



UNIVERSIDAD DE JAÉN
Centro de Estudios de Postgrado

Trabajo Fin de Máster

**APLICACIÓN DE TÉCNICAS DE
MUESTREO MEJORADO EN LA
LOCALIZACIÓN DE SITIOS DE
UNIÓN DE PROTEÍNAS
IMPLICADAS EN PATOGÉNESIS**

Alumno/a: Miriam Iliex Guirado López

Tutor/a: Prof. D. José Manuel Granadino Roldán

Dpto: Química Física y analítica

Julio, 2020

**APLICACIÓN DE TÉCNICAS DE MUESTREO MEJORADO EN LA
LOCALIZACIÓN DE SITIOS DE UNIÓN DE PROTEÍNAS IMPLICADAS
EN PATOGÉNESIS.**

Memoria del Trabajo Fin de Máster presentada por
Miriam Iliex Guirado López

Jaén, 1. de Julio de 2020

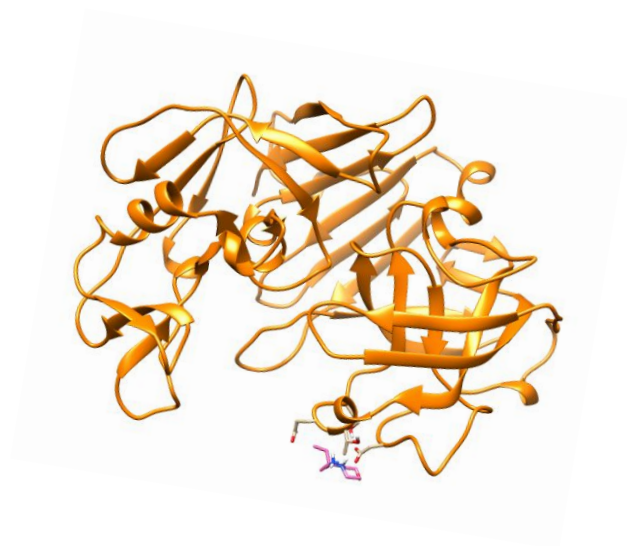
Fdo.: Miriam Iliex Guirado López

A handwritten signature in black ink, appearing to be 'Miriam Iliex Guirado López', written in a cursive style.

Tutor del Trabajo Fin de Máster

Fdo.: José Manuel Granadino Roldán
Profesor del Departamento de Química Física y
Analítica

EXERGO



***“Resulta difícil hacer predicciones,
especialmente acerca del futuro”***

Niels Bohr

DEDICATORIA



A mis padres porque sé que les hubiera encantado estar a mi lado durante esta nueva etapa.

AGRADECIMIENTOS



A Dios por todas sus bendiciones....

A mis padres por continuar cambiando su vida por mí; por dar todas sus energías para sacarme adelante y enseñarme a vivir; porque sé que el mayor sacrificio ha sido de ellos.

A mi hermano Luisy y a mi abuela Consuelo por aceptar el no verme a diario por venir tras un sueño.

A mi familia, la de sangre y la elegida por siempre hacerse presente.

A todos los profes del máster por brindarme sus conocimientos, en especial al coordinador Juan, su ayuda fue vital para el comienzo de este máster y a mi tutor "Chema" a quién cansé con dudas y lo hice repetir una y mil veces lo mismo hasta que logré entenderlo y aquí está profe: ya hemos terminado, gracias!!

A mis profesores de la UCLV quienes me ayudaron y me motivaron y han estado al tanto de mí en todos estos meses.

A mis amigos y amistades, esas con las que no han sido necesarias conversaciones diarias para mantener la amistad y siempre me han dado ánimo desde la distancia.

A Raddel: estuviste desde que surgió la idea de hacer un máster (creo que hasta fue tu idea) hasta el momento de redactar la última oración del TFM

A mis compañeros del máster, fue poco el tiempo que compartimos pero me alegra haberlos conocido, gracias por su ayuda. Mari, muchas gracias por estar al tanto de mí...

A mi Lagunillas' team, sin ustedes TODO hubiera sido más difícil. Mil gracias a los 3.

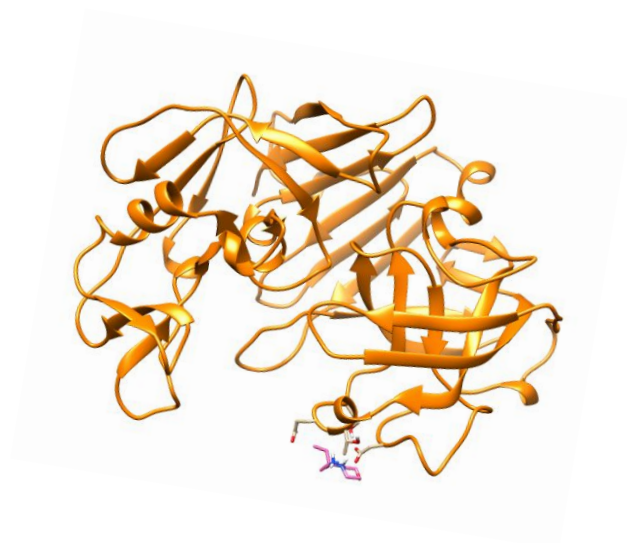
RESUMEN



RESUMEN

Los métodos teóricos de modelización molecular se usan para predecir y/o explicar los resultados experimentales. El diseño de fármacos basado en fragmentos (FBDD) ha dado paso al desarrollo de fármacos y a descubrir el sitio y la forma de unión de ligandos a la proteína. El método de dinámica molecular de fragmentos disueltos (fdMD) se aplica para mejorar el FBDD. Se tuvo como objetivo comprobar si se pueden mejorar los resultados mediante una modificación aplicada a la metodología fdMD centrada en la mejora del muestreo conformacional con el uso de H-REMD. Para ello se estudiaron 2 ligandos, F031 y F035, se probaron 3 metodologías diferentes y se realizó el cálculo de 3 descriptores. Resultó que, probadas las metodologías: el ligando F031 fue capaz de encontrar su sitio de unión (BS) experimental a diferencia del ligando F035; se logró distinguir entre BS correctos y falsos positivos; se encontraron posibles BS nuevos. Se concluyó que: con los métodos 2 y 3, el ligando F031 fue capaz de encontrar la pose correcta del ligando; en los casos en los que el ligando no encontró su pose experimental es necesario ampliar el tiempo de simulación de la dinámica y volver a calcular los descriptores; los descriptores propuestos podrían resultar efectivos para descartar con confianza falsos sitios de unión, permitiendo predecir mediante estos los sitios de unión en estudios posteriores donde no se conozca la experimental.

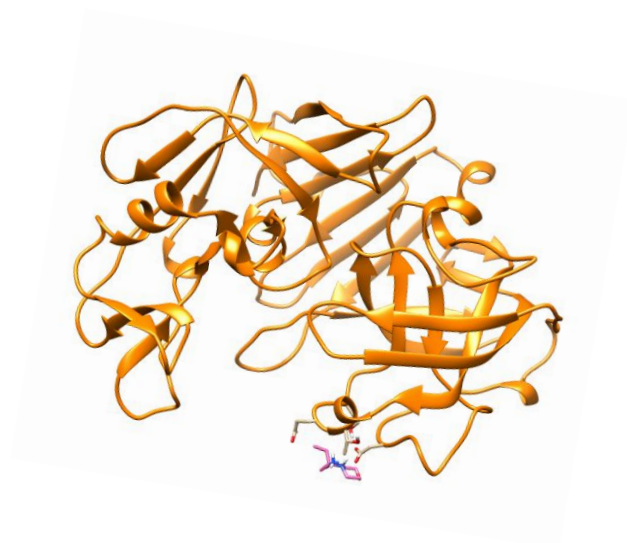
ABSTRACT



ABSTRACT

Theoretical molecular modeling methods are used to predict and/or explain experimental results. Fragment-based drug design (FBDD) has led to drug development and discovery of the site and manner of ligand binding to the protein. The Dissolved Fragment Molecular Dynamics (fdMD) method is applied to improve FBDD. The objective was to check if the results can be improved through a modification applied to the fdMD methodology focused on improving conformational sampling with the use of H-REMD. For this, 2 ligands F031 and F035 were studied, 3 different methodologies were tested and the calculation of 3 descriptors was performed. It turned out that, once the methodologies had been tested: the F031 ligand was able to find its experimental binding site (BS) unlike the F035 ligand; it was possible to distinguish between correct BS and false positives; possible new BSs were found. It was concluded that: with methods 2 and 3, ligand F031 was able to find the correct pose for the ligand; in cases where the ligand did not find its experimental pose, it is necessary to extend the simulation time of the dynamics and recalculate the descriptors; the proposed descriptors could be effective in confidently ruling out false binding sites, allowing the binding sites to be predicted in subsequent studies where the experimental one is not known.

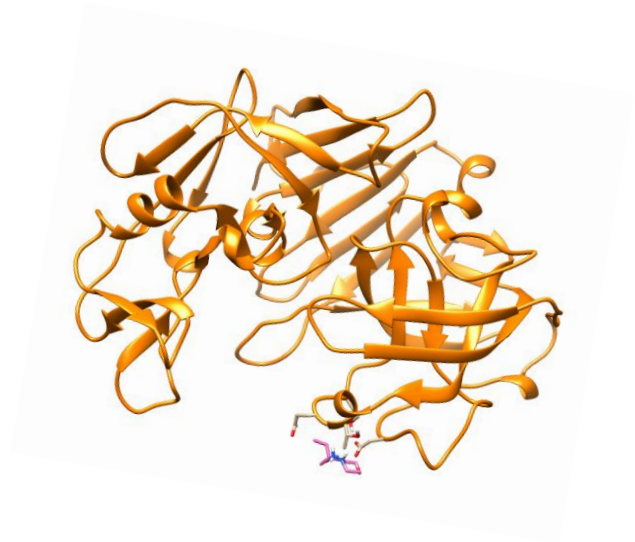
ÍNDICE



ÍNDICE	
DEDICATORIA	3
AGRADECIMIENTOS	5
RESUMEN	8
ABSTRACT	10
ÍNDICE	12
INTRODUCCIÓN	15
Dinámica Molecular	15
Dinámica Molecular Clásica(cMD)	16
Métodos para mejorar el muestreo conformacional.	16
Búsqueda de sitios de unión con FBDD.	17
El método fdMD.	18
Endotiapepsina como proteína modelo utilizada	19
OBJETIVOS	22
MATERIALES Y MÉTODOS	24
Campos de fuerza	24
Dinámica Molecular	26
Dinámica Molecular Clásica(cMD)	28
Dinámica molecular de intercambio de réplicas (REMD)	29
Dinámica molecular de intercambio de réplicas H-REMD	29
Energía libre de unión MMPB/GBSA	30
Dinámica molecular de fragmentos disueltos (fdMD)	31
Diagrama de flujo de fdMD.	31
Sistemas de proteína-ligando estudiados (9)	32
Preparación de proteínas(9)	33
Preparación y simulación de sistemas solvatados(9)	33
Descripción de los métodos utilizados	34

RESULTADOS	37
Ligando F031	38
Método 1	38
Método 2	38
Método 3	39
Listado de residuos de aminoácidos para cada clúster y su correspondiente figura 41	
Ligando F035	45
Método 1	45
Método 2	45
Método 3	45
Listado de residuos de aminoácidos para cada clúster y su correspondiente figura 47	
DISCUSIÓN	51
CONCLUSIONES	55
BIBLIOGRAFÍA	56
ANEXOS	61
CURRICULUM VITAE	63

INTRODUCCIÓN



INTRODUCCIÓN

Dinámica Molecular

Al realizar una investigación referente a un sistema biológico y/o químico resulta imperativa la gestión de métodos teóricos de modelización molecular, que serán los encargados de complementar o predecir los resultados obtenidos mediante técnicas experimentales para el mejor entendimiento de la estructura, la función y los mecanismos de reacción de estos sistemas. El análisis de los sistemas por rayos X permite el estudio en cuanto a propiedades físicas y estructurales de una molécula, pero la visualización de una única estructura molecular no es suficiente a la hora de dar solución a situaciones más complejas como pueden ser el movimiento de una proteína o su unión a un ligando, donde se hace necesario auxiliarse de la simulación.

“El principal objetivo de la modelización molecular, aplicada en distintas áreas como la bioquímica o la farmacología, consiste en describir las interacciones biomoleculares en base a las leyes generales de la química y de la física”.⁽¹⁾

Actualmente se hace necesario introducir aproximaciones para mantener la viabilidad de la simulación aplicando distintos niveles de teoría en función del sistema que va a ser analizado. Como norma general, a medida que aumenta el nivel teórico aumenta la calidad de los resultados obtenidos, pero también se incrementa notablemente el coste computacional.

