

UNIVERSIDAD DE ZARAGOZA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR Y
CELULAR



Universidad
Zaragoza

TRABAJO FIN DE GRADO

Título: Modelado molecular de las regiones flexibles del factor de inducción de apoptosis humano y de las interacciones con sus ligandos.

Memoria presentada por JAVIER MATA TREVIÑO,
Estudiante de Biotecnología, para optar al título de
Biotecnología por la Universidad de Zaragoza.
Septiembre de 2016

Profesores:
Milagros Medina Trullenque
Víctor Polo Ortiz

ÍNDICE:

1. Resumen.
2. Introducción.
 - 2.1. El factor de inducción de apoptosis mitocondrial.
 - 2.2. Estructura de la AIFM1h.
 - 2.3. Interacción de la AIFM1h con otras proteínas.
3. Objetivos.
4. Materiales y métodos.
 - 4.1. Ficheros informáticos.
 - 4.2. Programas informáticos.
 - 4.3. Servidores web.
5. Resultados.
 - 5.1. Búsqueda de una posible estructura secundaria del fragmento que no se observa en la estructura cristalina.
 - 5.2. Obtención de un fichero de coordenadas con el fragmento no observable en la estructura cristalina.
 - 5.2.1. Modelado manual de la región desordenada.
 - 5.2.2. Modelado por homología de la región desordenada de 4BV6.
 - 5.3. Realización de dinámica molecular de las estructuras obtenidas por Swiss-model, mediante el uso del programa Discovery Studio.
 - 5.4. Evaluación del acoplamiento de AIFM1h con otras proteínas con las que interacciona *in vivo*.
 - 5.4.1. Interacción con AIFM1h en estado oxidado.
 - 5.4.2. Interacción con AIFM1h en estado reducido.
6. Propuesta de una práctica para la asignatura de Bioinformática del grado de Biotecnología.
7. Conclusiones.
8. Bibliografía.
9. Anexos

I. RESUMEN:

La isoforma 1 del factor de inducción de apoptosis mitocondrial humano (AIFM1h) es una proteína relevante en dos procesos de vital importancia para el ser humano, la respiración mitocondrial y la apoptosis celular. Pero el desconocimiento de su estructura terciaria completa, su localización celular y sus regiones desordenadas hacen que sea muy difícil su estudio y caracterización en el laboratorio, por lo que se necesitan otros métodos para la obtención de información. El objetivo de este estudio es la obtención de datos teóricos por medio de la bioinformática que ayuden a completar la estructura terciaria de la proteína y su posible unión con algunas proteínas que se ha comprobado que interaccionan *in vivo* con ella.

Se ha demostrado que AIFM1h es una proteína singular en cuanto a su estructura terciaria se refiere, ya que ninguna otra proteína posee la conformación estructural que falta por caracterizar en su estructura terciaria. Por lo que se necesitará de herramientas de predicción de estructura terciaria y dinámica molecular para completarla y poder evaluar la unión con otras proteínas con las que interacciona *in vivo*.

I. ABSTRACT:

The isoform 1 of the apoptosis-inducing factor mitochondrial from human (AIFM1h) is a relevant protein in two processes which are very important for the human being, cellular respiration and apoptosis. But the ignorance of its complete tertiary structure, cellular location and disordered regions turn into the study and portrayal in the laboratory in a very complex process, so that other methods are necessary for getting more information. The objective of this research is the obtaining of theoretical data throughout bioinformatic tools that can help to find out the protein tertiary structure and its potential link with other proteins that has been proved to interact with AIFM1h *in vivo*.

It has been shown that AIFM1h is a singular protein according to its tertiary structure, because any other protein owns the structural shape which is not described in the tertiary structure of the protein. For this reason tertiary structure prediction tools and molecular dynamics will be needed to complete AIFM1h structure and can assess the link with other proteins which interact with AIFM1h *in vivo*

II. INTRODUCCIÓN:

1. El factor de inducción de la apoptosis mitocondrial (AIFM1).

La isoforma 1 del factor de inducción de apoptosis mitocondrial humano (AIFM1h) es una proteína de 613 aminoácidos, que contiene como cofactor una molécula de FAD. Su localización principal en células sanas es el espacio intermembrana de la mitocondria (IMS), pero cuando la célula recibe un estímulo apoptótico también puede encontrarse en el citosol o en el núcleo (Sevrioukova, 2011; Ferreira et al., 2014; Villanueva, 2015).

AIFM1h se sintetiza en el citosol como una apoproteína que contiene 613 aminoácidos, una secuencia de localización mitocondrial (MLS) en el extremo N-terminal y dos de localización nuclear (NLS). Tras su translocación al IMS, los primeros 54 aminoácidos del N-terminal se escinden proteolíticamente (AIFM1h $_{\Delta 1-54}$). En el IMS la hélice N-terminal de AIFM1h $_{\Delta 1-54}$ se inserta en la membrana mitocondrial interna y el resto de la proteína se pliega en el IMS adoptando una conformación en tres dominios.

Cuando la célula recibe un estímulo pro-apoptótico, se produce una ruptura proteolítica de AIFM1h $_{\Delta 1-54}$ en el residuo 101, dando lugar a la forma proapoptogénica, AIFM1h $_{\Delta 1-101}$. Esta se libera al citosol y posteriormente se transloca al núcleo, donde inducirá apoptosis mediante condensación de la cromatina y fragmentación del ADN a gran escala.

La unión de NADH con la AIFM1h produce la reducción de su cofactor FAD y la dimerización de la proteína, mientras que en su estado oxidado se encuentra como un monómero. Este hecho puede ser de gran relevancia en la mitocondria, en la cual hay una mezcla de monómero y dímero (ya que en la mitocondria hay una alta concentración de NADH) (Ferreira et al., 2014).

Por el momento se han propuesto dos funciones para esta proteína, una a nivel mitocondrial como oxidorreductasa y otra como factor de inducción de apoptosis a nivel nuclear, que dependen de su localización celular:

- La función mitocondrial parece estar relacionada con la actividad NADH oxidorreductasa de la proteína: cuando AIFM1h se encuentra localizada en el espacio intermembrana y anclada a la membrana interna acepta electrones del NADH a través de un mecanismo de transferencia de hidruro. No está claro quién es su aceptor de electrones final, pero no proporciona electrones a los complejos respiratorios mitocondriales, aunque parece jugar un papel clave en su biogénesis (Sevrioukova, 2011). Recientemente se ha implicado esta función con el importe a la mitocondria de diversas subunidades del complejo I y factores para el ensamblaje del complejo IV, a través de proteínas como MIA40 y CHCHD4 (Villanueva, 2015).
- La función apoptótica: cuando la célula recibe un estímulo apoptótico, se produce un corte proteolítico en AIFM1h que libera su forma soluble, AIFM1h $_{\Delta 1-101}$ al citosol, favoreciendo su translocación al núcleo, donde induce muerte celular programada independiente de caspasas (Sevrioukova, 2011; Medina et al., 2014; Villanueva, 2015).

2. Estructura de la AIFMh.

La forma soluble de la AIFMh se pliega adoptando una conformación en tres dominios; dominio de unión a FAD, dominio de unión a NAD y dominio de regulación de la apoptosis. El dominio de unión a FAD se encuentra dividido en dos partes, del aminoácido 127 al 300 y del 390 al 480, el dominio de unión a NAD entre los residuos 300-390 y el dominio de unión de la apoptosis del 490 al 613 (Sevrioukova, 2011; Ferreira et al., 2014; Villanueva, 2015). Los dominios de unión a FAD y NAD presentan una topología de Rossmann (consiste en una serie de láminas- β y hélices- α intercaladas que son las encargadas de interactuar con los cofactores) (Hanukoglu, 2015). Estas regiones se sitúan en los primeros aminoácidos de la cadena mientras que en el fragmento C-terminal se sitúa el dominio de regulación de la apoptosis (Sevrioukova, 2011).

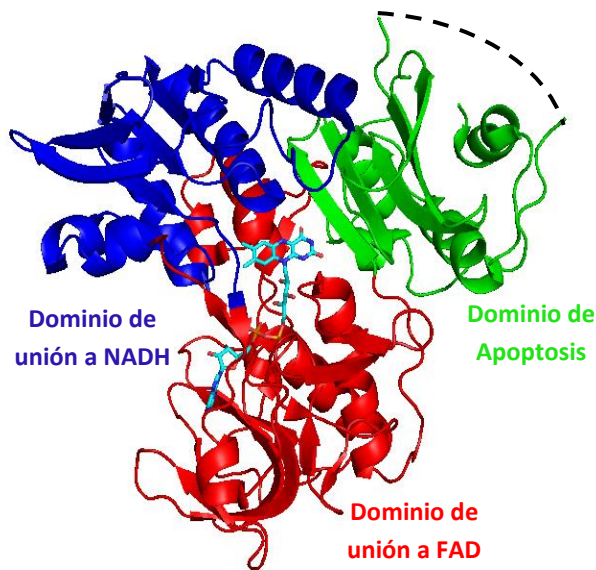


Figura 1: Estructura terciaria de la AIFM1h (PDB 4BV6), en la que se representan sus diferentes dominios: unión a FAD (rojo), unión a NADH (azul) y regulación de la apoptosis (verde).

En líneas negras discontinuas se muestra una región de la que no se obtiene información estructural debido a su flexibilidad, aminoácidos 546 hasta 558.

AIFM1h ha sido cristalizada tanto en el estado oxidado (correspondiente al fichero con código PDB 4BV6), como en el estado reducido por NADH (correspondiente al fichero 4BUR). Esto ha permitido analizar las diferencias estructurales inducidas por la unión del coenzima NADH, por el proceso redox y por la dimerización de la proteína (Ferreira et al., 2014).

Antes de centrarse en la comparación de la cadena proteica entre estas dos estructuras, ya que hay cambios conformacionales evidentes y de relativa importancia, hay que puntualizar que la estructura cristalina del estado oxidado solo posee el FAD, mientras que en la estructura del estado reducido se pueden observar además 2 moléculas de NADH por molécula de AIFM1h. El primer NADH se une a un sitio de unión preformado tanto en la estructura reducida como en la estructura oxidada y se propone que es relevante en la función redox de la proteína. Por el contrario el segundo NADH se une en una posición que se encuentra cerrada en la estructura en el estado oxidado y por tanto se cree que la unión con el primer NADH y la subsecuente reducción del cofactor flavínico, determinarán un cambio de conformación en la proteína que permita la unión de este segundo NADH.

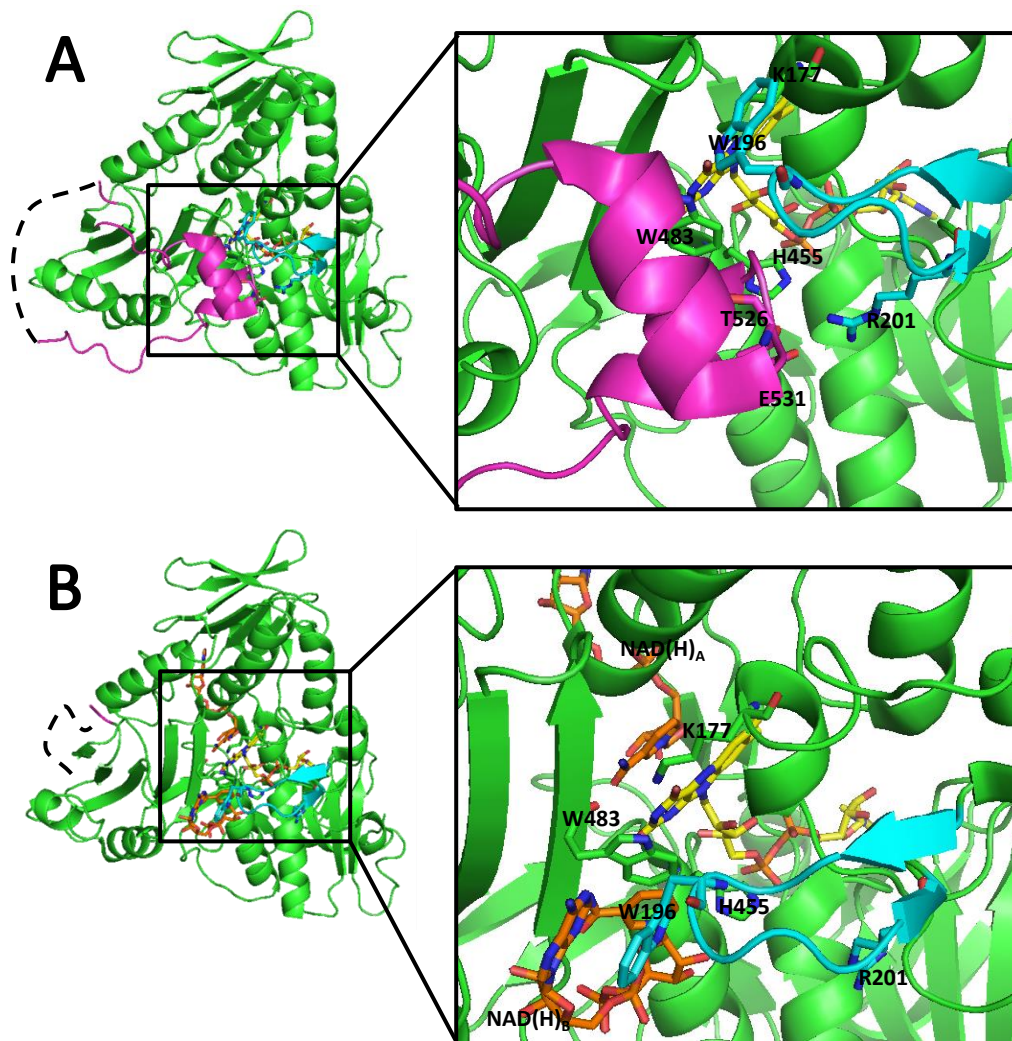


Figura 2: Cambios conformacionales en la AIFM1h consecuencia de la unión de NADH y la reducción de la flavina. (A) AIFM1h Δ_{1-127} en estado oxidado y (B) la AIFM1h Δ_{1-127} en estado reducido. El plegamiento terciario de la proteína se muestra en verde, con los residuos 510-560 de (A) en rosa, y que no se detectan en la estructura reducida como consecuencia de su desorden. Los residuos 192-202 se muestran en azul claro y se puede apreciar cómo se modifica su posición de (A) a (B). FADH y las dos moléculas de NADH se muestran como varillas con los carbonos en amarillo y naranja, respectivamente. Las regiones desordenadas se muestran como una serie de líneas discontinuas.

