



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Trabajo Fin de Grado

**ESTUDIO ASISTIDO MEDIANTE
SOFTWARE DE LA INTERACCIÓN
ENTRE UN FÁRMACO
ANTICANCERÍGENO Y SU DIANA
BIOLÓGICA (G6PDH).**

Alumno: Simón Pérez, María Inmaculada

Julio, 2021



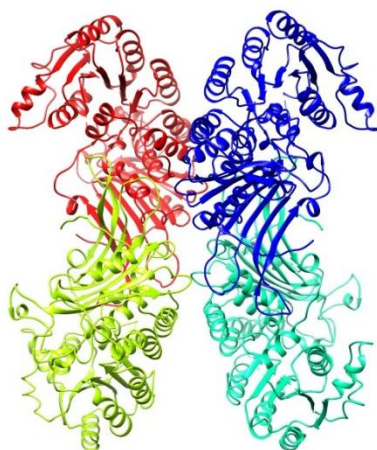
UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales



Facultad de
Ciencias Experimentales

Trabajo Fin de Grado

ESTUDIO ASISTIDO MEDIANTE SOFTWARE DE LA INTERACCIÓN ENTRE UN FÁRMACO ANTICANCERGÍGENO Y SU DIANA BIOLÓGICA. (G6PDH)



Alumno:

Simón Pérez, María Inmaculada.

Jaén, Julio, 2021.

ÍNDICE:

- 1. Resumen/Abstract.**
- 2. Introducción.**
 - 2.1. Glucosa-6-fosfato deshidrogenasa (G6PDH).**
 - 2.1.1. Importancia de la G6PDH.**
 - 2.1.2. Estructura cristalina de G6PDH humana.**
 - 2.1.3. Deficiencia de G6PDH.**
 - 2.2. Dinámica molecular.**
 - 2.2.1. Campos de fuerza.**
- 3. Objetivos.**
- 4. Materiales y métodos.**
 - 4.1. Amber.**
 - 4.2. Dinámica molecular de fragmentos disueltos (fdMD).**
 - 4.3. Preparación y dinámica molecular del sistema G6PDH.**
 - 4.3.1. Preparación del sistema de la proteína.**
 - 4.3.2. Minimización de la energía.**
 - 4.3.3. Calentamiento.**
 - 4.3.4. Equilibrado.**
 - 4.3.5. Producción y simulación MD.**
 - 4.4. Análisis de fdmd.**
 - 4.4.1. CCTRAJ y análisis de las trayectorias reactivas.**
 - 4.4.2. MMPBSA.py.**
 - 4.4.3. Cálculo de la energía libre.**
 - 4.4.3.1. Contribuciones de las interacciones en la energía.**
- 5. Resultados.**
 - 5.1. Distancia de enlace frente al tiempo.**
 - 5.2. ΔG .**
 - 5.2.1. ΔG total.**
 - 5.2.2. ΔG promedio.**
 - 5.3. Análisis de las instantáneas de cada ligando.**
- 6. Discusión.**
- 7. Conclusión.**
- 8. Bibliografía.**

1. Resumen

La glucosa-6-fosfato deshidrogenasa (G6PDH), es una enzima primordial involucrada en la primera reacción de la rama oxidativa de la ruta de las pentosas fosfato (PPP). Esta ruta juega un papel esencial en la proliferación celular, estando sobreactivada en muchos tipos de cáncer. Es la ruta necesaria para dar energía a las células, y las células cancerosas necesitan más energía que las sanas. Es importante encontrar un compuesto orgánico que pueda unirse a la ella e interferir su función. Un grupo experimental con el que colaboramos encontró un compuesto que se une e inhibe la G6PDH, pero no se conoce su sitio ni mecanismo de acción. Conocer esto ayudaría a plantear sustituciones que mejoren su mecanismo de acción. En este proyecto se plantea como objetivo predecir el sitio y conformación adoptada por dicho compuesto al unirse a la G6PDH. Dicho estudio se realiza con simulaciones de dinámica molecular, generando la película que muestre cómo el compuesto interactúa con G6PDH. Esto aportará ayuda al diseño de nuevos inhibidores de G6PDH para fines terapéuticos.

Abstract

Glucose-6-phosphate dehydrogenase (G6PDH) is a key enzyme involved in the first reaction of the oxidative branch of the pentose phosphate pathway (PPP). This pathway plays an essential role in cell proliferation, being over activated in many types of cancer. It is the pathway needed to give energy to cells, and cancer cells need more energy than healthy cells. It is important to find an organic compound that can bind to it and interfere with its function. An experimental group with which we collaborate found a compound that binds to and inhibits G6PDH, but its site and mechanism of action are not known. Knowing this would help to propose substitutions that improve its mechanism of action. The objective of this project is to predict the site and conformation adopted by this compound upon binding to G6PDH. This study is carried out with molecular dynamics simulations, generating a layer that shows how the compound interacts with G6PDH. This will help in the design of new G6PDH inhibitors for therapeutic purposes

.

2. INTRODUCCIÓN.

2.1. Glucosa-6-fosfato deshidrogenasa (G6PDH).

2.1.1. Importancia de la G6PDH

La glucosa-6-fosfato deshidrogenasa (G6PDH), es la molécula diana de estudio del presente Trabajo Fin de Grado. Es una enzima fundamental que se encuentra presente en las células de todos los seres vivos. Pertenecce al grupo de las oxidorreductasas deshidrogenasas, ya que cataliza reacciones de oxidación-reducción donde tiene lugar el proceso de transferencia de electrones. En las células humanas es la encargada de gobernar el conjunto de reacciones metabólicas por mediación de la ruta de las pentosas fosfato (PPP). (Obiol-Pardo *et al.*, 2014)

La G6PDH, al igual que muchas otras enzimas, depende de otras moléculas orgánicas pequeñas, denominadas coenzimas, para desempeñar su actividad catalítica. En el caso que tratamos, precisa del dinucleótido de nicotinamida y adenosín fosfato, NADPH. La enzima sobreactiva la ruta de las PPP. Esta ruta tiene lugar en el citoplasma y en ellas se pueden diferenciar dos fases, una oxidativa y otra no oxidativa. En la fase oxidativa es donde participa la G6PDH, catalizando el carbono-1 de la glucosa-6-fostato, formándose como producto 6-fosfoglucono-1-lactona, tal y como se muestra en la fig. 2.1. En este paso se obtiene NADPH por la reducción de la coenzima NADP⁺. El NADPH se emplea como donador de electrones en las rutas anabólicas y en la eliminación de especies reactivas de oxígeno para así proteger a las células del estrés oxidativo. (Ramos-Martínez, 2017).

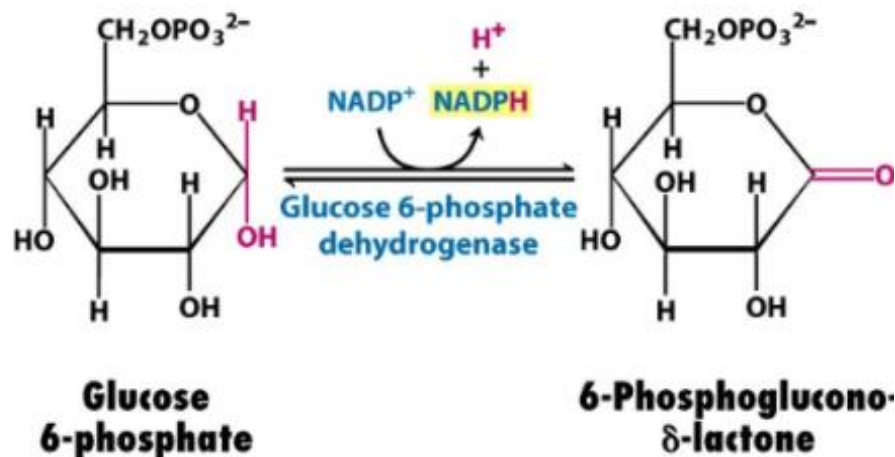


Figura 2.1. Primer paso de la ruta de las PPP donde participa la G6PDH.

La importancia de la G6PDH recae en que la ruta de las pentosas fosfato desempeña un papel primordial en la multiplicación celular, encontrándose sobreactivada en muchos tipos de cáncer. En varios estudios se ha demostrado que la enzima G6PDH se puede considerar como una nueva diana para el tratamiento contra el cáncer, inhibiendo la actividad del cáncer mediante dos procedimientos diferentes: silenciando la expresión génica de la G6PDH, o mediante compuestos inhibidores de su actividad enzimática. (Obiol-Pardo *et al.*, 2014; Ramos-Montoya *et al.*, 2006)

Se conoce la existencia de algunos inhibidores de la actividad de la enzima, entre ellos se encuentra la 6-aminocotinamida (6AN), que es un inhibidor competitivo propio de la G6PDH, pero posee unos efectos secundarios negativos, como daños en el sistema nervioso y deficiencia de vitamina B. Otros inhibidores son la combinación de fármacos como cefaperazona/sulbactam o ampicilina/sulbactam, con un mecanismo de inhibición competitivo y no competitivo, respectivamente, pero no se conoce mucha información acerca del sitio de unión a la enzima o el mecanismo de acción de dichos compuestos. También se ha descubierto que el metamizol interfiere en la actividad de la G6PDH y, por otro lado, que el metotrexato es un inhibidor bien conocido de la G6PDH, pero no es específico exclusivamente de ella e inhibe a otras enzimas que dependen también a su vez de la coenzima NADP^+ . (Obiol-Pardo *et al.*, 2014; Babiak *et al.*, 1998; Cifti *et al.*, 2001)

2.1.2. Estructura cristalina de G6PDH humana.

El monómero de la proteína, G6PDH, está formado por 514 aminoácidos, con un peso molecular de 59kDa. En la posición uno de la secuencia de aminoácidos se encuentra la metionina, que es la iniciadora de la secuencia madura.

La forma activada de la glucosa-6-fosfato deshidrogenasa humana existe en un equilibrio entre dímero y tetrámero, conocido como dimertetrámero. Dicha estructura se ve influenciada por el NADPH, ya que es primordial para mantener un pH óptimo necesario para la estimulación del crecimiento celular, además de ser necesario el control de la fuerza iónica debido a las fuerzas electrostáticas entre las superficies de contacto. Un pH elevado, entorno a $\text{pH} > 8$ y una fuerza iónica elevada conllevan a que el equilibrio tienda a favorecer al dímero. A estos elevados pHs, también se consiguen monómeros, pero son inactivos. Cuando se produce una disminución del pH, $\text{pH} < 6$, y fuerza iónica baja, el equilibrio se desplaza hacia el equilibrio del tetrámero. (Au *et al.*, 1999).

La primera estructura cristalina de la G6PDH se consiguió a partir de la enzima *Leuconostoc mesenteroides* (*L. mesenteroides*). Posteriormente, se obtuvo la primera estructura cristalina de la G6PDH humana determinada para la variante Canton Arg459Leu. Esta última es un tetrámero compuesto de dos dímeros.

En la fig. 2.2; se ve representado el tetrámero de la proteína G6PDH. Las cuatro subunidades de este tetrámero se nombran A, B, C, D y se muestran en azul oscuro, rojo, amarillo y azul celeste, respectivamente. Cada monómero está constituido por dos dominios, los cuales se nombran como dominio A y B, tal y como se muestra representado en la fig. 2.3. Uno de ellos es más pequeño y es el N-terminal, sitio de unión a coenzimas, conocido como huella dactilar. El segundo dominio, denominado $\beta + \alpha$, es más grande que el anterior y es el que incluye el sitio de unión al sustrato. Estos dos dominios están conectados entre sí por una hélice α , la cual contiene la secuencia de ocho residuos totalmente conservada RIDHYLGK. Este es el sitio por donde se produce la unión al sustrato y contiene además Lys205, la cual está implicada en la unión de la enzima al sustrato, que es esencial para la catálisis. (Au *et al.*,

1999). En ambas figuras (fig. 2.2; fig. 2.3) se puede apreciar que las moléculas de NADPH resaltadas de color verde.

