

# **Búsqueda de compuestos moduladores de STAU1 mediante herramientas computacionales para su aplicación en la distrofia miotónica tipo I**

**Andrea López Martínez**

Máster en Bioinformática y bioestadística

Área 3: Biología molecular

**Tutor: Emmanuel Fajardo**

**Profesor/a responsable de la asignatura: Laura Calvet Liñan**

2 de junio de 2022



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>Búsqueda de compuestos moduladores de STAU1 mediante herramientas computacionales para su aplicación en la distrofia miotónica tipo I</i>
Nombre del autor:	<i>Andrea López Martínez</i>
Nombre del consultor/a:	<i>Emmanuel Fajardo</i>
Nombre del PRA:	<i>Laura Calvet Liñan</i>
Fecha de entrega (mm/aaaa):	06/2022
Titulación:	<i>Máster en Bioinformática y bioestadística</i>
Área del Trabajo Final:	Biología molecular (área 3)
Idioma del trabajo:	Castellano
Número de créditos:	15
Palabras clave	<i>“DM1”, “modelo estructura”, “cribado virtual”</i>
<p><b>Resumen del Trabajo (máximo 250 palabras):</b> <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i></p>	
<p>La distrofia miotónica tipo I es una enfermedad neurodegenerativa que afecta a 1 de cada 8000 personas. Actualmente no tiene tratamiento y pese a que las estrategias terapéuticas más adecuadas para su cura serían las técnicas de edición génica, éstas aún no están listas para su la llegada a la clínica. Esto hace que las moléculas procedentes del reposicionamiento farmacéutico resulten terapias atractivas por sus aplicaciones a corto plazo.</p> <p>El objetivo de este trabajo es buscar compuestos que se puedan unir a la proteína STAU1, sobreexpresada en muestras de pacientes de DM1, para restaurar las vías de señalización alteradas por su desregulación.</p> <p>A nivel metodológico, se han utilizado servidores de modelado de proteína, se han identificado los lugares de unión y se han realizado dos cribados virtuales, uno basado en la estructura y otro basado en un ligando conocido.</p> <p>En este trabajo hemos obtenido dos modelos de la estructura de STAU1, hemos comprobado la unión de un ligando definido en una de las estructuras parciales y hemos obtenido dos bases de datos distintas sobre las que realizar un proceso de docking. Del docking, hemos obtenido tres listados distintos en los que contamos con más de 1500 compuestos que estudiar en futuros estudios.</p> <p>En conclusión, podemos confirmar que la metodología descrita nos ha permitido realizar un cribado basado en la estructura y otro basado en el ligando y obtener un conjunto de compuestos que se espera que se unan con alta afinidad a STAU1.</p>	

**Abstract (in English, 250 words or less):**

Myotonic dystrophy type I is a neurodegenerative disease that affects 1 in 8000 people. There is currently no cure and, although the most appropriate therapeutic strategies would be gene-editing techniques, these are not ready yet to be applied in large scale. This highlights the importance of repurposing small molecules already available on the market.

The aim of this work is to look for compounds that can bind STAU1 protein, which is overexpressed in samples from DM1 patients, to restore the signaling pathways altered by its deregulation.

Regarding the methodology, we have used protein-modelling servers and we have identified pockets on the protein structure. Lastly, we performed two virtual screenings of small molecules, one of them based on the protein structure and the other based on a known ligand.

In this work we have obtained two models of the STAU1 structure, we have checked the binding of a ligand present in one of the partial structures already described and we have obtained two different databases for docking. From the docking results, we have obtained three different lists in which we have more than 1500 compounds to consider in future studies.

In conclusion, we can confirm that the described methodology has allowed us to perform a structure-based and a ligand-based screening to obtain a set of compounds that are expected to bind with high affinity to STAU1.

# Índice

<b>1. Resumen</b> .....	<b>2</b>
<b>2. Introducción</b> .....	<b>3</b>
2.1. Contexto y justificación del Trabajo.....	3
2.2. Objetivos del Trabajo .....	5
2.3. Enfoque y método seguido .....	5
2.4. Planificación del Trabajo .....	6
2.5. Breve resumen de contribuciones y productos obtenidos .....	7
2.6. Breve descripción de otros capítulos de la memoria.....	7
<b>3. Estado del arte</b> .....	<b>9</b>
<b>4. Metodología</b> .....	<b>11</b>
4.1. Modelado por homología de la estructura 3D .....	11
4.2. Búsqueda de lugares de unión .....	12
4.3. Cribado virtual basado en la estructura.....	12
4.4. Cribado virtual basado en el ligando .....	13
<b>5. Resultados</b> .....	<b>16</b>
5.1. Modelo de la estructura 3D .....	16
5.2. Análisis de los sitios de unión .....	19
5.3. Cribado virtual basado en la estructura.....	22
5.4. Cribado virtual basado en el ligando .....	25
5.5. Análisis y comparación de moléculas seleccionadas.....	28
<b>6. Discusión</b> .....	<b>30</b>
<b>7. Conclusiones</b> .....	<b>31</b>
7.1. Conclusiones.....	31
7.2. Líneas de futuro .....	31
7.3. Seguimiento de la planificación.....	31
<b>8. Glosario</b> .....	<b>33</b>
<b>9. Bibliografía</b> .....	<b>34</b>

## Lista de figuras

- Figura 1. Alteraciones moleculares presentes en la distrofia miotónica tipo I y los principales abordajes terapéuticos (amarillo). Creado con Biorender.
- Figura 2. Vías de señalización alteradas en distrofia miotónica tipo I.
- Figura 3. Datos iniciales de staufen 1 presentes en UniProtKB.
- Figura 4. Interfaz de MTiOpenScreen con una de las listas de residuos utilizadas.
- Figura 5. Web de SwissSimilarity con una de las bases de datos marcada.
- Figura 6. Web de SwissSmilarity con la selección de la bases de datos de ZINC (drug-like).
- Figura 7. Resultados de la búsqueda de staufen 1 en PDB.
- Figura 8. Resultados de distintos modelos posibles en SwissModel.
- Figura 9. Modelo de staufen 1 seleccionado en SwissModel.
- Figura 10. Resultado del modelo de staufen 1 realizado con Phyre2.
- Figura 11. Templates en los que se basa el modelo de staufen 1 de Phyre2.
- Figura 12. Descripción de los tres primeros templates en los que se basa el modelo de staufen 1 de Phyre2.
- Figura 13. Modelo de staufen 1 realizado con SwissModel (izquierda) y Phyre2 (derecha).
- Figura 14. Lugares de unión identificados con fpocket sobre los templates 6SDW y 4DKK.
- Figura 15. Descripción de los ligandos CL y FLC de la estructura 4DKK de PDB.
- Figura 16. Resultado de la homología entre la secuencias de staufen 1.
- Figura 17. Descripción del ligando FLC consultada en PBD.
- Figura 18. Descripción del ligando CL consultada en PDB.
- Figura 19. Representación de la unión de Siramesine con el modelo de la proteína de Staufen1.
- Figura 20. Visualización de la unión del fármaco Glisindamide en el modelo de staufen1.
- Figura 21. Visualización del fármaco Brolaconazole y el modelo de la proteína staufen1
- Figura 22. Error encontrado en SwissDock al intentar realizar el proceso de docking con el ligando CL.
- Figura 23. Estructura de Glisindamide.
- Figura 24. Estructura de Hesperidin.
- Figura 25. Estructura de Radotinib.
- Figura 26. Estructura de Regorafenib.

## Lista de tablas

Tabla 1. Características de la interacción del ligando CL en el modelo de staufer 1 (Phyre2).

Tabla 2. Mejores resultados de docking para el pocket 1.

Tabla 3. Mejores resultados de docking para el pocket 2

Tabla 4. Inicio de la tabla de energías resultante del docking de staufer 1 y el ligando FLC en SwissDock.

Tabla 5. Compuestos similares a FLC de la base de datos de ZINC.

Tabla 6. Compuestos similares a FLC de la base de datos ChEMBL.

Tabla 7. Compuestos similares a FLC de la base de datos DrugBank.

Tabla 8. Resultados de docking de compuestos similares a FLC.

# 1. Resumen

La distrofia miotónica tipo I es una enfermedad neurodegenerativa que afecta a 1 de cada 8000 personas. Es una enfermedad altamente incapacitante causada por la repetición de tripletes CTG en el gen DMPK, que provoca una desregulación de múltiples vías de señalización, así como la alteración de un gran número de eventos de splicing. Actualmente no tiene tratamiento y pese a que las estrategias terapéuticas más adecuadas serían las técnicas de edición génica, su falta de desarrollo para la llegada a la clínica hace que las moléculas procedentes del reposicionamiento farmacéutico resulten terapias atractivas por sus aplicaciones a corto plazo.

El objetivo de este trabajo es buscar compuestos que se puedan unir a la proteína STAU1, sobreexpresada en muestras de pacientes de DM1, para restaurar las vías de señalización alteradas por esta proteína.

A nivel metodológico, se han utilizado servidores de modelado de proteína como *SwissModel* y *Phyre2*, se han identificado los lugares de unión mediante *fpocket* y se han realizado dos cribados virtuales con *MTiOpenScreen*, uno basado en la estructura y otro basado en un ligando conocido, para el que también se ha utilizado *SwissDock* y *SwissSimilarity*.

En este trabajo hemos obtenido dos modelos de la estructura de STAU1, hemos comprobado la unión del ligando FLC definido en una de las estructuras parciales y hemos obtenido dos bases de datos distintas sobre las que realizar un proceso de docking. Del docking hemos obtenido tres listados distintos (dos basados en los lugares de unión y otro basado en el ligando) en los que contamos con más de 1500 compuestos, así como de una selección de los mejores puntuados para valorar en futuros estudios.

En conclusión, podemos confirmar que la metodología descrita nos ha permitido realizar un cribado basado en la estructura y otro basado en el ligando y obtener un conjunto de compuestos que se espera que se unan con alta afinidad a STAU1.

Una de las limitaciones del estudio es la falta de conocimiento del efecto que tiene tanto el ligando como los compuestos descritos sobre la proteína STAU1, efecto que se espera estudiar próximamente *in vitro*. Con este trabajo hemos abierto la puerta al estudio de la modulación de STAU1 como estrategia terapéutica frente a la DM1.



## 2.Introducción

### 2.1. Contexto y justificación del Trabajo

La distrofia miotónica tipo I (DM1) o enfermedad de Steiner es una enfermedad neurodegenerativa causada por la repetición anormal de tripletes CTG en la región 3'UTR del gen de la distrofia miotónica proteína kinasa o *DMPK*. Es la distrofia muscular más común en adultos y tiene una incidencia de 1 cada 8000 personas alrededor del mundo, pero su incidencia llega a triplicarse en la zona norte de España<sup>1,2</sup>.

La DM1 es una enfermedad multisistémica compleja caracterizada por la gran variedad de tejidos afectados. Los síntomas más comunes son la debilidad muscular y la hiperexcitabilidad muscular o miotonía, cataratas, alteraciones cognitivas, gastrointestinales, problemas de conductividad cardíaca y resistencia a insulina, entre otros síntomas. Hay distintos subtipos definidos de la enfermedad dependiendo de la edad de aparición y el número de repeticiones presentes en *DMPK*, existiendo desde pacientes mínimamente afectados que no ven reducida su calidad de vida (DM1 adulta) hasta pacientes gravemente afectados con un grado de discapacidad alto (DM1 congénito) e incluso, muertes perinatales causadas por la gran amplitud de la repetición<sup>3,4</sup>.

El gen *DMPK* se encuentra en la región 19q3.3 del cromosoma 19 y en su estado basal, presenta entre 5 y 35 repeticiones CTG en su zona 3'UTR. En cambio, en muestras de pacientes con DM1, estas expansiones varían entre las 50 y más de 2000. Una característica significativa de la enfermedad es el fenómeno de la anticipación génica, en que cada generación está más afectada que la anterior dada la elongación del número de tripletes entre una y otra. De hecho, de unos padres con una pre-mutación en *DMPK*, entre 35 y 50 repeticiones, es altamente probable que en la siguiente generación aparezca algún individuo afectado, dado que además es una enfermedad dominante en la que un alelo afectado es suficiente para causarla<sup>5,6</sup>.