Como se puede apreciar en la figura 2-A, hay dos residuos (Glu531 y Thr526) de la región 510-560 (coloreada en rosa), que interactúan con Arg201 situada en un bucle pinzado por dos láminas- β (parte en azul claro, que contiene los residuos 192-202) (Sevrioukova, 2011; Villanueva, 2015). La imagen también muestra los aminoácidos que interactúan con el anillo de isoaloxacina del FAD, Lys177, His455 y Trp483, y que además contribuyen a estabilizar el coenzima en el interior del centro activo.

La figura 2-B muestra como el centro activo de la proteína se encuentra ocupado por dos moléculas de NADH que flanquean al FAD. La presencia de estas dos moléculas de NADH produce cambios sustanciales en la estructura proteica. Lo que mejor se puede apreciar es la desaparición de la región 510-560 (cambia su conformación de hélice por una conformación desordenada) (Sevrioukova, 2009; Villanueva et al., 2015), que junto a un reducido movimiento de la región 192-202 permiten la introducción del segundo NADH(B) (se cree que este movimiento es producido con la unión del primer NADH(A) a la proteína). Este

movimiento de las dos regiones nombradas anteriormente hace que deje de producirse la interacción entre Glu531 y Thr526 con Arg201. Lo que se mantiene es la interacción de los 3 aminoácidos nombrado anteriormente (Lys177, His455 y Trp483) con el FAD, que además al encontrarse tan cerca del NADH también interaccionan con él.

Cabe destacar también el cambio conformacional que se produce en el bucle de la región 192-202, en el que en estado oxidado Arg201 interacciona con la región 510-560 y se encuentra completamente desplazada en estado reducido. En cambio Trp196, que no parecía jugar ningún papel importante, en la figura 2-B se puede apreciar como interacciona con el anillo de nicotinamida del segundo NADH, contribuyendo a su estabilización (Sevrioukova, 2009; Villanueva et al., 2015).

Resaltar las regiones con líneas discontinuas que pertenecen a la región 546-558 en 4BV6 y la región 509-558 en 4BUR, que no se observan en la estructura cristalina. Este es un hecho fundamental, ya que a lo largo del trabajo se va a intentar resolver la estructura secundaria de estas regiones (mediante métodos bioinformáticos) debido a su importancia en la función de la proteína in vivo.

3. Interacción de la AIFM1h con otras proteínas.

Se ha descrito también la interacción de AIFM1h, con diversas proteínas, en los diferentes compartimentos celulares en los que se puede localizar (Sevrioukova, 2011; Ferreira et al., 2016)

- En la mitocondria se ha descrito que la forma reducida de AIFM1h forma un complejo con CHCHD4, una proteína implicada en el importe de diversas proteínas al espacio intermembrana de la mitocondria y en su plegamiento oxidativo (Mesecke et al., 2005). Entre estas proteínas se encuentran varias subunidades del complejo I de la cadena respiratoria y varios factores de ensamblaje del complejo IV. Esto conecta la actividad reductasa de AIFM1h con la biogénesis de los complejos de la cadena mitocondrial (Ferreira et al., 2016)
- En el citosol destaca la interacción de AIFM1h con dos proteínas que poseen una función contrapuesta entre ellas. Por un lado se encuentra la ciclofilina A (CypA) (Thali et al., 1994; Candé et al., 2003), cuya interacción con AIFM1h favorece el co-transporte óptimo de ambas proteínas al núcleo, favoreciendo la inducción de la apoptosis. Por su parte la interacción de AIFM1h con HSP70 (chaperona de bajo peso molecular que se une a diferentes proteínas para estabilizarlas, impedir su agregación o ayudarlas en el plegamiento) (Tavaria et al., 1996) induce la supervivencia celular al impedir su transporte al núcleo (Gurbuxani et al., 2003; Sevrioukova, 2011; Ferreira et al., 2016).
- En el núcleo la interacción del complejo AIFM1h: CypA con la histona H2AX (proteína que contribuye a la formación del nucleosoma y por consiguiente a la estructuración del DNA), forma el degradosoma, el cual uniría DNA produciendo condensación de la cromatina y la subsecuente degradación del DNA (Ferreira et al., 2016).

III. OBJETIVO DEL TRABAJO:

1. Predecir mediante la utilización de diversas herramientas informáticas la posible organización espacial de las regiones de la AIFM1h que no se observan en las estructuras cristalinas de AIFM1h_{ox} y AIFM1h_{red}:NAD⁺, para construir los modelos tridimensionales de ambas formas que contengan los aminoácidos no observados en las estructuras cristalográficas.
2. Evaluar el acoplamiento del coenzima en los modelos construidos.
3. Evaluar el acoplamiento de AIFM1h con otras proteínas con las que interacciona *in vivo*.
4. Diseñar una práctica de bioinformática consistente en el modelado molecular de las regiones que no se observan en las estructuras cristalográficas y la realización de un docking con algunas proteínas con las que interacciona *in vivo*.

IV. MATERIALES Y MÉTODOS:

1. Ficheros de proteínas.

Todos los ficheros de coordenadas de estructuras de proteínas han sido obtenidos del Protein Data Bank (PDB) <http://www.rcsb.org/pdb/home/home.do>:

- Ficheros de las coordenadas de AIFM1h:
 - 4BV6: estructura cristalográfica de AIFM1h en estado oxidado, contiene el cofactor FAD. Este fichero carece de la información estructural de 13 aminoácidos correspondientes a la región desde el aminoácido 546 hasta el aminoácido 558, ambos inclusive, la cual se considera flexible.
 - 4BUR: estructura cristalográfica de AIFM1h en estado reducido, de su cofactor flavínico, FAD_{rd}, y en complejo con dos moléculas del coenzima NAD/H. este fichero carece de la información estructural de 49 aminoácidos correspondientes a la región desde el aminoácido 509 hasta el 558, ambos incluidos.
 - 4LII: estructura cristalográfica de la AIFM1h en estado oxidado, similar al fichero 4BV6. Se utiliza porque el Swiss-model no identifica en su base de ficheros 4BV6.
- Ficheros de coordenadas de otras proteínas humanas:
 - 2CPL: fichero PDB de la estructura de la CypA.
 - 2K3J: fichero PDB de la estructura del CHCHD4.
 - 5FPD, 5FPE, 5FPM y 5FPN: ficheros PDB de la estructura de HSP70. Al visualizar su estructura terciaria en el PDB no presentan diferencias notables en la estructura terciaria por lo que en principio da igual cual utilizar. Se decide utilizar la 5FPM porque muestra la posición de un mayor número de residuos.
 - 1P3A y 1KX5: ficheros PDB de la estructura de H2AX unido a nucleosomas. Por tanto habrá que eliminar previamente del fichero las histonas que no sean la H2A y el DNA, ya que no son de interés para este estudio. Se decide utilizar el archivo 1KX5 debido a que está la histona completamente cristalizada y en 1P3A faltan unos fragmentos del N-terminal y del C-terminal.

2. Programas informáticos:

- PyMOL (<https://www.pymol.org/>): programa de visualización molecular de código abierto y auspiciado por los usuarios. Creado por Waren Lyford Delano y comercializado por Delano Scientific LLC, una compañía dedicada a la creación de herramientas computacionales accesibles universalmente para las comunidades científicas y educacionales. PyMOL es apropiado para producir imágenes 3D de alta calidad de moléculas pequeñas y de macromoléculas biológicas, además de permitir la determinación de ciertos parámetros, la comparación entre moléculas y algunas simulaciones básicas. En este trabajo PyMOL se ha utilizado para visualizar y comparar las moléculas objeto de estudio, producir imágenes y hacer algún modelado básico.
- Discovery Studio (<http://accelrys.com/products/collaborative-science/biovia-discovery-studio/>): software de pago que permite la simulación tanto de moléculas pequeñas como de complejos macromoleculares. Este producto posee un fuerte programa de colaboración académica y posee numerosas herramientas y algoritmos desarrollados por la comunidad científica que permiten al software realizar una serie de procesos y tareas muy diversas. En este estudio se utilizará su opción de modelado molecular para modelar la región 510-560 que aparece como dos hélices- α en 4BUR, después de la utilización del servidor Swiss-model.

3. Servidores web:

- Alineamientos de secuencia: los alineamientos de secuencia que se presentan se han realizado utilizando BLAST (Basic Local Alignment Search Tool) <http://blast.ncbi.nlm.nih.gov/Blast.cgi>. Los algoritmos de BLAST permiten la comparación de una secuencia problema (también denominada query) con un conjunto de secuencias almacenadas en una base de datos que puede elegir el usuario (Altschul, 1997; Schwede et al., 2014).
- Predictores de estructura secundaria. Estos servidores web predicen la estructura secundaria de una secuencia basándose en la secuencia de aminoácidos y en diversos algoritmos de predicción de estructura secundaria. En este caso se han utilizado varios servidores web: PHDsec (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_phd.html), Jpred4 (<http://www.compbio.dundee.ac.uk/jpred/>), Pfam (<http://pfam.xfam.org/>), PsiPred (<http://bioinf.cs.ucl.ac.uk/psipred/>) y Sam-t02 (<https://compbio.soe.ucsc.edu/HMM-apps/T02-query.html>). Entre ellos destaca PsiPred, que es uno de los muchos servidores que ofrece el Departamento de Biología Computacional del University College London (UCL) <http://bioinf.cs.ucl.ac.uk/introduction/>. Posee una red de trabajo neuronal propia que utiliza para procesar y analizar los datos obtenidos mediante un proceso de PSIBLAST (Buchan et al., 2013). Su página web también permite utilizar otros servidores como DISOPRED, utilizado para predecir regiones desordenadas de una proteína.
- Programas de modelado por homología. Los algoritmos que utilizan estos servidores permiten predecir una estructura terciaria a partir de una secuencia por modelado sobre una estructura homóloga con un importante grado de identidad (Mount, 2007). En este trabajo se ha utilizado Swiss-model (<http://swissmodel.expasy.org/interactive>), una

herramienta que además de utilizar el modelado por homología, incorpora la información de otros servidores (BLAST, PsiPred...) y bases de datos (PDB, Uniprot...), para la producción de modelos con mayor fiabilidad. A destacar su biblioteca de modelos (Swiss-Model Template Library, SMTL) que está conectada con el PDB y le permite disponer de cualquier información y actualización que se realiza en esta base de datos. También utiliza un sistema independiente denominado proyecto CAMEO (Continuous Automated Model EvaluatiOn, <http://cameo3d.org>) que continuamente realiza evaluaciones sobre la calidad de las predicciones siguiendo los criterios establecidos por la comunidad científica sobre la predicción de estructuras (Biasini et al., 2014).

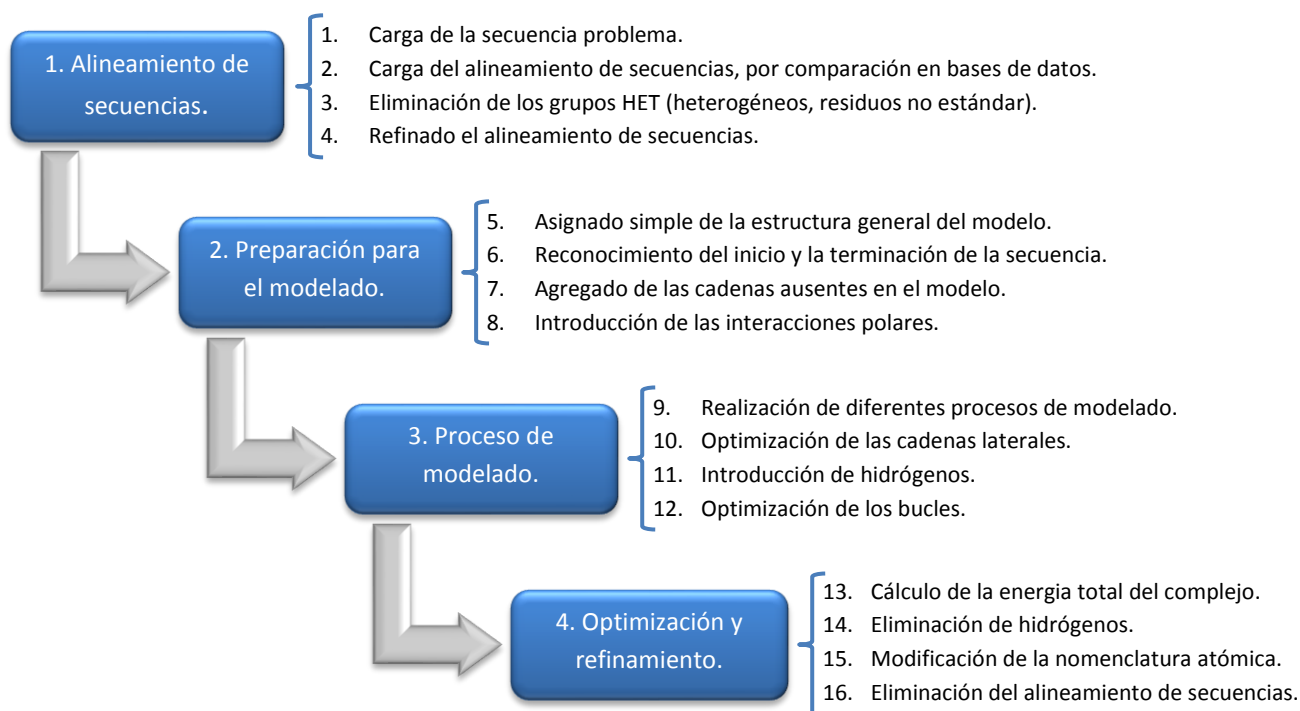


Figura 3: Proceso que realiza el programa Swiss-model para predecir la estructura terciaria de una secuencia a partir de los datos que ha extraído del PDB.

