



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

mec
MINISTERIO DE EDUCACIÓN Y CULTURA



INSTITUTO DE
INVESTIGACIONES BIOLÓGICAS
CLEMENTE ESTABLE



Identificación de los blancos de acción molecular de flavonoides mediante tamizaje virtual en librerías de estructura tridimensional de proteínas

Diego Carvalho Álvarez

Programa para el Desarrollo de las Ciencias Básicas – Maestría en Bioinformática
Universidad de la República - Facultad de Química
Ministerio de Educación y Cultura – Instituto de Investigaciones Biológicas
Clemente Estable

Montevideo – Uruguay
Noviembre de 2016



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

mec
MINISTERIO DE EDUCACIÓN Y CULTURA



INSTITUTO DE
INVESTIGACIONES BIOLÓGICAS
CLEMENTE ESTABLE



Identificación de los blancos de acción molecular de flavonoides mediante tamizaje virtual en librerías de estructura tridimensional de proteínas

Diego Carvalho Álvarez

Tesis de Maestría presentada al Programa para el Desarrollo de las Ciencias Básicas – Maestría en Bioinformática, como parte de los requisitos necesarios para la obtención del título de Magister en Bioinformática.

Directores de Tesis:

Dra. Margot Paulino Zunini

Dr. Fabio Polticelli

Codirector:

Dr. Juan A. Abin Carriquiry

Montevideo – Uruguay
Noviembre de 2016

Integrantes del Tribunal de Defensa de Tesis

Dr. Hugo Cerecetto

Dr. Sergio Pantano

Dr. Federico Iribarne

Montevideo – Uruguay
Noviembre de 2016

Agradecimientos

Quisiera agradecer a a todas aquellas personas que contribuyeron de diversas maneras en mi formación durante la realización de esta maestría, comenzando por mis orientadores Margot Paulino, Andrés Abin y Fabio Polticelli por compartir sus conocimientos, dedicarme su tiempo y apoyarme en este camino. A Florencia Arredondo y Federico Dajas por sus aportes en la interpretación de los resultados. A Carolina Mascayano por permitirme realizar una pasantía en su laboratorio.

Quiero agradecer también a mi familia por su apoyo y comprensión. A todos mis amigos y compañeros del CeBioinfo y del Departamento de Neuroquímica con los que comparto el trabajo día a día, especialmente a Marcela, Federico, Brenda, Vicente, Jorge, Emiliana, Nicolas, Daniel, Felicia y una larga lista así como a los compañeros de la USaCh: Francisco, Claudia, Beatriz, José, Nicole, Matías, Macarena, Pablo por hacerme sentir como en casa.

Agradezco también a las organizaciones y personas que hicieron posible esta tesis mediante la financiación de diversos rubros: a PEDECIBA Bioinformática, por permitirme realizar un posgrado gratuito en el país, a la ANII, por haber financiado 24 meses la realización de mi posgrado.

RESUMEN

Quercetina, uno de los flavonoides más abundantes en la dieta occidental, es ampliamente reconocido por sus propiedades antioxidantes y citoprotectoras por lo que se le atribuye un potencial tratamiento en patologías cuyo desarrollo involucra daño oxidativo, tales como las enfermedades neurodegenerativas [1, 2].

Trabajos precedentes en el Dpto. de Neuroquímica del IIBCE evidenciaron la actividad citoprotectora del flavonoide sobre cultivos neuronales frente al daño oxidativo [2]. Aunque su acción citoprotectora se ha atribuido clásicamente a propiedades antioxidantes directas, la evidencia acumulada indica que quercetina es capaz de interactuar con múltiples blancos y modular mecanismos de defensa celular entre otros [1].

Sin embargo, este mecanismo de acción no ha sido totalmente caracterizado. En el presente proyecto nos propusimos realizar el tamizaje reverso basado en la similitud y características farmacofóricas del ligando (*SHAFTS*) así como en la estructura del blanco (*LIBRA*, *idTarget*) para quercetina y un set asociado de flavonoides 98 % similares [3–5]. Esta estrategia permitió identificar un conjunto de proteínas candidato estructural y funcionalmente heterogéneo, donde se destaca la presencia de proteínas quinasa y poli-ADP-ribosa polimerasas vinculadas a procesos de sobrevivencia y muerte celular, así como a patologías del sistema nervioso [6, 7]. Los presentes candidatos ilustran un posible mecanismo de acción pleiotrópico y sinérgico para los efectos biológicos de quercetina.

Palabras claves:

Tamizaje reverso, anclaje molecular, flavonoides, quercetina.

Tabla de contenidos

1 Introducción	7
2 Fundamentos teóricos	8
3 Hipótesis – Objetivos	16
4 Metodología	17
5 Resultados y discusión	42
6 Consideraciones finales	84
7 Referencias bibliográficas	86
8 Apéndices. Información complementaria	111
9 Anexos. Material suplementario	121
10 Fin	150

Introducción

Las dietas ricas en fitoquímicos están correlacionadas con un amplio rango de efectos benéficos incluyendo un incremento de la longevidad, disminución de la incidencia de enfermedades cardiovasculares y una menor progresión de enfermedades cerebrovasculares [8–12]. Los flavonoides constituyen uno de las mayores familias de fitoquímicos encontradas en plantas. Junto con su contribución al sabor y color de frutas y vegetales, los flavonoides son ampliamente reconocidos por sus propiedades antioxidantes directas [2, 13–17]. En particular, varios flavonoides muestran efectos antiinflamatorios, anticancerígenos y neuroprotectores entre otros [18–30].

Este grupo polifenólico ubicuo comprende varias subclases tales como flavonas (incluyendo flavonoles), flavanonas, flavanos (incluyendo flavanoles), isoflavonas, chalconas y antocianidinas. Los flavonoles son los más abundantes y más ampliamente distribuidos en la naturaleza, siendo quercetina (3',4',3,5,7 pentahidroxi flavona) un flavonoide prototípico. En relación a los efectos neuroprotectores y anticancerígenos previamente citados, varios estudios sugieren que están mediados por la interacción con proteínas específicas involucradas en cascadas de señalización intracelular en lugar de mecanismos antioxidantes directos [31–40]. En este contexto una completa explicación del amplio rango de actividades biológicas de quercetina requiere la descripción de su espacio completo de blancos moleculares.

La predicción *in silico* de blancos de acción molecular es una alternativa rápida y accesible a ensayos de laboratorio húmedo tales como la cromatografía de afinidad o los microarreglos de proteínas [41]. *ReverseScreen3D*, *INVDOCK*, *Tarfisdock*, *idTarget* y *PharmMapper* son algunas de las herramientas *in silico* disponibles [5, 42–44]. Experiencias favorables con estas plataformas han logrado descubrir proteínas blanco para nuevos compuestos como xiloquetol, quinetina, así como fármacos aprobados como mebendazol [45–49]. Por una completa revisión de esta área referirse a Cereto-Massagué y col. (2015) [50].

En el presente estudio proponemos una estrategia de tamizaje jerárquico de proteínas blanco de quercetina basado en el arreglo secuencial de búsqueda por similitud de ligando, comparación de sitio de unión y anclaje molecular inverso (*inverse docking*). Se evaluó la relevancia del conjunto de proteínas blanco en relación a los efectos reportados para quercetina. Finalmente un conjunto seleccionado de blancos se analizó mediante dinámica molecular aportando evidencias favorables al anclaje de quercetina.

Fundamentos teóricos

• Flavonoides, estructura y función

Como se mencionó previamente, los flavonoides son compuestos polifenólicos ubicuos de las plantas, encontrándose en frutas, verduras, semillas y flores, así como en sus derivados. Desempeñan un papel importante en la biología vegetal, responden a la luz y controlan los niveles de las auxinas reguladoras del crecimiento y diferenciación de las plantas. Otras funciones incluyen un papel antifúngico y bactericida, confieren coloración lo que puede contribuir a los fenómenos de polinización y tienen una importante capacidad para fijar metales como el hierro y el cobre [51].

Estructuralmente los flavonoides son compuestos de bajo peso molecular que comparten un esqueleto común de difenilpiranos (C₆-C₃-C₆), integrado por dos anillos de fenilos (A y B), ligados a través de un anillo C de pirano (heterocíclico). Los átomos de carbono en los anillos C y A se numeran del 2 al 8, y los del anillo B desde el 2' al 6' (figura 1) [52]. Esta estructura básica permite una multitud de patrones de sustitución y variaciones en el anillo C, habiéndose descrito hasta el momento más de 10 000 flavonoides [53].

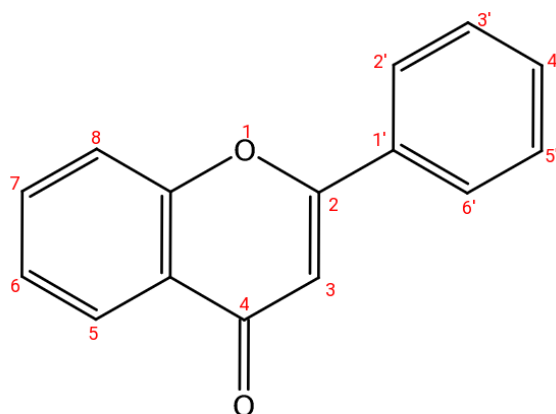


Figura 1. Estructura básica de un flavonoide típico.

Existen diversas familias de flavonoides: flavonas, flavonoles, flavan-3-oles, isoflavonas, flavanonas, y antocianidinas. Estas se diferencian por el grado de oxidación y sustitución del heterociclo. Dentro de cada familia, los distintos compuestos se diferencian por el número y localización de los grupos hidroxilo y/o metoxilo de sus ciclos bencénicos (anillos A y B). En la naturaleza la mayoría se encuentra conjugados a azúcares o a ácidos orgánicos, y si bien algunos pueden encontrarse en forma libre -agliconas- y otros están frecuentemente polimerizados [54]. La conjugación ocurre preferentemente en posición 3 y con menor frecuencia en la posición 7 del anillo A originando O-glicósidos con la D-glucosa.

En función de sus características estructurales se pueden clasificar en:

Tabla 1. Características estructurales y grupos funcionales de las principales familias de flavonoides. *La posición del anillo B respecto al anillo C es 2 en todos los casos, a excepción de las isoflavonas, que es 3.

Familia	Insaturación en anillo C	Grupo funcional en anillo C
flavanoles	-	3-hidroxi
flavanonas	-	4-oxo
flavonas	en 2-3	4-oxo
isoflavonas*	en 2-3	4-oxo
flavonoles	en 2-3	3-hidroxi y 4-oxo
antocianinas	en 1-2 y en 3-4	3-hidroxi

Alternativamente a la estructura 2-fenil-1,4 benzopirona descrita típica de los flavonoides se encuentran los isoflavonoides con estructura 3-fenil-1,4-benzopirona y neoflavonoides derivados de la 4-fenil-1,2-benzopirona (figura 2). Asimismo existe una multitud de compuestos relacionados entre los que se incluyen por ejemplo las chalconas, precursores de la síntesis de los flavonoides en las plantas que carecen del heterociclo central y los estilbenos como resveratrol. Al igual que los flavonoides estos compuestos también están asociados a numerosos efectos benéficos en los mamíferos, incluyendo aumento de la longevidad, anti-inflamación y antivirales, así como protección contra el cáncer, enfermedades cardiovasculares y aterosclerosis [2, 55].

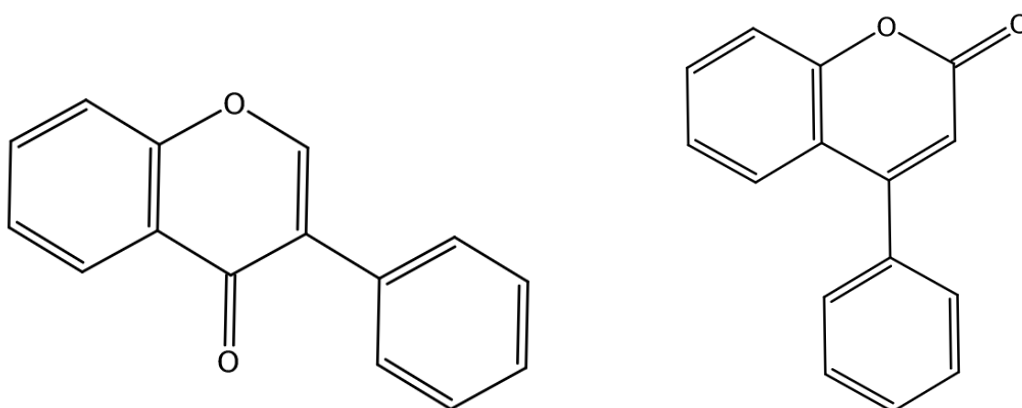


Figura 2. Estructura básica de un isoflavonoide y un neoflavonoide típicos.

◦ *Quercetina, efectos y mecanismos documentados*

Quercetina (3,3',4',5,7-pentahidroxi flavona) es uno de los flavonoides dietarios más abundantes y más estudiados [56, 57]. Entre los alimentos y bebidas con mayores niveles de quercetina encontramos a las cebollas, las manzanas, al té y al vino [58].

El consumo promedio de este flavonoide en una dieta occidental normal se sitúa entre 5 y 40 mg/día. Sin embargo, en dietas ricas en frutas y verduras el consumo diario de quercetina puede situarse entre 200 y 500 mg [58–62].

Quercetina es ampliamente reconocida por su espectro de efectos beneficiosos, tales como anti-cancerígenos, anti-inflamatorios, cardioprotectores, entre otros [63–66]. La evidencia acumulada indica que quercetina es una potencial herramienta terapéutica para el tratamiento o prevención de desórdenes neurodegenerativos como la enfermedad de Alzheimer, la enfermedad de Parkinson, y la esclerosis lateral amiotrófica entre otros [67–69]. Estas neuropatologías presentan diversas sintomatologías secundarias como producto de la degeneración de poblaciones neuronales vulnerables específicas incluyendo desde alteraciones motoras (como en las enfermedades de Huntington y Parkinson) hasta demencias severas (como en el Alzheimer) [70]. Aunque estos desórdenes presentan distintas características clínicas y neuropatológicas, comparten ciertas alteraciones que contribuyen a la disfunción y muerte neuronal. Entre estas alteraciones encontramos: defectos en el sistema proteosomal y agregación proteica, disrupción de la homeostasis del calcio, inflamación, excitotoxicidad, alteración de los procesos metabólicos y de la función mitocondrial, y aumento del daño oxidativo. La influencia de quercetina y otros flavonoides sobre estos aspectos de la patología ha sido extensamente estudiada tanto en modelos *in vitro* como *in vivo*.

En particular, numerosos estudios *in vitro* e *in vivo* mostraron efectos neuroprotectores de este flavonoide. Se incluyen desde modelos de muerte neuronal en líneas celulares y cultivos primarios frente a diversos agentes tóxicos (H_2O_2 , *tert-butil hidroperóxido*, *6-hidroxi-dopamina*, etc), hasta modelos experimentales de Parkinson, Alzheimer, hipoxia/isquemia, trauma, etc. revisado en [71, 72]. Por ejemplo, diversos estudios *in vitro* han mostrado que el tratamiento con quercetina puede aumentar la sobrevivencia de neuronas o líneas celulares derivadas de neuronas contra insultos oxidativos entre otras injurias [73–77]. En particular se ha observado que el tratamiento con flavonoides semejantes previene la muerte neuronal por el péptido β -amiloide, el cual juega un papel importante en el desarrollo de la enfermedad de Alzheimer. Asimismo algunos glicósidos de quercetina son capaces de mantener la integridad y funcionalidad de la sustancia negra compacta (SNpc) y del núcleo estriado de roedores tratados con *6-hidroxi-dopamina* (6-OHDA), una neurotoxina empleada para generar un modelo de enfermedad de Parkinson [78].

Estudios pre-clínicos *in vivo* de los efectos preventivos de quercetina sobre neurodegeneración se han enfocado en su mayoría sobre deficiencias cognitivas, isquemia y daño por trauma [71, 72]. Se ha mostrado por ejemplo mejoría en los procesos cognitivos, posiblemente mediante la protección de neuronas vulnerables a la muerte, así como a la mejora de la función neuronal existente [9, 79]. Evidencias experimentales también muestran un efecto neuroprotector de los flavonoides en la

isquemia cerebral. En este punto en particular, nuestro grupo de investigación ha demostrado que una preparación de quercetina/lecitina administrada 30 minutos antes de la oclusión permanente de la arteria cerebral media, reduce significativamente el daño motor y el volumen del infarto cerebral causado por la oclusión de la arteria durante 24 hs con respecto a las ratas isquémicas no tratadas [73, 80].

Varios mecanismos se han propuesto para explicar los beneficios de quercetina incluyendo su potente capacidad antioxidante directa y más recientemente la interacción amplia de con blancos intracelulares de cascadas de señalización y factores de transcripción, regulando la expresión de genes críticos [81, 82]. En este sentido trabajos recientes aportaron importantes evidencias sobre la capacidad de algunos flavonoides de estimular vías de respuesta adaptativa frente a situaciones de estrés celular, como el oxidativo, metabólico, inflamatorio (figura 3).

■ *Efectos y mecanismos antioxidantes*

El estrés oxidativo es uno de los elementos centrales de diversas patologías. Se define como una situación de desbalance entre la generación celular de pro-oxidantes y la actividad de los sistemas antioxidantes, que lleva al daño de componentes celulares (ácidos nucleicos, proteínas y lípidos) y a la disrupción del control y la señalización redox celular. Con el fin de mantener los niveles de las especies reactivas en proporciones no perjudiciales, los organismos aerobios desarrollaron múltiples mecanismos de defensa antioxidante. Gran parte de estas defensas actúan previniendo la formación de especies reactivas, neutralizándolas, o convirtiéndolas en sustancias menos reactivas [83]. Al respecto uno de los primeros indicios de la actividad de los flavonoides fue su acción antioxidante directa. Estudios estructura-actividad mostraron que el grupo catecol en posiciones 3' y 4' del anillo B, la presencia de una insaturación en los carbonos 2 y 3, un grupo hidroxilo libre en la posición 3 y un grupo carbonilo en la posición 4, aseguran una elevada potencia antioxidante [81, 84]. Estas características químicas les confieren la capacidad de donar hidrógeno/electrones a los radicales libres y de deslocalizar el radical resultante dentro de su estructura [81].

Quercetina presenta todas las características estructurales que expusimos previamente, necesarias para ser un efectivo secuestrador de radicales libres [85]. En este sentido, se propuso que la capacidad secuestradora de radicales libres de quercetina era la responsable de sus efectos beneficiosos [86]. Sin embargo existen limitantes cinéticas en la acción secuestradora de especies reactivas; se necesitan elevadas concentraciones del compuesto para atrapar efectivamente las especies reactivas, mientras que su concentración *in vivo* es limitada [87]. Por otra parte, las propiedades redox intrínsecas de este flavonoide hacen que luego de reaccionar con radicales libres, se convierta en un producto reactivo, con propiedades pro-oxidantes [88, 89]. Se postula que la actividad pro-oxidante de quercetina se debe fundamentalmente a la presencia del grupo catecol en posiciones 3' y 4' del anillo B.

Esta sustitución orto-di-hidroxi es susceptible a auto-oxidarse y convertirse en orto-semiquinonas y orto-quinonas/metilén quinonas, causando además la producción de especies reactivas.

Alternativamente quercetina y otros flavonoides serían capaces de evitar la generación de especies reactivas modulando la actividad de enzimas oxidasas como la *ciclooxigenasa*, *xantina oxidasa*, *5-lipooxigenasa*, entre otras [90–95]. Sin embargo, estudios previos en nuestro laboratorio han demostrado que la actividad antioxidante es una propiedad necesaria pero no suficiente para el efecto neuroprotector de quercetina [32].

Una aproximación interesante para restaurar la homeostasis redox involucra la activación del factor de transcripción *Nrf2* (*Nuclear factor erythroid 2-related factor 2*). Este miembro de la familia *Cap'n'Collar* de factores de transcripción de tipo cierre de leucina juega un rol central en la regulación del estado redox a través de la modulación de un amplio rango de respuestas adaptativas a fuentes intrínsecas y extrínsecas de estrés [96]. *Nrf2* activa la transcripción de más de 200 genes que están involucrados en el metabolismo de drogas, citoprotección e integridad celular tales como *NAD(P)H: quinona oxidoreductasa 1 (NQO1)*, *ligasa glutamato-cisteína* (anteriormente conocida como *gamma-glutamilcisteína sintasa*, *GCS*), *hemo-oxigenasa 1 (HO-1)*, entre otros [97]. El sistema de *Nrf2* está ampliamente expresado en el sistema nervioso central y es activado en respuesta a diferentes condiciones patológicas como daño cerebral agudo y enfermedades neurodegenerativas [98–101]. Específicamente, la alteración de esta vía ha sido relacionada con diversas patologías del sistema nervioso [102]. Sobre todo, numerosos estudios revelan que la activación de la vía de *Nrf2*, tanto por vías genéticas como químicas, confiere protección en diversos modelos *in vitro* e *in vivo* de patologías del sistema nervioso central a través de la sobre-expresión de genes de fase II [103–106]. En este sentido, hemos descrito previamente los efectos de quercetina contra un pretratamiento de H₂O₂ durante 24 h en cultivos primarios de células granulares de cerebelo [19]. Quercetina se internaliza rápidamente en las neuronas alcanzando el núcleo luego de su adición al cultivo. Previno la muerte neuronal por exposición a H₂O₂ aunque al momento de la adición del oxidante quercetina intracelular o metabolitos relacionados eran indetectables en los cultivos. El pre-tratamiento con quercetina causó la translocación nuclear de *Nrf2* y el incremento de la expresión de la enzima antioxidante *HO-1* así como de la subunidad catalítica de la enzima *ligasa glutamato-cisteína (GCS)*. Esta enzima es el paso limitante en la biosíntesis de glutatión un modulador central la homeostasis redox celular incluyendo la regulación de proteínas con grupos tioles y las defensas antioxidantes. En el paradigma experimental utilizado quercetina aumenta el nivel de glutatión total y la inhibición irreversible de la *GCS* suprime sus efectos protectores demostrando la importancia de este mecanismo.

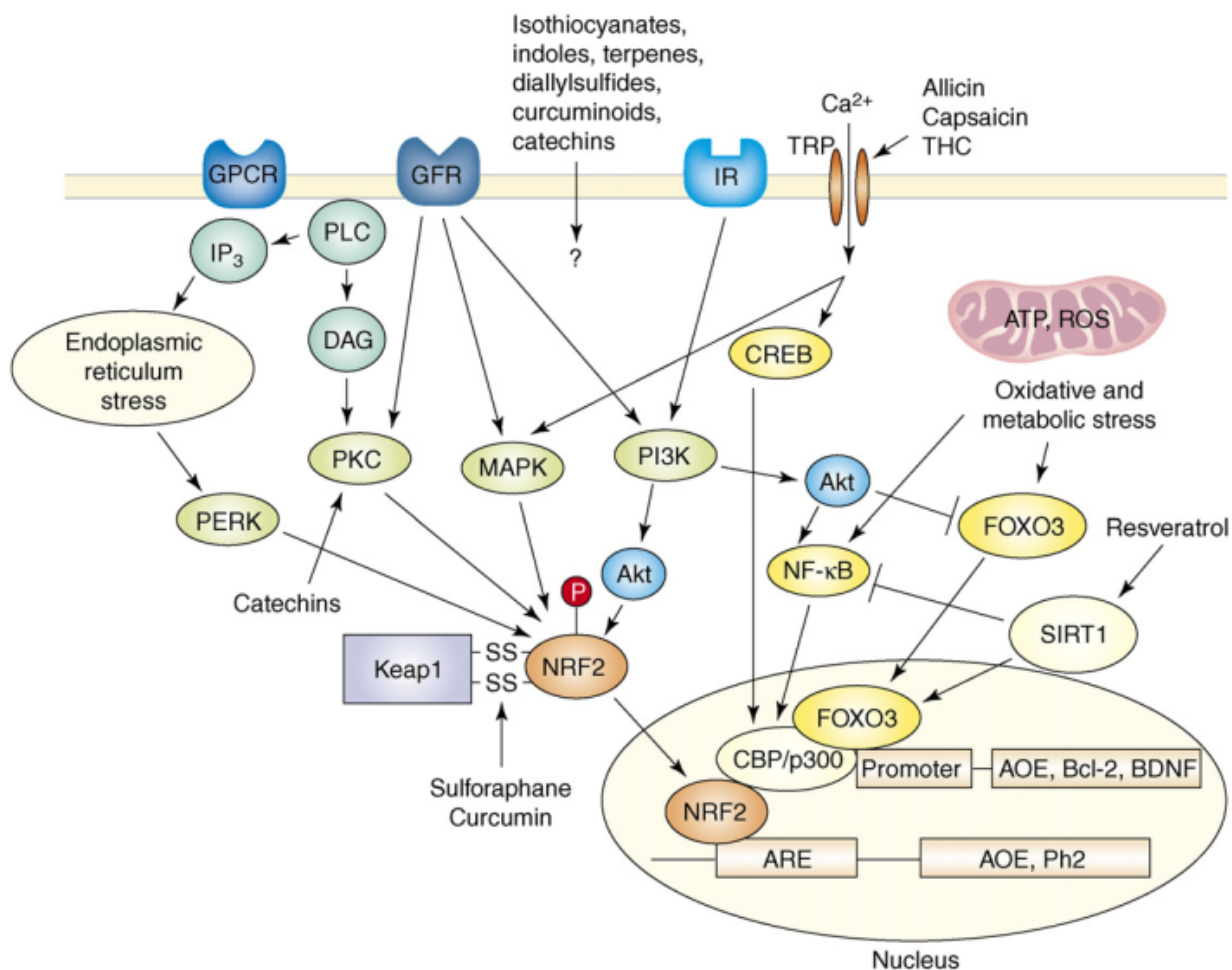


Figura 3. Mecanismos propuestos de activación por polifenoles de una respuesta adaptativa a través de Nrf2 (tomado de Mattson y col. 2006).

■ Efectos y mecanismos anti-inflamatorios

Los flavonoides en general son considerados como reguladores del sistema inmune y como antiinflamatorios lo que formaría parte de sus efectos *in vivo* como la protección contra el daño isquémico [107]. Uno de los mecanismos propuestos es la inhibición directa e indirecta de enzimas pro-oxidantes y pro-inflamatorias como *ciclooxigenasa*, *lipooxigenasa*, *fosfolipasa A2*, *hialuronidasa*, *enzima convertidora de angiotensina*, *mieloperoxidasa* (que produce el hipoclorito y otros pro-oxidantes) y *xantina oxidasa* (que produce los radicales superóxido) entre otras lo que conlleva una inhibición de la producción de mediadores pro-inflamatorios como los eicosanoides [2, 108].

Asimismo, ha sido demostrado la acción de los flavonoides sobre factores de transcripción como *NF-κB* (*factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas*) y *AP-1* (*proteína activadora 1*) [109]. La familia *NF-κB* de factores de transcripción consiste de cinco miembros, incluyendo *c-Rel*, *RelA* (*p65*), *RelB*, *NF-κB1* (*p50*), y *NF-κB2* (*p52*). La forma más abundante de *NF-κB* es un heterodímero compuesto por *RelA* y *NF-κB1*. En células en reposo, la mayoría de los dímeros permanecen inactivados en el citoplasma debido a la presencia de

proteínas inhibitorias (*IκB*). Frente a un estímulo, el complejo quinasa *IκB* (*IKK*) que consiste de *IKKα*, *IKKβ*, e *IKKγ/NEMO* fosforila *IκB* conduciendo a su ubiquitinación y degradación por el proteosoma. Esto permite a los dímeros de *NF-κB* entrar al núcleo e unirse a los sitios *κB* de los promotores/potenciadores de los genes blanco, que incluyen citoquinas, factores de crecimiento, moléculas de adhesión celular y proteínas anti-apoptóticas. La activación de *NF-κB* es importante para la proliferación celular normal y las respuestas inmunes. Una activación inapropiada o excesiva puede originar desórdenes inflamatorios como artritis autoinmune y favorecer el desarrollo y progresión del cáncer [110–114]. Sobre todo, la activación constitutiva de *NF-κB* ha sido observada en varios tipos de cánceres incluyendo linfomas, melanomas, mama, etc. [110, 115]. En adición *NF-κB* estimularía directamente la supervivencia celular, la progresión del ciclo celular, angiogénesis, invasión y metástasis [116]. Evidencia *in vitro* indica que quercetina y otros flavonoles son capaces de inhibir la activación del *NF-κB* inducida por agentes inflamatorios [117, 118]. En varios tipos celulares, incluyendo microglia, macrófagos, neuronas y células endoteliales el tratamiento con quercetina inhibe la producción de mediadores y enzimas pro-inflamatorias como *TNFα*, *IL-8*, *IL-1α* y *NOSi* [119]. En co-cultivos de microglia-neurona, este efecto resultó en una disminución de la muerte neuronal inducida por activación de la microglia [25, 120, 121]. En modelos *in vivo* el efecto anti-inflamatorio posee un rol en la supervivencia celular [122]. Resultados equivalentes fueron evidenciados para resveratrol y otros flavonoides como fisetina y kaempferol [2, 123, 124].

Una posible explicación de estos efectos anti-inflamatorios puede ser encontrada en la correlación entre el estrés oxidativo y la inflamación ya que las especies reactivas también exacerban procesos inflamatorios a través de la activación de factores de transcripción como *NF-κB* y la *proteína activadora 1* (*AP-1*). Otra posible explicación tras esta modulación es la inhibición de la degradación de la unidad inhibitoria (*IκBα*) de este factor de transcripción [119]. En esta línea, se ha evidenciado la desfosforilación de proteínas de señalización claves en esta cascada como *ERK*, *p38*, *JNK*, *Akt*, y *Src* lo que sugiere la acción directa de quercetina sobre esta vía [120].

■ *Efectos y mecanismos sobre la proliferación celular*

Apoptosis es el proceso de muerte celular caracterizado por encogimiento de la célula, condensación del núcleo, fragmentación del ADN y alteración de la membrana [125]. Las vías apoptóticas involucran la activación de la cascada de caspasas, conduciendo a la muerte celular. La familia *Bcl-2* consiste de dos grupos, el grupo anti-apoptótico (*Bcl-2* y *Bcl-xL*) y el grupo pro-apoptótico (*Bak*, *Bax* y *Bid*). La proteína anti-apoptótica *Bcl-2* se localiza en la membrana externa de la mitocondria e inhibe la liberación de citocromo c. En contraste, los factores pro-apoptóticos residen en el citosol. La translocación de *Bax* a la membrana mitocondrial lleva a la pérdida del potencial de membrana mitocondrial y un aumento de su permeabilidad. Se ha

evidenciado que el tratamiento concomitante con quercetina suprime la apoptosis inducida por H₂O₂, caracterizada por la fragmentación del ADN, activación de la cascada de caspasas y la expresión del gen *Bax* así como el aumento de la expresión de *Bcl-2* en células neuronales humanas *SH-SY5Y* [126]. Si bien estos efectos pueden atribuirse a mecanismos antioxidantes, se ha propuesto la modulación de proteínas de señalización como *quinasas c-Jun N-terminal (JNK)*. Alternativamente también pueden actuar manteniendo la homeostasis del calcio, ión implicado en la activación de las MAP-quinasas [40].

Paradójicamente quercetina también está asociada a efectos pro-apoptóticos. Como se mencionó anteriormente estudios epidemiológicos han reportado que el consumo de alimentos ricos en quercetina reduce el riesgo de cáncer gástrico en un 43 % y de cáncer de colon en un 32 % [127, 128]. El consumo de quercetina también se ha reportado reduce el riesgo de cáncer de pulmón en un 51 % e incluso en fumadores un 65 % [60]. La inducción de apoptosis en células cancerígenas durante diferentes etapas del ciclo celular sin afectar las células normales ha sido documentado en varios cánceres *in vivo* e *in vitro* [129, 130]. Por ejemplo, el tratamiento con quercetina condujo a una reducción significativa del tamaño de tumores [131]. Se cree su acción se basa en la inhibición de enzimas quinasas necesarias para el crecimiento del tumor [55]. Por ejemplo se ha propuesto la inhibición del receptor del factor de crecimiento epidérmico es parte del efecto anti-cancerígeno en células colorectales [132]. Asimismo, quercetina induciría la muerte neuronal vía un mecanismo involucrando la inhibición de *Akt/pKB* y *ERK* [39].

Hipótesis

La estructura química de la quercetina y flavonoides relacionados les permite interactuar de forma simultánea con múltiples componentes celulares y subsecuentemente producir una gran diversidad de efectos biológicamente relevantes.

Objetivo general

Predecir y caracterizar *in silico* potenciales blancos de acción molecular de quercetina y flavonoides similares.

Objetivos específicos

Desarrollar una estrategia de tamizaje reverso *in silico* para la identificación de proteínas blanco de quercetina y flavonoides similares.

Evaluar el vínculo entre el perfil de proteínas blanco y los efectos biológicos de quercetina.

Evaluar los requerimientos estructurales para la interacción de quercetina y flavonoides similares con sus proteínas blanco e identificar un posible patrón común.

Proponer un mecanismo de acción de quercetina.

Metodología

- Tamizaje virtual

El tamizaje o *virtual screening* es una aplicación de los métodos *in silico* para seleccionar compuestos prometedores de bases de datos químicas [133]. Puede ser considerado como la contraparte computacional de los métodos de evaluación biológica, tales como el high-throughput screening (HTS) [134]. En el descubrimiento de drogas, el uso de librerías de compuestos extensas y diversas para tamizaje computacional y biológico es una de las estrategias más extendidas [135]. Esto ha estimulado del uso del tamizaje virtual como un método efectivo para la evaluación de una variedad de colecciones de compuestos. Usualmente, el tamizaje virtual cae en dos tipos distintos: basados en ligando y basados en la estructura. [136, 137]

El tamizaje basado en el ligando recae en la exploración de descriptores moleculares tomados de compuestos con actividad conocida [138, 139]. En general, un conjunto de características mutuas de una serie de compuestos es identificada, y luego aplicados como filtros moleculares. Estos métodos de filtrado de bases de datos son usados para seleccionar compuestos para evaluación experimental y reducir el espacio químico a ser explorado en siguientes etapas. Varios programas permiten estimar descriptores moleculares para filtrar una base de datos. Estos programas son útiles en predecir importantes propiedades relacionadas a la actividad biológica, tales como solubilidad, estado de protonación y volumen molecular. [137]

Otra estrategia de tamizaje basado en el ligando es el uso de características estructurales tomadas de ligandos conocidos para generar modelos farmacofóricos. Estos modelos tridimensionales consisten en la compilación de propiedades estructurales consideradas relevantes para la actividad biológica. La generación de modelos farmacofóricos involucra los siguientes pasos típicos: (i) explorar el espacio conformacional de una serie de compuestos; (ii) identificar propiedades recíprocas; (iii) alinear los modelos de acuerdo a las propiedades identificadas; y (iv) generando el modelo farmacofórico. Un problema crítico que necesita ser resuelto por los algoritmos actuales de construcción de farmacóforo es el desarrollo de una herramienta efectiva de alineamiento molecular. [137]

En el tamizaje basado en la estructura, una base de compuestos es anclada en un sitio blanco seleccionado. Junto con la predicción del modo de unión, el tamizaje basado en la estructura provee de un ordenamiento de las moléculas ancladas. Esta jerarquía puede ser usada como único criterio para la selección de moléculas prometedoras, o puede ser combinada con otros métodos de evaluación. Los compuestos seleccionados pueden evaluarse experimentalmente para determinar su actividad biológica en el blanco investigado. En general el tamizaje basado en la estructura

consiste en los siguientes pasos: (i) preparación del blanco molecular; (ii) selección de la base de compuestos; (iii) *docking molecular*; y (iv) análisis post-anclaje. Una revisión rigurosa de la información disponible el blanco y ligandos conocidos, así como un análisis cuidadoso de las ventajas y errores en la selección de algoritmos de *docking*, es requerida para delinear las estrategias más apropiadas. [137]

Generalmente es el caso de que varias estructuras de un receptor dado están disponibles. Si el blanco se encuentra tanto en forma *apo* como *holo*, ambas debería ser consideradas en una estrategia de tamizaje virtual. Los cambios conformacionales, resultando de la interacción con ligandos, y la resolución estructural son detalles importantes a ser considerados en la selección de la estructura más adecuada. Luego, la estructura seleccionada sufre varios procedimientos para prepararla adecuadamente para los estudios de *docking*. En breve, la rutina de preparación consiste en agregar átomos de hidrógeno, eliminar moléculas de agua (con excepción de aquellas que se encuentren mediando interacciones importantes), especificando la protonación correcta y estados de tautomerización de los residuos del sitio de unión, y calculando las cargas parciales. [137, 138]

Otro paso importante es la preparación de una librería de compuestos. Las bases de datos que recopilan estructuras químicas son ampliamente usadas y comúnmente permiten la búsqueda y selección de compuestos de acuerdo a filtros químicos. Usualmente las bases de compuestos se guardan como anotaciones de línea, tales como *SMILES*, *SMARTS* e *InChI*, que son luego convertidos en estructuras tridimensionales. Convertir los archivos originales requiere la correcta asignación de la estereoquímica, cargas parciales y estados de ionización. Algunos ejemplos de base de datos comúnmente usadas en tamizaje virtual son *Zinc*, *PubChem*, *ChemSpider*, *ChEMBL*, *NuBBE*, *ChemBank*, *eMolecules*, *DrugBank*, *Binding DB*. [137, 140]

En la siguiente etapa, la base preparada es anclada en un sitio de unión. El algoritmo de búsqueda conformacional explora la superficie de energía potencial de cada molécula y los compuestos mejor posicionados son seleccionados como potenciales ligandos. Un tamizaje virtual involucra cientos de miles (o millones) de compuestos, con análisis post-anclaje para decidir que compuestos priorizar. Visualizar los complejos ligando-receptor predichos es muy útil para este propósito, permitiendo análisis de aspectos críticos, tales como la presencia de interacciones intermoleculares específicas. Otra faceta que puede evaluarse a través de la visualización de las estructuras predichas es si las soluciones se ajustan a requerimientos previamente determinados como una hipótesis de farmacóforo. [137]

- **Tamizaje reverso**

Las estrategias *in silico* para la predicción de blancos de acción pueden considerarse como un caso particular del tamizaje virtual. En este sentido, las herramientas

actuales pueden basarse en el ligando, en la estructura, utilizar aprendizaje automático y aproximaciones de redes. Por ejemplo en el primer caso, al buscar potenciales proteínas blanco de un ligando problema (*query*) este es comparado contra una librería de compuestos con interacción conocida [141]. Alternativamente la comparación de sitios de unión recurre a la similitud estructural con proteínas blanco conocidas. El anclaje inverso es otra estrategia basada en la estructura, que al contrario del anclaje directo clásicamente utilizado busca y compara una librería de proteínas en lugar de una librería de ligandos.

El presente trabajo siguió un esquema de tamizaje jerárquico, que toma ventaja de las herramientas más rápidas (similitud del ligando y sitio de unión) en sus primeras etapas para reducir el espacio de búsqueda y el número de falsos positivos previo al uso de una herramientas más costosa computacionalmente como el anclaje inverso [142]. Para cada etapa de la búsqueda se tomaron tres de los más recientes programas disponibles para la predicción *in silico* de proteínas blanco (figura 4) incluyendo a *LIBRA* (*LIGand Binding site Recognition Application*) desarrollado por el grupo del Dr. Fabio Polticelli [4].

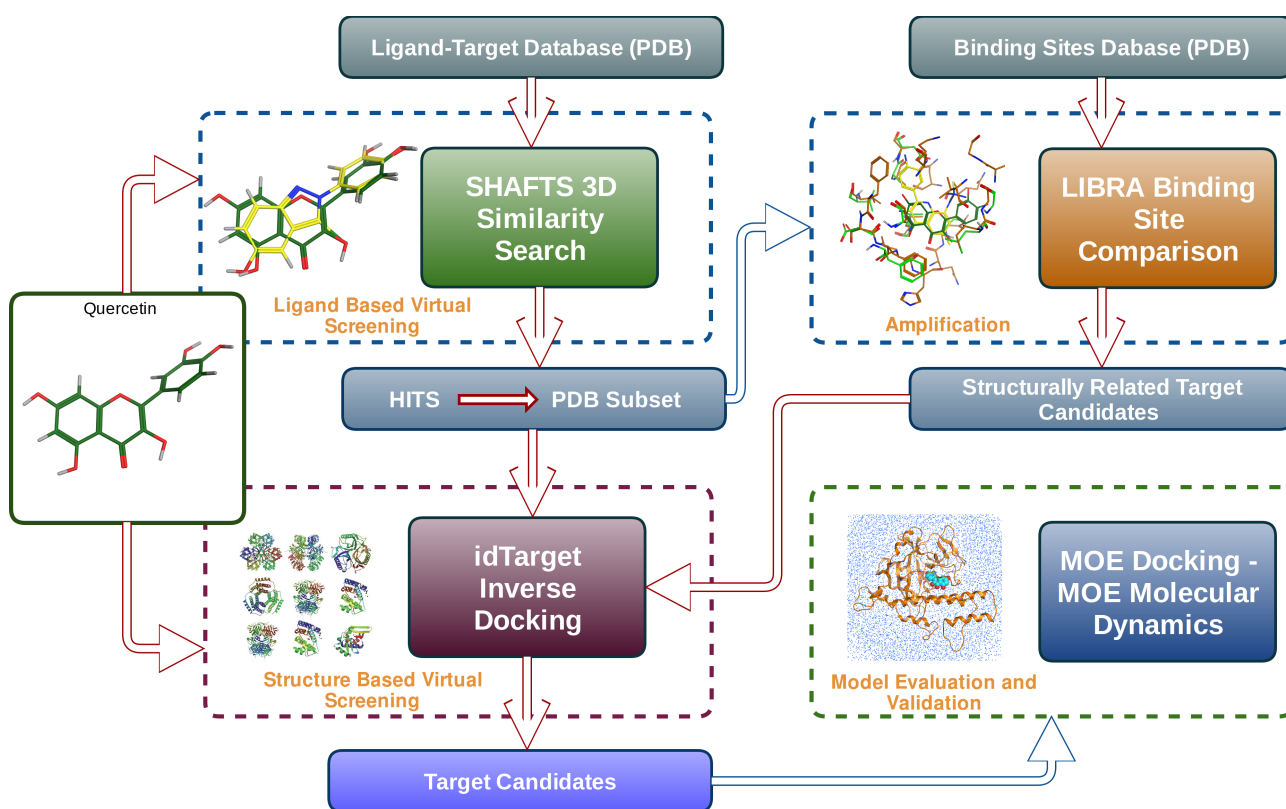


Figura 4. Proceso de tamizaje reverso. La estrategia involucra una búsqueda de similitud con el algoritmo *SHAFTS* (*SHApe Feature Similarity*) de ligandos co-cristalizados del *Protein Data Bank* (*PDB*). Una etapa de amplificación es agregada para identificar proteínas relacionadas por comparación de sitios de unión de *LIBRA* (*LIGand Binding site Recognition Application*). Finalmente, la proteínas candidatos son sometidas a *docking inverso* con el servidor *idTarget* (<http://idtarget.rcas.sinica.edu.tw/>).

◦ *Librería de ligandos asociados a proteínas blanco*

Una librería de ligandos contenidas en el *Protein Data Bank* fue desarrollada para el tamizaje basado en el ligando [143]. Primero, todos los ligandos contenidos en el *PDB* fueron descargados desde la versión 2014 de *LigandExpo*, un repositorio de estructuras co-cristalizadas y hetero-átomos en general asociada a un índice de proteínas [144]. *LigandExpo* acepta tanto consultas basadas en palabras claves (código *PDB* del ligando, nombre del compuesto y fórmula química) como en estructura química permitiendo pareos parciales.

Las estructuras moleculares fueron descargadas en formato *SMILES-stereo* y reconstruidas en el programa *MOE (Molecular Operating Environment)* [145]. Este formato de archivo resulta conveniente al estar correctamente indexado, y contener la totalidad de la estructura de la molécula incluyendo la isomería en contraposición de los registros cristalográficos en formato *.sdf* donde existen múltiples entradas redundantes correspondientes a moléculas parcialmente cristalizadas. Estas estructuras redundantes además de aumentar el tamaño de la base (aproximadamente 250 mil entradas) y con ello el costo computacional pueden generar resultados artificiales en el alineamiento.

Iniciando en 17969 ligandos del *PDB* se aplicó una serie de filtros en *MOE* con el objetivo de corregir la hidrogenación, remover ligandos de bajo peso molecular (sales, agua) y posibles entradas repetidas, remover moléculas con un peso molecular inferior al esqueleto hidrocarbonado de un flavonoide y seleccionar moléculas con propiedades tipo droga de acuerdo a las reglas de *Oprea*. Las reglas de *Oprea* son una versión refinada de la clásica regla de 5 de *Lipinski* para determinar la “*drogabilidad*” de una molécula [146]. Convenientemente estos filtros se encuentran implementados en *MOE* mediante herramientas de línea de comando (*sdfilter*, *sdsort*, y *sdwash*).

Asimismo se removieron moléculas conteniendo Se o As al no estar parametrizadas en el campo de fuerza *TAFF*. Las moléculas obtenidas fueron optimizadas en *MOE* con el campo de fuerza *Merck (MMFF91x)* hasta un gradiente de 0.01 kcal/mol y guardadas en formato *mol2*. La librería final contuvo 9924 moléculas distintas.

◦ *Tamizaje de quercetina basado en el ligando*

Los métodos de tamizaje reverso basados en el ligando toman el concepto de que compuestos estructuralmente similares poseen propiedades fisicoquímicas similares y posiblemente perfiles biológicos similares [147]. La esencia de los distintos métodos es entonces el criterio de similitud estructural utilizado. Por ejemplo las moléculas pueden ser representadas con descriptores (*chemical fingerprints*) y luego ser comparadas según su coeficiente de Tanimoto. Existen métodos que consideran

explícitamente las características tridimensionales de las moléculas o de su interacción con el blanco por medio de farmacóforos. Algunos criterios estadísticos también pueden introducirse para completar la selección de ligandos.

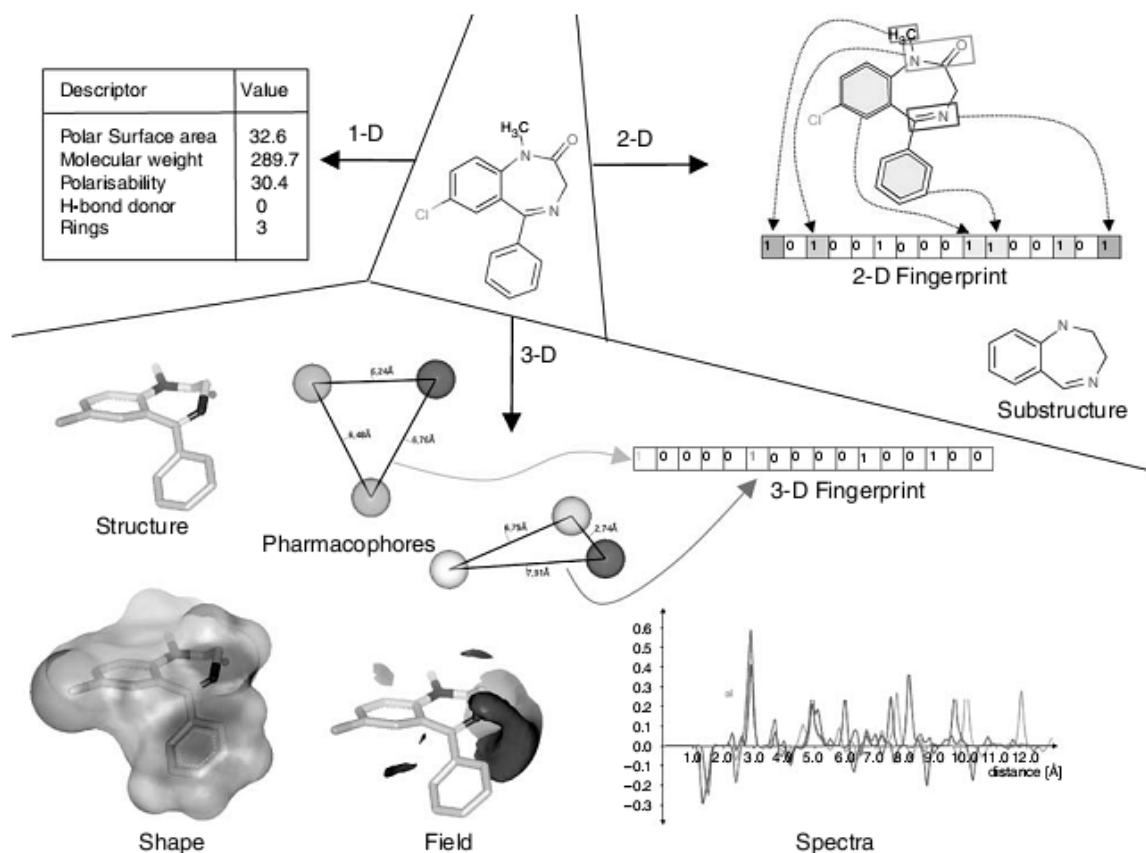


Figure 1 Examples of molecular descriptors for small-molecule-weight ligands

Figura 5. Ejemplos de algunos descriptores para moléculas pequeñas. Propiedades fisicoquímicas, forma, conectividad, elementos estructurales bi y tridimensionales, farmacóforos, campo electrostático, espectro, etc.

Algunas de las herramientas más conocidas de tamizaje reverso basado en el ligando incluyen *ReverseScreen3D*, *ChemMapper* y *TargetHunter*. *ReverseScreen3D* es un servidor de tamizaje reverso basado en la similitud bidimensional (por coeficiente de *Tanimoto*) y tridimensional entre un ligando propuesto y un conjunto de ligandos presentes en el *Protein Data Bank* [43]. Los ligandos se obtienen de estructuras proteicas previamente segregadas según su similitud de secuencia primaria con el objetivo de evitar la redundancia y sobre representación de proteína blanco. Al momento de su publicación la base contenía aproximadamente 8000 ligandos correspondientes a 6000 complejos proteicos, que se actualizan periódicamente. En un ensayo con ligando *4H-tamoxifeno* los autores recuperan 9 de 14 proteínas blanco conocidas en la franja del 5 % de candidatos mejor posicionados. *ChemMapper* es un servidor de tamizaje reverso que implementa el algoritmo *SHAFTS* de comparación estructural que se detallará a continuación. En adición al *Protein Data Bank* integra otras bases de actividad como *DrugBank* y *KEGG* (*Kyoto Encyclopedia of Genes and Genomes*). *TargetHunter* es otro servidor de tamizaje reverso que a diferencia de los

casos anteriores no recurre al *Protein Data Bank* por estructuras cristalográficas resueltas sino información de bases de actividad como *ChEMBL*, una de las bases de datos manualmente curada más grande de moléculas bioactivas y alrededor de 8000 blancos de acción. *TargetHunter* predice los blancos de acción para un compuesto en base al algoritmo *TAMOSIC*, (“*Targets Associated with its MOst Similar Counterparts*”) y al algoritmo *MCM* (“*Multiple-Category Model*”) e introduce además un criterio estadístico para la selección de los blancos [148, 149]. Otras herramientas utilizadas para tamizaje reverso basado en la similitud del ligando incluyen *HitPick* (<http://mips.helmholtz-muenchen.de/proj/hitpick>), *MedChem Studio*, *Schrödinger* y más recientemente *SEA* (*Similarity Ensemble Approach*) [150–153].

