

Búsqueda de fármacos para la MCADD a través del uso de la informática

Leyre Aramendía

Máster Oficial de Bioinformática y Bioestadística
Bases Moleculares para el Desarrollo de Fármacos

Marta Enciso Carrasco

María Jesús Marco Galindo

05/06/2018



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

B) GNU Free Documentation License (GNU FDL)

Copyright © 2018-Leire Aramendía

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

C) Copyright

© Leire Aramendía

Reservados todos los derechos. Está prohibido la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la impresión, la reprografía, el microfilme, el tratamiento informático o cualquier otro sistema, así como la distribución de ejemplares mediante alquiler y préstamo, sin la autorización escrita del autor o de los límites que autorice la Ley de Propiedad Intelectual.

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Búsqueda de fármacos para la MCADD a través del uso de la informática</i>
Nombre del autor:	<i>Leyre Aramendía Yerro</i>
Nombre del consultor/a:	<i>Marta Enciso Carrasco</i>
Nombre del PRA:	<i>Maria Jesús Marco Galindo</i>
Fecha de entrega (mm/aaaa):	05/06/2018
Titulación:::	<i>Máster en Bioinformática y Bioestadística</i>
Área del Trabajo Final:	<i>Bases Moleculares para el Desarrollo de Fármacos</i>
Idioma del trabajo:	<i>Español</i>
Palabras clave	<i>MCAD, farmacóforo, virtual screening.</i>
Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.</i>	
<p>La MCADD, deficiencia en acyl-CoA deshidrogenasa de cadena media es una enfermedad metabólica rara de carácter hereditario, cuya variación patogénica mayoritaria es la MCADK304E. Este intercambio de una lisina por un ácido glutámico en la posición 304 introduce 4 cargas negativas anormales en el core de la proteína, haciendo que los enfermos con esta alteración posean tan solo una actividad enzimática del 40%, poniendo en compromiso el desarrollo de un metabolismo normal. En este proyecto se han analizado bioinformáticamente 70.000 moléculas con el fin de encontrar una serie de candidatos a fármacos efectivos frente a la enfermedad. En función de las características estéricas/electrostáticas de los posibles ligandos se elaboró una lista de candidatos potenciales (hits), seleccionada mediante virtual screening. Los 100 compuestos con mejor puntuación fueron estudiados en detalle, escogiendo a 20 de ellos (leads) para, mediante algoritmos de docking, analizar su interacción con MCADK304E. A su vez se evaluaron sus propiedades farmacocinéticas mediante los criterios ADME-Tox, reduciendo hasta 14 el número de candidatos seleccionados para adentrarse en las fases intermedias del desarrollo de fármacos con garantía.</p> <p>Un proyecto ambicioso que combina diferentes herramientas informáticas para abordar problemas biológicos, dará a conocer el atractivo del área de las enfermedades raras y la necesidad de encontrar fármacos eficaces frente a ellas.</p>	

Abstract (in English, 250 words or less):

MCADD, the deficiency in medium chain acyl-CoA dehydrogenase is a metabolic orphan disease of hereditary character, whose major pathogenic variation is MCADK304E. This exchange of a lysine by a glutamic acid in position 304 introduces 4 abnormal negative charges in the core of the protein, making patients with this alteration possess only an enzymatic activity of 40%, compromising the development of their metabolism. In this project, 70,000 molecules have been analyzed bioinformatically in order to find few candidates for being an effective drug against the disease. Based on the steric / electrostatic characteristics of the possible ligands, a list of potential candidates (hits) was drawn up, selected by virtual screening. The 100 compounds with the best score were studied in detail, choosing 20 of them (leads) to, thanks of docking algorithms, analyze their interaction with MCADK304E. In turn, their pharmacokinetic properties were evaluated using the ADME-Tox criteria, reducing the number of candidates selected to enter the intermediate stages of drug development with a guarantee up to 14. An ambitious project that combines different computer approaches for solve biological problems and that will reveal the attractiveness of the area of orphan diseases and the need to find effective drugs for the patients.

Índice

1. Introducción	
1.1 Contexto y justificación del Trabajo.....	1
1.2 Objetivos del Trabajo.....	1
1.3 Enfoque y método seguido.....	2
1.4 Planificación del Trabajo	3
1.5 Breve resumen de productos obtenidos.....	4
1.6 Breve descripción de los otros capítulos de la memoria.....	4
2. MCAD: Estructura de la proteína. Efectos de la mutación.....	5
3. Identificación del farmacóforo.....	6
4. Virtual screening	
4.1 Pharmacophore-based ligand virtual screening.....	8
4.1.1 Ligand shared features pharmacophore.....	8
4.1.2 Ligand merged features pharmacophore.....	8
4.2 Reducción de candidatos.....	9
4.3 Protein-based virtual screening.....	10
5. Evaluación de propiedades ADME-Tox.....	11
6. Conclusiones	
6.1 Impresiones sobre el trabajo.....	12
6.2 Seguimiento de la planificación.....	12
6.3 Resultados obtenidos.....	13
6.4 Líneas de trabajo futuras.....	13
7. Glosario.....	14
8. Anexos.....	15
9. Bibliografía.....	27

Listado de figuras

Tabla 1. Distribución de objetivos y tareas.

Imagen 1. Diagrama de Gantt.

Imagen 2. Estructura secundaria (a) y cuaternaria de MCAD (b), con cada uno de los bolsillos de reacción sombreados en amarillo. Dicha reacción queda detallada a la derecha de la proteína completa (c).

Imagen 3. Selección del aminoácido E304 dentro de la proteína mutada.

Imagen 4. Zoom del bolsillo al que pertenece el aminoácido 304 de MCAD (A) vs MCADK304E (B).

Imagen 5. Distribución de las propiedades de la cavidad donde se aloja FAD dentro de MCAD / MCADK304E.

Imagen 6. Contactos de FAD dentro de su cavidad en MCAD.

Imagen 7. Farmacóforo determinado a través de las características comunes de FAD y riboflavina.

Imagen 8. Resultado del shared-features pharmacophore based ligand screening.

Imagen 9. Resultado del merged-features pharmacophore based ligand screening.

Imagen 10. Lista de 20 leads que analizar a través del protein-based virtual screening.

Imagen 11. Zoom out a la izquierda (MCADK304E cadenas A y B) y zoom in de la molécula experimental de docking (MCADK304E) con el ligando problema riboflavina (ROSA) superpuesta con la molécula original de MCADK304E que contiene el ligando (Riboflavina, VERDE) en su forma natural. A la izquierda, la información aportada por AutoDock.

Imagen 12. Resumen de resultados obtenidos para el protein-based virtual screening izquierda). Mejor docking analizado, correspondiente a la molécula 5882, con una energía de unión de -13,51.

Imagen 13. Lista de candidatos a fármaco contra MCADD.

1. Introducción

1.1 Contexto y justificación del Trabajo

En este TFM se abordará la búsqueda de un tratamiento farmacéutico para la deficiencia en acyl-CoA deshidrogenasa (MCADD) a través de la bioinformática^[11].

La MCADD es un defecto congénito transmitido por herencia autosómica recesiva que afecta a una media de 1:10.000 personas. Es, por tanto, una enfermedad hereditaria rara y su causa principal es la mutación c.985A>G en el gen ACADM, que resulta en el reemplazamiento de una lisina por un ácido glutámico en la posición 304 de la acyl-CoA deshidrogenasa de cadena media (MCAD)^[8]. La proteína mutada alcanza tan solo el 35-45% de su actividad enzimática, dando lugar a un fallo en la oxidación de ácidos grasos de cadena media. La consecuencia inmediata de este déficit de actividad es una rápida crisis metabólica, habitualmente acompañada de vómitos, hipoglucemia hipocetónica e incluso complicaciones como el coma o la muerte prematura si no tiene lugar una intervención médica adecuada^[22].

Hasta el momento no se ha identificado un fármaco eficaz frente a la enfermedad, sin embargo se ha encontrado la posibilidad de mejorar levemente los síntomas a través del aporte de cofactores, como la riboflavina^[5]. Es por ello que se ha decidido abrir una nueva vía de búsqueda, que permita encontrar un compuesto capaz de aumentar la calidad de vida de los pacientes.

Al tratarse de una enfermedad rara, donde la información experimental disponible es escasa y la inversión en investigación casi inexistente, se puede explotar y dar a conocer el potencial de la bioinformática. Facilitar la llegada de recursos económicos^[31] a empresas emergentes en el sector de las llamadas «orphan diseases» ha favorecido su crecimiento, llegando estas a revalorizarse hasta los 60 – 400 millones de dólares^[10]. El progreso se ha visto reflejado en la salida a mercado de 39 fármacos efectivos en el último año. El retorno económico para el sector alcanzó el pico de los 114 billones\$ en 2017, aumentando un 12% respecto a 2015 y con previsiones de doblarse en los próximos años^[4]. Es pues evidente tanto el atractivo como la necesidad social de introducirse en el sector, teniendo en cuenta que aún existen alrededor de 7000 enfermedades raras sin tratamiento alguno.

En resumen, este trabajo pretende encontrar una serie de compuestos capaces de postularse como fármacos efectivos para la MCADD, explotando las herramientas informáticas para abaratar los costes de producción farmacológica, realizando varias etapas del descubrimiento de fármacos a través de un ordenador. A su vez, se pretende visibilizar las *orphan diseases* y su potencial como nicho de mercado con el fin de favorecer la creación de nuevos fármacos para los pacientes con este tipo de enfermedades.

1.2 Objetivos del Trabajo

El objetivo de este trabajo no es otro que el de encontrar una serie de compuestos, candidatos a fármacos que puedan ser efectivos frente a la MCADD. Al ser un objetivo ambicioso y el tiempo disponible escaso, se ha dividido el proceso en tres fases, poniendo punto final a cada una de ellas con un hito relacionado con la reducción del número de compuestos con el que trabajar, que irá desde el total de los disponibles en las bases de datos hasta los más adecuados para continuar una investigación de carácter más específico para la enfermedad.

OBJETIVO GENERAL	OBJETIVO ESPECÍFICO	TAREA
Identificar el farmacóforo de MCAD	Obtener la estructura 3D de la molécula y ligandos	Búsqueda en bases de datos de visualización de moléculas
	Determinar el farmacóforo (Hito 3)	Similarity-based ligand screening: LiSiCA, Drugbank, scPDB.
		Utilización del software LigandScout
		Evaluación de los resultados
Buscar compuestos que interaccionen con MCAD	Seleccionar los ligandos que contengan en farmacóforo (Hits, Hito 5)	Pharmacophore-based ligand screening: LigandScout
		Evaluación de resultados
	Acotar la lista de compuestos a aquellos que interaccionen fielmente con la diana (Leads, Hito 7)	Protein based virtual screening: AutoDOCKtools
		Evaluación de resultados
Evaluar las propiedades ADME-Tox de los candidatos obtenidos en el objetivo anterior	Evaluar la toxicidad de los fármacos seleccionados	Búsqueda a través de bases de datos como LimTOX
	Anotar las propiedades farmacocinéticas (ADME) de los candidatos	Estudio de las características mediante pkCSM

Tabla 1. Distribución de objetivos y tareas

- Hito 1: Entrega PEC0. Elección del tema de trabajo.
- Hito 2: Entrega PEC1. Planificación del proyecto.
- **Hito 3:** Determinación del farmacóforo. Será la base para el ligand-based virtual screening.
- Hito 4: Entrega PEC 2. Desarrollo fase 1. Incluirá la fase preliminar de investigación y la identificación del farmacóforo.
- **Hito 5:** Consecución de la lista de moléculas que poseen el farmacóforo propuesto.
- Hito 6: Entrega PEC 3. Desarrollo fase 2. Se mostrarán los resultados del pharmacophore-based ligand virtual screening así como la preparación del protein-based virtual screening.
- **Hito 7:** Lista completa de candidatos.
- Hito 8: Entrega PEC 4. Memoria completa TFM.
- Hito 9: Entrega diapositivas.
- Hito 10: Entrega presentación oral.

1.3 Enfoque y método seguido

A la hora de abordar una enfermedad compleja como la MCADD, con sintomatología variada^[19], parece lógico escoger el epicentro de los problemas como diana. Al estar bien definida la causa de la patología, este proyecto se centrará en encontrar un compuesto capaz de disminuir los efectos de la mutación en la enzima defectuosa.

La principal característica que debe cumplir un fármaco para ser efectivo es la de interaccionar con la diana. El virtual screening es una herramienta con eficacia probada en numerosas investigaciones científicas^[14] para este fin. Se conocen varios formatos de esta disciplina, en función de si el screening se basa en el ligando o en la proteína en cuestión.

Ligand-based virtual screening es la modalidad más flexible, en la que no es necesario que se conozca la estructura completa de la proteína problema. A partir de la estructura primaria, algoritmos de machine learning, ligandos de unión conocidos o el farmacóforo se puede encontrar una serie de candidatos para una interacción efectiva.