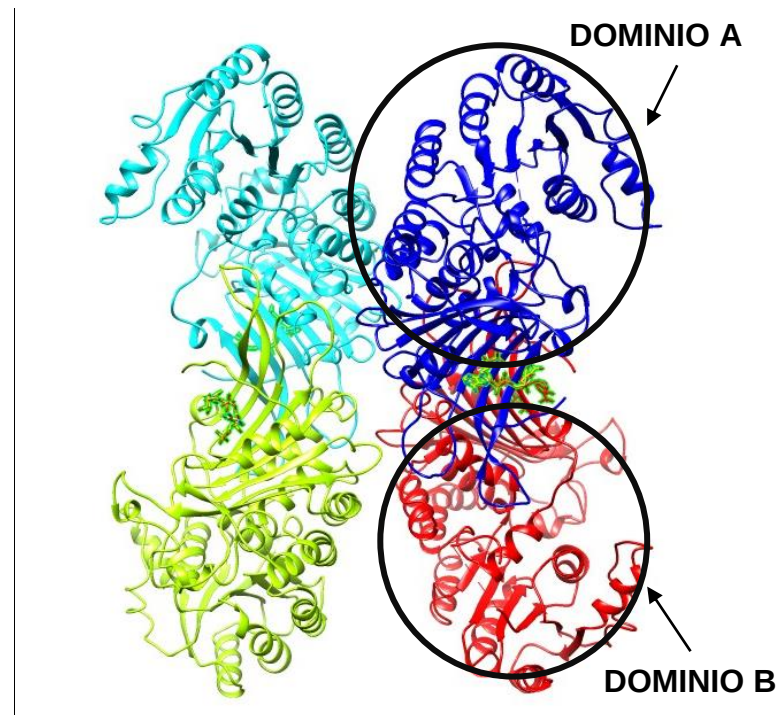


Figura 2.2. Tetrámero de la G6PDH.

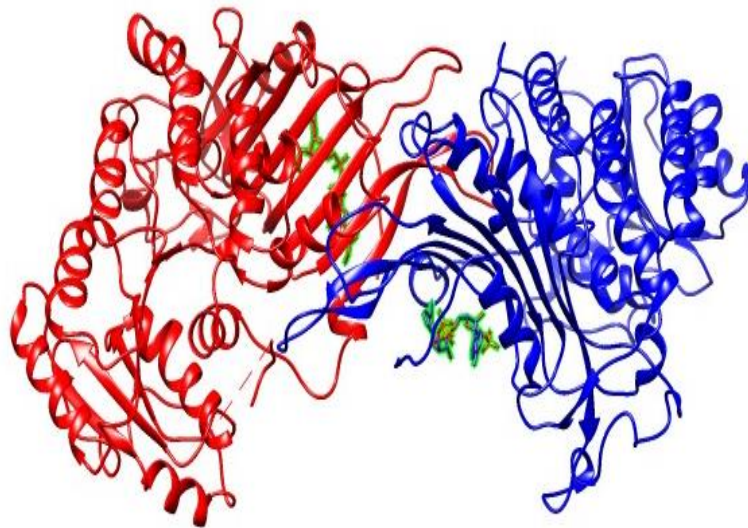


Figura 2.3. Dominio A y B del monómero.

2.1.3. Deficiencia de G6PDH.

A partir de diversos estudios se ha demostrado que el potencial de reducción-oxidación (redox) intracelular está determinado por las concentraciones de

oxidantes y reductores. El principal reductor intracelular a destacar es la coenzima NADPH, producida principalmente en la primera etapa de la ruta de las PPP, por la acción de la G6PDH. La coenzima NADPH es primordial para regular el pH óptimo necesario para el crecimiento celular. Los mecanismos de defensa antioxidantes dependen de la producción adecuada de esta coenzima para reducir los equivalentes durante el estrés oxidativo. (Tian *et al.*, 1998, 1999)

Hace tiempo que resulta importante comprender la relación entre las características bioquímicas de la deficiencia G6PDH y su base molecular. La deficiencia de G6PDH es la enzimopatía humana más común. Se estima que este defecto enzimático afecta a 400 millones de personas en todo el mundo. Hasta ahora se conocen unos 127 sitios de mutación. Esta deficiencia viene producida por diversos mecanismos genéticos asociados al gen del cromosoma X, y provocan mutaciones puntuales y sustituciones estructurales en la enzima que afectan a la transcripción, o estructura primaria de esta, conduciendo a una disminución de la actividad enzimática o pérdida de afinidad por el sustrato. (Bonilla *et al.*, 2007).

Gracias a la estructura tridimensional de la G6PDH de *L. mesenteroides* se pudo construir un modelo de esta enzima por sustitución molecular, obteniéndose una estructura dimérica, para poder entender la deficiencia de la enzima. Esta estructura de 3 Å de resolución, es la variante Canton Arg459Leu. En ella se encontró una molécula estructural de NADP⁺ en cada subunidad del tetrámero cerca de la interfaz del dímero. Para conocer la deficiencia grave hay que estudiar la estructura cerca de la interfaz del dímero y del sitio estructural del NADP⁺, que son las regiones que indican la importancia de la estabilidad de la enzima y de su actividad. (Au *et al.*, 1999).

El estudio de las variantes de la G6PDH muestra que la deficiencia, en la mayoría de los casos, se debe a la inestabilidad de la enzima. Esto ocurre debido a que las sustituciones de aminoácidos en diferentes lugares desestabilizan la estructura de la enzima.

2.2. Dinámica molecular

Los métodos de modelización molecular surgieron para estudiar, completar y predecir los resultados que se obtienen mediante técnicas experimentales y así conocer mejor la estructura, la función y el mecanismo de reacción de los sistemas que se estudian. La modelización molecular tiene como fin el estudio de la estructura y funciones moleculares con modelos y métodos computacionales.

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. (Bueren-Calabuig, 2014).

Los métodos computacionales con los que se estudian las estructuras y funciones de las moléculas incluyen muchas técnicas como son los cálculos de mecánica cuántica, dinámica molecular, métodos de Monte Carlo. Fundamentalmente podemos subrayar dos técnicas: la técnica de Monte Carlo, usada para crear una serie de estados posibles del sistema. Esta técnica se centra en una partícula del sistema, después define movimientos al azar de la partícula seleccionada, calcula su función de energía, y considera el movimiento válido que puede tener. La otra técnica clásica a destacar es la dinámica molecular (molecular dynamics, MD), destinada a analizar el comportamiento de un sistema en función del tiempo. El comportamiento de este sistema se calcula mediante ecuaciones del movimiento como las leyes de Newton ($F=ma$), donde se estudian las fuerzas que existen entre los átomos que componen el sistema. (Hinchliffe *et al.*, 2000; Jensen y Wiley, 2007; Lozano-Aponte y Scior, 2014; Orozco, 2014).

La MD se aplica en muchos campos, pero en este estudio se destaca su aplicación en el estudio del comportamiento de un ligando con su diana biológica, siendo esto un pilar fundamental en el diseño de nuevos fármacos.

Un proyecto de dinámica molecular se puede resumir en los siguientes pasos:

- Preparación del sistema: Se selecciona la proteína y a continuación se establece la configuración inicial (CI) del sistema que se va a usar

compuesta por un número determinado de átomos (N) y se definen las condiciones de equilibrio (CE).

- Equilibrado: Se deben de ir modificando los parámetros, como por ejemplo, la presión y el volumen, entre otros, manteniendo algunos constantes, hasta llevar al sistema al equilibrio, donde la energía permanece casi constante.
- Simulación MD: Cuando ya el sistema se ha equilibrado se especifica el tiempo de ejecución y se deja correr la simulación de la dinámica molecular. Este es el paso que requiere mayor tiempo. Se realizan unas mediciones y cálculos de las fuerzas del sistema con lo que después de se obtendrán aquellas propiedades fisicoquímicas del sistema que interesen.
- Análisis de los resultados: Finalmente, cuando ya se han obtenido los resultados se procede a su análisis obteniéndose así la información deseada. (Lozano-Aponte y Scior, 2014).

2.2.1. Campos de fuerza.

En MD se emplean un conjunto de funciones para expresar la energía de un sistema, conocidas como campo de fuerza. Estas funciones dependen de una serie de parámetros como pueden ser la masa atómica, términos energéticos, cargas parciales, etc.

Las moléculas se tratan como un grupo de átomos unidos por enlaces y que obedecen a una serie de funciones de energía potencial mecano-clásicas y se pueden aplicar a sistemas grandes. Estos métodos ignoran el movimiento de los electrones, calculando la energía del sistema como función de las posiciones nucleares.

Campo de fuerza es tanto la expresión matemática para describir la energía del sistema, como los parámetros que se utilizan: como distancia de enlace, fuerza de enlace etc. Definen la energía del sistema por medio de términos intramoleculares e intermoleculares tal y como se detalla en la ecuación 2.1. La función de energía queda definida por la expresión matemática de la ecuación 2.2, donde se tiene en cuenta la distancia de enlace, el ángulo de torsión, ángulo de enlace, interacciones electrostáticas y las de van der Waals.

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

Ecuación 2.1.

$$E = \text{Enlace (tensión)} + \text{Ángulo (flexión)} + \text{Diedro} + \text{Diedro impropio} + \text{Electrostático} \\ + \text{Van der Waals.}$$

Ecuación 2.2

Cada una de las componentes detalladas en el campo de fuerza se calcula de forma diferente. En general los campos de fuerza calculan la energía derivada de la tensión de enlace para átomos a dos enlaces de distancia, la energía de flexión para átomos a tres enlaces, la de torsión para átomos a cuatro enlaces, y la energía de van der Waals (E_{vdW}) y la energía electrostática (E_{el}) se calculan para átomos a cuatro enlaces o más de distancia. (Jensen y Wiley, 2007)

Cuando se va a realizar un trabajo con MD hay que elegir el campo de fuerza correcto que se ajuste al procedimiento que se va a seguir. En este trabajo se empleará el campo de fuerzas Amber (Assisted Model Building with Energy Refinement), concretamente su versión Amber ff14SB para el caso que estamos estudiando de la proteína, con la suite de programas Amber 18. Emplearemos este campo de fuerza porque está destinado para la simulación de proteínas y ácidos nucleicos. (Maier *et al.*, 2015; Case *et al.*, 2018).

La MD se utiliza para simular la estabilidad de los dímeros de proteínas y de los complejos fármaco-proteína a lo largo del tiempo. Gracias al empleo de esta técnica de simulación podemos calcular diferentes propiedades de las moléculas como pueden ser, entre otras, la energía libre del sistema, y además permite predecir la fuerza de interacción entre los posibles fármacos y sus dianas biológicas.

3. OBJETIVOS.

Un grupo experimental con el que colaboramos ha descubierto un compuesto que inhibe la G6PDH in vitro, pero no se conoce dónde se produce esta unión, ni con qué parte de la proteína está interactuando, esto ayudaría mucho a conocer, entender y mejorar su mecanismo de acción.

En este trabajo de investigación se tiene como objetivo principal el estudio del proceso de unión de la proteína G6PDH con un ligando, para proponer el mejor sitio de unión.

Para llevar a cabo este estudio, se plantea emplear la simulación de dinámica molecular, mediante los programas de la suite Amber, utilizando la estructura cristalina de G6PDH.

El análisis de la MD nos ayudará a identificar los posibles sitios donde se producen interacciones más fuertes del ligando con la G6PDH y elegir el mejor sitio de unión gracias a los resultados que se obtengan.

Encontrar un compuesto que pueda unirse e interferir en la función de la proteína glucosa-6-fosfato deshidrogenasa, será muy útil como tratamiento anticancerígeno.

Con simulaciones de ordenador vamos a poder generar la película que muestre cómo el compuesto que sabemos que se une a G6PDH establece dicha unión y cómo interactúan con la proteína. Esto permitirá plantear modificaciones a dicho compuesto que aumenten su afinidad para con la proteína.