La dinámica molecular (*Molecular dynamics*, MD) representa un tipo de simulación molecular computacional que facilita el análisis del comportamiento o la evolución de un sistema, ya sea físico, químico o biológico, a través del tiempo. Para ello se debe de calcular las fuerzas existentes entre los átomos que conforman el sistema auxiliándose de las ecuaciones del movimiento de Newton.⁽²⁾

A diferencia de otros métodos de simulación molecular, que permiten simular una etapa o "un estado" del sistema en estudio, la MD brinda la evolución de dicho sistema a través del tiempo, ya que se tienen en cuenta las velocidades de los átomos, y también es posible calcular la energía cinética del sistema. Si se analiza la interacción de un fármaco con su receptor, la MD permite observar los cambios conformacionales de dicho receptor y no únicamente del fármaco, lo cual es muy importante porque brinda información crucial para el diseño de nuevas

moléculas para uso clínico. Otra ventaja es que no sólo permite la simulación del sistema ligando-proteína, sino que además se puede incluir el disolvente que lo rodea de modo explícito (generalmente agua), e iones, regular la temperatura o la presión, o la incorporación de entidades químicas (metales, azúcares, cofactores, grupos prostéticos) que usualmente no son consideradas en otros métodos de simulación debido a sus limitaciones.⁽²⁾

La MD se emplea en múltiples campos; en el mundo de la biología se utiliza en el estudio de macromoléculas (proteínas, ácidos nucleicos, carbohidratos, lípidos) con el objetivo de conocer y estudiar los cambios conformacionales que experimentan estos sistemas, los canales iónicos y las membranas de las proteínas así como las simulaciones de plegamiento y desplegamiento que sufren estas.

Con MD se puede estudiar el comportamiento de un ligando frente a su diana molecular, siendo un eslabón clave en el diseño de nuevos fármacos; se puede estudiar un fluido y establecer su temperatura de fusión o evaporación; facilita el estudio de solubilidades y el paso de moléculas a través de membranas celulares, entre muchas otras cosas.⁽²⁾

Dinámica Molecular Clásica(cMD)

La cMD define la energía total del sistema con una parte intramolecular y otra intermolecular, a partir de la cual, aplicando las leyes de Newton, se puede determinar la fuerza sobre cada átomo y de ahí calcular las correspondientes velocidades. Los átomos se mueven a sus nuevas posiciones y se recalculan las velocidades, y así sucesivamente, de manera que se genera una trayectoria que describe la evolución de la posición y las velocidades de las partículas en el tiempo. De acuerdo con la aproximación ergódica, es clave hacer un buen muestreo para obtener promedios reales, por lo que existen diferentes métodos que contribuyen a optimizar dicho muestreo.

Métodos para mejorar el muestreo conformacional.

La mayoría de los procesos biológicos que ocurren en las células se ven influenciados por las interacciones proteína-ligando, las cuales, además, median el efecto de las moléculas del fármaco. En biología estructural y en biomedicina es importante una predicción lo más acertada posible, mediante métodos computacionales, de cómo ocurre la unión entre el ligando y la

molécula receptora y cuáles pueden ser los posibles sitios de unión que van a favorecer el posterior desarrollo de ligandos específicos.

Existen diferentes técnicas y métodos que mejoran el muestreo como lo es la dinámica molecular de intercambio de réplicas (REMD, *Replica Exchange Molecular Dynamics*). El método REMD fue introducido por primera vez en el estudio de biomoléculas por Okamoto y col.⁽³⁾. Es un método híbrido que combina simulaciones MD con el algoritmo de Monte Carlo para simular en paralelo al sistema a diferentes temperaturas e intercambiar esas simulaciones (réplicas) durante la simulación (T-REMD) ⁽⁴⁾. La naturaleza en paralelo del método REMD lo hace apropiado para su uso en clústeres de cálculo altamente paralelizados que están comúnmente disponibles hoy en día.

Los métodos de acoplamiento molecular (o *docking*) persiguen la obtención del sitio de unión y el modo de unión (*pose*) de un ligando en una macromolécula biológica. Dichos métodos a menudo descuidan la flexibilidad conformacional o la incluyen de una forma aproximada. Lo ideal sería un esquema de dos pasos, primero un *docking* rígido y luego un refinamiento flexible. Este esquema estará condicionado por el resultado de una búsqueda sistemática inicial, de la modelización y de la predicción.

En lugar de espaciar réplicas en todo el espacio de temperatura, también se puede espaciar réplicas en todo el "espacio hamiltoniano" (H-REMD). Es decir, cada réplica tiene una función de energía o hamiltoniano diferente, y se producen intentos de intercambio entre hamiltonianos adyacentes. Los programas sander y pmemd ⁽⁵⁾, utilizados en el presente trabajo, implementan H-REMD mediante el intercambio de coordenadas entre réplicas y evaluando la energía de esa nueva estructura.

Búsqueda de sitios de unión con FBDD.

Desde 1990 el filtrado de alto rendimiento (HTS, *High-throughput screening*) ha sido importante para el hallazgo de fármacos. El enorme espacio químico a explorar ha permitido el surgimiento, en los últimos años, de una alternativa para solucionar algunos inconvenientes del HTS: el diseño de fármacos basado en fragmentos (FBDD, *Fragment Based Drug Design*), técnica establecida que ha dado paso al desarrollo de más de 30 candidatos a fármacos.⁽⁶⁾

El FBDD se ha abordado desde una perspectiva teórica, inspirada en el método de estructuras de cristales solventes múltiples (MSCS, por sus siglas en inglés) de Dagmar Ringe ⁽⁷⁾. Estos métodos han evolucionado desde el uso de estructuras estáticas para representar la proteína hasta enfoques basados en MD que permiten la adaptación conformacional de la proteína a la presencia de las sondas, simulando la proteína en una solución de agua y moléculas co-disolventes (estas son, las mencionadas sondas).

Otro enfoque teórico es el seguido por Fabritiis y col.⁽⁸⁾ en el que un fragmento se coloca aleatoriamente en una caja solvatada y se realizan múltiples ejecuciones cortas de MD para obtener una MD global de hasta 100 μ s, con el objetivo final de predecir el sitio de unión y el modo de unión de un conjunto de fragmentos. Otra perspectiva del FBDD está dirigida a descubrir el sitio de unión y la forma de unión de ligandos de tamaño completo utilizando un elevado número de copias del ligando en la caja (64-96), lo cual requirió el uso de un término repulsivo en forma de un potencial de Lennard-Jones.

El método fdMD.

El FBDD fue utilizado inicialmente como técnica experimental y a pesar de constituir una muy buena técnica para el descubrimiento de fármacos presenta algunas desventajas como son la desnaturalización de proteínas o la agregación de ligandos. Para solventar estos problemas el grupo en el cual se ha desarrollado el presente Trabajo Fin de Máster propuso el llamado método de dinámica molecular de fragmentos disueltos (fdMD, *Fragment Dissolved Molecular Dynamics*), que es un método sistemático y semiautomático para ser aplicado en proyectos FBDD, destinado a encontrar el sitio de unión de un ligando pequeño con su proteína diana.⁽⁹⁾

Este método utiliza una caja previamente equilibrada que contiene una copia del ligando solvatado con moléculas de agua, y se utiliza un término potencial repulsivo de Lennard-Jones para evitar la agregación de los ligandos. Dicha caja puede usarse fácilmente para solvatar la diana de interés. Por lo tanto, la solvatación de la proteína con esta caja pre-equilibrada genera un sistema de simulación que contiene un número variable, aunque bajo, de copias del ligando, lo que evita el problema de la desnaturalización de la proteína. Este método tiene, por un lado, la ventaja de solvatar la diana con un bajo número de ligandos, evitando así la desnaturalización, y por otro, proporciona la opción de utilizar una base de datos de cajas solvatadas de diferentes ligandos para ser usadas en proyectos FBDD. Para validar la metodología se emplearon tres sets diferentes de proteína-ligando (set I, II y III), en particular

esta investigación está dirigida al set I, compuesto por 4 ligandos que se unen a diferentes zonas de la proteína diana endotiapepsina (Anexo 1). el trabajo se centra en dos de estos cuatro ligandos (ligando F031 y ligando F035) con el objetivo de estudiar modificaciones en el método y verificar si la metodología fdMD es lo suficientemente robusta como para encontrar el sitio de unión correcto.⁽⁹⁾

El método cuenta con secuencias de comandos para el análisis de resultados y la obtención de 6 descriptores que permiten descartar con seguridad sitios de unión espurios. Se propuso utilizar cuatro trayectorias de dinámica molecular **clásica** de 200 ns de longitud, las cuales se podrían ampliar cuando los resultados sean dudosos. Los resultados obtenidos de este modo demostraron que el número de réplicas utilizado es suficiente para identificar los verdaderos sitios de unión y descartar los falsos positivos. Se propusieron 6 descriptores: energía de enlace MMGBSA promedio, energía de enlace K_{DEEP} promedio, mejor energía de enlace MMGBSA, mejor energía de enlace K_{DEEP} ; se establecen además el total de trayectorias reactivas, o sea cuántas de las cuatro trayectorias terminan con un ligando interactuando con el *pocket* de unión analizado y el mayor tiempo que permanece el ligando en cada posible sitio de unión: tiempo de resistencia (RT, según siglas del inglés).⁽⁹⁾

En la siguiente investigación se pretende hacer algo similar pero en lugar de utilizar cMD utilizar H-REMD.

Endotiapepsina como proteína modelo utilizada

La proteasa aspártica, que se deriva del hongo *Endothia parasítica*, llamada endotiapepsina, fue elegida como diana para el set I del trabajo de validación del método fdMD. Esta elección se basó en la existencia de datos experimentales de esta proteína que permiten poder comprobar y evaluar la aplicación del método.

La endotiapepsina pertenece a la familia de las proteasas aspárticas. Se encuentra en hongos, vertebrados, plantas y retrovirus como el VIH⁽¹⁰⁾; juega un papel causal en numerosas enfermedades como la enfermedad de Alzheimer, la hipertensión y el VIH.

La endotiapepsina se ha utilizado como enzima modelo para estudios mecanicistas y para el descubrimiento de inhibidores de renina y β -secretasa⁽¹¹⁾. Es una molécula bilobulada, de peso molecular igual a 34 kDa⁽¹²⁾, consta de 330 residuos (con aproximadamente 170 residuos en

cada dominio⁽¹³⁾. Se divide en dos lóbulos topológicamente equivalentes que están relacionados por un doble eje. Solo 13 residuos son idénticos en ambos lóbulos, pero 86 están en posiciones topológicamente equivalentes. Esto indica que la estructura puede haber evolucionado por duplicación de genes, fusión y divergencia de secuencia.⁽¹⁴⁾

A partir de lo descrito anteriormente y teniendo en cuenta los objetivos de investigación que más adelante serán descritos, se plantea como **hipótesis** que: al optimizar el método fdMD, específicamente al mejorar el muestreo conformacional, se lograrán mejores resultados en la predicción tanto del sitio de unión como de la pose.

OBJETIVOS



OBJETIVOS

Este trabajo de investigación tiene como objetivo principal comprobar si se pueden mejorar los resultados, en cuanto a predicción de sitio de unión de ligandos, modo de unión, y detección de falsos sitios de unión, mediante el estudio de uno de los tres sets antes mencionados con una modificación aplicada a la metodología fdMD centrada en la mejora del muestreo conformacional con el uso de H-REMD. Derivándose los siguientes objetivos específicos:

1. Analizar las dinámicas moleculares H-REMD obtenidas para diferentes sistemas proteína-ligando para comprobar si están estabilizadas.
2. Clusterizar las mismas para determinar los diferentes sitios de unión visitados por los ligandos.
3. Realizar cálculos MMGBSA y asignar una energía libre de unión para cada clúster.
4. Proponer otros descriptores para, junto a la energía MMGBSA, poder concluir cuantitativamente sobre la bondad de las tres metodologías comparadas en el trabajo.

MATERIALES Y MÉTODOS



MATERIALES Y MÉTODOS

Campos de fuerza

El término mecánica molecular (MM) se fijó en la década de 1970 para describir la aplicación de la mecánica clásica a la determinación de estructuras moleculares de equilibrio. El método se conocía como el método Westheimer o el método de campo de fuerza. El nombre y el acrónimo, MM, ahora son conceptos firmemente establecidos.⁽¹⁵⁾

La idea es tratar el **“campo de fuerza como un conjunto de constantes que deben corregirse apelando al experimento o a un cálculo más riguroso”**⁽¹⁵⁾. En MM, es obligatorio considerar las interacciones no enlazantes, y las características intrínsecas a cada átomo teniendo en cuenta los átomos vecinos.

