



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA FÍSICA

**ESTUDIO DE LA INTERACCIÓN DE LA PROTEÍNA ALBUMINA
SÉRICA HUMANA CON DIVERSOS LIGANDOS UTILIZANDO
DINÁMICA MOLECULAR**

Trabajo de Integración Curricular:

Tipo: Proyecto de Investigación

Presentado para optar al grado académico de:

FÍSICO

AUTOR: IVÁN FABRICIO CHÁVEZ VELASCO

DIRECTORA: Biof. MARIA FERNANDA HEREDIA MOYANO

Riobamba – Ecuador

2022

© 2022, Iván Fabricio Chávez Velasco

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.

Yo, IVÁN FABRICIO CHÁVEZ VELASCO, declaro que el presente Trabajo de Integración Curricular es de mi autoría y los resultados de este son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este Trabajo de Integración Curricular; el patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 24 de noviembre del 2022



Iván Fabricio Chávez Velasco

0605674936

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA FÍSICA

El Tribunal del Trabajo de Integración Curricular certifica que: El Trabajo de Integración Curricular; tipo: Proyecto de Investigación, **ESTUDIO DE LA INTERACCIÓN DE LA PROTEÍNA ALBUMINA SÉRICA HUMANA CON DIVERSOS LIGANDOS UTILIZANDO DINÁMICA MOLECULAR**, realizado por el señor: **IVÁN FABRICIO CHÁVEZ VELASCO**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Integración Curricular, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal Autoriza su presentación.

	FIRMA	FECHA
Phd. Jiri Svozilik PRESIDENTE DEL TRIBUNAL		2022-11-24
Biof. María Fernanda Heredia Moyano DIRECTORA DEL TRABAJO DE INTEGRACIÓN CURRICULAR		2022-11-24
Mat. Luis Marcelo Cortez Bonilla ASESOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR		2022-11-24

DEDICATORIA

El presente proyecto de investigación se la dedico a mi madre, Mónica Velasco, quien estuvo presente en mi vida desempeñando el papel de progenitora, amiga, confidente y padre, lo que me permitió haber llegado hasta este momento tan importante de mi formación profesional. A mis abuelos: César y Norma quienes desde me pequeño me acogieron en su hogar con paciencia y solidaridad. A mi fuente de vitalidad, mi hermano Ariel, por compartir su tiempo y estar siempre a mí lado en los buenos y malos momentos, reviviendo mis ganas de seguir adelante. A mi padre, por enseñarme a vivir sin rencores, por demostrar que el pasado no determina tu futuro y sonreír a pesar de la adversidad.

Iván

AGRADECIMIENTO

Agradezco a mis padres, a mis abuelos y mi hermano por ser apoyo y fortaleza en aquellos momentos de dificultad y debilidad. A compañeros de clase que se volvieron amigos de vida e hicieron que las experiencias de este periodo de desarrollo se conviertan en bonitas anécdotas. También quiero agradecer a los docentes de la carrera de Física de la Escuela Superior Politécnica de Chimborazo por haber compartido sus conocimientos y experiencias, de manera especial a la biofísica María Fernanda Heredia, tutor de este proyecto de investigación. De igual manera, quiero expresar mi agradecimiento a la biofísica Daniela Barragán quien con su experiencia, conocimiento y motivación me orientó a culminar este trabajo con los mejores resultados posibles.

Iván

TABLA DE CONTENIDOS

ÍNDICE DE TABLAS^{ix}

ÍNDICE DE ILUSTRACIONES^x

ÍNDICE DE ANEXOS^{xii}

ÍNDICE DE ABREVIATURAS^{xiii}

RESUMEN^{xiv}

INTRODUCCIÓN¹

3

1. MARCO TEÓRICO REFERENCIAL³

1.1. Las proteínas³

1.1.1. *Interacciones proteína-ligando*⁴

1.1.2. *Acoplamiento del ligando*⁴

1.1.3. *Sitio de enlace*⁵

1.2. La proteína albumina sérica humana (HSA)⁵

1.2.1. *Estructura*⁶

1.2.2. *Sitios de Enlace*⁷

1.2.3. *El ligando ácido graso*⁷

1.2.3.1. *Cambios conformacionales del HSA en presencia de ácidos grasos*⁷

1.2.3.2. *Estructura del ácido cáprico*⁸

1.2.3.3. *Interacción del ácido cáprico con la HSA*⁸

1.2.4. *El ligando Warfarina*⁹

1.2.4.1. *Estructura conformacional*⁹

1.2.4.2. *Interacción de la Warfarina con la HSA*¹⁰

1.3. Dinámica molecular¹¹

1.3.1. *Principios básicos*¹²

1.3.1.1. *Integrando las ecuaciones de movimiento*¹³

1.3.1.2. *Condiciones iniciales (CI)*¹⁴

1.3.1.3. *Condiciones de frontera periódicas*¹⁵

1.3.2. *Integrador/ ensamble*¹⁶

1.3.2.1. *Ensamble NVT*¹⁷

- 1.3.2.2 *Ensamble NPT*19
- 1.3.1. *Topología molecular*20
 - 1.3.1.1. *Campo de fuerza*21
 - 1.3.1.2. *Campo de fuerza CHARMM*23
 - 1.3.1.3. *Restricción geométrica*25
 - 1.3.1.4. *Campo de Fuerza del ligando*26
 - 1.3.1.5. *Modelo para la molécula de agua*27
- 1.3.2. *Minimización de Energía*27
- 1.3.3. *Equilibrio del sistema*29
- 1.3.4. *Cantidades estadísticas simples*30
- 1.3.5. *Análisis estructurales y de trayectorias*30

32

- 2. **MARCO METODOLÓGICO**32
 - 2.1. **Recursos**32
 - 2.1.1. *Computacionales*32
 - 2.1.2. *Software*33
 - 2.1.3. *Virtuales*35
 - 2.1.4. *Python*36
 - 2.2. **Métodos matemáticos y algoritmos**37
 - 2.2.1. *Método del descenso más pronunciado*37
 - 2.2.2. *Integración del tiempo*38
 - 2.2.3. *Algoritmos de Acoplamiento*39
 - 2.3. **Metodología para la simulación de dinámica molecular**39
 - 2.3.1. *Archivos generados*42

44

- 3. **MARCO DE RESULTADOS Y DISCUSIÓN DE LOS RESULTADOS**44
 - 3.1. **Configuración y preparación de las muestras a simular**44
 - 3.1.1. *De la proteína HSA*44
 - 3.1.2. *Del complejo HSA-DKA*46

- 3.1.3. *Del complejo HSA-WFR*49
- 3.2. **Mediciones físicas de las muestras**52
 - 3.2.1. *Minimización de Energía*52
 - 3.2.2. *Equilibrio del sistema*53
- 3.3. **Cambios conformacionales**57

CONCLUSIONES64

RECOMENDACIONES66

1

3

11

ÍNDICE DE TABLAS

Tabla 1-1:3

Tabla 2-1:8

Tabla 3-1:9

Tabla 1-2:32

Tabla 2-2:33

Tabla 3-2:33

Tabla 4-2:34

Tabla 1-3: Características de la celda unitaria con la HSA solvatada y neutralizada45

Tabla 2-3: Características de la celda unitaria con la HSA – DKA solvatada y neutralizada.48

Tabla 3-3: Características de la celda unitaria con la HSA – WFR solvatada y neutralizada.51

Tabla 4-3: Resumen de las medidas físicas obtenidas en la fase de Minimización de energía, ensamble nvt y npt.57

Tabla 5-3: Promedios de las medidas analizadas en luego de la simulación de DM, para detectar cambios conformacionales62

ÍNDICE DE ILUSTRACIONES

Ilustración 1-1:3

Ilustración 2-1:6

Ilustración 3-1: Ilustración de un sistema en una simulación de dinámica molecular.12

Ilustración 4-1: Condiciones de contorno periódicas (CCP) con seguimiento a los átomos de la “super celda”, que se replica infinitamente en las tres direcciones.16

Ilustración 5-1:21

Ilustración 1-2: Algoritmo del método de descenso más pronunciado para optimizar una función no lineal38

Ilustración 2-2:42

Ilustración 1-3: Modelo de cinta de la proteína HSA sin ligandos con el pdb original.44

Ilustración 2-3: Modelo de cinta de la proteína HSA sin ligandos luego de la mutagénesis.45

Ilustración 3-3: Modelo de la celda unitaria con la HSA neutralizada y solvatada.46

Ilustración 4-3: Modelo de cinta de proteína HSA con diez ligandos DKA (amarillo). Visualización generada con el pdb original.47

Ilustración 5-3: Modelo de cinta de la proteína HSA con un solo ligando DKA (morado) luego del procesado y reconstrucción del sistema.47

Ilustración 6-3: Modelo de la celda unitaria del sistema HSA - DKA neutralizado y solvatado.49

Ilustración 7-3: Visualización generada con el pdb original. Modelo de cinta de dos proteínas HSA con un ligando de WFR (verde) cada una.50

Ilustración 8-3: Modelo de cinta de la proteína HSA con el ligando WFR (verde) luego del procesado y reconstrucción del sistema.50

Ilustración 9-3: Modelo de la celda unitaria del sistema HSA - WFR (verde) neutralizado y solvatado.51

Ilustración 10-3: Energía potencial vs Número de paso del método de descenso más pronunciado, durante la minimización de energía de cada sistema52

Ilustración 11-3: Temperatura vs tiempo de los tres sistemas luego del ensamble nvt. A) Datos crudos obtenidos y B) La temperatura promedio cada 5 segundos.54

Ilustración 12-3: Presión vs tiempo de los tres sistemas luego del ensamble npt. A) Datos crudos obtenidos y B) Presión promedio cada 5 segundos.55

Ilustración 13-3: Densidad vs tiempo de los tres sistemas luego del ensamble npt. A) Datos crudos obtenidos y B) Densidad promedio cada 5 segundos.56

Ilustración 14-3: Evolución del RMSD de los tres sistemas estudiados durante 10 ns de simulación DM. La estructura de referencia y de estudio es esqueleto total del sistema.58

- Ilustración 15-3:** Evolución temporal del RMSD de los tres sistemas estudiados durante 10 ns de simulación. El sistema es la referencia y la estructura de estudio la proteína.59
- Ilustración 16-3:** Evolución del RMSD de los complejos proteína-ligando, con 10 ns de simulación. La HSA es la referencia y la estructura de estudio el ligando.60
- Ilustración 17-3:** Evolución temporal del Rg de la HSA en los tres sistemas estudiados durante 10 ns de simulación.61
- Ilustración 18-3:** Evolución temporal del Rg del ligando en los complejos proteína-ligando estudiados durante 10 ns de simulación.62

ÍNDICE DE ANEXOS

ANEXO A:11

ANEXO B:11

ANEXO C:12

ANEXO D:12

ANEXO E:13

ANEXO F:14

ANEXO G:15

ANEXO H:16

ÍNDICE DE ABREVIATURAS

AMD	Del inglés Advanced Micro Devices
CCA	Condiciones de Contorno Aislado
CCP	Condiciones de Contorno Periódica
CEDIA	Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia
CF	Condiciones de Frontera
CHARMM	Del inglés Chemistry at Harvard Macromolecular Mechanics
CI	Condiciones Iniciales
CPU	Del inglés Central Processing Unit
DKA	Del inglés Dekanoic Acid
DM	Dinámica Molecular
EVN	Energía, Volumen y Número de partículas
GPU	Del inglés Graphic Processing Unit
GROMACS	Del inglés GRoningen MACHine for Chemical Simulations
HPC	Del inglés High Performance Computing
HSA	Del inglés Human Seric Albumin
MM	Mecánica Molecular
NPT	Número de partículas, Presión y Temperatura
NVT	Número de partículas, Volumen y Temperatura
PDB	Del inglés Protein Data Bank
PPID	Proteína Plegada e Intrínsecamente Desordenada
RAM	Del inglés Random Access Memory
RMN	Resonancia Magnética Nuclear
RMSD	Del inglés Root Media Standard Deviation
ROP	Del inglés Return-Oriented Programming
TIP3	Del inglés Transferable Intermolecular Potential 3
TMU	Del inglés Texture Mapping Unit
VMD	Del inglés Visual Molecular Dynamics
WFR	Warfarina

RESUMEN

El objetivo del presente Proyecto de Investigación fue estudiar el sistema complejo entre la proteína albumina sérica humana (HSA) y distintos ligandos usando dinámica molecular para analizar cambios conformacionales en la interacción proteína – ligando. Se obtuvo los archivos moleculares de una base de datos y mediante un algoritmo en Python se realizó el preprocesamiento de las estructuras a través de la mutagénesis. Luego, se utilizaron softwares para visualizar el mejor modelo generado y la dinámica molecular se ejecutó con GROMACS. Los complejos fueron modelados con el campo de fuerza CHARMM36 y CGenFF. Para la preparación de la muestra se aplicaron aspectos fisiológicos como: inclusión de protones al sistema, creación de una celda unitaria, establecer las propiedades de contorno y la inclusión de moléculas de agua. Se establecieron las propiedades dinámicas del complejo interactuando con ensambles de temperatura constante y con presión constante. Para cada fase se configuró el modelado a través de archivos mdp donde se asignan parámetros de integración de tiempo y algoritmos de acoplamiento. La simulación se ejecutó durante 10 nanosegundos y al final se extrajeron los resultados mediante archivos de datos. Se obtuvieron resultados físicos que describen cada sistema. También valores del RMSD y el radio de giro que describen la dinámica, demostrando diferencias en los valores que se producen durante la interacción de la proteína con cada tipo de ligando. Se mostró la existencia de valores muy distintos en aquellas mediciones que describen la estructura y la elasticidad de la interacción proteína ligando. Se concluyó que la HSA tiene cambios estructurales dependiendo del ligando con el que interactúa esto puede aumentar o disminuir sus propiedades físicas y se recomienda hacer una extensión del análisis de los ligandos en los sitios de enlace.

Palabras Clave: <DINÁMICA MOLECULAR>, <LIGANDO>, <PROTEÍNA>, <ALBUMINA SÉRICA HUMANA [HSA]>, <ÁCIDO DECANOICO [DCA]>, <WARFARIN [WRF]>, <ENSAMBLE>





D.B.R.A.I.
Ing. Cristian Castillo

0008-DBRA-UTP-2023

SUMMARY

The aim of this research project was to study the complex system between human serum albumin (HSA) protein and different ligands using molecular dynamics to analyse conformational changes in the protein-ligand interaction. Molecular files were obtained from a database and a Python algorithm was used to preprocess the structures through mutagenesis. Then, software was used to visualise the best model generated and molecular dynamics was run with GROMACS. The complexes were modelled with the CHARMMm36 and CGenFF force field. For the sample preparation, physiological aspects were applied such as: inclusion of protons into the system, creation of a unit cell, establishing the contour properties and the inclusion of water molecules. The dynamic properties of the complex interacting with constant temperature and constant pressure assemblies were established. For each phase, modelling was configured through mdp files where time integration parameters and coupling algorithms are assigned. The simulation was run for 10 nanoseconds and at the end the results were extracted through data files. Physical results describing each system were obtained. Also, RMSD and radius of gyration values describing the dynamics, showing differences in the values that occur during the interaction of the protein with each type of ligand. Very different values were shown for those measurements describing the structure and elasticity of the protein-ligand interaction. It was concluded that HSA has structural changes depending on the ligand with which it interacts that can increase or decrease its physical properties and it is recommended to make an extension of the analysis of the ligands in the binding sites.

Keywords: <MOLECULAR DYNAMICS>, <LIGAND>, <PROTEIN>, <HUMAN SERIC ALBUMIN [HSA]>, <DECANOIC ACID [DKA]>, <WARFARIN [WRF]>, <ENSAMBLE>.



Dra. Carmita Eulalia Rojas Castro. Mgs.

0602890469

INTRODUCCIÓN

La proteína albumina sérica humana (HSA) es la proteína más abundante en la sangre y sirve de transporte para algunos ácidos grasos esenciales y fármacos que ingresan al organismo del ser humano. En la actualidad existen varios estudios sobre la caracterización de la biomolécula HSA que permiten establecer su estructura molecular, a la vez que se ha profundizado sobre su interacción con distintos ligandos de manera experimental. Con las configuraciones moleculares proteína-ligando se puede avanzar en el estudio de las interacciones a través de modelos de dinámica molecular lo que ofrece productos importantes para la generación de nuevas drogas y fármacos. Además, en la Escuela Superior Politécnica de Chimborazo en la carrera de física no se han realizado estudios similares en este campo de estudio, es decir, en dinámica molecular y las interacciones ligando-proteína.

El avance de la tecnología computacional ha permitido la existencia de nuevas herramientas más potentes que han permitido realizar estudio de dinámica molecular cada vez más complejos. Estudios como el de la tesis publicada por la biofísica Daniela Barragán (2015) donde realiza una evaluación del campo de fuerza CHARMM27 para la simulación de ADN G-Cuádruple tetramoleculares y paralelos, para lo cual utilizo un computador de altas prestaciones que le permitieron obtener 400000 estructuras. El estudio permitió profundizar sobre el comportamiento del campo de fuerza usado y ofreció resultados favorables para las simulaciones de los modelos de dinámica molecular bajo condiciones fisiológicas.

La dinámica molecular se extiende en el estudio de otras biomoléculas tales como las proteínas, y entre otros, a través de modelos de campo de fuerza desarrollados a lo largo de estos años (Brooks y et al., 1983; Jorgensen y et al., 2002; Claudia y et al., 2010; Xiao y et al., 2011; Best y et al., 2012; Huang y et al.). Esto ha permitido extender los conocimientos para campos de fuerzas con biomoléculas de todo tipo, incluso aquellas que son del tipo fármaco o ácido graso con la inclusión de un campo de fuerza generalizado (Vanommeslaegue y Mackerell, 2012). Se pueden encontrar estudios donde se analizan los mecanismos moleculares para las interacciones proteína ligando, debido a su importancia en el entendimiento de procesos biológicos (Alonso y et al., 2006; Fu y et al., 2018).

Con lo que respecta a estudios de la estructura molecular de la proteína albumina sérica humana (HSA) y su sistema de interacción con distintos ligandos tenemos publicaciones que datan de 1999 (Peters, 1985; Sugio y et al., 1999; Artali y et al., 2005), ya que esta proteína es la más abundante en el plasma sanguíneo se han elaborado varios estudios de su interacción con

diferentes tipos de ligando usando técnicas experimentales (Liu y Hungwei, 2012;). De igual manera, se tiene como antecedente una tesis doctoral en física presentado por Da Silva (2017) donde a través de modelado computacional y un estudio teórico (mecánica molecular y mecánica cuántica) investiga la influencia de las correcciones de largo alcance en la energía de interacción entre el fármaco Ibuprofeno (IBU) y la proteína de albúmina de suero humano (HSA). Siguiendo con la línea experimental también existen estudios conformacionales con otros ligandos (Bhattacharya y et al., 2000; Petitpas y et al., 2001; Liu y Hungwei, 2012; Guglielmelli y et al.,2018), y de igual manera se ha extendido su análisis a través de técnicas de dinámica molecular (Fujiwara y Amisaki, 2006; Guglielmelli y et al.,2018).

El entendimiento de la interacción de la HSA con diferentes ligandos nos permite comprender cómo actúa esta biomolécula dentro de nuestro organismo y a su vez como podemos favorecer la realización de sus funciones. La dinámica molecular para entender el comportamiento de diferentes proteínas del ser humano ya ha sido utilizada con el fin de incrementar el conocimiento sobre los sistemas biológicos que luego pueden concluir en la creación de nuevos sistemas de tratamiento o medicamentos. De igual manera existen varios mecanismos en este campo de conocimiento para entender cómo funcionan las interacciones proteína-ligando. Tener un modelo computacional que permite describir y analizar la interacción de la HSA con diferentes ligandos genera bases para la creación de futuros fármacos sobre esta proteína.

Por los motivos antes expuestos, en el presente trabajo se desarrollará un estudio a través de dinámica molecular para caracterizar la proteína HSA y su interacción con diferentes ligandos. Mediante las herramientas de dinámica molecular, se va a analizar el sistema formado por una proteína que interactúa con el ligando en un sistema acuoso y cómo dicha interacción afecta las características estructurales de la proteína.

