molecular.

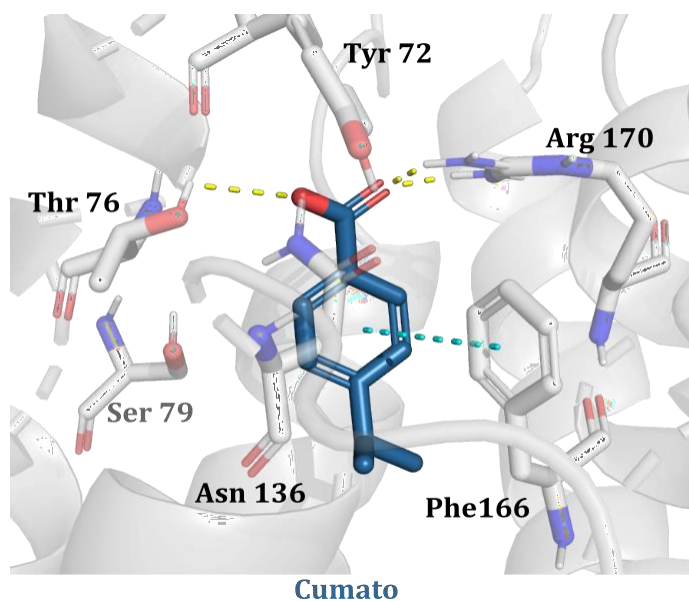
4.1.4. Docking molecular para CymR.

Para llevar a cabo el análisis de docking molecular específico en el sitio activo previamente definido (**Figura 4A**) se implementó un protocolo compuesto por tres fases distintas, diseñadas de manera rigurosa para una exploración detallada de las interacciones moleculares.

En primera instancia, se hizo un estudio de docking molecular entre el inhibidor de referencia cumato y la proteína CymR utilizando una caja de rejilla de 30 Å, centrando esta caja en las coordenadas del centro del sitio de unión (SiteMap) de la proteína (**Figura 4**). Es importante destacar que este sitio de unión comprende la hélice α_6 que desempeña un papel crítico en la función biológica de transducción de señales dentro de la familia de represores Rrf2 (Volbeda et al., 2017). Por lo tanto, este estudio inicial permitió evaluar la interacción entre el inhibidor de referencia cumato y la proteína CymR, con el propósito de identificar aminoácidos que puedan influir en la inhibición de la proteína,

El estudio permitió la identificación de interacciones entre el cumato y la proteína de interés. En detalle, se observaron interacciones tipo puente de hidrógeno con los aminoácidos Thr76 y Arg170. Además de una interacción tipo π - π con el aminoácido Phe166 (**Figura 5**). Estos aminoácidos pertenecen a las hélices α 3 y α 7 respectivamente. La hélice α 3 es esencial para el reconocimiento del ADN en la proteína CymR, por tanto, la interacción de los compuestos en estudio con esta hélice podría influir en el proceso de reconocimiento y afectar la capacidad de la proteína para unirse al ADN (Volbeda et al., 2017). Estos hallazgos proporcionan información valiosa sobre la naturaleza y a la fuerza de las interacciones moleculares que contribuyen a la afinidad y la actividad del compuesto en el sitio activo de la proteína CymR.

Figura 5. Visualización del inhibidor cumato para la proteína CymR.



Nota Elaboración propia (2023). Visualización del cumato (azul) en la proteína CymR. Cada conjunto de átomos de los residuos y del ligando poseen un código de color. Oxígenos: rojo; Nitrógenos: Azul. Las interacciones proteína ligando tipo puente de hidrógeno están*

representados por líneas discontinuas amarillas y las interacciones tipo π - π por líneas discontinuas azules.

En la segunda etapa del proceso de docking molecular, se realizó un screening virtual de la base de compuestos que constaba de 203 derivados del cumato (Apéndice 1). Este screening virtual incluyó un análisis de docking molecular con precisión estándar (SP), generando un conjunto de 5 poses por ligando. La selección de compuestos se basó en criterios que abarcaron Docking Score, propiedades farmacocinéticas y farmacodinámicas (ADME), consideraciones del valor comercial, así como interacciones específicas con residuos aminoacídicos relevantes. Como resultado de este proceso, se identificaron 55 compuestos prometedores.

En este sentido, los 55 compuestos se sometieron a un estudio de docking molecular de mayor rigor, utilizando el nivel de precisión extra-precisión (XP), lo que permitió obtener aproximadamente 15 poses por ligando. De este subconjunto más selectivo, se eligieron 10 derivados utilizando los mismos parámetros de selección previamente mencionados (**Tabla 1**).

Finalmente, estos 10 compuestos seleccionados fueron sometidos a un cálculo final de energía de afinidad con el objetivo de evaluar la estabilidad del complejo proteína-ligando y elegir a los candidatos más prometedores para un estudio posterior de dinámica molecular. Es importante destacar que estos compuestos exhiben valores significativos en los parámetros de energía de afinidad (MMGBSA) y Docking Score, por lo cual, respaldan la hipótesis de que los compuestos podrían ser candidatos prometedores como inhibidores debido a sus interacciones favorables con los residuos aminoacídicos clave en el sitio de unión de la proteína. Los resultados de este proceso se presentan de manera concisa y organizada en la **Tabla 1**.

Tabla 1. Puntuaciones de acoplamiento y energía de afinidad de los 10 compuestos mejores****calificados junto con sus propiedades ADME.**

Compuesto	Docking Score	MGBSA kcal/mol	MolW	Donador de H+	Aceptor de H+	LogP	Caco- 2	%absorción oral
Cumato	-8,243	-48,17	164,284	1	2	2,113	252	82
Ácido 3-(aminometil)-4,5-dihidroxibenzoico	-9,313	-48,38	183,163	5	5	-2,907	2	17
Ácido Gálico	-8,734	-27,56	164,284	1	2	2,113	252	82
Acido 4-bromo-3,5-dihidroxibenzoico	-8,545	-37,31	233,018	3	3,5	0,547	31	57
Ácido 3-bromo-2-metoxibenzoico	-8,493	-39,3	231,045	1	2,75	1,965	10	55
5-oxo-5,6,7,8-tetrahidronaftalen-2-carboxilato de metilo	-8,201	-34,4	204,255	0	4	1,475	935	89
Ácido 4-(2-oxoetil)benzoico	-8,145	-58,15	164,160	1	4	1,407	49	66
Ácido 4-amino-3-hidroxibenzoico	-7,637	-50,17	153,137	3,5	3,75	-0,135	27	53

Como se muestra en la (**Tabla 1**), el acoplamiento molecular sugiere como posibles inhibidores a ácido gálico (DS: -8,734; MGBSA: -27,56 kcal/mol), ácido 3-bromo-2-metoxibenzoico (DS: -8,493; MGBSA: -39,3 kcal/mol), y ácido 4-amino-3-hidroxibenzoico (DS: -7,637; MGBSA: -50,17 kcal/mol); asimismo, estos metabolitos no presentan violaciones a las reglas de Lipinski y a su vez muestran buenos valores de absorción.

El ácido gálico es un compuesto fenólico que se encuentra en diversas fuentes naturales como plantas, frutas y verduras. Además, se le atribuyen diversos efectos biológicos como actividad antiinflamatoria, antibiótica, protección cardiovascular, anticancerígena y antioxidante

Valores e intervalos de valores de las propiedades ADME basados en el manual de QikProp del software Schrodinger Inc.: Mol MW: 130,0-725,0; DonorHB: 0,0-6,0; AcptHB: 2,0-20,0; PlogPo/w: -2,0-6,5; QPPCaco (nm/sec): >25-25 -25%-; %absorción oral: Pobre menor a 25% y Alta mayor a 80%.

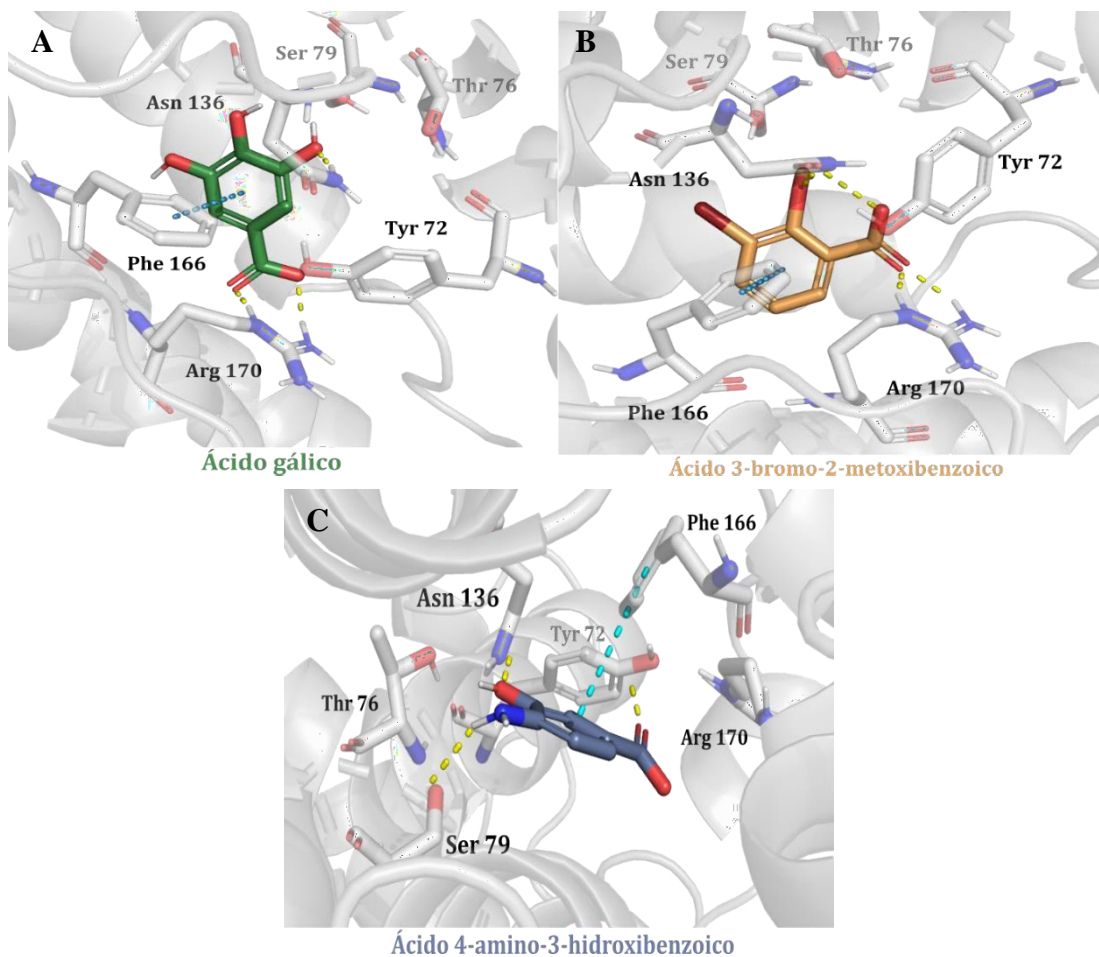
(Nobre et al., 2019; Salas et al., 2013). El ácido 3-bromo-2-metoxibenzoico es un ácido benzoico disustituido que se utiliza en la preparación de diversos compuestos biológicamente activos, como agonistas del adrenoreceptor α -2, los antagonistas del receptor Smoothened y los inhibidores de entrada del VIH-1 (Hattori et al., 1998).

El ácido 4-amino-3-hidroxibenzoico, también conocido como ácido para-aminosalicílico (PAS) es una molécula pequeña que inhibe la actividad del receptor del factor de crecimiento epidérmico (EGFR) y el factor de crecimiento epidérmico humano (HER2). También, se ha utilizado *In vitro* como agente antifúngico, especialmente contra *Candida albicans*, por lo cual, es un compuesto de constante estudio relacionado con microbiología y las infecciones fúngicas.

Además de su actividad antifúngica, se ha explorado el potencial del PAS como fármaco en el tratamiento para tratar el cáncer colorrectal metastásico (Meissner et al., 2006). Aunque su mecanismo de acción aún no se comprende por completo, se han sugerido hipótesis relacionadas con su capacidad para influir en los enlaces de hidrógeno intermoleculares y la permeabilidad del agua en células tumorales. Estos efectos podrían desempeñar un papel importante en la inhibición del crecimiento o la proliferación de las células cancerosas (Orii et al., 2004).

Como parte final del análisis, se examinaron las interacciones entre los tres mejores candidatos y su interacción con residuos aminoacídicos. El objetivo principal fue identificar patrones en las interacciones que pudieran proporcionar información relevante sobre la afinidad y eficacia de estos candidatos como inhibidores específicos de la proteína CymR (**Figura 6**). Además, se establecieron comparaciones con el inhibidor de referencia para evaluar la relevancia de las interacciones observadas.

Figura 6. Visualización del inhibidor cumato para la proteína CymR.



Nota Elaboración propia (2023). Cada conjunto de átomos de los residuos y del ligando poseen un código de color. Oxígenos: rojo; Nitrógenos: Azul. Los residuos aminoacídicos del bolsillo de*

unió están representados por color blanco. Los compuestos están nombrados de A-C: A. Ácido gálico, B. Ácido 3-bromo-2-metoxibenzoico, C. Ácido 4-amino-3hidroxibenzoico. Las interacciones proteína ligando tipo puente de hidrógeno están representados por líneas discontinuas amarillas, mientras que las interacciones tipo π - π por líneas discontinuas azules.

En primer lugar, se observa que tanto el ácido gálico y ácido 3-bromo-2-metoxibenzoico presentan interacciones tipo puente de hidrógeno con los aminoácidos Asn136 y Arg170, así como una interacción tipo π - π con el aminoácido Phe166 (**Figura 6A y 6B**). Es importante destacar que estos dos compuestos comparten interacciones clave con el inhibidor de referencia cumato, específicamente con los aminoácidos Phe166 y Arg170.

Por otra parte, el ácido 4-amino-3-hidroxibenzoico presenta un perfil de interacciones ligeramente que destaca por ciertas interacciones. En particular, se observan interacciones tipo puente de hidrógeno con los aminoácidos Tyr72, Ser79 y Asn136 junto con la interacción tipo π - π con el aminoácido Phe66 (**Figura 6C**).

Ahora bien, Aunque este compuesto no comparte las mismas interacciones tipo puente de hidrógeno con el inhibidor de referencia, es notable que todos los tres candidatos comparten una interacción específica con el aminoácido Asn136. Esto sugiere que Asn136 podría jugar un papel importante en la formación del complejo proteína-ligando para todos estos compuestos. Además, es relevante destacar que todos los candidatos comparten la interacción tipo π - π con el aminoácido Phe166, al igual que el compuesto de referencia cumato.

Finalmente, la evaluación de la afinidad en el sitio de unión identificado ha revelado resultados de considerable interés. En particular, hemos identificado la relevancia de ciertos aminoácidos como Tyr72 (hélice α 3), Ser79 (hélice α 3), Asn136 (hélice α 6), Phe166(hélice α 7) y Arg170(hélice α 7), cuyas interacciones desempeñan un papel fundamental en la estabilidad del

complejo proteína-ligando. Cabe destacar que la hélice $\alpha 3$ es esencial para el reconocimiento del ADN en la proteína CymR, mientras que, la hélice $\alpha 6$ que desempeña un papel crítico en la función biológica de transducción de señales dentro de la familia de represores Rrf2 (Volbeda et al., 2017). La interacción de los compuestos con estas hélices sugiere que podrían intervenir en el proceso de identificación y unión del ADN a la proteína, lo que podría resultar en una inhibición del sistema CymR-Vp16 (Gamez et al., 2021). Aunque no se ha reportado una función biológica específica para la hélice $\alpha 7$, se puede inferir esta hélice contribuye a la estabilidad y afinidad de los complejos en el sitio de unión identificado (Sethi et al., 2019). En este contexto, los resultados sugieren que el ácido gálico, ácido 3-bromo-2-metoxibenzoico y el ácido 4-amino-3-hidroxibenzoico podrían tener la capacidad de inhibir el sistema CymR-VP16.

Para corroborar y profundizar en la información obtenida, se realizó un estudio de dinámica molecular. Este análisis dinámico se hizo con el propósito de evaluar la estabilidad de las interacciones previamente identificadas mediante el enfoque de docking molecular. La dinámica molecular proporcionó una comprensión más completa de cómo estos candidatos interactúan con los residuos aminoacídicos a lo largo del tiempo, lo que es fundamental para validar la eficacia de los posibles inhibidores y su viabilidad como inhibidores de la proteína CymR.

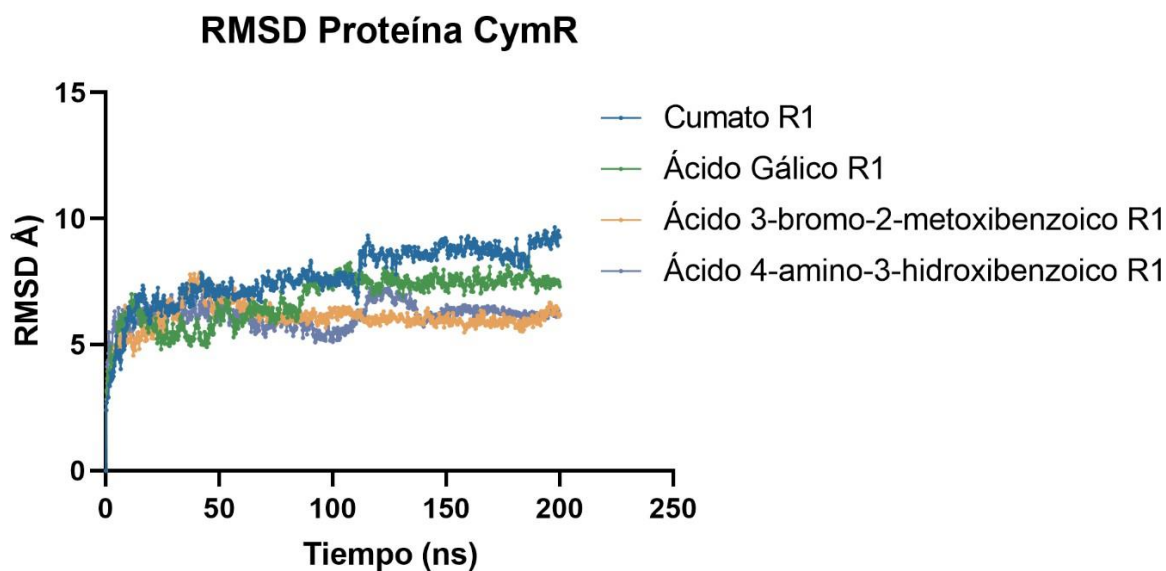
4.1.5. Dinámica Molecular CymR.

Para evaluar la estabilidad de los metabolitos seleccionados y propuestos por el docking molecular, se llevó a cabo el análisis de RMSD (Root Mean Square Deviation) durante un período de 200ns de simulación. El RMSD se emplea para medir las variaciones en la posición de los átomos de una proteína o ligando a lo largo del tiempo, por lo tanto, un valor bajo de RMSD se interpreta como una indicación de mayor estabilidad de la proteína o ligando (Kufareva & Abagyan, 2012).

En consecuencia, se analizó los resultados de dinámica molecular de la proteína CymR en conjunto con los compuestos: cumato, ácido gálico, ácido 3-bromo-2-metoxibenzoico y ácido 4-amino-3-hidroxibenzoico. Para simplificar el análisis de RMSD, se presentó el conjunto de datos obtenidos en la primera réplica. Las demás replicas estarán detalladas en la sección de anexos. No obstante, es fundamental destacar que, para el análisis de interacciones, se tendrán en cuenta todas las réplicas, puesto que, esto garantiza una evaluación más completa de las interacciones observadas.

Para empezar, se procedió a examinar el RMSD de la proteína en presencia de diversos ligandos (**Figura 7**). Este análisis permite comprender la estabilidad de la proteína, como del ligando en el bolsillo de unión. En la fase de equilibrio, el RSMD de la proteína en presencia de los compuesto cumato, ácido gálico, ácido 3-bromo-2-metoxibenzoico y el ácido 4-amino-3-hidroxibenzoico, se estabilizó aproximadamente en 6Å. Sin embargo, durante la fase de producción, se observaron variaciones en el RMSD de la proteína en función del ligando.

Figura 7. Valores de RMSD para Complejos CymR-Ligandos.

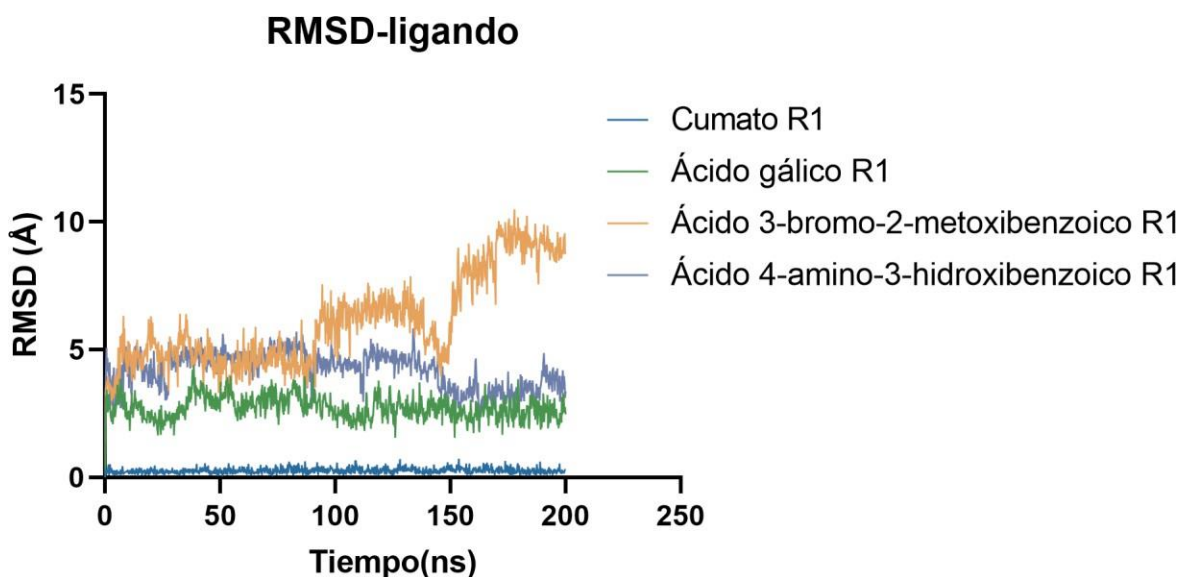


Nota* *Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para proteína CymR a lo largo del tiempo en relación con a ligando estudiad. Cada compuesto está representado por un color, Azul: Cumato, Verde: Ácido gálico, Naranja: Ácido 3-bromo-2-metoxibenzoico, Morado: Ácido 4-amino-3-hidroxibenzoico.*

Como se observa en la **Figura 7**, los valores de RMSD apenas experimentan ligeras variaciones para cada derivado de cumato, lo que sugiere una notable estabilidad y afinidad de la proteína CymR en relación con estos compuestos (Kufareva & Abagyan, 2012).

En este sentido, se procedió a examinar la estabilidad del ligando con respecto a la proteína y al sitio de unión identificado, mediante los datos obtenidos de RMSD (**Figura 8**). Se puede observar que los compuestos estudiados empiezan la simulación en el mismo valor de RMSD (5Å), hasta los 80ns., luego, se observaron variaciones en el RMSD para cada ligando.

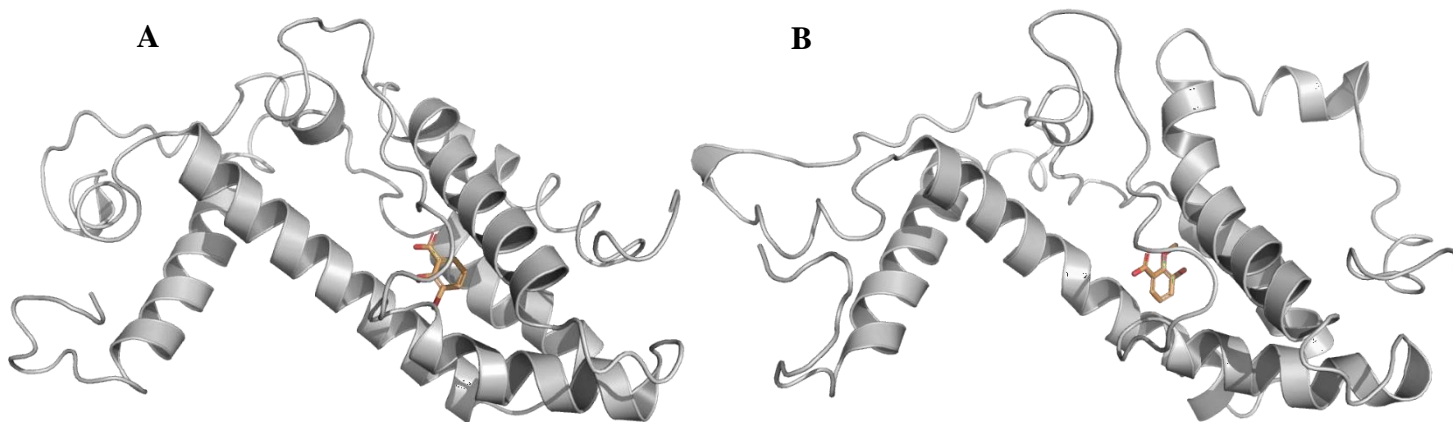
Figura 8. Valores de RMSD para los derivados de Cumato.



Nota Elaboración propia (2023). Imagen representativa de la variación del RMSD del ligando a lo largo del tiempo. Cada compuesto está representado por un color, Azul: Cumato, Verde: Ácido gálico, Naranja: Ácido 3-bromo-2-metoxibenzoico, Morado: Ácido 4-amino-3-hidroxibenzoico.*

Como se puede observar en la **Figura 8**, hasta los primeros 150ns, todos los compuestos estudiados muestran una notable estabilidad en sus respectivos valores de RMSD. Sin embargo, a partir de este punto, se destacan variaciones significativas en el RMSD del ácido 3-bromo-2-metoxibenzoico. Estas fluctuaciones sugieren un posible cambio de posición a la altura de los 150ns por parte del compuesto sin alejarse del sitio de unión y manteniendo su estabilidad durante el tiempo restante de la simulación (**Figura 9**). Estos resultados enfatizan la importancia de la dinámica a lo largo del tiempo en la comprensión de la estabilidad del complejo proteína-ligando.

Figura 9. Cambios posicionales del ácido 3-bromo-2-metoxibenzoico durante la réplica 1.



