



**Universidad Miguel**

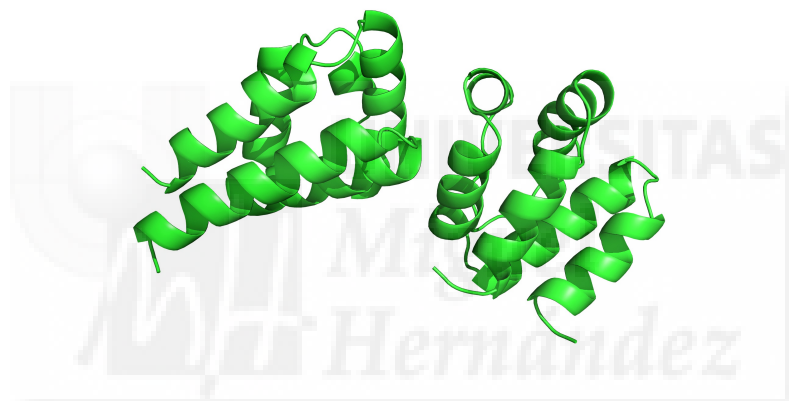
**Hernández**

**Campus de Elche**

**TRABAJO FIN DE GRADO**

**BIOTECNOLOGÍA (2016-2017)**

Búsqueda computacional de péptidos para la inhibición de las proteínas AIM2 y ASC del inflamasoma sin afectar a otros mecanismos relacionados.



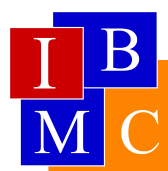
**Autor**

Abraham Muñoz Chicharro

**Tutores**

Gregorio Fernández Ballester

Magdalena Nikolaeva Koleva



Instituto de Biología Molecular y Celular.  
Facultad de Ciencias Experimentales.  
Departamento de Bioquímica y Biología Molecular

# ÍNDICE

1. RESUMEN.....	Pág. 3
2. INTRODUCCIÓN .....	Pág. 5
2.1. FAMILIAS PROTEICAS .....	Pág. 6
2.1.1. SUPERFAMILIA PROTEICA <i>DEATH DOMAIN</i> .....	Pág. 6
2.1.2. FAMILIA PYHIN.....	Pág. 7
2.1.3. FAMILIA NLR .....	Pág. 8
2.2. INTERACCIONES ESPECÍFICAS PROTEÍNA-PROTEÍNA (PPIs).....	Pág. 9
2.3. INFLAMASOMAS .....	Pág. 12
2.3.1. Inflamasoma AIM2.....	Pág. 12
2.3.2. Inflamasoma NLR.....	Pág. 13
2.4. EL INFLAMASOMA EN ENFERMEDADES HEPÁTICAS.....	Pág. 14
3. ANTECEDENTES Y OBJETIVOS .....	Pág. 17
4. MATERIALES Y MÉTODOS.....	Pág. 18
4.1. Modelos de interacción .....	Pág. 18
4.2. Visualización.....	Pág. 18
4.3. Medidas de Energía .....	Pág. 19
4.4. Minimización.....	Pág. 19
4.5. Mutación de ligandos.....	Pág. 20
4.6. Búsqueda de secuencias .....	Pág. 20
5. RESULTADOS.....	Pág. 22
5.1. Modelo de interacción ASC-PYD/ASC-PYD.....	Pág. 22
5.2. Modelo de interacción AIM2-PYD/ASC-PYD .....	Pág. 23
5.3. Matrices de puntuación específicas de posición (PSSM).....	Pág. 24
5.4. Selección de los mejores residuos por posición .....	Pág. 25
6. DISCUSIÓN .....	Pág. 31
6.1. Comparaciones entre secuencias de aminoácidos wild-type y mutados .....	Pág. 31
6.2. Interacción de los ligandos mutados con ASC-PYD.....	Pág. 35
6.3. Interacción de los ligandos mutados con AIM2-PYD .....	Pág. 36
6.4. Comprobación experimental de los resultados.....	Pág. 36
7. CONCLUSIÓN .....	Pág. 38
8. BIBLIOGRAFÍA .....	Pág. 39

## 1. RESUMEN

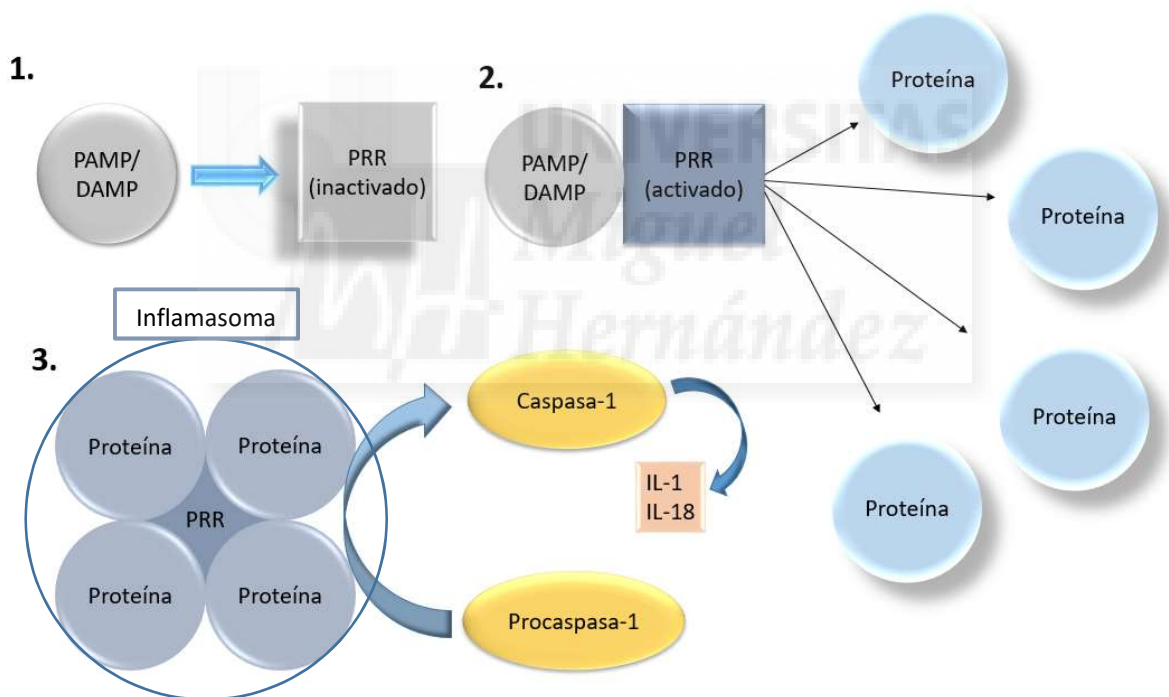
La inflamación es una respuesta coordinada del sistema inmune innato para frenar el avance de microorganismos patógenos. Está dirigida por un macrocomplejo peptídico llamado inflammasoma, constituido a su vez por familias de péptidos que interactúan entre sí. Sin embargo, hay condiciones en las que la inflamación no está bien regulada por el organismo, generando o bien una inflamación insuficiente, que deriva en una infección persistente, o bien una inflamación excesiva, que deriva en una enfermedad crónica o sistémica. Se ha desarrollado un trabajo de investigación *in silico* con el objetivo de encontrar péptidos que se unan e inhiban las proteínas que inician la formación del macrocomplejo. Para ello, ha sido necesario realizar modelos de interacción entre los dominios pyrin de ASC y AIM2 con acoplamiento molecular, minimización molecular y mutagénesis *in silico*, y con ellos estudiar aquellas regiones por las que interactúan las proteínas para seleccionar y aislar las secuencias peptídicas más convenientes sobre las que construir ligandos. En total se han hallado 4 ligandos de los cuales, mediante mutaciones *in silico*, se han obtenido un total de 1248 péptidos mutados. Finalmente, se han escogido los mejores y se han comparado con los ligandos silvestres, comprobando que hay un incremento en las interacciones hidrofóbicas y una disminución en las interacciones electrostáticas, que se resume en una menor especificidad de los ligandos mutados en comparación con los silvestres y una mayor afinidad. Con los datos generados se realizará un trabajo de experimentación *in vitro* para comprobar si los péptidos generados se unen a las zonas de interés.

Inflammation is a coordinate response of the innate immunity system to control the expansion of pathogenic microorganisms. It is managed by a peptidic macrocomplex called inflammasome, made in turn by peptide families which interact with each other. However, there are conditions where inflammation is not regulated adequately causing an insufficient inflammation, producing a persistent infections, or an excessive inflammation, producing chronic or systemic diseases. A working investigation *in silico* has been developed to find peptides that bind and inhibit the proteins in charge of starting the assembly of the macrocomplex. To do this, it has been necessary to predict pyrin domain interactions of AIM2 and ASC proteins with docking, molecular minimization and *in silico* mutagenesis, and with the predicted interactions, evaluate the binding regions to select and isolate the best peptidic sequences to use as templates to build ligands. 4 ligands have been found in total to use like templates. In total, with mutagenesis *in silico* over the 4 templates, 1248 mutated peptides have been generated, from which, the best has been chosen, and it has been compared with wild-type ligands. Finally, the interactions have been checked elucidating an increase of hydrophobic interactions while there is a decrease of electrostatic interactions. It means that the specificity of mutated ligands is lower than wild-type ligands and the specificity is higher. It is expected realize an *in vitro* research work in the future to verify if the generated peptides bind to the interest interaction zones.

**Palabras clave:** Inflammasoma, AIM2, ASC, *in silico*, mutagénesis, interacciones.

## 2. INTRODUCCIÓN

La inflamación es una respuesta protectora llevada a cabo por el sistema inmune innato que se activa debido a estímulos originados por una infección (antígenos) o por daño (células muertas) y está directamente regulado por el hospedador. Una inflamación insuficiente puede llevar a una infección persistente mientras que un exceso puede causar enfermedades inflamatorias crónicas o sistémicas. La función de inflamación del sistema inmune depende del reconocimiento de patrones moleculares asociados a patógenos (PAMPs), provenientes del patógeno invasor, y del reconocimiento de patrones moleculares asociados a daño (DAMPs), inducidos por estrés endógeno. Tanto los PAMPs como los DAMPs producen la activación de receptores de reconocimiento de patrones (PRRs), los cuales dan lugar a la formación de un macrocomplejo peptídico denominado inflamasoma, que desencadena la cascada de señalización en fases posteriores de la ruta de inflamación para la producción de interferón tipo I y citoquinas proinflamatorias que darán lugar a la inflamación tal como muestra la **figura 1**.



**Figura 1.** Esquema general de un inflamasoma. **1.** El PAMP o DAMP es reconocido por el PRR. **2.** El PRR interacciona con el PAMP o DAMP y se activa, dando lugar a su interacción con otras proteínas. **3.** Debido a las interacciones del PRR activado con las proteínas, se forma un macrocomplejo, al que llamamos inflamasoma, que activa la procaspasa-1 en caspasa-1, enzima responsable de la inflamación.

Muchas familias de PRRs son importantes para el complejo del inflamasoma (Guo et al. 2015), incluyendo proteínas de receptores tipo NOD (NLRs) y AIM2 (absent in melanoma 2). Ante ciertos estímulos, las proteínas NLRs o AIM2 pueden oligomerizar para activar las caspasas-1. Estas caspasas-1 activan a su vez las citoquinas proinflamatorias IL-1 e IL-18, que causan la inflamación.

Recientes avances en investigación han ayudado a entender mejor la activación macromolecular del complejo del inflamasoma al descubrirse que la maquinaria molecular del mismo está compuesta por una gran diversidad de proteínas que interaccionan entre ellas para dar lugar a una respuesta adecuada ante un agente externo.

## **2.1. FAMILIAS PROTEICAS**

Las familias proteicas son un conjunto de proteínas homólogas entre sí, es decir, guardan una similitud en la secuencia, y por tanto tienen una estructura parecida. Existe una gran diversidad de familias y todas ellas realizan una función específica en el organismo. Más concretamente, las familias proteicas Death Domain, PyHIN y NLR son responsables de la inflamación. Dichas familias interaccionan unas con otras para formar el macrocomplejo proteico del inflamasoma.

### **2.1.1. SUPERFAMILIA PROTEICA DEATH DOMAIN**

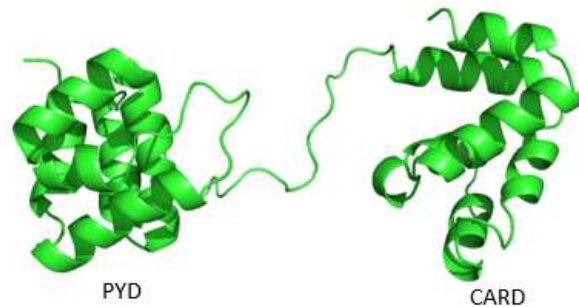
Según la base de datos *deathdomain* ([www.deathdomain.org](http://www.deathdomain.org)), la superfamilia *Death Domain* (en adelante DD) es una de las familias más grandes que existen en cuanto a módulos de interacción proteica, y juega un papel fundamental en las rutas de señalización de la apoptosis, inflamación, necrosis e inmunidad celular. Un ejemplo son los complejos de caspasa para la apoptosis e inflamación, que son ensamblados por las proteínas DD.