Por contra, para llevar a cabo el protein-based virtual screening es necesario conocer la estructura tridimensional de la proteína diana, pues se basa en algoritmos de docking. Estos buscan la mejor interacción entre una colección de posibles ligandos a través de puntuaciones de conformación, posición y orientación dentro del bolsillo de unión. Es, por tanto, un proceso más lento y complejo computacionalmente.

Se ha seleccionado un método que se apoya en ambas técnicas para encontrar los mejores fármacos y ambos se realizarán en paralelo.

El proyecto se apoyará en una programación basada en objetivos, con tareas específicas para ello, como se ha explicado en el punto anterior. Así pues, en una primera fase se determinará el farmacóforo de la enzima mutada. Esa información será el punto de partida para el ligand-based virtual screening. La segunda fase ocupará el grueso del trabajo y será en la que se obtendrán los resultados visibles del proyecto. Se partirá de una colección de en torno a 70.000 compuestos, que filtrados a través pharmacophore-based ligand screening, darán lugar a una lista con aquellas moléculas que contienen las características adecuadas para interactuar con MCAD, serán los hits. De entre esos ligandos candidatos se escogerán los que mejor puntuación obtengan y se someterán a los algoritmos de docking, para escoger únicamente los que presenten una buena interacción molecular con la proteína, los leads. Por último, se evaluarán sus capacidades ADME-Tox, haciendo que los candidatos finales sean efectivos y puedan adentrarse en el proceso de salida al mercado.

1.4 Planificación del Trabajo

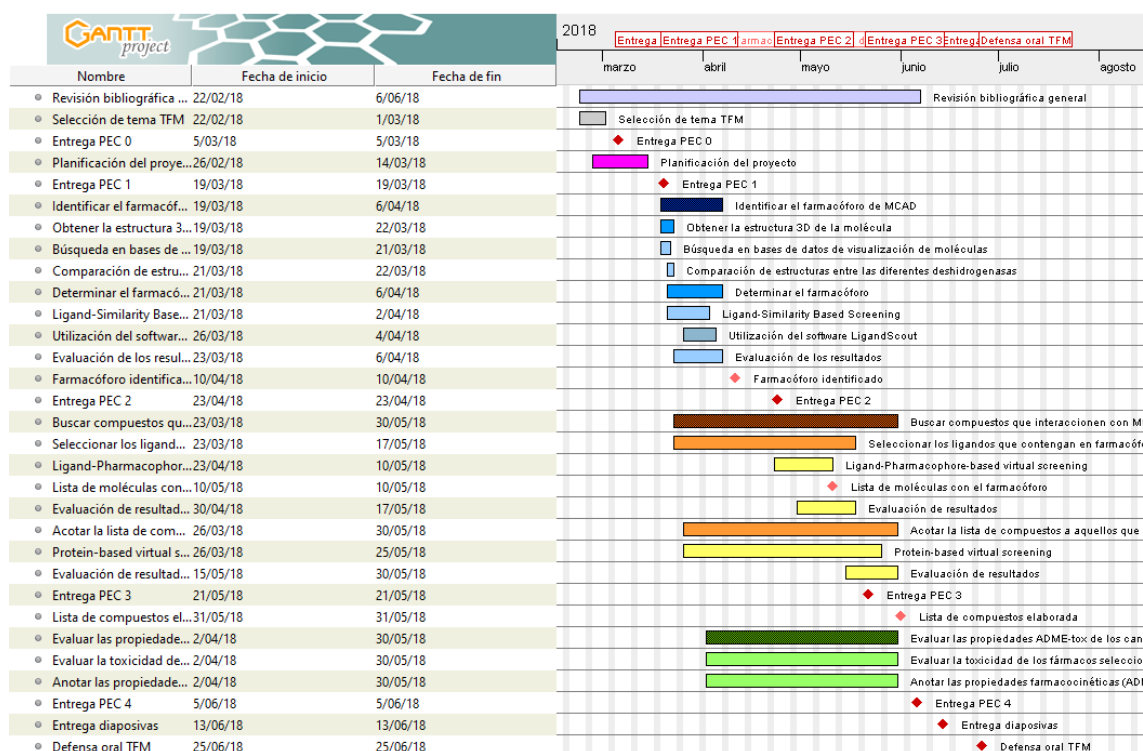


Imagen 1. Diagrama de Gantt.

1.5 Breve resumen de productos obtenidos

Tras finalizar cada una de las tareas programadas, se ha conseguido detallar una lista de 14 moléculas, fármacos potenciales para tratar a los pacientes de MCAD que sufren la mutación c.985 A>G. Dichos candidatos deberán ser sometidos a ensayos posteriores para determinar su validez farmacológica.

1.6 Breve descripción de los otros capítulos de la memoria

Para obtener la lista final de candidatos se ha elaborado un pipeline relacionado con el descubrimiento de fármacos por ordenador. En el resto de capítulos de la memoria se describirá cada una de las etapas, detallando los resultados obtenidos y la implicación que tiene dentro del global del trabajo.

- MCAD: Estructura de la proteína. Efectos de la mutación

Este primer capítulo recoge los datos recopilados en el estudio de la estructura de la proteína problema y los efectos que la mutación K304E causa en su estructura, funcionalidad, etc. Conocer cómo afecta esta variante a la MCAD silvestre ayudó a escoger una estrategia a través de la que abordar el proyecto.

- Identificación del farmacóforo

La forma de abordar el descubrimiento de fármacos fue a través de un cribado molecular basado en el conjunto de características estéricas y electrostáticas que los ligandos debían poseer para interactuar con la proteína. En este apartado se refleja el camino seguido para encontrar el farmacóforo de MCADK304E y los resultados obtenidos. Ésta será la base para el ligand-screening.

- Virtual Screening

El virtual screening es la herramienta que permitió reducir el número de compuestos candidatos hasta seleccionar los más adecuados para tratar la enfermedad. Se dividió en dos etapas, la primera de ellas basada en el farmacóforo de MCAD (ligand-based) y la segunda, en la interacción dentro del bolsillo de la propia proteína (protein based). Aquí se relatará la forma en la que se realizaron los experimentos de screening, detallando los criterios de selección de compuestos para cada uno de ellos. Se mostrarán los resultados finales del proyecto, mostrando cuáles son los candidatos a fármaco para MCADD.

- Evaluación de propiedades ADME-Tox

Por último, se evaluaron las propiedades de los candidatos para su actuación dentro del organismo. Estas son las conocidas como ADME-Tox e incluyen absorción, distribución, metabolización, excreción y toxicidad. Una molécula capaz de cumplir los criterios farmacológicos ADME y no presentar toxicidad puede considerarse como potencial fármaco efectivo.

2. MCAD: Estructura de la proteína. Efectos de la mutación.

La proteína silvestre MCAD está formada por 4 cadenas idénticas (homotetrámero), predominando la conformación α -hélice en la estructura secundaria de cada una de ellas[23].

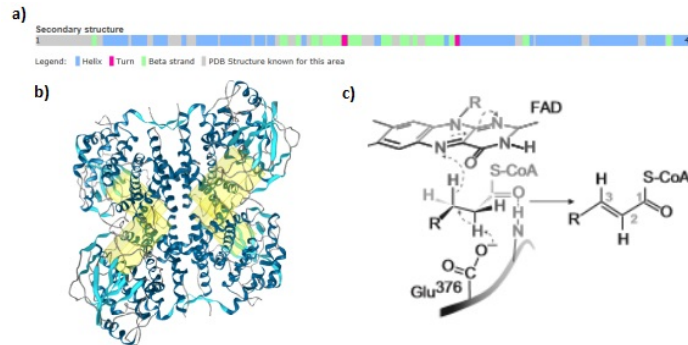


Imagen 2. Estructura secundaria (a) y cuaternaria de MCAD (b), con cada uno de los bolsillos de reacción sombreados en amarillo. Dicha reacción queda detallada a la derecha de la proteína completa (c).

Como proteína catalizadora del primer paso en la oxidación mitocondrial de ácidos grasos, posee dos ligandos principales. La flavoproteína electrón-transferente y el sustrato/electrón aceptor. Ambos se unen a MCAD dentro de un mismo bolsillo (imagen 2b), formando un complejo ternario (ETF-complex) a través del cual se lleva a cabo la α,β -deshidrogenación. La reacción que cataliza puede verse resumida en: *medium-chain acyl-CoA* + *electron-transfer flavoprotein* = *medium-chain trans-2,3-dehydroacyl-CoA* + *reduced electron-transfer flavoprotein* (imagen 1c)[27].

Estudios de simulación dinámica indican que la estabilidad del tetrámero se ve favorecida mediante la unión del FAD, por lo que este ligando mejoraría el ensamblaje de los monómeros[6] a la par que actuaría como sustrato de reacción. A través de estos datos se muestra el funcionamiento general de MCAD, la proteína silvestre. A continuación se expondrán las implicaciones de la mutación K304E en su funcionamiento.

Según la estructura cristalográfica de la proteína, el aminoácido 304 orienta su cadena lateral hacia una cavidad situada dentro del core interior del tetrámero, de unos 14 angstroms de diámetro, compartida por los 4 monómeros.

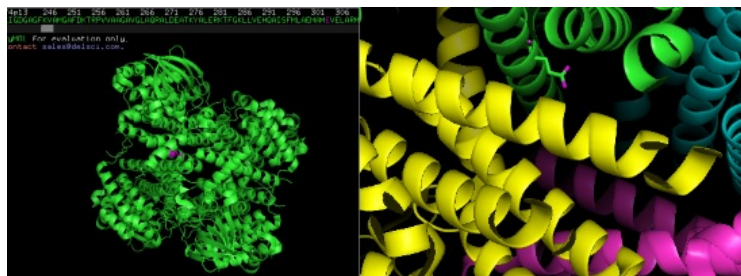


Imagen 3. Selección del aminoácido E304 dentro de la proteína mutada. Se muestra su ubicación exacta dentro del tetrámero.

No existen modificaciones dentro de la estructura secundaria de MCADK304E. Sin embargo, la mutación introduce 4 cargas negativas adicionales dentro del core, desestabilizando la estructura cuaternaria de la enzima. Respondiendo a esta inestabilidad, la proteína resultante comienza a desnaturalizar, formando agregados[30]. La consecuencia inmediata de este comportamiento anormal es una pérdida en actividad enzimática, que da lugar a los síntomas propios de la enfermedad [20][29].

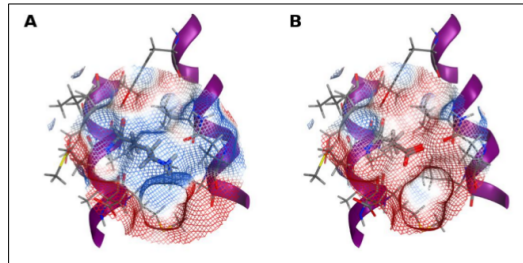


Imagen 4. Zoom del bolsillo al que pertenece el aminoácido 304 de MCAD (A) vs MCADK304E (B), con 4 cargas negativas extra. La nube electrostática positiva se muestra en azul y la negativa en rojo.

Esta modificación a nivel de carga podría afectar a la actividad catalítica debido a[9]:

- La exposición de una porción del cofactor o el sustrato a un entorno polar, haciendo que se reduzca la desolvatación del residuo catalítico E376, reduciendo su pKa.
- La incorrecta unión del sustrato al haberse incrementado las distancias entre el anillo de flavina y el enlace C α -C β , haciendo que no se alcancen los valores óptimos para la reacción.
- El acoplamiento incorrecto de FAD, debido a las modificaciones en su sitio de unión.

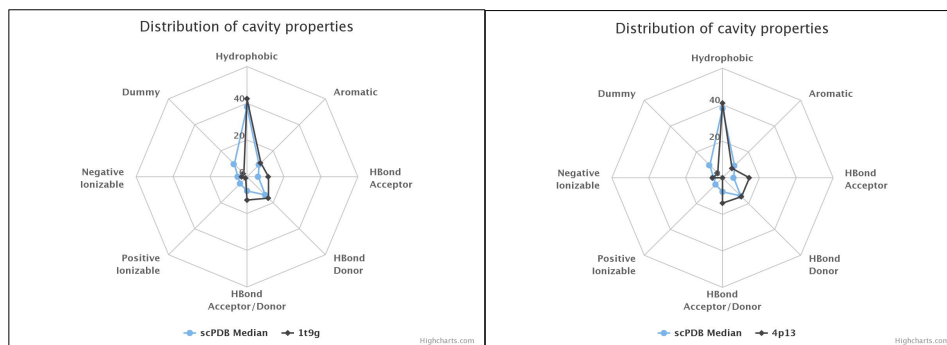


Imagen 5. Distribución de las propiedades de la cavidad donde se aloja FAD dentro de MCAD. A la izquierda, la enzima silvestre (1t9g), a la derecha la enzima mutante(4p13), donde se ve un aumento en la ionización negativa, entre otros parámetros.