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL

1.1. Las proteínas

Las proteínas están constituidas por sistemas macromoleculares, que a su vez están conformados por aminoácidos que se encuentran unidos por enlaces peptídicos. Es importante señalar que comúnmente cada proteína está constituida entre 50 a 2000 aminoácidos, por ello cada aminoácido está compuesta por un átomo de carbono central unido a un grupo carboxílico, un grupo amínico, un átomo de hidrógeno y por último una cadena lateral denotada por R, donde las propiedades químicas de dichas cadenas laterales designan a cada aminoácido su papel en la estructura y función (Da Silva, 2017a: p. 4).

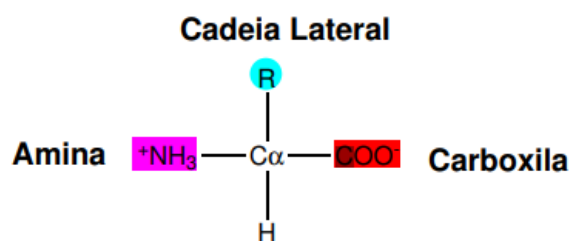


Ilustración 1-1: Esquema conformacional de un aminoácido.

Fuente: Cooper, 2017.

Son veinte los principales aminoácidos que se pueden clasificar de acuerdo con las propiedades de sus respectivas cadenas laterales, dichos aminoácidos normalmente se encuentran como residuos de proteínas, de esta forma tenemos los siguientes:

Tabla 1-1: Lista de aminoácidos esenciales y algunas de sus características.

Aminoácido	Sigla	Carga líquida de la cadena lateral	Polaridad
Alanina	GLY	Neutra	Apolar
Arginina	ARG	Positiva	
Asparagina	ASN	Neutra	Polar
Ácido Aspártico	ASP	Negativa	
Cisteína	CYS	Neutra	Apolar
Ácido Glutámico	GLU	Negativa	
Glutamina	GLN	Neutra	Polar

Glicina	GLY	Neutra	Apolar
Prolina	PRO	Neutra	Apolar
Serina	SER	Neutra	Polar
Tirosina	TYR	Neutra	Polar
Histidina	HIS	Positiva	
Isoleucina	ILE	Neutra	Apolar
Leucina	LEU	Neutra	Apolar
Lisina	LYS	Positiva	
Metionina	MET	Neutra	Apolar
Fenilalanina	PHE	Neutra	Apolar
Treonina	THR	Neutra	Polar
Triptófano	TRP	Neutra	Apolar
Valina	VAL	Neutra	Apolar

Fuente: Cooper, 2017.

Elaborado por: Chávez, I., 2022.

1.1.1. Interacciones proteína-ligando

Estas interacciones son fundamentales para casi todos los procesos que ocurren en los organismos vivos. La transmisión de señales mediada por ligandos a través de la complementariedad molecular es esencial para todos los procesos vitales, por ello estas interacciones químicas comprenden el reconocimiento biológico a nivel molecular. La evolución de la función de la proteína depende en parte del desarrollo de sitios altamente específicos diseñados para unirse a ligandos de moléculas pequeñas con afinidades adaptadas a las necesidades de la célula. La cooperación en la unión de ligandos es de vital importancia para la regulación de funciones biológicas competitivas. La regulación de procesos celulares a través de interacciones cooperativas proteína-ligando ocurre a través de mecanismos moleculares que involucran transiciones conformacionales de proteínas entre estados de baja y alta afinidad. Como consecuencia, Las interacciones de unión a ligandos se utilizan para cambiar las proteínas entre estados de diferente función (Fu y et al., 2018, p. 2).

1.1.2. Acoplamiento del ligando

Las estructuras tanto del ligando como de la proteína son necesarias para la aplicación de técnicas de acoplamiento. Si bien la variedad de estructuras conformacionales de las moléculas pequeñas puede ser relativamente fácil de predecir, la conformación de energía más baja obtenida puede no corresponder a la del ligando unido, por ello las estructuras de las proteínas presentan un desafío mayor. Muchas proteínas no tienen una estructura determinada experimentalmente y, por lo tanto, los estudios de acoplamiento no se pueden

realizar directamente. En algunos casos, se pueden usar técnicas computacionales para predecir la estructura tridimensional de una proteína siempre que la estructura de un homólogo de proteína este estrechamente relacionado.

Cabe añadir que es de gran importancia preparar cuidadosamente la estructura de la proteína antes del proceso de acoplamiento. Si bien las estructuras de la proteína unida a ligando pueden proporcionar los mayores enriquecimientos, los resultados finales pueden estar sesgados hacia tipos particulares de ligandos. Además, las técnicas de enhebrado de secuencias se pueden usar para generar modelos de estructuras de proteínas que, aunque no son tan buenas como las estructuras determinadas experimentalmente, se pueden usar como objetivos de acoplamiento (Alonso y et al., 2006, p. 536).

1.1.3. Sitio de enlace

El sitio de enlace es una posición donde la proteína se une a una molécula entrante que es comparativamente más pequeña en tamaño, llamada ligando. En las proteínas, los sitios de enlace son espacios en la estructura terciaria donde los ligandos se unen mediante fuerzas débiles (enlace no covalente). Solo unos pocos residuos participan realmente en la unión del ligando, mientras que los otros residuos de la proteína actúan como marco para proporcionar la conformación y la orientación correctas.

El sitio de enlace es un lugar de especificidad y afinidad química en la proteína que forma enlaces químicos con otras moléculas y ligandos, dicha afinidad ocasiona una fuerza de atracción química entre ellas. Como tal, puede haber competencia entre diferentes ligandos por el mismo sitio de enlace, y la reacción química resulta en un estado de equilibrio entre ligandos enlazantes y no enlazantes. Es importante señalar también que la saturación del sitio de enlace estará conformada por el número total de sitios de unión que están ocupados por ligandos por unidad de tiempo, sin embargo, los ligandos de alta afinidad tienen una fuerza intermolecular alta, pueden residir en el sitio de unión por más tiempo ya que tiene una concentración baja y hacen que los receptores cambien (Da Silva, 2017b: p. 8).

1.2. La proteína albumina sérica humana (HSA)

Es una proteína sintetizada por el hígado que se compone de 585 aminoácidos, tiene un alto contenido de cisteína y una alta proporción de aminoácidos cargados negativamente pero no de carbohidratos. El HSA se asocia con una amplia variedad de sustancias que incluyen hormonas y fármacos que poseen una región hidrófoba y, por lo tanto, la asociación de

hormona tiroidea con HSA puede considerarse inespecífica. El HSA constituye más de la mitad del contenido total de proteínas en suero y, por lo tanto, es el principal contribuyente al mantenimiento de la presión osmótica coloidal (Da Silva, 2017c: p. 87).

1.2.1. Estructura

La HSA es una proteína globular en forma de corazón con una serie repetitiva de seis subdominios helicoidales y se compone de 67 % de hélices α , 10 % de giros, 23 % de bobinas aleatorias y ninguna lámina β . Las estructuras de cristalografía de rayos X de alta resolución muestran tres dominios predominantes en la albúmina. Por lo general, se numeran como dominio I (1–195 aa), dominio II (196–383 aa) y dominio III (384–585 aa). Cada dominio se subdivide en dos subdominios, A y B. Existe un notable grado de secuencia y similitud estructural, así como una similitud de distribución de carga superficial entre HSA y sus contrapartes de bovino (BSA), equino (ESA), leporina (LSA), y fuentes caninas (CSA). A nivel de secuencia, HSA y BSA comparten un 76 % de identidad, mientras que, en general, las albúminas séricas de diversas fuentes comparten más del 62 % de identidad de secuencia. El análisis de dinámica molecular de HSA, BSA y CSA indica que el movimiento de los dominios I y III son clave para definir las propiedades de las albúminas. La BSA es estructuralmente algo más rígida que la HSA, mientras que la CSA es más flexible y posee sitios de unión al fármaco más grandes y accesibles al agua (Sugio y et al., 1999, p. 3).

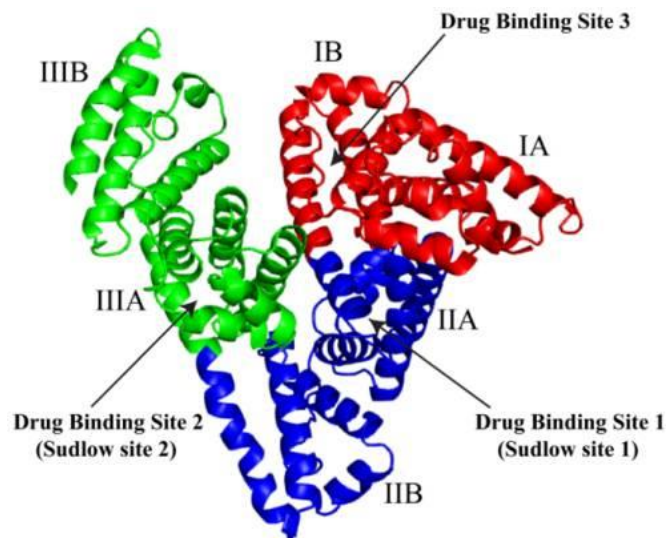


Ilustración 2-1: Organización de dominios de la albúmina sérica humana.

Dominio I (rojo), dominio II (azul) y dominio III (verde).

Fuente: Mishra y Heath, 2021.

1.2.2. Sitios de Enlace

Dos sitios de enlace importantes en la albúmina sérica humana son los sitios Sudlow I en el subdominio IIA y el Sudlow II ubicado IIIA. El sitio Sudlow I tiene una afinidad de unión preferencial por compuestos heterocíclicos voluminosos como azapropazona, fenilbutazona y Warfarina. El sitio II preferentemente se une a compuestos aromáticos como el ibuprofeno. La HSA también tiene siete sitios de enlace a ácidos grasos de cadena larga en sus tres dominios, los primeros cinco sitios de unión de ácidos grasos parecen tener los residuos de aminoácidos para facilitar las interacciones polares con la cabeza de carboxilato de una cadena de ácidos grasos. Cada sitio tiene una afinidad diferente por los ácidos grasos, por ello los sitios de unión de ácidos grasos se unen al resto carboxilato de ácidos grasos con interacciones electrostáticas/polares (Artali y et al., 2005, p. 486).

1.2.3. El ligando ácido graso

Los ácidos grasos son los componentes básicos de la grasa en nuestro cuerpo y en los alimentos que comemos. Durante la digestión, el cuerpo descompone las grasas en ácidos grasos, que luego pueden ser absorbidos por la sangre. El ácido graso es componente importante de los lípidos (componentes liposolubles de las células vivas) en plantas, animales y microorganismos, generalmente consta de una cadena lineal de un número par de átomos de carbono, con átomos de hidrógeno a lo largo de la cadena y en un extremo de la cadena y un grupo carboxilo (-COOH) en el otro extremo, es ese grupo carboxilo el que lo convierte en un ácido (ácido carboxílico). Además, si los ácidos grasos contienen uno o más dobles enlaces carbono-carbono se denominan insaturados, en cambio cuando no poseen estos enlaces entre carbonos son ácidos grasos saturados, y cuando contienen un solo doble enlace en la cadena alifática, se denominan monoinsaturados (LUMITOS AG, 2008, párr. 12).

1.2.3.1. Cambios conformacionales del HSA en presencia de ácidos grasos

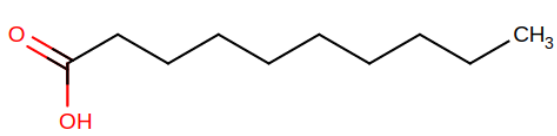
La HSA es la proteína más abundante en el plasma humano con concentraciones varias típicas. Además se une a ligandos como ácidos grasos, glucosa, fármacos, exceso de cobre y hormonas, lo que ayuda también en su transporte, distribución y metabolismo, pero también es una proteína alostérica que puede sufrir cambios conformacionales en respuesta a muchos factores, incluido el cambio de pH y la unión de ligandos, entre ellos la adición de ácidos grasos saturados de cadena larga como el ácido octanoico o el ácido palmítico (PA) a la HSA que puede introducir cambios en la conformación de las proteínas, estos ácidos se unen a la HSA en al menos siete sitios de unión diferentes que involucran los seis subdominios de la

proteína, incluidos los sitios I y II de Sudlow, con una afinidad particularmente alta por este último. Los cambios alostéricos inducidos por la unión de ácidos grasos alteran la capacidad de HSA para transportar otros metabolitos y fármacos, incluido la hormona tiroxina, la insulina y muchos fármacos de sulfonilurea (Fujiwara y Amisaki, 2006a: p. 731).

1.2.3.2. Estructura del ácido cáprico

El ácido cáprico o decanoico (DKA) es un sólido que pertenece a los ácidos grasos de cadena lineal, además son ácidos grasos de cadena alifática lineal. Las proteínas a las que se dirige el ácido cáprico incluyen furina, octanoiltransferasa, 3-oxoacil-[acil-carrier-protein] sintasa 1, proteína de unión a albúmina peptoestreptocócica y proteína putativa no caracterizada tcp14. A continuación se presentará un resumen de los componentes químicos (Wishart y et al., 2022, párr. 1-5).

Tabla 2-1: Características del ligando ácido decanoico DKA.

Nombre	ÁCIDO DECANÓICO
Identificadores	ácido decanoico
Formula	C10 H20 O2
Peso molecular	172.26
Tipo	NON-POLYMER
Isoméricas SMILES	CCCCCCCCC(=O)O
InChI	InChI=1S/C10H20O2/c1-2-3-4-5-6-7-8-9-10(11)12/h2-9H2,1H3, (H,11,12)
InChIKey	GHVNFZFCNZKVNT-UHFFFAOYSA-N
	

Fuente: Wishart y et al., 2022.

Elaborado por: Chávez, I., 2022.

1.2.3.3. Interacción del ácido cáprico con la HSA

En la HSA existen varios sitios de enlace donde puede interactuar con ácidos grasos de distinto tipo. Entre ellos tenemos ácidos grasos de cadena media donde se incluye el ácido cáprico o también llamado DKA, por su nombre en inglés, *decanoic acid*.

En un estudio cristalográfico presentado por Bhattacharya y et al. (2000, pp. 729-730) donde se analizan complejos HSA con ácidos grasos saturados de cadena media y cadena larga, se

identificó un total de siete sitios de unión ocupados por todos los ácidos grasos de cadena media y cadena larga. Adicional a estos, se encontró que los ácidos grasos de cadena media se unen en sitios adicionales de la proteína, lo que produce un total de 11 ubicaciones de unión distintas. Se observa que los dos principales sitios de unión de fármacos, en los subdominios IIA y IIIA, están ocupados por ácidos grasos y uno de ellos (en IIIA) parece coincidir con un sitio de unión de ácidos grasos de cadena larga de alta afinidad.

1.2.4. El ligando Warfarina

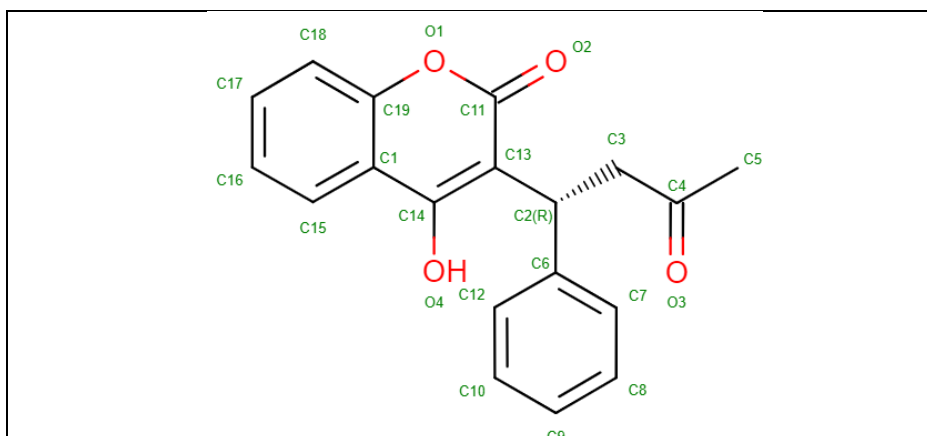
La Warfarina es un anticoagulante cumarínico que se utiliza en todo el mundo para el tratamiento y la prevención de la enfermedad tromboembólica. Sin embargo, la terapia con Warfarina puede ser difícil de manejar debido al estrecho índice terapéutico del fármaco y la amplia variabilidad interindividual en la respuesta del paciente. Este fármaco anticoagulante se sintetizó por primera vez en 1948 en el laboratorio de Karl Link en Wisconsin como parte de un programa para desarrollar rodenticidas más potentes basados en la estructura del dicumarol, gracias a ello se introdujo en la práctica clínica en la década de 1950 y, para el nuevo milenio, era uno de los veinte principales medicamentos y el más utilizado en todo el mundo para el tratamiento oral y la prevención de eventos tromboembólicos, incluida la fibrilación auricular. y trombosis venosa profunda (Rettie y Guying, 2006, pp. 223-224).

1.2.4.1. Estructura conformacional

La Warfarina consiste en una mezcla racémica de dos enantiómeros activos, formas R y S, cada uno de los cuales se elimina por vías diferentes. La Warfarina S es de dos a cinco veces más potente que el isómero R para producir una respuesta anticoagulante.

Tabla 3-1: Características del ligando Warfarina (WFR)

Nombre	R-WARFARINA
Identificadores	4-hidroxi-3-[(1R)-3-oxo-1-fenil-butil]cromen-2-ona
Fórmula	C 19 H 16 O 4
Peso molecular	308.33
Tipo	NO POLÍMERO
Isoméricas SMILES	CC(=O) C[C@H](c1ccccc1) C2=C(c3ccccc3OC2=O) O
InChI	InChI=1S/C19H16O4/c1-12(20)11-15(13-7-3-2-4-8-13)17-18(21)14-9-5-6-10-16(14) 23-19(17)22/h2-10,15,21H,11H2,1H3/t15-m/s1
InChIKey	PJVWTKQMONHTI-OAHLLOKOSA-N



Fuente: Feunang y et al., 2020

Elaborado por: Chávez, I., 2022.

1.2.4.2. Interacción de la Warfarina con la HSA

La Warfarina es un fármaco anticoagulante ampliamente utilizado en el tratamiento de la trombosis venosa y la embolia pulmonar y se transporta en la sangre casi exclusivamente a través de la albúmina sérica humana. Los efectos de la unión de la Warfarina al estado nativo de la albúmina se caracterizaron por absorción UV-vis, fluorescencia convencional y síncrona, calorimetría de titulación isotérmica, calorimetría diferencial de barrido y simulación de dinámica molecular. Los resultados generales indican que, en condiciones fisiológicas, la unión de Warfarina en el sitio DS1 de la albúmina promueve la estabilización local con efectos resultantes en la dinámica global de proteínas.

El aumento de la estabilidad de la proteína tiene un carácter tanto entálpico como entrópico. En condiciones desnaturizantes, el efecto estabilizador de la Warfarina se evidencia por un aumento tanto de la temperatura de fusión como de la entalpía de desdoblamiento de la albúmina con la relación molar fármaco/proteína. Más importante aún, la resistencia térmica aumenta debido al efecto selectivo en el lóbulo de la proteína específica que incluye el principal sitio de unión del fármaco.

La comparación del comportamiento térmico del complejo proteína-Warfarina con el de la presencia de un ligando típico del otro sitio principal de unión a la proteína, es decir, el sitio del fármaco DS2, proporciona información clave sobre los efectos de estabilización específicos del dominio sobre la albúmina. El efecto estabilizador de la Warfarina se evidencia por un aumento tanto de la temperatura de fusión como de la entalpía de despliegue de la albúmina con la relación molar fármaco/proteína. Más importante aún, la resistencia térmica aumenta debido al efecto selectivo en el lóbulo de la proteína específica que incluye el principal sitio de unión del fármaco (Petitpas y et al., 2001, p. 22805).