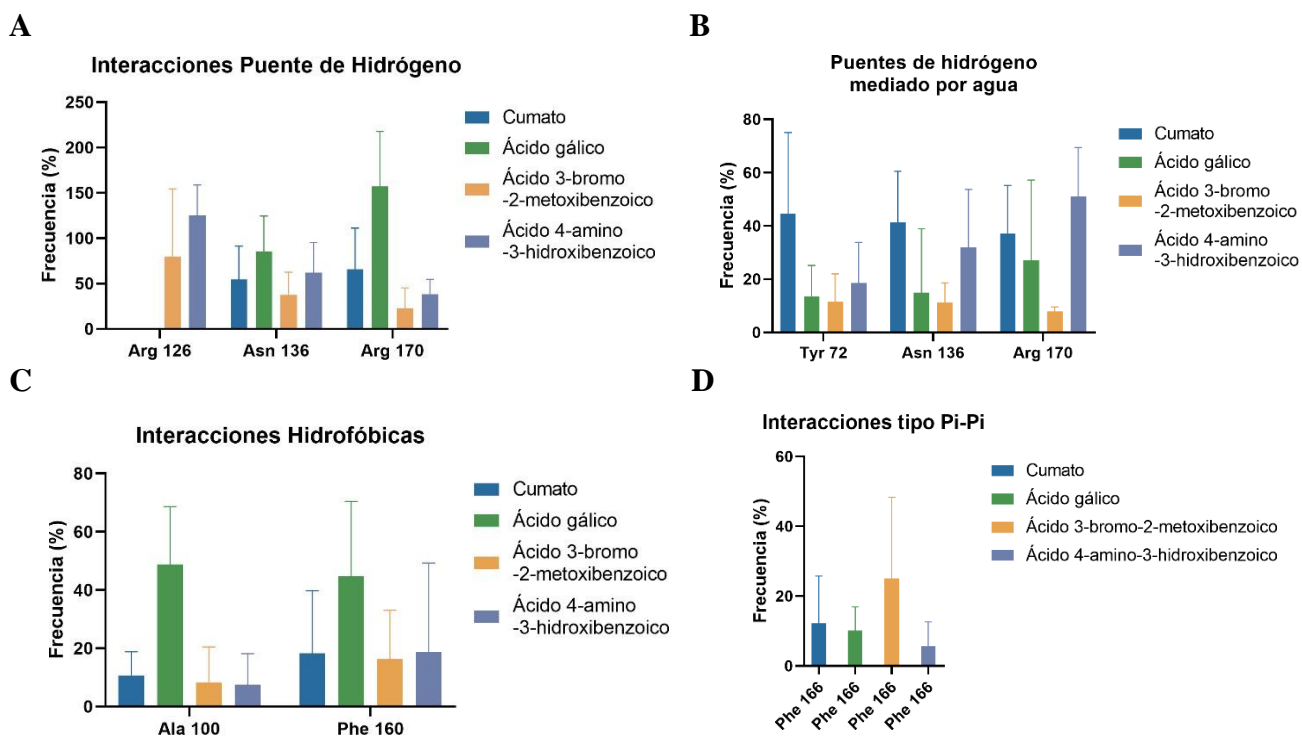
Nota Elaboración propia (2023). Ácido 3-bromo-2-metoxibenzoico coloreado de naranja. A) Posición inicial del ligando hasta los 150ns. B) Posición del ligando después de los 150ns.*

Por último, las interacciones no covalentes poseen un gran papel en la formación del complejo proteína-ligando. Estas interacciones pueden ser estudiadas a lo largo de la simulación de dinámica molecular mediante el reporte SID (Diagrama de Interacciones de la Simulación), con el objetivo de conocer en profundidad la estabilidad del complejo en el sitio de unión propuesto. Estas interacciones se clasifican en 4 tipos: Puentes de hidrógeno, puentes de hidrógeno mediado por agua, hidrofóbicas e iónicas (Kishk et al., 2020). Estas observaciones subrayan la importancia de evaluar las interacciones específicas entre los ligandos y la proteína a nivel molecular para comprender completamente la dinámica y la estabilidad de los complejos a lo largo del tiempo.

4.1.6. Análisis de interacciones proteína-ligando CymR.

Con el propósito de evaluar la estabilidad del complejo proteico formado, identificar integraciones clave en esta conformación y validar las interacciones previamente obtenidas en el proceso de docking molecular, se llevó a cabo un análisis detallado de los perfiles de interacciones entre cada compuesto y la proteína a lo largo de toda la simulación (**Figura 10**). Este análisis implicó el cálculo de frecuencias de las interacciones observadas durante el periodo de simulación.

Figura 10. Interacciones del Complejo CymR-Ligando a lo largo de la simulación.



Nota* *Elaboración propia (2023). A) Frecuencia de interacciones tipo puente de hidrógeno. B) Interacción de puentes de hidrógeno mediadas por agua. C) Interacciones hidrofóbicas. D) Interacciones tipo π - π . Cada compuesto está representado por un color, Azul: Cumato, Verde: Ácido gálico, Naranja: Ácido 3-bromo-2-metoxibenzoico, Morado: Ácido 4-amino-3-hidroxiibenzoico.*

Durante las simulaciones de dinámica molecular, se observó que los compuestos estudiados permanecieron en el bolsillo de unión propuesto para la proteína CymR, lo que sugiere una afinidad por este sitio de unión. En primer lugar, el ligando de referencia, cumato, mostró una frecuencia cercana a 50% en las interacciones tipo puente de hidrógeno con los aminoácidos Asn136 y Arg170 (**Figura 10A**). Además, al analizar los puentes mediados con agua que son redes de moléculas de agua que se unen a los aceptores o donores de puentes de hidrógeno cercanos en

las proteínas (Law & Daggett, 2010; Sabarinathan et al., 2011), se pudo identificar nuevamente la presencia de estos dos aminoácidos junto con la Tyr72 con una frecuencia cercana al 40% (**Figura 10B**).

Las hélices $\alpha 3$ y $\alpha 6$ son elementos cruciales en la proteína CymR debido a sus funciones biológicas específicas. La hélice $\alpha 3$ es esencial para el reconocimiento del ADN en la proteína CymR, mientras que, la hélice $\alpha 6$ que desempeña un papel crítico en la transducción de señales del grupo Fe-S (Volbeda et al., 2017) que desempeña una variedad de funciones de detección incluidas el control de los niveles celulares de hierro, el estrés nitrosativo y estrés oxidativo dentro de la familia de represores Rrf2 (Isabella et al., 2009; Volbeda et al., 2020). En este sentido el aminoácido Tyr72 que pertenece a la hélice $\alpha 3$, es esencial para el reconocimiento del ADN en la proteína CymR, por tanto, la interacción de los compuestos en estudio con esta hélice podría influir en el proceso de reconocimiento y afectar la capacidad de la proteína para unirse al ADN (Volbeda et al., 2017). De manera similar, la hélice $\alpha 6$ que alberga el aminoácido Asn136, desempeña un papel fundamental en la transducción de señales, por lo cual, cabe la posibilidad que los compuestos que establecen interacciones estables con esta región puedan inducir la supresión génica de la proteína CymR (Volbeda et al., 2017). Además, la Arg170 que pertenece a la hélice $\alpha 7$, podría contribuir a través de interacciones hidrofílicas a la estabilidad del ligando dentro del sitio de unión propuesto (Chung et al., 2009).

Adicionalmente, en término de interacciones hidrofóbicas se observa que los aminoácidos Ala100 y Phe160, con una frecuencia cercana al 20% (**Figura 10C**), contribuyen a la estabilidad del complejo. Lo anterior, es coherente, puesto que, se encuentran en las proximidades del sitio de unión del ADN (hélice $\alpha 6$), que es conocido por ser altamente hidrofóbico (**Figura 4**) (Volbeda et al., 2017). Por último, el aminoácido Phe168 que establece interacciones de tipo de tipo π - π con

el anillo benzoico del cumato con una frecuencia cercana a 20% (**Figura 10D**), también contribuyen con la estabilidad del complejo.

Al analizar los derivados del cumato, se puede observar en la **Figura 10**, que presentan un perfil de interacciones similar al cumato. Por ejemplo, el ácido gálico mostró una alta frecuencia de interacciones tipo puente de hidrógeno con los aminoácidos Asn136 y Arg170 (90% y 150% respectivamente). Aunque su frecuencia disminuye (20% y 30%) (**Figura 10B**) en presencia de agua como solvente, es importante destacar que el ácido gálico demuestra la capacidad de establecer interacciones estables con Asn136, un aminoácido que forma parte de la hélice $\alpha 6$, encargada de la transducción de señales (Volbeda et al., 2017). Además, el ácido gálico exhibe una afinidad con los aminoácidos Ala100 (50%) y Phe160(50%) mediante interacciones hidrofóbicas (**Figura 10C**). También, se observa una contribución a la estabilidad del complejo por parte de Phe168 gracia a su interacción tipo π - π (**Figura 10D**).

Por otro lado, el ácido 3-bromo-2-metoxibenzoico presentó frecuencias inferiores al 50% en las interacciones tipo puente de hidrógeno con los aminoácidos Asn136 y Arg170 (**Figura 10A**). Sin embargó, se observó que interactuaba con el aminoácido Arg126 (hélice $\alpha 6$) con una frecuencia cercana al 70% durante toda la simulación (**Figura 10A**). Estos resultados sugieren una mayor estabilidad con la proteína debido a su doble interacción con hélice $\alpha 6$, quien se encarga de la transducción de señales (Volbeda et al., 2017), de igual manera, las interacciones hidrofóbicas con Ala100 y Phe160 (**Figura 10C**) permiten un mejor acoplamiento del compuesto en el sitio de unión propuesto, al igual que su interacción tipo π - π con Phe168 (**Figura 10D**).

Por último, el ácido 4-amino-3-hidroxibenzoico destacó por su comportamiento similar al cumato. Sus interacciones tipo puente de hidrógeno con los aminoácidos Asn136 y Arg170 muestran frecuencias cercanas al 50% (**Figura 10A**). Cuando se incluye agua como intermediario,

la frecuencia de interacción con el aminoácido Arg170 aumentó al 60% y se observa una frecuencia de interacción cercana al 20% con Tyr72 (**Figura 10B**) presente en la hélice $\alpha 3$, encargada del reconocimiento del ADN. A pesar de que las interacciones hidrofóbicas con los aminoácidos Ala100 y Phe168 presentan una frecuencia menor al 20%, también contribuyen a la estabilidad del complejo (**Figura 10C**), al igual que su interacción con el aminoácido Phe168 mediante una interacción de tipo π - π . Estos análisis de acoplamiento y dinámica molecular sugieren que el ácido gálico, el ácido 3-bromo-2-metoxibenzoico y ácido 4-amino-3-hidroxibenzoico y el ácido 4-amino-3-hidroxibenzoico podría promover una estabilidad óptima en la proteína CymR.

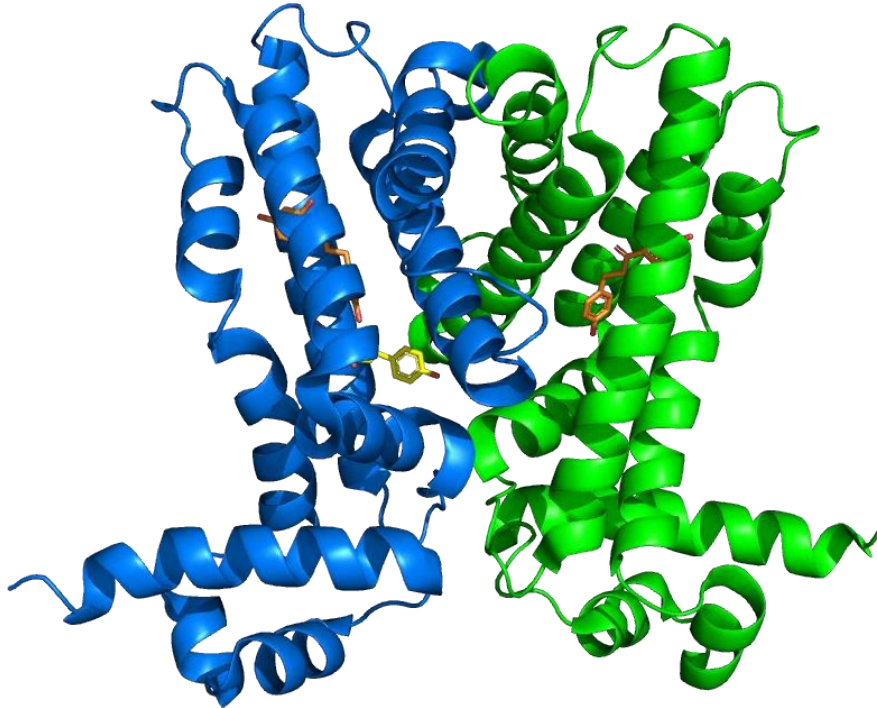
4.2. Proteína TtgR

4.2.1. Identificación de Target para TtgR.

Para la identificación de compuestos que puedan inhibir el sistema TtgR-Vp16, primero se hizo una caracterización de la proteína represora TtgR. En este sentido, se decidió no llevar a cabo un modelo de homología, puesto que se usó una estructura cristalina reportada bajo el código PDB 2UXI (**Figura 11**). Esta estructura cristalina presentó un porcentaje de identidad del 98% y un porcentaje de similitud del 99% con la secuencia del plásmido 907L. Para validar esta similitud se

empleó la herramienta de búsqueda de Homología Blast Homology Search y alineamiento con la herramienta ClustalW del software maestro Schrödinger.

Figura 11. Estructura cristalina de la proteína TtgR con código PDB: 2UXI.



Nota Elaboración propia (2023). Estructura cristalina bajo el código 2UXI a una resolución de 2.50 Å de la proteína TtgR. La estructura cristalina exhibe un dímero, con cadena proteica codificada por un sistema de colores diferentes. El color verde representa la cadena A y el color azul representa la cadena B. Asimismo, la estructura contiene moléculas de floretina representadas por el color naranja en el sitio de baja afinidad y color amarillo en el sitio de alta afinidad.*

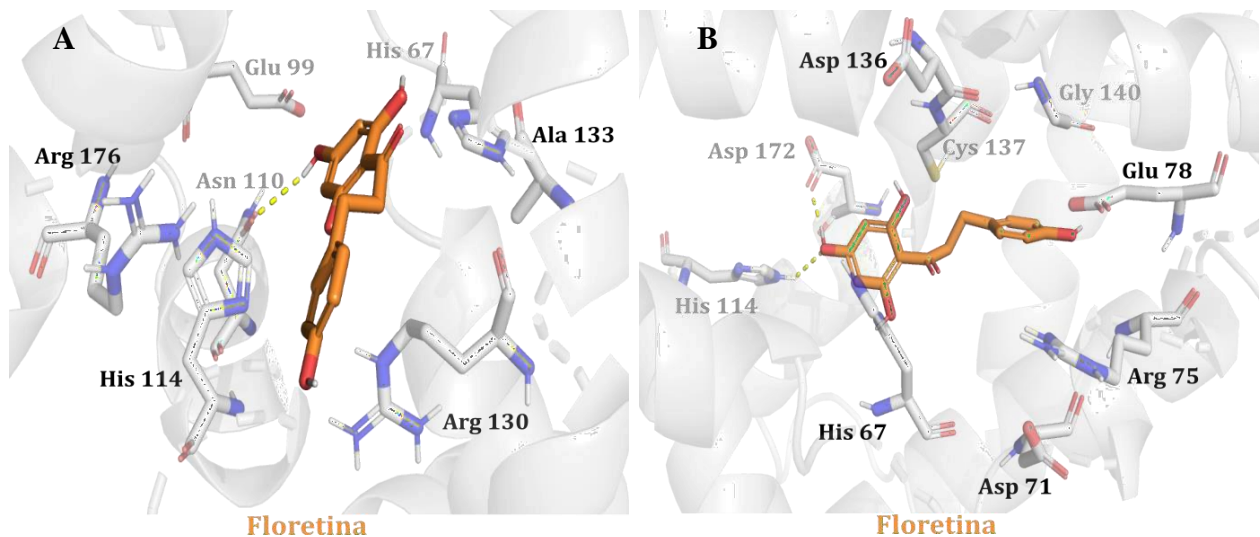
Como se muestra en la **Figura 11**, la estructura cristalina de TtgR consiste en un dímero, donde el bolsillo de unión puede acomodar tres ligandos de referencia. Es esencial resaltar que el monómero B (azul), que comprende la molécula de floretina causante de la inhibición (amarilla), es el monómero que posee el sitio de unión donde se genera la inhibición (Alguel et al., 2007a),

por lo cual, fue el foco principal de nuestro estudio y el único con el que trabajamos. Este monómero está compuesto por consta de 9 alfa hélices y se divide en dos dominios: El dominio de unión al ADN compuesto por $\alpha 1- \alpha 3$ y el segundo dominio de unión al compuesto por $\alpha 4- \alpha 9$ (Alguel et al., 2007b). Una vez identificado el monómero relevante y comprendida su estructura, el siguiente paso consistió en llevar a cabo estudio de docking molecular. A continuación, se examinaron los resultados obtenidos del docking molecular.

4.2.2. Docking molecular para TtgR.

Para llevar a cabo el análisis del docking molecular en el bolsillo de la proteína TtgR, es importante entender las características principales de los dos posibles (**Figura 12**) sitios de unión reportados en la literatura (Alguel et al., 2007b).

Figura 12. Visualización del inhibidor floretina para la proteína TtgR.



Nota* *Elaboración propia (2023). Visualización de la floretina en el bolsillo de unión de la proteína TtgR, A: Sitio de unión de alta afinidad, B: Sitio de unión de baja afinidad. Cada conjunto de átomos de los residuos y del ligando poseen un código de color. Oxígenos: rojo; Nitrógenos: Azul. La floretina está representada por color naranja*

El primer sitio, denominado sitio de alta afinidad (**Figura 12A**), se caracteriza por la presencia de aminoácidos cargados como His114, Arg130 y Arg170, los cuales pueden interactuar con compuestos de interés, contribuyendo a la estabilidad del complejo. Además, el aminoácido Asn110 destaca especialmente en el sitio de alta afinidad, puesto que, ocupa una posición central en el sitio activo y, en consecuencia, permite la unión del ligando a la proteína, siendo el principal contribuyente a la estabilidad del complejo. En término de interacciones hidrofóbicas, el sitio de unión incluye a los aminoácidos Leu66, Leu113 cal 134 y Phe168 (Alguel et al., 2007a).

El segundo sitio de unión, conocido como el sitio de baja afinidad (**Figura 12B**), también presenta aminoácidos cargados como His114 y Asp172 que contribuyen a la estabilidad del complejo. En este caso, el aminoácido Glu78 juega un papel crucial como precursor para la formación del complejo proteína-ligando en el sitio de baja afinidad (Alguel et al., 2007b).

En ambos sitios de unión, la floretina, se puede acoplar de manera óptima. Sin embargo, es importante mencionar que solo en el sitio de alta afinidad se observa la inhibición de la proteína (Alguel et al., 2007a). Por esta razón, se realizó un docking molecular en el sitio de alta afinidad de la proteína TtgR utilizando únicamente el monómero A, puesto que, las interacciones entre ligando y proteína se producen principalmente en este sitio.

4.2.3. Docking molecular para TtgR en el sitio de Alta afinidad.

Con la estructura cristalina de la proteína TtgR y los sitios de unión identificados, el siguiente paso consiste en llevar a cabo un docking molecular en el sitio de unión de alta afinidad centrando la caja en las coordenadas del ligando en la posición horizontal, el cual es causante de la inhibición (Alguel et al., 2007b). Para realizar este proceso, se realizó una lista de compuestos que constaba de 467 derivados de la floretina (Apéndice 2). Sin embargo, se encontró que la mayoría de los compuestos tenían un alto costo comercial o no estaban disponibles en el mercado.

Dado que nuestro objetivo es desarrollar un sistema de expresión génica regulado con un valor financiero asequible, se excluyeron aquellos compuestos cuyo precio superaba los \$100,00 USD por gramo, seleccionando únicamente 18 compuestos que eran económicamente viables. Estos 18 compuestos, fueron sometidos a un estudio de docking molecular con precisión extra (XP) generando 15 poses por ligando. Para ello, se Los resultados obtenidos en el docking molecular indican que los compuestos con potencial para inhibir la proteína son: quercetina, cloranfenicol, hesperetina, naringenina y tetraciclina (**Tabla 2**).

Tabla 2. Puntuaciones de acoplamiento y energía libre de unión de los compuestos mejores calificados en el sitio de unión de alta afinidad de proteína TtgR junto con sus propiedades ADME.

Compuesto	Docking Score	MGBSA Kcal/mol	MolW	Donador de H ⁺	Aceptor de H ⁺	LogP	Caco-2	%absorción oral
Floretina	-2,847	-17,67	274.273	2	3	2.054	113	76
Quercetina	-8,194	-45,95	302.240	4	5.250	0.363	18	52
Cloranfenicol	-8,493	--54,18	323.132	3	6.9	1.104	84	68
Hesperetina	-6.500	-45,93	302.283	2	4.750	1.770	127	75
Naringenina	-6,701	-31,94	272.257	2	4	1.631	128	74
Tetraciclina	-7.549	-49,29	444.440	3	8.250	- 0.817	4	20

Como se observa en la **Tabla 2**, la quercetina exhibe una afinidad y estabilidad en el sitio de alta afinidad de la proteína TtgR, respaldadas por los valores destacados de docking score: -8,1940 y MMGBSA: -45,95 kcal/mol. Además, durante la inspección visual de la molécula, se identificaron varias interacciones tipo puente de hidrógeno con el aminoácido Asn110 (**Figura 13A**), el cual es conocido por su importancia en la inhibición de la proteína y su fuerte unión con la floretina. Estas observaciones sugieren que la quercetina es un compuesto prometedor que posee una estabilidad óptima y se acopla eficazmente con el sistema TtgR-Vp16.

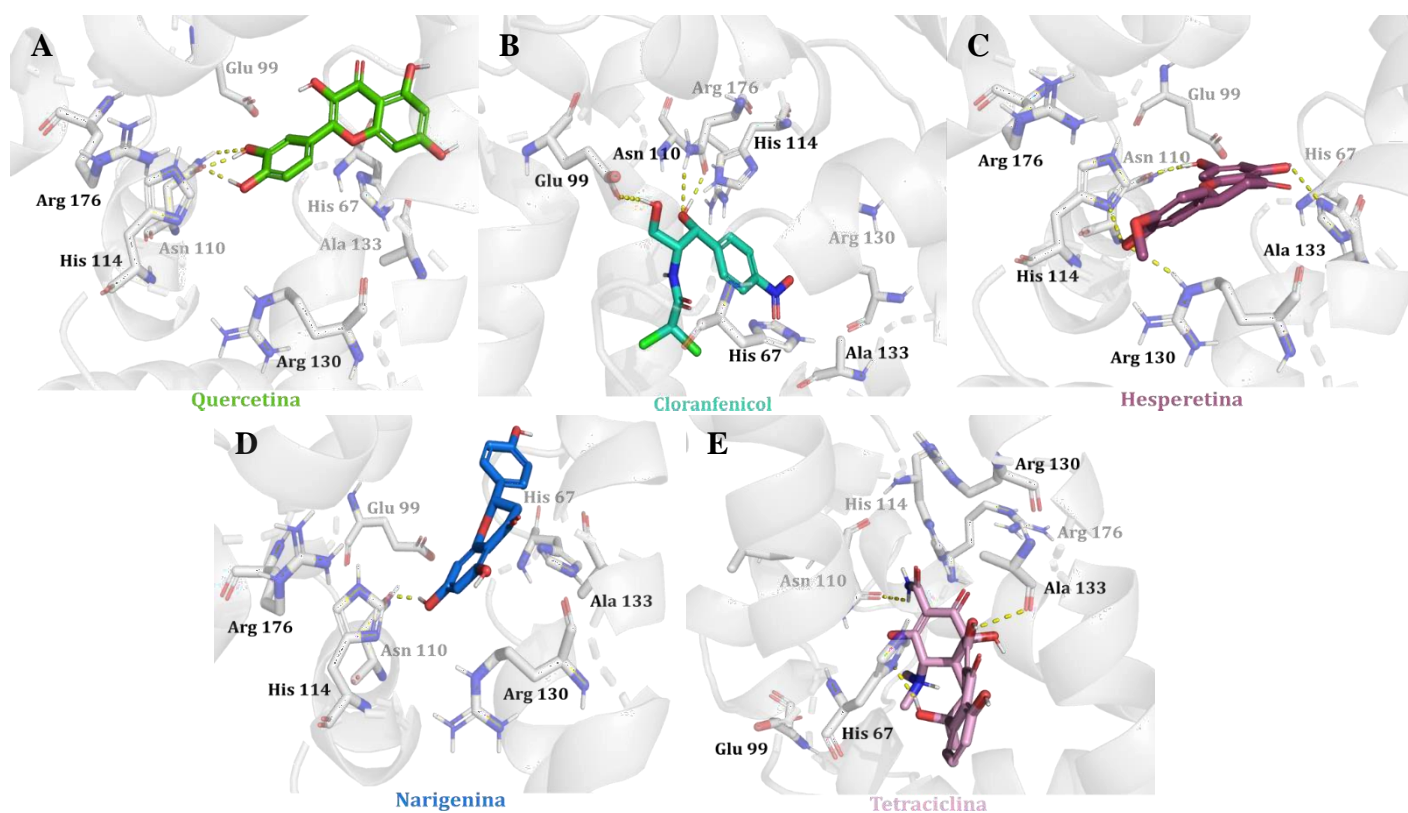
Por otro lado, compuestos como cloranfenicol y hesperetina, aunque no obtienen los puntajes más altos en Docking Score, demuestran valores interesantes en MGBSA: -54,18 kcal/mol y -45,93 kcal/mol respectivamente. Estos resultados sugieren la posibilidad de una interacción favorable con la proteína. Además, al analizar detenidamente, se observó que el cloranfenicol presenta interacciones tipo puente de hidrógeno con los aminoácidos Glu99, Asn110 y Arg176 (**Figura 13B**). De igual manera, la hesperetina presenta interacciones tipo puente de hidrógeno con los aminoácidos His67, Asn110, His114 y Arg130 (**Figura 13C**). En este sentido, se destacan principalmente los aminoácidos His114, puesto que, tiene la función de reconocimiento del ligando durante la formación del complejo, y el aminoácido Asn110, que desempeña un papel crucial en la conformación del complejo proteína-ligando (Alguel et al.,

2007b). Estos hallazgos podrían explicar los valores de afinidad obtenidos, respaldando la estabilidad del complejo.

Nota Elaboración propia (2023). Cada conjunto de átomos de los residuos y del ligando poseen un código de color. Oxígenos: rojo; Nitrógenos: Azul. Los compuestos están nombrados de A-E:*

A. Quercetina, B. Cloranfenicol, C. Hesperetina, D. Naringenina, E. Tetraciclina. Las

Figura 13. Interacciones Predichas en el sitio de alta afinidad por el Docking Molecular para TtgR.



interacciones proteína ligando tipo puente de hidrógeno están representados por líneas discontinuas amarillas y las interacciones tipo π - π por líneas discontinuas azules.

Por último, los compuestos naringenina y tetraciclina son los que presentan menor afinidad en el acoplamiento. La naringenina presentó valores de Docking Score: -6,701 y MMGBSA: -31,94 kcal/mol, y la tetraciclina presentó valores de Docking Score: -7,5940 y MMGBSA: -49,29 kcal/mol, **Tabla 2**. Como se observa en la **Figura 13D** y **Figura 13E**, la naringenina y la

tetraciclina tienden a situarse fuera del bolsillo de unión en comparación con los otros compuestos, por tal motivo, disminuyen su afinidad por este sitio de unión. Sin embargo, al analizar en detalle, se observó que ambos compuestos presentaban una interacción tipo puente de hidrógeno con el aminoácido Asn110, lo que sugiere que pueden acoplarse con la proteína TtgR, aunque su estabilidad puede verse afectada.

Finalmente, la evaluación de la afinidad presentada por los compuestos en este sitio de unión ha revelado resultados de considerable interés. En particular, hemos identificado la relevancia de ciertos aminoácidos, como His114 y Asn110, cuyas interacciones desempeñan un papel fundamental en la estabilidad del complejo-proteína ligando. Estos resultados coinciden con informes previos de la literatura científica (Alguel et al., 2007a). La quercetina destaca como el compuesto con mayor afinidad en el sitio de unión de alta afinidad, sin embargo, compuestos como hesperetina y cloranfenicol también tiene una afinidad notoria por este sitio de unión. Por otra parte, compuestos como la naringenina y tetraciclina, presentan limitaciones para su acoplamiento molecular, por lo que podría dificultar su estabilidad en el bolsillo de unión. En este contexto, los resultados sugieren que la quercetina, el cloranfenicol y la hesperetina podrían tener el potencial de inhibir el sistema TtgR-Vp16. Por lo tanto, se realizarán estudios de dinámica molecular en el sitio de alta afinidad, donde se reporta la inhibición para esta proteína. Estos estudios nos permitirán evaluarla estabilidad de los compuestos en el sitio de unión y determinar si podrían inhibir el sistema de interés TtgR-Vp16.

