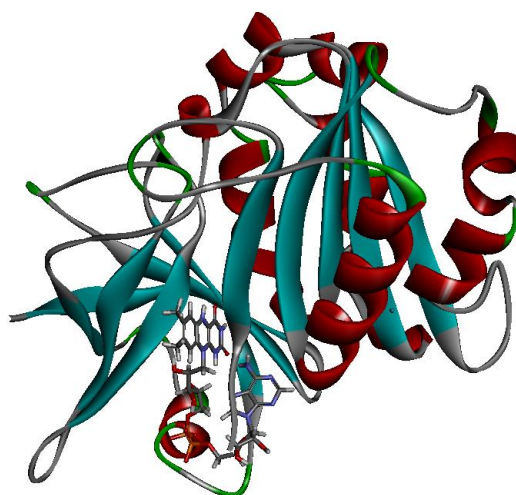


TRABAJO DE FIN DE GRADO

Universidad de Zaragoza

Grado en Química, Curso 2015/2016

Departamento de Bioquímica y Biología Molecular y Celular



Simulación computacional para la predicción de  
acoplamiento enzima-inhibidor: identificación del sitio de  
unión de inhibidores de flavoenzimas esenciales en  
patógenos

Acoplamiento de análogos del sustrato a la flavodoxina-NADP<sup>+</sup>  
reductasa de *Xanthomonas axonopodis* pv. *citri*

Autor: Jonás Pelegrín Costea

Directores: Milagros Medina Trullenque, Víctor Polo Ortiz

## **Agradecimientos**

Quisiera agradecer a los compañeros e integrantes del Departamento de Química Física Teórica de la Universidad su apoyo moral y técnico prestado a lo largo del curso, así como el ofrecimiento por parte de dicho departamento de los medios necesarios para llevar a cabo el trabajo.

Y por supuesto, agradecer a Milagros y Víctor, directores de este trabajo, tanto la oferta del mismo como su labor e implicación en él, pues sin ellos nada de esto habría sido posible. Gracias por estar presentes en este punto de inflexión de mi vida académica.

## Índice

1. Resumen/Abstract .....	1
2. Introducción .....	2
3. Objetivos y planteamiento del trabajo .....	7
4. Metodología y fundamentos .....	8
4.1 Superficies de energía Potencial .....	8
4.2 Mecánica Molecular .....	9
4.3 Campos de Fuerza .....	10
4.4 Acoplamiento Molecular .....	10
4.5 Software utilizado .....	11
5. Resultados y discusión .....	12
5.1 Identificación del sitio de unión del sustrato a partir de proteínas análogas..	12
5.2 Acoplamiento molecular de cuerpo rígido. Interacción de los residuos de <i>XacFPR</i> con el sustrato .....	15
5.3 Acoplamiento molecular flexible .....	18
6. Conclusiones y perspectivas.....	22
7. Bibliografía .....	23



## 1. Resumen

*Xanthomonas axonopodis* pv. *citri* es una bacteria responsable del cáncer de cítricos, una enfermedad que afecta a la calidad de estos. El genoma de este patógeno contiene el gen que codifica la enzima Flavodoxina-NADPH Reductasa (*XacFPR*), perteneciente a la familia de la Flavodoxina-NADPH Reductasa (FNR). A su vez, la *XacFPR* se clasifica en la subclase Ia de las FPRs, caracterizada por presentar una alanina enfrentada a la isoaloxacina del cofactor FAD de la proteína y una extensión más corta del dominio C-terminal con respecto a las FPRs de la subclase Ib.

*XacFPR* está involucrada en la respuesta a estrés oxidativo inducida por la planta para combatir al patógeno, por lo que la inhibición de su actividad puede ser un aspecto clave para frenar la supervivencia del patógeno. Por ello resulta de interés identificar cómo los posibles inhibidores interactuarían con la *XacFPR*. Sin embargo, teniendo en cuenta que hasta la fecha no existen modelos tridimensionales que representen la interacción de *XacFPR* con su sustrato, el coenzima NADPH, se hace necesario entender como dicho sustrato se une a esta enzima.

El objetivo que se planteó en el presente trabajo fue utilizar herramientas computacionales para simular la interacción de los componentes nucleotídicos del coenzima NADPH con *XacFPR*. Para ello realizamos el acoplamiento molecular entre la proteína y fragmentos del sustrato. Esta aproximación permite evaluar la energía de interacción entre distintas posiciones relativas de las dos moléculas, para así obtener aquella o aquellas posiciones más favorables.

## 1. Abstract

*Xanthomonas axonopodis* pv. *citri* is the pathogenic bacterium responsible for the severe citrus canker disease that affects citrus quality. The genome of this pathogen contains the gene that encodes a Flavodoxin-NADPH reductase (*XacFPR*) that belongs to the Ferredoxin-NADPH reductase (FNR) family. Within this family *XacFPR* is classified into the FPRs subclass Ia, characterized by a alanine facing the isoalloxazine ring of enzyme's FAD cofactor and a shorter C-terminal sequence when compare with the subclass Ib FPRs.

*XacFPR* is involved in the response of the pathogen to oxidative stress induced by the plant upon injection, so inhibition of *XacFPR* activity might be a key factor for preventing the pathogen survival. In this context, it is of interest to identify how potential inhibitors could bind to *XacFPR*. However, taking into account that at present there are not three-dimensional structural models for the interaction of *XacFPR* with its substrate, the NADPH coenzyme, it is necessary to start by identifying the key aspect of the binding of this substrate to *XacFPR*.

Thus, the aim of this work was to make use of computational tools to simulate the interaction between the nucleotide components of the NADPH coenzyme and *XacFPR*. For this purpose we docked the protein with different fragments of the substrate. This approach allows to evaluate interactions energies between several relative molecular poses of the substrate onto the protein and get the most favourable ones.

## 2. Introducción

*Xanthomonas axonopodis* pv. *citri* es una bacteria Gram-Negativa responsable del cáncer de cítricos, una enfermedad que provoca grandes pérdidas económicas a la industria de cítricos en todo el mundo, debido principalmente a la reducción de producción de frutos y al coste para su prevención y control. El patógeno penetra en la planta huésped a través del estoma y coloniza el apoplasto, mostrando puntos circulares en la superficie de las hojas a causa de la infección. Esto desemboca en lesiones necróticas en las hojas, tallos y frutas (figura 1).

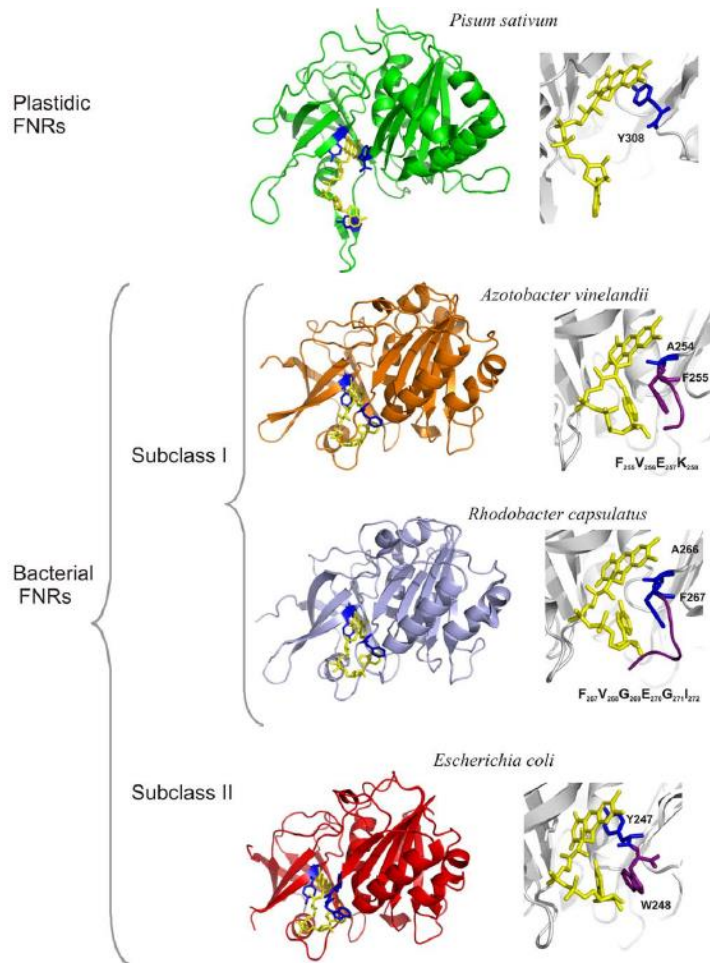


**Figura 1.** Síntomas del cáncer de cítricos en frutos y hojas infectados por la bacteria *Xanthomonas axonopodis* pv. *citri*

El metabolismo de las plantas cambia cuando se ve expuesto a patógenos, por ejemplo, produciendo especies reactivas de oxígeno (ROS), por lo que el patógeno se ve obligado a prevenir el estrés oxidativo generado para estabilizar y mantener la infección. Es aquí donde toman importancia las enzimas Flavodoxina-NADPH Reductasa (FPR) que contribuyen a la respuesta ante el estrés oxidativo en bacterias Gram-Negativas como *Escherichia coli* o *Pseudomonas putida*.<sup>1,2</sup>

Las FPRs pertenecen a la familia de la Ferredoxina-NADPH Reductasa (FNR). Las enzimas de esta familia son flavoenzimas redox presentes en numerosos organismos e involucradas en varias funciones metabólicas como fotosíntesis, asimilación de nitrógeno o desaturación de ácidos grasos. El grupo FAD permite a estas proteínas catalizar la transferencia electrónica entre el coenzima NADP(H) y proteínas de óxido-reducción como ferredoxinas, flavodoxinas o hemoxigenasas. Este cofactor está unido al dominio N-terminal de la proteína, mientras que el dominio C-terminal presenta el sitio de unión para el coenzima NAPD(H).