En este trabajo tomando quercetina como anzuelo buscamos dentro de la librería de ligandos por compuestos similares co-cristalizados en el *PDB*. La estructura molecular de quercetina fue construida en *MOE*, optimizada (*MMFF94x*, gradiente 0.01 kcal/mol) y guardada en formato mol2 para el tamizaje. El algoritmo *SHAFTS* (*SHApe-Feature Similarity*) implementado en su versión local (*stand-alone*) fue usado para tamizar quercetina contra los conformeros obtenidos de la librería de ligandos asociados a proteínas blanco. *SHAFTS* es un algoritmo para cálculo de similitud estructural y tamizaje virtual basado en el ligando [3, 154]. Ha demostrado un enriquecimiento satisfactorio de compuestos activos contra varias quinasas representativas en estudios de retrospectivos de tamizaje virtual [155–157]. Adopta una métrica de similitud que considera tanto la forma molecular como las características farmacofóricas (centro hidrofóbico, centro cargado positivamente o negativamente, donador o aceptor de enlace de hidrógeno y ciclos aromáticos). Un método de partición (*hashing*) de tripletes de características es usado para el rápido alineamiento de la estructura molecular. En primer lugar *SHAFTS* enumera todos los posibles tripletes de características de la molécula de consulta y de las conformaciones de moléculas blanco en la base de datos. Estos son guardados en una tabla de partición (*hashing*) donde la clave de búsqueda son el tipo de vértice y distancia entre los nodos de los tripletes. Todos los tripletes de la molécula de consulta son usados para comparar uno a uno los tripletes en la tabla de *hashing* de las moléculas blanco, y aquellas con la misma clave son definidas como un posible alineamiento. El puntaje de alineamiento estructural posee una componente asociada a la similitud geométrica (*shape score*) y una componente asociada a la similitud farmacofórica (*feature score*). El *shape score* se calcula como el solapamiento normalizado de la forma de las moléculas expresado como la suma del solapamiento de funciones gaussianas correspondientes a densidades atómicas individuales, de manera semejante al *score* de *docking ASE* de *MOE*:

$$V_{AB} = \sum_{i \in A} \sum_{j \in B} \int d\vec{r} \rho_i(\vec{r}) \rho_j(\vec{r})$$

, donde *i* y *j* iteran sobre todos los átomos A y B respectivamente, y *d_{ij}* es la distancia interatómica entre los átomos *i* y *j*. El *score* final es normalizado al rango [0,1] usando la métrica de coseno:

$$ShapeScore = \frac{V_{AB}}{\sqrt{V_A V_B}}$$

El *feature score* es calculado a partir del solapamiento de características farmacofóricas del mismo tipo. Una función exponencial se usa para normalizar entre [0,1] el solapamiento entre pares de características:

$$F_{AB} = \sum_{f \in F} \sum_{i \in A} \sum_{j \in B} \exp \left[-2.5 \left(\frac{d_{ij}}{R_f} \right)^2 \right]$$

, donde i y j iteran sobre todos los puntos farmacofóricos del mismo tipo en A y B respectivamente, d_{ij} es la distancia entre los puntos i y j , y R_f es un factor de tolerancia del solapamiento cuyo valor por defecto es 0.8 Å. De igual manera que *shape score*, el *feature score* final es normalizado a [0,1] usando la métrica de similitud del coseno:

$$FeatureScore = \frac{F_{AB}}{\sqrt{F_A F_B}}$$

El *score* final es la suma ponderada al rango [0,1] de *feature score* y *shape score* según el coeficiente de peso w con valor por defecto 1.0: $score = (shape\ score) + w*(feature\ score)$. Finalmente *SHAFTS* retorna una lista ordenada de identificadores moleculares asociados (*hits*) a un puntaje (*score*) de similitud contra la molécula de interés (*query*).

Debido a que *SHAFTS* utiliza una estrategia semirrígida para el alineamiento estructural, se realizó un análisis conformacional con el programa *Cindy* para generar un máximo de 100 conformeros de baja energía por molécula en la librería de ligandos asociados a proteínas del *PDB* [158]. Esta herramienta consiste de un algoritmo genético, en el que los conformeros son codificados a través de sus ángulos diedros. La población de conformeros evoluciona buscando generar conformeros energéticamente favorables y lo más diversos posibles para lo que se tiene en cuenta la desviación geométrica respecto de la conformación de partida.

SHAFTS fue luego implementado con sus condiciones por defecto. En principio, el algoritmo no distribuye el cálculo en todos los procesadores disponibles de una computadora, por lo que para acelerar el cálculo se subdividió la base de ligandos. Se ejecutaron búsquedas independientes que luego fueron integradas. Finalmente se recuperaron los identificadores correspondientes a proteínas asociadas a ligandos con un *score* de similitud con quercetina de al menos 1.5 (máximo 2.0) recurriendo al índice de ligandos contenidos en *LigandExpo* y rutinas (*scripts*) de *Bash* y *Python*. Depende fuertemente del tamaño del ligando.

A efectos comparativos esta etapa del tamizaje también fue realizada usando el servidor *ChemMapper*. Este servidor permite recurrir a otras bases de estructura como *DrugBank* además del *PDB*. En esta búsqueda se tomaron tres franjas de *score* definidas por el servidor, 0.8, 1.2 y 1.5.

- *Tamizaje de proteínas blanco por similitud de sitio de unión*

Delinear semejanzas en el sitio de unión de proteínas es otra posible ruta para encontrar nuevas proteínas blanco. Hasta el momento este ha sido el enfoque menos explotado. Se destaca por ejemplo la experiencia de Xie y col. del 2009 para explicar los efectos adversos de inhibidores de la proteína de transferencia de esteroides de colesterol (*CETP*) [159]. En una primera etapa caracterizan los sitios de unión de proteínas blanco conocidas partiendo de cristales o modelos por homología e identifican sitios semejantes con la herramienta *SOIPPA* (“*Profile-Profile Alignment*”) [160, 161]. En una siguiente etapa completan el tamizaje realizando *docking inverso* con *Surflex* y *eHits* y posterior optimización en *Autodock*. Se destaca también el uso de ligandos señuelos (*decoys*) obtenidos de las base de estructuras *ZINC* para normalizar el *score* de *docking*.

Una de las herramientas más conocidas es *PharmMapper*, un servidor de tamizaje reverso basado en el estudio de la similitud entre un ligando dado y un conjunto de modelos farmacofóricos obtenidos de las siguientes bases de datos *TargetBank*, *BindingDB*, *DrugBank* y *PDTD* [162]. Se ideó como una herramienta complementaria al *docking inverso*. Al momento de su publicación la base contenía aproximadamente 7000 farmacóforos basados en el receptor correspondientes a 1500 complejos ligando-proteína que se han venido actualizando con las estructuras más recientes. Los modelos farmacofóricos son generados mediante *LigandScout*, software que permite la extracción de características tridimensionales a partir de datos de estructura de complejos proteína–ligando [163]. Seis características farmacofóricas se tuvieron en cuenta en el proceso: centro hidrofóbico (H), centro cargado positivamente (P), centro cargado negativamente (N), aceptor de enlace de hidrógeno (HBA), dador de enlaces de hidrógeno (HBD), centro aromático (AR) y opcionalmente centro de interacción metálico (M). Todos los ligandos con peso molecular menor a 100, así como co-factores con peso molecular mayor a 600 son considerados “átomos del entorno” y no ligandos. *PharmMapper* realiza un alineamiento flexible del ligando problema contra cada farmacóforo de la base, reportando los blancos de acuerdo a un *score* de ajuste.

Asimismo una de las herramientas para la comparación e identificación de sitios de unión es *LIBRA* (*Ligand Binding site Recognition Application*). Este programa identifica semejanzas locales entre una estructura proteica y una colección de sitios funcionales y su entorno. Dada una proteína problema (*input*) *P*, *LIBRA* detecta el subconjunto más extenso de residuos similares entre la proteína *P* y no solamente los sitios funcionales conocidos sino también su ambiente local. A efectos de la predicción de sitios activos, esta aproximación requiere de la construcción de una base de datos de sitios conocidos, comenzando por el *CSA* (*Catalytic Site Atlas*) [164]. Para la predicción de sitios de unión de ligando, también fue generada una

base de datos extrayendo los residuos del entorno de ligando de 75460 estructuras de complejos proteicos depositados en el *PDB* (para Junio del 2015). La base final contiene 173240 sitios de unión en formato *PDB*. La similitud entre el subconjunto de residuos pertenecientes a la proteína problema y un sitio de unión junto con su entorno son definidos en *LIBRA* por un grafo formado por nodos y aristas de la siguiente manera: un residuo R_1 , perteneciente a la proteína problema, y un residuo R_2 perteneciente a la proteína conocida, forman un nodo si son idénticos o equivalentes, según una tabla de equivalencia. Si el nodo A es formado por pares de residuos (A_1, A_2) , el nodo B por el par (B_1, B_2) , d_1 es la distancia (A_1, B_1) , d_2 es la distancia (A_2, B_2) y δ es la distancia de error límite, existe entonces una arista que conecta A y B si $|d_1 - d_2| < \delta$. Una vez las relaciones entre la proteína problema y el sitio de unión conocido son definidas como un grafo, un subconjunto de residuos equivalentes es definidos como el conjunto de nodos tales que dos nodos cualquiera están conectados por una arista. La ventaja de esta aproximación es que permite implementar un algoritmo exacto para identificar el subconjunto más grande de residuos equivalentes.

En el presente trabajo se comparó el sitio de unión del ligando co-cristalizado de las proteínas seleccionadas mediante *SHAFTS* contra un conjunto representativo de proteínas contenido en el *PDB* usando *LIBRA* [4]. Específicamente se tomó para cada proteína identificada por *SHAFTS* en la etapa anterior se el entorno a 4.5 Å del ligando co-cristalizado como sitio de unión y se guardó en formato *PDB* previo al tamizaje en *LIBRA*. Los sitios similares fueron seleccionados de acuerdo a estos requerimientos: 100 % de residuos similares (de acuerdo a la matriz *blosum62*), motivo mínimo de 5 residuos, máximo *RMSD* de alineamiento estructural de 1 Å (Ångström) y ningún conflicto estérico entre el ligando de la proteína de prueba y la nueva proteína.

Y.Z. Chen y D.G. Zhi de la Universidad de Singapur fueron los primeros en consolidar un método de tamizaje reverso (del ligando a las proteínas) basado en el anclaje molecular (*docking*) del ligando: *INVDOCK* [42]. El método se implementó originalmente en una base de datos de cavidades proteicas del *Protein Data Bank* versión del 2001, alcanzando 2700 entradas en total. Utilizaron una versión de *docking* flexible con confórmeros de ligando y permitiendo rotar la cadena lateral de los residuos del sitio pero sin refinar la pose final. La función de *score* se basó en campo de fuerza *Amber*. Los candidatos a blanco son evaluados de acuerdo al *score* de *docking* y cuando es posible comparando contra otros ligandos que interaccionen con el mismo sitio. Del total de blancos identificados en los primeros ensayos con las moléculas 4H-tamoxifeno y vitamina E, un 50 % poseían evidencias experimentales. En trabajos recientes *INVDOCK* has sido empleado para identificar blancos de compuestos alcaloides entre otros [165]. Otros desarrollos de *docking inverso* incluyen los de Lauro y col. del 2011 con *AutoDock Vina*, Yang y col. del 2011 con *DOCK*, Y.Y. Li y col. del 2011 con *ICM*, Fan y col. del 2012 con *CDOCKER*, Lee y Kim del 2012 con *GOLD*, y especialmente el trabajo de Kun-Yi Hsin y col. del 2013

en el que desarrollan un método de *scoring* (puntaje) apoyado en aprendizaje automático (*machine learning*) [166–171].

◦ *Tamizaje basado en la estructura, docking inverso*

Alternativamente se han desarrollado algunos servidores *online* de *docking inverso* como *TarFisDock* y *idTarget*. *TarFisDock* realiza *docking inverso* utilizando el software *DOCK* contra una base de proteínas con valor terapéutico contenidas en la base *PDTD*. Al momento de su publicación la base contenía alrededor de 700 estructuras, superando las 1000 entradas para el 2013. La única fuente de flexibilidad considerada en el *docking* es el uso de estructuras redundantes para ciertas proteínas. Los candidatos son evaluados únicamente por el *score* de *docking* que considera únicamente interacciones de van der Waals y electrostáticas. En un estudio con los mismos ligandos empleados por *INVDOCK*, *4H-tamoxifeno* y *vitamina E* se evidencia que la franja del 10 % de blancos mejor valorados contiene el 50 % de las proteínas apoyadas experimentalmente. Tal como señalan estos autores, estos resultados están limitados por la función de *score* utilizada y la base de datos acotada. En una aplicación reciente *TarFisDock* es utilizado junto con *AutoDock* en la búsqueda de blancos para compuestos aromáticos [157, 172]. Song Ke y cols. utilizaron *TarFisDock* para generar un grupo preliminar de blancos *aril-aminopiridinas*, mejoran el anclaje con *GridDock* (basado en *AutoDock4*) y priorizan un nuevo conjunto de blancos de acuerdo a su correlación con la actividad anti-cancerígena [173].

Una de las herramientas más reciente de *docking inverso* es el servidor *idTarget*. En el presente trabajo la estructura optimizada de quercetina fue ingresada al servidor de *docking inverso idTarget* para el tamizaje de la lista de proteínas seleccionadas en las etapas precedentes. *IdTarget* realiza el *docking inverso* para una molécula de prueba contra todo el *PDB* o una lista acotada de estructuras [5]. Aplica la herramienta de *docking* *MEDock* junto a un algoritmo de tipo “*divide y conquistaras*” para analizar la superficie proteica en busca de posibles sitios de unión y la función de *score* *AutoDock4* optimizada para la predicción de afinidad de unión. Se trata de una función de *score* empírica en la que se describen los siguientes términos: interacción de *van der Waals*, interacción por enlace de hidrógeno, interacción electrostática, energía de desolvatación, y entropía torsional. La energía libre se predice según la siguiente formula:

$$\begin{aligned}
\Delta G_{\text{bind}} = & W_{\text{vdW}} \times \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{\text{H-bond}} \\
& \times \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{\text{estat}} \times \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \\
& + W_{\text{desol}} \times \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2/2\sigma^2)} \\
& + W_{\text{tor}} \times N_{\text{tors}}
\end{aligned}$$

Los coeficientes de peso W_i se derivan por regresión contra afinidades de unión reportadas en el *Ligand Protein Database (LPPDB)*. La energía de van der Waals es de la forma típica 6-12, donde los parámetros A_{ij} y B_{ij} son adoptados del campo de fuerza *Amber84*. El término de enlace de hidrógeno se basa en el potencial 12-10, ponderado por un término direccional $E(t)$. La interacción electrostática se calcula en base al potencial de Coulomb. El término de desolvatación se incluye calculando el volumen del entorno de a un átomo (V_i) ponderado por un parámetro de solvatación atómica (S_i) y un término exponencial con un factor de distancia σ . El término final representa la entropía torsional calculada simplemente como el número de enlaces rotables del ligando.

Una proteína blanco i para un ligando j es seleccionada considerando el perfil de afinidad de la proteína. Un *Z-score* es calculado de acuerdo a:

$$Z_{ij} = \frac{E_{ij} - E_i}{sd_i}$$

, donde E_{ij} es el *score* de *docking* del ligando j en el sitio i , E_i y sd_i son el centro y amplitud del perfil de afinidad de la proteína i .

Por defecto la búsqueda se realiza de manera no exhaustiva organizando el *PDB* en grupos homólogos. En contraposición se seleccionó el modo de búsqueda exhaustivo (*scanning mode*) del servidor restringiendo el tamizaje a las estructuras identificadas en la etapa anterior. Las proteínas a priorizarse como posibles blancos del ligando son aquellas con una energía de unión menor a -9.0 kcal/mol (250 nM de afinidad predicha) y un *Z-score* de 1.0 o menor de acuerdo al trabajo de Nikolić y col [174].

- **Anotación funcional de proteínas blanco, minería de texto**

El conjunto final de proteínas blanco fue anotado de acuerdo a su estructura, actividad y participación en enfermedades. El homólogo humano correspondiente fue considerado en proteínas pertenecientes a otros mamíferos. La información fue obtenida de diferentes fuentes incluyendo repositorios y herramientas de libre acceso

como *PDB*, *Uniprot*, *DAVID* (términos de ontología e información de la literatura), *GO TermMapper* (términos de ontología) y *FlyMine* (ontología y literatura) que brindan información complementaria [175–180]. Entre otros datos estas bases proporcionan términos de ontología que definen las propiedades de los productos génicos según su localización subcelular, función molecular y proceso biológico. En adición los blancos moleculares identificados por diferentes estrategias pueden ubicarse en redes o vías específicas asociadas a enfermedades. Se tiene entonces la base *KEGG*, que reúne datos sobre genomas, vías de señalización y metabolismos, enfermedades, drogas y otras sustancias.

El análisis automático de textos científicos también puede brindar información relevante al estudio masivo de proteínas como es el caso del tamizaje reverso o resultados de proyectos de secuenciación masiva. Cada año cientos de miles de artículos científicos son publicados en diversas áreas, una cantidad de información que vuelve difícil para un individuo derivar conclusiones que entrecrucen varios campos de estudio. Con el objetivo de extraer conocimiento del texto de manera masiva se han venido desarrollando varias estrategias que involucran alguna o varias de las siguientes etapas: i) extracción de información (*information extraction*), ii) reconocimiento de entidades (*named entity recognition*), iii) normalización (*normalization*) y iv) extracción de relaciones (*relation extraction*). Para la extracción de información se cuentan con diversos repositorios, por ejemplo el más conocido: *PubMed*, el servicio de la *National Library of Medicine* que incluye millones de resúmenes y artículos científicos completos. La información está disponible en varios formatos, tales como *.xml* o el formato más legible denominado *Medline*. A continuación el reconocimiento de entidades en el texto como organismos, macromoléculas, o ligandos es el paso más complejo dadas las características peculiares de los textos científicos. Algunas de las dificultades más frecuentes comienzan con el uso de sinónimos, por ejemplo para el caso de la proteína p53 la misma puede ocurrir también como *TP53*, *Hs.1845* y *TRP53*. Asimismo puede ser difícil distinguir cuando se hace referencia al gen, RNA, o la proteína. Los nombres de genes también pueden tener variaciones morfológicas (“*transcription factor*” y “*transcriptional factor*”), ortográficas (“*NF kappa B*” y “*NF-κB*”), combinatorias (“*homologue of actin*”, “*actin homologue*”), flexiva (“*antibody*”, “*antibodies*”). Algunos nombres se solapan incluso con el uso corriente del lenguaje como es el caso de ciertos genes de *drosophila* (ej. “*breathless*”, “*not*”, “*that*”) o se solapan entre si al derivarse de otras entidades (ej. “*IL-2 receptor*”). El uso de siglas genera problemas adicionales como es el caso de la proteína *Nrf2* que puede corresponder tanto a “*nuclear respiratory factor 2*” como a “*Nuclear factor (erythroid-derived 2)-like 2*”. Existen incluso casos en los que el nombre del gen coincide con el de otra entidad biológica no relacionada (*vertical polysemy*) como es el caso de *CA1* correspondiente tanto a “*Carbonic anhidrase 1*” como al área *CA1* del hipotálamo. Una vez las entidades de interés son reconocidas es necesario “mpearlas” a identificadores canónicos lo que se denomina “normalización”. En el caso de genes y proteínas los

mismos se encuentran en repositorios como *EntrezGene* y *Uniprot*. La normalización es un problema en sí mismo ya que incluye la determinación de la especie de origen [181]. Finalmente se tiene la extracción de relaciones explícitas o implícitas entre las entidades encontradas lo que permitiría generar nuevas hipótesis. En particular esto puede aplicarse a la reconstrucción de vías de señalización intracelular, modelos de patologías, definición de mecanismos de acción de nuevas moléculas o reposicionamiento de drogas conocidas. [182–186]

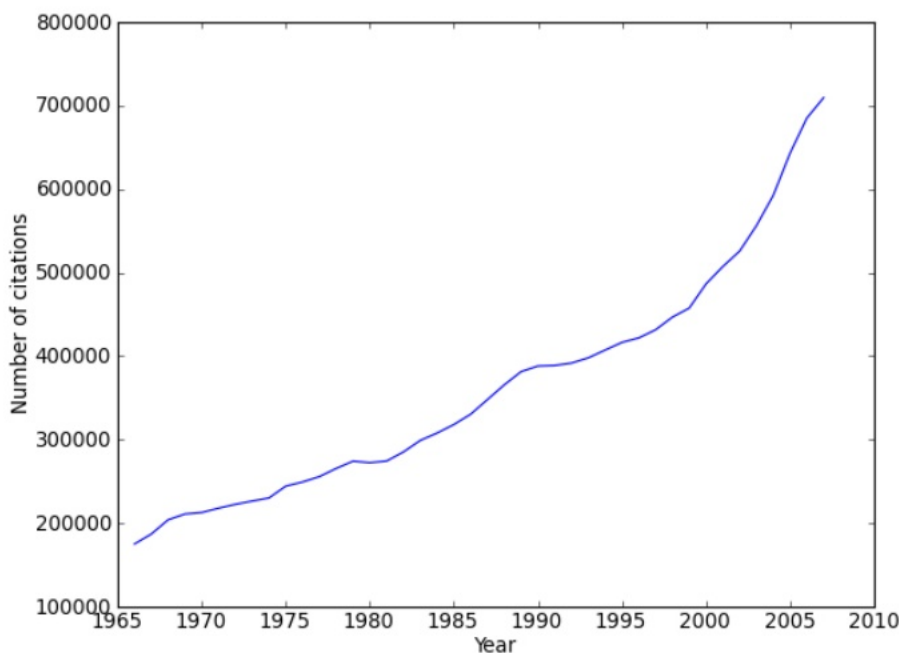


Figura 6. Número de publicaciones por año en la base de datos MEDLINE.

Pocas son las estrategias para demostrar un vínculo entre dos o más entidades biológicas identificadas. La más sencilla es recurrir a la proximidad en el texto, es decir las entidades que ocurren simultáneamente deberían estar vinculadas. En artículos de investigación biológica dos entidades pueden co-ocurrir por varias razones incluyendo funcional (ej. “todas las proteínas que catalizan la oxidación de Xs son Y, Z y V”), física (“todas las proteínas encontradas en la membrana son...”), por síntesis (“todas las proteínas sintetizadas en el...”), evolutiva (“todos los homólogos de X son...”), etc. sin que necesariamente participen del mismo proceso particular. Además una relación dada puede estar siendo negada o ser frecuentemente una especulación. Por estos motivos la co-ocurrencia puede no ser evidencia suficiente de un vínculo sin antes considerar el contexto. Se debe tener en cuenta que la minería de texto es una herramienta que acumula el ruido y sesgo de las fuentes originales. Debido a derechos de autor solo parte de las publicaciones científicas se encuentran disponibles para ser analizada, lo que restringe el trabajo a resúmenes. Existe además un sesgo del interés científico en ciertos tópicos, lo que resulta en una sobre-representación de ciertos artículos y por lo tanto de las entidades que en los mismos se puedan analizar. Sin embargo existen casos de éxito en los que por

ejemplo nuevas hipótesis fueron generadas al combinar información de fuentes en principio no relacionadas. Se tiene por ejemplo el caso de estudio de la enfermedad conocida como “*Complex Regional Pain Syndrome*” (CRPS), un desorden complejo de etiología y patogénesis parcialmente descrita [187]. Hettne y colaboradores identifican mediante minería de texto y análisis de redes de asociación a la proteína *NF-kappaB* como un posible mediador central de la iniciación y progresión de la enfermedad. En su trabajo utilizaron el software comercial *MedScan*, ahora parte de *PathwayStudio* (<https://www.elsevier.com/solutions/pathway-studio-biological-research>) para el procesamiento natural del lenguaje tomando como fuente resúmenes de artículos científicos contenidos en *Medline*.

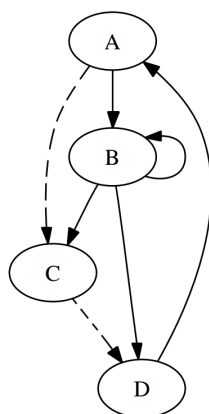


Figura 7. La minería de texto podría evidenciar relaciones “ocultas” (línea punteada) entre dos o más entidades.

En general todas estas estrategias recurren a lo que se denomina “*corpus*”, una colección de documentos que poseen entidades de interés señaladas por un experto humano. Se utilizan como conjunto de entrenamiento y referencia (“*benchmark*”) para comparar herramientas. Se tienen por ejemplo los corpus de *BioCreative*, *NLPBA*, *GENIA*, etc. [188–191]. Las mismas pueden distinguirse en basadas en diccionarios como *BioThesaurus*, *Whatizit* y otros; basadas en reglas y patrones como *ABGENE*; basadas en aprendizaje automático (herramientas de clasificación que utilizan información sintáctica y de contexto) como *NLProt* (implementa un algoritmo de *Support Vector Machines*, <http://cubic.bioc.columbia.edu/services/nlprot/>), basadas en el etiquetado de secuencia (consideran el orden de la secuencia de términos *-tokens-* en una sentencia y predicen la secuencia de marcas *-tags-* más probable) como *ABNER* (implementa un algoritmo de *Conditional Random Fields* <http://pages.cs.wisc.edu/~bsettles/abner/>), *BANNER* (<http://banner.sourceforge.net/>), *JNET*; basadas en Modelos Ocultos de Markov como *GENIA* (<http://www-tsujii.is.s.u-tokyo.ac.jp/GENIA/tagger/>), *OSCAR* (específico para el reconocimiento de entidades químicas, <http://sourceforge.net/projects/oscar3-chem/>) y finalmente estrategias híbridas [182, 184]. El portal de centro de minería de texto de la universidad de Manchester (<http://www.nactem.ac.uk/pathtext/>) y el portal de la Universidad de Jena (<http://julielab.de/JULIE+Lab.html>) ofrecen varias de estas herramientas, aunque su

aplicabilidad puede ser acotada y/o requieran de algún tipo de manipulación adicional para ser funcionales. Por otro lado, actualmente existen algunos servidores *online* dedicados a la minería de textos científicos, como ser: *PolySearch/PolySearch2*, *CoPub*, *PESCADOR*, *GenCLiP*, *PALM-IST*, *SciMiner* y *Slap* [192–199]. La estructura básica de las consultas que se pueden realizar a dichos servidores consisten: ‘dado una entidad X encontrar todas las entidades asociadas Y’s’, donde X e Y pueden ser una patología humana, gen/proteína, droga, metabolito, secuencia, etc. Son entonces estrategias basadas en diccionarios que utilizan una biblioteca de términos biológicos como referencia (“*Thesaurus*”).

Finalmente se destaca la base de datos *STRING* (*Search Tool for the Retrieval of Interacting Genes/Proteins*), que integra información de varias fuentes generando diagramas interactivos de interacción proteína-proteína en los que es posible identificar enriquecimientos funcionales [200]. La información recopilada incluye datos experimentales y predicciones basadas en minería de textos científicos (co-ocurrencia significativas de genes en *PubMed* y otros repositorios), características genómicas, y homología. La información presentada es ponderada en base a un *score* de confianza asociado a cada interacción.

En el presente trabajo se recurrió a los servidores *STRING*, *PolySearch/ PolySearch2*, *CoPub*, *PALM-IST*, *SciMiner* y *Slap* para complementar la anotación de los candidatos identificados en el tamizaje reverso e identificar otras proteínas asociadas en la literatura a los efectos de quercetina.

- **Docking de quercetina y flavonoides análogos**

Una librería de flavonoides análogos a quercetina fue obtenida de *PubChem* mediante una búsqueda de similitud estructural de al menos 98 % de coeficiente de Tanimoto. Las características estructurales de cada molécula son codificadas por *PubChem* originando una huella (*fingerprint*). Existen hasta 881 características estructurales en cada huella, indicando su presencia o ausencia, de manera que dos moléculas pueden ser compradas según: $Tanimoto = AB/(A + B - AB)$, donde *Tanimoto* es el coeficiente o puntaje de 0 a 1, *A* es el número de características presentes en la huella de *A*, *B* es el número de características presentes en la huella de *B* y *AB* es número de características comunes.

La estructura tridimensional de los compuestos resultantes fueron sometidas a minimización de energía con el campo de fuerza *MMFF94x* y análisis conformacional en *MOE*.

◦ Docking

La unión del ligando es un paso clave en las reacciones enzimáticas y en la modulación de vías de señalización intracelulares. La comprensión detallada de las interacciones entre pequeñas moléculas es por lo tanto esencial en el diseño de moléculas que se apliquen a un amplio rango de patologías como cáncer y enfermedades degenerativas. El anclaje o *docking molecular* una de las herramientas más frecuentemente usadas en el diseño racional de drogas a causa de su habilidad de predecir, con alto grado de fidelidad, la conformación de una molécula pequeña dentro de un sitio de unión apropiado en el blanco y brindar predicciones cuantitativas de la afinidad de unión en el complejo. Estudios de *docking* contra las proteínas blanco se desarrollaron para quercetina y la librería de flavonoides utilizando *MOE*. La estrategia de *docking* implementada en *MOE* involucra las siguientes etapas principales: (i) preparación del receptor (incluyendo la definición del sitio), (ii) exploración y (iii) posicionamiento de un extenso espacio conformacional representado varios modos de unión potenciales; (iv) predicción de la energía de interacción asociada con cada modo de unión (ver figura 8). [201]

Preparación del receptor: Un procedimiento de *docking* requiere de una estructura macromolecular de partida, típicamente datos cristalográficos obtenidos del *Protein Data Bank*. Los datos estructurales usualmente contienen errores por lo que usualmente se introducen etapas preparativas en las que se corrigen los segmentos terminales, se completan residuos o átomos ausentes y se elige posiciones alternas apropiadas para los residuos entre otros. Aunque algunos procedimientos de *docking* no lo requieran puede ser deseable optimizar la red de enlaces de hidrógeno del receptor y calcular cargas si por ejemplo se desea refinar mediante mecánica molecular la orientación final del ligando. Asimismo es recomendable realizar una minimización de energía con restricciones previo al inicio del *docking* para optimizar la geometría del receptor. Un paso adicional que puede tomarse es eliminar co-factores extraños o moléculas de agua libre. Aquellas moléculas del solvente que jueguen un papel estructural dentro de la proteína y co-factores relevantes deberían retenerse para la simulación.

Existen distintas estrategias para definir el sitio de unión del ligando. Típicamente se explora una porción o grilla de la proteína que suele corresponder con el sitio de unión del ligando co-cristalizado o un sitio activo reportado experimentalmente. Adicionalmente existen métodos para predecir sitios de unión sobre la superficie de una proteína como la herramienta *site finder* del paquete *MOE*. *Site finder* es un método geométrico que identifica zonas de la proteína no expuestas al solvente que concentren residuos interactuantes.

Búsqueda conformacional: La principal fuente de flexibilidad de un procedimiento de *docking* es el ligando. Para ello se recurre a un análisis conformacional a efectos de

muestrear estructuras óptimas del ligando, que correspondan con mínimos energéticos en la superficie de energía potencial. En la etapa de búsqueda conformacional, los parámetros estructurales de los ligandos, tales como los grados de libertad torsional, traslacional y rotacional, son modificados incrementalmente. Los algoritmos de búsqueda conformacional desarrollan esta tarea al aplicar métodos sistemáticos y estocásticos. El paquete *MOE* implementa estas dos estrategia y una adicional basada en el comportamiento dinámico del ligando (*LowModeMD*).

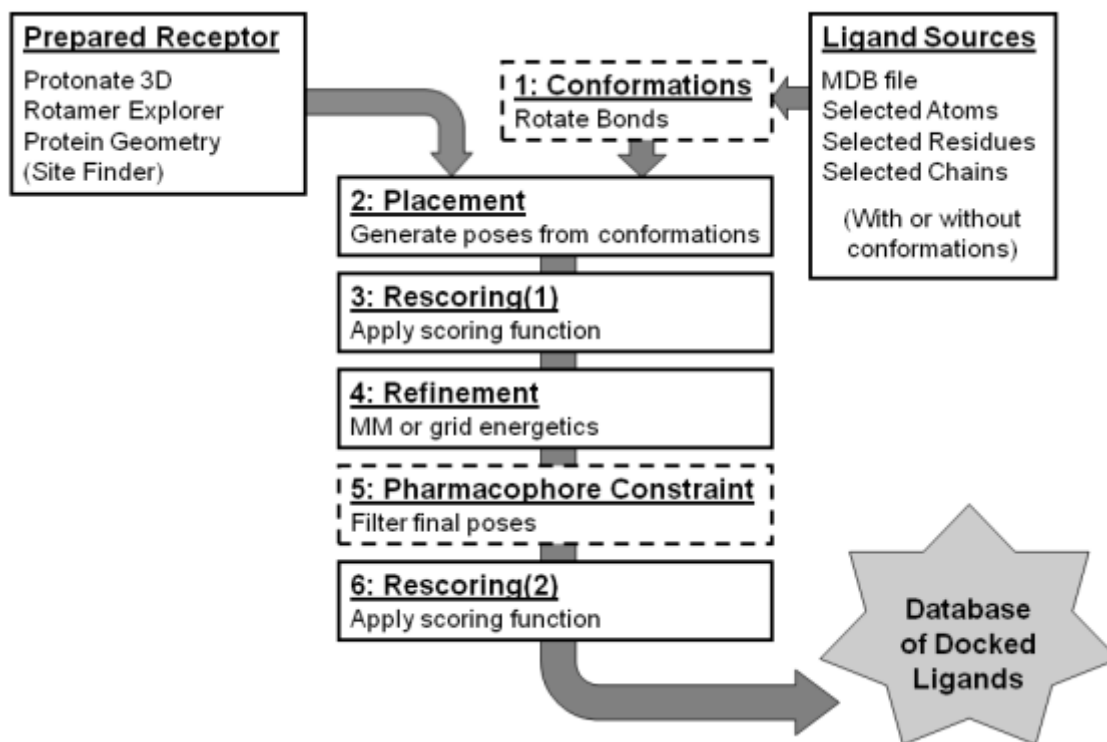


Figura 8. Etapas del algoritmo de *docking MOE* (*Molecular Operating Environment*).

Posicionamiento: Las estrategias de posicionamiento generan poses (modos de unión) a partir de conformaciones del ligando y son características de cada aplicación de *docking* molecular. *MOE* implementa dos estrategias principales. En la primer estrategia, *Alpha PMI*, las poses son generadas al alinear los momentos principales de inercia de un confórmero del ligando dado a un subconjunto aleatorio de alfa esferas en el sitio del receptor. Como se mencionó anteriormente las alfa esferas representan sitios de anclaje del receptor. Este método se ajusta bien a sitios estrechos. En la segunda estrategia, *Alpha Triangle* las poses son generadas por superposición de tripletes de átomos del ligando y tripletes de alfa esferas. En cada iteración se selecciona una conformación al azar. Un triplete aleatorio de átomos del ligando y un triplete aleatorio de alfa esferas son usados para determinar la pose. Un método derivado es *Triangle Matcher* en el que estos tripletes se determinan sistemáticamente. Alternativamente existe la estrategia de posicionamiento por farmacóforo. En este caso son conocidas las características tridimensionales de las

posibles interacciones con el receptor ya sea por una serie de ligandos o estructuras conocidas y se orienta la pose para satisfacer dichas características.

Evaluación de la energía de unión: los programas de anclaje molecular utilizan funciones de puntaje o *score* para estimar la energía de unión de los complejos ligando-proteína predichos. Las funciones de *score* se categorizan en tres grupos principales: basadas en campos de fuerza, empíricas, y basados en el conocimiento. Las funciones de *score* basadas en campos de fuerza estiman la energía de unión al sumar las contribuciones de enlace (estiramiento, flexión, torsión) y no enlazantes (electrostáticas y van der Waals) en una función maestra general. Este tipo de función de *score* aplica un método *ab-initio* para calcular la energía asociada a cada término de la función usando las ecuaciones de mecánica clásica. Las funciones de *score* empíricas son otro tipo de método de evaluación. Cada término de la función describe un evento físico involucrado en la formación del complejo ligando receptor. Esto incluye enlaces de hidrógeno, interacciones iónicas y apolares, así como desolvatación y efectos entrópicos. Como un primer paso en el desarrollo de un función empírica, una serie de complejos proteína-ligando con afinidades de unión conocidas es usado como conjunto de entrenamiento para desarrollar múltiples análisis de regresión. Entonces, las constantes de peso generadas por un modelo estadístico son usadas como coeficientes de ajuste de términos de la ecuación. Una tercera aproximación usada para evaluar la energía de unión ligando-receptor son las funciones de *score* basadas en el conocimiento. Este método usa potenciales de energía pareados de complejos ligando-receptor para obtener una función general. Estos potenciales son construidos al considerar la frecuencia con que dos átomos diferentes son encontrados dentro de una distancia dada en el conjunto de estructuras. Los diferentes tipos de interacciones observados en el conjunto son clasificadas y ponderadas de acuerdo a su frecuencia de ocurrencia. El *score* final es dado a la suma de interacciones individuales.

Varios tipos de funciones estrategias de *scoring* son implementadas en *MOE*. El *score* de *ASE* mide el ajuste geométrico del ligando al sitio. Este *score* es proporcional a la suma de las gaussianas $R_1 R_2 e^{-d^*/d/2}$ sobre todos los pares de átomos ligando-receptor o ligando-alfa esferas (dependiendo de cómo se defina el sitio de unión). R_1 y R_2 son los radios de los átomos en Å (-1.85 para alfa esferas) y d es la distancia entre el par en Å. Se utiliza una constante de proporcionalidad de 0.035 kcal/mol que puede ser ajustada.

El *score Affinity dG* estima la contribución de la entalpía a la energía libre de unión usando una función lineal:

$$G = C_{hb}f_{hb} + C_{ion}f_{ion} + C_{mtig}f_{mtig} + C_{hh}f_{hh} + C_{hp}f_{hp} + C_{aa}f_{aa}$$

, donde los términos f cuentan la fracción de contactos atómicos de distintos tipos y C son coeficientes que ponderan la contribución de cada término a la estimación de la afinidad. Los términos individuales son: hb , considera interacciones de enlace de hidrógeno de manera optimista (por ejemplo se asume que dos grupos hidroxilo interaccionen de la manera más favorable); ion , considera interacciones iónicas mediante un término tipo Coulomb, puede tanto contribuir como desfavorecer la afinidad de unión; $mliq$, de ligación a metales en la que se incluyen interacciones entre nitrógenos/sulfuros y metales de transición; hh , interacciones hidrofóbicas entre átomos de carbono por ejemplo, generalmente favorables; hp , interacciones entre átomos hidrofóbicos y polares, generalmente desfavorables; aa , interacción débil entre dos átomos cualquiera, generalmente favorable.

El *score Alpha HB* es una combinación lineal de dos términos. El primer término mide el ajuste geométrico del ligando al sitio de unión. El segundo término mide los efectos de enlaces de hidrógeno. Ambos términos se suman sobre todos los átomos del ligando. El primer término tiene un componente repulsiva y atractiva. La parte atractiva es sumada sobre todos los átomos del ligando. Cada átomo que se encuentra dentro de 3 Å del centro de una alfa esfera contribuye con $Ae^{-d^*/d/2}$ a este término, donde d es la distancia entre el átomo del ligando y la alfa esfera más próxima y A toma el valor de -0.6845. La componente repulsiva se suma sobre todos los solapamientos atómicos entre átomos del ligando y receptor. Para cada par solapante, la contribución es entre 0 y 1 dependiendo de la severidad del solapamiento. El segundo término mide el efecto de los enlaces de hidrógeno. Para donadores y aceptores de tipo no sp^3 , los enlaces de hidrógeno se proyectan desde el átomo. Si el sitio es ocupado por un átomo favorable, hay un *score* favorable de -2. De lo contrario, si este es ocupado por cualquier otro átomo, hay un *score* de +1. Para donadores y aceptores de tipo sp^3 , todos los átomos favorables dentro de 3.5 Å contribuyen al *score* con -1 mientras que otros átomos contribuyen con +1. Los metales en el receptor son tratados como aceptores pero con efecto tres veces mayor.

La función de *score London dG* estima la energía libre de unión de un ligando a partir de una pose dada. La forma funcional es la suma de términos:

$$\Delta G = c + E_{flex} + \sum_{h-bonds} C_{HB}f_{HB} + \sum_{m-lig} C_Mf_M + \sum_{atoms\ i} \Delta D_i$$

donde c representa el promedio de ganancia/perdida de entropía rotacional y traslacional; E_{flex} es la energía debido a la pérdida de flexibilidad del ligando (calculada desde la topología del ligando únicamente); f_{HB} mide las imperfecciones geométricas de los enlaces de hidrógeno y toma un valor entre [0,1]; C_{HB} es la energía de un enlace de hidrógeno ideal; f_M mide las imperfecciones geométricas de unión a metales y toma un valor de entre [0,1]; C_M es la energía de una unión ideal a metal; y D_i es la energía de desolvatación del átomo i . La diferencia en las energías de desolvatación es calculada de acuerdo a la siguiente fórmula:

$$\Delta D_i = c_i R_i^3 \{ \iiint_{u \notin A \cup B} |u|^{-6} du - \iiint_{u \in B} |u|^{-6} du \}$$

, donde A y B son el volumen de la proteína y/o el ligando con átomo i perteneciente al volumen B ; R_i es el radio de solvatación del átomo i (tomado como el parámetro sigma de van der Waals del campo de fuerza *OPLS-AA* más 0.5 Å); and c_i es el coeficiente de desolvatación del átomo i . Los coeficientes $\{c, c_{HB}, c_M, c_i\}$ se ajustaron a partir de aproximadamente 400 estructuras cristalográficas de complejos ligando-proteína disponibles con datos experimentales de pK_i . Los átomos están categorizados en una docena de tipos distintos para la asignación de coeficientes c_i . Las integrales triples se aproximan usando la formula generalizada de *Born*.

La función *GBVI/WSA dG* es una función de *score* basada en campo de fuerza que estima la energía libre de unión de un ligando a partir de una pose dada. Ha sido entrenada usando los campos de fuerza *MMFF94x* y *AMBER99* sobre 99 complejos ligando-proteína del conjunto de entrenamiento *SIE* [202]. La forma funcional es una sumatoria de términos:

$$\Delta G \approx c + \alpha [2/3 (\Delta E_{coul} + \Delta E_{sol}) + \Delta E_{vdw} + \beta \Delta SA_{weighted}]$$

donde c representa el promedio de ganancia/pérdida de entropía rotacional y traslacional; α, β son constantes determinadas junto con c durante el entrenamiento y dependientes del campo de fuerza; E_{coul} es el término electrostático de Coulomb con una constante dieléctrica de 1; E_{sol} es el término electrostático de solvatación calculado usando el modelo de solvatación *GB/VI*, E_{vdw} es la contribución de van der Waals a la unión; $SA_{weighted}$ es el área superficial, ponderada por la exposición, se penaliza las áreas expuestas.

MOE implementa dos etapas de *scoring* entre las que existe una etapa intermedia de refinamiento de la pose. En esta etapa se procede a la minimización energética del sistema utilizando el campo de fuerza configurado. En orden de acelerar el cálculo, solo los residuos o átomos (dependiendo de la definición del sitio de unión dada) dentro de una distancia dada (por defecto 6 Å) son considerados. Por defecto los átomos de la cadena principal se mantienen fijos durante el refinamiento. Alternativamente el movimiento de los átomos de las cadenas laterales puede restringirse recurriendo por ejemplo al factor de temperatura B de cada átomo o dejarse totalmente libre.

En resumen en el presente trabajo se aplicó el siguiente protocolo de *docking*. En primer lugar las estructuras cristalográficas se descargaron del *PDB* en su unidad biológica. Cada estructura fue optimizada antes del análisis, átomos de hidrógenos y cargas fueron asignadas con el campo de fuerza *Amber12*, átomos y bucles ausentes fueron completados automáticamente. A efectos del *docking* el sitio de unión se

definió de acuerdo al correspondiente ligando co-cristalizado y los co-factores fueron incluidos como parte del receptor. Se utilizaron parámetros por defecto incluyendo el método de posicionamiento *Triangle Matcher*, la función de *score London dG* y la función de *re-scoring GBVI/WSA-dG* para generar 10 poses por ligando. Debe tenerse en cuenta que para la estrategia de *docking*, el solvente no fue considerado explícitamente sino empíricamente a través de la función de *score*.

◦ Evaluación del patrón de interacción

El patrón de interacción entre las poses de *docking* obtenidas en *MOE* y los blancos seleccionados fueron evaluadas mediante el módulo *PLIF (Protein Ligand Interaction Fingerprint)* [203]. *PLIF* es un método simple para representar y analizar la interacción tridimensional proteína-ligando. El patrón de interacción es convertido en una cadena binaria y puede usarse como descriptor molecular o herramienta de filtro luego del *docking*. De manera de clasificar el patrón de interacción cada residuo de aminoácido de la proteína es distribuido en las siguientes categorías:

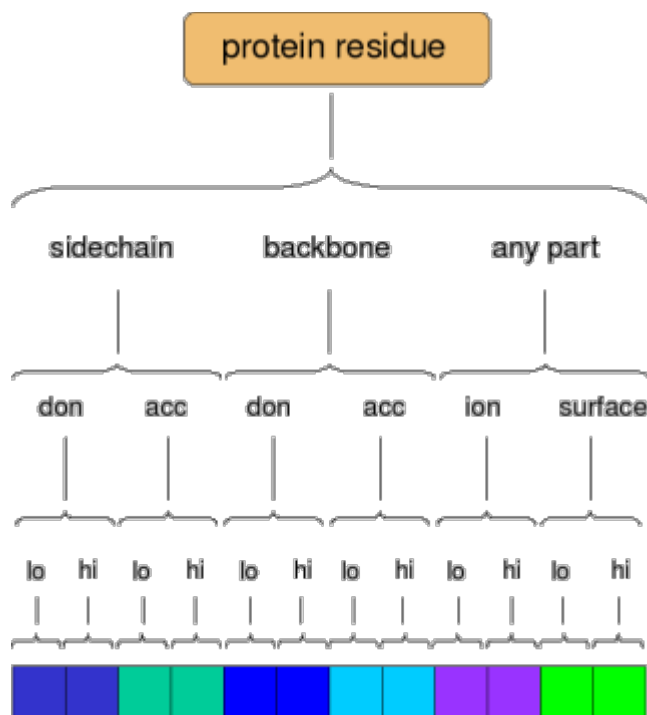


Figura 9. Tipos de interacción identificados por el módulo *PLIF (Protein Ligand Interaction Fingerprint)* de *MOE (Molecular Operating Environment)*.

Hay seis tipos de interacciones en las que un residuo puede participar: enlace de hidrógeno con la cadena lateral (dador o aceptor), enlace de hidrógeno con la cadena principal (dador o aceptor), interacciones iónicas y contactos superficiales. Las interacciones más potentes en cada categoría son consideradas. Si no se encuentra una interacción de una categoría particular, o ninguna pasa el criterio de aceptación, ningún carácter (*bit*) en la huella es generado. Si la interacción más fuerte pasa el

límite de interacción inferior, un carácter de bajo orden es indicado. Si la interacción más fuerte pasa el límite superior, entonces se indican tanto el carácter de bajo orden como el de alto. Entonces el patrón de bits de cada categoría puede tomar los valores 0 0, 0 1, 1 1, respectivamente. Los enlaces de hidrógeno entre átomos polares son calculados usando un método basado en estadísticas de contactos de proteínas, donde el par de átomos son puntuados de acuerdo a su distancia y orientación. El *score* se expresa como la probabilidad de ser un buen enlace de hidrógeno. Las interacciones iónicas son puntuadas al calcular el inverso de la distancia al cuadrado entre átomos con carga formal opuesta (ejemplo: oxígeno carboxilato y amina protonada) y expresado como porcentaje (100 % se corresponde a Å de distancia). Los contactos de superficie son determinados al calcular el área expuesta al solvente de un residuo, primero en ausencia de un ligando, y luego en presencia del ligando. La diferencia entre los dos valores indica el grado en el que el ligando ha apantallado al residuo del solvente, lo que se relaciona con una posible interacción hidrofóbica. El área expuesta al solvente es determinada al sumar 1.4 al radio de van der Waals de cada átomo y calcular la fracción de esta superficie total que no yace en el radio de alguna otra.

Para los complejos proteína ligando estudiados en el presente trabajo las siguientes interacciones fueron seleccionadas para analizar: enlace de hidrógeno con la cadena lateral, atracción iónica e interacción aromática.

- **Simulaciones de dinámica molecular de complejos quercetina-proteína blanco**

La flexibilidad del sitio de unión del blanco es un aspecto fundamental pero frecuentemente pasado por alto en el anclaje molecular. Enzimas y receptores pueden sufrir cambios conformacionales durante el proceso de reconocimiento [204]. En algunos casos estos rearrreglos estructurales son pequeños y el ligando se ajusta a un sitio de unión con poca movilidad. De otra manera, algunas proteínas sufren cambios significativos, que pueden involucrar elementos de estructura secundaria y terciaria.

Asimismo la mayoría de las funciones de *score* son limitadas al reproducir la energía de interacción absoluta de los complejos ligando-receptor. Problemas como desolvatación y efectos entrópicos son ejemplo de los desafíos que deben superar los actuales algoritmos de *docking*. El agua cristalográfica es uno de los mayores desafíos del *docking molecular* y el diseño de drogas. Estas moléculas están unidas al receptor y se observan en varias estructuras cristalográficas de una proteína particular. En aproximadamente el 65 % de las complejos cristalográficos proteína-ligando, al menos una molécula de agua está involucrada en el reconocimiento ligando receptor. Usualmente las aguas cristalográficas están localizadas en huecos profundos de la estructura del receptor y media enlaces de hidrógeno ente el ligando

y el sitio de unión de la proteína. En desarrollo de *docking*, estas moléculas pueden ser desplazadas por los ligandos designados o consideradas como parte de la estructura del receptor. La liberación de las aguas cristalográficas desde el sitio de unión es entrópicamente favorable, sin embargo el proceso puede causar una simultánea pérdida de entalpía. Para compensar por esta pérdida de entalpía, un motivo específico del ligando puede ser diseñado para imitar la red de interacción mostrada por el agua a través de la formación de enlace de hidrógeno equivalentes con la proteína. Alternativamente, las aguas estructurales pueden incluirse explícitamente en los experimentos de *docking*, permitiendo la formación de redes de enlaces de hidrógeno muy favorables entre el ligando y el sitio blanco. En este caso una variedad de métodos están disponibles para evaluar que moléculas de agua están fuertemente unidas, y entonces, se ajustan a este propósito. [201, 205, 206]

Estos problemas de flexibilidad pueden ser considerados por el uso de técnicas como la dinámica molecular. La dinámica molecular aplica las ecuaciones del movimiento de Newton, tal como se describen en la mecánica clásica, a la posición y velocidad específica de cada átomo en un sistema bajo estudio. Como resultado, la trayectoria y evolución temporal de un complejo ligando-receptor puede ser examinada. Tal trayectoria es importante por dos motivos. Primero, provee información configuracional y el momento para cada átomo, de las que propiedades termodinámicas del sistema pueden ser calculadas. Segundo, la trayectoria representa una exploración, u búsqueda, del espacio conformacional accesible a un sistema. El principio subyacente a la búsqueda conformacional usando dinámicas es que los átomos en una simulación eventualmente cubrirían su espacio conformacional completo. [207]

Inicialmente, una configuración específica es atribuida a los átomos con el propósito de reproducir la temperatura y presión de un sistema real. A partir del cómputo de las fuerzas actuando en cada partícula, es posible determinar la posición y velocidad de cada uno de estos átomos en un tiempo posterior. Estos cálculos son repetidos hasta que las trayectorias moleculares se integran en un intervalo de tiempo dado. Como la integración del movimiento es hecha numéricamente, para evitar inestabilidad, un paso de tiempo más corto que los movimientos más rápidos de la molécula debe ser usado. Esto se ubica normalmente entre 1 y 2 fs para simulaciones atomísticas, y es el principal cuello de botella del procedimiento.