La repetición anormal de tripletes CTG provoca la síntesis de un RNA mensajero "pegajoso" con múltiples estructuras de horquilla que se unen con alta afinidad a la proteína MBNL1. MBNL1 o muscleblind-like 1 es una proteína de la familia de muscleblind (junto con MBNL2 y MBNL3) que se encarga de la regulación del *splicing* alternativo en numerosos tejidos, pero es especialmente abundante en el músculo esquelético. El secuestro de la proteína MBNL1 en las repeticiones provoca la desregulación de la concentración y actividad de otras proteínas reguladoras del *splicing* como CELF1 o HNRNPA1, causando lo que se conoce como *esplíceopatía*<sup>7</sup>. La alteración de las proteínas responsables de *splicing* alteran la regulación de más de 100 eventos de *splicing* distintos, causando la pérdida de la función y la viabilidad celular, desregulando procesos como la autofagia, la función mitocondrial o el ciclo celular, tal y como está representado en la Figura 1.

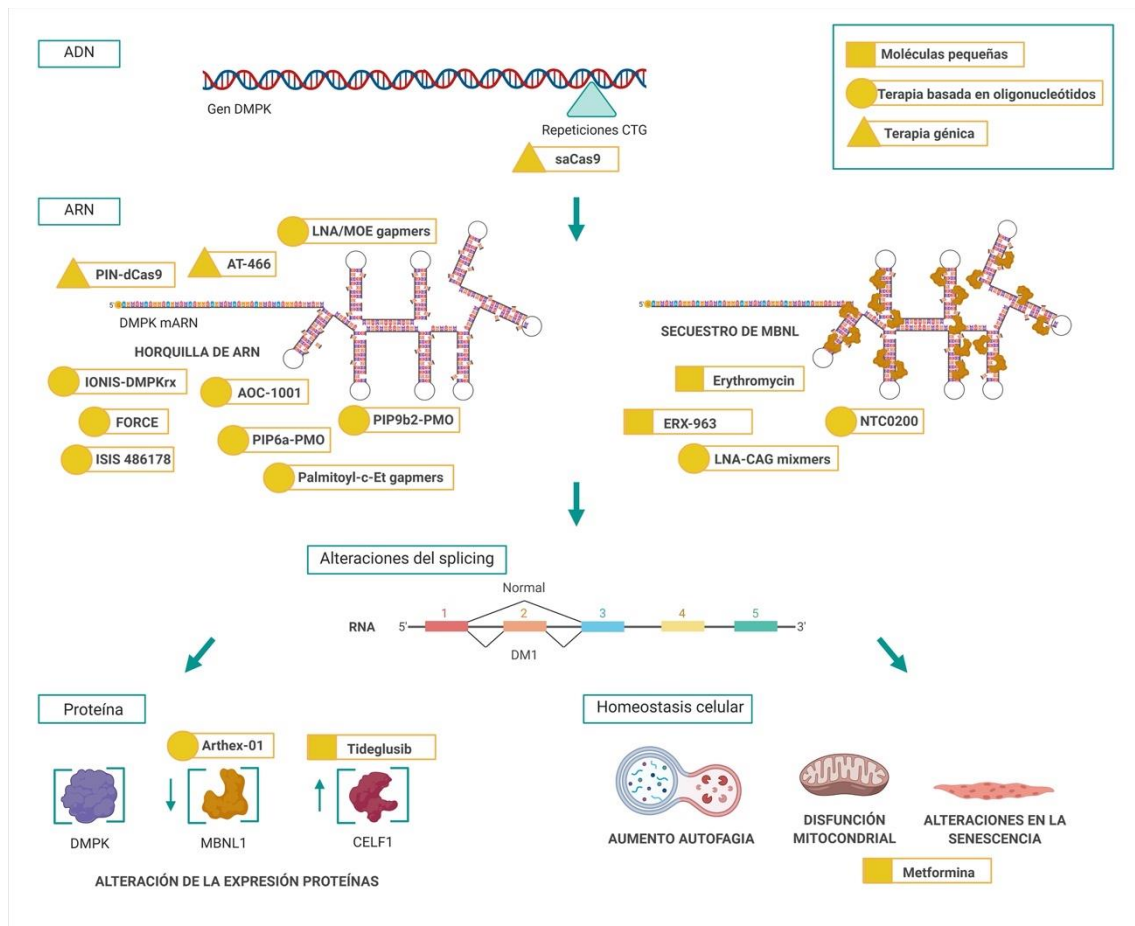


Figura 1. Alteraciones moleculares presentes en la distrofia miotónica tipo 1 y los principales abordajes terapéuticos (amarillo). Creado con Biorender.

Actualmente la enfermedad de Steiner no tiene cura y su tratamiento es básicamente sintomático y poco resolutivo. Habitualmente se utilizan fármacos que mejoran las alteraciones gastrointestinales, las cataratas o la resistencia a insulina, pero no paran o ralentizan el desarrollo de la enfermedad ni solventan los síntomas musculares o el deterioro cognitivo. Además, los pacientes de DM1 tienen un riesgo mayor de padecer ciertos tipos de cáncer, lo que puede provocar complicaciones adicionales<sup>8</sup>.

En los últimos años y gracias a los grandes avances en terapias basadas en oligonucleótidos o de edición génica, así como el aumento de estudios de reposicionamiento, ha incrementado el número de moléculas en fases preclínicas para el tratamiento de la DM1, tal y como está representado en la Figura 1. De igual forma, las terapias basadas en oligonucleótidos presentan grandes dificultades en su distribución a lo largo del cuerpo, siendo el tejido muscular uno de los tejidos más complejos para hacerlas llegar. Así también las terapias basadas en el sistema CRISPR/Cas9, pese a que son muy prometedoras porque podrían revertir la causa de la enfermedad, por el momento no están suficientemente desarrolladas para la aplicación *in vivo* en la gran mayoría de enfermedades<sup>9,10</sup>.

Una de las aproximaciones terapéuticas más translacionales en estos momentos son las estrategias de reposicionamiento terapéutico de compuestos, basadas en estudiar el efecto de librerías de compuestos ya aprobados para una

aplicación, en otra enfermedad distinta. Pese a que son variados los estudios de nuevas moléculas para la DM1, las proteínas o síntomas diana (definidos en la Figura 1) son muy parecidos en todos los casos y, por el momento, ninguna de las estrategias ha conseguido una mejora relevante en la gran mayoría de *hallmarks* de la enfermedad<sup>9,11</sup>.

Por lo tanto, en este trabajo nos centraremos en la búsqueda de nuevos fármacos para la distrofia miotónica utilizando como diana una proteína distinta a las habituales, definiendo así una nueva posible estrategia terapéutica con relativa rápida aplicación.

## 2.2. Objetivos del Trabajo

El objetivo principal del trabajo es la búsqueda de moduladores de la proteína staufer 1 mediante técnicas de cribado virtual de fármacos y así, establecer una nueva diana terapéutica frente a la DM1.

Este objetivo global se divide en los siguientes objetivos específicos:

- Objetivo 1: Definir la estructura 3D de la proteína staufer 1.
- Objetivo 2: Identificar los *pockets* o lugares de unión a compuestos de la proteína staufer 1.
- Objetivo 3: Realizar un cribado virtual de compuestos basado en la estructura de la proteína.
- Objetivo 4: Realizar un cribado virtual de compuestos basado en ligandos conocidos de la proteína.
- Objetivo 5: Comparar los resultados basados en la estructura y los resultados basados en el ligando para escoger los compuestos más adecuados.

## 2.3. Enfoque y método seguido

El primer paso es obtener la estructura completa de la proteína de interés, staufer 1. Para ello, se utilizó la base de datos Protein Data Bank (PDB)<sup>12</sup> y dado que no se encontraba la estructura completa, se utilizaron los servidores web SwissModel<sup>13</sup> y Phyre2<sup>14</sup> para crear dos modelos de la misma. Una vez obtenidos los modelos de la estructura, se escogió la más adecuada para los objetivos del trabajo. Además, se estudió la presencia y conservación de ligandos presentes en las estructuras o *templates* en las que se basan los modelos.

El siguiente paso para poder realizar el cribado basado en la estructura es conocer los *pockets* o lugares de unión a compuestos presentes en la proteína de interés. Esto se ha realizado gracias a 3DLigandSite<sup>15</sup> y fpocket<sup>16</sup>.

Una vez identificadas las cavidades, se utilizó MTiOpenScreen<sup>17</sup> para realizar el screening de compuestos. En este caso y dado el enfoque al reposicionamiento farmacéutico del trabajo, se utilizó la librería “Drugs” que propone el mismo servidor y que contiene fármacos aprobados en comercialización. De los resultados obtenidos, se han seleccionado los 25 con mejores resultados de docking.

Por otro lado, para el cribado basado en el ligando se realizó un proceso de docking del posible ligando de la estructura con SwissDock<sup>18</sup> y posteriormente, se realizó una búsqueda de esta estructura en SwissSimilarity<sup>19</sup>, utilizando distintas bases de datos para encontrar compuestos similares.

A partir de aquí, se obtuvo una reducida librería de compuestos que se estudió con MTiOpenScreen y se valoraron los resultados obtenidos.

Por último, se compararon los resultados obtenidos de ambos cribados y se valoraron las opciones más viables para el desarrollo futuro de este trabajo.

## 2.4. Planificación del Trabajo

Este trabajo se inició en septiembre de 2021 con la idea de finalizar en enero de 2022. La redacción de la memoria, presentación y defensa no se pudo completar a tiempo por diversos motivos. A continuación, se presenta el calendario inicial definido y seguido hasta la PEC4 y sus modificaciones en rojo:

HITO		INICIO	FIN
PEC 0: Definición del trabajo y objetivos		15/09/2021	22/09/2021
PEC 1: Definición del plan de trabajo		23/09/2021	04/10/2021
PEC 2: Estructura de la proteína y screening basado en la misma		05/10/2021	08/11/2021
PEC 3: Identificación del ligando y screening basado en éste		09/11/2021	06/12/2021
PEC 4: Redacción de la memoria		17/05/2022	02/06/2022
PEC 5	5A: Preparación de la presentación	30/05/2022	06/06/2022
	5B: Preparación de la defensa y defensa	13/06/2022	24/06/2022

	15/09-22/09	23/09-04/10	05/10 - 08/11	09/11 - 06/12
<b>TAREAS</b>				
PEC 0: Definición del trabajo y objetivos				
PEC 1: Definición del plan de trabajo				
PEC 2: Estructura de la proteína y screening basado en la misma				
PEC 3: Identificación del ligando y cribado basado en este				
PEC 4: Redacción de la memoria				
PEC 5: Presentación oral				
Preparación de la presentación				
Defensa pública (PEC 5B)				
	....	17/05 - 02/06	03/06 - 06/06	13/06 - 24/06
<b>TAREAS</b>				
PEC 0: Definición del trabajo y objetivos				
PEC 1: Definición del plan de trabajo				
PEC 2: Estructura de la proteína y screening basado en la misma				
PEC 3: Identificación del ligando y cribado basado en este				
PEC 4: Redacción de la memoria				
PEC 5: Presentación oral				
Preparación de la presentación				
Defensa pública (PEC 5B)				

## 2.5. Breve resumen de contribuciones y productos obtenidos

Los principales productos de este trabajo son:

- Memoria de TFM
- Modelo de la proteína de SwissDock (*model.pdb*).
- Modelo de la proteína de Phyre2 (*STAU1\_model\_Phyre2.pdb*).
- Listado de compuestos similares a FLC de SwissSimilarity (*Compuestos\_SwissSimilarity.xlsx*).
- Resultado del docking con FLC (*Staufen1\_FLC\_SwissDOCK.chimerax*).
- Resultado del docking basado en el ligando (*Ligand-based\_screening.csv*).
- Resultado del docking basado en la estructura (*Output\_pocket1.csv*, *Output\_pocket2.csv*).

