



Diseño computacional de fármacos para Niemann Pick tipo C

Daniel Sudrià Lopez
Máster en Bioinformática y Bioestadística
Bioinformática farmacéutica

Melchor Sánchez Martínez
María Jesús Marco Galindo

5 de Junio del 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/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Diseño computacional de fármacos para Niemann Pick tipo C</i>
Nombre del autor:	<i>Daniel Sudrià Lopez</i>
Nombre del consultor/a:	<i>Melchor Sánchez Martínez</i>
Nombre del PRA:	<i>María Jesús Marco Galindo</i>
Fecha de entrega (mm/aaaa):	06/2018
Titulación::	<i>Máster en Bioinformática y Bioestadística</i>
Área del Trabajo Final:	<i>Bioinformática Farmacéutica</i>
Idioma del trabajo:	<i>Castellano</i>
Palabras clave	<i>Colesterol, Niemann Pick, NPC</i>

Resumen del Trabajo (máximo 250 palabras): *Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.*

Niemann Pick-C (NP-C) es una enfermedad neurodegenerativa hereditaria que se manifiesta generalmente en la niñez. Tiene un cuadro clínico muy severo y causa la muerte a temprana edad. NP-C está causada por mutaciones en los genes de NPC1 y NPC2, proteínas lisosomales que transportan colesterol al citosól. Este trabajo final es computacional y tiene como objetivos: investigar el efecto de las mutaciones causantes de NP-C y obtener moléculas con potencial terapéutico para tratar NP-C. Los efectos de las mutaciones se investigaron mediante el acoplamiento de colesterol con modelos por homología de NPC1 mutado, con el resultado de que no tienen un efecto claro sobre la unión de colesterol. Aun así, se ha obtenido modelos para poder estudiar la actividad enzimática de NPC1 mutado. El diseño de moléculas con potencial terapéutico se realizó a partir de dianas terapéuticas: NPC1, NPC2, NPC1L1, LXR, OSBPL5; o fármacos con potencial terapéutico: ezetimibe, HP β CD, Miglustat. Se realizó un cribado virtual de ligandos contra cavidades de unión previamente descritas de las dianas; y un cribado virtual de ligandos por similitud con los fármacos. Finalmente, se analizó las propiedades farmacocinéticas (ADME-T) de los cinco mejores *hits* de cada cribado, para seleccionar los ligandos con potencial terapéutico. Del cribado de ligandos contra dianas, 2 moléculas tienen potencial terapéutico: 3,17-dimethyl-7,13,20,25-tetraoxatetraspiro[5.2.1.2.5¹⁴.2¹¹.2⁹.2⁶]pentacosan-10-one y N-(1-adamantyl)-5-oxo-1-phenylpyrrolidine-3-carboxamide. Además, del cribado por similitud se obtuvieron 4 moléculas derivadas de HP β CD con potencial terapéutico: 2,3:4,6-Di-O-cyclohexylidene-a-D-mannopyranose, (5R,13R)-4,9,12,17,21,24-hexaoxatetraspiro[4.0.5⁶.1.4¹³.0.5¹⁸.1⁵]tetracosane, (5S,13S)-4,9,12,17,21,24-hexaoxatetraspiro[4.0.5⁶.1.4¹³.0.5¹⁸.1⁵]tetracosane, (5R,13S)-

4,9,12,17,21,24-

hexaoxatetraspiro[4.0.5⁶.1.4¹³.0.5¹⁸.1⁵]]tetracosane.

Mayoritariamente se cumplieron los objetivos del trabajo al haber encontrado ligandos con potencial terapéutico contra NP-C.

Abstract (in English, 250 words or less):

Niemann Pick type C (NP-C) is a hereditary neurodegenerative disease with an early onset and severe symptomatology that causes death in few years, usually during childhood. It is caused by mutations on the genes of NPC1 and NPC2, which are lysosomal cholesterol transporters. The aims of this study were to computationally evaluate the effect of such mutations on cholesterol binding, and computationally design potential drugs for the treatment of NP-C.

NPC1 mutants were generated via homology modelling and their capability to bind cholesterol was assessed with docking experiments. Mutations were shown to have little effect on cholesterol binding, however, the models can be used as a tool to investigate the enzymatic activity of mutated NPC1. The potential-drug design was performed based on therapeutic targets: NPC1, NPC2, NPC1L1, LXR, OSBPL5; and preexisting drugs against NP-C: ezetimibe, HP β CD, Miglustat. In both cases the computational design started with a virtual screening of ligands against defined cavities of the targets, or based on similarity with the drugs. The screenings were followed by a pharmacokinetic (ADME-T) analysis of the top 5 *hits*. The project resulted in 2 ligands, from the screening against targets, with therapeutic potential: 3,17-dimethyl-7,13,20,25-

tetraoxatetraspiro[5.2.1.2.5¹⁴.2¹¹.2⁹.2⁶]]pentacosan-10-one and *N-(1-adamantyl)-5-oxo-1-phenylpyrrolidine-3-carboxamide*. And also, 4 ligands, from the similarity screening of HP β CD, with therapeutic potential: 2,3:4,6-Di-O-cyclohexylidene- α -D-mannopyranose,

(5R,13R)-4,9,12,17,21,24-*hexaoxatetraspiro[4.0.5⁶.1.4¹³.0.5¹⁸.1⁵]]tetracosane,* (5S,13S)-4,9,12,17,21,24-

hexaoxatetraspiro[4.0.5⁶.1.4¹³.0.5¹⁸.1⁵]]tetracosane, (5R,13S)-4,9,12,17,21,24-

hexaoxatetraspiro[4.0.5⁶.1.4¹³.0.5¹⁸.1⁵]]tetracosane.

In conclusion, the objectives of the study were mostly met having defined molecules with high probability of having therapeutic effects against NP-C.

Índice

1. Introducción.....	1
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
1.4.1 Tareas.....	3
1.4.2 Calendario.....	4
1.4.3 Hitos.....	5
1.5 Breve resumen de productos obtenidos	7
1.6 Breve descripción de los otros capítulos de la memoria.....	7
2. Materiales y Métodos	8
2.1 Identificación de <i>targets</i>	8
2.1.1 Mutantes de NPC1.....	8
2.1.2 Dianas terapéuticas	8
2.1.3 Ligandos	8
2.2 Modelaje 3D	8
2.3 <i>Alanine Scanning</i>	10
2.4 Análisis de Cavidades	10
2.4.1 <i>fpocket</i>	10
2.4.2 Bibliografía	12
2.5 Cribado de ligandos contra dianas	12
2.6 Cribado de ligandos por similitud	13
2.7 Acoplamiento de ligando a diana	14
2.8 Farmacocinética	15
2.8.1 <i>pkCSM</i>	15
2.8.2 <i>SwissADME</i>	20
3. Resultados	24
3.1 Identificación bibliográfica de objetivos de estudio.....	24
3.1.1 Mutaciones de NPC1	24
3.1.2 Dianas terapéuticas	25
3.1.3 Moléculas con efecto terapéutico.....	27
3.3 Cavidades	30
3.4 Cribado virtual de ligandos contra dianas.....	32
3.5 Cribado virtual de ligandos por similitud	33
3.6 Acoplamiento.....	35
3.6.1 Acoplamiento de colesterol a variantes de NPC1	35
3.6.2 Acoplamiento para evaluar el cribado por similitud.....	37
3.7 Farmacocinética	39
3.7.1 Cribado de ligandos contra dianas.....	39
3.7.2 Cribado de ligandos por similitud.....	44
4. Discusión.....	49
4.1 Bloque 1:	49
4.2 Bloque 2:	50
4.3 Bloque 3:	53
5. Conclusiones.....	55
6. Glosario	57
7. Bibliografía	58
8. Anexos	61

Lista de figuras

Figura 1: Calendario de realización de los bloques	4
Figura 2: Calendario de redacción de la memoria	5
Figura 3: Diagrama de Gantt de la realización de los bloques	6
Figura 4: Ejemplo de modelaje por homología con NPC1	9
Figura 5: Detección computacional de cavidades	11
Figura 6: Visualización de la unión de un ligando con NPC2	13
Figura 7: Ejemplo de los resultados obtenidos del cribado por similitud a partir de ezetimibe	14
Figura 8: Buscador principal de <i>PubChem</i>	15
Figura 9: Ejemplo de obtención de SMILES a partir del <i>PubChem SID</i>	16
Figura 10: <i>Output</i> de <i>SwissSimilarity</i>	20
Figura 11: Esquema <i>Boiled Egg</i>	21
Figura 12: Análisis químico de la molécula con <i>SwissADME</i>	22
Figura 13: Ejemplo de <i>output</i> de <i>SwissADME</i>	23
Figura 14: Mutaciones más comunes de NPC1	24
Figura 15: Modelos de NPC1 mutado generados	25
Figura 16: Gráfico de barras de los aminoácidos del sitio de unión ordenados según su importancia en el enlace	29
Figura 17: Resultado de los cribados por similitud	34
Figura 18: Acoplamiento de colesterol a las distintas versiones de NPC1	36
Figura 19: Sitio de unión de distintos ligandos a NPC1L1m2	38
Figura 20: Representación de los 4 ligandos similares a HP β CD con potencial terapéutico	48

Lista de tablas

Tabla 1: Características y parámetros de calidad de los modelos por homología generados	24
Tabla 2: Relación de estructuras cristalografiadas	26
Tabla 3: Los distintos modelos generados con sus estimadores de calidad	26
Tabla 4: Los tres ligandos seleccionados para el cribado por similitud	27
Tabla 5: Listado de cavidades definidas en la bibliografía	30
Tabla 6: cribados virtuales de ligandos contra dianas realizados	32
Tabla 7: Energía de unión de cada ligando con la diana expresada en kcal/mol	33
Tabla 8: Energías de unión (dG) de colesterol con NPC1 WT y NPC1 con las cinco mutaciones	35
Tabla 9: Energía de unión de ezetimibe y las 5 moléculas resultantes del cribado contra el sitio de unión de ezetimibe en NPC1L1m2	37
Tabla 10: Ligandos del cribado contra dianas que cumplen con los requisitos ADME-T especificados	40
Tabla 11: Propiedades farmacocinéticas de los ligandos del cribado virtual contra dianas que pasaron los requerimientos establecidos	40
Tabla 12: Ligandos del cribado contra dianas que cumplen con los requisitos ADME del segundo filtro	41
Tabla 13: Propiedades de los ligandos que cumplen las condiciones establecidas en el segundo filtro	42
Tabla 14: Ligandos del cribado contra NPC1L1 que cumplen los requerimientos del filtro	43
Tabla 15: Propiedades ADME-T de las moléculas del cribado contra NPC1L1 con mayor potencial terapéutico.	44
Tabla 16: Farmacocinética de los ligandos del cribado de HP β CD permeables a la BBB	45
Tabla 17: Análisis farmacocinético de los <i>hits</i> del cribado en base a ezetimibe.	47

1. Introducción

1.1 Contexto y justificación del Trabajo

Niemann Pick es una enfermedad rara de transmisión hereditaria autosómica recesiva, con un gran componente neurodegenerativo: disartria, ataxia, disfagia, distonias; y complicaciones hepáticas y del bazo: hepatoesplenomegalia (1). Tiene una prevalencia de 1/150.000 nacimientos, mayoritariamente se manifiesta durante la niñez y causa un rápido deterioro de los pacientes derivando en la muerte a los pocos años (2).

A nivel molecular Niemann Pick está causada por un defecto en el transporte intracelular de lípidos y se divide en cuatro subgrupos en función de la enzima o vía metabólica que se vea más afectada. Este proyecto se centra en Niemann Pick tipo C (NP-C) (3). Este subtipo está causado por mutaciones en los genes de NPC1 y NPC2 (4), proteínas encargadas de transportar colesterol desde el lumen lisosomal hasta el citosól (5). El bloqueo de esta vía impide que las células incorporen colesterol del espacio extracelular (proveniente de LDLs) al *pool* de colesterol activo en el citosól. Además, la acumulación de colesterol esterificado en los endosomas tardíos / lisosomas afecta su funcionamiento y dificulta el procesamiento de otros lípidos como la esfingomielina (6).

Aunque NP-C es una enfermedad muy grave y con origen molecular definido, no hay terapias específicas ni eficientes para su tratamiento. Es por esto, que en este trabajo se plantea hacer un cribado virtual para buscar moléculas con un potencial efecto terapéutico en NP-C, partiendo de dianas terapéuticas y fármacos contra NP-C.

1.2 Objetivos del Trabajo

El objetivo principal de este TFM es el diseño computacional de moléculas con potencial terapéutico para tratar NP-C basándose en dianas terapéuticas y fármacos con un efecto probado.

Objetivos generales:

1. Estudiar el efecto de las mutaciones causantes de NP-C: modelar la estructura 3D de las mutaciones más comunes de NPC1 y NPC2.
2. Cribado de ligandos contra dianas: con información de la conformación 3D de dianas terapéuticas, diseñar moléculas que puedan modificar su actividad.
3. Cribado de ligandos por similitud: conociendo la estructura de fármacos usados, o en estudio, identificar moléculas con potencial terapéutico.

Objetivos específicos:

1. Estudiar el efecto de las mutaciones causantes de NP-C:
 - 1.1 Obtener estructura 3D de NPC1 y NPC2.
 - 1.2 Definir residuos de NPC1 y NPC2 importantes para su función: *Alanine scanning* en las estructuras 3D de NPC1 y NPC2.
 - 1.3 Generar modelos 3D de homología de NPC1 y NPC2 con las mutaciones.
 - 1.4 Comparar el efecto de las distintas mutaciones en la interacción proteína-ligando.

2. Cribado de ligandos contra dianas:
 - 2.1 Obtener coordenadas de cavidades en las dianas (sitios de unión).
 - 2.2 Desarrollar *hits* a partir de un cribado de ligandos contra las cavidades de las dianas.
 - 2.3 Ordenar los ligandos validados según su farmacocinética (ADME: Administración, Distribución, Metabolismo y Excreción) y toxicidad, valorando especialmente su capacidad esperada de cruzar la barrera hemato-encefálica.

3. Cribado de ligandos por similitud:
 - 3.1 Encontrar ligandos similares a fármacos contra NP-C.
 - 3.2 Validar los ligandos con mayor índice de similitud mediante *docking* con su diana.
 - 3.3 Ordenar los ligandos validados según su farmacocinética y toxicidad, valorando especialmente su capacidad esperada de cruzar la barrera hemato-encefálica.

1.3 Enfoque y método seguido

Como se ha definido en apartados anteriores, este trabajo fin de máster consiste en diseñar fármacos para NP-C mediante herramientas computacionales. El proceso se trabaja con distintos *webservers* dado que se requiere un gran poder computacional solo disponible en centros de supercomputadoras.

El proyecto se divide en tres bloques. El primer bloque consiste en hacer modelos por homología de NPC1 y NPC2 con mutaciones causantes de patología para luego ver su efecto sobre la interacción con ligandos. El modelaje por homología se hace a partir de la estructura 3D de NPC1 y NPC2 mediante *SwissModel*, y para posterior valoración de los efectos se usa *MitAutoDock* para acoplar los modelos a ligandos.

Los bloques restantes tratan sobre el diseño de moléculas con potencial efecto terapéutico, pero desde dos enfoques distintos. Por un lado, se diseñan moléculas con potencial terapéutico a partir de dianas terapéuticas definidas por una búsqueda bibliográfica sobre NP-C (*PubMed*) y cuya estructura 3D está definida y accesible en bases de datos (*PDB*, *RSCPD9*). Primero se definen cavidades (sitios de unión) en la estructura de las dianas mediante *fpocket*, *RaptorX*, *CASTp*; y basándose en información extraída de la bibliografía. Después, se hace un cribado virtual de ligandos contra las

cavidades mediante *MitOpenScreen*. Finalmente se analiza la farmacocinética y toxicidad de los *hits* resultantes del cribado con *pkCSM*.

El último bloque consiste en diseñar moléculas con potencial terapéutico por similitud con fármacos para NP-C. Se parte de fármacos con efecto contra NP-C descritos en la bibliografía (*PubMed*) y en bases de datos (*DrugBank*, *PubChem*). La primera fase consiste en encontrar ligandos similares. Para realizar el cribado se usa *SwissSimilarity* con la base de datos ZINC. Los *hits* con mejores características se validan con *MitAutoDock* y su diana terapéutica (estructura 3D obtenida de PDB). Finalmente se analiza la farmacocinética y de los *hits* resultantes del cribado con *SwissADME*, y la toxicidad con *pkCSM*.

Los 3 bloques son independientes entre sí. El proyecto se ha planteado con esta estructura para poder avanzar en los distintos bloques de forma simultánea si es necesario y así facilitar el cumplimiento del calendario.

1.4 Planificación del Trabajo:

1.4.1 Tareas

Enumeración de las tareas a realizar en base a los objetivos marcados:

- 1.1 Obtener la estructura 3D de NPC1 y NPC2
- 1.2 Realizar un *alanine scanning* de NPC1 y NPC2
- 1.3 Generar modelos por homología de NPC1 y NPC2 con mutaciones causantes de patología
- 1.4 Hacer *docking* de ligandos con los modelos de NPC1 y NPC2 mutados

- 2.1 Obtener la estructura 3D de las dianas terapéuticas
- 2.2 Definir cavidades de la estructura de las dianas susceptibles de ser sitios de unión
- 2.3 Realizar un cribado virtual basado en las cavidades de las dianas
- 2.4 Hacer un listado de los *hits* encontrados según su energía de enlace
- 2.5 Evaluar la farmacocinética (ADME) y toxicidad de los cinco primeros *hits* de la lista
- 2.6 Valorar la capacidad de cruzar la barrera hemato-encefálica de las moléculas con potencial terapéutico con mejores características

- 3.1a Obtener los fármacos conocidos en formato SMILES
- 3.2 Realizar un cribado virtual para encontrar ligandos similares
- 3.3 Ordenar los ligandos obtenidos por similitud
- 3.4a Validar los *hits* con mayor índice de similitud mediante acoplamiento con su diana (si está descrita)
- 3.4b Evaluar la farmacocinética (ADME) y toxicidad de los *hits* con mayor índice de similitud
- 3.5 Valorar la capacidad de cruzar la barrera hemato-encefálica de las moléculas con potencial terapéutico con mejores características

A estas tareas hay que añadir las PEC de la asignatura:

- PEC 1: Plan de Trabajo (6 marzo a 19 marzo)
- PEC 2: Desarrollo del Trabajo fase 1 (20 marzo a 23 abril)
- PEC 3: Desarrollo del Trabajo fase 2 (24 abril a 21 mayo)
- PEC 4: Cierre de la memoria (22 mayo a 5 junio)
 - Materiales y Métodos
 - Introducción
 - Resultados
 - Discusión y Conclusiones
 - Ficha técnica
 - Bibliografía
 - Glosario
 - Abstract
- PEC 5a: Elaboración de la presentación (6 junio a 13 junio)
- PEC 5b: Defensa pública (14 junio a 25 junio)

1.4.2 Calendario

Los bloques se trabajan siguiendo el calendario expuesto a continuación (Fig 1). Los elementos no numerados son hitos integrados para evaluar el desarrollo del proyecto en sus distintas fases.

Name	Begin date	End date
◦ Bloque 1: NPC1 y NPC2 mutados	3/20/18	4/15/18
◦ 1.1 Obtener la estructura 3D de NPC1 y NPC2	3/20/18	3/20/18
◦ 1.2 Realizar un alanine scanning	3/21/18	3/25/18
◦ 1.3 Generar modelos por homología con mutaciones	3/26/18	4/1/18
◦ 1.4 Calcular cavidades de las estructuras mutadas	4/2/18	4/8/18
◦ 1.5 Hacer docking de ligandos con los modelos mutados	4/9/18	4/15/18
◦ Bloque 2: structure-based CADD	3/20/18	5/2/18
◦ 2.1 Obtener la estructura 3D de las dianas terapéuticas	3/20/18	3/20/18
◦ Necesidad de trabajar en paralelo	4/12/18	4/12/18
◦ 2.2 Calcular cavidades de la estructura de las dianas	4/4/18	4/10/18
◦ 2.3 Realizar un cribado virtual basado en las cavidades de las dianas	4/11/18	4/17/18
◦ 2.4 Hacer un listado de los hits encontrados	4/18/18	4/18/18
◦ 2.5 Evaluar ADME y toxicidad de los fármacos potenciales	4/19/18	4/25/18
◦ Informe desarrollo fase 1	4/23/18	4/23/18
◦ 2.6 Valorar la capacidad de cruzar BBB de los fármacos potenciales	4/26/18	5/2/18
◦ Bloque 3: ligand-based CADD	4/20/18	5/26/18
◦ 3.1a Obtener la estructura 3D de fármacos conocidos	4/20/18	4/20/18
◦ 3.1b Calcular la estructura 3D de los ligandos a partir de la forma 2D	4/20/18	4/26/18
◦ 3.2 Realizar un cribado virtual para encontrar ligandos similares	4/27/18	5/3/18
◦ 3.3 Ordenar los ligandos obtenidos por similitud	5/4/18	5/4/18
◦ 3.4 Validar los leads con mayor índice de similitud via docking	5/5/18	5/11/18
◦ 3.5 Hacer un listado de los hits encontrados	5/12/18	5/12/18
◦ 3.6 Evaluar ADME y toxicidad de los fármacos potenciales	5/13/18	5/19/18
◦ Informe desarrollo fase 2	5/21/18	5/21/18
◦ 3.7 Valorar la capacidad de cruzar BBB de los fármacos potenciales	5/20/18	5/26/18

Fig 1. Calendario de realización de los bloques.