Las FNRs se clasifican en dos grandes grupos (figura 2) según sus características estructurales y filogenéticas: FNRs plastídicas (presentes en tejidos fotosintéticos; exhiben una alta eficiencia catalítica) y FNRs bacterianas (denominadas FPRs; tienen baja eficiencia catalítica). Dentro de estas últimas podemos encontrar otros dos grupos, la subclase I (a la que pertenece la FPR de *Rhodobacter Capsulatus*) y la subclase II (representada por la FPR de *Escherichia coli*). La principal diferencia entre ambas subclases se encuentra en la región del carboxilo terminal, donde la *E. coli* FNR presenta una tirosina enfrentada a la isoaloxacina del FAD, mientras que las enzimas de la subclase I tienen una alanina en dicha posición y una extensión más larga del C-terminal.

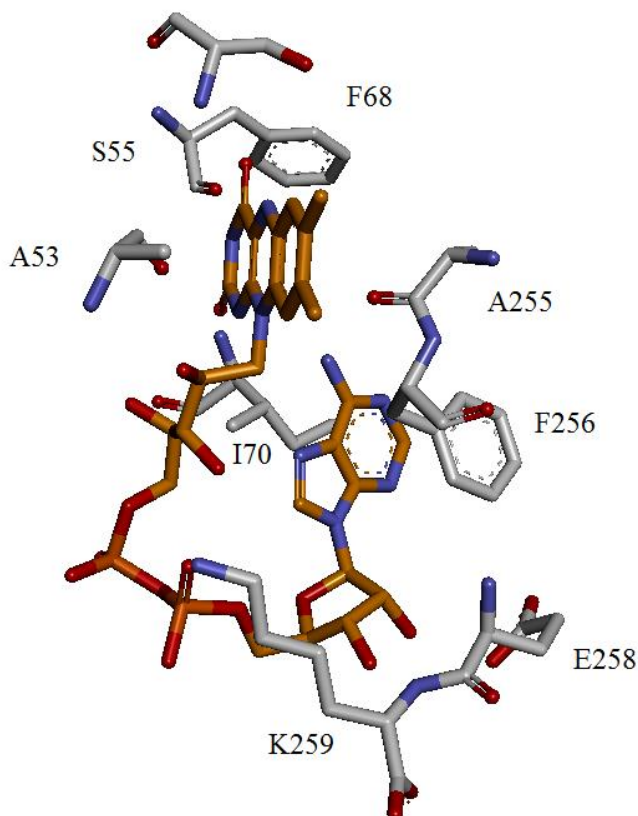


**Figura 2.** Clasificación de las FPRs según sus características estructurales. (Tondo et al., 2011)

El genoma de *Xanthomonas axonopodis* pv. citri contiene el gen que codifica la FPR (*XacFPR*), enzima involucrada en la respuesta ante el estrés oxidativo provocado por el organismo contra el patógeno, siendo esta enzima clave para la supervivencia del patógeno. Aunque esta enzima se asemeja mucho estructuralmente a las FPR de subclase I, su caracterización estructural reveló diferencias en el entorno de la isoaloxacina con respecto a la de *RcFPR*, por lo que se sugirió una subdivisión de la subclase I en subclase IA y IB. La subclase IA de FPRs, a la que pertenece la *XacFPR* posee una secuencia VEK C-terminal más corta, que permite a la isoaloxacina ser más accesible al solvente. En la figura anterior aparece la FPR del *Azotobacter vinelandii*, perteneciente también a la subclase IA y estructuralmente similar a la *XacFPR*.

La estructura de *XacFPR* se pliega en los dos dominios comunes a toda la familia de las FNR. El dominio N-terminal (residuos 4-98), donde se une el cofactor FAD como ya se ha indicado anteriormente, presenta 6 hojas  $\beta$  antiparalelas ( $\beta$ 1- $\beta$ 6) y una hélice  $\alpha$ 1. El dominio C-terminal (residuos 99-259) adopta un plegamiento típico de unión a nucleótidos con cinco láminas  $\beta$  paralelas rodeado por 9 hélices  $\alpha$ . El espacio entre estos dos dominios está ocupado por la hélice  $\alpha$ 3 del dominio C-terminal y con residuos del dominio N-terminal correspondientes a las hélices que conectan  $\beta$ 2 con  $\beta$ 3 y  $\beta$ 4 con  $\beta$ 5.<sup>3</sup>

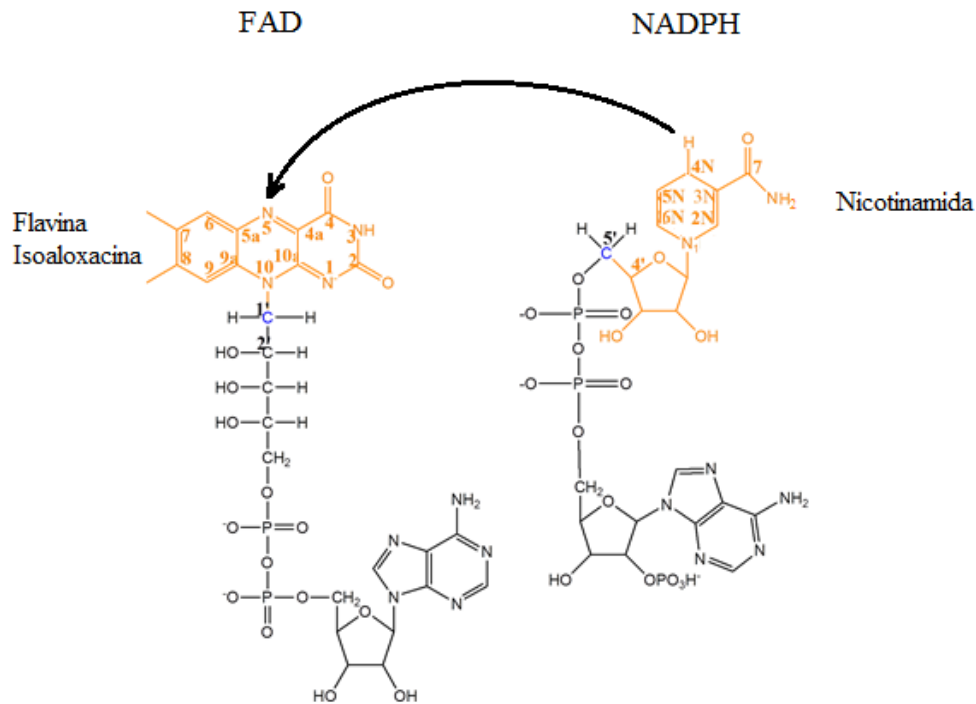
Con respecto con el entorno del FAD (figura 3), parece ser que el apilamiento del anillo de la adenina contra las cadenas laterales de Phe256 e Ile70 mantiene una conformación curvada que difiere de la disposición de las FNRs plastídicas, colocando los anillos de adenina e isoaloxacina próximos entre ellos, además de interacciones electrostáticas e hidrofóbicas adicionales. Los átomos O4 y O2 del anillo de isoaloxacina interaccionan con las amidas de Ser55 e Ile70, mientras que los átomos N3 y N10 están estabilizados por los átomos de Oxígeno de Phe68, Ala53 y Ala255. Los dos últimos residuos del C-terminal, Glu258 y Lys259, estabilizan la posición de la fracción ribitol del cofactor a través de interacciones polares, mientras que la cadena lateral de la Lys259 crea un puente salino con uno de los fosfatos del FAD.<sup>3,4</sup>



**Figura 3.** Entorno del FAD en *XacFPR*. En naranja (átomos de carbono) se representa el FAD, y en blanco los aminoácidos que fundamentalmente contribuyen a la interacción con la coenzima.

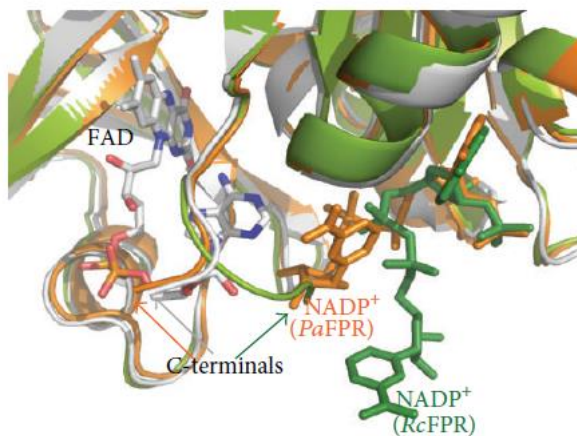
Durante el proceso de reducción de *XacFPR* por NADPH se espera una transferencia de hidruro directa desde el C4 del anillo de nicotinamida del NADPH al N5 del anillo de isoaloxacina de la flavina (figura 4). Para que esta transferencia se produzca de forma óptima los anillos de nicotinamida y flavina deben enfrentarse entre sí, aproximando sus posiciones C4 y N5 sin ningún impedimento entre ellas. Como ya se ha comentado, en *XacFPR* el FAD se encuentra en una conformación plegada con los anillos de adenina e isoaloxacina próximos entre sí, lo que supone un primer problema para la transferencia del hidrógeno ya que se tendrán que vencer las interacciones que confieren estabilidad a esta disposición para que la coenzima se despliegue, permitiendo el acceso de la nicotinamida del coenzima al anillo de flavina y así la transferencia. Además, los residuos del extremo carboxilo terminal del dominio C-terminal de las FPRs de subclase I también dificultan la entrada del anillo de nicotinamida del NADPH al sitio activo, por lo que es de esperar un reordenamiento conformacional para que tenga lugar la transferencia de hidruro.<sup>5</sup>





**Figura 4.** Estructuras de FAD y NADPH, indicando los átomos entre los que tiene que tener lugar la transferencia de hidruro.