1.3. Dinámica molecular

La dinámica molecular (DM) es una técnica de simulación por computadora en que átomos y moléculas interactúan por un periodo de tiempo, permitiendo una visualización del movimiento de las partículas. Esta técnica se concibió dentro de la física teórica, y ahora se usa ampliamente en el campo de la biofísica y la ciencia de materiales. Su campo de aplicación va desde superficies catalíticas hasta sistemas biológicos como las proteínas.

La simulación molecular se basa en usar métodos computacionales para modelar el comportamiento de las moléculas con la finalidad de realizar predicciones sobre estas interacciones. Los algoritmos de los métodos usados se basan en leyes físicas, químicas y bioquímicas de sistemas simples y complejos. Los sistemas simulados pueden utilizar tanto métodos clásicos como cuánticos, dependiendo de la complejidad, lo que desemboca en resultados con información sobre la estructura espacial, la dinámica y termodinámica del sistema molecular estudiado.

Actualmente el estudio de la interacción proteína-ligando tiene un interés investigativo para predecir y describir el funcionamiento de sistemas biomoleculares que regulan algunos procesos biológicos en el organismo del ser humano. La simulación molecular es una herramienta que ha permitido mejorar estos estudios mediante la implementación computacional de modelos físicos que describen las interacciones intermoleculares, además, el uso de herramientas computacionales permite realizar cálculos complejos que mejoran el conocimiento y posterior análisis de estos sistemas.

Los métodos computacionales son ventajosos sobre los métodos experimentales ya que permiten controlar rigurosamente parámetros y variables de entrada del sistema tales como: el número de partículas, la temperatura, la presión, volumen y otros; los experimentos producidos pueden ser visualizados y analizados a través de una interfaz gráfica (Barragán, 2015a: p. 5).

Las primeras simulaciones DM de gases simples se realizaron a fines de la década de 1950 por Alder y Wainwright, mientras que la primera simulación DM de una proteína se realizó a fines de la década de 1970 por el científico Andrew McCammon. Sin embargo, las simulaciones DM se han vuelto sustancialmente más populares y visibles en los últimos años, particularmente desde la perspectiva de los biólogos moleculares experimentales (Hollingsworth y Dror, 2018: p. 1129).

La simulación a través de DM puede calcular diferentes propiedades fisicoquímicas el sistema como la energía libre, entropía, solubilidad, viscosidad, presión, temperaturas de cambio de fase, y en sistemas biológicos, permite medir la fuerza de interacción entre posibles fármacos y sus dianas biomoleculares o receptores, e incluso, describir el comportamiento de una proteína y moléculas complejas bajo ciertas condiciones. Con este método se genera las trayectorias de las N partículas que conforman el sistema utilizando integración numérica directa de las ecuaciones de Newton, especificando adecuadamente el potencial de interacción interatómico, las condiciones iniciales y las de frontera (Lozano-Aponte y Scior, 2014: p. 86-87).

1.3.1. Principios básicos

Supongamos que tenemos un sistema configurado por N partículas donde su posición está determinado por un vector \vec{r} con componentes en $\hat{x}, \hat{y}, \hat{z}$. Las coordenadas de las partículas están dadas por \vec{r}_i tal que $i = 1, 2, \dots, N$. Con esta notación podemos describir la fuerza que interactúa sobre cada partícula usando las leyes de Newton:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\frac{\partial U}{\partial \vec{r}_i} \quad (1)$$

m_i es la masa de cada partícula.

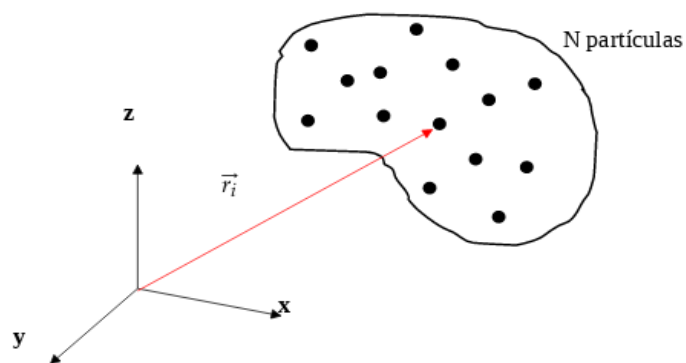


Ilustración 3-1: Ilustración de un sistema en una simulación de dinámica molecular.

Realizado por: Chávez, I. 2022.

La integración de la ecuación de Newton resultará en la trayectoria de cada partícula donde se obtiene las posiciones y velocidades en función del tiempo. Para estas operaciones existen

varios métodos que simplifican su resolución que podemos utilizar. Uno de estos métodos que simplifican las operaciones es el algoritmo de Verlet, computacionalmente se aplica sobre las posiciones y aceleraciones en un tiempo determinado t , también se toma en cuenta los mismos datos, pero en un paso de tiempo anterior es decir con $t + \Delta t$, estos datos sirven para calcular las nuevas posiciones $t + \Delta t$. Las velocidades no aparecen explícitamente en el algoritmo de Verlet. Se puede realizar una aproximación al dividir la diferencia en las posiciones.

$$v(t) = [r(t + \delta t) - r(t - \delta t)] / 2 \delta t \quad (2)$$

Para poder usar las ecuaciones clásicas de movimiento se debe mencionar que se realiza dos aproximaciones. La primera es la aproximación de Born-Oppenheimer el cual asume que el estado electrónico se acopla adiabáticamente al movimiento de los núcleos. La segunda es que el movimiento del núcleo está muy alejado del límite inferior de la incertidumbre de Heisenberg: $\Delta E \Delta t \gg \frac{\hbar}{2}$ (Li, 2005a: p. 568).

1.3.1.1. Integrando las ecuaciones de movimiento

Integrar las ecuaciones de movimiento en un sistema dinámico es una tarea de matemática aplicada basado en los principios físicos que rigen el movimiento de las partículas. Para empezar, debemos partir de las ecuaciones clásicas del movimiento, las cuales tienen la siguiente forma

$$\dot{q} = G(p, q), \dot{p} = F(p, q) \quad (3)$$

Donde $G(p, q) = \frac{dH}{dp}$ y $F(p, q) = -\frac{dH}{dq}$, y H es el Hamiltoniano, el cual está dado por

$$H = \sum_i \frac{p_i^2}{2m_i} + U(q) \quad (4)$$

Donde q representa las coordenadas del átomo i y p es el momento conjugado. Una receta fácil para implementar las ecuaciones de movimiento a un sistema de dinámica molecular es usar un método que se conoce como “salto generalizado” (*generalized leap-frog*). El método *leap-frog* consiste en avanzar el momento en la mitad de un paso de tiempo, luego, con el momento calculado también movemos las coordenadas hacia adelante un paso de tiempo

completo, recalculamos las fuerzas en las nuevas posiciones y, con estas nuevas fuerzas, avanzamos el momento un paso de tiempo completo (Hernandez, 2008, p.97).

El método de salto generalizado puede representarse simbólicamente de la siguiente manera

$$\begin{aligned}
 p\left(t + \frac{\Delta t}{2}\right) &= p(t) + \Delta t F\left[p\left(t + \frac{\Delta t}{2}\right), q(t)\right]/2 \\
 q(t + \Delta t) &= q(t) + \Delta t \{G\left[p\left(t + \frac{\Delta t}{2}\right), q(t)\right] + G\left[p\left(t + \frac{\Delta t}{2}\right), q(t + \Delta t)\right]\}/2 \\
 p(t + \Delta t) &= p\left(t + \frac{\Delta t}{2}\right) + \Delta t F\left[p\left(t + \frac{\Delta t}{2}\right), q(t + \Delta t)\right]
 \end{aligned} \quad (5)$$

Donde Δt es un paso de tiempo. Para el caso simple donde el Hamiltoniano es separable tal que partiendo de la ecuación (4) podemos reducir las ecuaciones anteriores para el paso generalizado de la siguiente manera

$$\begin{aligned}
 p\left(t + \frac{\Delta t}{2}\right) &= p(t) - \frac{\Delta t}{2} \frac{\partial U}{\partial q}(t) \\
 q(t + \Delta t) &= q(t) + \Delta t p\left(t + \frac{\Delta t}{2}\right) \\
 p(t + \Delta t) &= p\left(t + \frac{\Delta t}{2}\right) - \frac{\Delta t}{2} \frac{\partial U}{\partial q}(t + \Delta t)
 \end{aligned} \quad (6)$$

Estas ecuaciones se conocen simplemente como el algoritmo de salto. Usando las Ecs. (6) repetidamente, se puede trazar una trayectoria a partir de condiciones iniciales especificadas. Siempre que el paso de tiempo Δt sea lo suficientemente pequeño, este esquema conserva la energía con una precisión razonable y es reversible en el tiempo (Hernandez, 2008a: p.100-101).

1.3.1.2. Condiciones iniciales (CI)

Las ecuaciones de movimiento de Newton son ecuaciones diferenciales ordinarias de segundo orden y las condiciones iniciales nos brinda información sobre las posiciones y velocidades iniciales de cada partícula, $r^{3N}(t = 0)$ y $\dot{r}^{3N}(t = 0)$.

Generalmente las condiciones iniciales de los sólidos cristalinos son fáciles de establecer, sin embargo, para sistemas de interacción intermolecular se requiere más trabajo. Si se requiere obtener una configuración amorfa, una estrategia es extinguir el líquido durante la ejecución de la DM. Antes de ingresar las CI se debe visualizar el sistema, comprobando las

longitudes de enlace, los números de coordinación, etc. Si no tenemos cuidado podemos incursionar en fallas de la simulación debido a que algunos átomos pueden estar inicialmente demasiado cerca lo que genera fuerzas enormes (Li, 2005b: p. 566).

1.3.1.3. Condiciones de frontera periódicas

En la sección anterior se mencionó que una simulación DM trabaja con la integración de las ecuaciones de movimiento de Newton, sin embargo, un paso importante para la realización de estos cálculos es realizar una especificación para un potencial interatómico, con cierta condición inicial (CI) y condición de frontera (CF). Considere, por ejemplo, un sistema con N átomos en un volumen Ω . Se puede definir su energía interna como $E \equiv K + U$ donde K es la energía cinética y U es la energía potencial, para cada uno de los átomos, se definen como

$$K \equiv \sum_{i=1}^N \frac{1}{2} m_i |\dot{r}_i(t)|^2 \quad (7)$$

$$U = U(r^{3N}(t)) \quad (8)$$

Si el sistema está aislado entonces debemos notar que E es una cantidad que se mantiene constante durante todo el tiempo.

Respecto a las condiciones de frontera existen dos tipos importantes: condición de contorno aislada (CCA) y condición de contorno periódica (CCP). El CCA es ideal para estudiar grupos y moléculas, donde el sistema de partículas N está rodeado de vacío; estas partículas interactúan entre sí, pero se supone que están tan lejos de todo lo demás en el universo que no se producen interacciones con el exterior. Por otro lado, en el CCP se realiza un seguimiento explícito del movimiento de N partículas en la llamada súper celda, rodeada de imágenes periódicas infinitamente replicadas de sí misma. Por lo tanto, una partícula puede interactuar con partículas de esta súper celda y con partículas en súper celdas adyacentes (Li, 2005c: p.567).

Esto quiere decir que las condiciones de frontera periódica asumen que una la caja de simulación (i.e. la caja que contiene los átomos de la simulación) está rodeada por copias idénticas de ella misma en todas las direcciones. En las simulaciones de sistemas periódicos esta caja que contiene las características que se replican es llamada “súper celda” o celda unitaria. Adoptar una aproximación con CCP implica asumir que las imágenes periódicas de los átomos en la caja de simulación se mueven de la misma manera. (Hernández, 2008b: p. 102)

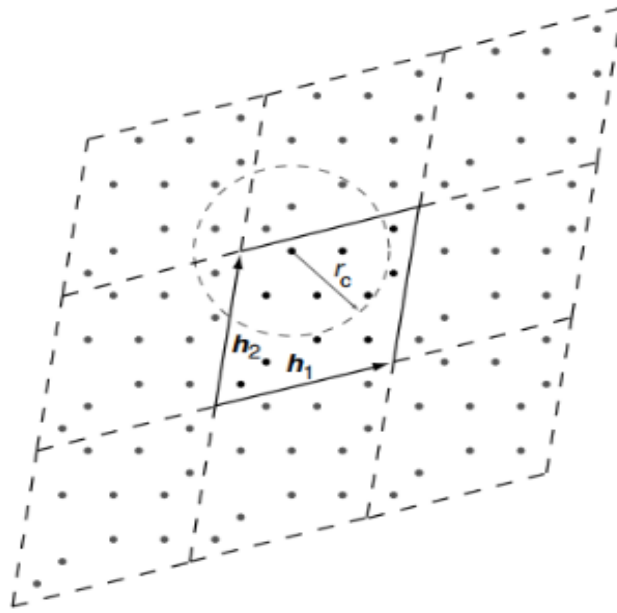


Ilustración 4-1: Condiciones de contorno periódicas (CCP) con seguimiento a los átomos de la “super celda”, que se replica infinitamente en las tres direcciones.

Realizado por: Li J. 2005.

1.3.2. Integrador/ ensamble

De la ecuación (1) se obtiene un conjunto de ecuaciones diferenciales ordinarias de segundo orden las cuales pueden ser no lineales. Existen varios métodos que permiten resolver estos problemas, sin embargo, son raramente utilizados en la práctica, ya que la existencia del Hamiltoniano permite realizar algoritmos de integración más precisos, entre los que destacan la familia de integradores predictores-correctores y la familia de integradores simplécticos (Li, 2005d: p.568-569).

Los ensambles son conceptos de física estadística que se refiere a la distribución de las condiciones iniciales. Cuando un sistema está configurado con un cierto ensamblaje luego sigue estrictamente una evolución determinada por la ecuación Hamiltoniana del movimiento con una energía que se conserva.

La teoría de ensambles se puede conectar a la dinámica molecular con la termodinámica y así obtener propiedades de interés. Esta teoría se basa principalmente en la generación de los microestados accesibles por el sistema. Existen dos tipos de ensambles que se pueden analizar los cuales son canónico y micro canónico, que a su vez pueden tener condiciones de

presión o temperatura constante, es decir, isobáricos o isotérmicos respectivamente (Barragán, 2015b: p.10-11).

En un método estándar de dinámica molecular las ecuaciones newtonianas del movimiento se fijan al volumen de la celda y la energía total es conservada, formando un ensamble micro canónico o también llamado ensamble EVN (Energía, volumen y numero de partículas). Con este método las condiciones de la simulación no se asemejan a las características que ocurren en un experimento real.

El físico japonés Suichi Nosé (1983, p. 265-266) propuso un método para simular una dinámica molecular configurando un ensamble con temperatura y volumen constante (NVT), y un segundo ensamble con la temperatura y la presión constante (NPT). El presente método devuelve una distribución canónica rigurosa para el momento y las coordenadas espaciales. Además, en el segundo ensamble se considera un volumen variable lo que le permite producir condiciones externas requeridas en los cambios estructurales que suceden durante la simulación.

1.3.2.1. Ensamble NVT

El sistema físico que se estudia en dinámica molecular está compuesto de N número de partículas o átomos, cada uno con sus propias coordenadas r_i , una masa m_i y energía potencial $U(r_i)$. A través de la implementación del lagrangiana se puede obtener una expresión estándar para el movimiento (ecuación 1) aplicado a cada uno de los átomos del sistema.

$$m_i \ddot{r}_i = \frac{-\partial U}{\partial r_i} \quad (9)$$

La idea de este método es suponer que nuestro sistema de interés está acoplado por un baño de calor controlando la temperatura, luego el sistema se puede describir mediante variables virtuales, coordenadas r_i , momentos p_i , el tiempo y un parámetro de acoplamiento s (Klein, 1986a: p. 458). El parámetro s actúa como un sistema externo lo que adiciona grados de libertad al sistema interno estudiado representado por una interacción $v_i = s\dot{r}_i$ la cual se puede interpretar como un cambio de calor entre los dos sistemas. De esta manera, la energía potencial queda expresada mediante $U_s = (f + 1)k_B T_{eq} \ln s$, donde f es el número de grados de libertad, k_B la constante de Boltzmann y T_{eq} un valor fijo de temperatura externa (Nosé,

1983b: 256-257). El lagrangiano para un sistema extendido de partículas y el externo será entonces

$$L_N = \sum_i \frac{m_i s^2 \dot{q}_i^2}{2} - U(q_i) + \frac{Q \dot{s}^2}{2} - U_s \quad (10)$$

La variable Q es introducida para construir una ecuación dinámica para s se puede interpretar como la “masa” para la variable s, pero tiene unidades de (energía * t²). A partir de la ecuación 10 se pueden resolver las ecuaciones de movimiento de Euler-Lagrange para obtener las expresiones de los momentos conjugados.

$$p_i = \frac{\partial L_N}{\partial \dot{r}_i} = m_i s^2 \dot{r}_i \quad (11)$$

$$p_s = \frac{\partial L_N}{\partial \dot{s}} = Q \dot{s} \quad (12)$$

Con las expresiones anteriores, Nosé (1983c: 257) postula el hamiltoniano que describe el sistema extendido con las condiciones que se han revisado anteriormente

$$H_N = \sum_i \frac{p_i^2}{2m_i s^2} + U(r_i) + \frac{\dot{p}_s^2}{2Q} + (f + 1)k_B T_{eq} \ln s \quad (13)$$

a su vez, define expresiones para el momento total

$$P = \sum_i p_i = \sum_i m_i s^2 \dot{r}_i \quad (14)$$

y el momento angular total del sistema,

$$M = \sum_i r_i \times p_i = \sum_i r_i \times m_i s^2 \dot{r}_i \quad (15)$$

El hamiltoniano de la ecuación 13 nos permite integrar el espacio de fase bajo la variable s para obtener una función de distribución para el sistema físico, particularmente, escogiendo la energía U_i asegura una partición de la función lo que deriva en una expresión compacta para función de distribución del sistema extendido (Nosé, 1983d: p.258). Luego, mediante hipótesis ergódica se relaciona el promedio del tiempo a lo largo de una trayectoria por el

promedio del ensamble. El promedio de cualquier cantidad estática expresada como una función de p_i/s y r_i a lo largo de una trayectoria determinada por las ecuaciones de movimiento (11) y (12) es exactamente un ensamble canónico (Klein, 1986b: p. 460). Por lo que, las ecuaciones definidas en esta sección para un ensamble NVT proveen un esquema para llevar a cabo cálculos de MD con temperatura constante.

1.3.2.2 Ensamble NPT

En el trabajo presentado por Nosé (1983e: 266) realiza una extensión de su método para simulación de DM para el caso en el que la temperatura y presión son constantes, para lo cual, considera una celda cúbica con aristas de longitud $L = V^{1/3}$ entonces sus coordenadas pueden ser representadas mediante $r_i = V^{1/3}x_i$, donde x_i es la escala de la coordenada con valores entre 0 y 1. Con estas consideraciones se postula un nuevo lagrangiano para este sistema extendido

$$L_N = \sum_i \frac{m_i s^2 V^{2/3} r_i^2}{2} - U(Lr_i) + \frac{Q\dot{s}^2}{2} - U_s + \frac{W\dot{V}^2}{2} - P_{ex}V \quad (16)$$

La configuración descrita por el lagrangiano en la ecuación 16 se puede imaginar como un sistema con un pistón acoplado capaz de dilatar de manera homogénea el sistema ajustando una presión P_{ex} que no se compensa por la presión dentro del sistema. Por lo tanto, se debe añadir una variable W para representar la energía cinética generada por la “masa” del pistón (Klein, 1986c: p. 460). Luego, se pueden resolver las ecuaciones de movimiento de Euler-Lagrange para obtener las expresiones de los momentos conjugados

$$p_i = \frac{\partial L_N}{\partial \dot{r}_i} = m_i L^2 \dot{r}_i \quad (17)$$

$$p_V = \frac{\partial L_N}{\partial \dot{s}} = W \dot{L} \quad (18)$$

A partir de estas ecuaciones se puede obtener el hamiltoniano

$$H_N = \sum_i \frac{p_i^2}{2m_i L^2} + U(Lr_i) + \frac{p_V^2}{2W} + PV \quad (19)$$

Y las ecuaciones dinámicas del sistema se definen como

$$r_i = -\left(\frac{\partial U}{\partial r_i}\right) \frac{1}{m_i L^2} - \left(\frac{2\dot{V}}{3V}\right) r_i \quad (20)$$

$$3VW\ddot{V} = \sum_i m_i L^2 r_i^2 - \sum_i r_i \left(\frac{\partial U}{\partial r_i}\right) \frac{1}{m_i L^2} - 3PV \quad (21)$$

Se demostró que usando los espacios de fase de las trayectorias que describen las ecuaciones (20) y (21) corresponden a un ensamble (N, P, H), es decir, que la entalpía del sistema extendido propuesto es conservada (Andersen, 1980; citado en Klein, 1986d: p. 461).