Un campo de fuerza está definido por la expresión matemática (Ecuación 1) usada para determinar la energía del sistema + los parámetros usados dentro de dicha expresión (los valores de equilibrio de los enlaces y ángulos, por ejemplo). Al proceso de encontrar dichos parámetros se le denomina parametrización. Todos los campos de fuerza determinan la energía del sistema mediante términos intramoleculares e intermoleculares (Ecuación 2).

$$E = \text{Enlace} + \text{Ángulo} + \text{Diedro Impropio} + \text{Cruzados} + \text{Electrostática} + \text{van der Waals}$$

Ecuación 1. Definición de campo de fuerza

$$E = \text{Intramolecular} + \text{Intermolecular}$$

Ecuación 2. Definición de campo de fuerza

Los campos de fuerza tienen parámetros para diferentes tipos de átomos. Otro problema es que las constantes de enlace, ángulo y torsionales no son completamente independientes, lo cual genera la necesidad de términos cruzados. Por tanto, no hay una forma correcta para la ecuación que describe un campo de fuerza...la precisión, versatilidad y velocidad de cálculo serán los factores necesarios para decidir cuál usar.

La presencia de términos cruzados en un campo de fuerza refleja el acoplamiento existente entre coordenadas internas. En principio habría que incluir términos cruzados entre todas las contribuciones al campo de fuerza, aunque generalmente solo son necesarios algunos para reproducir las características estructurales y/o dinámicas. Los términos cruzados incluyen dos

tipos de coordenadas internas (enlace-enlace, enlace-ángulo...) aunque se han usado también términos con más tipos de coordenadas (ángulo-ángulo-diedro, por ejemplo).

La presencia de términos cruzados se utiliza como un tipo de clasificación de los campos de fuerza, de modo que se puede hablar de campos de fuerza de clase I (solo términos armónicos, sin términos cruzados), clase II (términos anarmónicos y términos cruzados) y clase III (clase II que además tiene en cuenta efectos químicos tales como electronegatividad e hiperconjugación). También se clasifican atendiendo al ámbito de aplicación: orgánicos, biológicos, inorgánicos y universales. En esta investigación se utiliza un campo de fuerza que se describe más adelante y según estas clasificaciones pertenece a la clase I y es de tipo biológico.

Cuando se va a desarrollar un trabajo o investigación usando MM es de vital importancia elegir un campo de fuerza que se ajuste al procedimiento que se va a seguir, al tipo de sistema y al observable que se espera obtener. Este campo de fuerza va a obedecer al tipo de sistema para el que se diseña dicho campo de fuerza.

AMBER (de sus siglas en inglés, *Assisted Model Building with Energy Refinement*) es un campo de fuerza optimizado para la simulación de ácidos nucleicos y proteínas. Fue calibrado para longitudes y ángulos de enlace experimentales obtenidos de microondas, difracción de neutrones y estudios químico-cuánticos precisos. Los parámetros se refinaron luego con estudios de mecánica molecular en compuestos modelo tales como tetrahidrofurano, desoxiadenosina, fosfato de dimetilo, complejo unido por hidrógeno de 9-metiladenina-1-metiltilamina y otros.⁽¹⁵⁾

La siguiente ecuación muestra la expresión para la energía de un campo de fuerza tradicional utilizado en la mecánica molecular.

$$U = \sum_{\text{enlace}} U^{srt} + \sum_{\text{ángulo}} U^{bend} + \sum_{\text{torsiones}} U^{for} + \sum_{i,j} U_{ij}^{LJ} + \sum_{i,j} U_{ij}^{Coulomb}$$

Ecuación 3. Ecuación generalizada del campo de fuerza ⁽¹⁶⁾

Donde los tres primeros términos se corresponden con los términos de enlace y los dos últimos corresponden a términos no enlazantes, que son el término de Lennard-Jones y las contribuciones de Coulomb. El primero de estos términos no enlazantes describe la interacción de van der Waals y el segundo las interacciones electrostáticas. ⁽¹⁷⁾

Al desarrollar la ecuación anterior, teniendo en cuenta sólo términos armónicos y sin términos cruzados, se puede obtener:

$$U = \sum_{\text{enlace}} \frac{1}{2} k_b (r - r_0)^2 + \sum_{\text{ángulo}} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{\text{torsiones}} \frac{V_n}{2} [1 + \cos(n\theta - \delta)] + \sum_{LJ} 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) + \sum_{\text{Coulomb}} \frac{q_i q_j}{r_{ij}}$$

Ecuación 4. Ecuación generalizada del campo de fuerza desarrollada ⁽¹⁶⁾

Al desarrollar la ecuación se observa como los dos primeros términos, que corresponden a las energías de enlace y de ángulo, respectivamente, se representan como funciones armónicas simples que controlan la longitud y ángulo de los enlaces covalentes siendo r la distancia de enlace y θ el ángulo. El tercer término corresponde a la energía de torsión (diedro) y generalmente se representa por una función coseno donde θ es el ángulo de torsión, δ es la fase, n define el número de mínimo o máximo entre 0 y 2π y V_n determina la altura de la barrera de potencial a la torsión. El cuarto término corresponde al potencial de Lennard-Jones donde ϵ representa la profundidad del potencial, σ es la distancia (finita) en la que el potencial entre partículas es cero y r es la distancia entre partículas. El término final (Coulomb) es el que describe las interacciones electrostáticas; asignando una carga parcial a cada núcleo, y usando la ley de Coulomb se calcula la contribución a la energía total.

Dinámica Molecular

Las diferentes etapas de un proceso de simulación por dinámica molecular se resumen en el siguiente esquema (modificado de ⁽²⁾):

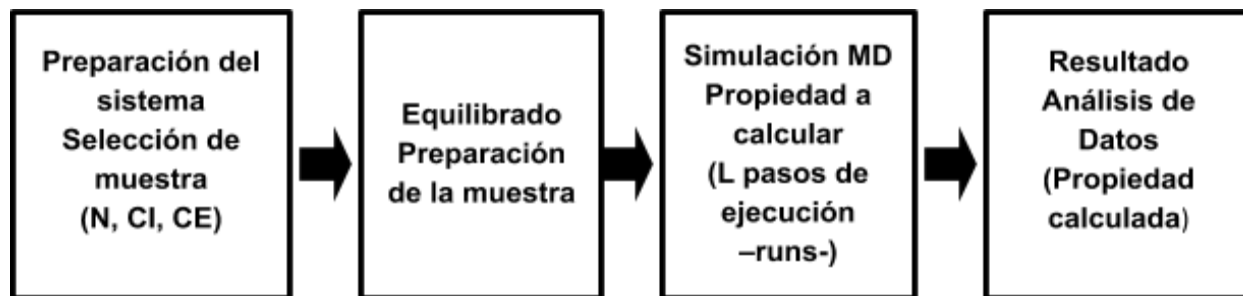


Figura 1. Etapas de un proceso de simulación por dinámica molecular

Primeramente se necesita establecer una configuración inicial (CI) del sistema a utilizar compuesta por N átomos y definir las condiciones de equilibrio (CE). Esta puede proceder de datos experimentales (cristalografía de rayos X, RMN), de modelos teóricos o de una combinación de ambos. Ya optimizada la estructura, se asignan de forma aleatoria velocidades

iniciales a cada uno de los átomos que componen el sistema a partir de la distribución de Maxwell-Boltzmann a una determinada temperatura. Posteriormente es acertado mantener fijas determinadas condiciones de simulación como son: número de partículas (N), volumen (V), temperatura (T), presión (P) o energía total del sistema (E), pH, entre otras. Se puede combinar más de un parámetro fijo, dando lugar a simulaciones microcanónicas (NVE), isotérmico-isobáricas (NPT) y canónicas (NVT), siendo NPT y NVT las más empleadas en MD de proteínas y ácidos nucleicos.⁽¹⁾

Uno de los parámetros importantes a decidir en una simulación de MD es el tiempo de simulación. El tiempo de simulación ha de ser del orden del tiempo en que se da el proceso a simular, que puede ir desde algunos pocos picosegundos a microsegundos. Se consiguen las propiedades fisicoquímicas de interés al efectuar el cálculo de fuerzas del sistema. Como regla general, a mayor tiempo de simulación, mayor será el espacio conformacional a explorar y por lo tanto los resultados obtenidos serán mucho más precisos, de acuerdo con la aproximación ergódica (Ecuación 5), que establece que el promedio tomado durante un largo período de tiempo y en una población más pequeña debería ser igual al promedio en menos tiempo y en una población más grande.

$$\langle A_{tiempo} \rangle = \frac{1}{M} \sum_i^M A(t_i)$$

Ecuación 5. Ecuación de la ergodicidad

El concepto de ergodicidad fue introducido por Ludwig Eduard Boltzmann⁽¹⁸⁾ en física estadística en su teoría cinética de los gases. Si un sistema para $t=0$ está en un punto del espacio de fase “razonable” su evolución en el tiempo probablemente visitará regiones relevantes del espacio de fase.

Es importante aclarar que los observables son medias en el espacio de fase, no medias temporales. La hipótesis ergódica establece:

$$\langle A_{espacio\ fase} \rangle = \langle A_{tiempo} \rangle$$

Ecuación 6. Hipótesis ergódica

Se cumple para $t \rightarrow \infty$, lo que implica que eventualmente el sistema pasará por todos los puntos importantes del espacio de fase.

Existen diferentes programas, tanto gratuitos como de pago, que permiten: cálculos de parámetros del campo de fuerza, realización de la simulación propiamente dicha, y análisis de los resultados obtenidos.⁽²⁾

Dinámica Molecular Clásica (cMD)

En cMD los cálculos comienzan con un primer conjunto de coordenadas para el sistema, que se obtiene generalmente a partir de datos de rayos X o RMN, y un conjunto inicial de velocidades que satisfacen la distribución de Maxwell-Boltzmann (Ecuación 7) para la temperatura objetivo, donde la temperatura y las velocidades se relacionan mediante la Ecuación 8. La influencia del disolvente puede determinarse mediante métodos implícitos: Generalized Born⁽¹⁹⁾, Poisson-Boltzmann⁽²⁰⁾, o explícitos, solvatando un sistema con un modelo de molécula de agua (TIP3P, TIP4P, etc.)⁽²¹⁾

$$v_i = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{1}{2}} \cdot e^{\left(-\frac{m_i \cdot v_i^2}{2\pi k_B T} \right)}$$

Ecuación 7. Distribución de Maxwell-Boltzmann

$$T = \frac{2}{3k_B} (E_k) = \frac{m}{3k_B} (v^2)$$

Ecuación 8. Relación entre la temperatura absoluta y su energía cinética

Donde m_i es la masa de la partícula, k_B es la constante de Boltzmann, T se refiere a la temperatura, v a la velocidad de la partícula y E_k la energía cinética. Las nuevas coordenadas y velocidades se obtienen resolviendo las ecuaciones de Newton (Ecuación 9) mientras que la temperatura se mantiene constante mediante el método de Langevin.^(5,22)

$$\frac{d\vec{v}_i(t)}{dt} = \frac{\vec{F}_i(t)}{m_i} \text{ siendo } F_i(t) = - \frac{\delta v(r_1(t), r_2(t), \dots, r_n(t))}{\delta r_i(t)}$$

Ecuación 9. Ley de Newton

Donde i , incluye todos los átomos del sistema ($i = 1, 2, 3, \dots, n$), siendo n el número total de átomos. Las coordenadas atómicas, r_i , y la velocidad, v_i , del átomo i , con la masa, m_i , se convierten en funciones del tiempo. El modelo matemático utilizado (campo de fuerza), se

describe cómo el valor de la energía potencial (V) dependiente de la disposición espacial de todos los átomos.⁽²³⁾

Dinámica molecular de intercambio de réplicas (REMD)

En las simulaciones REMD, varias copias (réplicas) del mismo sistema se simulan en paralelo utilizando, entre otras opciones, simulaciones MD a diferentes temperaturas, o a la misma temperatura utilizando diferentes hamiltonianos. El intercambio entre réplicas vecinas se intenta periódicamente con una probabilidad dada por el criterio de Metrópolis (Figura 2). De esta manera, REMD es capaz de superar fácilmente las barreras de alta energía y muestrear el espacio conformacional de modo más eficiente que mediante cMD.⁽⁴⁾