En la figura 5 se representa el entorno del FAD de *XacFPR* (gris) y la disposición del C-terminal para los complejos de la FPR del *Pseudomonas aeruginosa* con NADP<sup>+</sup> (*PaFPR*:NADP<sup>+</sup>, naranja) y *RcFPR*:NADP<sup>+</sup> (verde), así como la conformación que adopta el sustrato en cada caso. Se observa que en el complejo *RcFPR*:NADP<sup>+</sup> el anillo de nicotinamida sobresale hacia el solvente, dificultando la transferencia de hidruro en dicha conformación. En cambio en el complejo *PaFPR*:NADP<sup>+</sup> la nicotinamida adopta una conformación más cercana al anillo de isoaloxacina, debido probablemente a que su extremo C-terminal es más corto que el de la *RcFPR*. No obstante, este complejo tampoco muestra un acoplamiento catalíticamente competente entre la nicotinamida y el anillo de isoaloxacina para que tenga lugar la transferencia de hidruro.<sup>4</sup>



**Figura 5.** Superposición *cartoon* del sitio activo de *XacFPR* (gris), del complejo *PaFPR*:NADP<sup>+</sup> (naranja) y del complejo *RcFPR*:NADP<sup>+</sup> (verde). (Tondo et al., 2013)

A pesar de la gran homología estructural entre estas enzimas, *XacFPR* presenta diferencias significativas; tiene una menor longitud de la región C-terminal en comparación con *RcFPR*, haciendo su anillo de isoaloxacina más accesible al solvente. Como ya se ha indicado, las enzimas de la subclase I tienen una alanina en la posición equivalente a la tirosina que queda enfrentada al anillo de isoaloxacina en las de subclase II o en las enzimas plastídicas. Por tanto, la ausencia de este residuo aromático en las enzimas de subclase I exige cambios conformacionales en el sitio activo diferentes a los esperados en la subclase II o los enzimas plastídicos para alcanzar el complejo catalíticamente competente. Esto podría explicar la baja eficiencia como reductoras de las FPR comparadas con las enzimas plastídicas.

### 3. Objetivos y planteamiento del trabajo

La inhibición de la actividad reductasa de la *XacFPR* podría ser una cura efectiva del cáncer de cítricos causado por *X. axonopodis* pv. citri., por lo que en el Departamento de Bioquímica y Biología Molecular y Celular de nuestra Universidad se está investigando en la búsqueda de compuestos inhibidores de esta flavoenzima. En este proceso resulta de interés identificar cómo los posibles inhibidores interaccionarían con la proteína, para lo cual podrían obtenerse modelos mediante métodos computacionales de acoplamiento molecular. Sin embargo, antes es necesario testear cómo funcionan estos métodos en el centro activo de la enzima, y, ya que no existen modelos tridimensionales para la interacción de *XacFPR* con su sustrato fisiológico, el coenzima NADPH, es de interés describir primero su unión. Sin embargo, el estudio de la interacción de este compuesto presenta algunos aspectos limitantes en la modelización computacional por métodos de acoplamiento molecular, ya que como hemos visto anteriormente se requiere una reestructuración conformacional del centro catalítico de la enzima para la formación del complejo catalíticamente competente. Para solventarlo recurriremos a compartimentar el problema, y comenzar por estudiar la interacción de la unión de la mitad correspondiente al nucleótido de adenina del coenzima NADPH con *XacFPR*.

A continuación, se exponen los objetivos perseguidos en este trabajo:

- Familiarizarse y conocer los aspectos más importantes de las herramientas informáticas utilizadas en el posterior estudio computacional.
- Identificar el sitio de unión del sustrato mediante el estudio de proteínas análogas a *XacFPR*, con objeto de decidir cuál es la zona de interés para el estudio de acoplamiento molecular y dónde se va a requerir una reorganización de la proteína.
- Simular el acoplamiento molecular de cuerpo rígido entre *XacFPR* y 2'P-AMP y 2'P-ADP, e identificar las regiones de la proteína implicadas en el acoplamiento y aquellas que se espera que contribuyan con cambios conformacionales.
- Simular el acoplamiento molecular flexible entre *XacFPR* y 2'P-AMP y 2'P-ADP.

## 4. Metodología y fundamentos

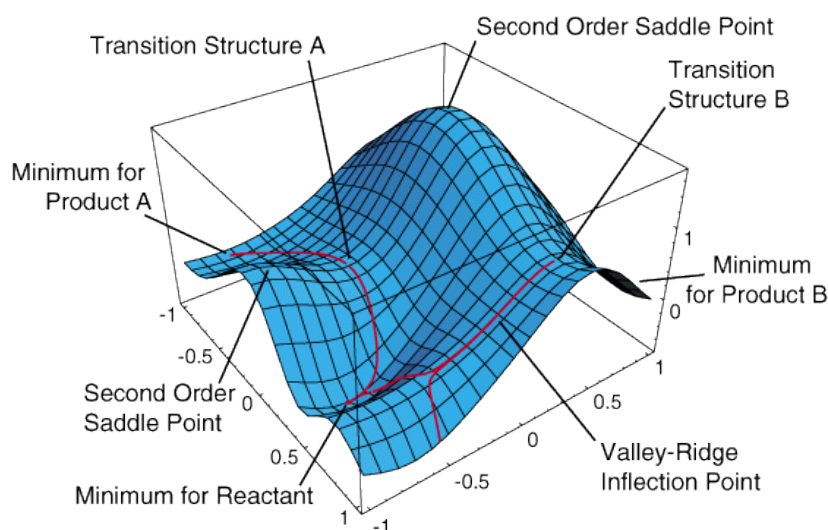
En este apartado se comentarán brevemente y de manera descriptiva los conceptos y fundamentos necesarios para la comprensión del trabajo, además de los métodos teóricos aplicados a los estudios desarrollados. Empezaremos introduciendo las superficies de energía potencial, para después poder abordar de manera más sencilla los métodos de mecánica molecular. Posteriormente introduciremos qué es el acoplamiento molecular y finalizaremos hablando sobre el programa utilizado para la realización de éste.

### 4.1 Superficies de energía potencial

Uno de los conceptos más importantes en modelización molecular es el de Superficie de Energía Potencial (SEP o, por sus siglas en inglés, PES), ya que es fundamental para explicar las propiedades de las moléculas o de un sistema molecular, así como su estabilidad, estructura o reacciones químicas en las que interviene.

La SEP deriva directamente de la aproximación de Born-Oppenheimer. Al resolver la ecuación de Schrödinger electrónica se halla la energía para una determinada configuración nuclear (energía potencial). Se conoce como hipersuperficie de potencial a ese valor de energía para todas las configuraciones nucleares posibles, pero la elevada dimensionalidad de la misma es el principal problema de la hipersuperficie de potencial, ya que se precisaría un número enorme de cálculos electrónicos para calcularla. La única alternativa consiste en diseñar estrategias que permitan evaluar el cálculo completo limitándose a analizar las zonas de interés químico.<sup>6</sup>

Los puntos de mayor interés dentro de la superficie de energía potencial (Figura 6) son los que representan estados estables o metaestables, como lo son los reactivos, productos o estructuras de transición. Estos puntos serán máximos y mínimos energéticos de la SEP, es decir, la primera derivada de la energía potencial será igual a cero.



**Figura 6.** Ejemplo de superficie de energía potencial y sus puntos de mayor interés.

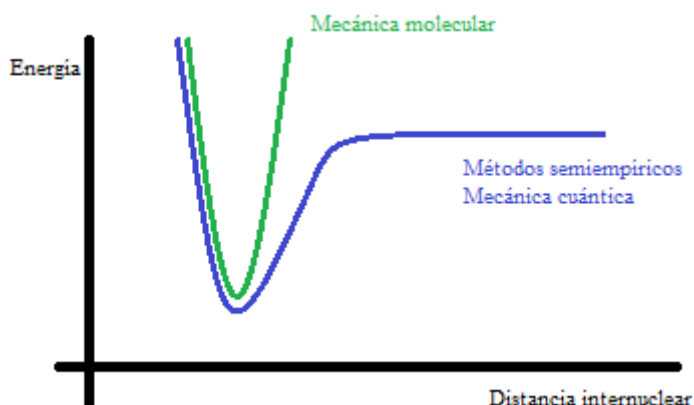
El camino para una reacción específica será el camino de menor energía que conecte a los reactivos con los productos, pasando por un punto de silla. Un punto silla o de ensilladura es un punto estacionario de la SEP que es un mínimo en todas las direcciones excepto en una; en la dirección que conecta reactivos con productos será un máximo.