Los datos experimentales indican que la desestabilización producida por la mutación es únicamente de tipo electrostático. Así pues, tras el análisis del efecto de la mutación sobre la estructura de la proteína, se decidió buscar una molécula capaz de llevar a cabo las funciones de FAD, adaptándose al nuevo entorno negativo de forma más eficiente. Esa molécula deberá ocupar el bolsillo de unión de ligando[1], actuando tanto como agente redox como de chaperona, mejorando así la estabilidad del tetrámero y aumentando la actividad enzimática de la MCAD en pacientes con este defecto congénito.

3. Identificación del farmacóforo.

Tras la revisión bibliográfica y la búsqueda en bases de datos se tomó la decisión de basar la búsqueda farmacológica en las propiedades de FAD y riboflavina para interactuar efectivamente con MCAD. El primero de ellos, por ser el ligando de la proteína silvestre, actuando como chaperona[6] y como agente de deshidrogenación[27]. El segundo, por existir evidencia de su funcionalidad como “fármaco” ante la enfermedad, compartiendo estructura con FAD y dando muestra de que el criterio seleccionado para encontrar un nuevo fármaco es consistente.

Así pues, en primer lugar, se estudiaron las características de ambos ligandos a nivel estérico y electrostático, lo que se denominó identificación del farmacóforo[2].

Se establecieron las claves del acoplamiento de FAD dentro del bolsillo de la proteína, prestando especial atención al N4, situado en los anillos de flavina y descrito como el encargado de soportar la reacción de deshidrogenación tras la unión del acyl-CoA. En la imagen 6 aparecen hasta 15 contactos que el ligando establece con la proteína, dando muestra de la complejidad de la unión.

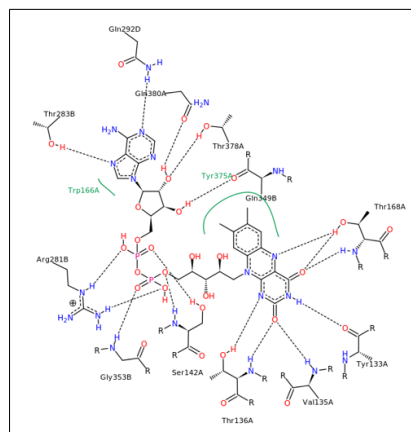


Imagen 6. Contactos de FAD dentro de su cavidad en MCAD.

En base a esta interacción se propuso realizar un tipo de farmacóforo apoyado en las propiedades intrínsecas de ambos ligandos de la proteína. Tanto FAD como riboflavina son capaces de interactuar con MCADK304, por lo que encontrar sus puntos en común nos permitió iniciar la búsqueda otras moléculas capaces de realizar su función. Se denominó shared-features pharmacophore.

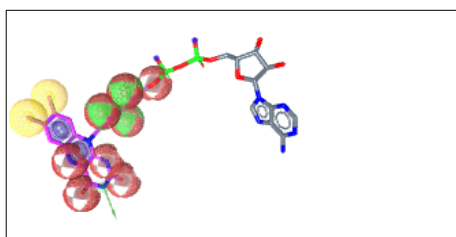


Imagen 7. Farmacóforo determinado a través de las características comunes de FAD y riboflavina. Cuenta con 2 estructuras aromáticas (azul), 2 contactos hidrofóbicos (amarillo), 7 aceptores (rojo) y 4 donadores (verde).

A su vez, basándose en la idea de que moléculas similares tienen funciones análogas, se realizó un ligand-based virtual screening a través de DrugBank[7] y scPDB[28] (anexo I), que proporcionó una serie de moléculas semejantes a FAD/riboflavina. Un proceso equivalente se llevó a cabo con LiSiCa[15], software de screening basado en el coeficiente de Tanimoto (anexos II y III). De entre los compuestos hallados, se escogieron los 10 con mayor coeficiente de similaridad y, a partir de ellos se creó el segundo modelo de farmacóforo, basado en su alineamiento molecular. Los compuestos que formaron este merged-features pharmacophore fueron: zinc ID[31] 53683566, 00391159, 53683526, 58632138, 02036848, 53682927, 85425702, 98208900, 96006025 y 53684299 (mostrado en el anexo VI).

Ligand Scout[16] es la aplicación seleccionada para crear los distintos farmacóforos, pues ofrece una amplia gama de herramientas de interacción a nivel molecular, adaptándose perfectamente a las exigencias del experimento. Así pues, se retrataron las características del farmacóforo de MCADK304E mediante el programa.

4. Virtual screening

Tras la determinación de los dos modelos de farmacóforo propuestos se llevaron a cabo experimentos de virtual screening. Esta metodología permitió encontrar aquellos compuestos que cumplían los criterios marcados en la primera fase del proyecto.

4.1 Pharmacophore-based ligand virtual screening

Se elaboró una librería inicial de compuestos como punto de partida. Tras utilizar 4 colecciones diferentes para realizar el screening de similaridad con LiSiCA (DrugBank approved, SureChembl, ChemMol y PDB), se eligieron SureChembl y PDB para arrancar el pharmacophore-based virtual screening. Son los conjuntos que más compuestos similares a FAD/Riboflavina mostraron, por lo que se fusionaron y pasaron a ser la colección escogida para este experimento.

LigandScout analizó más de 70.000 moléculas, filtradas a través de los modelos farmacológicos propuestos. Se denominaron *hits* aquellos compuestos que encajaron fielmente con, al menos, $\frac{3}{4}$ del total de los elementos que compone cada farmacóforo.

4.1.1 Ligand desheredar features pharmacophore

A partir de la disposición espacial de las 16 características que comparten FAD y riboflavina (imagen 7) se obtuvo una lista de 55 compuestos *hit*, con altos coeficientes de encaje farmacológico (superiores a 100 puntos).

Rank	Mark	Name	#	Matching Features	Pharmacophore-Fit Score	Max. Index
1		ZINC9820940	91	██████████	116.71	11
2		ZINC13943029	8	██████████	116.52	11
3		ZINC13939481	10	██████████	116.30	11
4		ZINC13943029	7	██████████	115.85	11
5		ZINC3060712	90	██████████	115.39	11
6		ZINC48019812	70	██████████	115.41	11
7		ZINC3060712	56	██████████	115.40	11
8		ZINC9820940	88	██████████	115.44	11
9		ZINC1063960	51	██████████	115.40	11
10		ZINC9820940	63	██████████	115.30	11
11		ZINC3060712	50	██████████	106.83	10
12		ZINC4339866	24	██████████	106.68	10

Imagen 8. Resultado del shared-features pharmacophore based ligand screening.

Son especialmente interesantes los resultados que posean tanto los 2 anillos aromáticos como el N aceptor, un criterio que se tendrá en cuenta a la hora de seleccionar positivamente las moléculas que serán efectivas en la reacción de deshidrogenación.

4.1.2 Ligand merged features pharmacophore

A partir de la fusión de las características que poseían los 10 ligandos con mejor puntuación en los test de similaridad se generó un farmacóforo de 7 características, 6 donadores y un aceptor. Al ser un criterio menos restrictivo que el anterior, aumentó el número de *hits*, con más de un 50% de positivos. Casi 40.000 moléculas superaron esta criba, teniendo que encontrar una nueva forma de acotar los resultados.

Rank	Mark	Name	#	Matching Features	Pharmacophore-Fit Score	Max. Index
1		ZINC43300747	20667	██████████	70.06	41823
2		ZINC79161821	25541	██████████	70.04	52196
3		ZINC33505425	33548	██████████	70.00	69016
4		ZINC79161823	25943	██████████	69.97	52564
5		ZINC17702310	45507	██████████	69.81	9058
6		ZINC17705309	4385	██████████	69.77	8882
7		ZINC4482384	15026	██████████	69.66	20687
8		ZINC9554496	24415	██████████	69.60	50263
9		ZINC79161826	29692	██████████	69.55	53332
10		ZINC43099585	28929	██████████	69.53	54807
11		ZINC0871813	6012	██████████	69.53	12161
12		ZINC95871250	5212	██████████	69.51	10581

Imagen 9. Resultado del merged-features pharmacophore based ligand screening.

4.2 Reducción de candidatos

Debido al tiempo y la capacidad computacional de la que se dispone para este proyecto, resulta inviable continuar el experimento con 40.000 *hits*. Por ello, se diseñó un nuevo sistema de cribado con el objetivo de reducir el número de moléculas que lanzar al protein-based virtual screening. A su vez, esta etapa sirvió como fase de análisis de resultados, puesto que en ella se valoró la eficacia de los compuestos obtenidos en base a los diferentes farmacóforos.

En primer lugar se ordenaron los candidatos en función de su score farmacológico. Cabe destacar que las moléculas seleccionadas a través del shared-pharmacophore poseen un score más alto, al hacer coincidir entre 11 y 16 características diferentes, por lo que se encuentran en los primeros puestos, con 60-100 puntos. Mientras, los representantes del filtrare por merged-pharmacophore obtienen valores en torno a los 70 puntos.

Basándose en el score, se seleccionaron los 100 primeros compuestos de la lista. Pese a partir de un conjunto de 40.000 no resultó traumático echar por tierra tal cantidad de moléculas, se debían analizar únicamente los mejores. La totalidad de compuestos eliminados procedían del merged-pharmacophore ligand screening, los cuales no poseían una de las características más relevantes que se buscaban, los anillos aromáticos.

Una vez obtenida esta pre-lista de 100 *hits*, se realizó un análisis más exhaustivo de sus propiedades. En este punto se llevaron a cabo estudios preliminares de las propiedades ADME-Tox a través de las reglas de Lipinski y se tuvieron en cuenta características propias de la patología que estábamos tratando, como podían ser:

- Poseer varios anillos, con enlaces capaces de participar en la reacción de deshidrogenación (similares a los de FAD/Riboflavina).
- Presentar un entorno electrostático positivo o neutro (ya que la mutación introduce cargas negativas que desestabilizan la molécula MCADK304E).

Los resultados de este estudio se presentan en los anexos IV y V y gracias a ellos se elaboró una lista manejable que lanzar a AutoDock.

	ZINC	Score	Benign F.	Rings	Charges	Lipinski
1	98209531	125,01	YES	YES	-5	
2	98209268	116,9	YES	YES	-5	
3	13543028	116,23	YES	2	-1	
4	5835253	115,92	YES	2	0	
5	98208793	115,82	YES	4	-2	
6	58660722	115,55	NO	NO	-3	NO
7	85425702	115,45	NO	YES	-3	NO
8	34390682	115,41	NO	NO	-2	YES
9	33821189	115,39	YES	NO	0	
10	40913495	115,27	YES	NO	0	
11	58638382	115,18	NO	2	-2	YES
12	95553067	115,13	YES	NO	0	
13	98208651	114,51	YES	2	-2	
14	73240563	85,54	YES	NO	1	
15	73240514	75,73	YES	YES	0	
16	84655552	75,62	YES	2	0	
17	33295967	66,05	NO	NO	1	YES
18	391159		YES	YES	0	
19	98208900		YES	NO	-2	
20	58639144	56,75	YES	NO	0	

Imagen 10. Lista de 20 *leads* que analizar a través del protein-based virtual screening.

Se seleccionaron de entre los 100 *hits* por poseer 2 o más características de entre las propuestas.

Benign F. = Uso farmacológico anterior. Rings = Anillos similares a los de flavoproteínas.

Charges = Carga formal a pH neutro. Lipinski = Cumplir 3 o más reglas de fármaco potencial.

Para escoger a los 20 *leads* del protein-based virtual screening se optó por escoger aquellas moléculas que cumplían al menos 2 de las 3 condiciones planteadas para el tratamiento de esta enfermedad. 17 de ellos fueron seleccionados por presentar efectos farmacológicos conocidos, un entorno electrostático neutro o una estructura de anillos aromáticos similar a la de flavoproteínas.

Los 3 restantes se añadieron por aparecer tanto en la lista procedente del shared-features pharmacophore como en la del merged-features pharmacophore. Con estos 20 leads se comenzó la nueva fase de cribado.

4.3 Protein-based virtual screening

El protein based virtual screening analiza la unión a nivel molecular entre una proteína y un ligando, evaluando cuál es la mejor conformación para hacer efectiva esa interacción y puntuando la calidad de la unión que se consigue. Se seleccionó AutoDock como software de análisis debido a que incluye varios algoritmos de docking y una enorme variabilidad para la selección de parámetros. Gracias a él se estudiará la interacción entre los leads y MCADK304E.