Andersen asevera que el método descrito en este apartado junto al ensamble NVT tienen potenciales usos en las simulaciones de DM en diferentes situaciones:

(I) Cuando es simulado solamente para un número pequeño de estados termodinámicos los cuales pueden ser usados para asegurar que los valores deseados de temperatura y/o presión sean archivados en cada simulación. (II) Estos métodos permiten las fluctuaciones de energía y densidad facilitando el estudio de procesos como la nucleación y separación de fase, en los cuales interviene una fluctuación considerable. (III) Permite la investigación de cómo un sistema responde ante rangos finitos de calefacción, enfriamiento, compresión y expansión. (IV) Los métodos permiten la simulación de procesos no equilibrados tales como reacciones químicas y cambios de fase que liberan una gran cantidad de energía. (Andersen, 1980, p. 2392-2393)

1.3.1. Topología molecular

La topología molecular es la estructura molecular desde un punto de vista físico y químico, donde se pueden entender la geometría y simetría del sistema de DM que se estudia. En una simulación esta información se almacena en un archivo de topología la cual contiene información necesaria para la simulación. Ciertas geometrías son compatibles con una topología dada, mientras que otras están “prohibidas” por razones topológicas.

Gutman y Polansky en su trabajo nos presenta una definición simple y práctica de topología molecular:

Suponga que una molécula M está compuesta por n átomos $A_1, A_2, A_3, \dots, A_n$, además para cualquier par de átomos A_i y A_j cumplen con las siguientes condiciones: (I) En la molécula M existe un enlace químico entre estos dos átomos y

(II) que estos dos átomos no sean químicamente enlazantes. (Gutman y Polansky, 1986a: p.12)

La topología molecular tiene conceptos análogos a la topología matemática, mediante sus definiciones podemos encontrar moléculas que son isotopológicas, es decir, moléculas que tienen el mismo valor topológico. En la Ilustración 3-1 podemos observar la representación química de enlace de varios compuestos moleculares diferentes pero que desde el punto de visto topológico son indistinguibles.

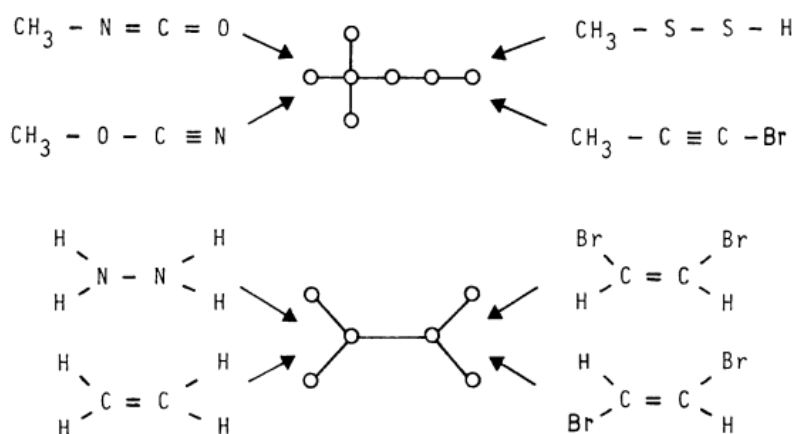


Ilustración 5-1: Ejemplos de moléculas que son isotopológicas.

Realizado por: Gutman, I. y Polansky, O., 1986.

En física, implementar este concepto resulta útil puesto que brinda la facultad de generar una misma medida física P para todas las moléculas que comparten un mismo valor topológico T , si dicha propiedad depende de la topología, es decir, si $P = f(T, a, b, c, \dots)$ donde a, b, c , son parámetros teóricos. Además, permite tener un marco de referencia para escoger las geometrías compatibles a una topología, y de igual manera, se puede determinar la simetría de la estructura molecular (Gutman y Polansky, 1986b: p.13-14).

1.3.1.1. Campo de fuerza

Los métodos de campo de fuerza, también conocidos como métodos de mecánica molecular (MM), se basan en ignorar las interacciones eléctricas de las moléculas para reducir los grados de libertad del sistema y solamente se calcula el movimiento de los núcleos.

La suposición fundamental en un campo de fuerza se realiza mediante la aproximación de Born-Oppenheimer, el cual permite separar los movimientos eléctricos de lo nucleares para poder escribir la energía del sistema solamente en función de las coordenadas de los núcleos.

Para completar la teoría se debe considerar que los campos de fuerzas satisfacen con las condiciones de aditividad y transferibilidad, es decir, que la energía del sistema puede ser escrita como la suma de sus potenciales y que la función de energía potencial de un grupo pequeño de moléculas puede aplicarse sobre un rango más alto de moléculas con grupos químicos similares (Monticelli y Tieleman, 2013, p. 2-3).

En dinámica molecular se requiere generalmente tener detalles de la estructura, propiedades e interacciones para los sistemas condensado-fase tales como moléculas orgánicas o una proteína en una solución. La calidad de los resultados depende de las expresiones englobadas en el campo de fuerza, la cual consiste en un conjunto de ecuaciones que modelan las energías potenciales y sus parámetros (Kaminski y Jorgensen, 1996: p. 18010).

Uno de estos modelos fue el conocido como OPLS (*Optimized potentials for liquids simulation*), el cual fue desarrollado y evaluado en el trabajo de Jorgensen y et al. (2002a: p. 11226-11227), donde definen los parámetros para cada una de las energías potenciales que intervienen en las interacciones de las moléculas, entre las que tenemos energía de las interacciones no enlazantes $E_{n.e.}$, la energía del estiramiento de enlaces E_{enlace} , la energía debido a un ángulo de flexión E_{angulo} y la energía de torsión de las moléculas $E_{torsion}$. En primer lugar, se define $E_{n.e.}$ del sistema mediante la suma de los términos de la fuerza de coulomb y la de Lennard-Jonnes, entonces la energía entre dos moléculas queda definida como

$$E_{ab} = \sum_i^{sobrea} \sum_j^{sobreb} \left[\frac{q_i q_j e^2}{r_{ij}} + 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) \right] f_{ij} \quad (22)$$

La misma expresión es usada para las interacciones no enlazantes intramoleculares entre todos los pares de átomos $i < j$ separados por tres o más enlaces. El término f_{ij} es un factor de escala con valor igual a 1 excepto para interacciones intramoleculares que toma el valor de la mitad. De igual manera, se establecen las ecuaciones para la energía de enlace y del ángulo de flexión

$$E_{enlace} = \sum_{enlaces} K_r (r - r|eq)^2 \quad (23)$$

$$(24)$$

$$E_{angulo} = \sum_{angulos} K_{\theta}(\theta - \theta|eq)^2$$

Para obtener los parámetros de energía de torsión se debe considerar una compensación debida a las energías del ángulo de flexión cuando θ es muy grande. En el mismo trabajo donde se postuló el modelo OPLS (Jorgensen y et al., 2002b: 11227) se postula una expresión para la energía de torsión y mediante una parametrización se encuentra una función que permite determinar el cambio de energía potencial durante una rotación ϕ sobre un enlace

$$E(\phi) = E_{n.e.}(\phi) + E_{enlace}(\phi) + E_{angulo}(\phi) + E_{torsion}(\phi) \quad (25)$$

En la ecuación (25) podemos observar que se considera cada una de las energías potenciales que aparecen durante una interacción molecular de una molécula sobre otra en función de un parámetro angular, es decir, de posición, por lo que cumple con las condiciones de campo de fuerza.

1.3.1.2. Campo de fuerza CHARMM

Se ha visto que el campo de fuerza es una herramienta integral para el estudio de fenómenos complejos en el área de simulaciones con biomoléculas. Existen diferentes programas que modelan el campo de fuerza para cálculos de dinámica molecular entre ellos tenemos CHARMM (*Chemistry at Harvard macromolecular mechanics*).

La primera simulación de dinámica molecular fue ejecutada en 1970 usando la primera versión de CHARMM y CHARMM FF. Durante el avance de las tecnologías este modelo ha incluido estrategias para optimizar los parámetros y añade aplicaciones potenciales en el descubrimiento de nuevos agentes terapéuticos. Aunque la presentación más antigua de este campo de fuerza se enfocaba en proteínas y ácidos nucleicos, en los últimos tiempos se ha extendido a lípidos y carbohidratos (Xiao y et al., 2011a: p. 167).

En el “*Journal of computational chemistry*” se presenta la primera versión de CHARMM de la mano de Bernard Brooks y sus colaboradores (1983, p. 190-193), anunciando el desarrollo de un programa flexible basado en el uso de funciones empíricas de energía para modelar los sistemas moleculares. La función de energía de CHARMM está formado por componentes que representan la energía interna, al igual que el modelo OPLS en la ecuación 25, y adicionalmente, se integran algunos términos:

- Las interacciones no enlazantes de Van der Waals E_{vdW} y del potencial electrostático E_{el} , el cual tiene varias expresiones dependiendo del tipo de sistema analizado
- Los enlaces de hidrógeno E_{hb} para lo cual se representa al átomo receptor anterior, el átomo receptor, al hidrógeno y el átomo pesado donador.
- Las restricciones para los armónicos de los átomos E_{c_r} y la restricción dihedral E_{c_ϕ} .

La versión de CHARMM27 fue desarrollada para el análisis de adn y arn, el cual consistía en una completa optimización de la versión de CHARMM. Este nuevo modelo incluye un mejor entendimiento de los mecanismos y energía libre de las roturas de Watson-Crick, lo que le permite realizar simulaciones que concuerdan con las observaciones experimentales, a diferencia de otros modelos como AMBER (Xiao y et al., 2011b: p. 176).

El nuevo modelo trae una notable mejora en la reproducción de datos experimentales, por esta razón, el mismo parámetro de optimización es implementado para la reproducción de nuevos modelos.

Modificando el campo de fuerza desarrollado en CHARMM27 (C27) se presenta un conjunto de ajustes a los parámetros de energía potencial para desarrollar CHARMM36 (C36). Los ajustes se realizaron debido a que en el anterior modelo los valores del área superficial por lípido son más pequeños que las estimaciones experimentales, por esta razón, en el nuevo modelo se modifican los parámetros de carga atómica parcial considerando la mecánica cuántica y los datos experimentales (Clauda y et al.,2010: p. 7831-7832).

Se observa una evolución importante en C36 con respecto a la versión C27 para el modelo de campos de fuerza, ya que existe una consistencia entre las simulaciones y los fenómenos reales que ocurren en una DM. La seguridad con la que trabaja C36 permite su utilidad en el desarrollo de simulaciones con biomoléculas heterogéneas, además de incluir estudios cuánticos a los modelos convencionales de la época. Aunque el modelo fue desarrollado para la mejora y optimización de lípidos, sentó las bases para un modelo de campo de fuerza optimizado para proteínas.

El biofísico computacional Robert Best junto a varios de sus colegas (2012: p. 3258-3261) presentan una mejora sobre los parámetros del campo de fuerza C36 con el fin de representar la energía potencial superficial de las proteínas. En el trabajo se implementa un nuevo potencial para el esqueleto de la proteína, se adiciona nuevos potenciales diédricos de cadena

lateral optimizados contra energías mecánicas cuánticas, y se añaden datos de RMN (Resonancia magnética nuclear) de proteínas desplegadas.

De igual manera para la optimización del esqueleto de la proteína al campo de fuerza CHARMM se le adiciona un término de potencial de ranura cúbico (CMAP) el cual es una función que depende de los dos términos de Fourier que representan los ángulos de torsión del esqueleto de la proteína, $V_{CMAP}(\varphi, \phi)$.

A pesar de las mejoras desarrolladas en la versión C36 se detectó un fallo en la generación de estructuras hélice alfa L para la arginina/serina inconsistente con los estudios RMN y los de dispersión de rayos X de ángulo pequeño.

Por esta razón, se implementa una mejora para el modelo de campo de fuerza enfocado en proteínas plegadas e intrínsecamente desordenadas (PDI), ya que estas estructuras fluctúan entre un conjunto de estados. En esta versión, se vuelve a definir un potencial CMAP para el esqueleto de la proteína basado en cálculos de reponderación y una mejor descripción para las interacciones específicas de los puentes salinos. El campo de fuerza modificado se puso a prueba con 15 péptidos y 20 proteínas en una simulación acumulativa de más de 500 microsegundos. Esta extensa simulación permitió concluir que el campo de fuerza mejorado es adecuado para simulaciones de DM con proteínas plegadas y desordenadas, el cual fue nombrado como CHARMM36m (Huang y et al., 2017, p. 175-176).

1.3.1.3. Restricción geométrica

En simulaciones de DM con sistemas de varios átomos la estructura molecular suele tener zonas rígidas de manera total o parcial. Las restricciones geométricas son impuestas para reforzar algunas características de la molécula o sistema que se está estudiando, algunos de los parámetros en los que se fijan restricciones son los ángulos de enlace, la longitud de enlace, entre otros.

Estas características imponen condiciones que permiten eliminar algunos modos vibracionales de la estructura lo que reduce el tiempo de integración numérica en cada paso de la simulación, lo que se traduce en un aumento de la eficiencia. Uno de los tipos de restricciones fundamentales es la restricción geométrica, donde se genera un modelo atómico general el cual resulta en un sistema parcialmente rígido. La forma más común de establecer restricciones geométricas es a través de una restricción de enlace donde la distancia entre dos

átomos d_{ij} se mantiene constante e igual a una longitud promedio de enlace, el cual se consigue de manera experimental (Rickaert, 1986a: 331). Podemos definir esta restricción a través de la siguiente expresión:

$$(r_i - r_j)^2 - d_{ij}^2 = 0 \quad (31)$$

Para tener un mejor modelo con restricciones que aseguren la rigidez de las moléculas se debe agregar una expresión para la rigidez lineal vectorial para el caso de ciertos átomos con menos cantidad de grados de libertad. Generalizando para todo tipo de moléculas (lineal, planar o tridimensional) se debe considerar un conjunto de átomos base y respecto a estos se conectan las restricciones rígidas de enlace, entonces se puede expresar la restricción de manera general para cualquier átomo (Rickaert, 1986b: 332).

$$r_k = r_1 + \sum_{i=2}^{n_b} a_{ik}(r_i - r_1) \quad (32)$$

Donde n_b representa los átomos no colineales cuando es igual a 2 o 3, mientras que los átomos coplanares están dados cuando es igual a 4. En la ecuación 31 las variables deben cumplir con la relación $k > n_b$. En el modelo se está usando los vectores $r_i - r_1$ como base de los espacios de 1,2 y 3 dimensiones y el coeficiente a_{ik} será fijado por la geometría molecular.

1.3.1.4. Campo de Fuerza del ligando

En MM el campo de fuerza se puede usar para el estudio de moléculas similares a fármacos de manera aislada o que interactúan con otros sistemas biológicos, por ejemplo, en los complejos proteína-ligando. El campo de fuerza visto en secciones anteriores funciona para sistemas macromoleculares, en este caso para el estudio de la proteína, sin embargo, aún se necesita establecer un modelo de campo de fuerza para moléculas de tipo ligando. En DM se observa que cada modelo para biomoléculas cuenta con su respectivo campo de fuerza general para moléculas tipo ligando o fármaco, lo mismo ocurre para CHARMM.

Los investigadores Vanommeslaegue y Mackerell (2012a: pp. 1-2) realizaron un trabajo con el objetivo de crear un campo de fuerza general aditivo para CHARMM, al cual lo llamaron CGenFF (CHARMM *general force field*), el cual utiliza las funciones aditivas de energía potencial para calcular la energía en función de sus coordenadas. De esta manera se considera

en el modelo una suma de los términos intramoleculares y los intermoleculares (ecuaciones 25-30).

El tipificado de átomos CGenFF primero asocia atributos a los átomos y enlaces en una molécula, como valencia, orden de enlace y pertenencia al anillo, entre otros. Luego de la tipificación de átomos, esta información se utiliza para asignar tipos de átomos CGenFF en función de un árbol de decisión programable, lo que permite una implementación sencilla de las reglas de tipificación de átomos (Vanommeslaegue y Mackerell, 2012a: pp. 8-9).

1.3.1.5. Modelo para la molécula de agua

Las soluciones acuosas y el agua tienen una gran importancia en los estudios de DM y en simulaciones de mecánica estadística. Por esta razón, es necesario implementar una función potencial intermolecular de la molécula de agua que contenga información razonable para el modelo del agua. Las funciones de potencial consideran monómeros de agua rígidos representados por 3, 4 o hasta 5 sitios de interacción.

Jorgensen y et al. (1983a: p. 926) estudiaron tres modelos del agua existentes en la época para luego desarrollar tres modelos nuevos y comparar todos los modelos. Todas las funciones de los modelos se basan en simulaciones de Monte Carlo, con un ensamble NPT a 25°C y 1 atmósfera. En los modelos propuestos por Jorgensen y et al. (1983b: p. 926) se parte de un modelo anterior de agua llamada TIPS3 (con 3 puntos de interacción), el cual consiste en cargas positivas sobre los hidrógenos y una carga negativa sobre el oxígeno, mediante esto se definía las interacciones de Coulomb entre todos los pares de cargas intermoleculares.

Las funciones de potencial consideran monómeros de agua rígidos representados por 3, 4 o hasta 5 sitios de interacción. Adicional al modelo se realiza una parametrización de los tres puntos lo que resulta en una mejora de la energía y la densidad del agua líquida, esta implementación está referenciada en el modelo llamado TIP3. La información proporcionada por las funciones potenciales son una base útil y valiosa para la simulación por computadora de muchos sistemas (Jorgensen y et al., 1983c: p. 926). Aunque vale recalcar que para sistemas con condiciones muy diferentes de temperatura y presión se necesita realizar una nueva parametrización.

1.3.2. Minimización de Energía

En las simulaciones de DM aplicando CCP se debe implementar una manera de resolver las interacciones entre las imágenes de las partículas que se replican, es decir, la celda unitaria del sistema. A estas relaciones se conoce como interacciones de largo alcance para lo cual tenemos dos convenciones un corte esférico caracterizado por un radio y la convención de mínima imagen. La primera convención es más útil para establecer límites de energía y presión, mientras que la segunda establece las fuerzas interatómicas para las cargas de los átomos.

En el estudio de los modelos de campo de fuerza se ha considerado los términos de las interacciones intermoleculares y para cada átomo se calcula su relación con el resto de los átomos que conforman el sistema. Esto produce términos que pueden ser despreciados debido al gran espacio que existe entre dos átomos, ya que este tipo de interacciones son inversamente proporcionales a la distancia, como se observa en las ecuaciones 22 y 26. Esta operación requiere establecer un límite para este tipo de interacciones y una corrección sobre la energía a través de un término de dispersión. En el proceso se obtienen términos negativos de dispersión debido a la naturaleza misma de las interacciones, las cuales actúan corrigiendo la energía y la presión.

La corrección de energía suele ser pequeña y la de presión es muy grande por lo que no se puede despreciar en ninguna circunstancia, sobre todo si se realizan simulaciones de ensamblajes NPT. Aunque, en principio, es posible parametrizar un campo de fuerza de modo que la presión esté cerca del valor experimental deseado sin corrección, dicho método hace que la parametrización dependa del corte. Para establecer los términos de dispersión se debe fijar un valor de corte r_c ; asumiendo que el sistema es homogéneo lo que permite establecer una expresión de la energía de dispersión (GROMACS development team, 2019a: párr. 1-2).