## 4.2 Mecánica molecular

La energía potencial y la configuración nuclear de un estado en una reacción química son los dos aspectos clave para el estudio de la misma. La forma más exacta de evaluar teóricamente la energía de los átomos o moléculas es mediante el uso de mecánica cuántica (QM), pero el uso de métodos QM para el tratamiento de sistemas complejos, como por ejemplo proteínas, hoy en día es computacionalmente inviable. Para el estudio de estos sistemas es útil el uso de métodos de mecánica molecular (MM).

Los métodos MM requieren menos poder de cómputo debido a que usan funciones de energía potencial parametrizadas a partir de sistemas moleculares más simples, y representan las moléculas como conjuntos de esferas (átomos), cada una de las cuales tendrá masa, carga y otras propiedades del átomo que representan como radio y volumen, y se conectan entre sí por medio de muelles. Son métodos empíricos y sus cálculos están basados en un modelo mecano-clásico, además aplican la aproximación de Born Oppenheimer según la cual el estado de la molécula depende de los núcleos (consideran la energía de interacción de núcleos), pero solo consideran a los electrones de manera implícita, por lo que no permiten describir la formación y/o ruptura de enlaces. Así, los métodos basados en MM no proporcionan una descripción de la estructura electrónica del sistema calculado.<sup>5</sup>

Los métodos de mecánica molecular no son tan precisos como los de mecánica cuántica o semiempíricos, pero en la zona de mínimos hay una equivalencia entre estos métodos, por lo que el potencial real y el calculado por un método MM serán muy parecidos (figura 7). Esta metodología es ideal para sistemas con un elevado número de átomos, que impiden utilizar métodos QM debido al gran número de conformeros a considerar (como ocurre en los sistemas biológicos) y es necesario considerar explícitamente el disolvente ya que no se puede utilizar la aproximación de fase gas o un modelo de disolvente continuo. Por lo tanto, la metodología MM es ideal para el estudio de sistemas biológicos, concretamente en procesos de reconocimiento molecular o dinámica de la proteína. En cuestiones de catálisis enzimática, estos métodos no se pueden utilizar para el cálculo de energías de activación, existiendo alternativas como la combinación de ambos QM/MM.



**Figura 7.** Ejemplo de curvas de potencial obtenidas con métodos MM y QM en función de la distancia internuclear.

### 4.3 Campos de Fuerza

En simulaciones de mecánica molecular se utilizan un conjunto de parámetros y ecuaciones que componen el denominado campo de fuerzas (*force fields*). Este concepto deriva de la hipótesis de que los enlaces tienen longitudes y ángulos de equilibrio a los que las moléculas se ajustan, y una desviación de esos valores implica un aumento de la energía del sistema. Así pues, se considera que los núcleos están sometidos a un campo de fuerzas que se compone la suma de los términos asociados a las tensiones de los enlaces, ángulos y torsiones, así como las interacciones de Van der Waals, electrostáticas y por formación de puentes de hidrógeno. Los campos de fuerza más recientes incluyen también términos específicos como son las interacciones por enlace de halógeno o efectos electrónicos debidos a la deslocalización electrónica.<sup>7</sup>

Los términos energéticos pueden dividirse en tres grupos. El primero de ellos, denominado campo de fuerza de valencia (*valence force field*) comprende aquellos términos que dependen directamente de los enlaces químicos, como la tensión, flexión o torsión. El segundo grupo incluye las interacciones a larga distancia, como las de van der Waals o electrostáticas. El último grupo se encarga de los términos entrecruzados, ya que en muchas ocasiones existen interacciones entre diferentes términos (por ejemplo, cuando un ángulo de enlace se acorta, las distancias del mismo tienden a aumentar para que los átomos no queden tan próximos entre sí).

Los primeros campos de fuerza fueron el EAS (1973), desarrollado por Engler, Aldose y Schleyer, y el MM1, propuesto por Allinger en 1974, quien más tarde mejoró su método con nuevas versiones. Hoy en día existen muchos tipos de campos de fuerza para sistemas biológicos, como el MM2, MM3, Amber o CHARMM entre otros, y para otros tipos de sistemas de interés como zeolitas o similares. Aunque los resultados obtenidos por los campos de fuerza son cualitativamente correctos o buenos en algunas propiedades específicas, su precisión cuantitativa está lejos de ser obtenida de manera correcta debido a problemas de fondo, como el de que las moléculas descritas no son polarizables (cada átomo posee una carga parcial fija, independiente de su entorno) o la ausencia de efectos colectivos.

### 4.4 Acoplamiento molecular

En el ámbito en el que se desarrolla este trabajo, el objetivo principal consiste en predecir la fortaleza y la especificidad con la que una molécula pequeña (como un sustrato) es capaz de unirse a una macromolécula biológica. Una de las posibles aplicaciones de estas predicciones es el desarrollo de nuevos sustratos capaces de inhibir enzimas, entre otras. El proceso de acoplamiento molecular permite evaluar la energía de interacción entre dos o varias posiciones relativas entre dos especies (ligando-receptor) utilizando métodos de la química computacional. Es muy utilizado en la etapa de reconocimiento molecular de diseño de fármacos, en la cual se puede emplear la mecánica molecular ya que, dado un tamaño total del sistema y un gran número de disposiciones relativas del fármaco con respecto del enzima, resultan los métodos de trabajo más rápidos y eficaces.<sup>7</sup>

La predicción de acoplamiento molecular se realiza en dos etapas denominadas “docking” y “scoring”. En la etapa de “docking” se generan un gran número de posibles configuraciones ligando-proteína, llamadas “poses” (posiciones). Primero, se identifican las cavidades de la proteína en la estructura nativa mientras que en el ligando se realiza un proceso de análisis conformacional para obtener los posibles conformeros del mismo. A continuación, se identifican aquellas posiciones en las cuales el ligando encaja satisfactoriamente en la cavidad propuesta,

utilizando criterios estrictamente geométricos. Esto se puede realizar considerando la proteína fija (o rígida) en su estructura nativa o ligeramente relajada (o flexible), lo cual aumenta considerablemente el número de posiciones válidas y el tiempo de cálculo de las mismas.

En la etapa de “scoring” se estima la afinidad entre receptor y ligando para cada una de las posiciones generadas en el “docking” utilizando diversas funciones que se pueden clasificar en tres familias: basadas en campos de fuerza, basadas en el conocimiento del sistema y las empíricas. Las funciones basadas en campos de fuerza calculan la energía libre de interacción “binding energy” proteína-ligando mediante una suma de interacciones de van der Waals y electrostáticas utilizando un campo de fuerzas dado. Los efectos del solvente se introducen mediante una función dieléctrica dependiente de la distancia. Las funciones de “scoring” basadas en el conocimiento representan la “binding energy” sumando interacciones entre pares de átomos de la proteína y el ligando. Los potenciales de interacción se obtienen de complejos proteína-ligando ya conocidos. Por último, las funciones de “scoring” empíricas calculan la energía de interacción mediante una suma de términos derivados de datos experimentales, en los cuales se tienen en cuenta varias contribuciones fundamentales: enlaces de hidrógeno, interacciones iónicas y lipofílicas y contribuciones de carácter entrópico, entre otras. Este último método es el empleado por el programa LIBDOCK utilizado en este trabajo.

La etapa de “scoring” por lo tanto resulta clave en la predicción del acoplamiento enzima-inhibidor, por ello se recomienda utilizar más de una función para que una posición se considere favorable y así se reduce el riesgo de obtener falsos positivos.<sup>8,9</sup>

#### 4.5 Software utilizado

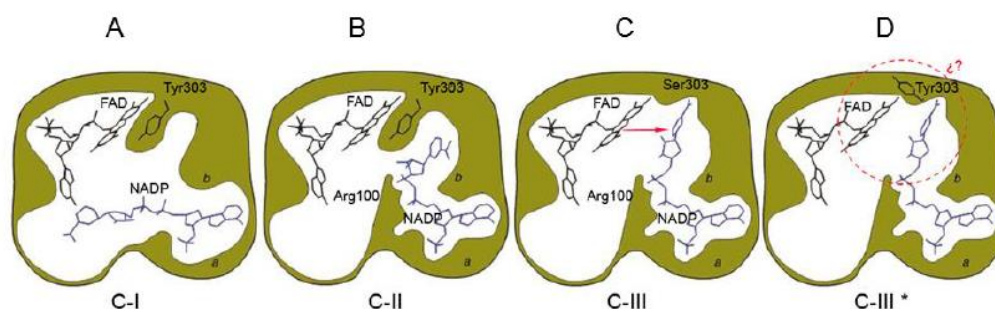
El programa utilizado a lo largo de este trabajo de fin de grado es el “BIOVIA Discovery Studio”, versión 4.5.0.15071. Este programa incluye una gran multitud de opciones y funciones, de las cuales nosotros utilizamos las relacionadas con las interacciones receptor-ligando, que son las que nos permiten realizar acoplamiento molecular y evaluar los resultados obtenidos.

Además del programa anteriormente citado, también se utilizó un software de internet de alineamiento múltiple de secuencias llamado Clustal Omega (Clustal O (1.2.1)).