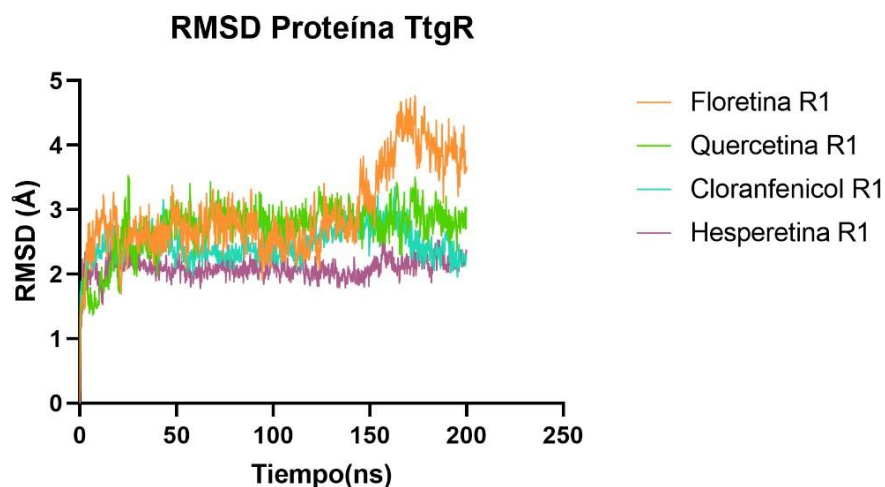
4.2.4. Resultados Dinámica Molecular para TtgR.

Para iniciar la presentación de los resultados obtenidos de los estudios de dinámica, es importante destacar que en este estudio se ha evaluado los derivados: floretina, quercetina, cloranfenicol y hesperetina, debido a su acoplamiento en el bolsillo de unión de la proteína TtgR.

Para una comprensión más clara y accesible de los resultados, los análisis de RMSD se centran exclusivamente en una sola réplica. Sin embargo, en el caso de los análisis de interacciones se han considerado las tres réplicas correspondientes. Esto permite una evaluación más completa y robusta de los derivados con el bolsillo de unión de alta afinidad de la proteína TtgR.

Para empezar, se procesado a examinar el RMSD de la proteína TtgR en presencia de los derivados mencionados anteriormente (**Figura 14**). Este análisis permite comprender de una mejor manera la estabilidad de los compuestos a lo largo de la simulación. De esta manera, se observó que, al inicio, la proteína en presencia de floretina exhibe un RMSD cercano a 3Å. Sin embargo, a medida que se avanza en la simulación, se observa un cambio de RMSD en los 150ns, lo que sugiere, un pequeño giro en una de las hélices de la proteína (**Figura 16**), por lo cual, la estabilidad del complejo TtgR-floretina podría estar siendo afectada.

Figura 14. Valores de RMSD para TtgR

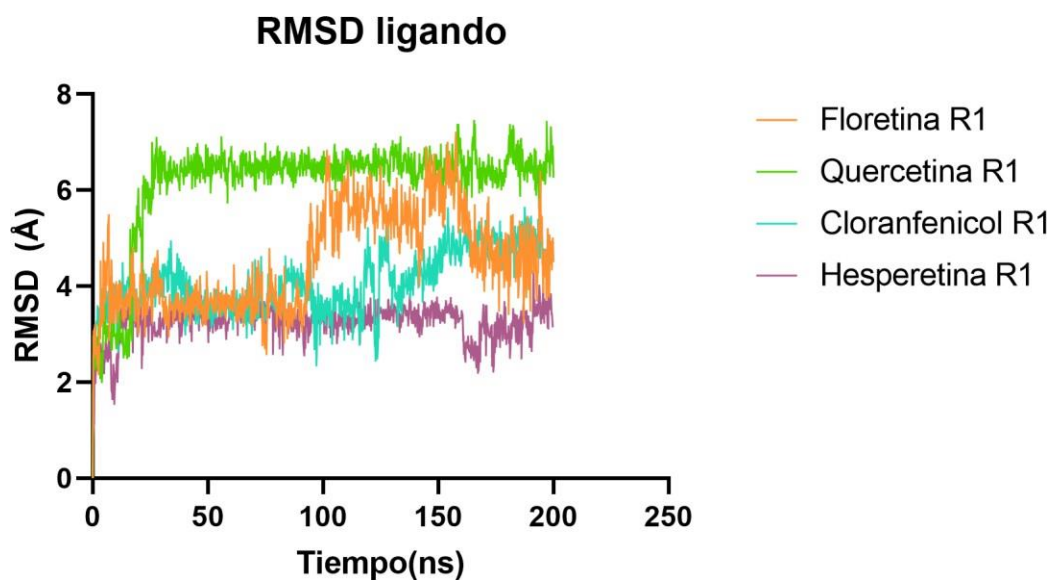


*Nota** Elaboración propia (2023). A) Imagen representativa de los valores de desviación cuadrática media RMSD para proteína TtgR en presencia de diferentes compuestos a lo largo del

tiempo. Cada compuesto está representado por un color, Naranja: Floretina, Verde: Quercetina, Ciano: Cloranfenicol, Morado: Hesperetina.

Al contrastar esta información con el análisis del RMSD de la floretina, se identificó una fluctuación cerca a los 100ns. Como se ilustra en la **Figura 15**, el RMSD de la floretina comienza en 4 Å y se mantiene estable hasta 100ns. Luego, experimenta una variación hasta alcanzar los 6Å. Finalmente, a los 150ns, se presenta una disminución del RMSD, regresando a los 4Å.

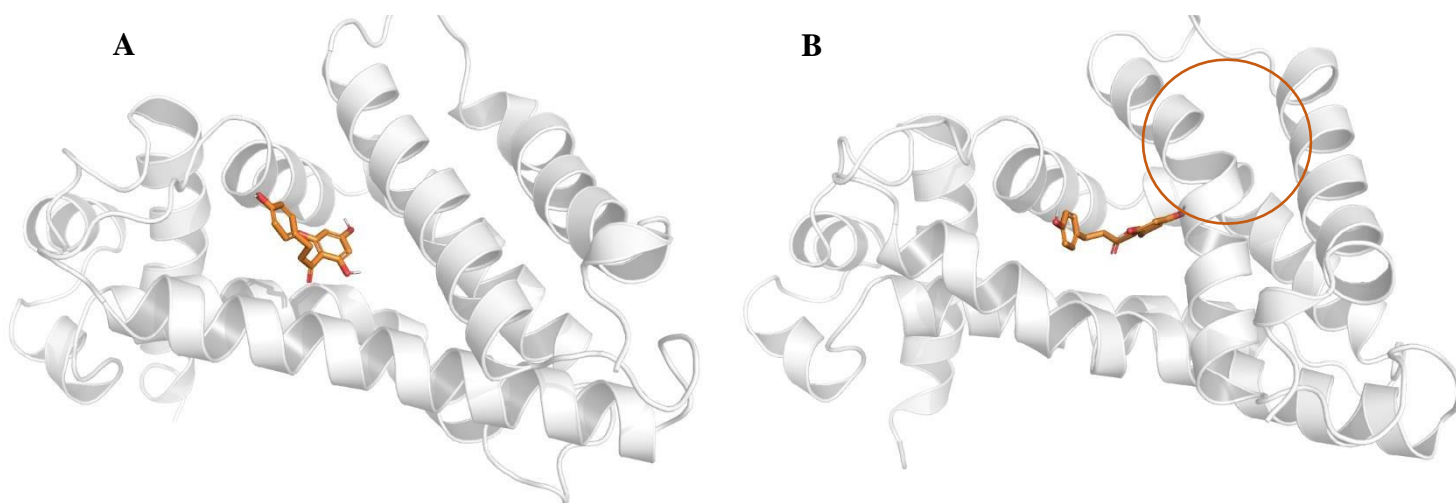
Figura 15. Valores de RMSD para la floretina y compuesto en estudio.



*Nota** Elaboración propia (2023). Variación del RMSD del ligando acoplado con la proteína a lo largo del tiempo de simulación con los compuestos. Cada compuesto está representado por un color, Naranja: Floretina, Verde: Quercetina, Ciano: Cloranfenicol, Morado: Hesperetina.

Estas fluctuaciones sugieren un posible cambio de posición a la altura de los 100ns por parte del compuesto sin alejarse del sitio de unión y manteniendo su estabilidad durante el tiempo restante de la simulación (**Figura 16**). Estos resultados enfatizan la importancia de la dinámica a lo largo del tiempo en la comprensión de la estabilidad del complejo proteína-ligando

Figura 16. Cambios posicionales de la floretina durante la réplica 1.



Nota Elaboración propia (2023). Floretina coloreado de naranja. A) Posición inicial del ligando hasta los 150ns. B) Posición del ligando después de los 150ns. El círculo naranja representa el ligero giro presentado por las alfas hélice de la proteína TtgR.*

Por último, al examinar el RMSD de la proteína en presencia de los otros ligandos, no se observan fluctuaciones destacadas, por el contrario, se mantiene estable en el intervalo 2-3Å. Este resultado sugiere que la proteína presenta una buena afinidad por estos derivados. Sin embargo, al analizar el RMSD de los compuestos, se evidencia que únicamente la quercetina presenta una estabilidad constante a lo largo de la simulación. Aunque, el RMSD del cloranfenicol y la hesperetina no experimenta fluctuaciones significativas, se puede observar ligeros cambios en el RMSD que podrían sugerir, que, en momentos puntuales de la simulación, los compuestos

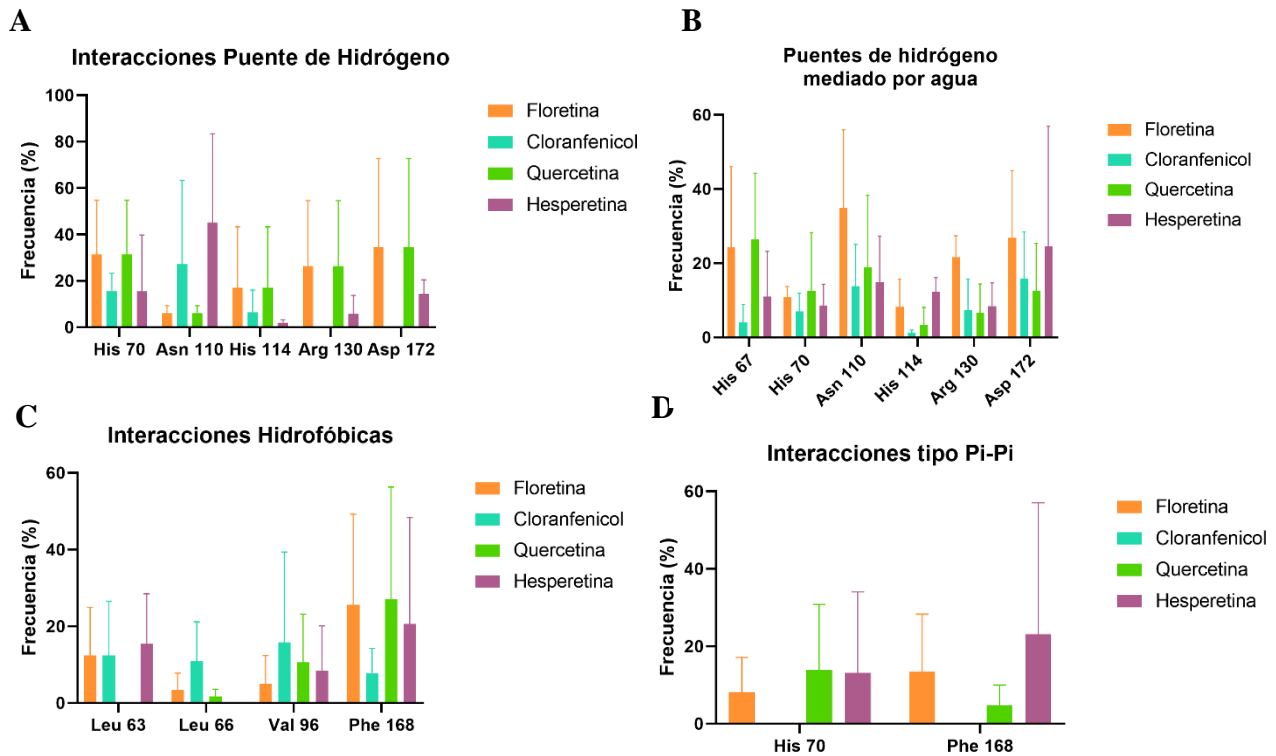
experimentan movimientos mínimos en el sitio de unión, lo que potencialmente afectaría la estabilidad del complejo. No obstante, para corroborar esta estabilidad se procede a realizar un análisis de las interacciones con los aminoácidos importantes del sitio de unión de la proteína TtgR a lo largo de la simulación de dinámica molecular.

4.2.5. Análisis de interacciones proteína-ligando TtgR.

La aplicación del análisis de Simulación de Interacciones Dinámicas (SID) en el estudio de TtgR representa una herramienta fundamental para evaluar la estabilidad de las interacciones de los compuestos con los aminoácidos a lo largo de la simulación. Este enfoque nos permite explorar en detalle cómo los compuestos interactúan dinámicamente con la proteína a lo largo del tiempo, proporcionando información valiosa sobre la naturaleza y la fuerza de estas interacciones. El principal objetivo de este análisis es identificar cualquier indicio de posible inhibición de TtgR por parte de los compuestos estudiados (**Figura 16**). Al examinar minuciosamente la evolución de las interacciones, se espera obtener información relevante cruciales que puedan contribuir a un mejor

entendimiento de la relación entre estos compuestos y la proteína TtgR, con implicaciones potenciales en la búsqueda de estrategias de inhibición.

Figura 17. Interacciones del Complejo TtgR-Ligando a lo largo de la simulación.



Nota Elaboración propia (2023). A) Frecuencia de interacciones tipo puente de hidrógeno. B) Interacción de puentes de hidrógeno mediadas por agua. C) Interacciones hidrofóbicas. D) Interacciones tipo π - π . Cada compuesto está representado por un color, Naranja: Floretina, Verde: Quercetina, Ciano: Cloranfenicol, Morado: Hesperetina.*

Durante las simulaciones de dinámica molecular, se observó que los compuestos permanecen en el bolsillo de unión de alta afinidad, lo que se alinea con los resultados previamente reportados y respalda la noción de que estos compuestos presentan una afinidad por el sitio de unión (Alguel et al., 2007b). Esta correspondencia ofrece un mayor argumento a la idea de que

dicha afinidad puede tener implicaciones relevantes en la posible inhibición del sistema TtgR-VP16.

Ahora bien, el ligando de referencia, floretina, exhibió frecuencias cercanas a 5%, 10% y 30% para las interacciones tipo puente de hidrógeno con los aminoácidos Asn110, His114 y Asp172 (**Figura 16A**), respectivamente. Estos tres aminoácidos son esenciales para la formación del complejo proteína ligando: la His114 desempeña un papel de reconocimiento del ligando antes de su posterior acople, el aminoácido Asn110 se encarga del anclaje del ligando con la proteína, mientras que Asp172 contribuye a la estabilidad mediante sus interacciones con los ligandos (Alguel et al., 2007b). Sin embargo, en presencia de agua, Asn110 y Asp172 aumenta su frecuencia de interacción hasta aproximadamente 40% (**Figura 16B**). De este modo, los puentes de hidrógeno mediados por agua pueden contribuir a la estabilidad del complejo proteína-ligando, aunque de forma limitada, lo que conlleva a una disminución de la flexibilidad por parte del aminoácido en dicha interacción, además, que sin la presencia de agua esta interacción podría no ser factible (Aquino et al., 2011).

En cuando a las interacciones hidrofóbicas, se puede observar que Leu63 y Phe168 (**Figura 16C**) son los aminoácidos que más contribuyen a la estabilidad del ligando con una frecuencia cercana al 20%, siendo Phe168 notable por su interacción de tipo π - π con la floretina (**Figura 16D**), por lo cual, existe una correspondencia con los patrones documentados en la literatura (Alguel et al., 2007b). Estos resultados sugieren que la floretina puede acoplarse a la proteína TtgR y muestra una afinidad fuerte por su sitio de unión. Sin embargo, la baja frecuencia de interacción de la floretina con Asn110 sugiere que este compuesto podría requerir la presencia de solvente para su acoplamiento, lo que podría limitar su capacidad de interactuar con la proteína. Lo anterior, supone que la floretina podría no ejercer una inhibición sobre el sistema TtgR-VP16.

Por otra parte, al analizar los compuestos se pudo encontrar que presentaban frecuencias similares a la floretina. En particular, la quercetina destacó como el compuesto más prometedor en términos de una posible inhibición del sistema TtgR-Vp16. A pesar de que se calcularon frecuencias por debajo del 20% en interacciones tipo puente de hidrógeno con el aminoácido Asn110 (**Figura 16A**), la quercetina exhibió una frecuencia de 40% con el aminoácido Asp172 (**Figura 16A**). Además, es relevante mencionar que se destacaron dos interacciones tipo puente de hidrógeno con los aminoácidos His70 y Arg130, con frecuencias cercanas al 30%, así como con His114 con una frecuencia del 20% (**Figura 16A**). Como se señalado previamente, las interacciones con los aminoácidos His114 y Asp172 desempeñan un papel fundamental en la estabilidad del complejo proteína-ligando (Alguel et al., 2007a). Las interacciones hidrofóbicas de la quercetina también presentaron la frecuencia más alta entre los compuestos, con una frecuencia cercana a 30% con el aminoácido Phe168. Estos patrones de interacción, en línea con los observado con la floretina, respaldan la noción de que la quercetina podría poseer potencial necesario para inhibir la proteína TtgR. Sin embargo, es importante enfatiza que estos resultados sugieren la buena estabilidad de la quercetina en el sistema, pero aún se requiere de una investigación más detallada para confirmar su capacidad de inhibición.

Para el caso del cloranfenicol, se observó una frecuencia cercana al 30% de interacciones tipo puente de hidrógeno con Aan110 (**Figura 16A**). No obstante, en presencia de agua, las interacciones con este aminoácido se disminuyen a una frecuencia de 18% (**Figura 16B**). Esto sugiere que el cloranfenicol tiende a interactuar con Asn110 directamente, sin intermediarios, lo que indica una buena estabilidad en el sitio de unión. Sin embargo, al observar las interacciones hidrofóbicas (**Figura 16C**) del cloranfenicol, se podría notar una afinidad limitada, puesto que, presentó la frecuencia más baja, por debajo del 10%. Esta limitación en las interacciones

hidrofóbicas podría restringir la capacidad de acoplamiento del cloranfenicol en el sitio de unión de la proteína TtgR. Además, es relevante destacar que el cloranfenicol no establece ninguna interacción de tipo π - π con ningún aminoácido (**Figura 16D**).

Estos resultados sugieren que, si bien el cloranfenicol presenta una buena probabilidad de interactuar con Asn110, el cual desempeña un papel fundamental en la formación del complejo (Alguel et al., 2007a), su afinidad en el bolsillo de unión puede verse afectada debido a su limitada interacción hidrofóbica, lo que podría hacerlo menos estable en comparación con la quercetina.

Finalmente, la hesperetina presentó la frecuencia más alta (40%) de interacciones tipo puente de hidrógeno con el aminoácido Asn110, sin embargo, con la Asp172 presentó una frecuencia por debajo del 20% (**Figura 16A**). En presencia del agua, se puede observar que la interacción con el aminoácido ASN110 disminuye por debajo del 20% (**Figura 16B**) y la interacción con ASP172 aumenta por encima de 20% (**Figura 16B**), lo que quiere decir, que la hesperetina es capaz de interactuar con Asn110 de manera independiente, sin embargo, requiere ayuda del solvente para interactuar con Asp172, por lo cual, la flexibilidad de este aminoácido es afectada y por lo tanto se crea una dependencia con el solvente, lo cual no es bueno, puesto que, la estabilidad del complejo se ve afectada.

En términos de interacciones hidrofóbicas, la hesperetina tiene la capacidad de interactuar con los aminoácidos Leu63 y Phe168 con frecuencias cercanas a 20% (**Figura 16C**), además, que presenta una interacción de tipo π - π con el aminoácido Phe168 con una frecuencia cercana a 20% (**Figura 16D**). Sin embargo, al comparar la capacidad de interacción entre la quercetina y la hesperetina, se puede observar que la hesperetina se limita a interactuar principalmente con Asn110, sin establecer interacciones con otros aminoácidos claves como His114 y Asp172, que también contribuyen significativamente en la estabilidad del complejo proteína-ligando (Alguel

et al., 2007a). Esto sugiere que, aunque la hesperetina muestra cierta afinidad y estabilidad por el bolsillo de unión de la proteína TtgR, su capacidad de acoplamiento puede verse afectada debido a la falta de interacción con His114 que se desempeña un papel fundamental de reconocimiento del ligando para la inhibición de la proteína (Alguel et al., 2007a).

En conclusión, los resultados obtenidos subrayan la ventaja de emplear métodos computacionales para el estudio eficiente y económico de proteínas y ligados, lo que ha permitido identificar posibles inhibidores, como el ácido 4-amino-3-hidroxibenzoico y la quercetina para dos sistemas de represión distintos: CymR-Vp16 y TtgR-Vp16, respectivamente. Estos hallazgos abren nuevas perspectivas en la búsqueda de compuestos inhibidores para estos sistemas de represión

5. Conclusiones.

Se logró identificar derivados del cumato y floretina que pueden interactuar con las proteínas CymR-VP16 y TtgR-VP16 mediante un estudio computacional.

De esta manera, los resultados del docking molecular permitieron identificar los compuestos ácido gálico, ácido 3-bromo-2-metoxibenzoico y ácido 4-amino-3-hidroxibenzoico como posibles inhibidores de la proteína CymR mediada por interacciones tipo puente de hidrógeno con los residuos aminoácidos Asn136 y Arg170 que pueden ser los causantes de la formación del complejo proteína-ligando. Asimismo, se pudo identificar los compuestos cloranfenicol, quercetina y hesperetina como posibles inhibidores de la proteína TtgR mediada por interacciones tipo puente de hidrógeno con los residuos Asn110 y His114 principal contribuyente en la formación del complejo proteína-ligando.

Por otra parte, los resultados de la dinámica molecular sugieren a los compuestos ácido 4-amino-3-hidroxibenzoico y la quercetina como los mejores posibles inhibidores de los complejos

CymR-VP16 y TtgR-VP16 dado su estabilidad en el sitio de unión y al presentar interacciones de tipo puente de hidrógeno con Asn136 y Arg170 en el primer caso, y His70, His114, Arg130 y Asp172 en el segundo.

6. Recomendaciones.

Inicialmente se recomienda realizar simulaciones mayores a los 200ns con el fin de confirmar la estabilidad dentro del sitio de unión y así establecer cambios conformacionales en presencia del inhibidor. De igual manera, se recomienda realizar ensayos *in vitro* con los compuestos seleccionados en el docking molecular con el fin de comprobar la supresión de la expresión génica en moscas *Drosophila melanogaster*.

Divulgación de Resultados.

Los resultados fueron presentados en el 34° Congreso Latinoamericano de Química CLAQ 2020.

El comité científico del 34° Congreso Latinoamericano de Química CLAQ 2020, el XVIII COLACRO, el X COCOCRO, el II SPAE y el IV C2B2.

Certifican que el trabajo titulado:

“Análisis *in silico* de derivados de Floretina, Cumato y Vanilato para suprimir expresión génica en *Drosophilla melanogaster* modificadas genéticamente”

Fue presentado por: Stelia Carolina Mendez Sanchez, Juan Nicolas Fuentes, Maria Lucia Otero, Luis Carlos Vesga

Presentado en modalidad oral

En el marco del 34° Congreso Latinoamericano de Química CLAQ 2020.

Llevado a cabo en la ciudad de Cartagena, Colombia del 11 al 15 de octubre de 2021.

Dr. Harold Ardila
Presidente Comité organizador

Dra. Elena Stashenko
Presidente Comité Científico
CLAQ 2020, XVIII COLACRO, X COCOCRO, II SPAE

Dra. Maricela Viola
Presidente Comité Científico IV C2B2

Referencias Bibliográficas

- Albanesi Daniela, Mansilla Maria Cecilia, Schujman Gustavo E., & de Mendoza Diego. (2005). Bacillus subtilis Cysteine Synthetase Is a Global Regulator of the Expression of Genes Involved in Sulfur Assimilation. *Journal of Bacteriology*, 187(22), 7631-7638. <https://doi.org/10.1128/JB.187.22.7631-7638.2005>
- Alguel, Y., Meng, C., Terán, W., Krell, T., Ramos, J. L., Gallegos, M.-T., & Zhang, X. (2007a). Crystal Structures of Multidrug Binding Protein TtgR in Complex with Antibiotics and

- Plant Antimicrobials. *Journal of Molecular Biology*, 369(3), 829-840.
<https://doi.org/10.1016/j.jmb.2007.03.062>
- Alguel, Y., Meng, C., Terán, W., Krell, T., Ramos, J. L., Gallegos, M.-T., & Zhang, X. (2007b). Crystal Structures of Multidrug Binding Protein TtgR in Complex with Antibiotics and Plant Antimicrobials. *Journal of Molecular Biology*, 369(3), 829-840.
<https://doi.org/10.1016/j.jmb.2007.03.062>
- Aquino, A. J. A., Tunega, D., Pašalić, H., Schaumann, G. E., Haberhauer, G., Gerzabek, M. H., & Lischka, H. (2011). Molecular Dynamics Simulations of Water Molecule-Bridges in Polar Domains of Humic Acids. *Environmental Science & Technology*, 45(19), 8411-8419.
<https://doi.org/10.1021/es201831g>
- Artemenko, N. (2008). Distance Dependent Scoring Function for Describing Protein–Ligand Intermolecular Interactions. *Journal of Chemical Information and Modeling*, 48(3), 569-574. <https://doi.org/10.1021/ci700224e>
- Bello, B., Resendez-Perez, D., & Gehring, W. J. (1998). Spatial and temporal targeting of gene expression in *Drosophila* by means of a tetracycline-dependent transactivator system. *Development*, 125(12), 2193-2202. <https://doi.org/10.1242/dev.125.12.2193>
- Bergeron, M., & Montay, G. (1997). The pharmacokinetics of quinupristin/dalfopristin in laboratory animals and in humans. *Journal of Antimicrobial Chemotherapy*, 39(suppl_1), 129-138. https://doi.org/10.1093/jac/39.suppl_1.129
- Bertram, R., & Hillen, W. (2008). The application of Tet repressor in prokaryotic gene regulation and expression. *Microbial biotechnology*, 1(1), 2-16.
- Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., & Shaw, D.