Vamos a plantear una versión reducida del compuesto mencionado, al que llamaremos modelo, para que sea abordable y poder llegar a saber por dónde se une utilizando la filosofía de fdMD (fragment dissolved Molecular Dynamics). (Privat *et al.*, 2020).

4. MATERIALES Y MÉTODOS.

4.1. Amber

Amber es una suite de programas de simulación molecular para biomoléculas. Está formada por un gran número de programas diseñados para poder simular sistemas químicos. En estas simulaciones es también el campo de fuerza que se utiliza. Amber se divide en dos partes: Amber y por otro lado AmberTools. En nuestro caso, se usa el campo de fuerzas ff14SB para la proteína, que es el sistema que estamos estudiando. Amber incluye programas destinados a la preparación de estructuras, como es el caso de LEaP, otros programas para el cálculo de energías, como Sander y programas de análisis como ptraj. (Case *et al.*, 2018).

El flujo de trabajo que se va a seguir en este trabajo es el que se muestra en la siguiente fig. 4.1.

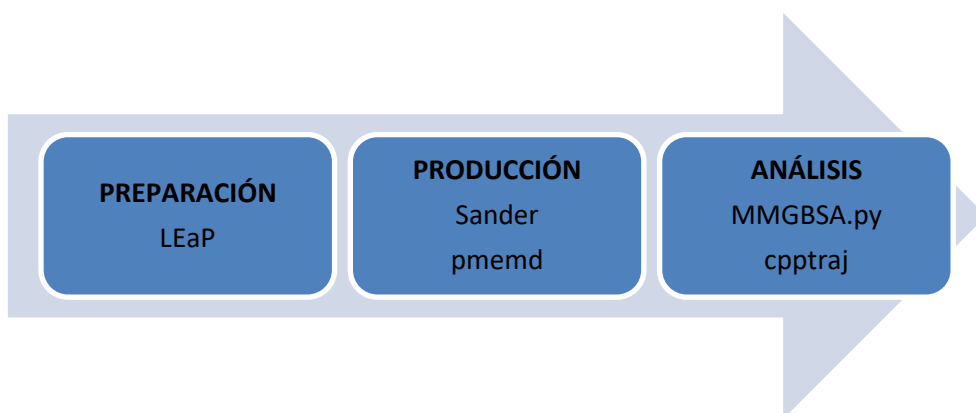


Figura 4.1. Diagrama del flujo de trabajo.

Con el programa LEaP se generan los ficheros de topología y coordenadas de cada átomo. La topología es el fichero que contiene los parámetros para cada átomo del campo de fuerza utilizado. Posteriormente se trabajará con Sander (*Simulated Annealing with NMR-Derived Energy Restraints*). Este es el programa de simulaciones que utilizaremos para el proceso de minimización, así como para el de calentamiento. Es capaz de trabajar con diferentes campos de fuerzas diseñados para biomoléculas u otro tipo de sistemas (por ejemplo, polímeros orgánicos) y con diferentes modelos para agua y otros disolventes. Otro programa que se usará es Pmemd. Se trata de una versión de Sander optimizado para ejecutarse en tarjetas gráficas (GPUs). Por último, para el análisis de los resultados se use el programa cpptraj, que es una herramienta para analizar la trayectoria obtenida. También se hará uso del script MMGBSA.py, que automatizará el cálculo de energía libre de unión según la metodología MMGBSA. (Case *et al.*, 2018; Roe *et al.*, 2013; Ansino, 2020; Kollman *et al.*, 2000).

4.2. Dinámica molecular de fragmentos disueltos (fdMD)

La mayor parte de los medicamentos que ingerimos a día de hoy son el resultado de un complejo procedimiento de investigación. El estudio y el desarrollo de los medicamentos se ha llevado a cabo siempre mediante métodos experimentales, pero en los últimos años se han desarrollado también

fármacos gracias a métodos computacionales. Estos métodos poseen los siguientes rasgos principales:

- Identificar compuestos ya existentes con una actividad biológica de interés y además crear nuevas estructuras con los efectos biológicos en los que estemos interesados.
- Dentro de las posibles opciones de moléculas que tengamos seleccionar aquella candidata que sea la más adecuada.
- Optimizar las propiedades deseadas de los fragmentos de las moléculas que disponemos.

El descubrimiento de fármacos basados en fragmentos, conocido como FBDD (Fragment Based Drug Design), es una técnica experimental usada para la búsqueda de nuevos fármacos. Esta técnica tiene como principal objetivo el estudio de los fármacos desde el punto de vista químico y consiste en fijar la relación existente entre la estructura química del fármaco y su actividad biológica.

El grupo en que se ha desarrollado este Trabajo Fin de Grado, propuso el llamado método de dinámica molecular de fragmentos disueltos, conocido con las siglas fdMD (Fragment Dissolved Molecular Dynamics), y con él poder solventar los problemas que FBDD pueda dar. Este es un método semiautomático y sistemático propuesto para ser aplicado en proyectos FBDD, con el objetivo principal de localizar el sitio de unión de una diana biológica con un ligando pequeño. (Privat *et al.*, 2020).

FdMD es el método con el que vamos a trabajar para el desarrollo de este proyecto. En él se emplea cajas de simulación previamente equilibradas en cuyo interior hay una copia del ligando solvatado con moléculas de agua. Se requiere emplear un término de repulsión de Lennard-Jones para evitar la agregación de los ligandos. Esta caja pre-equilibrada se puede usar para solvatar la diana de interés. Cuando se produce la solvatación de la proteína con esta caja previamente equilibrada se obtiene un sistema de simulación que contiene un número de copias del ligando, y así se evita la posible desnaturalización de la proteína. Se pueden destacar dos ventajas que presenta fdMD, por un lado, puede solvatar la diana biológica con un pequeño número de ligandos y así evitar la desnaturalización y por otro lado, aportar una base

de datos de cajas solvatadas de diferentes ligandos para a posteriori poder ser usadas en proyectos de FBDD.

Las seis etapas del método fdMD abarcan desde la preparación del fragmento y la proteína hasta el análisis de los resultados que se obtengan. (Privat *et al.*, 2020).

El método fdMD utiliza la suite de programas Amber18. En primer lugar, utilizando el módulo LEaP de Amber18, se solvata el ligando con moléculas de agua TIP3P. (William *et al.*, 1983).

Seguidamente, 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. De la misma forma se trabaja con la proteína, que se solvata con LEaP usando la caja de solvatación equilibrada con el ligando que se va a estudiar. Como resultado final, tras un protocolo de minimización y equilibrado, se obtiene un sistema final que constará de la proteína con un número (N) diferente de copias del ligando para el cual se quiere analizar su modo de unión a la proteína.

Las energías libres de enlace se calculan con el script MMGBSA.py, siguiendo el esquema MMGBSA (Molecular Mechanics Generalized Born Surface Area). (Miller *et al.*, 2012; Kollman *et al.*, 2000). En la figura de más abajo, fig. 4.2; se puede observar cómo se desarrolla el ciclo termodinámico. El método MMGBSA está planteado para valorar la energía libre de la unión a macromoléculas biológicas de ligandos pequeños. Estos cálculos están basados en simulaciones de dinámica molecular del complejo receptor-ligando, obteniéndose así la energía total del enlace. (Granadino-Roldán *et al.*, 2014).

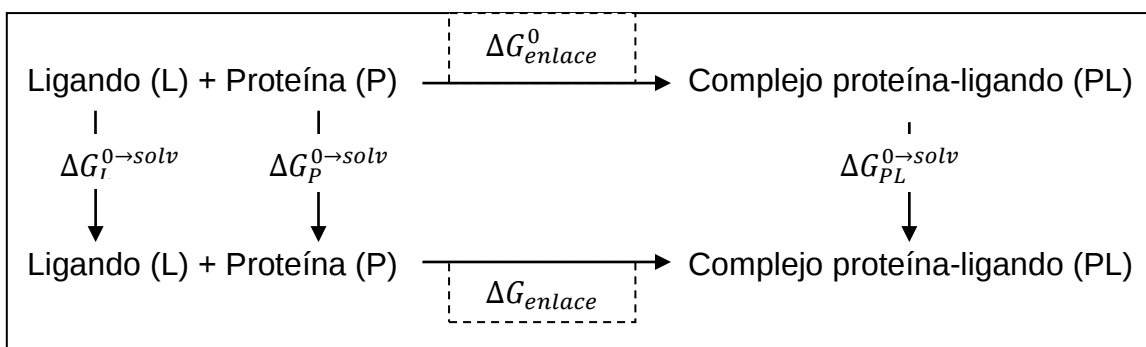


Figura 4.2. Ciclo termodinámico destinado para obtener una expresión de la energía total de enlace.

4.3. Preparación y dinámica molecular del sistema G6PDH.

4.3.1. Preparación del sistema de la proteína.

Inicialmente, se obtuvo la estructura de partida de la G6PDH del Protein Data Bank (PDB) (<http://www.rcsb.org>) con el código de entrada 1QKI. Se trata de una estructura cristalina con la mutación Canton y con NADP⁺ estructural. (Pettersen *et al.*, 2004).

Durante los pasos de preparación y de producción, se empleó el algoritmo de restricción SHAKE, que permite que el tiempo de integración sea de 4 fs, o sea, con el mismo número de pasos de cálculo se obtiene el doble de tiempo de dinámica molecular.

El procedimiento de fdMD se resume en seis pasos que van desde la preparación del sistema hasta el análisis de los resultados obtenidos derivados de los diversos cómputos. Los pasos 1 y 2 se corresponden a la preparación del fragmento de interés y de la preparación de la proteína, respectivamente. En el paso 1, el fragmento de interés se sumerge en una caja de moléculas de agua TIP3P equilibradas utilizando el módulo LEaP de Amber 18, después se minimiza la energía y el sistema se somete a una ejecución de MD, guardando la instantánea final para el paso 3. En el paso 2, se prepara la proteína. Se eliminaron de la estructura las moléculas orgánicas, mientras que se mantuvieron las aguas estructurales. La proteína se cargó en el módulo LEaP de la suite Amber18, donde se añadió, además de la caja cúbica de agua TIP3P equilibrada y el ligando, átomos de hidrógeno y un número apropiado de contraiones de sodio para que el sistema fuera eléctricamente neutro. En el paso 3, la proteína se disuelve utilizando varias copias de la caja que contiene el fragmento producido en el paso 1. En el paso 4, el sistema se somete a la minimización de la energía y al equilibrado, seguido de simulaciones de MD. En el paso 5, se eliminan las moléculas de agua y los contraiones de la trayectoria solvatada de proteína-ligando. Se genera un archivo pdb en la última instantánea de cada una de las N trayectorias, realizado con el script fdMD_OneTraj.cmd. En el paso 6 cada una de las N trayectorias se analiza utilizando el script fdMD_ReactiveTraj.py para así evaluar si estas trayectorias han sido reactivas o no, en el sentido de que el ligando interactúa con un posible sitio de unión durante un tiempo el tiempo de simulación.

4.3.2. Minimización de la energía.

Esta minimización se tiene que llevar a cabo mediante un proceso por etapas y es necesario para eliminar el posible estrés conformacional de la proteína.

Para conseguir dicha minimización, como cualquier proceso en Amber, son necesarios:

- 1) Fichero .top, se trata de un fichero de topología para el ligando, el mismo en todo momento. Este fichero proporciona los parámetros necesarios del campo de fuerza para cada uno de los átomos involucrados en el sistema. Se prepara a partir de un archivo .pdb gracias al módulo LEaP
- 2) Fichero con las coordenadas de todos los átomos.
- 3) Fichero de entrada, contiene las órdenes iniciales que el programa tiene que ejecutar. Se trata del fichero input.

La minimización se ejecuta en 5 etapas de minimización de 5000 ciclos cada una.