Nuevamente, las fuerzas actuando en un sistema son determinadas por un campo de fuerza dado, usualmente parametrizado por cálculos mecano-cuánticos y datos experimentales. Entre los diversos campos de fuerza disponibles, *AMBER*, *CHARMM* y *GROMOS* pueden ser resaltados por ser ampliamente usado en simulaciones de dinámica molecular [208–210].

Sin importar la utilidad de la dinámica molecular esta tiene limitaciones. Entre ellas, se puede señalar el alto costo computacional demandado para la simulación de sistemas grandes, que usualmente consisten en miles de átomos donde complejos ligando-receptor están bajo estudio. Algunos de los cambios conformacionales desarrollados por los receptores durante el reconocimiento molecular ocurren en escalas de tiempo que exceden la capacidad computacional disponible. Pura fuerza bruta, solo haciendo simulaciones más extensas, no es suficiente para extender el muestreo conformacional de los sistemas biomoleculares. La forma compleja de la superficie energética hace que la mayoría de las simulaciones exploren solo una pequeña región en torno del mínimo energético cercana a la conformación inicial. Con la disponibilidad de recursos una estrategia obvia es hacer series de simulaciones paralelas con varias conformaciones iniciales. Aunque esto puede ser eficiente, requiere de un conocimiento específico del sistema a simular, y no puede ser aplicado como una estrategia general. Esta aproximación es particularmente útil cuando varias estructuras cristalinas están disponibles (por ejemplo en enzimas reguladas alostéricamente). Otras aproximaciones han sido diseñadas para incrementar el muestreo del espacio en una sola simulación, como en metadinámicas por ejemplo, donde las conformaciones ya visitadas son penalizadas; ensambles ponderadas, donde simulaciones adicionales son iniciadas cuando nuevas conformaciones son visitadas; dinámicas moleculares aceleradas donde barreras energéticas son artificialmente reducidas [211–213]. Finalmente se tienen, los llamados *simulated annealing*, que han sido reemplazados por métodos de intercambio de réplica [214].

A pesar de sus limitaciones, la dinámica molecular es capaz de brindar contribuciones importantes al diseño de drogas, especialmente cuando se combina con otros métodos de modelado molecular, tales como el *docking molecular*. Usualmente el ligando estabiliza un subconjunto de varias posibles conformaciones del receptor, cambiando el equilibrio hacia estructuras de mínima energía [215]. En algunos casos, las simulaciones de dinámica molecular pueden producir conformaciones alternativas correspondiendo estas estructuras inducidas por el ligando. Asimismo, cuando no existen estructuras cristalográficas para un blanco molecular en particular (estructuras con sitios inaccesibles o pobremente definidos), la dinámica molecular puede ser aplicada para generar un conjunto de estructuras de *docking* [216]. De acuerdo con esto, estados conformacionales potenciales son muestreados por simulaciones de dinámica molecular basados en los datos cristalográficos disponibles, conformaciones accesibles (aquellas con sitios accesibles y bien definidos) pueden ser seleccionadas para *docking molecular* [207]. La dinámica molecular puede adicionalmente ser usada para estimar la estabilidad del complejo receptor ligando propuesto por *docking molecular* [217]. Cuando una conformación del ligando generado por dinámica molecular se desvía más de un valor dado de *RMSD* de la correspondiente solución de *docking*, el complejo ligando-receptor puede considerarse inestable [218].

En el presente trabajo se realizaron simulaciones de dinámica molecular con el método *Nosé Poincaré Anderson (NPA)* implementado en *MOE*. El método de *Nosé-Poincaré-Anderson* es una formulación de *Hamiltoniano* a tiempo real para la simulación de dinámica molecular isotérmica e isobárica. Pertenece a los llamados *Hamiltonianos extendidos* en los que grados de libertad adicionales son incluidos para que las trayectorias muestren de las condiciones termodinámicas deseadas. A diferencia del método *Nosé-Hoover-Anderson*, permite el uso de integradores simpléticos para las ecuaciones de movimiento [219]. Específicamente *MOE* utiliza el algoritmo generalizado *leapfrog*, que es irreversible en el tiempo y de segundo orden. Se eligió un ensamble con número constante de partículas, volumen y temperatura (*NVT*) a 300 K. Cada complejo fue neutralizado con iones (*NaCl*) e inmerso en una caja de agua periódica con un margen de 6 Å. Todo el sistema se minimizó a un gradiente de 1 kcal/mol *RMS* antes de iniciar la simulación. Previo a la simulación principal una etapa de 100 ps calentó gradualmente el sistema a 300 K.

Las simulaciones de dinámica molecular de *MOE* generan una serie de archivos periódicos de trayectoria con campos conteniendo información de medidas energéticas, de posición, velocidad, etc. así como las configuraciones muestreadas a intervalos regulares. Se destacan entonces los campos “*mol*”, correspondiente a la topología y configuración del sistema molecular; “*t*”, tiempo de muestreo (en picosegundos); “*U*”, energía potencial $U(r)$ del sistema atómico a tiempo t en kcal/mol; “*K*”, energía cinética de los átomos a tiempo t en kcal/mol; “*T*”, temperatura instantánea en Kelvin del sistema a tiempo t ; “*P*”, presión instantánea del sistema en kPa a tiempo t ; “*V*”, volumen instantáneo en angstrom cúbicos a tiempo t ; “*Ub*”, la energía potencial en kcal/mol de un conjunto específico de átomos seleccionados (por ejemplo el ligando) a tiempo t ; “*Ua*”, la energía potencial en kcal/mol del resto de los átomos del sistema (no seleccionados, por ejemplo proteína y solvente); “*Uab*”, la energía potencial de interacción en kcal/mol entre los átomos pertenecientes al conjunto a y pertenecientes al conjunto b ; “*pos*”, vector de posición a tiempo t ; “*vel*”, vector de velocidad a tiempo t . De acuerdo con esto se registró la energía potencial de interacción entre los átomos del ligando y el resto del sistema durante la simulación. Utilizando además una rutina de programación *SVL* de *MOE*, se descompuso el valor de la energía potencial de interacción asociada al receptor y asociada al solvente respectivamente. Asimismo para la etapa final de la simulación se calculó el patrón de contacto mediante el módulo *PLIF* detallado previamente en condiciones por defecto. La equilibración fue monitoreada por convergencia en términos de temperatura, energía, y desviación geométrica (*RMSD*) con respecto a los átomos del *backbone* de la estructura de partida.

Resultados y discusión

- Tamizaje reverso

- *Tamizaje reverso de quercetina basado en el ligando*

Tomando a quercetina como molécula de prueba y el algoritmo de *SHAFTS* cribamos localmente nuestra librería de ligandos con características tipo droga asociados a complejos del *PDB*. En paralelo, realizamos la misma búsqueda utilizando el servidor *ChemMapper*. Este servidor implementa *SHAFTS* y proporciona un *score* (*chemmapper score*) que pretende estimar la afinidad del ligando prueba en base a la afinidad conocida del ligando similar.

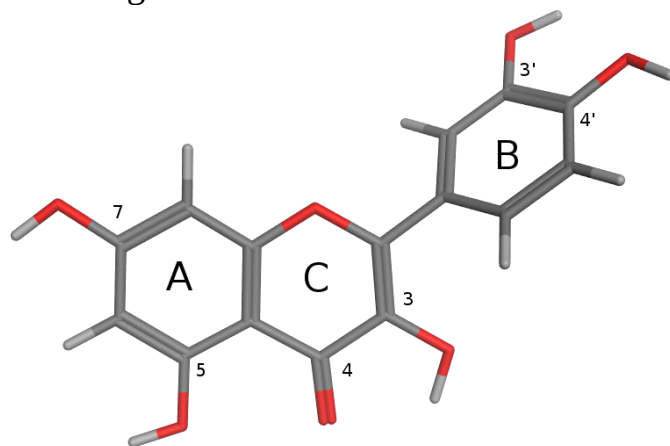


Figura 10. Estructura de quercetina optimizada por MMFF94x hasta un gradiente RMS de 0,01 kcal/mol en *MOE* (*Molecular Operating Environment*). Se indica la anotación ABC de los anillos y la numeración de los sustituyentes. Los átomos de oxígeno se representan en rojo.

Utilizando *ChemMapper* se realizaron tres búsquedas, correspondientes a los tres límites de *score* de similitud de ligando, de menor a mayor exigencia: 0.8, 1.2 y 1.5. En la primera franja, de ligandos con al menos un *score* de similitud de 0.8 se identifican 1000 estructuras; en la segunda franja, de ligandos con un *score* de al menos 1.2 se identifican 108 estructuras y finalmente en la franja de ligandos con un *score* de al menos 1.5 se identifican 10 estructuras. La anotación automática de las proteínas asociadas identificadas por *ChemMapper* fue revisada y corregida. Las listas generadas son redundantes al darse que varios ligandos pueden identificar la misma proteína y viceversa. Por otro lado se identifican también complejos entre ligandos y ácidos nucleicos que no fueron tenidos en cuenta. Con estas consideraciones se tiene que para el tamizaje con un *score* de al menos 0.8 se identifican aproximadamente 597 proteínas, de las que 200 son proteínas humanas, 32 son proteínas de ratón, 18 son proteínas de vaca y 17 proteínas de rata, entre otras. En la búsqueda con al menos un *score* de 1.2 se tienen aproximadamente 74 proteínas, de las que 33 son humanas y 3 de ratón entre otras. Finalmente en la franja

con un *score* de similitud de al menos 1.5 se tienen 20 proteínas de las que 8 son humanas. Si bien cada lista se incluye en la anterior el *chemmapper score* reportado por el servidor varía en cada caso, aún para la misma proteína en la misma búsqueda, y no se observa ninguna tendencia entre sí.

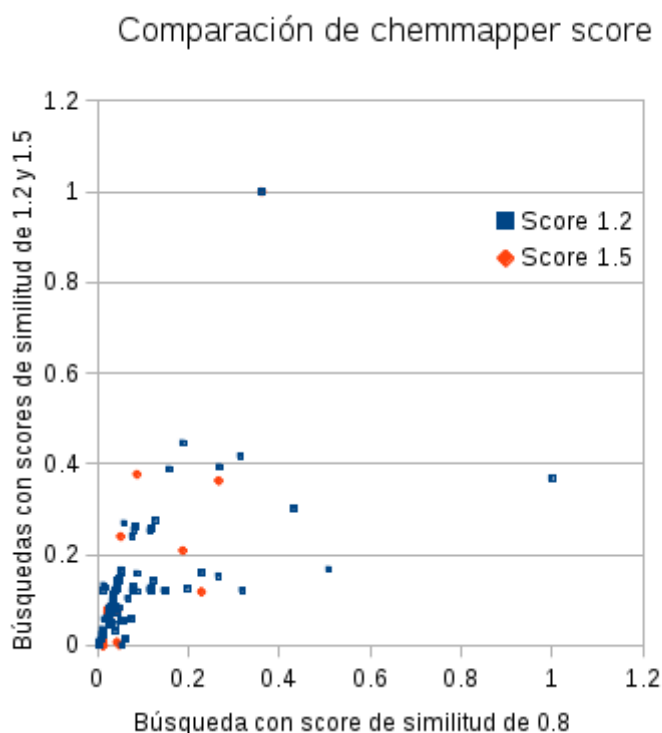


Figura 11. Comparación entre el *chemmapper score* de tamizajes realizados con diferente *score* de similitud (0.8, 1.2, 1.5).

El resultado obtenido utilizando *SHAFTS* de manera local en nuestra base de ligandos puede desglosarse de la misma manera, teniendo un máximo de 2700 ligandos similares quede reducirse a 2174 con un *score* de similitud de al menos 1.1, 631 con un *score* de al menos 1.2, 182 con un *score* de al menos 1.3, 53 con un *score* de al menos 1.4, 34 con un *score* de al menos 1.5, 18 con un *score* de al menos 1.6, 8 ligandos con un *score* de al menos 1.7 y 1 ligando con *score* de al menos 1.8. Como se mencionó en metodología la base de ligandos se fraccionó en 9 partes, y si bien la búsqueda se realizó seleccionando ligandos con una similitud de al menos 0.8 el máximo número de ligandos reportados es 300. Por lo tanto 0.8 resulta ser un punto de corte inespecífico, el resultado de cada búsqueda son los mejores 300 ligandos. Este mismo fenómeno ocurre también con el servidor *ChemMapper* donde el listado más largo posible es de 1000 entradas. Asimismo, el ligando de partida quercetina no es identificado con el máximo *score* de similitud. Esto se debe a una diferencia en la orientación del anillo B que reduce el *feature score*, es decir el solapamiento entre grupos OH del alineamiento no es el más óptimo. La fuente de esta diferencia estaría en el algoritmo de búsqueda conformacional.

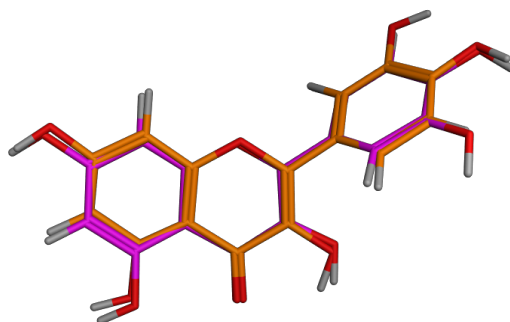
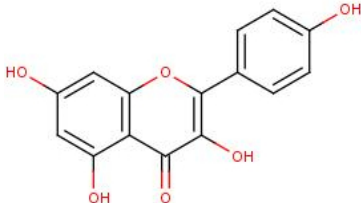
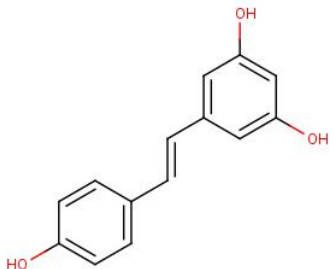


Figura 12. Alineamiento de quercetina según *SHAFTS*.

Las proteínas asociadas fueron recuperadas recurriendo al índice de la librería en *LigandExpo* y a rutinas de *Python* y *Bash*. Tomando como criterio de selección aquellos lignados con un *score* de similitud estructural de al menos 1.5 obteniendo una lista de 34 moléculas relacionadas a proteínas del *PDB*. Estos ligandos comparten la estructura típica de un flavonoide a excepción de aquellos identificados en el *PDB* como *FL9*, *PIT*, *RE2*, *SLX*, *STL* y *DEH* (ver ejemplos en tabla 2). El correspondiente conjunto de proteínas co-cristalizado incluye 25 proteínas de mamífero. Proteínas de plantas y otros organismos no fueron considerados para un análisis posterior.

Tabla 2. Ejemplos de ligandos similares a quercetina identificados por el algoritmo *SHAFTS* (*SHApe-Feature Similarity*) en la primera etapa del proceso de tamizaje reverso. Los ligandos son representados de acuerdo a su identificador del *Protein Data Bank*, nombre, *score* de similitud contra quercetina, estructura bidimensional y presencia en proteínas de mamífero.

PDB id	Nombre	Score de Similitud	Proteína de mamífero	Representación 2D
KMP	Kaempferol (3,5,7-trihydroxy-2-(4-hydroxyphenyl)-4h-chromen-4-one)	1,727	Si	
STL	Resveratrol (5-[(E)-2-(4-hydroxyphenyl)ethenyl]benzene-1,3-diol)	1,584	Si	

FL9	3',5'-dibromo-2',4,4',6'-tetrahydroxy aurone	1,523	Si	
SLX	(S)-scoulerine ((13aS)-3,10-dimethoxy-5,8,13,13a-tetrahydro-6H-isoquino[3,2-a]isoquinoline-2,9-diol)	1,518	-	
371	2-(3,4-dihydroxyphenyl)-8-(1,1-dioxidoisothiazolidin-2-yl)-3-hydroxy-6-methyl-4h-chromen-4-one	1,515	Si	
DEH	2,9-dihydroxy-3,10-dimethoxy-5,6-dihydroisoquino[3,2-a]isoquinolinium	1,503	-	

Si se compara el resultado de *SHAFTS* obtenido localmente con el obtenido por medio del servidor *ChemMapper* se aprecia primero que el *score* de similitud estructural es menor para cualquiera de los ligandos reportados por el servidor frente a la estrategia local; segundo, proteínas iguales no son reportadas a través del mismo ligando, y finalmente, varias proteínas identificadas en la estrategia local no son reportadas por el servidor. Se destaca la ausencia de la enzima xantina oxidasa co-cristalizada con quercetina.

Resultados complementarios fueron obtenidos utilizando una base de moléculas seleccionadas de *PubChem* 98 % similares a quercetina según su coeficiente de *Tanimoto*. Las mismas representan variantes en el patrón de hidroxilación de quercetina, tales como miricetina y robinetina. En resumen, las proteínas identificadas en el tamizaje con quercetina incluyen a todas las proteínas identificadas para el resto de las moléculas similares a excepción de la *proteína de shock térmico de 90 kDa (heat shock protein 90)*, el *receptor de estrógenos α* y la *piruvato deshidrogenasa quinasa*. Estas últimas proteínas se asocian a ligandos con un *score* de similitud cercano pero por fuera del valor de corte.

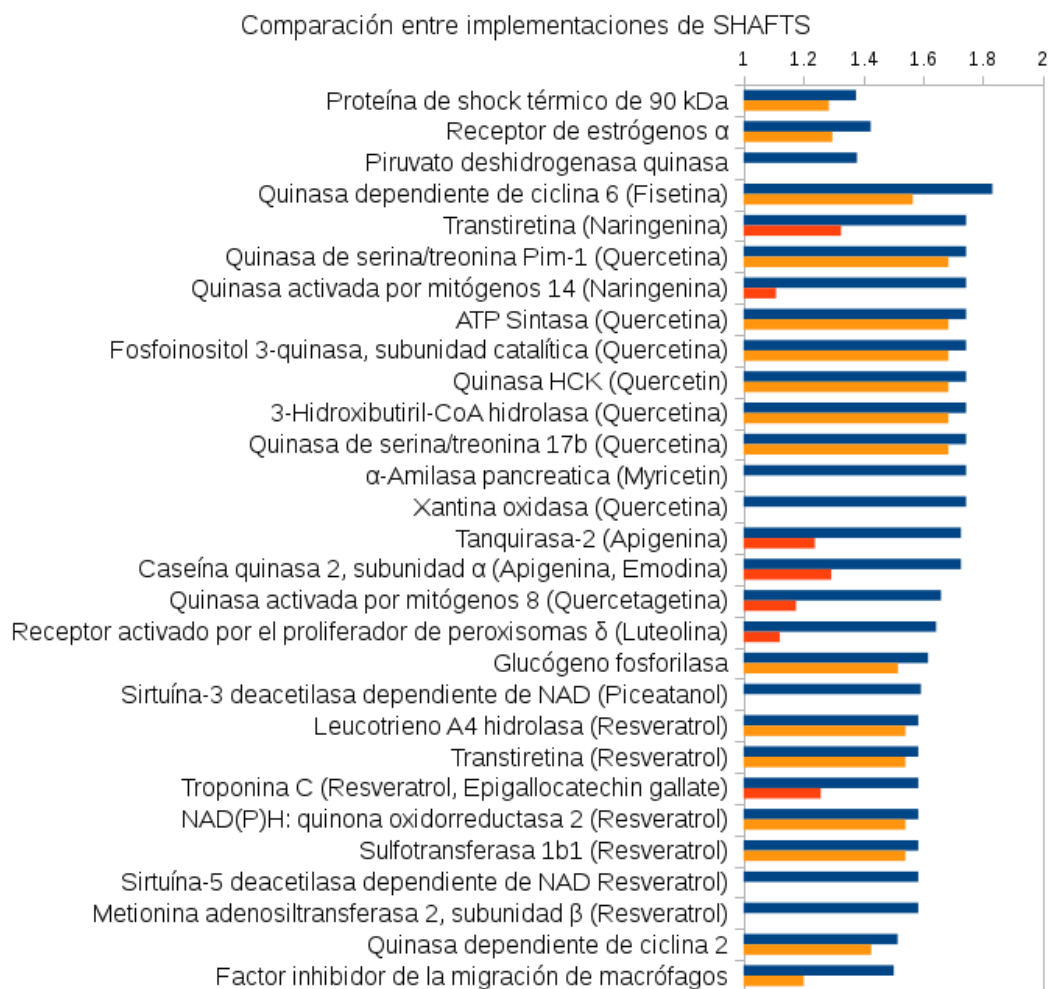


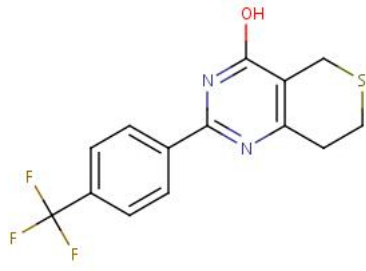
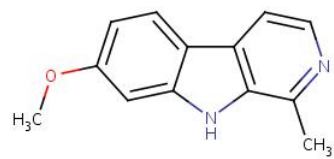
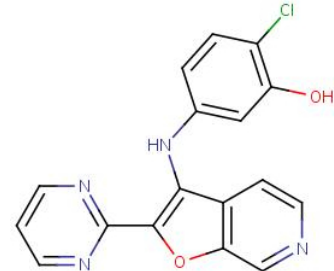
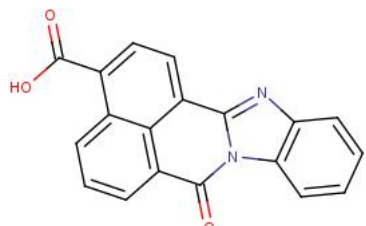
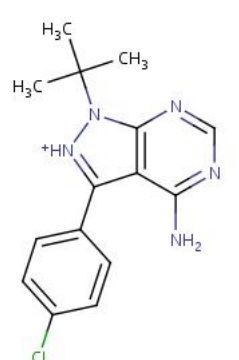
Figura 13. Comparación entre *SHAFTS* local (barras azules) y online (*ChemMapper*) (barras naranja) para quercetina. En rojo se indican proteínas identificadas por el servidor *ChemMapper* a través de un ligando diferente del encontrado por la estrategia local.

○ Comparación de sitios de unión

Basado en el análisis de *LIBRA* de similitud local de proteínas conducimos una etapa de amplificación de los blancos encontrados por tamizaje basado en el ligando. Es interesante notar que la mayoría de las proteínas identificadas también unen ligandos con cierta similitud a quercetina y pueden ser encontradas en la etapa previa de búsqueda en un rango de *score* de entre 1.0 y 1.3 (ver ejemplos en tabla 3). Este resultado muestra la concordancia entre ambas estrategias y sugiere que candidatos adicionales pueden haber sido inicialmente encontrados con un criterio de similitud menos exigente.

Tabla 3. Ejemplos de ligandos co-cristalizados de proteínas blanco identificadas en el *Protein Data Bank (PDB)* por tamizaje virtual reverso basado en comparación de sitios de unión con el algoritmo

LIBRA (*Ligand Binding site Recognition Application*). Los ligandos son representados de acuerdo de su identificador del PDB, nombre, score de similitud contra quercetina si se encuentra presente y estructura bi-dimensional.

PDB id	Nombre	Score de Similitud	Representación 2D
XAV	2-[4-(trifluoromethyl)phenyl]-7,8-dihydro-5H-thiopyrano[4,3-d]pyrimidin-4-ol	1,270	
HRM	7-methoxy-1-methyl-9h-beta-carboline	1,200	
FP3	2-chloro-5-{[2-(pyrimidin-2-yl)furo[2,3-c]pyridin-3-yl]amino}phenol	1,273	
609	7-oxo-7H-benzimidazo[2,1-a]benz[de]isoquinoline-3-carboxylic acid	1,168	
PP2	1-tert-butyl-3-(4-chloro-phenyl)-1h-pyrazolo[3,4-d]pyrimidin-4-ylamine	1,232	

Junto con el conjunto previo de candidatos tres clases principales de proteínas pueden distinguirse, *sulfotransferasas*, *poli-(ADP-ribosa) polimerasas* – *tanquirasas* y *quinasas*. La mayoría de estas quinasas comparten el mismo plegamiento de acuerdo a PFAM, el dominio proteína quinasa (PF00069) seguido del dominio tirosina quinasa (PF07714). Otro dominio quinasa presente en el conjunto de proteínas blanco es el dominio PI3 quinasa (PF00613) correspondiente a *fosfoinositol 3-quinasa*, conocido blanco de quercetina [220].

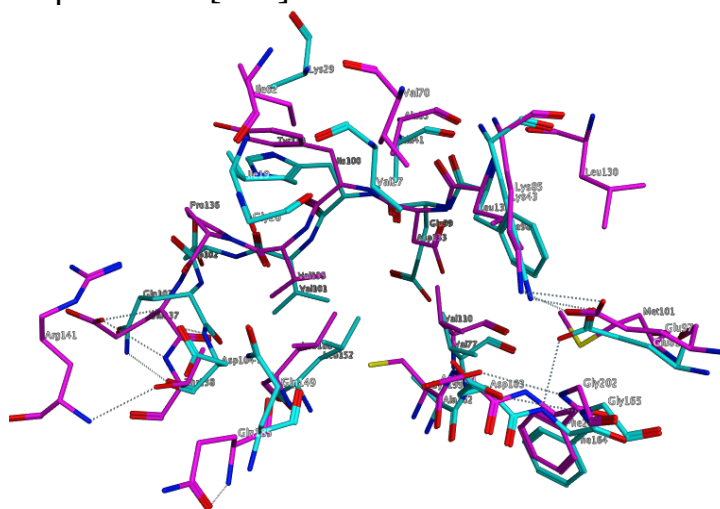


Figura 14. Superposición del sitio de unión de la quinasa dependiente de ciclina 6 (*PDB 1xo2*) y la glucógeno sintasa quinasa 3 β (*PDB 4j71*) (*RMSD* carbonos del sitio 1.768).

○ *Tamizaje basado en la estructura*

De las etapas previas de búsqueda se obtuvo una lista de 74 proteínas blanco. Este conjunto de proteínas fue sometido a tamizaje basado en la estructura contra quercetina utilizando *idTarget*. Este procedimiento permitió obtener una lista de blancos asociados a poses de *docking* de quercetina ordenada de acuerdo a su afinidad predicha (ver figuras 15 y 16). Para el conjunto de proteínas la afinidad predicha tiene un rango de entre 15.1 nM a 494.1 nM y muestra variaciones menores dependiendo del cristal en particular para una misma proteína. De la lista sin ordenar entregada a *idTarget* el servidor solo descartó una proteína de mamífero, la sirtuina 5, lo que confirma la robustez de la comparación estructural. Es remarcable también que todos los complejos de quercetina disponibles en el *PDB* al momento del ensayo fueron recuperados.

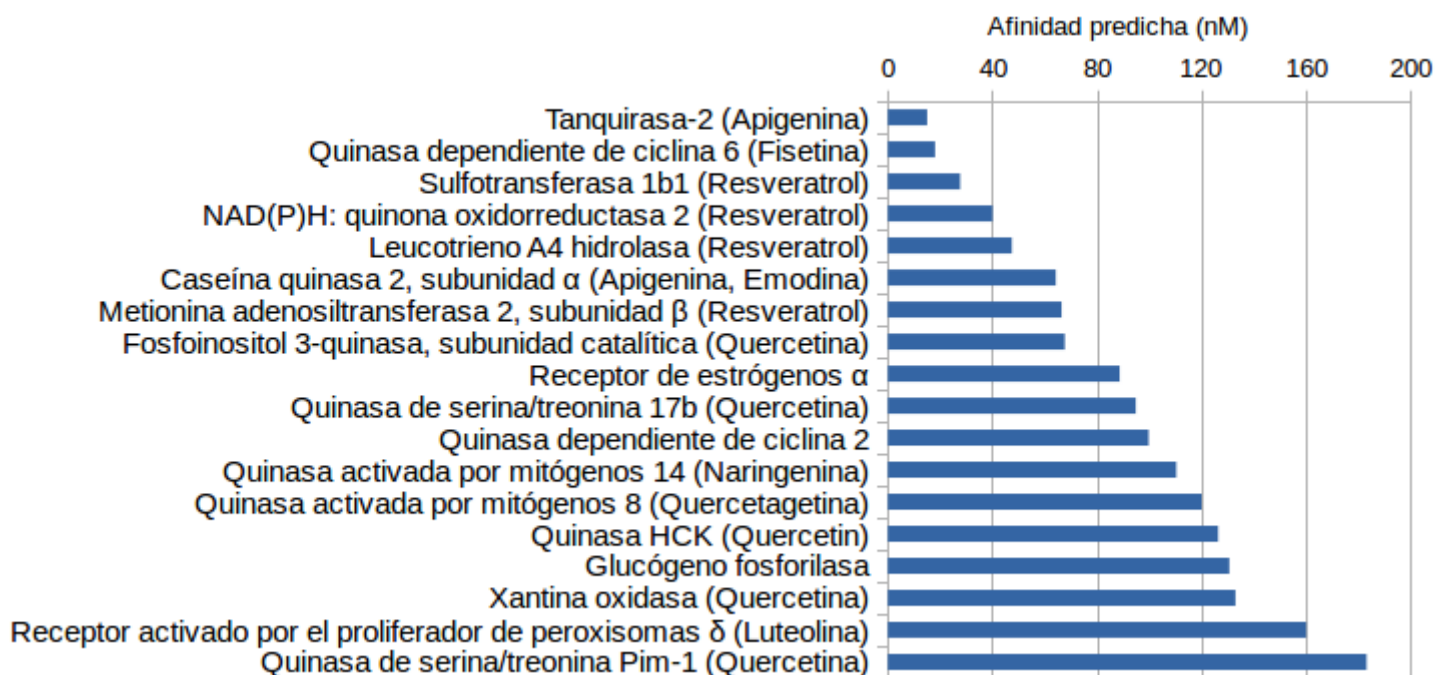


Figura 15. Proteínas blanco identificadas en el *Protein Data Bank (PDB)* por tamizaje virtual basado en similitud del ligando con el algoritmo *SHAFTS (SHApe Feature Similarity)* y *docking inverso* con el servidor *idTarget* (<http://idtarget.rcas.sinica.edu.tw/>). Se representan los candidatos con afinidad predicha menor a 200 nM para quercetina. Por detalles sobre entrada en el *PDB*, organismo de origen, ligando co-cristalizado, y *score* de similitud estructural con quercetina así como nombre de la proteína y *score* de *idTarget* (*Z-score* y afinidad predicha) ver tabla 3 del apéndice.

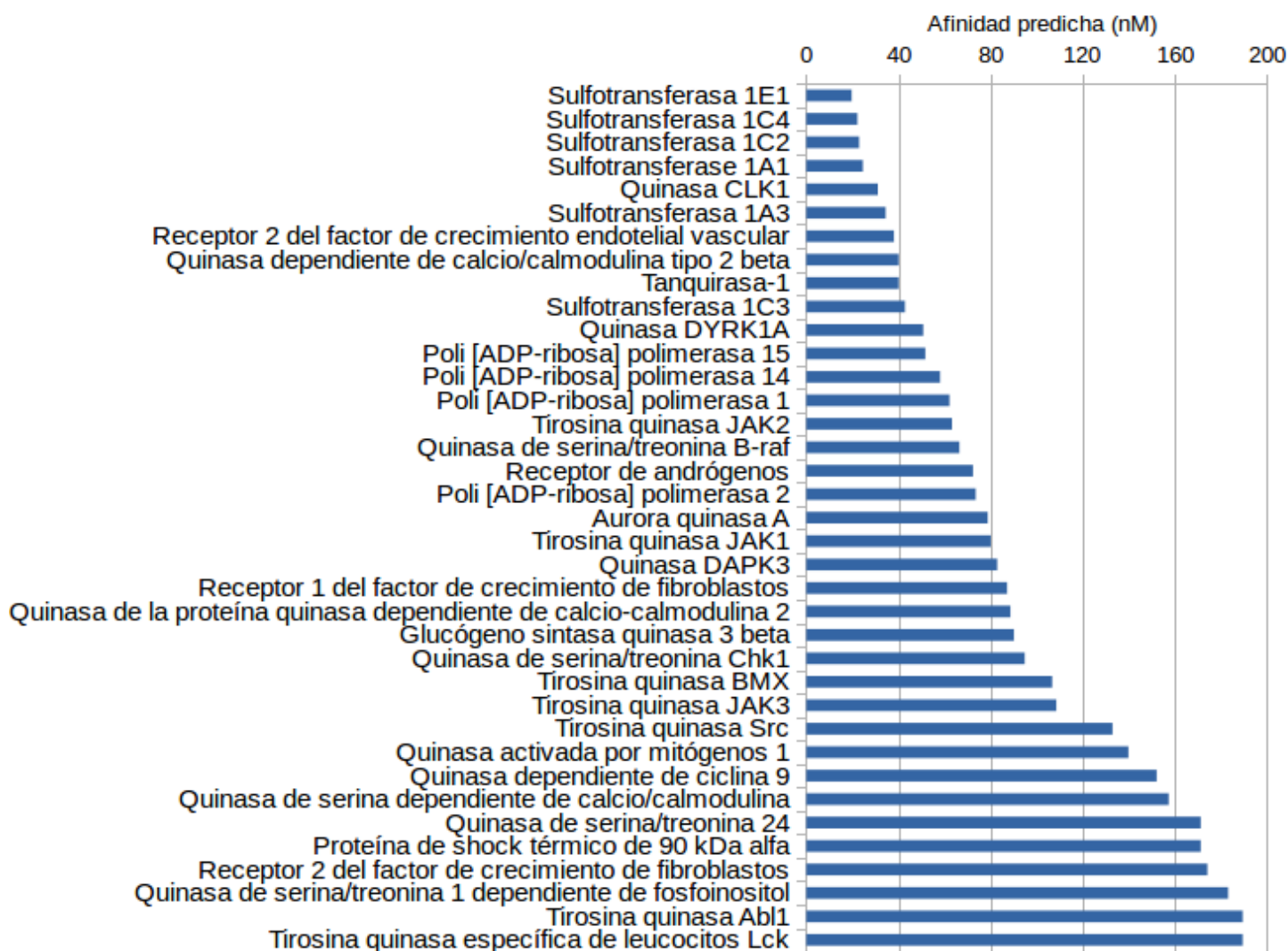


Figura 16. Proteínas blanco identificadas en el *Protein Data Bank* (PDB) por tamizaje reverso basado comparación del sitio de unión con el algoritmo *LIBRA* (*Ligand Binding Site Recognition Application*) y *docking inverso* con el servidor *idTarget* (<http://idtarget.rcas.sinica.edu.tw/>). Se representan los candidatos con afinidad predicha menor a 200 nM para quercetina. Por detalles sobre entrada en el PDB, organismo de origen, ligando co-cristalizado, y score de similitud estructural con quercetina así como nombre de la proteína y score de *idTarget* (Z-score y afinidad predicha) ver tabla 4 del apéndice.

De acuerdo a las poses de *docking* obtenidas por *idTarget* quercetina alcanza el sitio de unión conocido para cada proteína, el que generalmente es un sitio catalítico o alostérico (figura 17). Mientras que *idTarget* predice adecuadamente el sitio de unión, las poses de *docking* no siempre se asemejan a la orientación del ligando co-cristalizado. Frecuentemente quercetina aparece en el mismo plano pero con una orientación opuesta. Esta divergencia puede explicarse por la simetría interna de quercetina y no sería un fenómeno inusual si consideramos el caso del flavonoide apigenina. En los cristales reportados para la proteína caseína quinasa 2, el flavonoide muestra un modos de unión opuestos en el representante humano (*PDB 3amy*) y en un homologo vegetal extremadamente similar (*PDB 4dgm*) (figura 18).

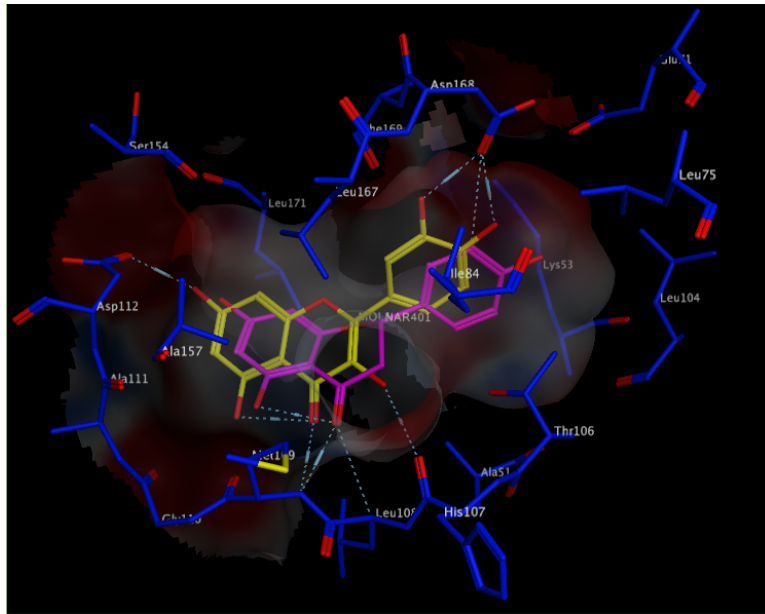


Figura 17. Resultado del *docking* de quercetina (amarillo) con la *proteína quinasa activada por mitógenos 14* (p38) co-cristalizada con naringenina (violeta) (4eh3) de acuerdo al servidor *idTarget* (<http://idtarget.rcas.sinica.edu.tw/>). Se representa la superficie electrostática en el entorno de quercetina a una distancia de van der Waals, las áreas mostradas en rojo corresponden a superficies cargadas negativamente y superficies blancas con carga neutra. Se observa un buen solapamiento entre las poses de naringenina y quercetina confirmando la robustez del método de *docking*.

Adicionalmente, algunas simulaciones de *docking* muestran ciertas peculiaridades, por ejemplo algunas poses de *sirtuina 3* solapan al péptido sustrato, algunas poses de la *glucógeno fosforilasa* se ubican en un sitio alostérico normalmente ocupado por un ligando endógeno AMP, las poses de las *sulfotransferasas* y *ATP sintasa* también se encuentran dentro del sitio del ATP próximo a un sitio alostérico. A excepción de las *sulfotransferasas*, estas proteínas se encuentran en el rango de *Z-score* y afinidad predicha menos favorable de la lista. Este conjunto de proteínas presentan sitios de unión amplios y expuestos en contraste con los candidatos mejor posicionados. De acuerdo al *score* de *idTarget* y tomando en cuenta estas características cualitativas, proteínas como el *factor de migración de los macrófagos*, *troponina C*, *transtiretina* y *ATP sintasa* entre otras no se deberían priorizar como proteínas blanco.

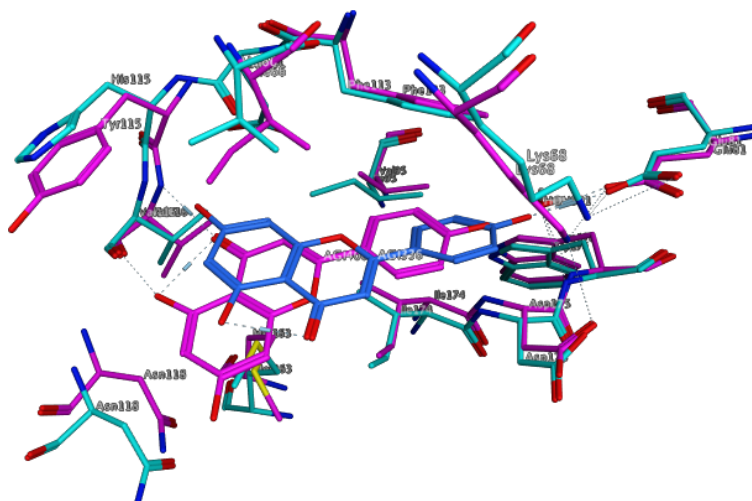


Figura 18. Superposición CK2 humana (*PDB 3amy*) y de maíz (*PDB 4dgm*) co-cristalizada con apigenina, representación del sitio de unión.

Un resultado interesante del presente estudio es que la mayoría de los ligandos endógenos, substratos o co-factores relacionados al conjunto de proteínas blanco son derivados de nucleótidos como ATP/ADP, FAD, NAD, piridoxal fosfato, PAPS (3'-fosfoadenosina-5'-fosfosulfato), HIBYL-CoA (3-hidroxitiril-CoA) y xantina. Las proteínas blanco difieren en largo de secuencia, plegamiento y estructura cuaternaria, lo que sugiere que cualquier similitud que comprenda a todo el conjunto será local y no global. En una primera instancia, la similitud estructural entre quercetina y ATP/ADP fue estudiada. El resultado de este análisis (figura 19) indica que quercetina puede estar actuando como un mimético del ATP. En este aspecto, estudios previos sugieren una correspondencia entre el motivo benzopirano del flavonoide y la porción adenina del ATP así como el motivo orto-dihidroxifenilo y los grupos fosfatos del ATP [221]. Otras semejanzas notables son entre quercetina y las hormonas estradiol y tiroxina (figura 20).

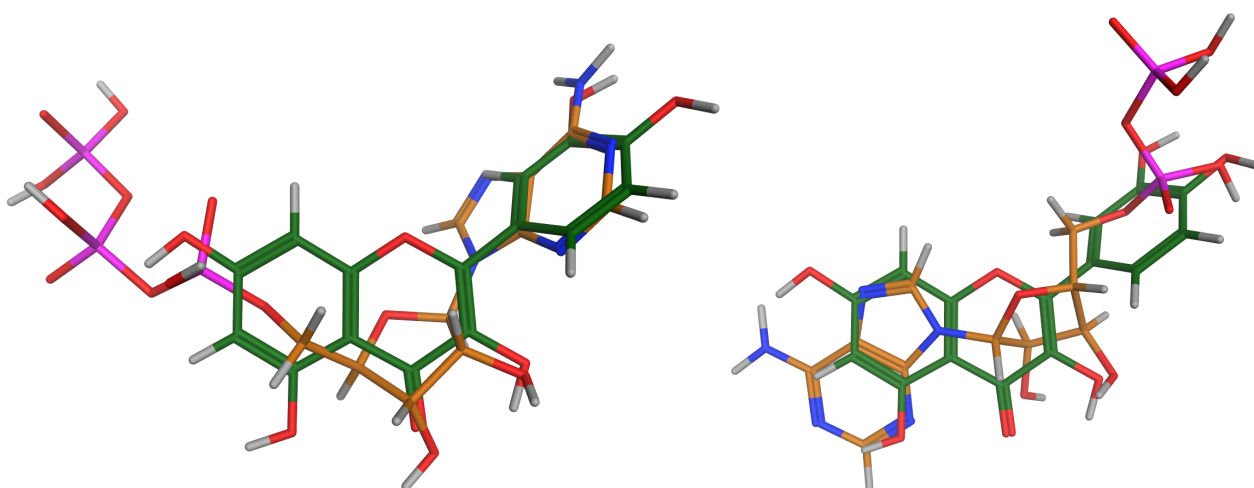


Figura 19. Alineamiento estructural entre ATP/ADP y quercetina dado por el algoritmo *SHAFTS* (*SHApe Feature Similarity*). El alineamiento muestra dos orientaciones diferentes de quercetina (*score* del alineamiento con ATP de 0,8693, *score* del alineamiento con ADP de 0,9790) que pueden solapar grupos funcionales de estos ligandos endógenos y sugieren una acción mimética de quercetina.

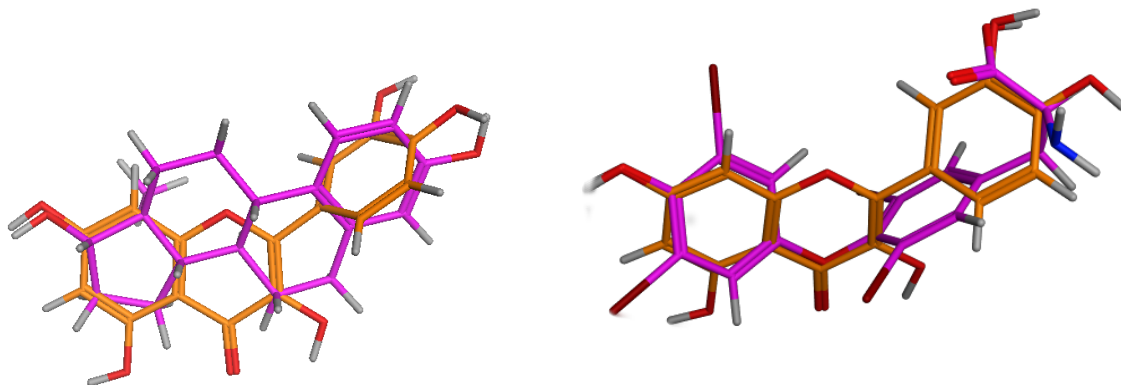


Figura 20. Alineamiento estructural entre estradiol, tiroxina y quercetina dado por el algoritmo *SHAFTS* (*SHApe Feature Similarity*). Score de similitud estradiol/quercetina: 1.148, score de similitud tiroxina/quercetina: 1.152.

- **Caracterización funcional**

El conjunto ordenado de blancos incluye proteínas con un amplio rango de plegamientos macromoleculares y actividades biológicas incluyendo enzimas, factores de transcripción y proteínas de transporte, y puede ser dividido en proteínas con una acción metabólica directa como *xantina oxidasa* y proteínas involucradas en vías de señalización intracelulares. Apoyando nuestros resultados algunas de las proteínas identificadas cuentan con ensayos *in vitro* reportando su interacción con quercetina en distintas bases de datos como *DrugBank*, *ChEMBL*, *BindingDB* y *SuperTarget* [222–225]. En este grupo se encuentran la *quinasa de serina/treonina PIM1*, la *quinasa de serina/treonina 17b*, *quinasa HCK*, *ATP sintasa*, *fosfoinositol 3-quinasa*, *3-hidroxitiril-CoA hidrolasa*, *caseína quinasa 2*, *xantina oxidasa*, *tirosina quinasa SRC* y el *receptor 2 del factor de crecimiento endotelial vascular*.

- **Breve reseña de las proteínas blanco**

Fosfoinositol 3-quinasa, subunidad catalítica: quinasa de serina/treonina, fosforila *fosfatidilinositol 4,5-bifosfato* para generar *fosfatidilinositol 3,4,5-trifosfato* (PIP3). PIP3 juega un rol clave reclutando proteínas con dominio PH a la membrana como *AKT1* y *PDK1*, activando cascadas de señalización involucradas en el crecimiento celular, supervivencia, proliferación, motilidad y morfología. Une la activación de receptores asociados a proteínas G a la producción de PIP3. Involucrada en la respuesta inflamatoria. (842 aa)

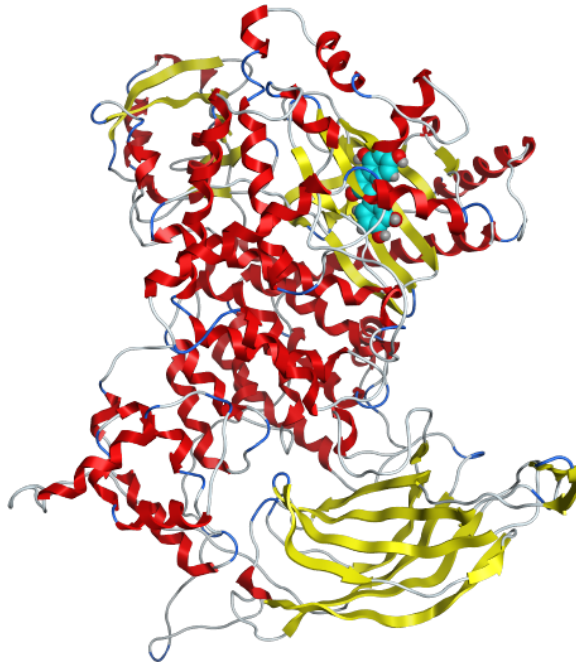


Figura 21. Representación de la estructura tridimensional del complejo *fosfoinositol 3-quinasa* /quercetina (verde) (PDB id: 1e8w).