## 5. Resultados y discusión

### 5.1 Identificación del sitio de unión del sustrato a partir de proteínas análogas

Estudios previos sobre el acoplamiento del coenzima NADP<sup>+</sup> a la FNR de *Anabaena* han planteado un mecanismo de reconocimiento del sustrato con la proteína que nos puede ayudar a abordar nuestro problema. Dicho mecanismo consta de varias etapas (figura 8). En primer lugar, se propone el reconocimiento de la fracción 2'P-AMP del coenzima (C-I), el cual produce una reorganización en la zona de unión dando lugar a un acoplamiento entre la mitad adenina del coenzima y diversos aminoácidos de la proteína (C-II).<sup>5</sup>

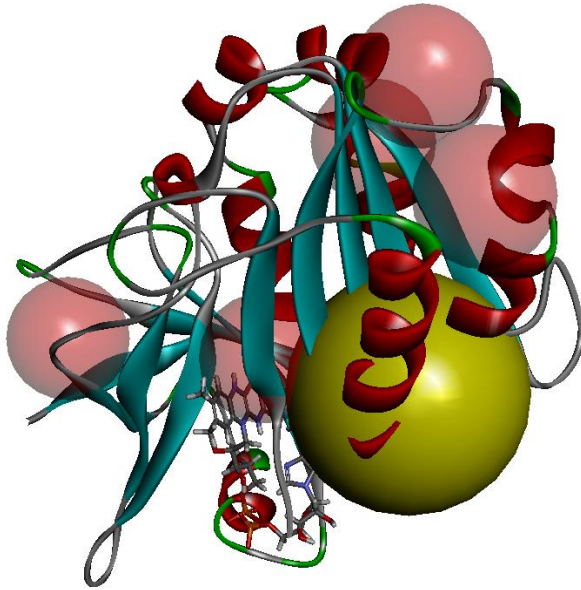


**Figura 8.** Hipótesis del mecanismo de reconocimiento y unión del NADP<sup>+</sup> en *AnFPR*. (Hermoso et al., 2002)

Uno de los puntos claves del mecanismo anteriormente citado es la reorganización inicial que sufre el sistema tras el reconocimiento de la fracción 2'P-AMP con objeto de dar lugar al C-II. Si bien la proteína a partir de la cual se planteó la anterior hipótesis presenta diferencias conformacionales con la *XacFPR*, como son que su FAD adquiere una conformación plegada y presenta una extensión en el C-terminal (figura 2), es muy factible que el mecanismo planteado para la formación de C-I y su transformación en C-II ocurra de forma similar en esta enzima. Por ello, los estudios realizados a lo largo de este proyecto utilizarán la fracción correspondiente al nucleótido de adenina del NADPH, que actúa como inhibidor competitivo del mismo, y que en este caso también se espera sea el punto inicial de reconocimiento.

Como se comentó en la introducción, para que se produzca la transferencia de hidruro entre el NADPH y el anillo de isoaloxacina éstos deben quedar enfrentados (figura 4), por lo que “a priori” predecir el sitio de unión del sustrato a la *XacFPR* no es un proceso complicado ya que el sustrato deberá colocarse en una posición relativamente cercana al coenzima. Sin embargo, el problema se encuentra en los cambios conformacionales que debe sufrir la proteína para alcanzar esa conformación. El programa Discovery Studio permite definir y editar un sitio de unión en una proteína a partir de unas “esferas de unión”, que nos facilitará la elección del mismo. El programa identifica las cavidades existentes en la proteína a partir de funciones geométricas y las ordena. En cualquier caso, la realización de esta etapa es estrictamente necesaria antes de realizar la simulación de acoplamiento molecular. En la figura 9 se muestran las esferas de los sitios de unión que el programa proporciona como más probables en *XacFPR*. La comparación de los sitios identificados por el programa con aquellos obtenidos empíricamente al evaluar la estructura de *XacFPR* en el contexto de las de otras enzimas de la familia (figura 5), sugiere la región identificada en amarillo en la figura 9 como aquella donde se unirá el nucleótido de adenina.



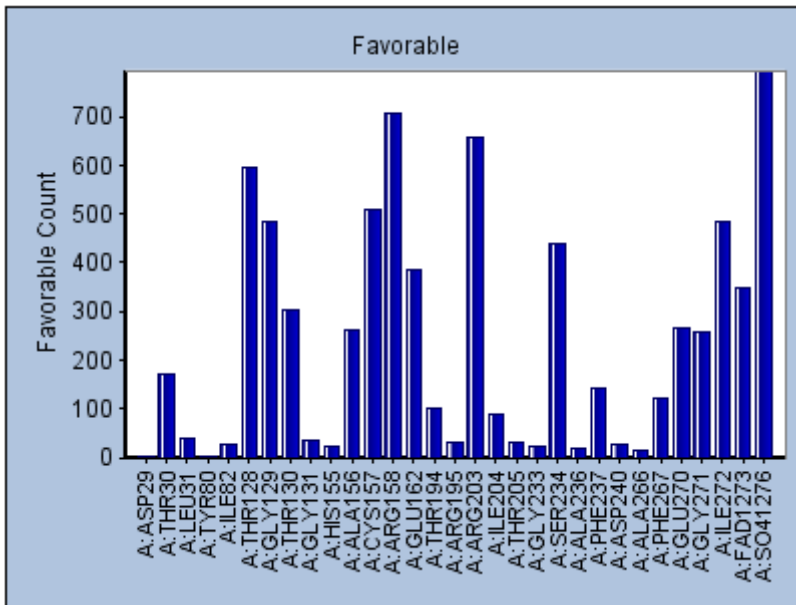


**Figura 9.** Estructura de la *XacFPR* con las esferas de los distintos sitios de unión identificados por Discovery Studio representadas en rojo. En amarillo se muestra la esfera del sitio de unión que a nosotros nos interesa.

Un primer paso antes de realizar cálculos con la *XacFPR* fue realizar un estudio con una proteína análoga, la FPR de *Rhodobacter Capsulatus*, tanto para establecer un primer contacto con el programa como para realizar posteriormente una comparativa de las interacciones entre los distintos aminoácidos con el sustrato. Se realizó el acoplamiento molecular de esa proteína con 2’P-AMP, y para ello primero se prepararon tanto la proteína (a partir de su PDB; código 2BGI), como el ligando (a partir del fragmento de NADPH). Posteriormente se realizó el acoplamiento molecular usando LIBDOCK.

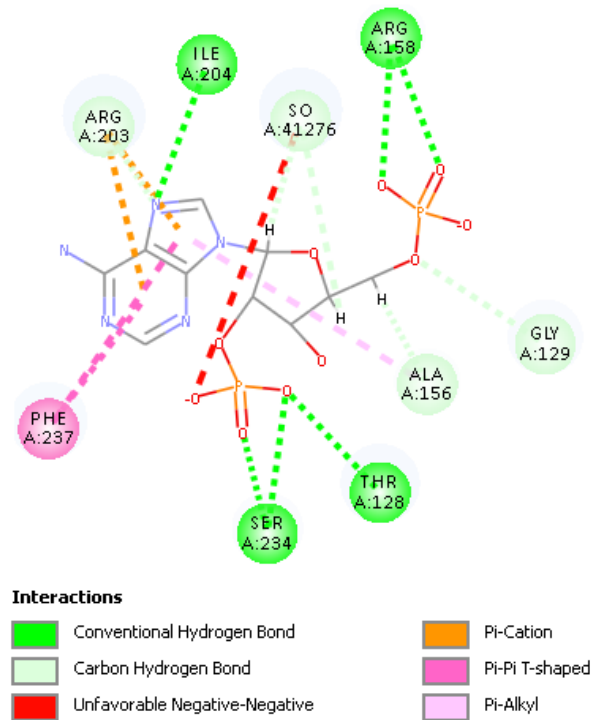
Tras la finalización de los cálculos se obtuvieron un total de 894 posiciones distintas del fragmento del sustrato en la *RcFPR*. Muchas de ellas son descartables por lógica bioquímica, ya que el grupo fosfato del sustrato utilizado que se une con la ribosa y la nicotinamida en el NADPH queda orientado de manera desfavorable con respecto al anillo de flavina, por lo que la transición de hidruro no se produciría. Aun así, con estos cálculos no queremos obtener la orientación final de la fracción 2’P-AMP sino realizar un primer aproximamiento sobre cómo se coloca y con qué partes de la proteína interacciona.

Después se utilizó la función “Analyze ligand poses” para saber qué interacciones se producen entre el nucleótido de adenina y la proteína, así como el número de estas, entre todas las posiciones obtenidas. En la figura 10 se muestra una gráfica del archivo de salida en la que se representa el número total de interacciones favorables para cada uno de los residuos con los que interacciona el fragmento. Excluyendo el grupo sulfato, podemos ver que el residuo que presenta mayor número de interacciones favorables es la Arginina 158 con 706 interacciones, seguida de la Arginina 203, la Tirosina 128 y la Cisteína 157.



**Figura 10.** Representación del número de interacciones favorables para cada residuo entre el fragmento 2'P-AMP y *RcFPR* de las 894 poses obtenidas.

En la siguiente figura se muestra el diagrama 2D de las interacciones producidas entre la fracción 2'P-AMP de una de las poses con distintos residuos. Podemos observar que la mayoría de los aminoácidos nombrados anteriormente como residuos con interacciones más favorables aparecen en el diagrama, así como algunos aminoácidos adyacentes a estos que también presentaban un gran número de interacciones favorables.



**Figura 11.** Diagrama 2D de las interacciones producidas entre una de las posiciones obtenidas de la fracción 2'P-AMP y la *RcFPR*.