El primer paso fue calibrar el software, encontrando los parámetros que es necesario modificar para conseguir una reproducción fiel de la realidad. Es necesario tener en cuenta que este tipo de experimentos no se realizan en condiciones naturales, puesto que la estructura cristalográfica de las proteínas no siempre se obtiene en disolución y la rigidez molecular no es algo que se dé en las formas silvestres. Así pues, se configuró AutoDock para obtener los mejores resultados en los análisis posteriores, donde se evaluaría la capacidad de los candidatos para interactuar con MCADK304E.

Se escogieron varios **enlaces flexibles dentro del ligando**, dando lugar a diferentes conformaciones que testar dentro de la región que alberga el bolsillo de FAD en la molécula de MCADK304E (cadenas A y B), en su **forma rígida**. El **algoritmo genético**^[3] fue el seleccionado para llevar a cabo el experimento y gracias a él se obtuvieron las mejores posiciones de unión ligando-molécula. El programa también aporta diferentes parámetros informativos, como la energía de unión, eficiencia del ligando, etc.

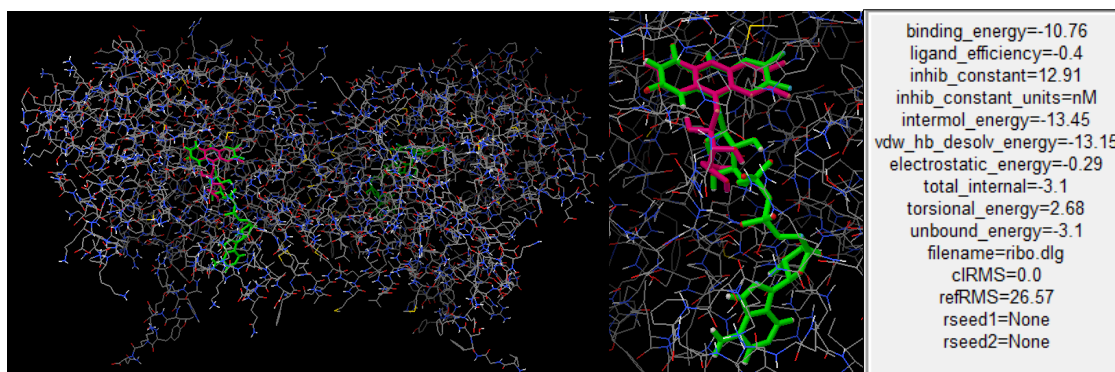


Imagen 11. Zoom out a la izquierda (MCADK304E cadenas A y B) y zoom in de la molécula experimental de docking (MCADK304E) con el ligando problema riboflavina (rosa) superpuesta con la molécula original de MCADK304E que contiene el ligando (FAD, verde) en su forma natural. A la izquierda, la información aportada por AutoDock.

Se ha considerado una buena aproximación a la realidad, obteniendo una buena superposición virtual sobre la estructura cristalográfica y con energías de unión negativas para el ligando. Así pues, será bajo estas premisas sobre las que se realicen todos los experimentos de docking para los *leads*.

Tras llevar a cabo el análisis de la interacción entre los 20 compuestos y MCADK304E, el resultado quedó recogido en el anexo VII y resumido en la imagen 10.

	ZINC	Binding ener.	Ligand eff.	Electrostatic
Ref. 1	FAD	-6,25	-0,12	-0,75
Ref. 2	RIBOFLAVIN	-10,76	-0,4	-0,29
1	98209531	3,12	0,07	-0,07
2	98209268	-4,82	-0,1	0,32
3	13543028	-7,09	-0,3	-0,29
4	5835253	-9,12	-0,41	-0,48
5	98208793	-13,51	-0,34	-1,53
6	58660722	-8,84	-0,25	0,07
7	85425702	1,27	0,02	-0,8
8	33295967	-13,17	-0,38	-0,9
9	34390682	-8,4	-0,3	-0,27
10	33821189	-6,51	-0,3	-0,48
11	40913495	-10,07	-0,35	-0,41
12	58638382	-8,48	-0,33	-0,84
13	95553067	-7,87	-0,36	-0,42
14	98208651	-10,66	-0,29	-0,82
15	73240563	-10,69	-0,34	-0,55
16	73240514	-9,4	-0,39	-0,33
17	84655552	-9,63	-0,42	-0,13
18	391159	-9,8	-0,52	-0,24
19	98208900	-3,86	-0,07	-0,45
20	58639144	-8,45	-0,32	-0,97

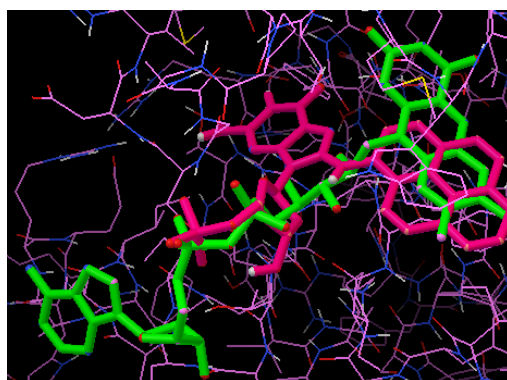


Imagen 12. Resumen de resultados obtenidos para el protein-based virtual screening (izquierda). Mejor docking analizado, correspondiente a la molécula 98208793, con una energía de unión de -13,51 (derecha).

Los resultados de docking se ordenaron en base a distintos tipos de energía, relacionados entre sí. Los compuestos con una afinidad menor que la de FAD o riboflavina no se consideraron aptos para continuar el estudio y por ello se eliminaron. Fueron 98209531, 98209268, 85425702, 33821189 y 98208900.

5. Evaluación de propiedades ADME-Tox

Tras la selección de candidatos en base a criterios moleculares (farmacóforo, docking) fue necesario analizar su capacidad de actuación como fármaco. Esto es evaluar su facilidad de absorción, distribución, metabolización y excreción dentro del organismo, así como las posibles reacciones adversas que pueda provocar. En esta última fase se utilizaron herramientas web como LimTox y pkCSM.

LimTox^{[17][18]} es una base de datos creada por el Centro Nacional de Investigaciones Oncológicas (CNIO) para buscar reacciones adversas y de toxicidad de una amplia gama de medicamentos. Proporciona conocimiento extraído a partir de la biblioteca biomédica PubMed, informes de la Agencia Europea del Medicamento (EMA, por sus siglas en inglés) y solicitudes de nuevos fármacos de la Administración de Alimentos y Medicamentos de Estados Unidos (FDA, por sus siglas en inglés).

pkCSM^{[24][25]} se seleccionó como medio complementario a LimTox para determinar las propiedades farmacocinéticas de las moléculas candidatas. Este software se apoya en el modelado gráfico basado en distancias. A través de la representación química gráfica de la molécula problema, algoritmos de machine learning ayudan a identificar características que la asocien con ciertos perfiles ADME-Tox.

Se analizaron aquellas moléculas que no habían presentado evidencia de haber sido utilizadas anteriormente como fármaco. Los resultados obtenidos se adjuntan en el anexo VIII. Tras esta evaluación se eliminaron dos compuestos, por presentar alta toxicidad (85425702) y una mala absorción y distribución (58660722).

En la imagen 11 se muestran los 14 candidatos, las moléculas que han conseguido superar todos los filtros propuestos para este experimento.

	ZINC	Score	Binding ener.	Ligand eff.	Electrostatic	ADME-Tox
1	98208793	115,82	-13,51	-0,34	-1,53	OK
2	33295967	66,05	-13,17	-0,38	-0,9	OK
3	73240563	85,54	-10,69	-0,34	-0,55	OK
4	98208651	114,51	-10,66	-0,29	-0,82	OK
5	40913495	115,27	-10,07	-0,35	-0,41	OK
6	391159		-9,8	-0,52	-0,24	OK
7	84655552	76,72	-9,63	-0,42	-0,13	OK
8	73240514	75,73	-9,4	-0,39	-0,33	OK
9	5835253	115,92	-9,12	-0,41	-0,48	OK
10	58638382	115,18	-8,48	-0,33	-0,84	OK
11	58639144	56,75	-8,45	-0,32	-0,97	OK
12	34390682	115,41	-8,4	-0,3	-0,27	OK
13	95553067	115,13	-7,87	-0,36	-0,42	OK
14	13543028	116,23	-7,09	-0,3	-0,29	OK

Imagen 13. Lista de candidatos a fármaco contra MCADD.

6. Conclusiones

6.1 Impresiones sobre el trabajo

La principal motivación para comenzar a trabajar fue la de explotar el potencial de la bioinformática en un plano de carácter investigador. La introducción al mundo de las enfermedades raras ha resultado ser muy gratificante a la par que ha permitido indagar sobre los diferentes recursos disponibles para el descubrimiento de fármacos. Se han utilizado tanto herramientas web (pkCSM, bases de datos) como basadas en el lenguaje de programación python (PyMol^[26], AutoDock) así como aplicaciones específicas del área farmacéutica (LigandScout). Una amplia gama de recursos informáticos aplicables a otros campos de investigación y ámbitos laborales que enriquecen los conocimientos del investigador.

6.2 Seguimiento de la planificación

Inicialmente se dedicaron varias semanas al diseño de un plan de trabajo. Este plan se basó en la consecución de objetivos parciales e hitos: Obtener la estructura 3D de la molécula y ligandos, determinar el farmacóforo de MCAD (hito), seleccionar los ligandos que contengan este farmacóforo (hito), acotar la lista de compuestos a aquellos que interaccionen fielmente con la proteína diana (hito), evaluar la toxicidad de los fármacos seleccionados y anotar las propiedades farmacocinéticas (ADME) de los candidatos. Conseguir cada uno de ellos significa avanzar un paso hacia la consecución del objetivo final, encontrar una serie de potenciales fármacos para MCADD. Así, si debido a una mala gestión del tiempo o cualquier otro contratiempo el proyecto debe finalizar en alguna de sus fases intermedias se hubiera podido elaborar dicha lista de candidatos, pero ésta hubiera sido menos específica en cuanto a criterios de selección.

Los plazos establecidos se han cumplido con bastante precisión, teniendo que modificar ciertas etapas debido a la necesidad de invertir más tiempo en la familiarización con las distintas herramientas. Este error supuso prolongar la duración de las tareas relativas a la experimentación y acortar los periodos de análisis de resultados. Especialmente en la etapa de protein-based virtual screening se ha perdido algo de información experimental, al evaluar únicamente los datos más relevantes. Además, se tuvo que acotar la fase de anotación de propiedades farmacocinéticas, por lo que sólo aquellos compuestos de los que no se conozca una actividad previa como fármaco fueron analizados pese a que hubiera sido interesante tener esta información para los 14 candidatos.

A pesar de este contratiempo se cree haber realizado una buena planificación y haber estructurado las tareas correctamente ha sido una de las claves para completar todas las fases del proyecto.

6.3 Resultados obtenidos

Una vez finalizado el proyecto, tras el análisis de resultados se puede afirmar que se han cumplido los objetivos propuestos. El resultado final ha sido una lista que contiene los 14 compuestos (de entre 70.000) con más probabilidad de adentrarse en las fases intermedias del desarrollo de fármacos y superarlas con éxito. Se dice que 1 de cada 10.000 moléculas pasa el desarrollo de fármacos, pues en este trabajo se han seleccionado 14 de 70.000, aquellas que más probabilidad se cree que tienen para ser las 7 estadísticamente correspondientes, dando muestra del poder de la bioinformática en este ámbito de trabajo.

Los criterios de selección se basaron en mitigar los efectos de la mutación K304E dentro de MCAD y mejorar su actividad enzimática. La solución propuesta para este problema fue encontrar un compuesto capaz de llevar a cabo la función que el ligando natural tenía en la enzima silvestre, pero en este caso, dentro de la proteína mutada. Se utilizó el virtual screening como herramienta, abordando la búsqueda de candidatos desde la perspectiva del ligando (ligand-based virtual screening) y desde la de MCADK304E (protein-based virtual screening). Ambas variantes de virtual screening se seleccionaron por conocerse el éxito que habían generado en otras investigaciones de tipo farmacológico, así que fueron trasladadas a la búsqueda de fármacos para MCADD.

El ligand based virtual screening ha dado como resultado una lista de 100 candidatos basada en las propiedades de FAD y riboflavina. Se ha considerado un acierto añadir a riboflavina como molde para el screening, puesto que al ser ya utilizado como fármaco ha aportado cierta seguridad al experimento, haciendo que aparezcan muchos compuestos farmacológicamente potenciales. Con un mayor conocimiento en programas de virtual screening se podría haber elaborado un farmacóforo más efectivo, que diera unos resultados más precisos y que recortara tiempo en un cribado posterior, pero para ello hubiera sido necesario más tiempo de dedicación al proyecto.