En paralelo se trabaja la redacción de la memoria según el calendario de la figura 2:

◦ Preparación de la memoria y la presentación	3/10/18	6/12/18
◦ Escritura del plan de trabajo	3/10/18	3/19/18
◦ Formalización del plan de trabajo	3/19/18	3/19/18
◦ Escritura del informe de la fase 1	4/17/18	4/23/18
◦ Escritura del informe de la fase 2	5/15/18	5/21/18
◦ Escritura de los Materiales y Métodos	5/1/18	5/7/18
◦ Escritura de la Introducción	5/8/18	5/14/18
◦ Escritura de los Resultados	5/15/18	5/21/18
◦ Escritura de la Discusión y las Conclusiones	5/22/18	5/28/18
◦ Escritura de la ficha técnica	5/30/18	5/30/18
◦ Escritura de la Bibliografía	5/30/18	5/30/18
◦ Escritura del Glosario	5/30/18	5/30/18
◦ Escritura del Abstract	5/31/18	6/4/18
◦ Entrega memoria	6/5/18	6/5/18
◦ Preparación de la presentación	6/6/18	6/12/18
◦ Entrega diapositivas presentación	6/13/18	6/13/18
◦ Defensa pública	6/14/18	6/24/18

Fig 2. Calendario de redacción de la memoria.

1.4.3 Hitos

Los hitos del TFM se han hecho coincidir mayoritariamente con las fechas de entrega de las PEC de la asignatura. La elección de estas fechas es lógica ya que separan el TFM en fases distintivas e inmóviles que requieren la entrega de un informe del proceso.

Hito 1: Formalización del plan de trabajo (19 marzo)

Hito 2: Valoración de la necesidad de trabajar sobre distintos objetivos simultáneamente (10 abril)

Hito 3: Informe del desarrollo de la fase 1: bloques 1 y 2 del TFM (23 abril)

Hito 4: Informe del desarrollo de la fase 2: bloques 2 y 3 del TFM (21 mayo)

Hito 5: Redacción de la memoria (5 junio)

Hito 6: Elaboración de la presentación (13 junio)

Hito 7: Defensa pública (25 junio)

Diagrama de Gantt de la calendarización de los bloques:

Las tareas del diagrama de Gantt (Fig 3) están ordenadas según el calendario presentado en el apartado 1.4.2

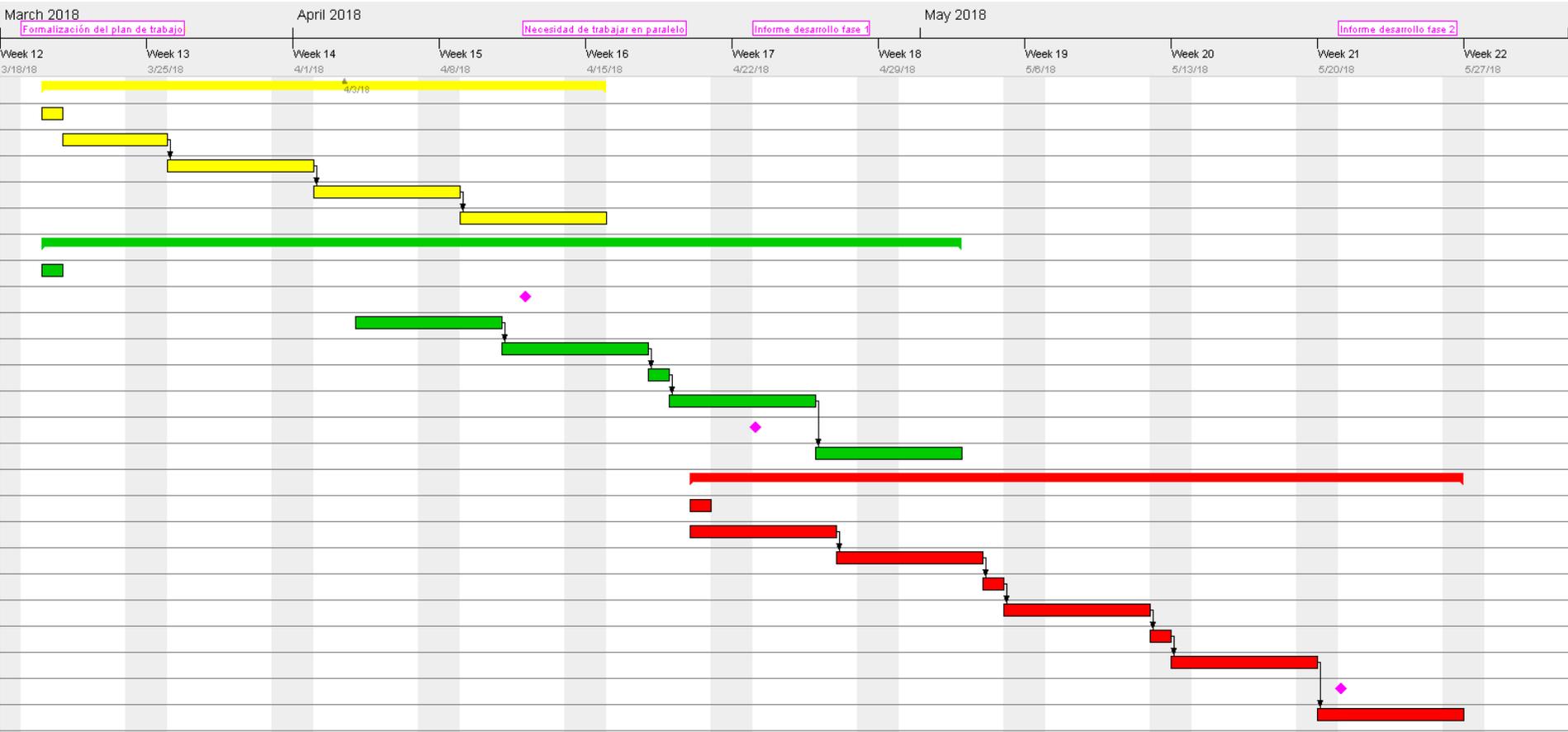


Fig 3. Diagrama de Gantt de los bloques. Las tareas del bloque 1 están representadas en color amarillo, las del bloque 2 en verde y las del bloque 3 en rojo. Los hitos están designados por rombos rosados.

1.5 Breve resumen de productos obtenidos

- Memoria
- Presentación
- Hoja de cálculo con los resultados del cribado de ligandos contra diana
- Hoja de cálculo con la descripción de los cinco *hits* con mayor puntuación de los cribados contra diana
- Hoja de cálculo con los resultados del análisis ADME-T con pkCSM de los *hits* del cribado contra dianas
- Hoja de cálculo con los resultados del cribado de ligandos por similitud
- Documento con el análisis farmacológico con *SwissADME* de los *hits* estudiados del cribado por similitud
- Hoja de cálculo con los resultados del análisis de toxicidad con pkCSM de los *hits* estudiados del cribado por similitud
- Hoja de cálculo con las energías del acoplamiento de colesterol a mutantes de NPC1
- Hoja de cálculo con las energías del acoplamiento de ezetimibe y ligandos derivados contra NPC1L1

1.6 Breve descripción de los otros capítulos de la memoria

2. Materiales y Métodos: descripción detallada de las herramientas empleadas para la resolución de este trabajo.

3. Resultados: enumeración y descripción de los resultados obtenidos para cada objetivo y los resultados finales de cada bloque.

4. Discusión: análisis del pipeline seguido en el trabajo y los resultados obtenidos.

5. Conclusión: breve síntesis de las ideas más importantes obtenidas del trabajo fin de máster.

6. Glosario: definición de los términos y acrónimos más relevantes utilizados dentro de la memoria.

7. Bibliografía: listado de la bibliografía consultada y citada en la memoria.

8. Anexos: listado de archivos de soporte del trabajo y resultados intermedios.

2. Materiales y Métodos

2.1 Identificación de *targets*

2.1.1 Mutantes de NPC1

En primera instancia se recurrió a la búsqueda bibliográfica de artículos para definir las mutaciones causantes de NP-C más comunes. Finalmente se usaron las cinco mutaciones más comunes de NPC1 en la [base de datos de mutaciones de NPC1 del Universitäts Klinikum Tübingen](#) (7), porqué aglutina información sobre mutaciones recogida de todas las fuentes.

2.1.2 Dianas terapéuticas

Mediante una búsqueda bibliográfica en *Pubmed* con términos como *Niemann Pick*, *NPC*, *colesterol transport*; se identificaron proteínas involucradas en el transporte de colesterol y con potencial de ser dianas terapéuticas para el tratamiento de NP-C. Los candidatos encontrados se investigaron en bases de datos como UniProt y RCSB. De estas bases de datos se extrajo más información de las interacciones entre las distintas proteínas y posibles dianas terapéuticas.

2.1.3 Ligandos

El mismo proceso de investigación bibliográfica se usó para definir fármacos con potencial terapéutico en NP-C a partir de los cuales diseñar moléculas con mayor potencial terapéutico. En esta búsqueda se hizo incidencia en estudios clínicos y de otras enfermedades relacionadas con el transporte de colesterol.

2.2 Modelaje 3D

Para poder realizar cribados virtuales se necesita un modelo 3D de la proteína a estudiar. Para modelar la estructura 3D de las dianas a estudiar de las que no se conocía la estructura se usó [SwissModel](#) (8,9).

SwissModel es un software en línea que, a partir de la secuencia de aminoácidos en formato *fasta*, construye modelos 3D usando como molde proteínas con secuencias similares. Cada modelo calculado incluye el archivo *pdb* resultante y un informe sobre la calidad del modelo. La calidad del modelo depende de la calidad de la estructura 3D usada como *template* y su cobertura y similitud con la secuencia a modelar.

Durante el proceso de modelaje, se puede hacer un paso intermedio para seleccionar los moldes que se desee (Fig 4A). En este trabajo, se ha usado

esta herramienta secundaria para diseñar modelos con una menor calidad en general, pero con una cobertura más completa y detallada de zonas de interés en la secuencia (principalmente, potenciales centros activos y/o de unión a fármacos). En caso necesario se ha usado más de un modelo por diana (Fig 4B).

A

Summary Templates **32** Models **1**

Template Results

Templates Quaternary Structure Sequence Similarity Alignment of Selected Templates More ▾

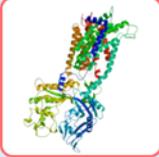
Sort	Name	Title	Coverage	GMQE	QSQE	Identity	Method	Oligo State	Ligands
<input type="checkbox"/>	3jd8.1.A	Niemann-Pick C1 protein		0.95	-	99.92	EM	monomer ✓	4 x NAG-NAG [Ⓞ] , 9 x NAG [Ⓞ] , 1 x CLR [Ⓞ]
<input checked="" type="checkbox"/>	5u74.1.A	Niemann-Pick C1 protein		0.72	-	100.00	X-ray, 3.3Å	monomer ✓	13 x NAG [Ⓞ] , 1 x MES [Ⓞ] , 1 x MAN [Ⓞ]
<input type="checkbox"/>	5u73.1.A	Niemann-Pick C1 protein		0.72	-	100.00	X-ray, 3.3Å	monomer ✓	13 x NAG [Ⓞ] , 1 x MAN [Ⓞ]
<input type="checkbox"/>	2hqc.1.A	Acriflavine resistance protein B		0.25	0.11	11.27	X-ray, 3.6Å	homo-trimer ✓	None

B

Summary Templates **32** Models **1**

Model Results

Order by: GMQE ▾



Model 01

Oligo-State Monomer

Ligands 1 x MES[Ⓞ], 1 x NAG[Ⓞ] ^

1 x 2-(N-MORPHOLINO)-ETHANESULFONIC ACID
Ligand 17 in contact with: Chain A : W381, L613, S617, D620, L684, I685, V1152, F1221, Y1225

1 x N-ACETYL-D-GLUCOSAMINE
Ligand 6 in contact with: Chain A : K417, P430, D551, N554, N556, N557

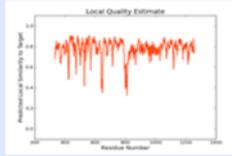
GMQE 0.67

QMEAN -3.17

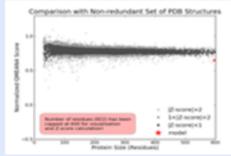
Global Quality Estimate

QMEAN		-3.17
Cβ		-1.55
All Atom		-0.88
Solvation		1.06
Torsion		-2.95

Local Quality Estimate



Comparison



Template	Seq Identity	Coverage
5u74.1.A	100.00%	

Description

Niemann-Pick C1 protein

Fig 4. Ejemplo de modelaje por homología con NPC1. **A** Selección de *template*. **B** Modelo resultante con: esquema de la estructura, ligandos incluidos y estimadores de calidad del modelo

El mismo proceso de una forma más simplificada se usó para generar modelos de NPC1 mutado. Se introdujo la variación de aminoácido de la mutación en la secuencia *fasta* de NPC1 y después se usó el archivo en formato *fasta*

generado para realizar un modelo 3D con la misma estructura de NPC1-5u74 como molde.

2.3 Alanine Scanning

Alanine scanning es una técnica que consiste en cambiar los aminoácidos de un sitio de unión uno a uno por alanina (aminoácido neutro) y re-evaluar la interacción con el ligando, para estimar la importancia de dicho aminoácido en la interacción (10). Los aminoácidos con menor energía de unión forman enlaces más estables y son más importantes para la unión. En este trabajo se usó [ABScan](#) (11) para estudiar qué aminoácidos de NPC1 y NPC2 eran más importantes para la unión a colesterol. *ABScan* requiere un archivo *pdb* que incluya la proteína, el sitio de unión al ligando y el ligando. Básicamente un *pdb* con la proteína y el ligando unidos. Además, se debe establecer el límite de tolerancia de la distancia de unión. Este valor determina el área alrededor del ligando en que se realizará el análisis. A mayor distancia, mayor área y más aminoácidos a analizar. En este proyecto se dejó el límite en el valor predeterminado: 4.5Å.

En segunda instancia se usó el *alanine scanning* para determinar los aminoácidos de cavidades activas de las proteínas. El resultado se usó para dirigir el cribado de ligandos contra sitios de unión de las dianas con evidencia experimental.

2.4 Análisis de Cavidades

El cribado virtual de ligandos se hace contra cavidades de la estructura de la diana. Para definir estas cavidades hay distintos métodos. A continuación, se explican los usados para este proyecto.

2.4.1 *fpocket*

[Fpocket](#) (12) es un *software* en línea de *Mobylye@RPBS* con el que se pueden calcular cavidades de una proteína (archivo *pdb*). Al principio del proyecto se usó *fpocket* para determinar cavidades para todas las dianas.

Fpocket genera varios archivos con las esferas alfa que conforman las cavidades en formato *pdb* o *pqr*, así como un *applet* para visualizar las cavidades en línea y scripts para visualizarlas en *PyMol* y *VMD*. De acuerdo a los *benchmarks*, *fpocket* detecta el centro activo de una enzima entre las 3 cavidades con mayor puntuación en un 97% de los casos. Por lo tanto, se usaron 3 cavidades por diana en los pasos subsiguientes, para asegurar que se trabajaba con el sitio activo de las dianas en, al menos, 1/3 cribados.

Para usar las cavidades derivadas del servidor, se escribió un script en R que extrae las coordenadas del centro de la cavidad y su límite [ANEXO: Cavidades_definidas_coordenadas.pdf], o los aminoácidos que la conforman. [ANEXO: Cavidades_definidas_aminoacidos.pdf].

Aun así, después de optimizar todo el proceso hubo un problema con el cribado de ligandos. Las cavidades generadas por *fpocket* eran demasiado grandes y el cribado generaba resultados inespecíficos. Añadiendo que el *applet* de visualización online no funcionaba, se decidió probar otras herramientas de generación de cavidades.

A parte de *fpocket* se probó [RaptorX](#) (13,14) y [CASTp](#) (15). Ambos programas encontraron un número menor de cavidades en las dianas. La posibilidad de ver la estructura de las cavidades permitió valorar la relevancia de las cavidades generadas. Un gran número de cavidades se sitúan al lado de cadenas secundarias, o en la zona transmembrana de la proteína (Fig 5A). Además, dos de las cinco dianas estudiadas son proteínas de membrana involucradas en el transporte de colesterol y los *softwares* detectaban todo su interior como una gran cavidad sin distinguir sitios de unión específicos (Fig 5B). Estas desventajas llevaron a plantear la definición de cavidades de una forma distinta como se explica en el siguiente apartado.

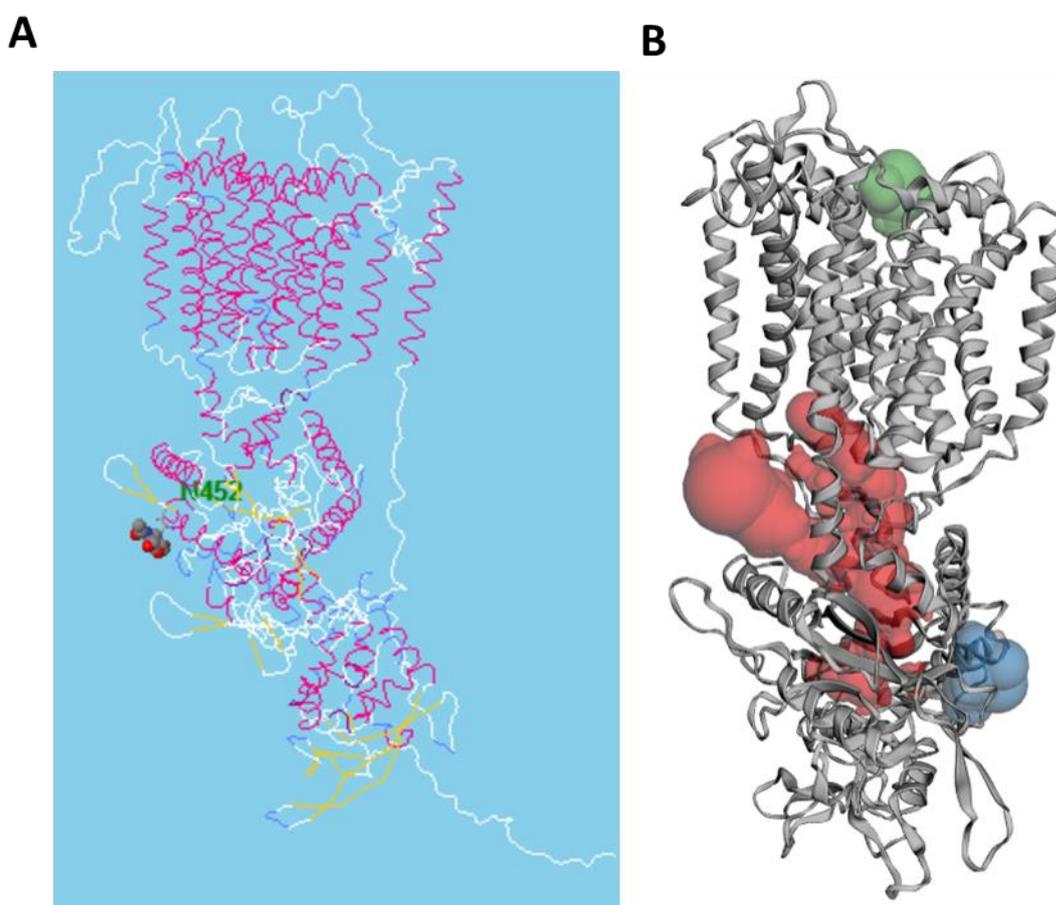


Fig 5. Detección computacional de cavidades. **A** análisis de cavidades de NPC15u74 con *RaptorX*, resulta en cavidades lejos de sitios de interés. **B** análisis de cavidades de NPC1-5u74 con *CASTp*, resulta en una cavidad extremadamente grande (cavidad roja).

2.4.2 Bibliografía

Otra forma de definir cavidades de unión es mediante la búsqueda bibliográfica de cavidades, o sitios de unión, descritas experimentalmente. Para ese fin, se emprendió una segunda búsqueda bibliográfica centrada en estudios de biología estructural de las dianas en artículos y bases de datos como *RSCPD*B y *UniProt*. Con este método se encontraron sitios de unión a ligandos de las distintas dianas. Dichos sitios de unión están normalmente definidos por los aminoácidos que realizan la interacción, por lo que el cribado de ligandos usando estos aminoácidos debería ser preciso. En este apartado también se incluyen los sitios de unión definidos mediante el *alanine scanning* descritos en el apartado 2.3, ya que se basan también en datos de recursos bibliográficos.

2.5 Cribado de ligandos contra dianas

[MitOpenScreen](#) (16) es otro servicio del portal *Mobylye@RPBS*. Esta herramienta sirve para hacer cribados de ligandos contra el modelo en formato *pdb* o *MOL2*. Para realizar el cribado hay que introducir distintos parámetros, en este proyecto se usaron:

- compound data-base (librería): diverse-lib
- compound library filters: compound is lead-like
- grid calculation mode: list of residues

Con estos parámetros se define que el cribado sea con la librería de moléculas general (no solo proteínas) con propiedades de *lead* y que el cribado se debería hacer específicamente contra la cavidad delimitada por los aminoácidos escritos. El formato de introducción de los aminoácidos sigue el ejemplo siguiente: _A_PHE_919_, _A_ALA_866_

Como resultados del cribado se obtiene un archivo csv con los ligandos probados ordenados en función de su fuerza de unión a la cavidad. De esta lista se seleccionaron los cinco primeros ligandos para seguir con el estudio. Los resultados del cribado también se pueden observar en el *applet* del servidor (Fig 6).



Fig 6. Visualización de la unión de un ligando con NPC2

2.6 Cribado de ligandos por similitud

El cribado de ligandos similares se realizó con [SwissSimilarity](#) (17). *SwissSimilarity* busca moléculas parecidas estructuralmente a partir de una molécula en formato SMILES (**S**implified **M**olecular-**I**nter **L**ine-**E**nter **S**ystem). Como indica el nombre, SMILES, consiste en una línea de texto que describe la estructura química de la molécula. Por ejemplo: CC(=O)OC1=CC=CC=C1C(=O)O corresponde a la estructura de la aspirina en formato SMILES.