Los datos están disponibles en [Google Drive](#).

## 2.6. Breve descripción de los otros capítulos de la memoria

- Capítulo 1: Resumen

Breve descripción y resumen del trabajo presentado.

- Capítulo 2: Introducción

Descripción de la distrofia miotónica tipo I y la problemática a responder. Organización y objetivos del trabajo.

- Capítulo 3: Estado del arte

Estrategias terapéuticas exploradas en la actualidad y qué puede aportar el trabajo presentado.

- Capítulo 4: Materiales y métodos

Descripción del trabajo desarrollado y métodos utilizados.

- Capítulo 5: Resultados

Descripción de productos obtenidos de los métodos descritos previamente.

- Capítulo 6: Discusión

Valoración de los resultados obtenidos y su valor y utilidad en el futuro desarrollo del trabajo. Próximos pasos a seguir.

- Capítulo 7: Conclusión

Ideas principales extraídas de los resultados descritos.

- Capítulo 8: Glosario
- Capítulo 9: Bibliografía
- Capítulo 10: Anexos

Link directo a los productos descritos en el apartado 2.5.

### 3. Estado del arte

La patogénesis de la enfermedad de Steiner es altamente compleja, así como su tratamiento. Pese a que los tratamientos que actúan directamente sobre la causa molecular, es decir, la expansión de CTG, serían los más adecuados, la aplicación de una terapia definitiva está lejos de la clínica aún. Es por eso por lo que la búsqueda de dianas distintas a las repeticiones, así como las estrategias de reposicionamiento son especialmente relevantes para la llegada de un tratamiento a la clínica de forma mucho más rápida.

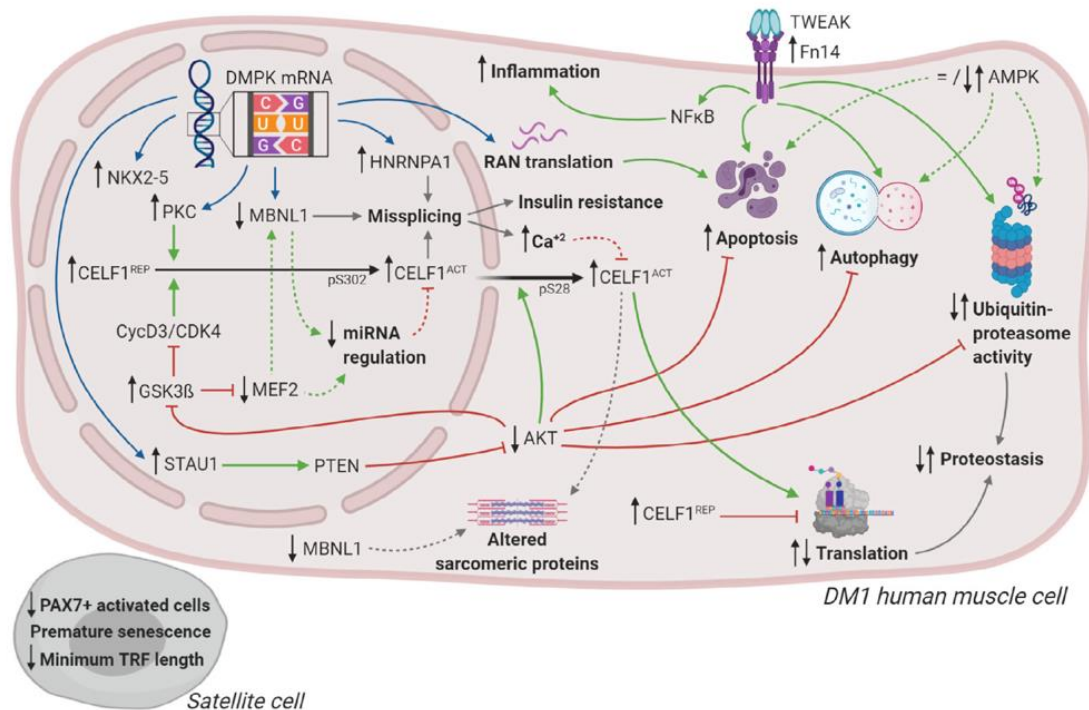


Figura 2. Vías de señalización alteradas en distrofia miotónica tipo I. Figura extraída de Ozimski (2020)<sup>20</sup>.

Tal y como observamos en la Figura 2, hay un gran número de vías de señalización, moléculas y procesos celulares implicados en la patogénesis de la distrofia miotónica.

En cuanto a las dianas alternativas a las repeticiones, las más exploradas son la sobreexpresión de MBNL1 gracias a la regulación de microARNs<sup>21</sup> o la estabilización de CELF1 (regulador antagónico de MBNL1 alterado en DM1) vía GSK3β gracias a un fármaco en reposicionamiento, tideglusib<sup>22</sup>. Pese a que ambas estrategias han mostrado efectos pre-clínicos prometedores, ninguna de las dos ha iniciado un ensayo clínico en pacientes por el momento.

#### STAU1

STAU1 es un gen ubicado en la región 20q13.3 del cromosoma 20 que codifica para la proteína de unión a ARN de doble cadena staufer 1 (STAU1 o staufer1). Esta proteína es miembro de la familia de proteínas de unión a ARN como MBNL1, CELF1 y HNRNPA1. Estas proteínas están caracterizadas por la presencia de múltiples lugares de unión a RNA. La peculiaridad de STAU1 es que también cuenta con un dominio de unión a microtúbulos, interaccionando

directamente con la tubulina para el transporte de ARNm a través del retículo endoplasmático rugoso hasta el lugar de traducción<sup>23</sup>.

En los últimos años, diversas publicaciones han descrito la sobreexpresión de STAU1 en distintos tipos de muestra de DM1. Inicialmente se describió su sobreexpresión en muestras de músculo esquelético de pacientes con DM1 y se debatió su rol en la patogénesis molecular de la enfermedad; altera la respuesta celular al estrés y actúa como regulador del *splicing*, así como regulador de la expresión de proteínas como MBNL1 y CELF1<sup>24,25</sup>. Más tarde, también se describió el efecto de la sobreexpresión de STAU1 en modelos murinos de la enfermedad, contribuyendo a la atrofia y desgaste muscular progresivos vía PTEN<sup>26</sup>. En estos trabajos se valoraba STAU1 como nueva diana terapéutica para la enfermedad de Steiner además de como potencial biomarcador.

Si nos fijamos en la Figura 2, observamos que la sobreexpresión de STAU1 presente en modelos de la enfermedad afecta directamente a la regulación de la vía de PTEN/AKT, proteínas conocidas por su alta implicación en los procesos carcinogénicos, que a la vez regulan múltiples procesos en la célula como la apoptosis, la autofagia, la proteostasis e incluso, la estabilización de CELF1. Además, se encuentra en un punto muy inicial de la patogénesis molecular, estando directamente relacionada con la expansión de tripletes y sin regular de forma directa (y única) procesos vitales para la homeostasis celular.

Por estos motivos, se valoró la modulación de STAU1 como posible estrategia terapéutica a explorar en este trabajo, intentando identificar moduladores de su actividad entre compuestos ya aprobados y comercializados para su uso.



## 4. Metodología

### 4.1. Modelado por homología de la estructura 3D

En Protein Data Bank (PDB) se comprobó que no figuraba la estructura completa de staufen 1 en la literatura. Se utilizó la base de datos de NCBI para obtener la secuencia de aminoácidos de la proteína y UniProtKB para obtener el código que se utilizará en los siguientes pasos. En este caso es el O95793, tal y como observamos en la Figura 3.

UniProtKB - O95793 (STAU1\_HUMAN)

Display Help video BLAST Align Format Add to basket History Add a publication Feedback

Entry Protein Double-stranded RNA-binding protein Staufen homolog 1  
Gene STAU1  
Organism Homo sapiens (Human)  
Status Reviewed - Annotation score: ●●●●● - Experimental evidence at protein level<sup>1</sup>

Function<sup>1</sup>

None

Function  
Names & Taxonomy  
Subcellular location  
Pathology & Biotech  
PTM / Processing  
Expression  
Interaction  
Structure

Binds double-stranded RNA (regardless of the sequence) and tubulin. May play a role in specific positioning of mRNAs at given sites in the cell by cross-linking cytoskeletal and RNA components, and in stimulating their translation at the site.

(Microbial infection) Plays a role in virus particles production of many viruses including of HIV-1, HERV-K, ebola virus and influenza virus. Acts by interacting with various viral proteins involved in particle budding process.

4 Publications

GO - Molecular function<sup>1</sup>

- double-stranded RNA binding Source: ProtInc
- protein phosphatase 1 binding Source: Ensembl
- RNA binding Source: UniProtKB

Figura 3. Datos iniciales de staufen 1 presentes en UniProtKB.

El modelo de la proteína se realizó mediante los servidores web SwissModel y Phyre2.

SwissModel se encarga de obtener un modelo de la estructura 3D de una proteína a partir de su secuencia de aminoácidos, obtenida gracias al código de UniProtKB, y a la similitud de ésta con unas secuencias cuya estructura 3D ha sido descrita. El servidor estudia diversos “templates” o moldes en los que se basará la estructura de la proteína. En este caso, hemos elegido manualmente aquellos *templates* en los que queremos que se base el modelo. A partir de ahí, se crean diversos modelos a los cuáles se les asignan diversas puntuaciones. Para cada modelo obtenemos las puntuaciones GMQE y QMEANDisCo, además del nivel de cobertura y porcentaje de identidad. Cuanto mayor sean estas puntuaciones, mayor será la calidad esperada del modelo. GMQE o Global Model Quality Estimate es un indicador de la calidad de los moldes o *templates* en los que se basa la estructura, contrastando el alineamiento de la secuencia aportada con la del molde y la propia estructura del molde elegido. QMEANDisCo es una puntuación asignada a cada residuo del modelo que sirve como indicador de la calidad del modelo general. La selección del modelo se realiza en base a los parámetros mencionados <sup>13</sup>.

De SwissModel obtuvimos el modelo en formato .pdb y un informe del modelo.

Dado que no hay una estructura completa descrita, se utilizó también el servidor web Phyre 2, en su versión *Intensive*, que nos permite obtener un modelo a través de distintos *templates* unidos. Además, si el modelo tiene un porcentaje

de confianza mayor al 90%, automáticamente obtenemos la predicción de los lugares de unión del sistema. Igual que SwissModel, la búsqueda de *templates* se basa en la secuencia de aminoácidos obtenida previamente, a partir de los cuáles predice la estructura secundaria de los residuos. Además, hace una predicción también de los “loops” o aquellas uniones entre regiones sin estructura secundaria definida por los *templates*.<sup>14</sup>

De Phyre2 se obtuvo el modelo en formato .pdb así como el informe completo que incluía dos lugares de unión.

Ambos modelos se visualizaron con PyMOL.

#### **4.2. Búsqueda de lugares de unión**

Los lugares de unión del modelo descrito se buscaron mediante 3DLigandSite. En este caso, el fichero estudiado ha sido el modelo obtenido de SwissDock directamente en el servidor. Por otro lado, Phyre2 ha utilizado el servidor fpocket para el mismo proceso.