El protein-based virtual screening ha demostrado ser una herramienta muy potente pero computacionalmente lenta. Lo que más atrajo la atención de este experimento es que riboflavina presentara una energía de unión mayor que el propio FAD, dado muestra de que compuestos más pequeños o farmacológicamente más probables pueden dar grandes resultados en el experimento. Bien es cierto que no se pueden extrapolar directamente los resultados a una unión en condiciones naturales, pero los resultados han sido bastante esperanzadores, especialmente para moléculas como 40913495 o 58638382 con altas energías de unión y la colocación espacial que permitiría llevar a cabo la reacción de deshidrogenación.

En definitiva, el camino escogido ha permitido alcanzar los objetivos dentro de los plazos marcados al inicio del proyecto. Sin embargo, con un mayor conocimiento en el ámbito computacional se podría haber aumentado la calidad de la lista obtenida.

6.4 Líneas de trabajo futuras

Utilizar el pipeline que se ha creado al realizar este trabajo puede ser de utilidad al abordar otras mutaciones, que puedan afectar a otras cavidades de la proteína. El intentar combatir una desestabilización electrostática es complicado, por lo que se podría trabar con defectos que afecten a nivel estructural o catalítico.

A su vez, sería interesante profundizar en la creación de farmacóforos más detallados, seleccionando las características más importantes en la interacción con la proteína y, por tanto, aportar mejores resultados.

7. Glosario

Aceptor: Un aceptor de electrones es una entidad química que es capaz de recibir electrones que le son transferidos desde otro compuesto y, por tanto, oxidarlo.

Chaperona: Las chaperonas son un grupo de proteínas que se encuentran tanto en procariotas como en eucariotas de manera abundante y en todos los compartimentos celulares, cuyas funciones principales son: favorecer el plegamiento de proteínas sintetizadas *de novo*, ayudar en el plegamiento correcto de proteínas después de un proceso de desnaturalización y colaborar en volver a plegar correctamente proteínas que han sido translocadas hacia algún compartimento celular.

EMA: Agencia Europea del Medicamento.

FDA: Food and Drug Administration.

Flavina: 7,8-dimetil-isoaloxazina, es una base nitrogenada cuya cadena principal es una sustancia heterocíclica nitrogenada de tres anillos y dos grupos oxo.

Flavoproteína: Las flavoproteínas se pueden definir, como enzimas que catalizan reacciones de óxido-reducción utilizando flavina mononucleótido (FMN) o flavina adenina dinucleótido (FAD) como coenzima.

MCAD: Acyl-CoA Deshidrogenasa de Cadena Media.

MCADD: Deficiencia en Acyl-CoA Deshidrogenasa de Cadena Media.

Orphan diseases: Enfermedades huérfanas. Es el término en inglés para referirse a las enfermedades raras, aquellas que por afectar a pocos pacientes se dejan de lado y no se presta interés a buscar una cura.

Pharmacophore-based ligand virtual screening: Cribado de ligandos naturales en base a la posesión de ciertas características estéricas y electrostáticas concretas, denominadas farmacóforo.

pKa: pH al que un compuesto está en un 50% en forma ionizada. Un aumento o descenso del pKa de una molécula puede alterar su estado de ionización para condiciones determinadas.

Protein-based virtual screening: Cribado de compuestos basado en la interacción (docking) de un ligando dentro de una proteína en cuestión. Se mide en términos de energía libre.

Similarity-based ligand virtual screening: Modalidad de screening basada en la premisa de que sustancias que comparten características estructurales tendrán funciones semejantes.

TFM: Trabajo de Fin de Máster.

8. Anexos

- **ANEXO I:** Resultados del screening de similaridad llevados a cabo con scPDB y Drugbank
- **ANEXO II:** Resultados del análisis LiSiCA para FAD
- **ANEXO III:** Resultados del análisis LiSiCA para Riboflavina
- **ANEXO IV:** Resultados del pharmacophore-based ligand virtual screening (parte 1, shared-features)
- **ANEXO V:** Resultados del pharmacophore-based ligand virtual screening (parte 2, merged-features)
- **ANEXO VI:** Resultados del protein-based virtual screening
- **ANEXO VII:** Resultados de los análisis ADME-Tox

ANEXO I

Propuestas de scPDB: Compuestos similares a FAD

Nombre
S-Adenosylhomocysteine
[(2R,3S,4R,5R)-5-(2-amino-6
Estradiol
GTP-gamma-S
(4R)-2-(2-ethoxyethyl)-4-ethyl
NADPH
2-[(4S)-4-[[5-(dimethylamino
Troloxerutin
dCF
DCF

Drugbank: Estructuras similares a FAD

Nombre	Score
Deamido-Nad+	0,831
GTA	0,823
Adenosine triphosphate	0,819
adenosine pentaphosphate	0,819
ADPR	0,817
Adenosine phosphate	0,799
NADH	0,769
cAMP	0,767
Guanosine-5'-Monophosphat	0,762
Tryptophanyl-5'amp	0,752
4-hydroxybenzoyl-CoA	0,745
Riboflavin	0,701

ANEXO II

LiSiCa plugin – PyMol: FAD

vs PDB			vs SureChEMBL		
ZINC	NAME	Score	ZINC	NAME	Score
85425702	FADH	0,93	4806433	5'-ADP-Na2	0,509
98208900	[[[2R,3S,4R,5R)-5-(0,814	12360703	ADP	0,509
53682927	NADH	0,702	3650334	Vitamin B2	0,509
98209472	[[[2R,3S,4R,5R)-5-(0,672	4806450	5'-ATP-Na2	0,5
98209260	[[[2R,3R,4R,5R)-5-(0,672	3871615	5'-ATP-Na2	0,5
98209273	5-chloro-N-[[4-(3,5-	0,672	3871614	5'-ATP-Na2	0,5
85534451	NADPH	0,672	3871612	5'-ATP-Na2	0,5
85629102	[[[2R,3S,4R,5R)-5-(0,656	3871613	5'-ATP-Na2	0,5
98208624	[[[2R,3S,4R,5R)-5-(0,656	8215728	dADP	0,491
98209251	[[[2R,3R,4R,5R)-5-(0,645	4096990	Propinol adenylate	0,481
98209475	[[[2R,3R,4R,5R)-5-(0,635	96328615	Adenylylimidodiphosph	0,481
85427559	NADP	0,629	96328614	Adenylylimidodiphosph	0,481

vs DrugBank			vs ChemMol		
ZINC	NAME	Score	ZINC	NAME	Score
53682927	NADH	0,702	3831423	Riboflavin	0,509
2036848	Riboflavin	0,509	2036848	Riboflavin	0,509
18456332	Adenosine triphosph	0,5	3831422	Riboflavin	0,509
3860156	Adenosine phosphat	0,434	3831424	Riboflavin	0,509
3927870	Fludarabine phosphate	0,426	29747209	ALPHA,BETA-MET	0,474
970363	Vidarabine	0,358	8215662	DATP	0,482
2169830	Adenosine phosphat	0,358	2126310	Vidarabine monoph	0,434
3798064	Cladribine	0,333	4806442	Adenylic acid	0,434
21981378	Tenofovir	0,333	3860156	Adenosine phosphat	0,434
3798247	CLOFARABINE	0,327	8613167	Adenylic acid	0,434
3823492	Nelarabine	0,321	13543718	arabine	0,426
3930376	ADEFOVIR DIPIVOXIL	0,318	1713574	2'-DEOXYADENOS	0,415

Conjunto de moléculas con mejor score obtenidas a partir del software LiSiCA para encontrar moléculas similares a FAD.

ANEXO III

LiSiCa plugin – PyMol: Riboflavina

vs PDB			vs DrugBank		
ZINC	NAME	Score	ZINC	NAME	Score
31976681	FMNH2	0,757	1035331	Ribavirin	0,517
58660829	7,8-dimethyl-2,4-dioxo-	0,632	1899	Penciclovir	0,5
58631698	(3R)-3-[7,8-dimethyl-2,	0,574	1505	Ganciclovir	0,5
13543028	6,7-dihydroxy-8-[[2S,3S	0,563	3802690	ENTECAVIR	0,469
58649829	[[[2R,3S,5R)-5-(3-carbam	0,543	13597823	Didanosine	0,467
8536413	formycin B	0,533	73711	Thiabendazole	0,464
58650380	[[[2R,3S,5R)-3-hydroxy-	0,531	3823492	Nelarabine	0,455
98209531	[[[Z)-3-[[[4aR)-7,8-dime	0,521	57146	Dyphylline	0,452
2047810	rCdRP	0,515	57147	Dyphylline	0,452
2047809	rCdRP	0,515	3872605	Fluvoxamine maleate	0,441
5892588	[[3,4-dihydroxy-5-(5-me	0,51	12466082	Fluvoxamine	0,441
98209268	[[[4R)-4-[[[4aR)-7,8-dime	0,5	4351	Ketotifen fumarate	0,441

vs ChemMol			vs SureChEMBL		
ZINC	NAME	Score	ZINC	NAME	Score
13813283	RO 08-2750	0,621	1602641	2-(2,4-diketo-7,8-di	0,655
1698103	10-Methylpyrimidol	0,571	33505288	2,5,6-trichloro-1-[[[2	0,645
66339769	5-Chloro-10-methyl	0,552	4096880	N1-(alpha-D-ribosyl	0,621
5117556	1-(beta-D-Ribofurac	0,552	33505297	2-bromo-5,6-dichlor	0,613
45322660	N-[4-[2-[4-cyano-N	0,548	88342484	1-[3-[(2R)-tetrahydri	0,6
16036162	docking.org/substance/	0,548	88342485	1-[3-[(2R)-tetrahydri	0,6
432659	N-(2-acetyl-5-methyl	0,548	59694510	8-[[[2S,3S,4R)-2,3,4	0,6
1470675	1-(3-methylbutyl)-1-	0,536	33505296	2-[2,5,6-trichloro-1-	0,594
5127789	INOSINE	0,533	33505285	1-[2,5,6-trichloro-1-	0,594
34523431	1-[3-(Dimethylamin	0,533	12354836	4-amino-7-[[[2R,3S,	0,581
8613160	INOSINE	0,533	59694487	6-methyl-8-[[[2S,3S,	0,581
8855117	INOSINE	0,533	88342761	1-[4-[(2R)-tetrahydri	0,581

Conjunto de moléculas con mejor score obtenidas a partir del software LiSiCA para encontrar moléculas similares a riboflavina.