- Servidores de docking proteína-proteína. Es un método que consiste en la conformación más probable que obtiene una proteína al estar unida con otra molécula (en este caso es unión proteína proteína), a fin de formar un complejo estable. En este caso se va a utilizar el servidor ClusPro (<https://cluspro.bu.edu/home.php>) (Comeau et al., 2003).

V. RESULTADOS Y DISCUSIÓN:

Las 2 estructuras de la proteína AIFM1h con las que se van a trabajar en este estudio presentan estas características:

1. AIFM1_{Δ1-101} con el cofactor FAD en estado oxidado y código PDB 4BV6, cuyo fichero PDB no contiene información estructural de 13 aminoácidos de estructura terciaria (desde el 546 hasta el 558 incluidos).
2. AIFM1_{Δ1-101rd}: NADH, con el cofactor FAD en estado reducido, dos moléculas de NADH por protómero y código PDB 4BUR, y cuyo fichero PDB no contiene información estructural de 49 aminoácidos de estructura terciaria (desde el 509 hasta el 558 incluidos).

A efectos prácticos de comparación con otras proteínas ambas corresponden a la misma secuencia de aminoácidos (Anexo 1), y por tanto ya que 4BV6 contiene un mayor trozo de secuencia de aminoácidos, muchos de los estudios que se detallan a continuación se realizarán con 4BV6 y posteriormente, cuando así se considere oportuno, se recalcularán con 4BUR.

1. Identificación de estructura secundaria y posibles conformaciones para los fragmentos que no se observan en las estructuras cristalinas.

En un inicio se utilizó esta aproximación para intentar obtener una estructura secundaria del fragmento a través de otras proteínas que posean una región parecida.

Se selecciona una secuencia de 71 aminoácidos (del 499 al 570) para la realización del BLAST. Se utilizan estos 71 aminoácidos, porque contienen la región presumiblemente desordenada y varios aminoácidos a cada lado, con objeto de que el programa posea suficiente información para obtener una alta homología entre la proteína problema y los posibles moldes. El servidor web utilizado en este proceso es: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGETYPE=BlastSearch&LINK_LOC=blasthome. Este servidor ofrece dos opciones, que proporcionan los siguientes resultados:

- A) **Búsqueda en todas las bases de datos de proteínas existentes.** Se identifican un gran número de proteínas con una elevada homología (superior al 85%), que resultan ser todas AIFs de diferentes especies. En ningún caso se encuentra la información deseada, por lo tanto se desecha la información aportada por esta búsqueda.
- B) **Búsqueda de secuencias homólogas con estructuras depositadas en el PDB.** Esta búsqueda proporciona unos 30 resultados (un número muy reducido en comparación a la opción A), pero todas las proteínas que presentan una homología elevada son AIFM1, de ratón o humanas. Todas ellas carecen en su estructura del fragmento objeto de estudio y por lo tanto no pueden ser de utilidad para su modelado. Las proteínas identificadas en esta búsqueda con menor identidad de secuencia (50% más o menos), son fragmentos de proteínas que en ningún caso contienen en su estructura una región equivalente al fragmento objeto de estudio.

A la vista de estos datos se obtiene la primera conclusión del estudio, el PDB no contiene ninguna estructura que pueda servir como modelo para predecir ni la estructura secundaria y ni la conformación de la región 499-570.

Por tantos se decide utilizar como alternativa predictores de estructura secundaria. Se introduce la secuencia: STPAVPQAPVQGE, que corresponde a los 13 aminoácidos que faltan en el fichero PDB 4BV6. En algunas ocasiones se debe utilizar un mayor número de aminoácidos, porque cada predictor necesita un número mínimo de aminoácidos que supera los 13 aminoácidos de interés. En cada caso se indicará cual es el número seleccionado para la predicción.

En general los resultados obtenidos son similares a los del BLAST, porque compara con otras proteínas depositadas en las bases de datos y selecciona AIFs. A destacar el Jpred2 y el PsiPred, que son programas con redes de trabajo neuronales propias y que aportan resultados más definidos.

Jpred 2 (secuencia utilizada: ITIPPSTPAVPQAPVQGEDYGKGVIFYLRD): este servidor busca en

las bases de datos existentes y además realiza su propia predicción. Lo primero que realiza es una comparación para identificar la homología que hay entre las diferentes proteínas AIF (figura 4) y después ejecuta los algoritmos propios del programa para poder predecir la estructura secundaria (figura 5).

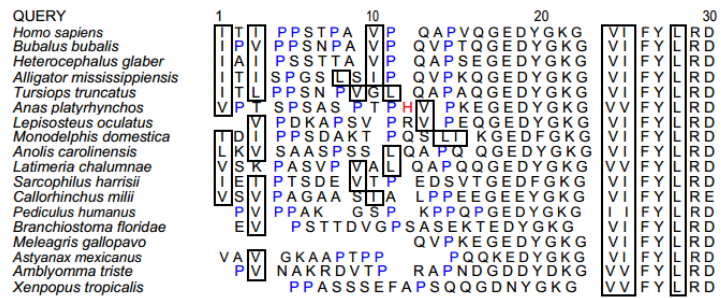


Figura 4: Comparación de las secuencias de AIFM de diferentes organismos depositados en la base de datos de UniProt. Los aminoácidos en azul indican completa identidad, los rojos conservación de las propiedades químicas en comparación con otros en la misma posición y los recuadrados son los utilizados como base para realizar la predicción.

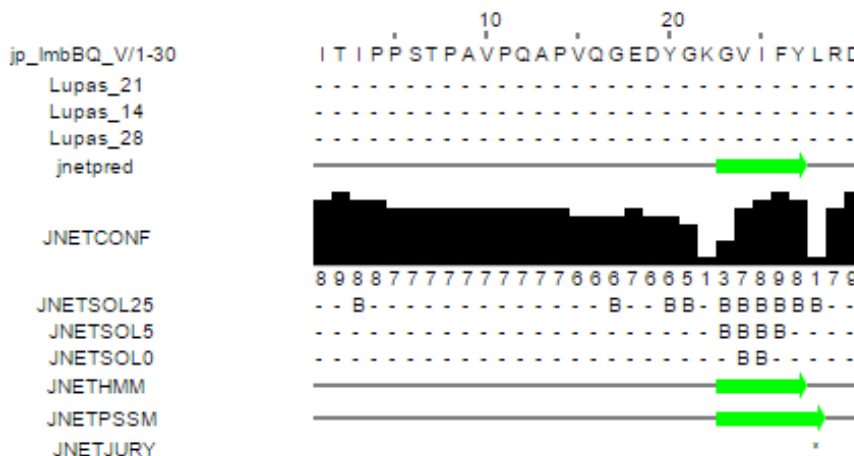


Figura 5: En la parte superior se presenta el resultado del algoritmo propio del programa (Lupas), el cual no ha podido resolver ningún tipo de estructura secundaria en nuestra región de interés. En la parte inferior aparece el Jnet, que utiliza un sistema de BLAST, el cual la región GKGVIIFY se pliega en una lámina-β. Esta información la obtiene de la comparación de AIFMs depositadas en el UniProt y que concuerda con la conformación que adopta esa zona en la proteína en estado reducido (tanto en 4BV6 como en 4BUR se puede apreciar).

PsiPred: como este programa ofrece muchas opciones referentes a estructura secundaria y terciaria, se realizan diferentes predicciones, para poder comparar los resultados. Una vez obtenidos los datos se realiza una comparación entre estos y lo que se sabe de la estructura secundaria de la AIFM1.

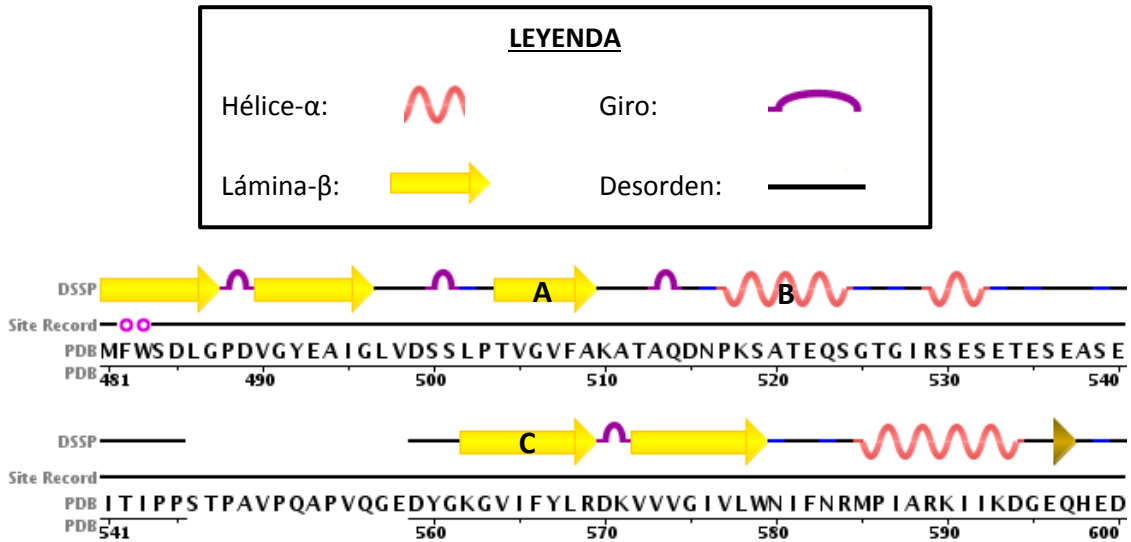


Figura 6: Estructura secundaria del fichero 4BV6 (depositado en el PDB), en la que se puede apreciar como hay una larga zona carente de estructura secundaria definida (que contiene los aminoácidos que no se observan en la estructura 4BUR) y una zona vacía que se corresponde a los 13 aminoácidos que no se observan en el fichero 4BV6.

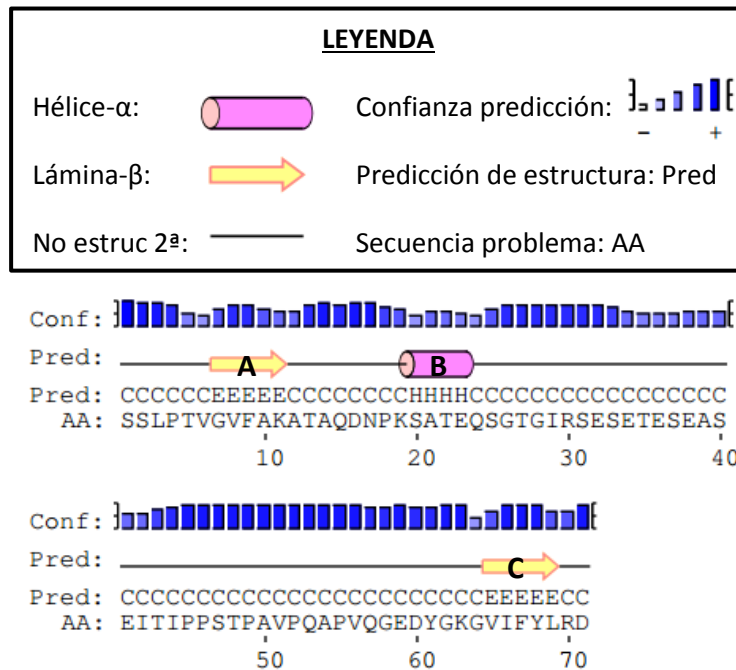


Figura 7: Predicción del PsiPred de la estructura secundaria de la región 499-570 de AIFM1h (71 aminoácidos) que contiene las regiones no observadas en 4BV6 y 4BUR.

Se introduce la secuencia de 71 aminoácidos y se filtran tanto las regiones de poca complejidad, como las hélices súper-enrolladas (figura 7). Como se puede apreciar hay una secuencia de gran longitud entre los aminoácidos 525 y 560, que corresponde a una región sin estructura definida y en la cual se incluyen tanto los 13 aminoácidos que no se observan en 4BV6 como los 49 que no se observan en 4BUR. El predictor también identifica un par de láminas- β y una hélice- α , presentes en el fichero de 4BV6 (figura 6). Este último hecho corrobora la fiabilidad de la predicción realizada.

También se realiza una predicción de estructura secundaria con sólo 18 aminoácidos PPSTPAVPQAPV QGEDYG (subrayados los 13 aminoácidos de interés), para corroborar que para esta región se obtiene el mismo resultado que con la predicción de 71 aminoácidos.

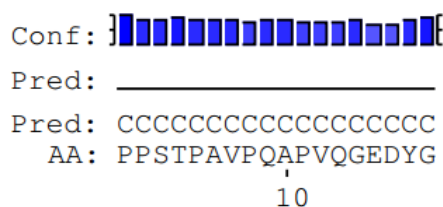


Figura 8: Estructura secundaria de los 18 aminoácidos, predicha por el PsiPred (misma leyenda que en figura 7).