En la primera etapa, se mantienen fijos los átomos de proteína y receptor, y se minimiza el disolvente y contraiones. En la segunda, tercera y cuarta etapa se mantienen fijos los átomos de los aminoácidos. En la quinta etapa se permite el movimiento de todos los átomos

4.3.3. Calentamiento.

Se produjo un aumento gradual de la temperatura. En esta etapa de calentamiento se usa un comando indicando de este que este paso de calentamiento se realice a volumen constante (ntb=1).

Como el calentamiento supone dar mucha energía, si hay un enlace que esté muy tensionado puede que se aumente la energía más en dicha parte y aumente exponencialmente y lleve a la “explosión”. De no hacer dicha subida de la temperatura gradualmente se produciría la desnaturalización del sistema.

4.3.4. Equilibrado.

Cuando se alcanza la temperatura deseada de 300K, la presión se mantiene constante (ntb=2). Al variar ahora el volumen, el sistema puede aumentar su densidad hasta alcanzar su estado de equilibrio.

4.3.5. Producción y simulación MD.

En este punto ya queda completado el sistema para que se desarrolle la simulación de MD.

Mientras que el proceso de simulación está en curso se va indicando el tiempo de duración de cada paso de integración con el comando $dt=0.004$. Esto quiere decir que tendrán un tiempo de 4fs por paso.

En la simulación, con el comando $dt=0.004$, se indica que la unidad mínima de paso temporal, conocida como timestep es de 0.004 ps (4fs), y con el comando $ntwx=5000$ se indica que cada 5000 timestep se escribe el estado del sistema. Al multiplicar ambos datos (0.004×5000) se obtiene un valor de 20 ps, esto quiere decir, que cada 20 ps se va a reflejar el estado del sistema en la trayectoria. A esta descripción del estado instantáneo del sistema se le llama snapshot. Como cada 5000 ps se describe el estado del sistema, y se toma una snapshot cada 20 ps, quedan 250 unidades de tiempo con las cuales se constituye la trayectoria de la simulación, conocida también como frame. Se conseguirá el estado del sistema, y por tanto todas las variables que se van a estudiar cómo es la de la energía con estas frames.

4.4. Análisis de fdmd

4.4.1. CPPTRAJ y análisis de las trayectorias reactivas.

Con el programa cpptraj se analizaron los posibles enlaces de hidrógeno, tanto intermoleculares como intramoleculares, establecidos durante la simulación, averiguándose la distancia de estos enlaces en cada foto de la trayectoria.

Las trayectorias se analizaron utilizando 4 scripts de acuerdo con el protocolo de la metodología fdMD.

Mediante el primer script `fdMD_OneTraj.cmd`, se obtuvieron las trayectorias que contenían solo a la proteína con uno de los ligandos, eliminando las moléculas de agua de la trayectoria y dividiéndola esta en N trayectorias proteína-ligando, siendo N el número de ligandos en la caja de simulación.

Cada una de las N trayectorias se analizan empleando el script `fdMD_ReactiveTraj.py` para evaluar si son reactivos o no, es decir, si el ligando había interactuado con un posible sitio de unión durante el tiempo de simulación preestablecido, que en nuestro caso fue de 20ns.

Se recogen todos los datos obtenidos para ejecutar un cálculo en MMGBSA. Posteriormente, se utilizan los scripts 3 y 4, `fdMD_MMGBSA_Send.cmd` y `fdMD_MMGBSA_Analyze.py`, respectivamente, con estas trayectorias reactivas para generar las entradas necesarias para ejecutar el siguiente paso, los cálculos MMGBSA y para extraer y graficar los datos de dichos cálculos. (Miller *et al.*, 2012).

Así, los archivos de salida que se obtuvieron permitieron visualizar los posibles sitios de unión, el tiempo que el ligando permanece en cada posible bolsillo de unión y la energía de unión MMGBSA. (Privat *et al.*, 2020).

4.4.2. *MMGBSA.py*

Se emplearon en este punto los scripts 3 y 4 `fdMD_MMGBSA_Send.cmd` y `fdMD_MMGBSA_Analyze.py`, respectivamente, con estas trayectorias reactivas para generar las entradas necesarias para ejecutar el siguiente paso, los cálculos MMGBSA y para extraer y graficar los datos de dichos cálculos. Así, los archivos de salida que se obtuvieron nos permitieron visualizar los posibles sitios de unión, el tiempo que el ligando permanece en cada posible bolsillo de unión y la energía de unión MMGBSA. Además, la última instantánea de cada trayectoria se utilizó para calcular la energía de unión con KDEEP con el fin de comparar ambos valores y proponer los posibles sitios de unión. (Privat *et al.*, 2020).

4.4.3. *Cálculo de la energía libre*

En dinámica molecular, uno de los objetivos es poder calcular la energía libre de Gibbs de los sistemas. En este trabajo final, los cálculos de energía libre de enlace utilizaron MMGBSA y de área superficial de Poisson-Boltzmann de mecánica cuántica (MMPBSA) tal y como se efectúan en el paquete de Amber. El cálculo de la energía libre de Gibbs obtenida con el ciclo termodinámico, se obtiene con la ecuación 4.1. Dicha energía está estimada a partir de las energías libres de solvatación de los reactivos, proteína (P), ligando (L) o del complejo proteína-ligando (PL).

$$\Delta G_{enlace} = \Delta G_{enlace}^0 + \Delta G_{PL}^{0 \rightarrow solv} - \Delta G_L^{0 \rightarrow solv} - \Delta G_P^{0 \rightarrow solv}$$

Ecuación 4.1.

Con MMGBSA, calculamos la energía libre para un proceso, es decir, la unión del ligando a la proteína. Esto se calcula a partir de la entalpía y de la entropía según la siguiente ecuación 4.2:

$$\Delta G_{enlace}^0 = \Delta H_{enlace}^0 - T\Delta S_{enlace}^0$$

Ecuación 4.2.

La contribución de la entalpía se puede expresar tal y como se muestra en la ecuación 4.3. Esta energía se obtiene teniendo en cuenta la entalpía estándar interna (ΔH_{int}^0), la contribución referida a las fuerzas de van der Waals (ΔH_{vdW}^0) y por último, las electrostáticas (ΔH_{elec}^0).

$$\Delta H_{enlace}^0 = \Delta H_{int}^0 - \Delta H_{vdW}^0 + \Delta H_{elec}^0$$

Ecuación 4.3. Cálculo

Agrupando las ecuaciones anteriores, 4.2 y 4.3 se puede calcular la variación de la energía libre de Gibbs como la suma de las energías libres de cada estado.

$$\Delta G = \Delta H_{int}^0 - \Delta H_{vdW}^0 + \Delta H_{elec}^0 - \Delta G^{solv} - T\Delta S$$

Ecuación 4.4.

En esta última expresión, ecuación 4.4, quedan representados una serie de términos. Los tres primeros referidos a enlace, interacciones de van der Waals y a las interacciones electrostáticas. Los dos últimos términos se deben a la energía libre de solvatación (ΔG^{solv}) y al término entrópico. (Gohlke y Case, 2004). Las interacciones de van der Waals y electrostáticas son las que aportan mayor contribución a la energía total del sistema. Dichas contribuciones se detallan a continuación. (Granadino-Roldán *et al.*, 2014).

4.4.3.1. Contribuciones de las interacciones en la energía.

Como ya se ha dicho, dentro de las contribuciones que aportan más energía estás las electrostáticas. Esta interacción energética es debida a todas las cargas de los átomos, tanto las cargas inducidas como las permanentes. En la

ecuación 4.5 se muestra la ecuación de la ley de Coulomb, la más empleada para calcular estas interacciones:

$$E_{el}(R_{AB}) = Q^A Q^B / \epsilon R_{AB}$$

Ecuación 4.5.

Donde R_{AB} es la distancia entre los átomos A y B, Q^A y Q^B las cargas de los átomos A y B, y ϵ la constante dieléctrica.

Las contribuciones electrostáticas habitualmente se modelan haciendo uso de cargas puntuales centradas en los átomos, lo que se conoce como cargas parciales, y con ellas se representa la densidad de carga molecular. En una representación molecular real es necesario contar tanto con los momentos dipolares, y así dejar incluidos los pares no compartidos, asimismo se debe tener en cuenta a los momentos cuadrupolares, incluyendo así los enlaces π . Cuando se asignan cargas parciales, se suelen alcanzar cargas demasiado grandes para aquellos átomos que no están en la superficie de la molécula. Añadiendo un término de penalización distinto de cero, se soluciona el problema de las cargas parciales grandes para los átomos que no están en la superficie. Esto se lleva a cabo con Amber. (Cisneros *et al.*, 2014; Carazo, 2015).

A diferencia de las interacciones electrostáticas, las de van der Waals no son producidas por las cargas de los átomos, se deben a la energía que se describe entre la atracción o repulsión que se da entre los átomos que no están directamente unidos. Normalmente, estas interacciones se calculan con la ecuación 4.6, conocida como ecuación de Lennard-Jones:

$$E_{vdW}(R_{AB}) = (C_1/R^{12}) - (C_2/R^6)$$

Ecuación 4.6.

En dicha ecuación R_{AB} es el radio de van der Waals, y C_1 y C_2 como constantes. El parámetro R_{AB} de esta ecuación se obtiene como una suma de radios atómicos, asumiéndose los átomos como esféricos.

La energía debida a estas interacciones de van der Waals (E_{vdw}) toma el valor cero para distancias interatómicas grandes. En distancias interatómicas medias se produce una atracción dipolo inducido entre las nubes electrónicas de los

átomos, y por último, es muy repulsiva en distancias cortas, debido al solapamiento que se produce de las nubes electrónicas de los dos átomos, debido a que los electrones se repelen entre sí porque son de cargas opuestas. La magnitud R_{AB} de la ecuación 4.6 de Lennard-Jones se obtiene como la sumatoria de los radios atómicos del sistema. En las contribuciones que se están tratando en este apartado, se entienden a los átomos como esféricos, pero esto no siempre es así: tiene como excepción a los átomos de hidrógeno con una distribución electrónica que difiere mucho de ser esférica, y por otro lado, tampoco se consideran átomos esféricos a aquellos que tienen pares de electrones sin compartir. Pero estas dos excepciones no son de mucha importancia para las interacciones de van der Waals, ya que son átomos que participan en los enlaces de hidrógeno y este tipo de enlace tiene tendencia electrostática. (Jensen y Wiley, 2007; Carazo, 2015).

5. RESULTADOS.

Se obtuvieron dos dinámicas de simulación molecular para explorar los posibles sitios de unión entre ligando y proteína. Como se ha comentado anteriormente, cada una de las N trayectorias que se obtuvieron se analizaron con el script `fdMD_ReactiveTraj.py` para evaluar si eran reactivas o no.

En la caja de simulación de partida, tenemos desde el 970 al 1027 ligandos, y tal y como se ha comentado con el script (`fdMD_ReactiveTraj.py`) sólo nos quedamos con tres de todos ellos, ya que se observó que estos tres eran los reactivos. Los ligandos reactivos con los que nos quedamos y de los cuáles hacemos este estudio, los numeramos como 980, 996 y 1022. Con el script `fdMD_ReactiveTraj.py` se analizó estas tres trayectorias reactivas, midiendo la distancia a la que se encontraba el átomo central, denominado C99, respecto de la proteína. Esta distancia la consideramos desde la última instantánea, comparándola con todas las anteriores durante el tiempo que se desarrolla la simulación. Las trayectorias se consideran reactivas si por un lado, el átomo C99 tiene al menos un átomo de la proteína a una distancia próxima a la proteína en la última instantánea y por otro lado, que esta distancia se mantenga en ese sitio durante al menos el 90% de ocupación en las

instantáneas durante los últimos ns del tiempo de simulación. Este mismo script es el que se encarga de eliminar los archivos relacionados con las trayectorias no reactivas y además genera un archivo PDB de cada última instantánea de estos tres que se utilizará para obtener las energías de enlace.

A continuación, mostramos los resultados obtenidos para la distancia del C99 entre la última foto y las anteriores, y eso nos ayuda a predecir si el ligando ha encontrado un sitio de unión y ha estado ahí estable al menos al final de la simulación.