Quinasa dependiente de ciclina 2: quinasa de serina/treonina involucrada en el control del ciclo celular; esencial para la meiosis pero dispensable para la mitosis. Dispara la duplicación de los centrosomas y el ADN. Actúa en la transición G1-S y promueve el inicio de la síntesis de ADN, modula la progresión de G2, controla el tiempo de entrada en mitosis/meiosis al controlar la activación de *ciclinaB/CDK1* por fosforilación. Involucrada en la iniciación y mantenimiento del ciclo celular durante la diferenciación celular, previene la proliferación celular y regula negativamente la diferenciación, pero es requerida para la proliferación de tipos celulares específicos como nuevas neuronas, eritrocitos y células hematopoyéticas. (298 aa)

Quinasa dependiente de ciclina 6: quinasa de serina/treonina involucrada en el control del ciclo celular y diferenciación, promueve la transición G1/S. (326 aa)

Quinasa dependiente de ciclina 9: quinasa de serina/treonina involucrada en la regulación positiva de la transcripción. (372 aa)

Quinasa DAPK3 (death-associated protein kinase 3): quinasa de serina/treonina involucrada en la regulación de la apoptosis, autofagia, transcripción, traducción, reorganización del citoesqueleto de actina, movilidad celular, contracción de la fibra muscular lisa, y mitosis, particularmente citoquinesis. Regula tanto señales de muerte celular por apoptosis de tipo I y II, dependiendo del contexto celular. La primera es dependiente de caspasas, mientras que la última es independiente y caracterizada por acumulación de vesículas autofágicas. (454 aa)

Quinasa de serina/treonina 17b: actúa como un regulador positivo de la apoptosis. (372 aa)

Quinasa de serina/treonina Chk1: necesaria para la activación del arresto celular y reparación del ADN. (476 aa)

Quinasa de serina/treonina 24 (quinasa MST3): promueve la apoptosis en respuesta a estímulos de estrés y activación de caspasas. Media la apoptosis inducida por estrés oxidativo al modular la fosforilación de *JNK1-JNK2* (*MAPK8* y *MAPK9*), p38 (*MAPK11*, *MAPK12*, *MAPK13* y *MAPK14*). (443 aa)

Quinasa de serina/treonina PIM1: involucrada en la supervivencia y proliferación celular, confiere una ventaja selectiva a la tumorigénesis. Entre otros mecanismos ejerce su actividad oncogénica al fosforilar e inhibir proteínas proapoptóticas (*BAD*, *FOXO3*, *MAP3K5*). (313 aa)

Quinasa de serina/treonina 1 dependiente de fosfoinositol: juega un rol central en la transducción de señales por insulina al provocar la activación por fosforilación de *PKB/AKT1*, propagando la señal a blancos debajo de la cascada y controlando la supervivencia y proliferación celular, así como la captación y reserva de glucosa y aminoácidos. Regula negativamente al receptor del factor de crecimiento tumoral (*TGF-beta*). Activa la actividad transcripcional de *PPAR-gamma* y promueve la diferenciación de adipocitos. Activa la vía de *NF-kappaB* al fosforilar *IKKB*. (556 aa)

Quinasa quinasa quinasa activada por mitógenos 5 (MAP3K5): quinasa de serina/treonina que actúa como componente esencial de la cascada de MAPK. Juega un rol importante en respuestas celulares a cambios en el ambiente, particularmente en la transducción de la señal apoptótica dependiente de la activación de caspasas dependiente de mitocondria. Requerida para la respuesta inmune innata. (1374 aa)

Quinasa activada por mitógenos 1 (ERK2): quinasa de serina/treonina que actúa como componente esencial de la cascada de MAP quinasa junto a *MAPK3/ERK1*. Dependiendo del contexto celular esta cascada media diversas funciones biológicas como crecimiento celular, adhesión, supervivencia y diferenciación a través de la regulación de la transcripción, traducción, rearrreglos del citoesqueleto. (360 aa)

Quinasa 2 activada por MAPK: quinasa de serina/treonina involucrada en la producción de citoquinas, endocitosis, reorganización del citoesqueleto, migración celular, control del ciclo celular, remodelado de la cromatina, respuesta al daño del ADN y regulación transcripcional. Siguiendo a un estrés es fosforilada y activada por *MAPK14*. Involucrada en la respuesta inflamatoria al regular la producción de *IL-6* y del factor de necrosis tumoral (*TNF*) postranscripcionalmente. (400 aa)

Quinasas activadas por mitógenos 8, 9 y 10 (JNK1, JNK2 y JNK3): quinasas de serina/treonina involucradas en procesos de proliferación, diferenciación, migración, transformación y apoptosis. Participan en la cascada de quinasas *JNK* (*c-Jun N-terminal kinase*) activada por estímulos extracelulares como citoquinas pro-inflamatorias y estrés físico. Promueven la apoptosis al fosforilar factores clave como *TP53*. (JNK1: 427 aa, JNK2: 424 aa, JNK3: 464 aa)

Quinasa activada por mitógenos 14 (MAPK 14): quinasa de serina/treonina que actúa como componente esencial de la vía de las MAPK, es una de las cuatro quinasas p38 que tienen un rol importante en las cascadas de señalización iniciadas por estímulos extracelulares como citoquinas pro-inflamatorias y estrés físico conduciendo a la activación directa de factores de transcripción. En conjunto

fosforilan un amplio rango de proteínas, estimándose de 200 a 300 sustratos. (360 aa)

Aurora quinasa A: quinasa de serina/treonina que contribuye a la regulación de la progresión del ciclo celular. Asociada con el centrosoma durante la mitosis, juega un rol en el establecimiento del uso mitótico, duplicación del centrosoma, alineamiento de los cromosomas y citoquinesis. Requerida para la activación de *CDK1* en el centrosoma. (403 aa)

Caseína quinasa 2, subunidad α : subunidad catalítica de la quinasa de serina/treonina activada constitutivamente que fosforila un largo número de sustratos conteniendo residuos acídicos C-terminales a serina o treonina fosforilada. Regula numerosos procesos celulares como la progresión del ciclo celular, apoptosis y transcripción, así como infecciones virales. Actuaría como nodo regulatorio que integra y coordina numerosas señales. Durante la mitosis funciona como parte del punto de control dependiente de p53. (391 aa)

Quinasas asociadas a Rho 1 y 2 (ROCK1 y ROCK2): quinasas de serina/treonina, actúan como reguladores clave del citoesqueleto de actina y polaridad celular, involucradas en la regulación de la contracción del músculo liso, formación de fibras de estrés y adhesión focal, retracción de neuritas, adhesión celular y motilidad. (ROCK1: 1354 aa, ROCK2: 1388 aa)

Quinasa asociada a CDC7: quinasa de serina/treonina, regularía la transición G1/S del ciclo celular y/o la replicación del ADN. (574 aa)

Serina/treonina quinasa B-Raf: miembro de la familia de quinasas Raf de transducción de señales de crecimiento, participa en la regulación de la cascada de las MAP quinasas, afectando la división y diferenciación celular. Mutaciones en el gen de esta proteína se asocian a varios cánceres. (766 aa)

Glucógeno sintasa quinasa 3 beta (GSK-3 β): serina/treonina quinasa constitutivamente activa que actúa como regulador negativo en el control hormonal de la homeostasis de la glucosa, señalización *Wnt* y regulación de factores de transcripción y microtúbulos, al fosforilar e inactivar varias proteínas. (420 aa)

Quinasa dependiente de calcio/calmodulina tipo 2 beta: quinasa de serina/treonina, funciona autonómicamente luego de la unión a calcio/calmodulina y autofosforilación, en neuronas actúa cascada abajo de los receptores de *NMDA* para promover la formación de dendritas y sinapsis. (666 aa)

Quinasa de serina dependiente de calcio/calmodulina (CASK): pseudoquinasa de serina/treonina, se ha descrito como proteína de anclaje con un rol en el anclaje de proteína a membranas sinápticas y tráfico de canales iónicos. Contribuye al desarrollo neuronal, une proteínas de superficie neuronal como la proteína precursora de amiloide. Mediaría la interacción de la matriz extracelular con el citoesqueleto de actina. (921 aa)

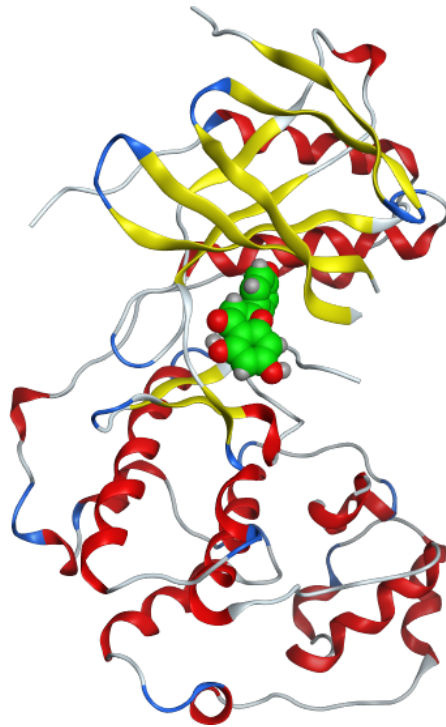


Figura 22. Representación de la estructura tridimensional del complejo *quinasa activada por mitógenos 14/quercetina* (verde) (PDB id: 4eh3).

Quinasa CLK1: quinasa de especificidad dual actuando tanto en residuos de serina/treonina como tirosina, en el núcleo fosforila proteínas ricas en serina/arginina involucradas en el procesamiento del pre-ARNm. (526 aa)

Quinasa DYRK1A (*dual specificity tyrosine-phosphorylation-regulated kinase 1A*): fosforila residuos de serina, treonina y tirosina en sustratos como *FOXO1* y *SIRT1*. Jugaría un rol en la regulación de funciones nucleares en la proliferación. (763 aa)

Quinasa HCK: tirosina quinasa que se encuentra en las células hematopoyéticas y juega un rol en la regulación de la respuesta inmune innata, sobrevivencia y proliferación celular, adhesión y migración. Durante los procesos de fagocitosis, media la movilización de lisosomas y activación de la *NADPH oxidasa*. Juega un rol en la liberación de moléculas inflamatorias. (526 aa)

Quinasa de la proteína quinasa dependiente de calcio-calmodulina 2: tirosina quinasa que tendría un rol en el desarrollo de neuritas. (588 aa)

Tirosina quinasa JAK1 (*quinasa de Janus 1*): involucrada en la vía de señalización por interferón alfa/beta/gamma. Quinasa pareja del receptor de interleuquina 2. (1154 aa)

Tirosina quinasa JAK2 (*quinasa de Janus 2*): involucrada en procesos de crecimiento celular, desarrollo, diferenciación y modificación de histonas. Media eventos de señalización esenciales a la respuesta inmune innata y adaptativa. Participa en la traducción de señales por receptores de tipo II como *IFN-alfa/beta/gamma*, etc. (1132 aa)

Tirosina quinasa JAK3 (*quinasa de Janus 3*): involucrada en procesos de crecimiento celular, desarrollo y diferenciación. Media eventos esenciales a la respuesta inmune innata y adaptativa y juega un rol en la hematopoyesis durante el

desarrollo de las células T. Participa en la traducción de señales por receptores de tipo I como *IFN-alfa/beta/gamma*, etc. (1124 aa)

Tirosina quinasa *BMX*: participa en la regulación de la reorganización de actina, migración celular, supervivencia y proliferación celular, adhesión y apoptosis. También juega un rol en la respuesta adaptativa contra estrés extracelular como por ejemplo en células endoteliales. (675 aa)

Tirosina quinasa *Src*: activada por varias clases de receptores celulares incluyendo receptores de la respuesta inmune, integrinas, receptores tirosina quinasa, receptores acoplados a proteínas G, etc. Participa en vías de señalización que controlan la respuesta inmune, adhesión celular, propagación del ciclo celular, apoptosis, migración y transformación. (536 aa)

Tirosina quinasa *Abl1*: juega un rol en procesos de crecimiento celular y supervivencia tales como remodelado del citoesqueleto en respuesta a estímulos externos, motilidad y adhesión celular, endocitosis, autofagia, respuesta al daño celular y apoptosis. (1149 aa)

Tirosina quinasa específica de leucocitos *Lck*: juega un rol esencial en la selección y maduración de células T en desarrollo en el timo y en la función de células T maduras, por ejemplo en el reconocimiento de antígeno. (509 aa)

Tirosina quinasa *ITK-TSK (interleukin-2-inducible T-cell kinase)*: juega un rol esencial en la regulación de la respuesta inmune adaptativa, específicamente en la proliferación y diferenciación de células T. (620 aa)

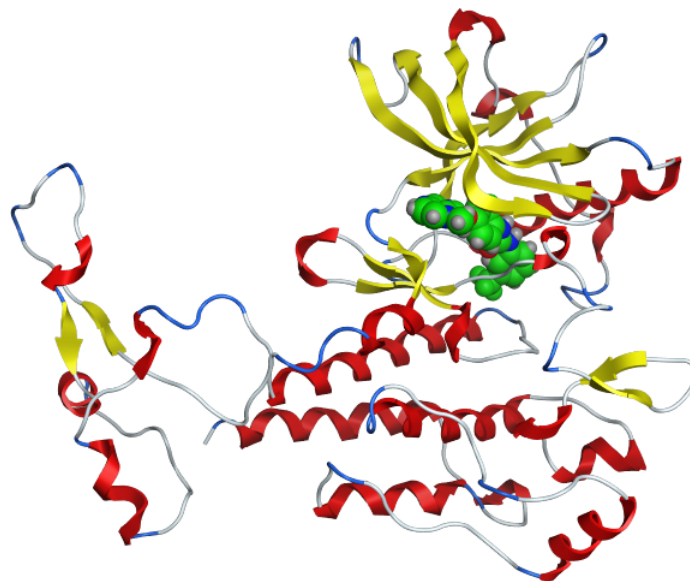


Figura 23. Representación de la estructura tridimensional del complejo *receptor 2 del factor de crecimiento endotelial vascular* (dominio quinasa)/inhibidor pirrolopirimidina (verde) (PDB id: 3vhe).

Receptor 2 del factor de crecimiento endotelial vascular: receptor tirosina quinasa, su actividad se restringe principalmente a células del endotelio vascular donde juega un rol esencial en la regulación de la angiogénesis. Varios inhibidores de VEGFR están en investigación para el tratamiento de varios cánceres. (1356)

Receptores 1 y 2 del factor de crecimiento de fibroblastos (FGFR1 y FGFR2): tirosinas quinasas que actúan como receptores de factores de crecimiento de fibroblastos y juegan un rol esencial en el desarrollo embrionario, diferenciación y migración celular. (FGFR1: 853 aa, FGFR2: 822 aa)

Receptor tirosina quinasa A3 de efrina (EPHA3): une promiscuamente ligandos unidos a membrana de la familia de las efrinas en células adyacentes, conduciendo a señales bidireccionales en las células vecinas. Regula la adhesión celular, la organización del citoesqueleto y migración. (983 aa)

Receptor tirosina quinasa B2 de efrina (EPHB2): une promiscuamente ligandos unidos a membrana de la familia de las efrinas en células adyacentes, conduciendo a señales bidireccionales en las células vecinas. Funciona en el guiado del axón durante el desarrollo. (987 aa)

(Endoplasmina): chaperona molecular que funciona en el procesamiento y transporte de proteínas secretadas. Al asociarse con CNPY3 es requerida para el plegamiento adecuado de receptores tipo Toll. Funciona en la degradación asociada al retículo endoplásmico. Possee actividad ATPasa. (804 aa)

(Proteína de shock térmico de 90 kDa alfa): chaperona molecular que promueve la maduración, mantenimiento estructural y regulación apropiada de proteínas blanco específicas involucradas en el control del ciclo celular y transducción de señales. Sufre un ciclo funcional esencial asociado a su actividad ATPasa. Este ciclo probablemente induce cambios conformacionales en las proteínas “cliente” causando entonces su activación. Uno lipopolisacáridos bacterianos y media una respuesta inflamatoria asociada, incluyendo la secreción de TNF por monocitos. (732 aa)

Tanquirasas 1 y 2: poli-ADP-ribosiltransferasas involucradas en varios procesos como la vía de señalización *Wnt*, extensión de los telómeros y tráfico vesicular. Actúan como un activadores de la vía *Wnt* al mediar la poli-ADP-ribosilación (*PARsylation*) de *AXIN1* y *AXIN2*, dos componentes claves de la destrucción del complejo beta-catenina (las proteínas poli-ADP-ribosiladas son marcadas para su ubiquitinación y degradación). (1327 aa, 1166 aa).



Figura 24. Representación de la estructura tridimensional del complejo *tanquirasa 2/luteolina* (verde) (PDB id: 4hkn).

Poli-(ADP-ribosa) polimerasa 1: involucrada en la vía de reparación por escisión de bases (*BER*), catalizando la poli-ADP-ribosilación de un número limitado de proteínas aceptoras involucradas en la arquitectura de la cromatina y metabolismo del ADN. Las *PARPs* pueden activarse en células que experimenten estrés o daño al ADN. La *PARP* activada puede depletar la célula de *ATP* en un intento por reparar el ADN dañado, lo que conduce a lisis y muerte celular (necrosis) [226]. También tienen la habilidad de inducir apoptosis mediante un mecanismo independiente de caspasas. (1014 aa)

Poli-(ADP-ribosa) polimerasa 2: involucrada en la vía de reparación por escisión de bases (*BER*), catalizando la poli-ADP-ribosilación de un número limitado de proteínas aceptoras involucradas en la arquitectura de la cromatina y metabolismo del ADN. Esta modificación sigue al daño del ADN y aparece como paso obligatorio en la vía de señalización y detección que conduce a la reparación de las rupturas de la hebra. (583 aa).

Poli-(ADP-ribosa) polimerasa 3: involucrada en la vía de reparación por escisión de bases (*BER*), catalizando la poli-ADP-ribosilación de un número limitado de proteínas aceptoras involucradas en la arquitectura de la cromatina y metabolismo del ADN. Infiere negativamente en la progresión G1/S del ciclo celular sin interferir con la duplicación del centrosoma. (540 aa)

Poli-(ADP-ribosa) polimerasa 10: se requeriría para la mantención de la progresión del ciclo celular (1025 aa).

Poli-(ADP-ribosa) polimerasa 12: (701 aa)

Poli-(ADP-ribosa) polimerasa 14: (1801 aa)

Poli-(ADP-ribosa) polimerasa 15: represor transcripcional. (678 aa)

Poli-(ADP-ribosa) polimerasa 16: mono-ADP-ribosiltransferasa que juega un rol en la respuesta a proteínas desplegadas. (323 aa)

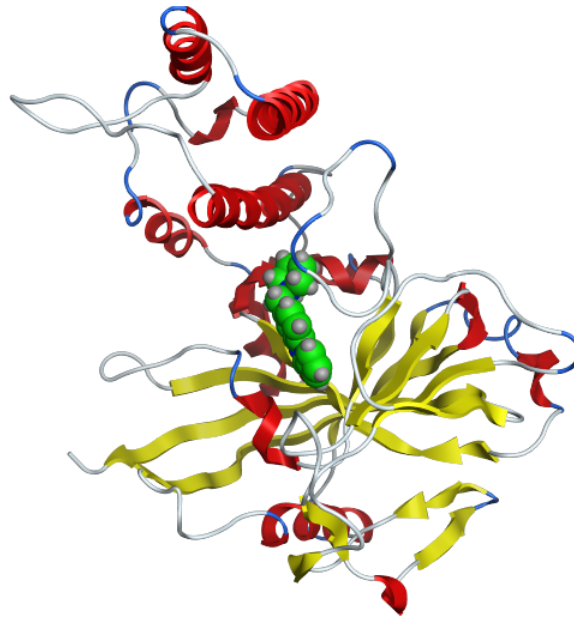


Figura 25. Representación de la estructura tridimensional del complejo *poli-(ADP-ribosa) polimerasa 1/inhibidor quinoxalina* (verde) (PDB id: 3gjlw).

Sirtuína-3 deacetilasa dependiente de NAD: activa o desactiva proteínas blanco mitocondriales al desacetilar residuos de lisina. Contribuye a la regulación del metabolismo energético. (399 aa)

Receptor activado por el proliferador de peroxisomas δ : receptor nuclear que une agentes proliferadores de peroxisomas como drogas hipolipidemicas y ácidos grasos. Una vez activado por un ligando, se une a regiones específicas del ADN donde modula la transcripción de genes blancos tales como acil-CoA oxidasa. Participa de la homeostasis intestinal al suprimir las respuestas pro-inflamatorias mediadas por NF-kappaB. (505 aa)

(Receptor de estrógenos alfa): receptor nuclear de hormonas. Las hormonas esteroideas y sus receptors están involucrados en la regulación de la expresión génica y afectan la proliferación y diferenciación de los tejidos blanco. La transactivación dependiente del ligando involucra tanto la unión directa del homodímero a elementos de respuesta a estrógeno del ADN o asociación con otros factores de transcripción como AP-1. Existe una represión mutua entre el *receptor de estrógenos* y *NF-kappaB* dependiente de tejido. (595 aa)

(Receptor de andrógenos): receptor de hormonas esteroideas que son factores de transcripción activados por ligando que regulan la expresión génica y afecta la proliferación y diferenciación de tejidos blanco. (920 aa)

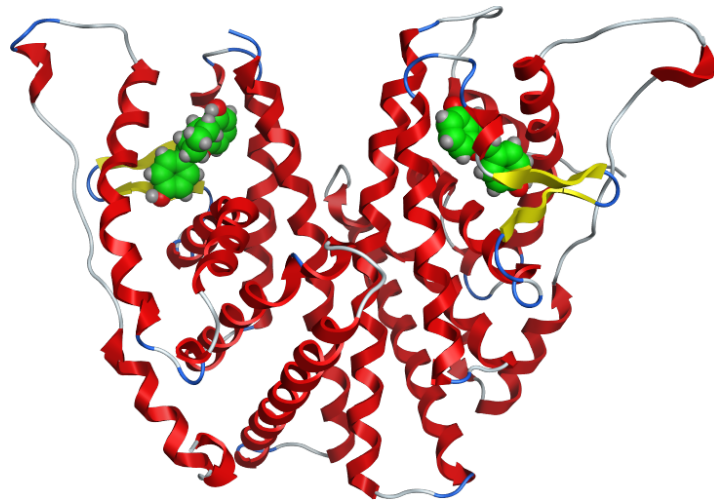


Figura 26. Representación de la estructura tridimensional del complejo *receptor de estrógenos alfa* (dímero)/agonista benzopirano (verde) (*PDB id: 2pog*).

NAD(P)H: quinona oxidorreductasa 2: funciona como quinona reductasa en conexión con reacciones de conjugación de hidroquinonas involucradas en vías de detoxificación así como procesos biosintéticos tales como la decarboxilación de residuos glutamato en la síntesis de protrombina. (231 aa)

Xantina oxidasa: enzima clave en la degradación de las purinas, cataliza la oxidación de hipoxantina a xantina y la oxidación de xantina a ácido úrico. Contribuye a la generación de especies reactivas del oxígeno. (1333 aa)

Leucotrieno A4 hidrolasa: epóxido hidrolasa que cataliza el paso final de la síntesis del mediador pro-inflamatorio leucotrieno B4. También tiene actividad aminopeptidasa. (611 aa)

3-Hidroxitiril-CoA hidrolasa: hidroliza el catabolito *3-hidroxitiril-CoA*, participaría en el catabolismo de valina. (386 aa)

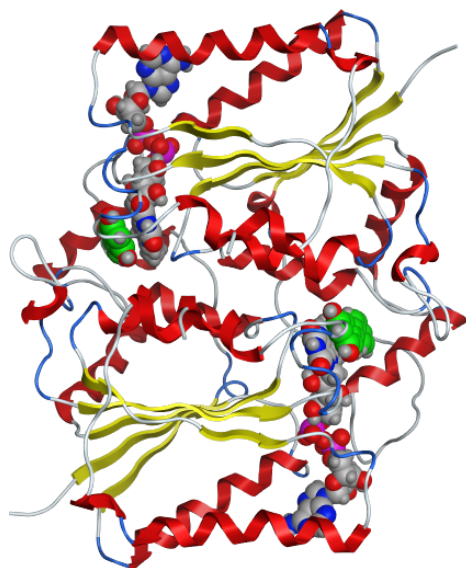


Figura 27. Representación de la estructura tridimensional del complejo *NAD(P)H: quinona oxidorreductasa 2* (dímero)/resveratrol (verde) y cofactor FAD (gris) (*PDB id: 1sg0*).

Sulfotransferasa 1A1: utiliza 3'-fosfoadenosina-5'-fosfosulfato (PAPS) como donador de sulfonato para catalizar la conjugación de catecolaminas, compuestos fenólicos y neurotransmisores. También tiene actividad estrógeno sulfotransferasa, responsable de la sulfatación y activación de minoxidil. La sulfatación incrementa la solubilidad de la mayoría de los compuestos y por lo tanto su excreción renal, pero también puede resultar en la activación de metabolitos. Media la activación de hidroxilamidas a productos de unión al ADN, participando entonces como un modulador del riesgo de cáncer. (295 aa)

Sulfotransferasa 1A3: cataliza la conjugación de aminas fenólicas (neurotransmisores como dopamina y serotonina) así como drogas fenólicas. (295 aa)

Sulfotransferasa 1B1: cataliza la conjugación de varias hormonas, neurotransmisores, drogas y compuestos xenobióticos. Sulfata dopamina, hormonas tiroideas y pequeños fenoles entre otros. (296 aa)

Sulfotransferasa 1C2: cataliza la conjugación de varias hormonas, neurotransmisores, drogas y compuestos xenobióticos. Puede estar involucrada en la activación de hidroxil aminas carcinogénicas. (307 aa)

Sulfotransferasa 1C3: tienen baja actividad sulfotransferasa hacia sustratos con grupos alcohol (*in vitro*), catalizaría la conjugación de compuestos xenobióticos. (304 aa)

Sulfotransferasa 1C4: cataliza la conjugación de drogas, xenobióticos, hormonas y neurotransmisores. Puede estar involucrada en la activación de hidroxil aminas carcinogénicas. (302 aa)

Sulfotransferasa 1E1 (estrógeno sulfotransferasa): cataliza la conjugación de estradiol y estrona. Tendría un rol en la regulación de la actividad del receptor de estrógenos al metabolizar estradiol libre. Sulfata otras moléculas similares a excepción de cortisol, testosterona y dopamina. (294 aa)

Metionina adenosil transferasa 2, subunidad β : subunidad regulatoria de la sintetasa correspondiente. (334 aa)

Factor inhibidor de la migración de macrófagos (MIF): citoquina pro-inflamatoria involucrada en la respuesta inmune contra patógenos bacterianos. La expresión de MID en los sitios de inflamación sugiere un rol en la regulación de la función de los macrófagos. Contrarresta la actividad anti-inflamatoria de los glucocorticoides. Tiene actividad tautomerasa de fenil piruvato y dopacromo (*in vitro*), pero el sustrato fisiológico es desconocido. (155 aa)

α -Amilasa pancreática: cataliza la hidrólisis de polisacáridos. (511 aa)

Glucógeno fosforilasa: enzima importante en el metabolismo de los carbohidratos. (842 aa)

Transtiretina: proteína de unión a hormonas tiroideas, transportaría tiroxina por el flujo sanguíneo al cerebro. (147 aa)

Troponina C: proteína de unión a calcio central en la regulación de la contracción muscular. (161 aa)

ATP sintasa: complejo mitocondrial (complejo V) que produce ATP a partir de ADP en presencia de un gradiente de protones membranal que es producido por la cadena de transporte de electrones. Consiste de dos dominios estructurales, F(1) conteniendo el núcleo catalítico extramembranal, y F(0) conteniendo el canal de protones, unidos por un tallo central y periférico. Durante la catálisis, la síntesis de ATP en el dominio catalítico se acopla a un mecanismo rotatorio de las subunidades centrales para la translocación de protones.

(Piruvato deshidrogenasa quinasa): quinasa que juega un rol en la regulación del metabolismo de glucosa y ácidos grasos las subunidades de la piruvato deshidrogenasa. (407 aa)

◦ *Minería de texto y otras fuentes de anotación automática*

La información estructural y funcional, especialmente sobre la participación de las proteínas blanco en procesos biológicos y enfermedades puede asociarse a los efectos de quercetina y los flavonoides en general (tabla 4). Recurriendo a las anotaciones de ontología funcional (GO), se observa que el conjunto de blancos está significativamente enriquecido en funciones asociadas a muerte y proliferación celular. Por otro lado, las vías *KEGG* resaltan su participación en cáncer y neurodegeneración. El rol de algunas proteínas blanco en la fisiología neuronal también es notable, dada su participación en la señalización por neurotrofinas, diferenciación neuronal y función sináptica. Estos se corresponden con efectos ya reportados de los flavonoides [227–229].

Tabla 4. Asociación entre proteínas blanco identificadas en el *PDB*, enfermedades y procesos biológicos vinculados. Las proteínas blanco se han separado de acuerdo al organismo de origen y anotado según su entrada correspondiente en Uniprot (con su homólogo humano en paréntesis), nombre, clase del *PDB*, enfermedad y proceso biológico asociado. Las proteínas co-cristalizadas con quercetina se muestran remarcadas. La información fue recopilada de bases de datos de libre acceso como *PDB*, Uniprot y repositorios de ontología génica.

Uniprot id	Nombre	Enfermedad relacionada	Proceso biológico relacionado	Clase (PDB)
Homo sapiens				
Q00534	Quinasa dependiente de ciclina 6	Cáncer		Ciclo celular/transferasa
Q9H2K2	Tanquirasa 2			Transferasa/inhibidor de transferasa
P49888	Sulfotransferasa 1E1 (estrógeno sulfotransferasa)	Cáncer		Transferasa
O75897	Sulfotransferasa 1C4			Transferasa
O00338	Sulfotransferasa 1C2	Cáncer		Transferasa

P50225	Sulfotransferasa 1A1	Cáncer		Transferasa
O43704	Sulfotransferasa 1B1			Transferasa
P09960	Leucotrieno A4 hidrolasa	Enfermedades inmunes		Hidrolasa
P49759	Quinasa CLK1		Proliferación	Transferasa
PODMM9	Sulfotransferasa 1A3	Cáncer		Transferasa
P35968	Receptor 2 del factor de crecimiento endotelial vascular	Cáncer	Muerte celular/proliferación	Transferasa/inhibidor de transferasa
P16083	NAD(P)H: quinona oxidoreductasa 2	Cancer, enfermedad de Parkinson		Oxidoreductasa
Q13554	Quinasa dependiente de calcio/calmodulina tipo 2 beta	Cáncer	Muerte celular/ asociado a neurona	Transferasa
O95271	Tanquirasa 1			Transferasa/inhibidor de transferasa
Q6IMI6	Sulfotransferasa 1C3			Transferasa
Q13627	Quinasa DYRK1A			Transferasa
Q460N3	Poli-[ADP-ribosa] polimerasa 15			Transferasa
Q460N5	Poli-[ADP-ribosa] polimerasa 14			Transferasa
P09874	Poli-[ADP-ribosa] polimerasa 1	Enfermedad de Alzheimer		Transferasa/inhibidor de transferasa
O60674	Tirosina quinasa JAK2	Cáncer	Muerte celular/proliferación/ asociado a neurona	Transferasa/inhibidor de transferasa
P68400	Caseína quinasa 2, subunidad alfa	Enfermedad de Parkinson		Transferasa
Q9NZL9	Metionina adenosil transferasa 2, subunidad β			Oxidoreductasa
P15056	Quinasa de serina/treonina B-raf	Cáncer	Muerte celular/proliferación/ asociado a neurona	Transferasa/inhibidor de transferasa
Q9UGN5	Poli-[ADP-ribosa] polimerasa 2		Muerte celular	Transferasa
O14965	Aurora quinasa A	Cáncer	Muerte celular/ asociado a neurona	Transferasa/inhibidor de transferasa
P23458	Tirosina quinasa JAK1	Cáncer	Proliferación	Transferasa/inhibidor de transferasa
O43293	Quinasa DAPK3	Cáncer	Muerte celular/ asociado a neurona	Transferasa
P11362	Receptor 1 del factor de crecimiento de fibroblastos	Cáncer	Muerte celular/proliferación/ asociado a neurona	Transferasa/inhibidor de transferasa
Q96RR4	Quinasa de la proteína quinasa dependiente de calcio-calmodulina 2	Desorden de Ansiedad		Transferasa/inhibidor de transferasa
(P10275)	(Receptor de andrógenos)		(Proliferación)	(Hormona /Receptor de Factor de Crecimiento)
(P03372)	(Receptor de estrógenos alfa)		(Proliferación)	(Proteína de Unión a Lipidos)
P49841	Glucógeno sintasa quinasa 3 beta	Enfermedad de Alzheimer, cáncer	Muerte celular/ asociado a neurona	Transferasa
O94768	Quinasa de serina/treonina 17b		Muerte celular	Transferasa
O14757	Quinasa de serina/treonina Chk1	Cáncer	Proliferación	Transferasa
P24941	Quinasa dependiente de ciclina 2	Cáncer		Transferasa
P51813	Tirosina quinasa BMX		Muerte celular/proliferación	Transferasa/inhibidor de transferasa
P52333	Tirosina quinasa JAK3		Muerte celular/proliferación	Transferasa
Q16539	Quinasa activada por mitógenos 14	Enfermedad de Parkinson, esclerosis lateral amiotrófica	Muerte celular/ asociado a neurona	Transferasa/inhibidor de transferasa
P45983	Quinasa activada por mitógenos 8	Cáncer, Enfermedad de	Muerte celular/ asociado	Transferasa/inhibidor de

		Parkinson	a neurona	transferasa
P08631	Quinasa HCK	Enfermedades inmunes, enfermedad de Parkinson	Muerte celular	Transferasa
P12931	Tirosina quinasa Src	Cáncer	Muerte celular/proliferación/ asociado a neurona	Tirosina-proteína quinasa
P50750	Quinasa dependiente de ciclina 9		Proliferación	Transferasa
O14936	Quinasa de serina dependiente de calcio/calmodulina (CASK)	Microcefalia	Proliferación	Transferasa
P37231	Receptor activado por el proliferador de peroxisomas gamma	Diabetes, cáncer, enfermedad de Alzheimer, enfermedad de Huntington	Muerte celular	Regulador de la transcripción
Q9Y6E0	Quinasa de serina/treonina 24 (quinasa MST3)		Muerte celular/ asociado a neurona	Transferasa
P21802	Receptor 2 del factor de crecimiento de fibroblastos	Cáncer	Muerte celular/proliferación/ asociado a neurona	Transferasa
P11309	Quinasa de serina/treonina PIM1		Muerte celular	Transferasa
O15530	Quinasa de serina/treonina 1 dependiente de fosfoinositol	Cáncer	Muerte celular/ asociado a neurona	Transferasa
P00519	Tirosina quinasa Abl1	Cancer	Muerte celular/proliferación/ asociado a neurona	Transferasa
P06239	Tirosina quinasa específica de leucocitos Lck	Diabetes	Muerte celular/proliferación/ asociado a neurona	Transferasa
Q08881	Tirosina quinasa ITK-TSK	Enfermedades inmunes	Proliferación	Transferasa/inhibidor de transferasa
Q9H0J9	Poli-[ADP-ribosa] polimerasa 12			Transferasa
P45984	Quinasa activada por mitógenos 9	Cáncer	Muerte celular/ asociado a neurona	Transferasa
Q8N5Y8	Poli-[ADP-ribosa] polimerasa 16		Muerte celular	Transferasa/inhibidor de transferasa
Q6NVY1	3-Hidroxiubutiril-CoA hidrolasa			Hidrolasa
P29320	Receptor tirosina quinasa A3 de efrina (EPHA3)		Asociado a neurona	Transferasa
P14174	Factor inhibidor de la migración de macrófagos (MIF)	Enfermedades inmunes, enfermedad de Alzheimer	Muerte celular	Isomerasa
P63316	Troponina C		Muerte celular	Proteína contractil
(P07900)	(Proteína de shock térmico de 90 kDa)		(Asociado a neurona)	(Chaperona)
Q13464	Quinasas asociadas a Rho 1	Cáncer	Muerte celular/ asociado a neurona	Transferasa
Q99683	Quinasa quinasa quinasa activada por mitógenos 5	Esclerosis lateral amiotrófica	Muerte celular/ asociado a neurona	Transferasa/inhibidor de transferasa
P49137	Quinasa 2 activada por MAPK		Asociado a neurona	Transferasa
P53779	Quinasa activada por mitógenos 10	Encefalopatía epiléptica	Asociado a neurona	Transferasa/inhibidor de transferasa
O00311	Quinasa asociada a CDC7		Proliferación	Transferasa
P04746	Alfa-amilasa	Enfermedad de Alzheimer, enfermedad de Parkinson		Hidrolasa/inhibidor de hidrolasa
P02766	Transtiretina	Amiloidosis		Proteína de transporte
(Q15119)	(Piruvato deshidrogenasa quinasa)			(Transferasa/inhibidor de transferasa)
Q9Y6F1	Poli-[ADP-ribosa] polimerasa 3	Cáncer		Transferasa/inhibidor de transferasa
Q9NTG7	Sirtuína-3 deacetilasa dependiente de NAD	Envejecimiento		Hidrolasa/inhibidor de hidrolasa

Q53GL7	Poli-[ADP-ribosa] polimerasa 10		Proliferación	Transferasa
Mus musculus				
P54763 (P29323)	Receptor tirosina quinasa B2 de efrina (EPHB2)	Cáncer		Transferasa
Rattus norvegicus				
P63086 (P28482)	Quinasa activada por mitógenos 1	Enfermedad de Alzheimer, cáncer	Muerte celular/proliferación/ asociado a neurona	Transferasa
Canis lupus familiaris				
P41148 (P14625)	Endoplasmina			
Oryctolagus cuniculus				
P00489 (P11217)	Glucógeno fosforilasa			Transferasa/inhibidor de transferasa
Sus scrofa				
O02697 (P48736)	Fosfoinositol 3-quinasa, subunidad catalítica	Envejecimiento, cáncer	Muerte celular/ asociado a neurona	Fosfoinositol 3- quinase gamma
Bos taurus				
P80457 (P47989)	Xantina oxidasa		Muerte celular	Oxidoreductasa
Q28021 (O75116)	Quinasas asociadas a Rho 2	Cancer	Muerte celular/ asociado a neurona	Transferasa
P19483 (P25705)	ATP sintasa	Enfermedad de Alzheimer, enfermedad de Parkinson, enfermedad de Huntington		Hidrolasa

De manera complementaria se realizó una consulta a varios servidores de minería de texto especializados en la recuperación de entidades asociadas (vía co-ocurrencia) a uno a varios términos de búsqueda dados: *PolySearch/PolySearch2*, *GenCLiP*, *PALM-IST*, *SciMiner*, y *Slap*. Se utilizaron términos de búsqueda relativos a sobrevida neuronal, muerte neuronal y quercetina (molécula de origen vegetal con actividad antioxidante y protectora). A partir de los resultados de los servidores se generaron listas de *hits* correspondientes a entidades proteicas que fueron integradas y normalizadas manualmente. Para ser reportada, cada asociación debe originarse de un número significativo de artículos. Adicionalmente se conservaron solamente los hits presentes en el resultado de al menos dos servidores. Se destaca en primer lugar el conjunto de proteínas simultáneamente a procesos de muerte neuronal/neuroprotección y efectos de quercetina (tabla 5).

Tabla 5. Proteínas identificadas por análisis automático de texto (*text mining*) como asociadas simultáneamente a procesos de muerte neuronal, neuroprotección y quercetina.

Uniprot ID	Gene	Nombre
P10415	BCL2	Regulador de la apoptosis Bcl-2
P42574	CASP3	Caspasa-3

P09601	HMOX1	Hemo oxigenasa 1 (HO-1)
P05231	IL6	Interleuquina-6
P01375	TNF	Factor de necrosis tumoral (TNF-alpha)
Q07817	BCL2L1	Regulador de la apoptosis Bcl-X
P19838	NFKB1	Factor nuclear kappa B (NF-kappaB)
P31749	AKT1	Proteína quinasa B (AKT)
P05412	JUN	Proteína activadora 1 (AP-1)
P35354	PTGS2	Ciclooxigenasa-2 (COX-2)
P48736	PIK3CG	Fosfoinositol 3-quinasa
P45983	MAPK8	Quinasa activada por mitógeno 8 (MAPK8, JNK1)
P15692	VEGFA	Factor de crecimiento endotelial vascular A
P99999	CYCS	Citocromo c
P34932	HSPA4	Proteína de choque térmico Hsp70-4
P01584	IL1B	Interleuquina-1 beta
P04179	SOD2	Superóxido dismutasa [Mn], mitocondrial
P04040	CAT	Catalasa
P37231	PPARG	Receptor activado por el proliferador de peroxisomas δ
Q96EB6	SIRT1	Sirtuína-1 deacetilasa dependiente de NAD
P35228	NOS2	Óxido nítrico sintasa, inducible
P00441	SOD1	Superoxido dismutasa 1 [Cu-Zn]
Q07812	BAX	Regulador de la apoptosis BAX
Q16539	MAPK14	Quinasa activada por mitógeno 14 (p38)
P29475	NOS1	Óxido nítrico sintasa, constitutiva
P49768	PSEN1	Presenilina-1 (PS-1)
O00206	TLR4	Receptor tipo Toll-4 (TLR4)
P03372	ESR1	Receptor de estrógenos alfa
Q16665	HIF1A	factor 1 inducible por hipoxia-alfa (HIF-1-alpha)
P09874	PARP1	Poli-[ADP-ribosa] polimerasa 1
Q16236	NFE2L2	Factor nuclear eritroide-2 (Nrf2)
P00533	EGFR	Receptor del factor de crecimiento epidermal
Q92731	ESR2	Receptor de estrógenos beta
P11511	CYP19A1	Citocromo P450 19A1
P28482	MAPK1	Quinasa activada por mitógeno 1 (ERK 2)
P15559	NQO1	NAD(P)H: quinona oxidoreductasa 1
P09429	HMGB1	Amfoterina HMGB1
P40763	STAT3	Transductor de señal y activador de la transcripción 3 (STAT3)
P42345	MTOR	Quinasa de serina/treonina mTOR
P04637	TP53	Proteína p53
O15530	PDPK1	Quinasa de serina/treonina 1 dependiente de fosfoinositol
P05771	PRKCB	Proteína quinasa C, beta
Q13131	PRKAA1	Proteína quinasa A
P55211	CASP9	Caspasa-9
P24385, Q9H014	CCND1	Ciclina D1 (BCL-1)

Varias de las proteínas candidato presentadas en el tamizaje reverso se identifican dentro de este conjunto. Estas incluyen receptores como el receptor de estrógenos y proteínas de señalización como proteínas quinasas y la enzima PARP-1, los que se ubicarían cascada arriba y en el desarrollo de varias vías de señalización. Al respecto un diagrama de interacción proteína-proteína (STRING) de todos los blancos reportados ilustra múltiples asociaciones funcionales entre las proteínas candidatos, especialmente quinasas y poli-(ADP-ribosa) polimerasas apoyando su participación

en procesos biológicos relacionados. Dentro de las proteínas no identificadas se encuentran efectores finales de vías de señalización, como ser interleuquinas, caspasas y factores de transcripción.

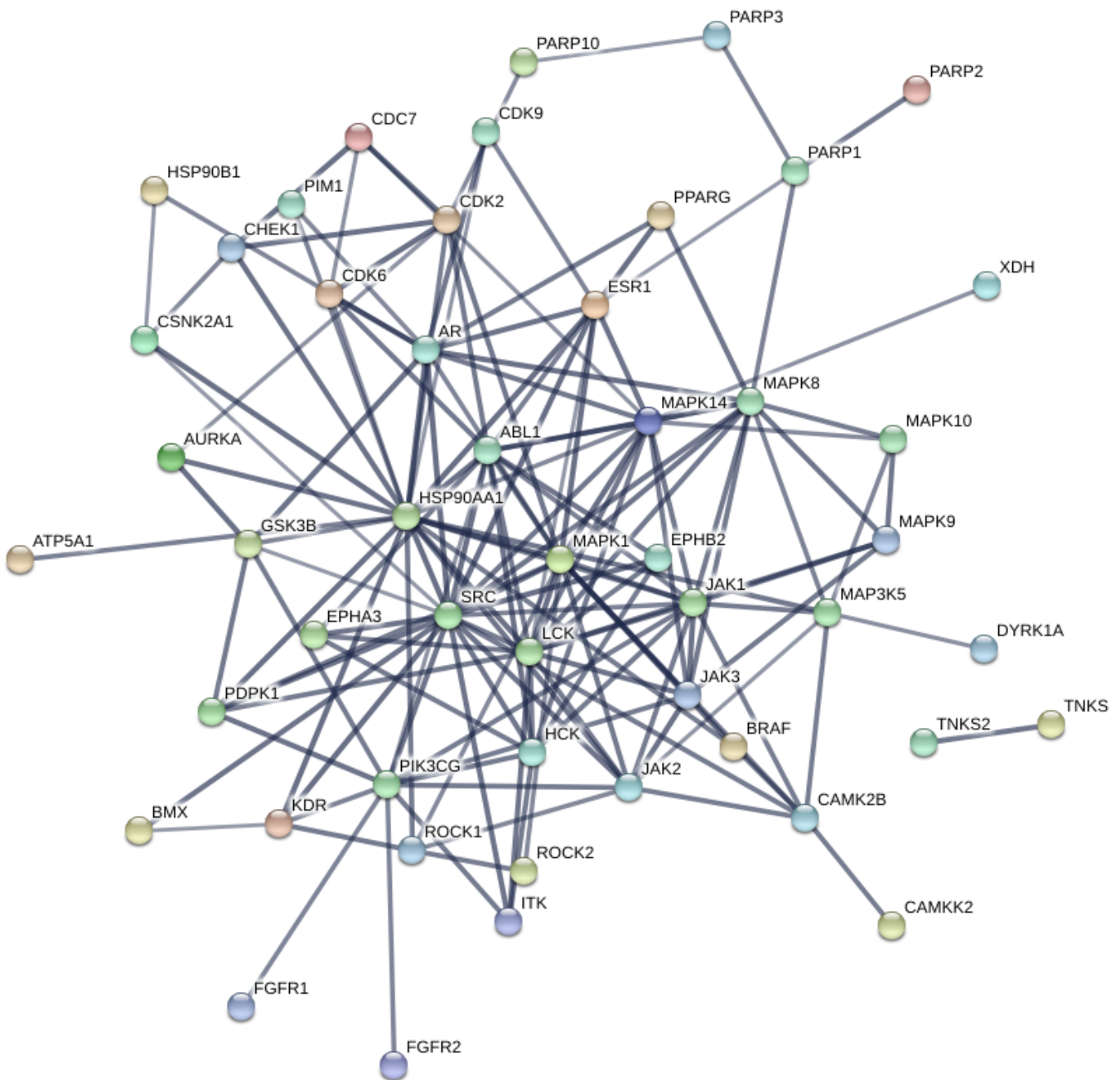


Figura 28. Representación de la red de interacción proteína-proteína de los candidatos a blanco. La base de datos STRING establece asociaciones funcionales entre pares de proteínas basado en evidencia experimental documentada, homología e inferencia por minería de texto incluyendo modificaciones proteicas, agregación, vecindad genómica y co-expresión. Solo asociaciones con alta confianza son presentadas (probabilidad mayor a 0,7).

Las proteínas quinasas deben ser destacadas, un gran número de las que han sido reconocidas como prometedores blancos terapéuticos para tumores y neurodegeneración incluyendo MAP quinasas, receptores tirosina quinasas y quinasas

relacionadas a la vía de señalización de PI3 quinasa/Akt/mTOR [230–237]. Por ejemplo, la MAPK14 (p38) es una proteína clave involucrada en la apoptosis neuronal y la inflamación, cuya inhibición ha sido propuesta como tratamiento contra el daño isquémico y la neurodegeneración [235]. En concordancia con nuestros resultados, estudios previos han mostrado que el tratamiento con quercetina conduce a efectos pro-apoptóticos en modelos de cáncer a través de la modulación de estas vías de quinasas [238]. Entre las proteínas que no han sido asociadas previamente a efectos de quercetina, se encuentran por ejemplo los receptores de factores de crecimiento como el FGFR (*fibroblast growth factor receptors*). Estos receptores tirosina quinasa juegan un rol importante en el desarrollo embrionario, proliferación, diferenciación, migración y están asociados a varios cánceres humanos [239]. La acción sobre este receptor tendría un efecto inhibitorio y sinérgico sobre vías de supervivencia dependientes de *fosfoinositol 3-quinasa (PI3K)-Akt*. Una situación análoga es presentada por otro receptor tirosina quinasa, el receptor 1 del factor de crecimiento tipo insulina (*insulin-like growth factor 1 receptor*) que ha sido implicado en los efectos anticancerígenos del flavonoide luteolina [240].

Un grupo novedoso es el de las *tanquirasas*, un conjunto particular de *poli-ADP-ribosa polimerasas* que difieren en su estructura global y funciones. Se ha sugerido que los inhibidores de tanquirasas pueden mejorar efectos anti-cancerígenos si se combinan con otros inhibidores de quinasas como los de *MEK*, *receptor del factores de crecimiento epidérmico* o *fosfoinositol 3-quinasa* [241]. La proteína de señalización *PARP-1* es otro ejemplo de una asociación con proteínas quinasas. *PARP-1* facilita respuestas inflamatorias al promover la expresión génica por el factor *NF-kappaB* y promueve a muerte celular asociada a la mitocondria en tejidos lesionados [7]. Está ampliamente aceptado que *PARP-1* señala a la vía de las MAPK al modular la fosforilación de *ERK1/2*, *p38* y *quinasas c-Jun N terminal*. Hasta ahora los estudios sobre la relación entre *PARP-1* y MAPK sugieren que se estimularían recíprocamente para propagar respuestas de estrés a largo plazo. Este entrecruzamiento entre proteínas refuerza el concepto de un mecanismo sinérgico y a varios niveles subyacente a los efectos biológicos de los flavonoides.

Otro punto relevante es el pleiotropismo de varias proteínas blanco tanto en cáncer como en neuroprotección. La *glucógeno sintasa quinasa 3 beta* (GSK-3 β) es un ejemplo perfecto de un promotor blanco involucrado tanto en la supervivencia del cáncer como también promueve la exportación del núcleo y degradación de *Nrf2* (*Nuclear Factor-erythroid 2 (NF-E2) p45-related Factor-2*), un factor de transcripción que inicia respuestas antioxidantes y citoprotectoras [238, 242, 243]. Asimismo GSK-3 β estimula la actividad de *NF-kappaB* y como consecuencia activa la cascada inflamatoria. Recientemente, se ha asociado la inhibición de *GSK-3beta* con la actividad protectora del flavonoide baicaleína y el fármaco fasudil [236] [244]. Estudios previos también han mostrado que la inhibición de *GSK-3beta* en neuronas del hipocampo induce la formación de múltiples axones, una actividad que

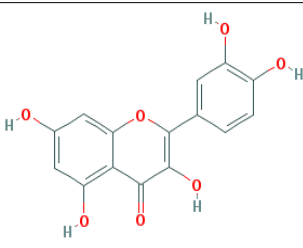
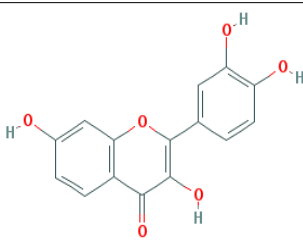
podría relacionarse a efectos neuritogénicos de quercetina [245]. Asimismo, esta proteína ilustra un posible efecto sinérgico de quercetina; por ejemplo en modelos de isquemia cerebral, la inhibición de MAPK14 promueve la sobrevivencia celular mediante la regulación de dos vías intrínsecas, activación de ERK1/2 y la inhibición de *GSK-3beta*.

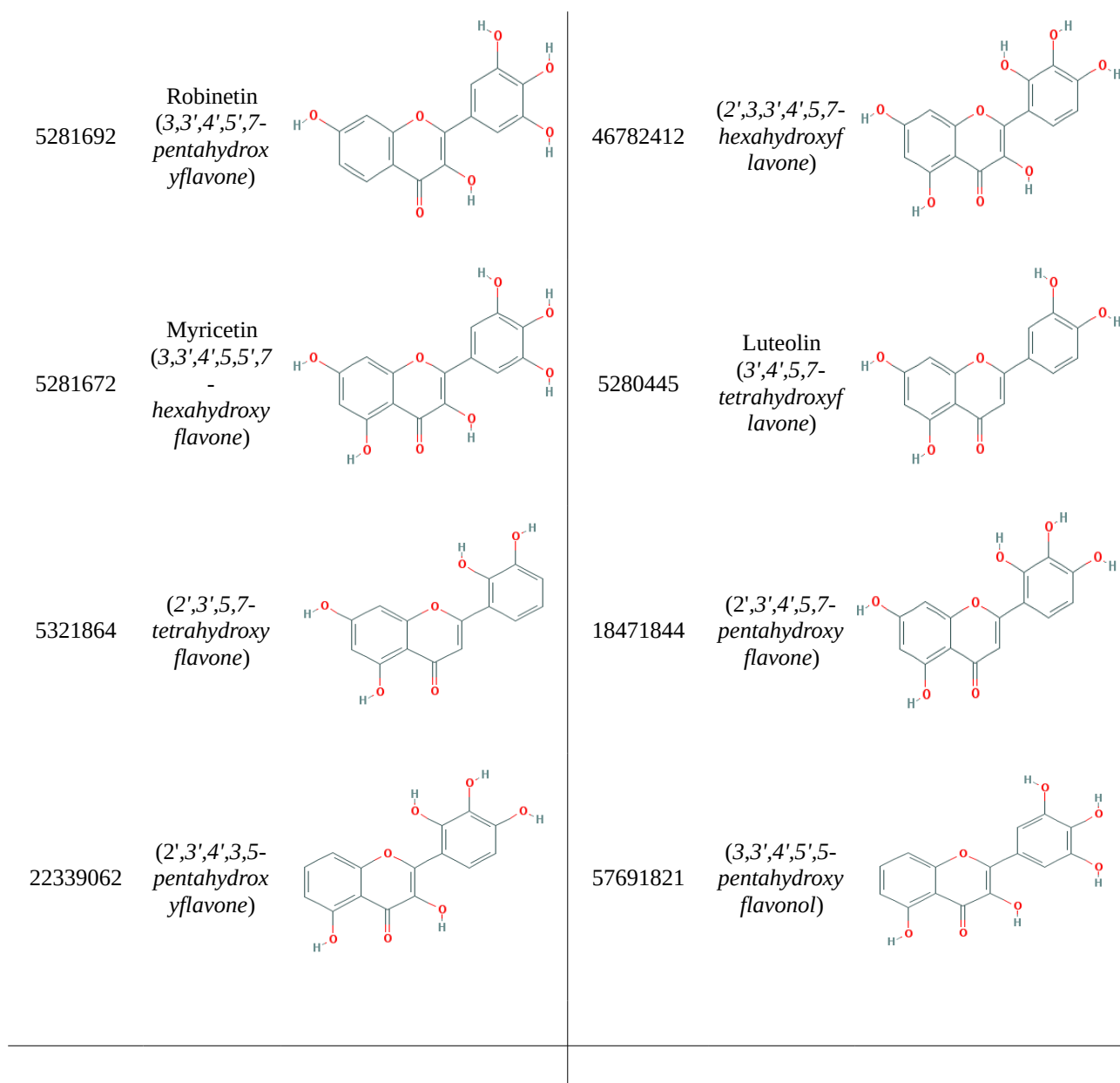
Finalmente, el metabolismo de los polifenoles también es insinuado por la presencia de enzimas oxidoreductasas y sulfotransferasas como *sulfotransferasa 1B1* y *NAD(P)H: quinona oxidoreductasa 2 (NQO2)*. La última funciona aparentemente como quinona reductasa en conexión con reacciones de conjugación de hidroquinonas de vías de detoxificación. La inhibición de *NQO2* por quercetina (K_d experimental de 50 nM), resveratrol ($K_d = 30$ nM) y otros polifenoles se ha sugerido tiene un rol en la inducción de defensas antioxidantes que pueden subyacer a sus efectos quimiopreventivos a largo plazo [246]. Como la actividad de esta enzima se limita a la excreción de polifenoles o también a la producción de moléculas biológicamente activas aún no está claro.