La relación entre la función y la estructura es una característica clave que define a la superfamilia de proteínas DD. La clasificación de las proteínas en las distintas subfamilias se realiza principalmente por la homología de secuencias. Estos módulos de interacción de dominios están evolutivamente conservados en muchos organismos pluricelulares incluyendo mamíferos, insectos (como *Drosophila melanogaster*) y gusanos (*Caenorhabditis elegans*). En el genoma humano existen varios dominios de interacción clasificados en 4 subfamilias:

- Subfamilia Death Domain (DD)
- Subfamilia Death effector domain (DED)
- Subfamilia Caspase recruitment domain (CARD)
- Subfamilia Pyrin domain (PYD)

Los PYD tienen aproximadamente 95 residuos los cuales no presentan una estructura terciaria conservada. No obstante, todos los dominios comparten una estructura secundaria muy similar, compuesta de 6 hélices al igual que el resto de las subfamilias anteriormente citadas.

Una de las proteínas pertenecientes a esta familia clave para la formación del inflamasoma es la proteína ASC (de Alba 2009), compuesta por un dominio CARD y un dominio PYD (en adelante ASC-PYD), tal y como muestra la **figura 2**.

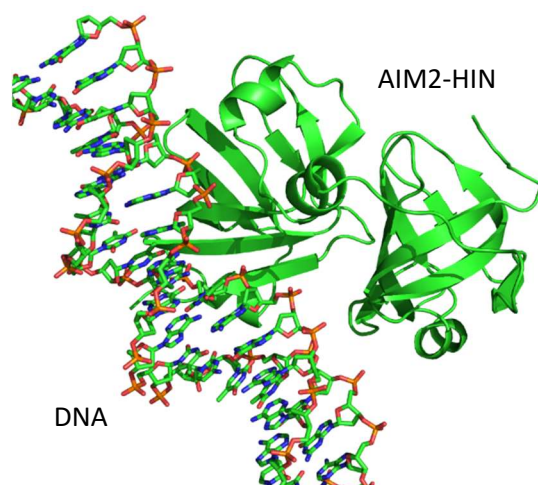


**Figura 2.** Proteína ASC. A la izquierda está representado el dominio PYD y a la derecha el dominio CARD. Los dominios proteicos no tienen una estructura terciaria conservada pero están compuestos por 6 hélices alfa. (Estructura cristalográfica 2KN6.PDB)

### 2.1.2. FAMILIA PYHIN

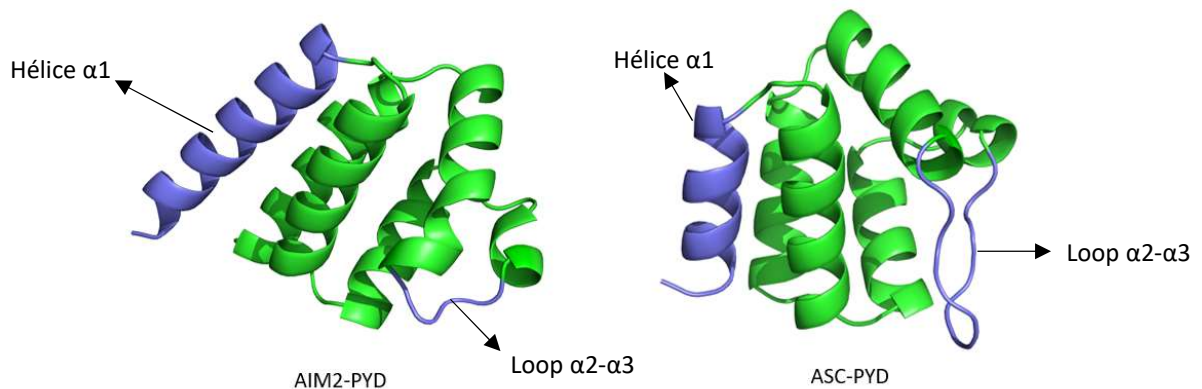
Es una familia proteica presente en mamíferos a la cual pertenecen las proteínas HIN-200, contenidas en el gen IFI16. Son proteínas caracterizadas por contener un dominio PYD y un dominio HIN, cuya conformación inicial es la de autoinhibición. De esta forma, se evita que el inflamasoma quede constitutivamente activo.

AIM2 es una proteína perteneciente a esta familia. Dicha proteína está compuesta por dos dominios: un dominio HIN (en adelante AIM2-HIN) y un dominio PYD (en adelante AIM2-PYD). El dominio AIM2-HIN está compuesto por dos superestructuras secundarias de barriles  $\beta$ , capaz de unirse a ADN extracelular (Jin et al. 2012) tal como puede observarse en la **figura 3**.



**Figura 3.** Proteína AIM2-HIN. En la estructura se observan dos dominios HIN interaccionando con DNA citosólico. (Estructura cristalina 3RN2.PDB).

El dominio PYD de AIM2 (AIM2-PYD) comparte muchos de los residuos clave con el resto de dominios PYD aunque el loop  $\alpha 2-\alpha 3$  es más corto y muestra motivos únicos entre el amino y el carboxilo terminal. También cabe destacar que en el modelo cristalográfico la hélice  $\alpha 6$  es más larga que en el resto de dominios PYDs. (Chu et al. 2015). Para ver estas diferencias se ha comparado el cristal 4O7Q.PDB (Lu et al. 2014), que contiene el dominio AIM2-PYD, y el cristal 1UCP.PDB (Liepinsh et al. 2003), que contiene el dominio ASC-PYD, en la **figura 4**.



**Figura 4.** Comparación de AIM2-PYD y ASC-PYD. En azul se muestra las hélices  $\alpha 6$  y los loops  $\alpha 2-\alpha 3$ . A la izquierda está representada la estructura 4O7Q.PDB y a la derecha la estructura 1UCP.PDB.

La mayoría de los residuos cargados de AIM2-PYD están expuestos y muestran una distribución distinta al resto de dominios PYD debido a los residuos que se encuentran en el interior de la hélice  $\alpha 2$ . Los residuos ácidos predominan en las hélices  $\alpha 1$  y  $\alpha 2$  mientras que los residuos alcalinos están presentes en las hélices  $\alpha 5$  y  $\alpha 6$ . Cerca de la hélice  $\alpha 2$  se encuentran los residuos hidrofóbicos F27 y F28, únicos en el dominio AIM2-PYD y altamente conservados entre las distintas especies. Es posible que todas estas características contribuyan a la función específica de la proteína AIM2, que es la de activar el inflamasoma.

### 2.1.3. FAMILIA NLR

Las proteínas pertenecientes a la familia NLR son reguladores del sistema inmune y, como tal, también están asociados a desórdenes inmunológicos.

Esta familia proteica presenta algunas características básicas tales como un dominio efector en el N-terminal, un dominio NACHT, y una región de unión a ligando en el C-terminal, que está compuesto por secuencias repetitivas ricas en leucinas (LRRs) según (Jenny P. et al. 2008).

El dominio NACHT es un dominio NTPasa que se encuentra en el sitio de unión a nucleótidos (en adelante NBS) de ciertas proteínas como NLRP3 o NLRC4 y está formado por tres subdominios

helicoidales unidos entre sí. Las secuencias LRRs son secuencias ricas en leucina, las cuales aportan una estructura versátil de unión a lipopéptidos y proteínas. Por tanto, actúan como PRRs en el sistema inmune innato.

Las proteínas de todas estas familias interactúan entre sí con sus respectivos dominios para poder ensamblar el complejo del inflamasoma.

## 2.2. INTERACCIONES ESPECÍFICAS PROTEÍNA-PROTEÍNA (PPIs)

Existen evidencias sobre proteínas homotípicas, es decir, proteínas que sólo interactúan con otras proteínas de su misma subfamilia. A este tipo de interacciones se las denomina interacciones específicas proteína-proteína (PPIs).

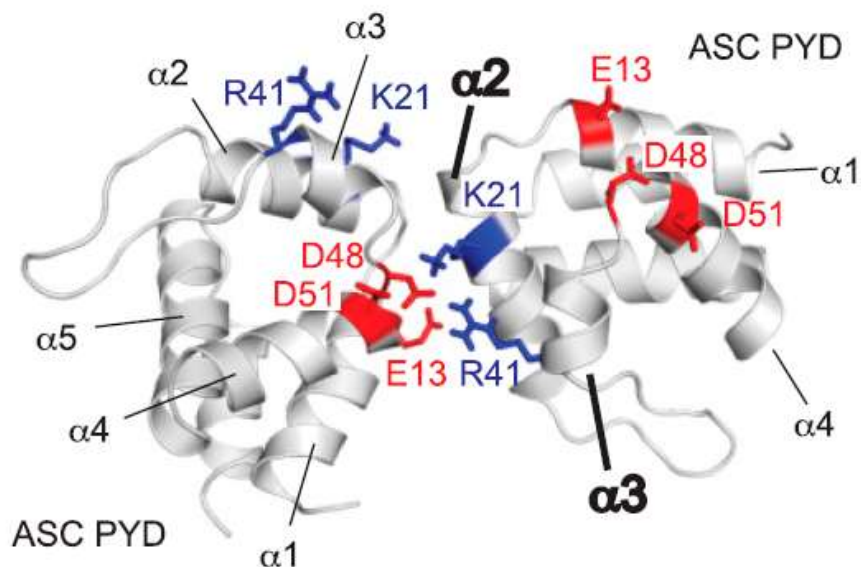
Un ejemplo son las proteínas pertenecientes a la superfamilia DD. Las PPIs entre las distintas proteínas de esta familia proteica, como interacciones entre PYDs, son imprescindibles para la determinación de eventos en la ruta de formación de los distintos inflamasomas y, por tanto, son un objeto de estudio para conocer e inhibir dicha ruta.

### INTERACCIÓN ASC-PYD/ASC-PYD

Se ha determinado que los dominios PYD de ASC tienen residuos básicos en las hélices  $\alpha 2$  y  $\alpha 3$  y residuos ácidos en las hélices  $\alpha 1$  y  $\alpha 4$ , clave para la interacción ASC-PYD/ASC-PYD (Vajjhala et al. 2012) tal como se muestra en la **figura 5**.

Estos dominios PYD interactúan entre sí para formar una estructura que da lugar a un complejo proteico donde se soportan los dominios CARD para activar la procaspasa, la cual a su vez activará la IL-1 y la IL-18, que son responsables directas de la inflamación y la fiebre.





**Figura 5.** Interacción entre dominios ASC-PYD. Los residuos D48, D51 (hélice  $\alpha 4$ ) y E13 (hélice  $\alpha 1$ ) del primer dominio ASC-PYD interactúan electrostáticamente con los residuos K21 (hélice  $\alpha 2$ ) y R41 (hélice  $\alpha 3$ ) del segundo dominio ASC-PYD. En la interacción, los residuos básicos del primero y los residuos ácidos del segundo quedan libres, permitiendo a su vez la interacción con otros dominios ASC-PYD pudiendo así desencadenar una oligomerización. Figura tomada de (Vajjhala et al. 2012).

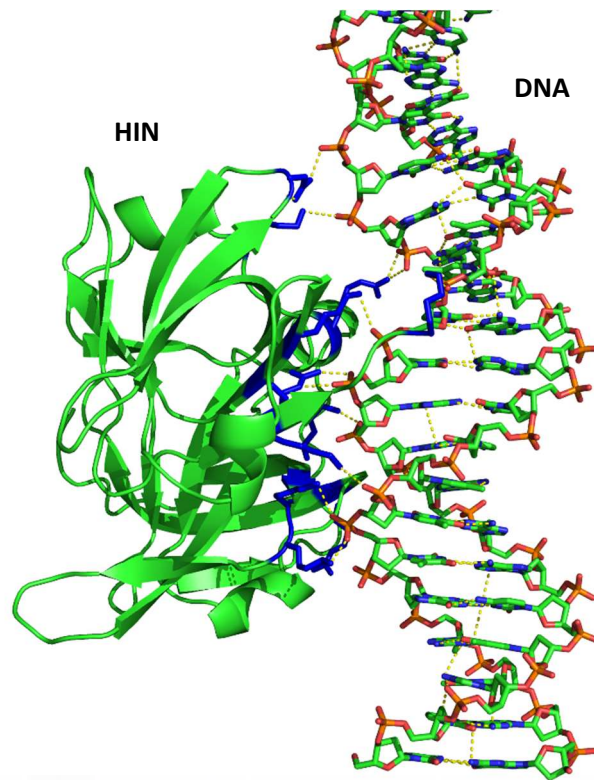
Existen estructuras cristalográficas que muestran las interacciones ASC-PYD/ASC-PYD. Dichas estructuras contienen un complejo de varios ASC-PYD interactuando electrostáticamente entre sí mediante los aminoácidos K21 y R41 de uno de los dominios ASC-PYD con los aminoácidos E13, D48 y D51 del otro. Dichas interacciones se pueden visualizar en la estructura cristalográfica 3J63.PDB (Lu et al. 2014)

#### INTERACCIÓN AIM2-PYD/AIM2-HIN

La hélice ácida  $\alpha 2$  del dominio PYD se une a la zona alcalina del dominio HIN, quedando inicialmente la proteína AIM2 autoinhibida. Actualmente no existen estructuras que describan esta interacción.

#### INTERACCIÓN AIM2-HIN/DNA

Existen modelos estructurales que establecen la interacción que tiene lugar entre AIM2-HIN y el DNA. Dicha interacción es electrostática y ocurre entre los aminoácidos básicos del dominio AIM2-HIN y los grupos fosfato del DNA tal como se muestra en la **figura 6**.



**Figura 6.** Interacción entre el dominio AIM2-HIN y el DNA. Los residuos de HIN que participan en dicha interacción (K160, K162, K163, R204, R244, K251, N287, K309, R311, K335, K340) están representados en azul. (Estructura 3RN2.PDB).