3DLigandSite es un servidor web que se dedica a predecir lugares de unión de ligandos. Utiliza una predicción de estructura de la proteína para construir un modelo, muy parecido a lo que hacen los otros dos servidores descritos (SwissModel y Phyre 2). La diferencia es que se basa explícitamente en los ligandos unidos a los *templates* y su unión al model aportado (en nuestro caso) y predecir el lugar concreto. Pese a que puede resultar muy parecido a SwissModel, éste es mucho más conservador a la hora de mantener los ligandos de los *templates*, ya que, si no tienen una identidad alta, quedan descartados del modelo. En cambio, 3DLigandSite explora estas conformaciones de forma menos conservadora, lo que permite tener un abanico más amplio de ligandos a estudiar<sup>15</sup>.

fpocket, accesible a través de la página de la Ressource Parisienne en Bioinformatique Structurale, es un paquete de código abierto que identifica “pockets” o aquellas zonas de la proteína donde es altamente probable que se una algún ligando<sup>16</sup>. Se ha utilizado este servidor sobre el modelo obtenido con Phyre2.

Por último, ante la falta de ligandos descritos mediante SwissModel y 3DLigandSite, se utilizó blastp, de NCBI, para estudiar la similitud de una secuencia parcial descrita de staufen 1 perteneciente a la estructura 4DKK de PDB y la secuencia FASTA de la isoforma escogida.

#### **4.3. Cribado virtual basado en la estructura**

El cribado basado en la estructura se ha realizado con MTiOpenScreen, pese a que previamente se intentó utilizar DockBlaster sin éxito, debido a la gran cantidad de tiempo que se demora el proceso.

MTiOpenScreen es un servidor web en el que se puede realizar el proceso de docking a la vez que el screening de distintos compuestos en base a las energías de unión de estos y gracias a AutoDock Vina. En el servidor podemos escoger entre realizar un proceso de docking ciego o bien marcar los sitios de unión predichos en la molécula. En este caso, hemos realizado tres procesos distintos basados en los lugares de unión predichos en el modelo de Phyre2 con fpocket y a través de 3DLigandSite. En cuanto a los compuestos, para este cribado hemos elegido la lista “Drugs-lib” que cuenta con más de 50.000 compuestos y fue obtenida de distintas bases de datos de compuestos de PubChem BioAssay Database<sup>17</sup>. En la Figura 4 podemos observar un ejemplo de su visualización.

The screenshot displays the MTiOpenScreen web interface with the following sections:

- Protein receptor**: Receptor file (Receptor\_3DStructure) [🔗](#)
- STAU1\_modelling.pdb (not detected)** [📄](#)  
This file cannot be displayed in your browser. Click on the "save" link to download it.
- Compounds database**: Select library for the compound (use the database filters below) Drugs-lib
- Compound library filters**: Is lead-like?  Yes
- Grid calculation**: Mode List of residues
- Calculate grid from a list of residues (list of residues mode)**:  
List of residues separated by commas (i.e. "A\_PHE\_919\_...A\_ALA\_866\_...") [🔗](#)  
\_ILE\_207\_...LEU\_210\_...ALA\_211\_...LYS\_220\_...GLU\_221\_...PRO\_222\_...TYR\_224\_...MET\_240\_...VAL\_242\_...LYS\_256\_...LYS\_257\_...LYS\_260\_...ALA\_263\_...ALA\_264\_...MET\_267

Figura 4. Interfaz de MTiOpenScreen con una de las listas de residuos utilizadas.

#### 4.4. Cribado virtual basado en el ligando

El cribado basado en el ligando se realizó gracias a la lista de compuestos obtenidos de SwissSimilarity, a partir del ligando FLC (descrito en apartados posteriores). SwissSimilarity es un servidor web que permite realizar un cribado rápido de compuestos basado en el ligando, ya sea en su propia estructura 2D basada en los enlaces presentes en la propia molécula, o también en base a su conformación 3D y predicción del farmacóforo. El servicio cuenta con diversas bases de datos de compuestos (cuenta con más de dos millones en total) entre las que puedes filtrar entre aquellos que están en comercialización, los bioactivos, en ensayos clínicos e incluso nutraceuticos<sup>19</sup>. Para la selección de compuestos, se escogieron la clase de compuestos “Drugs” y se seleccionaron dos librerías, *DrugBank* y *ChEMBL approved drugs*, ambas con el método de screening combinado 2D y 3D. En la Figura 5 observamos uno de los métodos de screening descritos.



Don't know where to start? Try with an example **Diclofenac**, **Propranolol** or **Nilotinib**.

1 - Enter a molecule in SMILES format

No SMILES available? Draw a molecule **using the sketcher**

2 - Select a class of compounds

Please, select a class of compounds here:

3 - Select compound library and screening method

Put the mouse over a radio button to see the corresponding computation time

	2D							3D	2D & 3D	
	FP2	ECFP4	MHFP6	Pharmacophore	ERG	Scaffold	Generic Scaffold	Electroshape	E3FP	Combined
DrugBank	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ChEMBL approved drugs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
ChEMBL drug candidates in clinics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4 - Submit

**START SCREENING**

Reset form

Figura 5. Web de SwissSimilarity con una de las bases de datos marcada.

En una segunda selección, más enfocada al reposicionamiento, se eligió el conjunto de compuestos comercializados presentes en la base de datos ZINC, aquellos que cumplen las condiciones “drug-like”. En este caso el proceso de cribado también fue combinado (2D y 3D), como podemos observar en la Figura 6.



Don't know where to start? Try with an example **Diclofenac**, **Propranolol** or **Nilotinib**.

### 1 - Enter a molecule in SMILES format

No SMILES available? Draw a molecule **using the sketcher**

### 2 - Select a class of compounds

Please, select a class of compounds here:

### 3 - Select compound library and screening method

Put the mouse over a radio button to see the corresponding computation time

	2D							3D	2D & 3D	
	FP2	ECFP4	MWFP6	Pharmacophore	ERG	Scaffold	Generic Scaffold	Electroshape	ESPP	Combined
ZINC (Drug-like)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
ZINC (Lead-like)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ZINC (Fragments)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Asinex	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
AsisChem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ChemBridge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Enamine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Maybridge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Otava	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
SPECS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 4 - Submit

**START SCREENING**

Reset form

Figura 6. Web de SwissSmilarity con la selección de la bases de datos de ZINC (drug-like).

El proceso de docking para cada molécula elegida se ha realizado también con MTiOpenScreen. Para este objetivo, se unieron todos los compuestos seleccionados en una misma base de datos en formato *.mol2* y se envió al servidor.

# 5. Resultados

## 5.1. Modelo de la estructura 3D

La estructura 3D completa de la proteína staufen 1 no está descrita en la literatura, como comprobamos en la base de datos PDB, pero existen estructuras parciales de la misma descritas mediante distintos métodos.

The screenshot shows the RCSB PDB search results for 'staufen 1'. The interface includes a navigation bar with options like 'Deposit', 'Search', 'Visualize', 'Analyze', 'Download', 'Learn', 'More', 'Documentation', and 'Careers'. A sidebar on the left provides filters for scientific name, taxonomy, experimental method, polymer entity type, refinement resolution, release date, symmetry type, and SCOP classification. The main content area displays four search results, each with a 3D model icon, a '3D View' button, and detailed information including the PDB ID, title, authors, release date, method, organisms, and macromolecules.

PDB ID	Title	Released	Method	Organisms	Macromolecule
6SDY	Solution structure of Staufen1 dsRBD4 - hARF1 SBS dsRNA complex.	2020-01-15	SOLUTION NMR	Homo sapiens	Double-stranded RNA-binding protein Staufen homolog 1 (protein), hARF1 SBS dsRNA (nucleic acid)
6SDW	Solution structure of Staufen1 dsRBD3+4 - hARF1 SBS dsRNA complex.	2020-01-15	SOLUTION NMR	Homo sapiens	Double-stranded RNA-binding protein Staufen homolog 1 (protein), hARF1 SBS dsRNA (nucleic acid)
4DKK	The X-ray Crystal Structure of the Human STAU1 SSM-RBD'5 Domain-Swapped Dimer	2013-03-20	X-RAY DIFFRACTION 1.701 Å	Homo sapiens	Double-stranded RNA-binding protein Staufen homolog 1 (protein), CL, FLC
6HTU	Structure of hStau1 dsRBD3-4 in complex with ARF1 RNA	2018-11-21	X-RAY DIFFRACTION 2.888 Å	Homo sapiens	Double-stranded RNA-binding protein Staufen homolog 1 (protein), RNA (19-MER) (nucleic acid)

Figura 7. Resultados de la búsqueda de staufen 1 en PDB.

Dada la falta del modelo completo y gracias a su secuencia de aminoácidos, se realizó el modelo de la proteína mediante los softwares SwissModel y Phyre2.

El resultado del modelado con SwissModel fue:



Figura 8. Resultados de distintos modelos posibles en SwissModel.

El modelo seleccionado fue el siguiente:



Figura 9. Modelo de stau1 seleccionado en SwissModel.

El modelo se seleccionó en función de los datos de “coverage” o cobertura de la secuencia, QMEANDisCo y la identidad. Como podemos observar en la Figura 8, el modelo 02 tienen una mayor cobertura, pero todos los demás criterios son más bajos, por lo que se ha escogido el modelo 01. Igualmente, podemos comprobar que la identidad de la secuencia es muy baja (15,96%) y la puntuación QMEANDisCo tampoco es excesivamente buena, ya que preferiríamos que fuera >50% o incluso >60%. Dado que el modelo no es especialmente bueno, se decidió probar otro servidor web para el modelado, Phyre2.

El modelo obtenido con Phyre2 fue el siguiente:

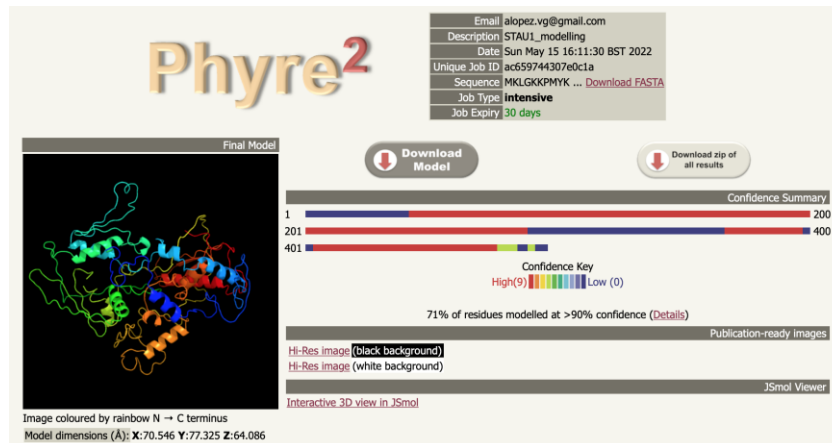


Figura 10. Resultado del modelo de staufer 1 realizado con Phyre2.