- E. (2006). Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, 43-43. <https://doi.org/10.1109/SC.2006.54>
- Brooks, B. R., Brooks III, C. L., Mackerell Jr., A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., ... Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30(10), 1545-1614. <https://doi.org/10.1002/jcc.21287>
- Case, D. A., Cheatham III, T. E., Darden, T., Gohlke, H., Luo, R., Merz Jr., K. M., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16), 1668-1688. <https://doi.org/10.1002/jcc.20290>
- Chung, J. Y., Hah, J.-M., & Cho, A. E. (2009). Correlation between Performance of QM/MM Docking and Simple Classification of Binding Sites. *Journal of Chemical Information and Modeling*, 49(10), 2382-2387. <https://doi.org/10.1021/ci900231p>
- Damm, W., Frontera, A., Tirado-Rives, J., & Jorgensen, W. L. (1997). OPLS all-atom force field for carbohydrates. *Journal of Computational Chemistry*, 18(16), 1955-1970. [https://doi.org/10.1002/\(SICI\)1096-987X\(199712\)18:16<1955::AID-JCC1>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1096-987X(199712)18:16<1955::AID-JCC1>3.0.CO;2-L)
- De Vivo, M., Masetti, M., Bottegoni, G., & Cavalli, A. (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry*, 59(9), 4035-4061. <https://doi.org/10.1021/acs.jmedchem.5b01684>

- Deans, T. L., Cantor, C. R., & Collins, J. J. (2007). A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell*, *130*(2), 363-372.
- Di, L., & Kerns, E. (2015). *Drug-like properties: Concepts, structure design and methods from ADME to toxicity optimization*. Academic press.
- Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, *9*(1), 71. <https://doi.org/10.1186/1741-7007-9-71>
- Eaton, R. W. (1997). p-Cymene catabolic pathway in *Pseudomonas putida* F1: Cloning and characterization of DNA encoding conversion of p-cymene to p-cumate. *Journal of Bacteriology*, *179*(10), 3171-3180.
- Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., & Mee, R. P. (1997). Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design*, *11*(5), 425-445. <https://doi.org/10.1023/A:1007996124545>
- Fischer, B., Fukuzawa, K., & Wenzel, W. (2008). Receptor-specific scoring functions derived from quantum chemical models improve affinity estimates for in-silico drug discovery. *Proteins: Structure, Function, and Bioinformatics*, *70*(4), 1264-1273. <https://doi.org/10.1002/prot.21607>
- Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004a). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, *47*(7), 1739-1749. <https://doi.org/10.1021/jm0306430>

- Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004b). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, *47*(7), 1739-1749. <https://doi.org/10.1021/jm0306430>
- Fussenegger, M., Morris, R. P., Fux, C., Rimann, M., von Stockar, B., Thompson, C. J., & Bailey, J. E. (2000). Streptogramin-based gene regulation systems for mammalian cells. *Nature Biotechnology*, *18*(11), 1203-1208. <https://doi.org/10.1038/81208>
- Fussenegger, M., Schlatter, S., Dätwyler, D., Mazur, X., & Bailey, J. E. (1998). Controlled proliferation by multigene metabolic engineering enhances the productivity of Chinese hamster ovary cells. *Nature biotechnology*, *16*(5), 468-472.
- Gamez, S., Vesga, L. C., Mendez-Sanchez, S. C., & Akbari, O. S. (2021). Spatial control of gene expression in flies using bacterially derived binary transactivation systems. *Insect Molecular Biology*, *30*(5), 461-471. <https://doi.org/10.1111/imb.12717>
- Gitzinger, M., Kemmer, C., Fluri, D. A., Daoud El-Baba, M., Weber, W., & Fussenegger, M. (2012). The food additive vanillic acid controls transgene expression in mammalian cells and mice. *Nucleic acids research*, *40*(5), e37-e37.
- Gohlke, H., Hendlich, M., & Klebe, G. (2000). Knowledge-based scoring function to predict protein-ligand interactions¹¹Edited by R. Huber. *Journal of Molecular Biology*, *295*(2), 337-356. <https://doi.org/10.1006/jmbi.1999.3371>
- Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., & Tjian, R. (1993). Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. *Cell*, *75*(3), 519-530. [https://doi.org/10.1016/0092-8674\(93\)90386-5](https://doi.org/10.1016/0092-8674(93)90386-5)

- Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89(12), 5547. <https://doi.org/10.1073/pnas.89.12.5547>
- Halgren, T. A. (2009). Identifying and Characterizing Binding Sites and Assessing Druggability. *Journal of Chemical Information and Modeling*, 49(2), 377-389. <https://doi.org/10.1021/ci800324m>
- Hanson, S. M., Newstead, S., Swartz, K. J., & Sansom, M. S. (2015). Capsaicin interaction with TRPV1 channels in a lipid bilayer: Molecular dynamics simulation. *Biophysical journal*, 108(6), 1425-1434.
- Harger, M., & Ren, P. (2019). Virial-based Berendsen barostat on GPUs using AMOEBA in Tinker-OpenMM. *Results in chemistry*, 1, 100004.
- Hattori, T., Takeda, A., Suzuki, K., Koike, N., Koshiishi, E., & Miyano, S. (1998). Accelerating effect of meta substituents in the ester-mediated nucleophilic aromatic substitution reaction. *Journal of the Chemical Society, Perkin Transactions 1*, 22, 3661-3672.
- Heffelfinger, G. S., & Swol, F. van. (1994). Diffusion in Lennard-Jones fluids using dual control volume grand canonical molecular dynamics simulation (DCV-GCMD). *The Journal of chemical physics*, 100(10), 7548-7552.
- Hendrichs, J., & Robinson, A. (2009). Chapter 243—Sterile Insect Technique. En V. H. Resh & R. T. Cardé (Eds.), *Encyclopedia of Insects (Second Edition)* (Second Edition, pp. 953-957). Academic Press. <https://doi.org/10.1016/B978-0-12-374144-8.00252-6>
- Hollingsworth, S. A., & Karplus, P. A. (2010). *A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins.*

- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, 51(1), 69-82. <https://doi.org/10.1021/ci100275a>
- Huang, N., Kalyanaraman, C., Bernacki, K., & Jacobson, M. P. (2006). Molecular mechanics methods for predicting protein–ligand binding. *Phys. Chem. Chem. Phys.*, 8(44), 5166-5177. <https://doi.org/10.1039/B608269F>
- Huang, S.-Y., & Zou, X. (2006). An iterative knowledge-based scoring function to predict protein–ligand interactions: I. Derivation of interaction potentials. *Journal of Computational Chemistry*, 27(15), 1866-1875. <https://doi.org/10.1002/jcc.20504>
- Huey, R., Morris, G. M., & Forli, S. (2012). Using AutoDock 4 and AutoDock vina with AutoDockTools: A tutorial. *The Scripps Research Institute Molecular Graphics Laboratory*, 10550(92037), 1000.
- Hullo Marie-Françoise, Auger Sandrine, Soutourina Olga, Barzu Octavian, Yvon Mireille, Danchin Antoine, & Martin-Verstraete Isabelle. (2007). Conversion of Methionine to Cysteine in *Bacillus subtilis* and Its Regulation. *Journal of Bacteriology*, 189(1), 187-197. <https://doi.org/10.1128/JB.01273-06>
- Iftimie, R., Minary, P., & Tuckerman, M. E. (2005). Ab initio molecular dynamics: Concepts, recent developments, and future trends. *Proceedings of the National Academy of Sciences*, 102(19), 6654-6659. <https://doi.org/10.1073/pnas.0500193102>
- Ioakimidis, L., Thoukydidis, L., Mirza, A., Naeem, S., & Reynisson, J. (2008). Benchmarking the Reliability of QikProp. Correlation between Experimental and Predicted Values. *QSAR & Combinatorial Science*, 27(4), 445-456. <https://doi.org/10.1002/qsar.200730051>

- Isabella, V. M., Lapek Jr, J. D., Kennedy, E. M., & Clark, V. L. (2009). Functional analysis of NsrR, a nitric oxide-sensing Rrf2 repressor in *Neisseria gonorrhoeae*. *Molecular Microbiology*, *71*(1), 227-239. <https://doi.org/10.1111/j.1365-2958.2008.06522.x>
- Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics*, *55*(2), 351-367. <https://doi.org/10.1002/prot.10613>
- Janin, J. (1995). Protein-protein recognition. *Progress in Biophysics and Molecular Biology*, *64*(2), 145-166. [https://doi.org/10.1016/S0079-6107\(96\)00001-6](https://doi.org/10.1016/S0079-6107(96)00001-6)
- Jennings, P., Limonciel, A., Felice, L., & Leonard, M. O. (2013). An overview of transcriptional regulation in response to toxicological insult. *Archives of Toxicology*, *87*(1), 49-72. <https://doi.org/10.1007/s00204-012-0919-y>
- John, B., & Sali, A. (2003). Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic acids research*, *31*(14), 3982-3992.
- Johnston, A. W. B., Todd, J. D., Curson, A. R., Lei, S., Nikolaidou-Katsaridou, N., Gelfand, M. S., & Rodionov, D. A. (2007). Living without Fur: The subtlety and complexity of iron-responsive gene regulation in the symbiotic bacterium *Rhizobium* and other α -proteobacteria. *BioMetals*, *20*(3), 501-511. <https://doi.org/10.1007/s10534-007-9085-8>
- Joint Center for Structural Genomics. (2006). *CRYSTAL STRUCTURE OF A TETR-FAMILY TRANSCRIPTIONAL REGULATOR (ECA1819) FROM PECTOBACTERIUM ATROSEPTICUM AT 1.64 Å RESOLUTION*. <https://doi.org/10.2210/pdb2HYT/pdb>
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). *Comparison of simple potential functions for simulating liquid water*. *79*(2), 926-935. <https://doi.org/10.1063/1.445869>

- Kaczmarczyk, A., Vorholt, J. A., & Francez-Charlot, A. (2013). Cumate-Inducible Gene Expression System for Sphingomonads and Other Alphaproteobacteria. *Applied and Environmental Microbiology*, 79(21), 6795-6802. <https://doi.org/10.1128/AEM.02296-13>
- Kishk, S. M., Kishk, R. M., Yassen, A. S. A., Nafie, M. S., Nemr, N. A., ElMasry, G., Al-Rejaie, S., & Simons, C. (2020). Molecular Insights into Human Transmembrane Protease Serine-2 (TMPS2) Inhibitors against SARS-CoV2: Homology Modelling, Molecular Dynamics, and Docking Studies. *Molecules*, 25(21). <https://doi.org/10.3390/molecules25215007>
- Konteaty, Z. D., Klon, A. E., Zou, J., & Meshkat, S. (2011). Chapter Fourteen—Computational Approach to De Novo Discovery of Fragment Binding for Novel Protein States. En L. C. Kuo (Ed.), *Fragment-Based Drug Design* (Vol. 493, pp. 357-380). Academic Press. <https://doi.org/10.1016/B978-0-12-381274-2.00014-5>
- Kramer, B. P., & Fussenegger, M. (2005). Hysteresis in a synthetic mammalian gene network. *Proceedings of the National Academy of Sciences*, 102(27), 9517-9522.
- Kufareva, I., & Abagyan, R. (2012). Methods of Protein Structure Comparison. En A. J. W. Orry & R. Abagyan (Eds.), *Homology Modeling: Methods and Protocols* (pp. 231-257). Humana Press. https://doi.org/10.1007/978-1-61779-588-6_10
- Lanig, H., Othersen, O. G., Beierlein, F. R., Seidel, U., & Clark, T. (2006). Molecular dynamics simulations of the tetracycline-repressor protein: The mechanism of induction. *Journal of molecular biology*, 359(4), 1125-1136.
- Law, P. B., & Daggett, V. (2010). The relationship between water bridges and the polyproline II conformation: A large-scale analysis of molecular dynamics simulations and crystal structures. *Protein Engineering, Design and Selection*, 23(1), 27-33. <https://doi.org/10.1093/protein/gzp069>

- Leelananda, S. P., & Lindert, S. (2016). Computational methods in drug discovery. *Beilstein journal of organic chemistry*, 12(1), 2694-2718.
- Lemak, A., & Balabaev, N. (1994). On the Berendsen thermostat. *Molecular Simulation*, 13(3), 177-187.
- Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., & Friesner, R. A. (2011). The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins: Structure, Function, and Bioinformatics*, 79(10), 2794-2812. <https://doi.org/10.1002/prot.23106>
- Liu, J., & Wang, R. (2015). Classification of Current Scoring Functions. *Journal of Chemical Information and Modeling*, 55(3), 475-482. <https://doi.org/10.1021/ci500731a>
- Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, 27(3), 221-234. <https://doi.org/10.1007/s10822-013-9644-8>
- Martin, O., & Schomburg, D. (2008). Efficient comprehensive scoring of docked protein complexes using probabilistic support vector machines. *Proteins: Structure, Function, and Bioinformatics*, 70(4), 1367-1378. <https://doi.org/10.1002/prot.21603>
- McGuire, S. E., Roman, G., & Davis, R. L. (2004). Gene expression systems in *Drosophila*: A synthesis of time and space. *TRENDS in Genetics*, 20(8), 384-391.
- Meissner, M., Varwig, D., Beier, C., Jacobi, V., Kaufmann, R., & Gille, J. (2006). Dystrophic calcinosis cutis after subcutaneous administration of para-aminosalicylic acid for treatment of pulmonary tuberculosis. *JDDG: Journal der Deutschen Dermatologischen Gesellschaft*, 4(6), 489-491.

- Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011). Molecular docking: A powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2), 146-157.
- Muhammed, M. T., & Aki-Yalcin, E. (2019). Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chemical biology & drug design*, 93(1), 12-20.
- Mullick, A., Xu, Y., Warren, R., Koutroumanis, M., Guilbault, C., Broussau, S., Malenfant, F., Bourget, L., Lamoureux, L., Lo, R., Caron, A. W., Pilotte, A., & Massie, B. (2006). The cumate gene-switch: A system for regulated expression in mammalian cells. *BMC Biotechnology*, 6(1), 43. <https://doi.org/10.1186/1472-6750-6-43>
- Nobre, D. A. C., Macedo, W. R., Silva, G. H., Lopes, L. S., & Jaimes, E. H. L. (2019). Aplicación y efecto antioxidante del ácido gálico sobre la calidad de semillas de trigo. *Revista de Ciências Agrárias*, 42(1), 22-29.
- Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of molecular biology*, 302(1), 205-217.
- Orii, C., Takenaka, S., Murakami, S., & Aoki, K. (2004). A novel coupled enzyme assay reveals an enzyme responsible for the deamination of a chemically unstable intermediate in the metabolic pathway of 4-amino-3-hydroxybenzoic acid in *Bordetella* sp. Strain 10d. *European Journal of Biochemistry*, 271(15), 3248-3254. <https://doi.org/10.1111/j.1432-1033.2004.04258.x>
- Páll, S., Abraham, M. J., Kutzner, C., Hess, B., & Lindahl, E. (2015). Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. En S. Markidis & E.

- Laure (Eds.), *Solving Software Challenges for Exascale* (pp. 3-27). Springer International Publishing.
- Phuc, H. K., Andreasen, M. H., Burton, R. S., Vass, C., Epton, M. J., Pape, G., Fu, G., Condon, K. C., Scaife, S., Donnelly, C. A., Coleman, P. G., White-Cooper, H., & Alphey, L. (2007). Late-acting dominant lethal genetic systems and mosquito control. *BMC Biology*, 5(1), 11. <https://doi.org/10.1186/1741-7007-5-11>
- Rajagopalan, S., Teter, S. J., Zwart, P. H., Brennan, R. G., Phillips, K. J., & Kiley, P. J. (2013). Studies of IscR reveal a unique mechanism for metal-dependent regulation of DNA binding specificity. *Nature Structural & Molecular Biology*, 20(6), 740-747. <https://doi.org/10.1038/nsmb.2568>
- Ramos Juan L., Martínez-Bueno Manuel, Molina-Henares Antonio J., Terán Wilson, Watanabe Kazuya, Zhang Xiaodong, Gallegos María Trinidad, Brennan Richard, & Tobes Raquel. (2005). The TetR Family of Transcriptional Repressors. *Microbiology and Molecular Biology Reviews*, 69(2), 326-356. <https://doi.org/10.1128/mmbr.69.2.326-356.2005>
- Rapp, C. S., Schonbrun, C., Jacobson, M. P., Kalyanaraman, C., & Huang, N. (2009). Automated site preparation in physics-based rescoring of receptor ligand complexes. *Proteins: Structure, Function, and Bioinformatics*, 77(1), 52-61. <https://doi.org/10.1002/prot.22415>
- Robinson, S. W., Afzal, A. M., & Leader, D. P. (2014). Bioinformatics: Concepts, methods, and data. *Handbook of Pharmacogenomics and stratified medicine*, 259-287.
- Roos, K., Wu, C., Damm, W., Reboul, M., Stevenson, J. M., Lu, C., Dahlgren, M. K., Mondal, S., Chen, W., Wang, L., Abel, R., Friesner, R. A., & Harder, E. D. (2019). OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *Journal of Chemical Theory and Computation*, 15(3), 1863-1874. <https://doi.org/10.1021/acs.jctc.8b01026>

- Sabarinathan, R., Aishwarya, K., Sarani, R., Vaishnavi, M. K., & Sekar, K. (2011). Water-mediated ionic interactions in protein structures. *Journal of Biosciences*, 36(2), 253-263. <https://doi.org/10.1007/s12038-011-9067-4>
- Salas, M. G., Cruz, A. Z., Belmares, S. Y. S., Valdivia, B., Herrera, R. R., González, C., & Chávez, J. M. (2013). Actividad anticancerígena del ácido gálico en modelos biológicos in vitro. *Revista Científica de la Universidad Autónoma de Coahuila*, 5(9), 5-11.
- Sanjeevi, M., Hebbar, P. N., Aiswarya, N., Rashmi, S., Rahul, C. N., Mohan, A., Jeyakanthan, J., & Sekar, K. (2022). Methods and applications of machine learning in structure-based drug discovery. En *Advances in Protein Molecular and Structural Biology Methods* (pp. 405-437). Elsevier.
- Santiago, C., Nguyen, K., & Schapira, M. (2011). Druggability of methyl-lysine binding sites. *Journal of Computer-Aided Molecular Design*, 25(12), 1171-1178. <https://doi.org/10.1007/s10822-011-9505-2>
- Santos, L. H., Ferreira, R. S., & Caffarena, E. R. (2019). Integrating molecular docking and molecular dynamics simulations. *Docking screens for drug discovery*, 13-34.
- Schneider, D. (2000). Using *Drosophila* as a model insect. *Nature Reviews Genetics*, 1(3), 218-226. <https://doi.org/10.1038/35042080>
- Schwartz, C. J., Giel, J. L., Patschkowski, T., Luther, C., Ruzicka, F. J., Beinert, H., & Kiley, P. J. (2001). IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. *Proceedings of the National Academy of Sciences*, 98(26), 14895-14900. <https://doi.org/10.1073/pnas.251550898>

- Sethi, A., Joshi, K., Sasikala, K., & Alvala, M. (2019). Molecular docking in modern drug discovery: Principles and recent applications. *Drug discovery and development-new advances*, 2, 1-21.
- Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: A software program for pK_a prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design*, 21(12), 681-691. <https://doi.org/10.1007/s10822-007-9133-z>
- Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., & Farid, R. (2006). Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *Journal of Medicinal Chemistry*, 49(2), 534-553. <https://doi.org/10.1021/jm050540c>
- Simossis, V. A., & Heringa, J. (2005). PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic acids research*, 33(suppl_2), W289-W294.
- Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. J. (2014). Computational methods in drug discovery. *Pharmacological Reviews*, 66(1), 334-395. <https://doi.org/10.1124/pr.112.007336>
- Stebbins, M. J., & Yin, J. C. P. (2001). Adaptable doxycycline-regulated gene expression systems for *Drosophila*. *Gene*, 270(1), 103-111. [https://doi.org/10.1016/S0378-1119\(01\)00447-4](https://doi.org/10.1016/S0378-1119(01)00447-4)
- Tan Anjiang, Fu Guoliang, Jin Li, Guo Qihong, Li Zhiqian, Niu Baolong, Meng Zhiqi, Morrison Neil I., Alphey Luke, & Huang Yongping. (2013). Transgene-based, female-specific lethality system for genetic sexing of the silkworm, *Bombyx mori*. *Proceedings of the National Academy of Sciences*, 110(17), 6766-6770. <https://doi.org/10.1073/pnas.1221700110>

- Tanous, C., Soutourina, O., Raynal, B., Hullo, M.-F., Mervelet, P., Gilles, A.-M., Noirot, P., Danchin, A., England, P., & Martin-Verstraete, I. (2008). The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in *Bacillus subtilis*. *Journal of Biological Chemistry*, 283(51), 35551-35560.
- Thomas, D. D., Donnelly, C. A., Wood, R. J., & Alphey, L. S. (2000). Insect Population Control Using a Dominant, Repressible, Lethal Genetic System. *Science*, 287(5462), 2474-2476. <https://doi.org/10.1126/science.287.5462.2474>
- Tuckerman, M. E., & Martyna, G. J. (2000). Understanding Modern Molecular Dynamics: Techniques and Applications. *The Journal of Physical Chemistry B*, 104(2), 159-178. <https://doi.org/10.1021/jp992433y>
- Venken, K. J., Simpson, J. H., & Bellen, H. J. (2011). Genetic manipulation of genes and cells in the nervous system of the fruit fly. *Neuron*, 72(2), 202-230.
- Vesga, L. C., Kronenberger, T., Tonduru, A. K., Kita, D. H., Zattoni, I. F., Bernal, C. C., Bohórquez, A. R. R., Mendez-Sánchez, S. C., Ambudkar, S. V., Valdameri, G., & Poso, A. (2021). Tetrahydroquinoline/4, 5-Dihydroisoxazole Molecular Hybrids as Inhibitors of Breast Cancer Resistance Protein (BCRP/ABCG2). 16(17), 2686-2694. <https://doi.org/10.1002/cmdc.202100188>
- Volbeda, A., Crack, J. C., Brun, N. E. L., & Fontecilla-Camps, J. C. (2020). Structure–Function Relationships of the NsrR and RsrR Transcription Regulators. En *Encyclopedia of Inorganic and Bioinorganic Chemistry* (pp. 1-23). <https://doi.org/10.1002/9781119951438.eibc2704>
- Volbeda, A., Dodd, E. L., Darnault, C., Crack, J. C., Renoux, O., Hutchings, M. I., Le Brun, N. E., & Fontecilla-Camps, J. C. (2017). Crystal structures of the NO sensor NsrR reveal how its

- iron-sulfur cluster modulates DNA binding. *Nature Communications*, 8(1), 15052.
<https://doi.org/10.1038/ncomms15052>
- Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., & Hou, T. (2019). End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews*, 119(16), 9478-9508.
<https://doi.org/10.1021/acs.chemrev.9b00055>
- Wang, Y., Shaikh, S. A., & Tajkhorshid, E. (2010). Exploring transmembrane diffusion pathways with molecular dynamics. *Physiology*.
- Williams, M., & Daviter, T. (2016). Protein-ligand interactions. *Biophys. Physicobiol*, 13, 85.
- You, L., Cox, R. S., Weiss, R., & Arnold, F. H. (2004). Programmed population control by cell-cell communication and regulated killing. *Nature*, 428(6985), 868-871.
<https://doi.org/10.1038/nature02491>
- Yu, W., & MacKerell, A. D. (2017). Computer-Aided Drug Design Methods. En P. Sass (Ed.), *Antibiotics: Methods and Protocols* (pp. 85-106). Springer New York.
https://doi.org/10.1007/978-1-4939-6634-9_5

Apéndices

Apéndice 1 Base de datos de compuestos para CymR

PubChem ID	Nombre del metabolito	Masa molecular	Score de Biodisponibilidad	Donadores de puentes de hidrógeno	Aceptores puentes de hidrógeno	iLOGP	Absorción gastrointestinal (GI Absortion)
54671203	Doxiciclina	414.41	0.11	6	8	1.93	Baja
54675776	Tetraciclina	444.43	0.11	6	9	1.19	Baja
10820	Cumato	164.2	0.85	1	2	1.85	Alta
370	Ácido Gálico	170.12	0.56	4	5	0.21	Alta
72	B-Resorcilato	154.12	0.56	3	4	0.66	Alta
7428	Galato de metilo	184.15	0.55	3	5	0.97	Alta
12081	4-Bromobenzoato de metilo	215.04	0.55	0	2	2.33	Alta
54675864	3,5-Dibromo-4-Hidroxibenzoato	294.9	0.85	1	3	1.56	Alta
86023	Ácido 4-Bromo-3,5-Dihidroxibenzoico	233.02	0.56	3	4	1.14	Alta
68385	Clorhidrato de Trimetobenzamida	424.92	0.55	1	6	0	Alta
2760954	Ácido 4-(Terc-Butoxicarbonil) Benzoico	251.28	0.56	2	4	2.21	Alta
18749	Ácido 3,5-Dicloro-4-Hidroxibenzoico	207.01	0.85	2	3	1.31	Alta
137566	Ácido 4-Amino-3-Hidroxibenzoico	153.14	0.56	3	3	0.52	Alta
75076	3,5-Dihidroxibenzoato De Metilo	168.15	0.55	2	4	1.33	Alta
65083	Ácido 3-Amino-4-Hidroxibenzoico	153.14	0.56	3	3	0.6	Alta
256687	4-(Bromometil)Benzoato De Metilo	229.07	0.55	0	2	2.42	Alta
765380	Ácido 4-Metoxi-1-Naftoico	202.21	0.85	1	3	1.71	Alta
242	Benzoato	121.11	0.85	0	2	1.12	Alta
69065	Ácido 1,4-Naftalenodicarboxílico	216.19	0.85	2	4	1.14	Alta
255720	Ácido 2-Amino-3-Metoxibenzoico	167.16	0.85	2	3	0.8	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

69273	4-Yodobenzoato De Metilo	262.04	0.55	2	2	2.31	Alta
19829	Ácido 3,4-Dihidroxi-5-Metoxibenzoico	184.15	0.56	3	5	1.26	Alta
1406520	4-propan-2-ilbenzoato de fenilo	240.3	0.55	0	2	3.1	Alta
219219	1,5-diacetilnaftaleno	212.24	0.55	0	2	2.04	Alta
291203	Benzoato de piridin-4-ilo	199.21	0.55	0	3	2.06	Alta
348912	Ácido 4-etil-1-naftoico	200.23	0.85	1	2	1.8	Alta
390995	6-hidroxi-2-naftoato de metilo	202.21	0.55	1	3	2.02	Alta
240439	2-hidroxitereftalato de dimetilo	210.18	0.55	1	5	2.35	Alta
245116	Ácido 3,5-diamino-4-metilbenzoico	166.18	0.56	3	2	0.78	Alta
299486	2-yodo-3-metoxibenzoato de metilo	292.07	0.55	0	3	2.44	Alta
253598	Ácido 3-bromo-2-metoxibenzoico	231.04	0.85	1	3	1.24	Alta
282745	6,7-dimetoxinaftaleno-2-carboxilato de metilo	246.26	0.55	0	4	2.84	Alta
3898848 6	[(2R)-2-etilhexilo] 4-(dimetilamino)benzoato	277.4	0.55	0	2	3.79	Alta
3898848 7	[(2S)-2-etilhexilo] 4-(dimetilamino)benzoato	277.4	0.55	0	2	3.79	Alta
7004829	Ácido 4-[(1S)-1-bromoetil]benzoico	229.07	0.85	1	2	1.86	Alta
7004831	Ácido 4-[(1R)-1-bromoetil]benzoico	229.07	0.85	1	2	1.81	Alta
522695	Ácido 4-(dimetilamino)-1-naftoico	215.25	0.85	1	2	1.86	Alta
458962	Ácido 5-nitronaftaleno-1-carboxílico	217.18	0.56	1	4	1.17	Alta
5896894	Ácido 2-oxo-1,2-dihidrobenzo[cd]indol-6-carboxílico	213.19	0.85	2	3	1.02	Alta
4437804	Ácido 3-isopropil-2-metoxibenzoico	194.23	0.85	1	3	1.9	Alta
212595	3-formil-4-hidroxibenzoato de metilo	180.16	0.55	1	4	1.35	Alta
12065	Ácido 3,5-diiodo-4-hidroxibenzoico	389.91	0.85	2	3	1.51	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

3729087	4-(dibromometil)benzoato de metilo	307.97	0.55	0	2	2.63	Alta
1328832 6	Ácido 4-nitro-1-naftoico	217.18	0.56	1	4	0.98	Alta
3692708	1-metil-3,4-dihidro-2H-quinolina-6-carboxilato de metilo	205.25	0.55	0	2	2.5	Alta
1486286	Ácido 4-(aminooximetil)benzoico	167.16	0.85	2	4	1.29	Alta
8631043 9	Ácido 4-[(2R)-2-etilhexoxi]carbonilbenzoico	278.34	0.85	1	4	3.04	Alta
8631044 1	Ácido 4-[(2S)-2-etilhexoxi]carbonilbenzoico	278.34	0.85	1	4	3.04	Alta
145162	Ácido 2-bromo-3-metoxibenzoico	231.04	0.85	1	3	1.96	Alta
458219	Ácido 2-yodo-3-metoxibenzoico	278.04	0.85	1	3	1.56	Alta
9451992	2-oxo-1,2,3,4-tetrahydroquinolina-6-carboxilato de metilo	205.21	0.55	1	3	1.97	Alta
579631	Ácido 6-(metoxicarbonil)nicotínico	181.15	0.56	1	5	1.15	Alta
1984915 8	Ácido 3-amino-2-metoxibenzoico	167.16	0.85	2	3	1.06	Alta
1230900 7	Ácido 3-metoxi-2-(metilamino)benzoico	181.19	0.85	2	3	1.29	Alta
1481529 1	3- HIDROXI -2- METOXIBENZOATO DE METILO	182.17	0.55	1	4	2.01	Alta
8481931 1	Ácido 4-fluoro-3,5-dihidroxibenzoico	172.11	0.56	3	5	0.68	Alta
1906997 1	Ácido 6-hidroxi-5,6,7,8-tetrahidronaftaleno-2-carboxílico	192.21	0.85	2	3	1.58	Alta
1273035 7	2-formil-3-metoxibenzoato de metilo	194.18	0.55	0	4	1.74	Alta
1807125 0	Ácido 4-[(2,2-Dimetilpropanamido)metil]benzoico	235.28	0.85	2	3	1.96	Alta
1026228 8	Ácido 8-metoxiquinolona-5-carboxílico	203.19	0.85	1	4	1.48	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