#### INTERACCIÓN AIM2-PYD/ASC-PYD

Aunque actualmente no hayan estructuras cristalográficas que describan esta interacción, existen estudios de *docking* que han propuesto a los residuos ácidos de la hélice  $\alpha 2$  de AIM2-PYD y los residuos hidrofóbicos adyacentes a dicha hélice, como residuos clave que interaccionan con el dominio ASC-PYD. Esta hipótesis fue confirmada con un estudio experimental de mutagenicidad. También pueden ocurrir otros modos de interacción entre la hélice ácida  $\alpha 1$  de AIM2-PYD con las hélices básicas  $\alpha 2$ - $\alpha 3$  de ASC-PYD y/o la hélice básica  $\alpha 5$  de AIM2-PYD con la hélice ácida  $\alpha 4$  de ASC-PYD (Jin et al. 2013).

Todas estas interacciones son responsables de la formación y funcionamiento del inflamasoma, como se verá más detalladamente en el siguiente apartado, y su estudio es esencial para poder entender todas las rutas y procesos que llevan a activar la inflamación.

## 2.3. INFLAMASOMAS

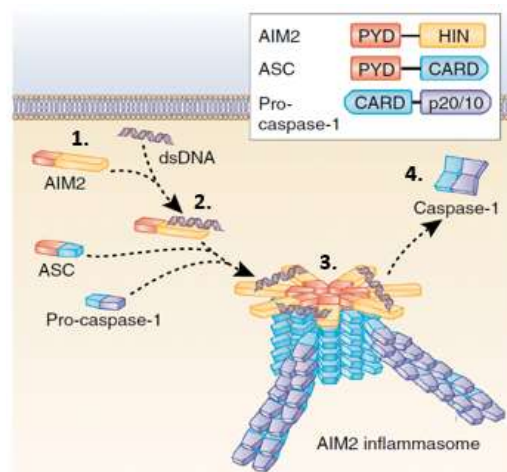
Los inflammasomas son complejos proteicos macromoleculares capaces de activar la inflamación. En el caso de los inflammasomas hepáticos estudiados, los más relevantes son el AIM2 y el NLR.

Los nombres de los inflammasomas denotan la proteína sobre la que se soportan. Así, en el caso de los inflammasomas NLRs, éstos se forman por uno o dos miembros de la familia NLR. En el caso de la proteína AIM2, ésta es la responsable de formar el inflammasoma AIM2.

Los inflammasomas NLRs y AIM2 se han asociado a una serie de enfermedades autoinmunes y autoinflamatorias como la fibrosis hepática o enfermedades hepáticas asociadas al alcoholismo (ALDs), por ello son objetos de estudio en diversas investigaciones.

### 2.3.1. Inflammasoma AIM2

La proteína AIM2 puede formar un inflammasoma que reclute a la caspasa-1. Esta proteína está constitutivamente inhibida debido a que su dominio HIN interacciona con su dominio PYD. Cuando el dominio HIN se une a su estímulo, el ADN citosólico (el cual puede encontrarse en el medio extracelular durante una infección por patógenos), deja libre el dominio PYD, quedando así la proteína activada. Una vez el dominio PYD presenta su zona de unión libre, se inicia una oligomerización de proteínas ASC que da lugar a la formación del inflammasoma. Cuando el complejo proteico está ensamblado, los dominios CARD de las proteínas ASC interaccionan con los dominios CARD de las procaspasas-1, quedando éstas activadas en caspasas-1 (Haitao G. et al. 2015) tal y como muestra la **figura 7**.



**Figura 7.** Ruta canónica del inflammasoma AIM2. **1.** Los dominios HIN y PYD de AIM2 interaccionan entre sí, quedando la proteína AIM2 inicialmente autoinhibida. **2.** El dominio HIN se une a su estímulo, el ADN citosólico y el dominio PYD se libera, quedando la proteína activada. **3.** El dominio PYD interacciona con el dominio PYD de las proteínas ASC y comienza una oligomerización que da lugar al complejo proteico del inflammasoma. Cuando el complejo proteico queda ensamblado, los dominios CARD de las proteínas ASC interaccionan con los dominios CARD de las pro-caspasas-1. **4.** La interacción lleva a la activación de procaspasas-1 en caspasas-1. Figura tomada de Haitao G. et al. 2015.

Las diferencias estructurales entre las interacciones de ASC-PYD con otros dominios PYD determinan el mecanismo de interacción PYD-PYD resultante, el cual está implicado probablemente en el ensamblaje del inflamasoma y, consecuentemente, en la respuesta inmune innata. Aunque está bien estudiada la importancia de los PYD en el ensamblaje del inflamasoma, se ha determinado recientemente el mecanismo molecular subyacente, revelando un mecanismo de oligomerización único responsable de la activación de caspasa-1.

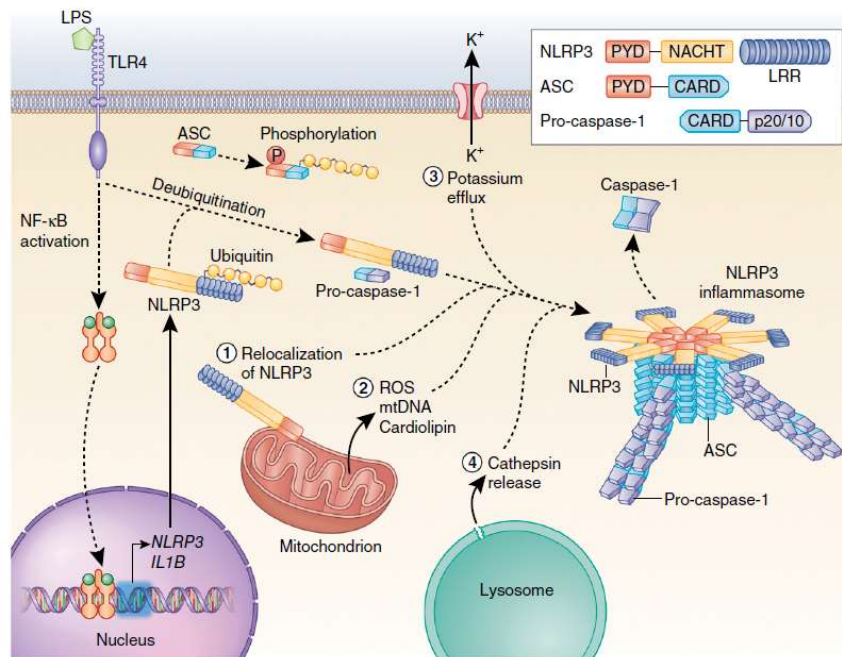
### 2.3.2. Inflamasoma NLR

Este inflamasoma se ensambla con proteínas de la familia NLR. Se piensa que estas proteínas inician o regulan los mecanismos de defensa del hospedador a través de plataformas de señalización que desencadenan la activación de caspasas y del factor nuclear potenciador de las cadenas kappa de células B activadas (NF- $\kappa$ B). También se ha visto que las proteínas NLRs que contienen un PYD en el N-terminal se asocian con proteínas que contienen un dominio PYD y un dominio CARD (PYCARD) como ASC, asociadas a apoptosis.

A pesar de su papel fundamental en la coordinación de rutas de señalización en el sistema inmune, su modo de acción está todavía sin descubrir.

#### ➤ Inflamasoma NLRP3

La proteína NLRP3 contiene un dominio PYD, un dominio NACHT de unión a nucleótidos (NBS) y un motivo de repeticiones ricas en leucinas (LRR). Esta proteína interacciona con las proteínas PYCARD, las cuales contienen un dominio de reclutamiento de caspasas (Haitao G. et al. 2015) tal como muestra la **figura 8**.



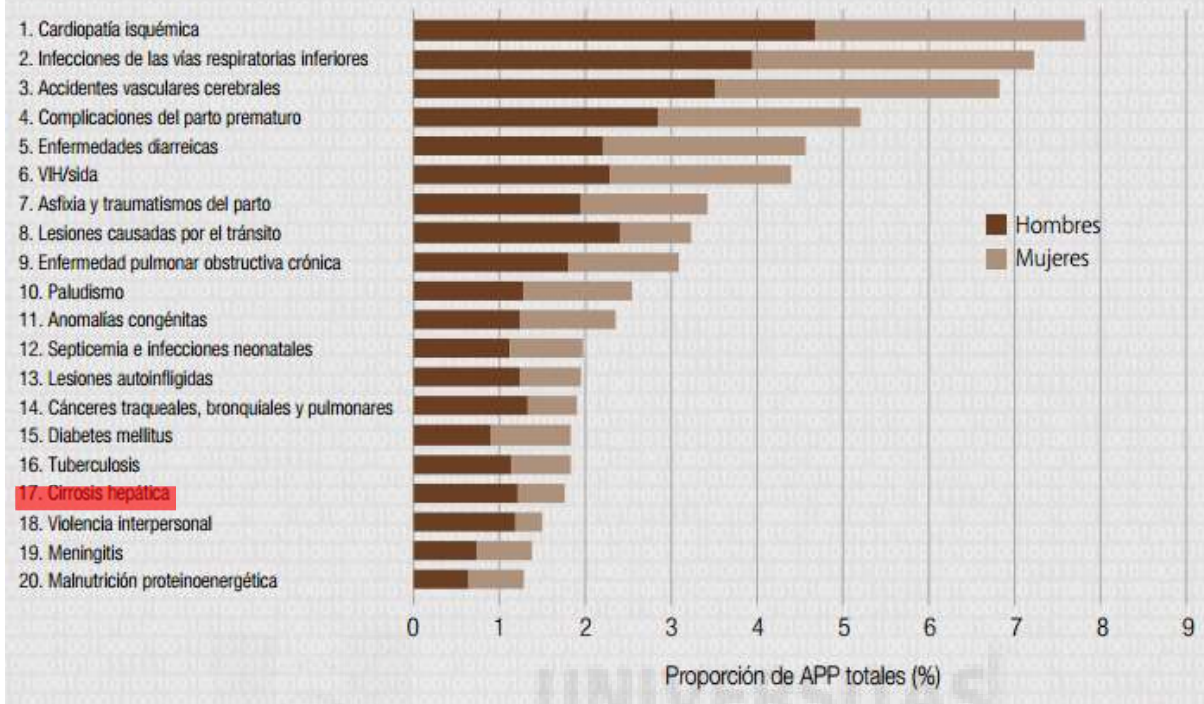
**Figura 8.** Ruta canónica del inflamasoma NLRP3. La lipoproteína se une a la LRR (TLR4), el cual activa el NF-κβ. La activación lleva a una sobreexpresión, en el núcleo, de proteína NLRP3, la cual está inicialmente ubiquitinada. Para que pueda llevarse a cabo la formación del inflamasoma es necesario que la proteína NLRP3 sea desubiquitinada y que, paralelamente, la proteína ASC sea ubiquitinada y fosforilada. Además, la proteína NLRP3 debe recibir una segunda señal de activación bien sea por la mitocondria (liberación de ROS, ADN mitocondrial o cardiolipina), el lisosoma (liberación de cathepsina) o por la expulsión de K<sup>+</sup> del citosol por parte de los canales de potasio. Una vez está formado el inflamasoma, la procaspasa-1 sufre proteólisis y se transforma en caspasa-1, responsable directa de la inflamación. Figura tomada de Haitao G. et al. 2015.

## 2.4. EL INFLAMASOMA EN ENFERMEDADES HEPÁTICAS

La inflamación también puede desencadenarse por DAMPs sin que haya infección por un patógeno. Estos DAMPs activan la proteína AIM2 al igual que lo haría una molécula de ADN citosólico y da lugar a una enfermedad inflamatoria crónica.

Una de las enfermedades inflamatorias crónicas más frecuentes en el mundo es la cirrosis hepática. Según la Organización Mundial de la Salud, la cirrosis hepática se encuentra entre las 20 primeras causas de muerte en todo el mundo tal y como se muestra en la **figura 9**.

## Las 20 primeras causas de muerte prematura de hombres y mujeres a nivel mundial, 2012



**Figura 9.** Gráfica en la que se representan las 20 primeras causas de muerte en 2012 según la OMS. La cirrosis hepática, enfermedad producida por una inflamación crónica, se encuentra en el puesto 17 ([http://apps.who.int/iris/bitstream/10665/112817/1/WHO\\_HIS\\_HSI\\_14.1\\_spa.pdf](http://apps.who.int/iris/bitstream/10665/112817/1/WHO_HIS_HSI_14.1_spa.pdf) --> Fecha última entrada (09/06/2017)).

La cirrosis hepática es consecuencia de una inflamación crónica del hígado y ésta se desencadena por DAMPs. De forma normal, los DAMPs permanecen ocultos del entorno extracelular. No obstante, cuando hay daños en el tejido, son liberados al medio y reconocidos por el inflamasoma, que se activa y da lugar a la inflamación. En el caso de los hepatocitos, cuando existe daño tisular en el hígado, se liberan DAMPs tales como ácido úrico, ATP, péptidos beta-amiloideos, DNA y cristales de colesterol.

El reconocimiento de los DAMPs por el inflamasoma deriva en una inflamación crónica que se presenta en varias enfermedades tales como **enfermedades hepáticas alcohólicas (ALDs)** producidas por un consumo excesivo de alcohol que da lugar a una esteatosis hepática. Si la situación se prolonga, se puede alcanzar una esteatohepatitis que dé lugar a una cirrosis. Los estadios avanzados son irreversibles y tienen una alta tasa de mortalidad por fallo hepático.

Asimismo, el **daño hepático inducido por una sobredosis de paracetamol** da lugar a una muerte en cadena de hepatocitos que liberan DAMPs al medio, produciendo la activación del inflamasoma, lo que da lugar a una inflamación crónica.