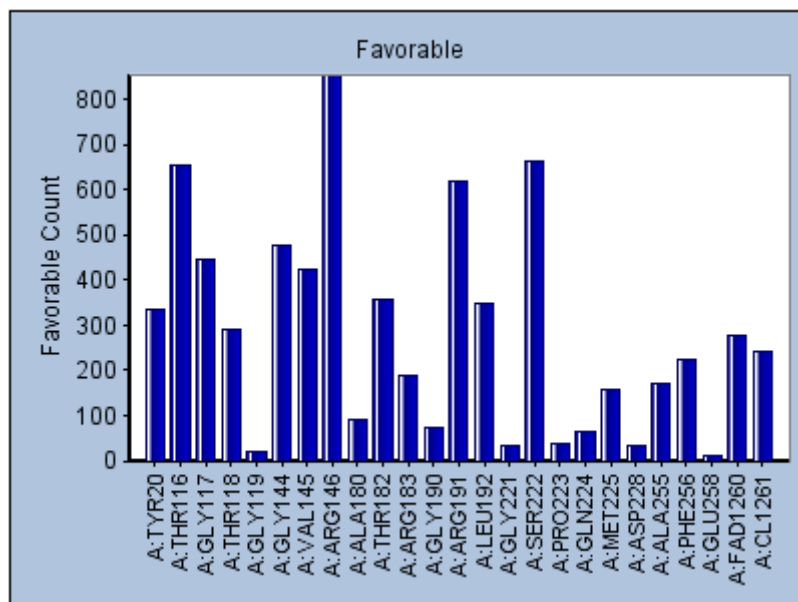
## 5.2 Acoplamiento molecular de cuerpo rígido. Interacción de los residuos de XacFPR con el sustrato

Una vez identificado el sitio de unión del sustrato, realizamos cálculos de acoplamiento molecular con nuestra proteína; el estudio se realizó con la XacFPR, utilizando como sustratos 2’P-AMP y más adelante 2’P-ADP (análogo al anterior, pero con un grupo fosfato más en posición 5’). El proceso para la realización del mismo es similar al del apartado anterior: partimos de la estructura de la proteína (a partir del PDB; código 4B4D) y del fragmento 2’P-AMP o 2’P-ADP, y tras ser preparadas por el programa realizamos el acoplamiento molecular, de nuevo usando LIBDOCK

En el primer caso se obtuvieron un total de 895 posiciones diferentes. De la función “Analyze ligand poses” obtuvimos los aminoácidos que presentaban mayor número de interacciones favorables. Éstos están representados en la figura 12, junto con el recuento de otro tipo de interacciones que pueden ser de interés. En la figura 13 queda representada la gráfica del número de interacciones favorables respecto de cada uno de los residuos.

Top 5 Residues with Favorable Interactions (5)							
Residue	Favorable	Unfavorable	HydrogenBond	Charge	Hydrophobic	Halogen	Other
A:ARG146	855	65	724	762	271	0	0
A:SER222	662	25	653	0	1	0	12
A:THR116	653	19	623	0	46	0	0
A:ARG191	617	35	436	106	291	0	0
A:GLY144	477	15	477	0	0	0	0

**Figura 12.** Tabla con el número de interacciones existentes en todas las posiciones obtenidas tras el acoplamiento molecular de XacFPR y 2’P-AMP para los 5 aminoácidos con mayor número de interacciones favorables.

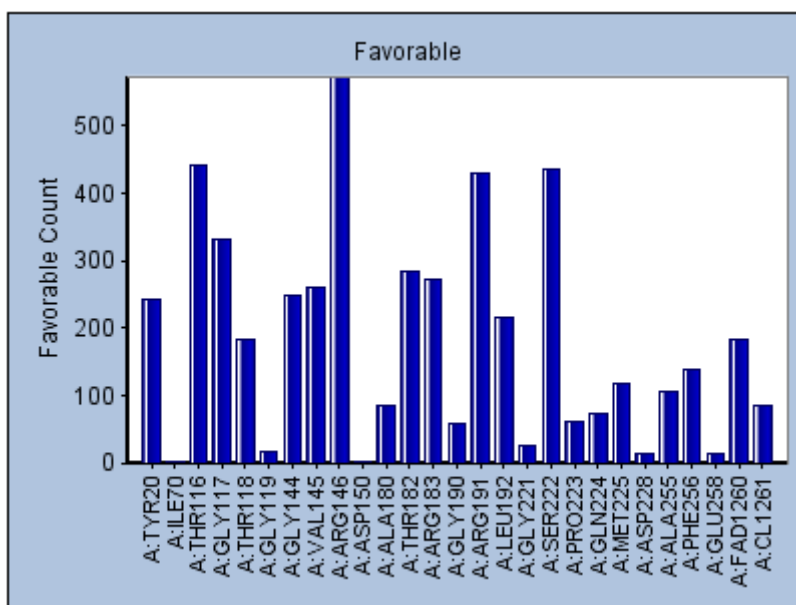


**Figura 13.** Representación del número de interacciones favorables para cada residuo entre el fragmento 2’P-AMP y XacFPR.

Para la simulación con 2'P-ADP el número de posiciones obtenidas fueron 600 y los residuos con interacciones más favorables fueron ARG146, THR116, SER222, ARG191 y GLY117 tal y como se muestra en la tabla de la figura 14.

Top 5 Residues with Favorable Interactions (5)							
Residue	Favorable	Unfavorable	HydrogenBond	Charge	Hydrophobic	Halogen	Other
A:ARG146	572	62	510	523	244	0	0
A:THR116	441	15	414	0	48	0	0
A:SER222	436	25	434	0	1	0	3
A:ARG191	429	10	308	128	183	0	0
A:GLY117	330	1	281	0	68	0	0

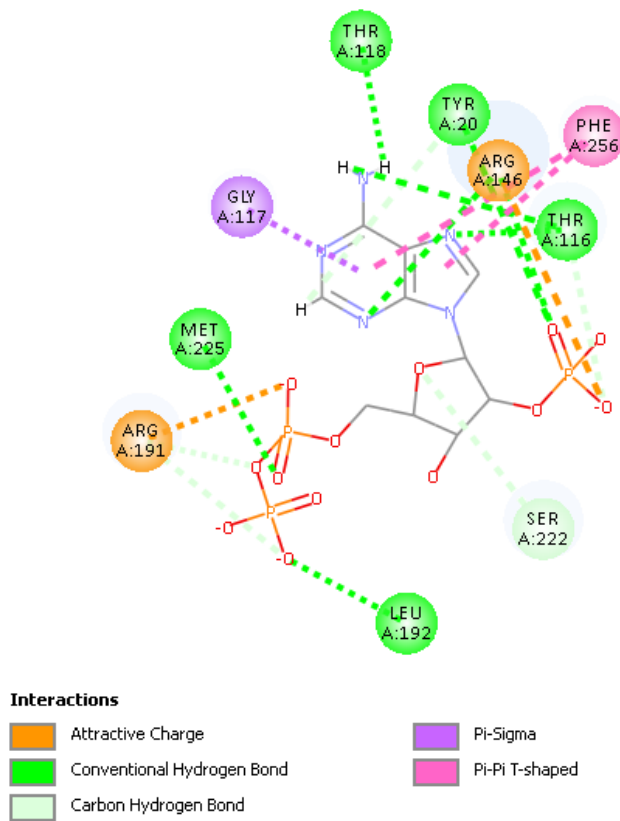
**Figura 14.** Tabla con el número de interacciones existentes en todas las posiciones obtenidas tras el acoplamiento molecular de *XacFPR* y 2'P-ADP para los 5 aminoácidos con mayor número de interacciones favorables.



**Figura 15.** Representación del número de interacciones favorables para cada residuo entre el fragmento 2'P-ADP y *XacFPR*.

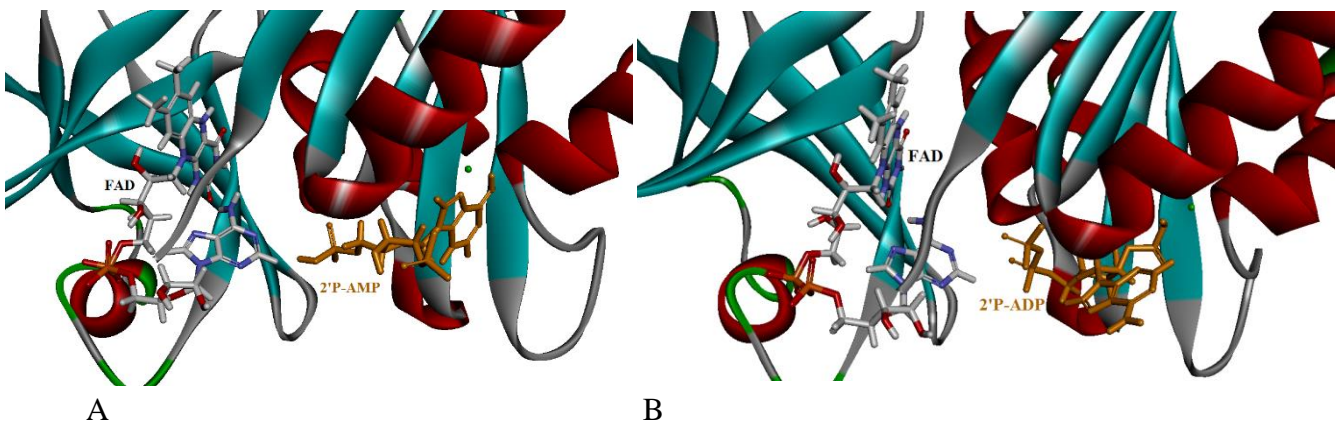
A la vista de estos resultados podemos comprobar que, para ambos nucleótidos de adenina, los 4 aminoácidos que mayor número de interacciones favorables presentan con la *XacFPR* la Arginina 146, Treonina 116, Serina 222 y Arginina 191. Además, los aminoácidos adyacentes a estos también presentan un número considerable de interacciones favorables como podemos ver en las gráficas de las figuras 13 y 15.