Por último se realiza una predicción de estructuras desordenadas con los 71 aminoácidos (499-570) para ver si al no presentar ningún tipo de estructura secundaria este fragmento, se debe a una estructura desordenada (figura 9).

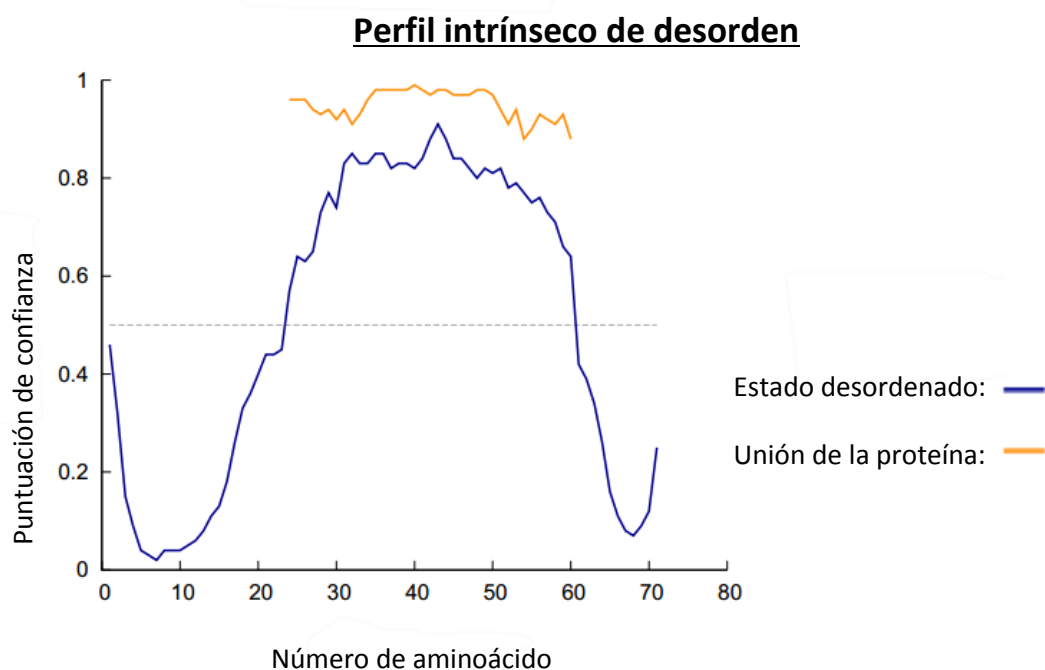


Figura 9: Entre las posiciones 25 y 60 (justo la posición de los 49 y 13 aminoácidos no vivibles en las estructuras cristalinas), el predictor DISOPred asigna un valor elevado a la función de confianza, sugiriendo que esta debe ser una región desordenada, que no presenta una estructura secundaria definida.

El conjunto de estos resultados indica que el fragmento 524-560 forma parte de una región que tiende a encontrarse desordenada y presenta alta flexibilidad, una de las razones por la que probablemente no se pueda observar en la estructura cristalina.

2. Producción de un fichero de coordenadas para AIFM1h que incluya el fragmento no observable en las estructuras cristalinas resueltas.

2.1. Modelado manual de la región desordenada.

Debido a que el fragmento ausente en 4BV6 es una secuencia corta, solo 13 aminoácidos, se planteó como primera opción modelar esos 13 residuos utilizando un programa de visualización de estructuras que permite cierto modelado geométrico (PyMOL). El protocolo seguido es:

- Producción del modelo manual de la cadena de 13-1 aminoácidos. En este caso se elige como aminoácido extra Pro545, que es el último que se visualiza en la estructura antes de la región desordenada (región 546-558)
- Superposición con la opción "*pair fitting*" de dicha Pro545 del fragmento generado sobre la de 4BV6.
- Una vez que se encuentran las prolina apareadas, se elimina la prolina del fragmento y se sustituye el enlace con la opción "*create bond*" que ofrece PyMOL. De esta forma se crea el enlace peptídico entre Pro545 de la proteína y Ser546 del fragmento.
- A continuación se crea el enlace entre el otro extremo del fragmento con el otro aminoácido de la cadena (558-559), creando el enlace Glu558-Asp559. Al estar tan separados ambos aminoácidos esta acción genera una distancia de enlace "imposible" (figura 10).
- Se utiliza la opción "*sculpting*", que permite mover los ángulos y distancias de enlace con objeto de obtener estructuras geoméricamente correctas.

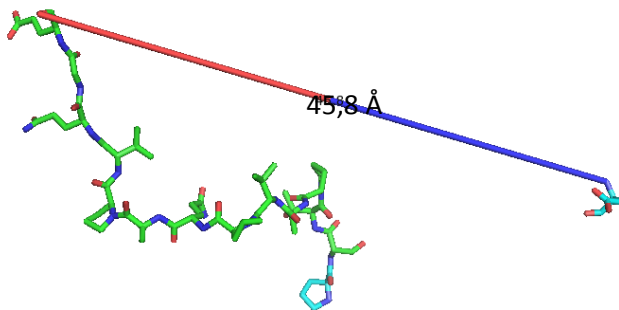


Figura 10: Modelo producido con PyMOL de la cadena polipeptídica de los 13 aminoácidos de interés (coloreada en CPK y con los carbonos en verde). Se muestra el acoplamiento a Pro545, así como el acoplamiento con un "enlace" de 45,8 Å entre Glu558-Asp559. Pro545 y Asp559 correspondientes a 4BV6 se muestran con los carbonos coloreados en azul claro.

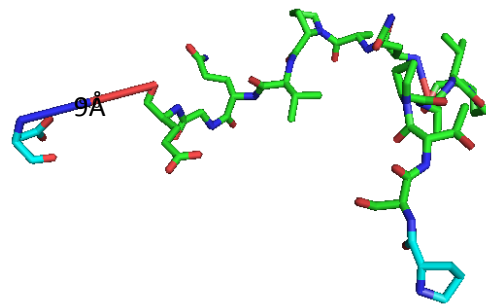


Figura 11: Cadena polipeptídica de los 13 aminoácidos de interés, tras ser sometido el modelo a la opción "*sculpting*" de PyMOL, lo que ha producido el acercamiento del enlace Glu554-Asp559 hasta los 9 Å.

La comparación de la figura 9 y 10 permite apreciar que aunque la utilización del "*sculpting*" aproxima considerablemente las posiciones 558 y 559. En la mejor estructura obtenida esta distancia de enlace sigue siendo demasiado elevada, aunque dicho modelo contiene ya todos los átomos de AIFM1h y podría mejorarse sometiendo este fichero a una simulación de dinámica molecular.

Este resultado indica que, aunque PyMOL es una herramienta buena para la visualización de proteínas y permite cierto grado de modelado, la opción “*sculpting*” se queda limitada a la hora de modelar la zona desordenada de 4BV6, por lo que tampoco resulta factible utilizarlo para 4BUR (ya que este fichero posee 49 aminoácidos que no se observan en la estructura cristalina).

2.2. Modelado por homología de la región desordenada de 4BV6 y 4BUR.

Como aproximación complementaria se utiliza el modelado por homología, usando el servidor Swiss-model, que permite obtener un modelo de estructura terciaria completa de cualquier proteína a partir de un molde, siendo esta una estructura o modelo teórico.

El buscador de Swiss-model identifica la estructura de AIFM1h contenida en el PDB 4LII (como ya se ha explicado anteriormente se utiliza este fichero en lugar de 4BV6, ya que Swiss-model no lo reconoce) como un buen molde para modelar AIFM1h oxidada, a pesar de no poseer tampoco los 13 residuos desordenados.

Una vez obtenido el modelo utilizando como molde 4LII se comprueba si se parece a la estructura depositada en el fichero 4BV6 (figura 12) ya que si fuesen muy diferentes no se podría utilizar ya que esto podría falsear los resultados posteriores. Como se puede apreciar, las imágenes se superponen con bastante similitud, (RMSD = 0.161 que lo emite PyMOL cuando se realiza el alineamiento de secuencias y que cuanto más bajo es este número mejor superponen ambas estructuras). Las únicas diferencias que se pueden destacar son las resultantes de la superposición de algunas estructuras de lámina- β , pero se puede deber a que el modelado finaliza con una etapa de dinámica y además a que el fichero de coordenadas ha perdido asignaciones obligatorias introducidas para el PDB cristalográfico. También podría deberse a que son ficheros PDB diferentes, pero esto es menos probable (se podrá confirmar esta hipótesis cuando se realice a comparación de la 4BUR).

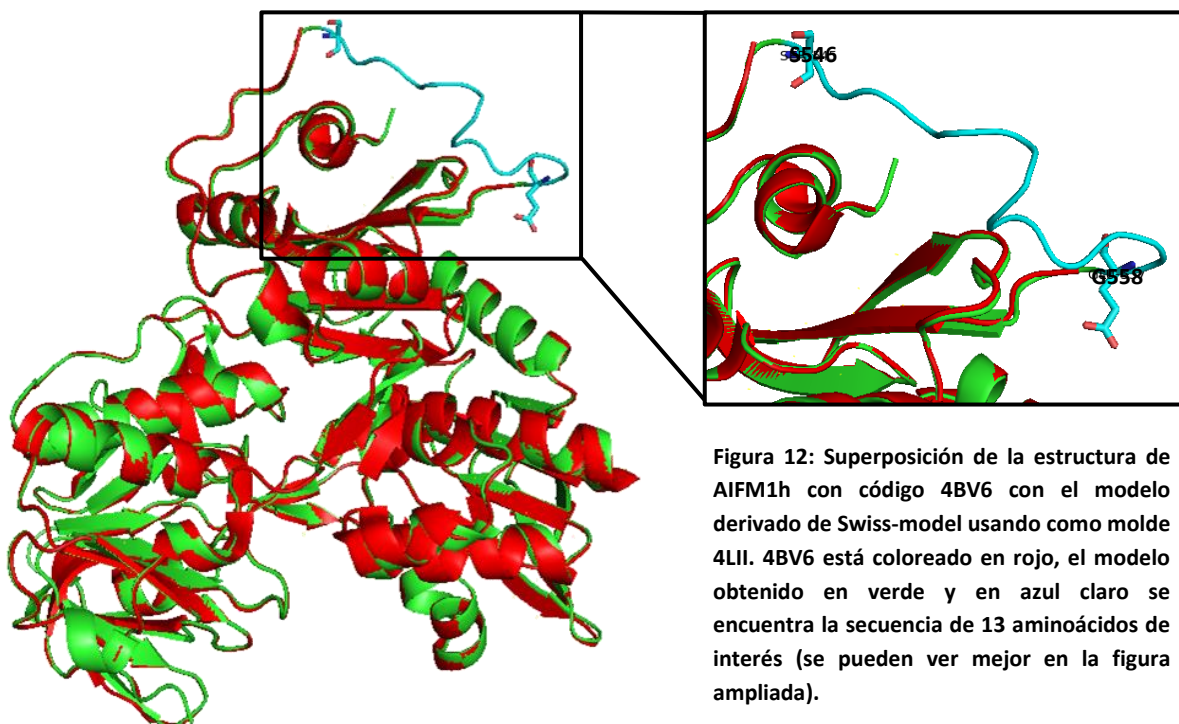


Figura 12: Superposición de la estructura de AIFM1h con código 4BV6 con el modelo derivado de Swiss-model usando como molde 4LII. 4BV6 está coloreado en rojo, el modelo obtenido en verde y en azul claro se encuentra la secuencia de 13 aminoácidos de interés (se pueden ver mejor en la figura ampliada).

Una vez obtenido el modelo de AIFM1h en estado oxidado, se procede a la obtención del modelo de la proteína en estado reducido. Como en este caso Swiss-model sí que reconoce el fichero 4BUR, se elige ese ya que es el que se ha estado utilizando a lo largo de todo el estudio. Es preciso recordar que este fichero no dispone de información sobre 49 aminoácidos.

Una vez Swiss-model ha realizado su predicción se procede a comparar el fichero 4BUR original con la predicción. Como se puede observar en la figura 13 la superposición es muy adecuada (RMSD = 0.262) exceptuando los 49 aminoácidos que lógicamente al no estar en el archivo original es imposible que coincidan. También sucede lo mismo que en la comparación anterior, y es que algunas láminas- β no encajan a la perfección, por tanto se descarta que se deba a la diferencia de ficheros en el caso anterior y será por lo comentado anteriormente.