1149285 5	[(2R)-2-etilhexilo] 4-nitrobenzoato	279.33	0.55	0	4	3.19	Alta
1165916 1	[(2S)-2-etilhexilo] 4-nitrobenzoato	279.33	0.55	0	4	3.19	Alta
3374705 0	Ácido 8-metoxiquinolina-4-carboxílico	203.19	0.85	1	4	1.36	Alta
3757638 9	4-[(2,2-dimetilpropanoilamino)metil]-N-metilbenzamida	248.32	0.55	2	2	2.47	Alta
583760	Ácido 4-etilbenzoico, éster isopropílico	192.25	0.55	0	2	2.95	Alta
589438	Etanona, 1-[4-(1-metiletetil)fenil]-	160.21	0.55	0	1	2.26	Alta
1293616 97	[(2R)-2-etilhexilo] 4-cianobenzoato	259.34	0.55	0	3	3.09	Alta
1352933 4	Éster metílico del ácido 4-yodometil-benzoico	276.07	0.55	0	2	2.4	Alta
1384470 4	4-acetilbenzoato de terc-butilo	220.26	0.55	0	3	2.67	Alta
9297829 2	4-[(R)-metilsulfinil]benzoato de metilo	198.24	0.55	0	3	2.05	Alta
1447343 5	(-)-4-(1-Metoxietil)ácido benzoico	180.2	0.85	1	3	1.74	Alta
1530911 6	5-oxo-5,6,7,8-tetrahidronaftaleno-2-carboxilato de metilo	204.22	0.55	0	3	2.27	Alta
1092280 3	2-bromo-3-metoxibenzoato de metilo	245.07	0.55	0	3	2.4	Alta
1093251 6	Ácido 5-metoxinaftaleno-1-carboxílico	202.21	0.85	1	3	1.81	Alta
1155008 6	Ácido 4-(1-oxopropan-2-il)benzoico	178.18	0.85	1	3	1.28	Alta
1176965 4	4-oxocromano-7-carboxilato de metilo	206.19	0.55	0	4	2.12	Alta
5343258 7	4-amino-3-metilbenzoato de terc-butilo	207.27	0.55	1	2	2.55	Alta
6098499 1	4-propan-2-ilbencenocarboximidato de metilo	177.24	0.55	1	2	2.59	Alta
1137507	4-(hidroxiamino)benzoato de metilo	459.56	0.55	2	3	3.81	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

10262208	2-cloro-3-metoxibenzoato de metilo	200.62	0.55	0	3	2.27	Alta
45081407	4-(Propan-2-il)benzoilo fluoruro	166.19	0.55	0	2	2.22	Alta
23498173	fluoruro de 4-(difluorometil)benzoilo	174.12	0.55	0	4	1.68	Alta
45081539	4-(fluorocarbonil)benzoato de metilo	182.15	0.55	0	4	1.9	Alta
12520613	N-oxobenzamide	135.12	0.55	0	3	1.61	Alta
66663605	(6S)-6-Amino-5,6,7,8-tetrahidronaftaleno-2-ácido carboxílico	191.23	0.55	2	3	1.59	Alta
86316640	Ácido (6R)-6-amino-5,6,7,8-tetrahidronaftaleno-2-carboxílico	191.23	0.55	2	3	1.59	Alta
12274276	4-(metilamino)benzoato de propan-2-ilo	193.24	0.55	1	2	2.48	Alta
93501005	(5S)-5-amino-5,6,7,8-tetrahidronaftaleno-2-carboxilato de metilo	205.25	0.55	1	3	2.32	Alta
92134902	(5R)-5-amino-5,6,7,8-tetrahidronaftaleno-2-carboxilato de metilo	205.25	0.55	1	3	2.32	Alta
45094190	2-etoxi-3-metoxibenzoato de metilo	210.23	0.55	0	4	2.64	Alta
45096600	2-(4-isopropilfenil)-1-propen-1-ona	174.24	0.55	0	1	2.61	Alta
18000903	2-[4-(Propan-2-il)fenil]prop-2-enal	174.24	0.55	0	1	2.41	Alta
70815235	4-[(1S)-1-hidroxiopropil]benzoato de terc-butilo	236.31	0.55	1	3	3.02	Alta
93510404	4-[(1R)-1-hidroxiopropil]benzoato de terc-butilo	236.31	0.55	1	3	3.02	Alta
86316960	4-[(1S)-1-bromoetil]benzoato de metilo	243.1	0.55	0	2	2.64	Alta
86316961	4-[(1R)-1-bromoetil]benzoato de metilo	243.1	0.55	0	2	2.64	Alta
12475261	1-(p-tolil)propano-1,2-diona	162.19	0.55	0	2	1.37	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

12493228	4-metoxibenzoato de terc-butilo	208.25	0.55	0	3	2.93	Alta
12730356	Ácido 2-formil-3-metoxibenzoico	180.16	0.85	1	4	1.13	Alta
45156962	Éster metílico del ácido 4-tiocarboxibenzoico	196.22	0.55	0	3	2.03	Alta
13144281	6-cianopiridina-3-carboxilato de metilo	162.15	0.55	0	4	1.58	Alta
13432708	Ácido benzoico, 3,5-dimercapto-	186.25	0.56	1	2	1.44	Alta
45812856	N-{{4-(metoximetil)fenil}metil}-2,2-dimetilpropanamida	235.32	0.55	1	2	2.76	Alta
61308278	3-amino-4-bromobenzoato de terc-butilo	272.14	0.55	1	2	2.67	Alta
83529762	N'-amino-4-propan-2-ilbencenocarboximidamida	177.25	0.55	2	1	1.55	Alta
67894539	Ácido 3-[(2S)-2-aminopropil]benzoico	179.22	0.55	2	3	1.79	Alta
22819068	Ácido 3-[(2R)-2-aminopropil]benzoico	179.22	0.55	2	3	1.79	Alta
11321881	Ácido 1,5-naftalenodicarboxílico	216.19	0.85	2	4	1.07	Alta
15893378	5-nitronaftaleno-1-carboxilato de metilo	231.2	0.55	0	4	2.07	Alta
46311082	3-bromo-2-metoxibenzoato de metilo	245.07	0.55	0	3	2.41	Alta
53212146	Ácido 1-(dimetilamino)isoquinolina-4-carboxílico	216.24	0.85	1	3	1.8	Alta
69450511	Ácido 4-ciclohexanocarbonilbenzoico	232.28	0.85	1	3	2.06	Alta
62234050	1-(4-Propan-2-ilfenil)but-3-in-1-ona	186.25	0.55	0	1	2.58	Alta
82125790	1-(4-Propan-2-ilfenil)-2-sulfaniletanona	194.29	0.55	0	1	2.17	Alta
20083401	Ácido 4-(piperidin-4-ilmetil)benzoico	219.28	0.55	2	3	2.11	Alta
15887808	Ácido 5-clorosulfonilnaftaleno-1-carboxílico	270.69	0.56	1	4	1.28	Alta
10899634	Ácido 4-formilnaftaleno-1-carboxílico	200.19	0.85	1	3	1.27	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

8226319 9	Ácido 4-propano-2- iloxinaftaleno-1-carboxílico	230.26	0.85	1	3	1.84	Alta
9542239 4	Ácido 8-propano-2- iloxiquinolina-5-carboxílico	231.25	0.85	1	4	1.87	Alta
1790994 8	1-(4-Propan-2-ilfenil)prop- 2-en-1-ona	174.24	0.55	0	1	2.48	Alta
1097835 0	Ácido 4- (bromometil)naftaleno-1- carboxílico	265.1	0.85	1	2	1.87	Alta
1025470 79	Ácido 3-aliloxi-4-isopropil- benzoico	220.26	0.85	1	3	2.52	Alta
7489280 6	Ácido 1H-indazol-3,7- dicarboxílico	206.15	0.56	3	5	0.17	Alta
1351248 7	(4- Isopropilfenil)(oxo)acetalde hído	176.21	0.55	0	2	1.43	Alta
1255975 9	Ácido 2-(metiltio)-3- metoxibenzoico	198.24	0.85	1	3	1.57	Alta
8352839 8	Ácido 8-metoxiisoquinolina- 4-carboxílico	203.19	0.85	1	4	0.95	Alta
5389153 3	6-isopropilnicotinato de metilo	179.22	0.55	0	3	2.43	Alta
9660802 6	Ácido 2- (ciclohexilmetil)pirimidina- 5-carboxílico	220.27	0.85	1	4	2.06	Alta
5978000 5	Ácido benzoico, 4- (metiltio)-, éster isopropílico	210.29	0.55	0	2	2.87	Alta
1875237 4	Ácido 3-propilbenzoico	164.2	0.85	1	2	1.94	Alta
8363787 6	Ácido 8-etilquinolina-4- carboxílico	201.22	0.85	1	3	1.67	Alta
4420316 1	3-amino-4,5- dihidroxibenzoato de metilo	183.16	0.55	3	4	1.22	Alta
7684688 3	Ácido 1H-indol-7- carboxílico, 3- (aminocarbonil)-	204.18	0.56	3	3	0.89	Alta
7684688 4	Ácido 3-acetil-1H-indol-7- carboxílico	203.19	0.85	2	3	1.32	Alta
5341850 5	Ácido 5-hidrazinil-1- naftoico	202.21	0.56	3	3	1.11	Alta
2231339 2	Cromano-7-carboxilato de metilo	192.21	0.55	0	3	2.49	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

5134617 1	Ácido 4-amino-3,5-dihidroxibenzoico	169.13	0.56	4	4	0.23	Alta
6705017 8	Ácido 3-hidroxi-4-(metoxicarbonil)benzoico	196.16	0.56	2	5	1.6	Alta
1487705 9	Éster metílico del ácido 3-cloro-2-metoxibenzoico	200.62	0.55	0	3	2.37	Alta
91682	1,3-propanodiona, 1-[4-(1-metiletil)fenil]-3-fenil-	266.33	0.55	0	2	2.51	Alta
6952219 9	5-nitroisoquinolina-1-carboxilato de metilo	232.19	0.55	0	5	1.6	Alta
8633885 9	Ácido 4-[(1S)-1-cloroetil]benzoico	184.62	0.85	1	2	1.72	Alta
8633886 1	Ácido 4-[(1R)-1-cloroetil]benzoico	184.62	0.85	1	2	1.72	Alta
8481928 9	Fenilpirimidina-5-carboxilato	200.19	0.55	0	4	1.88	Alta
6764273 7	fenilpirimidina-2-carboxilato	200.19	0.55	0	4	1.88	Alta
4509657 9	cianuro de 4-(propan-2-il)benceno-1-carbonilo	173.21	0.55	0	2	2.01	Alta
6138013 2	1-(4-Propan-2-ilfenil)but-3-en-1-ona	188.27	0.55	0	1	2.67	Alta
5527250 0	Ácido (7S)-7-amino-5,6,7,8-tetrahidronaftaleno-2-carboxílico	191.23	0.55	2	3	1.42	Alta
6666353 0	Ácido (7R)-7-amino-5,6,7,8-tetrahidronaftaleno-2-carboxílico	191.23	0.55	2	3	1.42	Alta
5527632 5	Ácido (R)-4-(1-(metilamino)etil)benzoico	179.22	0.55	2	3	1.65	Alta
1110697 7	Cloruro de 4-(diclorometil)benzoilo	223.48	0.55	0	1	2.12	Alta
1306652 75	Ácido (3S)-3-metil-1,2,3,4-tetrahidroisoquinolina-6-carboxílico	191.23	0.55	2	3	1.51	Alta
8401404 1	Ácido 2-metoxicarbonilpirimidina-5-carboxílico	182.13	0.56	1	6	0.98	Alta
2215872 5	Ácido 4,8-quinolindicarboxílico	217.18	0.56	2	5	0.84	Alta
1840737 4	Ácido 3,4,5-triamino-benzoico	167.17	0.56	4	5	0.19	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

6692053 8	Ácido 3-amino-4,5-dihidroxibenzoico	169.13	0.56	4	4	0.04	Alta
5398156 8	Ácido 4-(metilamino)naftaleno-1-carboxílico	201.22	0.85	2	2	1.48	Alta
1665561 0	Ácido 2-mercapto-3-metoxibenzoico	184.21	0.56	1	3	1.45	Alta
2293714 5	Ácido 4-(2-oxoetil)benzoico	164.16	0.85	1	3	0.95	Alta
5422244 5	3-yodo-2-metoxibenzoato de metilo	292.07	0.55	0	3	2.39	Alta
1127533 6	Éster metílico del ácido 4-(1-metiletetil)benzoico	176.21	0.55	0	2	2.58	Alta
1267807 7	Ácido a,a-dibromo-p-toluico	293.94	0.85	1	2	1.83	Alta
2362585 6	2-cianopirimidina-5-carboxilato de metilo	163.13	0.55	0	5	1.46	Alta
542787	4-nitrosobenzoato de metilo	165.15	0.55	0	4	2.14	Alta
2302428 7	Ácido 2-metoxi-3-(metilamino)benzoico	181.19	0.85	2	3	1.45	Alta
5413957 1	1-(5-aminonaftalen-1-il)etanon	185.22	0.55	1	1	1.68	Alta
6860971 6	3,4-dihidroxi-5-metilbenzoato de metilo	182.17	0.55	2	4	1.83	Alta
2194627 3	Ácido 5-formilnaftaleno-1-carboxílico	200.19	0.85	1	3	1.29	Alta
1310317 89	Ácido 4-[(2R)-1-hidroxipropan-2-il]benzoico	180.2	0.85	2	3	1.46	Alta
8934399 7	(S)-4-(1-(Metilamino)etil)ácido benzoico	179.22	0.55	2	3	1.65	Alta
1937098 3	Ácido 4-(diclorometil)benzoico	179.22	0.55	2	3	1.65	Alta
6845845 2	Ácido-3-amino-5-hidroxi-4-yodo-benzoico,	279.03	0.56	3	3	0.98	Alta
5783064 4	2-Fluoro-1-[4-(1-metiletetil)fenil]etanon	180.22	0.55	0	2	2.04	Alta
5375095 3	3-(hidroximetil)-2-metoxibenzoato de metilo	196.2	0.55	1	4	2.17	Alta
6941898 9	Éster isopropílico del ácido 4-cloro-3-hidroxi-benzoico	214.65	0.55	1	3	2.54	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

4411938 6	Ácido 3-borono-2-metoxibenzoico	195.97	0.56	3	5	0	Alta
2271186 3	Ácido 4-carboxinaftaleno-1-borónico	216	0.56	3	4	0	Alta
2734369	Ácido 4-metoxicarbonilfenilborónico	179.97	0.55	2	4	0	Alta
5321671 3	Ácido 5-borono-1-naftoico	216	0.56	3	4	0	Alta
8465438 8	1-(Ftalazin-5-il)etanona	172.18	0.55	0	3	1.47	Alta
8465838 0	Ácido 3-(2-fluoropropil)benzoico	182.19	0.85	1	3	1.78	Alta
8466041 0	Ácido 3-(2,2-difluoroetil)benzoico	186.16	0.85	1	4	1.52	Alta
8466528 3	Ácido 7-fluoro-5,6,7,8-tetrahidronaftaleno-2-carboxílico	194.2	0.85	1	3	1.7	Alta
1268449 07	Clorhidrato de ácido 4-[(Piperidin-3-il)metil]benzoico	255.74	0.55	2	3	0	Alta
9117382 4	Ácido 3-(hidroximetil)-2-metoxibenzoico	182.17	0.85	2	4	1.14	Alta
8610580 6	Ácido 3-amino-4-cloro-5-hidroxibenzoico	187.58	0.56	3	3	0.93	Alta
9190271 1	Ácido 3-(aminometil)-4,5-dihidroxibenzoico	186.16	0.55	4	5	0.7	Alta
1039471 75	4-hidroxil-3-metilbenzoato de propan-2-ilo	194.23	0.55	1	3	2.34	Alta
1212357 83	4-(dimetilsilil)benzoato de metilo	194.3	0.55	0	2	2.75	Alta
1038302 65	4-bromo-3-hidroxibenzoato de propan-2-ilo	259.1	0.55	1	3	2.67	Alta
9166385 8	6-cianopiridazina-3-carboxilato de metilo	163.13	0.55	0	5	1.26	Alta
1067664 80	Ácido 1-(metilamino)isoquinolina-4-carboxílico	202.21	0.85	2	3	1.29	Alta
1067677 93	Ácido 1-metilsulfanilisoquinolina-4-carboxílico	219.26	0.56	1	3	1.78	Alta
1067673 95	Ácido 1-(hidroxiamino)isoquinolina-4-carboxílico	204.18	0.56	3	4	0.88	Alta

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

131081665	Ácido (3R)-3-metil-1,2,3,4-tetrahydroisoquinolina-7-carboxílico	191.23	0.55	2	3	1.72	Alta
131002589	Ácido (3S)-3-metil-1,2,3,4-tetrahydroisoquinolina-7-carboxílico	191.23	0.55	2	3	1.72	Alta
129949658	Ácido 3,4-dihidroxi-5-yodobenzoico	280.02	0.56	3	4	0.96	Alta
163041246	4-nitrosobenzoato de propan-2-ilo	193.2	0.55	0	4	2.62	Alta

Apéndice 2. Base de datos de compuestos para TtgR

PubChem ID	Nombre del metabolito	Masa molecular	Donadores de puentes de hidrógeno	Aceptores de puentes de hidrógeno	iLOG P	%Absorción gastrointestinal (GI Absortion)
425	GU17	256.257	1.000	2.250	1.437	81.006
2483	2',3,4,4'-tetrahydroxy Chalcone	272.257	2.000	3.000	0.935	71.675
2535	3-(3,4-Dihydroxyphenyl)-1-(2-hydroxy-4-methoxyphenyl)prop-2-en-1-one	286.284	3.000	4.250	0.803	79.244
4788	Phloretin	274.273	2.000	3.250	0.882	75.344
14611	2-Propen-1-one, 1-(hydroxyphenyl)-3-phenyl-	224.259	0.000	1.750	2.166	100.000
25861	4,2'-Dihydroxychalcone	240.258	1.000	2.500	1.551	90.020
25862	2',4',3-Trihydroxy-4-methoxychalcone	286.284	1.000	2.250	1.903	84.428
89340	3-(4-Hydroxy-phenyl)-1-phenyl-prop-2-en-1-one	224.259	1.000	2.750	1.557	94.581
94240	1-(4-Hydroxy-phenyl)-3-phenyl-prop-2-en-1-one	224.259	0.000	1.750	2.021	100.000
97766	3-(3,4-Dihydroxyphenyl)-1-(2-hydroxyphenyl)-2-propen-1-one	256.257	3.000	4.250	0.392	78.208
97894	1-(5-Chloro-2-hydroxyphenyl)-3-(4-chlorophenyl)prop-2-en-1-one	293.149	0.000	1.750	3.084	100.000

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

103608	2',4,4',6'-Tetrahydroxy-3-methoxychalcone	302.283	2.000	3.000	1.154	63.925
118586	2',4',4'-Trihydroxy-3-methoxychalcone	286.284	1.500	4.000	1.170	70.753
121065	5'-Chloro-2'-hydroxychalcone	258.704	0.000	1.750	2.656	100.000
125157	CID 125157	300.310	3.000	4.500	1.108	79.783
130238	1-(2,4-Dihydroxy-3-methoxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	286.284	2.500	5.000	0.599	67.970
134561	1,3-Bis-(4-hydroxyphenyl)-2-propen-1-one	240.258	1.000	2.500	1.448	85.602
147608	Hesperetin dihydrochalcone	304.299	2.000	3.250	1.203	64.826
154102	1-(2,4-Dihydroxy-6-methoxyphenyl)-3-phenylprop-2-en-1-one	270.284	1.000	2.500	1.578	89.016
155802	2-Propen-1-one, 3-(4-hydroxyphenyl)-1-(2,4,6-trihydroxyphenyl)-	272.257	2.000	3.000	0.791	74.194
162957	2',4',4'-Trihydroxy-6'-methoxychalcone	286.284	2.000	3.250	0.906	76.738
165210	1-(2,4-Dihydroxyphenyl)-3-(3,4-dimethoxyphenyl)prop-2-en-1-one	300.310	2.333	3.833	1.484	75.309
166795	Chalcone, 2',4'-dihydroxy-4-methoxy-	270.284	1.000	2.500	1.853	91.785
177052	3-(4-Chlorophenyl)-1-(4-hydroxyphenyl)-2-propen-1-one	258.704	0.000	1.750	2.449	100.000
179533	Dihydropashanone	302.326	3.333	5.083	1.108	79.581
213726	1-(2,4-Dihydroxyphenyl)-3-(2-hydroxyphenyl)prop-2-en-1-one	256.257	1.000	2.250	1.458	86.445
244230	4-Hydroxy-4'-chlorochalcone	258.704	1.000	2.750	1.955	96.911
261164	CID 261164	288.256	2.000	2.750	0.835	66.253
261706	2-Propen-1-one,3-(4-hydroxy-3-methoxyphenyl)-1-(2-hydroxyphenyl)-	270.284	2.000	3.500	1.258	90.361
276225	1-(2-Hydroxy-5-methyl-phenyl)-3-phenyl-prop-2-en-1-one	238.285	0.000	1.750	2.673	100.000
298192	4-[1,1,2,2,3,3-Hexafluoro-3-(4-hydroxyphenyl)propyl]phenol	336.233	2.000	1.500	1.473	90.405
344530	2-Propen-1-one, 1-(2,4-dihydroxyphenyl)-3-phenyl-	240.258	0.000	1.500	2.008	92.694
380962	4,4'-Dihydroxy-3,3'-dimethoxychalcone	300.310	2.000	3.500	1.861	76.503
442255	Asebogenin	288.299	3.000	4.500	0.792	70.201
459987	2,4,2'-Trihydroxy-5'-methyl chalcone	270.284	2.000	3.250	1.420	85.590
460715	3-(2,4-Dihydroxyphenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	256.257	2.000	3.250	1.073	82.083
460718	CID 460718	256.257	1.000	2.250	1.477	86.553
466245	CID 466245	258.704	0.000	1.750	2.656	100.000
468135	4,5,4'-Trihydroxychalcone	256.257	2.000	3.250	0.833	75.474
503271	2,4,6-Trihydroxy-3-(3-hydroxy-3-phenylprop-2-enoyl)-5-methylbenzaldehyde	314.294	0.000	3.000	2.038	88.262

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

592216	2-Propen-1-one, 1-(2-hydroxy-4-methoxyphenyl)-3-(4-hydroxyphenyl)-	270.284	2.000	3.500	1.309	88.601
597332	1-(2,4-Dihydroxy-5-methoxyphenyl)-3-phenylprop-2-en-1-one	270.284	1.000	2.500	1.776	91.896
597670	1-(2,4-Dihydroxy-6-methoxyphenyl)-3-(2-hydroxyphenyl)prop-2-en-1-one	286.284	2.000	3.250	1.295	80.722
606469	2-Propen-1-one, 1-(2,6-dihydroxy-4-methoxyphenyl)-3-phenyl-, (E)-	270.284	2.000	3.500	1.334	94.062
617249	1-(2-Hydroxyphenyl)-3-(4-methylphenyl)prop-2-en-1-one	238.285	0.000	1.750	2.387	100.000
617944	4-Hydroxyl-4'-fluorochalcone	242.249	1.000	2.750	1.772	95.841
637162	2',4'-Dihydroxy-4,6'-dimethoxydihydrochalcone	302.326	3.000	3.500	1.264	75.352
638276	2'-Hydroxychalcone	224.259	0.000	1.750	2.166	100.000
638278	Trihydroxychalcone	256.257	1.000	2.250	1.437	81.006
641785	Cardamonin	270.284	1.000	2.500	1.578	89.016
795426	1-(3-Amino-1,2,4-triazol-4-yl)-3-phenylpropan-1-one	216.242	2.000	3.500	0.225	68.326
805755	1-(2,6-Dihydroxy-4-methylphenyl)-2-phenoxyethanone	258.273	1.000	3.750	1.163	93.059
820823	1-(4-Hydroxy-2-methoxyphenyl)-3-(4-hydroxyphenyl)propan-1-one	272.300	2.000	3.500	1.514	84.080
854214	2-(4-Fluorophenoxy)-1-(2,4,6-trihydroxyphenyl)ethanone	278.236	1.000	3.500	1.116	84.441
854223	2-(4-Methoxyphenoxy)-1-(2,4,6-trihydroxyphenyl)ethanone	290.272	2.000	4.500	0.629	68.632
854793	1-(2,4-Dihydroxy-6-methylphenyl)-2-phenoxyethan-1-one	258.273	0.000	2.500	1.775	90.687
899213	2-(4-Bromophenoxy)-1-(2,4-dihydroxyphenyl)ethanone	323.142	0.000	2.500	2.010	92.063
1226045	2',4',6'-Trihydroxydihydrochalcone	258.273	1.000	2.500	1.560	87.653
2066126	2-(4-Chlorophenoxy)-1-(2,4-dihydroxyphenyl)ethanone	278.691	0.000	2.500	2.007	92.046
2777833	CID 2777833	238.285	0.000	1.750	2.281	100.000
2778084	CID 2778084	293.149	0.000	1.750	3.045	100.000
2778558	1-(5-Fluoro-2-hydroxyphenyl)-3-phenylprop-2-en-1-one	242.249	0.000	1.750	2.493	100.000
2780755	4-Fluoro-2'-hydroxychalcone	242.249	0.000	1.750	2.423	100.000
2780925	6'-Methoxy-2',4'-dihydroxychalcone	270.284	2.000	3.500	1.047	85.904
2794955	2-Methyl-2'-hydroxychalcone	238.285	0.000	1.750	2.214	100.000
2888278	3-(2,4-Dichlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	293.149	1.000	2.750	2.567	100.000
2888668	2-Propen-1-one, 1-(4-hydroxyphenyl)-3-(4-methylphenyl)-	238.285	0.000	1.750	2.217	100.000
2926587	4-Chloro-5'-fluoro-2'-hydroxychalcone	276.694	0.000	1.750	2.902	100.000