Cuando se analiza un sistema con CCP se establece una convención de mínima imagen o también conocido como imagen más cercana el cual contiene la ventaja de mantener neutralizada la carga local. Una condición necesaria para esto es que la dimensión lineal del sistema (celda unitaria) sea mayor al rango de correlaciones estáticas.

La regla de mínima imagen falla al crear funciones de correlación dinámica, incluso para sistemas con acoplamiento relativamente menor. Se podría implementar el mismo proceso que en la sección anterior sin embargo no funcionaría en el caso de las propiedades dinámicas electrostáticas. En el caso de sistemas de Coulomb la mejor manera de proceder es reemplazar la fuerza de Coulomb entre dos partículas por una efectiva obtenida a través de

la suma de las interacciones entre todas las imágenes periódicas. Para implementar este procedimiento se puede adoptar como objetivo los pares de fuerza, haciendo uso de la bien definida técnica de sumas de Ewald (Hansen, 1986, p. 92- 93). La suma de Ewald se introdujo por primera vez como un método para calcular las interacciones de largo alcance de las imágenes periódicas en los cristales. La idea es convertir la suma única de convergencia lenta en dos términos de convergencia rápida y un término constante (GROMACS development team, 2022, párr. 1-4).

1.3.3. Equilibrio del sistema

En las simulaciones de DM, a menudo surge la necesidad de mantener parámetros como la temperatura o la presión en lugar de la energía y el volumen. Existe un método para realizar el acoplamiento a un baño externo con temperatura o presión constante con tiempo ajustables constantes para el acoplamiento. El método es extensible a otras variables y gradientes, y puede aplicarse también a moléculas poliatómicas que implican restricciones internas.

El sistema debe acoplarse con el calor baño externo el cual esta caracterizado por una temperatura constante T_o . Por lo que se insertan nuevos términos estocásticos y de fricción en la ecuación de movimiento.

$$m_i \dot{v}_i = F_i - m_i \gamma_i v_i + R(t) \quad (43)$$

F_i Es la fuerza del sistema, $R(t)$ me define la variable estocástica Gaussiana y γ_i es una constante de amortiguación que determina la fuerza de acoplamiento que existe entre el sistema y el baño. Físicamente, esta ecuación representa la frecuencia de las colisiones con las partículas que conforman el gas ideal, con su respectiva T_o . Considerando que la fricción es constante para todas las partículas podemos definir un $\gamma = \gamma_i$ y también se define una temperatura cuando el sistema está bajo la influencia del acoplamiento estocástico T . De esta manera los grados de libertad pueden en principio ser acoplados con el baño usando diferentes constantes de fricción (Berendsen y et al., 1984a: 3684-3685).

$$\frac{dE_k}{dt} = \sum_{i=1}^{3N} v_i F_i - 2\gamma \left(\frac{3N}{2} kT_o - E_k \right) \quad (44)$$

El primer término en la ecuación 44 es exactamente la derivada temporal de la energía potencial, mientras que el segundo término adicional describe el acoplamiento global al baño

de calor. El acoplamiento a un baño de presión constante se puede realizar según el mismo principio, pero ahora se agrega un término adicional a las ecuaciones de movimiento que efectúa un cambio de presión (Berendsen y et al., 1984a: 3684-3686). De esta manera se puede definir los términos adicionales para cada acoplamiento

$$\left(\frac{dT}{dt}\right)_{\text{baño}} = \frac{(T_o - T)}{\tau_T} \quad (45)$$

$$\left(\frac{dP}{dt}\right)_{\text{baño}} = \frac{(P_o - P)}{\tau_P} \quad (46)$$

En las ecuaciones 45 y 46, en cada uno de los términos extra en el acoplamiento de temperatura y presión, se ha introducido la constante de tiempo τ que en este caso es igual a

1.3.4. Cantidades estadísticas simples

Las mediciones en DM significan establecer un promedio de tiempo de las propiedades físicas sobre la trayectoria del sistema, ya que usualmente son funciones de las coordenadas y velocidades de las partículas. Por ejemplo, si una propiedad física A estuviera en función de las posiciones y velocidades de las N partículas, entonces el promedio de esa propiedad durante un tiempo T se puede obtener mediante

$$\langle A \rangle = \frac{1}{N_T} \sum_t^{N_T} A(t) \quad (47)$$

Donde t representa un instante o paso del número total de pasos existentes. Algunas propiedades físicas que siguen este tipo de cantidades estadísticas son: la energía potencial, energía cinética, energía total, temperatura, volumen, densidad y presión (Ercolessi, 1997, pp. 21-22).

1.3.5. Análisis estructurales y de trayectorias

Al ejecutar las simulaciones de DM se obtiene diferente tipo de información con respecto a la dinámica y la estructura del sistema estudiado. Una de las mediciones que comúnmente se realiza para el análisis de las trayectorias luego de una simulación DM es la desviación de la raíz cuadrada media o RMSD, por sus siglas en inglés (*root mean square deviation*).

El RMSD de las coordenadas atómicas es la opción más popular para cuantificar las diferencias en las estructuras y dinámicas macromoleculares. Se usa comúnmente en análisis de trayectorias de dinámica molecular (MD) para establecer el período de equilibrio y la calidad de las simulaciones biomoleculares y para agrupar conformaciones similares (Sargsyan, 2017, pp. 1518-1520). Estos valores representan las fluctuaciones temporales de una molécula respecto a una estructura de referencia. La expresión del RMSD es

$$RMSD = \sqrt{\frac{\sum_i^N \|r_i - \hat{r}_i\|^2}{N}} \quad (48)$$

x_i y x_r definen la partícula i de la estructura estudiada y de la estructura de referencia, respectivamente. Otra medición que se puede evaluar es la del radio de giro, ya que brinda información sobre la flexibilidad de la molécula con el sistema y la expresión se puede definir usando los mismos conceptos que en la ecuación anterior (Barragán, 2015e: p. 31).

$$Rg = \left(\frac{\sum r_i^2 m_i}{\sum m_i} \right)^{1/2} \quad (49)$$

Donde m_i es la masa de cada átomo y r la distancia de cada átomo respecto al centro de masas.

CAPÍTULO II

2. MARCO METODOLÓGICO

El presente proyecto de investigación se realiza principalmente con el uso de herramientas computacionales para elaborar simulaciones físicas de sistemas moleculares. En primer lugar, la construcción de las muestras o modelos moleculares se hace en un computador personal a través del software gromacs y otras herramientas que permiten ejecutar de la mejor manera las simulaciones, las cuales fueron virtuales y locales.

La ejecución de algunas fases de las simulaciones se realizó a través de un clúster el cual se configuró para realizar los comandos de dinámica molecular. Esto permite usar un computador de altas prestaciones de manera virtual a través de conexión de internet desde un computador personal. Los resultados se almacenan en una memoria en la nube proporcionado por la institución que opera el clúster. La información generada se puede descargar a través del internet.

El análisis de los datos generados por la DM se realiza mediante programación en Python, en un entorno de trabajo instalado y adaptado a las necesidades del proyecto de investigación. A través de estos recursos de computación se generan las imágenes y datos necesarios para presentar los resultados y la discusión de estos. A continuación, se realiza la descripción de todos los recursos usados en el proyecto, además de una base teórica de los métodos matemáticos escogidos en la resolución de la DM. Al final de esta sección se realiza una descripción más específica de los métodos usados en cada fase del proyecto.

2.1. Recursos

2.1.1. Computacionales

El desarrollo de los modelos moleculares y la redacción del presente trabajo de investigación se realizó en una computadora portátil. A continuación, se presentan sus características

Tabla 1-2: Características del computador personal

Marca	Lenovo
Modelo	V15-ADA
Procesador	AMD Ryzen3
Disco Duro	1 Tb

Número de núcleos	64
Número de hilos	128
Reloj de aumento máximo	Hasta 3.4 GHz
Reloj base	2.25 GHz
Caché L3 total	256 MB
RAM	8 Gb
Sistema operativo	Ubuntu 22.04
Número de sockets	1P/2P

Elaborado por: Chávez, I., 2022.

Las simulaciones del presente trabajo de investigación se realizaron en un clúster proporcionado por la Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia (CEDIA), llamado comúnmente supercomputador (HPC), tiene grandes prestaciones ideal para el desarrollo de algoritmos de DM incluyendo el procesamiento de los datos a través de una red de alta velocidad.

Tabla 2-2: Características del HPC de CEDIA

Capacidad	10 PFLOP en AI 2 nodos NVIDIA DGX A100 2TB de memoria RAM +130 TB almacenamiento Usuarios ilimitados por institución 16 GPUs NVIDIA A100 compartidos Almacenamiento por 1 año
Disponibilidad	Soporte 16x7 Monitorización en tiempo real

Fuente: CEDIA, 2022.

Elaborado por: Chávez, I., 2022.

La capacidad que tiene cada GPU que conforma el HPC para realizar operaciones computacionales tienen las siguientes características

Tabla 3-2: Características del GPU A100

Núcleos	6912
TMU	432
ROP	160
Memoria	40GB
Tipo de memoria	HBM2e
Ancho de Bus	5120 bits

Fuente: techpowerup.com, s.f.

Elaborado por: Chávez, I., 2022.

2.1.2. Software

VMD. - software diseñado para el modelado, visualización y los análisis biológicos, proporciona diferentes formas de representar una molécula, además se puede utilizar para animar y analizar la trayectoria de una simulación de dinámica molecular (MD). En particular, puede actuar como una interfaz gráfica para un programa de MD externo al mostrar y animar una molécula en proceso de simulación en una computadora (Theoretical and computational biophysics group, 2008, párr. 1).

AVOGADRO. - es un avanzado editor y visualizador de moléculas diseñado para uso multiplataforma en química computacional, modelado para moléculas, bioinformática, ciencia de los materiales, entre otros. Ofrece un renderizado de alta calidad y una gran variedad de complementos en su potente arquitectura (Avogadro Chemistry, s.f, párr. 1).

GROMACS. - es uno de los códigos de software libre y de código abierto más utilizados en química, utilizado principalmente para simulaciones dinámicas de biomoléculas. Proporciona un amplio conjunto de tipos de cálculo, herramientas de preparación y análisis. Se admiten varias técnicas avanzadas para los cálculos de energía libre y para generar distintos campos de fuerza especializado para cada tipo de biomolécula, que, a su vez, cuentan con su propio campo de fuerza general (Abraham y et al., 2015, pp. 19-20).

Tabla 4-2: Programas para generar parámetros compatibles a cada campo de fuerza

AMBER	Antechamber	Parametriza moléculas usando GAFF
	acpype	Una interfaz de Python para AnteChamber, escrita con topologías de GROMACS
CHARMM	CGenFF	Servidor oficial de CHARMM para generar el campo de fuerza
GROMOS87/ GROMOS96	PRODRG 2.5 ATB	Un servidor automático para generar topología Nuevo servidor para la generación de topología, usando GROMOS96 54A7
OPLS-AA	TopolBuild	Convierte un archivo Tripos .2mol en una topología
	TopolGen	Un Script que permite convertir todo archivo .pdb en un archivo topológico
	LigParGen	Un servidor del grupo Jorgensen para producir topologías OPLS

Fuente: MD tutorials, 2020

Realizado por: Chávez, I., 2022

ANACONDA. - es una distribución de código abierto de los lenguajes de programación R y Python para la ciencia de datos que tiene como objetivo simplificar la administración y la implementación de paquetes. Su administración incluye paquetes adecuados para Linux, macOS y Windows (Rondón, 2022, párrs. 2-3).

2.1.3. Virtuales

PROTEIN DATA BANK. - es el único archivo mundial de datos estructurales de macromoléculas biológicas. Incluye datos obtenidos por cristalografía de rayos X y espectrometría de resonancia magnética nuclear (RMN) presentados por biólogos y bioquímicos de todo el mundo, cuya misión es mantener un único Archivo de Banco de Datos de Proteínas de datos estructurales macromoleculares que esté libre y públicamente disponible para la comunidad mundial (Berman y et al., 2000, párrs. 1-2).

IMAGE CONTAINERS. - es un archivo estático e inalterable que incluye un código ejecutable para que pueda ejecutar un proceso aislado en la infraestructura de tecnología de la información (TI). La imagen se compone de bibliotecas del sistema, herramientas del sistema y otras configuraciones de plataformas que un programa de software necesita para ejecutarse en una plataforma de contenedores como Docker o CoreOS Rkt . La imagen comparte el kernel del sistema operativo de su máquina host (Earls, 2018, párrs. 1-2).

ENROOT. - es una herramienta simple pero poderosa para convertir imágenes de contenedores/OS tradicionales en entornos limitados sin privilegios. Utiliza las mismas tecnologías subyacentes que los contenedores, pero elimina gran parte del aislamiento que brindan de manera inherente al mismo tiempo que preserva la separación del sistema de archivos (NVIDIA corporation, 2022, párrs. 1-2).

CONEXION HPC. – tiene la capacidad de realizar cálculos y procesar datos a altas velocidades, además aprovecha el poder de los clústeres en la resolución de problemas que requieren computación masiva, lo cual permite procesar grandes cantidades de datos a una mayor velocidad que una computadora normal, generando también conocimientos más rápidos (IBM, s.f., párr. 1).

CHARMM36.ff. - El campo de fuerza de proteína CHARMM36 aditivo de todos los átomos se usa ampliamente en modelos y simulaciones moleculares. Con respecto al aditivo proteico CHARMM FF, se presentó una actualización reciente que produjo el CHARMM36 (C36) FF, realizando principales mejoras sobre CHARMM22/CMAP que incluyen potenciales de CMAP de columna vertebral refinados y nuevos parámetros diédricos de cadena lateral, además tiene mayor precisión en la generación de conjuntos conformacionales de la columna vertebral del polipéptido para péptidos y proteínas intrínsecamente desordenados (Mackerell Lab, s.f., párr. 21).

CGENFF. - realiza la tipificación de átomos, la asignación de parámetros y cargas por analogía de forma totalmente automatizada. La tipificación de átomos se realiza mediante un árbol de decisión programable determinista, mientras que la asignación de parámetros enlazados se basa en la sustitución de tipos de átomos en la definición del parámetro deseado. Se asocia una penalización a cada sustitución y se elige el parámetro existente con la penalización total más baja como aproximación al parámetro deseado, esta puntuación de penalización se devuelve al usuario como una medida de la precisión de la aproximación (Mackerell Lab, 2018, párr. 1).

2.1.4. Python

Python es un lenguaje de programación interpretado, que tiene variables dinámicas, es decir que no necesita que sus variables sean declaradas con anterioridad y durante la ejecución del código puede variar su tipo entre entero (int), decimal (float) o cadena (str). Es multiplataforma, orientado a objetos y además tiene una amplia curva de aprendizaje por lo que es considerado un buen lenguaje para empezar en el mundo de la programación.

JUPYTER NOTEBOOK. - es una aplicación web de código abierto que permite compartir y crear documentos. Se puede considerar también como una aplicación cliente-servidor, ya que la aplicación inicia el servidor en la máquina local y abre la interfaz del portátil en el navegador web desde donde se puede editar y ejecutar. El cuaderno se guarda como archivo ipynb y se puede exportar como archivos html, pdf y LaTeX (Jupyter org, 2015, párr 1).

NUMPY. - biblioteca que pertenece a Python utilizada principalmente para trabajar con matrices, también tiene funciones para trabajar en el dominio del álgebra lineal, la transformada de Fourier, entre otras. Es importante añadir que es un proyecto de código abierto y se lo puede usar libremente (Vanderplas, 2016a: p. 33).

MATPLOTLIB. - es una biblioteca multiplataforma de visualización de datos y trazado gráfico para Python y su extensión numérica NumPy. Como tal, ofrece una alternativa viable de código abierto a MATLAB. Los desarrolladores también pueden usar las API (interfaces de programación de aplicaciones) de matplotlib para incrustar gráficos en aplicaciones GUI. Un script Python matplotlib está estructurado de modo que unas pocas líneas de código son todo lo que se requiere en la mayoría de los casos para generar un gráfico de datos visuales (Vanderplas, 2016b: p. 217).

MODELLER. - . El programa construye modelos basados en restricciones: restricciones derivadas de la homología que se extraen de la alineación de la plantilla y el objetivo; restricciones estereoquímicas, que incluyen longitudes de enlace y ángulos de enlace, que se obtienen mediante el campo de fuerza de la mecánica molecular CHARMM, y ángulos diédricos y distancias atómicas no enlazadas, que se obtienen de un conjunto representativo de todas las estructuras proteicas conocidas; y, por último, y también opcionalmente, cualquier restricción que pueda agregar el usuario, es decir, reticulación o estructura secundaria predicha. También implementa el modelado comparativo de estructuras proteicas mediante la satisfacción de restricciones espaciales, y puede realizar muchas tareas adicionales, incluido el modelado de Novo de bucles en estructuras proteicas, la optimización de varios modelos de estructuras proteicas con respecto a una función objetivo definida de forma flexible, alineación múltiple de secuencias y/o estructuras de proteínas, agrupamiento, búsqueda de bases de datos de secuencias, comparación de estructuras de proteínas, entre otros (Sali, 2022, párr. 1).

2.2. Métodos matemáticos y algoritmos

En las etapas de dinámica molecular es necesario utilizar algunos métodos matemáticos para resolver los problemas matemáticos y computacionales que aparecen en la ejecución de las simulaciones. Durante los últimos siglos han parecido distintos nuevos métodos que permiten minimizar la función, realizar una integración temporal y diferentes algoritmos específicos para una etapa en la dinámica molecular. GROMACS integra varios de estos algoritmos para resolver los cálculos en cada etapa de la DM.

2.2.1. Método del descenso más pronunciado

Es un método para realizar la minimización de una función no lineal y se propuso por primera vez de la mano de Cauchy en el cual buscaba las raíces de la función descrita por n variables. Por esta razón este método ha jugado un papel muy importante en el desarrollo de la teoría de optimización. Una mejor descripción del método se encuentra en el artículo escrito por Meza (2010, pp. 719-720) con el nombre del paso más pronunciado, donde además nos brinda un vistazo del pseudocódigo que describe el algoritmo para trabajar computacionalmente con esta metodología y hallar la optimización de una función no lineal:

Supongamos que nos gustaría encontrar el mínimo de una función $f(x)$, $x \in \mathbb{R}^n$ y $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Denotaremos el gradiente de f por $g_x = g(x_k) = \nabla f(x_k)$. La idea general detrás de la mayoría de los métodos de minimización es calcular un paso a lo largo de una dirección de búsqueda dada, d_k . Para el método de descenso más pronunciado, la dirección de búsqueda está dada por $d_k = -\nabla f(x_k)$. El algoritmo para el método de descenso más pronunciado puede ser escrito como

```

Given an initial  $x_0, d_0 = -g_0$ , and a convergence
tolerance  $tol$ 
for  $k = 0$  to  $maxiter$  do
  Set  $\alpha_k = \operatorname{argmin} \phi(\alpha) = f(x_k) - \alpha g_k$ 
   $x_{k+1} = x_k - \alpha_k g_k$ 
  Compute  $g_{k+1} = \nabla f(x_{k+1})$ 
  if  $\|g_{k+1}\|_2 \leq tol$  then
    converged
  end if
end for

```

Ilustración 1-2: Algoritmo del método de descenso más pronunciado para optimizar una función no lineal

Realizado por: Meza, 2010.

2.2.2. Integración del tiempo

En una simulación de DM es necesario realizar integraciones con respecto al tiempo de ecuaciones diferenciales no lineales bastante complejas y con una cantidad considerable de elementos, debido al número de partículas que intervienen en el sistema. Realizar estos cálculos requieren de poder un poder computacional considerable, por lo que han surgido métodos computacionales que permiten realizar este proceso, entre ellos: el algoritmo de Verlet y el integrador *Leap-frog*.