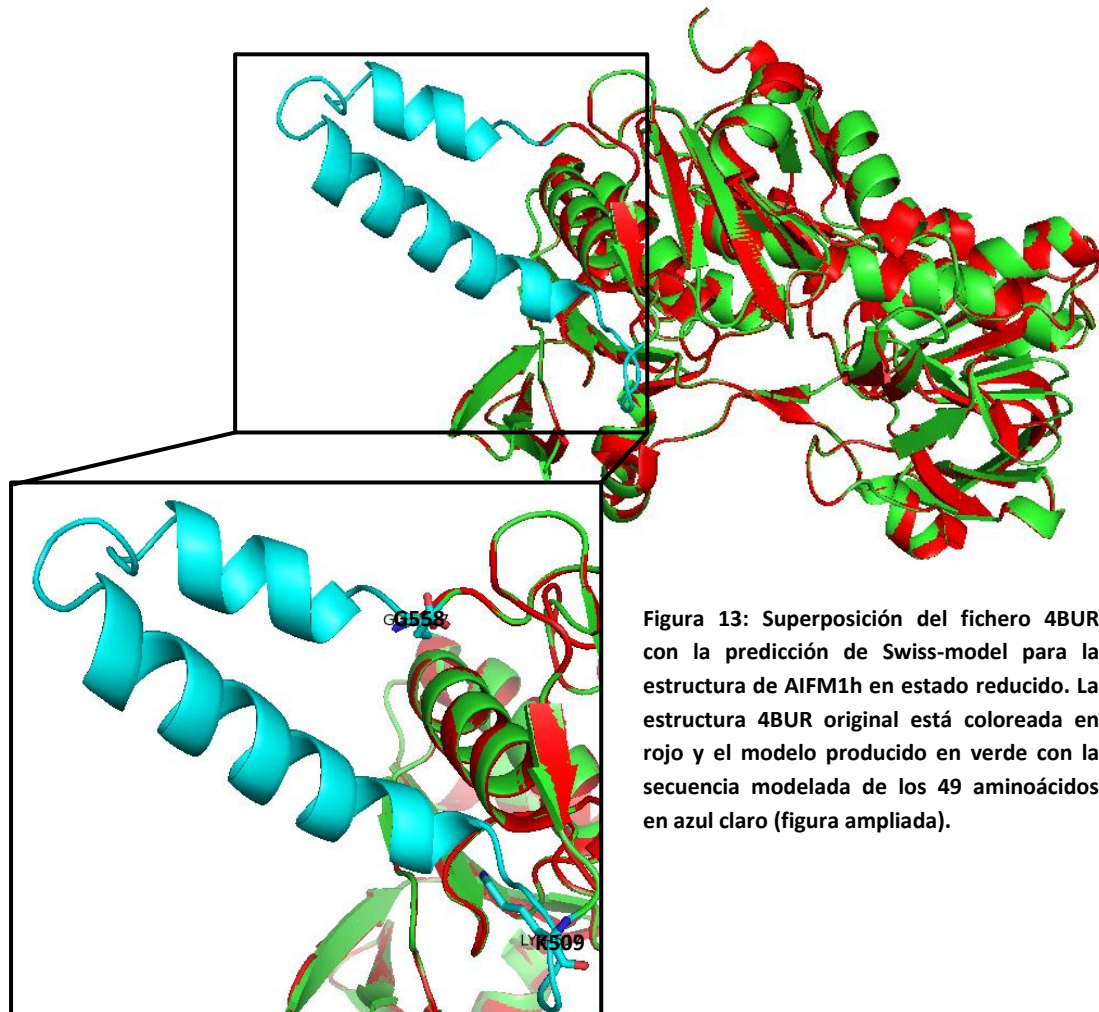


Figura 13: Superposición del fichero 4BUR con la predicción de Swiss-model para la estructura de AIFM1h en estado reducido. La estructura 4BUR original está coloreada en rojo y el modelo producido en verde con la secuencia modelada de los 49 aminoácidos en azul claro (figura ampliada).

El Swiss-model no es un programa especializado en hacer dinámica. Este hecho se puede observar en el caso de la AIFM1h en estado reducido, en la que las 2 hélices- α que ha predicho el programa, se encuentra muy alejada del resto de la proteína, algo que no parece real. Por esa razón ha de minimizarse esta estructura y someterse a una simulación de dinámica molecular. Para ello se va a emplear el programa informático Discovery Studio, que, entre otros, permite realizar este tipo de cálculos.

3. Simulación de dinámica molecular de las estructuras obtenidas en el modelado por homología para encontrar la estructura optima que represente a la proteína.

A través de la interfaz de Discovery Studio, se ha utilizado la herramienta de dinámica molecular NAMD (Not (just) Another Molecular Dynamics program) creada para simulaciones biomoleculares. Emplea el modelo de campos de fuerza CHARMM y diversas plataformas y algoritmos para realizar la simulación.

Tras 7 procesos de dinámica NAMD, siendo el resultado el que se muestra en la figura 14. Se ha simulado durante 70 ps (10 ps cada una de las simulaciones), tiempo demasiado corto como para que se produzca el equilibrio y se pueda obtener una estructura más realista. Habría que realizar más simulaciones de dinámica molecular para poder alcanzar unos resultados más fiables.

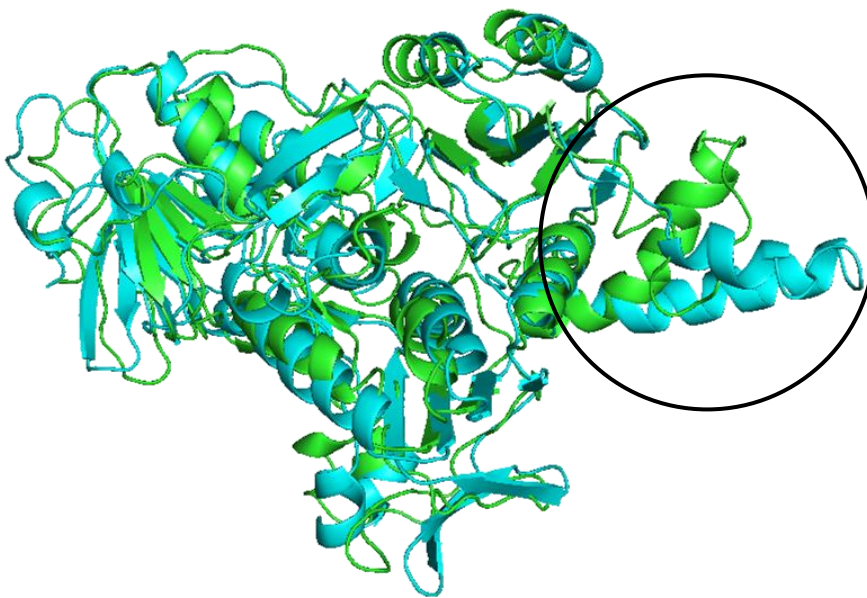


Figura 14: comparación entre el modelo producido por Swiss-model (azul claro) y el modelo después de realizar MD (verde) (70ps). Se puede apreciar como al flexibilizar la molécula varían la región de dos hélices cambia, desapareciendo la hélice más corta en favor de una región desordenada y flexionándose la hélice de mayor longitud hacia el resto de la proteína.

4. Evaluación del acoplamiento de AIFM1h con otras proteínas con las que interacciona *in vivo*.

Se ha utilizado el servidor ClusPro para simular las interacciones proteína-proteína. Se ha seguido siempre el mismo patrón, como AIFM1h es la de proteína de interés será el receptor y las otras cuatro proteínas serán los ligandos, permitiendo de esta forma que AIFMh1 esté fija y el resto de proteínas se mueven alrededor de su superficie.

Después de que el programa realice los 8 trabajos (ya que se estudia la interacción por separado de las 4 proteínas tanto con la forma oxidada como con la forma reducida de AIFM1h), el programa obtiene diferentes tipos de simulaciones: favoreciendo interacciones electrostáticas, interacciones hidrofóbicas, interacciones de Van der Waals más electrostáticas o un simulación balanceada en la que tienen en cuenta todo lo anterior. En este estudio se elige la opción balanceada ya que en el artículo de Camacho et al., 2003, se sugiere utilizarla ya que es la que más fiabilidad posee.

ClusPro genera una gran cantidad de poses para los acoplamientos, pero muchos de ellos son muy parecidos, por eso los agrupa en "clusters" realizando una media. El programa genera así

unos 30 clusters que corresponden a posiciones diferentes, ordenándolas de mayor a menor repetición, según su propio algoritmo. Los primeros clusters son los que contienen poses que se repiten m-as a menudo, aunque no tienen por qué ser los de menor energía.

4.1. Interacción con AIFM1h en estado oxidado

- Interacción entre AIFM1h y CypA: Experimentalmente se ha sugerido que la interacción debe producirse en el núcleo entre las regiones 328-397 de AIFM1h y 55-150 de CypA (Candé et al., 2003). Sin embargo la única estructura (figura 15) que se acerca a esta disposición es la obtenida en el cluster 28 que se corresponde a un número muy alto y además tampoco interaccionan en su totalidad las regiones nombradas anteriormente. Destaca que en esta simulación se produce una interacción de la región desordenada (Gly556-Tyr559.)

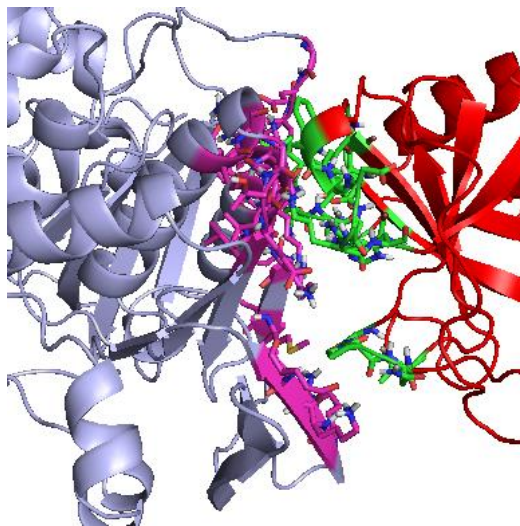


Figura 15: interacción entre AIFM1h y CypA, coloreadas en azul y rojo respectivamente, correspondientes al cluster número 28.

Las regiones que interaccionan de AIFM1h: Ala323-Val330, Ile345-Arg358, Arg387-Ala397 y Gly556-Tyr559; con los carbonos coloreados en rosa.

Las regiones que interaccionan de CypA: Arg55-Phe60, Ala103-Thr107, Trp121-His126 y Ser147-Lys151; con los carbonos coloreados en verde

Además de la caracterización del cluster que mayor similitud posee con los datos experimentales, también se describe la interacción del cluster 0 (figura 16), por ser el que mayor número de veces se repite y uno de los que menor energía posee (Anexo 3). Se puede apreciar cómo se produce la interacción entre la zona C-terminal de AIFM1h y CypA, cuyas regiones de interacción coinciden bastante con el cluster 28.

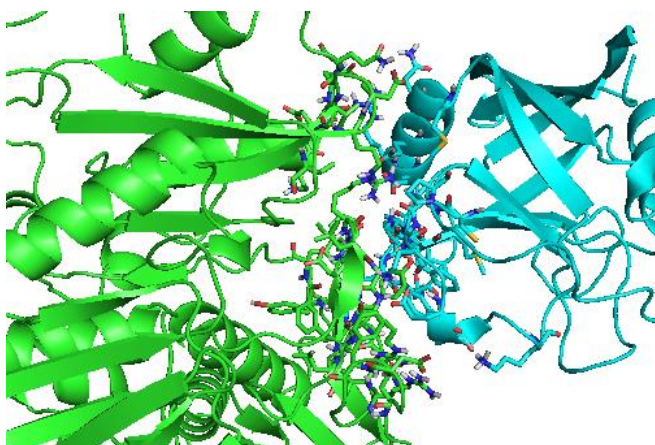


Figura 16: interacción entre AIFM1h y CypA, coloreadas en azul y rojo respectivamente, correspondientes al cluster número 0.

Las regiones que interaccionan de AIFM1h: Asp415-Leu427, Phe442-Arg449, Pro475-Trp477, Arg568-Val573 y Lys592-Glu596 .

Las regiones que interaccionan de CypA: Ile57-Met61, Cys115-Lys125 y Met136-Asn149.

- Interacción entre AIFM1h y HSP70: se conoce que la interacción entre estas dos proteínas se da en el citoplasma, en el que HSP70 se une a AIFM1h en su región 150-268, de esta forma HSP70 bloquea el reconocimiento de la NLS de AIFM1h y evita que migre al núcleo reteniéndola en el medio citoplasmático (Gurbuxani et al., 2003)

Por esta razón se va a buscar la interacción entre estas dos regiones (figura 16), siendo la que mejor se adapta al número 0, pero no es ni de lejos una buena interacción, ya que solo interactúa con una pequeña parte de esta región: Rg151-Pro155 y Lys199-Tyr211; e interactúa en mayor medida con otras zonas.

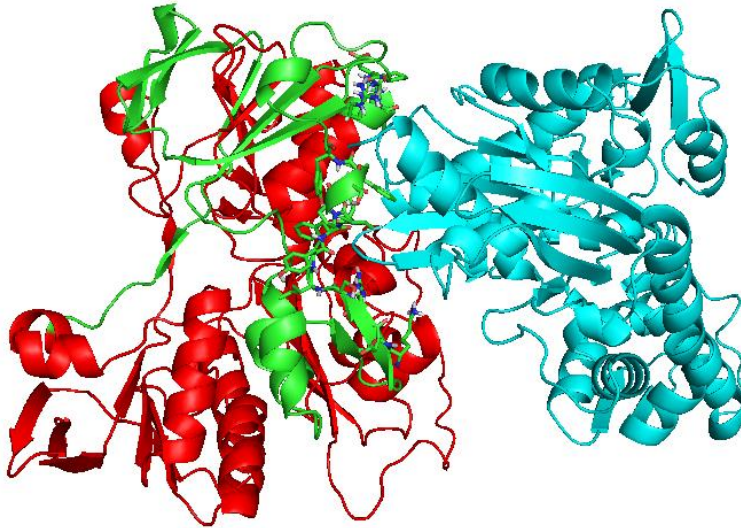


Figura 17: interacción entre AIFM1h y HSP70, coloreadas en azul y rojo respectivamente.

La zona coloreada en verde es la zona en la que teóricamente según la bibliografía interactúa, que se puede apreciar como solo interactúan dos pequeñas regiones de esa zona (carbonos en verde.)

- Interacción entre AIFM1h y CHCHD4: al ser CHCHD4 una proteína que interactúa con AIFM1h en la mitocondria, es posible que no contengamos toda la información para la simulación, ya que solo se posee la estructura terciaria de AIFM1h después de escindirse de la membrana de la mitocondria y abandonarla (que es donde interactúa con esta proteína). Por ello se va a evaluar la interacción que realizan en una zona que posiblemente no sea la real, se ha elegido el cluster 1 porque aunque posee menos repeticiones que el 0 posee unas energías menores (anexo 4).