Otras afecciones pueden ser la **lesión por reperusión isquémica** que consiste en una privación de oxígeno a los hepatocitos seguido de una exposición repentina a oxígeno, lo que también da lugar a

una muerte masiva de hepatocitos que liberan DAMPs al medio activando el inflamasoma. Esto ocurre en pacientes a los que se les ha realizado un trasplante de hígado, han perdido mucha sangre, sufren un *shock* por fallo cardíaco o padecen una sepsis.

Por último, hay que destacar a la **hepatitis B**, que al tratarse de una infección crónica, da lugar a la liberación continuada de PAMPs que activan constantemente el inflamasoma.

Cuando tienen lugar las PPIs en ausencia de un patógeno, se ensambla el inflamasoma sin un objetivo de defensa frente a un agente externo, lo que conlleva a que las funciones de ciertos órganos, como el hígado, se vean comprometidas.

Por lo tanto, dado que los eventos que ocurren en los receptores para las señales de apoptosis, inflamación e inmunidad están relacionados con enfermedades como el cáncer y los desórdenes inmunitarios, la investigación de las interacciones proteína-proteína adquieren una gran importancia.



### 3. ANTECEDENTES Y OBJETIVOS

El estudio de proteínas de la superfamilia DD facilitará la comprensión de eventos de señalización mediados por éstas, así como las rutas de señalización de la apoptosis, la inflamación y la inmunidad celular. Además, puede ayudar en el descubrimiento de nuevos fármacos, tales como péptidos que interfieran en las interacciones proteína-proteína entre las subfamilias de la superfamilia DD.

En este trabajo se utilizarán métodos computacionales para el estudio de las interacciones entre las proteínas ASC-ASC y ASC-AIM2, y el diseño de posibles fármacos al disponer de la existencia de estructuras cristalográficas con buena resolución. De esta forma pueden obtenerse buenos resultados y, además, de una forma rápida y económica.

El objetivo general es abordar la inhibición del inflammasoma hepático AIM2 a través de las distintas rutas de señalización en las que participan las interacciones PPIs. Para ello es necesario recopilar y entender la información existente sobre dichas interacciones que son un objeto de estudio principal para abarcar el problema de la inflamación crónica.

En el caso del inflammasoma hepático AIM2, hay dos tipos distintos de PPIs:

1. Interacción ASC-PYD/ASC-PYD
2. Interacción ASC-PYD/AIM2-PYD

Basándose en estos PPIs, los objetivos específicos de este trabajo son:

- Seleccionar secuencias peptídicas que sean capaces de interferir en la interacción entre las proteínas que forman parte del inflammasoma de AIM2.
  - Interacciones PYD-PYD de ASC
  - Interacciones PYD-PYD entre AIM2 y ASC
- Mejorar la afinidad y especificidad de las secuencias seleccionadas mediante mutagénesis *in silico* y calcular la energía de interacción.



## 4. MATERIALES Y MÉTODOS

Para el estudio de las regiones de interacción entre las distintas proteínas y la obtención de los péptidos de interés se han llevado a cabo procedimientos *in silico* de construcción de modelos de los complejos silvestres (*wild-type*), identificación de las zonas de interacción mediante puentes de hidrógeno e interacciones hidrofóbicas y polares, selección de péptidos que podrían actuar como ligandos y mutagénesis *in silico* de los mismos para mejorar su afinidad y especificidad, procedimientos que se describen a continuación en mayor profundidad.

### 4.1. Modelos de interacción

Se han modelado las interacciones AIM2-PYD/ASC-PYD y ASC-PYD/ASC-PYD para establecer qué zonas presentan una mayor área de interacción y, mediante dichas zonas, poder establecer los ligandos de interés.

En el modelo cristalográfico 3J63.PDB está representada la oligomerización de varios dominios ASC-PYD que dan lugar al inflammasoma.

En la estructura cristalográfica 4O7Q.PDB está representado el dominio AIM2-PYD y en el 1UCP.PDB un único dominio ASC-PYD. Con estos dos cristales y la información experimental obtenida de la interacción AIM2-PYD/ASC-PYD (Jin et al. 2013) se realizó un acoplamiento virtual o *docking* para modelar dicha interacción. Para ello, se utilizó el software Gramm-X de Vakser Lab (<http://vakser.compbio.ku.edu/resources/gramm/grammx/>), basado en el potencial de aproximación de Lennard-Jones para realizar el acoplamiento molecular.

### 4.2. Visualización

Se visualizaron los modelos cristalográficos con PyMOL v1.6.0.0, un software que renderiza y anima estructuras moleculares 3D.

Para visualizar las zonas de interacción en el modelo ASC-PYD/AIM2-PYD se usó la opción “show > polar contacts > find polar contacts > within selections”, las cuales coinciden con las descritas por el artículo (Vajjhala et al. 2012): la hélice  $\alpha_1$  y la hélice  $\alpha_4$  (región negativa) de un dominio interaccionan con la hélice  $\alpha_2$  y hélice  $\alpha_3$  (región positiva) del otro.

Para aislar los ligandos de ASC-PYD/ASC-PYD, se visualizó el cristal 3J63.PDB y, con el mismo software, que permite eliminar o añadir secuencias peptídicas del cristal que contiene el ensamblaje de los

dominios PYD que forman el inflammasoma, se seleccionaron dos dominios ASC-PYD que interactúan entre sí y se eliminó el resto de la estructura.

### 4.3. Medidas de Energía

Con el software FoldX v2.8, que evalúa los efectos de las mutaciones en la estabilidad, plegamiento y dinámica de las proteínas y ácidos nucleicos, se midió la energía de interacción del modelo AIM2-PYD/ASC-PYD, la cual resultó ser negativa, es decir, tiene mayor estabilidad que ambas proteínas por separado:

$$1. E_{\text{interacción}} = E_{\text{complex}} - (E_{\text{lig}} + E_{\text{dom}}).$$

Cuanto más negativa sea esta energía, más fuerte será la interacción entre ligando y receptor. Para medir las energías de interacción, FoldX v2.8 realiza una estimación cuantitativa de la importancia de las interacciones que contribuyen a la estabilidad de las proteínas y complejos proteicos y, mediante datos empíricos obtenidos de experimentos de ingeniería proteica, establece las distintas variaciones de energía.

Esta estimación cuantitativa para hacer un cálculo aproximado de la estabilidad de las proteínas se basa en la suma de todas las interacciones de sus aminoácidos entre sí:

$$2. E_{\text{estabilidad}} = \Sigma \text{ de términos}$$

Donde  $\Sigma$  de términos es el sumatorio de las energías de interacciones electrostáticas, puentes de hidrógeno, puentes Cys-Cys, interacciones de Van der Waals, interacciones dipolo-dipolo, interacciones dipolo-dipolo inducido, solvatación polar, solvatación hidrofóbica, entropía de esqueleto peptídico, entropía de cadena lateral, etc.

### 4.4. Minimización

Se aplicó minimización molecular con Yasara v15.5, un software de modelado y simulación moleculares, para reducir la energía del complejo, es decir, buscar su conformación más estable. El programa funciona con un PVL (Portable Vector Language), un sistema de desarrollo que permite visualizar proteínas y realizar simulaciones en tiempo real con unos campos de fuerza muy precisos en ordenadores estándar.

Primero se relajó todo el complejo AIM2-PYD/ASC-PYD con el algoritmo "Steepest descent minimization" (simulation > temperature control > steepest descent minimization) y, tras fijarlo y dejar

libres las cadenas laterales responsables de la interacción, se aplicó el algoritmo “Simulated annealing minimization” (simulation > temperature control > simulated annealing minimization).

Después se volvió a medir la energía de interacción con FoldX v2.8 resultando ser más baja que la anterior, lo que significa que se trata de una conformación más estable que la del *wild-type*.

#### **4.5. Mutación de ligandos**

Se realizaron una serie de mutaciones virtuales en los ligandos con el objetivo de aumentar la afinidad y la especificidad por la proteína diana a inhibir.

El protocolo de mutagénesis se inicia con la mutación de los residuos en cada posición de los distintos ligandos en alaninas. A continuación, se construye una matriz de puntuaciones por cada complejo ligando-proteína, en la que puede verse qué aminoácidos desestabilizan la interacción proteína-ligando y qué aminoácidos la favorecen al sustituir cada posición del ligando poly-Alanina por cada uno de los 20 aminoácidos y medir la energía tras cada mutación.

Para la construcción de los ligandos poly- Alanina, las respectivas mutaciones, la medida de las energías de interacción y la construcción de las matrices de puntuación, se emplearon archivos de comandos (*scripts*) basados en Python, un lenguaje de programación libre. Dichos *scripts* ejecutan el programa FoldX con los parámetros de entrada necesarios para realizar todas las tareas mencionadas.

A partir de los datos de energía obtenidos con FoldX se generan unas matrices, representativas de la energía de interacción de un residuo concreto del ligando contra el receptor, permaneciendo el resto de posiciones del ligando como alaninas. Es por eso que no se contemplan los posibles choques estéricos con residuos adyacentes del ligando. Este hecho se considera en un siguiente paso, construyéndose los ligandos mediante combinaciones de los residuos más favorables de cada posición, volviéndose a medir la energía de interacción ligando-proteína.

#### **4.6. Búsqueda de secuencias**

Los aminoácidos que se seleccionan de cada posición del ligando son los que menor energía de interacción presentan para el complejo, es decir, mejorarán supuestamente la afinidad y/o especificidad del ligando con la proteína diana.

Del análisis de las matrices de puntuación, se observan posiciones restrictivas, que se mutarán por aminoácidos que disminuyan la energía del complejo, y posiciones tolerantes en los que es indiferente el aminoácido a mutar. La existencia de posiciones tolerantes se explica porque los ligandos tienen conformación de  $\alpha$ -hélice. Por tanto, hay residuos que quedan hacia el lado opuesto de la zona de

interacción. La importancia de estas posiciones tolerantes reside en que, si no se tienen en cuenta a la hora de realizar combinaciones, éstas disminuyen considerablemente, disminuyendo a su vez la carga de cálculo computacional.

Para empezar con las mutaciones de los ligandos, se seleccionan los mejores aminoácidos por posición a partir de las matrices.

Luego, para llevar a cabo todas las combinaciones favorables en todas las posiciones salvo las tolerantes, se utiliza nuevamente un script basado en Python, el cual ejecuta a FoldX para llevar a cabo las siguientes tareas:

- 1) Realizar las combinaciones de los aminoácidos seleccionados sobre el ligando para generar todos los complejos.
- 2) Medir la energía de interacción de los complejos.

Al finalizar estas tareas se genera un archivo Excel con todos los complejos generados. Es importante ordenar de menor a mayor estos complejos por orden de energía de interacción para saber cuáles son los mejores candidatos, ya que cuanto menor sea la energía de interacción, mayor será la afinidad del ligando por el receptor. Para terminar, se comparan estos complejos con el WT y se seleccionan los que mejoren la energía de interacción.

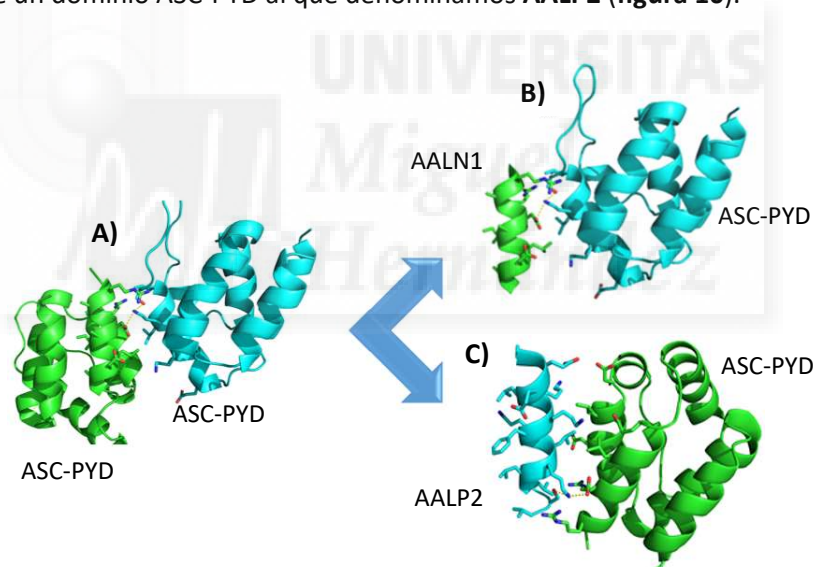
## 5. RESULTADOS

Para llevar a cabo la selección de péptidos capaces de interferir en las interacciones proteína-proteína implicados en el inflammasoma hepático, se aislaron péptidos a partir de las estructuras conocidas de interacción.

### 5.1. Modelo de interacción ASC-PYD/ASC-PYD

A partir del complejo cristalizado de interacción del dominio ASC-PYD/ASC-PYD se seleccionó la hélice  $\alpha 1$  (secuencia 2-GRARDAILDALLEN-14) implicada en la interacción y se eliminó el resto del dominio al que pertenece, ya que dicha hélice tiene una mayor superficie de interacción que la hélice  $\alpha 4$ . De esta forma se obtuvo un modelo estructural de un ligando inicial que interacciona con la **región positiva** de un dominio ASC-PYD al que denominamos **AALN1 (figura 10)**.