Ambos métodos no usan correctores sobre las coordenadas, son fáciles de usar y requieren menos memoria. Estos métodos producen las mismas trayectorias, por lo tanto, son equivalentes. Tienen una ventaja importante que les permite ser reversibles en el tiempo, lo que esencialmente implica que no tengan desvíos (Barendsen y Van Gunsteren, 1986, p. 56). El esquema que sigue el algoritmo de Verlet es el siguiente

$$x_{n+1} = 2x_n - x_{n-1} + h^2 f(x_n) \tag{50}$$

El algoritmo no contiene velocidades explícitas, pero se pueden observar a través de la aproximación

$$v_n = (x_{n+1} - x_{n-1})/2h \quad (51)$$

Otra formulación del mismo algoritmo es el esquema de *leap-frog*, el cual consiste en

$$v_{n+\frac{1}{2}} = v_{n-\frac{1}{2}} + hf(x_n) \quad (52)$$

$$x_{n+1} = x_n + hv_{n+\frac{1}{2}} \quad (53)$$

La equivalencia entre estos dos esquemas hace que todas sus propiedades sean las mismas.

2.2.3. Algoritmos de Acoplamiento

El método de simulación por computadora de DM es una gran herramienta para estudiar propiedades dinámicas de moléculas en soluciones. Existe un interés por estudiar las propiedades biomoleculares de los sistemas como función de la temperatura y la presión, en lugar de usar el volumen y la energía como variables.

En 1984, Berendsen y et al. (pp. 3687-3689) publican un método para realizar el acoplamiento de la DM con un baño de externo, el cual puede tener la temperatura o la presión constante durante el tiempo, con valores ajustables. Los casos fueron presentados a través de un integrador *leap-frog* para el caso de restricciones generales de con acoplamiento de temperatura y presión. Se evaluaron los resultados del método y se concluyó que es ideal para generar simulaciones de DM no equilibradas.

Con la introducción de la DM a presión constante se lograron grandes avances y en particular con la aparición del método de Parrinello-Rahman, en el que se permite que la caja cambie de forma para adaptarse a una nueva estructura. El método Parrinello-Rahman ahora se describe en los libros de texto y también se usa ampliamente en diferentes variantes (Martonak y et al., 2003, p.1).

2.3. Metodología para la simulación de dinámica molecular

Existen diferentes etapas durante la creación de la simulación por DM en el presente trabajo debido a que se crearon distintos archivos con características biomoleculares diferentes. En

general el proceso se puede resumir en una primera etapa de preparación del sistema. Luego, la etapa de ensamblaje para que el sistema se equilibre. El tercer paso es definir la ejecución de la dinámica molecular y finalmente la extracción de los resultados de la DM.

Los archivos PDB de los sistemas que se estudian en el presente trabajo se obtienen del PDB, cada uno con su respectivo código. Cada archivo tiene una sección de residuos faltantes, que representa aquellas partes de la proteína que no se pudo determinar su estructura mediante el experimento que se haya realizado en la obtención de su estructura (cristalografía de rayos X). Estos faltantes se detectan gracias a la cadena de aminoácidos la cual también se dispone en el PDB mediante un archivo fasta. Por lo que, a cada estructura se le aplicó un preprocesamiento para completar esto “*missing residues*” a través de la librería de Python *modeller*. Se elaboró un programa que automatiza todo el proceso de comparación entre la secuencia generada por *modeller* y la secuencia verdadera para obtener como salida un fichero alignment.ali (Anexo A) el cual se utiliza para completar la cadena usando los modelos disponibles en la librería (Anexo B, C y D). La salida del código serán los modelos favorables para el sistema junto a su puntuación molpdf, de los cuales se escoge el de menor valor para la preparación de la muestra.

La topología se realiza utilizando el campo de fuerza CHARMM36m para la proteína, los ligandos se configuran usando CGenFF y las moléculas de agua tienen el modelo TIP3. Para la reconstrucción del complejo proteína-ligando se realiza un proceso previo a la molécula del ligando. Se le agrega moléculas de hidrógenos al archivo con el ligando aislado, esto se realiza mediante el software de Avogadro y se guarda con un archivo tipo mol2. Este último archivo es el que se utiliza para generar la topología en CGenFF. El modelo debe tener condiciones fisiológicas, es decir la celda unitaria tiene que pasar por un proceso de neutralización, solvatación y equilibrio.

En primera instancia, La celda unitaria del sistema se crea a través del comando de gromacs editconfig, el cual será un dodecaedro rómbico con la molécula centralizada en la caja y se solvata la proteína con el comando *solvate*, en la entrada del soluto se ingresa el complejo generado anteriormente y para la estructura del solvente es spc216.

La neutralización se realiza mediante los comandos *grompp* y *genion*, donde se escoge al calcio como ion y el sodio como protón. Para el parámetro de corte se utilizará el algoritmo de Verlet y con un integrador que usa el método del descenso más pronunciado configurado en el archivo “ions.mdp”.

La energía del sistema es minimizada con el comando *grompp* e ingresando un nuevo archivo “em.mdp”, este proceso tiene una parada de minimización de 1000 KJ/mol, un ancho de paso de energía de 0.01 y el máximo de 50000 pasos. (Anexo E). Se ejecuta el proceso de simulación con el comando *mdrun -v -deffnm* lo que nos permite obtener los archivos con el mismo nombre que ingresamos en *grompp*.

El equilibrio del solvente y los iones alrededor de la proteína se realiza con un ensamble NVT mediante los comandos *grompp* y *mdrun*, ingresando un nuevo archivo “nvt.mdp” en el cual se ha configurado un integrador tipo *leap-frog*, con 50000 pasos, un ancho de 2 femtosegundos, las energías, velocidades y coordenadas se guardan en cada picosegundo de tiempo. Las fuerzas electrostáticas de Coulomb serán de tipo PME para electrostática de alto rango y el termostato se basa en el modelo de Barendsen. Finalmente, las velocidades y temperaturas se asignan basándose en la distribución de Maxwell (Anexo F).

Los grupos de acoplamiento varían dependiendo de que sistema se analice. Cuando solo se encuentra la proteína HSA debemos configurar los grupos “*Protein Non-Protein*”. Los grupos de acoplamiento del sistema HSA – DKA son “*Protein_DKA Water SOD*”. Finalmente, para la HSA – WFR se acoplan los grupos “*Protein_WFR Water SOD*”.

En los sistemas donde existen ligandos es necesario crear restricciones de posición de la proteína y el ligando, ya que ambos fueron rejuntables al crear la topología del sistema. En primer lugar, se crea una restricción sobre la topología del ligando sin considerar a los hidrógenos utilizando el comando de *gromacs make_ndx*, sobre este índice usamos el comando *genrestr*, el cual devuelve un archivo “dka_posre.itp” que contiene la información de restricción, y la añadimos en la topología del sistema con *#ifdef POSRES_LIG*. Luego, utilizando el archivo gro de todo el sistema se crea un índice que nos permita seleccionar los grupos atómicos que conforman la proteína y el ligando.

El proceso de equilibrio continúa con un ensamble NPT usando los mismos comandos que el paso anterior, pero en este caso se ingresa un archivo “npt.mdp”, el cual tiene todos sus parámetros con las mismas características que el archivo para el equilibrio nvt, pero se agregan los parámetros de acoplamiento de presión. En este estudio, el acoplamiento se realizó mediante el método de Parrinello-Rahman del tipo isotrópico, con una constante de tiempo de 2 picosegundos y una presión de referencia de 1 bar (Anexo G).

La ejecución de la DM de los tres sistemas propuestos en este estudio se configuró con el integrador de tipo *leap-frog*, con un ancho de paso de 2 femtosegundos y 5 millones de pasos, dando un tiempo de simulación de 10 nanosegundos. Los parámetros de electrostática,

baróstato y acoplamiento de presión tienen que ser el mismo que se utilizó durante la fase de minimización y equilibrio del sistema. Los datos obtenidos se almacenan cada 10 picosegundos de simulación ejecutada (Anexo H). En cada uno de los archivos .mdp, para cada grupo, se debe ingresar el valor de constante temporal (0.1 ps) y la temperatura de referencia (300 K).

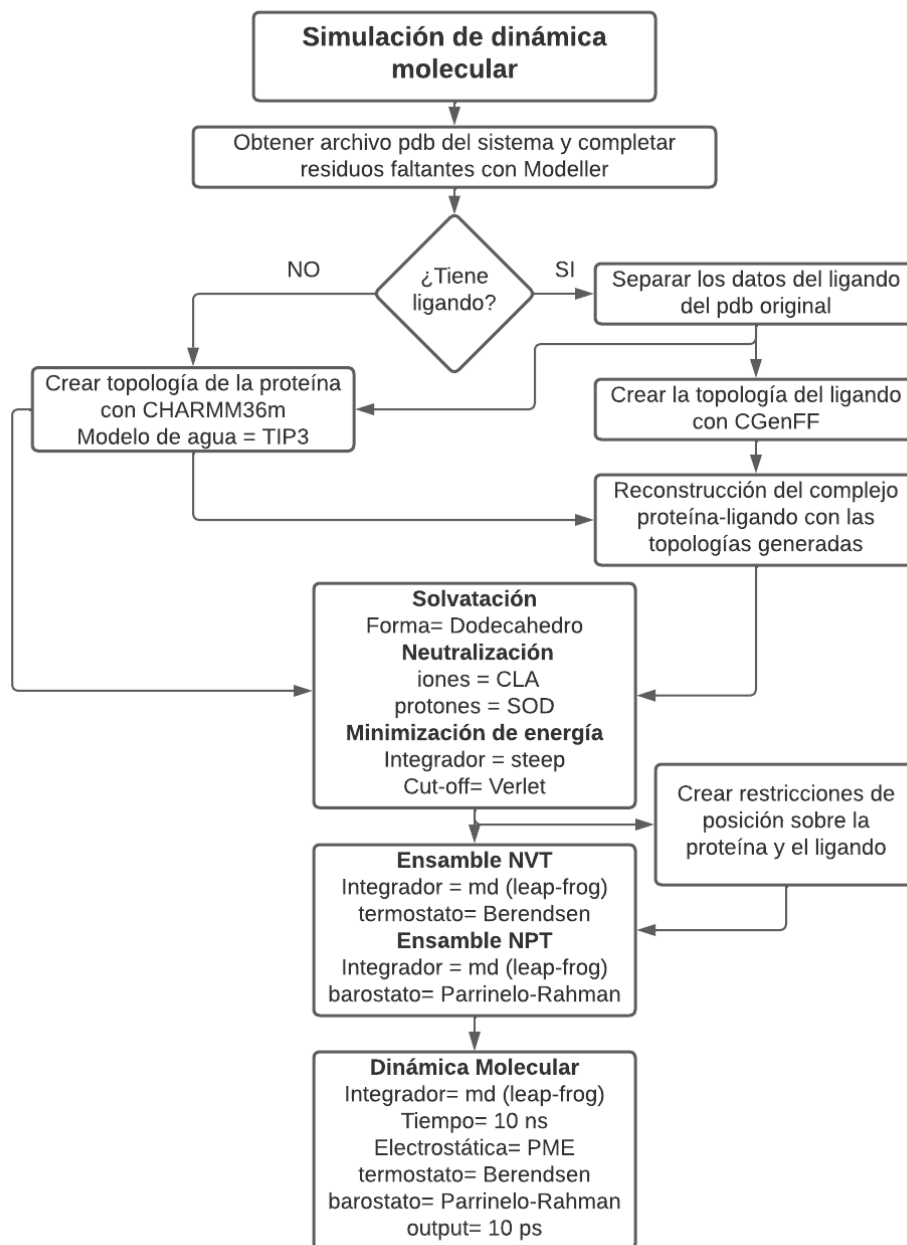


Ilustración 2-2: Proceso metodológico usado en la simulación de dinámica molecular

Realizado por: Chávez, I., 2022

2.3.1. Archivos generados

Durante el proceso de preparación y ejecución de una DM en cada etapa se crean varios archivos de distintos tipos o es necesario generar algunos de estos para tener la entrada de los comandos en gromacs. A continuación, se presenta un resumen de lo que significa la extensión de cada archivo y la información que almacena.

- . **gro** contiene la estructura molecular en formato GROMOS87
- . **rtp** brindan información de la topología de una molécula o sección de un sistema en específico. Se importan dentro del archivo de topología general del sistema
- . **top** Es el archivo de la topología de todo el sistema. Es un archivo ascii que se lee usando *gmx grommp* y crea una topología binaria
- . **tpr** el archivo contiene la estructura inicial de la simulación, también la topología de la molécula y todos sus parámetros
- . **trr** contiene datos sobre las trayectorias moleculares de la simulación, tales como velocidades, coordenadas, fuerzas y energías.
- . **edr** archivo donde se guardan las energías
- . **log** tan solo es un archivo de texto donde se imprimen la información generada y configuraciones ejecutadas por gromacs, también contiene una explicación más detallada de los errores de ejecución, en caso de haberlos.
- . **cpt** tiene el estado completo de una simulación, para ejecutar desde "*check-point*"
- . **xtc** es un archivo completo para las trayectorias en formato más portátil
- . **xvg** archivo de salida de los comandos de gromacs que analizan el sistema post ejecución de la DM y es de software libre
- . **mol2** es una representación completa y portátil de una molécula creada en SYBYL
- . **prm** es usado para establecer e ingresas los datos de los parámetros moleculares
- . **ndx** es un archivo de índice que define dentro de la estructura grupos de moléculas proveniente de una base de datos con la información molecular del sistema.

CAPÍTULO III

3. MARCO DE RESULTADOS Y DISCUSIÓN DE LOS RESULTADOS

3.1. Configuración y preparación de las muestras a simular

El proceso de preparación de cada modelo estudiado tiene diferencias en algunos pasos, por lo que, a continuación, se presenta por separado los resultados de las muestras obtenidos.

3.1.1. De la proteína HSA

Las coordenadas iniciales del HSA sin ligando se obtiene con el código 1AO6. El archivo pdb de la proteína se deriva de la estructura cristalina en rayos X con una resolución de 2.5 angstroms. La Ilustración 1-3 representa la disposición tridimensional del archivo pdb descargado, mientras que la Ilustración 2-3 es el mismo sistema luego de hace la mutagénesis de los residuos faltantes. Se puede identificar que la estructura y los dominios de la segunda molécula están mejor definidos y no hay zonas de la proteína donde la cadena sea discontinúa.

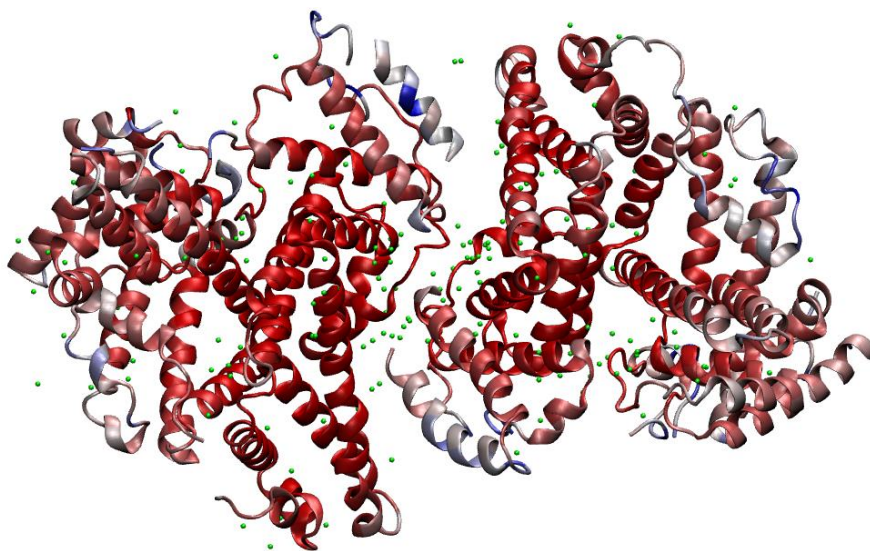


Ilustración 1-3: Modelo de cinta de la proteína HSA sin ligandos con el pdb original.

Realizado por: Chávez, I., 2022.



Ilustración 2-3: Modelo de cinta de la proteína HSA sin ligandos luego de la mutagénesis.

Realizado por: Chávez, I., 2022.

El archivo escogido tuvo una puntuación molpdf de 194.59, el cual se cambia de nombre para trabajar los pasos de preparación de la muestra (“HSAfilled.pdb”). Las características de la celda unitaria creado con gromacs luego de la solvatación y neutralización se pueden observar en la Tabla 1-3. Las coordenadas de nuevo centro son distintas al centro inicial del sistema lo que indica que la proteína ha sido desplazada. Esto ocurre porque hemos especificado que la HSA quede ubicada en el centro de la celda unitaria. El sistema tiene una carga neta de -15, lo que explica porque durante el proceso de neutralización se agregaron en posiciones aleatorias 15 protones de sodio (Ilustración 3-3).

Tabla 1-3: Características de la celda unitaria con la HSA solvatada y neutralizada.

Coordenadas	x	y	z
Tamaño del sistema (nm)	8,024	7,714	7,816
Centro (nm)	1,081	1,617	1,174
Vectores de la Celda (nm)	8,025	7,714	7,815
Variación (nm)	2,931	2,240	2,734
Nuevo Centro (nm)	4,012	3,857	3,908
Ángulos	90°	90°	90°
Diámetro (nm)	9,224		
Volumen (nm ³)	483.768		
Densidad (g/l)	1054,46		
Átomos del soluto	9233		

Realizado por: Chávez, I., 2022.

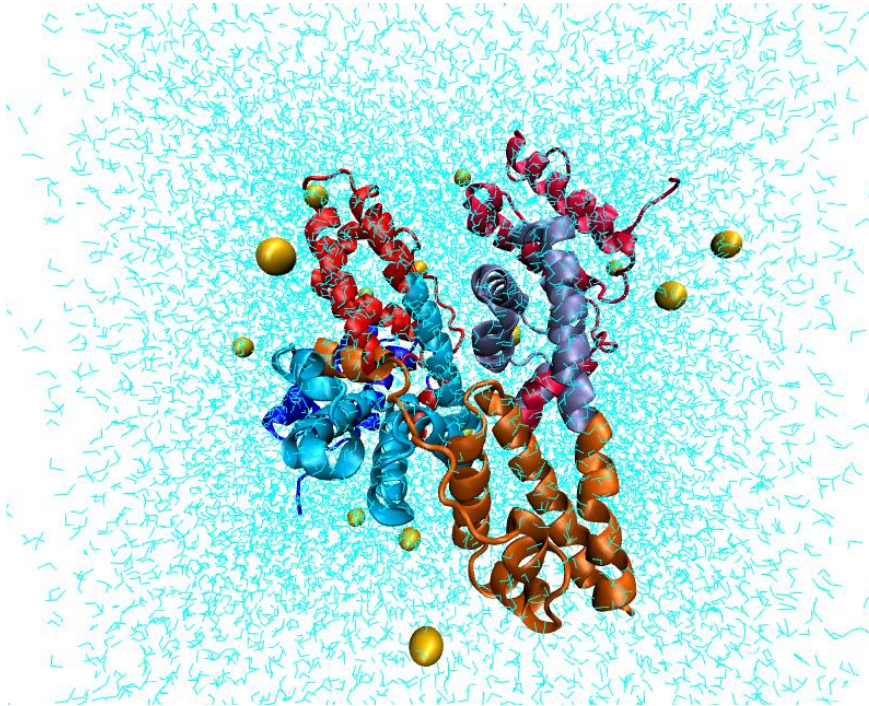


Ilustración 3-3: Modelo de la celda unitaria con la HSA neutralizada y solvatada.

Realizado por: Chávez, I., 2022.

3.1.2. Del complejo HSA-DKA

Para el archivo PDB del complejo ligando-proteína del HSA con el ácido cáprico se ingresa el código 1E7E, el cual se basa en un análisis cristalográfico con resolución de 2.5 angstroms. El sistema consta de diez moléculas del ligando DKA distribuidas alrededor de la estructura de la proteína; se puede observar en la Ilustración 4-3. Utilizando el software VMD se puede notar que el complejo tiene residuos HOH, mientras que analizando el archivo pdb en un editor de texto verificamos que falta ubicar algunos residuos, átomos y heteroátomos. Luego del procesamiento con modeller se generó un modelo con puntuación de 22.833 molpdf y los HOH han sido retirados del complejo. Sin embargo, las moléculas del ligando también se han eliminado por lo que hemos integrado al sistema un solo ligando DKA, manteniendo una proporción de una proteína y un ligando, el resultado obtenido se puede visualizar en la Ilustración 5-3 y se utilizará en las siguientes secciones de la DM.