ANEXO IV

PHARMACOPHORE SHARED FEATURES LIGAND-BASED VIRTUAL SCREENING

PHARMA COPHORE		2036848	Riboflavina	168,15										
		8215434	FAD	166,07										
ZINC	NAME	Score	Benign F.	Rings	Charges	Lipinski	<500 times	<10 accept.	<5 don.	-2<LogP<5	<5 rotatable			
1	31975681	FMNH2	135,84	NO	YES	-3	NO	465	13	5	-4,05	7		
2	98209531	(2R,3-((4aR)-7,7	125,01	YES	YES	-5								
3	98209538	(4R)-4-((4aR)-7	116,9	YES	YES	-5								
4	13543028	6,7-dihydroxy-4	116,71	YES	2	-1								
5	98208940	[(2R,3S,4S,5E)-	116,68	YES		/								
6	58650780	2'-5'-phosphonol	116,3	NO	NO	-4	NO	555	19	5	-4,05	9		
7	13543028	5,7-dihydroxy-4	116,23	YES	2	-1								
8	8536413	kanmycin B	116,2	YES		-1								
9	58650780	2'-5'-phosphonol	116,17	NO	NO	-4	NO							
10	34951251	(2R,4R,5R,6R)-	116,14	NO	NO	-5	NO	555	19	5	-4,05	9		
11	58633405	[(2R,3S,5S)-5-G	116,05	NO	NO	-2	NO	411	15	5	-4,02	11		
12	58633075	[(2R,3S,4R,5E)-	115,98	NO	NO	-2	NO	252	13	7	-4,2	7		
13	36748846	5-(2,6,8-trioxo-)	115,93	NO	NO	-3	NO	465	15	5	-3,45	12		
14	31516909	N(6)-L-dioxonyl	115,91	NO	NO	-3	NO	489	17	5	-2,09	8		
15	5875253	2-(5,6-dichloro-	115,92	YES	2	0								
16	98208793	(2R,3S,5R)-5-G	115,82	YES	4	-2								
17	15477104	[9-(2R,3R,4S,5	115,82	NO	NO	-4	NO	424	15	2	-3,39	6		
18	12502442	[9-(2R,3R,4S,5	115,74	NO	NO	-4	NO	440	14	2	-2,5	6		
19	58519942	[(2R,3S,4R,5R)	115,73	NO	NO	-1	NO	679	15	4	1,67	20		
20	58631698	(3R)-3-(7,8-dim	115,69	NO	YES	-3	NO							
21	33587197	[(2R,3S,5R)-5-G	115,64	NO	NO	-2	YES	424	12	4	-1,6	4		
22	16051847	1D-enjo-inositol	115,62	NO	NO	1	NO	445	13	12	-4,96	6		
23	58650741	9-(2R,3S,4R)-2	115,55	YES	NO	-1								
24	58650722	2-methylthio-N	115,55	NO	NO	-3	NO	525	14	5	-2,51	9		
25	12501842	3-(2,6,8-trioxo-)	115,54	NO	NO	-3	NO	437	15	5	-4,2	10		
26	58638485	[(2R,3S,5R)-5-G	115,5	NO	NO	-2	YES	259	12	4	-2,64	4		
27	12501842	3-(2,6,8-trioxo-)	115,49	NO	NO	-3	NO	437	15	5	-4,2	10		
28	36521966	Misostatin	115,48	YES	NO	-1								
29	85425702	FADH	115,45	NO	YES	-3	NO	794	24	9	-2,21	13		
30	98209518	[(2R,3R,4R,5R)	115,45	YES	NO	-4								
31	53683398	2-METHYLTHIO	115,44	NO	NO	-2	YES	461	12	3	1,99	9		
32	13543028	5,7-dihydroxy-4	115,44	YES	2	-1								
33	4096207	Adenylosuccinic	115,43	NO	NO	-4	NO	459	16	3	-3,05	9		
34	34390682	N6-ISOPENTEN	115,41	NO	NO	-2	YES	415	12	3	0,51	8		
35	33871288	(2R,3R,4S,5S,6	115,39	YES	NO	0								
36	58639065	[(2R,3S,4R,5R)	115,39	NO	NO	-4	NO	525	18	5	-4,09	8		
37	16051613	[(2R,3R,4S,5S)	115,31	NO	NO	-4	NO	338	15	4	-4,39	9		
38	40913495	(2R,3R,4S,5R)-	115,27	YES	NO	0								
39	95910503	(2R)-4-hydroxy	115,23	NO	NO	0	YES	343	10	5	-1,45	3		
40	48015812	CHEBL39774	115,18	NO	NO	-2	YES	375	13	4	-1,89	5		
41	58638382	(2S,3R,5R)-5-G	115,18	NO	2	-2	YES	387	13	4	-2,75	4		
42	95953067	(2R,3R,4S,5R)-	115,13	YES	NO	0								
43	32304140	3-(2,6,8-trioxo-)	115,12	NO	NO	-3	NO	437	15	5	-4,2	10		
44	58627004	[(2R,3S,5R)-3-h	115,1	NO	NO	-2	YES	259	12	3	-1,93	5		
45	79921150	2-(2S,4R,5R,6'	115,08	NO	NO	0	NO	383	12	9	-4,46	6		
46	87493176	methylidoprotein	115,01	NO	YES	-2	NO	731	23	10	-3,42	9		
47	98208661	(2S,3R,5R)-5-G	114,51	YES	2	-2								
48	4096223	Adenosine 2',5'	96,56	NO	NO	-4	NO	423	15	3	-2,19	6		
49	58632852	[(2R,3R,4R,5R)	96,01	NO	NO	-2	YES	423	13	3	-1,11	8		
50	95557410	(3S)-3-[(1-2-flu	85,54	YES	NO	-1								
51	73240563	2,5-dichloro-3-	85,54	YES	NO	1								
52	73240514	4-(6-methoxy-3-	75,73	YES	YES	0								
53	84955552	4-(6-methoxy-3-	75,62	YES	2	0								
54	33296967	(NE)-N-(5-(2R)-	66,05	NO	NO	1	YES	472	9	1	3,11	5		
55	71329080	3-(4-methylpiper	65,92	NO	NO	1	YES	446	7	2	4,39	6		

Lista de los 55 compuestos candidatos obtenidos a partir del ligand-based virtual screening - shared features pharmacophore.

ANEXO V

PHARMACOPHORE MERGED FEATURES LIGAND-BASED VIRTUAL SCREENING

	ZINC	NAME	Score	PHARMACOPHORE						<500 umes	<10 accept.	<5 don.	-2*LogP <5	<5 rotatable
				Benign F.	Rings	Charges	Lipinski							
56	53683566	Guanosino-3'		NO	NO	-1	YES	344	12	4	-2,07	1		
57	992259	lanillovin		YES	YES	0								
58	53683526	Phosphoatrin		NO	NO	-3	NO	521	10	6	-9,96	0		
59	2036848	Riboflavin		YES	YES	0								
60	53682927	NADH		NO	NO	-2	NO	603	21	0	-3,59	11		
61	96006025	GTA		NO	NO	-2	NO	704	27	9	-5,6	12		
62	53684299	Desamido-Nad+		NO	NO	-2	NO	652	21	6	-5,4	11		
63	58632138	Adenosino pos		NO	NO	-5	NO	652	24	5	-4,94	12		
64	85425702	FADH		NO	YES	-3	NO	704	24	9	-2,21	13		
65	98208500	[(2R,3S,4R,5		YES	NO	-2								
66	43100747	[(2R,3S,4R,5I	70,06	NO	NO	-1	YES	414	12	3	-1,41	6		
67	79161421	[(2R,3S,4R,5I	70,04	NO	NO	-2	YES	471	12	4	-0,12	4		
68	33505425	[(2R,3S,4R,5I	70	NO	NO	-2	YES	300	13	6	-1,72	4		
69	79161423	[(2R,3S,4S,5F	69,97	NO	NO	-2	YES	471	12	4	-0,12	4		
70	17105310	[(2R,3S,4S,5F	69,81	NO	NO	-1	YES	300	12	4	-1,25	5		
71	17105309	[(2R,3S,4S,5F	69,77	NO	NO	-1	YES	300	12	4	-1,25	5		
72	44482284	2-[(2S,2S,4S,5S	69,66	YES	NO	-1	YES							
73	59554496	[(2R)-2-hydrox	69,6	NO	NO	0	YES	304	7	2	1,75	6		
74	79161426	[(2R,3R,4S,5I	69,55	NO	NO	-2	YES	471	12	4	-0,12	4		
75	43009585	(2S)-2-amino-4	69,53	NO	NO	0	NO	400	11	6	-0,26	9		
76	3871613	5'-ATP-Na2	69,53	NO	NO	-4	NO	523	10	4	-3,54	0		
77	95871250	8-6-Aminohex	69,51	NO	NO	-1	NO	400	14	0	-0,99	11		
78	87515504	TNP-ATP(5-)	56,91	NO	NO	-4	NO	714	27	3	-3,09	10		
79	98208240	[(2R,3R,4S,5I	56,83	YES	NO	-3								
80														
81	31414003	[(2R,3R,4R,5I	56,82	NO	NO	-2	YES	347	11	3	-1,56	4		
82	58650238	[(2R,3S,4R,5I	56,77	NO	NO	-1	NO	417	14	6	-2,16	7		
83	13536720	[(2R,3S,4R,5I	56,76	NO	NO	-1	NO	477	14	6	-1,71	10		
84	98208707	5-(3a,4,5,6a)	56,76	YES	NO	-2								
85	64480332	N-(5S)-5-amio	56,76	NO	NO	0	NO	400	14	0	-1,71	11		
86	14967079	Not available	56,76	NO	NO	-1	NO	400	16	7	-3,22	0		
87	31976707	(2S)-2-amino-2	56,76	NO	NO	0	NO	444	14	0	-1,81	7		
88	33295982	[(2R,3S,4R,5I	56,76	NO	NO	-3	NO	514	15	3	-0,76	9		
89	58660722	2-methylblio-2	56,76	NO	NO	-3	NO	525	17	5	-2,51	9		
90	58660594	[(2R,3S,4R,5I	56,76	NO	NO	-1	YES	443	14	5	-1,19	7		
91	58626689	[(2R,3S,4R,5I	56,75	NO	NO	-1	NO	500	17	7	-0,72	15		
92	34232883	[(2R,3S,4R,5I	56,75	NO	NO	-2	YES	425	13	3	-0,49	7		
93	58639144	6-amino-9-(2I	56,75	YES	NO	0								
94	3860156	Adenosino ph	56,75	NO	NO	-2	YES	345	12	4	-1,52	4		
95	58633115	(2R,3R,4S,5R	56,75	NO	NO	-1	YES	370	10	4	-0,3	4		
96	31539932	glycyl-AMIN	56,74	NO	NO	-1	NO	403	14	6	-2,56	7		
97	98208702	[(2R,3S,4R,5I	56,74	YES	NO	-1								
98	72106195	[(2R,3S,4R,5I	56,74	NO	NO	-2	YES	413	12	3	0,25	6		
99	58650767	[(2R,3S,4R,5I	56,74	NO	NO	1	NO	401	16	11	-3,45	11		
100	34390682	N6-ISOPENTE	56,74	NO	NO	-2	YES	415	12	3	0,51	0		

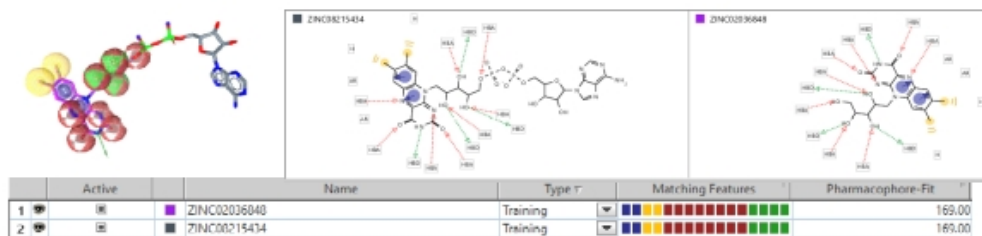
Lista de los 45 compuestos seleccionados a partir del ligand-based virtual screening – merged features pharmacophore.

Para completar los 100 candidatos se escogieron las 10 moléculas a partir de las cuales se creó el merged-pharmacophore y los 35 compuestos con mejor puntuación en el ligand merged features screening.

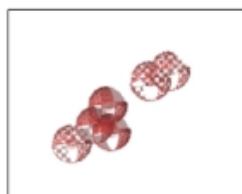
ANEXO 6

FARMACÓFOROS DESCRITOS

A) SHARED FEATURES PHARMACOPHORE



B) MERGED FEATURES PHARMACOPHORE



ZINC

53683566
391159
53683526
2036848
53682927
96006025
53684299
58632138
85425702
98208900