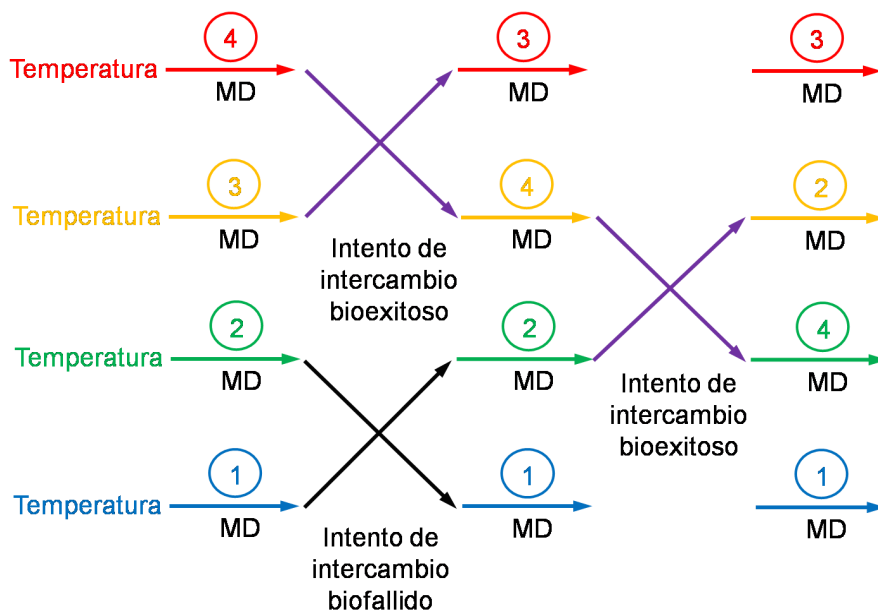


Figura 2. Representación de una simulación REMD, modificado de Qi ⁽⁴⁾

Dinámica molecular de intercambio de réplicas H-REMD

En las simulaciones H-REMD, durante las réplicas se modifican diferentes términos de los campos de fuerza. Los intercambios que tienen lugar entre simulaciones controladas por diferentes campos de fuerza permiten que la exploración conformacional se realice a través de diferentes réplicas, logrando al final un aceleramiento en dicho muestreo conformacional. H-REMD puede superar a T-REMD porque se centra en un grado de libertad relevante de un sistema, en lugar de afectar a todos los átomos del sistema como ocurre en T-REMD.⁽²⁴⁾

En otros estudios de acoplamiento basados en H-REMD, las interacciones de Coulomb y Lennard-Jones entre el ligando y el receptor se escalan a lo largo de la escalera de réplica en disolvente implícito o explícito. Sin embargo, el debilitamiento de las interacciones de atracción entre el ligando y el receptor en dicho enfoque puede dar como resultado el muestreo de muchos estados insignificantes en las réplicas superiores.⁽²⁴⁾

Energía libre de unión MMPB/GBSA

La metodología MMPB(GB)SA (*Molecular Mechanics Poisson-Boltzmann /Generalized Born Surface Area*) es un método propuesto para estimar la energía libre de la unión de pequeños ligandos a macromoléculas biológicas. Estos cálculos se basan en simulaciones de dinámica molecular del complejo receptor-ligando y están a medio camino entre los cálculos empíricos y los métodos de perturbación química, con un coste computacional más elevado..^(25,26)

Uno de los métodos computacionales más utilizados en el diseño de fármacos es el *docking* seguido del *scoring*, mediante los cuales se predice el modo de unión del fármaco, seguido de una estimación de la afinidad de unión. Estos métodos son rápidos, pero no particularmente precisos; pueden usarse para predecir modos de unión y diferenciar ligandos.⁽²⁶⁾

MM/GBSA supera a la MM/PBSA, debido a la rapidez del primer método que intenta reproducir en un menor tiempo los resultados de MM/PBSA. Por ejemplo, el método MM/GBSA ha sido ampliamente utilizado en cálculos de energía libre: un *rescoring* usando MM/GBSA produce mejores resultados que un *docking* sin posterior *rescoring* para el conjunto de datos DUD-E (Directory of Useful Decoys, Enhanced).⁽²⁷⁾

En este método, la energía libre de unión, $\Delta G_{\text{unión}}$, se estima a partir de las energías libres de los reactivos y el producto de la reacción (Ecuación 10).⁽²⁶⁾

$$\Delta G_{\text{enlace}} = \langle G_{\text{PL}} \rangle - \langle G_{\text{P}} \rangle - \langle G_{\text{L}} \rangle$$

Ecuación 10. Energía libre de unión estimada partir de las energías libre de los reactivos

En MM/PB(GB)SA, la energía libre de un estado, es decir, proteína (P), ligando (L) o complejo proteína-ligando (PL), se estima a partir de la siguiente expresión:^(28,29)

$$G = E_{enlace} + E_{el} + E_{vdW} + G_{pol} + G_{np} - TS$$

Ecuación 11. Suma de la energía libre de cada estado

Donde los tres primeros términos son términos energéticos de MM: enlace, interacciones electrostáticas y de van der Waals. Los términos Gpol y Gnp representan las contribuciones polares (Gpol) y no polares (Gnp) a la energía libre de solvatación. El término entrópico final es estimado por un análisis de las frecuencias de vibración.⁽³⁰⁾

El término de Gpol se obtiene de la resolución de la ecuación 12 mientras que el término no polar se estima a partir de una relación lineal con el área de superficie accesible al solvente (SASA).⁽³⁰⁾

$$G_{pol} = -\frac{1}{2} \sum_{i,j} \frac{q_i q_j}{f_{GB}(r_{ij}, R_i, R_j)} \left(1 - \frac{\exp[-k f_{GB}]}{\epsilon} \right)$$

Ecuación 12. Aproximación de Gpol mediante el modelo generalizado de Born

Dinámica molecular de fragmentos disueltos (fdMD)

Diagrama de flujo de fdMD.

El método fdMD usando REMD utiliza la suite de programas Amber18⁽⁵⁾. Como paso previo el ligando se solvata con moléculas de agua TIP3P⁽²¹⁾ utilizando el módulo LEaP de Amber18⁽⁵⁾. La caja solvatada se somete a minimización seguida de MD, y la instantánea final de esa ejecución de MD se exporta para ser utilizada como caja de solvatación. Así, la proteína se solvata con LEaP usando la caja de solvatación equilibrada, de modo que el sistema final constará de la proteína y un número diferente (N) de ligandos asociados a la proteína. Después de un protocolo de minimización y equilibrado, se obtiene una REMD de 25 ó 50 ns de longitud, según el método que se ensaye.

Cada trayectoria se analiza por medio de 3 *scripts*. El primero (fdMD_OneTraj.cmd) elimina las moléculas de agua y los contraiones de la trayectoria, divide la trayectoria global en N trayectorias de proteína-ligando (cada trayectoria está formada por la proteína y un solo ligando) y escribe la última instantánea de cada una de las N trayectorias en formato PDB. Los *scripts* 2 y 3 (fdMD_MMGBSA_Send.cmd y fdMD_MMGBSA_Analize.py, respectivamente)

generan los ficheros de entrada necesarios para ejecutar un cálculo MMGBSA usando la suite Amber18 y obtener y graficar los datos de esos cálculos.

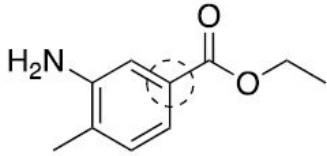
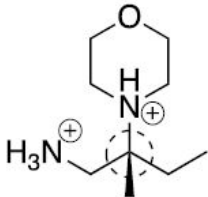
Posteriormente se realiza el clustering, en este caso de tipo dbscan.sh. Se realizó en 2 partes: step 1 y step 2. Con el primer *script* cpptraj_clustering_dbscan_step1.sh se generan las gráficas kdist.dat las cuales permiten determinar los parámetros epsilon y minpoints (ver tablas 2 y 3) necesarios para ejecutar el segundo script cpptraj_clustering_dbscan_step2.sh con el que finalmente quedan hecho los *clustering*.

Una vez terminado se examinan los posibles sitios de unión ocupados, mediante los archivos de salida resultantes de los *scripts* (info.res, clúster.res y result.res). Para la predicción de sitios de unión se propuso el uso de tres descriptores. Los descriptores propuestos al usar H-REMD en fdMD fueron *cluster members* (el número total de miembros que puebla cada clúster), la energía libre de unión ΔG (kcal/mol) considerando como una trayectoria a todos los miembros que pueblan cada clúster y la energía de enlace K_{DEEP} (kcal/mol) del representante de cada uno de los clústeres obtenidos, que se calculó con la ayuda de <https://www.playmolecule.org/>; del total de clústeres que se obtienen por cada método solamente se toma como representativos los clústeres de poses con población mayor o igual al 20% del clúster más poblado de cada método y para cada ligando.

Sistemas de proteína-ligando estudiados ⁽⁹⁾

Para testear el método fdMD con H-REMD se emplearon los ligandos F031 y F035 que forman parte del primer set de sistemas proteína-ligando (ver Tabla 1). De los átomos de carbono central se seleccionó uno y se renombró como C99 al que se le aplicó un término de repulsión necesario para evitar la agregación de ligandos. El set I se estudia para comprobar si fdMD es lo suficientemente robusto a la hora de identificar el sitio de unión correcto sin importar cuál sea la estructura de la proteína empleada.

Tabla 1. Sistemas de proteína-ligando estudiado, modificado de Privat y col. ⁽⁹⁾

SET I	Ligando ^a	Proteína	Etiqueta	PDB ID	Ligandos ^b
		ENDOTIAPEPSINA	F031	5DPZ ⁽³¹⁾	9
			F035	4Y3W ⁽³¹⁾	12

a. El círculo de línea discontinua indica el átomo C99 (ver más arriba)

b. Número de ligandos en la caja de simulación.

Preparación de proteínas⁽⁹⁾

La estructura de la proteína se descargó del Banco de datos de proteínas (los códigos PDB se muestran en la Tabla 1). La proteína se cargó en el módulo LEaP de la suite Amber18 ⁽⁵⁾, donde se agregaron contraiones, cuando fue necesario, y la proteína se solvata utilizando moléculas de agua TIP3P⁽²¹⁾. El campo de fuerza utilizado fue ff14SB y la proteína solvatada se minimizó con pmemd utilizando 10000 ciclos de minimización con el algoritmo *steepest descent* (SD) con un *cut off* de 10 Å para interacciones no enlazantes, mientras que las interacciones electrostáticas fueron modeladas con el método Ewald ⁽³²⁾ de malla de partículas.

Preparación y simulación de sistemas solvatados⁽⁹⁾

Como punto de partida para las simulaciones de fdMD se tomó la estructura minimizada de la proteína, sin agua y sin contraiones. La proteína fue solvatada con la caja de solvatación de

ligando equilibrada correspondiente, se eliminaron las moléculas existentes a una distancia menor de 1.0 Å de la proteína, se seleccionó la distancia mínima desde el borde de la caja, teniendo en cuenta el tamaño del ligando y la forma de la proteína, por lo que se colocó una cantidad razonable de ligandos en la caja de simulación, utilizando valores de 11 a 20 Å. Finalmente, se agregaron contraiones cuando fue necesario; esto solo se realiza para la metodología 4 (ver más abajo descripción de los métodos).

El campo de fuerza ff14SB se utilizó para cálculos posteriores y ParmEd para modificar la topología con el objetivo de utilizar el reparto de masa de hidrógeno (HMR) para poder utilizar períodos de integración de 4 fs y agregar un potencial de repulsión de Lennard-Jones, poniendo a cero la parte atractiva ($B_{C99-C99} = 0$), entre los átomos C99 de los ligandos, con $= 0.01 \text{ kcal} \cdot \text{mol}^{-1}$ y $R_{\min} = 23.0 \text{ Å}$.

Descripción de los métodos utilizados

Se utilizaron 3 métodos, dependiendo del número de ligandos y el método para colocarlos en la caja de simulación.

Método 1: El ligando se coloca aleatoriamente y se realizan 8 réplicas que empiezan con la misma posición del ligando. El método 1 se emplea como un control.

Método 2: Se utiliza Autodock Vina⁽³³⁾ para realizar un docking del ligando utilizando parámetros por defecto y usando la proteína completa como caja de búsqueda. Las estructuras ya tienen aplicado el potencial de Lennard-Jones repulsivo en C99 para evitar la aglomeración. Del docking obtenido se seleccionan las 8 poses mejores, cada réplica está definida por la variación en los términos del hamiltoniano. En el método 2, los ligandos ya están en BSs de la proteína y se emplea para ver si un *docking* previo es mejor que colocar los ligandos aleatoriamente como se hace en el método 3.

Método 3: Es el método fdMD propiamente dicho. Se toma como estructura a trabajar las estructuras solvatadas y pre-equilibradas de fdMD, ya con el potencial repulsivo en C99 de Lennard-Jones para evitar la aglomeración. Así se multiplican las opciones, puesto que se trataría de que cada réplica tendría entre 9 y 26 ligandos a la vez.