En este caso se ha obtenido un modelo en el que el 71% de los residuos está modelado con una confianza >90%, resultado bastante aceptable y, sobre todo, mejor en comparación con el obtenido con SwissModel. La diferencia en este caso es que el modelo de la proteína se ha obtenido en base a la secuencia de distintos *templates* y no de una única secuencia conocida. Los *templates* utilizados para este modelo son los siguientes:

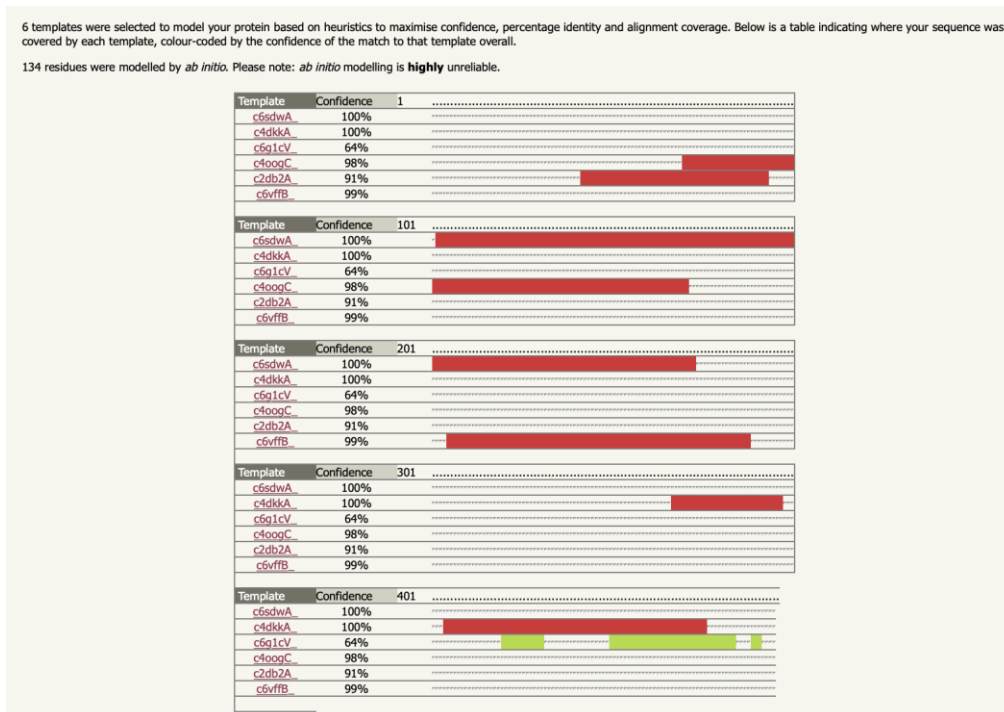


Figura 11. Templates en los que se basa el modelo de staufer 1 de Phyre2.

En rojo observamos las regiones de cada *template* seleccionadas para el modelo y la confianza con la que se representan, basada en el porcentaje de identidad y la cobertura del alineamiento. Si nos fijamos en la Figura 12, en la que se detallan cuatro de los *templates* utilizados, vemos que los tres primeros son estructuras parciales de la proteína de interés observados en PDB, por lo que podemos comprobar que los resultados del modelo son fiables.



#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c6sdwA			100.0	100	<b>PDB header:</b> rna binding protein <b>Chain:</b> A; <b>PDB Molecule:</b> double-stranded rna-binding protein staufen homolog 1; <b>PDBTitle:</b> solution structure of stau1 dsrbd3+4 - harf1 sbs dsrna complex. <b>PDB Entry:</b> <a href="#">PDBe</a> <a href="#">RCSB</a> <a href="#">PDBj</a> 
2	c4dkkA			100.0	100	<b>PDB header:</b> rna binding protein <b>Chain:</b> A; <b>PDB Molecule:</b> double-stranded rna-binding protein staufen homolog 1; <b>PDBTitle:</b> the x-ray crystal structure of the human stau1 ssm-rbd'5 domain-2 swapped dimer <b>PDB Entry:</b> <a href="#">PDBe</a> <a href="#">RCSB</a> <a href="#">PDBj</a> 
3	c6htuA			99.9	97	<b>PDB header:</b> rna binding protein <b>Chain:</b> A; <b>PDB Molecule:</b> double-stranded rna-binding protein staufen homolog 1; <b>PDBTitle:</b> structure of hstau1 dsrbd3-4 in complex with arf1 rna <b>PDB Entry:</b> <a href="#">PDBe</a> <a href="#">RCSB</a> <a href="#">PDBj</a> 
4	d1uhza			99.9	61	<b>Fold:</b> dsRBD-like <b>Superfamily:</b> dsRNA-binding domain-like <b>Family:</b> Double-stranded RNA-binding domain (dsRBD) <b>PDB entry:</b> <a href="#">PDBe</a> <a href="#">RCSB</a> <a href="#">PDBj</a> 
5	c5n8IA			100.0	36	<b>PDB header:</b> rna binding protein <b>Chain:</b> A; <b>PDB Molecule:</b> risc-loading complex subunit tarbp2; <b>PDBTitle:</b> structure of trbp dsrbd 1 and 2 in complex with a 19 bp sirna (complex2 b) <b>PDB Entry:</b> <a href="#">PDBe</a> <a href="#">RCSB</a> <a href="#">PDBj</a> 

Figura 12. Descripción de los tres primeros templates en los que se basa el modelo de staufen 1 de Phyre2.

Dado que, en este caso, la confianza en el modelo es superior y éste es más completo que el obtenido con SwissModel, se ha elegido el modelo de Phyre2 para pasos posteriores. En la Figura 13 podemos observar ambos modelos.

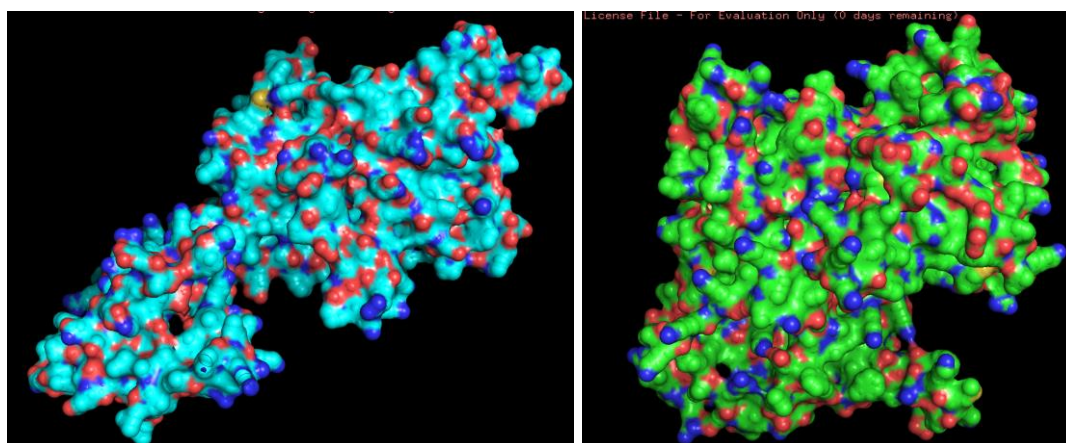


Figura 13. Modelo de staufen 1 realizado con SwissModel (izquierda) y Phyre2 (derecha).

## 5.2. Análisis de los sitios de unión

Una vez tenemos el modelo, podemos estudiar sus lugares de unión y ligandos, en el caso de que estén descritos.

Antes de descartar completamente el modelo resultante de SwissModel, realizamos una búsqueda de lugares de unión con 3DLigandSite, ya que no se había observado ningún ligando en el modelo presentado. De esta búsqueda

tampoco obtuvimos ningún resultado por lo que descartamos definitivamente el modelo.

En cuanto al modelo obtenido con Phyre2, en la Figura 12 observamos el botón de *Run Investigator* para poder detectar “pockets” o lugares de unión mediante *fpocket*. Para este objetivo, realizamos un *Run Investigator* de los tres primeros *templates*. Además, este mismo modelo se utilizó para realizar una búsqueda en 3DLigandSite también.

Por último, en una búsqueda inicial y como observamos en la Figura 5, la estructura parcial de *staufen1*, 4DKK en PDB (*template* sobre el que se ha realizado un *Run Investigator* también), presenta dos ligandos; el ion cloruro o CL y el anion citrato o FLC, que serán contemplados en resultados posteriores.

En el análisis con *fpocket* realizado a través de Phyre2 observamos los siguientes *pockets* o lugares de unión:

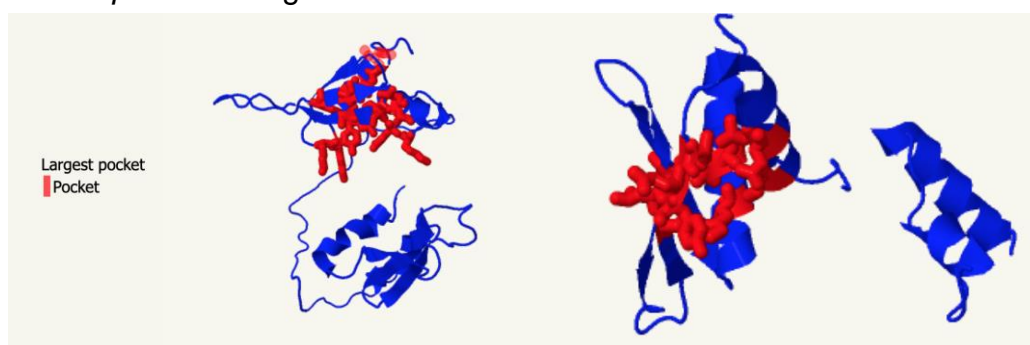


Figura 14. Lugares de unión identificados con *fpocket* sobre los *templates* 6SDW y 4DKK.

Los lugares de unión descritos se encuentran en los *templates* 6SDW y 4DKK, previamente descritos como estructuras parciales de *staufen 1*. Para el cribado basado en la estructura, se extrajeron las secuencias de estos lugares de unión y se utilizaron en el formato correcto en MTiOpenScreen (apartado 5.3).

Por otro lado, se utilizó también 3DLigandSite para detectar lugares de unión y ligandos en el modelo de Phyre2 y, como esperábamos, uno de los ligandos descritos fue CL, presente en el modelo parcial 4DKK.

Residue	AA	Min. Lig Dist	Lig Contacts	Conservation	RSA	Probability
#NM_Cluster_1	CL_2_4dkk_A					
207	ILE	1.6	1.0	0.67	28.83	1.0
208	SER	4.96	1.0	0.67	39.84	0.41
260	LYS	3.44	1.0	0.72	37.65	0.95

Tabla 1. Características de la interacción del ligando CL en el modelo de *staufen 1* (Phyre2).

Previamente hemos mencionado que el modelo 4DKK presenta dos ligandos en PDB, CL y FLC (Figura 15).

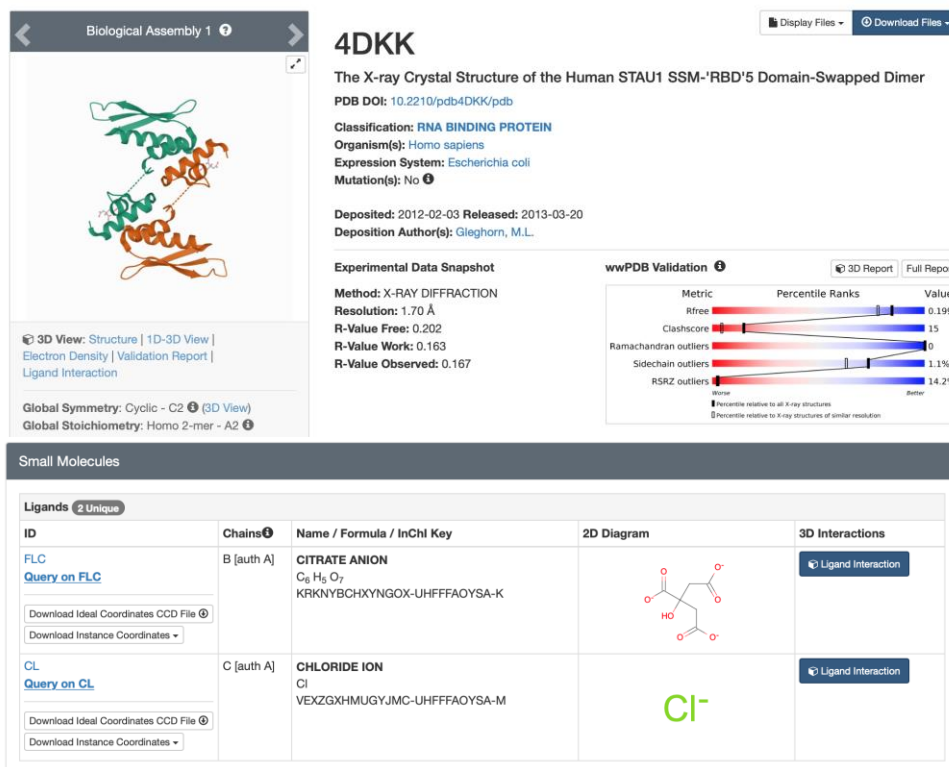


Figura 15. Descripción de los ligandos CL y FLC de la estructura 4DKK de PDB.