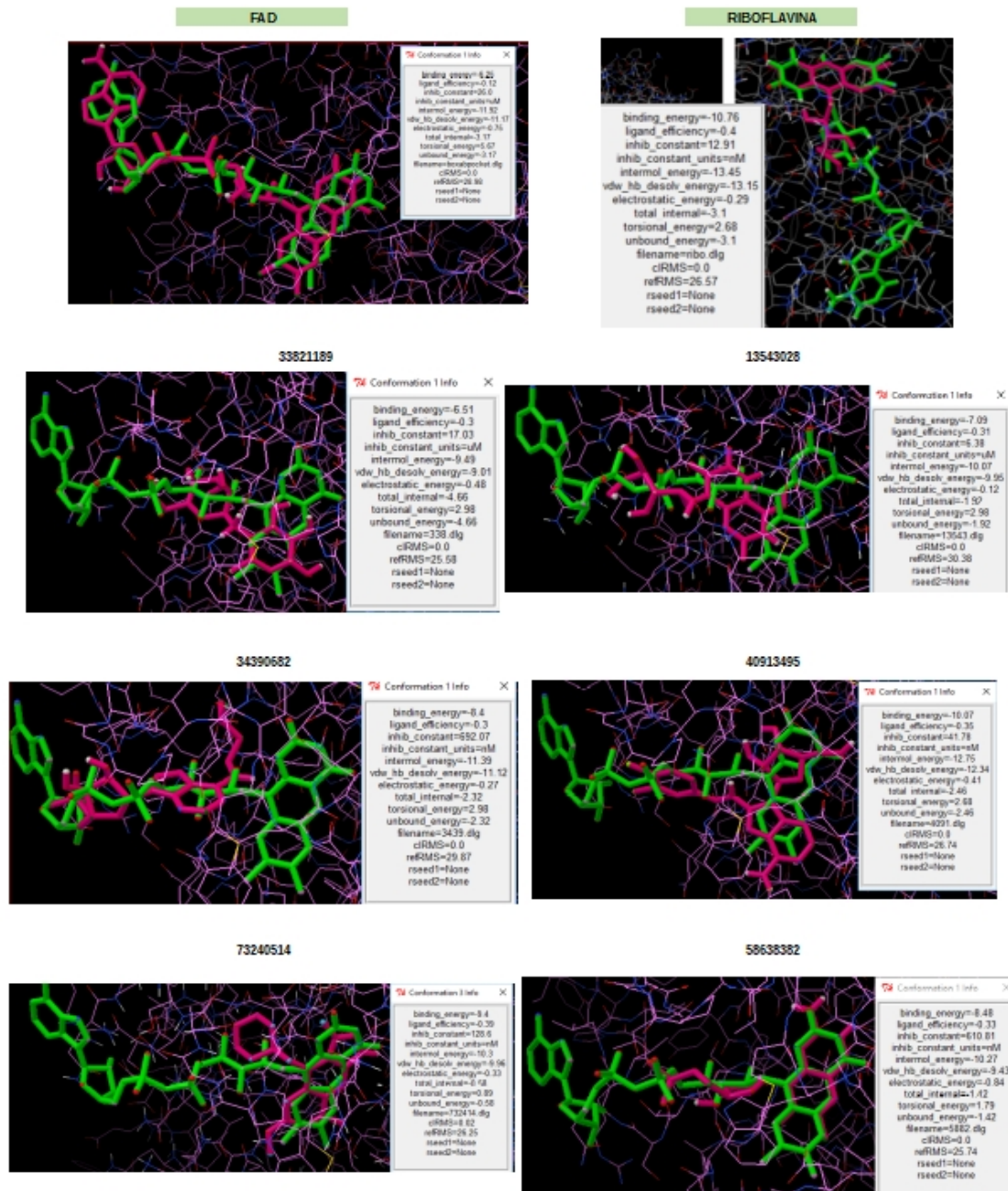
NAME

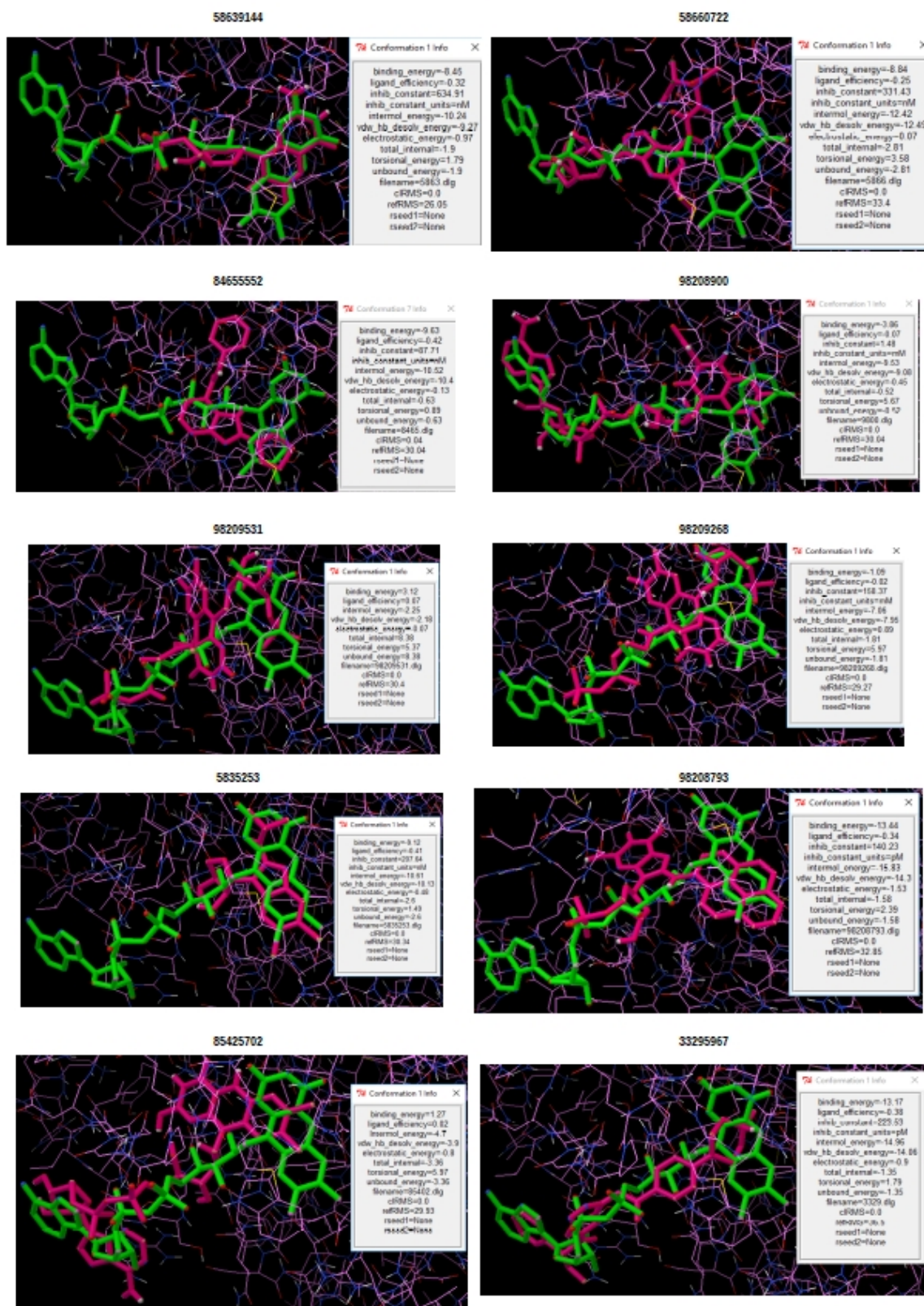
Guanosine 3',5'-monophosphate
lumiflavina
Phosphoaminophosphonic Acid-Adenylate Ester
Riboflavin
NADH
GTA
Desamido-Nad+
adenosine pentaphosphate
FADH

[*]([R,35,4R,5R]-5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl)methoxy-hydroxy-phosphoryl[*]

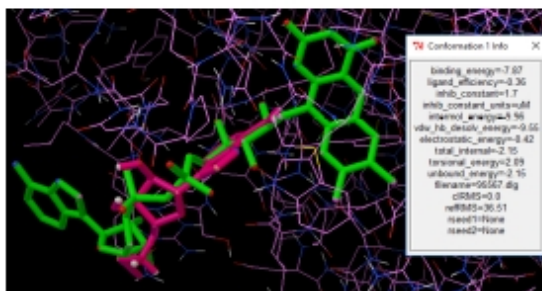
ANEXO VII

Resultados obtenidos en el protein-based virtual screening realizado con AutoDock. FAD y Riboflavina se analizaron para calibrar el programa y efectuar un docking eficiente. A continuación se muestran las diferentes moléculas candidatas. Aparece tanto la interacción dentro del bolsillo FAD (verde) como la información referente al experimento.

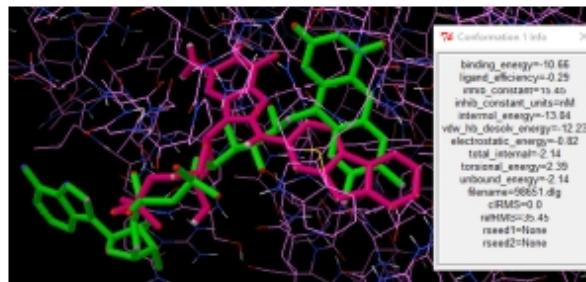




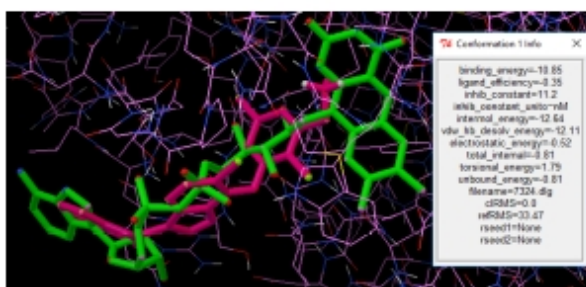
95553067



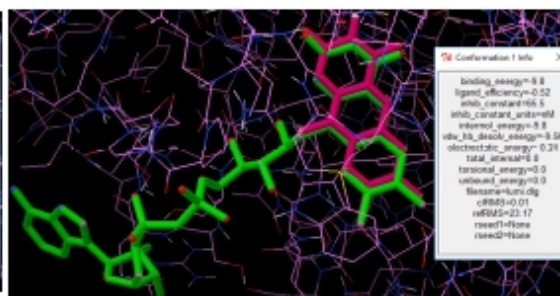
98208651



73240563



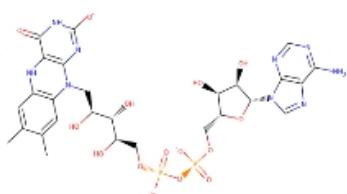
391159



ANEXO VIII

Resultados obtenidos a través de pkCSM para las moléculas de las que no se conocía actividad farmacológica demostrada.

Molecule Depiction



85425702

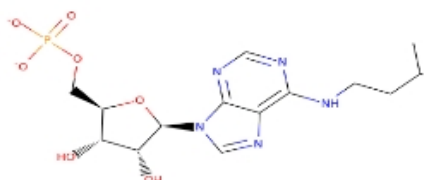
Molecule properties:

Descriptor	Value
Molecular Weight	784.549
LogP	-3.23396
#Rotatable Bonds	13
#Acceptors	23
#Donors	8
Surface Area	296.831

Property	Model Name	Predicted Value	Unit
Absorption	Water solubility	-2.892	Numeric (log mol/L)
Absorption	Caco2 permeability	-1.802	Numeric (log Papp in 10 ⁻⁶ cm/s)
Absorption	Intestinal absorption (human)	7.638	Numeric (% Absorbed)
Absorption	Skin Permeability	-2.735	Numeric (log Kp)
Absorption	P-glycoprotein substrate	Yes	Categorical (Yes/No)
Absorption	P-glycoprotein I inhibitor	No	Categorical (Yes/No)
Absorption	P-glycoprotein II inhibitor	No	Categorical (Yes/No)
Distributor	VDss (human)	0.4	Numeric (log L/kg)
Distributor	Fraction unbound (human)	0.381	Numeric (Fu)
Distributor	BBB permeability	-3.211	Numeric (log BB)
Distributor	CNS permeability	-6.375	Numeric (log PS)
Metabolism	CYP2D6 substrate	No	Categorical (Yes/No)
Metabolism	CYP3A4 substrate	No	Categorical (Yes/No)
Metabolism	CYP1A2 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C19 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C9 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2D6 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP3A4 inhibitor	No	Categorical (Yes/No)
Excretion	Total Clearance	-0.183	Numeric (log ml/min/kg)
Excretion	Renal OCT2 substrate	No	Categorical (Yes/No)
Toxicity	AMES toxicity	Yes	Categorical (Yes/No)
Toxicity	Max. tolerated dose (human)	0.431	Numeric (log mg/kg/day)
Toxicity	hERG I inhibitor	No	Categorical (Yes/No)
Toxicity	hERG II inhibitor	Yes	Categorical (Yes/No)
Toxicity	Oral Rat Acute Toxicity (LD50)	2.482	Numeric (mol/kg)
Toxicity	Oral Rat Chronic Toxicity (LOAEL)	2.751	Numeric (log mg/kg_bw/day)
Toxicity	Hepatotoxicity	No	Categorical (Yes/No)
Toxicity	Skin Sensitisation	No	Categorical (Yes/No)
Toxicity	<i>T.Pyriiformis</i> toxicity	0.285	Numeric (log ug/L)
Toxicity	Minnow toxicity	8.489	Numeric (log mM)

Pharmacokinetic Properties

Molecule Depiction



34390682

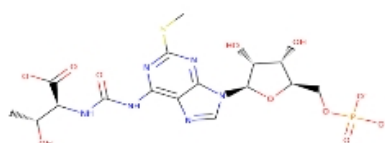
Molecule properties:

Descriptor	Value
Molecular Weight	415.343
LogP	-1.2512
#Rotatable Bonds	8
#Acceptors	12
#Donors	3
Surface Area	160.338

Property	Model Name	Predicted Value	Unit
Absorption	Water solubility	-2.473	Numeric (log mol/L)
Absorption	Caco2 perm eability	-0.263	Numeric (log Papp in 10 ⁻⁶ cm/s)
Absorption	Intestinal absorption (hum an)	55.02	Numeric (% Absorbed)
Absorption	Skin Permeability	-2.735	Numeric (log Kp)
Absorption	P-gly coprotein substrate	Yes	Categorical (Yes/No)
Absorption	P-gly coprotein I inhibitor	No	Categorical (Yes/No)
Absorption	P-gly coprotein II inhibitor	No	Categorical (Yes/No)
Distribution	VDss (human)	0.617	Numeric (log L/kg)
Distribution	Fraction unbound (hum an)	0.638	Numeric (Fu)
Distribution	BBB perm eability	-1.534	Numeric (log BB)
Distribution	CNS permeability	-3.399	Numeric (log PS)
Metabolism	CYP2D6 substrate	No	Categorical (Yes/No)
Metabolism	CYP3A4 substrate	No	Categorical (Yes/No)
Metabolism	CYP1A2 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C19 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C9 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2D6 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP3A4 inhibitor	No	Categorical (Yes/No)
Excretion	Total Clearance	0.418	Numeric (log ml/min/kg)
Excretion	Renal OCT2 substrate	No	Categorical (Yes/No)
Toxicity	AMES toxicity	No	Categorical (Yes/No)
Toxicity	Max. tolerated dose (hum an)	0.881	Numeric (log mg/kg/day)
Toxicity	hERG I inhibitor	No	Categorical (Yes/No)
Toxicity	hERG II inhibitor	No	Categorical (Yes/No)
Toxicity	Oral Rat Acute Toxicity (LD50)	2.489	Numeric (mol/kg)
Toxicity	Oral Rat Chronic Toxicity (LOAEL)	2.4	Numeric (log mg/kg_bw/day)
Toxicity	Hepatotoxicity	Yes	Categorical (Yes/No)
Toxicity	Skin Sensitisation	No	Categorical (Yes/No)
Toxicity	<i>T.Pyriiformis</i> toxicity	0.285	Numeric (log ug/L)
Toxicity	Minnow toxicity	2.805	Numeric (log m M)