• Docking de quercetina y análogos

Los motivos de hidroxilación de quercetina puede constituir una característica fundamental para su interacción con proteínas blanco. En este aspecto un librería de quercetina y flavonoides análogos fue construida mediante una búsqueda por similitud estructural en *PubChem* con un coeficiente de Tanimoto de 98 % (tabla 6). La librería resultante incluye flavonoides conocidos como miricetina, fisetina, luteolina así como compuestos sintéticos que comparten el mismo esqueleto pero diferente patrón de hidroxilación.

Tabla 6. Análogos de quercetina obtenidos por búsqueda estructural en *PubChem*. Las moléculas fueron anotadas por su identificador *PubChem*, nombre y estructura bi-dimensional.

PubChem CID	Name	2D Representation	PubChem CID	Name	2D Representation
5280343	Quercetin (3,3',4',5,7-pentahydroxyflavone)		5281614	Fisetin (3,3',4',7-tetrahydroxyflavone)	



En una primera instancia quercetina y sus análogos fueron sometidos a *docking molecular* con *MOE*. Este incorpora la función de *score* basada en campo de fuerza *GBVI/WSA dG* que estima la energía libre de unión de una pose dada así como el efecto de la solvatación. *MOE* permite la representación explícita de co-factores y sustratos en la vecindad del sitio de unión lo que supone una mejora respecto de *idTarget*. Una comparación entre el *score* de *idTarget* y *MOE* es presentada en la figura 29 evidenciando una correlación significativa entre ambos sistemas.

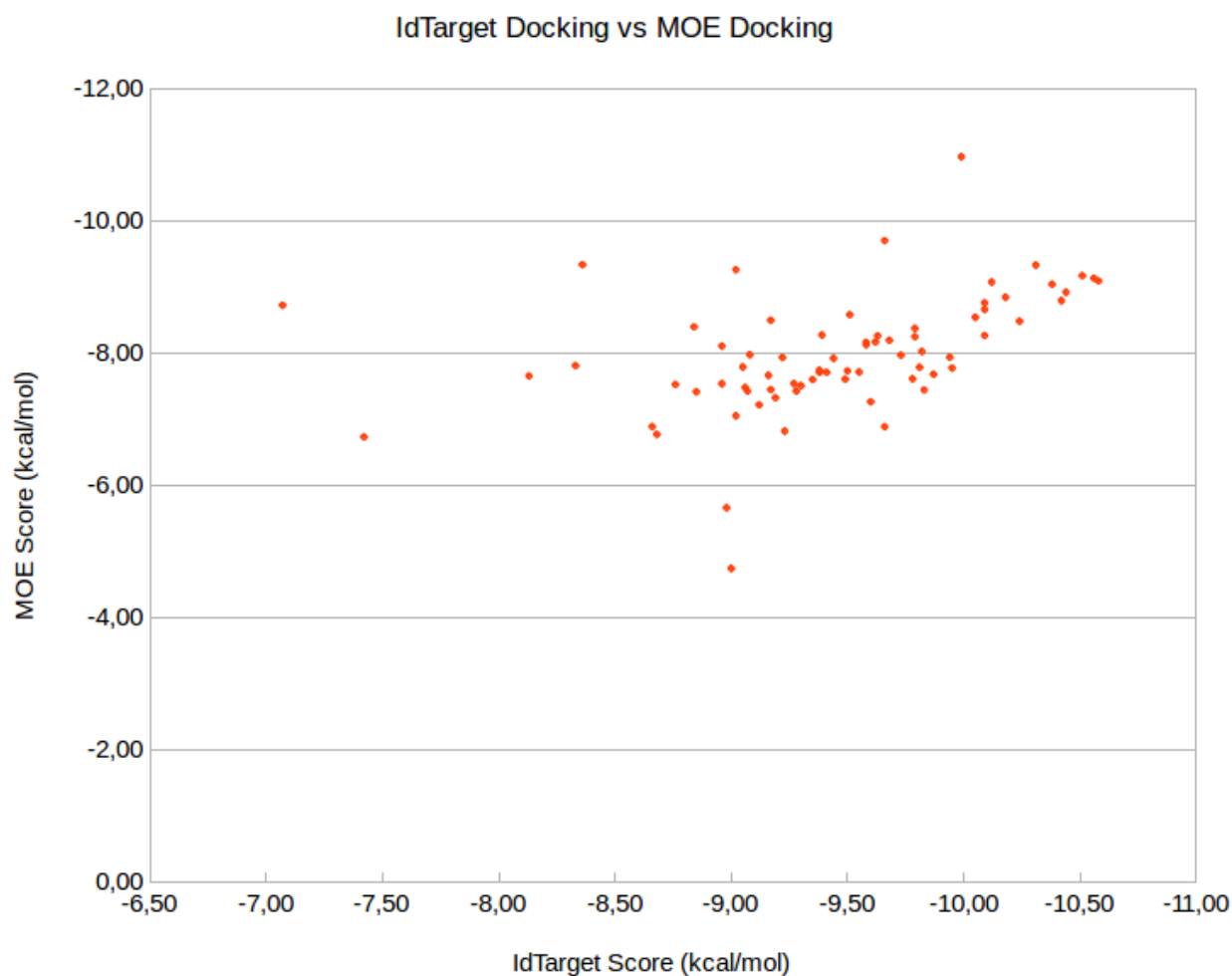


Figura 29. Comparación entre el score de *idTarget* y *MOE* (*Molecular Operating Environment*). Una sola estructura cristalográfica fue considerada para cada proteína en *MOE*. El mejor score obtenido en *MOE* fue comparado contra el mejor score obtenido en *idTarget*.

La existencia de alguna característica general de las proteínas blanco fue evaluada por medio de su patrón de interacción con las poses de *docking*. Los resultados fueron resumidos para cada grupo proteico reconocido por el análisis de *LIBRA* como quinasas, poli ADP-ribosa polimerasa – tanquirasas y sulfotransferasas (figura 30). Estos patrones buscan ser independientes de las hidroxilaciones específicas de quercetina al utilizar una librería de flavonoides análogos. Pocos residuos reúnen la mayor parte de las interacciones con los flavonoides anclados y fueron característicos de cada grupo de proteínas. Sobre todo los resultados confirman la preferencia por amino ácidos con grupos donadores o aceptores de enlace de hidrógeno así como aminoácidos aromáticos que establecen dichos contactos con los flavonoides. Por ejemplo, el par glutamato/aspartato prevalece tanto en quinasas como en *PARPs* junto con la participación de residuos apolares o aromáticos como valina y tirosina.

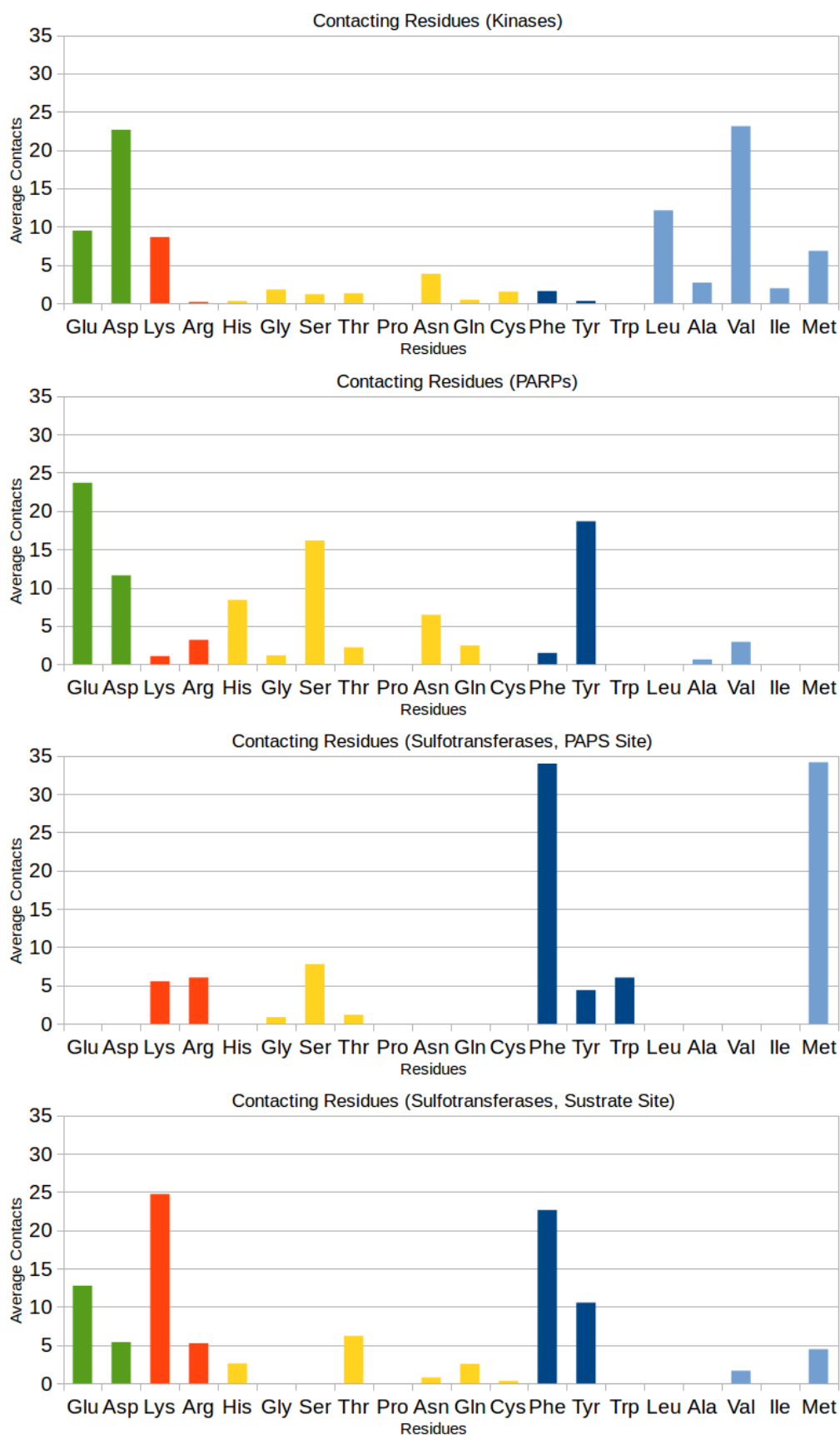


Figura 30. Patrón de interacción proteína-ligando entre las proteínas blanco identificadas en el *Protein Data Bank* y flavonoides. Los contactos fueron evaluados con el módulo *PLIF* (*Protein Ligand Interaction Fingerprint*) de *MOE* (*Molecular Operating Environment*). Se promediaron los

contactos por tipo de aminoácido y presentados por grupo de proteína reconocido por *LIBRA*. A: proteínas quinasas, B: poli ADP-ribosa polimerasas – tanquirasas, C: sulfotransferasas (sitio de unión a PAPS), D: sulfotransferasas (sitio de unión alostérico).

El amplio rango de interacciones de quercetina y otros compuestos naturales es una característica destacable. Este tipo de compuestos provienen de rutas biosintéticas complejas en las plantas, y se ha propuesto esto pueda explicar su potencialidad para interactuar con múltiples macromoléculas. Asimismo es notable la semejanza en los efectos biológicos reportados para quercetina, resveratrol y otros flavonoides con características farmacofóricas similares. Los *scores* de *docking* mostraron poca variación entre los flavonoides seleccionados (figura 31), apoyando un mecanismo de acción similar para estos compuestos.

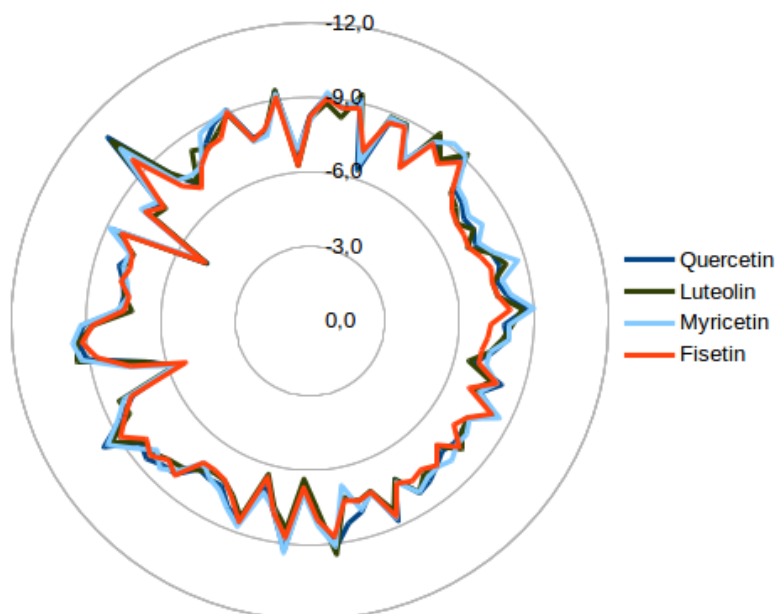


Figura 31. Comparación de los *score* de *docking* de MOE (*Molecular Operating Environment*) de los flavonoides quercetina, fisetina, miricetina, y luteolina contra las proteínas blanco identificadas en el *Protein Data Bank (PDB)*.

- **Simulaciones de dinámica molecular de complejos quercetina blanco**

El presente tamizaje virtual mostró un grupo de proteínas blanco que no han sido previamente reportadas, con un espectro de actividades en concordancia con los mecanismos de acción propuestos para quercetina. Asimismo, estos nuevos candidatos pueden ser de interés para el descubrimiento de drogas y validación de nuevos blancos. Tomando en consideración su relevancia biológica evaluamos un grupo de seleccionado de proteínas por medio de dinámica molecular. Esta es una aproximación más realista que permite evaluar el comportamiento dinámico de un

sistema y la influencia explícita de moléculas del solvente tal como ocurriría en un organismo vivo.

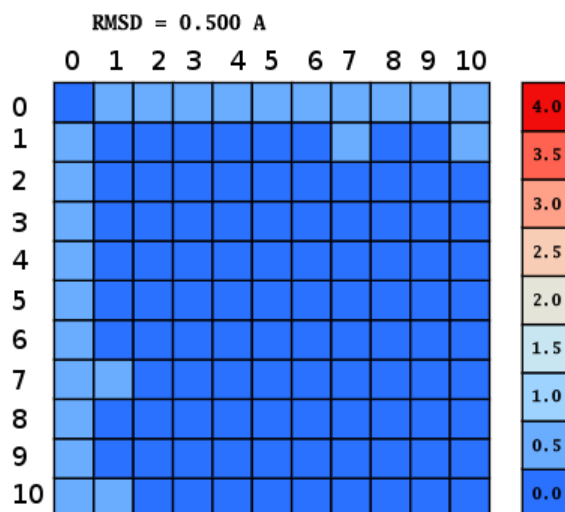


Figura 32. Ejemplo de desviación geométrica (*RMSD*, *Root Mean Square Deviation*) entre complejos de quercetina-poli-(*ADP-ribosa*) polimerasa 1 (*PARP1*) a lo largo una simulación de dinámica molecular. Para cada complejo se corrieron tres simulaciones de 10 ns analizadas en etapas de 1 ns. El último paso de cada etapa fue comparado con la estructura inicial (complejo energéticamente minimizado).

Los siguientes complejos con quercetina fueron generados: glúcogeno sintasa quinasa 3 beta (*GSK-3b*) a partir del cristal *1j1b*, la quinasa activada por mitogeno 14 (*MAPK14*) a partir del cristal *3s3i*, la *NAD(P)H*: quinona oxidoreductasa 2 (*NQO2*) a partir del cristal *1sg0* y la poli *ADP-ribosa* polimerasa 1 (*PARP*) a partir del cristal *3gfw*. Siempre que fue posible se tomó el cristal identificado en el tamizaje reverso, ya que esta se aproximaría mejor al modo de unión del flavonoide. Para la *MAPK14* un cristal diferente al obtenido en el tamizaje fue elegido debido a una bucle ausente en la estructura que no podría ser correctamente completado en *MOE*. Se corrieron tres simulaciones de 10 ns para cada sistema en una caja de aguas periódica a 300 K. Los gráficos de *Ramachandran* indicaron que los ángulos diedros phi/psi de cada proteína permanecieron bajo rangos permitidos con pocas excepciones ubicadas en la vecindad.

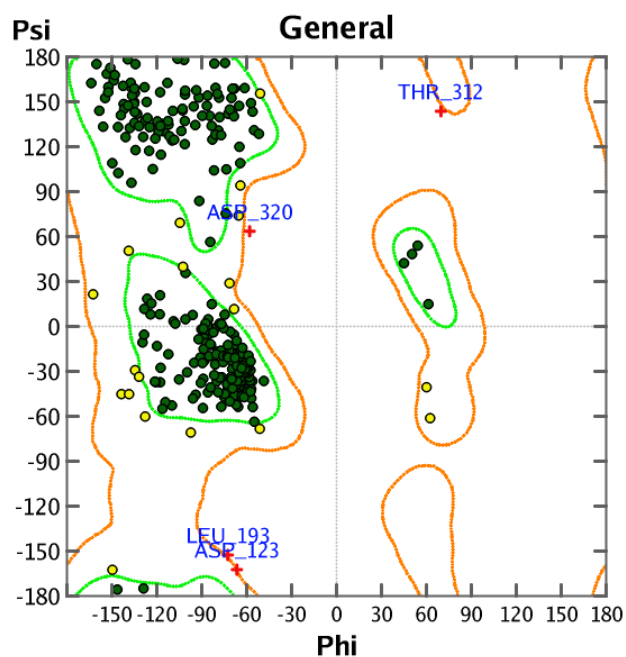


Figura 33. Ejemplo de gráfico de *Ramachandran* (psi/phi) del complejo quercetina-*poli*-(*ADP*-*ribosa*) *polimerasa* 1 (*PARP1*) luego de 10 ns de simulación de dinámica molecular. Para cada complejo se corrieron tres simulaciones de 10 ns analizadas en etapas de 1 ns. La estructura correspondiente al último paso (*frame*) de la simulación fue energéticamente minimizada (gradiente *RMS* 1 kcal, campo de fuerza *Amber12*) previo al cálculo del gráfico.

Cada sistema mostró una divergencia inicial del punto de partida de 1 Å en torno a los átomos del esqueleto de la proteína (tabla 7). Etapas comparadas dentro de la simulación poseen en general una divergencia de 0,5 Å. La energía potencial total del sistema (figura 34) y la energía potencial de interacción (figura 35) fueron medidas a lo largo de la simulación. El valor promedio en el último ns de la simulación se tomó como representativo (tabla 7).

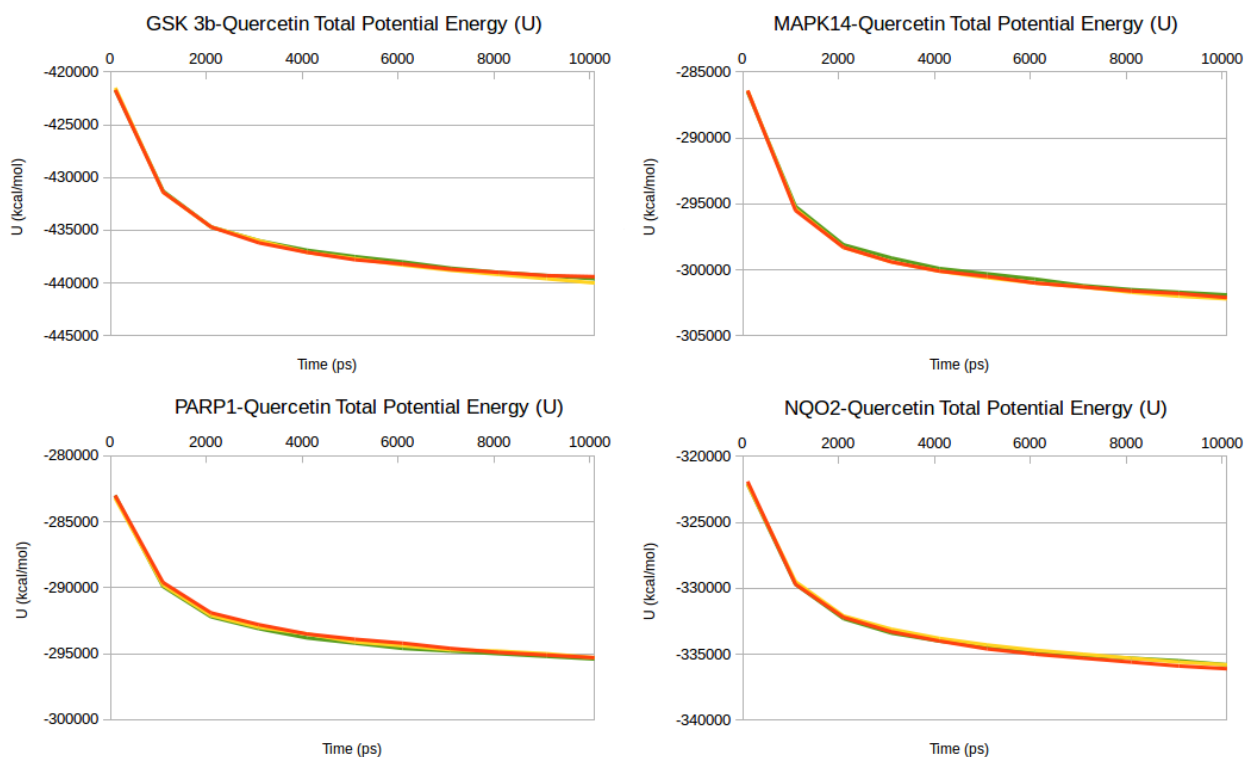


Figura 34. Energía potencial total (U, kcal) de complejos quercetina-proteína blanco. Se corrieron tres simulaciones de 10 ns para cada sistema en etapas de 1 ns con una etapa de calentamiento previa hasta 300 K. El muestreo se realizó cada 0,5 ps. Los valores graficados corresponden al promedio de cada etapa.

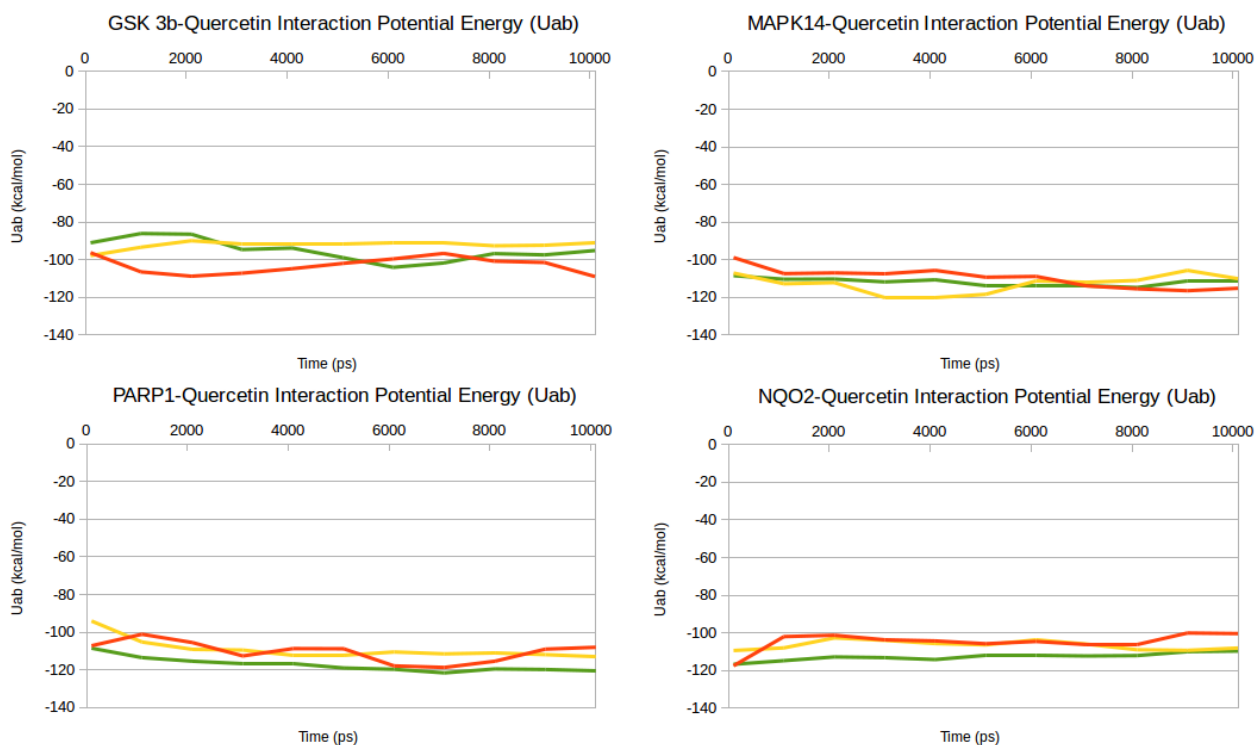


Figura 35. Energía potencial de interacción (Uab, kcal) de complejos quercetina-proteína blanco. Se corrieron tres simulaciones de 10 ns para cada sistema en etapas de 1 ns con una etapa de

calentamiento previa hasta 300 K. El muestreo se realizó cada 0,5 ps. Los valores graficados corresponden al promedio de cada etapa.

Los gráficos de energía potencial total indican un comportamiento homogéneo del sistema y la energía potencial de interacción muestran a quercetina estabilizada en el sitio de unión. Cierta divergencia se observa entre las simulaciones para un sistema dado. En este sentido se evaluaron las interacciones a lo largo de la última etapa de la simulación. Cada sistema mostró un patrón característico que se corresponde con la energía de interacción medida, el que involucra residuos con grupos donadores o aceptores de enlace de hidrógeno como glutamato, aspartato y serina. Figuras 36 a 39 muestran estos patrones y la distribución espacial de los contactos. El receptor contribuyó con la mayor parte de la energía de interacción a través de grupos aceptores y dadores de enlace de hidrógeno y en menor medida residuos aromáticos así como también hubo una participación explícita de moléculas de agua en la unión. Se destaca que las moléculas del solvente contribuyeron con alrededor de 30 % de la energía de interacción en el caso de los complejos de *PARP1* y *MAPK14*. Este porcentaje aumenta hasta alrededor de 40 % en el caso de los complejos de *GSK-3 beta* y *NQO2*. Este resultado sugiere que existe un remplazo de interacciones en favor de moléculas del solvente con distintas consecuencias. Por ejemplo un interacción con enlace de hidrógeno con el *backbone* del residuo de aspartato 700 en el sitio de unión de *GSK-3b* y una menor exposición al solvente explicarían la divergencia en la energía con otras simulaciones de la misma proteína. En este sentido las interacciones con Glu 71, Asp 168 y Ser 154 en el sitio de unión de *MAPK14* apoyan el mejor desempeño de su primera simulación mientras que en las simulaciones de *PARP1* la intervención de moléculas de agua parece favorecer la unión. Si bien la energía potencial de interacción aporta información respecto del efecto del solvente esta no es una estimación de la energía libre de interacción, no se espera entonces que el ordenamiento establecido entre las proteínas blanco sea el mismo dado por el *score* de *docking* (tabla 7). En este sentido, un método más robusto para predecir la energía libre de unión como las dinámicas moleculares dirigidas deben implementarse para mejorar el resultado.

Tabla 7. Energía potencial de interacción de una selección de complejos quercetina-proteína blanco. “Uab” y “Protein” representa la energía de interacción con el entorno incluyendo y excluyendo las moléculas del solvente respectivamente. El valor promedio de tres simulaciones (última etapa) fue comparado contra el *score* de *MOE* (*Molecular Operating Environment*) y *idTarget*. Cada complejo se anotó con la desviación geométrica observada (*RMSD*) tomada entre el último paso de la simulación y el inicio (estructura energéticamente minimizada). SD corresponde a desviación estándar.

Uniprot id	Nombre	Score de docking (kcal)		Energía potencial de interacción (Uab) (kcal) (SD)	Energía potencial de interacción (proteína) (kcal) (SD)	RMSD (Å)
		MOE	idTarget			
P16083	NAD(P)H: quinona oxidoreductasa 2	-8,76	-10,09	-106,1 (5,0)	-65,1 (3,1)	0,985
P09874	Poli [ADP-ribosa] polimerasa 1	-7,45	-9,83	-113,8 (6,3)	-82,9 (2,3)	1,021
P49841	Glucógeno sintasa quinasa 3beta	-7,27	-9,61	-98,4 (9,4)	-57,4 (8,0)	0,940

Otra característica destacable de las interacciones de quercetina es la participación que tienen los distintos grupos hidroxilo de la molécula, siendo las sustituciones en posición 3', 4', y 7 aquellas que generalmente hacen contactos con el receptor. Es interesante que estas características son compartidas por otros polifenoles como resveratrol que poseen efectos biológicos similares.

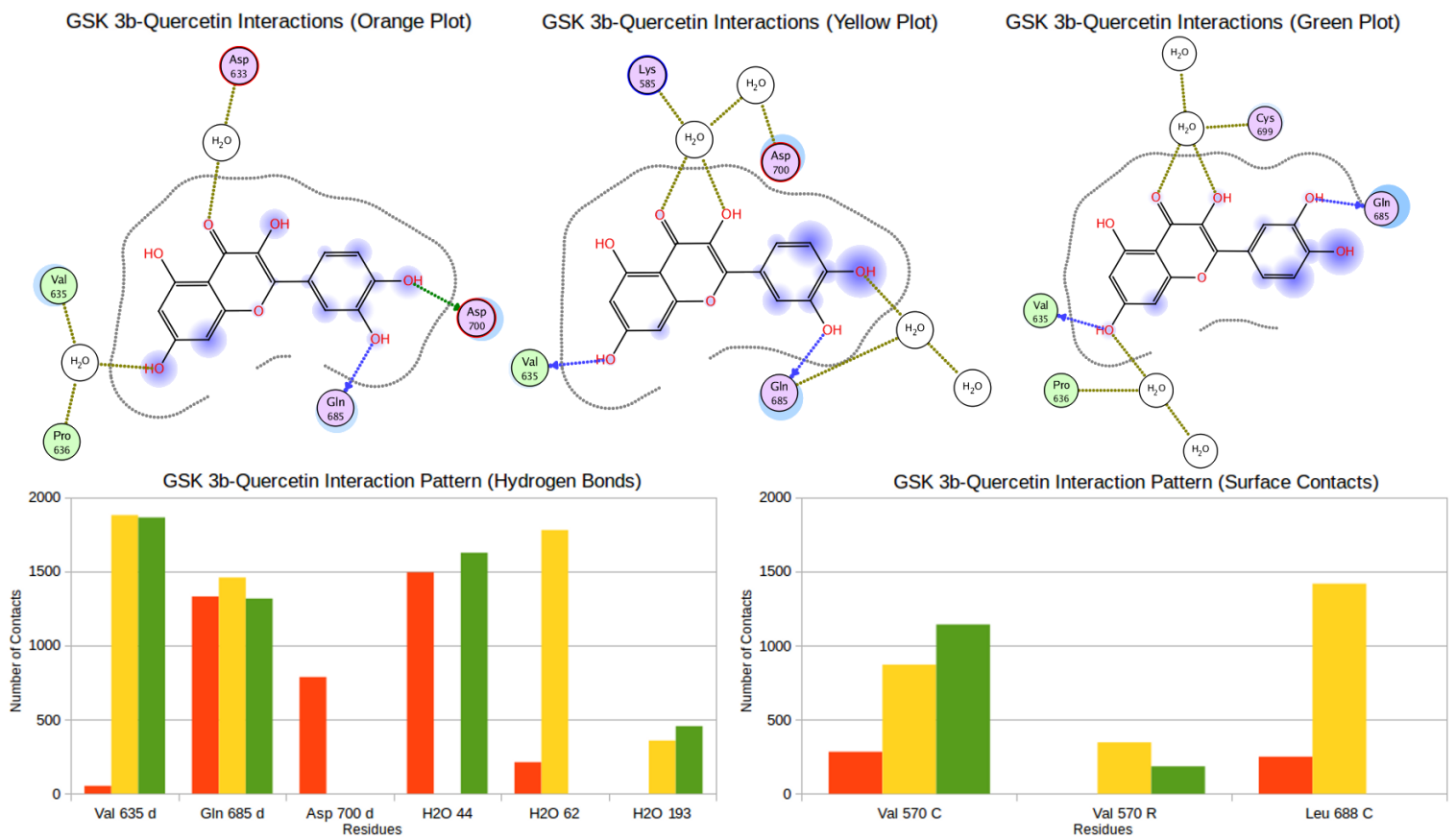


Figura 36. Patrón de interacción proteína-ligando del complejo GSK-3beta-quercetina al final de simulaciones de dinámica molecular (10 ns) junto a una representación bidimensional del sitio de unión. Círculos verdes representan zonas apolares, círculos purpura representan zonas polares, círculos rojos residuos de aminoácidos ácidos y círculos azules representan aminoácidos básicos. Contactos con la proteína son presentados por una medial luna azul alrededor de los aminoácidos. Flechas azules representan aceptores de la cadena principal de la proteína, verdes representan aceptores y dadores de las cadenas laterales. Anillos aromáticos con un “+” describen una interacción pi-cación, 2 anillos aromáticos un interacción de apilamiento. Áreas con fondo azul están expuestas al ligando. Líneas punteadas purpura representan contactos con el metal. En los gráficos las residuos de la proteína se anotaron por el número y tipo de interacción, siendo “d” dador de enlace de hidrógeno con átomos de la cadena principal, “D” donador de enlace de hidrógeno con átomos de la cadena lateral, “a” aceptor de enlaces de hidrógeno (cadena principal),

“A” aceptor de enlace de hidrógeno (cadena lateral), “C” contacto de superficie, “R” contacto aromático. Interacciones fueron separadas de acuerdo de si involucran enlaces de hidrógenos o contactos de superficie (incluyendo aromáticos). Moléculas de agua en posiciones equivalentes se cuantificaron como la misma. Interacciones presentes en menos de 300 etapas no fueron presentadas.

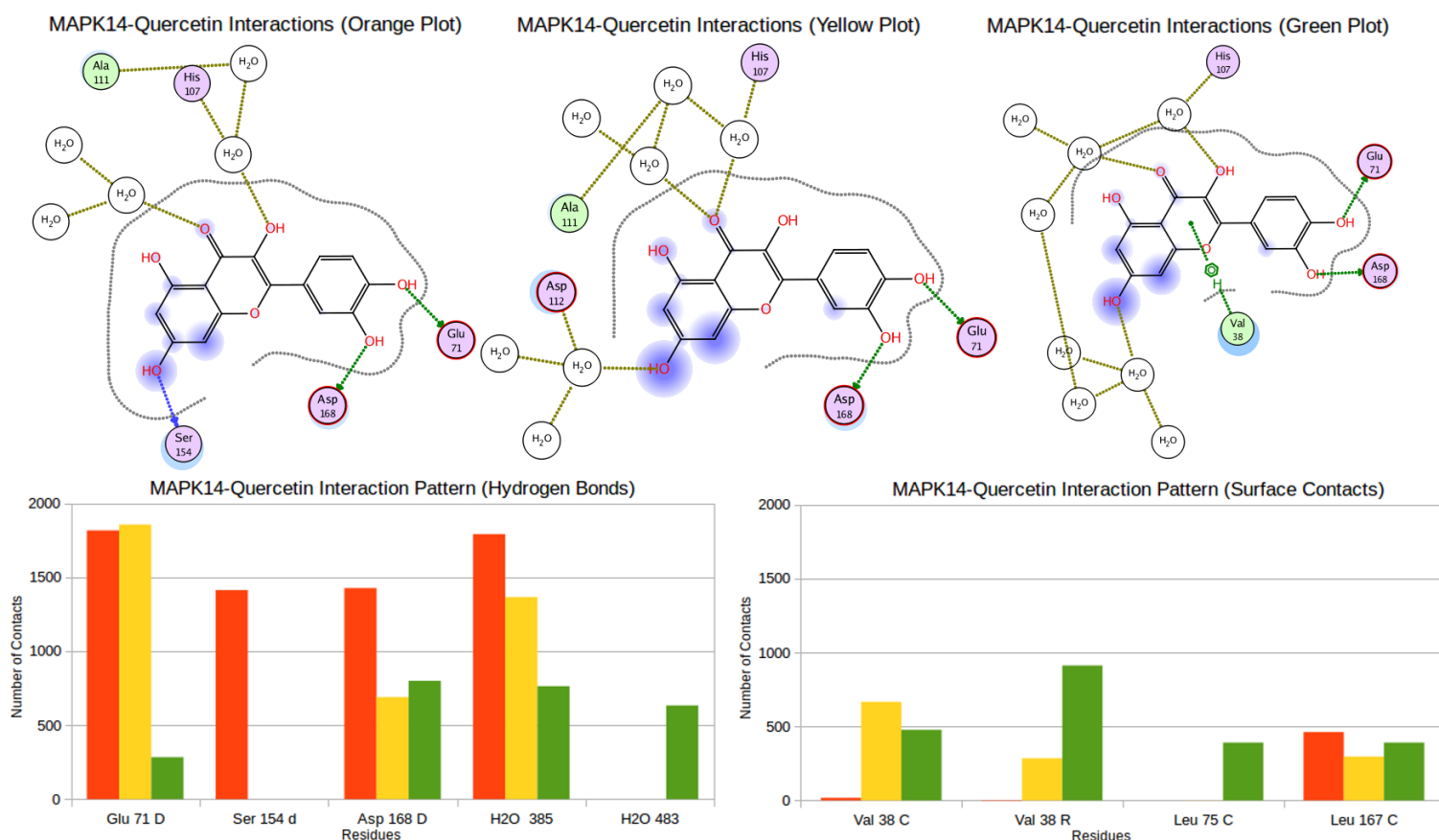


Figura 37. Patrón de interacción proteína-ligando del complejo MAPK14-quercetina al final de simulaciones de dinámica molecular (10 ns) junto a una representación bidimensional del sitio de unión. Círculos verdes representan zonas apolares, círculos purpura representan zonas polares, círculos rojos residuos de aminoácidos ácidos y círculos azules representan aminoácidos básicos. Contactos con la proteína son presentados por una medial luna azul alrededor de los aminoácidos. Flechas azules representan aceptores de la cadena principal de la proteína, verdes representan aceptores y dadores de las cadenas laterales. Anillos aromáticos con un “+” describen una interacción pi-cación, 2 anillos aromáticos un interacción de apilamiento. Áreas con fondo azul están expuestas al ligando. Líneas punteadas purpura representan contactos con el metal. En los gráficos las residuos de la proteína se anotaron por el número y tipo de interacción, siendo “d” dador de enlace de hidrógeno con átomos de la cadena principal, “D” donador de enlace de hidrógeno con átomos de la cadena lateral, “a” aceptor de enlaces de hidrógeno (cadena principal), “A” aceptor de enlace de hidrógeno (cadena lateral), “C” contacto de superficie, “R” contacto aromático. Interacciones fueron separadas de acuerdo de si involucran enlaces de hidrógenos o contactos de superficie (incluyendo aromáticos). Moléculas de agua en posiciones equivalentes se cuantificaron como la misma. Interacciones presentes en menos de 300 etapas no fueron presentadas.

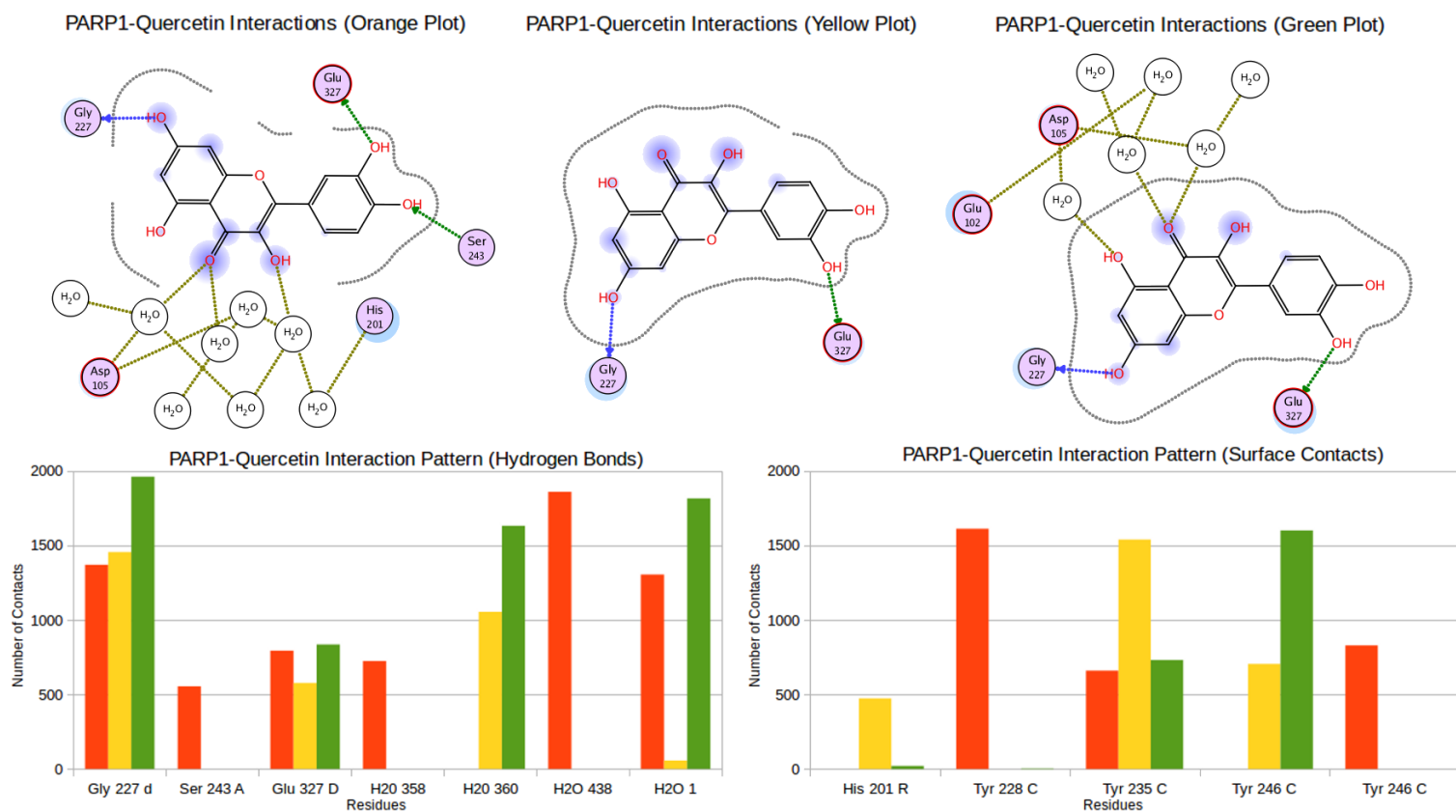


Figura 38. Patrón de interacción proteína-ligando del complejo PARP1-quercetina al final de simulaciones de dinámica molecular (10 ns) junto a una representación bidimensional del sitio de unión. Círculos verdes representan zonas apolares, círculos púrpura representan zonas polares, círculos rojos residuos de aminoácidos ácidos y círculos azules representan aminoácidos básicos. Contactos con la proteína son presentados por una medial luna azul alrededor de los aminoácidos. Flechas azules representan aceptores de la cadena principal de la proteína, verdes representan aceptores y dadores de las cadenas laterales. Anillos aromáticos con un “+” describen una interacción pi-cación, 2 anillos aromáticos un interacción de apilamiento. Áreas con fondo azul están expuestas al ligando. Líneas punteadas púrpura representan contactos con el metal. En los gráficos las residuos de la proteína se anotaron por el número y tipo de interacción, siendo “d” dador de enlace de hidrógeno con átomos de la cadena principal, “D” donador de enlace de hidrógeno con átomos de la cadena lateral, “a” aceptor de enlaces de hidrógeno (cadena principal), “A” aceptor de enlace de hidrógeno (cadena lateral), “C” contacto de superficie, “R” contacto aromático. Interacciones fueron separadas de acuerdo de si involucran enlaces de hidrógenos o contactos de superficie (incluyendo aromáticos). Moléculas de agua en posiciones equivalentes se cuantificaron como la misma. Interacciones presentes en menos de 300 etapas no fueron presentadas.

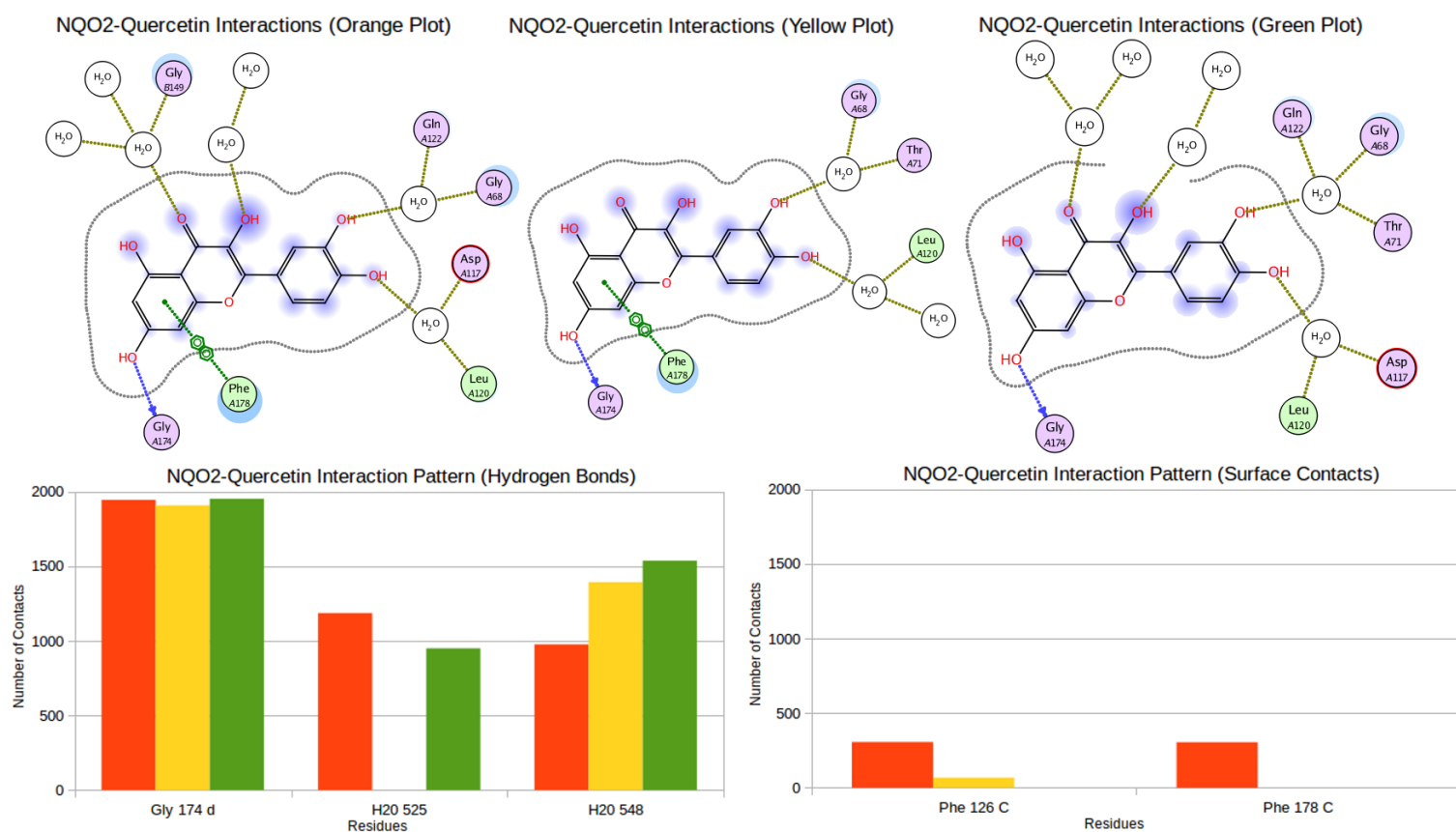


Figura 39. Patrón de interacción proteína-ligando del complejo PARP1-quercetina al final de simulaciones de dinámica molecular (10 ns) junto a una representación bidimensional del sitio de unión. Círculos verdes representan zonas apolares, círculos purpura representan zonas polares, círculos rojos residuos de aminoácidos ácidos y círculos azules representan aminoácidos básicos. Contactos con la proteína son presentados por una medial luna azul alrededor de los aminoácidos. Flechas azules representan aceptores de la cadena principal de la proteína, verdes representan aceptores y dadores de las cadenas laterales. Anillos aromáticos con un “+” describen una interacción pi-cación, 2 anillos aromáticos un interacción de apilamiento. Áreas con fondo azul están expuestas al ligando. Líneas punteadas purpura representan contactos con el metal. En los gráficos las residuos de la proteína se anotaron por el número y tipo de interacción, siendo “d” dador de enlace de hidrógeno con átomos de la cadena principal, “D” donador de enlace de hidrógeno con átomos de la cadena lateral, “a” aceptor de enlaces de hidrógeno (cadena principal), “A” aceptor de enlace de hidrógeno (cadena lateral), “C” contacto de superficie, “R” contacto aromático. Interacciones fueron separadas de acuerdo de si involucran enlaces de hidrógenos o contactos de superficie (incluyendo aromáticos). Moléculas de agua en posiciones equivalentes se cuantificaron como la misma. Interacciones presentes en menos de 300 etapas no fueron presentadas.

Consideraciones finales

En este trabajo se desarrolló un tamizaje jerárquico utilizando *SHAFTS-LIBRA* e *idTarget* para identificar los potenciales blancos de acción de quercetina. El tamizaje recuperó cada blanco de quercetina descrito en el *PDB* y extendió la lista a nuevos blancos putativos. La mayoría de los blancos predichos caen sobre blancos clínicos convencionales con efectos antitumorales y neuroprotectores, y algunos de ellos han sido establecidos también por evidencias experimentales. En adición, algunas de estas proteínas blanco están involucradas en diferentes etapas dentro de las mismas vías de señalización o en vías interconectadas, apoyando un mecanismo de acción sinérgico para quercetina.

Las simulaciones de dinámica molecular mostraron un comportamiento homogéneo de los complejos de quercetina con el flavonoide estabilizado en el sitio de unión según su energía de interacción. La principal contribución a la energía de interacción fue a través de residuos del receptor con grupos aceptores o dadores de enlace de hidrógeno y en menor medida motivos aromáticos. Cada sistema mostró un patrón de interacción característico en el que las sustituciones de quercetina en posición 3', 4' y 7 fueron las generalmente involucradas en contactos con el receptor. Adicionalmente las moléculas del solvente contribuyeron con un porcentaje importante de la energía de interacción y mediando explícitamente la unión. La exposición al solvente se corresponde con la divergencia entre complejos, resultando en interacciones más favorables para sitios menos amplios como el de *MAPK14* y *PARP1*.

Sin embargo algunas dificultades aún deben solucionarse. Aparte de las limitaciones del *score* de *docking* para el ordenamiento de proteínas, la baja representación o ausencia de ciertas proteína en el *PDB* puede ocasionar falsos negativos [247, 248]. Asimismo, para refinar estos resultados otros repositorios de interacciones como *DrugBank* pueden ser incorporados. Sobre todo, las interacciones de más complejos ligando-proteínas deben evaluarse por medio de simulaciones de dinámica molecular para asegurar la flexibilidad de las proteínas y considerar el efecto del solvente.