5.1. Distancia de enlace frente al tiempo.

En los siguientes gráficos, 5.1; 5.2; 5.3, quedan representadas las distancias a las que se encuentra el átomo central C99 con respecto de la proteína, desde donde se encuentra en la última instantánea y en todas las anteriores, durante el tiempo que dura la simulación para cada uno de los ligados reactivos.

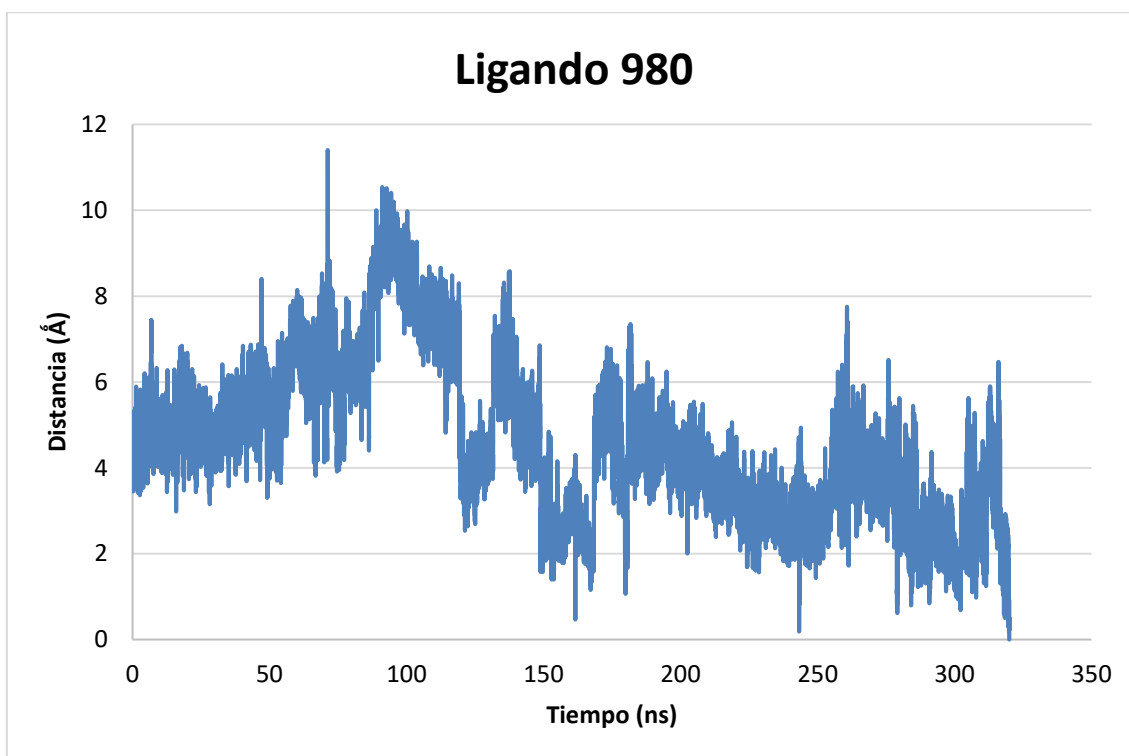


Gráfico 5.1. Distancia de C99 desde durante el tiempo que dura la simulación para el ligando 980.

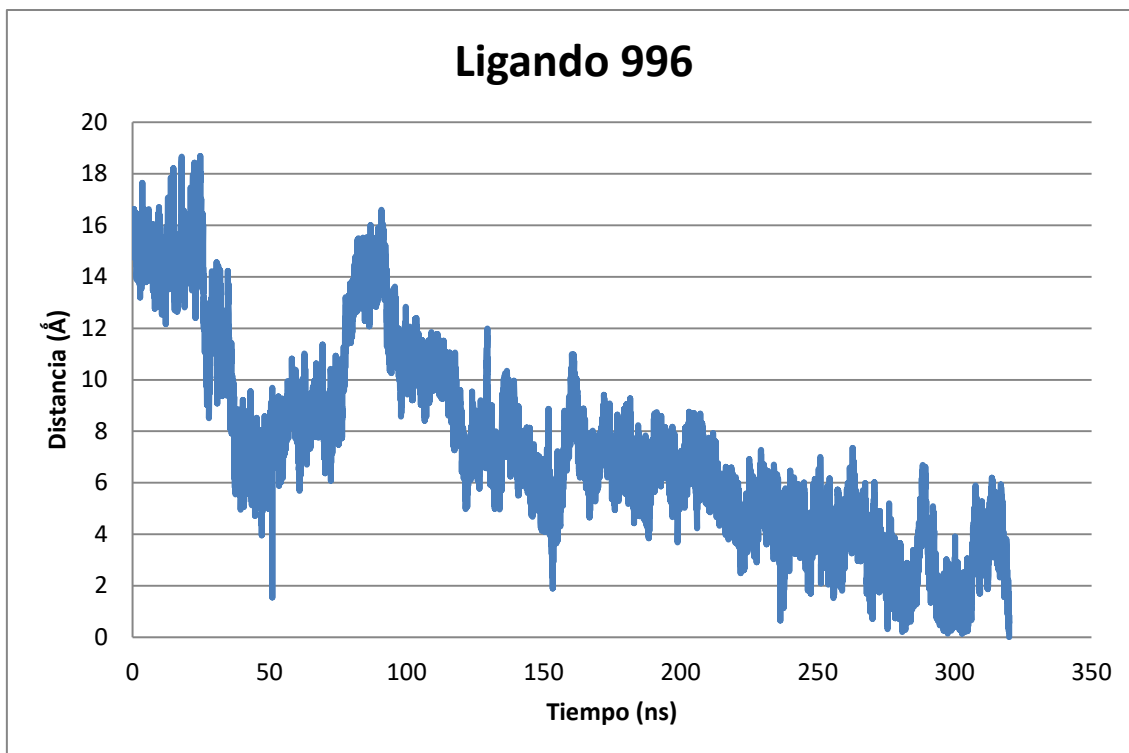


Gráfico 5.2. Distancia de C99 durante el tiempo que dura la simulación para el ligando 996.

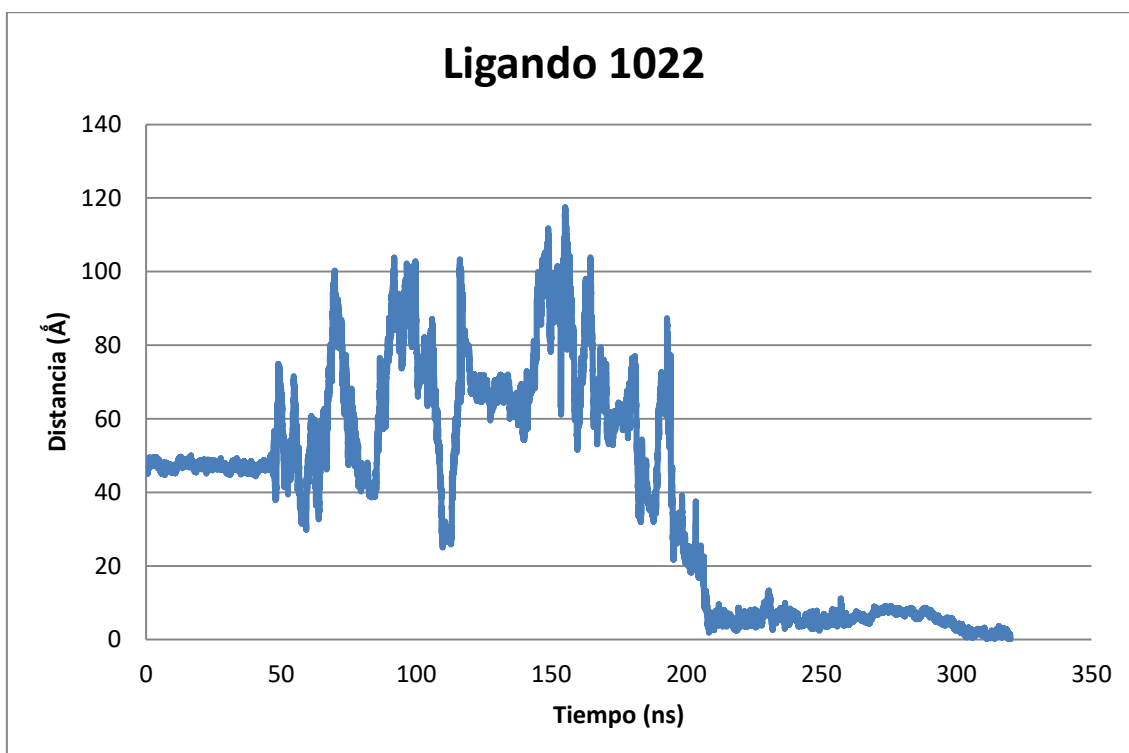


Gráfico 5.3. Distancia de C99 durante el tiempo que dura la simulación para el ligando 1022.

5.2. ΔG

5.2.1. ΔG total

En este apartado se pueden observar los gráficos 5.4, 5.5 y 5.6 correspondientes a las representaciones gráficas que se han realizado de las energías totales de interacción, en Kcal/mol, de cada una de las copias de los ligandos en función del tiempo de simulación, en ns de toda la trayectoria.