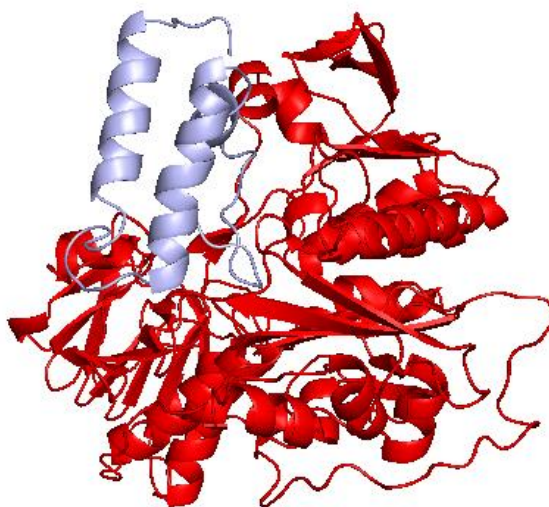


Figura 18: interacción entre AIFM1h y CHCHD4, coloreadas en rojo y azul respectivamente.

La región que interactúa de AIFM1h es la de los residuos 300-550, mientras que de CHCHD4 interactúa prácticamente toda, al ser una proteína pequeña y estar embebida en AIFM1h

- Interacción entre AIFM1h y H2AX: Hasta la fecha no hay información de esta interacción. Se ha elegido el cluster 0 por ser el que menor energía de estabilización posee y el representado por más poses. Además es curioso que posea la misma energía de estabilización tanto siendo la media de las 115 conformaciones como de la de menor energía, eso quiere decir que esas 115 conformaciones serán casi idénticas unas de otras. Resaltar que esta interacción (figura 19) se produce en una región justo al lado de la zona en la que interaccionan AIFM1h y CypA, algo de gran importancia, ya que H2AX se une al complejo AIFM1h-CypA. La zona e unión de H2AX se sitúa cerca de esa otra zona, pero sin llegar a taparla, por lo que podría representar con bastante probabilidad el sitio de unión factible en el que interaccionaría tanto con AIFM1h como con Cyp A.

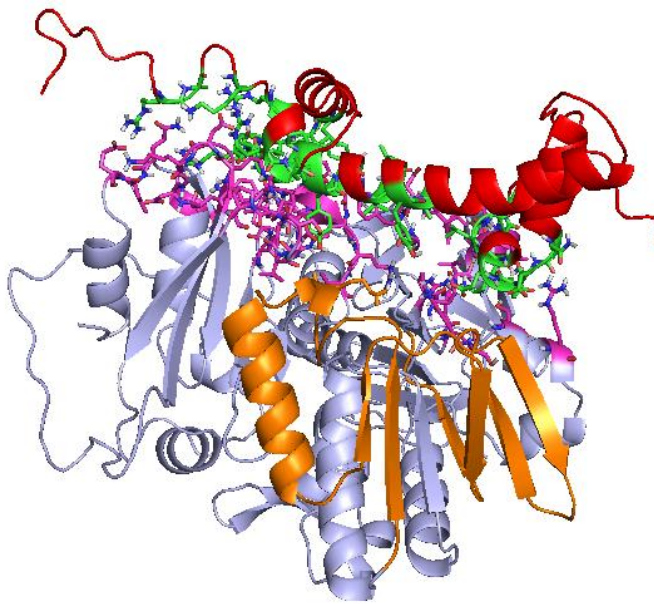


Figura 19: interacción entre AIFM1h y H2AX, coloreadas en azul y rojo respectivamente.

Las regiones que interaccionan de AIFM1h: Arg265-Arg272, Glu401-Gly419, Tyr443-Val452, Trp478, Gly494-Pro502, Leu567-Val573 y Lys592-Glu598; con los carbonos coloreados en rosa.

Las regiones que interaccionan de H2AX: Arg11-Arg21, Tyr39-Leu58 y Asn89-Arg99; con los carbonos coloreados en verde.

Además destaca la región naranja que es la zona de unión de AIFM1h con CypA.

A la vista de estos resultados se generan dudas sobre la fiabilidad de estas simulaciones, ya que en los dos casos que se ha obtenido experimentalmente la zona de unión (CypA y HSP70), apenas coinciden con las simulaciones. Esto puede deberse a varios factores: la fiabilidad del programa (que se presupone buena), a que son justo las dos proteínas más grandes que se han seleccionado para estos estudios (por lo tanto el error puede ser mayor) o a que AIFM1h en estado oxidado no es el mejor receptor, ya que las interacciones que realiza con estas proteínas se realizan inicialmente en el núcleo y citoplasma en el que ya posee sus 2 moléculas de NADH unidas y por tanto posee un cambio conformacional.

Por lo que se va a realizar una simulación tanto con CypA y HSP70 unidas a AIFM1h en estado reducido después de haber sido minimizado por dinámica molecular.

4.2 Interacción con AIFM1h en estado reducido.

En este caso se van a analizar las interacciones con AIFM1h en estado reducido. Se centrará sobre todo en las interacciones con HSP70 como con CypA (en el caso de CHCHD4 y H2AX son muy parecidas tanto en uno como en otro y además no se poseen datos experimentales con los que comparar) porque se produce una modificación relevante de los resultados, haciendo que se parezcan más a los resultados obtenidos en los estudios experimentales. Cuando se estudie la interacción con CypA también se analizará la relación con H2AX.

- Interacción entre AIFM1h y HSP70: es muy sorprendente lo que se ha producido en esta proteína al cambiar la estructura oxidada por la proteína. Mientras que en la estructura oxidada la HSP70 solo poseía el cluster 0 que se acercaba levemente a la región 150-268, en el caso de la proteína reducida se han visualizado 5 clusters correspondientes a los números: 2, 4, 5, 10, 11 (todos muy repetidos y con bajas energías de estabilización). Se ha elegido el cluster 2 (figura 20), por ser el que menores energías posee (anexo 8).

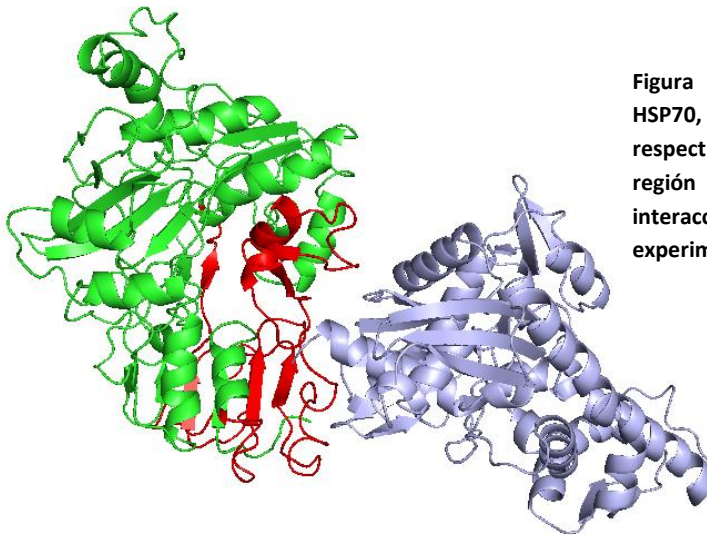


Figura 20: interacción entre AIFM1h y HSP70, coloreadas en verde y azul respectivamente. La región roja es la región 150-268 de AIFM1h que interacciona con HSP70 según los datos experimentales

- Interacción entre AIFM1h y CypA: ocurre algo parecido al caso anterior, no solo el cluster elegido mejora el contacto con la región de AIFM1h asignada experimentalmente (región 328-397), sino que es un cluster menor, en este caso es el cluster 9 (figura 21).

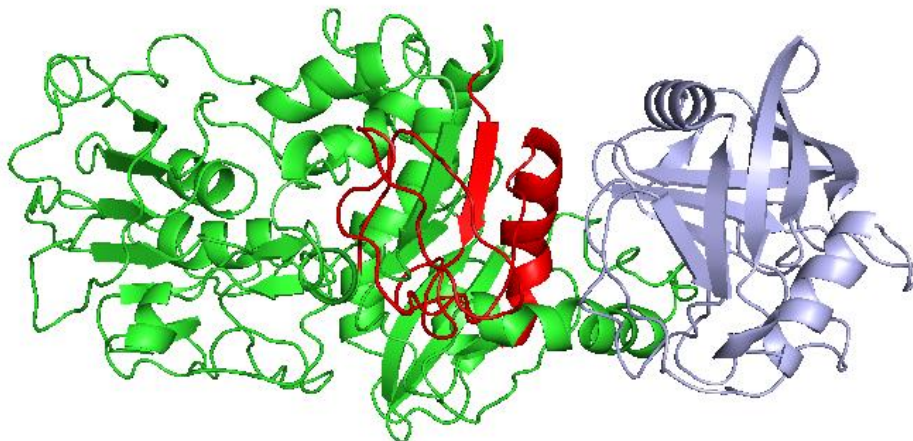


Figura 21: interacción entre AIFM1h y CypA, coloreadas en verde y azul respectivamente. La región roja es la región 328-397 de AIFM1h que interacciona con CypA según los datos experimentales.

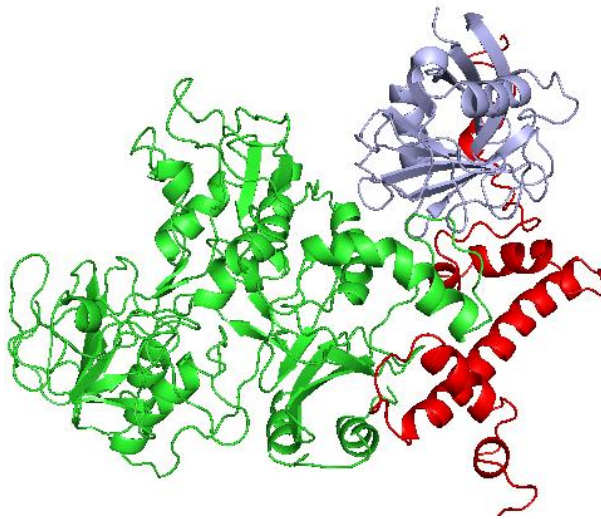


Figura 22: interacción entre AIFM1h, CypA y H2AX en verde, azul y rojo respectivamente.

Que es la interacción que se ha propuesto para la formación del degradosoma y posterior compactación del DNA que conducirá a la apoptosis celular.

Como se puede apreciar en la figura 22, se da la interacción del complejo AIFM1h:CypA con H2AX, en el que las tres proteínas se mantienen en contacto entre sí. Se han seleccionado el cluster 9 de CypA y el cluster 0 de H2AX.

A la vista de estos resultados evaluar de manera positiva la dinámica molecular y posterior simulación de la interacción proteína-proteína de AIFM1h con HSP70 y CypA, que distan de los resultados obtenidos en la simulación con AIFM1h oxidada y que son mucho más parecidos con los resultados obtenidos de manera experimental.

VI. Propuesta de una sesión práctica para la asignatura de Bioinformática del grado de Biotecnología.

Título: Estudio computacional del Factor de Inducción de la Apoptosis Mitocondrial humano (AIFM1h), modelado de sus regiones flexibles y evaluación de las interacciones con otras proteínas.

Para la realización de esta práctica entraremos en contacto con los conceptos básicos sobre el manejo de diversas bases de datos de secuencias y de estructuras de proteínas, programas de visualización y modelado de estructuras y servidores que nos permiten analizar y simular distintas propiedades estructurales de proteínas.

1. Modelado por homología de la región desordenada de 4BV6 y 4BUR.

Las estructuras cristalográficas 4BV6 y 4BUR no contienen las coordenadas de todos los átomos, presentando una región entre las posiciones 546-558 y 509-558, respectivamente, de desorden. En esta práctica utilizaremos Swiss-model (<https://swissmodel.expasy.org/interactive>), para obtener un modelo de estructura terciaria de la proteína que contenga todos sus átomos.

- Introduce una secuencia larga de aminoácidos (unos 50-100) en el “template” del servidor para que la búsqueda de la proteína sea eficiente. Como puedes apreciar el servidor ofrece muchas opciones de ficheros PDB. Elige el 4BUR para el estado reducido de la proteína y el fichero 4LII para el estado oxidado. Una vez el servidor los procese y obtenga un modelo descárgalo.
- Los archivos obtenidos en Swiss-model, no funcionan bien en PyMOL, por el tipo de formato, por lo que haremos con el mismo programa una copia de la estructura completa (seleccionando todos los aminoácidos y coenzimas y utilizando la opción “copy to object”) y lo guardaremos en otro nuevo archivo.
- A continuación, haz una superposición en PyMOL de los ficheros PDB de las proteínas con los de los modelos obtenidos con Swiss-model con objeto de introducir en los modelos los cofactores y coenzimas (Swiss-model realiza el modelado sin ningún coenzima ni cofactor).

2. Evaluación de la interacción de los modelos de AIFMh con otras proteínas.

Descargar los ficheros de coordenadas de algunas de las proteínas que interactúan con la AIFM1h. Descarga los siguientes ficheros PDB: 2CPL (Ciclofilina A), 2K3J (CHCHD4), 5FPM (HSP70) y 1P3A (histona H2AX). En algunos casos deberás modificar estos PDB, debido a que las proteínas de interés están formando complejos con otras que no necesitas.

- Describe a que proteína pertenecen cada uno de los ficheros ¿Qué función posee cada una de estas proteínas? ¿De qué manera y donde podría interactuar con AIFM1h?
- Utiliza el ClusPro, (<https://cluspro.bu.edu/home.php>) para evaluar el acoplamiento e interacción que se produce entre la AIFMh y el resto de proteínas tanto en el estado reducido como en el oxidado. Se te recomienda que te registres en la página de ClusPro y que crees una cuenta que te facilite la descarga y el trabajo.
- Una vez que se hayan realizado los trabajos, descarga los que sean simulaciones balanceadas (primera de las 4 opciones que aparecen en la parte superior).
- Como ves se obtienen unos 30 modelos de interacción de la proteína AIFM1h con las diferentes proteínas, elige uno de los 10 primeros y describe las interacciones que se dan entre AIFM1h y las diferentes proteínas.