Dentro del formato SMILES hay distintas variaciones. En este proyecto se trabajó siempre con **SMILES canónico**, que es el formato SMILES estandarizado para su uso computacional. Por lo tanto, en lo que resta de trabajo, siempre que se indique formato SMILES se hará referencia al SMILES canónico.

Para realizar el cribado de ligandos por similitud, el formato SMILES de los fármacos a estudiar se obtuvo de la base de datos en línea *PubChem* (18).

El cribado se realizó con los parámetros siguientes:

- **Library to screen:** Zinc Drug-like, ZINC15 es la librería pública con más ligandos (10.639.400) que ofrece el programa (19). Al seleccionar esta librería se asegura que se compruebe el mayor número de ligandos posibles y que la información de los compuestos resultantes del cribado esté disponible en línea. Además, todos los ligandos de la librería están disponibles comercialmente.

- **Method:** Combined. *SwissSimilarity* da la opción de realizar el cribado por similitud con distintos métodos y algoritmos. Seleccionando *Combined* se usa una combinación de los métodos disponibles:
 - o **FP2 fingerprints:** similitud 2D, en base a la estructura química
 - o **Electroshape:** similitud 3D en base a la estructura, no requiere alineamiento
 - o **Spectrophores:** similitud 3D en base a la estructura, no requiere alineamiento
 - o **Shape-IT:** similitud 3D en base a la estructura, requiere alineamiento
 - o **Align-IT:** similitud 3D en base a farmacóforos, requiere alineamiento

Los resultados del cribado se almacenan en un csv con los compuestos ordenados según su similitud con la molécula de entrada. Las fórmulas estructurales de las moléculas se presentan en la web ordenadas según su parecido al ligando de entrada (Fig 7).

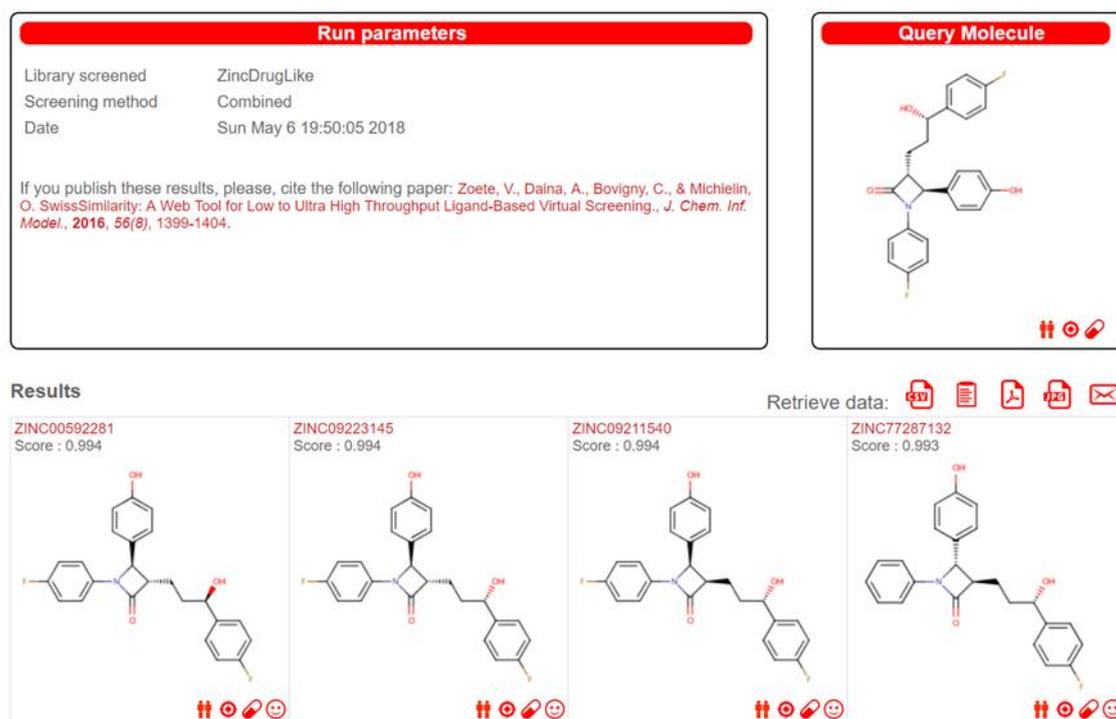


Fig 7. Ejemplo de los resultados obtenidos del cribado por similitud a partir de ezetimibe.

2.7 Acoplamiento de ligando a diana

Para estudiar las capacidades de *docking* se usó el software [MitAutoDock](#) (20). Esta herramienta permite estudiar la unión de hasta 10 ligandos a una proteína. El formulario requiere los ligandos en formato *mol2* o *sdf*, y la diana en formato *mol2* o *pdb*. En caso de usar proteínas grandes, es recomendable usar las opciones avanzadas del programa e introducir una *grid* para reducir la superficie de acoplamiento. La *grid* funciona de forma idéntica a la de *MitOpenScreen* (Apartado 2.5). En este proyecto se usó la opción de

establecer una *grid* para delimitar el acoplamiento a las cavidades de interés definidas.

Para el bloque 1, se usó el *docking* para estudiar la unión de colesterol a los distintos mutantes de NPC1.

Para el bloque 3 se usó el *docking* para validar los resultados del cribado de similitud a partir de fármacos con una diana molecular clara y definida.

2.8 Farmacocinética

Las propiedades farmacocinéticas (ADME-T) de los *hits* obtenidos de los cribados se analizaron usando dos herramientas en línea distintas: [pkCSM](#) (21) para los *hits* del cribado contra dianas terapéuticas y [SwissADME](#) (22) para los *hits* del cribado por similitud. En un principio estaba planeado usar solo *pkCSM*, pero el software de cribado por similitud ofrece la posibilidad de realizar un análisis farmacocinético directamente. A consecuencia, se usó esta herramienta (*SwissADME*) para el análisis ADME de los resultados del cribado por similitud. El análisis de toxicidad se realizó igualmente con *pkCSM*. Hubiera sido deseable poder combinar ambos métodos de análisis de farmacocinética en los resultados de todos los cribados, pero debido a la falta de tiempo, no se pudo realizar un análisis de los resultados con tanta profundidad.

2.8.1 *pkCSM*

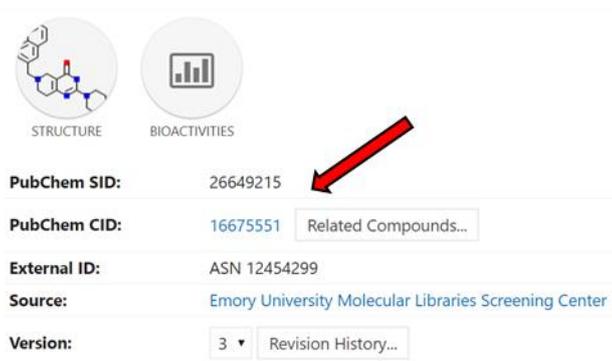
Los ligandos resultantes del cribado de *MitOpenScreen* están identificados según su *PubChemSID* (Substance Identifier). El *PubChemSID* es una codificación interna de *PubChem* para identificar todas las sustancias agregadas a la base de datos. Usando el *PubChemSID*, se pueden buscar las sustancias en la página de [PubChem](#) (Fig 8).



Fig 8. En el buscador principal de *PubChem* podemos buscar compuestos o sustancias cambiando la lengüeta del menú seleccionada (flechas rojas). En este ejemplo se pueden buscar compuestos introduciendo el CID.

Cuando la estructura química de una sustancia pasa los procesos de estandarización de la base de datos, se le asocia un *PubChemCID* (Compound identifier). En la ficha del compuesto estandarizado se puede obtener información sobre su estructura, como la fórmula SMILES canónica y el nombre en formato IUPAC (formato estándar de nomenclatura de compuestos químicos en base a su estructura química). Des de la página de la sustancia, se puede acceder a la ficha de *PubChem* del compuesto (Fig 9A). Y accediendo a la ficha del compuesto se puede obtener su fórmula SMILES (Fig 9B).

A



STRUCTURE BIOACTIVITIES

PubChem SID: 26649215

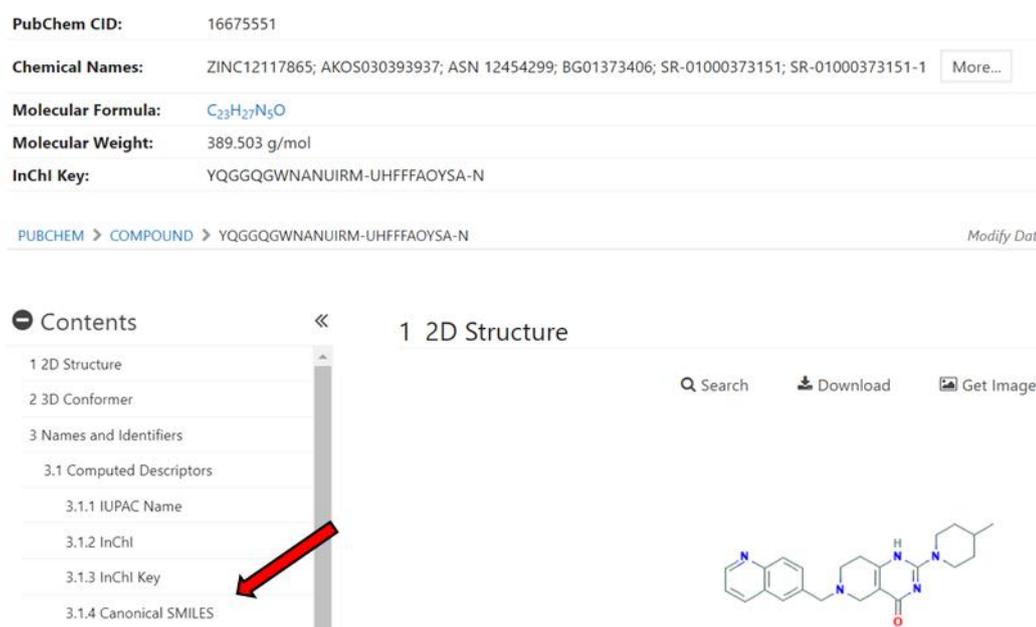
PubChem CID: [16675551](#) Related Compounds...

External ID: ASN 12454299

Source: [Emory University Molecular Libraries Screening Center](#)

Version: 3 Revision History...

B



PubChem CID: 16675551

Chemical Names: ZINC12117865; AKOS030393937; ASN 12454299; BG01373406; SR-01000373151; SR-01000373151-1 More...

Molecular Formula: $C_{23}H_{27}N_3O$

Molecular Weight: 389.503 g/mol

InChI Key: YQGGQGWNANUIRM-UHFFFAOYSA-N

PUBCHEM > COMPOUND > YQGGQGWNANUIRM-UHFFFAOYSA-N Modify Dat

Contents

- 1 2D Structure
- 2 3D Conformer
- 3 Names and Identifiers
 - 3.1 Computed Descriptors
 - 3.1.1 IUPAC Name
 - 3.1.2 InChI
 - 3.1.3 InChI Key
 - 3.1.4 Canonical SMILES

1 2D Structure

Search Download Get Image

Fig 9. Ejemplo de obtención de SMILES a partir del PubChem SID del ligando 26649215. **A** Página de la sustancia con *PubChem SID* 26649215, vemos el enlace a la página del compuesto 16675551 (flecha roja). **B** Página del compuesto *PubChem CID* 16675551, vemos en el menú el apartado con el ligando en formato SMILE (flecha roja).

pkCSM permite realizar un análisis de: **Administración, Distribución, Metabolismo, Excreción, Toxicidad**, o todo el análisis **ADME-T**, a partir de una fórmula SMILES, o un archivo con un listado de moléculas en formato SMILES. En este caso, usamos una lista de los cinco mejores hits del cribado de ligandos contra cada sitio de unión de las dianas. *pkCSM* retorna una hoja de cálculo con los siguientes parámetros ADME-T calculados para cada molécula:

Absorción

-Permeabilidad CaCo-2: mide la absorción de la mucosa intestinal de fármacos administrados oralmente. Valores > 0.90 tendrían una permeabilidad CaCo-2 alta.

-Absorción intestinal (humano): predicción del porcentaje de moléculas que serán absorbidas en el intestino delgado (principal área de absorción de fármacos orales). $<30\%$ se considera mala absorción.

-Solubilidad en agua: solubilidad en agua del compuesto a 25°C expresada en (log mol/L). Los compuestos liposolubles son absorbidos con mayor dificultad.

-Substrato de glicoproteína P: proteína de membrana encargada de la expulsión de sustancias tóxicas de la célula. Sus substratos son eliminados rápidamente.

-Inhibidores de glicoproteína P I y II: inhibidores de la glicoproteína P pueden causar interacciones con otros fármacos incrementando su vida media en el cuerpo. Hay que tenerlo presente en situaciones de poli medicación.

-Permeabilidad de la piel: permeabilidad de compuestos en la piel expresada como Log Kp (cm/h). Compuestos con un Log Kp > -2.5 tienen una permeabilidad en piel baja.

Distribución:

-VDss humano: el volumen de distribución es el volumen teórico que una dosis de fármaco / molécula necesitaría para distribuirse uniformemente por todo el cuerpo (células y plasma sanguíneo). Es una medida de la tendencia de una molécula a permanecer en sangre o ser adquirida por los tejidos. Un VDss alto indica que la molécula es cuestión se distribuye más en los tejidos que en sangre. VDss bajo: < 0.71 L/kg (log VDss < -0.15). VDss alto: > 2.81 L/KG (log VDss > 0.45).

-Fracción libre (humano): las moléculas son transportadas en sangre en forma soluble o unidas a proteínas del plasma sanguíneo. La fracción de moléculas solubles son más accesibles a tejidos y células debido a su mayor capacidad de difusión.

-Permeabilidad de la BBB (barrera hemato-encefálica): la capacidad de una molécula de cruzar la BBB y distribuirse por el sistema nervioso central (CNS) se mide con logBB (ratio de concentración en cerebro y plasma de la molécula). Moléculas con logBB > 0.3 tienen la capacidad de cruzar la BBB y moléculas con logBB < -1 con dificultad la cruzarán.

-Permeabilidad del (CNS): el producto de la permeabilidad de la BBB y el área de superficie (logPS) es un parámetro de distribución en el CNS más directo ya que omite el efecto de la distribución sistémica de la molécula (VDss). logPS > -2 es indicativo de compuestos que penetran el CNS, y se considera que compuestos con logPS < -3 no podrán penetrar el CNS.

Metabolismo:

-Inhibidor de CYP450: el citocromo P450 es el mayor agente detoxificador del cuerpo. No es recomendable tomar fármacos que lo inhiban directamente o por competición ya que pueden causar efectos adversos variados y peligrosos. Se consideran inhibidores compuestos que en concentraciones menores de 10µM inhiban la actividad de CYP450 en un 50%.

-Sustrato de CYP2D6/CYP3A4: es importante saber si una molécula será sustrato de CYP450 para tener en cuenta posibles efectos de inhibición competitiva que pueda causar. CYP2D6 y CYP3A4 son las principales isoformas de CYP450, por lo que se estudia la capacidad de compuestos de ser su sustrato.

Excreción:

-Substrato de OCT2 renal: OCT2 es un transportador renal involucrado en la excreción de moléculas. Es importante establecer si fármacos son sustratos de OCT2, porque es indicativo de su aclaramiento renal y de posibles interacciones con otros compuestos excretados por la misma vía.

-Aclaramiento total: se mide por la constante de proporcionalidad CL_{tot}, determinada principalmente por el aclaramiento hepático (metabolismo del fármaco) y el aclaramiento renal (excreción del fármaco). Sirve para determinar dosis efectivas de los fármacos y *pkCSM* expresa su predicción en logCL_{tot} (log ml/min/kg).

Toxicidad:

-LD 50 de rata: la dosis letal media (LD50) es una medida de toxicidad general: concentración de compuesto necesaria para causar letalidad en el 50% de los casos (concentración en mol/kg). Se usa como guía para determinar el límite del abanico de concentraciones a probar en ensayos clínicos.

-Toxicidad *P. piriformis*: bacteria protozoa que se usa de forma similar a las ratas en el punto anterior para delimitar el rango de toxicidad. Se expresa en función del pIGC50, que es el logaritmo negativo de la concentración de compuesto necesaria para inhibir el crecimiento de la bacteria: $\log \mu\text{g/ml}$. Toxicidad si $\text{pIGC50} < -0.5 \log \mu\text{g/ml}$.

-Toxicidad AMES: el test AMES sirve para valorar si un compuesto tiene potencial mutagénico en bacterias. Un resultado positivo indicaría que el compuesto tiene potencial cancerígeno.

-Toxicidad Minnow: modelo como el de la rata, pero estudiando la LD50 en el pez Flathead Minnow. Los resultados son en formato LC50, valores de $\text{LC50} < -0.3$ son considerados como toxicidad aguda.

-Dosis máxima tolerada: estimación de la dosis máxima tolerada recomendada (MRTD) en humanos. $\text{MRTD} \leq 0.477 \log \text{mg/kg/día}$ se considera bajo, $\text{MRTD} > 0.477 \log \text{mg/kg/día}$ se considera alto.

-Hepatotoxicidad: que un fármaco no afecte el hígado es muy importante, porque el hígado es el órgano detoxificador del cuerpo. Disminución de la actividad hepática puede causar toxicidad por otros fármacos que en condiciones normales no tendrían ningún efecto negativo.

-Toxicidad oral crónica (rata): efecto de la exposición oral crónica a dosis bajas del fármaco en ratas. Mide LOAEL (mínima dosis tóxica) y NOAEL (máxima dosis segura) expresadas en $\log \text{mg/kg}_{\text{bw}}/\text{día}$.

-Sensibilización de la piel: evaluación de los efectos tópicos del fármaco. Se considera que hay sensibilización si aparecen reacciones alérgicas o dermatitis.

-Inhibidor de hERG I / II: la inhibición de estos canales de potasio es la principal causa de síndrome del QT largo. Fármacos inhibidores de hERG I y II pueden causar complicaciones cardíacas.

El último paso del proceso es filtrar los ligandos según sus características *ADME-T* para acabar obteniendo los posibles *leads*. Dado que en este proyecto se está estudiando una enfermedad neurodegenerativa hereditaria, se estima que el tratamiento deberá ser crónico, fácil de subministrar y el fármaco deberá llegar al CNS. Por lo tanto, se prioriza:

- la capacidad de absorción intestinal de los ligandos.
- la capacidad de cruzar la barrera hemato-encefálica y difundir en el CNS de los ligandos.
- la baja toxicidad de los ligandos.

La enfermedad que se estudia en este trabajo está causada por defectos en el transporte de lípidos, por lo que se espera que la mayoría de sitios de unión de dianas sean de carácter hidrofóbico. Es por esto que se presupone que los ligandos encontrados tengan baja solubilidad en agua y distribución sistémica. A consecuencia, no se hace tanto énfasis en su distribución sistémica. Finalmente, añadir que la distribución sistémica de fármacos hidrofóbicos se puede mejorar con vehículos lipofílicos, pero este proceso está fuera del alcance de este proyecto.

2.8.2 SwissADME

El estudio farmacocinético con *SwissADME* se puede realizar directamente des del *output* de *SwissSimilarity* (Fig 10). Es por ello que se usa *SwissADME* para analizar las propiedades farmacocinéticas (ADME) de los cinco mejores *hits* del cribado por similitud de cada ligando. La toxicidad de los *hits* se evalúa con *pkCSM*.

Results

Retrieve data:

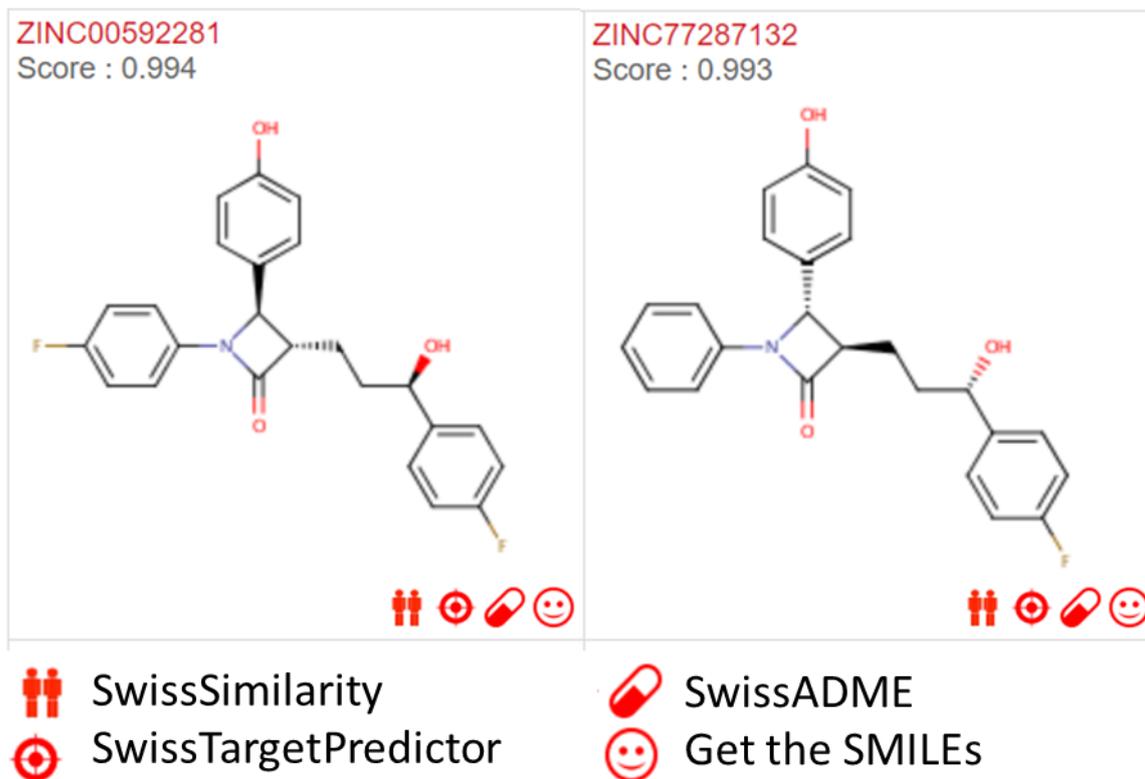


Fig 10. *Output* de *SwissSimilarity*. Se pueden descargar los resultados del cribado como CSV, archivo de texto, PDF, JPG, o exportar directamente al email. Además, clicando los respectivos iconos, se puede ir directamente a otro análisis del ligando. En este trabajo se ha usado la opción *SwissADME* para hacer el estudio farmacocinético de los *hits* del cribado.