Pharmacokinetic Properties

Molecule Depiction



58660722

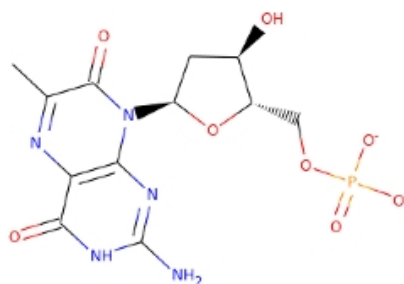
Molecule properties:

Descriptor	Value
Molecular Weight	535.408
LogP	-4.3665
#Rotatable Bonds	9
#Acceptors	16
#Donors	5
Surface Area	200.288

Property	Model Name	Predicted Value	Unit
Absorption	Water solubility	-2.891	Numeric (log mol/L)
Absorption	Caco2 permeability	-0.346	Numeric (log Papp in 10 ⁻⁶ cm/s)
Absorption	Intestinal absorption (human)	0.047	Numeric (% Absorbed)
Absorption	Skin Permeability	-2.735	Numeric (log Kp)
Absorption	P-glycoprotein substrate	Yes	Categorical (Yes/No)
Absorption	P-glycoprotein I inhibitor	No	Categorical (Yes/No)
Absorption	P-glycoprotein II inhibitor	No	Categorical (Yes/No)
Distribution	VDss (human)	0.031	Numeric (log L/kg)
Distribution	Fraction unbound (human)	0.468	Numeric (Fu)
Distribution	BBB permeability	-1.926	Numeric (log BB)
Distribution	CNS permeability	-3.916	Numeric (log PS)
Metabolism	CYP2D6 substrate	No	Categorical (Yes/No)
Metabolism	CYP3A4 substrate	No	Categorical (Yes/No)
Metabolism	CYP1A2 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C19 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C9 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2D6 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP3A4 inhibitor	No	Categorical (Yes/No)
Excretion	Total Clearance	-0.219	Numeric (log ml/min/kg)
Excretion	Renal OCT2 substrate	No	Categorical (Yes/No)
Toxicity	AMES toxicity	No	Categorical (Yes/No)
Toxicity	Max. tolerated dose (human)	0.473	Numeric (log mg/kg/day)
Toxicity	hERG I inhibitor	No	Categorical (Yes/No)
Toxicity	hERG II inhibitor	No	Categorical (Yes/No)
Toxicity	Oral Rat Acute Toxicity (LD50)	2.481	Numeric (mol/kg)
Toxicity	Oral Rat Chronic Toxicity (LOAEL)	2.544	Numeric (log mg/kg_bw/day)
Toxicity	Hepatotoxicity	Yes	Categorical (Yes/No)
Toxicity	Skin Sensitisation	No	Categorical (Yes/No)
Toxicity	<i>T.Pyriformis</i> toxicity	0.285	Numeric (log ug/L)
Toxicity	Minnow toxicity	7.017	Numeric (log mM)

Pharmacokinetic Properties

Molecule Depiction



58638382

Molecule properties:

Descriptor	Value
Molecular Weight	387.245
LogP	-3.13588
#Rotatable Bonds	4
#Acceptors	12
#Donors	3
Surface Area	144.208

Property	Model Name	Predicted Value	Unit
Absorption	Water solubility	-2.459	Numeric (log mol/L)
Absorption	Caco2 permeability	0.823	Numeric (log Papp in 10 ⁻⁶ cm/s)
Absorption	Intestinal absorption (human)	39.944	Numeric (% Absorbed)
Absorption	Skin Permeability	-2.735	Numeric (log Kp)
Absorption	P-glycoprotein substrate	No	Categorical (Yes/No)
Absorption	P-glycoprotein I inhibitor	No	Categorical (Yes/No)
Absorption	P-glycoprotein II inhibitor	No	Categorical (Yes/No)
Distribution	VDss (human)	0.57	Numeric (log L/kg)
Distribution	Fraction unbound (human)	0.522	Numeric (Fu)
Distribution	BBB permeability	-1.434	Numeric (log BB)
Distribution	CNS permeability	-3.934	Numeric (log PS)
Metabolism	CYP2D6 substrate	No	Categorical (Yes/No)
Metabolism	CYP3A4 substrate	No	Categorical (Yes/No)
Metabolism	CYP1A2 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C19 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2C9 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP2D6 inhibitor	No	Categorical (Yes/No)
Metabolism	CYP3A4 inhibitor	No	Categorical (Yes/No)
Excretion	Total Clearance	-0.181	Numeric (log ml/min/kg)
Excretion	Renal OCT2 substrate	No	Categorical (Yes/No)
Toxicity	AMES toxicity	No	Categorical (Yes/No)
Toxicity	Max. tolerated dose (human)	0.409	Numeric (log mg/kg/day)
Toxicity	hERG I inhibitor	No	Categorical (Yes/No)
Toxicity	hERG II inhibitor	No	Categorical (Yes/No)
Toxicity	Oral Rat Acute Toxicity (LD50)	2.645	Numeric (mol/kg)
Toxicity	Oral Rat Chronic Toxicity (LOAEL)	2.734	Numeric (log mg/kg_bw/day)
Toxicity	Hepatotoxicity	Yes	Categorical (Yes/No)
Toxicity	Skin Sensitisation	No	Categorical (Yes/No)
Toxicity	<i>T.Pyriformis</i> toxicity	0.285	Numeric (log ug/L)
Toxicity	Minnow toxicity	5.245	Numeric (log mM)

9. Bibliografía

- [1] A novel drug target site for medium chain acyl-CoA dehydrogenase deficiency: Implications for future drug therapy. H. Kang, J. Vockley, A.-W. Mohsen (2012) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA meeting. The American Society of Human Genetics web.
- [2] Acoplamiento inverso y mapeo de farmacóforo como herramientas para encontrar nuevos blancos farmacológicos de compuestos naturales. Maldonado Rojas W., K. Noguera Oviedo, J. Olivero (2012) Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales 36 (140): 411-420, 2012. ISSN0370-3908.
- [3] Aprendizaje Automático: Algoritmos genéticos. Dr. Alejandro Guerra Hernández Universidad Veracruzana (2004). http://di002.edv.uniovi.es/~alguero/eaac/eaac_archivos/09-10/Trabajos%20para%20evaluación/Quintarios/Artículos%20proporcionados/clase08.pdf
- [4] - Biopharmadive online. <https://www.biopharmadive.com/news/trends-rare-disease-orphan-drug-development/439866/>
- [5] Cofactors and metabolites as potential stabilizers of mitochondrial acyl-CoA dehydrogenases. Tania G, Lucas et al (2011) Biochimica et biophysica acta 1821 1658-1663.
- [6] Development of Chemical Chaperone for Medium Chain Acyl-CoA Dehydrogenase Deficiency. Al-Walid A. Mohsen. Childrens Hospital of Pittsburg (of UPMC) research projects. <http://www.chp.edu/research/areas/genetics/projects/mohsen/chemical-chaperone>
- [7] Drugbank database → <https://www.drugbank.ca>
- [8] Fatty Acid Beta-Oxidation Disorders: A Brief Review. Vijay A. Vishwanath. Ann Neurosci 2016;23:51–55 . DOI: 10.1159/000443556
- [9] In vitro & in silico: Two complementary approaches to elucidate the molecular mechanisms of Medium-Chain Acyl-CoA dehydrogenase deficiency (MCADD). Cátia Alexandra Marques Bonito Ferreira (2015) Master Thesis. Universidade of Lisboa.
- [10] Incentives for Starting Small Companies Focused on Rare and Neglected Diseases. Sean Ekins et al. Pharm Res. 2016 April ; 33(4): 809–815. doi:10.1007/s11095-015-1841-9.
- [11] Información web sobre el área 26 TFM UOC. <https://sites.google.com/a/uoc.edu/tfmbyb/home/area26>
- [12] La deficiencia de MCAD Deficiencia de acil-CoA deshidrogenasa de ácidos grasos de cadena media MCADD. ORPHANET (2017). https://www.orpha.net/data/patho/Han/Int/es/DeficienciaMCAD_Es_es_HAN_ORPHA42.pdf
- [13] Ligand-Based and Structure-Based Virtual Screening. Val Gillet University of Sheffield (2009)
- [14] Ligand based pharmacophore model development for the identification of novel antiepileptic compound. Naz Khan H, Kulsoom Sand Rashid (2012) H. Epilepsy research 98, 62-71.
- [15] Ligand-based virtual screening interface between PyMOL and LiSiCA. Athira Dilip, Samo Lešnik, Tanja Štular, Dušanka Janežič and Janez Konc (2016). J Cheminform. doi: 10.1186/s13321-016-0157-z
- [16] LigandScout is the most beautiful molecular design software available today . Alexander Varnek (2017) . Inte :ligand. http://www.inteligand.com/download/LigandScout_Articles_Brochure_2017.pdf
- [17] LimTox: a web tool for applied text mining of adverse event and toxicity associations of compounds, drugs and genes. Cañada A, Capella-Gutierrez S, Rabal O, Oyarzabal J, Valencia A, Krallinger M. Nucleic Acids Res. 2017 Jul 3;45(W1):W484-W489. doi: 10.1093/nar/gkx462.
- [18] LimTox website → <http://limtox.bioinfo.cnio.es/>
- [19] Loss of the Mitochondrial Fatty Acid β -Oxidation Protein Medium-Chain Acyl-Coenzyme A Dehydrogenase Disrupts Oxidative Phosphorylation Protein Complex Stability and Function. Sze Chern Lim et al. SCIENTIFIC REPORTS | (2018) 8:153 | DOI:10.1038/s41598-017-18530-4
- [20] Medium-Chain Acyl-Coenzyme a Dehydrogenase Deficiency. Matern, D.; Rinaldo, P.; Pagon, R. A.; Bird, T. D.; Dolan, C. R.; Stephens, K.; Adam, M. P. (1993). Reviewed 2015. GeneReviews[web version]

- [21] Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. Wilcken, B.; Hammond, J.; Silink, M. (1994). *Archives of Disease in Childhood*. 70 (5): 410–412. doi:10.1136/adc.70.5.410.
- [22] ORPHANET: El portal sobre enfermedades raras y medicamentos huérfanos. <http://www.orpha.net/consor/cgi-bin/index.php>
- [23] PDB database → <http://www.rcsb.org>
- [24] pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. Douglas E, V. Pires†, Tom L. Blundell† and David B. Ascher†. *J Med Chem*. 2015 May 14; 58(9): 4066–4072. doi:10.1021/acs.jmedchem.5b00104
- [25] pkCSM website → <http://biosig.unimelb.edu.au/pkcsml/>
- [26] Pymol → <https://en.wikipedia.org/wiki/PyMOL>
- [27] Riboflavin and beta-oxidation Flavoenzymes. Barbara Henriques, João V Rodrigues, Cláudio M.Gomes (2012). *Research gate chap.37*. Doi:10.1039/9781849734714.
- [28] ScPDB database → <http://www.bioinfo-pharma.u-strasbg.fr/scPDB/>
- [29] The Domain-Specific and Temperature-Dependent Protein Misfolding Phenotype of Variant Medium-Chain acyl-CoA. Johanna M. Jank et al (2014) <https://doi.org/10.1371/journal.pone.0093852.g001>.
- [30] Unveiling the Pathogenic Molecular Mechanisms of the Most Common Variant (p.K329E) in Medium-Chain Acyl-CoA Dehydrogenase Deficiency by in Vitro and in Silico Approaches. Bonito CA, Nunes J, Leandro J, Louro F, Leandro P, Ventura FV, Guedes RC (2016). *Biochemistry* Dec 27;55(51):7086-7098. doi: 10.1021/acs.biochem.6b00759.
- [31] What Is the Purpose of the Orphan Drug Act? Matthew Herder. *PLoS Med*. 2017 Jan; 14(1): e1002191. doi:10.1371/journal.pmed.1002191
- [32] ZINCdock database → <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360656/>