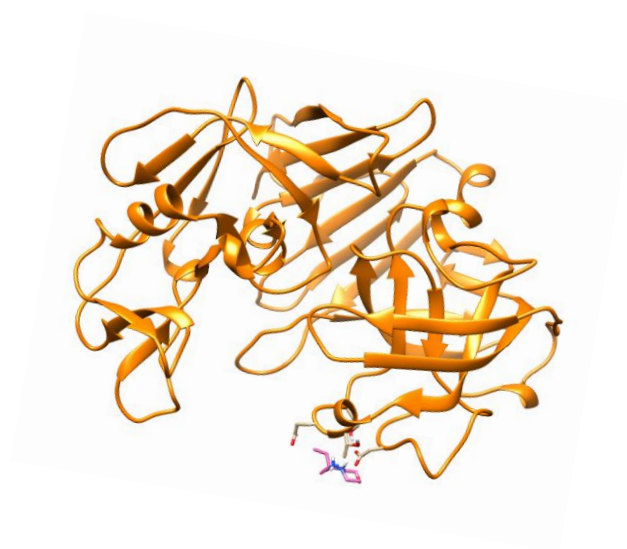
Para el método 1 los cálculos REMD se ejecutan con cut-off de 10 Å, 4 fs de tiempo de integración, $n_{\text{stlim}} = 125$ (número de pasos antes de cada cambio en las réplicas), y n_{umexchg}

= 100000 (intentos de intercambio entre réplicas, ver Figura 2), con lo tenemos 50 ns de producción en cada réplica.

Para los métodos 2 y 3 los cálculos REMD se ejecutan con cut-off de 10 Å, 4 fs de integración, $n_{\text{stim}} = 125$, y $n_{\text{exchg}} = 50000$, con lo tenemos 25 ns de producción en cada réplica.

El propósito del proyecto en resumen es realizar H-REMD utilizando como parámetros a variar en las réplicas el potencial de Lennard-Jones y el radio de van der Waals para las interacciones proteína-ligando. El objetivo de esta técnica es acelerar el muestreo conformacional y predecir el *binding site* y la *binding pose* de ligandos. Para los ligandos estudiados F031 y F035 se probó la metodología fdMD mediante un sistema de ligando-proteína donde la estructura de la proteína era la estructura experimental del ligando unido a esa proteína (marcado con el nombre del ligando).

RESULTADOS



RESULTADOS

A la hora del análisis de los resultados con el software Chimera⁽³⁴⁾ se visualizan los sitios a donde se dirige el ligando y además se obtiene el listado de residuos de aminoácidos de la proteína que se encuentran a una distancia menor que 3.6 Å de ligando para cada uno de los clústeres representativos, 13 en total (ver tablas 2 y 3). Luego mediante el análisis de los tres descriptores (*cluster members*, ΔG y K_{DEEP}) se le da una puntuación a cada BS del listado para comprobar si los mejores resultados corresponden al sitio de unión experimental y así poder sugerir, en nuevos casos problemas donde no se conozca el sitio de unión experimental que con el cálculo de estos descriptores se es capaz de predecir cuál es el BS correcto donde se une un determinado ligando a su correspondiente proteína.

Es necesario aclarar que el BS se refiere al sitio de unión del ligando con la proteína diana y que la **pose** es la conformación adoptada por el ligando al interactuar con el BS; un ligando puede encontrar el BS experimental pero con una conformación diferente a la pose experimental, para ello se calcula la RMSD mínima (*root mean square deviation*, de sus siglas en inglés), medida de la exactitud que puede tener el *docking* a la hora de predecir una pose óptima. Cuando este valor es menor a 2 Å se considera que la pose es correcta ^(5,35). En los casos en los que el ligando encuentra su sitio experimental se calculó también la RMSD mínima para saber si el ligando no solo es capaz de llegar al sitio de unión correcto, sino que además lo hace en la pose correcta, indicando entonces que el acoplamiento reproduce correctamente la forma de unión y las interacciones presentes entre el ligando y la proteína en ambos casos.

Según el total de veces que llega un ligando a un sitio determinado se obtienen, para cada uno de los ligandos en cada método, diferentes clústeres representativos. A partir del clúster original se obtiene un conjunto de clústeres representativos que indican las zonas donde más tiempo permanece el ligando. Para determinar dichos clústeres se selecciona el elemento con mayor número de *snapshots* del clúster obtenido y todos aquellos elementos cuyo valor de *snapshots* sea mayor o igual al 20% ^(5,35) de dicho elemento.

Cuando el número de miembros del clúster representativo es elevado significa que el ligando ha estado por esa zona durante tiempo suficiente para considerar que las interacciones que

tienen lugar en ese sitio son buenas, o sea las energías del sistema son bajas y se toma esta zona como zona de interés.

La energía de unión (ΔG) es uno de los descriptores para medir la afinidad existente entre el ligando y la proteína una vez formado el complejo. ΔG se expresa en kcal/mol y tendrá un valor negativo que se traduce en una mayor afinidad del sistema proteína-ligando.

La energía de enlace K_{DEEP} es la encargada de predecir la afinidad del complejo ligando-proteína basándose en las redes neuronales convolucionales 3D de última generación⁽³⁶⁾. Varios autores apuntan a K_{DEEP} como el mejor indicador para un sistema de *scoring* por la velocidad, el rendimiento y la facilidad de obtener este parámetro.

Ligando F031

Método 1

Para el ligando F031 método 1 en lugar de encontrar la pose experimental se observa como los dos representativos de los clústeres más poblados (Tabla 2) indican nuevos BSs. Estos BS 1 y 2 (Figuras 3 y 4, respectivamente) al ser diferentes del BS experimental⁽³¹⁾ para el ligando F031 son considerados como falsos positivos.

Método 2

El ligando F031 con el método 2 encuentra su pose experimental, S1' (Figura 5) y al analizar los valores de los descriptores (Tabla 2) se observa un solo clúster representativo, muy poblado y además los valores energéticos de ΔG y K_{DEEP} son valores bajos demostrando que los enlaces del complejo ligando-proteína son estables y con una alta afinidad. En este caso donde el ligando encuentra su BS experimental se comprueba los resultados del *docking* calculando la RMSD mínima y se obtiene un valor igual a 0.86 Å, así que efectivamente el ligando con esta metodología no solo es capaz de encontrar su BS experimental, sino que se une a él con la orientación correcta. Por tanto, se podría afirmar que para el ligando F031, la metodología 2 es lo suficientemente robusta para favorecer la unión entre el ligando y la proteína en el sitio de unión y en la pose correcta.

Método 3

Para el ligando F031 método 3 se obtuvieron cuatro clústeres representativos. El primer clúster (Tabla 2) es el único que encuentra el BS experimental, S1' (Figura 6), es el más poblado con 200 miembros por lo que el ligando ha estado por esa zona durante tiempo suficiente para considerar que las interacciones que tienen lugar en ese sitio son buenas, cuando se comparan sus valores energéticos estos son valores bajos por lo que el complejo debería ser estable y con una alta afinidad. En este caso también se calculó la RMSD mínima (1.32 Å) para comprobar que la metodología 3, al igual que la 2, no solo encuentra el BS experimental, sino que además el ligando llegue a su sitio experimental en la forma correcta.

El segundo clúster representativo (Tabla 2) está compuesto por miembros que describen al ligando en un sitio de unión que, aunque difiere del experimental, es un BS ya descrito previamente por Radeva y col.⁽³¹⁾, S6 (Figura 7). Podría para este caso aumentarse el tiempo de la dinámica para saber si el ligando permanece en el sitio S6 o se adentra en la proteína hasta llegar a su sitio de unión experimental S1' (Anexo 1), por el momento se clasifica como un falso positivo, este resultado concuerda con Privat y col.⁽⁹⁾ donde descartan al *pocket* S6 y lo clasifican como un sitio espurio.

Los 2 últimos clústeres (Tabla 2) que se obtienen con este método corresponden con nuevos sitios de unión definidos en el trabajo como BS 3 y 4 (Figuras 8 y 9, respectivamente). Al comparar los valores de los descriptores de estos dos BS con los valores de los clústeres anteriores cuantitativamente parecen peores por lo que serían sitios menos favorecidos que se clasifican entonces como falsos positivos.

Al constar esta metodología con cuatro clústeres representativos y que uno de ellos termine con el ligando en el BS experimental (S1') permite confrontar los resultados de cada descriptor para cada clúster y comprobar si en el caso donde no se conozca el sitio de unión experimental comparando los valores obtenidos para los descriptores se es capaz de distinguir cuál del total de clústeres representativos que se tenga corresponderá al BS experimental y cuáles representarán sitios espurios. Si se analizan los valores correspondientes a la metodología 4 (Tabla 2) se observa, efectivamente, que el mejor valor del descriptor *cluster members* coinciden con el BS experimental, S1' (Figura 6) y que además el valor de ΔG sale bastante bien; o sea de los tres descriptores propuestos 2 nos indican que S1' es el sitio más favorecido; por lo que parece indicar que bastaría con comparar los valores de los descriptores propuestos,

en la metodología fdMD, para predecir el BS experimental (que sería el sitio de unión más favorecido) y los sitios que representan falsos positivos. Señalar que el descriptor *cluster members* es el más robusto ya que sale siempre mejor para la pose experimental pero para poder hacer cualquier afirmación se hace obligatorio alargar el tiempo de simulación de la dinámica y ampliar el set estudiado para poder tener una mayor replicabilidad en los resultados y poder validar el método.

Tabla 2. Resultados obtenidos para el ligando F031 con cada metodología.

Método 1	Pocket	<i>Cluster members</i>	ΔG (kcal/mol)	K_{DEEP} (kcal/mol)
Experimental S1'	BS 1	811	-9.51	-4.87
epsilon=1.5	BS 2	330	-9.74	-4.72
minpoints=2				
Método 2	Pocket	<i>Cluster members</i>	ΔG (kcal/mol)	K_{DEEP} (kcal/mol)
Experimental S1'	Experimental S1'	943	-9.69	-6.05
epsilon=1.5				
minpoints=4				
Método 3	Pocket	<i>Cluster members</i>	ΔG (kcal/mol)	K_{DEEP} (kcal/mol)
Experimental S1'	Experimental S1'	200	-6.56	-5.26

epsilon=1.5	S6	118	-6.65	-6.48
minpoints=4	BS 3	51	-1.57	-5.76
	BS 4	45	-3.02	-5.49

Listado de residuos de aminoácidos para cada clúster y su correspondiente figura

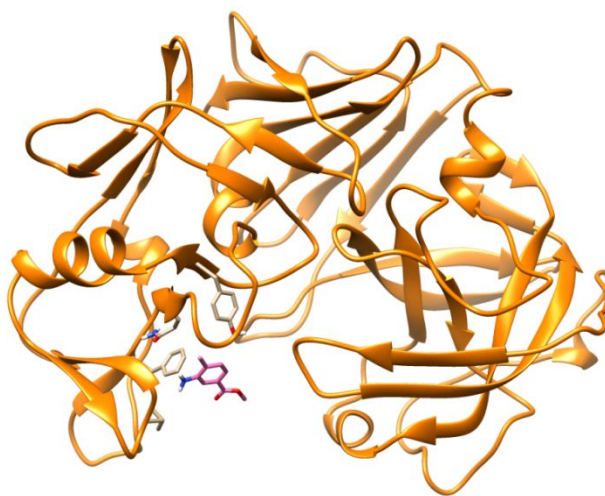


Figura 3. Sitio de interacción del ligando F031 con la endotiapepsina por el método 1 (BS 1)

BS 1: TYR226 VAL248 GLY249 GLY250 PHE291 GLN295



Figura 4. Sitio de interacción del ligando F031 con la endotiaepsina por el método 1 (BS 2)

BS 2: ALA187 VAL188 SER189 THR190 GLU196 TRP197
 THR198 ASP215 GLY216

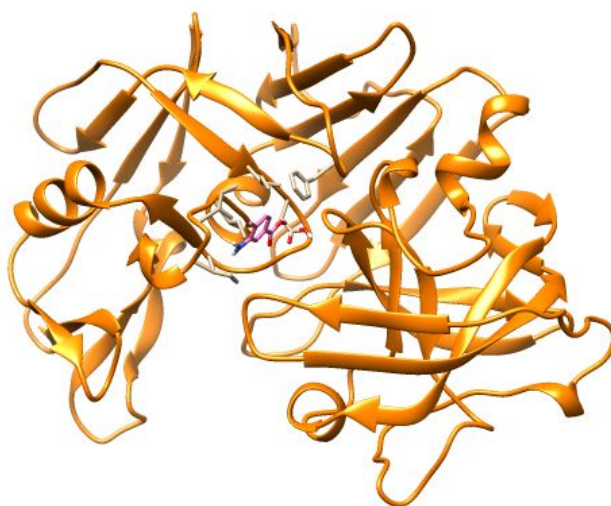


Figura 5. Sitio de interacción del ligando F031 con la endotiaepsina por el método 2
 (Experimental S1') ⁽³¹⁾