El análisis de *SwissADME* es más focalizado en la biodisponibilidad y la química médica que no el de *pkCSM*, y consta de dos partes:

1. Gráfico *Boiled egg* (23): este gráfico esquematiza la predicción de la distribución de la molécula (Fig 11). Es muy útil para este proyecto ya que hace énfasis en la capacidad de la molécula de cruzar la BBB y de ser, o no, expulsada del CNS por la glicoproteína P.

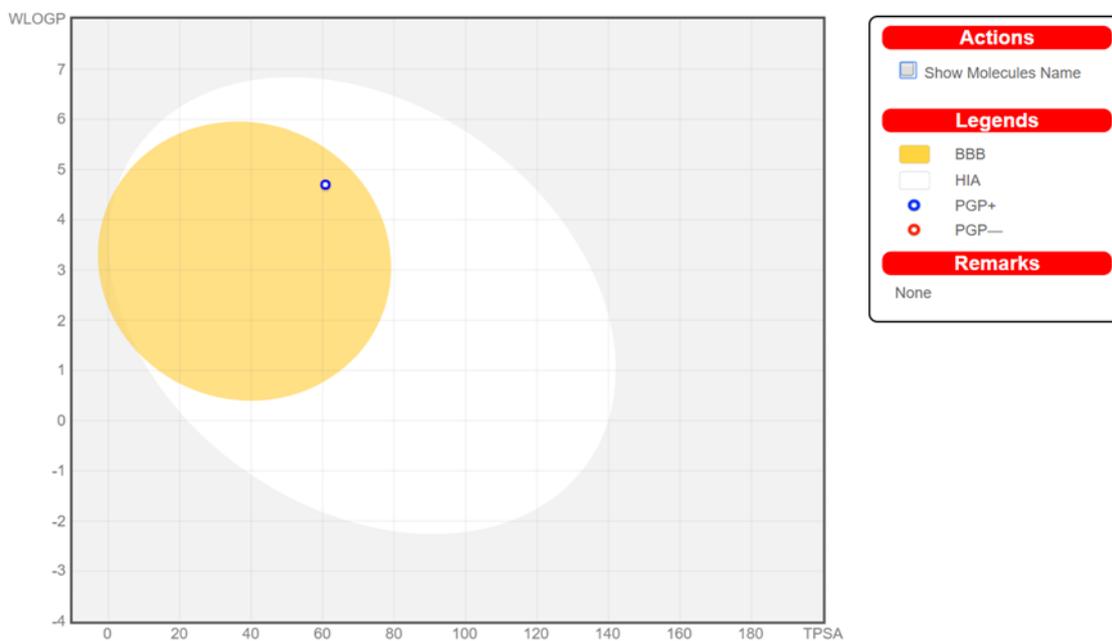
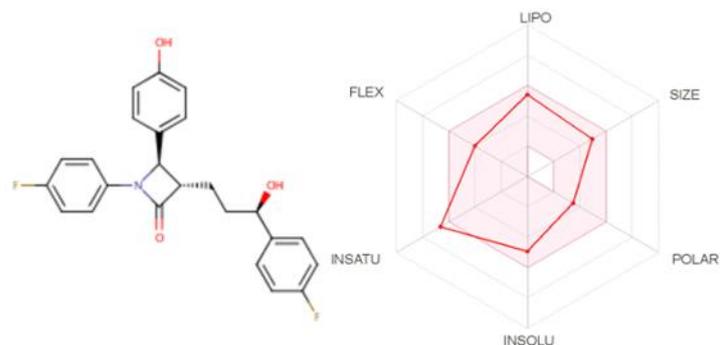


Fig 11. Esquema *Boiled Egg*. El punto representa la molécula de estudio. Si está dentro del área blanca es absorbida en el sistema gastro-intestinal, y si además está dentro del área amarilla, puede permear a través de la BBB. El color del borde del punto determina si la molécula es susceptible de ser un sustrato (azul), o no (rojo), de la glicoproteína P.

2. Resumen de las propiedades farmacocinéticas de la molécula. En primera instancia hay un resumen de las cualidades físico-químicas de la molécula, detalles de su estructura química y un gráfico descriptor (*bioavailability radar*) (Fig 12). Las cualidades fisicoquímicas listadas, y las propiedades ADME descritas en el siguiente apartado, pueden descargarse en formato CSV.



SMILES Oc1ccc(cc1)[C@H]1[C@H](CC[C@H](c2ccc(cc2)F)O)C(=O)N1c1ccc(cc1)F

Physicochemical Properties	
Formula	C ₂₄ H ₂₁ F ₂ NO ₃
Molecular weight	409.43 g/mol
Num. heavy atoms	30
Num. arom. heavy atoms	18
Fraction Csp ³	0.21
Num. rotatable bonds	6
Num. H-bond acceptors	5
Num. H-bond donors	2
Molar Refractivity	112.97
TPSA	60.77 Å ²

Fig 12. Análisis químico de la molécula. Los gráficos de los ejes son respectivamente: FLEX – flexibilidad, LIPO – liposolubilidad, SIZE – tamaño, POLAR – polaridad, INSOLU – insolubilidad, INSATU – insaturación. Los detalles numéricos de cada condición se pueden consultar pasando el ratón por encima. TPSA: suma de la superficie de todos los átomos polares de una molécula.

A parte de las cualidades químicas, el informe también contiene medidas de (Fig 13):

- **Lipofilia:** se mide con el coeficiente de partición entre n-Octanol y agua ($\log P_{OW}$). $\log P_{OW} > 4.5$ es indicativo de alta lipofilia y posible acumulación de la molécula en tejidos. Hay distintas formas de calcularlo, en *SwissADME* se usan 5 modelos distintos y la media de estos (*Consensus log P_{OW}*). En este proyecto nos fijaremos en la media y si hay grandes discrepancias entre modelos.
- **Solubilidad en agua:** se mide con el logaritmo de la solubilidad molar en agua de la molécula ($\log S$). *SwissADME* incluye 3 modelos distintos para calcular $\log S$. Queda a elección del usuario usar uno, u otro. En este proyecto usaremos el consenso de los 3 modelos.
- **Farmacocinética:** distintos parámetros de absorción, distribución y metabolismo.
 - GI absorption:* absorción gastrointestinal.
 - BBB permeant:* capacidad de cruzar la barrera hemato-encefálica.
 - P-gp substrate:* sustrato, o no, de la glicoproteína P.

CYP1A2/2C19/2C9/2D6/3A4 inhibitor: capacidad de la molécula de interferir con la actividad de los citocromos detoxificadores del hígado.

Log K_p : absorción cutánea.

- **Druglikeness:** valoración de la capacidad de la molécula de actuar como un fármaco en relación a biodisponibilidad y distribución. Este apartado consta de 5 modelos distintos para predecir tal condición de una forma cualitativa; y *bioavailability score*, que es un método semicuantitativo que predice la probabilidad de que la molécula tenga > 10% de biodisponibilidad oral en rata o permeabilidad CaCo2 mesurable. Este apartado se trata como el de *solubilidad en agua*: por consenso.
- **Química médica:** este apartado contiene información sobre el posible efecto de la molécula. Los algoritmos *PAINS* y *Brenk* contrastan la estructura de la molécula con bases de datos de patrones estructurales conocidos por su alta reactividad, dar falsos positivos, ser inespecíficos, inestable metabólicamente... *Leadlikeness* considera si la molécula podría ser tratada como lead y ser optimizada físicoquímicamente. En caso negativo, pasando el ratón por encima del área se puede leer las características que no cumplen con las restricciones de *leadlikeness*. Por último, *synthetic accessibility* analiza la estructura y devuelve una estimación de la dificultad de su producción sintética (1 = fácil de sintetizar, 10 = difícil de sintetizar).

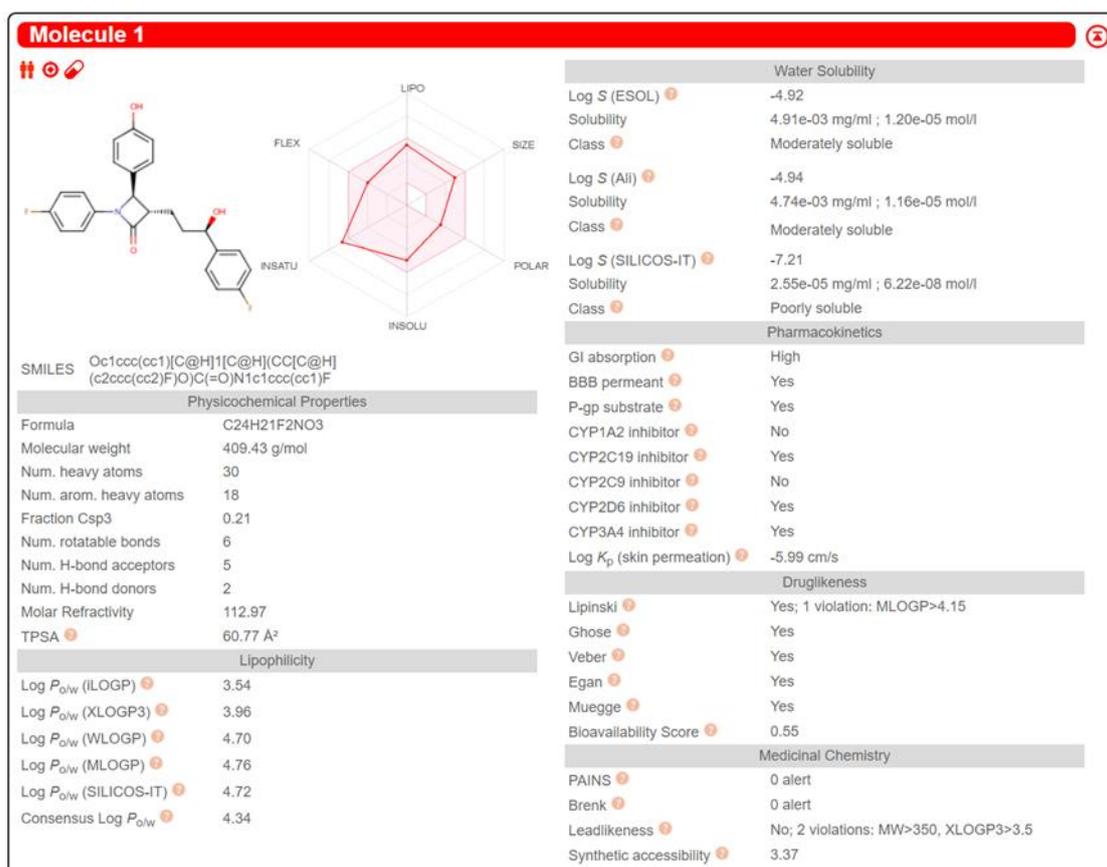


Fig 13. Ejemplo del output de SwissADME.

3. Resultados

3.1 Identificación bibliográfica de objetivos de estudio

3.1.1 Mutaciones de NPC1

De las mutaciones puntuales de NPC1 causantes de NP-C, se analizaron las cinco con mayor prevalencia (Fig 14): Debido a complicaciones de calendario, este apartado se centró solamente en las mutaciones más representativas, las que afectan a NPC1, y se dejó el estudio del efecto de las mutaciones en NPC2 para futuros trabajos.

Genomic position Exon:

cDNA position Classification ≥

AA position Number of patients ≥

Variant	dbSNP id	Gene	Type	cDNA change	AA change	Exon/intron	Classification	# patients	# controls
chr18:21115443 T>C	rs28942105	NPC1	missense_variant	c.3467A>G	p.Asn1156Ser	22	5 (pathogenic)	7	0
chr18:21116700 A>G	rs80358259	NPC1	missense_variant	c.3182T>C	p.Ile1061Thr	21	5 (pathogenic)	67	5
chr18:21118528 G>C	rs80358257	NPC1	missense_variant	c.3019C>G	p.Pro1007Ala	20	5 (pathogenic)	16	2
chr18:21118573 C>A	rs80358254	NPC1	missense_variant	c.2974G>T	p.Gly992Trp	20	5 (pathogenic)	12	1
chr18:21121319 T>G	rs80358253	NPC1	missense_variant	c.2324A>C	p.Gln775Pro	15	5 (pathogenic)	5	0

5 variants (after filtering)

Fig 14. Cinco mutaciones más comunes de NPC1 causantes de NP-C.

El modelaje por homología para obtener la estructura 3D de NPC1 mutadas (Fig 15) se realizó a partir del mismo archivo sobre el que se aplicaron las mutaciones (Tabla 1). Se escribió un *script* en R para modificar el archivo *pdb* original (NPC1-5u74) e introducirle las mutaciones. [ANEXO: Mutaciones.pdf]

Modelo	Template	Cobertura	Rango	Seq. Identity	GMQE	QMEAN
G992W	5u74.1.A	1.00	334-1255	99.92	0.67	-3.26
I1061T	5u74.1.A	1.00	334-1255	99.92	0.67	-3.26
N1156S	5u74.1.A	1.00	334-1255	99.92	0.67	-3.26
P1007A	5u74.1.A	1.00	334-1255	99.92	0.67	-3.26
Q775P	5u74.1.A	1.00	334-1255	99.92	0.67	-3.26

Tabla 1 Características y parámetros de calidad de los modelos por homología generados. Como todos los modelos tienen la misma base, y solo un cambio de base, los parámetros de calidad son iguales.

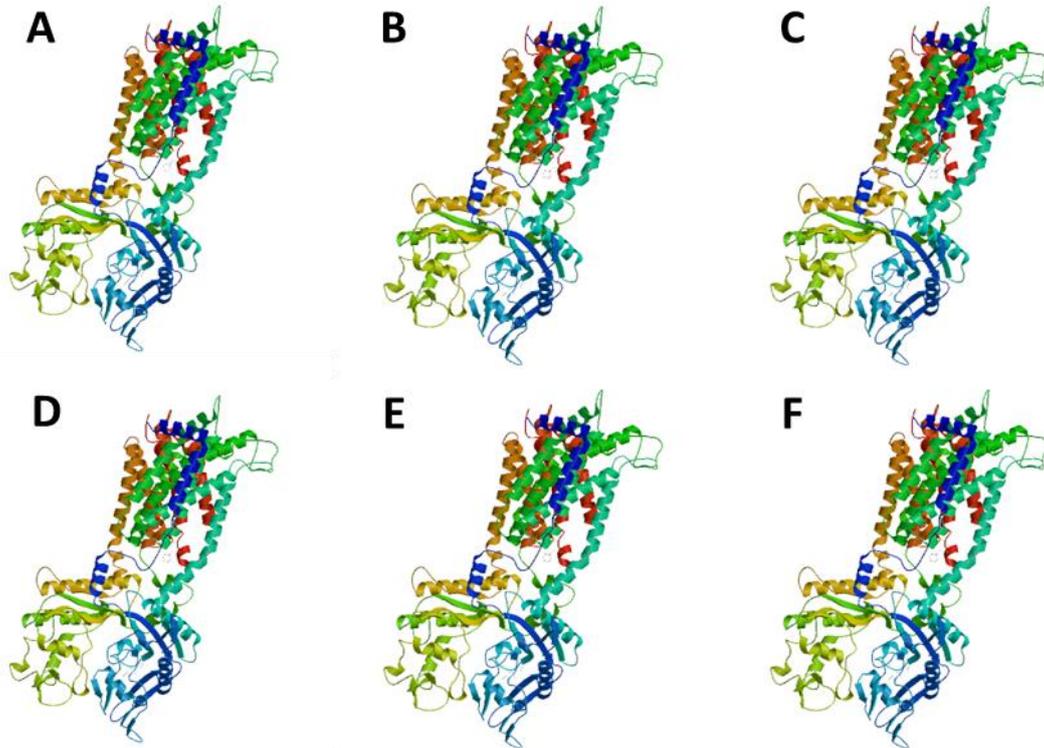


Fig 15. Modelos de NPC1 mutado generados por homología a partir de NPC1-5u74.
A G992. **B** I061T. **C** N1156S. **D** P1007A. **E** Q775P. **F** WT

3.1.2 Dianas terapéuticas

En la búsqueda bibliográfica se identificaron cinco proteínas, relacionadas con el transporte de colesterol, como potenciales dianas terapéuticas:

-NPC1: proteína transmembrana de la membrana lisosomal que transporta colesterol del lumen del LE/LY al citosól (24).

Tiene tres dominios en la cara luminal del lisosoma (CTD, MDL, NTD) que le permiten interactuar con NPC2 y colesterol (25).

NPC2 se une al dominio MDL de NPC1 (24) y entrega colesterol al dominio NTD de NPC1, que reconoce colesterol por el extremo hidroxil (26). El dominio CTD es el encargado de mantener NTD y MDL en una conformación óptima y, curiosamente, es el dominio en que se acumulan más mutaciones causantes de NP-C (27).

-NPC2: proteína soluble lisosomal que captura colesterol por el extremo isoocil y lo presenta a NPC1(24,28).

-OSBPL5 (o ORP5): proteína transmembrana mayoritariamente localizada en el retículo endoplasmático, que puede estar involucrada en el último paso de transporte de colesterol al citosól con NPC1 (29).

-NPC1L1: proteína transmembrana localizada mayoritariamente en la membrana apical de enterocitos, que media la absorción de colesterol durante la digestión (30).

-LXR (tipo β): receptor nuclear expresado ubicuamente que actúa como sensor de oxi-esteroles y modifica el metabolismo de colesterol y otros lípidos en consecuencia (31,32). En este trabajo, cuando se mencione LXR se hará referencia a LXR β a no ser que se especifique lo contrario.

Tres de las proteínas seleccionadas tenían estructuras cristalografiadas disponibles en UniProt y RCSPDB. En la Tabla 2 se detallan sus características.

Diana	Código PDB	Referencia
NPC1	5U73 / 5U74	Li et al. 2017
NPC1-NTD + colesterol	3GKJ	Kwon et al. 2009
NPC1 + esteroles (cryoEM)	3JD8	Li X et al. 2016
NPC1-MDL	5F1B	Wang et al 2016
NPC2 bovino	1NEP	Friedland et al. 2003
NPC1-MDL+NPC2	5KWY	Li X et al. 2016
NPC1L1 (parcial)	3QNT	Kwon et al. 2011
LXR + agonista	3KFC	Bernotas et al 2010

Tabla 2. Relación de estructuras cristalografiadas, su código PDB y el artículo donde se describen.

Las estructuras 3D de OSBPL5 y NPC1L1 fueron generadas vía modelaje por homología usando *SwissModel* como se detalla en el apartado 2.2. En ambos casos se generaron dos modelos. Para OSBPL5 se usaron dos modelos que cubrían distintas partes de la proteína. Y para NPC1L1 se generó un modelo general de buena calidad y otro de menor calidad, pero con mayor resolución en el área con sitios de unión. Adicionalmente, se creó un modelo de NPC2 a partir de 1nep.pdb (NPC2 bovino sin colesterol). Se generó el modelo, porque la unión de colesterol causa un cambio conformacional en K25, M79, K123 y Q146 de NPC2 (24). Con ambos modelos se pudo realizar el cribado de ligandos en las dos conformaciones. En la Tabla 3 se describen de forma detallada los modelos generados:

Modelo	Template	Cobertura	Rango	Seq Identity	GMQE	QMEAN
OSBPL5m2	5u78.4.A	0.13	127-243	87.18	0.06	-1.19
OSBPL5m3	4ph7.1.A	0.36	369-696	38.75	0.24	-0.81
NPC1L1m1	3jd8.1.A	0.89	33-1292	41.67	0.61	-8.32
NPC1L1m2	5u74.1.A	0.89	336-1296	41.67	0.52	-4.42
NPC2(no-Chol)	1nep.1.A	0.85	20-147	80.47	0.79	-0.74

Tabla 3. Los distintos modelos con sus estimadores de calidad. GMQE (*global model quality estimator*) es una medida de la calidad general del modelo que depende en gran medida de la cobertura del modelo y varía de 0 a 1 (máximo es 1). QMEAN valora la semejanza entre el modelo y los datos experimentales. Va de valores negativos hasta cero (valor máximo) y -4.0 se considera el límite de calidad.

Los modelos generados por homología de OSBPL5 y NPC2 sin colesterol son de calidad óptima. En cambio, los modelos generados a partir de la secuencia de NPC1L1 tienen un QMEAN muy bajo, por debajo del límite de calidad. Se usarán igualmente para este trabajo, pero se debería optimizar el modelo para próximos proyectos.

3.1.3 Moléculas con efecto terapéutico

Con la búsqueda bibliográfica se identificaron tres ligandos a estudiar (Tabla 4):

Ezetimibe: antagonista de NPC1L1 que reduce un 50% la absorción intestinal de colesterol, y un 20% la concentración de colesterol en sangre (33). Con menor concentración de colesterol extracelular (LDL, VLDL), las células absorberían menos colesterol y los lisosomas estarían menos estresados por acumulación de lípidos.