Sin embargo, en 3DLigandSite solo hemos podido observar el ligando CL. Se decidió estudiar la homología de la secuencia utilizada de staufen 1 y verificar si coincidía con la secuencia del modelo 4DKK y por lo tanto, confirmar si estructuralmente es posible que el ligando FLC también estuviera presente. La secuencia de la estructura 4DKK corresponde a la isoforma a de la proteína, mientras que la utilizada en este trabajo es la secuencia de UniProtKB O95793 (referencia de staufen 1), por lo tanto, al estudiar la homología de estas secuencias con BLASTP, muestra un 97% de identidad, coincidiendo con una sección posterior de la secuencia escogida (362 a 476). De hecho, si nos fijamos en la Figuras 11 y 12, efectivamente esa zona del modelo de la proteína utiliza como *template* la estructura 4DKK.

ref|NP\_059348.2|:1-496 double-stranded RNA-binding protein Staufen homolog 1 isoform a [Homo sapiens]

Sequence ID: Query\_53337 Length: 496 Number of Matches: 1

Range 1: 362 to 476 [Graphics](#)

[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
231 bits(590)	4e-80	Compositional matrix adjust.	111/115(97%)	112/115(97%)	0/115(0%)
Query 1		GPLGSKATVTAMIARELLYGGTSPTAETILKNNISSGHVPHGPLTRPSEQLDYLSRVQGF			60
P		+KATVTAMIARELLYGGTSPTAETILKNNISSGHVPHGPLTRPSEQLDYLSRVQGF			
Sbjct 362		APNPAKATVTAMIARELLYGGTSPTAETILKNNISSGHVPHGPLTRPSEQLDYLSRVQGF			421
Query 61		QVEYKDFPKNNKNEFVSLINCSSQPPLISHGIGKDVESCHDMAALNLIKLLSELD			115
Sbjct 422		QVEYKDFPKNNKNEFVSLINCSSQPPLISHGIGKDVESCHDMAALNLIKLLSELD			476

Figura 16. Resultado de la homología entre la secuencias de staufen 1.

Por lo tanto, en apartados posteriores se trabajará con la presencia de los ligandos CL y FLC, de forma que también podremos estudiar si los deberíamos contemplar como ligandos del modelo o no.

La información disponible de los ligandos en PDB es la siguiente:

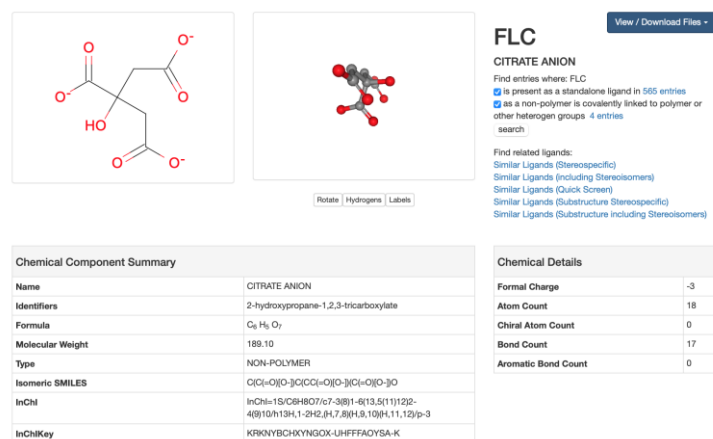


Figura 17. Descripción del ligando FLC consultada en PDB.

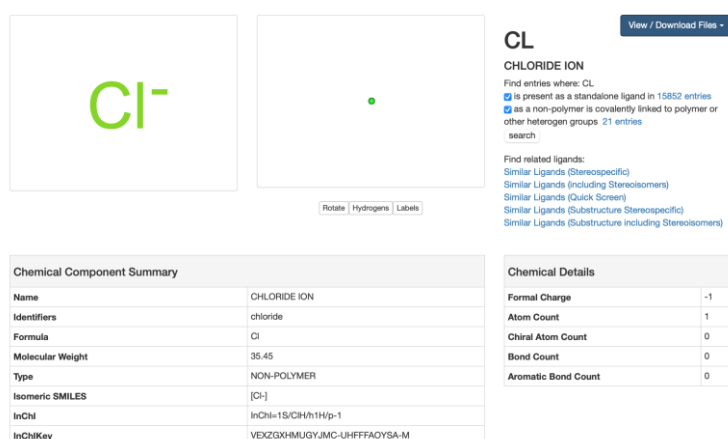


Figura 18. Descripción del ligando CL consultada en PDB.

### 5.3. Cribado virtual basado en la estructura

El cribado virtual basado en la estructura se realizó con MTiOpenScreen gracias al modelo de staufen 1 obtenido de Phyre2, a los dos *pockets* seleccionados gracias a fpocket y al lugar de unión de CL descrito con 3DLigandSite. Las secuencias de aminoácidos de los *pockets* en el formato requerido en MTiOpenScreen son las siguientes:

Pocket 1 (fpocket):

\_A\_ILE\_207\_\_, \_A\_LEU\_210\_\_, \_A\_ALA\_211\_\_, \_A\_LYS\_220\_\_, \_A\_GLU\_221\_\_  
 \_\_, \_A\_PRO\_222\_\_, \_A\_TYR\_224\_\_, \_A\_MET\_240\_\_, \_A\_VAL\_242\_\_, \_A\_LYS\_2  
 56\_\_, \_A\_LYS\_257\_\_, \_A\_LYS\_260\_\_, \_A \_\_ALA\_263\_\_, \_A \_\_ALA\_264\_\_, \_A  
 \_\_MET\_267\_\_

Pocket 2 (fpocket):

\_A\_SER\_409\_\_,\_A\_LEU\_412\_\_,\_A\_ASP\_413\_\_,\_A\_VAL\_423\_\_,\_A\_GLU\_424\_\_,\_A\_TYR\_425\_\_,\_A\_SER\_438\_\_,\_A\_LEU\_439\_\_,\_A\_ILE\_440\_\_,\_A\_HIS\_461\_\_,\_A\_ALA\_464\_\_,\_A\_ALA\_465\_\_

Pocket 3 (3DLigandSite): \_A\_ILE\_207\_\_,\_A\_SER\_208\_\_,\_A\_LYS\_260\_\_

El proceso de docking se realizó siguiendo las instrucciones descritas en el apartado de materiales y métodos. Se han seleccionado los 25 compuestos con mejor resultado de docking y la energía de cada unión, para cada *pocket*. Los resultados de docking se han adjuntado en la carpeta de Google Drive mencionada en el trabajo

Los 25 compuestos con un mejor resultado de docking para el *pocket 1* son los siguientes:

Compuesto	Energy	nRot
Siramesine_ZINC000001533883	-12.7	6
Capmatinib_ZINC000043195321	-12.5	4
Tegobuvir_ZINC000100057121	-12.3	6
Tasosartan_ZINC000013444037	-11.9	4
Prulifloxacin_ZINC000001269201	-11.8	4
Hesperidin_ZINC000252008997	-11.8	15
Radiprodil_ZINC000028363953	-11.5	5
Sanguinarium_ZINC000000000706	-11.4	0
Lidoflazine_ZINC000022034381	-11.4	9
Ripisartan_ZINC000000538309	-11.4	6
Guadecitabine_ZINC000043203165	-11.4	12
Palosuran_ZINC000034375693	-11.3	7
Hesperidin_ZINC000038144570	-11.3	15
Ziprasidone_ZINC000000538550	-11.3	4
Ci-988_ZINC000014952331	-11.3	13
Regorafenib_ZINC000006745272	-11.2	6
Glisindamide_ZINC000000537799	-11.2	6
Altiratinib_ZINC000113198271	-11.2	8
Etalocib_ZINC000003930629	-11.1	14
Tak-593_ZINC000095579919	-11.1	6
Dihydroergocristine_ZINC000003947494	-11.1	6
Radotinib_ZINC000059749972	-11.1	7
Bisantrene_ZINC000001577736	-11.1	6
Doramapimod_ZINC000024044436	-11.0	8
Bifeprunox_ZINC000052971454	-11.0	4

Tabla 2. Mejores resultados de docking para el pocket 1.

Como ejemplo de la visualización de la unión de un compuesto en el pocket descrito, hemos utilizado el fármaco Siramesine, con la mejor energía de unión a la proteína.

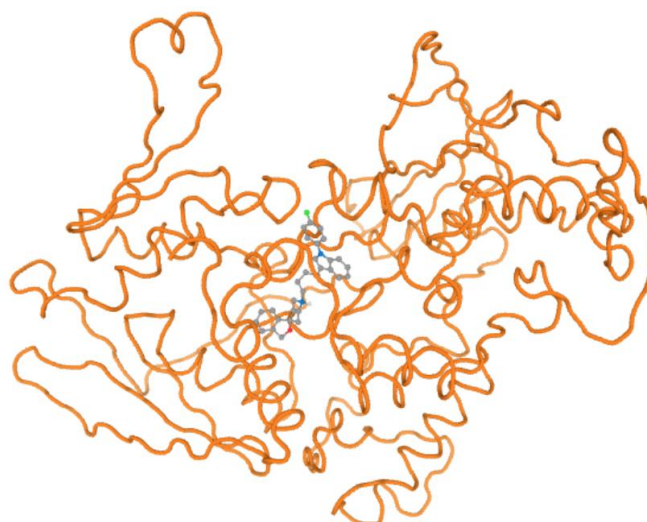


Figura 19. Representación de la unión de Siramesine con el modelo de la proteína de Staufen1.

Los 25 compuestos con un mejor resultado de docking para el pocket 2 son los siguientes:

Compound	Energy	nRot
Glisindamide_ZINC000000537799	-12.3	6
Radotinib_ZINC000059749972	-12.3	7
Rebastinib_ZINC000063933734	-12.2	7
Quizartinib_ZINC000043204002	-11.9	8
Pf-00477736_ZINC000019862643	-11.8	5
Hesperidin_ZINC000100204058	-11.7	15
Metrenperone_ZINC000000537997	-11.6	5
Regorafenib_ZINC000006745272	-11.6	6
Bafetinib_ZINC000022940637	-11.5	9
Hesperidin_ZINC000104889096	-11.5	15
Xaliproden_ZINC000000577115	-11.5	5
Fluazuron_ZINC000002570819	-11.5	5
Lumacaftor_ZINC000064033452	-11.5	5
Gepotidacin_ZINC00022327887	-11.4	5
Risperidone_ZINC000000538312	-11.3	4
Peliglitazar_ZINC000003935841	-11.3	12
Paliroden_ZINC000035826853	-11.3	6
Hesperidin_ZINC000008382287	-11.3	15
Clopimozide_ZINC000004215428	-11.3	7
Cinperene_ZINC000002018994	-11.3	5
Tulopafant_ZINC000002016001	-11.2	5
Cenisertib_ZINC000063298059	-11.2	6
Imiglitazar_ZINC000049756486	-11.2	11
Talniflumate_ZINC000000601275	-11.2	6
Difluanine_ZINC000019364169	-11.2	10

Tabla 3. Mejores resultados de docking para el pocket 2.

En este caso, para el ejemplo de visualización del fármaco unido a la proteína se ha escogido Glisindamide, con la mejor energía de unión de los compuestos estudiados.