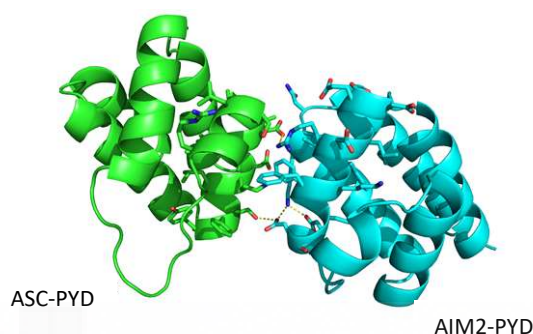
Por otro lado, se seleccionó la hélice  $\alpha 2$  (secuencia 17-AEELKFKLKLKLLSV-30) y se eliminó el resto del dominio al que pertenece, obteniéndose así, un modelo de ligando inicial que interacciona con la **región negativa** de un dominio ASC-PYD al que denominamos **AALP2 (figura 10)**.



**Figura 10.** Representación de los ligandos AALN1 y AALP2 basados en el complejo de interacción ASC-PYD/ASC-PYD. **A)** Complejo de estructura conocida ASC-PYD/ASC-PYD. **B)** Complejo derivado AALN1/ASC-PYD donde el ligando interacciona con la región positiva del dominio ASC-PYD. **C)** Complejo AALP2/ASC-PYD donde el ligando interacciona con la región negativa del ligando.

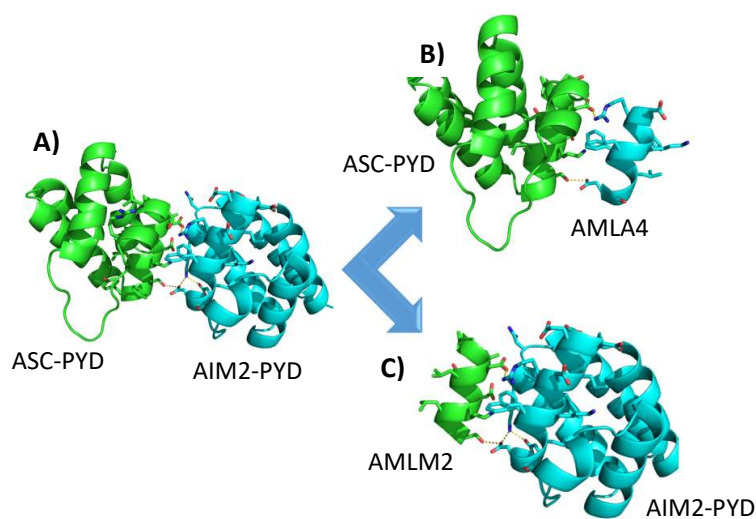
## 5.2. Modelo de interacción AIM2-PYD/ASC-PYD

Para abordar el estudio de interacción entre AIM2-PYD/ASC-PYD, fue necesario construir un modelo de interacción mediante acoplamiento (*docking*) proteína-proteína entre el PYD de AIM2 y el PYD de ASC ya que no se disponen de estructuras cristalográficas de la interacción. Se restringió la interacción a los aminoácidos D51 y D54 (hélice  $\alpha$ 4) de ASC-PYD y los aminoácidos K64, R67 y K71 (hélice  $\alpha$ 5) de AIM2-PYD con el software Gramm-X Vakser Lab para llevar a cabo dicho acoplamiento. Una vez finalizado el *docking*, se obtuvieron 5 modelos de los cuales se escogió aquel en el que interaccionaban F27 y F28 de AIM2-PYD, imprescindibles para la unión a ASC-PYD, que se ilustra en la **figura 11**.



**Figura 11.** Modelo de interacción entre el dominio ASC-PYD y el dominio AIM2-PYD. El modelo ha sido obtenido mediante acoplamiento (*docking*) proteína-proteína.

Para aislar los ligandos *wild-type* de ASC-PYD/AIM2-PYD, se visualiza el complejo AIM2-PYD/ASC-PYD con PyMOL v1.6.0.0 y se construye un modelo de interacción AIM2-PYD/ligando eliminando toda el dominio ASC-PYD salvo la hélice  $\alpha$ 4 (secuencia 48-DALDLTKLVS-58), al que se denomina **AMLA4**, y un modelo de interacción ASC-PYD/ligando eliminando todo el dominio AIM2-PYD salvo la hélice  $\alpha$ 5 (secuencia 23-DRFKFFLSD-31), al que se denomina **AMLM2** (**figura 12**).

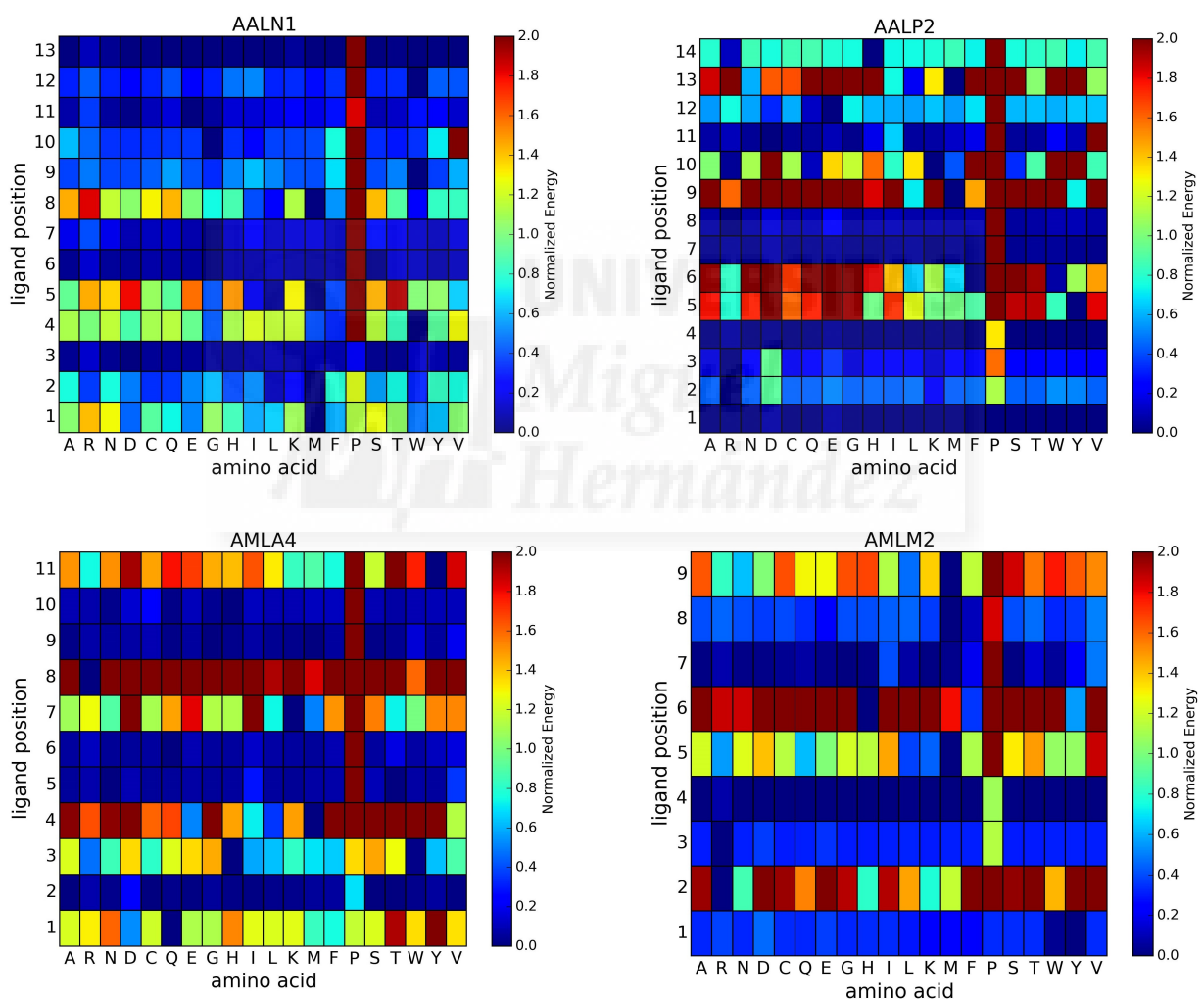


**Figura 12.** Obtención de los ligandos AMLA4 y AMLM2 basados en el complejo de interacción ASC-PYD/AIM2-PYD. **A)** Complejo completo ASC-PYD/AIM2-PYD. **B)** Complejo AMLA4/AIM2-PYD. **C)** Complejo AMLM2/ASC-PYD

A partir de estos 4 complejos (AALN1, AALP2, AMLA4, AMLM2) se determinan las características de interacción entre los ligandos (péptidos) y los dominios mediante el cálculo de las matrices de puntuación específica de posición (PSSM).

### 5.3. Matrices de puntuación específicas de posición (PSSM)

En las matrices de la **figura 13** se muestra la energía de interacción normalizada de los 20 aminoácidos naturales para cada posición de los 4 ligandos. El color azul indica una energía de interacción más baja y por tanto más favorable (interacción ligando-dominio más fuerte), mientras que el rojo indica mayor energía de interacción por repulsión de cargas o choque estérico, lo que inestabiliza el complejo y, por tanto, es menos favorable.



**Figura 13.** Matriz de puntuación de los distintos ligandos. La leyenda de color representa la energía normalizada: cuanto mayor sea el tono de azul en la casilla, mayor fuerza de interacción entre ligando-dominio, y cuanto mayor sea el tono de rojo, menor fuerza de interacción entre ligando-dominio. En el eje Y se representan las posiciones del ligando peptídico y en el X los 20 aminoácidos naturales. Las casillas azules representan las posiciones que puede ocupar un determinado aminoácido aportando estabilidad al complejo. Las casillas rojas indican las posiciones que no debería ocupar un determinado aminoácido puesto que inestabiliza el complejo.

#### 5.4. Selección de los mejores residuos por posición

Una vez obtenidas las matrices, se escogieron aquellos aminoácidos a sustituir en los ligandos que favoreciesen la interacción proteína-ligando tal y como muestra la **tabla 1**.

**Tabla 1.** Tablas de combinación de mutaciones en los distintos ligandos. La columna “Posición cristal” indica la posición del aminoácido en el cristal del que se ha derivado. La columna “Posición Matriz” indica la posición del aminoácido en el complejo donde se han llevado a cabo las mutaciones. La columna WT indica los aminoácidos *wild-type* por los que está compuesto el ligando. La columna Mut. indica los aminoácidos que sustituyen a los aminoácidos *wild-type*. Las posiciones subrayadas son las posiciones tolerantes, aquellos aminoácidos que no intervienen en la interacción.

AALN1				AALP2			
Posición cristal	Posición Matriz	WT	Mut.	Posición cristal	Posición Matriz	WT	Mut.
2	1	G	MDEW	<u>17</u>	<u>1</u>	<u>A</u>	<u>A</u>
3	2	R	RK	<u>18</u>	<u>2</u>	<u>E</u>	<u>E</u>
<u>4</u>	<u>3</u>	<u>A</u>	<u>A</u>	<u>19</u>	<u>3</u>	<u>E</u>	<u>E</u>
5	4	R	WFM	<u>20</u>	<u>4</u>	<u>L</u>	<u>L</u>
6	5	D	ILM	21	5	K	KY
<u>7</u>	<u>6</u>	<u>A</u>	<u>A</u>	22	6	K	FW
<u>8</u>	<u>7</u>	<u>I</u>	<u>I</u>	<u>23</u>	<u>7</u>	<u>E</u>	<u>E</u>
9	8	L	ILMWF	<u>24</u>	<u>8</u>	<u>K</u>	<u>K</u>
<u>10</u>	<u>9</u>	<u>D</u>	<u>D</u>	25	9	L	MY
<u>11</u>	<u>10</u>	<u>A</u>	<u>A</u>	26	10	K	RQKMS
<u>12</u>	<u>11</u>	<u>L</u>	<u>L</u>	<u>27</u>	<u>11</u>	<u>L</u>	<u>L</u>
<u>13</u>	<u>12</u>	<u>E</u>	<u>E</u>	<u>28</u>	<u>12</u>	<u>L</u>	<u>L</u>
<u>14</u>	<u>13</u>	<u>N</u>	<u>N</u>	29	13	S	NLM
				30	14	V	RH

AMLA4				AMLM2			
Posición cristal	Posición Matriz	WT	Mut.	Posición cristal	Posición Matriz.	WT	Mut.
48	1	D	QD	23	1	D	DYW
<u>49</u>	<u>2</u>	<u>A</u>	<u>A</u>	24	2	R	R
50	3	L	HWR	25	3	F	FR
51	4	D	ELM	<u>26</u>	<u>4</u>	<u>K</u>	<u>K</u>
<u>52</u>	<u>5</u>	<u>L</u>	<u>L</u>	27	5	F	KLMR
<u>53</u>	<u>6</u>	<u>I</u>	<u>I</u>	28	6	F	HFY
54	7	D	KMLT	<u>29</u>	<u>7</u>	<u>L</u>	<u>L</u>
55	8	K	R	30	8	S	SM
<u>56</u>	<u>9</u>	<u>L</u>	<u>L</u>	31	9	D	MLD
<u>57</u>	<u>10</u>	<u>V</u>	<u>V</u>				
58	11	S	RFY				



Para obtener el mínimo número posible de combinaciones y reducir así el tiempo de cálculo, se ha dejado en las posiciones tolerantes (filas de casillas azules representadas en las matrices) el aminoácido *wild-type*. En el resto de posiciones, se realizan combinaciones con los mejores Aa seleccionados para cada uno de los 4 complejos. El número de complejos derivados de los 4 complejos iniciales se muestran en la **tabla 2**.

**Tabla 2.** Número de complejos obtenidos por ligando.

AALN1 → 360 complejos	AALP2 → 240 complejos
AMLA4 → 216 complejos	AMLM2 → 432 complejos

A continuación, se miden las energías de los nuevos complejos con FoldX, de tal forma que quedan consideradas todas las posiciones a la vez, y se ordenan según su energía de interacción. Para finalizar, se comparan las energías con el *wild-type*.