HP β CD: ciclodextrina usada como coadyuvante para mejorar la distribución de fármacos liposolubles (34). Pese a su gran tamaño, es capaz de facilitar el transporte a través de la BBB (35) y estudios han demostrado su eficacia en reducir el estrés lipídico de lisosomas en NP-C (36,37).

Miglustat: fármaco para el tratamiento de la enfermedad de Gaucher (38), que cursa con acumulación de glicoesfingolípidos. Aunque la acumulación de glicoesfingolípidos es secundaria en NP-C, Miglustat puede aliviar sus síntomas (6).

Ligando	PubChem CID	SMILES
Ezetimibe	150311111	<chem>C1=CC(=CC=C1[C@@H]2[C@H](C(=O)N2C3=CC=C(C=C3)F)CC[C@@H](C4=CC=C(C=C4)F)O)O</chem>
HPBCD	14049689	<chem>CC(COC[C@@H]1[C@@H]2[C@@H]([C@H]([C@H](O1)O[C@@H]3[C@H](O[C@@H]([C@@H]([C@H]3O)O)O[C@@H]4[C@H](O[C@@H]([C@@H]([C@H]4O)O)O[C@@H]5[C@H](O[C@@H]([C@@H]([C@H]5O)O)O[C@@H]6[C@H](O[C@@H]([C@@H]([C@H]6O)O)O[C@@H]7[C@H](O[C@@H]([C@@H]([C@H]7O)O)O[C@@H]8[C@H](O[C@H](O2)[C@@H]([C@H]8O)O)COCC(C)O)COCC(C)O)COCC(C)O)COCC(C)O)COCC(C)O)O)O</chem>
Miglustat	51634	<chem>CCCCN1C[C@@H]([C@H]([C@@H]([C@H]1CO)O)O)O</chem>

Tabla 4. Los tres ligandos seleccionados junto con su PubChem CID y la fórmula SMILES que se usó para el cribado por similitud.

3.2 Alanine scanning

El *Alanine scanning* se realizó con estructuras de las dianas que incluyeran ligandos:

- **NPC1:** 3jd8.pdb y 3gkj.pdb, incluyen NPC1 unido a colesterol
- **NPC2:** 5kwy.pdb incluye NPC2 unido a colesterol
- **LXR:** 3kfc.pdb incluye LXR unido a 61X (agonista de LXR)

El análisis por duplicado de la unión de NPC1 y colesterol se usó para comprobar la robustez del método. Como se puede ver en la Figura 16A/B. Ambas estructuras presentan un sitio de unión a colesterol parecido: los aminoácidos presentan energías de unión parecidas. Y más importante, coinciden en que F108 es el aminoácido más importante para la correcta unión a colesterol, y que F203, W27, Q79, M193, L83 también son importantes para la interacción. En el caso de NPC2, los aminoácidos F85, Y119 y L124 son los más importantes para el reconocimiento y captura de colesterol (Fig 16C). En LXR los aminoácidos encargados de reconocer los ligandos son F329, W457, F271, H435, L345 y M312 (Fig 16D).

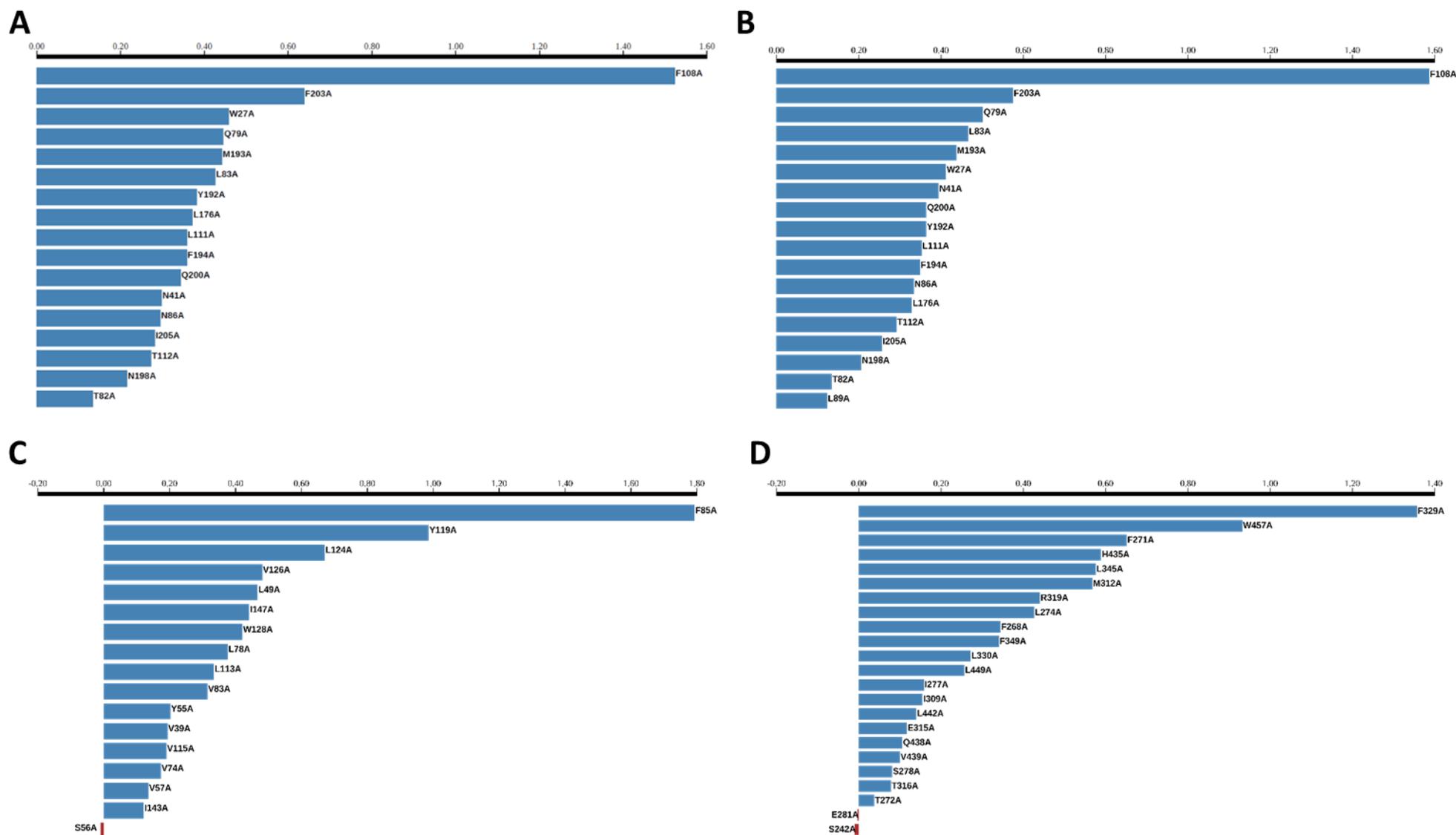


Fig16. Gráfico de barras de los aminoácidos del sitio de unión ordenados según su importancia en el enlace, medido en ddG. **A** Sitio de unión de NPC1-3jd8 a colesterol. **B** Sitio de unión de NPC1-3gkj a colesterol. **C** Sitio de unión de NPC2-5kwy a colesterol. **D** Sitio de unión de LXR-3kfc al ligando 61X.

3.3 Cavidades

Las cavidades generadas con *fpocket* no se usaron en los experimentos definitivos como se explica en el apartado 2.4, por lo que se incluyen como archivo anexo [ANEXO: archivo *pdb* para cada cavidad]. En la Tabla 5 (continua en la página siguiente) se enumeran detalladamente las cavidades definidas a partir de la bibliografía que se usaron en los experimentos siguientes.

Cavidad	Aminoácidos	Origen
NPC1: interacción con colesterol	_A_TRP_27_,_A_ASN_41_,_A_GLN_79_,_A_THR_82_,_A_LEU_83_,_A_ASN_86_, _A_PHE_108_,_A_LEU_111_,_A_THR_112_,_A_LEU_176_,_A_MET_193_, _A_PHE_194_,_A_ASN_198_,_A_GLN_200_,_A_ALA_201_,_A_PHE_203_,_A_ILE_205_	<i>Alanine scanning de 3jd8</i>
NPC1: interacción con colesterol (todos los aminoácidos)	_A_TRP_27_,_A_ASN_41_,_A_GLN_79_,_A_THR_82_,_A_LEU_83_,_A_ASN_86_, _A_PRO_90_,_A_PHE_108_,_A_LEU_111_,_A_THR_112_,_A_LEU_176_,_A_CYS_177_, _A_TYR_192_,_A_MET_193_,_A_PHE_194_,_A_ASN_198_,_A_GLN_200_, _A_ALA_201_,_A_PRO_202_,_A_PHE_203_,_A_ILE_205_	<i>Visualizador 3D de RSCPDB: 3jd8</i>
NPC1: interacción con colesterol (aminoácidos en contacto con colesterol)	_A_TRP_27_,_A_GLN_79_,_A_THR_82_,_A_LEU_83_,_A_ASN_86_,_A_PRO_90_, _A_PHE_108_,_A_LEU_111_,_A_LEU_176_,_A_TYR_192_,_A_MET_193_,_A_PHE_194_, _A_GLN_200_,_A_ALA_201_,_A_PRO_202_,_A_PHE_203_	<i>Visualizador 3D de RSCPDB: 3jd8</i>
NPC1: interacción con NPC2 (MDL)	_A_PHE_504_,_A_GLU_502_,_A_GLN_421_,_A_PHE_503_,_A_TYR_506_, _A_PRO_424_,_A_TYR_423_	(24)
NPC1: plano de interacción CTD-NTD	_A_GLY_910_,_A_GLY_911_,_A_GLY_913_,_A_MET_912_,_A_VAL_234_, _A_THR_235_,_A_ALA_236_	(27)

Tabla 5. Listado de las cavidades descritas en la bibliografía, la secuencia de aminoácidos que las componen, y el recurso de donde se obtuvo la secuencia.

Cavidad	Aminoácidos	Origen
NPC2: unión a colesterol	_A_VAL_39_, _A_LEU_49_, _A_TYR_55_, _A_SER_56_, _A_VAL_57_, _A_VAL_74_, _A_LEU_78_, _A_VAL_83_, _A_PHE_85_, _A_LEU_113_, _A_VAL_115_, _A_TYR_119_, _A_LEU_124_, _A_VAL_126_, _A_TRP_128_, _A_ILE_143_, _A_ILE_147_	<i>Alanine scanning de 5kwy</i>
NPC2: interacción con colesterol	_A_MET_79_, _A_VAL_83_, _A_VAL_86_, _A_PRO_120_	(25)
LXR: unión a ligando	_A_SER_242_, _A_PHE_268_, _A_PHE_271_, _A_THR_272_, _A_LEU_274_, _A_ALA_275_, _A_ILE_277_, _A_SER_278_, _A_GLU_281_, _A_ILE_309_, _A_MET_312_, _A_GLU_315_, _A_THR_316_, _A_ARG_319_, _A_PHE_329_, _A_LEU_330_, _A_LEU_345_, _A_PHE_349_, _A_HIS_435_, _A_GLN_438_, _A_VAL_439_, _A_LEU_442_, _A_LEU_449_, _A_TRP_457_	<i>Alanine scanning de 3kfc</i>
NPC1L1: Entrada de la cavidad de unión a esteres	_A_ILE_105_, _A_MET_188_, _A_ASN_211_, _A_LEU_213_	(39)
NPC1L1: cavidad de unión a esteres	_A_PHE_35_, _A_SER_56_, _A_GLN_95_, _A_SER_98_, _A_LEU_99_, _A_SER_102_, _A_THR_106_, _A_HIS_124_, _A_ASN_127_, _A_THR_128_, _A_TYR_156_, _A_ALA_214_, _A_PRO_215_, _A_LEU_216_, _A_ILE_218_	(39)
NPC1L1: unión de ezetimibe	_A_F_532_, _A_M_543_	(40)
OSBPL5: union a colesterol	_A_MET_93_, _A_PHE_103_, _A_ARG_147_, _A_LYS_150_, _A_PHE_152_, _A_ISO_249_	(41)

Tabla 5 (continuación). Listado de las cavidades descritas en la bibliografía, la secuencia de aminoácidos que las componen, y el recurso de donde se obtuvo la secuencia.

3.4 Cribado virtual de ligandos contra dianas

El cribado virtual de ligandos contra dianas se realizó usando los parámetros detallados en métodos 2.4.1.1. y con los modelos 3D del apartado 3.1.2. La *grid* se definió en cada caso con las cavidades definidas en el apartado 3.3. Por lo tanto, se realizaron los 12 cribados virtuales mostrados en la Tabla 6:

Diana	Modelo	Cavidad
NPC1	3jd8	Definida por AlaScan
NPC1	3jd8	Unión a colesterol
NPC1	3jd8	Interacción con NPC2
NPC1	3jd8	Interacción CTD-NTD
NPC2	5kwy	Definida por AlaScan
NPC2	5kwy	Interacción con NPC1
NPC2	NPC2(no-Chol)	Unión a colesterol
NPC2	5kwy	Unión a colesterol
LXR	3kfc	Unión a ligando
NPC1L1	NPC1L1m1	Entrada de la cavidad de unión a esteres
NPC1L1	NPC1L1m1	Unión a esteres
OSBPL5	OSBPL5m2	Union a colesterol

Tabla 6. Cribados virtuales de ligandos contra dianas realizados.

Aunque NPC1-5u74 tiene mayor resolución, el cribado de ligandos de NPC1 se realizó con NPC1-3jd8 como se ve en la Tabla 6. Esto es debido a que, por alguna razón desconocida, el cribado de ligandos con NPC1-5u74 resulta en error. Y a que, además, la mayoría de cavidades están descritas en NPC1-3jd8 mismo por lo que es más óptimo usar esta versión. Solamente el cribado contra la zona de interacción de los dominios CTD y NTD de NPC1 puede verse afectado por el modelo usado, ya que NPC1-5u74 está supuestamente hecho para proveer mayor resolución de esta región (27).

Una situación parecida se da con el cribado de ligandos contra NPC1L1. En este caso se usa el modelo con menor calidad porque su rango incluye los aminoácidos que definen la cavidad de interés.

Y en el caso de OSBPL5 se usa el modelo 2 para el cribado, porque cubre el rango de aminoácidos de la cavidad de interés.

En la Tabla 7 a continuación se muestra una relación de los 5 mejores *hits* de cada cribado y su energía de unión. Los resultados completos del cribado contra dianas se encuentran en el archivo adjunto [[Cribados_contra_Diana.ods](#)].

Cribado	TOP1	TOP2	TOP3	TOP4	TOP5
NPC1 - AlaScan	-12.8	-12.5	-12.4	-12.4	-12.4
NPC1 - colesterol	-12.8	-12.4	-12.2	-12.2	-12.1
NPC1 - NPC2	-9.2	-9.1	-9.1	-9.1	-8.9
NPC1 - CTD/NTD	-7.8	-7.8	-7.7	-7.6	-7.6
NPC2 - AlaScan	-12.1	-12	-12	-11.7	-11.6
NPC2 - NPC1	-12.6	-12.3	-12	-11.8	-11.8
NPC2 5kwy - colesterol	-12.4	-12.4	-11.9	-11.9	-11.7
NPC2 1nep - colesterol	-7.5	-7.5	-7.4	-7.4	-7.3
LXR - ligando	-13.6	-12.8	-12.7	-12.6	-12.6
NPC1L1 - entrada cavidad	-8.3	-8.1	-8	-7.9	-7.9
NPC1L1 - esterés	-11.2	-10.8	-10.7	-10.7	-10.7
OSBPL5 - colesterol	-9.5	-9.2	-9.1	-9	-9

Tabla 7. Energía de unión de cada ligando con la diana expresada en kcal/mol. La unión es más fuerte cuando más negativa es la energía (dG). AlaScan refiere a cavidad definida mediante *alanine scanning*. NPC2 5kwy – colesterol es el cribado de ligandos contra la cavidad de unión a colesterol en el modelo de NPC2 con la conformación del brazo [K25, M79, K123, Q146] tras unir colesterol. NPC2 1nep – colesterol es el cribado de ligandos contra la cavidad de unión a colesterol del modelo de NPC2 en la conformación nativa antes de capturar colesterol.

3.5 Cribado virtual de ligandos por similitud

El cribado de ligandos por similitud se realizó en base a los tres fármacos descritos en el apartado 3.1.3 usando los parámetros descritos en el apartado 2.6 de Materiales y Métodos.

La Figura 17 (página siguiente) muestra el esquema molecular de las cinco moléculas con mayor índice de similitud con cada uno de los fármacos probados. La información de los ligandos generada por cada cribado se encuentra en el archivo adjunto [[Cribados_por_Similitud.ods](#)].

Las moléculas generadas por el cribado por similitud de ezetimibe y Miglustat tienen una estructura parecida, pero las derivadas de HP β CD parecen derivadas de las unidades estructurales de HP β CD.

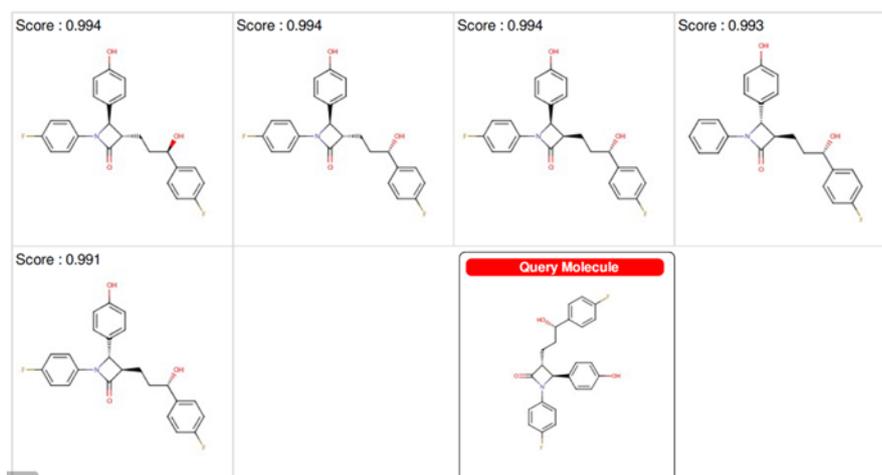
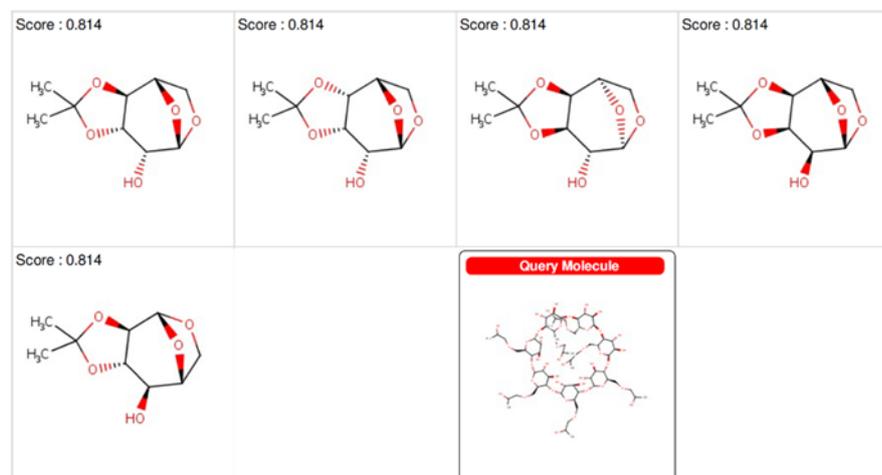
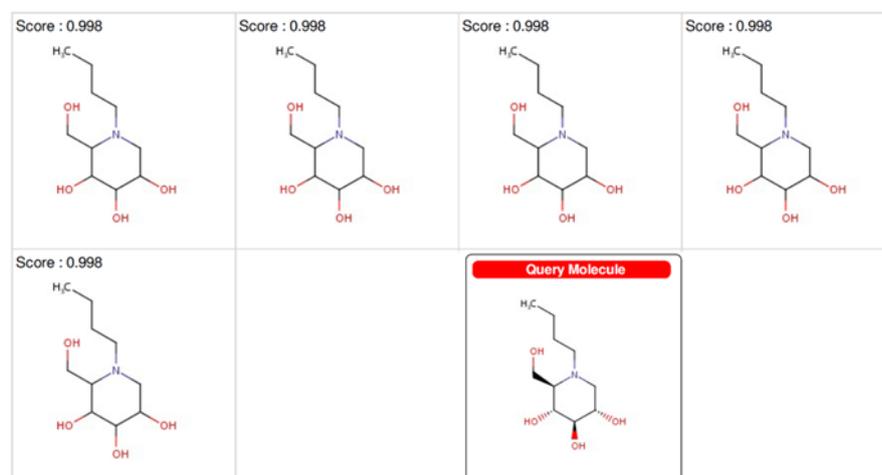
A**B****C**

Fig 17. Resultado de los cribados por similitud. **A** Hits con mayor índice de similitud con ezetimibe (query molecule). **B** Hits con mayor índice de similitud con HPβCD (query molecule). **C** Hits con mayor índice de similitud con Miglustat (query molecule).

3.6 Acoplamiento

Se realizaron dos experimentos de acoplamiento por separado. Uno para el bloque 1: acoplamiento de colesterol a variantes de NPC1. Otro para el bloque 3: acoplamiento para validar el cribado por similitud. Los resultados se detallan a continuación:

3.6.1 Acoplamiento de colesterol a variantes de NPC1

Primero se hizo un estudio de acoplamiento de colesterol a la cavidad de unión a colesterol definida en 3jd8 de NPC1-WT y NPC1 con mutaciones (Tabla 8).