Experimental S1': GLY37 GLY80 PHE194 ILE217 ASH219 ILE 300
 GLY301 ILE302 ILE304

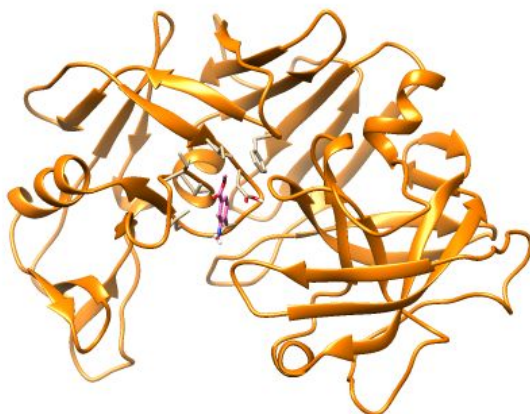


Figura 6. Sitio de interacción del ligando F031 con la endotiapepsina por el método 3
 (Experimental S1') ⁽³¹⁾

Experimental S1': GLY80 PHE194 ILE217 ASH219 THR222 ILE300
 ILE302 ILE304

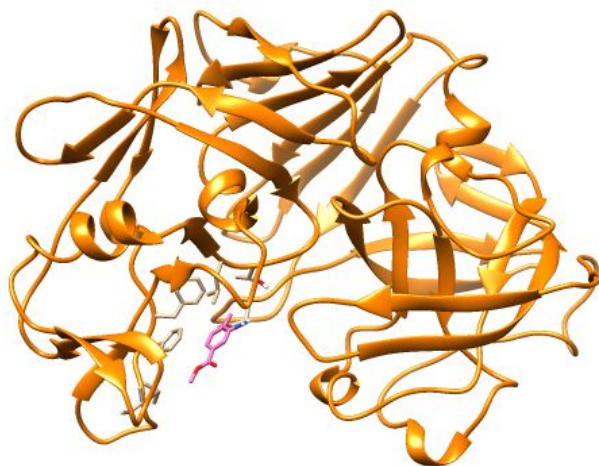


Figura 7. Sitio de interacción del ligando F031 con la endotiapepsina por el método 3 (S6) ⁽³¹⁾

S6: ASH15 THR223 LEU224 VAL248 PHE280 ILE283 PHE291

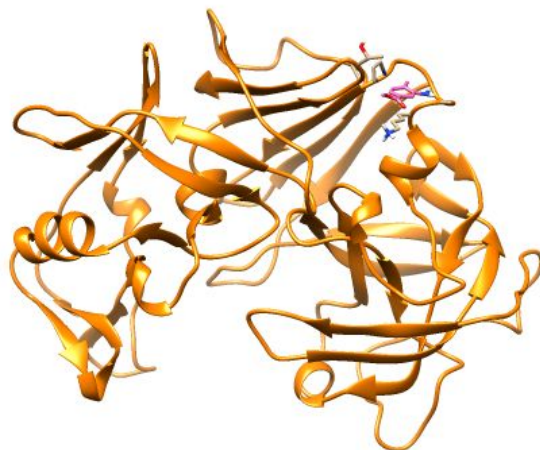


Figura 8. Sitio de interacción del ligando F031 con la endotiapepsina por el método 3 (BS 3)

BS 3: LYS149 ALA150 PRO155 GLY319 ALA320 THR321

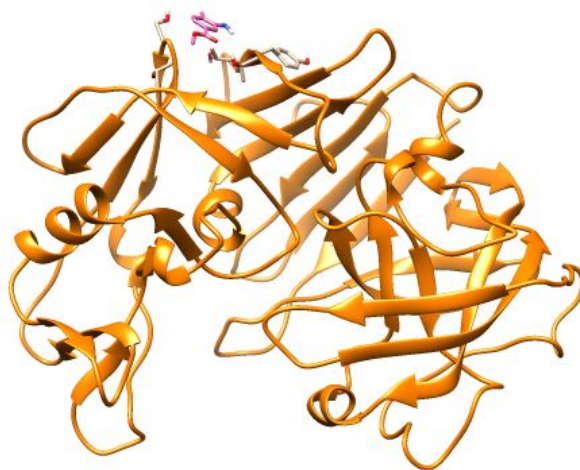


Figura 9. Sitio de interacción del ligando F031 con la endotiapepsina por el método 3 (BS 4)

BS 4: THR184 TYR185 THR186 THR200 GLY267 SER268

Ligando F035

Método 1

Para el segundo ligando estudiado, F035, en el método 1 se obtienen 2 representativos del clúster (Tabla 3) bastante poblados indicando que el ligando permanece por esas zonas por un buen tiempo por lo que se debe considerar para futuras investigaciones de BS nuevos; pero por el momento se consideran como falsos positivos a los BS 5 y 6 (Figuras 10 y 11, respectivamente) ya que no coinciden con los sitios experimentales S3 y S6 (Anexo 1) para este ligando.

Método 2

Para el ligando F035, método 2, se determinan también dos clústeres (Tabla 3) lo suficientemente poblados pero ninguno coincide con el BS experimental (Anexo 1), por lo que S1 y BS 5 (Figura 12 y 13 respectivamente) son considerados como sitios espurios. Se debe señalar que el primer clúster representativo es el que tiene un mayor valor del descriptor *cluster members* para este ligando en general, por tanto el sitio de unión S1 (Figura 12) es donde por más tiempo ha permanecido el ligando y además a este sitio de unión le corresponde el mejor valor de K_{DEEP} , siendo entonces el sitio de unión S1 el más favorecido.

Método 3

Para el ligando F035 método 3 se obtiene también dos clústeres representativos (Tabla 3) con valores que a simple vista nos pueden parecer adecuados en cuanto a número de miembros del clúster y los valores energéticos pero cuando se observa su listado de residuos y sus figuras 14 y 15 y se comparan con los datos experimentales se ve que el ligando se dirige a BS nuevos no a los sitios experimentales por tanto una vez más se consideran a los BS 6 y 7 (Figuras 14 y 15, respectivamente) como falsos positivos.

Tabla 3. Resultados obtenidos para el ligando F035 con cada metodología.

Método 1	Pocket	<i>Cluster members</i>	ΔG (kcal/mol)	KDEEP (kcal/mol)
Experimental S3 y S6	BS 5	663	-36.2	-4.37
epsilon=1	BS 6	980	-41.72	-3.79
minpoints=4				
Método 2	Pocket	<i>Cluster members</i>	ΔG (kcal/mol)	K_{DEEP} (kcal/mol)
Experimental S3 y S6	S1	2142	-27.83	-5.55
epsilon=1.5	BS 5	2062	-28.25	-4.1
minpoints=4				
Método 3	Pocket	<i>Cluster members</i>	ΔG (kcal/mol)	K_{DEEP} (kcal/mol)
Experimental S3 y S6	BS 6	921	-23.07	-3,76
epsilon=1.5	BS 7	479	-15.0	-4.1
minpoints=4				

Para el ligando F035 no se calculó la RMSD porque en ningún caso el ligando encuentra su sitio de unión experimental.

Cuando comparamos los resultados de los clústeres entre ambos métodos para este ligando (Tabla 3) el mejor valor del descriptor *cluster members* y el de K_{DEEP} corresponden en el método 2, al sitio S1 por tanto los descriptores parecen indicar que este sitio es el más favorecido y si se aumentara el tiempo de simulación de la dinámica molecular es muy probable que el ligando encuentre su BS experimental S3 que se encuentra muy próximo a S1 (Anexo1). De acá se puede concluir que los descriptores propuestos serían capaces de predecir el sitio de unión correcto con un tiempo de simulación mayor y que la metodología 2 parece ser la más adecuada.

Listado de residuos de aminoácidos para cada clúster y su correspondiente figura

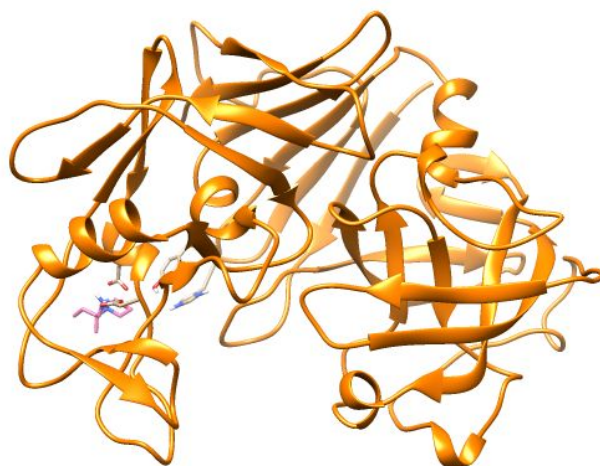


Figura 10. Sitio de interacción del ligando F035 con la endotiapepsina por el método 1 (BS 5)

BS 5: TYR163 HIP164 ASP276 ASP279

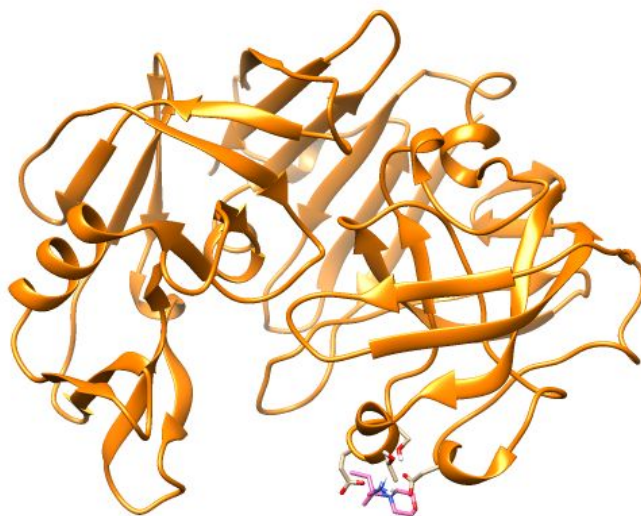


Figura 11. Sitio de interacción del ligando F035 con la endotiapepsina por el método 1 (BS 6)

BS 6: ASP54 GLY55 THR117 GLU118 SER120

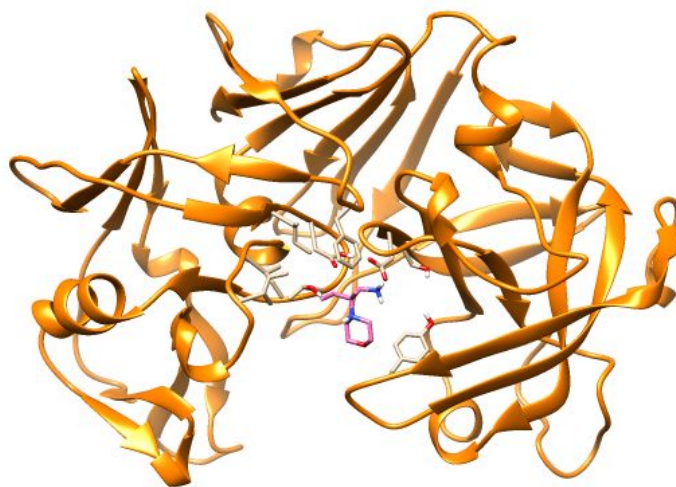


Figura 12. Sitio de interacción del ligando F035 con la endotiapepsina por el método 2 (S1)⁽³¹⁾

S1: ASP35 GLY37 SER38 SER78 TYR79 GLY80 PHE194 ILE217
 ASH219 THR222 ILE302 ILE304

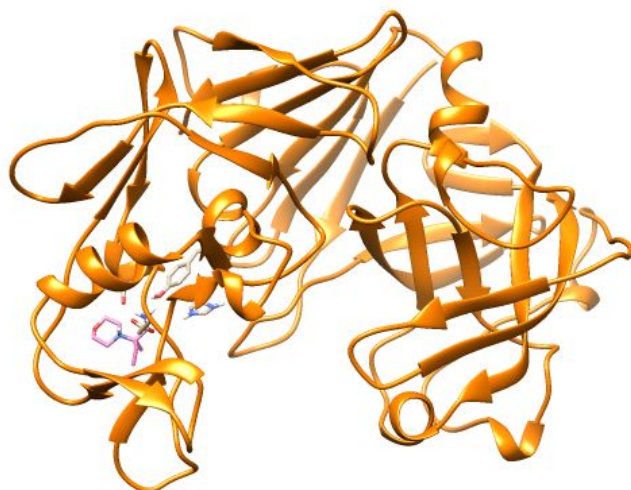


Figura 13. Sitio de interacción del ligando F035 con la endotiapepsina por el método 2 (BS 5)