Para ayudarte en esta tarea se anexa dos comandos que puedes introducir en PyMOL, los cuales te permiten hacer que se seleccionen los aminoácidos que interactúan en una distancia de 5 armstrongs:

- `select interface_lig0, byres (lig.000.00 near_to 5 of rec).`
- `select interface_rec0, byres (rec near_to 5 of lig.000.00).`

*El comando está diseñado para el caso en el que decidas utilizar el ligando 000.00 en el PyMOL, si decides utilizar otro deberás cambiarlo por el número correspondiente.

VII. Conclusiones

1. El PDB no contiene ninguna estructura depositada que pueda servir como modelo para predecir ni la estructura secundaria y ni la conformación de la región 499-570 de AIFM1h.
2. El fragmento 525-560 de AIFM1h se predice como una región desordenada de alta flexibilidad.
3. PyMOL es un buen visor de macromoléculas, pero su función “*sculpting*” no resulta útil para simular de forma eficiente cambios estructurales grandes en una proteína o modelar regiones.
4. Swiss-model permite la obtención de modelos teóricos completos de la estructura terciaria de proteínas que no se encuentran depositadas en la bibliografía, pero no realiza una dinámica molecular que permita obtener modelos que se aproximen a las posibles conformaciones de la proteína en solución.
5. La utilización de las herramientas de dinámica molecular de la plataforma Discovery studio permiten generar una estructura terciaria que se acerque más a la realidad, sin embargo se requieren más ciclos de dinámica para alcanzar la estructura de equilibrio.
6. Las simulaciones de la interacción proteína-proteína de AIFM1h con HSP70 y CypA son muy diferentes si se realizan con la estructura oxidada y la reducida, siendo en el caso de la interacción con AIFM1h en estado reducido en el que más se asemejan a los datos experimentales.
7. El sistema objeto de estudio y la metodología empleada resultan adecuados para proponer un práctica de bioinformática de modelado y acoplamiento molecular.

VII. Conclusions

1. PDB server doesn't contain any deposited structure which may be used as model to predict no the secondary structure neither region 499-577 conformation of AIFM1h.
2. AIFM1h's 525-560 fragment is predicted as a disordered region with high flexibility.
3. PyMOL is a good macromolecular visualization system, but its “*sculpting*” function isn't useful for large structural changes simulation or regions modelling
4. Swiss-model allows for getting theoretical complete models of the protein tertiary structure which are not found in bibliography, but it doesn't perform molecular dynamics that makes possible to obtain models which are close to the possible conformations of the solvated protein.
5. The use of molecular dynamics tools from Discovery studio let to generate a tertiary structure which approach to the real conformation, however more dynamics cycle are required to reach the balanced structure.
6. Interaction protein-protein simulations of AIFM1h with HSP70 and CypA are very different comparing the oxidized state and reduced state structures, being the simulation in the reduced state the one that agrees with the available experimental data.
7. The biochemical systems and the methodologies used in the present study have resulted adequate to propose an academic project for the subject in bioinformatics bases on molecular modeling and docking.

VIII. BIBLIOGRAFÍA:

- Altschul, S. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17), pp.3389-3402.
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T., Bertoni, M., Bordoli, L. and Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42(W1), pp.W252-W258.
- Buchan, D., Minnici, F., Nugent, T., Bryson, K. and Jones, D. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, 41(W1), pp.W349-W357.
- Candé, C., Vahsen, N., Kouranti, I., Schmitt, E., Daugas, E., Spahr, C., Luban, J., Kroemer, R., Giordanetto, F., Garrido, C., Penninger, J. and Kroemer, G. (2004). AIF and cyclophilin A cooperate in apoptosis-associated chromatinolysis. *Oncogene*, 23(8), pp.1514-1521.
- Comeau, S., Gatchell, D., Vajda, S. and Camacho, C. (2003). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, 20(1), pp.45-50.
- Ferreira, P., Villanueva, R., Martínez-Júlvez, M., Herguedas, B., Marcuello, C., Fernandez-Silva, P., Cabon, L., Hermoso, J., Lostao, A., Susin, S. and Medina, M. (2014). Structural Insights into the Coenzyme Mediated Monomer–Dimer Transition of the Pro-Apoptotic Apoptosis Inducing Factor. *Biochemistry*, 53(25), pp.4204-4215.
- Ferreira, P., Villanueva, R., Martínez-Júlvez, M. and Medina, M. (2016). AIFM1 (Apoptosis Inducing Factor, Mitochondrion-associated, 1). "Encyclopedia of Signaling Molecules, 2nd Edition". ISBN: 978-1-4419-0462-1. Fecha: 2016. S. Choi Ed. Springer
- Gurbuxani, S., Schmitt, E., Cande, C., Parcellier, A., Hammann, A., Daugas, E., Kouranti, I., Spahr, C., Pance, A., Kroemer, G. and Garrido, C. (2003). Heat shock protein 70 binding inhibits the nuclear import of apoptosis-inducing factor. *Oncogene*, 22(43), pp.6669-6678.
- Hanukoglu, I. (2015). Proteopedia: Rossmann fold: A beta-alpha-beta fold at dinucleotide binding sites. *Biochemistry and Molecular Biology Education*, 43(3), pp.206-209.
- Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K. and Herrmann, J. (2005). A Disulfide Relay System in the Intermembrane Space of Mitochondria that Mediates Protein Import. *Cell*, 121(7), pp.1059-1069.
- Mount, D. (2007). Using the Basic Local Alignment Search Tool (BLAST). *Cold Spring Harbor Protocols*, 2007(7), p.pdb.top17.

- Sevrioukova, I. (2009). Redox-Linked Conformational Dynamics in Apoptosis-Inducing Factor. *Journal of Molecular Biology*, 390(5), pp.924-938.
- Sevrioukova, I. (2011). Apoptosis-Inducing Factor: Structure, Function, and Redox Regulation. *Antioxidants & Redox Signaling*, 14(12), pp.2545-2579.
- Tavaría, M., Gabriele, T., Kola, I. and Anderson, R. (1996). A hitchhiker's guide to the human Hsp70 family. *Cell Stress Chaper*, 1(1), p.23.
- Thali, M., Bukovsky, A., Kondo, E., Rosenwlrth, B., Walsh, C., Sodroski, J. and Göttlinger, H. (1994). Functional association of cyclophilin A with HIV-1 virions. *Nature*, 372(6504), pp.363-365.
- Villanueva, R. (2015). *The reductase activity of the human Apoptosis Inducing Factor: molecular bases and implications in the enzyme apoptotic function*. Doctorado. Universidad de Zaragoza. Universidad de Ciencias, Departamento de Bioquímica y Biología Molecular y Celular.
- Villanueva, R., Ferreira, P., Marcuello, C., Usón, A., Miramar, M., Peleato, M., Lostao, A., Susin, S. and Medina, M. (2015). Key Residues Regulating the Reductase Activity of the Human Mitochondrial Apoptosis Inducing Factor. *Biochemistry*, 54(33), pp.5175-5184.

Anexo 1: Secuencia de la proteína AIFM1h

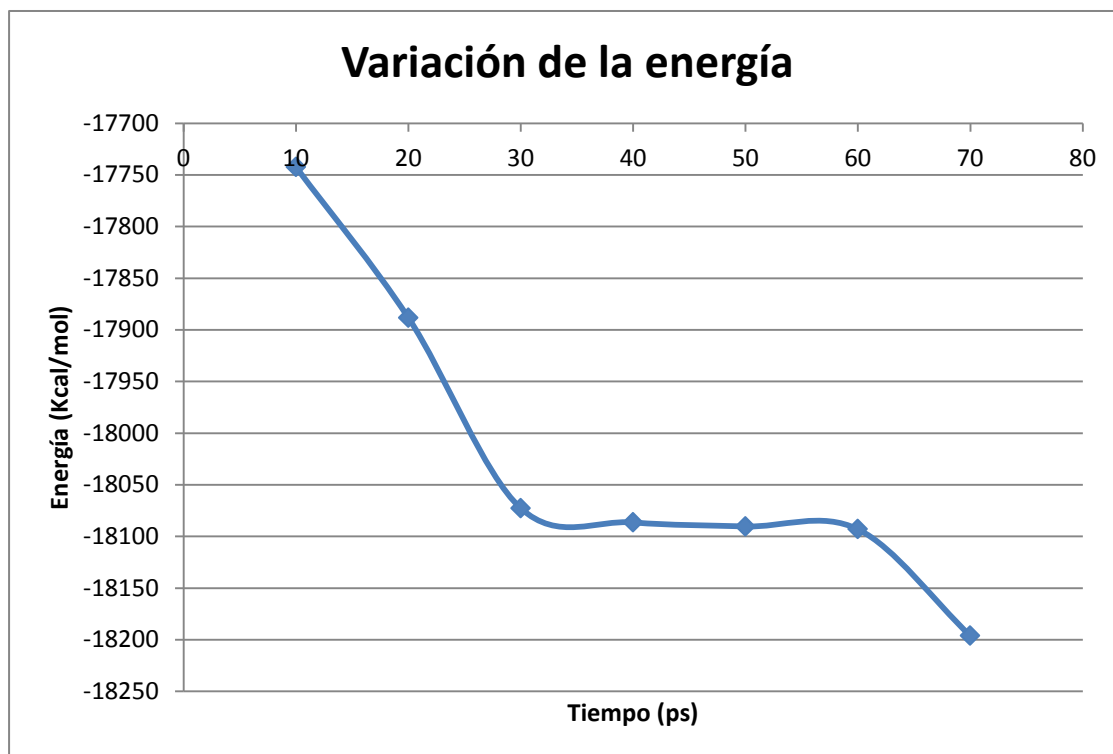
10	20	30	40	50
MFRCGGLAAG	ALKQKLVPLV	RTVCVRSRQ	RNRLPGNLFQ	RWHVPLELQM
60	70	80	90	100
TRQMASSGAS	GGKIDNSVLV	LIVGLSTVGA	GAYAYKTMKE	DEKRYNERIS
110	120	130	140	150
GLGLTPEQKQ	KKAALSASEG	EEVPQDKAPS	HVPFLLIGGG	TAAFAAARSI
160	170	180	190	200
RARDPGARVL	IVSEDPPELY	MRPPLSKELW	FSDDPNVTKT	LRFKQWNGKE
210	220	230	240	250
RSIYFQPPSF	YVSAQDLPHI	ENGGVAULTG	KKVVQLDVRD	NMVKLNDGSQ
260	270	280	290	300
ITYEKLIAT	GGTPRSLSAI	DRAGAEVKSR	TTLFRKIGDF	RSLEKISREV
310	320	330	340	350
KSITIIIGGGF	LGSELACALG	RKARALGTEV	IQLFPEKGNM	GKILPEYLSN
360	370	380	390	400
WTMEKVRREG	VKVMPNAIVQ	SVGVSSGKLL	IKLKDGRKVE	TDHIVAAGVL
410	420	430	440	450
EPNVELAKTG	GLEIDSDFGG	FRVNAELQAR	SNIWVAGDAA	CFYDIKLGRR
460	470	480	490	500
RVEHHDHAVV	SGRLAGENMT	GAAPYWHQS	MFWSDLGPDV	GYEAIGLVDS
510	520	530	540	550
SLPTVGVFAK	ATAQDNPKSA	TEQSGTGIRS	ESETESEASE	ITIPPSTPAV
560	570	580	590	600
PQAPVQGEDY	GKGVIFYLRD	KVVVGIVLWN	IFNRMPIARK	IIKDGEQHED
610				
LNEVAKLFNI	HED			

Secuencia no observable en 4BV6:

Región 546-558: STPAVPQAPVQGE.

Secuencia no observable en 4BUR:

Región 509-558: KATAQDNPKSATEQSGTGIRSESETESEASEITIPPSTPAVPQAPVQGE.

Anexo 2: Datos sobre las dinámicas realizadas en Discovery Studio.

Anexo 3: Tabla de energía de las interacciones de CypA con AIFM1h en estado oxidado.

Cluster	Miembros	Representante	Energía
0	81	Media	-508.9
		Energía más baja	-543.1
1	54	Media	-483.8
		Energía más baja	-572.4
2	54	Media	-526.1
		Energía más baja	-526.1
3	51	Media	-569.7
		Energía más baja	-569.7
4	48	Media	-481.4
		Energía más baja	-552.6
5	43	Media	-489.2
		Energía más baja	-585.0
6	37	Media	-549.1
		Energía más baja	-549.1
7	36	Media	-506.6
		Energía más baja	-520.3
8	34	Media	-555.2
		Energía más baja	-555.2
9	31	Media	-549.5
		Energía más baja	-569.1
10	27	Media	-476.0
		Energía más baja	-577.3
11	27	Media	-498.1
		Energía más baja	-498.1
12	24	Media	-508.0
		Energía más baja	-508.0
13	24	Media	-492.6
		Energía más baja	-511.4
14	23	Media	-507.0
		Energía más baja	-565.6
15	19	Media	-535.8
		Energía más baja	-535.8
16	19	Media	-493.6
		Energía más baja	-493.6
17	19	Media	-477.0
		Energía más baja	-526.6
18	18	Media	-547.4
		Energía más baja	-547.4
19	18	Media	-480.0
		Energía más baja	-509.9
20	17	Media	-554.0
		Energía más baja	-574.0
21	14	Media	-518.3
		Energía más baja	-518.3

22	13	Media	-507.6
		Energía más baja	-507.6
23	12	Media	-533.6
		Energía más baja	-533.6
24	11	Media	-539.6
		Energía más baja	-539.6
25	11	Media	-527.2
		Energía más baja	-527.2
26	11	Media	-524.4
		Energía más baja	-524.4
27	11	Media	-521.0
		Energía más baja	-521.0
28	10	Media	-566.6
		Energía más baja	-566.6
29	10	Media	-513.1
		Energía más baja	-513.1

Anexo 4: Tabla de energía de las interacciones de HSP70 con AIFM1h en estado oxidado.