WT	G992	I1067T	N1156S	P1007A	Q775P
-8.53	-8.71	-8.97	-8.67	-9.18	-9.58
-7.78	-8.32	-8.28	-7.64	-8.84	-9.55
-7.61	-8.24	-8.04	-7.61	-8.48	-8.47
-7.04	-8.16	-7.91	-7.54	-8.31	-8.12
-6.95	-8.16	-7.89	-7.31	-7.92	-8.06
-6.69	-8	-7.61	-7.3	-7.59	-8.03
-6.67	-7.89	-7.09	-6.77	-7.29	-7.69
-6.53	-6.71	-7.01	-6.5	-6.85	-6.64
-6.28	-6.65	-6.1	-6.44	-6.55	-6.6
-5.91	-5.23	-6.09	-6.33	-6.27	-6.11

Tabla 8. Energías de unión (dG) de colesterol con NPC1 WT y NPC1 con las cinco mutaciones. Se muestran los resultados de las diez mejores posiciones.

La energía de unión del acoplamiento de colesterol a las distintas versiones de NPC1 no varía excesivamente, y curiosamente NPC1-WT es la proteína que une al ligando con menor fuerza. Para más información referir al archivo adjunto [[Acoplamiento_Colesterol_MutantesNPC1.ods](#)].

Adicionalmente se generaron los modelos del acoplamiento (Fig 18). En los modelos se puede observar como los mutantes N1156S y P1007A unen colesterol en una zona distintivamente distinta de las demás versiones de NPC1.

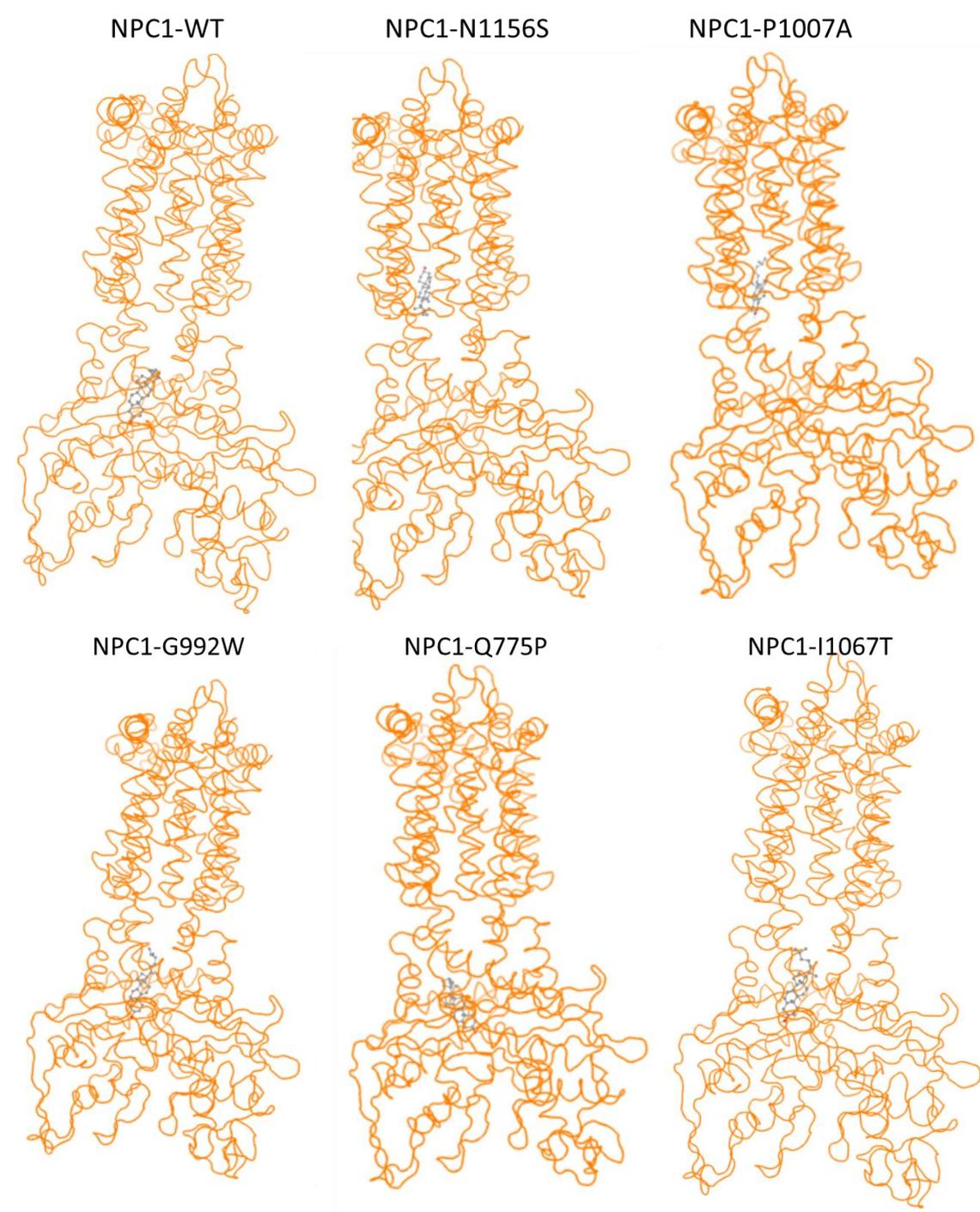


Fig 18. Esquemática del acoplamiento de colesterol (gris) a las distintas versiones de NPC1 (naranja).

3.6.2 Acoplamiento para evaluar el cribado por similitud

Por otra parte, se hizo un experimento de acoplamiento virtual con ezetimibe y las 5 moléculas resultantes del cribado por similitud a ezetimibe contra NPC1L1m2 para valorar su capacidad de interacción con la diana (Tabla 9). Para más detalles sobre la interacción. Consultar el archivo adjunto [[Acoplamiento_Ezetimibe_Derivados_NPC1L1.ods](#)]

Ezetimibe	Top 1	Top 2	Top 3	Top 4	Top 5
-2.03	-5.44	-2.6	-2.39	-4.31	-3.8
-1.05	-5.21	-1.91	-2.26	-1.94	-2.38
-0.73	-3.94	-0.96	-1.41	-1.61	-1.94
0.08	-3.67	1.28	-0.25	-1.51	-0.91
1.95	-0.84	1.73	0.25	-0.83	0.84
2.22	-0.27	2.24	0.64	-0.66	3.4
3.17	0.54	2.32	1.39	-0.2	3.8
5.26	4.36	5.36	2.83	1.39	7.71
9.61	8.3	5.74	8.37	3.29	15.45
80.95	78.14	137.06	34.28	38.72	18.86

Tabla 9. Energía de unión de ezetimibe y las 5 moléculas resultantes del cribado contra el sitio de unión de ezetimibe en NPC1L1m2.

La energía de enlace de las distintas posiciones de todos los ligandos contra NPC1L1m2 es muy baja. De acuerdo con estos resultados, los mejores candidatos serían el Top 1 (ZINC00592281) y el Top 4 (ZINC77287132). Como se ve en la Figura 19, los ligandos localizan en la misma zona. Situación de esperar ya que el sitio de acoplamiento está definido solo por dos aminoácidos.

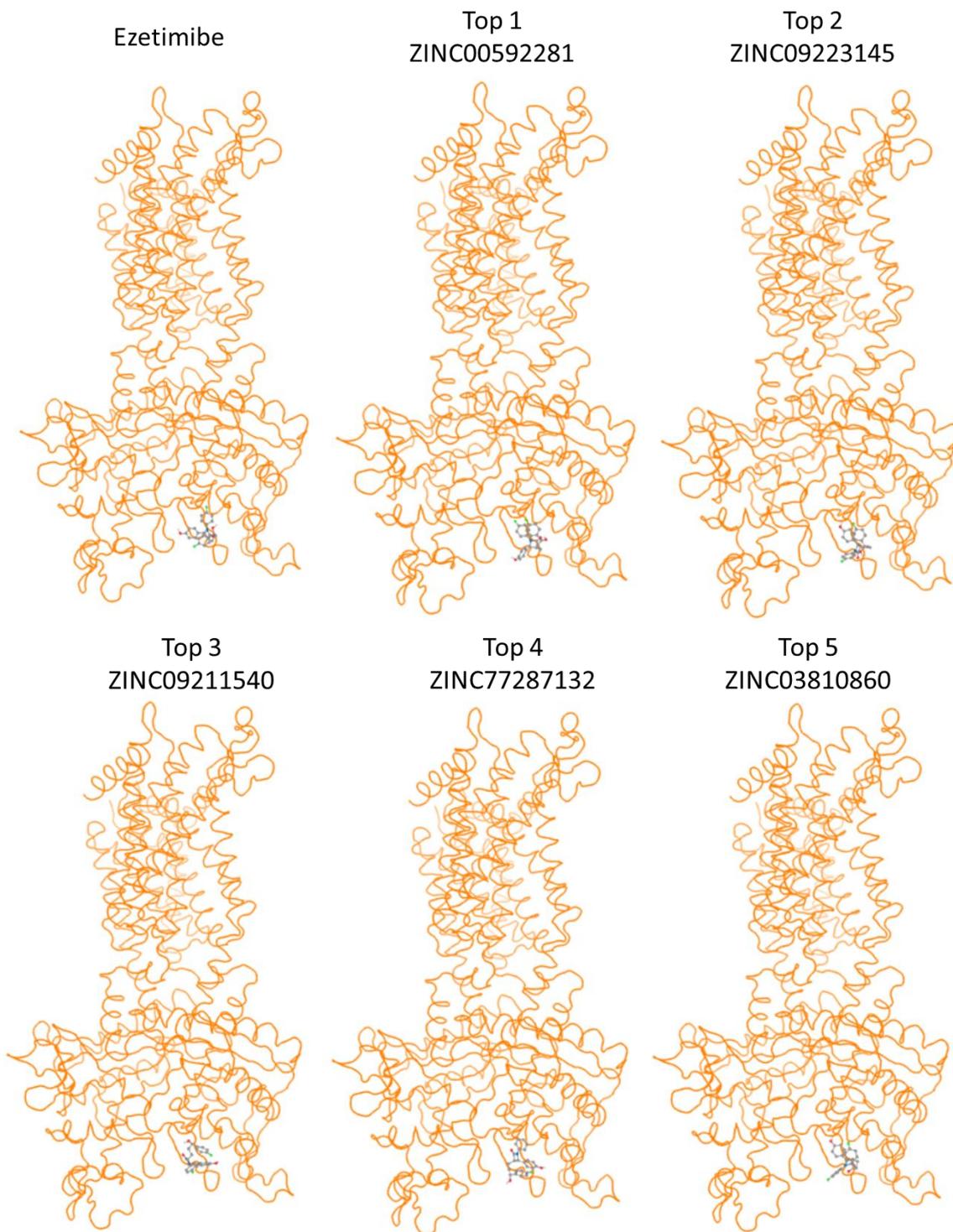


Fig 19. Sitio de unión de los distintos ligandos a NPC1L1m2 (naranja).

3.7 Farmacocinética

Los resultados del cribado virtual de ligandos contra dianas y los del cribado virtual de ligandos por similitud se procesaron por separado.

3.7.1 Cribado de ligandos contra dianas

El análisis de las propiedades ADME-T se realizó sobre los cinco ligandos con mayor puntuación en el cribado, recogidos en el archivo adjunto [[top5_SMILES.ods](#)]. Los resultados del análisis farmacocinético se encuentran en el archivo adjunto [[ADMET_Cribado_contra_Diana.ods](#)]. El archivo consta de una hoja por diana, con cinco moléculas por sitio de unión, y una hoja con los resultados de cada filtraje aplicado.

En primera instancia se aplicó un filtro para identificar los ligandos con más potencial terapéutico, es decir, los que tenían una buena absorción y distribución, y muy baja toxicidad. Para identificar los ligandos que cumplían tales requisitos, se aplicó el filtro siguiente:

- Absorción intestinal > 90%
- Permeabilidad a la BBB ≥ -1
- Permeabilidad al CNS ≥ -3
- Sustrato de la glicoproteína P: NO
- Toxicidad AMES: NO
- Inhibidor de hERG I/II: NO
- Hepatotoxicidad: NO
- Inhibidor de CYP450: NO

Como se puede observar, con el filtro aplicado se prioriza la baja toxicidad de las moléculas, que se espera que tengan una mínima capacidad de permear a través de la BBB y llegar al CNS. De todos los cribados, dos ligandos pasaron este filtro (Tabla 10). Como se detalla en la tabla 11, que recoge las propiedades farmacocinéticas de los dos ligandos, se espera que ambos ligandos tengan buenas propiedades de absorción y distribución. Es más preocupante su posible efecto tóxico. Aun que se haya filtrado los ligandos por toxicidad en sistemas específicos con los filtros de AMES, inhibición de hERG I/II y hepatotoxicidad, los ligandos resultantes tienen una dosis máxima tolerada esperada muy baja. Los valores obtenidos indican una dosis máxima de la escala de $\mu\text{g}/\text{kg}/\text{día}$, que podría ser un factor negativo en el proceso de optimización de estas moléculas hacia potenciales fármacos de NP-C.

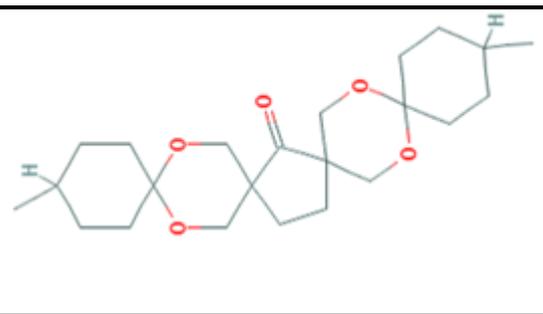
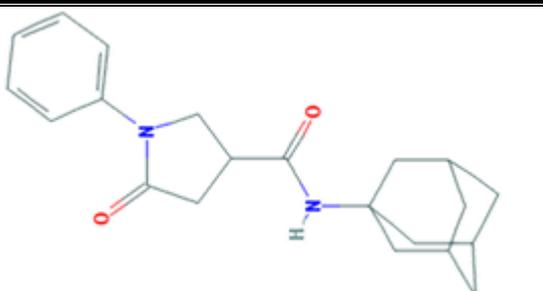
Ligando contra el sitio de unión definido por AlaScan de NPC1	
PubChem SID: 24805285	
Nomenclatura IUPAC: 3,17-dimethyl-7,13,20,25-tetraoxatetraspiro[5.2.1.2.5 ¹⁴ .2 ¹¹ .2 ⁹ .2 ⁶]pentacosan-10-one	
Ligando contra el sitio de unión a colesterol de NPC2	
PubChem SID: 17472624	
Nomenclatura IUPAC: N-(1-adamantyl)-5-oxo-1-phenylpyrrolidine-3-carboxamide	

Tabla 10. Ligandos del cribado contra dianas que cumplen con los requisitos ADME-T especificados. AlaScan refiere a cavidad definida mediante *alanine scanning*

Diana	NPC1	NPC2
Sitio de unión	Definido por AlaScan	Unión a colesterol
PubChem SID	24805285	17472624
Absorción intestinal	97.712%	94.606%
VDss	0.39	0.579
Fracción libre	0.265	0.068
Permeabilidad de la BBB	-0.022	0.269
Permeabilidad del CNS	-2.81	-2.119
Dosis máxima tolerada	-0.375	-0.273
Toxicidad oral aguda	2.183	2.698
Toxicidad oral crónica	0.166	1.439

Tabla 11. Propiedades farmacocinéticas de los ligandos del cribado virtual contra dianas que pasaron los requerimientos establecidos. AlaScan refiere a cavidad definida mediante *alanine scanning*. Donde, VDss está expresado como log VDss (volumen de distribución de estado estacionario). Permeabilidad de la BBB expresada como log BB (logaritmo de la ratio de concentración del cerebro al plasma sanguíneo). Permeabilidad del CNS expresado como Log PS (producto de la permeabilidad de la barrera hemato-encefálica y el área de superficie). Dosis máxima tolerada expresada en log (mg/kg/día). Toxicidad oral aguda (rata) expresada como el LD50 (mol/kg). Toxicidad oral crónica (rata) expresada en log (mg/kg_bw/día).

Con el primer análisis de farmacocinética se vio que los parámetros de toxicidad son muy exclusivos. Por ejemplo, una molécula puede tener distintos grados de hepatotoxicidad, y este análisis no lo contempla. Por lo tanto, se realizó otro filtraje de ligandos, pero solo teniendo en cuenta las propiedades ADME. Dado que excluir la toxicidad de los requerimientos haría el filtraje mucho más permisivo, se decidió incrementar los requerimientos de distribución de la molécula a los valores óptimos:

- Absorción intestinal > 90%
- **Permeabilidad a la BBB ≥ 0.3**
- **Permeabilidad al CNS ≥ -2**
- Sustrato de la glicoproteína P: NO
- **VDss > 0**

Cabe recordar que en el anterior filtraje se evaluaron moléculas con una capacidad mínima de cruzar la BBB. En el segundo filtraje se buscan moléculas con un potencial óptimo de difundir en el CNS. La posible ventana terapéutica de estos ligandos debería ser menor, pero su posible efecto terapéutico mayor.

Dos ligandos cumplen todos los requerimientos (Tabla 12) y la predicción de sus propiedades farmacocinéticas se resume en la Tabla 13.

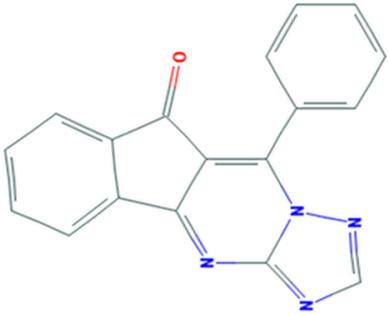
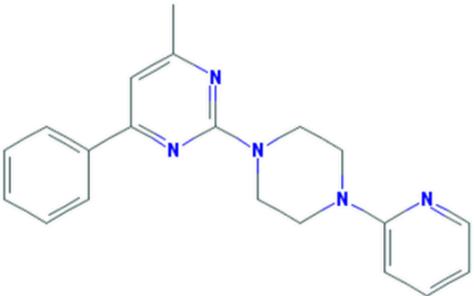
Ligando contra la unión CTD-NTD de NPC1	
PubChem SID: 847986	
Nomenclatura IUPAC: 10-Phenyl-1,3,4,10a-tetraaza-cyclopenta[b]fluoren-9-one	
Ligando contra la unión CTD-NTD de NPC1	
PubChem SID: 844825	
Nomenclatura IUPAC: 4-methyl-6-phenyl-2-(4-pyridin-2-yl)piperazin-1-yl)pyrimidine	

Tabla 12. Ligandos del cribado contra dianas que cumplen con los requisitos ADME del segundo filtro.

Diana	NPC1	NPC1
Sitio de unión	Unión CTD-NTD	Unión CTD-NTD
PubChem SID	847986	844825
Absorción intestinal	100%	99.969%
VDss	0.167	0.099
Fracción libre	0.291	0.199
Permeabilidad de la BBB	0.442	0.496
Permeabilidad del CNS	-1.954	-1.826
Dosis máxima tolerada	0.392	-0.372
Toxicidad oral aguda	2.296	2.465
Toxicidad oral crónica	0.882	0.79

Tabla 13. Resumen de las propiedades de absorción, distribución, toxicidad de los ligandos que cumplen las condiciones de absorción y distribución establecidas en el segundo filtro. Dónde, VDss está expresado como log VDss (volumen de distribución de estado estacionario). Permeabilidad de la BBB expresada como log BB (logaritmo de la ratio de concentración del cerebro al plasma sanguíneo). Permeabilidad del CNS expresado como Log PS (producto de la permeabilidad de la barrera hemato-encefálica y el área de superficie). Dosis máxima tolerada expresada en log (mg/kg/día). Toxicidad oral aguda (rata) expresada como el LD50 (mol/kg). Toxicidad oral crónica (rata) expresada en log (mg/kg_bw/día).

En este caso, los ligandos también tienen una baja dosis máxima tolerada predicha. Además, ambos son potencialmente hepatotóxicos y el ligando SID: 847986 también parece tener potencial mutagénico (positivos para toxicidad AMES).

Se podría rebajar los requerimientos del filtraje para que salieran más moléculas prometedoras, pero esto muy probablemente solo serviría para desperdiciar recursos ya que se proseguirían los experimentos con más candidatos, pero con peores cualidades ADME-T. Aun así, se incluyen los resultados del cribado y del análisis de farmacocinética precisamente por si en un futuro se desea analizar los datos de otra forma, incluyendo más ligandos en el estudio, o priorizando otras cualidades farmacocinéticas.

Finalmente, hay que puntualizar que los ligandos de NPC1L1 tienen requisitos farmacocinéticos distintos. NPC1L1 está expresado en la mucosa intestinal, por lo que los ligandos no deben cruzar la BBB y tampoco precisan de una buena distribución sistémica. En consecuencia, el análisis farmacocinética de los ligandos contra NPC1L1 solo requeriría filtrar por toxicidad: toxicidad AMES, hepatotoxicidad, inhibición de hERG I/II. Lamentablemente, ningún ligando contra NPC1L1 cumple todos los requerimientos del filtro. No obstante, hay tres ligandos con características prometedoras (Tabla 14)., pero de los que se predice hepatotoxicidad o toxicidad AMES (Tabla 15).