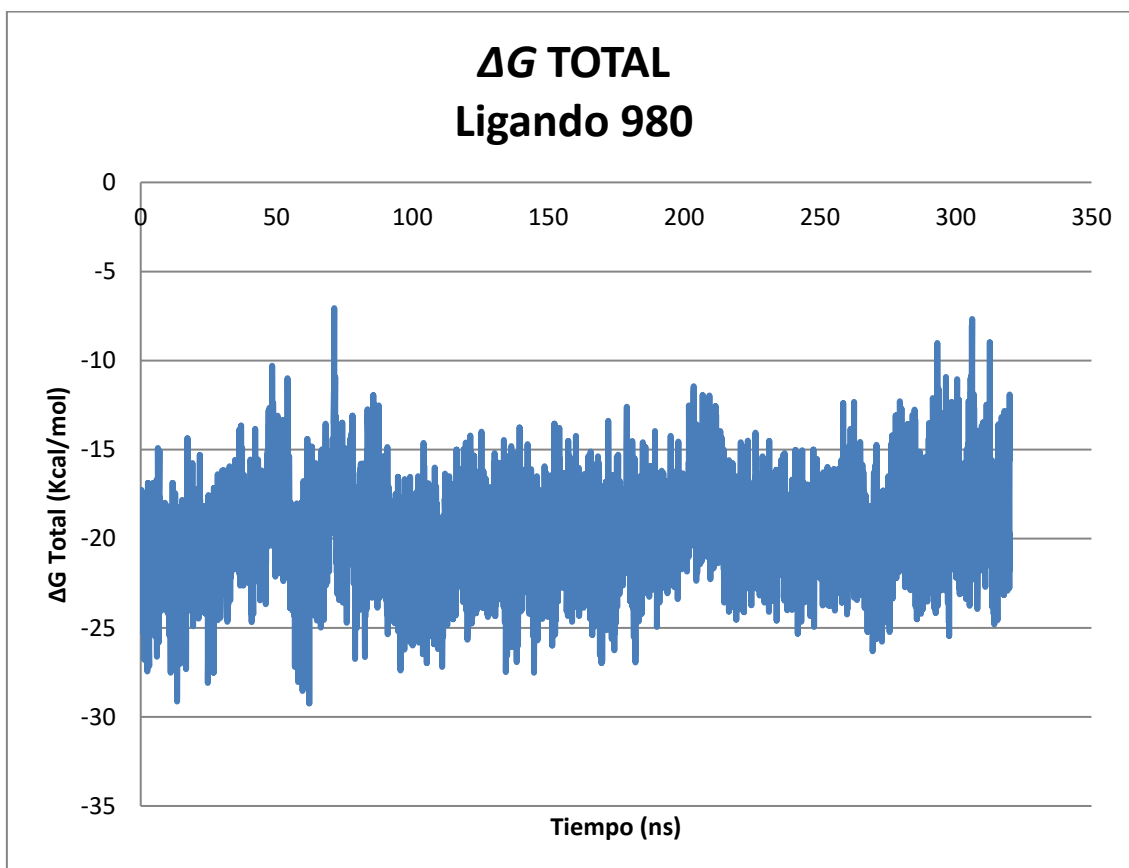


Gráfico 5.4. ΔG total frente al tiempo del ligando 980

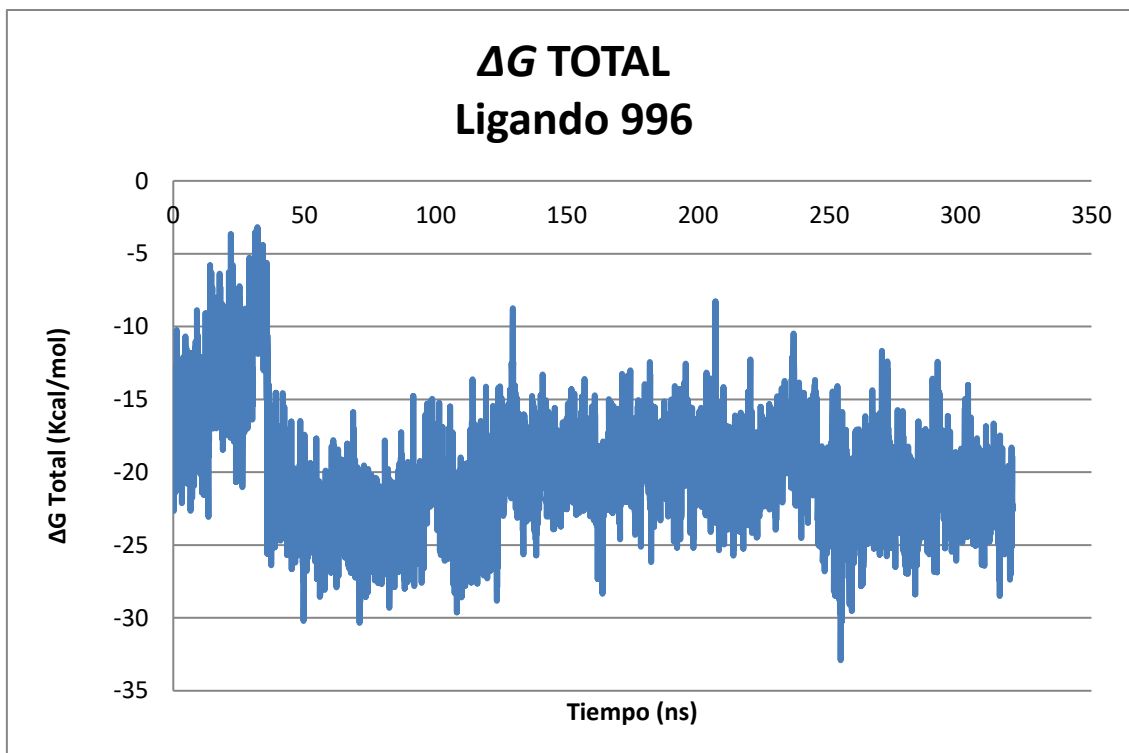


Gráfico 5.5. ΔG total frente al tiempo del ligando 996

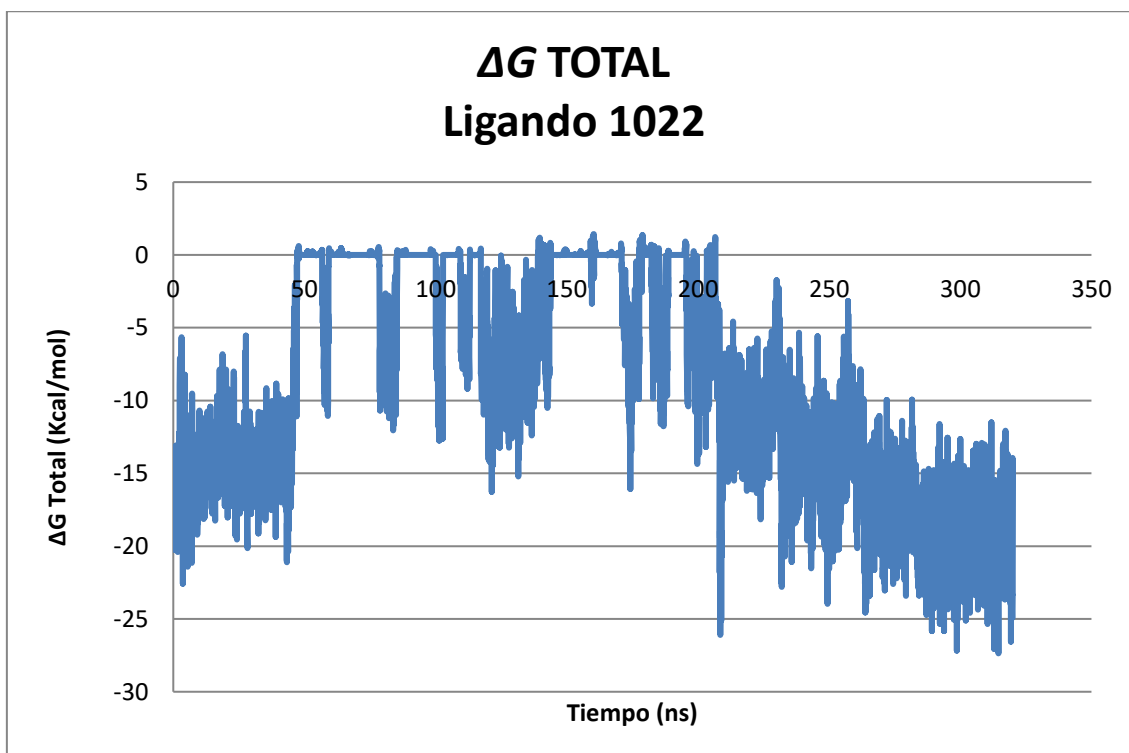


Gráfico 5.6. ΔG total frente al tiempo del ligando 1022

5.2.2. ΔG promedio.

En las siguientes tablas se muestra la energía libre promedio, por un lado obtenida con las contribuciones de van der Waals (VDWAALS), por otro lado, con las contribuciones electrostáticas (EEL), y por último la energía total media de cada uno de los ligandos.

Tabla 5.1. ΔG promedio del ligando 980

Método	Media (Kcal/mol)
VDWAALS	-29,9710
EEL	-19,6416
DELTA TOTAL	-19,8487

Tabla 5.2. ΔG promedio del ligando.

Método	Media (Kcal/mol)
VDWAALS	-29,3722
EEL	-13,2716
DELTA TOTAL	-20,1737

Tabla 5.3. ΔG promedio del ligando 1022.

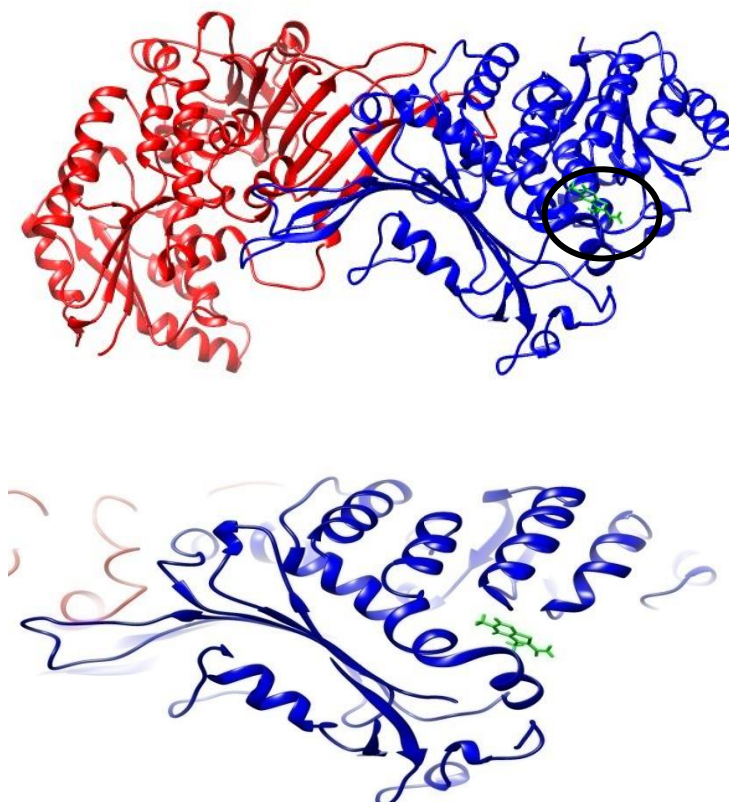
Método	Media (Kcal/mol)
VDWAALS	-13,0714
EEL	-6,2644
DELTA TOTAL	-8,9709

5.3. Análisis de las instantáneas de cada ligando

A continuación, se muestran las instantáneas iniciales y finales de cada uno de los ligandos seleccionados, poniendo la vista en el sitio donde se encuentran al inicio y luego al final de la simulación, para ver si se ha mantenido en el mismo sitio, o si por lo contrario ha ido moviéndose hasta encontrar un mejor sitio dónde poder interaccionar con la proteína y donde permanezca unido por más tiempo de la simulación que estamos ejecutando. En las figuras que se muestran a continuación, fig. 5.1; 5.2; 5.3; se corresponden con las imágenes de la simulación de cada uno de los tres ligandos copia reactivos (980, 996 y 1022). En estas figuras, además se ha hecho una diferenciación entre la imagen al principio y al final de la simulación, siendo la letra A para el inicio y B para el final. En estas imágenes queda capturado el residuo de cada copia de los ligandos que permanece unido a la proteína. Estas imágenes han sido obtenidas con el programa UCSF Chimera.

Instantáneas del Ligando 980: En la fig 5.1. (A) se muestra la posición que tiene este ligando reactivo al inicio de la simulación y en (B) al final de la simulación.

A



B

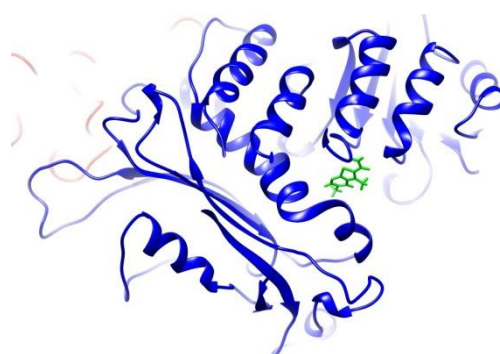
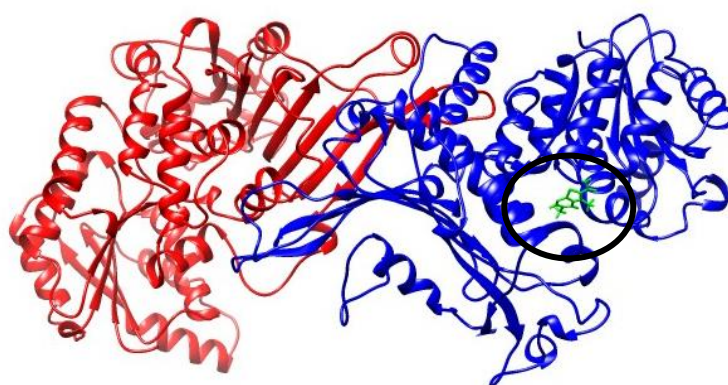
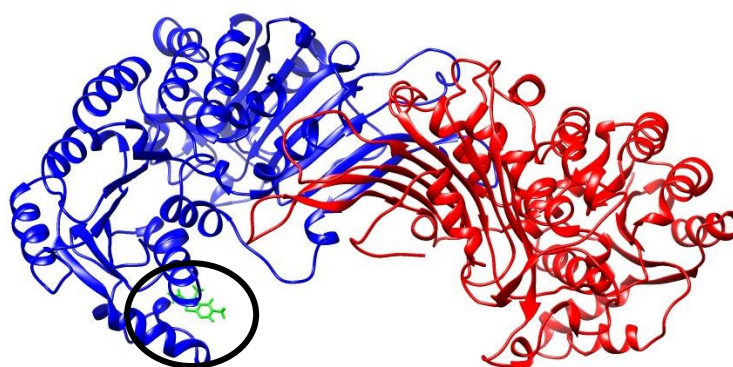


Figura 5.1. Situación del ligando 980.