La traslación del conjunto de potenciales blancos a un efecto biológico dado ofrece varios desafíos. Por ejemplo se ha señalado que los metabolitos de una droga también pueden ser capaces de unirse a blancos. Al respecto existen numerosas estrategias para predecir e identificar metabolitos *in silico* e *in vitro* como *MetaSite*, o microsomas de hígado [249]. Asimismo varias proteínas en el conjunto de blancos ofrecen otros sitios de interacción además del sitio activo que pueden ser explotados, tal como el caso de las proteínas quinasa. Otro grado de libertad a ser considerado son los patrones de expresión (periodo, distribución tisular y subcelular, concentración) y así como la concentración y afinidad del ligando endógeno. Esto se

relaciona con la capacidad que tiene cada blanco de producir un efecto biológico al ser modulado por una molécula pequeña, es decir su “drogabilidad” [250].

La evidencia acumulada sobre los efectos y mecanismos de acción de quercetina establece que mientras quercetina inhibe la apoptosis, promueve mecanismos de defensa dependientes del factor *Nrf2* e inhibe la respuesta inflamatoria mediada por *NF-kappaB* como base de su efecto neuroprotector *in vitro* e *in vivo* también promueve la senescencia y apoptosis de células cancerígenas [251, 252]. Paradójicamente existe un solapamiento entre las proteínas blanco que se involucran en cada tipo de efecto. Por ejemplo la inhibición directa de *fosfoinositol 3-quinasa (PI3K)* en células cancerígenas es reconocida como un mecanismo pro-apoptótico de quercetina y se ha propuesto subyace al efecto neurotóxico de quercetina a dosis altas [82]. Por otro lado, la activación de vías de sobrevida dependientes de *PI3K* estaría involucrada en los efectos neuroprotectores de quercetina y otros flavonoides semejantes [244, 245, 253, 254]). Esta situación implica la participación de una proteína de señalización cascada arriba de la vía, tal como un receptor. En el marco de la tesis el receptor de estrógenos alfa fue evidenciado como un posible blanco [231, 255, 256]. Finalmente las diferencias entre los modelos neuronal y de cáncer; tal como dosis, duración y tipo de tratamiento (pre-tratamiento) así como el metabolismo y mecanismo de sobrevida de cada célula forman parte también de la diferencia de efectos observados.

En conclusión, reportamos una nueva aproximación que provee información valiosa para futuros ensayos *in vitro* e *in vivo* de los flavonoides así como para el desarrollo de drogas. Sobre todo, los resultados presentes amplían el entendimiento de los mecanismos de acción de quercetina y pueden generar nuevas hipótesis sobre su potencial terapéutico lo que permitiría su reposicionamiento en futuros desarrollos.

Referencias bibliográficas

1. Dajas F, Abin-Carriquiry JA, Arredondo F, et al. (2013) Neuroprotective Actions of Flavones and Flavonols : Mechanisms and Relationship to Flavonoid Structural Features. *Cent Nerv Syst Agents Med Chem* 13:1–6. doi: 10.2174/1871524911313010005
2. Bisht K, Wagner K-H, Bulmer AC (2010) Curcumin, resveratrol and flavonoids as anti-inflammatory, cyto- and DNA-protective dietary compounds. *Toxicology* 278:88–100. doi: 10.1016/j.tox.2009.11.008
3. Liu X, Jiang H, Li H (2011) SHAFTS: a hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening. *J Chem Inf Model* 51:2372–85. doi: 10.1021/ci200060s
4. Viet Hung L, Caprari S, Bizai M, et al. (2015) LIBRA: LIgand Binding site Recognition Application. *Bioinformatics*. doi: 10.1093/bioinformatics/btv489
5. Wang J, Chu P, Chen C-M, Lin J (2012) idTarget: a web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. *Nucleic Acids Res* 40:W393-9. doi: 10.1093/nar/gks496
6. Kanninen K, White AR, Koistinaho J, Malm T (2011) Targeting Glycogen Synthase Kinase-3 β for Therapeutic Benefit against Oxidative Stress in Alzheimer's Disease: Involvement of the Nrf2-ARE Pathway. *Int J Alzheimers Dis* 2011:985085. doi: 10.4061/2011/985085
7. Ba X, Garg NJ (2011) Signaling mechanism of poly(ADP-ribose) polymerase-1 (PARP-1) in inflammatory diseases. *Am J Pathol* 178:946–955. doi: 10.1016/j.ajpath.2010.12.004
8. Burr ML (1995) Explaining the French paradox. *J R Soc Health* 115:217–9.
9. Commenges D, Scotet V, Renaud S, et al. (2000) Intake of flavonoids and risk of dementia. *Eur J Epidemiol* 16:357–63.
10. Hollman PCH, Geelen A, Kromhout D (2010) Dietary Flavonol Intake May Lower Stroke Risk in Men and Women. *J Nutr* 140:600–604. doi: 10.3945/jn.109.116632
11. Kuriyama S, Shimazu T, Ohmori K, et al. (2006) Green tea consumption and mortality due to cardiovascular disease, cancer, and all causes in Japan: the

Ohsaki study. *JAMA J Am Med Assoc* 296:1255–1265. doi: 10.1001/jama.296.10.1255

12. McCullough ML, Peterson JJ, Patel R, et al. (2012) Flavonoid intake and cardiovascular disease mortality in a prospective cohort of US adults. *Am J Clin Nutr* 95:454–64. doi: 10.3945/ajcn.111.016634
13. Benavente-García O, Castillo J (2008) Update on uses and properties of citrus flavonoids: new findings in anticancer, cardiovascular, and anti-inflammatory activity. *J Agric Food Chem* 56:6185–205. doi: 10.1021/jf8006568
14. Cao G, Sofic E, Prior RL (1997) Antioxidant and prooxidant behavior of flavonoids: structure-activity relationships. *Free Radic Biol Med* 22:749–760.
15. Fuhrman B, Aviram M (2001) Flavonoids protect LDL from oxidation and attenuate atherosclerosis. *Curr Opin Lipidol* 12:41–8.
16. Kale A, Gawande S, Kotwal S (2008) Cancer phytotherapeutics: role for flavonoids at the cellular level. *Phytother Res* 22:567–77. doi: 10.1002/ptr.2283
17. Lotito SB, Frei B (2006) Consumption of flavonoid-rich foods and increased plasma antioxidant capacity in humans: cause, consequence, or epiphenomenon? *Free Radic Biol Med* 41:1727–1746. doi: 10.1016/j.freeradbiomed.2006.04.033
18. Arab L, Liebeskind DS (2010) Tea, flavonoids and stroke in man and mouse. *Arch Biochem Biophys* 501:31–6. doi: 10.1016/j.abb.2010.03.015
19. Arredondo F, Echeverry C, Abin-Carriquiry J a, et al. (2010) After cellular internalization, quercetin causes Nrf2 nuclear translocation, increases glutathione levels, and prevents neuronal death against an oxidative insult. *Free Radic Biol Med* 49:738–47. doi: 10.1016/j.freeradbiomed.2010.05.020
20. Calderone V, Chericoni S, Martinelli C, et al. (2004) Vasorelaxing effects of flavonoids: investigation on the possible involvement of potassium channels. *Naunyn Schmiedebergs Arch Pharmacol* 370:290–8. doi: 10.1007/s00210-004-0964-z
21. Cazarolli LH, Zanatta L, Alberton EH, et al. (2008) Flavonoids: cellular and molecular mechanism of action in glucose homeostasis. *Mini Rev Med Chem* 8:1032–8.

22. Cherry JJ, Rietz A, Malinkevich A, et al. (2013) Structure based identification and characterization of flavonoids that disrupt human papillomavirus-16 E6 function. *PLoS One* 8:e84506. doi: 10.1371/journal.pone.0084506
23. Davis JM, Murphy EA, McClellan JL, et al. (2008) Quercetin reduces susceptibility to influenza infection following stressful exercise. *AJP Regul Integr Comp Physiol* 295:R505–R509. doi: 10.1152/ajpregu.90319.2008
24. García-Saura MF, Galisteo M, Villar IC, et al. (2005) Effects of chronic quercetin treatment in experimental renovascular hypertension. *Mol Cell Biochem* 270:147–55.
25. Hämäläinen M, Nieminen R, Vuorela P, et al. (2007) Anti-inflammatory effects of flavonoids: genistein, kaempferol, quercetin, and daidzein inhibit STAT-1 and NF-kappaB activations, whereas flavone, isorhamnetin, naringenin, and pelargonidin inhibit only NF-kappaB activation along with their inhibitory eff. *Mediators Inflamm* 2007:45673. doi: 10.1155/2007/45673
26. Jung SH, Kang KD, Ji D, et al. (2008) The flavonoid baicalin counteracts ischemic and oxidative insults to retinal cells and lipid peroxidation to brain membranes. *Neurochem Int* 53:325–37. doi: 10.1016/j.neuint.2008.09.004
27. Maher P, Dargusch R, Ehren JL, et al. (2011) Fisetin lowers methylglyoxal dependent protein glycation and limits the complications of diabetes. *PLoS One* 6:e21226. doi: 10.1371/journal.pone.0021226
28. Reed J (2002) Cranberry flavonoids, atherosclerosis and cardiovascular health. *Crit Rev Food Sci Nutr* 42:301–16. doi: 10.1080/10408390209351919
29. Spencer JPE, Vauzour D, Rendeiro C (2009) Flavonoids and cognition: the molecular mechanisms underlying their behavioural effects. *Arch Biochem Biophys* 492:1–9. doi: 10.1016/j.abb.2009.10.003
30. Vidya Priyadarsini R, Senthil Murugan R, Maitreyi S, et al. (2010) The flavonoid quercetin induces cell cycle arrest and mitochondria-mediated apoptosis in human cervical cancer (HeLa) cells through p53 induction and NF- κ B inhibition. *Eur J Pharmacol* 649:84–91. doi: 10.1016/j.ejphar.2010.09.020
31. Dajas F (2012) Life or death: Neuroprotective and anticancer effects of quercetin. *J Ethnopharmacol* 143:383–96. doi: 10.1016/j.jep.2012.07.005

32. Echeverry C, Arredondo F, Abin-Carriquiry JA, et al. (2010) Pretreatment with natural flavones and neuronal cell survival after oxidative stress: a structure-activity relationship study. *J Agric Food Chem* 58:2111–5. doi: 10.1021/jf902951v
33. Huang W-W, Chiu Y-J, Fan M-J, et al. (2010) Kaempferol induced apoptosis via endoplasmic reticulum stress and mitochondria-dependent pathway in human osteosarcoma U-2 OS cells. *Mol Nutr Food Res* 54:1585–1595. doi: 10.1002/mnfr.201000005
34. Lee KW, Kang NJ, Heo Y-S, et al. (2008) Raf and MEK protein kinases are direct molecular targets for the chemopreventive effect of quercetin, a major flavonol in red wine. *Cancer Res* 68:946–55. doi: 10.1158/0008-5472.CAN-07-3140
35. Lee KW, Bode AM, Dong Z (2011) Molecular targets of phytochemicals for cancer prevention. *Nat Rev Cancer* 11:211–8. doi: 10.1038/nrc3017
36. Lee Y-K, Hwang J-T, Kwon DY, et al. (2010) Induction of apoptosis by quercetin is mediated through AMPK α 1/ASK1/p38 pathway. *Cancer Lett* 292:228–36. doi: 10.1016/j.canlet.2009.12.005
37. Mandel S, Weinreb O, Amit T, Youdim MBH (2004) Cell signaling pathways in the neuroprotective actions of the green tea polyphenol (-)-epigallocatechin-3-gallate: implications for neurodegenerative diseases. *J Neurochem* 88:1555–1569.
38. Pandey AK, Hazari PP, Patnaik R, Mishra AK (2011) The role of ASIC1a in neuroprotection elicited by quercetin in focal cerebral ischemia. *Brain Res* 1383:289–99. doi: 10.1016/j.brainres.2011.01.085
39. Spencer JPE, Rice-Evans C, Williams RJ (2003) Modulation of pro-survival Akt/protein kinase B and ERK1/2 signaling cascades by quercetin and its in vivo metabolites underlie their action on neuronal viability. *J Biol Chem* 278:34783–93. doi: 10.1074/jbc.M305063200
40. Williams RJ, Spencer JPE, Rice-Evans C (2004) Flavonoids: antioxidants or signalling molecules? *Free Radic Biol Med* 36:838–49. doi: 10.1016/j.freeradbiomed.2004.01.001
41. Hart CP (2005) Finding the target after screening the phenotype. *Drug Discov Today* 10:513–9. doi: 10.1016/S1359-6446(05)03415-X

42. Chen YZ, Zhi DG (2001) Ligand-protein inverse docking and its potential use in the computer search of protein targets of a small molecule. *Proteins* 43:217–26.
43. Kinnings SL, Jackson RM (2011) ReverseScreen3D: a structure-based ligand matching method to identify protein targets. *J Chem Inf Model* 51:624–34. doi: 10.1021/ci1003174
44. Li H, Gao Z, Kang L, et al. (2006) TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res* 34:W219-24. doi: 10.1093/nar/gkl114
45. Sridhar A, Saremy S, Bhattacharjee B (2014) Elucidation of molecular targets of bioactive principles of black cumin relevant to its anti- tumour functionality - An Insilico target fishing approach. 10:
46. Dakshanamurthy S, Issa NT, Assefnia S, et al. (2012) Predicting new indications for approved drugs using a proteochemometric method. *J Med Chem* 55:6832–48. doi: 10.1021/jm300576q
47. Kumar SP, Parmar VR, Jasrai YT, Pandya HA (2015) Prediction of protein targets of kinetin using in silico and in vitro methods: a case study on spinach seed germination mechanism. *J Chem Biol* 8:95–105. doi: 10.1007/s12154-015-0135-3
48. Su J, Chang C, Xiang Q, et al. (2014) Xyloketal B, a marine compound, acts on a network of molecular proteins and regulates the activity and expression of rat cytochrome P450 3a: a bioinformatic and animal study. *Drug Des Devel Ther* 8:2555–602. doi: 10.2147/DDDT.S73476
49. Wu X, Chen X, Dan J, et al. (2016) Characterization of anti-leukemia components from *Indigo naturalis* using comprehensive two-dimensional K562/cell membrane chromatography and in silico target identification. *Sci Rep* 6:25491. doi: 10.1038/srep25491
50. Cereto-Massagué A, Ojeda MJ, Valls C, et al. (2015) Tools for in silico target fishing. *Methods* 71:98–103. doi: 10.1016/j.ymeth.2014.09.006
51. Formica J V, Regelson W (1995) Review of the biology of Quercetin and related bioflavonoids. *Food Chem Toxicol an Int J Publ Br Ind Biol Res Assoc* 33:1061–1080.

52. Kühnau J (1976) The flavonoids. A class of semi-essential food components: their role in human nutrition. *World Rev Nutr Diet* 24:117–91.
53. Pollastri S, Tattini M (2011) Flavonols: old compounds for old roles. *Ann Bot* 108:1225–33. doi: 10.1093/aob/mcr234
54. Middleton E, Kandaswami C, Theoharides TC (2000) The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease, and cancer. *Pharmacol Rev* 52:673–751.
55. Niedzwiecki A, Roomi M, Kalinovsky T, Rath M (2016) Anticancer Efficacy of Polyphenols and Their Combinations. *Nutrients* 8:552. doi: 10.3390/nu8090552
56. Hertog MG, Feskens EJ, Hollman PC, et al. (1993) Dietary antioxidant flavonoids and risk of coronary heart disease: the Zutphen Elderly Study. *Lancet* 342:1007–11.
57. de Vries JH, Janssen PL, Hollman PC, et al. (1997) Consumption of quercetin and kaempferol in free-living subjects eating a variety of diets. *Cancer Lett* 114:141–4.
58. Hertog MG, Kromhout D, Aravanis C, et al. (1995) Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study. *Arch Intern Med* 155:381–6.
59. Rimm EB, Katan MB, Ascherio A, et al. (1996) Relation between intake of flavonoids and risk for coronary heart disease in male health professionals. *Ann Intern Med* 125:384–389.
60. Knekt P, Järvinen R, Seppänen R, et al. (1997) Dietary flavonoids and the risk of lung cancer and other malignant neoplasms. *Am J Epidemiol* 146:223–30.
61. Kimira M, Arai Y, Shimoi K, Watanabe S (1998) Japanese intake of flavonoids and isoflavonoids from foods. *J Epidemiol* 8:168–75.
62. Johannot L, Somerset SM (2006) Age-related variations in flavonoid intake and sources in the Australian population. *Public Health Nutr* 9:1045–54.
63. Bischoff SC (2008) Quercetin: potentials in the prevention and therapy of disease. *Curr Opin Clin Nutr Metab Care* 11:733–40. doi: 10.1097/MCO.0b013e32831394b8
64. Chen C, Zhou J, Ji C (2010) Quercetin: a potential drug to reverse multidrug resistance. *Life Sci* 87:333–338. doi: 10.1016/j.lfs.2010.07.004

65. Kressler J, Millard-Stafford M, Warren GL (2011) Quercetin and endurance exercise capacity: a systematic review and meta-analysis. *Med Sci Sports Exerc* 43:2396–404. doi: 10.1249/MSS.0b013e31822495a7
66. Mendoza EE, Burd R (2011) Quercetin as a systemic chemopreventative agent: structural and functional mechanisms. *Mini Rev. Med. Chem.* 11:
67. Gilgun-Sherki Y, Melamed E, Offen D (2001) Oxidative stress induced-neurodegenerative diseases: the need for antioxidants that penetrate the blood brain barrier. *Neuropharmacology* 40:959–75.
68. Langley B, Ratan RR (2004) Oxidative stress-induced death in the nervous system: cell cycle dependent or independent? *J Neurosci Res* 77:621–9. doi: 10.1002/jnr.20210
69. Simonian NA, Coyle JT (1996) Oxidative stress in neurodegenerative diseases. *Annu Rev Pharmacol Toxicol* 36:83–106. doi: 10.1146/annurev.pa.36.040196.000503
70. Olesen J, Baker MG, Freund T, et al. (2006) Consensus document on European brain research. *J Neurol Neurosurg Psychiatry* 77 Suppl 1:i1-49.
71. Dajas F, Abin-Carriquiry JA, Arredondo F, et al. (2015) Quercetin in brain diseases: Potential and limits. *Neurochem Int* 89:140–8. doi: 10.1016/j.neuint.2015.07.002
72. Kelsey NA, Wilkins HM, Linseman DA (2010) Nutraceutical antioxidants as novel neuroprotective agents. *Molecules* 15:7792–814. doi: 10.3390/molecules15117792
73. Dajas F, Rivera-Megret F, Blasina F, et al. (2003) Neuroprotection by flavonoids. *Braz J Med Biol Res* 36:1613–20.
74. Jakubowicz-Gil J, Rzeski W, Zdzisinska B, et al. (2008) Cell death and neuronal arborization upon quercetin treatment in rat neurons. *Acta Neurobiol Exp (Wars)* 68:139–46.
75. Silva B, Oliveira PJ, Dias A, Malva JO Quercetin, kaempferol and biapigenin from *Hypericum perforatum* are neuroprotective against excitotoxic insults. *Neurotox Res* 13:265–79.

76. Ansari MA, Abdul HM, Joshi G, et al. (2009) Protective effect of quercetin in primary neurons against Abeta(1-42): relevance to Alzheimer's disease. *J Nutr Biochem* 20:269–75. doi: 10.1016/j.jnutbio.2008.03.002
77. Ossola B, Kääriäinen TM, Männistö PT (2009) The multiple faces of quercetin in neuroprotection. *Expert Opin Drug Saf* 8:397–409. doi: 10.1517/14740330903026944
78. Magalingam KB, Radhakrishnan AK, Haleagrahara N (2015) Protective Mechanisms of Flavonoids in Parkinson's Disease. *Oxid Med Cell Longev* 2015:314560. doi: 10.1155/2015/314560
79. Williams RJ, Spencer JPE (2012) Flavonoids, cognition, and dementia: actions, mechanisms, and potential therapeutic utility for Alzheimer disease. *Free Radic Biol Med* 52:35–45. doi: 10.1016/j.freeradbiomed.2011.09.010
80. Rivera F, Costa G, Abin A, et al. (2008) Reduction of ischemic brain damage and increase of glutathione by a liposomal preparation of quercetin in permanent focal ischemia in rats. *Neurotox Res* 13:105–14.
81. Rice-Evans CA, Miller NJ, Paganga G (1996) Structure-antioxidant activity relationships of flavonoids and phenolic acids. *Free Radic Biol Med* 20:933–956.
82. Spencer JPE (2007) The interactions of flavonoids within neuronal signalling pathways. *Genes Nutr* 2:257–273. doi: 10.1007/s12263-007-0056-z
83. Halliwell B GJ (1998) *Free radicals in biology and medicine*, Third. Oxford University Press, Oxford
84. Bors W, Heller W, Michel C, Saran M (1990) Flavonoids as antioxidants: determination of radical-scavenging efficiencies. *Methods Enzymol* 186:343–355.
85. Silva MM, Santos MR, Caroço G, et al. (2002) Structure-antioxidant activity relationships of flavonoids: a re-examination. *Free Radic Res* 36:1219–27.
86. Formica J V, Regelson W (1995) Review of the biology of Quercetin and related bioflavonoids. *Food Chem Toxicol* 33:1061–80.
87. Forman HJ, Davies KJA, Ursini F (2014) How do nutritional antioxidants really work: nucleophilic tone and para-hormesis versus free radical scavenging in

vivo. *Free Radic Biol Med* 66:24–35. doi:
10.1016/j.freeradbiomed.2013.05.045

88. Metodiewa D, Jaiswal AK, Cenas N, et al. (1999) Quercetin may act as a cytotoxic prooxidant after its metabolic activation to semiquinone and quinoidal product. *Free Radic Biol Med* 26:107–116.
89. Boots AW, Kubben N, Haenen GRMM, Bast A (2003) Oxidized quercetin reacts with thiols rather than with ascorbate: implication for quercetin supplementation. *Biochem Biophys Res Commun* 308:560–5.
90. Laughton MJ, Evans PJ, Moroney MA, et al. (1991) Inhibition of mammalian 5-lipoxygenase and cyclo-oxygenase by flavonoids and phenolic dietary additives. Relationship to antioxidant activity and to iron ion-reducing ability. *Biochem Pharmacol* 42:1673–1681.
91. Sanz MJ, Ferrandiz ML, Cejudo M, et al. (1994) Influence of a series of natural flavonoids on free radical generating systems and oxidative stress. *Xenobiotica* 24:689–99. doi: 10.3109/00498259409043270
92. Hoult JR, Moroney MA, Payá M (1994) Actions of flavonoids and coumarins on lipoxygenase and cyclooxygenase. *Methods Enzymol* 234:443–54.
93. Chang WS, Lee YJ, Lu FJ, Chiang HC (1993) Inhibitory effects of flavonoids on xanthine oxidase. *Anticancer Res* 13:2165–2170.
94. Nijveldt RJ, van Nood E, van Hoorn DE, et al. (2001) Flavonoids: a review of probable mechanisms of action and potential applications. *Am J Clin Nutr* 74:418–425.
95. Paulino M, Alvareda E, Iribarne F, et al. (2016) Toward the understanding of the molecular basis for the inhibition of COX-1 and COX-2 by phenolic compounds present in Uruguayan propolis and grape pomace. *J Biomol Struct Dyn* 1–15. doi: 10.1080/07391102.2015.1124808
96. Uruno A, Motohashi H (2011) The Keap1-Nrf2 system as an in vivo sensor for electrophiles. *Nitric oxide Biol Chem / Off J Nitric Oxide Soc* 25:153–160. doi: 10.1016/j.niox.2011.02.007
97. Wakabayashi N, Slocum SL, Skoko JJ, et al. (2010) When NRF2 talks, who's listening? *Antioxid Redox Signal* 13:1649–1663. doi: 10.1089/ars.2010.3216

98. Sussan TE, Jun J, Thimmulappa R, et al. (2008) Disruption of Nrf2, a key inducer of antioxidant defenses, attenuates ApoE-mediated atherosclerosis in mice. *PLoS One* 3:e3791. doi: 10.1371/journal.pone.0003791
99. Calkins MJ, Johnson DA, Townsend JA, et al. (2009) The Nrf2/ARE pathway as a potential therapeutic target in neurodegenerative disease. *Antioxid Redox Signal* 11:497–508.
100. Copple IM (2012) The Keap1–Nrf2 Cell Defense Pathway – A Promising Therapeutic Target? *Curr Concepts Drug Metab Toxicol* Volume 63:43–79.
101. Kwak M, Kensler TW (2010) Targeting NRF2 signaling for cancer chemoprevention. *Toxicol Appl Pharmacol* 244:66–76. doi: 10.1016/j.taap.2009.08.028
102. Ramsey CP, Glass CA, Montgomery MB, et al. (2007) Expression of Nrf2 in neurodegenerative diseases. *J Neuropathol Exp Neurol* 66:75–85.
103. Kanninen K, Malm TM, Jyrkkänen H-K, et al. (2008) Nuclear factor erythroid 2-related factor 2 protects against beta amyloid. *Mol Cell Neurosci* 39:302–313.
104. Vargas MR, Johnson DA, Sirkis DW, et al. (2008) Nrf2 activation in astrocytes protects against neurodegeneration in mouse models of familial amyotrophic lateral sclerosis. *J Neurosci Off J Soc Neurosci* 28:13574–13581.
105. Yang L, Calingasan NY, Thomas B, et al. (2009) Neuroprotective effects of the triterpenoid, CDDO methyl amide, a potent inducer of Nrf2-mediated transcription. *PLoS One* 4:e5757–e5757.
106. Dang J, Brandenburg L-O, Rosen C, et al. (2012) Nrf2 expression by neurons, astroglia, and microglia in the cerebral cortical penumbra of ischemic rats. *J Mol Neurosci* MN 46:578–584. doi: 10.1007/s12031-011-9645-9
107. González R, Ballester I, López-Posadas R, et al. (2011) Effects of flavonoids and other polyphenols on inflammation. *Crit Rev Food Sci Nutr* 51:331–62. doi: 10.1080/10408390903584094
108. Pietta PG (2000) Flavonoids as antioxidants. *J Nat Prod* 63:1035–1042.
109. Kopitar-Jerala N (2015) Innate Immune Response in Brain, NF-Kappa B Signaling and Cystatins. *Front Mol Neurosci*. doi: 10.3389/fnmol.2015.00073
110. Wu D, Wu P, Zhao L, et al. (2015) NF-κB Expression and Outcomes in Solid Tumors. *Medicine (Baltimore)* 94:e1687. doi: 10.1097/MD.0000000000001687

111. Grilli M, Memo M (1999) Possible role of NF-kappaB and p53 in the glutamate-induced pro-apoptotic neuronal pathway. *Cell Death Differ* 6:22–7. doi: 10.1038/sj.cdd.4400463
112. Małek R, Borowicz KK, Jargiełło M, Czuczwar SJ Role of nuclear factor kappaB in the central nervous system. *Pharmacol Rep* 59:25–33.
113. Pizzi M, Sarnico I, Lanzillotta A, et al. (2009) Post-ischemic brain damage: NF-kappaB dimer heterogeneity as a molecular determinant of neuron vulnerability. *FEBS J* 276:27–35. doi: 10.1111/j.1742-4658.2008.06767.x
114. Zhang H, Sun S-C (2015) NF-κB in inflammation and renal diseases. *Cell Biosci* 5:63. doi: 10.1186/s13578-015-0056-4
115. Park M, Hong J (2016) Roles of NF-κB in Cancer and Inflammatory Diseases and Their Therapeutic Approaches. *Cells* 5:15. doi: 10.3390/cells5020015
116. Shin SY, Woo Y, Hyun J, et al. (2011) Relationship between the structures of flavonoids and their NF-κB-dependent transcriptional activities. *Bioorg Med Chem Lett* 21:6036–41. doi: 10.1016/j.bmcl.2011.08.077
117. Kim E-K, Kwon K-B, Song M-Y, et al. (2007) Flavonoids Protect Against Cytokine-Induced Pancreatic β-Cell Damage Through Suppression of Nuclear Factor κB Activation. *Pancreas* 35:e1–e9. doi: 10.1097/mpa.0b013e31811ed0d2
118. Ruiz PA, Braune A, Hölzlwimmer G, et al. (2007) Quercetin inhibits TNF-induced NF-kappaB transcription factor recruitment to proinflammatory gene promoters in murine intestinal epithelial cells. *J Nutr* 137:1208–15.
119. Boots AW, Haenen GRMM, Bast A (2008) Health effects of quercetin: from antioxidant to nutraceutical. *Eur J Pharmacol* 585:325–37. doi: 10.1016/j.ejphar.2008.03.008
120. Kao T-K, Ou Y-C, Raung S-L, et al. (2010) Inhibition of nitric oxide production by quercetin in endotoxin/cytokine-stimulated microglia. *Life Sci* 86:315–21. doi: 10.1016/j.lfs.2009.12.014
121. Zhang ZJ, Cheang LCV, Wang MW, Lee SM-Y (2011) Quercetin exerts a neuroprotective effect through inhibition of the iNOS/NO system and pro-inflammation gene expression in PC12 cells and in zebrafish. *Int J Mol Med* 27:195–203. doi: 10.3892/ijmm.2010.571

122. Kinaci MK, Erkasap N, Kucuk A, et al. (2012) Effects of quercetin on apoptosis , NF- κ B and NOS gene expression in renal ischemia / reperfusion injury. 249–254. doi: 10.3892/etm.2011.382
123. Zheng LT, Ock J, Kwon B-M, Suk K (2008) Suppressive effects of flavonoid fisetin on lipopolysaccharide-induced microglial activation and neurotoxicity. *Int Immunopharmacol* 8:484–94. doi: 10.1016/j.intimp.2007.12.012
124. Zhang F, Liu J, Shi J-S (2010) Anti-inflammatory activities of resveratrol in the brain: role of resveratrol in microglial activation. *Eur J Pharmacol* 636:1–7. doi: 10.1016/j.ejphar.2010.03.043
125. Elmore S (2007) Apoptosis: a review of programmed cell death. *Toxicol Pathol* 35:495–516. doi: 10.1080/01926230701320337
126. Suematsu N, Hosoda M, Fujimori K (2011) Protective effects of quercetin against hydrogen peroxide-induced apoptosis in human neuronal SH-SY5Y cells. *Neurosci Lett* 504:223–7. doi: 10.1016/j.neulet.2011.09.028
127. STEINMETZ KA, POTTER JD (1996) Vegetables, Fruit, and Cancer Prevention. *J Am Diet Assoc* 96:1027–1039. doi: 10.1016/S0002-8223(96)00273-8
128. Steinmetz KA, Potter JD (1996) Vegetables, fruit, and cancer prevention: a review. *J Am Diet Assoc* 96:1027–39. doi: 10.1016/S0002-8223(96)00273-8
129. Zhang X-A, Zhang S, Yin Q, Zhang J (2015) Quercetin induces human colon cancer cells apoptosis by inhibiting the nuclear factor-kappa B Pathway. *Pharmacogn Mag* 11:404. doi: 10.4103/0973-1296.153096
130. Wang G, Zhang J, Liu L, et al. (2012) Quercetin potentiates doxorubicin mediated antitumor effects against liver cancer through p53/Bcl-xl. *PLoS One* 7:e51764. doi: 10.1371/journal.pone.0051764
131. Srivastava S, Somasagara RR, Hegde M, et al. (2016) Quercetin, a Natural Flavonoid Interacts with DNA, Arrests Cell Cycle and Causes Tumor Regression by Activating Mitochondrial Pathway of Apoptosis. *Sci Rep* 6:24049. doi: 10.1038/srep24049
132. Richter M, Ebermann R, Marian B (1999) Quercetin-induced apoptosis in colorectal tumor cells: possible role of EGF receptor signaling. *Nutr Cancer* 34:88–99. doi: 10.1207/S15327914NC340113

133. Lyne PD (2002) Structure-based virtual screening: an overview. *Drug Discov Today* 7:1047–55.
134. Stahura FL, Bajorath J (2004) Virtual screening methods that complement HTS. *Comb Chem High Throughput Screen* 7:259–69.
135. Kodadek T (2011) The rise, fall and reinvention of combinatorial chemistry. *Chem Commun* 47:9757. doi: 10.1039/c1cc12102b
136. Oprea TI, Matter H (2004) Integrating virtual screening in lead discovery. *Curr Opin Chem Biol* 8:349–58. doi: 10.1016/j.cbpa.2004.06.008
137. Westermaier Y, Barril X, Scapozza L (2015) Virtual screening: an in silico tool for interlacing the chemical universe with the proteome. *Methods* 71:44–57. doi: 10.1016/j.ymeth.2014.08.001
138. Geppert H, Vogt M, Bajorath J (2010) Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation. *J Chem Inf Model* 50:205–16. doi: 10.1021/ci900419k
139. Sliwoski G, Kothiwale S, Meiler J, Lowe EW (2014) Computational methods in drug discovery. *Pharmacol Rev* 66:334–95. doi: 10.1124/pr.112.007336
140. Ekins S, Mestres J, Testa B (2007) In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br J Pharmacol* 152:9–20. doi: 10.1038/sj.bjp.0707305
141. Klabunde T (2007) Chemogenomic approaches to drug discovery: similar receptors bind similar ligands. *Br J Pharmacol* 152:5–7. doi: 10.1038/sj.bjp.0707308
142. Kumar A, Zhang KYJ (2015) Hierarchical virtual screening approaches in small molecule drug discovery. *Methods* 71:26–37. doi: 10.1016/j.ymeth.2014.07.007
143. Berman HM (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242. doi: 10.1093/nar/28.1.235
144. Feng Z, Chen L, Maddula H, et al. (2004) Ligand Depot: a data warehouse for ligands bound to macromolecules. *Bioinformatics* 20:2153–5. doi: 10.1093/bioinformatics/bth214

145. Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada H 2R7 (2011) Molecular Operating Environment (MOE), 2011.10.
146. Oprea TI, Tropsha A, Faulon J-L, Rintoul MD (2007) Systems chemical biology. *Nat Chem Biol* 3:447–50. doi: 10.1038/nchembio0807-447
147. Martin YC, Kofron JL, Traphagen LM (2002) Do structurally similar molecules have similar biological activity? *J Med Chem* 45:4350–4358.
148. Nidhi, Glick M, Davies JW, Jenkins JL (2006) Prediction of biological targets for compounds using multiple-category Bayesian models trained on chemogenomics databases. *J Chem Inf Model* 46:1124–1133. doi: 10.1021/ci060003g
149. Wang L, Ma C, Wipf P, et al. (2013) TargetHunter: an in silico target identification tool for predicting therapeutic potential of small organic molecules based on chemogenomic database. *AAPS J* 15:395–406. doi: 10.1208/s12248-012-9449-z
150. Liu X, Vogt I, Haque T, Campillos M (2013) HitPick: a web server for hit identification and target prediction of chemical screenings. *Bioinformatics* 29:1910–2. doi: 10.1093/bioinformatics/btt303
151. Xu H, Tao Y, Lu P, et al. (2013) A computational drug-target network for yuanhu zhitong prescription. *Evid Based Complement Alternat Med* 2013:658531. doi: 10.1155/2013/658531
152. Chen H, Yao K, Nadas J, et al. (2012) Prediction of molecular targets of cancer preventing flavonoid compounds using computational methods. *PLoS One* 7:e38261. doi: 10.1371/journal.pone.0038261
153. Wang Z, Liang L, Yin Z, Lin J (2016) Improving chemical similarity ensemble approach in target prediction. *J Cheminform* 8:20. doi: 10.1186/s13321-016-0130-x
154. Lu W, Liu X, Cao X, et al. (2011) SHAFTS: a hybrid approach for 3D molecular similarity calculation. 2. Prospective case study in the discovery of diverse p90 ribosomal S6 protein kinase 2 inhibitors to suppress cell migration. *J Med Chem* 54:3564–74. doi: 10.1021/jm200139j

155. Bai F, Liu H, Tong L, et al. (2012) Discovery of novel selective inhibitors for EGFR-T790M/L858R. *Bioorg Med Chem Lett* 22:1365–70. doi: 10.1016/j.bmcl.2011.12.067
156. Kong X, Qin J, Li Z, et al. (2012) Development of a novel class of B-Raf(V600E)-selective inhibitors through virtual screening and hierarchical hit optimization. *Org Biomol Chem* 10:7402–17. doi: 10.1039/c2ob26081f
157. Zheng R, Chen T, Lu T (2011) A comparative reverse docking strategy to identify potential antineoplastic targets of tea functional components and binding mode. *Int J Mol Sci* 12:5200–12. doi: 10.3390/ijms12085200
158. Liu X, Bai F, Ouyang S, et al. (2009) Cyndi: a multi-objective evolution algorithm based method for bioactive molecular conformational generation. *BMC Bioinformatics* 10:101. doi: 10.1186/1471-2105-10-101
159. Xie L, Li J, Xie L, Bourne PE (2009) Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput Biol* 5:e1000387. doi: 10.1371/journal.pcbi.1000387
160. Xie L, Bourne PE (2007) A robust and efficient algorithm for the shape description of protein structures and its application in predicting ligand binding sites. *BMC Bioinformatics* 8 Suppl 4:S9. doi: 10.1186/1471-2105-8-S4-S9
161. Xie L, Bourne PE (2008) Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. *Proc Natl Acad Sci U S A* 105:5441–5446. doi: 10.1073/pnas.0704422105
162. Liu X, Ouyang S, Yu B, et al. (2010) PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. *Nucleic Acids Res* 38:W609-14. doi: 10.1093/nar/gkq300
163. Wolber G, Langer T (2005) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 45:160–169. doi: 10.1021/ci049885e
164. Furnham N, Holliday GL, de Beer TAP, et al. (2014) The Catalytic Site Atlas 2.0: cataloging catalytic sites and residues identified in enzymes. *Nucleic Acids Res* 42:D485-9. doi: 10.1093/nar/gkt1243

165. Yan X, Kang H, Feng J, et al. (2016) Identification of Toxic Pyrrolizidine Alkaloids and Their Common Hepatotoxicity Mechanism. *Int J Mol Sci* 17:318. doi: 10.3390/ijms17030318
166. Lauro G, Romano A, Riccio R, Bifulco G (2011) Inverse virtual screening of antitumor targets: pilot study on a small database of natural bioactive compounds. *J Nat Prod* 74:1401–7. doi: 10.1021/np100935s
167. Yang L, Wang K, Chen J, et al. (2011) Exploring off-targets and off-systems for adverse drug reactions via chemical-protein interactome--clozapine-induced agranulocytosis as a case study. *PLoS Comput Biol* 7:e1002016. doi: 10.1371/journal.pcbi.1002016
168. Li YY, An J, Jones SJM (2011) A computational approach to finding novel targets for existing drugs. *PLoS Comput Biol* 7:e1002139. doi: 10.1371/journal.pcbi.1002139
169. Fan S, Geng Q, Pan Z, et al. (2012) Clarifying off-target effects for torcetrapib using network pharmacology and reverse docking approach. *BMC Syst Biol* 6:152. doi: 10.1186/1752-0509-6-152
170. Lee M, Kim D (2012) Large-scale reverse docking profiles and their applications. *BMC Bioinformatics* 13 Suppl 1:S6. doi: 10.1186/1471-2105-13-S17-S6
171. Hsin K-Y, Ghosh S, Kitano H (2013) Combining machine learning systems and multiple docking simulation packages to improve docking prediction reliability for network pharmacology. *PLoS One* 8:e83922. doi: 10.1371/journal.pone.0083922
172. Olivero-verbel J, Cabarcas-Montalvo M, Ortega-Zúñiga C (2010) Theoretical targets for TCDD: a bioinformatics approach. *Chemosphere* 80:1160–6. doi: 10.1016/j.chemosphere.2010.06.020
173. Ke S, Barata T, Solmajer T, et al. (2012) Target fishing and docking studies of the novel derivatives of aryl-aminopyridines with potential anticancer activity. *Bioorg Med Chem* 20:5220–8. doi: 10.1016/j.bmc.2012.06.051
174. Nikolić AR, Petri ET, Klisurić OR, et al. (2015) Synthesis and anticancer cell potential of steroidal 16,17-seco-16,17a-dinitriles: identification of a selective inhibitor of hormone-independent breast cancer cells. *Bioorg Med Chem* 23:703–11. doi: 10.1016/j.bmc.2014.12.069

175. Consortium TU (2014) UniProt: a hub for protein information. *Nucleic Acids Res* 43:D204–D212. doi: 10.1093/nar/gku989
176. Dennis G, Sherman BT, Hosack DA, et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4:P3.
177. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30.
178. Boyle EI, Weng S, Gollub J, et al. (2004) GO::TermFinder - Open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20:3710–3715. doi: 10.1093/bioinformatics/bth456
179. Harris M a, Clark J, Ireland a, et al. (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32:D258–D261. doi: 10.1093/nar/gkh036
180. Lyne R, Smith R, Rutherford K, et al. (2007) FlyMine: an integrated database for *Drosophila* and *Anopheles* genomics. *Genome Biol* 8:R129. doi: 10.1186/gb-2007-8-7-r129
181. Harmston N, Filsell W, Stumpf MPH (2012) Which species is it? Species-driven gene name disambiguation using random walks over a mixture of adjacency matrices. *Bioinformatics* 28:254–60. doi: 10.1093/bioinformatics/btr640
182. Harmston N, Filsell W, Stumpf MPH (2010) What the papers say: text mining for genomics and systems biology. *Hum Genomics* 5:17–29. doi: 10.1186/1479-7364-5-1-17
183. Li C (2014) Biological network evaluation and relation discovery from scientific literature. 162.
184. Agarwal P, Searls DB (2008) Literature mining in support of drug discovery. *Brief Bioinform* 9:479–492. doi: 10.1093/bib/bbn035
185. Yang Y, Adelstein SJ, Kassis AI (2012) Target discovery from data mining approaches. *Drug Discov Today* 17 Suppl:S16-23. doi: 10.1016/j.drudis.2011.12.006

186. Hahn U, Cohen KB, Garten Y, Shah NH (2012) Mining the pharmacogenomics literature--a survey of the state of the art. *Brief Bioinform* 13:460–94. doi: 10.1093/bib/bbs018
187. Hettne KM, de Mos M, de Bruijn AGJ, et al. (2007) Applied information retrieval and multidisciplinary research: new mechanistic hypotheses in complex regional pain syndrome. *J Biomed Discov Collab* 2:2–17. doi: 10.1186/1747-5333-2-2
188. Fluck J, Madan S, Ansari S, et al. (2016) Training and evaluation corpora for the extraction of causal relationships encoded in biological expression language (BEL). *Database (Oxford)*. doi: 10.1093/database/baw113
189. Rinaldi F, Ellendorff TR, Madan S, et al. (2016) BioCreative V track 4: a shared task for the extraction of causal network information using the Biological Expression Language. *Database (Oxford)*. doi: 10.1093/database/baw067
190. Borrajo L, Romero R, Iglesias EL, Redondo Marey CM (2011) Improving imbalanced scientific text classification using sampling strategies and dictionaries. *J Integr Bioinform* 8:176. doi: 10.2390/biecoll-jib-2011-176
191. Hoehndorf R, Ngonga Ngomo A-C, Pyysalo S, et al. (2011) Ontology design patterns to disambiguate relations between genes and gene products in GENIA. *J Biomed Semantics* 2:S1. doi: 10.1186/2041-1480-2-S5-S1
192. Cheng D, Knox C, Young N, et al. (2008) PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites. *Nucleic Acids Res* 36:W399–W405. doi: 10.1093/nar/gkn296
193. Liu Y, Liang Y, Wishart D (2015) PolySearch2: a significantly improved text-mining system for discovering associations between human diseases, genes, drugs, metabolites, toxins and more. *Nucleic Acids Res* 43:W535–W542. doi: 10.1093/nar/gkv383
194. Frijters R, van Vugt M, Smeets R, et al. (2010) Literature mining for the discovery of hidden connections between drugs, genes and diseases. *PLoS Comput Biol* 6:e1000943. doi: 10.1371/journal.pcbi.1000943
195. Barbosa-Silva A, Fontaine J-F, Donnard ER, et al. (2011) PESCADOR, a web-based tool to assist text-mining of biointeractions extracted from PubMed queries. *BMC Bioinformatics* 12:435. doi: 10.1186/1471-2105-12-435

196. Wang J-H, Zhao L-F, Lin P, et al. (2014) GenCLiP 2.0: a web server for functional clustering of genes and construction of molecular networks based on free terms. *Bioinformatics* 30:2534–6. doi: 10.1093/bioinformatics/btu241
197. Mandloi S, Chakrabarti S (2015) PALM-IST: Pathway Assembly from Literature Mining - an Information Search Tool. *Sci Rep* 5:10021. doi: 10.1038/srep10021
198. Hur J, Schuyler AD, States DJ, Feldman EL (2009) SciMiner: web-based literature mining tool for target identification and functional enrichment analysis. *Bioinformatics* 25:838–40. doi: 10.1093/bioinformatics/btp049
199. Chen B, Ding Y, Wild DJ (2012) Assessing drug target association using semantic linked data. *PLoS Comput Biol* 8:e1002574. doi: 10.1371/journal.pcbi.1002574
200. Kuhn M, Szklarczyk D, Pletscher-Frankild S, et al. (2014) STITCH 4: integration of protein-chemical interactions with user data. *Nucleic Acids Res* 42:D401-7. doi: 10.1093/nar/gkt1207
201. Sousa SF, Ribeiro AJM, Coimbra JTS, et al. (2013) Protein-ligand docking in the new millennium--a retrospective of 10 years in the field. *Curr Med Chem* 20:2296–314.
202. Naïm M, Bhat S, Rankin KN, et al. (2007) Solvated Interaction Energy (SIE) for Scoring Protein–Ligand Binding Affinities. 1. Exploring the Parameter Space. *J Chem Inf Model* 47:122–133. doi: 10.1021/ci600406v
203. Deng Z, Chuaqui C, Singh J (2004) Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J Med Chem* 47:337–44. doi: 10.1021/jm030331x
204. Lin J-H (2011) Accommodating protein flexibility for structure-based drug design. *Curr Top Med Chem* 11:171–8.
205. Klebe G (2006) Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov Today* 11:580–94. doi: 10.1016/j.drudis.2006.05.012
206. Bissantz C, Kuhn B, Stahl M (2010) A medicinal chemist's guide to molecular interactions. *J Med Chem* 53:5061–84. doi: 10.1021/jm100112j

207. Salsbury FR (2010) Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Curr Opin Pharmacol* 10:738–44. doi: 10.1016/j.coph.2010.09.016
208. Cornell WD, Cieplak P, Bayly CI, et al. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J Am Chem Soc* 117:5179–5197. doi: 10.1021/ja00124a002
209. Brooks BR, Bruccoleri RE, Olafson BD, et al. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217. doi: 10.1002/jcc.540040211
210. Christen M, Hünenberger PH, Bakowies D, et al. (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem* 26:1719–1751. doi: 10.1002/jcc.20303
211. Micheletti C, Laio A, Parrinello M (2004) Reconstructing the Density of States by History-Dependent Metadynamics. *Phys Rev Lett* 92:170601. doi: 10.1103/PhysRevLett.92.170601
212. Dickson A, Brooks CL (2014) WExplore: hierarchical exploration of high-dimensional spaces using the weighted ensemble algorithm. *J Phys Chem B* 118:3532–42. doi: 10.1021/jp411479c
213. Abdul-Wahid B, Feng H, Rajan D, et al. (2014) AWE-WQ: Fast-Forwarding Molecular Dynamics Using the Accelerated Weighted Ensemble. *J Chem Inf Model* 54:3033–3043. doi: 10.1021/ci500321g
214. Kubitzki MB, de Groot BL (2007) Molecular Dynamics Simulations Using Temperature-Enhanced Essential Dynamics Replica Exchange. *Biophys J* 92:4262–4270. doi: 10.1529/biophysj.106.103101
215. Durrant JD, McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biol* 9:71. doi: 10.1186/1741-7007-9-71
216. Harvey MJ, De Fabritiis G (2012) High-throughput molecular dynamics: the powerful new tool for drug discovery. *Drug Discov Today* 17:1059–1062. doi: 10.1016/j.drudis.2012.03.017
217. Alonso H, Bliznyuk AA, Gready JE (2006) Combining docking and molecular dynamic simulations in drug design. *Med Res Rev* 26:531–568. doi: 10.1002/med.20067

218. Chen Y-C (2015) Beware of docking! *Trends Pharmacol Sci* 36:78–95. doi: 10.1016/j.tips.2014.12.001
219. Bond SD, Leimkuhler BJ, Laird BB (1999) The Nosé–Poincaré Method for Constant Temperature Molecular Dynamics. *J Comput Phys* 151:114–134. doi: 10.1006/jcph.1998.6171
220. Walker EH, Pacold ME, Perisic O, et al. (2000) Structural determinants of phosphoinositide 3-kinase inhibition by wortmannin, LY294002, quercetin, myricetin, and staurosporine. *Mol Cell* 6:909–19.
221. Teillet F, Boumendjel A, Boutonnat J, Ronot X (2008) Flavonoids as RTK inhibitors and potential anticancer agents. *Med Res Rev* 28:715–745. doi: 10.1002/med.20122
222. Davies M, Nowotka M, Papadatos G, et al. (2015) ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* doi: 10.1093/nar/gkv352
223. Günther S, Kuhn M, Dunkel M, et al. (2008) SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic Acids Res* 36:D919–22. doi: 10.1093/nar/gkm862
224. Law V, Knox C, Djoumbou Y, et al. (2014) DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 42:D1091-7. doi: 10.1093/nar/gkt1068
225. Liu T, Lin Y, Wen X, et al. (2007) BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res* 35:D198-201. doi: 10.1093/nar/gkl999
226. Yu S-WS-W, Andrabi SA, Wang H, et al. (2006) Apoptosis-inducing factor mediates poly(ADP-ribose) (PAR) polymer-induced cell death. *Proc Natl Acad Sci U S A* 103:18314–9. doi: 10.1073/pnas.0606528103
227. Blasina MF, Vaamonde L, Morquio A, et al. (2009) Differentiation induced by *Achyrocline satureioides* (Lam) infusion in PC12 cells. *Phytother Res* 23:1263–9. doi: 10.1002/ptr.2607
228. Chen P-Y, Wu M-J, Chang H-Y, et al. (2015) Up-Regulation of miR-34a Expression in Response to the Luteolin-Induced Neurite Outgrowth of PC12 Cells. *J Agric Food Chem* 63:4148–59. doi: 10.1021/acs.jafc.5b01005

229. Spencer JPE (2009) Flavonoids and brain health: multiple effects underpinned by common mechanisms. *Genes Nutr* 4:243–50. doi: 10.1007/s12263-009-0136-3
230. Granville CA, Memmott RM, Gills JJ, Dennis PA (2006) Handicapping the race to develop inhibitors of the phosphoinositide 3-kinase/Akt/mammalian target of rapamycin pathway. *Clin Cancer Res* 12:679–89. doi: 10.1158/1078-0432.CCR-05-1654
231. Han D, Scott EL, Dong Y, et al. (2015) Attenuation of mitochondrial and nuclear p38 α signaling: a novel mechanism of estrogen neuroprotection in cerebral ischemia. *Mol Cell Endocrinol* 400:21–31. doi: 10.1016/j.mce.2014.11.010
232. Hynes NE, Lane HA (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer* 5:341–54. doi: 10.1038/nrc1609
233. Schroeter H, Boyd C, Spencer JPE, et al. MAPK signaling in neurodegeneration: influences of flavonoids and of nitric oxide. *Neurobiol Aging* 23:861–80.
234. Sebolt-Leopold JS, Herrera R (2004) Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nat Rev Cancer* 4:937–47. doi: 10.1038/nrc1503
235. Zhang X, Zhang L, Wang G, et al. (2015) Suppression of mitochondrial fission in experimental cerebral ischemia: The potential neuroprotective target of p38 MAPK inhibition. *Neurochem Int* 1–8. doi: 10.1016/j.neuint.2015.06.010
236. Zhao Y-F, Zhang Q, Xi J-Y, et al. (2015) Multitarget intervention of Fasudil in the neuroprotection of dopaminergic neurons in MPTP-mouse model of Parkinson's disease. *J Neurol Sci* 353:28–37. doi: 10.1016/j.jns.2015.03.022
237. Zhou H-Y, Huang S-L (2012) Current development of the second generation of mTOR inhibitors as anticancer agents. *Chin J Cancer* 31:8–18. doi: 10.5732/cjc.011.10281
238. Russo GL, Russo M, Spagnuolo C, et al. (2014) Quercetin: a pleiotropic kinase inhibitor against cancer. *Cancer Treat Res* 159:185–205. doi: 10.1007/978-3-642-38007-5_11

239. Dieci MV, Arnedos M, Andre F, Soria JC (2013) Fibroblast growth factor receptor inhibitors as a cancer treatment: from a biologic rationale to medical perspectives. *Cancer Discov* 3:264–79. doi: 10.1158/2159-8290.CD-12-0362
240. Fang J, Zhou Q, Shi X, Jiang B (2007) Luteolin inhibits insulin-like growth factor 1 receptor signaling in prostate cancer cells. *Carcinogenesis* 28:713–23. doi: 10.1093/carcin/bgl189
241. Lehtiö L, Chi NW, Krauss S (2013) Tankyrases as drug targets. *FEBS J* 280:3576–3593. doi: 10.1111/febs.12320
242. Jain AK, Jaiswal AK (2007) GSK-3 β acts upstream of Fyn kinase in regulation of nuclear export and degradation of NF-E2 related factor 2. *J Biol Chem* 282:16502–16510.
243. Yoshino Y, Ishioka C (2015) Inhibition of glycogen synthase kinase-3 β induces apoptosis and mitotic catastrophe by disrupting centrosome regulation in cancer cells. *Sci Rep* 5:13249. doi: 10.1038/srep13249
244. Qi Z, Xu Y, Liang Z, et al. (2015) Baicalein alters PI3K/Akt/GSK3 β signaling pathway in rats with diabetes-associated cognitive deficits. *Int J Clin Exp Med* 8:1993–2000.
245. Salcedo-Tello P, Ortiz-Matamoros A, Arias C (2011) GSK3 Function in the Brain during Development, Neuronal Plasticity, and Neurodegeneration. *Int J Alzheimers Dis* 2011:189728. doi: 10.4061/2011/189728
246. Buryanovskyy L, Fu Y, Boyd M, et al. (2004) Crystal structure of quinone reductase 2 in complex with resveratrol. *Biochemistry* 43:11417–26. doi: 10.1021/bi049162o
247. Warren GL, Andrews CW, Capelli A-M, et al. (2006) A critical assessment of docking programs and scoring functions. *J Med Chem* 49:5912–31. doi: 10.1021/jm050362n
248. Xie L, Bourne PE (2005) Functional coverage of the human genome by existing structures, structural genomics targets, and homology models. *PLoS Comput Biol* 1:e31. doi: 10.1371/journal.pcbi.0010031
249. Cruciani G, Carosati E, De Boeck B, et al. (2005) MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *J Med Chem* 48:6970–6979. doi: 10.1021/jm050529c

250. Kandoi G, Acencio ML, Lemke N (2015) Prediction of Druggable Proteins Using Machine Learning and Systems Biology: A Mini-Review. *Front Physiol.* doi: 10.3389/fphys.2015.00366
251. Boly R, Gras T, Lamkami T, et al. (2011) Quercetin inhibits a large panel of kinases implicated in cancer cell biology. *Int J Oncol* 38:833–42. doi: 10.3892/ijo.2010.890
252. Ghosh A, Sarkar S, Mandal AK, Das N (2013) Neuroprotective role of nanoencapsulated quercetin in combating ischemia-reperfusion induced neuronal damage in young and aged rats. *PLoS One* 8:e57735. doi: 10.1371/journal.pone.0057735
253. Popkie AP, Zeidner LC, Albrecht AM, et al. (2010) Phosphatidylinositol 3-kinase (PI3K) signaling via glycogen synthase kinase-3 (Gsk-3) regulates DNA methylation of imprinted loci. *J Biol Chem* 285:41337–47. doi: 10.1074/jbc.M110.170704
254. Wang Y, Zhang ZZ, Wu Y, et al. (2013) Quercetin postconditioning attenuates myocardial ischemia / reperfusion injury in rats through the PI3K / Akt pathway. *46:861–867.*
255. Chen J, Hu R, Ge H, et al. (2015) G-protein-coupled receptor 30-mediated antiapoptotic effect of estrogen on spinal motor neurons following injury and its underlying mechanisms. *Mol Med Rep* 12:1733–40. doi: 10.3892/mmr.2015.3601
256. Zhu C, Wang S, Wang B, et al. (2015) 17 β -Estradiol up-regulates Nrf2 via PI3K/AKT and estrogen receptor signaling pathways to suppress light-induced degeneration in rat retina. *Neuroscience* 304:328–339. doi: 10.1016/j.neuroscience.2015.07.057
257. Maier JA, Martinez C, Kasavajhala K, et al. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* 11:3696–3713. doi: 10.1021/acs.jctc.5b00255
258. Gerber PR, Müller K (1995) MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J Comput Aided Mol Des* 9:251–68.
259. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem* 17:490–512.