Las mutaciones realizadas dan como resultado una serie de secuencias peptídicas con una energía de interacción proteína-ligando más baja que la respectiva con el ligando *wild-type*, tal como muestran las tablas con los 5 complejos de menor energía de cada uno de los 4 complejos principales (**tabla 3**).

**Tabla 3.** Tabla de los mejores complejos mutantes para cada uno de los 4 complejos principales mutados. La columna  $\Delta G$  contiene las energías de interacción de los distintos ligandos. En la columna  $\Delta\Delta G$  está la diferencia entre la energía de interacción del mutante y la energía de interacción del *wild-type*. **A)** AALN1/ASC-PYD. **B)** AALP2/ASC-PYD. **C)** AMLA4/AIM2-PYD. **D)** AMLM2/ASC-PYD.

**A)**

Nombre	Secuencia	$\Delta G$ (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
<b>AALN1 (wt)</b>	GRARDAILDALLEN	2.61	-
AALN1_59.pdb	MKAWMAIWDALLEN	-4.75	-7.36
AALN1_44.pdb	MRAMMAIWDALLEN	-4.74	-7.35
AALN1_89.pdb	MKAMMAIWDALLEN	-4.73	-7.34
AALN1_14.pdb	MRAWMAIWDALLEN	-4.67	-7.28
AALN1_79.pdb	MKAMIAIWDALLEN	-4.42	-7.03

**B)**

Nombre	Secuencia	$\Delta G$ (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
<b>AALP2 (wt)</b>	AEELKKFKLKLLSV	-4.64	-
AALP2_186.pdb	AEELYWFKMRLLMH	-10.38	-5.74
AALP2_207.pdb	AEELYWFKMSLLLR	-10.22	-5.58
AALP2_185.pdb	AEELYWFKMRLLMR	-10.00	-5.36
AALP2_191.pdb	AEELYWFKMQLLMR	-9.84	-5.20
AALP2_209.pdb	AEELYWFKMSLLMR	-9.82	-5.18

C)

Nombre	Secuencia	$\Delta G$ (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
<b>AMLA4 (wt)</b>	DALDLTDKLVLS	1.78	-
AMLA4_162.pdb	DAWLLTMRLVY	-8.02	-9.80
AMLA4_177.pdb	DAWMLTLRLVY	-7.74	-9.52
AMLA4_60.pdb	QAWLLTTRLVY	-7.59	-9.37
AMLA4_63.pdb	QAWMLTKRLVY	-7.06	-8.84
AMLA4_51.pdb	QAWLLTKRLVY	-6.70	-8.48

D)

Nombre	Secuencia	$\Delta G$ (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
<b>AMLM2 (wt)</b>	DRFKFFLSD	-4.45	-
AMLM2_94.pdb	DRRKLHLM	-7.11	-2.66
AMLM2_310.pdb	WRFKLHLM	-7.05	-2.60
AMLM2_307.pdb	WRFKLHLSM	-6.43	-1.98
AMLM2_311.pdb	WRFKLHML	-6.38	-1.93
AMLM2_312.pdb	WRFKLHLMD	-5.99	-1.54

El complejo con menor energía de interacción (mayor estabilidad) derivado de los complejos AALN1, AALP2, AMLA4 y AMLM2 son AALN1\_59, AALP2\_186, AMLA4\_162 y AMLM2\_94 respectivamente. El péptido AALN1\_59 mejora la energía en 7.36 kcal/mol respecto al *wild-type*, el péptido AALP2\_186 la mejora en 5.74 kcal/mol y los péptidos AMLA4\_162 y AMLM2\_94 mejoran la energía de interacción en 9.8 kcal/mol y 2.66 kcal/mol respectivamente. Estas variaciones de energía se deben a los cambios en las interacciones ligando-proteína originados a su vez por los cambios en la secuencia de aminoácidos. En la **tabla 4** se indican las interacciones del ligando mutado con su receptor.

**Tabla 4.** Interacción entre los residuos del ligando y del receptor. En la columna "Ligando" están indicados los aminoácidos que interaccionan con el receptor. En la columna "Receptor" están indicados los aminoácidos que interaccionan con el ligando. En la columna "Interacción" están indicados los tipos de interacción que realizan los aminoácidos del ligando con los del receptor. A) AALN1\_59/ASC-PYD. B) AALP2\_186/ASC-PYD. C) AMLA4\_162/AIM2-PYD. D) AMLM2\_94/ASC-PYD.

A)

**AALN1\_59/ASC-PYD**

Ligando	Receptor	Interacción
E12	K22	Puente de hidrógeno (electrostático)
W4	L25	Hidrofóbica
W8	L25	Hidrofóbica

**B) AALP2\_186/ASC-PYD**

Ligando	Receptor	Interacción
Y5	D51	Puente de hidrógeno (electrostático)
R10	D6	Puente de hidrógeno (electrostático)
Esqueleto	R3	Puente de hidrógeno (electrostático)
Esqueleto	R5	Puente de hidrógeno (electrostático)
W6	L9, L50	Hidrofóbica
M9	L9, L50	Hidrofóbica

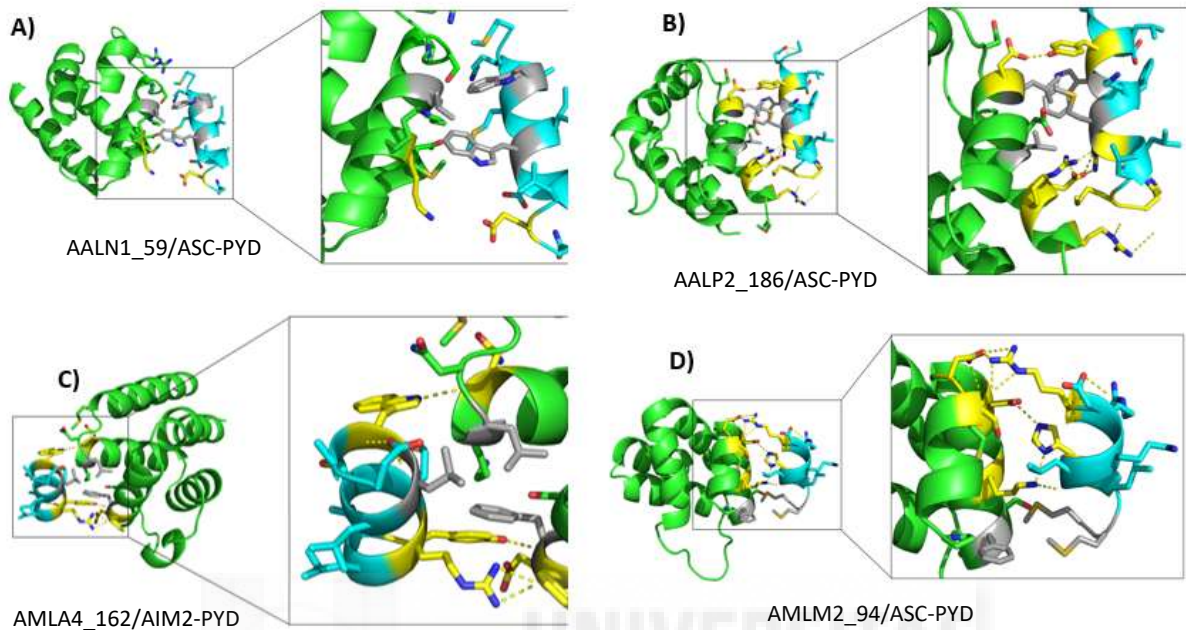
**C) AMLA4\_162/AIM2-PYD**

Ligando	Receptor	Interacción
W3	Esqueleto	Puente de hidrógeno (electrostático)
R8	Esqueleto	Puente de hidrógeno (electrostático)
Y11	Esqueleto	Puente de hidrógeno (electrostático)
L5	F28, L72	Hidrofóbica

**D) AMLM2\_94/ASC-PYD**

Ligando	Receptor	Interacción
H6	D51, K55	Puente de hidrógeno (electrostático)
R2	D48, esqueleto	Puente de hidrógeno (electrostático)
Esqueleto	K55	Puente de hidrógeno (electrostático)
M8	F59	Hidrofóbica
M9	F59	Hidrofóbica

Estas interacciones de los ligandos mutados con sus respectivos receptores, representadas en la **figura 14**, pueden ser clave para evitar que se forme el inflammasoma, ocupando las zonas de interacción de las proteínas activadoras.



**Figura 14.** Interacción entre los receptores y el mejor péptido derivado de cada uno de los 4 complejos iniciales. Las zonas grises representan interacciones hidrofóbicas y las amarillas representan las interacciones electrostáticas. **A)** Complejo AALN1\_59/ASC-PYD. Interacciones electrostáticas: E12 de AALN1\_59 con K22 de ASC-PYD. Interacciones hidrofóbicas: W4 y W8 de AALN1\_59 con L25 de ASC-PYD. **B)** Complejo AALP2\_186/ASC-PYD. Interacciones electrostáticas: Y5 de AALP2\_186 con D51 de ASC-PYD, R10 de AALP2\_186 con D6 de ASC-PYD, esqueleto de AALP2\_186 con R3 y R5 de ASC-PYD. Interacciones hidrofóbicas: W6 y M9 de AALP2\_186 con L9 y L50 de ASC-PYD. **C)** Complejo AMLA4\_162/AIM2-PYD. Interacciones electrostáticas: W3, R8 e Y11 de AMLA4\_162 con el esqueleto de AIM2-PYD. Interacciones hidrofóbicas: L5 de AMLA4\_162 con L28 y L72 de AIM2-PYD. **D)** Complejo AMLM2\_94/ASC-PYD. Interacciones electrostáticas: H6 de AMLM2\_94 con D51 y K55 de ASC-PYD, R2 de AMLM2\_94 con D48 y el esqueleto de ASC-PYD, esqueleto de AMLM2\_94 con K55 de ASC-PYD. Interacciones hidrofóbicas: M8 y M9 de AMLM2\_94 con F59 de ASC-PYD.

En base a los resultados obtenidos y mirando los 5 mejores péptidos en cada uno de los complejos, se han analizado las secuencias obtenidas para acotar lo máximo posible los péptidos propuestos susceptibles de interferir en las interacciones proteína-proteína de los inflammasomas. La **tabla 6** recoge en la columna R2 los aminoácidos seleccionados que han dado lugar a los siguientes péptidos propuestos.

**Tabla 6.** Tabla de combinación de mutaciones. La columna “Posición cristal” indica la posición del aminoácido en el cristal del que se ha derivado. La columna “Posición Matriz” indica la posición del aminoácido en la matriz donde se ha llevado a cabo las mutaciones. La columna “WT” indica los aminoácidos *wild-type* por los que está compuesto el ligando. La columna “R2” indica los aminoácidos que podrían aumentar la afinidad del péptido por el receptor y se han propuestos en función de los 5 mejores complejos mutantes anteriormente mostrados. Las posiciones subrayadas indican las posiciones tolerantes.

AALN1				AALP2			
Posición cristal	Posición Matriz	WT	R2	Posición cristal	Posición Matriz	WT	R2
2	1	G	M	<u>17</u>	<u>1</u>	<u>A</u>	<u>A</u>
3	2	R	RK	<u>18</u>	<u>2</u>	<u>E</u>	<u>E</u>
<u>4</u>	<u>3</u>	<u>A</u>	<u>A</u>	<u>19</u>	<u>3</u>	<u>E</u>	<u>E</u>
5	4	R	WM	<u>20</u>	<u>4</u>	<u>L</u>	<u>L</u>
6	5	D	IM	21	5	K	Y
<u>7</u>	<u>6</u>	<u>A</u>	<u>A</u>	22	6	K	W
<u>8</u>	<u>7</u>	<u>I</u>	<u>I</u>	<u>23</u>	<u>7</u>	<u>E</u>	<u>E</u>
9	8	L	W	<u>24</u>	<u>8</u>	<u>K</u>	<u>K</u>
<u>10</u>	<u>9</u>	<u>D</u>	<u>D</u>	25	9	L	M
<u>11</u>	<u>10</u>	<u>A</u>	<u>A</u>	26	10	K	RSQ
<u>12</u>	<u>11</u>	<u>L</u>	<u>L</u>	<u>27</u>	<u>11</u>	<u>L</u>	<u>L</u>
<u>13</u>	<u>12</u>	<u>E</u>	<u>E</u>	<u>28</u>	<u>12</u>	<u>L</u>	<u>L</u>
<u>14</u>	<u>13</u>	<u>N</u>	<u>N</u>	29	13	S	LM
				30	14	V	RH

AMLA4				AMLM2			
Posición cristal	Posición Matriz	WT	R2	Posición cristal	Posición Matriz.	WT	R2
48	1	D	QD	23	1	D	DW
<u>49</u>	<u>2</u>	<u>A</u>	<u>A</u>	24	2	R	R
50	3	L	W	25	3	F	FR
51	4	D	LM	<u>26</u>	<u>4</u>	<u>K</u>	<u>K</u>
<u>52</u>	<u>5</u>	<u>L</u>	<u>L</u>	27	5	F	L
<u>53</u>	<u>6</u>	<u>I</u>	<u>I</u>	28	6	F	H
54	7	D	KMLT	<u>29</u>	<u>7</u>	<u>L</u>	<u>L</u>
55	8	K	R	30	8	S	SM
<u>56</u>	<u>9</u>	<u>L</u>	<u>L</u>	31	9	D	MLD
<u>57</u>	<u>10</u>	<u>V</u>	<u>V</u>				
58	11	S	Y				

## 6. DISCUSIÓN

Para poder predecir *in silico* el comportamiento de dos proteínas que interactúan entre sí, hay que simplificarlo lo máximo posible. Esto es, en lugar de tener en cuenta toda la proteína y realizar mutaciones sobre ella, se realiza primero una eliminación sobre la proteína en la que se va a trabajar dejando sólo la zona de interacción.