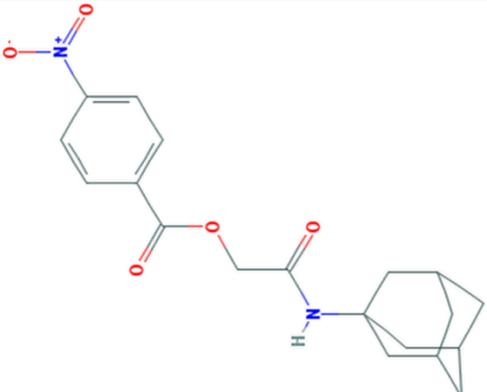
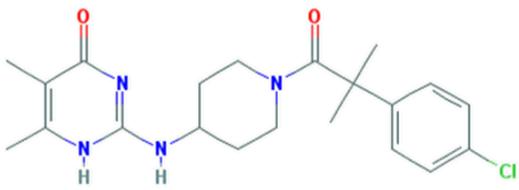
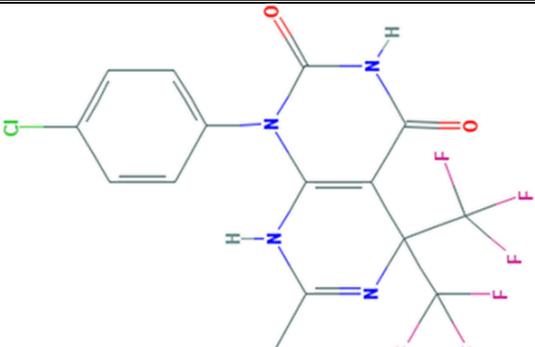
Ligando contra la abertura de la cavidad de unión a ligandos de NPC1L1	
PubChem SID: 85268288	
Nomenclatura IUPAC: [2-(1-adamantylamino)-2-oxoethyl] 4-nitrobenzoate	
Ligando contra la cavidad de unión a ligandos de NPC1L1	
PubChem SID: 26614584	
Nomenclatura IUPAC: 2-[[1-[2-(4-chlorophenyl)-2-methylpropanoyl]piperidin-4-yl]amino]-5,6-dimethyl-1H-pyrimidin-4-one	
Ligando contra la cavidad de unión a ligandos de NPC1L1	
PubChem SID: 24295645	
Nomenclatura IUPAC: 1-(4-chlorophenyl)-7-methyl-5,5-bis(trifluoromethyl)-8H-pyrimido[4,5-d]pyrimidine-2,4-dione	

Tabla 14. Ligandos del cribado contra NPC1L1 que cumplen la mayoría de los requerimientos del filtro. Para más detalles referirse a la Tabla 15.

Como muestra la Tabla 15, la ventana terapéutica de estas moléculas con potencial farmacológico es muy estrecha, la dosis máxima tolerable es cercana a cero (negativo en escala logarítmica). No obstante, el ligando 85268288 se espera que requiera de transportadores en sangre para distribuirse (fracción libre = 0), y el ligando 24295645 tiene un volumen de distribución cercano a cero también. Estas características de mala distribución podrían aprovecharse en un proceso de optimización para reducir la propagación de los ligandos por el cuerpo y así mitigar su efecto tóxico.

Diana	NPC1L1	NPC1L1	NPC1L1
Sitio de unión	Abertura de la cavidad	Cavidad de unión a ligandos	Cavidad de unión a ligandos
Ligando	85268288	26614584	24295645
Absorción intestinal	95.795%	94.299%	87.781%
VDss	0.204	0.216	-0.533
Fracción libre	0	0.239	0.141
Permeabilidad de la BBB	-0.595	-1.168	-1.581
Permeabilidad del CNS	-2.465	-3.246	-3.337
Dosis máxima tolerada	-0.249	-0.24	-0.037
Hepatotoxicidad	NO	SI	SI
Toxicidad AMES	SI	NO	NO
Toxicidad oral aguda	2.796	2.424	2.092
Toxicidad oral crónica	2.115	0.952	1.046

Tabla 15. Propiedades ADME-T de las moléculas del cribado de ligandos contra NPC1L1 con mayor potencial terapéutico. Dónde, VDss está expresado como log VDss (volumen de distribución de estado estacionario). Permeabilidad de la BBB expresada como log BB (logaritmo de la ratio de concentración del cerebro al plasma sanguíneo). Permeabilidad del CNS expresado como Log PS (producto de la permeabilidad de la barrera hemato-encefálica y el área de superficie). Dosis máxima tolerada expresada en log (mg/kg/día). Toxicidad oral aguda (rata) expresada como el LD50 (mol/kg). Toxicidad oral crónica (rata) expresada en log (mg/kg_bw/día).

3.7.2 Cribado de ligandos por similitud

El análisis farmacocinético de los 5 primeros *hits* de cada cribado por similitud con *SwissADME* se empezó con el esquema *Boiled Egg*, en el archivo adjunto [[ADME_BoiledEgg_Cribado_Similitud.pdf](#)], y resultó que ninguno de los 15 ligandos estudiados cruzaba la barrera hemato-encefálica. Como consecuencia, se amplió el estudio de la permeabilidad de la BBB al top 50 de los cribados por similitud de HP β CD y Miglustat. Cabe recordar que la diana de ezetimibe es NPC1L1 que se expresa en la membrana apical de los enterocitos, por lo que fármacos derivados de ezetimibe no deben cruzar la BBB.

El esquema *Boiled Egg* de los ligandos 38 a 41 del cribado en base a HP β CD muestra que son permeables a la BBB. Por lo tanto, se continuó el estudio farmacocinético con ellos, obteniéndose los resultados mostrados en la Tabla 16:

Ligando	ZINC77291447	ZINC43198724	ZINC43198723	ZINC43198722
Absorción intestinal	Alta	Alta	Alta	Alta
Solubilidad H ₂ O	Soluble	Soluble	Soluble	Soluble
Substrato P-gp	Si	Si	Si	Si
Inhibidor CYP450	No	No	Inhibidor de CYP3A4	No
Druglikeness	Si	Si	Si	Si
Química médica	Correcta	Correcta	Correcta	Correcta
Leadlikeness	Si	Si	Si	Si
Toxicidad AMES	No	No	No	No
Inhibidor de hERG I/II	No	No	No	No
Hepatotoxicidad	Si	Si	Si	Si
Dosis máxima tolerada	0.059	-0.137	-0.137	-0.137
Toxicidad oral aguda	2.62	2.317	2.317	2.317
Toxicidad oral crónica	1.079	1.062	1.062	1.062

Tabla 16. Farmacocinética de los ligandos del cribado por similitud de HP β CD que son permeables a la BBB, dónde: Dosis máxima tolerada expresada en log (mg/kg/día). Toxicidad oral aguda (rata) expresada como el LD50 (mol/kg). Toxicidad oral crónica (rata) expresada en log (mg/kg_bw/día).

Este resultado se considera positivo pese a que los ligandos sean hepatotóxicos y tengan una dosis máxima tolerada baja. Esto es debido a que se espera que puedan cruzar la BBB. HP β CD, el compuesto del que derivan, no puede cruzarla y no se sabe mediante que mecanismo realiza su efecto terapéutico directo sobre la enfermedad. Moléculas similares, pero con la capacidad de cruzar al CNS tienen un gran potencial como terapia para NP-C.

Sin embargo, ningún ligando del top 50 del cribado por similitud en base a Miglustat resultó permeable a la BBB. Debido a la baja puntuación de similitud de los ligandos a partir del número 50, se detuvo el análisis de permeabilidad a través de la BBB. Por falta de tiempo no se pudo profundizar en el análisis de la metodología empleada para determinar si los resultados negativos fueron fruto de un error humano y se concluyó aquí su estudio.

En cambio, la permeabilidad de la BBB no es ningún problema para los ligandos del cribado por similitud a ezetimibe como se ha explicado anteriormente. Por lo tanto, los ligandos del cribado en base a ezetimibe se evaluaron directamente en función de su farmacocinética (Tabla 17). Desafortunadamente, los cinco ligandos son potencialmente tóxicos ya que resultaron positivos para el test de inhibición de hERG II, el de hepatotoxicidad, y el de inhibición de varias isoformas de CYP450. Con estas predicciones de

toxicidad, no se pueden calificar los ligandos similares a ezetimibe como moléculas con potencial terapéutico.

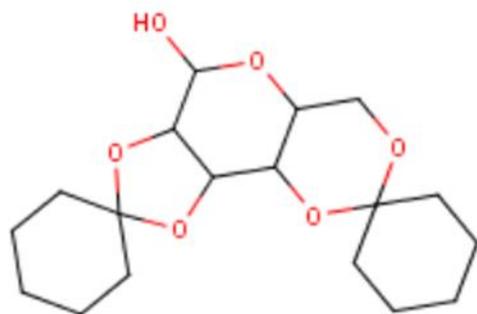
Para más detalle sobre el análisis de farmacocinética de los ligandos del cribado por similitud referirse al archivo adjunto [[ADME-Cribado_por_Similitud.ods](#)]. Y para más detalles sobre el análisis de toxicidad de los mismos ligandos, consultar [[Toxicidad_Cribado_por_Similitud.ods](#)].

Ligando	ZINC00592281	ZINC09223145	ZINC09211540	ZINC77287132	ZINC03810860
Absorción intestinal	Alta	Alta	Alta	Alta	Alta
Solubilidad H ₂ O	Moderada	Moderada	Moderada	Moderada	Moderada
Substrato P-gp	Si	Si	Si	Si	Si
Inhibidor CYP450	Inhibidor de CYP2C19/CYP2D6 CYP3A4	Inhibidor de CYP2C19/CYP2D6 CYP3A4	Inhibidor de CYP2C19/CYP2D6 CYP3A4	Inhibidor de CYP2C19/CYP2C9 CYP2D6/CYP3A4	Inhibidor de CYP2C19/CYP2D6 CYP3A4
Druglikeness	Si	Si	Si	Si	Si
Química médica	Correcta	Correcta	Correcta	Correcta	Correcta
Leadlikeness	No	No	No	No	No
Toxicidad AMES	No	No	No	No	No
Inhibidor de hERG I/II	Inhibidor de hERG II	Inhibidor de hERG II			
Hepatotoxicidad	Si	Si	Si	Si	Si
Dosis máxima tolerada	0.056	0.056	0.056	0.068	0.056
Toxicidad oral aguda	2.633	2.633	2.633	2.569	2.633
Toxicidad oral crónica	1.197	1.197	1.197	1.249	1.197

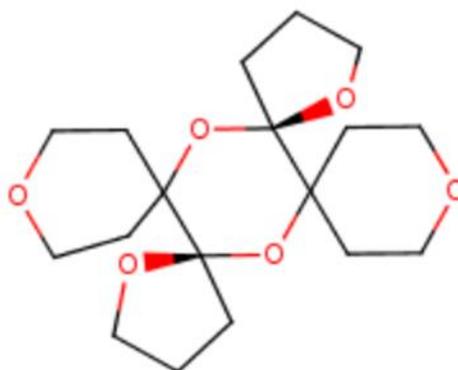
Tabla 17. Análisis farmacocinético de los *hits* del cribado por similitud en base a ezetimibe. Todos los compuestos son inhibidores de hERG II, interaccionan con isoformas de CYP450.

En conclusión, del cribado por similitud solo cuatro ligandos similares a HP β CD (Fig 20) cumplen los requisitos establecidos para considerarlos moléculas con potencial terapéutico para NP-C.

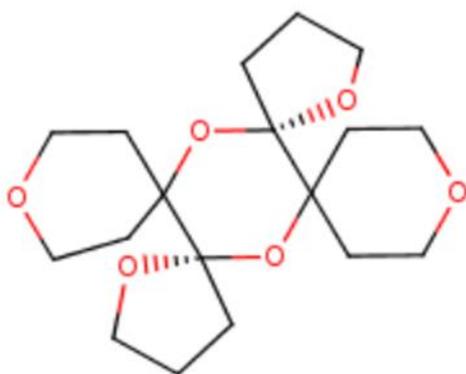
A ZINC77291447
Score : 0.784



B ZINC43198724
Score : 0.784



C ZINC43198723
Score : 0.784



D ZINC43198722
Score : 0.784

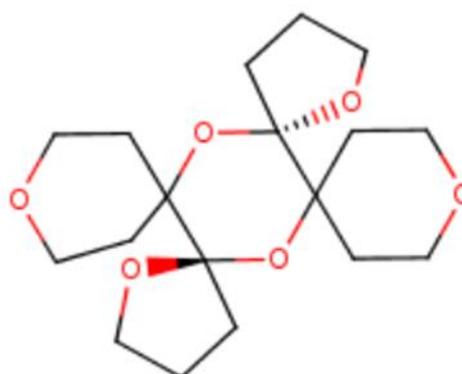


Fig 20. Representación de los cuatro ligandos similares a HP β CD con potencial terapéutico. De cada molécula consta: el identificador en la base de datos ZINC12, la puntuación de similitud del cribado a HP β CD y el esquema molecular del ligando.

A 2,3:4,6-Di-O-cyclohexylidene- α -D-mannopyranose

B (5R,13R)-4,9,12,17,21,24-hexaoxatetrspiro[4.0.5^{6}.1.4^{13}.0.5^{18}.1^{5}]tetracosane

C (5S,13S)-4,9,12,17,21,24-hexaoxatetrspiro[4.0.5^{6}.1.4^{13}.0.5^{18}.1^{5}]tetracosane

D (5R,13S)-4,9,12,17,21,24-hexaoxatetrspiro[4.0.5^{6}.1.4^{13}.0.5^{18}.1^{5}]tetracosane

4. Discusión

La discusión de los resultados se divide en los tres bloques del trabajo:

4.1 Bloque 1:

Las mutaciones de NPC1 que se analizaron debían cumplir dos requisitos: ser una mutación puntual y causar NP-C. Este tipo de mutaciones normalmente se describen en artículos de tipo clínico, por lo que hay mucha información sobre el cuadro clínico que causan, pero poca sobre sus efectos en la proteína. Estudios moleculares previos sobre mutaciones en NPC1 muestran que la mayoría de mutaciones causantes de NP-C (50%) se encuentran en el dominio CTD (27). Este dominio es el encargado de mantener los otros dos dominios (NTD y MDL) en una conformación óptima para su función. Mutaciones puntuales en este dominio se espera que puedan inducir desalineamientos de las estructuras proteicas.

En este proyecto se planteó estudiar el efecto de las mutaciones más comunes sobre la estructura de NPC1 y su función: unión a colesterol. Las mutaciones en cuestión: G992W - I1061T - N1156S - P1007A - Q775P, se sitúan en el dominio CTD o en las regiones transmembrana próximas con la excepción de Q775P que localiza en el dominio SSD de unión a esteroides.

Analizando las energías de unión del experimento de acoplamiento, parece ser que los mutantes y NPC1-WT unen colesterol con igual intensidad. Pero, al observar con detenimiento el modelo de interacción que *MitAutoDock* genera, se aprecian diferencias claras. En los mutantes N1156S y P1007A, el colesterol se une preferentemente a una zona mucho más interior de la cavidad. Este cambio de zona de unión debido a la mutación podría dificultar el último paso de translocación del colesterol al citosól y explicar cómo estas mutaciones causan NP-C. Interferir con la translocación de colesterol al otro lado de la bicapa lipídica podría causar un bloqueo en el transporte de colesterol al citosól en su último paso. Para poder validar esta hipótesis hace falta entender todo el proceso de eflujo controlado de colesterol del lisosoma al citosól. Y también habría que validar los resultados de *docking* con otro *software*.

Algunas de las mutaciones más comunes crean problemas durante el plegamiento de NPC1 e inducen su degradación antes de que transloque a la membrana lisosomal, como por ejemplo I1061T (4,42), o Q775P (43). Está claro que no hacía falta estudiar el *docking* de colesterol en este caso, porque NP-C sería causado por la falta de proteína en el lisosoma y no por la baja actividad de NPC1. Pero, se creyó oportuno generar el modelo por homología y aprovechar para ver el acoplamiento para estudiar los efectos de dichas mutaciones en la estructura de NPC1. Según los resultados obtenidos, la estructura de NPC1-I1061T y NPC1-Q775P no varía mucho de la WT, hecho que no explica su degradación prematura. Puede ser que el modelo no sea lo suficientemente robusto como para que se vean los efectos de un cambio de un aminoácido. También podría ser que las mutaciones en este caso solo

creen un gran cambio de estructura durante el plegamiento de la proteína, pero no sobre la predicción de la proteína completa y plegada.

Igualmente, los modelos generados de NPC1 con mutaciones podrán usarse en un futuro para analizar los efectos de las mutaciones sobre la actividad enzimática de NPC1 con más detalle. Por ejemplo, se podría hacer experimentos de acoplamiento entre los mutantes de NPC1 y moléculas con potencial terapéutico contra NPC1 como las que se han obtenido en este trabajo. Basándose en los resultados obtenidos en este trabajo, sería muy interesante ver los efectos de las mutaciones que han causado diferencias de *docking*, N1156S y P1007A, sobre la unión de ligandos. Este sencillo experimento podría aportar información sobre los defectos funcionales de NPC1 que causan NP-C.

En proyectos futuros también se debería investigar las mutaciones en NPC2. En este trabajo, no se ha tratado el tema por la dificultad de encontrar información sobre la incidencia de las distintas mutaciones. Probablemente debido a lo raras que son (solo el 5% de los casos), es difícil hacer estudios poblacionales óptimos. Pero esto no hace el estudio de NPC2 menos importante, justo al contrario. ¿Por qué hay tanta diferencia entre la prevalencia de mutaciones en un gen u otro? Podría ser que NPC2 tenga mayor importancia que NPC1 para la viabilidad o el desarrollo de las células. Por ejemplo, al ser soluble, podría tener una función más amplia que NPC1 y donar lípidos a varias proteínas de membrana del lisosoma. En este contexto, mutaciones en el gen de NPC2 podrían ser letales en estadios embrionarios. Esta teoría se podría contrastar con estudios de actividad enzimática de NPC2 mutado: si las mutaciones de NPC2 inhiben completamente su actividad y se encuentran en homocigosis, entonces se rechaza la hipótesis de que mutaciones de NPC2 puedan ser letales durante el desarrollo embrionario.

4.2 Bloque 2:

Las dianas para cribado se eligieron de dos formas: proteínas directamente implicadas en la vía afectada (NPC1, NPC2, OSBPL5) y proteínas involucradas en el transporte de colesterol a nivel sistémico (NPC1L1, LXR). Un fármaco contra el primer grupo sería lo más deseable, ya que incidiría directamente sobre la raíz del problema y seguramente presentaría menos efectos adversos. El principal problema de este grupo es que las cualidades farmacocinéticas esperadas de los ligandos son muy específicas: que el ligando fuera liposoluble y cruzara la barrera hemato-encefálica (BBB), pero a la vez no causara toxicidad ni se acumulase en tejidos. El segundo grupo en cambio, presenta dianas más accesibles, y por lo tanto los ligandos tienen menos restricciones respecto a sus cualidades farmacocinéticas.

Por todo lo expuesto en el párrafo anterior, en un principio se esperaba que el paso más complicado o tedioso del proyecto fuera el análisis farmacocinético. En cambio, resultó serlo la detección de cavidades. La idea inicial era definir 3 cavidades de cada diana con *fpocket* y realizar el cribado sobre éstas. De esta forma, todas las dianas eran tratadas de forma

homogénea. Pero el cribado de ligandos contra dianas daba errores, *MitOpenScreen* no reconocía las cavidades. Como hay un paso intermedio manual de traducción de las cavidades en *pdb* a formatos que reconozca *MitOpenScreen* (aminoácidos que forman la cavidad y coordenadas del centro y límite de la cavidad), al principio se pensó que la definición de cavidades por aminoácidos no fuera suficientemente precisa. Por lo que se extrajeron las coordenadas de las cavidades y se volvieron a realizar los cribados, obteniéndose resultados parecidos. El problema resultó ser el tamaño de las cavidades generadas por *fpocket*. El programa definía cavidades tan grandes que, o bien los cribados de ligandos no eran específicos, o bien no se podían realizar los cribados. Finalmente, tras probar otros *softwares* como *RaptorX* y *CASTp*, se llegó a la conclusión que este tipo de programas dan resultados muy dispares. Además, NPC1 y NPC1L1 tienen una gran cavidad interior necesaria para el transporte de lípidos, que confundía los programas de detección de cavidades.

Al mismo tiempo, había problemas con la estructura de NPC1-5u74. Esta debería ser la versión cristalografiada más reciente y precisa de NPC1 humano, pero *MitOpenScreen* y *MitAutoDock* no la reconocían.

Al final, coincidiendo con las pruebas de otros *softwares* de detección de cavidades, se decidió volver a la literatura. Por un lado, se definieron cavidades basadas en la literatura, y por el otro se detectó una mutación en la secuencia de NPC1-5u74 y se cambió por NPC1-3jd8 (versión anterior, pero de alta definición igual).

Las cavidades definidas en la literatura deberían ser robustas, porque han sido definidas experimentalmente. Pero, también pueden estar incompletas, ya que en los artículos de biología molecular se definen aminoácidos que interaccionan, no cavidades de interacción tridimensionales. En ese sentido, el *alanine scanning* resultó ser de gran ayuda ya que permite extraer todos los aminoácidos de un sitio de unión y además los ordena según su importancia en la interacción.

El segundo punto de tensión en este bloque fue el análisis de farmacocinética. Ya se esperaba que pocos *hits* resultaran óptimos, pero solo dos cumplieron los requisitos: 24805285 contra NPC1 y 17472624 contra NPC2. Aunque ya se sabía que era difícil identificar moléculas con tales características, se esperaba obtener al menos un *hit* por diana. Un compuesto liposoluble, que cruce la BBB, se distribuya con facilidad y no cause efectos adversos ni interacciones no es fácil de diseñar, pero en este caso hay un problema añadido: la enfermedad es genética, y por lo tanto el tratamiento probablemente será crónico. Por esto se puso tanto énfasis en la toxicidad al definir los requisitos, porque las moléculas liposolubles se acumulan con facilidad en los tejidos y pueden causar toxicidad rápidamente.