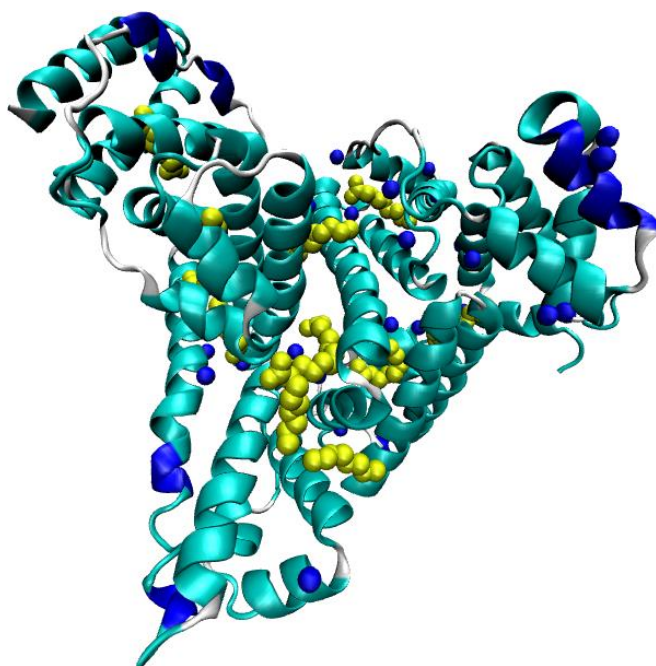


Ilustración 4-3: Modelo de cinta de proteína HSA con diez ligandos DKA (amarillo). Visualización generada con el pdb original.

Realizado por: Chávez, I., 2022.

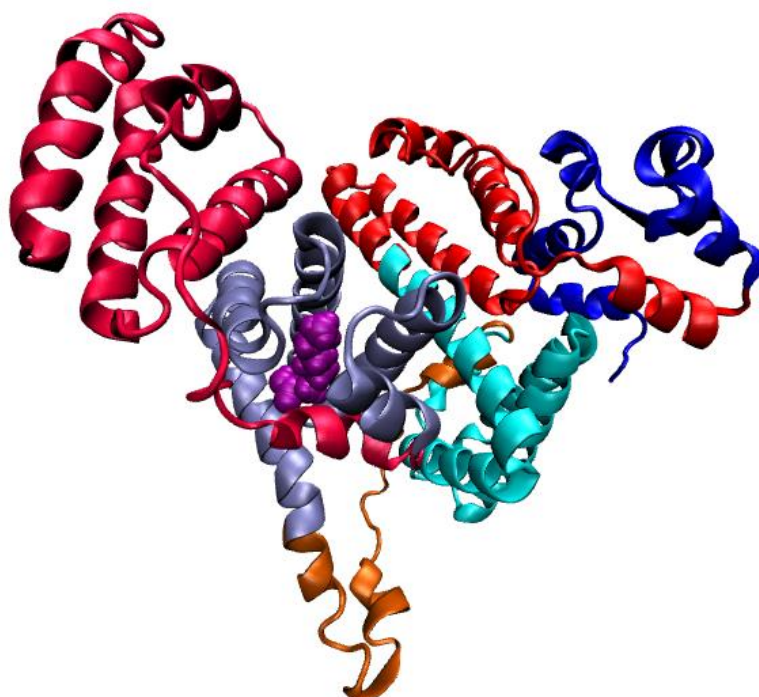


Ilustración 5-3: Modelo de cinta de la proteína HSA con un solo ligando DKA (morado) luego del procesado y reconstrucción del sistema.

Realizado por: Chávez, I., 2022.

El proceso de construcción de la celda, solvatación y neutralización de la carga del sistema se realiza de la misma manera que en la sección anterior. La primera característica por notar es que la carga neta del sistema no varía en nada, sigue siendo de -15, por lo que gromacs reemplazo 15 residuos del solvente por átomos de sodio y neutralizando el sistema. Las características de la celda unitaria para el complejo HSA-DKA formado están resumidos en la Tabla 2-3 y su representación en VMD en la Ilustración 6-3.

Al igual que en el sistema de la sección anterior el centro fue modificado para centralizar la proteína en la celda unitaria. El volumen de la celda se ha reducido un valor considerable lo que indica que la proteína ha modificado su disposición espacial en una estructura un poco más compacta por efecto de la interacción con el ligando. La densidad también ha aumentado su valor y esto puede deberse a que en el sistema tenemos un número mayor de átomos de soluto, ya que el ligando forma parte del grupo de soluto del complejo, además que el tamaño de la celda unitaria tiene menos volumen.

Tabla 2-3: Características de la celda unitaria con la HSA – DKA solvatada y neutralizada.

Coordenadas	x	y	z
Tamaño del sistema (nm)	8,661	5,230	9,132
Centro (nm)	2,514	0,941	2,002
Vectores de la Celda (nm)	8,661	5,230	9,132
Variación (nm)	1,817	1,674	2,564
Nuevo Centro (nm)	4,331	2,615	4,566
Ángulos	90°	90°	90°
Diámetro (nm)	9,271		
Volumen (nm ³)	413,679		
Densidad (g/l)	1059,51		
Átomos del soluto	9529		
Átomos del solvente	32100		

Realizado por: Chávez, I., 2022.

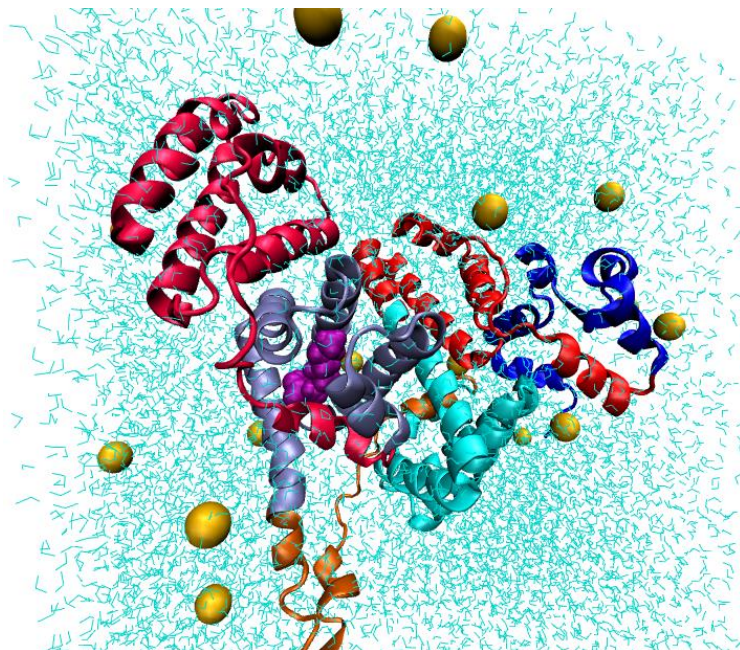


Ilustración 6-3: Modelo de la celda unitaria del sistema HSA - DKA neutralizado y solvatado.

Realizado por: Chávez, I., 2022.

3.1.3. Del complejo HSA-WFR

El complejo ligando-proteína del HSA con la Warfarina se consigue ingresando el código 2bxd en el PDB, el cual está generado a partir de un estudio con difracción de rayos X y una resolución de 3.05 angstroms. El sistema original consta de dos moléculas de la proteína con un ligando de Warfarina cada una, esto se puede observar en la Ilustración 7-3, utilizando el software VMD se puede notar que el complejo no tiene residuos HOH, el pdb del sistema nos indica que faltan ubicar algunos residuos en el esqueleto de la proteína. Se procede con un preprocesado utilizando el mismo algoritmo que en la sección anterior obteniendo un modelo con una puntuación de -41.643 molpdf. En el nuevo pdb, los ligandos han sido retirados del complejo, por lo que debemos volver a integrarlos sin embargo la estructura no se modifica, además ahora tenemos una sola proteína. El resultado obtenido se observa en la Ilustración 8-3 y es la que se utiliza en los pasos siguientes de la DM.

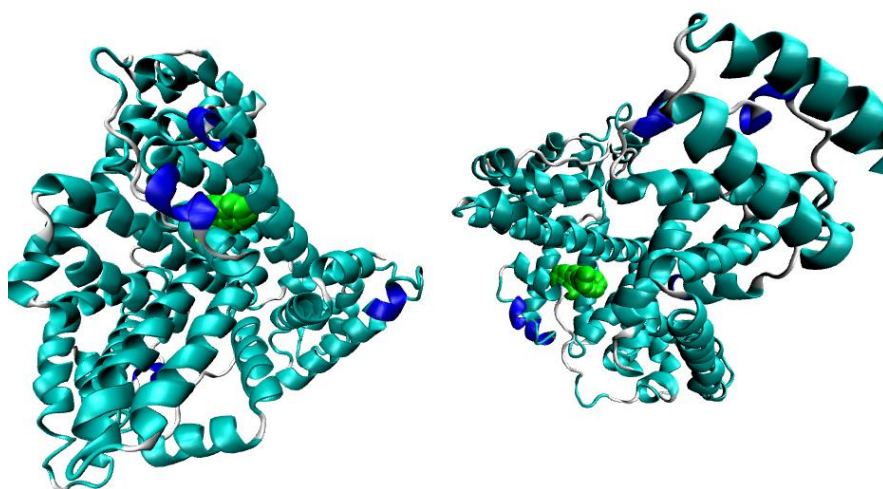


Ilustración 7-3: Visualización generada con el pdb original. Modelo de cinta de dos proteínas HSA con un ligando de WFR (verde) cada una.

Realizado por: Chávez, I., 2022.

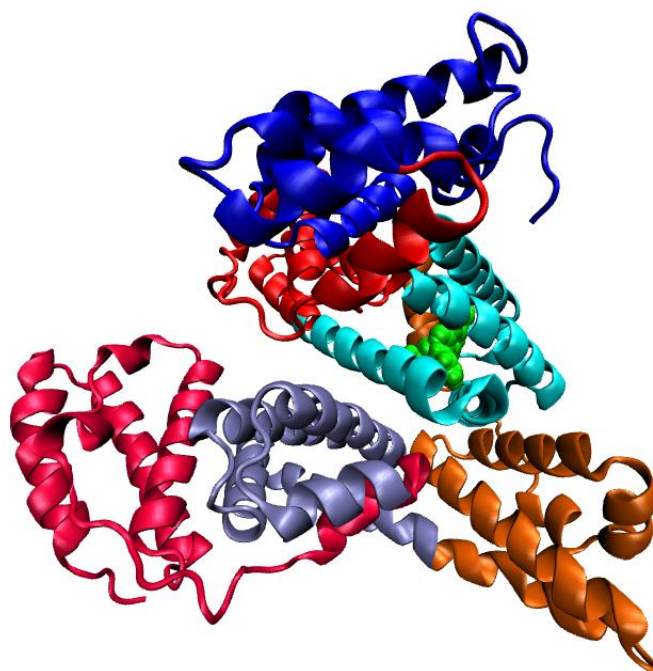


Ilustración 8-3: Modelo de cinta de la proteína HSA con el ligando WFR (verde) luego del procesado y reconstrucción del sistema.

Realizado por: Chávez, I., 2022.

Con el nuevo modelo se construye la celda unitaria y se realiza la solvatación de esta. En el proceso de conseguir una carga neutra del sistema se observó que fueron agregados 15 protones de sodio en posiciones aleatorias. Esto indica que la inclusión de los ligandos en la

HSA no modifica la carga neta del sistema, ya que en los tres se han agregado la misma cantidad de protones. Las características del complejo HSA-WFR formado luego de la neutralización están resumidos en la Tabla 3-3 y su representación en VMD en la Ilustración 9-3. Los datos indican que el volumen y la densidad tiene valores similares a los observados en el complejo HSA-DKA a pesar de que existan más átomos de soluto, debido a que la molécula del ligando WFR tiene mayor número de átomos que el ácido graso DKA.

Tabla 3-3: Características de la celda unitaria con la HSA – WFR solvatada y neutralizada.

Coordenadas	x	y	z
Tamaño del sistema (nm)	6,248	9,171	7,215
Centro (nm)	-0,020	0,101	0,057
Vectores de la Celda (nm)	6,248	9,171	7,216
Variación (nm)	3,144	4,485	3,551
Nuevo Centro (nm)	3,124	4,585	3,608
Ángulos	90°	90°	90°
Diámetro (nm)	9,263		
Volumen (nm ³)	413,41		
Densidad (g/l)	1061,47		
Átomos del soluto	10720		
Átomos del solvente	30678		

Realizado por: Chávez, I., 2022.

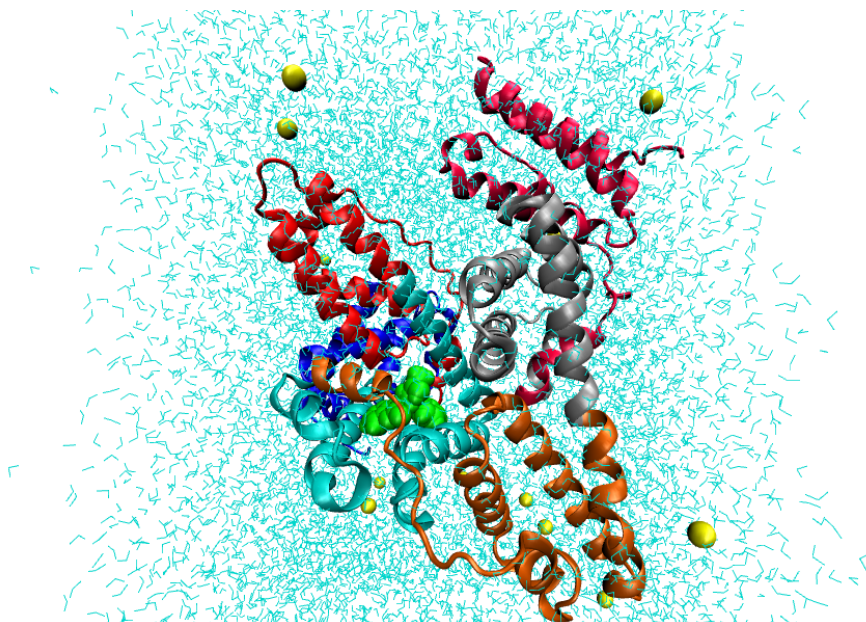


Ilustración 9-3: Modelo de la celda unitaria del sistema HSA - WFR (verde) neutralizado y solvatado.

Realizado por: Chávez, I., 2022.

3.2. Mediciones físicas de las muestras

Luego de los resultados de las estructuras de cada muestra se puede trabajar las fases de minimización de energía y ensambles, los cuales ofrecen nuevas características de como los complejos se configuran según sus características físicas. Durante cada fase existen medidas físicas estadísticas simples que pueden ser extraídas, analizadas y comparadas entre sí. A continuación, se presentan las distintas medidas físicas durante la preparación de la DM.

3.2.1. Minimización de Energía

En la fase de minimización de energía potencial, cada sistema converge a una energía mínima, evitando fallos estructurales ocasionados por la aleatoriedad en las posiciones de los átomos de hidrógeno. A continuación, se representa el proceso de minimización de cada sistema modelado con respecto a la ejecución de cada paso de simulación.

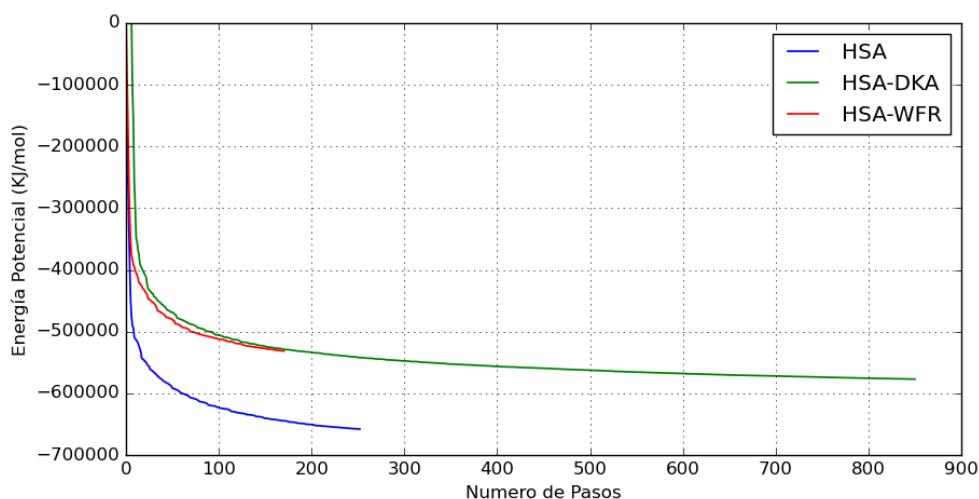


Ilustración 10-3: Energía potencial vs Número de paso del método de descenso más pronunciado, durante la minimización de energía de cada sistema

Realizado por: Chávez, I., 2022.

Para los tres sistemas se utilizó el mismo integrador y se observa que tienen energías negativas, esto puede deberse a que la carga neta de todos fue negativa, tal y como fue analizado en la sección anterior.

La energía del sistema HSA sin ligandos converge a la energía establecida en el paso 253, lo que fue relativamente rápido. Sin embargo, el sistema que converge con más velocidad fue el de HSA-WFR ya que se detuvo en el paso 171, casi la mitad de los pasos que el HSA. Y

finalmente, la convergencia de energía del HSA-DKA demuestra ser la que más pasos requiere, siendo incluso tres veces más que la HSA con un total de 851 pasos.

La energía potencial obtenida en la HSA resulta ser la menor de los tres sistemas con un valor de $-6,581 \times 10^5$ (el orden de 10^5 concuerda con la teoría), el valor de fuerza máxima reportada en el sistema es de $9,577 \times 10^2$ con una norma de 47,15. Seguida por el complejo HSA-DKA energía potencial de $-5,770 \times 10^5$, la fuerza máxima fue de $9,734 \times 10^2$ y una norma igual a 61,99. Finalmente, la que mayor energía potencial reporta es el complejo HSA-WFR, con $-5,316 \times 10^5$, una fuerza máxima de $9,734 \times 10^2$ y una norma igual a 61,99.

3.2.2. *Equilibrio del sistema*

En la fase de equilibrio del sistema primero se realiza el ensamble NVT con los tres complejos, donde se mantiene constante el volumen y la temperatura, esta última se asigna manualmente, en este caso fue de 300 K.

Se observa que durante el ensamble los sistemas presentan oscilaciones entre niveles de temperatura más alto y bajo que 300K, esto se debe al movimiento oscilatorio de los átomos que realizan para estabilizar el sistema, de tal manera que se calientan y enfrían para mantener una temperatura constante.

El comportamiento de los picos cuando la proteína esta sin ligando se asemeja más al complejo HSA-DKA, esto se puede observar de mejor manera en la ilustración 10-3 B) ya que mientras avanza el tiempo se puede observar cambios de temperaturas similares. Cuando nos enfocamos en un punto de tiempo donde se presenta un pico de la proteína HSA, por ejemplo, en $t=20\text{ps}$, se observa que en el mismo punto existe un pico en la HSA-DKA y con la misma dirección, mientras que en el complejo HSA-WFR el pico se ubica al lado contrario. Al pasar los 40 ps de simulación, se tiene un cambio en el comportamiento de los sistemas, puesto que el HSA y el HSA-DKA empiezan a reducir sus picos de temperatura lo que indica una mejor estabilidad de temperatura.

Una mejor estabilización de la temperatura de los sistemas durante el ensamble se produjo mediante el cambio del comportamiento de ambos sistemas. El HSA a partir de los 40 ps de simulación cambia la dirección de sus oscilaciones y luego de esto se estabiliza. El HSA-DKA realiza este proceso a los 20 ps de simulación. El HSA-WFR no cambio mucho la

manera en que avanza la temperatura con el tiempo puesto que desde el inicio realizo el cambio de oscilación para llegar a niveles estables de temperatura.