BS 5: TYR163 HIP164 ASP276 ASP279

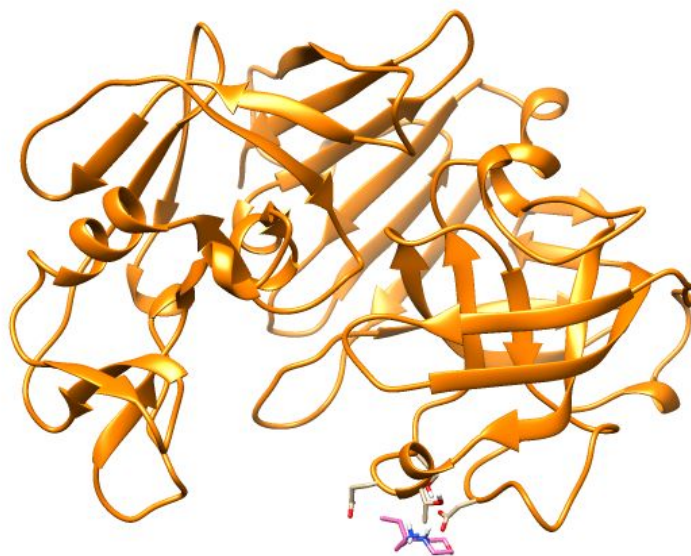


Figura 14. Sitio de interacción del ligando F035 con la endotiapepsina por el método 3 (BS 6)

BS 6: ASP54 GLY55 THR117 GLU118 SER120

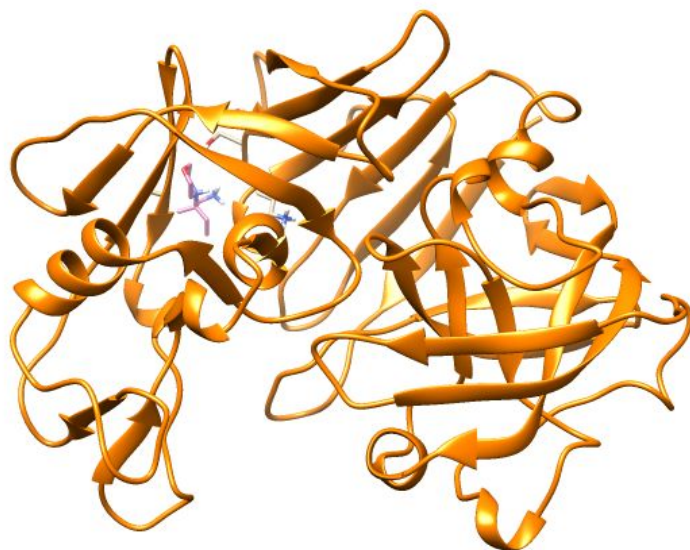


Figura 15. Sitio de interacción del ligando F035 con la endotiapepsina por el método 3 (BS 7)

BS 7: THR180 VAL272 SER329 LYS330

DISCUSIÓN



DISCUSIÓN

De los datos experimentales ⁽³¹⁾ conocemos que el ligando F031 se une a la endotiapepsina en el sitio S1' y el ligando F035 en los sitios S3 y S6. Estos resultados fueron corroborados a través del método fdMD utilizando cMD ⁽⁹⁾. Sin embargo, mediante el método fdMD utilizando H-REMD, después de analizar cada uno de los 13 clústeres representativos que fueron encontrados, se concluye que el ligando F031 encuentra la pose experimental con 2 de las metodologías ensayadas (método 2 y método 3), a diferencia del ligando F035, que no encuentra ninguna de sus poses experimentales. Estos resultados sugieren la necesidad de ampliar la duración de la dinámica molecular en todas las metodologías para comprobar si: con un incremento en el tiempo de simulación ambos ligandos serían capaces de encontrar su BS experimental ya que si se analizan los resultados obtenidos y las condiciones de la dinámica del método fdMD utilizando cMD se observa como con cMD se logró predecir con exactitud y confianza para cada ligando su sitio experimental y diferenciar los falsos positivos gracias a que el tiempo de simulación de la dinámica fue de 200 ns y en los casos dudosos se extendió hasta 1µs, mucho mayor que lo que se simuló con H-REMD, donde el tiempo máximo fue de 50ns.

Según lo obtenido en cada una de las tres metodologías, que difieren en cuanto al número de ligandos y su colocación en la caja de simulación, (descripción de las metodologías en la sección de materiales y métodos) tenemos: con respecto a la metodología 1, que el ligando F031 encuentra 2 nuevos BS no definidos anteriormente en la literatura^(11,12) y con algunos valores similares de los descriptores a otras metodologías en donde el ligando encontró su pose experimental pero aún así fueron considerados falsos positivos. Para el ligando F035 también encuentra 2 BSs nuevos que, mediante el análisis de los descriptores y conociendo la experimental, fueron considerados como falsos positivos. Podemos concluir que el método 1 no fue capaz de encontrar el sitio de unión experimental.

Con las metodologías 2 y 3 el ligando F031 encuentran la experimental, por lo que se concluye que mejorando el muestreo conformacional se obtienen mejores resultados en la predicción de sitios de unión como en la forma en la que se une el ligando la endotiapepsina y permite entonces pasar a otras etapas sucesivas del diseño de fármacos que terminan con la síntesis química de medicamentos que van a ser capaces de unirse a la diana terapéutica y lograr el

efecto farmacológico deseado de forma más rápida y más estable en el organismo, lo que permite una mayor duración del efecto farmacológico.

Las metodologías 2 y 3, para el ligando F035, tenían 2 representativos por cada método, pero en ningún caso a pesar de lo poblado de sus clústeres y sus valores de energías el ligando llega a su sitio de unión experimental por lo que ambas metodologías solo arrojaron falsos positivos. En la metodología 2 a uno de los clúster representativos le corresponden los mejores valores de *cluster members* y K_{DEEP} y el ligando en este caso llega a un sitio de unión muy próximo a uno de sus experimentales por lo que una vez más se demuestra que si se ampliara el tiempo de simulación de la dinámica los ligandos serían capaces de encontrar sus respectivos BSs.

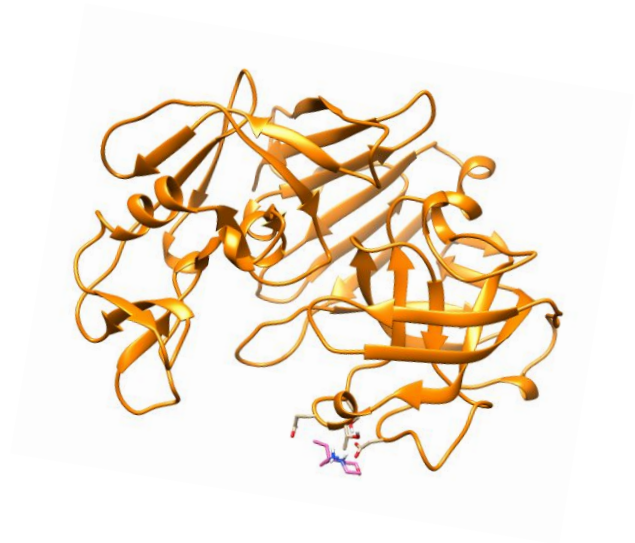
Podemos concluir que para el caso en estudio los descriptores *cluster members*, ΔG y K_{DEEP} son útiles a la hora de puntuar los resultados del *docking* y corroborar si el ligando ha encontrado su BS experimental o si es un falso positivo, aunque se recomienda en investigaciones futuras en donde no se conozca el sitio de unión experimental utilizar además de estos otros descriptores que permitan discriminar entre sitios correctos y espurios y faciliten y agilicen mucho más el proceso de *scoring* y aumenten la fiabilidad del método fdMD para el diseño racional de fármacos. Estos descriptores pueden ser la variabilidad topológica, distribución de carga eléctrica, superficie polar, entre otros.

El análisis de los resultados que se fueron obteniendo tanto del *docking* como del *scoring*, la visualización de las figuras y la lista de los 13 residuos de aminoácidos de todos los BS (Figuras de la 3 a la 15) incentivó a una búsqueda bibliográfica sobre los sitios de unión previamente definidos en la literatura no solo por Radeva y col.⁽³¹⁾ sino también por Jumde y col.⁽¹¹⁾ y por Huschmann y col.⁽¹²⁾ y se encontró que los BS declarados como “nuevos” durante la investigación: BS 1 y 2 (Figura 3 y 4 respectivamente), BS 5 (Figura 10 y 13), BS 6 (Figuras 11 y 14) y BS 7 (Figura 15) no han sido previamente definidos, por lo que se sugiere investigar más a fondo estos BSs ya que sí podrían ser BSs nuevos; para estos sitios los valores de los descriptores son valores correctos, los clústeres están bien poblados y además presenta valores energéticos que indican estabilidad del complejo.

Con el método 1 ambos ligandos se dirigen a BS no descritos anteriormente: BS 1 y BS2 (Figura 3 y 4 respectivamente) para el ligando F031 y BS 5 (Figura 10 y 13) y BS 6 (Figuras 11 y 14) para el ligando F035. Si posteriormente se verificará que estos BSs son nuevos

podríamos afirmar que la metodología 1 es efectiva a la hora de predecir nuevos sitios de unión de ligandos a proteínas, aspecto positivo para el diseño posterior de nuevos fármacos. Incluso en el caso del ligando F035 con la metodología 2 uno de los dos representativos corresponde al BS 5, también encontrado por el ligando F035 en el método 1 y para el método 3 también ocurre similar, uno de los representativos corresponde al BS 6 también encontrado por el ligando en el método 1. Este comportamiento nos incentiva a seguir apostando por estas metodologías y perfeccionarlas hasta lograr una optimización del método fdMD para perfeccionar y agilizar finalmente el diseño racional de fármacos.

CONCLUSIONES



CONCLUSIONES

1. Con el análisis de las dinámicas moleculares H-REMD, se concluye que con los métodos 2 y 3, el ligando F031 no solo es capaz de llegar el sitio de unión correcto, sino que además lo hace en la pose correcta.
2. Después de clusterizar las dinámicas H-REMD se concluye que: el ligando F031 encuentra su pose experimental S1' con 2 metodologías, a diferencia del ligando F035 que no encuentra su pose experimental S3 y S6. De ahí que se recomienda ampliar el tiempo de simulación de la dinámica molecular y calcular nuevamente los descriptores en función de conocer con exactitud el final de estos ligandos, y así determinar si los sitios de unión visitados por los ligandos diferentes al experimental se pueden clasificar como *binding sites* nuevos o sitios espurios.
3. Además, de la energía MMGBSA se podría concluir que, tanto la población de miembros que tenga cada clúster así como la energía de enlace K_{DEEP} , son útiles para puntuar los resultados del *docking* y corroborar si el ligando ha encontrado su BS experimental o si es un falso positivo. Además de estos descriptores, en el caso en los que no se conozca la pose experimental, también se pueden utilizar otros descriptores como la variabilidad topológica, distribución de carga eléctrica, superficie polar, entre otros, que permitan discriminar entre sitios correctos y espurios y faciliten mucho más el proceso de *scoring*.

BIBLIOGRAFÍA



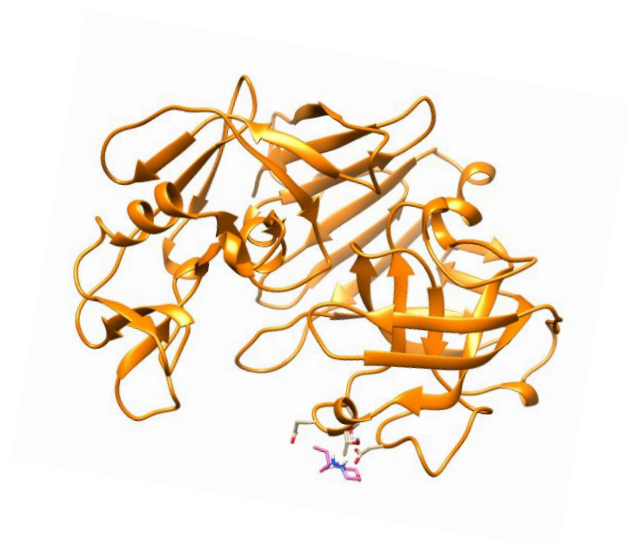
BIBLIOGRAFÍA

1. Bueren-Calabuig J.A. Dinámica molecular. Bioinformática con ñ [Internet]. 1.^a ed. Carolina del Sur, Estados Unidos: CreateSpace; 2014. p. 425-441. Disponible en: <http://doi.org/10.5281/zenodo.1066360>
2. Lozano-Aponte J, Scior T. ¿ Qué sabe Ud. Acerca de... Dinámica Molecular? Revista mexicana de ciencias farmacéuticas. 2014;45(1):86-8.
3. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chemical physics letters. 1999;314(1-2):141-51
4. Qi R, Wei G, Ma B, Nussinov R. Replica Exchange Molecular Dynamics: A Practical Application Protocol with Solutions to Common Problems and a Peptide Aggregation and Self-Assembly Example. Peptide Self-Assembly: Springer; 2018. p. 101-19.
5. D.A. Case, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, J. Swails, R.C.Walker, J.Wang, H.Wei, R.M.Wolf, X.Wu, L. Xiao, D.M. York and P.A. Kollman (2018), AMBER 2018, University of California, San Francisco.
6. Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, Jhoti H. Twenty years on: the impact of fragments on drug discovery. Nature reviews Drug discovery. 2016;15(9):605.
7. Allen KN, Bellamacina CR, Ding X, Jeffery CJ, Mattos C, Petsko GA, et al. An experimental approach to mapping the binding surfaces of crystalline proteins. The Journal of Physical Chemistry. 1996;100(7):2605-11.
8. Ferruz N, Harvey MJ, Mestres J, De Fabritiis G. Insights from fragment hit binding assays by molecular simulations. Journal of chemical information and modeling. 2015;55(10):2200-5.
9. Privat C, Granadino-Roldán JM, Bonet J, Santos M, Rubio-Martinez J. Fragment Dissolved Molecular Dynamics: A systematic and efficient method to locate binding sites. Journal of Chemical Theory and Computation. 2020. "in press"
10. Mondal M, Radeva N, Fanlo-Virgós H, Otto S, Klebe G, Hirsch AK. Fragment Linking and Optimization of Inhibitors of the Aspartic Protease Endothiapepsin: Fragment-Based