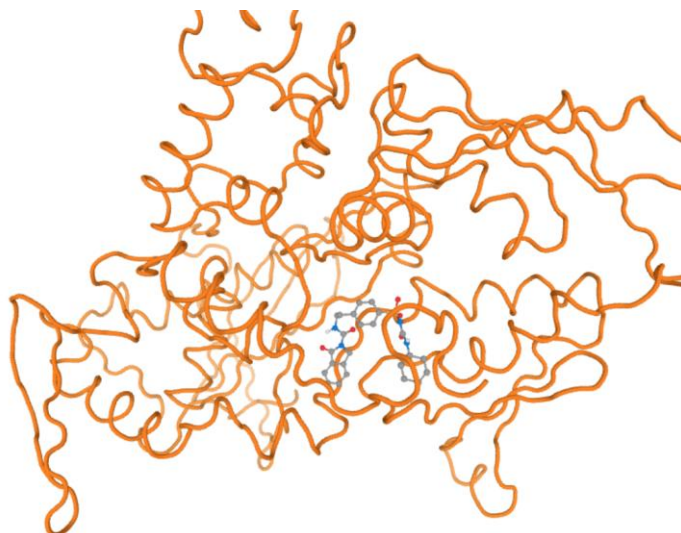


Figura 20. Visualización de la unión del fármaco Glisindamide en el modelo de staufen1.

En el caso del *pocket* 3, el docking no ha resultado exitoso, ya que los resultados de las energías de todos los compuestos propuestos han sido igual a 0, es decir, no hay compuestos que se unan a ese lugar. Por ese motivo, no se presenta la tabla de resultados.

En la visualización de uno de los compuestos, observamos que el *pocket* con menos residuos definidos, al que se une el ion cloruro, no resulta una buena estructura de la proteína sobre la que realizar el estudio de docking, ya que ni la molécula mejor puntuada se une a la proteína de interés.

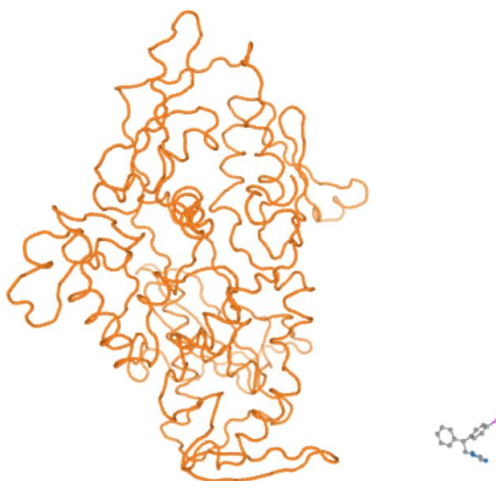


Figura 21. Visualización del fármaco Brolaconazole y el modelo de la proteína staufen1

#### 5.4. Cribado virtual basado en el ligando

En el desarrollo de este trabajo hemos identificado un ligando del modelo, el ion cloruro o CL e hipotetizamos que podría haber otro, el anión citrato o FLC. Con estos dos compuestos se realizará un cribado virtual basado en ambos ligandos. El primer paso ha sido utilizar SwissDock para comprobar las energías de las distintas interacciones del ligando con el modelo.

El ligando CL nos presentó una dificultad, y es que al ser el ion cloruro y al intentar añadir hidrógenos sin éxito, no se pudo realizar el proceso de docking.

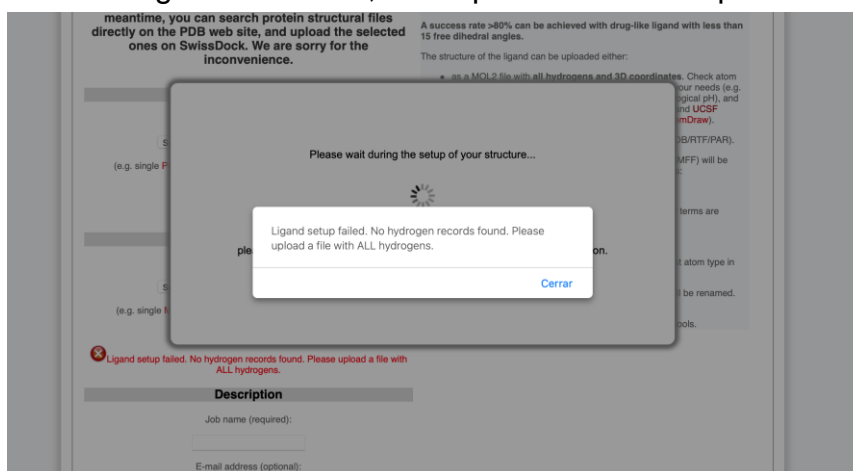


Figura 22. Error encontrado en SwissDock al intentar realizar el proceso de docking con el ligando CL.

Dada la simplicidad del ligando y la imposibilidad de realizar el proceso de docking, se descartó para la continuación del trabajo. Por lo tanto, nos centramos en el posible ligando FLC.

El resultado de docking con el ligando FLC fue el siguiente:

Cluster	Element	FullFitness (kcal/mol)	Estimated $\Delta G$ (kcal/mol)
0	0	-3121.16	-12.15
1	0	-3120.56	-12.09
2	0	-3119.32	-12.23
3	0	-3118.68	-10.87
4	0	-3117.64	-12.47
5	0	-3117.46	-11.84
6	0	-3117.33	-11.66
7	0	-3117.32	-14.00
8	0	-3116.51	-11.55
9	0	-3116.45	-11.66
10	0	-3116.05	-11.39
11	0	-3115.93	-11.65
12	0	-3115.35	-12.12
13	0	-3115.24	-9.74
14	0	-3114.98	-11.03
15	0	-3114.64	-9.41
16	0	-3114.58	-11.97
17	0	-3114.52	-11.16
18	0	-3114.10	-11.89
19	0	-3113.98	-11.01

Tabla 4. Inicio de la tabla de energías resultante del docking de staufen 1 y el ligando FLC en SwissDock.

En los resultados del docking observamos 42 clusters, es decir, 42 posiciones donde se puede unir el ligando FLC. La energía de Gibbs estimada para cada elemento en todas estas posiciones es negativa, lo que indica que es un proceso exergónico y, por lo tanto, espontáneo, con lo que asumimos que FLC es un ligando del modelo de staufen 1 y procedemos a realizar la selección de compuestos.



A continuación, se utilizó SwissSimilarity para encontrar compuestos similares a FLC en cuanto a estructura, gracias a su fórmula en formato SMILES. Se utilizaron tres bases de datos de compuestos distintas, compuestos comerciales depositados en ZINC (enfocado al reposicionamiento y conocer más datos de los compuestos) y aquellos definidos como fármacos en las bases de datos de DrugBank y aquellos aprobados en ChEMBL.

Compuestos similares a FLC en la selección de ZINC:

ZINC ID	Similarity Score	SMILES
ZINC000002572383	0.998	C[C@H](C(O)=O)C@@(O)(CC(O)=O)C(O)=O
ZINC000002572383	0.998	C[C@H](C(O)=O)C@@(O)(CC(O)=O)C(O)=O

Tabla 5. Compuestos similares a FLC de la base de datos de ZINC.

Compuestos similares a FLC en la selección de ChEMBL:

ChEMBL ID	Similarity Score	SMILES
CHEMBL1261	1.000	OC(=O)CC(O)(CC(O)=O)C(O)=O
CHEMBL1455497	0.993	OC(CC(O)=O)C(O)=O

Tabla 6. Compuestos similares a FLC de la base de datos ChEMBL.

Compuestos similares a FLC en la selección de DrugBank:

DrugBank ID	Similarity Score	Usual Name	SMILES
DB15556	1.000	Zinc glycinate citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB11110	1.000	Magnesium citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB04272	1.000	Citric acid	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB06784	1.000	Gallium citrate Ga-67	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB09125	1.000	Potassium citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB09154	1.000	Sodium citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB09275	1.000	Bismuth subcitrate potassium	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB14526	1.000	Chromic citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB11093	1.000	Calcium citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB09501	1.000	Ferric ammonium citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB11154	1.000	Zinc citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB11233	1.000	Silver citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB13582	1.000	Ferric (59Fe) citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB13771	1.000	Ferric sodium citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB13995	1.000	Ferric pyrophosphate citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB14495	1.000	Manganese citrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB14520	1.000	Tetra ferric tricitrate decahydrate	OC(=O)CC(O)(CC(O)=O)C(O)=O
DB01727	1.000	Isocitric Acid	O[C@@H]([C@H](CC(O)=O)C(O)=O)C(O)=O
DB03499	0.993	D-Malic acid	O[C@H](CC(O)=O)C(O)=O
DB03343	0.993	Malate Like Intermediate	O[C@H](CC(O)=O)C(O)=O
DB15555	0.993	Magnesium malate	OC(CC(O)=O)C(O)=O
DB12751	0.993	Malic acid	OC(CC(O)=O)C(O)=O

Tabla 7. Compuestos similares a FLC de la base de datos DrugBank.

Como podemos observar, la descripción SMILES de los 17 primeros compuestos es igual, ya que estos responden a conjugaciones del anión citrato con diferentes iones. Por lo tanto, el código ZINC de todos estos elementos será el mismo.

Los compuestos presentados se unieron en una misma base de datos en formato .mol2 para poder utilizarlos en MTiOpenScreen para el proceso de docking con el modelo de la proteína. Además, se obtuvieron los códigos ZINC de todos ellos gracias a su SMILES.

Los compuestos con mejores energías de docking fueron los siguientes:

Compuesto	Energía	nRot
ZINC000000895175	-6.1	6
ZINC000003869734	-6.0	6
ZINC000003869734	-6.0	6
ZINC000000895180	-6.0	6
ZINC000000895081	-5.8	6
ZINC000002572383	-5.8	6
ZINC000000895176	-5.8	6
ZINC000002572383	-5.8	6
ZINC000000901307	-5.7	6
ZINC000002572371	-5.7	6
ZINC000000901307	-5.6	6
ZINC000000895179	-5.5	6
ZINC000000895081	-5.5	6
ZINC000002572371	-5.4	6
ZINC000000895264	-5.1	4
ZINC000000895074	-4.9	4

Tabla 8. Resultados de docking de compuestos similares a FLC.

## 5.5. Análisis y comparación de las moléculas seleccionadas

No hay coincidencias entre las moléculas seleccionadas en el screening basado en la estructura y aquellas basadas en el ligando. Por otro lado, entre los compuestos del screening basado en la estructura, hay cuatro que coinciden entre los 25 mejores de cada lugar de unión, cuyas estructuras son:

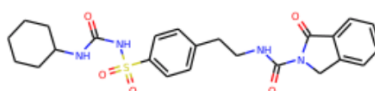


Figura 23. Estructura de Glisindamide.

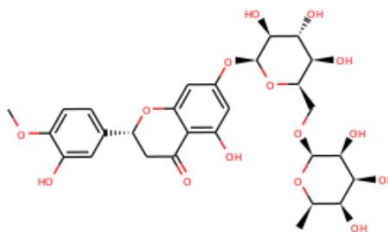


Figura 24. Estructura de Hesperidin.

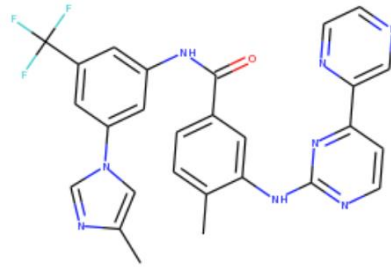


Figura 25. Estructura de Radotinib.

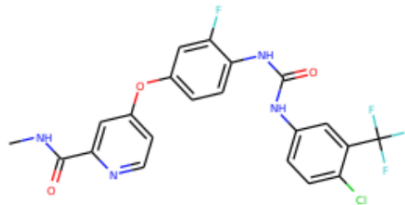


Figura 26. Estructura de Regorafenib.

Para ampliar un poco más la información sobre ellos, se han buscado estas moléculas en PubChem y se ha obtenido información sobre su indicación farmacológica:

- Glisindamide: es un compuesto del grupo de la sulfonilurea con actividad antihiper glucémica. Tiene una alta capacidad de unión al receptor de sulfonilurea tipo I (SUR1) y se ha descrito también actividad agonista al receptor activador de la proliferación de peroxisoma (PPARgamma).
- Hesperidin: es un compuesto natural derivado del hesperetin cuyo principal rol es mutágeno.
- Radotinib: es un compuesto que actualmente se encuentra en investigación para algunos subtipos de leucemias.
- Regorafenib: es un inhibidor de distintas kinasas utilizado como terapia para diversos tipos de cáncer como el cancer colorrectal metastásico refractorio, el carcinoma hepatocelular o los tumores gastrointestinales.