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

2926703	1-(5-Chloro-2-hydroxyphenyl)-3-(4-fluorophenyl)prop-2-en-1-one	276.694	0.000	1.750	2.913	100.000
2926733	4-Chloro-2'-hydroxy-5'-methylchalcone	272.730	0.000	1.750	3.082	100.000
2927265	2-Propen-1-one, 1-(2-hydroxy-5-methylphenyl)-3-(4-methylphenyl)-	252.312	0.000	1.750	2.956	100.000
3060385	C15H11BrO3	319.154	0.000	2.750	1.991	94.909
3086055	3-(4-Methoxyphenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	288.299	2.000	3.500	1.295	73.142
3122803	3-(4-Hydroxyphenyl)-1-thiophen-2-ylprop-2-en-1-one	230.281	1.000	2.250	1.662	95.025
3700993	3-(4-Nitrophenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	303.271	1.000	4.500	0.485	66.447
3734728	1-(5-Chloro-2-hydroxyphenyl)-3-(4-methylphenyl)prop-2-en-1-one	272.730	0.000	1.750	2.877	100.000
3977068	2'-Hydroxy-4,3',5'-trichlorochalcone	327.594	0.000	1.750	3.502	100.000
3995255	2-Propen-1-one, 1-(2,5-dihydroxyphenyl)-3-(4-hydroxyphenyl)-	256.257	1.000	2.250	1.507	81.545
4021103	3-(4-Hydroxy-3-methoxyphenyl)-1-(2-hydroxy-5-methylphenyl)prop-2-en-1-one	284.311	1.000	2.500	2.480	96.018
4106424	3-(4-Aminophenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	273.288	3.000	3.500	0.061	45.508
4171778	3-(3-Fluorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	242.249	0.000	1.750	2.493	100.000
4473888	1-(5-Chloro-2-hydroxyphenyl)-3-(2-chlorophenyl)-2-propen-1-one	293.149	0.000	0.750	2.973	100.000
4483657	3-(4-Hydroxyphenyl)-1-(4-methylphenyl)prop-2-en-1-one	238.285	1.000	2.750	1.756	95.745
4525387	4'-Amino-4-hydroxychalcone	239.273	2.000	2.750	0.604	68.588
4538437	3-(2,6-Dichlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	293.149	1.000	3.000	2.000	100.000
4811014	2-Methyl-4'-hydroxychalcone	238.285	0.000	1.750	2.039	100.000
4811016	3-Methyl-4'-hydroxychalcone	238.285	0.000	1.750	2.316	100.000
4811018	3-(2,5-Dimethylphenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	252.312	0.000	1.750	2.986	100.000
5057077	2,4,4'-Trihydroxychalcone	256.257	1.000	2.250	1.597	79.837
5179897	CID 5179897	300.310	3.000	4.500	0.819	72.774
5217675	CID 5217675	256.257	0.000	1.250	2.058	88.284
5280960	Naringenin chalcone	272.257	2.000	3.000	0.791	74.194
5281222	Butein	272.257	2.000	3.000	0.935	71.675
5281294	Okanin	288.256	2.000	2.750	0.916	66.861
5282361	4-Hydroxychalcone	224.259	1.000	2.750	1.557	94.581
5282362	4'-Hydroxychalcone	224.259	0.000	1.750	2.040	100.000
5316793	Pinostrobin chalcone	270.284	2.000	3.500	1.334	94.062
5318608	Isonobavachalcone	298.295	2.000	5.500	0.709	78.974
5319688	2'-O-Methylisoliquiritigenin	270.284	2.000	3.500	1.293	86.190
5322052	(E)-2,4,4'-Trihydroxychalcone	256.257	1.000	2.250	1.579	79.733

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

5337942	(2E)-3-(4-fluorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	242.249	0.000	1.750	2.423	100.000
5353470	Calythropsin	286.284	3.000	4.250	0.803	79.244
5355846	1-(2-Hydroxyphenyl)-3-(thiophen-2-yl)prop-2-en-1-one	230.281	0.000	1.250	2.241	100.000
5356279	(2E)-1-(5-chloro-2-hydroxyphenyl)-3-(2,4-dichlorophenyl)prop-2-en-1-one	327.594	0.000	0.750	3.433	100.000
5356337	4,5'-Dichloro-2'-hydroxychalcone	293.149	0.000	1.750	3.084	100.000
5356341	(E)-1-(4-chlorophenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	258.704	1.000	2.750	1.955	96.911
5357166	(2E)-1-(5-chloro-2-hydroxyphenyl)-3-(thiophen-2-yl)prop-2-en-1-one	264.726	0.000	1.250	2.729	100.000
5357218	1-(2-Hydroxyphenyl)-3-(4-hydroxyphenyl)-2-propen-1-one	240.258	1.000	2.500	1.569	90.122
5357587	(E)-1-(3-chloro-5-fluoro-2-hydroxyphenyl)-3-(4-hydroxy-3-methoxyphenyl)prop-2-en-1-one	322.720	1.500	4.250	2.033	100.000
5372946	3-(4-Chlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	258.704	0.000	1.750	2.594	100.000
5374076	Flemichapparín	270.284	1.000	2.500	1.776	91.896
5374122	2,2',4'-Trihydroxy-6'-methoxychalcone	286.284	2.000	3.250	1.005	79.381
5376916	2'-Hydroxy-4-methylchalcone	238.285	0.000	1.750	2.392	100.000
5376979	2',4'-Dihydroxychalcone	240.258	0.000	1.500	1.989	92.585
5377024	1-(4-Fluorophenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	242.249	1.000	2.750	1.789	95.942
5377516	2'-Hydroxy-5'-chlorochalcone	258.704	0.000	1.750	2.656	100.000
5387866	5'-Chloro-4,2'-dihydroxychalcone	274.703	1.000	2.500	2.042	92.892
5461154	2',3,4,4',6'-Pentahydroxychalcone	288.256	3.000	3.750	0.262	64.577
5467477	4,4'-Dihydroxychalcone	240.258	1.000	2.500	1.466	85.704
5708670	2'-Hydroxy-5'-methylchalcone	238.285	0.000	1.750	2.654	100.000
5708870	1-(2-Hydroxy-4-methylphenyl)-3-phenylprop-2-en-1-one	238.285	0.000	1.750	2.299	100.000
5708898	(E)-3-(2,4-dichlorophenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	293.149	0.000	1.750	3.045	100.000
5708964	5'-Fluoro-2'-hydroxychalcone	242.249	0.000	1.750	2.493	100.000
5709317	1-(2-Hydroxy-6-methoxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	270.284	2.000	3.500	1.047	85.904
5709318	2',3,4-Trihydroxychalcone	256.257	3.000	4.250	0.392	78.208
5711223	2',4'-Dihydroxy-4-methoxychalcone	270.284	1.000	2.500	1.853	91.785
5712115	1-(2-Hydroxyphenyl)-3-(2-methylphenyl)prop-2-en-1-one	238.285	0.000	1.750	2.214	100.000
5712162	3-(4-Chlorophenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	258.704	0.000	1.750	2.449	100.000
5728928	(2E)-3-(4-fluorophenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	242.249	0.000	1.750	2.260	100.000
5729224	(E)-3-(2,4-dichlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	293.149	1.000	2.750	2.567	100.000

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

5729327	(2E)-1-(4-hydroxyphenyl)-3-(4-methylphenyl)prop-2-en-1-one	238.285	0.000	1.750	2.231	100.000
5736169	(E)-3-(4-chlorophenyl)-1-(5-fluoro-2-hydroxyphenyl)prop-2-en-1-one	276.694	0.000	1.750	2.921	100.000
5736181	(2E)-1-(5-fluoro-2-hydroxyphenyl)-3-(4-fluorophenyl)prop-2-en-1-one	260.240	0.000	1.750	2.731	100.000
5736187	(E)-1-(5-fluoro-2-hydroxyphenyl)-3-(thiophen-2-yl)prop-2-en-1-one	248.271	0.000	1.250	2.497	100.000
5736188	(E)-1-(5-chloro-2-hydroxyphenyl)-3-(4-fluorophenyl)prop-2-en-1-one	276.694	0.000	1.750	2.913	100.000
5736195	(2E)-3-(4-chlorophenyl)-1-(2-hydroxy-5-methylphenyl)prop-2-en-1-one	272.730	0.000	1.750	3.082	100.000
5736261	(2E)-1-(2-hydroxy-5-methylphenyl)-3-(4-methylphenyl)prop-2-en-1-one	252.312	0.000	1.750	2.975	100.000
5743475	2'-Hydroxy-6'-methoxychalcone	254.285	1.000	2.750	1.712	100.000
5761068	3-(4-Hydroxyphenyl)-1-(thiophen-2-yl)prop-2-en-1-one	230.281	1.000	2.250	1.662	95.025
5797148	(2E)-3-(2-fluorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	242.249	0.000	1.750	2.672	100.000
5811533	2,2',4'-Trihydroxychalcone	256.257	1.000	2.250	1.458	86.445
5815007	2',4'-Dihydroxy-4-(dimethylamino)chalcone	283.326	4.000	2.000	1.687	77.853
5908055	(2E)-1-(5-chloro-2-hydroxyphenyl)-3-(4-methylphenyl)prop-2-en-1-one	272.730	0.000	1.750	2.882	100.000
5953849	2',4'-Dihydroxy-3,4-dimethoxychalcone	300.310	2.333	3.833	1.484	75.309
5960113	4,5'-Dichloro-2'-hydroxy-4'-methylchalcone	307.176	0.000	0.750	3.652	100.000
5973690	(E)-3-(4-hydroxy-3-methoxyphenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	270.284	2.000	3.500	1.258	90.361
5980988	4,2',5'-Trihydroxychalcone	256.257	1.000	2.250	1.524	81.648
5985833	5'-Chloro-2'-hydroxy-4'-methylchalcone	272.730	0.000	0.750	3.262	100.000
5989122	4,2'-Dihydroxy-3-methoxy-5'-methylchalcone	284.311	2.000	3.500	1.548	92.063
6027568	3-Fluoro-2'-hydroxychalcone	242.249	0.000	1.750	2.493	100.000
6092405	5'-Chloro-4,2'-dihydroxy-3-methoxychalcone	304.729	2.000	3.500	1.676	92.810
6123887	4-Hydroxy-4'-methylchalcone	238.285	1.000	2.750	1.756	95.745
6140247	1-(4-Aminophenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	239.273	2.000	2.750	0.604	68.588
6143779	2,6-Dichloro-2'-hydroxychalcone	293.149	1.000	3.000	2.000	100.000
6145740	3'-Chloro-2'-hydroxy-4'-methylchalcone	272.730	0.000	1.750	2.829	100.000
6213003	(E)-1-(3,4-dimethylphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	252.312	1.000	2.750	2.321	100.000
6213964	4'-Hydroxy-3-methylchalcone	238.285	0.000	1.750	2.235	100.000

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

6252826	3'-Chloro-2'-hydroxychalcone	258.704	0.000	1.750	2.603	100.000
6253344	Helichrysetin	286.284	2.000	3.250	0.923	76.835
6270253	(2E)-3-(2,4-dichlorophenyl)-1-(2-hydroxy-5-methylphenyl)prop-2-en-1-one	307.176	0.000	0.750	3.473	100.000
6293081	2',4'-Dihydroxy-4',6'-dimethoxychalcone	300.310	3.000	4.500	0.804	72.687
6304520	2',3',4'-Trihydroxychalcone	256.257	0.000	1.250	2.040	88.176
6435370	5-(3-(5-Bromo-2,4-dihydroxyphenyl)-3-oxo-1-propenyl)-2-hydroxybenzoic acid	379.163	1.000	5.250	0.779	69.442
6438092	Homobutein	286.284	1.500	4.000	1.170	70.753
6438580	2',6'-Dihydroxy-4,4'-dimethoxychalcone	300.310	3.000	4.500	1.108	79.783
6439186	Kukulkanin A	300.310	2.500	5.250	1.022	78.793
6439187	Kukulkanin B	286.284	2.500	5.000	0.599	67.970
6441981	2-Propen-1-one, 1-(4-ethoxy-2-hydroxy-6-methoxyphenyl)-3-(4-hydroxyphenyl)-	314.337	2.000	4.500	1.180	87.845
6442675	Echinatin	270.284	2.000	3.500	1.419	85.532
6447568	Acrylophenone, 5'-bromo-2',4'-dihydroxy-3-(p-methoxyphenyl)-	349.180	1.000	3.750	1.928	94.542
6447571	Acrylophenone, 5'-bromo-2',4'-dihydroxy-3-(3,4-dihydroxyphenyl)-	351.153	2.000	4.250	0.951	73.953
6447575	2',4'-Dihydroxy-5'-bromochalcone	319.154	0.000	2.750	1.991	94.909
6447958	Acrylophenone, 5'-acetyl-2',4'-dihydroxy-3-phenyl-	282.295	0.000	3.500	1.617	88.297
6447960	1-(5-Acetyl-2,4-dihydroxyphenyl)-3-(4-methoxyphenyl)-2-propen-1-one	312.321	1.000	4.500	1.508	87.657
6447962	Acrylophenone, 5'-acetyl-2',4'-dihydroxy-3-(p-hydroxy-m-methoxyphenyl)-	328.321	1.000	4.250	1.538	80.052
6447964	Acrylophenone, 5'-acetyl-2',4'-dihydroxy-3-(p-(dimethylamino)phenyl)-	325.363	4.000	4.000	1.572	75.072
6447966	Acrylophenone, 5'-bromo-2',4'-dihydroxy-3-(p-hydroxy-m-methoxyphenyl)-	365.180	1.000	3.500	1.898	86.582
6447967	Acrylophenone, 5'-bromo-2',4'-dihydroxy-3-(p-(dimethylamino)phenyl)-	362.222	4.000	3.250	1.863	81.199
6474196	2,2',4'-Trihydroxy-5'-methylchalcone	270.284	2.000	3.250	1.350	83.705
6474292	2,2',4'-Trihydroxychalcone	256.257	2.000	3.250	1.073	82.083
6474295	Pinocembrin chalcone	256.257	1.000	2.250	1.459	86.450
6474679	2'-Hydroxy-4'-chlorochalcone	258.704	0.000	1.750	2.575	100.000
6474895	4',3,4'-Trihydroxychalcone	256.257	2.000	3.250	0.833	75.474
6478367	2,4,6-trihydroxy-3-[(Z)-3-hydroxy-3-phenyl-prop-2-enoyl]-5-methylbenzaldehyde	314.294	0.000	3.000	2.038	88.262

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

6537040	4,2'-Dihydroxy-4'-methoxychalcone	270.284	2.000	3.500	1.309	88.601
6610901	2-(4-Iodophenoxy)-1-(2,4,6-trihydroxyphenyl)ethanone	386.142	2.000	2.500	1.623	87.407
7152435	2',4'-Dihydroxy-3',6'-dimethoxychalcone	300.310	2.500	5.250	0.810	77.548
7514812	3-Chloro-4'-hydroxychalcone	258.704	0.000	1.750	2.530	100.000
7514813	(E)-3-(3-fluorophenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	242.249	0.000	1.750	2.348	100.000
7515046	4'-Fluoro-4-hydroxy-3,5-dimethylchalcone	270.303	1.000	2.750	3.078	100.000
7946917	(E)-1-(5-chlorothiophen-2-yl)-3-(4-hydroxyphenyl)prop-2-en-1-one	264.726	1.000	3.250	1.675	95.095
8826820	(E)-1-(2,4-difluorophenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	260.240	1.000	2.750	1.993	100.000
8832860	3-Ethoxy-4,4'-dihydroxychalcone	284.311	0.000	2.500	2.019	89.770
8855762	(1z,2e)-1-(2-Fluorophenyl)-3-(4-hydroxyphenyl)-2-propen-1-one	242.249	1.000	2.750	1.530	94.426
8856720	3',4'-Dichloro-4-hydroxychalcone	293.149	1.000	2.750	2.288	100.000
9207702	1-(2,4-Dihydroxyphenyl)-2-(4-methoxyphenyl)sulfanylethanone	290.333	1.000	2.000	2.107	93.306
9448548	(2E)-1-(4-Hydroxy-3-methoxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	270.284	2.000	3.500	1.480	86.672
9838073	1,3-Bis(2,4-dihydroxyphenyl)propane	260.289	4.000	3.250	0.680	76.436
9860386	2',4'-Dihydroxy-2,4-difluoro-chalcone	276.239	0.000	1.500	2.274	94.250
9904479	2',4',6',3-Tetrahydroxy-4-methoxychalcone	302.283	2.000	3.000	1.154	63.927
9971453	(E)-1-(2,4-dihydroxyphenyl)-3-naphthalen-2-ylprop-2-en-1-one	290.318	1.000	2.500	2.276	95.761
9972132	2,3,4,6-Tetrahydroxy-5-(3-phenylpropanoyl)benzaldehyde	302.283	1.000	4.250	0.956	77.841
9993268	Larrein	270.284	1.500	4.250	1.222	92.921
10022754	4-Iodo-2',4',6'-trihydroxychalcone	382.154	3.000	2.250	1.669	87.682
10038861	1,3-Bis(3,4-dihydroxyphenyl)prop-2-en-1-one	272.257	2.000	3.000	0.890	69.417
10062515	(E)-1-(2,3-Dihydroxy-4-methoxyphenyl)-3-(4-hydroxyphenyl)-2-propene-1-one	286.284	2.000	3.250	1.214	82.087
10062573	4-Chloro 2'-hydroxy-4',6'-dimethylchalcone	286.757	0.000	1.750	2.878	100.000
10084706	1-Propanone, 3-phenyl-1-(2,3,4,6-tetrahydroxyphenyl)-	274.273	1.000	2.250	1.559	82.947
10107266	2,4,2',4'-Tetrahydroxychalcone	272.257	2.000	3.000	0.814	74.328
10108086	2-Chloro-2'-hydroxy-4',6'-dimethylchalcone	286.757	0.000	1.750	3.184	100.000
10221174	1-Propanone, 1-(2,4-dihydroxyphenyl)-3-hydroxy-3-(4-hydroxyphenyl)-	274.273	1.000	2.250	1.393	78.206
10266895	4,2'-Dihydroxy-4',6'-dimethoxydihydrochalcone	302.326	4.000	4.500	0.774	72.482

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

10356276	N-(3-chlorophenyl)-2,4,6-trihydroxybenzamide	279.679	2.000	2.250	1.392	86.208
10377524	2'-Hydroxy-4',6'-dimethylchalcone	252.312	0.000	1.750	2.450	100.000
10445922	3'-Methyl-2',4',6'-trihydroxydihydrochalcone	272.300	1.000	2.500	2.475	95.826
10491149	2',4'-Dihydroxy-4-methylchalcone	254.285	0.000	1.500	2.293	94.363
10495842	4-Bromo-2',4'-dihydroxychalcone	319.154	0.000	1.500	2.506	95.608
10515090	4,2'-Dihydroxy-5'-methylchalcone	254.285	1.000	2.500	2.006	92.685
10540319	2',4'-Dihydroxy-4-chlorochalcone	274.703	0.000	1.500	2.496	95.552
10565177	2,4-Dimethyl-2'-hydroxy-5'-chloro-trans-chalcone	286.757	0.000	1.750	2.886	100.000
10586364	Benzyl 2,6-dihydroxybenzoate	244.246	0.000	2.500	1.653	95.266
10591624	beta-(4-Hydroxyphenyl)-2',6'-dihydroxy-3',4'-dimethoxypropiophenone	318.326	4.333	5.833	0.464	67.465
10612087	2',4'-Dihydroxy-3-methoxychalcone	270.284	1.000	2.500	1.919	92.173
10614151	2',4'-Dihydroxy-2,4-dimethoxychalcone	300.310	2.000	3.500	1.421	75.138
10618629	2'-Hydroxy-2,4-dichloro-3',5'-dichlorochalcone	362.039	0.000	0.750	3.818	100.000
10657726	1-(2,4-Dihydroxyphenyl)-3-(4-pyridyl)-2-propene-1-one	241.246	0.000	2.500	1.338	84.943
10657779	1-(2,6-Dihydroxyphenyl)-3-phenylpropan-1-one	242.274	1.000	2.750	1.691	96.763
10661294	(R)-1-(3,4-Dihydroxyphenyl)-3-(2,4,6-trihydroxyphenyl)-2-propanol	292.288	3.000	3.000	0.786	57.521
10683323	2,4-Dimethyl-2',4'-dihydroxy-trans-chalcone	268.312	0.000	1.500	2.328	94.571
10732214	3,4,2',4',alpha-Pentahydroxydihydrochalcone	290.272	2.000	3.000	1.088	59.001
10752529	2',6'-Dihydroxychalcone	240.258	1.000	2.500	1.519	95.144
10752664	2-trans-Cinnamyl-1,3,5-benzenetriol	242.274	3.000	2.250	1.264	87.553
10755967	4-Chloro-2',4',6'-trihydroxychalcone	290.702	1.000	2.250	1.860	88.796
10778685	2',4',6'-Trihydroxy-4-fluorochalcone	274.248	1.000	2.250	1.710	87.922
10779899	2,2',4'-Trihydroxy-5'-chloro-trans-chalcone	290.702	2.000	3.250	1.468	84.393
10779900	2',3,4'-Trihydroxy-5'-chloro-trans-chalcone	290.702	3.000	4.250	0.775	80.447
10801179	2,4,4'-Trihydroxydihydrochalcone	258.273	1.000	2.250	1.622	79.981
10802100	2',3,4,5'-Tetrahydroxychalcone	272.257	2.000	3.000	1.029	72.619
10804786	3,4-Dichloro-2',4'-dihydroxychalcone	309.148	0.000	1.500	2.917	100.000
10806690	4-Bromo-2',4',6'-trihydroxychalcone	335.153	1.000	2.250	1.905	89.059
10824319	1-(2,4-Dihydroxyphenyl)-3-(thiophen-2-yl)propanone (9)	246.280	0.000	1.000	2.143	93.486
10826041	1-(2,6-Dihydroxyphenyl)-3-(4-methoxyphenyl)propan-1-one	272.300	2.000	3.750	1.442	95.306

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

10826895	2-Methoxy-4,4'-dihydroxy-6-methylchalcone	284.311	2.000	3.500	1.599	86.581
10827074	2',4'-Dihydroxy-4-(methylthio)chalcone	286.345	1.600	2.000	2.088	93.163
10850872	3-Fluoro-2',4'-dihydroxy-4-methoxychalcone	288.275	1.000	2.500	2.206	93.852
10853051	5'-Bromo-4,2'-dihydroxychalcone	319.154	1.000	2.500	2.067	93.042
10900865	4-[3-(4-Hydroxyphenyl)propyl]resorcinol	244.290	3.000	2.250	1.384	82.946
10977690	1-(2-Hydroxyphenyl)-3-(3-methyl-2-thienyl)-2-propene-1-one	244.308	0.000	1.250	2.486	100.000
11022479	Neosakuranetin	286.284	3.000	4.250	0.695	81.979
11022960	4,4'-Dihydroxy-2',6'-dimethoxychalcone	300.310	3.000	4.500	0.921	69.708
11035487	(E)-1-(4-hydroxyphenyl)-3-(4-iodophenyl)prop-2-en-1-one	350.155	2.000	1.750	2.366	100.000
11161591	(E)-3-phenyl-1-(2,3,4,6-tetrahydroxyphenyl)prop-2-en-1-one	272.257	1.000	2.000	1.470	81.811
11414940	2',4',6'-Trihydroxy-3'-methoxy-trans-chalcone	286.284	2.500	5.000	0.700	73.875
11415131	(E)-3-(3,4-dichlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	293.149	1.000	2.750	2.301	100.000
11449431	2',4'-Dihydroxy-4,6'-dimethoxychalcone	300.310	2.000	3.500	1.319	74.538
11565326	1-(2,4,6-Trihydroxyphenyl)-3-(4'-aminophenyl)-1-propanol	275.304	4.000	2.500	0.142	45.978
11572837	Benzamide, 2,4-dihydroxy-N-[(4-hydroxy-3-methoxyphenyl)methyl]-	289.287	2.500	3.000	1.442	72.050
11594429	(E)-1-(4-chlorophenyl)-3-(4-hydroxy-3,5-dimethylphenyl)prop-2-en-1-one	286.757	1.000	2.750	3.246	100.000
11695370	(E)-3-(3,4-dimethoxyphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	316.310	3.000	4.250	1.007	70.847
11778601	Stercurensin	284.311	1.000	2.500	2.429	100.000
11778945	3-(3,4-Dihydroxyphenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	290.272	3.000	4.000	0.284	52.357
11817084	(E)-1-(2-Hydroxyphenyl)-3-(3-thienyl)-2-propene-1-one	230.281	1.000	1.250	2.246	100.000
11973879	2',4'-Dihydroxy-5',6'-dimethoxychalcone	300.310	3.833	6.583	0.543	77.041
12044949	(E)-3-(4-ethoxyphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	300.310	1.000	3.250	1.628	87.441
12110863	Nubigenol	290.272	2.000	4.250	0.510	60.147
12133006	2',3',4',5',6'-Pentahydroxychalcone	288.256	1.000	1.750	1.447	76.971
12136838	(E)-2-deuterio-1-(2-hydroxyphenyl)-3-phenylprop-2-en-1-one	224.259	0.000	1.750	2.166	100.000
12147207	(E)-3-(2,4-dichlorophenyl)-1-(2,4-dihydroxyphenyl)prop-2-en-1-one	309.148	0.000	0.500	2.834	100.000
12223617	Benzoic acid, 2,4-dihydroxy-6-methyl-, phenylmethyl ester	258.273	1.000	2.500	1.755	81.667

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

12227424	2-(Phenylthio)-1-(2,4,6-trihydroxyphenyl)ethanone	276.306	1.000	2.000	1.528	87.529
12313213	Neoplathymenin	288.256	2.000	2.750	0.871	64.776
12318347	(E)-1-(2,4-dihydroxyphenyl)-3-naphthalen-1-ylprop-2-en-1-one	290.318	1.000	2.500	2.404	96.900
12365230	1-(2,6-Dihydroxyphenyl)-3-(4-hydroxy-3-methoxyphenyl)propan-1-one	288.299	2.500	5.250	0.821	74.573
12365231	1-(2,6-Dihydroxyphenyl)-3-(4-hydroxyphenyl)propan-1-one	258.273	2.000	3.500	1.029	84.545
12430810	2'-Hydroxy-beta-phenylpropiolophenone	222.243	0.000	1.750	1.907	100.000
12519991	2-[3-(4-Hydroxyphenyl)propyl]phenol	228.290	2.000	1.500	2.100	95.478
12541852	2',4'-Dihydroxy-6'-hydroxy-alpha-phenoxyacetophenone	260.246	1.000	3.500	0.850	82.880
12883590	2',4,6'-Trihydroxychalcone	256.257	2.000	3.250	0.925	83.326
12954609	(2E)-1-(4-aminophenyl)-3-(4-chlorophenyl)prop-2-en-1-one	257.719	1.000	2.000	1.713	83.424
13487165	1-(2-Hydroxy-5-methylphenyl)-3-phenylpropan-1-one	240.301	0.000	1.750	2.960	100.000
13643061	(E)-2',6'-Dihydroxy-beta-(4-methoxyphenyl)acrylophenone	270.284	2.000	3.500	1.358	94.204
13939318	Loureirin D	288.299	3.000	3.250	0.874	62.642
13990811	2',4',6'-Trihydroxy-4-methoxychalcone	286.284	2.000	3.250	1.270	85.342
14430410	Myrigalone P	286.327	1.000	2.500	3.145	100.000
14440596	2,4,6-Trihydroxychalcone	256.257	2.000	3.250	1.144	82.497
14440598	(E)-2,4,4',6'-Tetrahydroxychalcone	272.257	2.000	3.000	0.983	73.214
14452794	(E)-2',6'-Dihydroxy-beta-(4-chlorophenyl)acrylophenone	274.703	1.000	2.500	1.987	100.000
14592782	2-(3-(4-Hydroxyphenyl)propyl)benzene-1,3,5-triol	260.289	3.000	2.250	1.208	79.528
14632193	4,2',4',alpha-Tetrahydroxydihydrochalcone	274.273	2.000	3.250	1.039	77.756
14756267	Benzyl 2,4,6-trihydroxybenzoate	260.246	0.000	2.250	1.568	86.360
14756268	Benzyl 3-chloro-2,4,6-trihydroxybenzoate	294.691	0.000	2.250	1.996	91.409
14793017	(e)-1-(2-Hydroxy-4,5-dimethylphenyl)-3-p-tolylprop-2-en-1-one	266.339	0.000	1.750	3.291	100.000
14977371	3,4,4'-Trihydroxy-3'-methoxychalcone	286.284	3.000	4.250	0.982	77.361
14977377	2',3,4'-Trihydroxy-5'-methylchalcone	270.284	3.000	4.250	0.615	79.510
15290682	2,2',4'-Trihydroxy-4-methoxychalcone	286.284	2.000	3.250	1.276	85.378
15385488	2',4',4'-Trihydroxy-3',3'-methoxychalcone	316.310	3.000	6.750	0.451	71.369
15680254	2',3,4'-Trihydroxychalcone	256.257	1.000	2.250	1.507	81.545
15680255	2,3,4,2',4'-Pentahydroxychalcone	288.256	3.000	3.750	0.415	65.180