Sin embargo, cuando se quitaron los requerimientos de toxicidad del filtro y se aplicó un filtro ADME más restrictivo, se obtuvieron dos nuevos ligandos contra NPC1: 847986, 844825. Este grupo de ligandos debería tener unas características farmacocinéticas superiores, con mayor probabilidad de efecto terapéutico. No obstante, habría que validar su toxicidad experimentalmente y mediante otros *softwares*. Ligandos con una ventana terapéutica tan estrecha

difícilmente se pueden usar como fármaco de una terapia crónica, pero podrían ser optimizados.

De los ligandos contra NPC1, el ligando más seguro (baja toxicidad), 24805285, reconoce la cavidad de unión a colesterol y los otros dos: 847986 y 844825, se unen a la estructura de unión entre los dominios CTD y NTD de NPC1. En el contexto de NP-C, 24805285 podría estabilizar la cavidad de unión a colesterol de NPC1 y favorecer su actividad, o bloquearla y disminuir aún más, si es posible, el transporte de colesterol. Para evaluar más en profundidad su potencial terapéutico, habría que realizar experimentos *in vitro* y ver qué clase de efecto tiene. Por otro lado, las otras dos moléculas tienen perspectivas prometedoras como potenciales fármacos. Al reconocer la interacción entre los dos dominios luminales de NPC1 sin interactuar directamente con los sitios de unión a colesterol, podrían estabilizar la estructura de NPC1 sin interferir en el transporte de colesterol directamente. Dado que la mayoría de mutaciones de NPC1 se dan en el dominio CTD, se cree que las mutaciones causan un cambio conformacional de los dominios luminales de NPC1 que dificultan no solo el reconocimiento de colesterol unido a NPC2 sino también su transporte al interior de la membrana (27). Es por esto, que estos dos ligandos son tan prometedores, porque podrían estabilizar la estructura de NPC1 en su conformación más óptima sin interferir en su unión a colesterol o NPC2. En un futuro, debería hacerse experimentos de *docking* con estos ligandos contra los modelos de NPC1 mutado, para estudiar el efecto de la interacción bajo estas condiciones. Si la interacción ligando – NPC1 mutado sigue siendo fuerte, se podría verificar *in vitro* si hay algún tipo de rescate de la función.

El otro ligando que cumple los requisitos marcados de molécula con potencial terapéutico para NP-C es 17472624, que reconoce el sitio de unión a colesterol de NPC2. Primero hay que destacar que probablemente este ligando solo serviría a la minoría de pacientes con mutaciones en NPC2. Después añadir que, como la primera molécula activa contra NPC1 descrita, habría que realizar experimentos *in vitro* para valorar correctamente el efecto del ligando.

Los ligandos obtenidos son prometedores, pero hay un problema, y es que NPC1, NPC2 y todo fármaco contra ellos deben coexistir en el lumen lisosomal. Por lo tanto, habría que comprobar que los ligandos seleccionados pueden mantener su estructura en condiciones ácidas, u optimizarlos para que lo hagan. Estas propiedades no se cubren con los programas informáticos comunes, por lo que deberían desarrollarse nuevos softwares o estudiar la reacción al cambio de pH de las moléculas *in vitro*.

Finalmente, también se ha analizado los ligandos del cribado contra NPC1L1 a parte, porque tienen requerimientos distintos a los demás ya que su diana se encuentra en la mucosa del intestino delgado. La mejor característica para un ligando efectivo contra NPC1L1 sería que no fuera absorbido en el intestino. De esta forma solo bloquearía el receptor y sería excretado con las heces. El resultado del cribado en este caso, no han sido ligandos con mala absorción sino ligandos con mala distribución. Aun que podría investigarse más estos ligandos para optimizarlos, sería mucho más sencillo si simplemente no fueran

absorbidos. No habría complicaciones de toxicidad o interacción con otros fármacos.

Después de obtener resultados parciales durante el proyecto, se planteó hacer un análisis más manual de los resultados del cribado contra NPC1L1 y buscar ligandos con baja absorción intestinal, pero debido a restricciones del calendario fue imposible de realizar. Resta para futuros experimentos centrar el análisis del cribado en la baja absorción intestinal.

En conclusión, se han obtenido 2 ligandos con características prometedoras y 2 más con potencial activo pero una estrecha ventana terapéutica. Además, el análisis farmacocinético de los cribados se podría optimizar para cada diana en concreto y probablemente se obtendrían muchas más moléculas con potencial terapéutico.

4.3 Bloque 3:

El bloque 3 ha sido mucho más fácil de realizar en comparación que el bloque 2. Que *SwissSimilarity* y *SwissADME* estén relacionados ha facilitado y acelerado el análisis de la farmacocinética de los ligandos. Estas ventajas se aprovecharon, no solo para agilizar el proceso, sino también para familiarizarse con otro programa de predicción de propiedades farmacocinéticas. Al final, todos los programas hacen predicciones y pueden ser más fuertes en un sentido u otro. Por eso, se acostumbra a consultar distintos programas que realicen la misma función para luego tomar una decisión informada y en consenso con los distintos resultados obtenidos. Durante este trabajo no hubo tiempo, pero en el siguiente paso, se analizaría los resultados de los cribados con los dos programas de análisis farmacocinético en paralelo, para tener una idea más clara de las propiedades ADME-T de los *hits*.

Lamentablemente, con el análisis de *SwissADME* no se encontraron ligandos similares a Miglustat con capacidad de cruzar la BBB. Ampliar el estudio a los primeros 50 *hits* tampoco cambió nada y además los ligandos ya no tenían una puntuación de similitud óptima y se tuvo que dejar el estudio aquí. Habría que volver a realizar todo el pipeline, porque un resultado tan negativo es muy inesperado. Puede que en algún momento se usara un archivo erróneo o se seleccionara un parámetro equivocado, porque Miglustat tiene efectos probados en el CNS, y moléculas similares deberían poder cruzar la BBB también.

Con el estudio de compuestos similares a HP β CD también se tuvo que ampliar los ligandos estudiados hasta el top50. Pero en este caso hubo resultados positivos y se encontraron 4 compuestos capaces de cruzar la BBB y con una similitud con HP β CD de 0.784. Como no se sabe con seguridad si HP β CD cruza la BBB, estos resultados son muy esperanzadores, ya que permitirán estudiar con más facilidad el efecto de HP β CD en NP-C. De los 4 compuestos, ZINC77291447, ZINC43198724 y ZINC43198722 son los más seguros ya que no interaccionan con isoformas de CYP450, y por lo tanto podrían tomarse en paralelo con otros fármacos sin complicaciones. La baja dosis máxima tolerada predicha que tienen puede ser una complicación, pero al ser sustancias

liposolubles, se pueden tomar en bajas dosis y se acumularán en los tejidos. Con un buen estudio farmacocinético y una dosificación controlada, tienen potencial para convertirse en fármacos efectivos.

Por otro lado, con los ligandos similares a ezetimibe, si encontráramos un ligando que no fuera absorbido sería perfecto. Los resultados del cribado de ligandos similares a ezetimibe son compuestos activos pero muy tóxicos y con interacciones múltiples con citocromos hepáticos. Estos compuestos, por lo tanto, no pueden usarse como fármaco y no tienen potencial terapéutico. Estos mismos ligandos serían útiles si se combinaran con moléculas que no fueran absorbibles. Por ejemplo, se podrían modificar con técnicas de nanotecnología para aumentar su tamaño, o añadirles una cola inerte en el extremo no activo que impidiera su absorción.

En este cribado se pudo realizar una validación más al conocer la diana del ligando original: acoplamiento de los ligandos con NPC1L1. En teoría este debería ser el experimento más valioso para determinar el potencial de los *hits* como fármacos, ya que se está probando su interacción contra la diana específica. Pero, las energías de unión obtenidas en este caso fueron muy débiles (valores de dG poco negativos), incluso en el caso de ezetimibe. Que ezetimibe tenga valores de dG poco negativos puede ser indicativo de que la cavidad utilizada para el acoplamiento no sea la correcta, o este solo parcialmente definida. Habría que volver a realizar una búsqueda bibliográfica, para ver si se encuentran artículos en que se valide la cavidad utilizada, o en que se defina con más detalle.

Como conclusión final de los bloques 2 y 3, se han obtenido ligandos prometedores en ambos bloques, y se ha definido métodos para mejorar los resultados obtenidos. Añadir, que aun que los bloques partan de información distinta (ligando – diana), los métodos usados y los pasos a seguir son parecidos y se pueden combinar. Por ejemplo, un análisis ADME-T cruzado con *SwissADME* y *pkCSM* daría resultados mucho más robustos y fiables que solo usando uno de los métodos. Esto sería muy importante en proyectos que van más allá de lo planteado en este trabajo y buscan optimizar *leads* y hacer experimentos *in vitro*. Cuando más fiable la información de inicio, más rápido y fácil se adquiere el objetivo.

5. Conclusiones

Conclusiones generales

Con el diseño de moléculas con potencial terapéutico aplicado en este trabajo se estableció la base para generar fármacos potenciales a partir de ocho moléculas. No obstante, sus resultados distan mucho de la aplicación clínica o investigación experimental. Hace falta re-validar los resultados con diferentes programas, y especializar la metodología en algunos casos para optimizar los *hits* obtenidos a *leads*.

El diseño de moléculas con un efecto terapéutico a partir de dianas o fármacos son métodos muy parecidos, difieren en el objeto de estudio inicial, pero la lógica detrás de ambos procesos, y sus pasos son muy parecidos.

Aun que se trabaje con modelos computacionales y cribados virtuales, no hay que olvidar que se trabaja en el campo de la biología, y que se puede recurrir a la extensa bibliografía de estudios experimentales para dirigir de forma manual este tipo de procesos informáticos.

Cumplimiento de los objetivos

Los objetivos se han logrado mayoritariamente. Se ha logrado identificar moléculas con potencial terapéutico y analizar el efecto de las mutaciones de NPC1, pero se podría mejorar y adaptar las metodologías utilizadas. En general, no ha habido tiempo suficiente para poder profundizar debidamente en los distintos objetivos del trabajo. Un buen ejemplo es el caso del cribado con Miglustat, donde no ha habido tiempo de contrastar si en algún punto del pipeline ha habido un error de procedimiento.

Planificación del trabajo y metodología emprada

En general se ha seguido la planificación del curso. Se dio un breve retraso debido a las complicaciones para determinar las cavidades de las dianas en el bloque 2. La metodología prevista no fue adecuada, y se tuvo que probar diversas alternativas. Como consecuencia se invirtió más tiempo del esperado en esta etapa del proyecto. Aun así, como se previó que podría haber complicaciones debido a la variedad de métodos utilizados, se planteó el calendario de forma que se pudiera trabajar en paralelo en los distintos bloques. Esta distribución permitió recuperar tiempo después del informe de fase 1 cuando se detectó el problema.

Líneas de trabajo futuro

Bloque 1:

Investigar el efecto de las mutaciones de NPC2 sobre la proteína
Hacer un análisis más exhaustivo de las mutaciones de NPC1 y posible investigación de su efecto en la proteína

Bloque 2:

Investigar el resultado de los cribados más a fondo y analizar más *hits*.
“Personalizar” la selección de ligandos del cribado en función de su farmacocinética para obtener más *hits* a optimizar

Bloque 3:

Repetir el cribado por similitud de Miglustat
Realizar un análisis ADME-T más extenso de los ligandos similares a HP β CD
Investigar en detalle el sitio de unión de ezetimibe, para definir mejor el sitio de acoplamiento.

6. Glosario

NP-C: enfermedad de Niemann Pick tipo C

BBB: barrera hemato-encefálica

CNS: sistema nervioso central

NPC1: proteína Niemann Pick C 1

NPC2: proteína Niemann Pick C 2

NPC1L1: proteína Niemann Pick C 1 Like 1

LXR: Liver X Receptor

OSBPL5: Oxy-Sterol Binding Protein-Like 5

HP β CD: hydroxypropyl β ciclodextrin

ADME: absorción, distribución, metabolismo, excreción

ADME-T: absorción, distribución, metabolismo, excreción – toxicidad

7. Bibliografía

1. JE W, , Guffon N, Rohrbach M, Hwu WL, Korenke GC, Bembi B, Luzy C, Giorgino R SF. Natural history of Niemann-Pick disease type C in a multicentre observational retrospective cohort study. *Mol Genet Metab*. 2009;98(3):250–4.
2. Vanier MT. Niemann–Pick diseases. In: *Handbook of Clinical Neurology*. 2013. p. 1717–21.
3. Vanier M, Millat G. Niemann-Pick disease type C. *Clin Genet* [Internet]. 2003;64(4):269–81. Available from: <http://doi.wiley.com/10.1034/j.1399-0004.2003.00147.x>
4. Fernandez-Valero EM, Ballart A, Iturriaga C, Lluch M, Macias J, Vanier MT, et al. Identification of 25 new mutations in 40 unrelated Spanish Niemann-Pick type C patients: Genotype-phenotype correlations. *Clin Genet*. 2005;68(3):245–54.
5. Pentchev PG, Comly ME, Kruth HS, Vanier MT, Wenger DA, Patel S, et al. A defect in cholesterol esterification in Niemann-Pick disease (type C) patients. *Proc Natl Acad Sci U S A* [Internet]. 1985;82(23):8247–51.
6. Lyseng-Williamson KA. Miglustat: A review of its use in Niemann-Pick disease type C. *Drugs*. 2014;74(1):61–74.
7. Institute of Medical Genetics and Applied Genomics UKT. NPC-db2 Niemann-Pick Type C Database [Internet]. [cited 2018 Jan 6]. Available from: <https://medgen.medizin.uni-tuebingen.de/NPC-db2/search.php>
8. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014;42(W1):252–8.
9. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* [Internet]. 2018;(June):1–8. Available from: <http://doi.org/10.1093/nar/gky427>
10. Kortemme T, Kim DE, Baker D. Computational Alanine Scanning of Protein-Protein Interfaces. *Sci Signal* [Internet]. 2004;2004(219):pl2-pl2. Available from: <http://stke.sciencemag.org/cgi/doi/10.1126/stke.2192004pl2>
11. Chandra N. ABS-Scan Webserver for insilico binding site alanine scanning mutagenesis [Internet]. [cited 2018 Jan 6]. Available from: <http://proline.biochem.iisc.ernet.in/abscan/index.php>
12. Vincent Le Guilloux PS, Tuffery and P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics*. 2009;10(168).
13. Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc*. 2012;7(8):1511–22.
14. Wang Z, Xu J. Predicting protein contact map using evolutionary and physical constraints by integer programming. *Bioinformatics* [Internet]. 2013;29(13):i266-73. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23812992>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3694661>
15. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: Computed atlas of surface topography of proteins with structural and

- topographical mapping of functionally annotated residues. *Nucleic Acids Res.* 2006;34(WEB. SERV. ISS.):116–8.
16. Trott O, Olson A. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem.* 2010;31(2):455–61.
 17. Zoete V, Daina A, Bovigny C, Michielin O. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. *J Chem Inf Model.* 2016;56(8):1399–404.
 18. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. *Nucleic Acids Res.* 2016;44(D1):D1202–13.
 19. Sterling T, Irwin JJ. ZINC 15 - Ligand Discovery for Everyone. *J Chem Inf Model.* 2015;55(11):2324–37.
 20. GARRETT M. MORRIS, RUTH HUEY, WILLIAM LINDSTROM, MICHEL F. SANNER, RICHARD K. BELEW, DAVID S. GOODSSELL AJO. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem.* 2009;30:2785–91.
 21. Pires DEV, Blundell TL, Ascher DB. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem.* 2015;58(9):4066–72.
 22. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep [Internet].* 2017;7(January):1–13. Available from: <http://dx.doi.org/10.1038/srep42717>
 23. Daina A, Zoete V. A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem.* 2016;1117–21.
 24. Li X, Saha P, Li J, Blobel G, Pfeffer SR. Clues to the mechanism of cholesterol transfer from the structure of NPC1 middle luminal domain bound to NPC2. *Proc Natl Acad Sci [Internet].* 2016;113(36):10079–84. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1611956113>
 25. Li X, Wang J, Coutavas E, Shi H, Hao Q, Blobel G. Structure of human Niemann–Pick C1 protein. *Proc Natl Acad Sci [Internet].* 2016;113(29):8212–7. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1607795113>
 26. Kwon HJ, Abi-Mosleh L, Wang ML, Deisenhofer J, Goldstein JL, Brown MS, et al. Structure of N-Terminal Domain of NPC1 Reveals Distinct Subdomains for Binding and Transfer of Cholesterol. *Cell [Internet].* 2009;137(7):1213–24. Available from: <http://dx.doi.org/10.1016/j.cell.2009.03.049>
 27. Li X, Lu F, Trinh MN, Schmiedege P, Seemann J, Wang J, et al. 3.3 Å structure of Niemann–Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol transport. *Proc Natl Acad Sci [Internet].* 2017;114(34):201711716. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1711716114>
 28. Vanier MT, Millat G. Structure and function of the NPC2 protein. *Biochim Biophys Acta - Mol Cell Biol Lipids.* 2004;1685(1–3):14–21.
 29. Du X, Kumar J, Ferguson C, Schulz TA, Ong YS, Hong W, et al. A role for oxysterol-binding protein-related protein 5 in endosomal cholesterol trafficking. *J Cell Biol.* 2011;192(1):121–35.

30. Lin Jia, Jenna L. Betters and LY. Niemann-Pick C1-Like 1 (NPC1L1) Protein in Intestinal and Hepatic Cholesterol Transport. *Annu Rev Physiol*. 2011;73:239–59.
31. Komati R, Spadoni D, Zheng S, Sridhar J, Riley KE, Wang G. Ligands of therapeutic utility for the liver X Receptors. *Molecules*. 2017;22(1):1–24.
32. Sugizaki T, Watanabe M, Horai Y, Kaneko-Iwasaki N, Arita E, Miyazaki T, et al. The niemann-Pick C1 like 1 (NPC1L1) inhibitor ezetimibe improves metabolic disease via decreased liver X receptor (LXR) activity in liver of obese male mice. *Endocrinology*. 2014;155(8):2810–9.
33. Davis HR, Veltri EP. Zetia: Inhibition of Niemann-Pick C1 Like 1 (NPC1L1) to Reduce Intestinal Cholesterol Absorption and Treat Hyperlipidemia. *J Atheroscler Thromb*. 2007;14:99–108.
34. Wu WM1, Wu J BN. Effect of 2-hydroxypropyl-beta-cyclodextrin on the solubility, stability, and pharmacological activity of the chemical delivery system of TRH analogs. *Pharmazie*. 2002;57(2):130–4.
35. Yaksh TL1, Jang JD, Nishiuchi Y, Braun KP, Ro SG GM. The utility of 2-hydroxypropyl-beta-cyclodextrin as a vehicle for the intracerebral and intrathecal administration of drugs. *Life Sci*. 1991;48(7):623–33.
36. Liu B. Therapeutic potential of cyclodextrins in the treatment of Niemann–Pick type C disease. *Clin Lipidol [Internet]*. 2012;7(3):289–301. Available from: <http://www.futuremedicine.com/doi/10.2217/clp.12.31>
37. Tanaka Y, Yamada Y, Ishitsuka Y, Matsuo M, Shiraishi K, Wada K, et al. Efficacy of 2-Hydroxypropyl- β -cyclodextrin in Niemann–Pick Disease Type C Model Mice and Its Pharmacokinetic Analysis in a Patient with the Disease. *Biol Pharm Bull*. 2015;844(38):844–51.
38. Hollak CE, Hughes D, van Schaik IN, Schwierin B BB. Miglustat (Zavesca) in type 1 Gaucher disease: 5-year results of a post-authorisation safety surveillance programme. *Pharmacoepidemiol drug Saf*. 2009;18(9):770–7.
39. Kwon HJ, Palnitkar M, Deisenhofer J. The structure of the NPC1L1 N-terminal domain in a closed conformation. *PLoS One*. 2011;6(4).
40. Weinglass AB, Kohler M, Schulte U, Liu J, Nketiah EO, Thomas A, et al. Extracellular loop C of NPC1L1 is important for binding to ezetimibe. *Proc Natl Acad Sci U S A*. 2008;105(32):11140–5.
41. Suchanek M, Hynynen R, Wohlfahrt G, Lehto M, Johansson M, Saarinen H, et al. The mammalian oxysterol-binding protein-related proteins (ORPs) bind 25-hydroxycholesterol in an evolutionarily conserved pocket. *Biochem J [Internet]*. 2007;405(3):473–80. Available from: <http://biochemj.org/lookup/doi/10.1042/BJ20070176>
42. Mark E. Gelsthorpe,1, Nikola Baumann, Elizabet Millard, Sarah E. Gale, S. Joshua Langmade JES and DSO. Niemann-Pick Type C1 I1061T Mutant Encodes a Functional Protein That Is Selected for Endoplasmic Reticulum-associated Degradation Due to Protein Misfolding. *J Biol Chem*. 2008;283:8229–36.
43. Millat G, Marçais C, Tomasetto C, Chikh K, Fensom AH, Harzer K, et al. Niemann-Pick C1 Disease: Correlations between NPC1 Mutations, Levels of NPC1 Protein, and Phenotypes Emphasize the Functional Significance of the Putative Sterol-Sensing Domain and of the Cysteine-Rich Luminal Loop. *Am J Hum Genet [Internet]*. 2001;68(6):1373–85. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0002929707610489>

8. Anexos

- Cavidades determinadas con *fpocket* (formato *pdb*):
 - 3 cavidades por diana
- *Script* de R para traducir las estructuras de las cavidades en una línea de aminoácidos: Cavidades_definidas_aminoacidos.pdf
- *Script* de R para traducir las estructuras de las cavidades en coordenadas: Cavidades_definidas_coordenadas.pdf
- *Script* de R para crear los 5 modelos *pdb* de NPC1 mutado: Mutaciones.pdf