260. William L. Jorgensen *, David S. Maxwell and, Tirado-Rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. doi: 10.1021/JA9621760
261. Yamanishi Y, Araki M, Gutteridge A, et al. (2008) Prediction of drug-target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* 24:i232-40. doi: 10.1093/bioinformatics/btn162
262. van Laarhoven T, Nabuurs SB, Marchiori E (2011) Gaussian interaction profile kernels for predicting drug-target interaction. *Bioinformatics* 27:3036–43. doi: 10.1093/bioinformatics/btr500
263. Science C, Gönen M (2012) Predicting drug-target interactions from chemical and genomic kernels using Bayesian matrix factorization. *Bioinformatics* 28:2304–10. doi: 10.1093/bioinformatics/bts360
264. Wang Y, Zeng J (2013) Predicting drug-target interactions using restricted Boltzmann machines. *Bioinformatics* 29:i126–i134. doi: 10.1093/bioinformatics/btt234
265. Wang Y, Chen S, Deng N, Wang Y (2013) Drug repositioning by kernel-based integration of molecular structure, molecular activity, and phenotype data. *PLoS One* 8:e78518. doi: 10.1371/journal.pone.0078518
266. Wang Y, Liu Z, Li C, et al. (2012) Drug target prediction based on the herbs components: the study on the multitargets pharmacological mechanism of qishenkeli acting on the coronary heart disease. *Evid Based Complement Alternat Med* 2012:698531. doi: 10.1155/2012/698531
267. Takarabe M, Kotera M, Nishimura Y, et al. (2012) Drug target prediction using adverse event report systems: a pharmacogenomic approach. *Bioinformatics* 28:i611–i618. doi: 10.1093/bioinformatics/bts413
268. Ye H, Liu Q, Wei J (2014) Construction of drug network based on side effects and its application for drug repositioning. *PLoS One* 9:e87864. doi: 10.1371/journal.pone.0087864

APÉNDICES

Apéndice 1

Campo de fuerza de MOE

Varios métodos han sido desarrollados en las pasadas décadas para modelar la estructura de macromoléculas biológicas especialmente proteínas. En el marco de esta tesis se destacan los modelos y herramientas de mecánica molecular. Siguiendo la aproximación de Born-Oppenheimer, la mecánica molecular asume que la energía potencial de un sistema molecular es función de las coordenadas nucleares, donde los electrones son considerados implícitamente. Esto posibilita modelar desde sistemas pequeños hasta complejos con miles a millones de átomos.

A la forma funcional y el conjunto de parámetros utilizado para el cálculo de la energía potencial de un sistema se denomina campo de fuerza. En los campos de fuerza más extendidos cada átomo es simulado como una partícula con al menos masa, radio y carga neta asignada. Una función típica tiene entonces términos para describir la contribución covalente (átomos enlazados) y no covalente (interacciones de largo alcance electrostáticas y van der Waals) a la energía potencial. Los términos enlazantes consisten usualmente en funciones armónicas que no permiten la ruptura del enlace o los más costosos potenciales de Morse que ofrecen una descripción más realista. Las interacciones de van der Waals son usualmente calculadas con un potencial de Lennard-Jones (van der Waals) y las interacciones electrostáticas mediante la ley de Coulomb. Otros términos considerados son los términos cruzados que describen el acoplamiento entre variables internas como enlaces y ángulos, o términos que aseguren la planaridad de sistemas aromáticos (diedros impropios, “out of plane”). En adición a la forma funcional los campos de fuerza definen tipos de átomos asociados a un subconjunto de parámetros. Un conjunto de parámetros típico, usualmente empírico, incluye valores de masa atómica, radio de van der Waals y carga parcial, así como constantes de fuerza y punto de equilibrio para largo de enlace, ángulos, diedros, etc.

Se tiene por ejemplo la forma funcional del campo de fuerza del programa *MOE*:

$$E(x) = E_{str} + E_{ang} + E_{stb} + E_{oop} + E_{tor} + E_{vdw} + E_{ele} + E_{sol} + E_{res}$$

donde x son las posiciones atómicas. Cada uno de los términos de energía es en si mismo una sumatoria que involucra tipos particulares de interacciones atómicas.

E_{str} corresponde a la energía del estiramiento del enlace:

$$E_{str} = w_{str} \sum_{i-j} k_{ij} (r_{ij} - L_{ij})^2 + k'_{ij} (r_{ij} - L_{ij})^3 + k''_{ij} (r_{ij} - L_{ij})^4$$

donde la suma se extiende sobre todos los enlaces $i-j$; w_{str} es un coeficiente de ponderación; k_{ij} , k'_{ij} y k''_{ij} son constantes de fuerza y L_{ij} es el largo del enlace en su punto de equilibrio. Los potenciales de Morse pueden ser aproximados suficientemente bien al expandirlos en series de Taylor entorno al punto de equilibrio y truncarlos en el término cuartico. De esta manera el término lineal será cero en el equilibrio y el término constante puede despreciarse.

E_{ang} es la energía de flexión del enlace:

$$E_{ang} = w_{ang} \sum_{i-j-k} k_{ijk} d_{ijk}^2 + k'_{ijk} d_{ijk}^3 + k''_{ijk} d_{ijk}^4$$

donde la suma se extiende sobre todos los ángulos de enlace $i-j-k$; w_{ang} es un coeficiente de ponderación; k_{ijk} , k'_{ijk} y k''_{ijk} son constantes de fuerza; A_{ijk} es el ángulo de equilibrio (radianes), d_{ijk} puede ser $a_{ijk} - A_{ijk}$ (forma angular) o $\cos(a_{ijk}) - \cos(A_{ijk})$ (forma coseno) dependiendo de la forma funcional elegida para las interacciones, y

$$a_{ijk} = \cos^{-1} \frac{x_{ji}^T x_{jk}}{r_{ji} r_{jk}}, \quad x_{ji} = x_j - x_i$$

Para moléculas lineales, $k_{ijk} (1 + \cos a_{ijk})$ es usado en lugar de la expresión cuartica. Para octaedros y geométricas cuadradas planas $k_{ijk} (1 - \cos (4a_{ijk}))$ es usado en lugar de la expresión cuartica.

E_{stb} es un término conjugado de estiramiento-flexión del enlace:

$$E_{stb} = w_{stb} \sum_{i-j-k} (k_{ijk} (r_{ij} - L_{ij}) + k_{kji} (r_{jk} - L_{jk})) d_{ijk}$$

donde la suma se extiende sobre todos los ángulos de enlace $i-j-k$; w_{stb} es un coeficiente de ponderación, k_{ijk} y k_{kji} son constantes de fuerza estiramiento-flexión y el resto de los términos son como los presentes en las funciones de estiramiento y de torsión del enlace. Estos términos conjugados no son utilizados para ángulos de enlace lineales o casi lineales.

E_{oop} es la energía por fuera del plano:

$$E_{oop} = w_{oop} \sum_{i;jkl} k_{i;jkl} X_{i;jkl}^2$$

donde la suma de términos se extiende sobre todos los átomos i con tres vecinos j , k e l ; w_{oop} es un coeficiente de ponderación, $k_{i;jkl}$ es la constante de fuerza por fuera del plano, y $X_{i;jkl}$ es el ángulo de Wilson (el ángulo entre el enlace il y el plano ijk). En algunos campos de fuerza la ecuación de energía fuera del plano es la misma que para las energías de torsión excepto que en lugar de usar átomos $i-j-k-l$ unidos en secuencia, la suma se extiende sobre todos los átomos i con tres vecinos unidos j , k y l .

E_{tor} es la energía de torsión del ángulo diedro:

$$E_{tor} = w_{tor} \sum_{i-j-k-l} \sum_{n=0}^6 k_{n;ijkl} (1 + \cos(nT_{ijkl} - P_{n;ijkl}))$$

donde la suma se extiende sobre todos los conjuntos de cuatro átomos unidos consecutivamente $i-j-k-l$; w_{tor} es un coeficiente de ponderación, $k_{n;ijkl}$ son constantes de fuerza, $P_{n;ijkl}$ son parámetros de fase de ángulo y:

$$T_{ijkl} = \cos^{-1} \frac{(x_{ji} \times x_{jk})^T (x_{jk} \times x_{lk})}{|x_{ji} \times x_{jk}| |x_{jk} \times x_{lk}|}, \quad x_{ji} = x_j - x_i$$

E_{vdw} es la energía de van der Waals:

$$E_{vdw} = w_{vdw} \sum_{i<j} e_{ij} \left[\frac{(1+a)R_{ij}}{r_{ij} + aR_{ij}} \right]^{n_{ij}} \left[\frac{n_{ij} (1+b)R_{ij}^{m_{ij}}}{m_{ij} r_{ij}^{m_{ij}} + bR_{ij}^{m_{ij}}} - \frac{m_{ij} + n_{ij}}{m_{ij}} \right] s(r_{ij}) T_{ij} I_{ij}^{vdw}$$

donde w_{vdw} es un coeficiente de ponderación, y e_{ij} , R_{ij} , m_{ij} y n_{ij} son parámetros del campo de fuerza. Las constantes a y b previenen la división por 0 y son llamadas constantes de amortiguamiento. I^{vdw} es un factor de escala de la interacción definida como 0 para interacciones 1-2 y 1-3, un valor dependiente del conjunto de parámetros

para interacciones 1-4, y 1 para otras interacciones. Aunque compleja, con una apropiada selección de parámetros esta forma puede reproducir los potenciales usuales 12-6, 12-10 y 9-6. s es una función de suavización (*cutoff*):

$$s(r) = \begin{cases} 1 & \text{if } r < r_0 \\ 1 - p((r - r_0)/(r_1 - r_0)) & \text{if } r \in [r_0, r_1] \\ 0 & \text{if } r > r_1 \end{cases}$$

donde $p(x) = x^3(6x^2 - 15x + 10)$. Esta función polinómica tiene las propiedades de $p(1)=1$, $p(0)=0$, $p'(0)=p'(1)=0$ y $p''(0)=p''(1)=0$. Al definir los parámetros de corte r_0 y r_1 , una variedad de funciones de afinamiento que son continuas en su derivadas primera y segunda pueden ser creadas.

T_{ij} es un factor de escala de la interacción usado para escalar interacciones no enlazantes. Asociado con cada átomo i hay un estado con valor T_i . Dados dos átomos i y j , el factor T_{ij} es definido:

$$T_{ij} = \begin{cases} T_{like} & \text{if } T_i = T_j \\ T_{wild} & \text{if } T_i \neq T_j \text{ and } (T_i = 0 \text{ or } T_j = 0) \\ T_{unlike} & \text{otherwise} \end{cases}$$

donde T_{like} , T_{unlike} y T_{wild} son parámetros de estado que pueden ser ajustados para habilitar/deshabilitar interacciones no enlazantes entre moléculas.

E_{ele} es la energía electrostática:

$$E_{ele} = \begin{cases} \frac{w_{ele} e^2}{4\pi\epsilon_0 d} \sum_{i < j} q_i q_j \left[\frac{1}{r_{ij} + b_{ele}} \right] s(r_{ij}) T_{ij} I_{ij}^{ele} & \text{Coulomb} \\ \frac{w_{ele} e^2}{4\pi\epsilon_0 d} \sum_{i < j} q_i q_j \left[\frac{1}{(r_{ij} + b_{ele})^2} \right] s(r_{ij}) T_{ij} I_{ij}^{ele} & \text{Distance Dependent Dielectric} \\ \frac{w_{ele} e^2}{4\pi\epsilon_0 d} \sum_{i < j} q_i q_j \left[\frac{1}{r_{ij} + b_{ele}} - \frac{\alpha r_{ij}^2}{R_c^3} - \frac{(1 - \alpha)}{R_c} \right] s(r_{ij}) T_{ij} I_{ij}^{ele} & \text{Reaction Field, } \alpha = \frac{d - d_x}{d + 2d_x} \end{cases}$$

donde w_{ele} es un coeficiente de ponderación, d es una constante dieléctrica en el interior del soluto, d_x es la constante dieléctrica del solvente, s y T son las mismas que en la función de van der Waals, q_i es la carga parcial sobre el átomo i , b_{ele} es una constante de amortiguamiento para prevenir denominadores de valor 0, R_c es una distancia de corte para interacciones no enlazantes. I_{ele} , de manera similar a I_{vdw} , es un factor de escala de la interacción definido como 0 para interacciones 1-2 y 1-3, un valor dependiente del conjunto de parámetro para interacciones 1-4, y 1 para otras interacciones.

E_{sol} es la energía implícita de solvatación usada el modelo de la generalizada de Born (GB/VI):

$$E_{sol} = -w_{sol}W(d^{-1} - d_x^{-1})\frac{e^2}{4\pi\epsilon_0}\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\frac{q_iq_j\sqrt{G_iG_j}}{\sqrt{y_{ij} + \exp(-y_{ij})/4}}s(r_{ij})T_{ij}, \quad y_{ij} = r_{ij}^2G_iG_j$$

donde w_{sol} y W son coeficientes de ponderación, d es la constante dieléctrica en el interior del soluto, d_x es la constante dieléctrica del solvente, s y T son las energías de van der Waals. G_i es la energía interna del átomo i , se define como:

$$G_i = -\frac{1}{2}\left[\frac{1}{R_i^3} - V_i\right]^{1/3}, \quad V_i \approx \iiint_{x \text{ resolute}} |x - x_i|^{-6} \delta(|x - x_i| > R_i) dx$$

donde R_i es el radio del átomo i del modelo de solvente y V_i en un aproximación a la integral de Born.

E_{res} es la energía de restricción. Las restricción son términos extra agregados a la función de potencia para imponer condiciones geométricas. Pueden aplicarse restricciones a la distancias, ángulos y diedros así como limitar el sistema a una determinada región del espacio.

El campo de fuerza de *MOE* es capaz de incluir una amplia familia de modelos entre los que se tiene *Amber12:EHT*, parametrizado para proteínas y ácidos nucleicos vía *Amber12* y parametrizado para moléculas pequeñas con *EHT (2D Extended Hueckel Theory)*; *CHARMM27*, parametrizado para proteínas y ácidos nucleicos (no se incluyen por defecto pequeñas moléculas); *MMFF94*, parametrizado para pequeñas moléculas orgánicas; *OPLS-AA*, parametrizado para proteínas y algunas moléculas orgánicas [209, 257–260].

Apéndice 2

Otras estrategias de tamizaje reverso, métodos probabilísticos y métodos basados en aprendizaje automático

Si bien su fundamento es el mismo que los métodos anteriores se distinguen de los mismos por utilizar específicamente herramientas de aprendizaje automático que no necesitan recurrir explícitamente a datos estructurales de ningún tipo.

Distinguen tres fuentes de información que denominan como “espacio químico” en el se describen las propiedades de los ligandos, “espacio de blancos” en el que se describen las propiedades de las proteínas blanco y finalmente la red de interacciones proteína-ligando. Esta información es utilizada para entrenar un algoritmo de aprendizaje automático que en general se limita a determinar nuevos pares proteína-ligando dentro de la propia base de partida. Su utilidad sería entonces la de identificar nuevas aplicaciones a moléculas con blancos de acción conocidos, lo que se denomina reposicionamiento, aunque hasta el momento solo hay trabajos de desarrollo del método. Los distintos métodos se distinguen por la fuente de información, y principalmente por el algoritmo de aprendizaje automático utilizado.

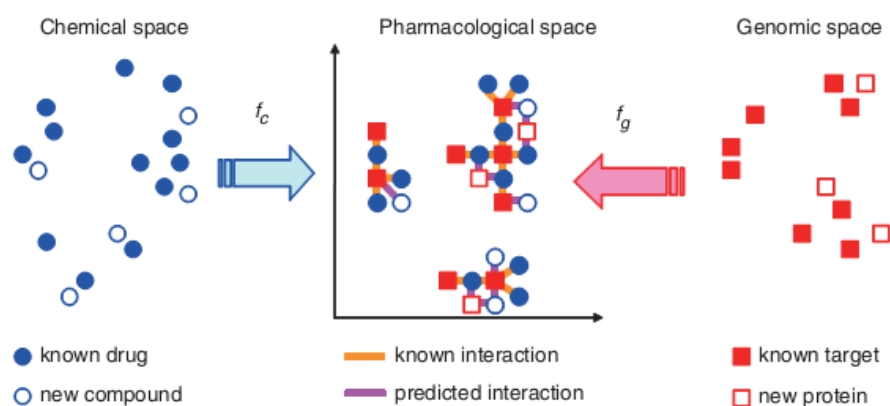


Figura 1. Esquema de búsqueda utilizado por los algoritmos de predicción de blancos basados en aprendizaje automático.

Yoshihiro Yamanishi y col. fueron los pioneros en el desarrollo de este tipo de estrategia [261]. En este trabajo el “espacio químico” se considera como similitud estructural de las moléculas, el “espacio genómico” o “espacio de blancos” como similitud de secuencia proteica y el “espacio farmacológico” como la interacción registrada en bases de datos como *KEGG*, *GLIDA*, *SuperTarget* y *Matador*, *BRENDA*

y *ChEMBL*. En particular, recopilan información de interacciones droga-blanco para enzimas, canales iónicos, receptores acoplados a proteínas G, y receptores nucleares.

Statistics for the drug–target interaction networks

Statistics	Enzyme	Ion channel	GPCR	Nuclear receptor
No. of drugs	445	210	223	54
No. of target proteins	664	204	95	26
(Total in human genome)	(2741)	(292)	(757)	(49)
No. of drug–target interactions	2926	1476	635	90

Figura 2. Base de datos de interacción utilizadas por los algoritmos de identificación de blancos basados en aprendizaje automático.

La similitud entre dos compuestos se calcula como el tamaño de todas las subestructuras iguales entre los mismos empleando el algoritmo SIMCOMP. Aplicando esta operación para todos los pares de compuestos se construye un matriz de similitud S_c , considerada el “espacio químico”. La similitud de secuencia se calcula como el *score* normalizado del alineamiento local (*Smith-Waterman score*) entre dos proteínas. Nuevamente aplicando esta operación para todos los pares de proteínas se construye una matriz de similitud denotada como S_g , considerada como el “espacio genómico”. Como limitante los autores reconocen que pueden existir blancos de acción con escasa similitud de secuencia que interaccionen con las mismas drogas. Finalmente se calculan perfiles de interacción droga-blanco como una cadena de caracteres “0” o “1”. Dadas una proteína y un blanco el objetivo del método es predecir si es posible una asociación entre los mismos. Para ello los autores emplearon tres algoritmos de aprendizaje automático: “*nearest profile*”, “*weighted profile*” y “*bipartite graph learning*”.

Apoyándose en el desarrollo de Yamanishi y col. han aparecido varios otros métodos que aplican los siguientes algoritmos: “*gaussian interaction profile kernel*”, “*bayesian matrix factorization*”, “*boltzman machines*”, “*support vector machine*”, entre otros. [262–265]. Asimismo han aparecido algunos métodos similares que utilizan datos sobre efectos biológicos como efectos adversos reportados como “*drugCIPHER*”, “*adverse event report system*” y “*drug side effect similarity inference*” [266–268].

Apéndice 3

Resultados preliminares del servidor PharmMapper

PharmMapper es un servidor de tamizaje reverso basado en farmacóforos del sitio de unión (<http://59.78.96.61/pharmmapper/>). Utilizando la estructura optimizada de quercetina (formato *.mol2*) se realizó una primera búsqueda de sitios de unión de proteínas que poseen características complementarias a las del flavonoide en el *Protein Data Bank*. De manera análoga al servidor *idTarget* los resultados son presentados de acuerdo a un *score* de ajuste de la molécula al sitio de unión (*fit score*) y un *Z'-score* asociado calculado en base a la distribución de *fit-scores* para un sitio dado. El propósito de este *Z'-score* es ponderar sitios grandes con muchos puntos farmacofóricos que pueden originar falsos positivos. Tomando las proteínas asociadas a un *Z'-score* mayor o igual a 0 y un *fit score* mayor o igual a 3.9 se tiene el siguiente conjunto de candidatos (ver tabla 1).

Tabla 1. Proteínas blanco identificadas en el *Protein Data Bank (PDB)* por tamizaje reverso basado en farmacóforo del sitio de unión con el servidor *PharmMapper* (<http://59.78.96.61/pharmmapper/>). No se presentan cristales redundantes de la misma proteína. Las proteínas candidato fueron segregadas de acuerdo a su organismo de origen y anotadas con su entrada correspondiente en el PDB, ligando co-cristalizado y su similitud estructural con quercetina si se dispone (*SHAFTS*), así como nombre de la proteína y score de *PharmMapper* (*Z'-score* y *fit-score*). Proteínas co-cristalizadas con quercetina o identificadas por otras herramientas de tamizaje se presentan remarcadas. Los score de similitud enmarcados en paréntesis se corresponden a ligandos semejantes a quercetina presentes en un cristal alternativo al identificado por *PharmMapper*. *GWVI/WSA dG score* es el score de docking asignado en *MOE (Molecular Operating Environment)* al alineamiento de quercetina en el sitio de unión asignado por *PharmMapper*.

PDB id	Ligando	Similitud	Nombre	Z-score	Fit-score	GBVI/WSA dG score (kcal/mol)
Homo sapiens						
1XO2	FSE	1.832	Cyclin-dependent kinase 6	2.5	4.4	-9.1
1I5R	HYC (EQI)	(1.172)	Estradiol 17-beta-dehydrogenase 1	1.7	5.0	-6.8
1NJS	KEU	-	Trifunctional purine biosynthetic protein adenosine-3	1.0	4.5	-6.6
5P21	GNP (KOB)	(1.090)	GTPase HRas (H-Ras-1)	0.8	5.6	-
2BDJ	HET (PSN)	(1.140)	Proto-oncogene tyrosine-protein kinase Src	0.6	4.0	-
2IW8	4SP (371)	(1.515)	Cyclin-dependent kinase 2	0.6	4.1	-7.8
1PYN	941 (F16)	(1.217)	Tyrosine-protein phosphatase non-receptor type 1	0.6	4.0	-5.8
1UDW	CTP	-	Uridine-cytidine kinase 2	0.5	5.1	-6.7

3C4F	C4F	1.229	Fibroblast growth factor receptor 1	0.3	4.0	-7.8
3EWH	K11 (BPK)	(1.090)	Vascular endothelial growth factor receptor 2	0.2	4.1	-7.0
2POG	WST	1.425	Estrogen receptor (ER) (ER-alpha)	0.2	3.8	-8.1
1D2S	DHT (EST)	(1.136)	Sex hormone-binding globulin	0.2	4.2	-6.7
1M4H	(6IP)	1.27	Beta-secretase 1	0.1	4.2	-5.5
1DB1	VDX	-	Vitamin D3 receptor (VDR)	0.0	4.3	87.1
1XAN	FAD (ELI)	1.136	Glutathione reductase, mitochondrial	0,0	3.8	-4.8
2ZB0	GK3 (NAR)	(1.712)	Mitogen-activated protein kinase 14	4.0	0.0	-6.8
1I9Q	IOF	1.131	Carbonic anhydrase 2	4.0	0.0	-
1W22	NHB (MCM)	1.136	Histone deacetylase 8	3.9	0.0	-6.2
Rattus norvegicus						
1HJ1	AOE (EST)	(1.136)	Estrogen receptor beta (ER-beta)	0.4	4.1	-
Mus musculus						
1AQU	EST	1.136	Estrogen sulfotransferase	1.3	4.9	-6.6
1REJ	B1L	1.140	cAMP-dependent protein kinase catalytic subunit alpha	0.5	4.3	-7.7
Sus scrofa						
1E90	MYC	1.744	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform (PI3-kinase subunit gamma)	0.3	3.8	-7.3
1JS3	142	1.227	Aromatic-L-amino-acid decarboxylase	0.3	3.8	-6.2
Bos taurus						
1TVY	GDU	-	Beta-1,4-galactosyltransferase 1	0.1	4.2	-6.2
1XH4	R69 (NU3)	1.315	cAMP-dependent protein kinase catalytic subunit alpha (PKA C-alpha)	0.1	3.9	-6.9
1QDQ	074	-	Cathepsin B	0.0	3.9	-5.9

ANEXOS

Anexo 1

Figuras suplementarias

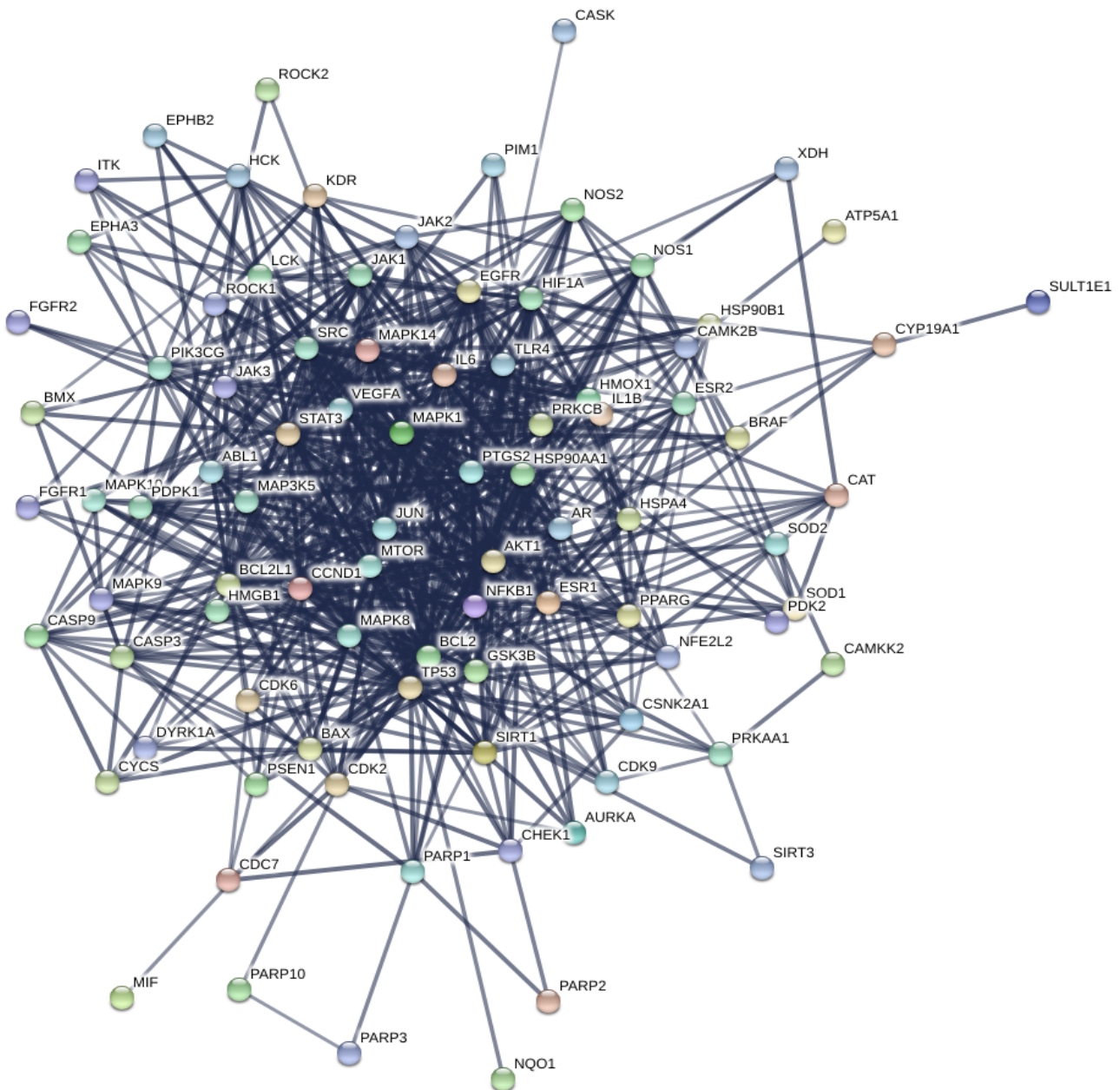


Figura 1. Representación de la red de interacción proteína-proteína de los candidatos a blanco junto a asociadas a queretina muerte celular y neuroprotección en la literatura. La base de datos STRING establece asociaciones funcionales entre pares de proteínas basado en evidencia experimental documentada, homología e inferencia por minería de texto incluyendo modificaciones proteicas, agregación, vecindad genómica y co-expresión. Solo asociaciones con alta confianza son presentadas (probabilidad mayor a 0,7).

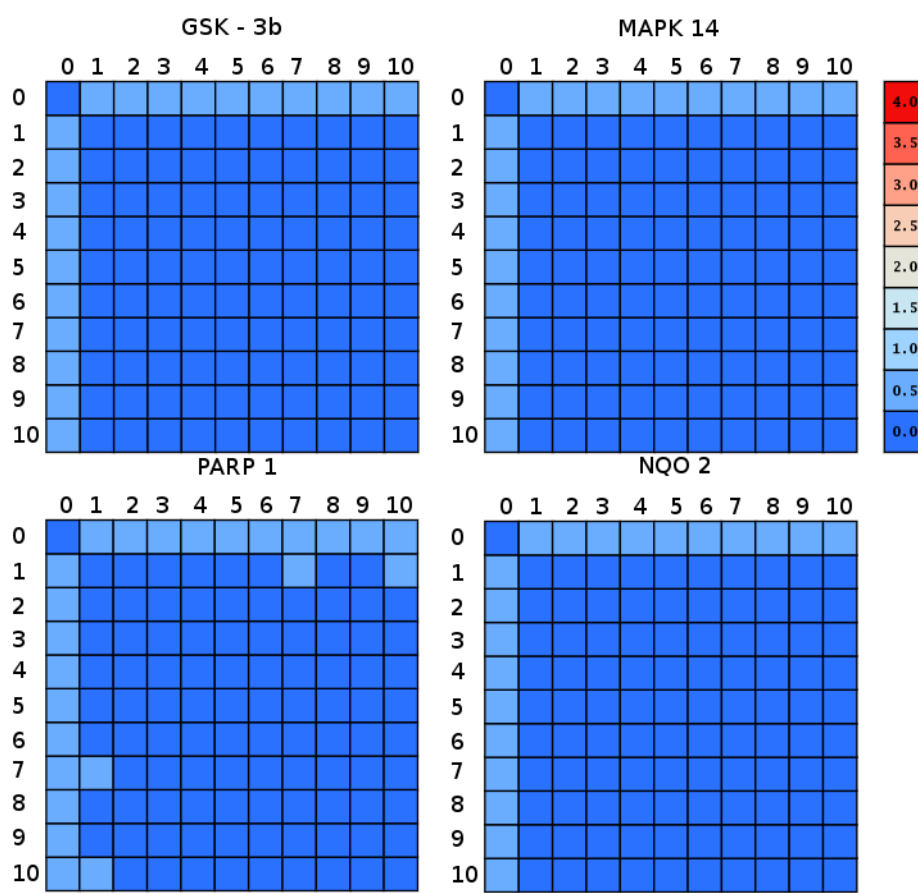


Figura 2. Desviación geométrica (RMSD, Root Mean Square Deviation) entre complejos de quercetina a lo largo de simulaciones de dinámica molecular. Para cada complejo se corrieron tres simulaciones de 10 ns analizadas en etapas de 1 ns. El último paso de cada etapa fue comparado con la estructura inicial (complejo energéticamente minimizado).

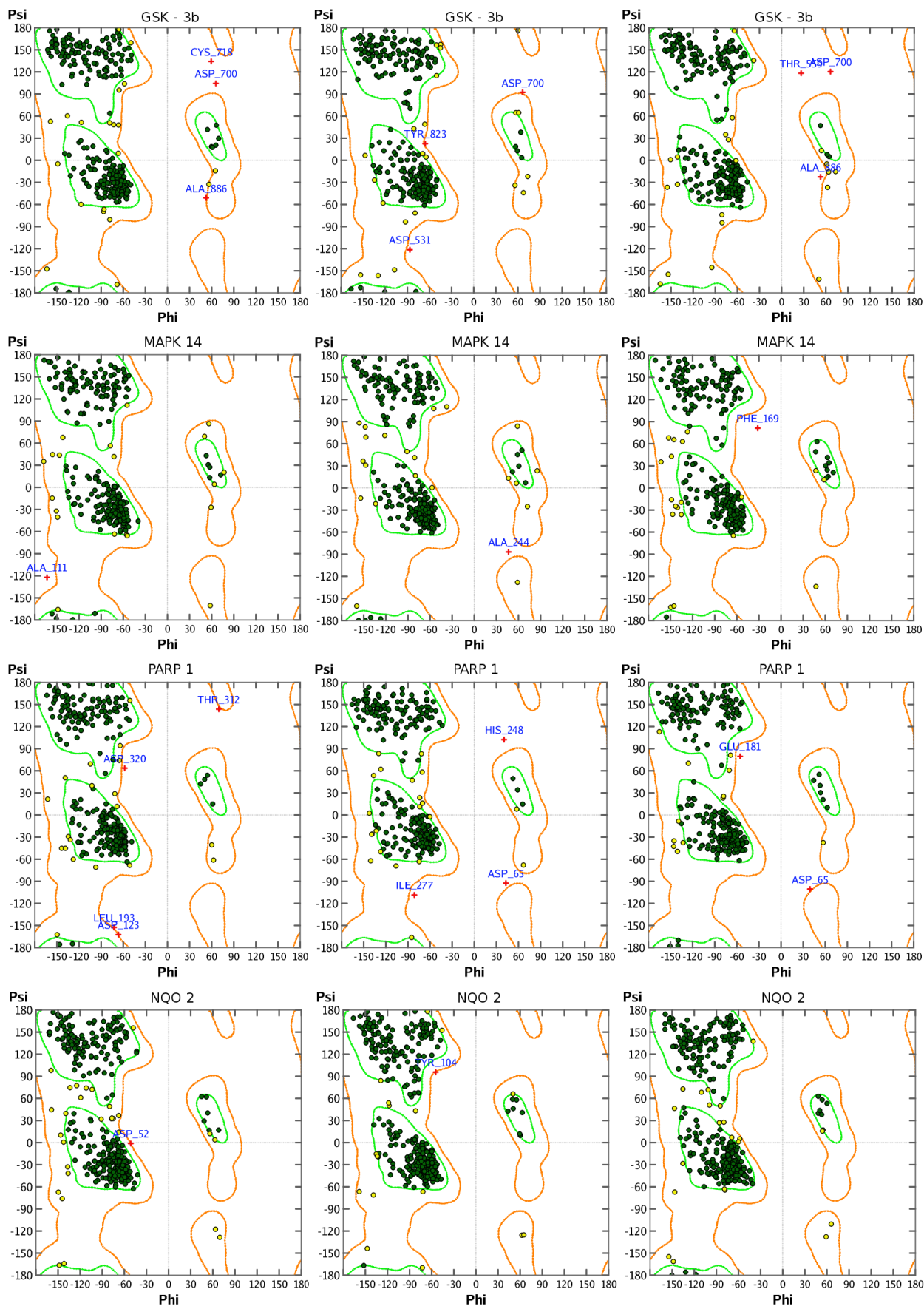


Figura 3. Gráficos de *Ramachandran* (psi/phi) de complejos quercetina-proteína blanco luego de simulaciones de dinámica molecular. Para cada complejo se corrieron tres simulaciones de 10 ns y la última etapa de cada simulación fué energeticamente minimizada (gradiente RMS 1 kcal, campo de fuerza *Amber12*) previo al cálculo del gráfico.

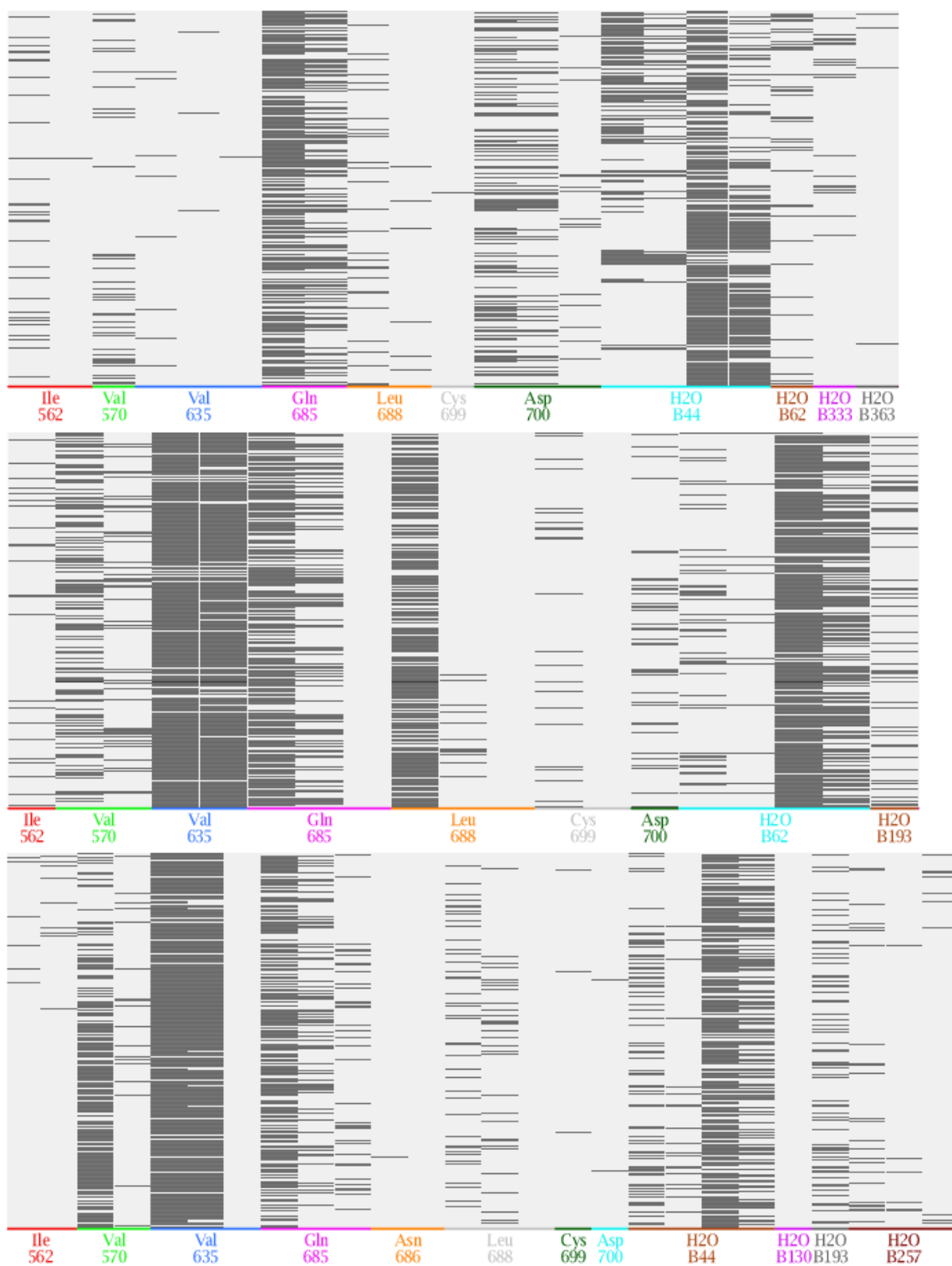


Figure 4. Patrón de interacción proteína ligando de complejos GSK3b-quercetina durante la última etapa (décimo ns) de tres simulaciones de dinámica molecular independientes.

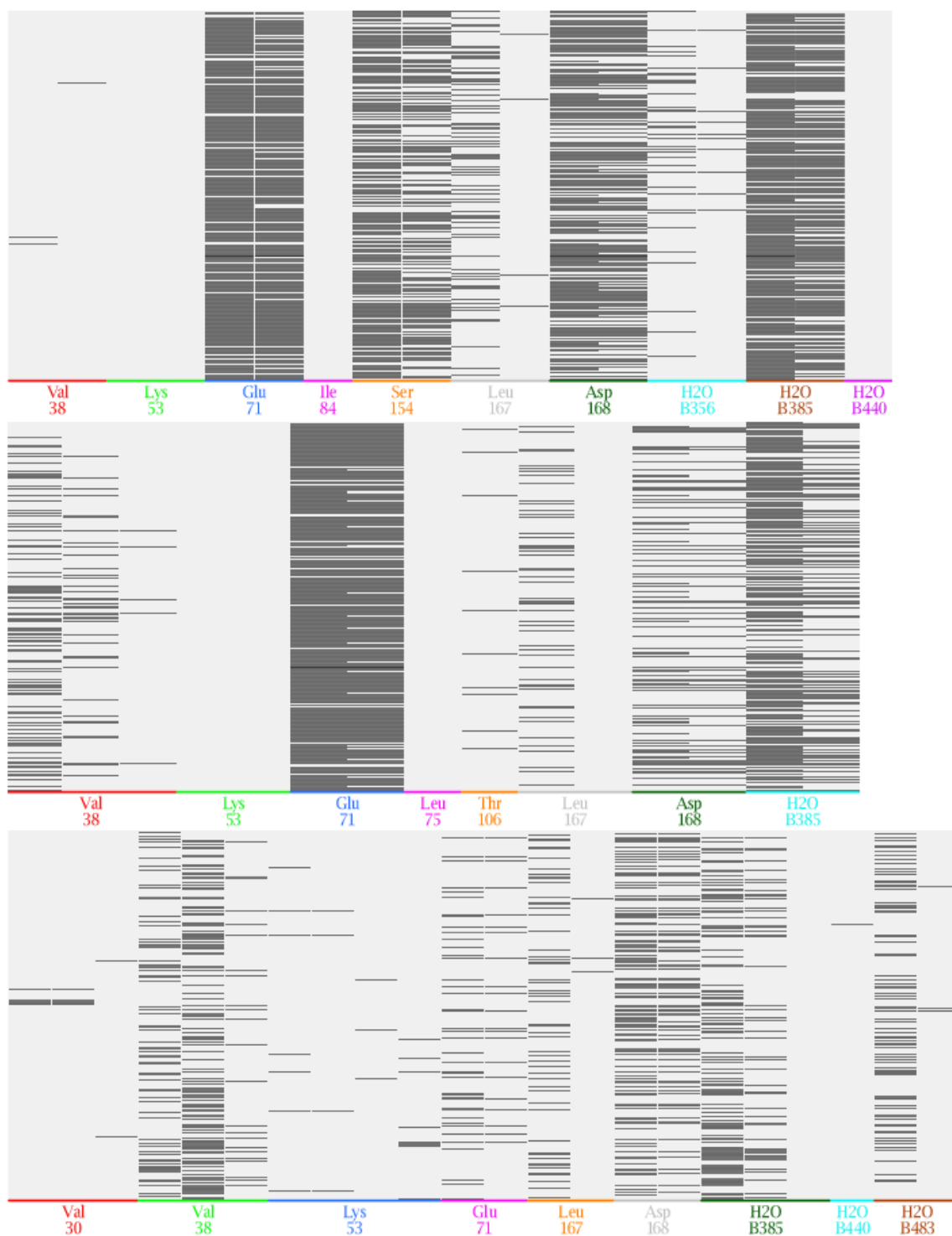


Figura 5. Patrón de interacción proteína ligando de complejos MAPK14-quercetina durante la última etapa (décimo ns) de tres simulaciones de dinámica molecular independientes.



Figura 6. Patrón de interacción proteína ligando de complejos PARP1-quercetina durante la última etapa (décimo ns) de tres simulaciones de dinámica molecular independientes.

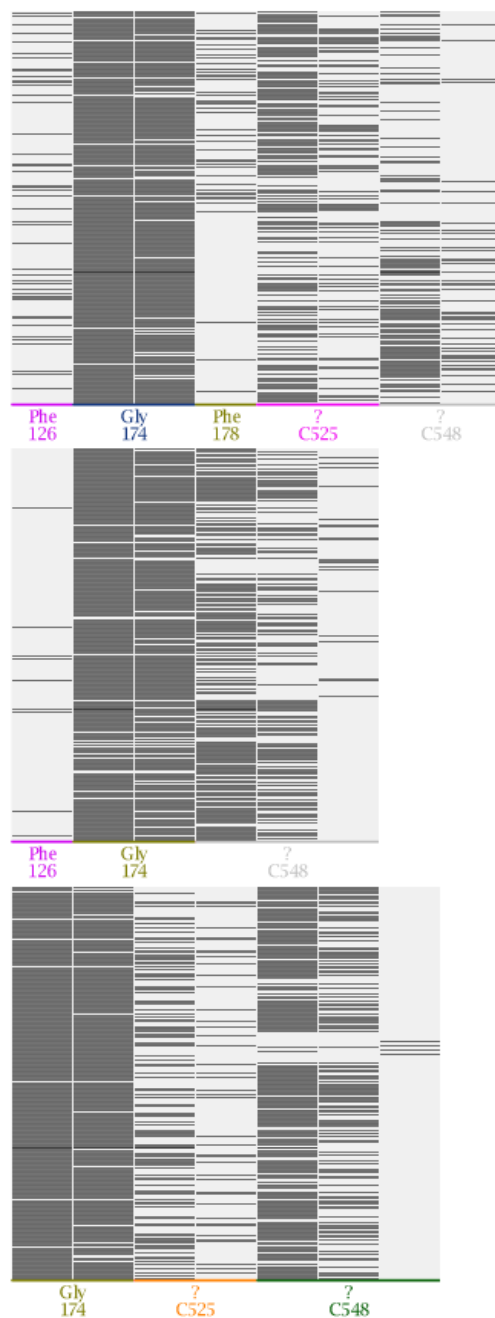
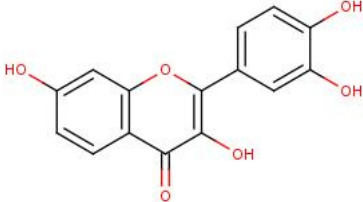
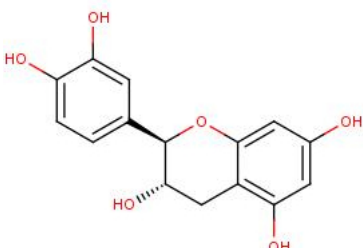
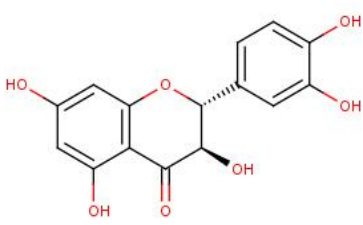
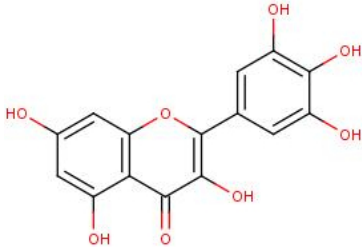
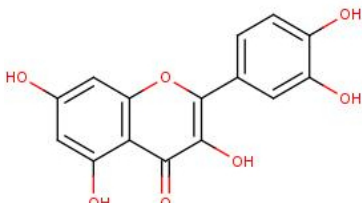


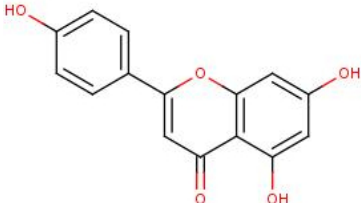
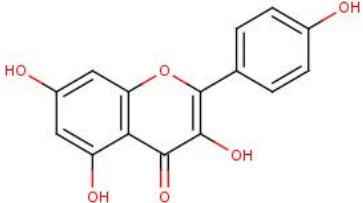
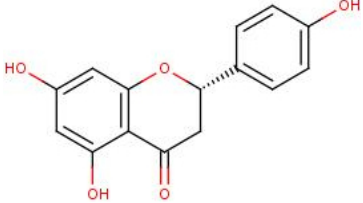
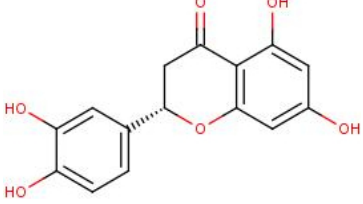
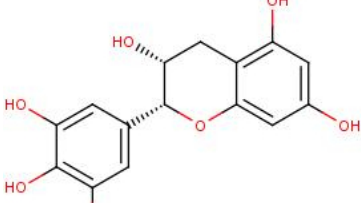
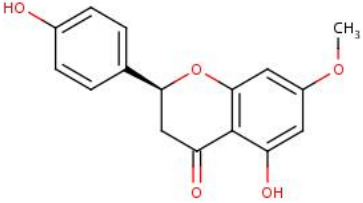
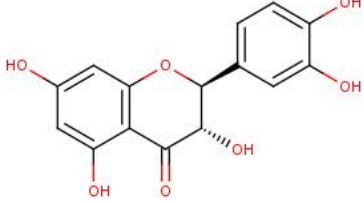
Figura 7. Patrón de interacción proteína ligando de complejos NQO2-quercetina durante la última etapa (décimo ns) de tres simulaciones de dinámica molecular independientes.