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

16086812	N-(4-hydroxybenzyl)-2,4-dihydroxybenzamide	259.261	2.500	1.750	1.393	80.454
19006761	1,3-Bis(4-hydroxyphenyl)propan-1-one	242.274	1.000	2.500	1.590	86.404
19540635	beta-(2-Pyridyl)-2',4'-dihydroxyacrylophenone	241.246	0.000	2.500	1.379	86.163
19540636	(E)-1-(2,4-dihydroxyphenyl)-3-[4-(trifluoromethyl)phenyl]prop-2-en-1-one	308.256	0.000	1.500	2.736	96.957
19540655	4-Fluoro-2',4'-dihydroxychalcone	258.248	0.000	1.500	2.307	94.443
19540750	1-(2,4-Dihydroxyphenyl)-3-(5-methylfuran-2-yl)-2-propene-1-one	244.246	0.000	3.500	1.114	87.507
19910385	4-[(E)-3-phenylprop-2-enyl]benzene-1,3-diol	226.274	2.000	1.500	1.918	94.414
19910460	2-[(E)-3-phenylprop-2-enyl]benzene-1,3-diol	226.274	2.000	1.500	1.966	100.000
20560968	4-[3-Oxo-3-(2,4,6-trihydroxyphenyl)propyl]phenyl dihydrogen phosphate	354.252	2.000	20.500	-4.965	21.985
20980077	3-(Cyclohexa-2,4-dien-1-YL)-1-(2,4,6-trihydroxyphenyl)prop-2-EN-1-one	258.273	1.000	2.250	1.546	86.957
21117850	4,2',4',6'-Tetrahydroxy-3-methoxychalcone	302.283	2.000	3.000	1.154	63.925
21712961	(E)-1,3-bis(4-aminophenyl)prop-2-en-1-one	238.288	3.000	3.000	-0.189	51.872
21722044	3'-Methylflavokawin	314.337	3.000	4.500	1.719	79.449
21722167	4,2',4'-Trihydroxy-3-methoxydihydrochalcone	288.299	1.500	4.000	1.357	71.818
23651847	Bis-(4-hydroxybenzyl)sulfide	246.323	3.600	2.000	1.412	90.045
24970524	Helichryson B	288.299	2.500	5.250	0.792	75.028
25200812	CID 25200812	288.256	3.000	3.750	0.368	66.020
25219913	(2E)-1-(4-aminophenyl)-3-(4-hydroxy-2-methoxyphenyl)prop-2-en-1-one	269.299	3.000	3.750	0.587	68.588
25243927	(Z)-1-(2,4-dihydroxyphenyl)-3-hydroxy-3-(4-hydroxyphenyl)prop-2-en-1-one	272.257	1.000	2.000	1.439	77.927
27111968	(E)-3-(4-hydroxy-3,5-dimethylphenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	268.312	1.000	2.500	2.720	95.860
42607560	4,2',3',4'-Tetrahydroxychalcone	272.257	1.000	2.000	1.364	75.876
42607593	Triangularin	284.311	2.000	3.500	2.189	100.000
42607636	3,4,2',4',alpha-Pentahydroxychalcone	288.256	3.000	3.750	0.378	67.404
42607705	4,2',4'-Trihydroxy-6'-methoxydihydrochalcone	288.299	3.000	3.250	0.845	64.555
42607706	4,2',6'-Trihydroxy-4'-methoxy-3'-methyldihydrochalcone	302.326	3.000	4.500	1.614	76.423
44404713	(E)-3-[4-(dimethylamino)phenyl]-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	299.326	5.000	2.750	1.013	70.876

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

44410790	N-(3,4,5-trihydroxybenzyl)-2,4-dihydroxybenzamide	291.260	4.500	3.250	0.266	47.980
44428608	2',3,4,4',5,6'-Hexahydroxychalcone	304.256	4.000	4.500	-0.217	43.113
44428616	2,2',4,4'-Tetrahydroxy-6'-methoxychalcone	302.283	3.000	4.000	0.616	55.442
44557158	Anemarchalconyn	238.242	1.000	2.500	1.139	82.550
44577024	3-Hydroxyhelichrysetin	302.283	3.000	4.000	0.343	53.958
44589166	(E)-3-(3-bromophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	335.153	1.000	2.250	1.982	90.699
44589167	(E)-3-(1,3-benzodioxol-5-yl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	300.267	1.000	4.250	0.996	83.741
44611968	4-(3-(2,4-Dihydroxyphenyl)propyl)-2-methylbenzene-1,3-diol	274.316	4.000	3.250	1.447	83.738
46224961	(E)-1-(2,4-dihydroxy-6-methoxyphenyl)-3-(4-methylphenyl)prop-2-en-1-one	284.311	1.000	2.500	1.832	90.504
46224987	3-Methoxyhelichrysetin	316.310	2.000	3.250	1.350	66.942
46225019	(E)-3-(4-chlorophenyl)-1-(2,4-dihydroxy-6-methoxyphenyl)prop-2-en-1-one	304.729	1.000	2.500	1.978	91.355
46225020	(E)-1-(2,4-dihydroxy-6-methoxyphenyl)-3-(4-fluorophenyl)prop-2-en-1-one	288.275	1.000	2.500	1.811	90.382
49859660	1-(3,4-Dihydroxyphenyl)-3-(2,4,6-trihydroxyphenyl)-1,3-propanedione	304.256	1.000	4.000	0.346	50.206
52920300	(E)-3-phenyl-1-(2,4,6-trihydroxy-3,5-dimethylphenyl)prop-2-en-1-one	284.311	1.000	2.250	2.956	100.000
53321909	4-Hydroxybenzyl 2,4,6-Dihydroxybenzoate	276.245	0.000	2.000	1.457	77.433
53323255	4-Hydroxybenzyl 2,4-Dihydroxybenzoate	260.246	0.000	2.250	1.580	81.813
53405565	1-(2,4-Dihydroxyphenyl)-3-(3-methoxyphenyl)prop-2-en-1-one	270.284	1.000	2.500	1.919	92.173
53422685	1-(2,4-Dihydroxyphenyl)-3-(4-methoxycyclohexa-2,4-dien-1-yl)prop-2-en-1-one	272.300	1.000	2.500	1.991	92.597
53549908	trans 1-(3-Thienyl) 3-(4-hydroxyphenyl) 2-propene 1-one	230.281	2.600	2.000	1.388	92.637
53759355	3-(4-Methoxyphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	286.284	2.000	3.250	1.253	85.244
53951317	1-Phenyl-3-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	256.257	2.000	3.250	1.127	82.399
54037829	4,4'-Diaminochalcone	238.288	3.000	3.000	-0.222	51.680
54281601	1-(2,6-Dihydroxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	256.257	2.000	3.250	0.925	83.326
54350948	3-(5-Hydroxy-4-methoxy-2-methylphenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	318.326	2.000	3.250	1.288	65.811
54490564	2,4,4',6-Tetrahydroxy-chalcone	272.257	2.000	3.000	1.000	73.313

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

54570259	2',4,4',6-Tetrahydroxychalcone	272.257	2.000	3.000	0.814	74.328
56647587	1-(5-Fluoro-2-hydroxyphenyl)-3-phenylpropan-1-one	244.265	0.000	1.750	2.721	100.000
57048077	1-Phenyl-3-(2,4,6-trihydroxyphenyl)propan-1-one	258.273	2.000	3.500	1.041	82.540
57391970	(E)-3-(4-chlorophenyl)-1-(2,4-dihydroxy-3-methylphenyl)prop-2-en-1-one	288.730	0.000	1.500	3.370	100.000
57395415	(E)-1-(2,4-dihydroxy-3-methylphenyl)-3-(4-methoxyphenyl)prop-2-en-1-one	284.311	1.000	2.500	2.657	100.000
57395416	(E)-1-(2,4-dihydroxy-3-methylphenyl)-3-(4-fluorophenyl)prop-2-en-1-one	272.275	0.000	1.500	3.199	100.000
57400130	(E)-3-(2-chlorophenyl)-1-(2,4-dihydroxyphenyl)prop-2-en-1-one	274.703	0.000	0.500	2.425	95.138
57402432	(E)-1-(2,4-dihydroxyphenyl)-3-[4-hydroxy-2-(methoxymethoxy)phenyl]prop-2-en-1-one	316.310	4.000	5.250	0.609	64.874
58178431	3-(4-Methylphenyl)propanoyl-phloroglucinol	272.300	1.000	2.500	1.743	88.725
60112525	Qucecetin chalcone	304.256	1.000	4.000	0.430	53.881
66691743	3-(3-Methylphenyl)propanoyl-phloroglucinol	272.300	1.000	2.500	1.747	88.750
66691967	3-(4-Fluorophenyl)propanoyl-phloroglucinol	276.264	1.000	2.500	1.793	89.020
66704345	1-(4-Chlorophenyl)-3-(4-hydroxy-3,5-dimethylphenyl)prop-2-en-1-one	286.757	1.000	2.750	3.264	100.000
66738878	3-(4-Hydroxy-3-methoxyphenyl)-1-(4-hydroxyphenyl)prop-2-en-1-one	270.284	1.000	2.500	1.747	87.914
67251446	4,2'-Dihydroxydihydrochalcone	242.274	1.000	2.500	1.715	90.976
70792582	1-(2,4,6-Trihydroxyphenyl)-2-hydroxy-3-(4-hydroxyphenyl)-2-propene-1-one	288.256	1.000	4.250	0.393	72.453
71036891	(Z)-2-amino-1-(2,4-dihydroxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	271.272	5.000	2.750	-0.173	56.567
71352575	2-((2S)-2-Hydroxy-3-(3,4-dihydroxyphenyl)propyl)phloroglucinol	292.288	3.000	3.000	0.786	57.521
71361874	1-(2,4-Dihydroxyphenyl)-3-(3-hydroxyphenyl)prop-2-en-1-one	256.257	1.000	2.250	1.507	81.545
71438154	2',4',5',3,4-Pentahydroxychalcone	288.256	2.000	2.750	0.871	64.776
71441129	3-Phenyl-1-(2,4,6-trihydroxy-3-methoxyphenyl)prop-2-en-1-one	286.284	2.500	5.000	0.685	73.785
71442179	1-(2-Hydroxy-5-methylphenyl)-3-(thiophen-2-yl)prop-2-en-1-one	244.308	0.000	1.250	2.806	100.000
71457360	2,2',4',6'-Tetrahydroxychalcone	272.257	2.000	3.000	1.168	78.108
71462729	(E)-3-(3,5-difluorophenyl)-1-(2,4-dihydroxyphenyl)prop-2-en-1-one	276.239	0.000	1.500	2.633	96.351

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

71625781	2',4'-Dihydroxy-3'-methyl-6'-methoxy-omega-(4-hydroxyphenyl)propiofenone	302.326	3.000	3.250	1.692	72.323
71728402	2,4,4',6-Tetrahydroxydibenzoylmethane	288.256	1.000	4.250	0.410	70.062
71740915	(E)-1-(3-chloro-6-hydroxy-2,4-dimethylphenyl)-3-(4-chlorophenyl)prop-2-en-1-one	321.202	1.000	-0.250	3.840	100.000
71740916	(E)-1-(3-chloro-6-hydroxy-2,4-dimethylphenyl)-3-(2-chlorophenyl)prop-2-en-1-one	321.202	1.000	0.000	3.033	100.000
71817312	(E)-1-(4-amino-2-hydroxyphenyl)-3-(4-ethylphenyl)prop-2-en-1-one	267.327	1.000	1.750	2.285	82.244
71819497	(E)-1-(4-amino-2-hydroxyphenyl)-3-(5-methylthiophen-2-yl)prop-2-en-1-one	259.322	1.000	2.250	1.168	75.707
71819498	(E)-1-(4-amino-2-hydroxyphenyl)-3-(4-methylsulfanylphenyl)prop-2-en-1-one	285.360	2.600	2.250	1.206	75.931
71819500	(E)-1-(4-amino-2-hydroxyphenyl)-3-(4-fluorophenyl)prop-2-en-1-one	257.264	1.000	1.750	1.452	77.368
71819501	(E)-1-(4-amino-2-hydroxyphenyl)-3-(4-chlorophenyl)prop-2-en-1-one	273.718	1.000	1.750	1.619	78.348
72079736	4,4'-Dihydroxy-3,5-dimethylchalcon	268.312	1.000	2.500	2.736	95.958
72945334	(E)-1-(5-chloro-2,4-dihydroxyphenyl)-3-(2,4-dihydroxyphenyl)prop-2-en-1-one	306.702	2.000	3.000	1.295	79.026
73064893	CID 73064893	316.310	3.333	4.583	0.874	68.594
73097563	2',4'-Dihydroxy-6'-methoxy-3'-methylchalcone	284.311	1.000	2.500	2.429	100.000
73193562	1-(2,4-Dihydroxy-3-methoxyphenyl)-3-(4-hydroxy-3-methoxyphenyl)prop-2-en-1-one	316.310	2.500	5.000	1.010	70.939
73310377	3-(2,4-Dihydroxyphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	288.256	3.000	3.750	0.225	69.652
73830004	1-(2-Hydroxy-4,6-dimethoxy-3-methylphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	314.337	3.000	4.500	1.719	79.449
74819202	2',3',4',4'-Tetrahydroxychalcone	272.257	1.000	2.000	1.382	75.980
75026045	3-(2,3-Dihydroxyphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	288.256	3.000	3.750	0.513	67.753
75026046	2',4',6',3,4,5-Hexahydroxychalcone	304.256	4.000	4.500	-0.217	43.113
75310994	2',3',4',6',4-Pentahydroxychalcone	288.256	2.000	2.750	0.754	69.276
75954323	3',5'-Dimethyl-2',4',6'-trihydroxychalcone	284.311	1.000	2.250	2.956	100.000
76315185	(E)-3-(2-bromophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	335.153	1.000	2.250	2.187	91.014
76315186	(E)-3-(4-methylphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	270.284	1.000	2.250	1.680	87.744
76318835	(E)-3-(3-chlorophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	290.702	1.000	2.250	1.954	90.421

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

76322491	(E)-3-(3,4-dimethylphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	284.311	1.000	2.250	2.303	91.391
76329687	2-Fluoro-2',4',6'-trihydroxychalcone	274.248	1.000	2.250	2.016	89.767
76329688	(E)-3-(2,4-dichlorophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	325.148	1.000	2.250	2.565	93.191
76333354	(E)-3-(3-fluorophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	274.248	1.000	2.250	1.782	88.807
76333355	(E)-3-(2-chlorophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	290.702	1.000	2.250	2.162	90.834
78385139	2-Hydroxy-5-[3-(4-hydroxyphenyl)prop-2-enoyl]-4-methoxybenzaldehyde	298.295	2.000	5.500	0.693	78.884
85126404	1-(2-Hydroxy-4,6-dimethylphenyl)-3-phenylprop-2-en-1-one	252.312	0.000	1.750	2.450	100.000
85177930	2',4'-Dihydroxy-4-chloro chalcone	274.703	0.000	1.500	2.496	95.552
85231517	1-(2,4-Dihydroxyphenyl)-3-(2,4-dimethylphenyl)prop-2-en-1-one	268.312	0.000	1.500	2.328	94.571
85257122	3-(4-Chlorophenyl)-1-(2,4,6-trihydroxyphenyl)-2-propen-1-one	290.702	1.000	2.250	1.860	88.796
85370109	1-(2,6-Dihydroxy-4-methoxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	286.284	3.000	4.250	0.695	81.979
85394271	alpha,2',4',6'-Tetrahydroxy-beta-(3,4-dihydroxyphenyl)propiophenone	306.271	2.000	4.000	0.425	53.129
85420425	3-Phenyl-1-(2,3,4,6-tetrahydroxyphenyl)prop-2-en-1-one	272.257	1.000	2.000	1.470	81.811
85840766	alpha,beta-Dihydro-2',3,4',6'-tetrahydroxychalcone	274.273	2.000	3.250	0.883	75.348
86142500	1-Phenyl-3-(2,4,6-trihydroxyphenyl)propane-1,3-dione	272.257	1.000	4.500	0.547	79.212
86289291	2-[3-(3,4-Dihydroxyphenyl)prop-2-enoyl]-3,5-dihydroxyphenolate	288.256	3.000	3.750	0.262	64.577
86289307	2',4,4',6'-Tetrahydroxychalcone(1-)	272.257	2.000	3.000	0.782	74.142
86289426	Isoliquiritigenin(1-)	256.257	1.000	2.250	1.419	80.904
89351474	2',4,4'-Trihydroxy-3,5-dimethyl-trans-chalcone	284.311	2.000	3.250	2.066	89.607
90788180	3'-Formyl-6',4-dihydroxy-2'-methoxy-5'-methylchalcone	312.321	2.000	5.500	1.432	87.173
90842264	1-(3,4,5-Trihydroxyphenyl)-3-(2,4,6-trihydroxyphenyl)-2-propanol	308.287	3.000	2.750	0.866	52.160
91527531	2',4,4',5'-Tetrahydroxychalcone	272.257	1.000	2.000	1.424	74.536
91586951	4,2',5'-Trihydroxy-4'-methoxychalcone	286.284	2.000	3.250	1.264	80.690
91872142	4',6'-Dimethylchalconaringenin	268.312	1.000	2.500	1.839	91.703

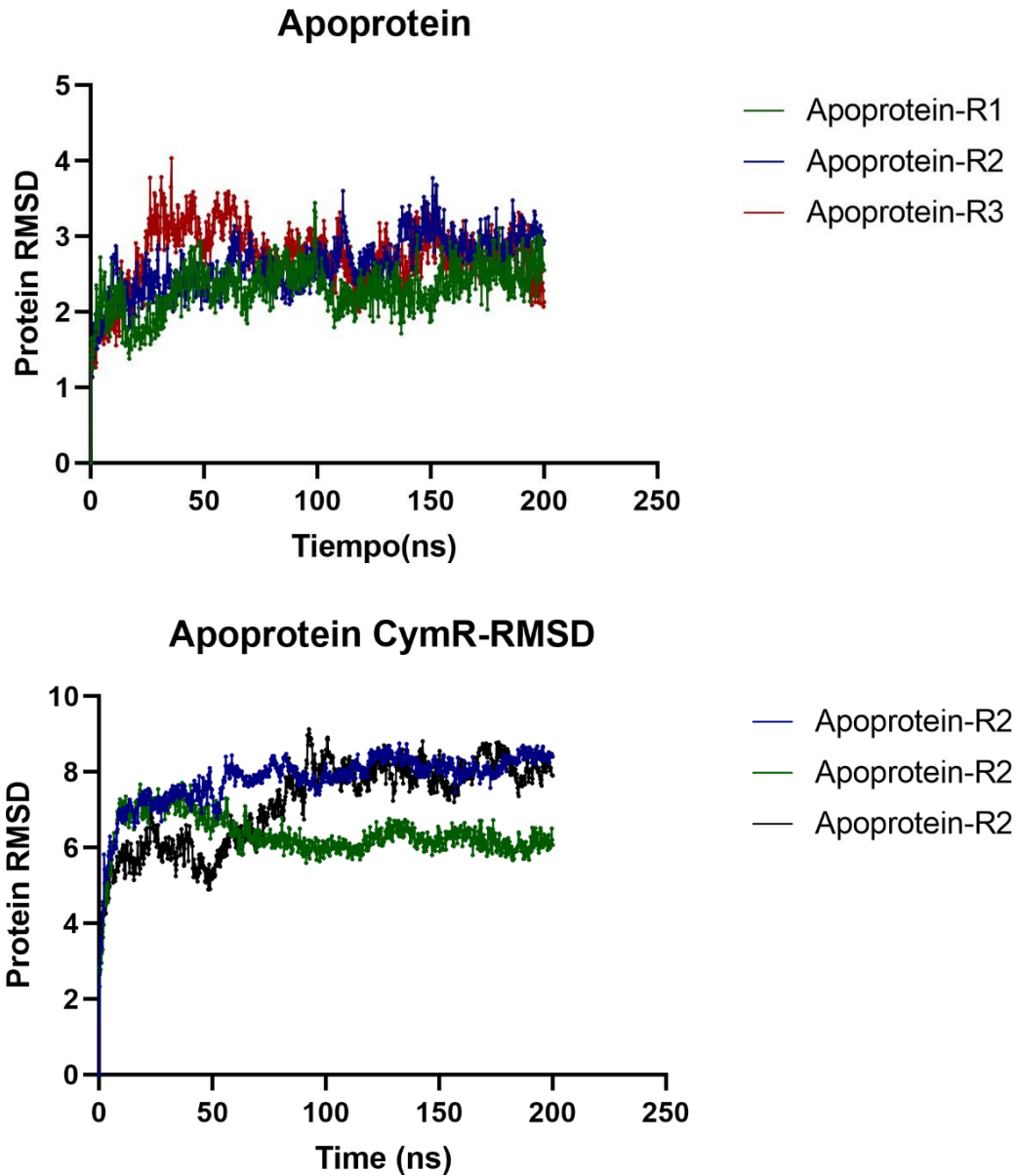
Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

91971254	4-[(E,3R)-3-(2,4-dihydroxyphenyl)-3-hydroxyprop-1-enyl]benzene-1,2-diol	274.273	3.000	2.000	0.954	58.825
10111085 6	1-(2,4-Dihydroxyphenyl)-3-(2-hydroxyphenyl)-1-propanone	258.273	1.000	2.500	1.586	87.805
10112747 4	5-[(S)-2-Hydroxy-3-(2,4,6-trihydroxyphenyl)propyl]pyrogallol	308.287	3.000	2.750	0.851	52.074
10114592 3	(E)-omega-(3-Furyl)-2',4',6'-trihydroxyacrylophenone	246.219	1.000	4.250	0.471	80.604
10114592 4	(E)-omega-(3-Thienyl)-2',4',6'-trihydroxyacrylophenone	262.280	2.000	1.750	1.671	87.628
10133852 9	(E)-3'-Formyl-4',6'-dihydroxy-5'-methylchalcone	282.295	0.000	3.500	2.029	91.840
10145839 4	beta,2',4',6'-Tetrahydroxy-beta-(4-hydroxyphenyl)propiophenone	290.272	2.000	3.250	0.629	58.390
10147905 8	2',4'-Dihydroxy-2,3-dihydrochalcone	242.274	0.000	1.500	1.944	92.321
10151056 5	1-(4-Hydroxyphenyl)-3-(2,4,6-trihydroxyphenyl)propane-2-ol	276.288	3.000	3.250	0.876	64.571
10152543 4	3-Fluoro-alpha,beta-dihydro-2',4',6'-trihydroxychalcone	276.264	1.000	2.500	1.794	89.024
10160719 6	3'-Formyl-2',4',6'-trihydroxychalcone	284.268	0.000	3.250	1.188	83.562
10188530 7	(E)-1-(2,4-dihydroxyphenyl)-3-(6,6-dimethyl-2-bicyclo[3.1.1]hept-2-enyl)prop-2-en-1-one	284.354	1.000	2.500	3.718	100.000
10192699 4	4',6'-Dichloro-2'-hydroxychalcone	293.149	0.000	0.750	2.989	100.000
10193255 6	1-(2-Hydroxyphenyl)-3-(4-methylphenyl)-2-propyn-1-one	236.270	0.000	1.750	2.215	100.000
10199834 2	(2,4-Dihydroxyphenyl)[3alpha-(3,4-dihydroxyphenyl)oxiran-2beta-yl] ketone	288.256	1.500	3.750	0.543	59.025
10208018 0	(E)-1-(2,4-Dihydroxyphenyl)-3-(2-methoxy-4-hydroxyphenyl)-2-propene-1-one	286.284	2.000	3.250	0.995	77.255
10208135 5	2',3,4',6'-Tetrahydroxychalcone	272.257	2.000	3.000	0.886	75.576
10208135 6	2',4',6'-Trihydroxy-4-aminochalcone	271.272	3.000	3.250	0.027	57.649
10219914 2	4-[(2S)-2-Hydroxy-3-(2,4,6-trihydroxyphenyl)propyl]phenol	276.288	3.000	3.250	0.876	64.571
10247099 2	Cilicicone b	322.271	2.000	4.000	0.196	49.745
11840037 3	Methyl 2-[2-oxo-2-(2,4,6-trihydroxyphenyl)ethoxy]benzoate	318.282	2.000	6.500	0.255	64.371
12704393 7	(E)-3-(2,5-difluorophenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	292.238	1.000	2.250	2.342	92.137
12932040 3	Naringenin dibenzoylmethane tautomer	288.256	1.000	4.250	0.403	70.022
12962990 5	4,4'-Dihydroxy-3,5,3'-trimethylchalkon	282.338	1.000	2.500	3.412	100.000

Análisis *in silico* para suprimir la expresión génica *Drosophila melanogaster*

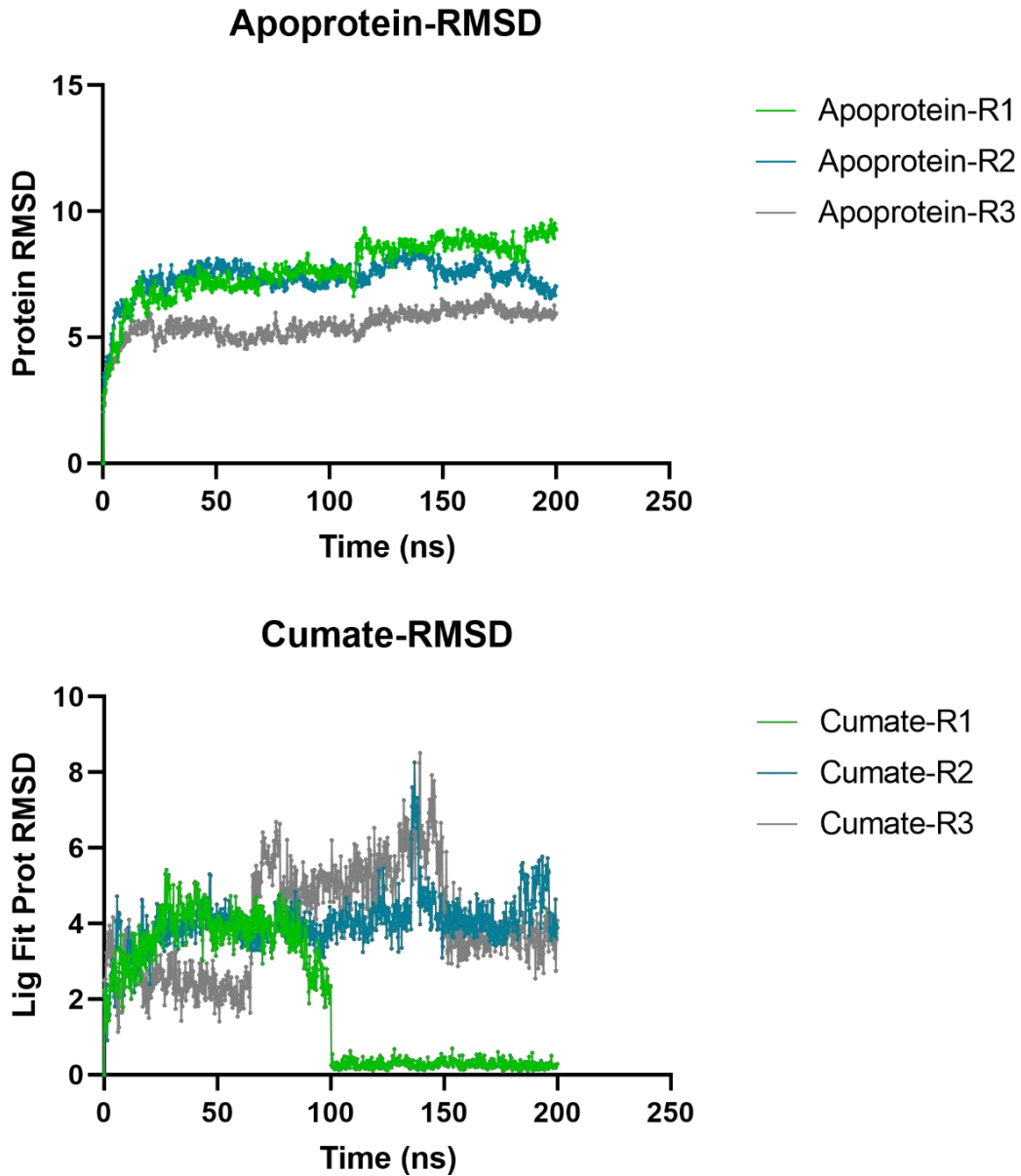
12964215 8	Tetraoxychalkone	272.257	0.000	3.500	0.979	89.481
12964761 9	4,4'-Dihydroxy-3,5,3',5'-tetramethylchalkon	296.365	1.000	2.500	3.830	100.000
12968559 7	2',3',4,4',5',6'-Hexahydroxychalkone	304.256	2.000	2.500	0.753	51.607
12968577 6	3-Cyclohexa-2,4-dien-1-yl-1-(2,4-dihydroxyphenyl)prop-2-en-1-one	242.274	0.000	1.500	2.027	92.808
12968655 8	2',6',4-Tri-hydroxy-4'-methoxydihydrochalkone	288.299	3.000	4.250	0.773	82.432
12983512 2	2-(4-Chlorophenoxy)-1-(2,4,6-trihydroxyphenyl)ethanone	294.691	1.000	3.500	1.281	85.404
12988306 4	2-Bromo-2',4',6'-trihydroxychalkone	335.153	1.000	2.250	2.187	91.014
13183636 7	3-(4-Methoxyphenyl)-1-(2,3,4,6-tetrahydroxyphenyl)propan-1-one	304.299	2.000	3.250	1.281	68.358
13183776 1	1-(2,4-Dihydroxy-3-methoxyphenyl)-3-(4-hydroxyphenyl)propan-1-one	288.299	2.500	5.000	0.849	69.402
13227816 6	O-methoxy phloretin	304.299	3.000	5.250	0.723	64.160
13249110 0	1-(2-Hydroxy-4-fluorophenyl)-3-(4-methylphenyl)propane-1-one	258.292	0.000	1.750	2.824	100.000
13413867 2	Isoliquiritigenin metabolite M2	272.257	1.000	2.000	1.406	74.432
13486818 2	(E)-1-(5-bromo-2,4-dihydroxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one	335.153	1.000	3.500	1.507	83.734
13666280 0	1-(2-Hydroxyphenyl)-3-phenyl-2-propen-1-one hydrazone	238.288	4.000	1.750	1.143	92.471
13763197 5	3-(4-Propoxyphenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	316.353	1.000	3.500	2.132	91.001
13764041 4	3-(4-Ethoxyphenyl)-1-(2,4,6-trihydroxyphenyl)propan-1-one	302.326	1.000	3.500	1.688	88.404
13766025 1	(E)-1-(2,6-dihydroxyphenyl)-3-(3-methoxyphenyl)prop-2-en-1-one	270.284	2.000	3.500	1.415	94.536
15471078 0	3-(3-Chlorophenyl)-1-(2-hydroxyphenyl)prop-2-en-1-one	256.688	0.000	1.750	2.319	100.000
15593260 6	1-(4-Hydroxy-2,3,5,6-tetramethylphenyl)-3-phenylpropan-1-one	282.382	0.000	0.750	4.549	100.000
15689183 9	(E)-3-(3,4-dihydroxyphenyl)-1-(4-hydroxy-2-methylphenyl)prop-2-en-1-one	270.284	2.000	3.250	0.789	75.219
16300326 0	3-(2,3,4-Trihydroxyphenyl)-1-(2,3,6-trihydroxyphenyl)prop-2-en-1-one	304.256	4.000	4.500	-0.159	51.561
16462169 3	1-(2,5-Dihydroxyphenyl)-3-(2-fluoro-4-hydroxyphenyl)propan-1-one	276.264	1.000	2.250	1.702	82.661

Apéndice 3. RMSD de las proteínas TtgR y CymR sin ligando.