Instantáneas del ligando 996: En la fig 5.2. (A) se muestra la posición que tiene este ligando reactivo al inicio de la simulación y en (B) al final de la simulación.

A



B

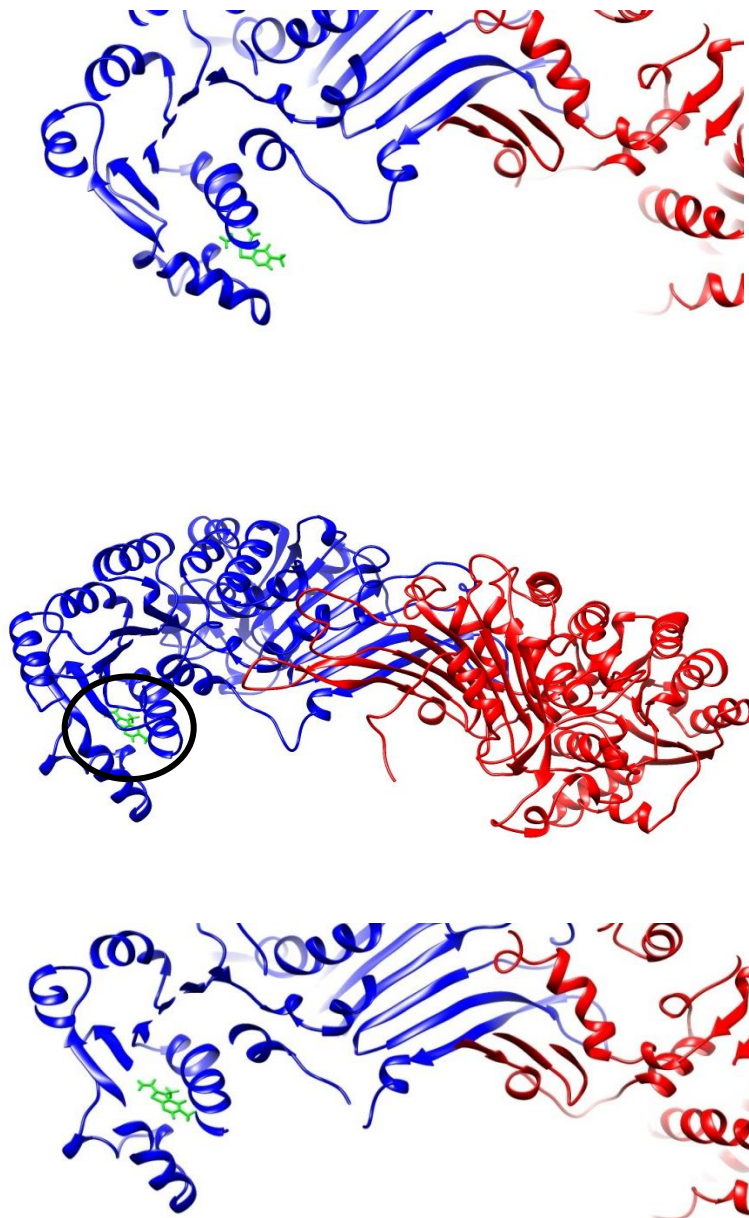
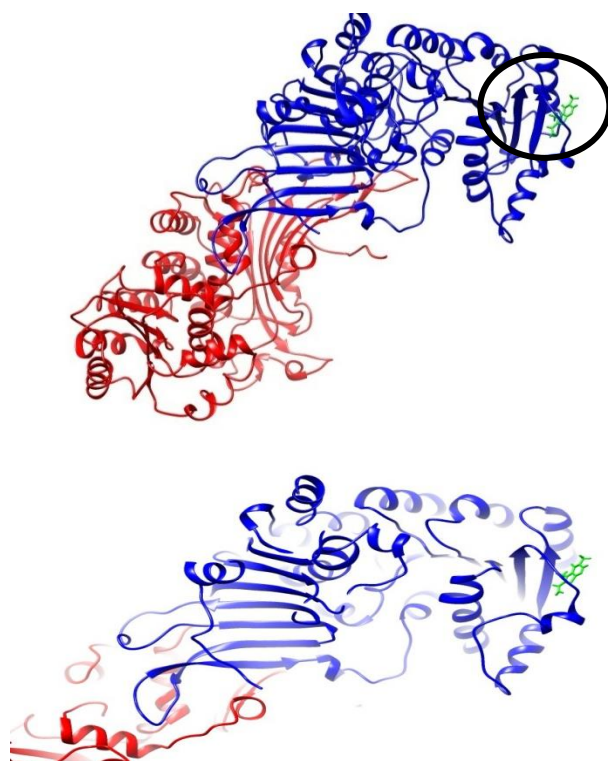


Figura 5.2. Situación del ligando 996.

Instantáneas del ligando 1022: En la fig 5.2. (A) se muestra la posición que tiene este ligando reactivo al inicio de la simulación y en (B) al final de la simulación.

A



B

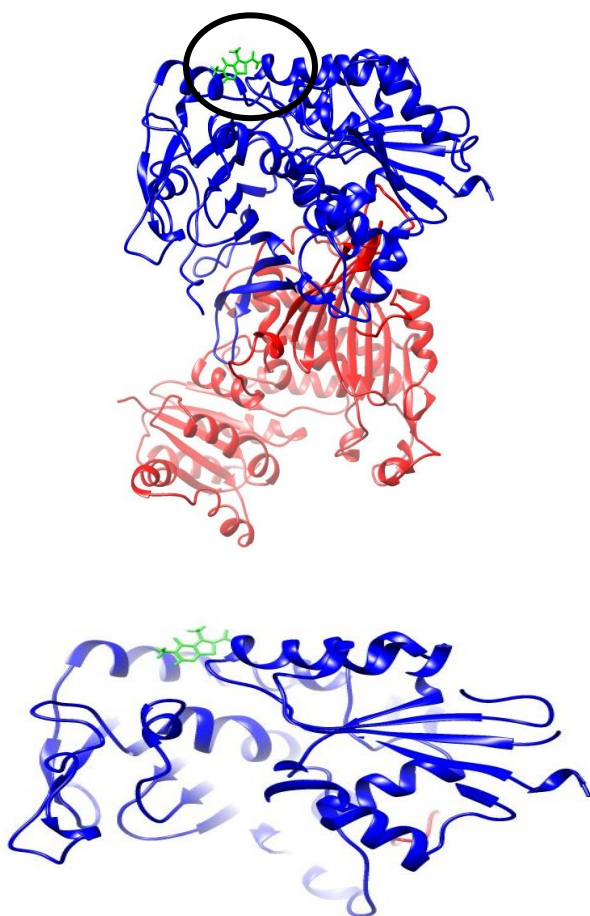


Figura 5.3. Situación del ligando 1022.

6. DISCUSIÓN.

En este apartado se va a ir haciendo un estudio detallado de todos los resultados obtenidos y que se han ido reflejando con anterioridad en este proyecto. Nos centraremos en el estudio de cuatro aspectos diferentes de los que se han obtenido unos resultados tras realizarse la simulación molecular. Los cuatro aspectos en los que nos vamos a fijar son:

- Estudios de la distancia a la que se encuentra el átomo central C99 desde su posición en la última foto, con respecto a todas las anteriores. Esto es de gran utilidad, ya que ayuda a decidir si el ligando ha encontrado un sitio de unión y si ha estado ahí al menos al final de la simulación.
- Se realiza además un estudio de los valores que se obtienen de energía libre para cada ligando en función del tiempo en el que se produce la simulación. ΔG es la energía de la interacción, y persigue encontrar el menor valor de esta energía de los tres ligandos activos con los que finalmente únicamente nos hemos quedado. Cuanto menor sea el valor de ΔG quiere decir que la unión que se establece entre el ligando y la proteína es más fuerte. Se suele acudir al estudio de ΔG frente al tiempo, para conocer si la dinámica molecular ha convergido o no.
- Por otro lado, se contemplan los resultados obtenidos de las interacciones de van der Waals y electrostáticas.
- Asimismo, ha sido de gran importancia visualizar con UCSF Chimera cómo se han ido posicionando los ligandos en la proteína a lo largo de la simulación.

El estudio de todo esto tiene como fin poder conocer por dónde se une el ligando al que llamamos modelo y la proteína y en posteriores estudios llegar a poder plantear modificaciones a dicho compuesto que aumenten su afinidad para con la proteína.

Tras estudiar la distancia a la que se encuentra el átomo central, C99, desde su posición final, con respecto a todas las anteriores, hemos obtenido los gráficos 5.1, 5.2, 5.3. En ellos se representa la distancia en Å del átomo C99 desde el tiempo inicial (0 ns) hasta el final (320 ns).

El gráfico 5.1 es el que se corresponde con el ligando 980. Podemos ver cómo el átomo C99 se ha ido encontrando a diferentes distancias de la proteína. En esta copia de ligando, a simple vista, se observa que el ligando se ha mantenido muy próximo al mismo sitio de unión. No obstante, hay que decir, que ha tenido momentos en los se ha alejado un poco de dicha posición pero conforme la simulación se ha ido acercando al final del tiempo, esta distancia se va acortando y va tomando una posición más próxima con la proteína.

En el gráfico 5.2 podemos ver lo mismo comentado anteriormente solo que con la copia de ligando 996. De este se puede señalar, en comparación con el otro, que ha tenido más sitios de unión diferentes en las que se ha posicionado hasta ir recortando distancia con la proteína conforme la simulación ha transcurrido.

El ligando copia 1022 es el que aparece representado en el gráfico 5.3. Se considera que este ligando es el que ha tenido sitios de unión más dispares en comparación con los otros dos anteriores. Alrededor de los primeros 50 ns de la simulación, este ha estado a la misma distancia de la proteína, pero desde aquí hasta aproximadamente los 200 ns el ligando se ha estado moviendo alrededor de la proteína hasta que se ha unido a ella en un sitio de unión mucho más cercano que el sitio al que se unía al comienzo de la simulación.

Ahora vamos a hacer el estudio de los datos obtenidos para la energía libre total de cada uno de los ligando copia con respecto del tiempo. Cabe mencionar, que si el valor de la energía se aproxima o se hace 0 quiere decir que el ligando copia se ha encontrado próximo al agua

El gráfico 5.4 muestra los resultados para el 980. Se puede visualizar una consecución de resultados muy parecidos de la energía. Este ligando no parece ser que se haya aproximado mucho al agua ya que no presenta valores muy próximos al 0, lo que quiere decir que ha estado más cercano a la proteína.

En el gráfico 5.5 tenemos los resultados para el 996. Este al principio de la simulación, en torno a los primeros 20 ns, ha estado más alejado de la proteína, incluso se puede decir que ha estado por agua debido a que se dan valores de 0 kcal/mol. Pero tras darse estos primeros 20 ns, a este ligando copia le ocurre lo mismo que al anterior, mantiene unos valores de energía más o menos constantes sin sufrir variaciones de valores muy notables.

Pero en el gráfico 5.6 sí que se puede discutir diferencias significativas con respecto al 980 y 996. Este muestra más resultados de energía con valores de 0 o próximos a 0 durante un mayor tiempo de simulación, es decir, que se ha encontrado durante mayor tiempo alejado de la proteína y más próximo al agua. Al final si se observa como la energía del sistema va disminuyendo, indicando así que se va produciendo una interacción con la proteína más fuerte.