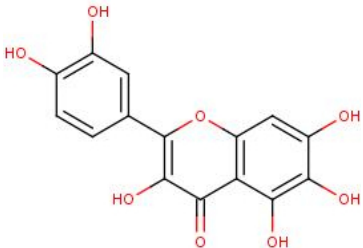
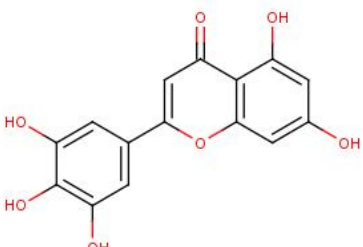
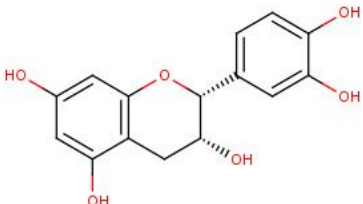
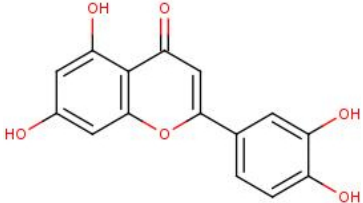
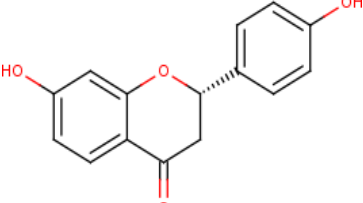
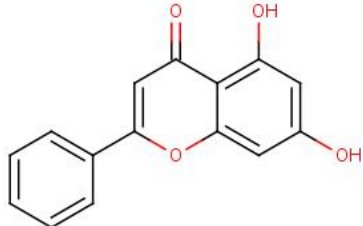
Anexo 1

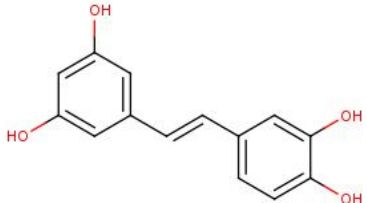
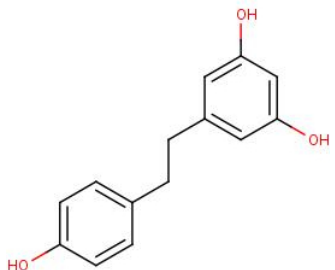
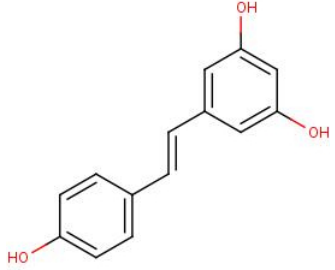
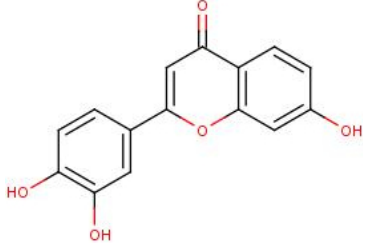
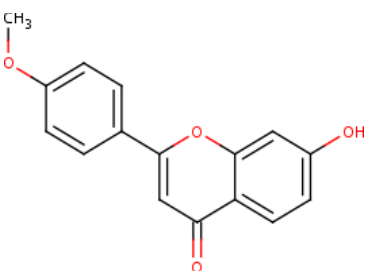
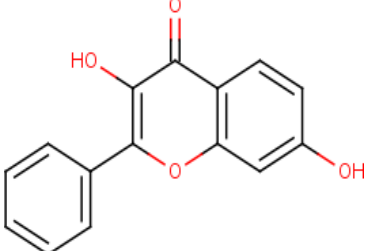
Tablas suplementarias

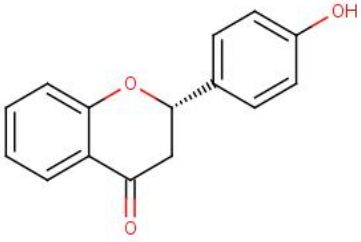
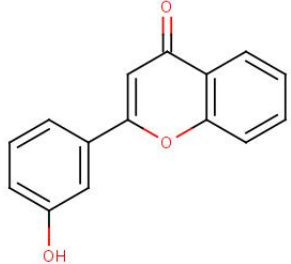
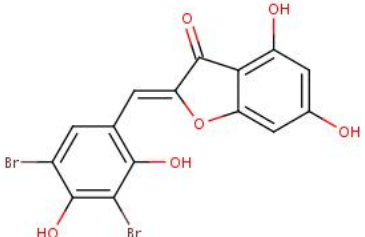
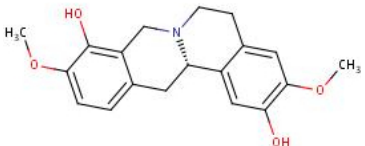
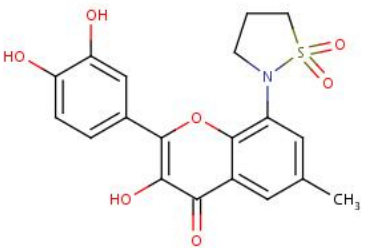
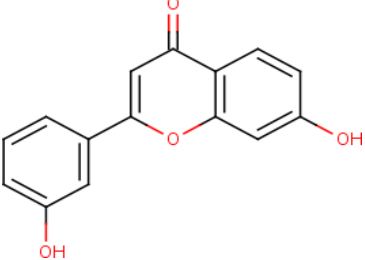
Tabla 1. Librería de ligandos similares a quercetina identificados por el algoritmo *SHAFTS* (*SHApe-Feature Similarity*) en la primera etapa del proceso de tamizaje reverso. Los ligandos son representados de acuerdo a su identificador del *Protein Data Bank (PDB)*, nombre, *score* de similitud contra quercetina, estructura bidimensional y presencia en proteínas de mamífero.

PDB id	Name	Similarity score	Mammalian protein	2D Representation
FSE	Fisetin (<i>3,7,3',4'-tetrahydroxyflavone</i>)	1,832	Yes	
KXN	(<i>2R,3S</i>)-2-(<i>3,4</i> -dihydroxyphenyl)-3,4-dihydro-2H-chromene-3,5,7-triol	1,785	-	
DQH	(<i>2r,3r</i>)-trans-dihydroquercetin	1,767	-	
MYC	Myricetin (<i>3,3',4',5,5',7-hexahydroxyflavone</i>)	1,744	Yes	
QUE	Quercetin (<i>3,5,7,3',4'-pentahydroxyflavone</i>)	1,744	Yes	

AGI	Apigenin (5,7-dihydroxy-2-(4-hydroxyphenyl)-4H-chromen-4-one)	1,727	Yes	
KMP	Kaempferol (3,5,7-trihydroxy-2-(4-hydroxyphenyl)-4H-chromen-4-one)	1,727	Yes	
NAR	Naringenin ((2S)-5,7-dihydroxy-2-(4-hydroxyphenyl)-2,3-dihydro-4H-chromen-4-one)	1,712	Yes	
ERD	(2s)-2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-2,3-dihydro-4H-chromen-4-one	1,699	-	
EGT	Epigallocatechin (2-(3,4,5-trihydroxy-phenyl)-chroman-3,5,7-triol)	1,695	-	
SAK	Sakuranetin ((2S)-5-hydroxy-2-(4-hydroxyphenyl)-7-methoxy-2,3-dihydro-4H-chromen-4-one)	1,671	-	
DH2	(2S,3S)-2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydro-4H-chromen-4-one	1,669	-	

MYU	Quercetagetin (<i>2-(3,4-dihydroxyphenyl)-4h-chromene-3,5,6,7-tetraol</i>)	1,660	Yes	
MYF	<i>Pentahydroxyflavone</i>	1,646	Yes	
28E	Epicatechin (<i>((2R,3R)-2-(3,4-dihydroxyphenyl)-3,4-dihydro-2H-chromene-3,5,7-triol)</i>)	1,645	-	
LU2	Luteolin (<i>3',4',5,7-tetrahydroxyflavone</i>)	1,644	Yes	
DFV	<i>5-deoxyflavanone</i>	1,643	-	
57D	Chrysin (<i>5,7-dihydroxy-2-phenyl-4H-chromen-4-one</i>)	1,617	Yes	

PIT	Piceatannol (4-[(E)-2-(3,5-dihydroxyphenyl)ethenyl]benzene-1,2-diol)	1,593	Yes	
RE2	Dihydroresveratrol (5-[2-(4-hydroxyphenyl)ethyl]benzene-1,3-diol)	1,584	Yes	
STL	Resveratrol (5-[(E)-2-(4-hydroxyphenyl)ethenyl]benzene-1,3-diol)	1,584	Yes	
15Z	2-(3,4-dihydroxyphenyl)-7-hydroxy-4H-chromen-4-one	1,569	Yes	
15W	7-hydroxy-2-(4-methoxyphenyl)-4H-chromen-4-one	1,542	Yes	
HHF	3,7-dihydroxy-2-phenyl-4H-chromen-4-one	1,537	-	

DFL	4'-hydroxyflavanone	1,526	Yes	
20B	2-(3-hydroxyphenyl)-4H-chromen-4-one	1,524	Yes	
FL9	3',5'-dibromo-2',4,4',6'-tetrahydroxy aurone	1,523	Yes	
SLX	(S)-scoulerine ((13a <i>S</i>)-3,10-dimethoxy-5,8,13,13a-tetrahydro-6 <i>H</i> -isoquino[3,2- <i>a</i>]isoquinoline-2,9-diol)	1,518	-	
371	2-(3,4-dihydroxyphenyl)-8-(1,1-dioxidoisothiazolidin-2-yl)-3-hydroxy-6-methyl-4 <i>H</i> -chromen-4-one	1,515	Yes	
F94	7-hydroxy-2-(3-hydroxyphenyl)-4 <i>H</i> -chromen-4-one	1,512	Yes	

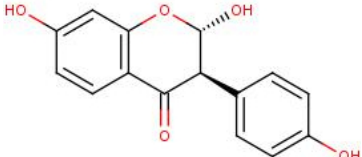
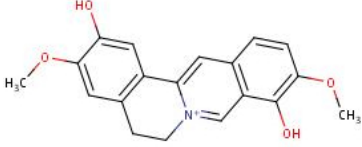
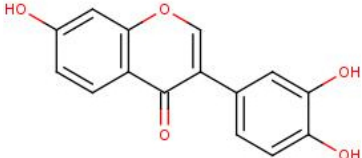
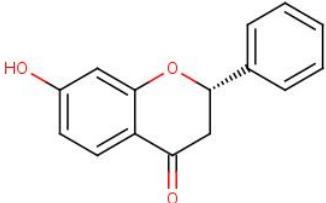
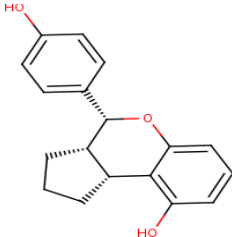
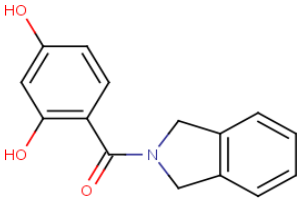
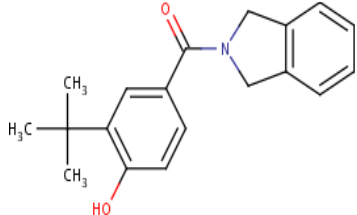
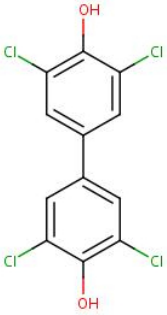
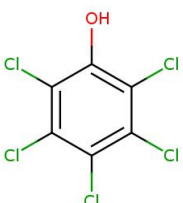
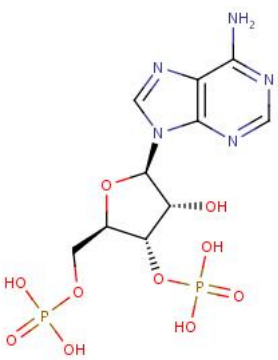
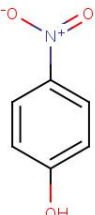
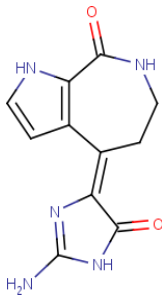
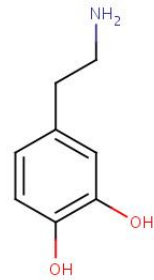
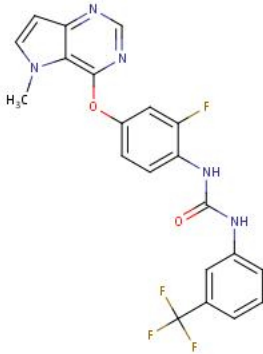
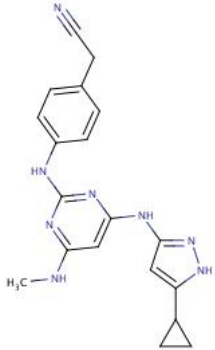
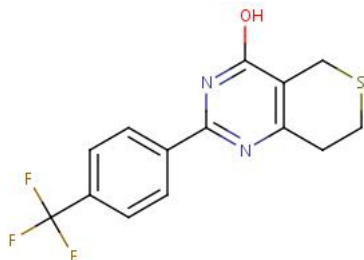
2HI	2,7,4'-trihydroxyisoflavanone	1,509	-	
DEH	2,9-dihydroxy-3,10-dimethoxy-5,6-dihydroisoquino[3,2-a]isoquinolinium	1,503	-	
47X	3-(3,4-dihydroxyphenyl)-7-hydroxy-4H-chromen-4-one	1,502	Yes	
DDC	7-hydroxy-2-phenyl-chroman-4-one	1,501	-	
(WST)	((3aS,4R,9bR)-4-(4-hydroxyphenyl)-1,2,3,3a,4,9b-hexahydrocyclopenta[c]chromen-9-ol)	(1,425)	Yes	
(PFT)	(1,3-dihydroisoindol-2-yl)-(2,4-dihydroxyphenyl)methanone	(1,379)	Yes	
(T5M)	(3-tert-butyl-4-hydroxy-phenyl)-(1,3-dihydroisoindol-2-yl)methanone	(1,376)	Yes	

Tabla 2. Ligandos co-cristalizados de proteínas blanco identificadas en el *Protein Data Bank (PDB)* por tamizaje virtual reverso basado en comparación de sitios de unión con el algoritmo *LIBRA (Ligand Binding site Recognition Application)*. Los ligandos son representados de acuerdo de su identificador del *PDB*, nombre, *score* de similitud contra quercetina si se encuentra presenta y estructura bi-dimensional.

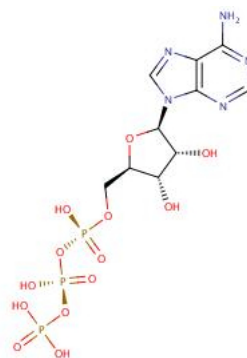
PDB id	Name	Similarity score	2D Representation
PCQ	<i>3,5,3',5'-tetrachloro-biphenyl-4,4'-diol</i>	-	
PCI	<i>Pentachlorophenol</i>	-	
A3P	<i>Adenosine-3'-5'-diphosphate</i>	-	
NPO	<i>p-nitrophenol</i>	-	

DBQ	Debromohymenialdisine ((4Z)-4-(2-amino-5-oxo-1,5-dihydro-4H-imidazol-4-ylidene)-4,5,6,7-tetrahydropyrrolo[2,3-c]azepin-8(1H)-one)	-	
LDP	L-Dopamine (4-(2-aminoethyl)benzene-1,2-diol)	1,193	
42Q	1-{2-fluoro-4-[(5-methyl-5H-pyrrolo[3,2-d]pyrimidin-4-yl)oxy]phenyl}-3-[3-(trifluoromethyl)phenyl]urea	-	
5CP	[4-({4-[(5-cyclopropyl-1H-pyrazol-3-yl)amino]-6-(methylamino)pyrimidin-2-yl}amino)phenyl]acetonitrile	-	
XAV	2-[4-(trifluoromethyl)phenyl]-7,8-dihydro-5H-thiopyrano[4,3-d]pyrimidin-4-ol	1,270	

EHB	<i>(1Z)</i> -1-(3-ethyl-5-hydroxy-1,3-benzothiazol-2(3H)-ylidene)propan-2-one	1,212	
HRM	7-methoxy-1-methyl-9h-beta-carboline	1,200	
3RA	5-(1,3-benzodioxol-5-ylmethyl)-2-(phenylamino)-4H-imidazol-4-one	-	
AR6	Adenosine-5-diphosphoribose	-	
GJW	7-(pyrrolidin-1-ylmethyl)pyrrolo[1,2-a]quinoxalin-4(5H)-one	-	
MEW	2-methylquinazolin-4(3H)-one	-	
1J5	<i>trans</i> -4-{2-[(1R)-1-hydroxyethyl]imidazo[4,5-d]pyrrolo[2,3-b]pyridin-1(6H)-yl}cyclohexanecarbonitrile	-	

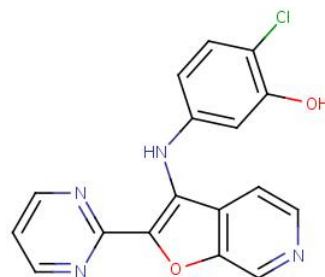
ATP Adenosine-5'-triphosphate

-



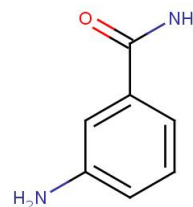
FP3 2-chloro-5-{[2-(pyrimidin-2-yl)furo[2,3-c]pyridin-3-yl]amino}phenol

1,273



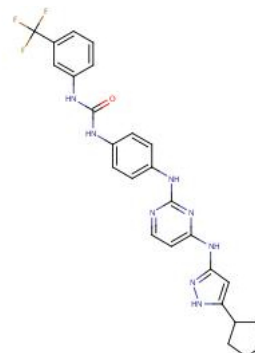
3AB 3-aminobenzamide

-



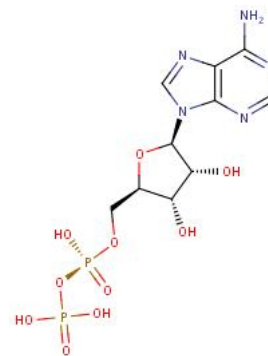
CJ5 1-[4-[[4-[(5-cyclopentyl-1H-pyrazol-3-yl)amino]pyrimidin-2-yl]amino]phenyl]-3-[3-(trifluoromethyl)phenyl]urea

-

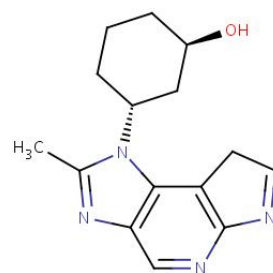


ADP Adenosine-5'-diphosphate

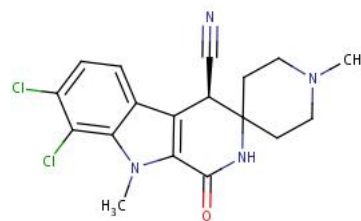
-



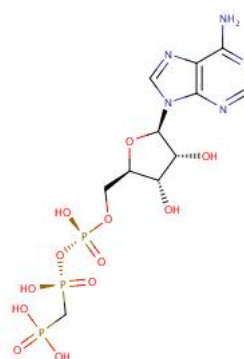
0Q2 (1R,3R)-3-(2-methylimidazo[4,5-d]pyrrolo[2,3-b]pyridin-1(8H)-yl)cyclohexanol



7CP (4R)-7,8-dichloro-1',9-dimethyl-1-oxo-1,2,4,9-tetrahydrospiro[beta-carboline-3,4'-piperidine]-4-carbonitrile

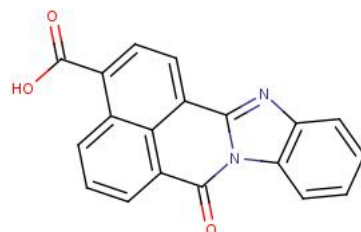


ACP Adenosine-5'-[beta, gamma-methylene]triphosphate

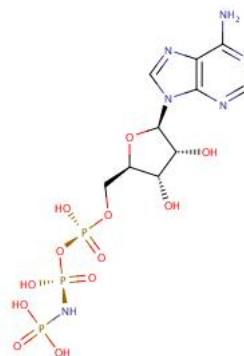


609 7-oxo-7H-benzimidazo[2,1-a]benz[de]isoquinoline-3-carboxylic acid

1,168

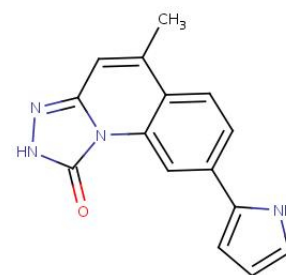


ANP Phosphoaminophosphonic acid-adenylate ester



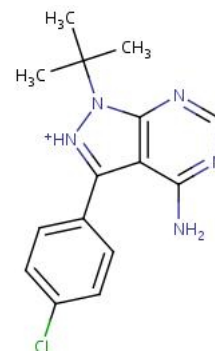
YEX 5-methyl-8-(1*h*-pyrrol-2-yl)
[1,2,4]triazolo[4,3-*a*]quinolin-1(2*h*)-one

-



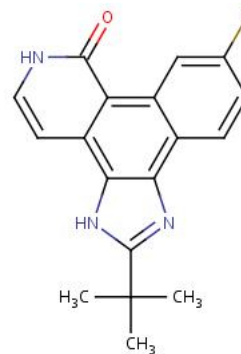
PP2 1-*tert*-butyl-3-(4-chloro-phenyl)-1*h*-
pyrazolo[3,4-*d*]pyrimidin-4-ylamine

1,232



IZA 2-*tert*-butyl-9-fluoro-3,6-dihydro-7*h*-
benz[*h*]-imidaz[4,5-*f*]isoquinoline-7-one

1.106



AMP Adenosine monophosphate

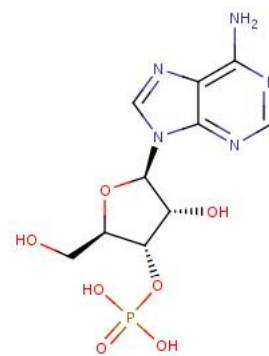
-



3AM

3'-adenylic acid

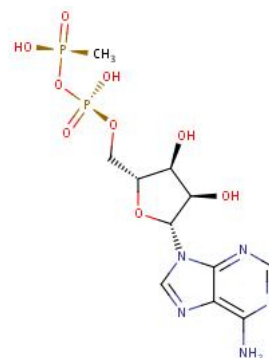
-



M33

5'-O-[(S)-hydroxy{[(S)-hydroxy(methyl)phosphoryl]oxy}phosphoryl]adenosine

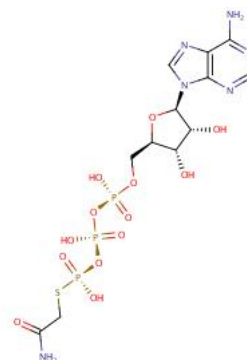
-



112

Thiophosphoric acid o-((adenosylphospho)phospho)-s-acetamidyl-diester

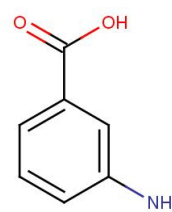
-



GAB

Gabaculine (3-aminobenzoic acid)

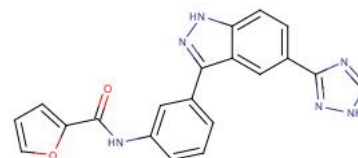
-



35F

N-{3-[5-(1H-1,2,4-triazol-3-yl)-1H-indazol-3-yl]phenyl}furan-2-carboxamide

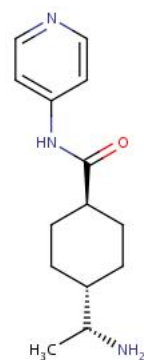
1,083



Y27

(r)-trans-4-(1-aminoethyl)-n-(4-pyridyl) cyclohexanecarboxamide

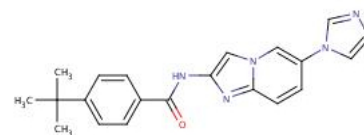
-



IM6

4-tert-butyl-N-[6-(1H-imidazol-1-yl)imidazo[1,2-a]pyridin-2-yl]benzamide

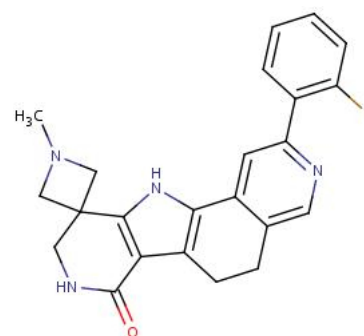
-



L8I

2'-(2-fluorophenyl)-1-methyl-6',8',9',11'-tetrahydrospiro[azetidine-3,10'-pyrido[3',4':4,5]pyrrolo[2,3-f]isoquinolin]-7'(5H)-one

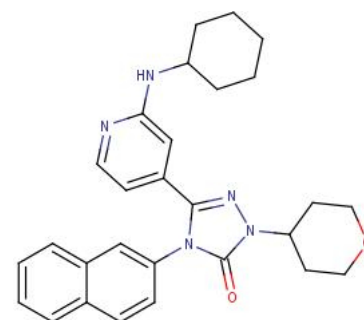
-



589

5-[2-(cyclohexylamino)pyridin-4-yl]-4-naphthalen-2-yl-2-(tetrahydro-2H-pyran-4-yl)-2,4-dihydro-3H-1,2,4-triazol-3-one

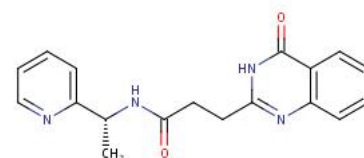
-



5ME

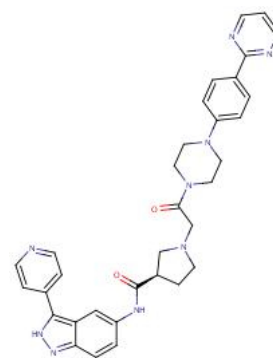
3-(4-oxo-3,4-dihydroquinazolin-2-yl)-N-[(1R)-1-(pyridin-2-yl)ethyl]propanamide

-



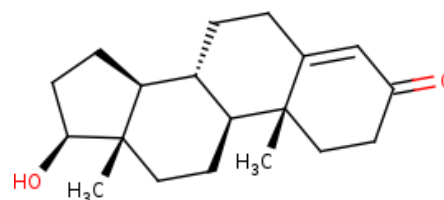
38Z

(3R)-1-(2-oxo-2-(4-[4-(pyrimidin-2-yl)phenyl]piperazin-1-yl)ethyl)-N-[3-(pyridin-4-yl)-2H-indazol-5-yl]pyrrolidine-3-carboxamide



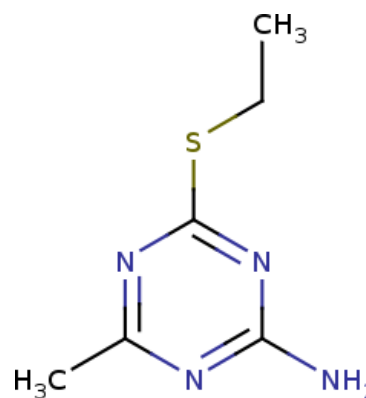
TES

Testosterone



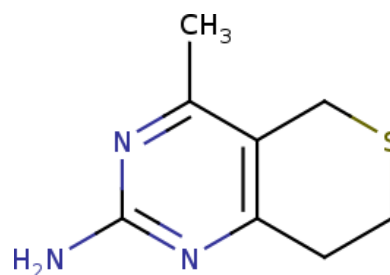
B2J

4-(ethylsulfanyl)-6-methyl-1,3,5-triazin-2-amine



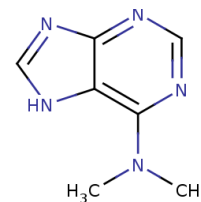
MO8

4-methyl-7,8-dihydro-5H-thiopyrano[3,4-e]pyrimidin-2-amine



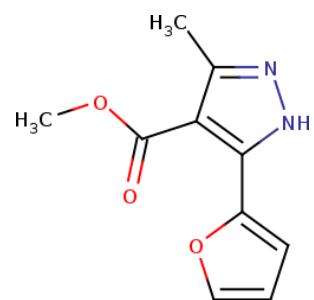
42C

N,N-dimethyl-7H-purin-6-amine



37D methyl 3-(furan-2-yl)-5-methyl-2H-pyrazole-4-carboxylate

1,11



RDE 2-(3-amino-2,5,6-trimethoxy-phenyl)ethyl 5-chloro-2,4-dihydroxy-benzoate

-

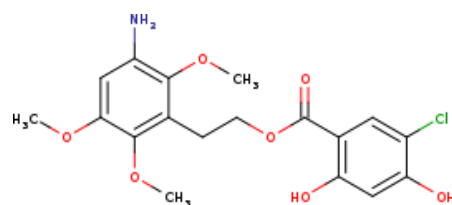


Tabla 3. Proteínas blanco identificadas en el *Protein Data Bank (PDB)* por tamizaje virtual basado en similitud del ligando con el algoritmo *SHAFTS (SHApe Feature Similarity)* y *docking* inverso con el servidor *idTarget (http://idtarget.rcas.sinica.edu.tw/)*. Las proteínas candidato fueron segregadas según su organismo de origen y anotadas con su correspondiente entrada en el PDB, ligando co-cristalizado, y *score* de similitud estructural con quercetina así como nombre de la proteína y *score* de *idTarget* (*Z-score* y afinidad predicha). Las proteínas co-cristalizadas con quercetina se muestran remarcadas. *Los *scores* para los cristales 4hlg y 4hnh resultan de un anclaje en un sitio de unión alternativo..

PDB id	Ligand	Similarity	Name	Z-score	ΔG (kcal/mol)	Predicted affinity (nM)
Homo sapiens						
1xo2	FSE (Fisetin)	1,832	Cyclin dependent kinase 6	-0,09	-10,56	18,1
4hl5	15W	1,542		-0,10	-10,58	17,5
4hlf	15Z	1,569		0,00	-10,43	22,6
4hlg*	20B	1,524		1,99	-7,57	2800
4hnh*	F94	1,512	Tankyrase-2	1,78	-7,88	1700
4kzq	DFL	1,526		-0,17	-10,67	15,1
4hkk	AGI (Apigenin)	1,727		0,11	-10,27	29,6
4hkn	LU2 (Luteolin)	1,644		0,16	-10,20	33,3
3ckl	STL (Resveratrol)	1,584	Sulfotransferase family cytosolic 1B member 1	0,08	-10,31	27,7
1sg0	STL (Resveratrol)	1,584	NRH dehydrogenase [quinone] 2	0,24	-10,09	40,1
3ftu	RE2 (Dihydroresveratrol)	1,584		0,31	-9,99	47,5
3ftx	RE2 (Dihydroresveratrol)	1,584	Leukotriene A-4 hydrolase	0,36	-9,91	54,3
3fts	STL (Resveratrol)	1,584		0,13	-10,25	30,6
3amy	AGI (Apigenin)	1,727	Casein kinase 2 subunit alpha	0,43	-9,81	64,3
2ydx	STL (Resveratrol)	1,584	Methionine adenosyltransferase 2 subunit beta	0,45	-9,79	66,5
(2pog)	(WST)	(1,425)	(Receptor de estrógenos alfa)	(0,57)	(-9,62)	(88,7)
3lm5	QUE (Quercetin)	1,744	Serine/threonine-protein kinase 17B	0,59	-9,58	94,9
2duv	371	1,515	Cell division protein kinase 2	0,61	-9,55	99,8

4eh3	NAR (Naringenin)	1,744	Mitogen-activated protein kinase 14	0,66	-9,49	110,4	
3v3v	MYU (Quercetagenin)	1,660	Mitogen-activated protein kinase 8	0,69	-9,44	120,1	
2hck	QUE (Quercetin)	1,744	Hematopoietic cell kinase HCK	0,71	-9,41	126,4	
3sz1	LU2 (Luteolin)	1,644	Peroxisome proliferator-activated receptor gamma	1,16	-8,77	372,3	
4jaz	STL (Resveratrol)	1,584		0,81	-9,27	160,1	
2o65	MYF	1,646	Serine/threonine-protein kinase PIM-1	0,89	-9,16	192,7	
2o63	MYC (Myricetin)	1,744		0,93	-9,10	213,3	
2o64	MYU (Quercetagenin)	1,660		0,86	-9,19	183,2	
2o3p	QUE (Quercetin)	1,744		0,88	-9,17	189,5	
4lmu	QUE (Quercetin)	1,744	3-Hydroxyisobutyryl-CoA hydrolase	0,91	-9,12	206,2	
3bpt	QUE (Quercetin)	1,744		-1,03	-9,02	244,1	
3l5r	47X	1,502	Macrophage migration inhibitory factor	1,00	-9,00	252,5	
2l98	STL (Resveratrol)	1,584	Troponin C, slow skeletal and cardiac muscles	1,01	-8,98	261,2	
(3k99)	(PFT)	(1,379)	(Proteína de shock térmico de 90 kDa)	(1,01)	(-8,98)	(261,2)	
(2xhx)	(T5M)	(1,376)		(1,25)	(-8,63)	(471,5)	
4gqr	MYC (Myricetin)	1,744	Pancreatic alpha-amylase	1,22	-8,68	433,4	
1thc	FL9	1,523	Transthyretin	1,40	-8,42	672,1	
4des	57D (Chrysin)	1,617		1,31	-8,55	539,7	
4der	AGI (Apigenin)	1,727		1,83	-7,80	1900	
4det	KMP (Kaempferol)	1,727		1,60	-8,14	1100	
4dew	LU2 (Luteolin)	1,644		2,10	-7,42	3600	
4deu	NAR (Naringenin)	1,744		1,44	-8,36	743,8	
1dvs	STL (Resveratrol)	1,584		1,96	-7,61	2600	
(4mp7)	(PFT)	(1,379)		(Piruvato deshidrogenasa quinasa)	(1,46)	(-8,34)	(769,3)
4hd8	PIT (Piceatannol)	1,593		NAD-dependent protein deacetylase Sirtuin-3	1,60	-8,13	1100
Oryctolagus cuniculus							
3ebo	57D (Chrysin)	1,617	Glycogen phosphorylase, muscle form	0,73	-9,39	130,7	
Sus scrofa							
1e90	MYC (Myricetin)	1,744	Phosphatidylinositol 3-kinase catalytic subunit	0,45	-9,78	67,7	
1e8w	QUE (Quercetin)	1,744		0,45	-9,78	67,7	
Bos taurus							

3nvy	QUE (Quercetin)	1,744	Xanthine dehydrogenase/oxidase	0,73	-9,38	133,7
2jj2	QUE (Quercetin)	1,744		1,44	-8,36	743,8
2jj1	PIT (Piceatannol)	1,593	ATP synthase	0,00	-8,96	270,1
2jiz	STL (Resveratrol)	1,584		1,47	-8,32	795,7

Tabla 4. Proteínas blanco identificadas en el *Protein Data Bank (PDB)* por tamizaje reverso basado comparación del sitio de unión con el algoritmo *LIBRA (Ligand Binding Site Recognition Application)* y *docking* inverso con el servidor *idTarget (http://idtarget.rcas.sinica.edu.tw/)*. Las proteínas candidato fueron segregadas de acuerdo a su organismo de origen y anotadas con su entrada correspondiente en el *PDB*, ligando co-cristalizado y su similitud estructural con quercetina si se dispone, así como nombre de la proteína y *score* de *idTarget (Z-score* y afinidad predicha). Proteínas co-cristalizadas con quercetina se presentan remarcadas.

PDB id	Ligand	Similarity	Name	Z-score	ΔG (kcal/mol)	Predicted affinity (nM)
Homo sapiens						
1g3m	PCQ	-	Estrogen sulfotransferase	-0,05	-10,51	19,7
2gwh	PCI (Pentachlorophenol)	-	Sulfotransferase 1C4	-0,01	-10,44	22,2
3bfx	A3P	-	Sulfotransferase 1C2	1,48	-10,42	23,0
1ls6	NPO (p-Nitrophenol)	-	Sulfotransferase 1A1	0,04	-10,38	24,6
1z57	DBQ (Debromohymenialdi sine)	-	Dual specificity protein kinase CLK1	0,13	-10,24	31,1
2a3r	LDP (L-Dopamine)	1,193	Sulfotransferase 1A3	0,18	-10,18	34,5
3vhe	42Q	-	Vascular endothelial growth factor receptor 2	0,22	-10,12	38,1
3bhh	5CP	-	Calcium/calmodulin- dependent protein kinase type 2 beta	-0,93	-10,09	40,1
3uh4	XAV	1,270	Tankyrase-1	0,24	-10,09	40,1
2reo	A3P (Adenosine-3'- 5'-diphosphate)	-	Sulfotransferase 1C3	1,61	-10,05	42,9
3anq	EHB	1,212	Dual specificity tyrosine- phosphorylation-regulated kinase 1A	0,45	-9,78	67,7
3anr	HRM	1,200		0,34	-9,95	50,8
4aze	3RA	-		0,45	-9,79	66,5
3v2b	AR6 (Adenosine-5- diphosphoribose)	-	Poly [ADP-ribose] polymerase 15	0,34	-9,94	51,7
4abk	AR6 (Adenosine-5- diphosphoribose)	-	Poly [ADP-ribose] polymerase 14	0,39	-9,87	58,1
3gjw	GJW	-	Poly [ADP-ribose] polymerase 1	0,61	-9,55	99,8
4gv7	MEW	-		0,42	-9,83	62,2
4iva	1J5	-	Tyrosine-protein kinase JAK2	0,43	-9,82	63,3
4fvq	ATP (Adenosine-5'- triphosphate)	-		0,79	-9,30	152,2

3prf	FP3	1,273	Serine/threonine-protein kinase B-raf	0,45	-9,79	66,5
2am9	TES (Testosteron)	-	Androgen receptor	0,48	-9,74	72,4
3kcz	3AB (3-aminobenzamide)	-	Poly [ADP-ribose] polymerase 2	0,49	-9,73	73,6
4j8m	CJ5	-		1,39	-8,44	649,8
1ol6	ATP (Adenosine-5'-triphosphate)	-	Aurora kinase A	0,52	-9,69	78,8
1mq4	ADP (Adenosine-5'-diphosphate)	-		0,54	-9,66	82,9
4ei4	0Q2	-	Tyrosine-protein kinase JAK1	0,52	-9,68	80,1
3bhy	7CP	-	Death-associated protein kinase 3	1,3	-9,66	82,9
3gqi	ACP	-	Fibroblast growth factor receptor 1	0,56	-9,63	87,2
2zv2	609	1,168	Calcium/calmodulin-dependent protein kinase kinase 2	0,57	-9,62	88,7
1j1b	ANP	-		0,84	-9,23	171,3
4ekk	ANP	-	Glycogen synthase kinase-3 beta	0,57	-9,61	90,2
1o6l	ANP	-		0,58	-9,60	91,7
1j1b	ANP	-		0,92	-9,11	209,7
2yex	YEX	-	Serine/threonine-protein kinase Chk1	0,59	-9,58	94,9
3sxs	PP2	1,232	Cytoplasmic tyrosine-protein kinase BMX	0,64	-9,51	106,8
3lxl	IZA	1.106	Tyrosine-protein kinase JAK3	0,65	-9,50	108,6
2src	ANP	-	Proto-oncogene tyrosine-protein kinase Src	0,73	-9,38	133,0
4imy	AMP (Adenosine monophosphate)	-	Cell division protein kinase 9	0,79	-9,30	152,2
3blq	ATP (Adenosine-5'-triphosphate)	-		0,87	-9,18	186,3
3c0i	3AM	-	Peripheral plasma membrane protein CASK	0,8	-9,28	157,4
3a7h	ATP (Adenosine-5'-triphosphate)	-	Serine/threonine-protein kinase 24	0,84	-9,23	171,3
3a7j	ADP (Adenosine-5'-diphosphate)	-		0,96	-9,06	228,2
3b24	B2J	-		1,10	-8,85	325,3
3ft5	MO8	-		1,28	-8,60	496,0
3hyz	42C	-	Heat shock protein hsp 90-alpha	0,84	-9,23	171,3
3hz1	37D	1,111		1,15	-8,78	366,1
4egi	B2J	-		1,09	-8,86	319,8
3b2t	M33	-	Fibroblast growth factor	0,84	-9,22	174,2

			receptor 2			
4rrv	ATP (Adenosine-5'-triphosphate)	-	Phosphoinositide-dependent kinase-1	0,97	-9,04	236,0
4aw0	ATP (Adenosine-5'-triphosphate)	-	Tyrosine-protein kinase ABL1	0,86	-9,19	183,2
2g1t	112	-	Tyrosine-protein kinase LCK	0,88	-9,17	189,5
1qpc	ANP	-	Tyrosine-protein kinase ITK/TSK	0,88	-9,17	189,5
4m15	ADP (Adenosine-5'-diphosphate)	-	Poly [ADP-ribose] polymerase 12	0,91	-9,12	206,2
2pqf	GAB (Gabaculine)	-	Mitogen-activated protein kinase 9	0,94	-9,08	220,6
3e7o	35F	1,083	Poly [ADP-ribose] polymerase 16	0,41	-9,06	228,2
4f0d	3AB	-	Ephrin type-A receptor 3	0,96	-9,05	232,1
2qoc	ANP	-	Rho-associated protein kinase 1	0,98	-9,02	244,1
2etr	Y27	-	Mitogen-activated protein kinase kinase 5	1,02	-8,96	270,1
3vw6	IM6	-	MAP kinase-activated protein kinase 2	1,02	-8,96	270,1
3m2w	L8I	-	Mitogen-activated protein kinase 10	1,1	-8,85	325,3
1ny3	ADP (Adenosine-5'-diphosphate)	-	Cell division cycle 7-related protein kinase	1,14	-8,79	359,9
3oy1	589	-	Poly [ADP-ribose] polymerase 3	1,11	-8,84	330,8
4f9a	ADP (Adenosine-5'-diphosphate)	-	Poly [ADP-ribose] polymerase 10	1,16	-8,76	378,6
4gv2	5ME	-		1,46	-8,33	782,4
3hkv	3AB (3-Aminobenzamide)	-		2,34	-7,07	6600
Mus musculus						
1jpa	ANP	-	Ephrin type-B receptor 2	1,23	-8,66	448,3
Rattus norvegicus						
4qta	38Z	-	Mitogen-activated protein kinase 1	0,75	-9,35	139,9
4gt3	ATP (Adenosine-5'-triphosphate)	-		0,82	-9,26	162,8
Canis lupus familiaris						
2fyp	RDE	-	Endoplasmin	1,25	-8,63	471,5

Bos taurus

2h9v	Y27	-	Rho-associated protein kinase 2	0,95	-9,07	224,4
------	-----	---	------------------------------------	------	-------	-------

Tabla 5. Proteínas identificadas por análisis automático de textos como asociadas a procesos de muerte neuronal, neuroprotección y efectos de quercetina. Se remarca en sombreado el tipo de asociación. * Agregada por homología. ** Agregadas por asociación al correspondiente factor de crecimiento.

Uniprot ID	Name	Neuronal survival	Neuronal death	Quercetin
Q96EB6	NAD-dependent protein deacetylase sirtuin-1			
P15559	NADPH:quinone oxidoreductase			
Q8IXJ6	NAD-dependent protein deacetylase sirtuin-2			
P15692	Vascular endothelial growth factor A			
P09038	Fibroblast growth factor 2			
P05019	Insulin-like growth factor I (IGF-I)			
Q9NTG7	NAD-dependent protein deacetylase sirtuin-3, mitochondrial*			
P37231	Peroxisome proliferator-activated receptor gamma			
Q16539	Mitogen-activated protein kinase 14 (p38)			
P45983	Mitogen-activated protein kinase 8 (JNK1)			
P48736	Phosphatidylinositol 3-kinase catalytic subunit			
P25705	ATP synthase, mitochondrial			
P47989	Xanthine dehydrogenase/oxidase			
Q6NVY1	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial			
P08631	Tyrosine-protein kinase HCK			
P11362	Fibroblast growth factor receptor 1**			
P21802	Fibroblast growth factor receptor 2**			
O15530	Phosphoinositide-dependent kinase-1			
P09874	Poly [ADP-ribose] polymerase 1			
P28482	Mitogen-activated protein kinase 1 (Extracellular signal-regulated kinase 2, ERK2)			
P49841	Glycogen synthase kinase-3 beta (GSK-3 beta)			
P53779	Mitogen-activated protein kinase 10 (JNK3)			
Q99683	Apoptosis signal regulating kinase 1; MAP/ERK kinase kinase 5; ASK1			
P00519	Tyrosine-protein kinase ABL1			
P35968	Vascular endothelial growth factor receptor 2**			
P12931	Proto-oncogene tyrosine-protein kinase Src			
P08069	Insulin-like growth factor 1 receptor **			
P11511	Cytochrome P450 19A1			
P08183	Multidrug resistance protein 1			
P04798	Cytochrome P450 1A1			
P10632	Cytochrome P450 2C8			
P15121	Aldehyde reductase			
Q9UNQ0	ATP-binding cassette sub-family G member 2			
P10415	Apoptosis regulator Bcl-2			
P42574	Caspase-3			
P09601	Heme oxygenase 1 (HO-1)			
P05231	Interleukin-6			
P01375	Tumor necrosis factor (TNF-alpha)			
Q07817	Bcl-2-like protein 1 (Bcl2-L-1) (Apoptosis regulator Bcl-X)			
P19838	Nuclear factor NF-kappa-B			

P31749	Protein kinase B (Proto-oncogene c-Akt)
P05412	Proto-oncogene c-Jun (Activator protein 1)
P99999	Cytochrome c
P34932	Heat shock 70 kDa protein 4
P01584	Interleukin-1 beta
P04179	Superoxide dismutase [Mn], mitochondrial
P35228	Nitric oxide synthase, inducible
P00441	Superoxide dismutase 1 [Cu-Zn]
P29475	Nitric oxide synthase, brain (Constitutive NOS)
P35354	Prostaglandin G/H synthase 2 (Cyclooxygenase-2) (COX-2)
P04040	Catalase
Q07812	Apoptosis regulator BAX (Bcl-2-like protein 4) (Bcl2-L-4)
P49768	Presenilin-1 (PS-1)
O00206	Toll-like receptor 4 (hToll)
P03372	Estrogen receptor (ER) (ER-alpha)
Q16236	Nuclear factor erythroid 2-related factor 2
P05771	Protein kinase C, beta
Q16665	Hypoxia-inducible factor 1-alpha (HIF-1-alpha)
P00533	Epidermal growth factor receptor
Q92731	Estrogen receptor beta
P09429	High mobility group protein B1, Amphoterin
P40763	Signal transducer and activator of transcription 3
Q13131	Protein kinase A; cAMP dependent protein kinase
P42345	Serine/threonine-protein kinase mTOR
P04637	Tumor protein p53
P55211	Caspase-9
P24385, Q9H014	Cyclin d1, B-cell lymphoma 1 protein (BCL-1)

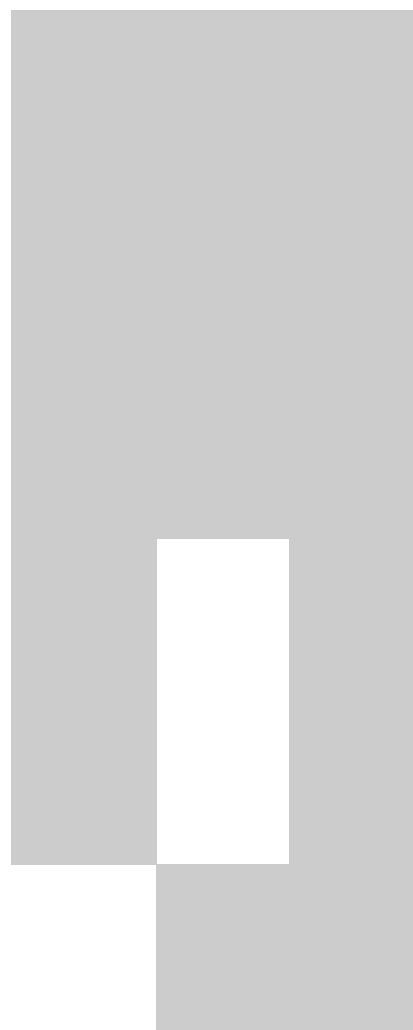


Tabla 6. Energía potencial de interacción de complejos quercetina-proteína blanco seleccionados. “Uab” and “Protein” representan la energía de interacción de quercetina con el entorno incluyendo y excluyendo las moléculas del solvente respectivamente. El valor promedio de la última etapa de simulación es comparado con los *scores* de *docking* de *MOE* e *idTarget*. Los complejos fueron anotados el valor de desviación geométrica (*RMSD*), correspondiente a la estructura del último pasod e simulación contra la estructura inicial (estructura energéticamente minimizada).

Uniprot id	Name	Docking Score (kcal)		Potential Interaction Energy (Uab) (kcal)	Potential Interaction Energy (Protein) (kcal)	RMSD (Å)
		MOE	idTarget			
P16083	NRH dehydrogenase [quinone] 2	-8,76	-10,09	-100,4	-63,9	0,985
				-108,1	-68,6	0,819
				-109,7	-62,7	0,792
P09874	Poly [ADP-ribose] polymerase 1	-7,45	-9,83	-108,0	-82,8	0,884
				-113,0	-85,2	0,954
				-120,5	-80,7	1,021
P49841	Glycogen synthase kinase-3 beta	-7,27	-9,61	-109,0	-65,9	0,908
				-91,0	-50,1	0,940
				-95,1	-56,3	0,855
Q16539	Mitogen-activated protein kinase 14	-7,27	-9,49	-115,2	-85,2	0,721
				-110,1	-68,6	0,729
				-111,6	-76,8	0,701