En base a los resultados obtenidos, la energía de interacción de los ligandos mutados es más baja que la de los complejos *wild-type*. Esto sugiere que se han encontrado, mediante procedimientos *in silico*, ligandos con mayor afinidad a las proteínas diana debido a cambios en la secuencia de aminoácidos. Estos cambios han dado lugar a una mayor estabilidad al optimizar los choques estéricos entre aminoácidos como triptófanos, fenilalaninas y prolina, la posición de los aminoácidos hidrofóbicos para establecer las interacciones hidrofóbicas y la repulsión de aminoácidos con misma carga. De esta forma se ha dado una mayor estabilidad a los diferentes complejos.

En el **complejo proteico AIM2-PYD/ASC-PYD**, el dominio AIM2-PYD se une al dominio ASC-PYD para iniciar la oligomerización de dominios ASC-PYD que da lugar al macrocomplejo del inflammasoma. Si los ligandos mutados interactúan con la zona predicha de su dominio correspondiente, cabe esperar que dichos ligandos compitan con las secuencias *wild-type*. Si esto sucede, se interrumpirían las interacciones que se producen de forma natural y se impediría que se formase el inflammasoma, evitando así que se active la inflamación.

### 6.1. Comparaciones entre secuencias de aminoácidos *wild-type* y mutados.

Los cambios significativos en el ligando mutado AALN1\_59 son en las posiciones 4 y 5. Disminuye la polaridad respecto a su péptido wt AALN1 por las transformaciones en las posiciones 4 y 5 al introducir un triptófano y una metionina respectivamente. Además, en la posición 8 se transforma L a W, un aminoácido mucho más grande, más hidrofóbico y de carácter aromático. En AALP2\_186 también hay una reducción en la polaridad por transformaciones en las posiciones 5, 6 y 13, y más concretamente, una eliminación de cargas positivas en las posiciones 5 y 6 al transformarse esas posiciones en Y y W respectivamente. En AMLA4\_162 hay una significativa reducción de la polaridad por introducir L y M en las posiciones 4 y 7. También hay un cambio de L a W, al igual que en AALN1\_59, lo que puede derivar a una interacción específica debido a su tamaño, aunque se trate de una interacción hidrofóbica. Por último, en el péptido AMLM2\_94, si bien aumenta la carga positiva en la posición 3, disminuye la polaridad en las posiciones 8 y 9 al introducir aminoácidos hidrofóbicos, en este caso metioninas, cuya cadena lateral tiene muchos grados de libertad y por lo tanto rellena huecos y establece fuertes interacciones hidrofóbicas. Por otro lado, en la posición 5 hay una reducción de

tamaño al sustituirse una F por una L, lo que probablemente reduce posibles choques y optimice interacciones de van der Waals.

Por tanto, los cambios en las secuencias mutadas han sido mayoritariamente sustituciones de aminoácidos polares por aminoácidos hidrofóbicos, como muestra la **tabla 7**, lo que incrementa considerablemente el número de interacciones hidrofóbicas y reduce el número de interacciones electrostáticas. Esto da como resultado un aumento en la afinidad y una disminución en la especificidad.

**Tabla 7.** Tabla comparativa de cada una de las secuencias *wild-type* con su respectivo ligando mutado de menor energía de interacción. La columna “Pos” indica la posición en la secuencia peptídica. En la columna “WT” está la secuencia peptídica del ligando *wild-type*. En la columna Mut. se encuentra la secuencia derivada del péptido con menor energía de interacción. Las posiciones en negrita son las posiciones que han sufrido cambios de aminoácido.

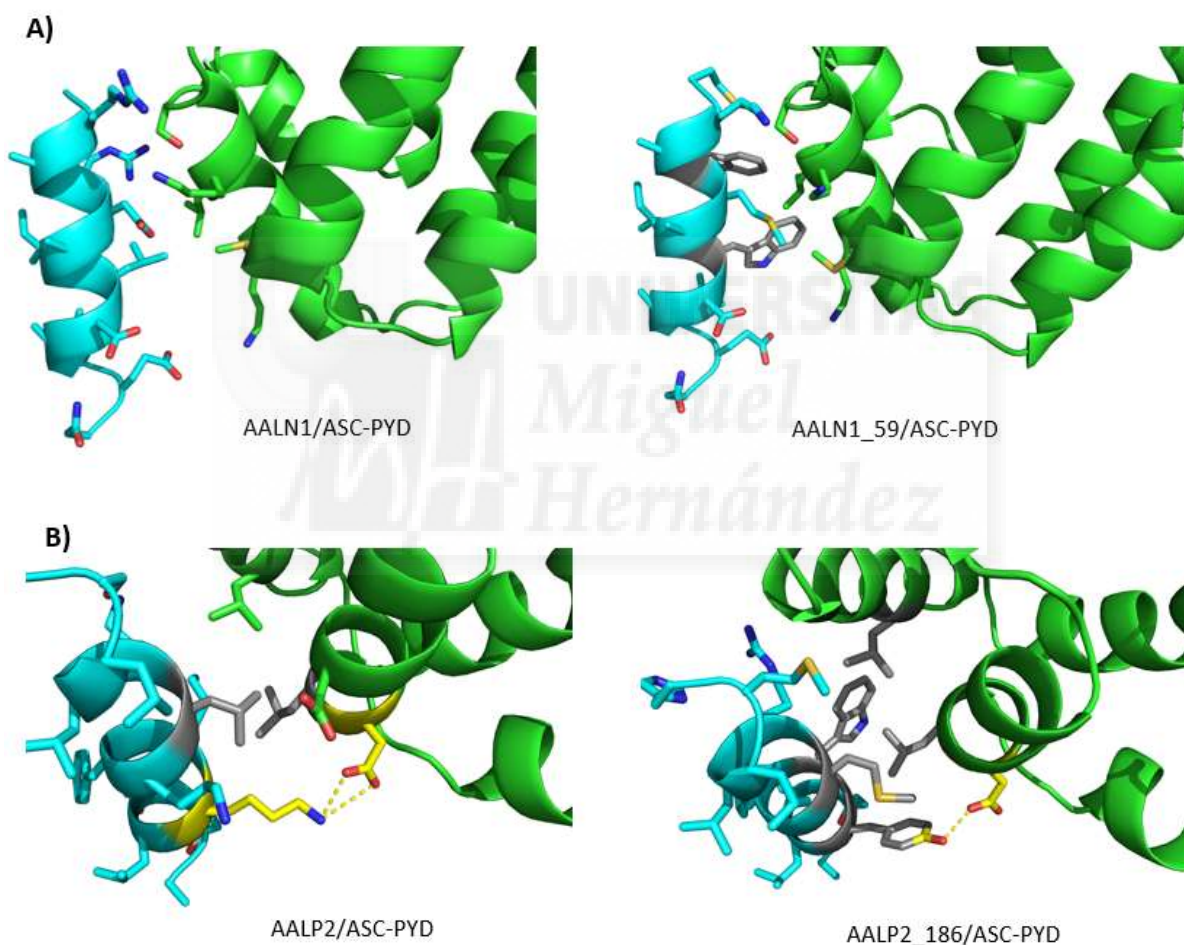
AALN1 vs AALN1_59			AALP2 vs AALP2_186		
Pos.	WT	Mut	Pos.	WT	Mut
<b>1</b>	<b>G</b>	<b>M</b>	1	A	A
<b>2</b>	<b>R</b>	<b>K</b>	2	E	E
3	A	A	3	E	E
<b>4</b>	<b>R</b>	<b>W</b>	4	L	L
<b>5</b>	<b>D</b>	<b>M</b>	5	K	Y
6	A	A	6	K	W
7	I	I	7	F	F
<b>8</b>	<b>L</b>	<b>W</b>	8	K	K
9	D	D	9	L	M
10	A	A	10	K	R
11	L	L	11	L	L
12	E	E	12	L	L
13	N	N	13	S	M
			14	V	H

AMLA4 vs AMLA4_162			AMLM2 vs AMLM2_94		
Pos.	WT	Mut	Pos.	WT	Mut
1	D	D	1	D	D
2	A	A	2	R	R
<b>3</b>	<b>L</b>	<b>W</b>	<b>3</b>	<b>F</b>	<b>R</b>
<b>4</b>	<b>D</b>	<b>L</b>	4	K	K
5	L	L	5	F	L
6	T	T	6	F	H
<b>7</b>	<b>D</b>	<b>M</b>	7	L	L
<b>8</b>	<b>K</b>	<b>R</b>	8	S	M
9	L	L	9	D	M
10	V	V			
<b>11</b>	<b>S</b>	<b>Y</b>			

Los cambios de R4 a W4 y L8 a W8 en el ligando mutado AALN1\_59 son interesantes desde un punto de vista estérico. Estos dos triptófanos quedan orientados hacia el interior, encajados en un bolsillo hidrofóbico, aislados del solvente. Debido a esta característica y a su tamaño, pueden ser clave para una interacción específica ligando-proteína. En la interacción de AALP2\_186 también hay un cambio interesante desde un punto de vista estérico. Al haber mutado L6 a W6, éste también queda orientado hacia el interior, en un bolsillo hidrofóbico. No obstante, hay un cambio de K5 a Y5, que disminuye la interacción electrostática por puente de hidrógeno.

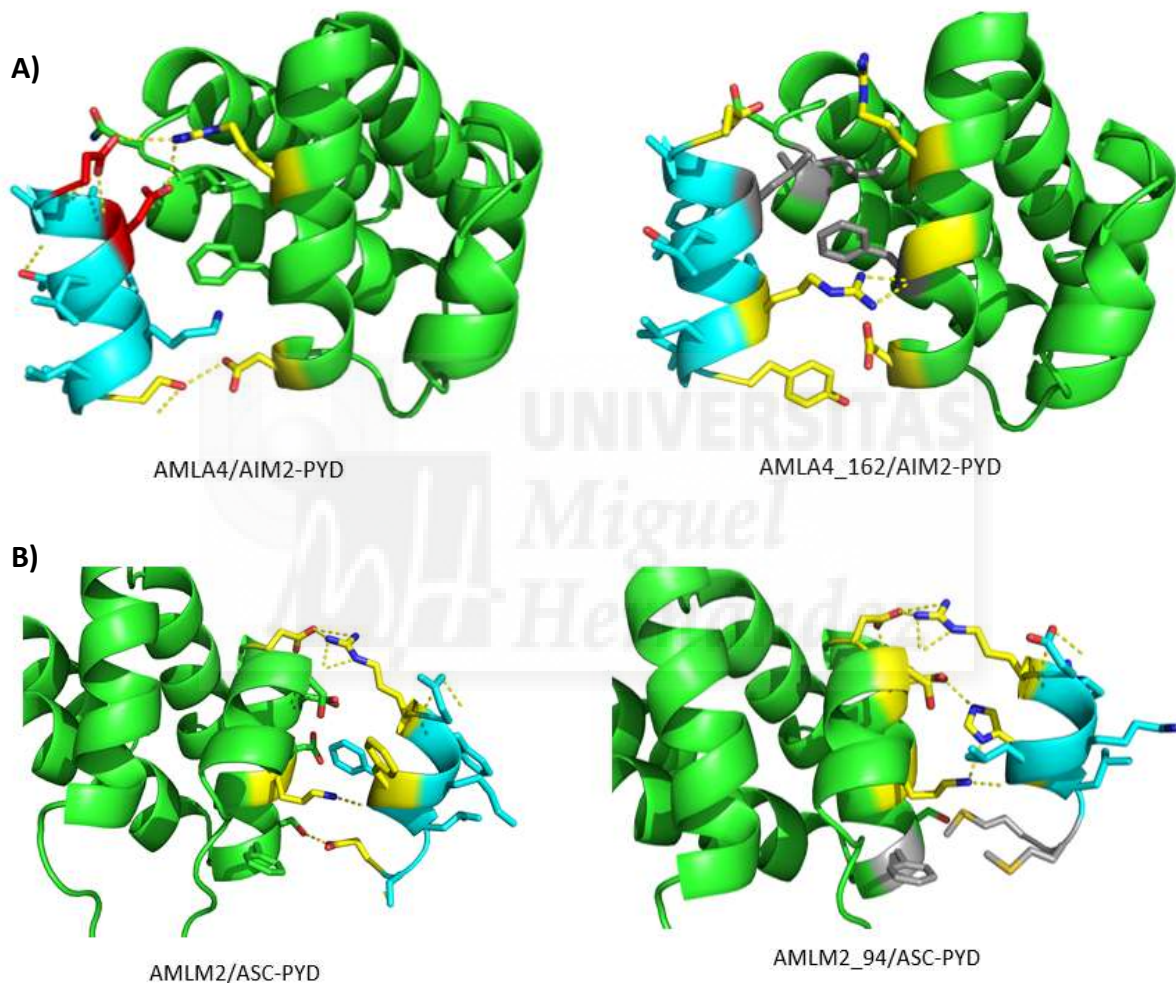
En la **figura 15** se representan todos estos cambios y cómo interaccionan los péptidos mutados de forma distinta con la hélice positiva ( $\alpha 2$ ) y las hélices negativas ( $\alpha 1$  y  $\alpha 4$ ) respectivamente.