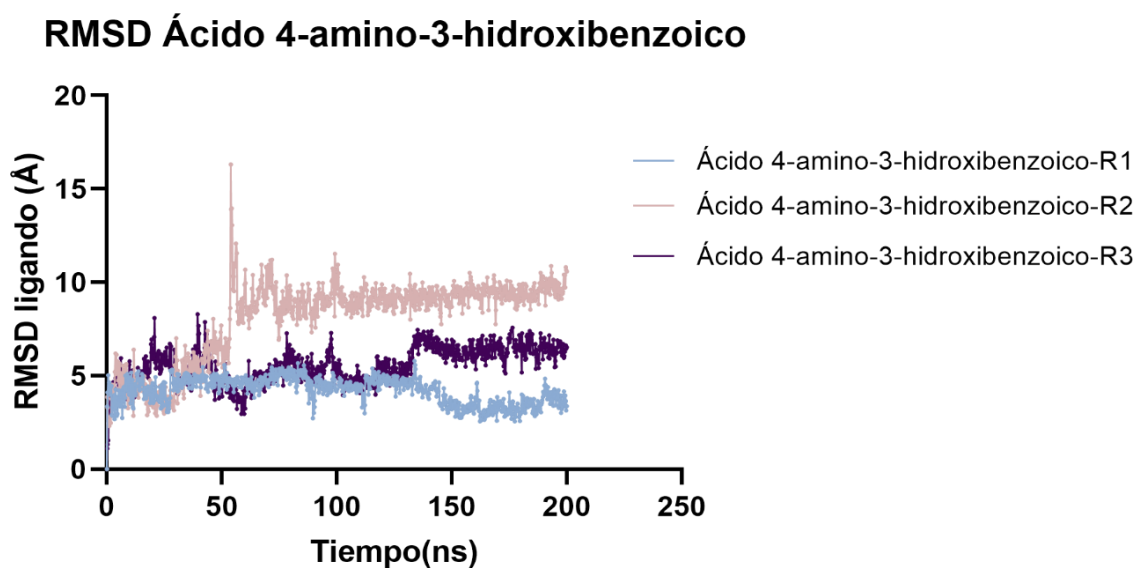
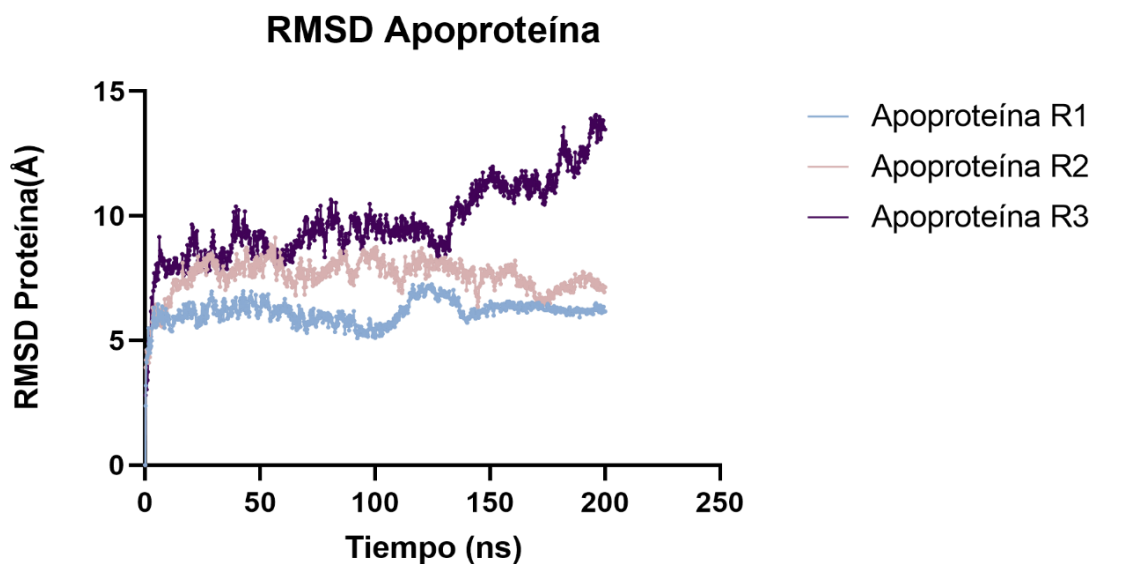
*Nota** Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para las proteínas TtgR y CymR sin ligando. Cada color representa una réplica

Apéndice 4. RMSD de la proteína CymR sin ligando y con Cumato.



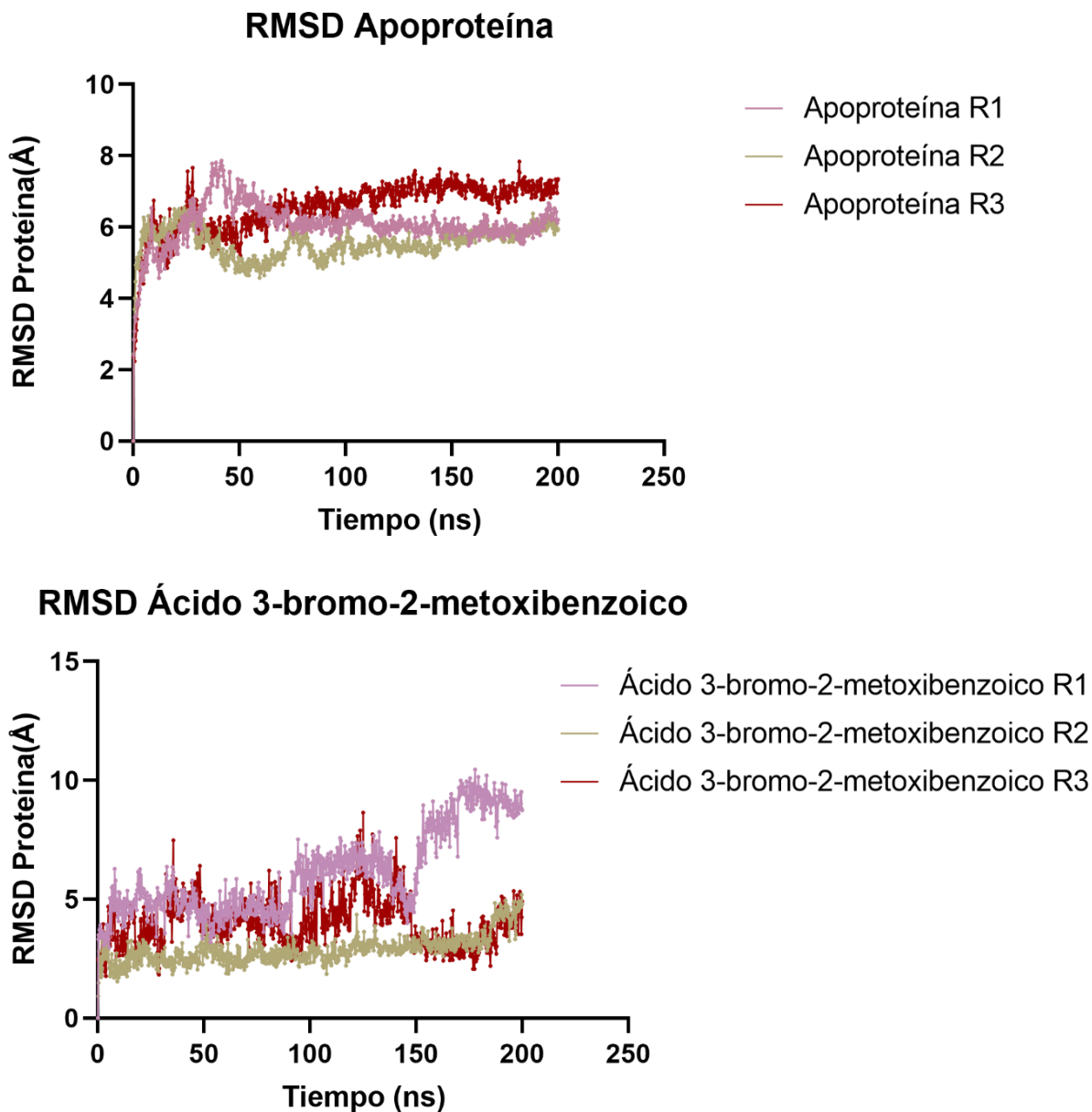
*Nota** Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína CymR sin ligando y con Cumato. Cada color representa una réplica

Apéndice 5. RMSD de la proteína CymR sin el ligando y con Ácido 4-amino-3-hidroxibenzoico



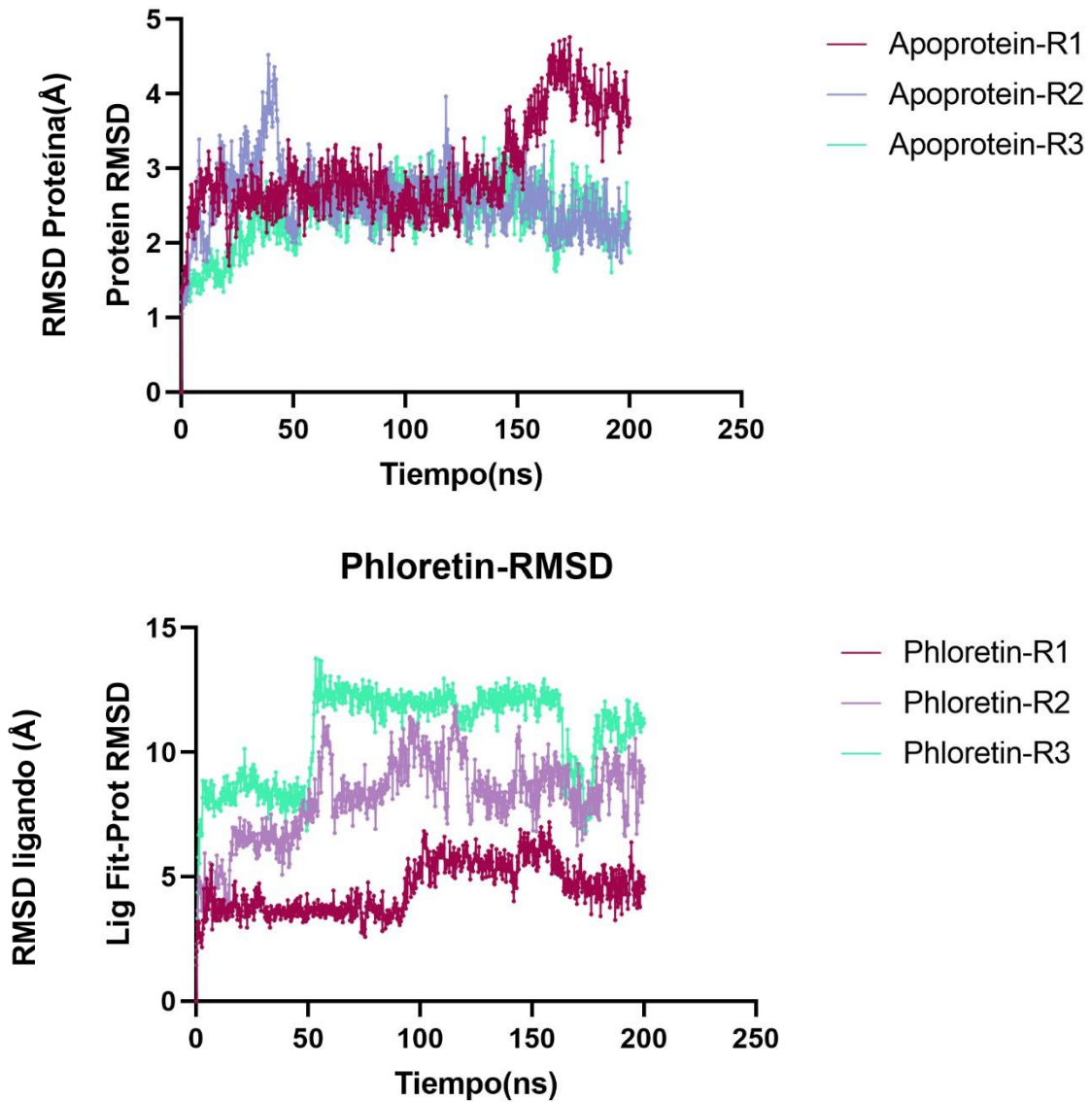
*Nota** Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína CymR sin ligando y Ácido 4-amino-3-hidroxibenzoico. Cada color representa una réplica

Apéndice 6. RMSD de la proteína CymR sin el ligando y con Ácido 3-bromo-2-metoxibenzoico



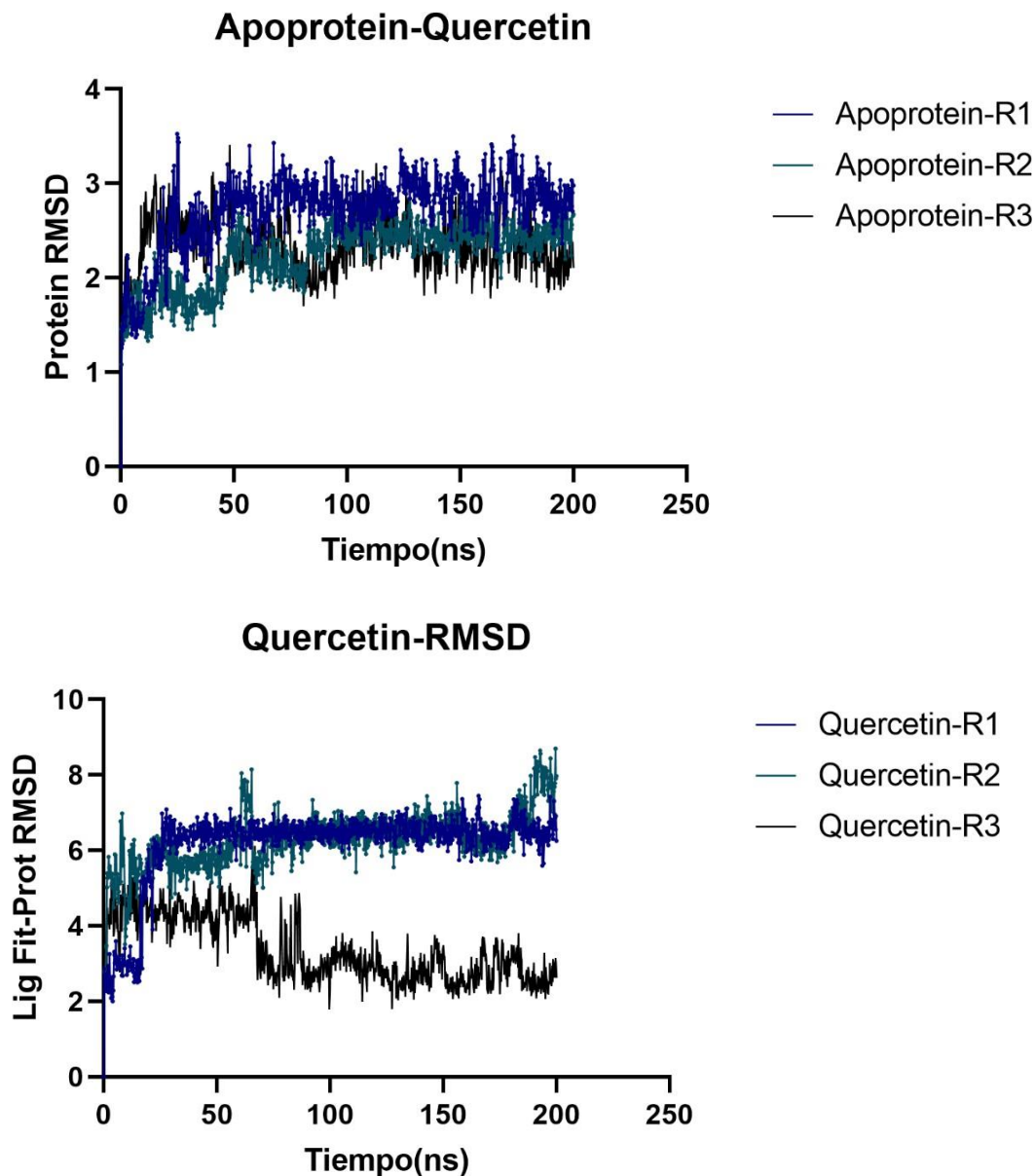
*Nota** Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína CymR sin ligando y Ácido 3-bromo-2-metoxibenzoico. Cada color representa una réplica

Apéndice 7. RMSD de la proteína TtgR sin el ligando y Floretina.



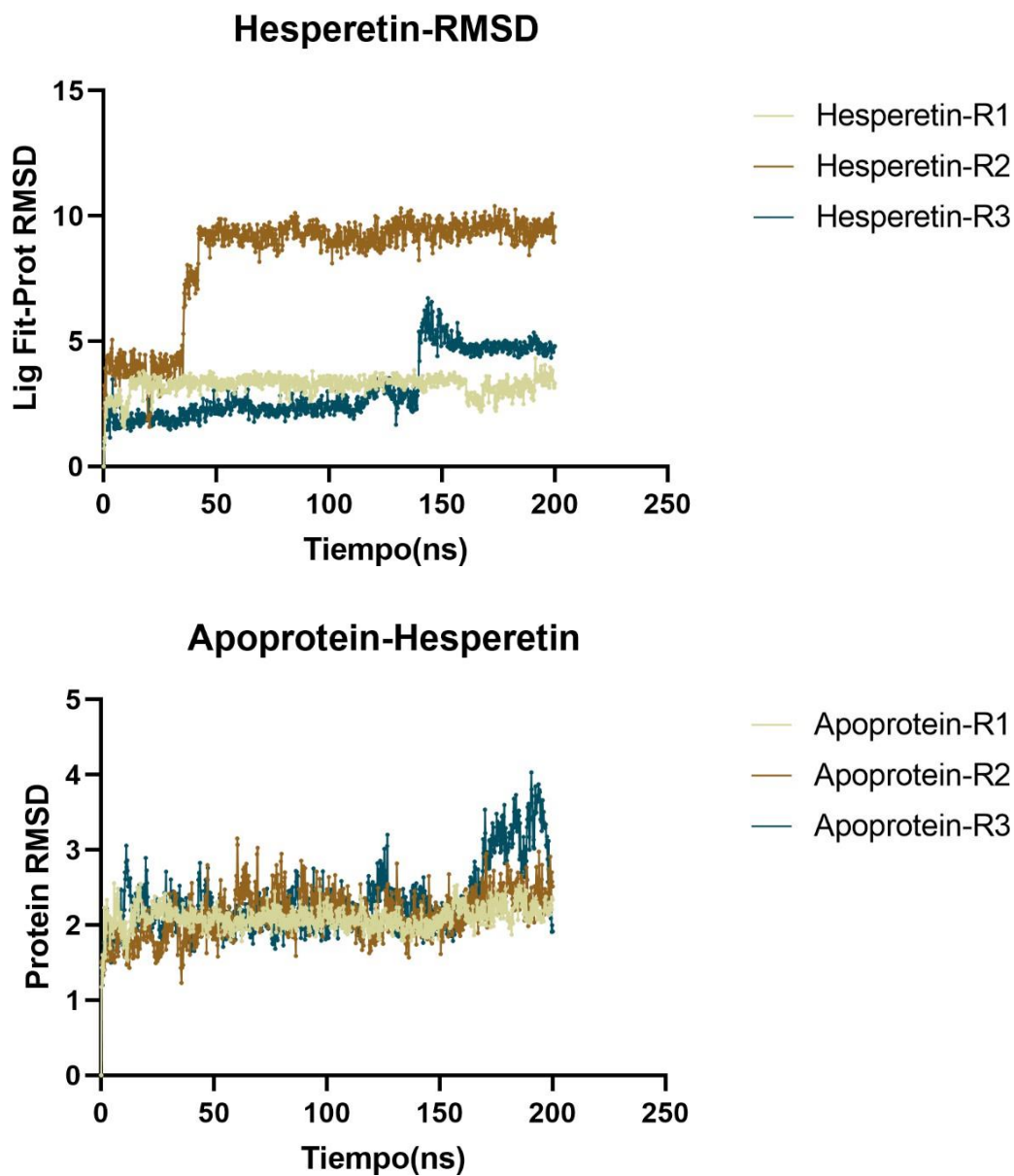
Nota Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína TtgR sin ligando y Floretina. Cada color representa una réplica*

Apéndice 8. RMSD de la proteína TtgR sin el ligando y Quercetina.



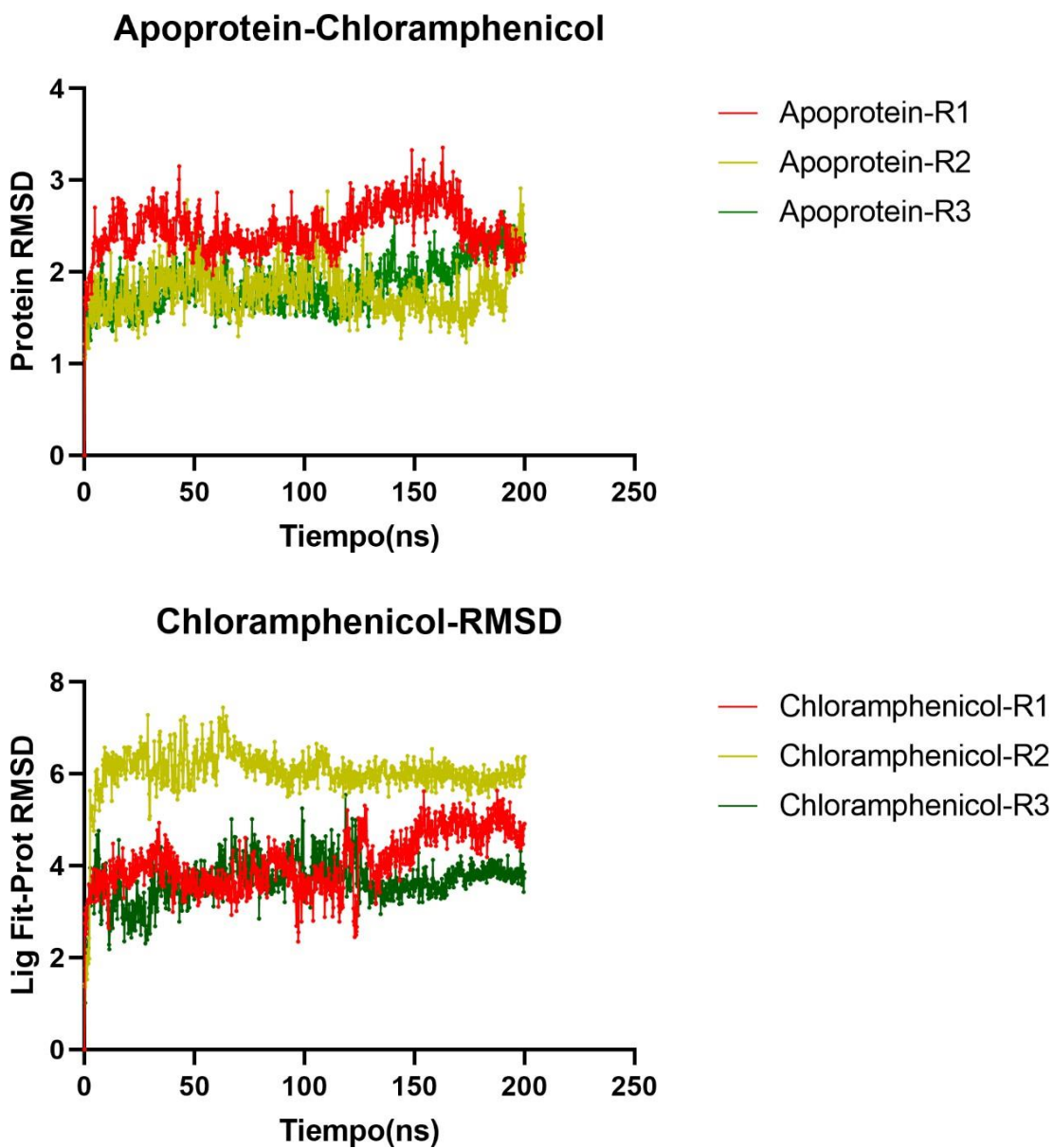
*Nota** Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína TtgR sin ligando y Quercetina. Cada color representa una réplica

Apéndice 9. RMSD de la proteína TtgR sin el ligando y Hesperetina.



Nota Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína TtgR sin ligando y Hesperetina. Cada color representa una réplica*

Apéndice 10. RMSD de la proteína TtgR sin el ligando y Hesperetina.



Nota Elaboración propia (2023). Imagen representativa de los valores de desviación cuadrática media RMSD para la proteína TtgR sin ligando y Cloranfenicol. Cada color representa una réplica*