Cluster	Miembros	Representante	Energía
0	92	Media	-687.3
		Energía más baja	-774.1
1	79	Media	-609.6
		Energía más baja	-751.5
2	63	Media	-594.6
		Energía más baja	-680.2
3	62	Media	-656.1
		Energía más baja	-678.1
4	52	Media	-664.1
		Energía más baja	-775.4
5	50	Media	-580.9
		Energía más baja	-685.8
6	50	Media	-661.3
		Energía más baja	-699.0
7	46	Media	-636.5
		Energía más baja	-671.4
8	39	Media	-708.3
		Energía más baja	-708.3
9	32	Media	-609.7
		Energía más baja	-655.0
10	31	Media	-594.9
		Energía más baja	-644.8
11	30	Media	-711.8
		Energía más baja	-711.8
12	26	Media	-590.7
		Energía más baja	-688.4
13	19	Media	-672.6
		Energía más baja	-672.6
14	17	Media	-571.7
		Energía más baja	-712.5
15	14	Media	-603.6
		Energía más baja	-603.6
16	13	Media	-635.9
		Energía más baja	-635.9
17	12	Media	-661.0
		Energía más baja	-661.0
18	12	Media	-644.3
		Energía más baja	-644.3
19	11	Media	-681.3
		Energía más baja	-681.3
20	11	Media	-588.7
		Energía más baja	-702.5
21	10	Media	-725.2
		Energía más baja	-725.2

22	9	Media	-634.2
		Energía más baja	-677.3
23	9	Media	-615.9
		Energía más baja	-629.1
24	7	Media	-596.9
		Energía más baja	-700.5
25	4	Media	-580.4
		Energía más baja	-580.4
26	1	Media	-591.7
		Energía más baja	-591.7

Anexo 5: Tabla de energía de las interacciones de CHCHD4 con AIFM1h en estado oxidado.

Cluster	Miembros	Representante	Energía
0	78	Media	-704.6
		Energía más baja	-704.6
1	73	Media	-737.8
		Energía más baja	-737.8
2	57	Media	-737.8
		Energía más baja	-841.5
3	56	Media	-607.8
		Energía más baja	-707.4
4	54	Media	-630.8
		Energía más baja	-662.7
5	51	Media	-587.0
		Energía más baja	-699.2
6	48	Media	-651.6
		Energía más baja	-682.1
7	45	Media	-584.1
		Energía más baja	-719.8
8	44	Media	-692.0
		Energía más baja	-692.0
9	42	Media	-751.3
		Energía más baja	-775.3
10	38	Media	-730.7
		Energía más baja	-730.7
11	30	Media	-602.3
		Energía más baja	-637.0
12	27	Media	-583.9
		Energía más baja	-724.6
13	27	Media	-702.3
		Energía más baja	-702.3.
14	25	Media	-658.0
		Energía más baja	-658.0
15	25	Media	-651.1
		Energía más baja	-655.2
16	20	Media	-780.7
		Energía más baja	-780.7
17	18	Media	-594.1
		Energía más baja	-656.1
18	17	Media	-592.5
		Energía más baja	-631.4
19	14	Media	-659.8
		Energía más baja	-659.8
20	14	Media	-647.5
		Energía más baja	-647.5
21	14	Media	-622.3
		Energía más baja	-622.3

22	14	Media	-590.1
		Energía más baja	-629.2
23	13	Media	-598.3
		Energía más baja	-682.5
24	12	Media	-589.9
		Energía más baja	-638.9
25	4	Media	-600.2
		Energía más baja	-623.9

Anexo 6: Tabla de energía de las interacciones de H2AX con AIFM1h en estado oxidado.

Cluster	Miembros	Representante	Energía
0	115	Media	-885.6
		Energía más baja	-885.6
1	64	Media	-711.0
		Energía más baja	-787.6
2	49	Media	-853.1
		Energía más baja	-853.1
3	41	Media	-721.2
		Energía más baja	-835.6
4	40	Media	-838.8
		Energía más baja	-838.8
5	27	Media	-704.1
		Energía más baja	-801.0
6	27	Media	-819.3
		Energía más baja	-819.3
7	26	Media	-819.3
		Energía más baja	-819.3
8	24	Media	-725.2
		Energía más baja	-780.3
9	22	Media	-745.3
		Energía más baja	-791.5
10	20	Media	-859.1
		Energía más baja	-859.1
11	20	Media	-844.1
		Energía más baja	-844.1
12	20	Media	-736.9
		Energía más baja	-865.4
13	18	Media	-721.5
		Energía más baja	-763.7
14	17	Media	-746.6
		Energía más baja	-812.4
15	16	Media	-774.9
		Energía más baja	-774.9
16	15	Media	-739.8
		Energía más baja	-877.0
17	15	Media	-783.3
		Energía más baja	-790.5
18	14	Media	-809.1
		Energía más baja	-868.6
19	14	Media	-798.0
		Energía más baja	-798.0
20	14	Media	-779.0
		Energía más baja	-779.0
21	14	Media	-704.6
		Energía más baja	-742.4

22	13	Media	-704.8
		Energía más baja	-731.9
23	13	Media	-840.4
		Energía más baja	-840.4
24	13	Media	-761.4
		Energía más baja	-761.4
25	13	Media	-708.8
		Energía más baja	-856.5
26	13	Media	-696.3
		Energía más baja	-793.8
27	12	Media	-794.0
		Energía más baja	-794.0
28	11	Media	-736.1
		Energía más baja	-736,1
29	10	Media	-715,0
		Energía más baja	-788,0

Anexo 7: Tabla de energía de las interacciones CypA con AIFM1h en estado reducido.

Cluster	Miembros	Representante	Energía
0	147	Media	-586.0
		Energía más baja	-784.3
1	53	Media	-543.5
		Energía más baja	-728.6
2	50	Media	-520.5
		Energía más baja	-611.5
3	48	Media	-524.0
		Energía más baja	-597.1
4	46	Media	-520.1
		Energía más baja	-599.5
5	40	Media	-597.1
		Energía más baja	-597.1
6	37	Media	-527.5
		Energía más baja	-582.9
7	32	Media	-556.1
		Energía más baja	-589.6
8	28	Media	-592.4
		Energía más baja	-592.4
9	27	Media	-581.6
		Energía más baja	-581.6
10	27	Media	-580.3
		Energía más baja	-596.7
11	26	Media	-522.8
		Energía más baja	-584.5
12	25	Media	-519.3
		Energía más baja	-640.5
13	25	Media	-522.8
		Energía más baja	-701.8
14	25	Media	-556.2
		Energía más baja	-567.1
15	25	Media	-550.6
		Energía más baja	-572.4
16	23	Media	-522.8
		Energía más baja	-582.0
17	23	Media	-565.6
		Energía más baja	-584.1
18	22	Media	-657.3
		Energía más baja	-657.3
18	18	Media	-552.6
		Energía más baja	-595.4
20	17	Media	-647.7
		Energía más baja	-647.7
21	17	Media	-553.0
		Energía más baja	-553.0
22	16	Media	-539.1

		Energía más baja	-611.6
23	13	Media	-555.9
		Energía más baja	-610.8
24	12	Media	-575.3
		Energía más baja	-575.3
25	11	Media	-524.5
		Energía más baja	-548.5
26	10	Media	-528.3
		Energía más baja	-565.8
27	10	Media	-596.0
		Energía más baja	-596.0
28	10	Media	-546.1
		Energía más baja	-560.6
29	9	Media	-536.6
		Energía más baja	-551.1

Anexo 8: Tabla de energía de las interacciones de HSP70 con AIFM1h en estado reducido.

Cluster	Miembros	Representante	Energía
0	78	Media	-636.3
		Energía más baja	-794.1
1	71	Media	-750.6
		Energía más baja	-834.9
2	70	Media	-676.9
		Energía más baja	-700.8
3	43	Media	-623.1
		Energía más baja	-808.8
4	43	Media	-605.6
		Energía más baja	-690.6
5	39	Media	-624.2
		Energía más baja	-846.9
6	37	Media	-670.7
		Energía más baja	-693.7
7	26	Media	-714.9
		Energía más baja	-715.2
8	25	Media	-613.2
		Energía más baja	-658.7
9	24	Media	-697.7
		Energía más baja	-697.7
10	23	Media	-596.1
		Energía más baja	-698.2
11	22	Media	-698.4
		Energía más baja	-698.4
12	21	Media	-629.2
		Energía más baja	-762.7
13	21	Media	-746.4
		Energía más baja	-746.4
14	18	Media	-654.5
		Energía más baja	-670.1
15	18	Media	-645.0
		Energía más baja	-645.0
16	17	Media	-596.0
		Energía más baja	-645.2
17	17	Media	-681.1
		Energía más baja	-681.1
18	16	Media	-710.4
		Energía más baja	-714.5
18	16	Media	-602.7
		Energía más baja	-681.2
20	15	Media	-630.1
		Energía más baja	-655.3
21	15	Media	-598.6
		Energía más baja	-793.9

22	15	Media	-648.7
		Energía más baja	-660.9
23	14	Media	-671.6
		Energía más baja	-671.6
24	13	Media	-634.2
		Energía más baja	-673.3
25	12	Media	-622.4
		Energía más baja	-682.0
26	12	Media	-660.4
		Energía más baja	-753.0
27	12	Media	-701.3
		Energía más baja	-701.3
28	12	Media	-665.7
		Energía más baja	-665.7
29	12	Media	-656.2
		Energía más baja	-656.6

Anexo 9: Tabla de energía de las interacciones de CHCHD4 con AIFM1h en estado reducido.

Cluster	Miembros	Representante	Energía
0	100	Media	-787.0
		Energía más baja	-815.0
1	70	Media	-659.3
		Energía más baja	-848.8
2	70	Media	-704.0
		Energía más baja	-775.6
3	58	Media	-742.6
		Energía más baja	-819.5
4	49	Media	-660.0
		Energía más baja	-770.7
5	38	Media	-792.2
		Energía más baja	-792.2
6	37	Media	-659.9
		Energía más baja	-827.8
7	36	Media	-677.5
		Energía más baja	-718.1
8	35	Media	-651.7
		Energía más baja	-708.0
9	34	Media	-671.7
		Energía más baja	-697.9
10	33	Media	-819.1
		Energía más baja	-819.1
11	33	Media	-650.3
		Energía más baja	-743.8
12	28	Media	-662.5
		Energía más baja	-722.6
13	26	Media	-679.4
		Energía más baja	-679.4
14	23	Media	-700.0
		Energía más baja	-722.4
15	21	Media	-641.5
		Energía más baja	-759.2
16	21	Media	-641.9
		Energía más baja	-690.9
17	21	Media	-704.6
		Energía más baja	-704.6
18	21	Media	-629.5
		Energía más baja	-687.5
18	20	Media	-705.0
		Energía más baja	-705.0
20	19	Media	-647.1
		Energía más baja	-690.4
21	18	Media	-726.3
		Energía más baja	-726.3

22	16	Media	-671.8
		Energía más baja	-671.8
23	13	Media	-703.4
		Energía más baja	-703.4
24	13	Media	-682.2
		Energía más baja	-682.2
25	13	Media	-656.7
		Energía más baja	-656.7
26	11	Media	-656.7
		Energía más baja	-716.8
27	10	Media	-703.8
		Energía más baja	-703.8
28	9	Media	-635.7
		Energía más baja	-672.4

Anexo 10: Tabla de energía de las interacciones de H2AX con AIFM1h en estado reducido.

Cluster	Miembros	Representante	Energía
0	72	Media	-901.3
		Energía más baja	-912.6
1	61	Media	-959.3
		Energía más baja	-959.3
2	58	Media	-861.1
		Energía más baja	-986.3
3	48	Media	-892.1
		Energía más baja	-892.1
4	44	Media	-752.9
		Energía más baja	-865.8
5	41	Media	-758.8
		Energía más baja	-910.7
6	40	Media	-790.7
		Energía más baja	-883.7
7	35	Media	-761.1
		Energía más baja	-821.3
8	27	Media	-866.6
		Energía más baja	-866.6
9	26	Media	-833.9
		Energía más baja	-833.9
10	25	Media	-839.4
		Energía más baja	-839.4
11	23	Media	-817.9
		Energía más baja	-825.1
12	23	Media	-933.9
		Energía más baja	-933.9
13	23	Media	-831.9
		Energía más baja	-831.9
14	22	Media	-836.1
		Energía más baja	-875.0
15	21	Media	-797.5
		Energía más baja	-814.8
16	19	Media	-922.9
		Energía más baja	-922.9
17	17	Media	-938.0
		Energía más baja	-938.0
18	16	Media	-733.4
		Energía más baja	-822.3
18	14	Media	-832.6
		Energía más baja	-832.6
20	13	Media	-916.8
		Energía más baja	-916.8
21	13	Media	-860.0
		Energía más baja	-860.0

22	13	Media	-848.5
		Energía más baja	-848.5
23	13	Media	-823.8
		Energía más baja	-823.8
24	13	Media	-751.2
		Energía más baja	-826.4
25	12	Media	-790.9
		Energía más baja	-836.4
26	12	Media	-827.2
		Energía más baja	-850.0
27	11	Media	-738.1
		Energía más baja	-750.0
28	10	Media	-742.7
		Energía más baja	-772.7
29	10	Media	-740.0
		Energía más baja	-834.2