**Figura 15.** Comparación entre las interacciones de los ligandos principales (AALN1 y AALP2) y sus ligandos derivados de menor energía de interacción (AALN1\_59 y AALP2\_186). **A)** AALN1 y AALN1\_59. Los cambios de R4 a W4 y L8 a W8 provocan que ambos triptófanos queden orientados hacia el interior, aislados del solvente **B)** AALP2 y AALP2\_186. El triptófanos generado por el cambio de L6 a W6 queda orientado hacia el interior. El cambio de K5 a Y5 permite formar una región hidrofóbica pero disminuye la interacción electrostática por puente de hidrógeno.



La mutación de D4 a L4 en el ligando AMLA4\_186 evita la repulsión electrostática entre las cargas negativas que había previamente entre D1 y D4, lo que mejora la estabilidad del péptido. Además, este cambio aumenta la interacción hidrofóbica con F28, esencial para la interacción AIM2-PYD/ASC-PYD. En el péptido mutado AMLM2\_94 también aumenta el número de interacciones hidrofóbicas por los cambios de S8 y D9 a M8 y M9 respectivamente. No obstante, hay un cambio de F6 a H6 que da lugar al establecimiento de dos puentes de hidrógeno. Todos estos cambios están representados en la **figura 16**.



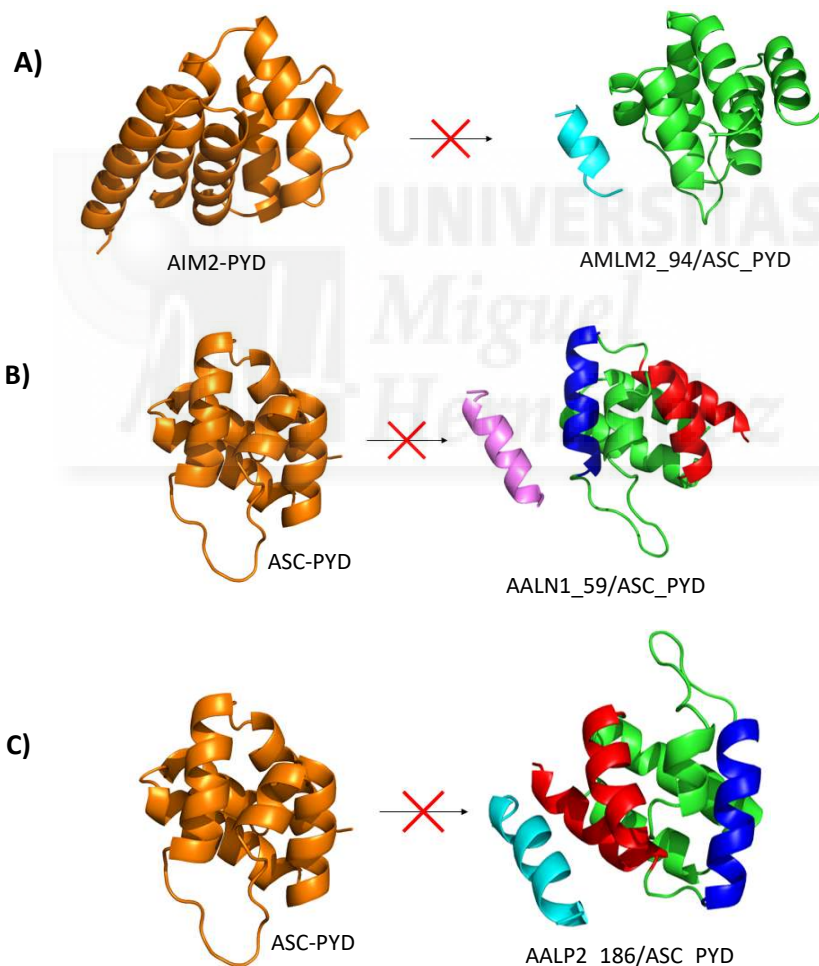
**Figura 16.** Comparación entre las interacciones de los ligandos principales (AMLA4 y AMLM2) y sus ligandos derivados de menor energía de interacción (AMLA4\_162 y AMLM2\_94). **A)** AMLA4 y AMLA4\_162. El cambio de D4 a L4 elimina la repulsión electrostática inicial entre D1 y D4, y aumenta la interacción hidrofóbica con F28. **B)** AMLM2 y AMLM2\_94. Se mantiene el número de interacciones hidrofóbicas y aumenta el número de interacciones electrostáticas por el cambio de F6 por H6 y los cambios de K8 y S11 a R8 e Y11 respectivamente.

## 6.2. Interacción de los ligandos mutados con ASC-PYD

A la luz de los cambios de secuencia mencionados, si se utilizan péptidos derivados de los ligandos AALN1, AALP2, y AMLM2, hipotéticamente éstos se unirán al dominio ASC-PYD impidiendo la interacción de este dominio con el dominio AIM2-PYD o con otros dominios ASC-PYD de tal modo que:

- Los ligandos derivados de **AALN1** se unirán a la hélice positiva del dominio ASC-PYD impidiendo que AIM2-PYD interactúe con ASC-PYD y los derivados de **AALP2** se unirán a las hélices negativas.
- Los ligandos derivados de AMLM2 se unirán a la hélice  $\alpha_4$  y al loop  $\alpha_3$ - $\alpha_4$  de ASC-PYD causando el mismo efecto.

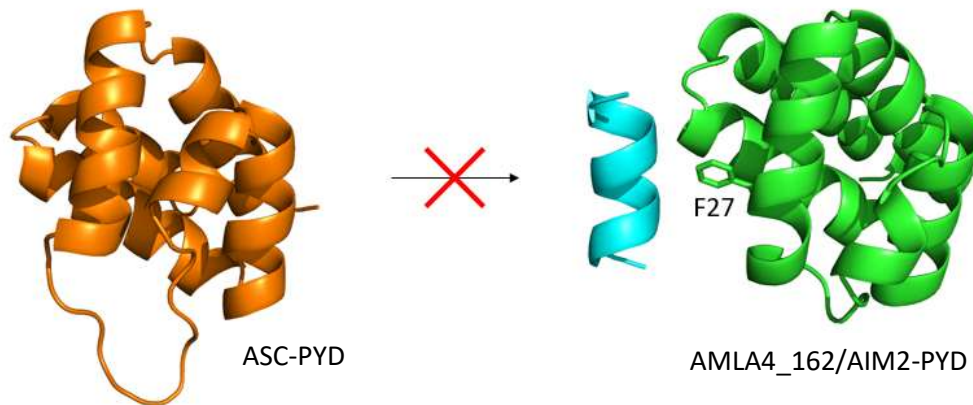
Los secuestros de las proteínas diana se ilustran en la **figura 17**.



**Figura 17.** Ligandos mutados interaccionando con el ligando ASC-PYD. **A)** AMLM2\_94 interacciona con ASC-PYD impidiendo la unión de AIM2. **B)** AALN1\_59 interacciona con la hélice positiva  $\alpha_2$  de ASC-PYD impidiendo la unión de otro ASC-PYD. **C)** AALP2\_186 interacciona con las hélices negativas  $\alpha_1$  y  $\alpha_4$  de ASC-PYD impidiendo la unión de otro ASC-PYD.

### 6.3. Interacción de los ligandos mutados con AIM2-PYD

Si se utilizan péptidos derivados del ligando **AML4**, cabe esperar que éstos se unan al dominio AIM2-PYD bloqueando el acceso al residuo F27, clave para la interacción ASC-PYD/AIM2-PYD, tal como muestra la **figura 18**.



**Figura 18.** Ligando mutado interaccionando con AIM2-PYD. Como el ligando ocupa la zona de interacción, el dominio ASC-PYD no puede unirse al dominio AIM2-PYD.

### 6.4. Comprobación experimental de los resultados

En base a los resultados *in silico* obtenidos, los ligandos mutados tienen una menor energía de interacción (mayor afinidad) por su proteína correspondiente que el *wild-type*, por lo que se espera que se den las interacciones previamente mencionadas desplazando aquellas que se dan de forma natural.

No obstante todo el trabajo se ha realizado *in silico*. Durante la preparación y aplicación de los procedimientos no se ha llevado a cabo ningún tipo de estudio en el laboratorio, por lo que no se tiene la certeza de que los péptidos propuestos compitan con las proteínas *wild-type* durante la formación del inflamasoma. En colaboración con el Dr. José Manuel González Navajas del Centro de Investigación Biomédica en Red (CIBERehd) y Fundación FISABIO, del Hospital General Universitario de Alicante, se investiga abordar la problemática de la activación constitutiva de la inflamación hepática y su posible tratamiento con péptidos miméticos de las regiones de interacción.

Si bien se espera que los péptidos predichos *in silico* interfieran en la formación de los complejos *in vitro*, no hay certeza de que una vez formadas las interacciones entre los distintos dominios, éstas se rompan y el inflamasoma se desarme.

Además, los péptidos suelen ser inestables por administración oral, se suelen degradar una vez alcanzan el estómago, y por lo tanto es difícil que alcancen su diana por la dificultad de los péptidos en penetrar la membrana celular. Así que sería necesario generar estructuras derivadas químicamente de los péptidos o bien buscar sustancias químicas con cierto parecido a los ligandos peptídicos para poder ser administrados por vía oral o intravenosa para tratar las enfermedades hepáticas relacionadas con la inflamación crónica. Un primer paso podría ser el de generar peptidomiméticos, pseudopéptidos y péptidos cíclicos, y a partir de éstos, realizar un segundo paso basado en encontrar moléculas orgánicas pequeñas ya que suelen interactuar con una alta afinidad con la superficie de interacción.

Las ventajas de utilizar peptidomiméticos, pseudopéptidos y péptidos cíclicos residen en que son resistentes a proteasas, por lo que pueden administrarse vía oral y pueden ser buenos candidatos en interferir en las interacciones proteína-proteína que da lugar al ensamblaje del inflammasoma. No obstante, si se derivan moléculas más pequeñas a partir de éstas, las ventajas aumentan ya que las moléculas pequeñas son más fáciles de sintetizar y más económicas (Fernandez-Ballester et al. 2011).



## 7. CONCLUSIÓN Y PROYECCIONES FUTURAS

El objetivo de este trabajo ha sido la predicción *in silico* de péptidos capaces de bloquear ciertas PPIs descritas en la bibliografía. Dichas PPIs son interacciones entre ASC-PYD/ASC-PYD y AIM2-PYD/ASC-PYD que inician la formación del inflamasoma y, para bloquearlos, se han llevado a cabo una serie de estrategias:

1. Se han construido los modelos de interacción.
2. Se ha aislado de cada interacción las regiones peptídicas interaccionantes y se han obtenido un total de 4 ligandos.
3. Con los 4 péptidos obtenidos se han hecho mutaciones *in silico* que han dado lugar a un total de 1248 ligandos derivados.

Todos estos ligandos derivados tienen una energía de interacción menor que los *wild-type*. De cada uno de los 4 ligandos iniciales, se ha escogido el ligando derivado de menor energía (mayor afinidad) y se ha comparado cada uno de ellos con la interacción *wild-type*.

Las comparaciones entre las interacciones de los péptidos mutados y los *wild-type* revelan que hay un incremento en las interacciones hidrofóbicas mientras que disminuyen las interacciones electrostáticas. Esto lleva a deducir que existe un incremento en la afinidad del ligando mutado pero su especificidad por la zona de unión disminuye. En definitiva, se han obtenido unos péptidos más afines a las proteínas responsables de la formación del inflamasoma pero menos específicos, por lo que hay que comprobar de manera experimental, *in vitro*, si estos péptidos se unen a la zona deseada de la proteína e impiden que se produzca la inflamación o, por el contrario, hay que continuar en la búsqueda de otros péptidos que alcancen dicho objetivo.

Es necesario realizar estudios *in vitro* para refinar los datos y confirmar las hipótesis generadas de los trabajos *in silico*. Se espera que en un futuro, estos estudios se lleven a cabo en el hospital universitario de Elche por el Dr. Jose Manuel González Navajas.

## 8. BIBLIOGRAFÍA

Chu, L. H., et al. (2015). "An updated view on the structure and function of PYRIN domains." Apoptosis **20**(2): 157-173.

de Alba, E. (2009). "Structure and interdomain dynamics of apoptosis-associated speck-like protein containing a CARD (ASC)." J Biol Chem **284**(47): 32932-32941.

Guo, H., et al. (2015). "Inflammasomes: mechanism of action, role in disease, and therapeutics." Nat Med **21**(7): 677-687.

Jin, T., et al. (2012). "Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor." Immunity **36**(4): 561-571.

Jin, T., et al. (2013). "Structure of the absent in melanoma 2 (AIM2) pyrin domain provides insights into the mechanisms of AIM2 autoinhibition and inflammasome assembly." J Biol Chem **288**(19): 13225-13235.

Liepinsh, E., et al. (2003). "The death-domain fold of the ASC PYRIN domain, presenting a basis for PYRIN/PYRIN recognition." J Mol Biol **332**(5): 1155-1163.

Lu, A., et al. (2014). "Crystal structure of the F27G AIM2 PYD mutant and similarities of its self-association to DED/DED interactions." J Mol Biol **426**(7): 1420-1427.

Lu, A., et al. (2014). "Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes." Cell **156**(6): 1193-1206.

Vajjhala, P. R., et al. (2012). "Multiple binding sites on the pyrin domain of ASC protein allow self-association and interaction with NLRP3 protein." J Biol Chem **287**(50): 41732-41743.

Haitao, J. B., J. P., et al. (2015). "Inflammasomes: mechanism of action, role in disease, and therapeutics". Nature Medicine **21**, 677–687.

Fernandez-Ballester et al. (2011). "In silico-based direct evolution of peptides and peptidomimetics in drug discovery". Current Topics in Pharmacology. 15, 2. 35-55.