## 6. Discusión

En este trabajo se han descrito dos modelos de la proteína STAU1 obtenidos mediante servidores web de modelado de estructuras. Estos modelos, especialmente el obtenido en Phyre2, mucho más completo y con lugares de unión definidos, nos ha permitido poder realizar un screening de compuestos de unión al modelo de la proteína de interés, obteniendo una lista de 1500 moléculas para dos de los lugares de unión descritos. Hemos podido visualizar los 25 compuestos con mejores puntuaciones de docking entre estas dos listas y, además, observar que algunos de ellos son comunes para ambos lugares de unión.

Por otro lado, a través de la visualización de algunas estructuras parciales de STAU1 hipotetizamos que el ligando FLC podía unirse también al modelo creado de la proteína. A través de un estudio de docking hemos demostrado que también se puede unir a la molécula modelo, lo que nos ha permitido obtener una lista muy restringida de compuestos basada en su estructura, dadas también las específicas búsquedas que se han realizado con SwissSimilarity.

En un inicio, la idea de realizar un cribado de compuestos en base a un ligando conocido de la proteína resultaba más atractiva, dado que ya contábamos con una interacción demostrada y descrita en la literatura. El resultado, pero, ha sido más limitado, dado que la simplicidad del ligando y la falta de bibliografía sobre su efecto sobre la actividad de la proteína no nos ha permitido realizar el cribado específico que esperábamos. Por otro lado, el cribado basado en la estructura nos ha permitido estudiar un base de datos mucho más extensa de la cual hemos podido obtener compuestos con efecto conocido sobre la homeostasis celular, lo que nos puede permitir realizar una valoración o hipotetizar sobre los efectos que pueden tener en células DM1.

En resumen, pese a que se esperaba tener un cribado de fármacos más específico basado en el ligando de unión a la proteína, hemos obtenido una librería limitada de compuestos que, además, presentan peores puntuaciones de docking que la obtenida mediante la estructura. Por esta razón, pese a que no quedarán descartados los compuestos similares al ligando, sí que se valoran de forma más positiva los resultados de los cribados correspondientes a los dos lugares de unión basados en la estructura para futuros ensayos. Así, hemos obtenido una lista de fármacos a nuestro alcance que se podrían ensayar en modelos pre-clínicos de la enfermedad mediante técnicas *in vitro* y así estudiar su papel modulador de staufen 1 y los beneficios (esperados) sobre la patogénesis de la DM1.

## 7. Conclusiones

### 7.1. Conclusiones

- Se han obtenido dos modelos distintos de la estructura de la proteína STAU1 gracias a su secuencia de aminoácidos y distintas plantillas almacenadas en PDB.
- Se ha obtenido una lista de 16 ligandos similares a FLC, de los cuáles se ha podido estudiar su unión al modelo de la proteína STAU1 mediante docking.
- Se han obtenido dos bases de datos de 1500 compuestos que se unen a dos lugares de unión distintos de la proteína STAU1, de los cuáles se han podido visualizar los 25 con mejores puntuaciones de cada lista y además, se ha comprobado que cuatro de ellos son comunes; Glisindamide, Hesperidin, Radotinib y Regorafenib. De estos cuatro, además, se han descrito sus aplicaciones para poder valorarlas en futuros experimentos.
- Pese a que son resultados muy iniciales, las técnicas utilizadas nos han permitido lograr los objetivos planteados; obtener un listado de compuestos de unión a la proteína de interés para estudiar en futuros ensayos *in vitro*.

### 7.2. Líneas de futuro

En un futuro, sería realmente interesante poder evaluar las propiedades farmacocinéticas y la toxicidad de los 25 primeros compuestos resultantes de cada cribado basado en la estructura, así como de los compuestos resultantes del cribado basado en el ligando. Para esto se podría utilizar VEGA QSAR y pkCSM, de los que he conocido su existencia a partir de la realización de este trabajo.

Además, una vez realizado este estudio, se valora adquirir algunos de estos compuestos para realizar ensayos *in vitro* en células musculares de pacientes de DM1 y observar así su efecto sobre la actividad de STAU1 y la expresión de proteínas reguladas por esta última, así como valorar otros *hallmarks* de la enfermedad.

### 7.3. Seguimiento de la planificación

Siendo crítica, me habría gustado poder planificar y realizar el trabajo con los conocimientos que tengo ahora mismo y no con mis conocimientos iniciales. Considero que mi búsqueda inicial y planteamiento del trabajo carecía de preparación, ya que, si hubiera valorado realmente todos los aspectos conflictivos que podían surgir, como la falta de ligandos en la estructura o la

variedad de resultados entre servidores de modelado, el planteamiento de los objetivos habría sido diferente. Además, no contaba con la suficiente información sobre softwares o servidores web y sus requerimientos y utilidades, lo que me ha hecho ir aprendiendo “sobre la marcha” y a base de error-acierto, que en muchos casos han retrasado el trabajo debido a las grandes demoras para obtener los resultados de algunos servidores como DockBlaster.

Si volviera a repetir el trabajo, sin duda pediría más ayuda al tutor, así como bibliografía, consultas de “work-in-progress” y comentaría el workflow con tiempo suficiente.

## 8. Glosario

**STAU1:** proteína de unión a ARN de doble cadena staufen 1. También se menciona como staufen 1

**DM1:** distrofia miotónica tipo 1

**Screening:** se utiliza indistintamente como cribado. Nos permite seleccionar compuestos en función a su interacción con la proteína diana.

**Docking:** o acoplamiento molecular. Estudios que analizan las energía necesaria y posición de la unión de un compuesto en un lugar de unión.

**3'UTR:** 3' untranslated region o región 3' sin traducir

**CTG:** Citosina Timina Guanina

**DMPK:** distrofia miotónica proteína kinasa

**FLC:** anion citrato

**CL:** ion cloruro

**MBNL1:** muscleblind-like protein

**CELF1:** CUGBP Elav-like family member 1

**Splicing:** simplificación de splicing alternativo, proceso por el cuál se obtienen mRNA maduros.

**NCBI:** base de datos del National Center for Biotechnology Information

**UniProtKB:** base de datos de Universal Protein Resource.

**PyMOL:** es un sistema de visualización de moléculas.

**SMILES:** simplified molecular-input line-entry system. Descriptivo de la formula de moléculas de amplio uso.

**ZINC:** es una base de datos gratuita de moléculas comerciales.

**DrugBank:** base de datos de fármacos con información sobre los mismos y sus estructuras diana.

**ChemBL:** base de datos de moléculas bioactivas con propiedades farmacológicas.

**PubChem:** base de datos de moléculas de diversos tipos.

## 9. Bibliografía

1. Day, J. W. & Ranum, L. P. W. Genetics and molecular pathogenesis of the myotonic dystrophies. *Current Neurology and Neuroscience Reports* (2005) doi:10.1007/s11910-005-0024-1.
2. Yum, K., Wang, E. T. & Kalsotra, A. Myotonic dystrophy: disease repeat range, penetrance, age of onset, and relationship between repeat size and phenotypes. *Curr Opin Genet Dev* **44**, 30–37 (2017).
3. Ballester-Lopez, A. *et al.* The need for establishing a universal CTG sizing method in myotonic dystrophy type 1. *Genes (Basel)*. **11**, 1–9 (2020).
4. Ballester-Lopez, A. *et al.* A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype. *Hum. Mutat.* **41**, 420–431 (2020).
5. Brook, J. D. *et al.* Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **69**, 385 (1992).
6. Savic Pavicevic, D. *et al.* Molecular genetics and genetic testing in myotonic dystrophy type 1. *Biomed Res Int* **2013**, 391821 (2013).
7. López-Martínez, A., Soblechero-Martín, P., De-La-puente-ovejero, L., Nogales-Gadea, G. & Arechavala-Gomez, V. An Overview of Alternative Splicing Defects Implicated in Myotonic Dystrophy Type I. *Genes (Basel)*. **11**, 1109 (2020).
8. Empanaza, J. I. *et al.* Cancer Phenotype in Myotonic Dystrophy Patients: Results from a Meta-analysis. *Muscle Nerve* (2018) doi:10.1002/mus.26194.
9. Pascual-Gilabert, M., López-Castel, A. & Artero, R. Myotonic dystrophy type 1 drug development: A pipeline toward the market. *Drug Discov. Today* **26**, 1765–1772 (2021).
10. Hammond, S. M. *et al.* Delivery of oligonucleotide- based therapeutics: challenges and opportunities. *EMBO Mol. Med.* **13**, 1–23 (2021).
11. Konieczny, P. *et al.* Myotonic dystrophy: candidate small molecule therapeutics. *Drug Discov Today* **22**, 1740–1748 (2017).
12. Burley, S. K. *et al.* Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. in *Physiology & behavior* (eds. Wlodawer, A., Dauter, Z. & Jaskolski, M.) vol. 1607 627–641 (Springer New York, 2017).
13. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
14. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
15. Wass, M. N., Kelley, L. A. & Sternberg, M. J. E. 3DLigandSite: Predicting ligand-binding sites using similar structures. *Nucleic Acids Res.* **38**, 469–473 (2010).
16. Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: An open source



- platform for ligand pocket detection. *BMC Bioinformatics* **10**, 1–11 (2009).
17. Labbé, C. M. *et al.* MTiOpenScreen: A web server for structure-based virtual screening. *Nucleic Acids Res.* **43**, W448–W454 (2015).
  18. Grosdidier, A., Zoete, V. & Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* **39**, 270–277 (2011).
  19. 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.* **56**, 1399–1404 (2016).
  20. Ozimski, L. L., Sabater- Arcis, M., Bargiela, A. & Artero, R. The hallmarks of myotonic dystrophy type 1 muscle dysfunction. *Biol. Rev.* **in press**, brv.12674 (2020).
  21. Cerro-Herreros, E. *et al.* Preclinical characterization of antagomiR-218 as a potential treatment for myotonic dystrophy. *Mol. Ther. - Nucleic Acids* **26**, 174–191 (2021).
  22. Wang, M. *et al.* Correction of Glycogen Synthase Kinase 3 $\beta$  in Myotonic Dystrophy 1 Reduces the Mutant RNA and Improves Postnatal Survival of DMSXL Mice. *Mol. Cell. Biol.* **39**, 1–20 (2019).
  23. DesGroseillers, L. & Lemieux, N. Localization of a human double-stranded RNA-binding protein gene (STAU) to band 20q13.1 by fluorescence in situ hybridization. *Genomics* **36**, 527–9 (1996).
  24. Ravel-Chapuis, A. *et al.* Staufen1 impairs stress granule formation in skeletal muscle cells from myotonic dystrophy type 1 patients. *Mol. Biol. Cell* **27**, 1728–1739 (2016).
  25. Bondy-Chorney, E., Crawford Parks, T. E., Ravel-Chapuis, A., Jasmin, B. J. & Cote, J. Staufen1s role as a splicing factor and a disease modifier in Myotonic Dystrophy Type I. *Rare Dis* **4**, e1225644 (2016).
  26. Crawford Parks, T. E., Marcellus, K. A., Péladeau, C., Jasmin, B. J. & Ravel-Chapuis, A. Overexpression of Staufen1 in DM1 mouse skeletal muscle exacerbates dystrophic and atrophic features. *Hum. Mol. Genet.* **3**, 3–5 (2020).

## 10. Anexos

Los datos anexos se encuentran en el siguiente [link](https://drive.google.com/drive/folders/1YS1wm08qD-eRizBfEPGMc-1knFsX8NuW?usp=sharing) de Google Drive: <https://drive.google.com/drive/folders/1YS1wm08qD-eRizBfEPGMc-1knFsX8NuW?usp=sharing>