La siguiente figura es una representación 2D de una de las posiciones relativas de la fracción 2'P-ADP, en la que podemos ver, entre otras, las interacciones con los aminoácidos nombrados anteriormente.



**Figura 16.** Diagrama 2D de las interacciones producidas entre una de las posiciones obtenidas de la fracción 2'-P-ADP y la *RcfFPR*.

Aunque el número de posiciones relativas obtenidas tras el acoplamiento molecular se ve reducido casi un 33% al usar el nucleótido con un grupo fosfato más, este número sigue siendo muy elevado e incluye de nuevo muchas posiciones que no permitirían un correcto acoplamiento del coenzima NADPH para la transferencia de hidruo. Dos ejemplos de la estructuras de *XacFPR* con el fragmento 2'-P-AMP y 2'-P-ADP se muestran en la figura 16. Ambas podrían ser orientaciones favorables para la transferencia de hidruo desde el NADPH ya que los grupos fosfato de la posición 5' quedan orientados hacia el FAD de la proteína.



**Figura 17.** Ejemplo del acoplamiento molecular entre *XacFPR* y 2'-P-AMP (figura A) o 2'-P-ADP (figura B). En blanco se representa el FAD (átomos de carbono) y en naranja el correspondiente nucleótido de adenina.

### 5.3 Acoplamiento molecular flexible

Se realizó una simulación de acoplamiento molecular entre la fracción 2'P-ADP y *XacFPR* otorgando flexibilidad a ciertas zonas de la proteína que se esperan que sufran una reorganización estructural durante el proceso de transferencia de hidruro. Para elegir a qué residuos de la proteína le damos flexibilidad nos basamos tanto en la información bioquímica estructural y funcional de este sistema y sus análogos, como en resultados obtenidos a lo largo de este trabajo.

Sabemos que los residuos del extremo del dominio C-terminal de las FPRs deben sufrir un reordenamiento estructural con objeto de favorecer el acoplamiento y la transferencia de hidruro entre el anillo de nicotinamida del NADPH y el de la flavina del FAD, por ello es de gran importancia otorgar flexibilidad a esta zona.

Las interacciones obtenidas tras analizar las distintas poses en los cálculos realizados en los apartados anteriores también nos indican que los aminoácidos que contribuían a la interacción pueden también ser clave en el amoldamiento de la proteína.

La figura 18 muestra el alineamiento múltiple realizado con Clustal Omega. de las secuencias de aminoácidos pertenecientes a la FNR de la *Anabaena* y a las FPRs de *Escherichia coli*, *Rhodobacter capsulatus*, *Xanthomonas axonopodis pv. citri* y *Pseudomonas aureginosa*. Para facilitar la visualización del alineamiento, la secuencia de la *Anabaena* fue acortada omitiendo algunos de los aminoácidos iniciales. Este alineamiento nos permite comparar secuencias de aminoácidos de distintas proteínas para conocer las analogías y diferencias entre ellas, así como para hacernos a la idea de qué aminoácidos ocupan posiciones equivalentes entre dos o varias proteínas. Así pues, gracias al alineamiento podemos saber qué aminoácidos de *XacFPR* se corresponden con los de otras proteínas de interés, como puede ser la FNR de *Anabaena*, para saber si se puede esperar una reorganización de esa región y otorgarle flexibilidad.

El análisis conjunto de estos aspectos englobó a un conjunto de regiones de la proteína a las que posteriormente se le otorgaría flexibilidad. Estos aminoácidos se muestran marcados de color amarillo en la figura 19 y listados a continuación:

- Región de Treonina 116 a Glicina 119
- Región Valina 145 a Fenilalanina 147
- Región de Serina 222 a Leucina 226.
- Región de Alanina 255 a Lisina 259 (extremo carboxilo terminal).

```

AnFNR -TQAKAKH-----ADVPVNL YRPNAPFIGKVISNEPLVKEGGIGIVQHI--
EcFPR -----MADW-----VTGKVTKVQNWTD
RcFPR MFPAAPSHAPVRRFFIAPSGLKDETTVNETTPIAPAKVLP-----DAQTVTSVRHWTD
XacFPR -----MSSAF-----GAETVLEVRHWTD
PaFPR -----MSNL-----YTERVLSVHHWTD
                                         : *:.

AnFNR -KFDL--T-GGNLKYIEGQSIGIIPPQVDKNGKPEKLRLYSIASTRHGDDVDDKTISLCV
EcFPR ALFSLTVH-APVLPFTAGQFT---KLGLE-IDGERVQRAYSIVNSPDNPDLEFYLVTD---
RcFPR TLFSFRVTRPQTLRFRSGEFV---MIGLLDDNGKPIMRAYSIASPAWDEEELFYLSIK---
XacFPR AYFSFTTTTRDAGFRFENGQFV---MIGLE-TETRPLLRAYSIASANWEEHLEFFSIK---
PaFPR TLFSFKTTRNPGLRFKGTQFV---MIGLE-VDGRPLMRAYSIASPNYEEHLEFFSIK---
      *.:      : : *:      * :      * ** ..      .: : .

AnFNR RQLEYKHPESGETVYGCSTYLTHIEPGSEVKITG-PVGKEMLLPDDPEANVIMLATGTG
EcFPR -----VPDGKLSRLAALKPGDEVQVWSEAAAGFFVLDEVPHCETLWMLATGTA
RcFPR -----VPDGPLTSRLQHIKVGQIILRPKPVGTLVLDALLPGKRLWFLATGTG
XacFPR -----VPDGPLTSRLQHIQPGDKVLVGKKPTGTLISDLHPGRNLYLLGTGTG
PaFPR -----VPDGPLTSRLQHLKEGDELMVSRKPTGTLVHDDLLPGKHLVLLSTGTG
      . * : * : : *.: : : . * : : : :*.***.

AnFNR IAPMRTYLWRMFKAERAANPEYQFKGFSWL VFGVPTTPNILYKEELEEI-----QQK
EcFPR IGPYLSILQLG-KLDR-----FKNLVLVHAARYAADLSYPLMQEL-----EKR
RcFPR IAPFASLMREP-EAYEK-----FDEVIMMHTCRTVAELEYGRQLVEALQEDPLIGEL
XacFPR LAPWLSIIKDP-ETYER-----FDKVLITQGVRFVQDLAYRDYFERELPQHEFLGDL
PaFPR MAPFLSVIQDP-ETYER-----YEKVLVHGVRVWSELAYADFITKVLPEHEYFGDQ
      :.* : : : : : : : : * : : .

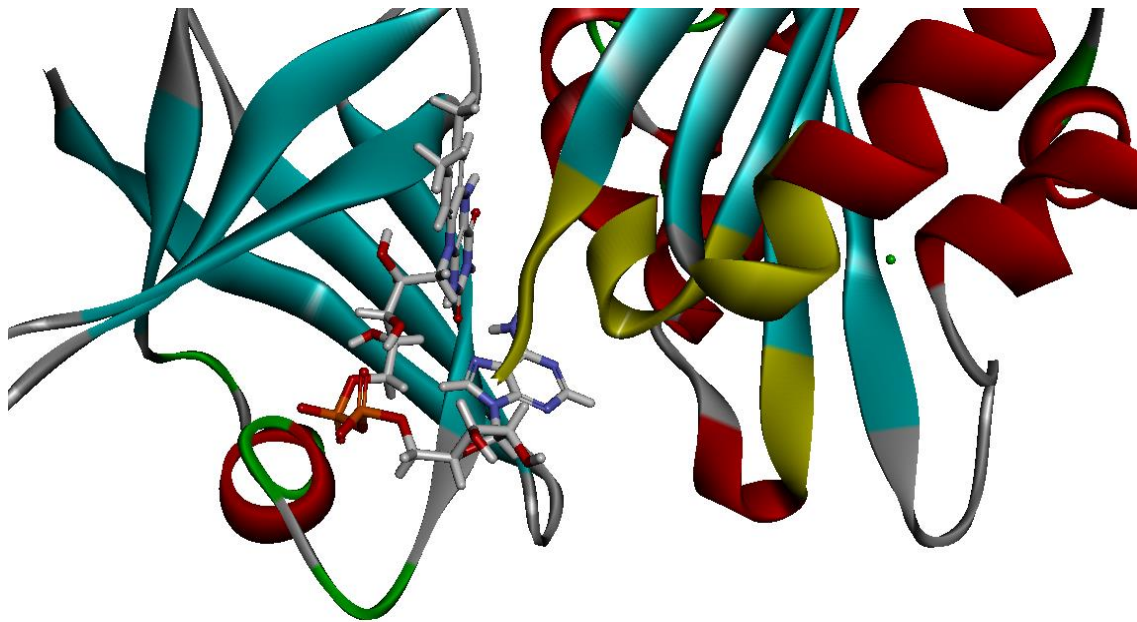
AnFNR YPDNFRITYAISREQKN-PQGGRMYIQDRVAEH-ADELWQLIKNQKTHYICGLRGMEEG
EcFPR YEGKLRITVWSRETAAGSLTGRIPALIESGELESTIGLP-MNKETSHVMLCGNPQMVRD
RcFPR VEGKLYYPTTTREEFH--HMGRITDNLASGKVFEDLGITPMNPETDRAMVCGSLAFNVD
XacFPR LREKLLYYPAVTRETFA--NQGRLELMADGRMQQTLGLPTLDPANDRFMICGSPQMLAD
PaFPR VKEKLIYYPLVTREPFR--NQGRQTDLMRSGKLFEDIGLPPMNPQDDRAMICGSPSMLEE
      ::      :**      **      ..      : :      : :**      :

AnFNR IDAALSAAAAKEGVTWSDYQKDLKAGRWVETV-----
EcFPR TQQLLKETRQMTKHL-----RRRPGHMTAEHYW-----
RcFPR VMKVLESYGLREGA-----NSEPREFVVEKAFVGEGI
XacFPR LRSLLDSRGFQTSR-----IGTPGHYVFERAFVEK--
PaFPR TSAVLDSFGLKISPR-----MGEPGDYLIERAFVEK--
      * . *