Podemos superponer los gráficos 5.1. con 5.4, 5.2 con 5.5 y el 5.3 con 5.6 y podemos ver como se relacionan los valores de los tres ligandos copia y ver como se llega a una misma idea de que el que más tiempo ha tardado en interaccionar con la proteína ha sido el 1022, por todo lo que anteriormente se ha detallado.

Las ΔG promedio quedan plasmadas en las tablas 5.1, 5.2, 5.3. A partir del snapshot 1, la copia del ligando 980 está estable en el binding site, es decir permanece estable durante toda la simulación en el sitio de unión a la proteína. El ligando 996, permanece estable a partir del snapshot 2350. Y por último, el 1022 se tiene en cuenta a partir del snapshot 10400. Como se ha comentado anteriormente se persigue localizar la energía más baja que es la que lleva relacionada una interacción ligando proteína más fuerte. Pero se tiene que tener en cuenta las interacciones que se establecen tanto de van der Waals como las electrostáticas, ya que sabemos que son las contribuciones más importantes para la energía.

Tras conocer los datos conseguidos de los enlaces de hidrógeno que se establecían entre el ligando y la proteína, se ha comprobado que estos o no se establecían o lo hacían con un porcentaje de ocupación bajo, lo más que se ha producido ha sido de un 18%, por ejemplo, en el 980. Debido a esto, para este estudio, no se han considerado los enlaces de hidrógeno. Al no darse esta interacción o en muy poca proporción, el valor obtenido para las interacciones electrostáticas de cada uno de los ligandos reactivos, puede ser menor que el valor de las de van der Waals, ya que en estas interacciones quedan implícitas las contribuciones de los enlaces de hidrógeno, las cuales son insignificantes en este trabajo. Los ligandos 980 y 996 presentan, curiosamente unos valores muy parecidos entre ellos. La diferencia se hace más notable con el 1022 que

posee unas energías inferiores a los otros dos, haciéndose notar nuevamente que se encuentra en un sitio de unión más distante de la proteína.

Por último, gracias a las imágenes obtenidas con UCSF Chimera, visualizamos dónde se ha encontrado el ligando al inicio y fin de la simulación para cada uno de los ligando reactivos, representado esto en las fig 5.1; 5.2; 5.3. Es muy curioso ver cómo el ligando reactivo 1022 ha mostrado un mayor desplazamiento del sitio de unión, encontrándose al comienzo de la simulación próximo a la proteína en una determinada posición y luego se ha ido alejando alrededor de la proteína hasta que se ha desplazado a otra posición diferente y se ha vuelto a acercarse a la proteína a una menor distancia de esta conforme el tiempo de simulación ha ido concluyendo.

En los otros dos ligandos no es tan notable esta variación de la posición de unión con la proteína, al inicio y fin de la simulación, se considera que el ligando 980 y 996 encuentran desde el inicio un hueco en el que posicionarse y producir la unión con la proteína y no tener que buscar otros sitios posibles más estables.

7. CONCLUSIÓN.

En el presente trabajo, presentamos como con un modelo computacional semiautomatizado denominado fdMD, se generan cajas de simulación de pequeños fragmentos solvatados que se utilizan para solvatar a su molécula diana y de esta forma poder conocer los posibles sitios de unión que se establecen entre el ligando y la proteína.

Después de estudiar los diferentes resultados que se han obtenido, no se puede llegar a una conclusión final de cuál es el mejor sitio de unión a la G6PDH. Se han ido viendo las diferentes interacciones y se ha planteado cuáles parecían ser más fuertes o menos fuertes de los ligandos copia con la proteína, según lo obtenido. Se sabe tras realizar todo el estudio de este Trabajo Fin de Grado que el análisis de la dinámica molecular nos ha ayudado a identificar los posibles sitios donde se producen las interacciones más fuertes del ligando con la G6PDH, pero no podemos concluir con la elección del mejor sitio de unión con los resultados que se han obtenido.

Este problema se podría solventar con un mayor tiempo de simulación de la dinámica molecular, que facilitaría poder ver con más seguridad la interacción del ligando con la proteína, y si realmente se queda el ligando en un sitio u otro con mayor fiabilidad y durante más tiempo de simulación.

Esto no quiere decir que la dinámica mandada no haya funcionado, todo lo contrario, de todos los ligandos de los que partíamos de un inicio se ha concluido únicamente con tres ligandos copia que se seleccionaron porque eran reactivos y los tres muestran que son muy adecuados a unirse con la proteína pero con esta simulación que se ejecutó no es suficiente para poder dar una conclusión final de este trabajo

8. BIBLIOGRAFÍA.

ANSINO, O. M. (2020) *Estudio asistido mediante software de la interacción entre un fármaco anticancerígeno y su diana biológica*, Trabajo Fin de Grado, Universidad de Jaén.

AU, S.WN., GOVER, S., LAM, V. MS., ADAMS, M.J. (1999) "Human glucose-6-phosphate dehydrogenase: the crystalstructure reveals a structural NADP+ molecule and provides insights into enzyme deficiency", *Journal of Molecular Graphics and Modelling*. 8, pp. 293-303.

BABIAK, R.M., CAMPELLO, A.P., CARNIERI, E.G., OLIVEIRA, M.B. (1998) "Methotrexate: pentosecycle and oxidative stress, Cell Biochemistry and Function."

BONILLA, J. F., M.D., M. SC., SÁNCHEZ, M. C., QUIM, L., CHUAIRE, L., M.Sc (2007); "Glucosa-6-fosfato deshidrogenasa (G6PD). Respuesta de los hematíes y otras células humanas a la disminución en su actividad" *Colombia Médica*.

BUEREN-CALABUIG, J.A. (2014) "Dinámica molecular" *Bioinformática con ñ*. 17. <http://doi.org/10.5281/zenodo.1066360>

CARAZO, B. J. (2015) *Diseño de fármacos mediante modelización molecular: aplicación a la leucemia mielocítica aguda*, Trabajo Fin de Grado, Universidad de Jaén.

CASE,D.A., BEN-SHALON, I.Y., BROZELL, S.R., CERUTTI, D.S., CHEATHAM, T.E., CRUZEIRO, V.W.D., DARDEN, T.A., DUKE, R.E., GHOREISHI, D., GILSON, M.K., GOHLKE, H., GOETZ, A.W., GREENE, D., HARRIS, R, HOMEYER, N., HUANG, Y., IZADI, S., KOVALENKO, A., KURTZMAN, T., LEE, T.S., LEGRAND, S., LI, P., LIN, C., LIU, J., LUCHKO, T., LUO, R., MERMELSTEIN, D.J., MERZ,K.M., MIAO, Y., MONARD, G., NGUYEN, C., NGUYEN, H., OMELYN, I., ONUFRIEV, A., PAN, F., QI, R., ROE, D.R., ROITBERG, A., SAGUI, C., SCHOTT-VERDUGO, S., SHEN, J.,

C.L., SMITH, J., SALOMONFERRER, R., SWAILS, J., WALKER, R.C., WANG, J., WEI, H., WOLF, R.M., WU, X., XIAOI, L., YORK, D.M. and KOLLMAN, P.A. (2018), "AMBER 2018", *University of California, San Francisco*.

CIFTI, M., BUYUKOKUROGLU, M.E., KUFREVIOGLU, O.I. (2001) "Effect of cefaperazone/sulbactam and ampicillin/sulbactam on the in vitro activity of human erythrocyte glucose-6-phosphate dehydrogenase" *Journal of Basic and Clinical Physiology and Pharmacology*, 12, pp.305-313.

CIFTI, M., OZMEN, I., BUYUKOKUROGLU, M.E., PENCE, S. KUFREVIOGLU, O.I. (2001) "Effects of metamizol and magnesium sulfate on enzyme activity of glucose-6-phosphate dehydrogenase from human erythrocyte in vitro and rat erythrocyte in vivo" *Clinical Biochemistry*. 34, pp. 297-302.

CISNEROS, G.A., KARTTUNEN, M. REN, P., SAGUI, C. (2014) "Classical electrostatic for biomolecular simulations", *Chemical Reviews*, 114 (2014) 779-814

GOHLKE, H y CASE, D.A. (2004) "Converging free energy estimates: MM-PB (GB) SA studies on the protein-protein complex Ras-Raf". *Journal of computational chemistry*, 25(2), pp.238-50

GRANADINO-ROLDÁN, J.M., OBIOL-PARDO, C., PINTO, M., GARZON, A., RUBIO-MARTÍNEZ, J. (2014) "Molecular dynamics analysis of the interaction between the human BCL6 BTB domain and its SMRT, NcoR and BCOR corepressors: The quest for a consensus dynamic pharmacophore" *Journal of Molecular Graphics and Modelling*.

HINCHLIFE, A., WILEY, J., SONS. (2000) "Modelling Molecular Structures" 2^a Edición (2000)

JESEN, F., WILEY, (2007) "Introduction to computational chem". 2^a Ed.

LOZANO-APONTE, J., SCIOR, T. (2014): “¿ Qué sabe Ud. Acerca de... Dinámica Molecular? What do you know about... Molecular Dynamics?”

KOLLMAN, P.A., MASSOVA, I., REYES, C., KUHN, B., HUO, S., CHONG, L., CHEATHAM, T.E. (2000). “Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models”. *Accounts of Chemical Research*, 33(12), 889-897.

MAIER, J. A., MARTINEZ, C., KASAVAJHALA, K., WICKSTROM, L., HAUSER, K. E., SIMMERLING, C. (2015).” ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB”. *Journal of chemical theory and computation*. 11, pp. 3696-3713.

MILLER, B.R., McGEE, T. D., SWAILS, J. M., HOMEYER, N., GOHLKE, H., y ROITBERG, A. E. (2012). “MMPBSA.py: An Efficient Program for End-State Free Energy Calculations”. *Journal of Chemical Theory and Computation*.

OBIOL-PARDO, C., ALCARRAZ-VIZÁN, G., DÍAZ-MORALLI, S., CASCANTE, M., RUBIO-MARTINEZ, J (2014); “Design of an interface peptide as new inhibitor of human glucose-6-phosphate dehydrogenase”. *Journal of Molecular Graphics and Modelling*.

OROZCO, M. (2014) “Theoretical view of protein dynamics” *Chem Soc*. 43

PDB (Protein Data Bank) www.rcsb.org/pdb/

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C., y FERRIN, T. E. (2004). “UCSF Chimera—a visualization system for exploratory research and analysis”.

PRIVAT, C., GRANADINO-ROLDÁN, J.M., BONET, J., SANTONS, M., PÉREZ, J.J., RUBIO-MARTÍNEZ, J. (2020): “Fragment dissolved molecular dynamics: a systematic and efficient method to locate binding sites”.

RAMOS-MARTINEZ, J.I. (2017). "The regulation of the pentose phosphate pathway". *Archives of Biochemistry and Biophysics*.

RAMOS-MONTOYA, W. W., LEE, S., BASSILIAN, S. LIM, R.V.TREBUKHINA, M.V., KAZHYNA, C.J., CIUDAD, V., NOE, J.J., CENTELLES, M., CASCANTE. (2006). "Pentose phosphate cycle oxidative and nonoxidative balance: a new vulnerable target for overcoming drug resistance in cáncer" *International Journal of Cancer* 119, pp. 2733-3274.

ROE, D.R., CHEATHAM III, T. E. (2013). "PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data". 9(7), 3084-3095.

TIAN, W.N., BRAUNSTEIN, L. D., PANG, J., STUHLEMEIER, K. M., XI, Q. C., TIAN, X., STANTON, R. C. (1998). " Importance of glucose-6-phosphate dehydrogenase activity for cell growth" *J. Biol. Chem*, 273, pp. 10609-17.

TIAN, W. N., BRAUNSTEIN, L. D., KIRA, A., PANG, J., ROSE, M., TIAN, X., STANTON, R. C. (1999): "Importance of glucose-6-phosphate dehydrogenase activity in cell death".

WILLIAMNS, J., CHANDRASEKHAR, J. MADURA, J.D., IMPEY, R.W., KLEIN, M.L. (1983) "Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*". 79(2):926-35.