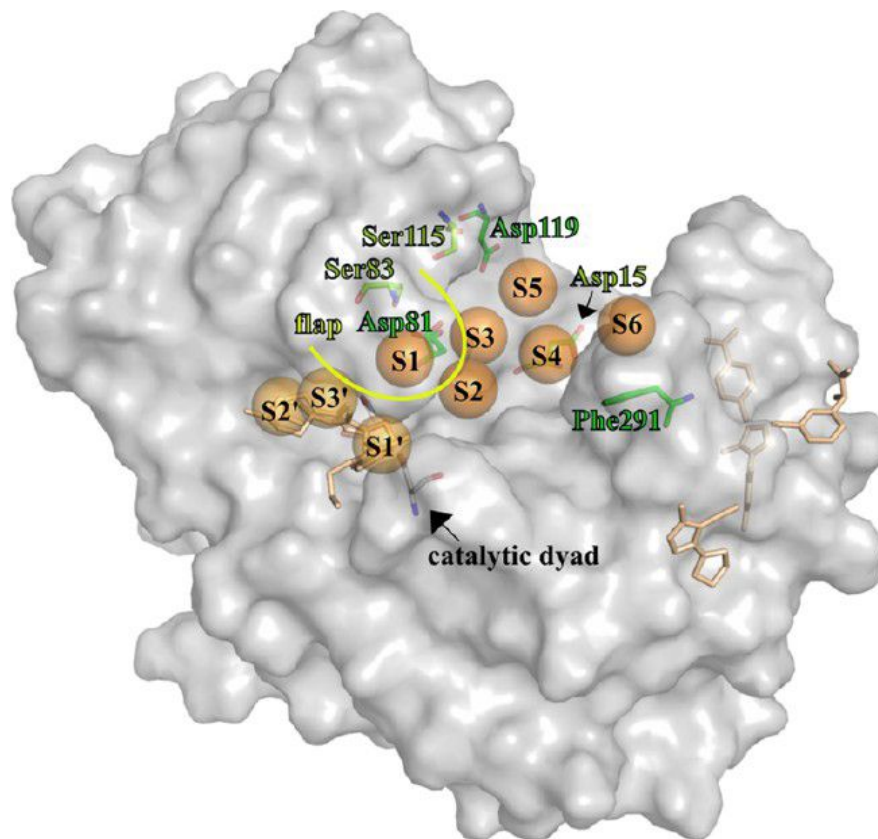
- Drug Design Facilitated by Dynamic Combinatorial Chemistry. *Angewandte Chemie International Edition*. 2016;55(32):9422-6.
11. Jumde VR, Mondal M, Gierse RM, Unver MY, Magari F, van Lier RC, et al. Design and synthesis of bioisosteres of acylhydrazones as stable inhibitors of the aspartic protease endothiapepsin. *ChemMedChem*. 2018;13(21):2266-70.
 12. Huschmann FU, Linnik J, Sparta K, Ühlein M, Wang X, Metz A, et al. Structures of endothiapepsin–fragment complexes from crystallographic fragment screening using a novel, diverse and affordable 96-compound fragment library. *Acta Crystallographica Section F: Structural Biology Communications*. 2016;72(5):346-55.
 13. Coates L, Tuan H-F, Tomanicek S, Kovalevsky A, Mustyakimov M, Erskine P, et al. The catalytic mechanism of an aspartic proteinase explored with neutron and X-ray diffraction. *Journal of the American Chemical Society*. 2008;130(23):7235-7.
 14. Cooper J, Foundling S. The structure of a synthetic pepsin inhibitor complexed with endothiapepsin. *European Journal of Biochemistry*. 1987;169(1):215-21.
 15. Hinchliffe A. *Molecular modelling for beginners*: John Wiley & Sons; 2005.
 16. Hidalgo-Ruíz J-L. Estudio teórico de imidas tetracíclicas y pirazolonas como inhibidores de K-Ras. 2016.
 17. De Vivo M, Masetti M, Bottegoni G, Cavalli A. Role of molecular dynamics and related methods in drug discovery. *Journal of medicinal chemistry*. 2016;59(9):4035-61.
 18. Toth PG, Marsalek P, Pokora O. Ergodicity and parameter estimates in auditory neural circuits. *Biological cybernetics*. 2018;112(1-2):41-55.
 19. Onufriev AV, Case DA. Generalized Born implicit solvent models for biomolecules. *Annual review of biophysics*. 2019;48:275-96.
 20. Li J, Ying J, Xie D. On the analysis and application of an ion size-modified Poisson–Boltzmann equation. *Nonlinear Analysis: Real World Applications*. 2019;47:188-203.
 21. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*. 1983;79(2):926-35.
 22. Schlick T. *Topics in Nucleic Acids Structure. Molecular Modeling and Simulation*: Springer; 2002. p. 147-98.
 23. Gavathiotis E, Suzuki M, Davis ML, Pitter K, Bird GH, Katz SG, et al. BAX activation is initiated at a novel interaction site. *Nature*. 2008;455(7216):1076-81.

24. Granadino JM. Accelerated fully flexible ligand-receptor docking including explicit solvent using Hamiltonian replica exchange. "comunicación personal"
25. Maffucci I, Hu X, Fumagalli V, Contini A. An efficient implementation of the Nwat-MMGBSA method to rescore docking results in medium-throughput virtual screenings. *Frontiers in chemistry*. 2018;6:43.
26. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*. 2015;10(5):449-61.
27. Zhang X, Perez-Sanchez H, C Lightstone F. A comprehensive docking and MM/GBSA rescoring study of ligand recognition upon binding antithrombin. *Current topics in medicinal chemistry*. 2017;17(14):1631-9.
28. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*. 2000;33(12):889-97.
29. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate– DNA helices. *Journal of the American Chemical Society*. 1998;120(37):9401-9.
30. Gohlke H, Case DA. Converging free energy estimates: MM-PB (GB) SA studies on the protein–protein complex Ras–Raf. *Journal of computational chemistry*. 2004;25(2):238-50
31. Radeva N, Krimmer SG, Stieler M, Fu K, Wang X, Ehrmann FR, et al. Experimental active-site mapping by fragments: hot spots remote from the catalytic center of endothiapepsin. *Journal of medicinal chemistry*. 2016;59(16):7561-75.
32. Darden T, York D, Pedersen L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of chemical physics*. 1993;98(12):10089-92.
33. Olson AJ, Trott O. AutoDock Vina. California:Scripps Research;2014
34. UCSF Chimera 1.14. California:Resource for Biocomputing, Visualization, and Informatics (RBVI);2019
35. Arcon JP. Determinantes moleculares de la interacción droga-proteína: uso de cosolventes como moléculas de prueba. 2018.
36. Jiménez J, Skalic M, Martinez-Rosell G, De Fabritiis G. K deep: Protein–ligand absolute binding affinity prediction via 3d-convolutional neural networks. *Journal of chemical information and modeling*. 2018;58(2):287-96

ANEXOS



ANEXOS



Anexo 1. Representación de superficie EP excluida del disolvente con todos los bolsillos de unión definidos indicados por esferas (PDB 4Y5L).⁽³¹⁾

CURRICULUM VITAE



CURRICULUM VITAE

PERFIL PROFESIONAL

Poseo habilidades integrales en salud, servicios farmacéuticos, educación sanitaria y análisis clínico que me permiten realizar funciones en los servicios farmacéuticos institucionales y comunitarios, en la industria farmacéutica, biotecnológica y cosmetológica, la educación, la investigación y la regulación farmacéutica.

FORMACIÓN ACÁDEMICA

Máster en Biotecnología y Biomédica

Universidad de Jaén, España. Curso (2019-2020)

Licenciada en Ciencias Farmacéuticas.

Universidad Central "Marta Abreu" de Las Villas. Septiembre del 2013-junio del 2018.

FORMACIÓN COMPLEMENTARIA

Curso de idioma inglés. Centro Diocesano de Formación Dra. Felicia Pérez.

Obispado de Santa Clara, Cuba. Septiembre del 2018 -Julio del 2019.

Curso de buenas prácticas iniciales.

Gestión del capital humano. Planta de producción NOVATEC. Empresa MEDSOL, Organización Superior de Dirección Empresarial (OSDE) BioCubaFarma, La Habana, Cuba. 2017

EXPERIENCIA LABORAL

Noviembre del 2019-Actualidad (Jaén, España)	Farmacéutica en la Parafarmacia de la Perfumería PRIMOR, Calle Roldán y Marín, 7, 23001, Jaén
Septiembre del 2018-Setiembre del 2019 (Camajuaní, Villa Clara, Cuba)	Especialista en Servicios Farmacéuticos y Plantas Medicinales en Unidad Empresarial Presupuestada Sectorial de Salud Pública Camajuaní, Villa Clara, Cuba
Enero-febrero del 2018 (Villa Clara, Cuba)	Práctica Laboral de Farmacia Clínica y Hospitalaria en el Hospital Universitario Dr. Celestino Hernández Robau.
Junio-julio del 2017 (La Habana, Cuba)	Práctica Laboral de Producción de Medicamentos en la planta de producción NOVATEC vinculada a la Empresa MEDSOL, perteneciente a la Organización Superior de Dirección Empresarial (OSDE) BioCubaFarma.
Junio-julio del 2016 (Camajuaní, Villa Clara, Cuba)	Práctica Laboral de Farmacia Dispensarial y Dependiente en la Farmacia Comunitaria 634.
Junio-julio del 2015 (Camajuaní, Villa Clara, Cuba)	Práctica Laboral en los Servicios farmacéuticos de la Farmacia Comunitaria del municipio Camajuaní, Villa Clara, Cuba

HABILIDADES

- Trabajo en laboratorios de análisis químico, tecnología farmacéutica, biomedicina y otros.

- Destrezas de comunicación, empresariales y de gestión.
- Capacidad de trabajar en equipo.
- Conocimientos de contabilidad para gestionar el stock de medicamentos y realizar pedidos.
- Conocimientos adecuados de los medicamentos y de materias primas; de la tecnología farmacéutica y del control físico, químico, biológico y microbiológico de los medicamentos.
- Preparación adecuada en temas del metabolismo y del efecto de los medicamentos y de la acción de los tóxicos.
- Poseo estudios que me permiten saber evaluar los datos científicos relativos a los medicamentos para poder proporcionar, sobre esta base, información apropiada.
- Conocimiento adecuado de las condiciones legales y otras en materia de ejercicio de las actividades farmacéuticas.

PUBLICACIONES

Título Devenir farmacéutico en San Antonio de las Vueltas, 1867-2018

Autores principales Msc. Urbano Monteagudo Romero

Autores Lic. Miriam Iliex Guirado López, Kenia Rodríguez González,
Xiomara Cadenas Pérez

Publicado Revista Cubana de Farmacia Vol.52, No. 4 (2019) Editorial
Ciencias Médicas ISSN 1561-2988

IDIOMAS

Idioma	Habla	Lee	Escribe
Español	correctamente	correctamente	correctamente
Inglés	bien	bien	bien

INFORMÁTICA

Buen manejo de ordenadores Windows, tratamientos de textos en internet, correo electrónico y Microsoft Office (Excel, Word, Power Point, Endnote), programas estadísticos para el procesamiento de datos, bases de datos, simulaciones de dinámica molecular.

OTROS DATOS DE INTERÉS

- ☐ Miembro de la Sociedad Cubana de Farmacología.
- ☐ Vinculación desde el 2015 hasta el 2018 con el grupo de investigación del departamento de Farmacia de la Universidad Central “Marta Abreu” de Las Villas, Villa Clara, Cuba que tiene como objetivo fundamental el estudio de la composición química y la actividad farmacológica de extractos obtenidos de plantas aromáticas cultivadas en Cuba.