```

**Figura 18.** Alineamiento de secuencias de distintas proteínas; en orden, la FNR de *Anabaena* y las FPR de *Escherichia coli*, *Rhodobacter Capsulatus*, *Xanthomonas a. pv. citri* y *Pseudomonas aureginosa*. En negro se remarcan los aminoácidos a los que se les otorgó flexibilidad en *XacFPR* y los que serían los aminoácidos equivalentes a estos en el resto de proteínas. Alineamiento realizado con Clustal Omega.





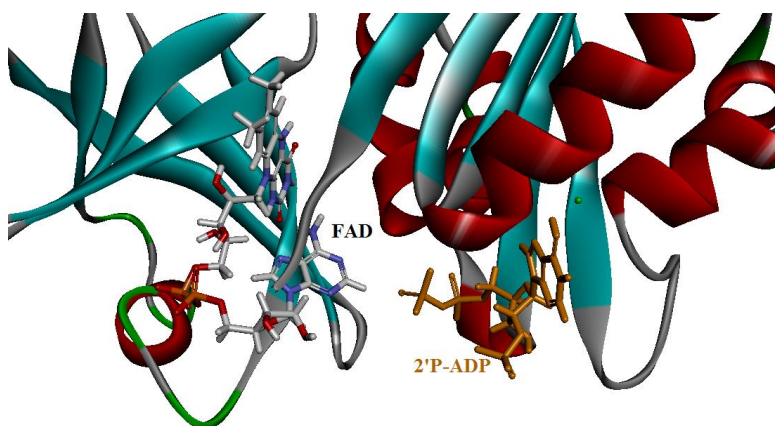
**Figura 19.** Estructura de *XacFPR* que muestra en color amarillo las regiones de la proteína a las que se les otorgó flexibilidad en el acoplamiento molecular. El FAD queda representado en modelo de varillas con los carbonos en color blanco.

El propio programa ofrece por defecto distintas regiones a las que otorgar flexibilidad, pero ninguna de ellas incluía todas las zonas que de interés en este estudio. La más similar se correspondía con el sitio de unión del nucleótido de adenina por lo que incluía los aminoácidos cercanos a dicha zona, pero excluía toda la región carboxilo terminal (esfera amarilla de la figura 9).

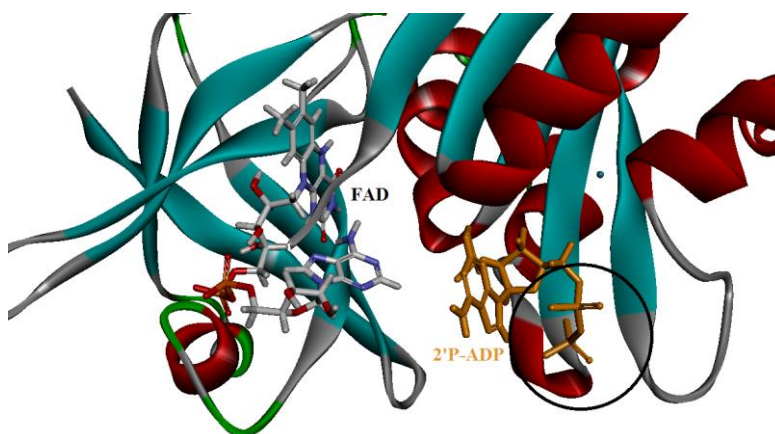
Se realizó el acoplamiento molecular con 2'P-ADP, utilizando primero la zona de flexibilidad identificada en nuestras observaciones y después la proporcionada por el programa. En el primero se obtuvieron 113 posiciones relativas distintas del ligando, tras un tiempo de cálculo de 13 horas, 16 minutos y 30 segundos. En cambio, solo se obtuvieron 2 posiciones distintas en el acoplamiento molecular dando flexibilidad a la zona del sitio de unión predeterminada por el programa, con un tiempo de cálculo de 3 horas, 18 minutos y 14 segundos. Sin embargo, en este último caso estas dos posiciones tienen los grupos fosfato de la posición 5' orientados hacia la parte exterior de la proteína y no hacia el anillo de flavina.

En las figuras 20 y 21 se muestra un ejemplo para el acoplamiento del nucleótido de adenina 2'P-ADP con *XacFPR* otorgándole flexibilidad a los aminoácidos marcados en la figura anterior y dándole flexibilidad a la zona predeterminada por el programa respectivamente. Nótese en este segundo caso (figura 21) la orientación que toman grupos fosfato del nucleótido de adenina (rodeados por un círculo negro en la figura), ya que lo hacen de manera desfavorable con respecto del FAD de la proteína, lo que no favorecería la transferencia de hidruro. Esta orientación se presentaba, con ligeros cambios, para las dos poses obtenidas en dicho estudio.





**Figura 20.** Ejemplo del acoplamiento molecular entre *XacFPR* y 2'-P-ADP otorgándole flexibilidad a los aminoácidos que propusimos. En blanco se representa el FAD (átomos de carbono) y en naranja el nucleótido de adenina.



**Figura 21.** Ejemplo del acoplamiento molecular entre *XacFPR* y 2'-P-ADP otorgándole flexibilidad a la zona predeterminada por el programa en el sitio de unión. En blanco se representa el FAD (átomos de carbono) y en naranja el nucleótido de adenina.

Debido a las grandes diferencias entre los resultados obtenidos entre los dos estudios llevados a cabo en este apartado podemos decir que, tanto por el número de poses resultantes como por la orientación de las mismas, la región del extremo del dominio C-terminal y la reorganización que pueda sufrir son factores claves para estudiar la interacción de la unión de la mitad correspondiente al nucleótido de adenina del coenzima NADPH, o del mismo NADPH, con *XacFPR*.

## 6. Conclusiones y perspectivas

El nucleótido de adenina, 2'P-AMP y 2'P-ADP se unirán en *XacFPR* en una región equivalente a la identificada en otros miembros de la familia.

Los residuos ARG146, SER222, THR116, y ARG191, así como residuos adyacentes a estos forman parte del sitio de unión del nucleótido de adenina en *XacFPR*, contribuyendo a su acoplamiento en una orientación fisiológica a la proteína, ya que son los aminoácidos que mayor número de interacciones favorables presentan con dicho nucleótido.

La flexibilidad del extremo C-terminal de *XacFPR* es un factor clave para el correcto acoplamiento de la mitad 2'P-ADP del coenzima NADPH. Esto está avalado por los diferentes resultados obtenidos entre los dos estudios de acoplamiento molecular flexible realizados en este trabajo.

Las estructuras aquí obtenidas para el acoplamiento de *XacFPR* con 2'P-ADP son un buen modelo de partida para, utilizando métodos de dinámica molecular, plantear simulaciones que permitan evaluar el grado de reorganización de la proteína en el centro activo para así alojar la parte del nucleótido de nicotinamida del NADPH, con objeto de que sea posible la transferencia de hidruro.

## 7. Bibliografía

1. M. L. Tondo, M. Musumeci, M.L. Delprato, E. Ceccarelli, E. Orellano, PLOS ONE 2011, volume 6, issue 11.
2. S. Petrocelli, M.L. Tondo, L. Daurelio, E. Orellano, PLOS ONE 2012, volume 7, issue 7.
3. A. Sánchez-Azqueta, D. Catalano-Dupuy, A. López-Rivero, M.L. Tondo, E. Orellano, E. Ceccarelli, M. Medina, Biochimica et Biophysica Acta 2014, 1837, 1730-1738.
4. M.L. Tondo, R. Hurtado-Guerrero, E. Ceccarelli, M. Medina, E. Orellano, M. Martínez-Júlvez, BioMed Research International 2013, ID906572.
5. I. Lans, Tesis doctoral 2012.
6. Química cuántica (J. Bertran Rusca, V. Branchadell, M. Moreno, M. Sodupe). Editorial Síntesis, segunda edición, junio 2002.
7. Essentials of Computational Chemistry (C. J. Cramer). Editorial John Wiley and sons, segunda edición, 2004
8. P. Ferrara, H. Gohlke, D. J. Price, G. Klene, C. L. Brooks, J. Med. Chem 2004, 47, 3032-3047.
9. T. Fraczek, A. Siwek, P. Paneth, J. Med. Chem 2013, 53, 3326-3342