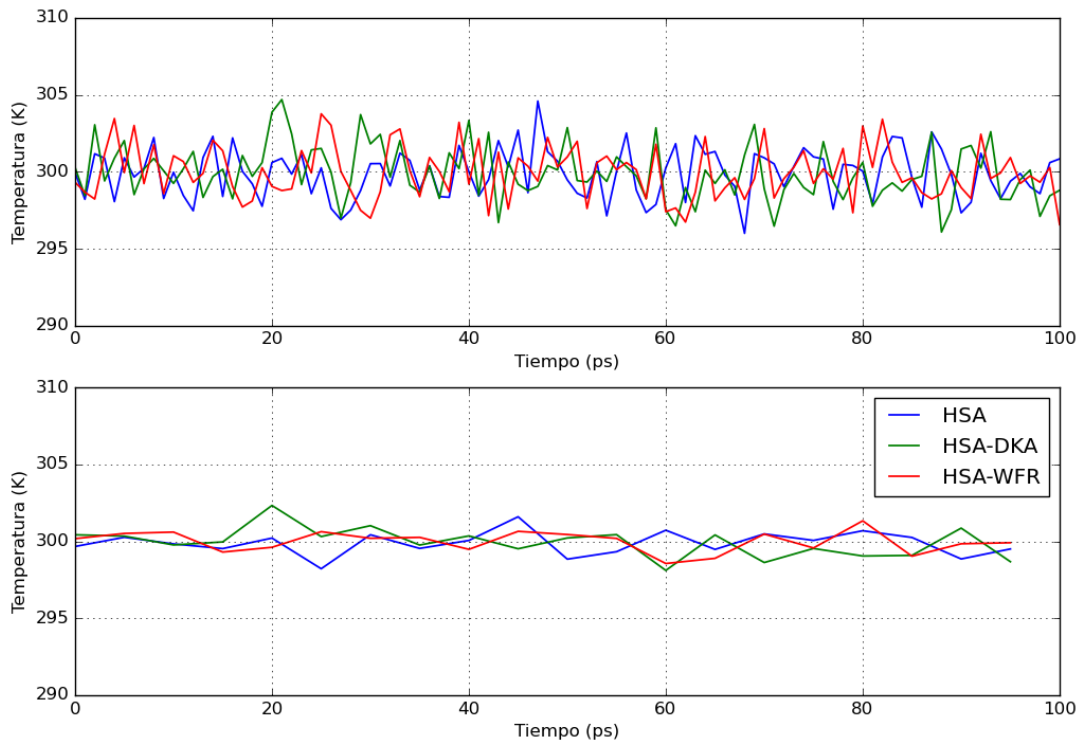


Ilustración 11-3: Temperatura vs tiempo de los tres sistemas luego del ensamble nvt. A) Datos crudos obtenidos y B) La temperatura promedio cada 5 segundos.

Realizado por: Chávez, I., 2022.

La segunda fase de equilibrio del sistema incluye mantener estable la presión y la temperatura mientras que el volumen puede variar para adaptarse a las características físicas del sistema. Como se requiere que los complejos tengan condiciones normales entonces se ajusta el equilibrio con un valor de 1 bar de presión. En la ilustración 11-3 tenemos en la imagen A) como evoluciona el ensamble en el tiempo, se puede observar que inicialmente existe una presión negativa para e HSA-DKA, sin embargo, en promedio la presión resulta ser positiva. De igual manera que en el análisis anterior de la temperatura, tenemos oscilaciones de presión que van entre niveles superiores e inferiores de un nivel establecido, en este caso de 1 bar, pero con amplitudes bastante grandes de -500 a 500 bares.

En promedio, las presiones si se mantienen en los niveles esperados, puesto que obteniendo una media cada 5 segundos de simulación se tiene que la presión de los tres sistemas oscila con amplitudes inferiores y cercanos a 1 bar, tal y como se observa en la ilustración 11-3 B). A diferencia que, en la temperatura, no podemos encontrar una similitud en el comportamiento de estos sistemas complejos, cada uno tiene periodos distintos donde alterna

sus niveles de presión de valores positivos a negativos. Lo que sí se puede diferenciar es que el HSA-DKA resulta tener mayor amplitud en sus oscilaciones, lo que concuerda con los análisis anteriores y sigue siendo el complejo menos estable de los tres propuestos.

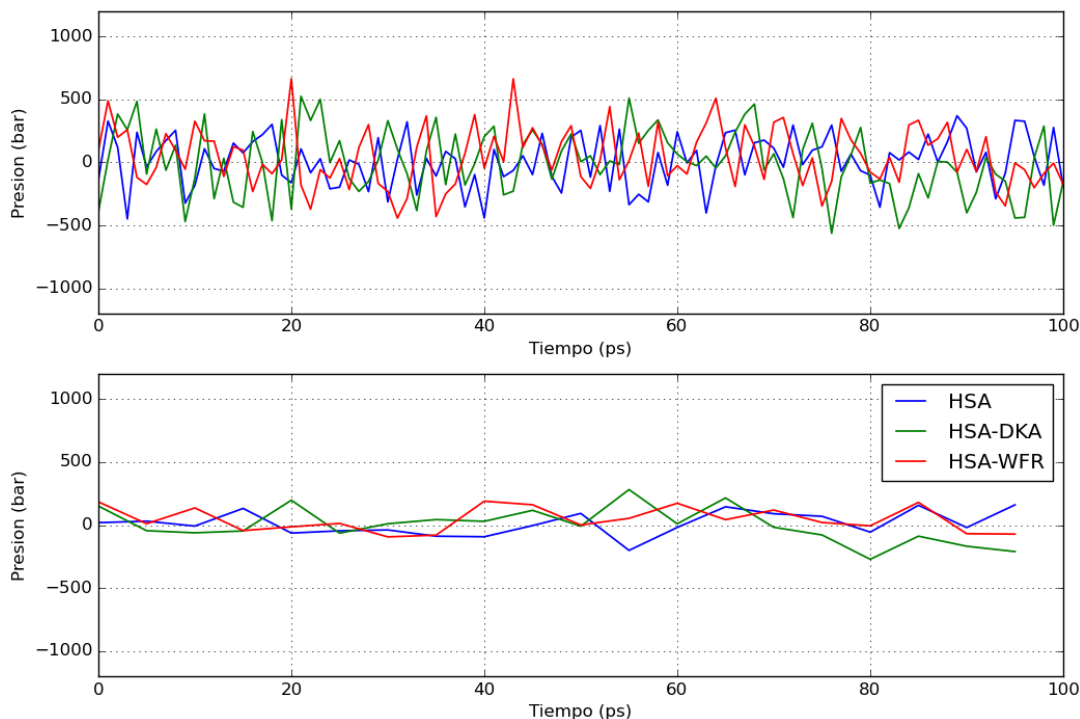


Ilustración 12-3: Presión vs tiempo de los tres sistemas luego del ensamble npt. A) Datos crudos obtenidos y B) Presión promedio cada 5 segundos.

Realizado por: Chávez, I., 2022.

En el último ensamble los volúmenes no son constantes ya que los sistemas se deforman para mantener el equilibrio de las condiciones. Esto puede observarse en como la densidad de cada complejo se altera con el tiempo, puesto que este parámetro físico es indirectamente proporcional al volumen de la celda unitaria (ilustración 12-3 A). Hemos observado que la celda con mayor volumen fue la de la HSA por lo que en teoría debe tener la menor densidad de los tres. Por otro lado, el volumen de los complejos proteína-ligando analizados en el presente trabajo mantuvieron la proporción de una proteína y un ligando, lo que a su vez resultó en una similitud en el volumen de sus celdas unitarias.

El ligando WFR tiene un mayor número de átomos que el ligando DKA, esto incluso se puede notar comparando el valor de átomos de soluto en las tablas 2-3 y 3-3, esto influye en la densidad puesto que indica una mayor cantidad de materia el cual es proporcional a la densidad, por esta razón en la ilustración 12-3 B) se puede notar claramente que el sistema

de mayor densidad es el del HSA-WFR, seguida del HSA-DKA y finalmente el HSA sin ligandos.

Se debe aclarar que en el tiempo la masa de los complejos es constante y como los tres sistemas fueron antes equilibrados con un ensamble nvt, entonces las oscilaciones de densidad no se alteran en gran medida, de hecho, podemos observar que es una medida física muy estable durante todo el dominio de tiempo.

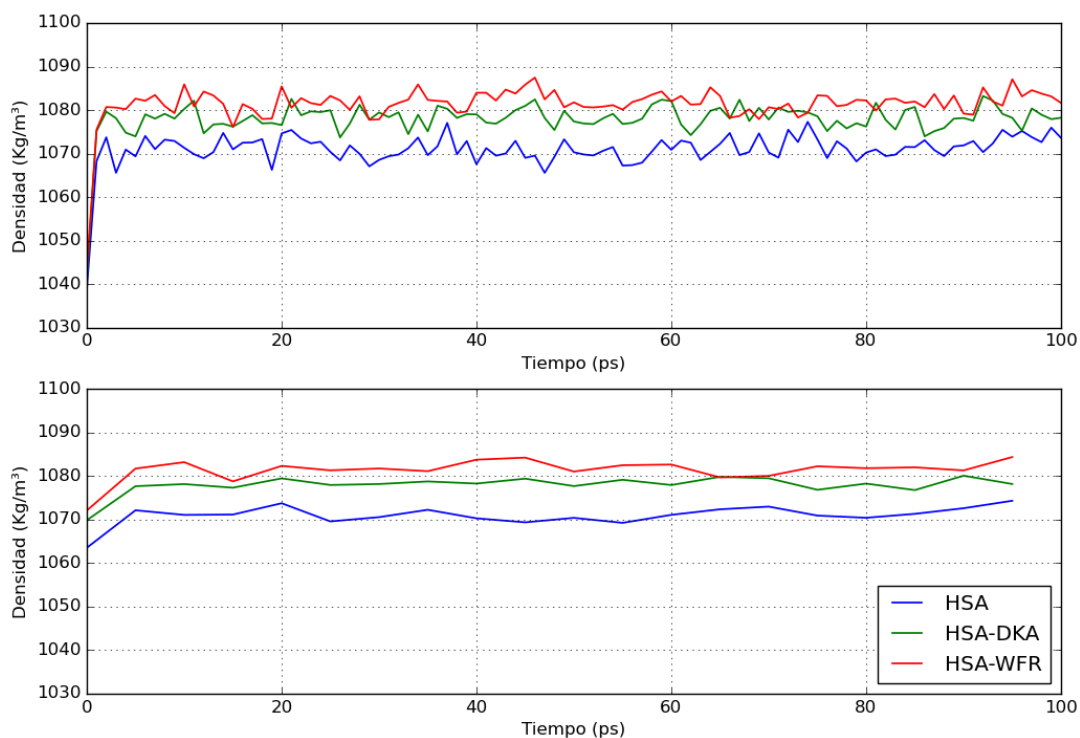


Ilustración 13-3: Densidad vs tiempo de los tres sistemas luego del ensamble npt. A) Datos crudos obtenidos y B) Densidad promedio cada 5 segundos.

Realizado por: Chávez, I., 2022.

Los valores físicos obtenidos en las fases de minimización de energía, ensamble nvt y ensamble npt de los tres complejos proteína-ligando que se presentaron en las gráficas de esta sección han sido promediados. Los resultados obtenidos juntos a sus errores se resumen en la siguiente tabla, lo que permite corroborar algunas observaciones discutidas con anterioridad.

Tabla 4-3: Resumen de las medidas físicas obtenidas en la fase de Minimización de energía, ensamble nvt y npt.

	HSA	HSA-DKA	HSA-WFR
Átomos	48563	41629	41398
Átomos de solvente	39330	32100	10720
Átomos del soluto	9233	9529	30678
Energía Potencial ($\times 10^5$ KJ)	-6.581	-5.77	-5.316
Numero de Pasos	253	851	171
Temperatura (K)	299.89 ± 4.29	299.93 ± 4.30	299.94 ± 3.59
Presión (bar)	11.253 ± 409	1.4659 ± 544	43.394 ± 553
Densidad (Kg/m^3)	1070.9 ± 19.2	1077.9 ± 20.9	1081.4 ± 22.1

Realizado por: Chávez, I., 2022.

3.3. Cambios conformacionales

Para analizar cómo cambia la estructura de cada complejo se analiza los valores de las trayectorias durante la ejecución de la simulación. La DM molecular fue simulada durante diez nanosegundos, luego con los datos almacenados y procesados por gromacs se ejecutan comandos para el cálculo de las mediciones. A continuación, se presenta los análisis estructurales y de trayectorias ejecutados en el presente trabajo.

El RMSD provee una medida directa de la desviación estructural de las coordenadas iniciales, así como las fluctuaciones atómicas en el transcurso de la simulación de DM. En todos los análisis de los complejos los valores de rmsd están en nanómetros y el tiempo en picosegundos. Como se vio en la sección 1.4, en la ecuación 48, la definición para el cálculo de los valores rmsd requiere de una estructura de referencia y una de estudio.

En primera instancia se ha calculado como la estructura del complejo se desvía del esqueleto inicial de todo el sistema (*backbone-system*), esto incluye a las moléculas del solvente, los protones agregados, la proteína y el ligando.

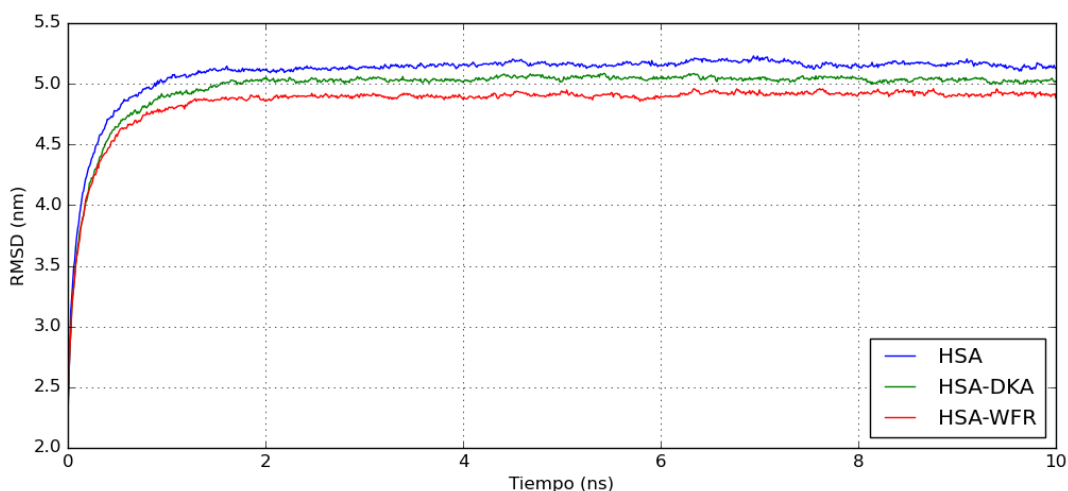


Ilustración 14-3: Evolución del RMSD de los tres sistemas estudiados durante 10 ns de simulación DM. La estructura de referencia y de estudio es esqueleto total del sistema.

Realizado por: Chávez, I., 2022.

En la ilustración 13-3 se observa que todos los sistemas parten desde el mismo valor de desviación de cero nanómetros, esto se debe a que la referencia y el estudio incluye todas las moléculas junto a sus posiciones iniciales. En $t=0$ ns no va a existir una desviación ya que apenas la simulación se está ejecutando por lo que todos tienen un incremento inicial del rmsd hasta que llega a un valor de estabilización luego de 1 ns. Desde este punto se puede notar que los tres complejos tienen un comportamiento similar de fluctuación no muy pronunciado, sin embargo, los sistemas con ligando tienen una menor desviación del rmsd. Un valor menor de la desviación indica que estos sistemas tienden a ser estables. Esto indica que la proteína HSA puede mantener su estabilidad durante su interacción con los ligandos, favoreciendo el transporte de estos. Se puede inferir que el complejo con menor desviación es el que tiene una estructura más compacta puesto que sus distancias son menores, en este caso el sistema HSA-WFR.

Para establecer como en la proteína cambian de posiciones sus átomos durante la ejecución de la DM, se establece como estructura de estudio a la HSA en cada uno de los tres sistemas, comparándolo con el esqueleto inicial (*backbone-protein*).

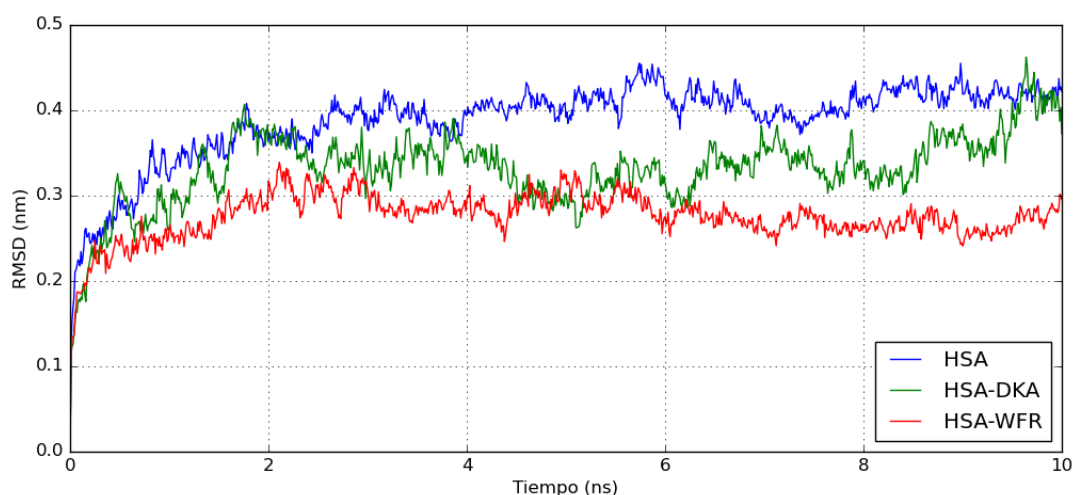


Ilustración 15-3: Evolución temporal del RMSD de los tres sistemas estudiados durante 10 ns de simulación. El sistema es la referencia y la estructura de estudio la proteína.

Realizado por: Chávez, I., 2022.

Al igual que en anterior análisis se puede observar que las tres curvas parten de un mismo valor, puesto que inicialmente las proteínas aún no interactúan con la simulación de DM, por lo que aún no existe trayectorias de las moléculas. La HSA tiene una desviación inicial de aproximadamente 0.1 nm en los tres sistemas, luego existe un ascenso de la rmsd durante los primeros 1.9 nanosegundos de simulación, hasta llegar a valores de estabilidad y existen fluctuaciones de distancia. El complejo HSA se mantiene en un incremento del rmsd hasta los 2.5 ns aproximadamente donde estabiliza la posición de sus moléculas alrededor de 0.4 nanómetros, luego no presenta una mayor fluctuación. De la misma manera, la proteína del complejo HSA-WFR tiene poca variación en sus fluctuaciones luego de haber llegado a una medida de rmsd estable, en aproximadamente 0.3 nanómetros, se puede observar que no afecto en gran medida la estabilidad de la estructura, pero si la volvió más compacta comparado con la HSA sin ligando. Por otro lado, la HSA-DKA es el complejo que más fluctúa la posición de la proteína durante la simulación, esto impide que podamos inferir un valor promedio rmsd solo observando la ilustración 14-3, ya que existen picos bastante pronunciados, por ejemplo, en $t=5$ ns se tiene valores de rmsd cercanos a 0.28 nm, luego existe un incremento hasta 0.48 nm, donde vuelve a descender.

Con estas observaciones, se nota que es necesario saber que sucede con los ligandos en cada complejo para poder establecer cómo afectan a la proteína. Se ha realizado un último análisis de las desviaciones de distancia ahora estableciendo como estructura de estudio a los ligandos y la referencia es la proteína HSA (*protein-ligand*).

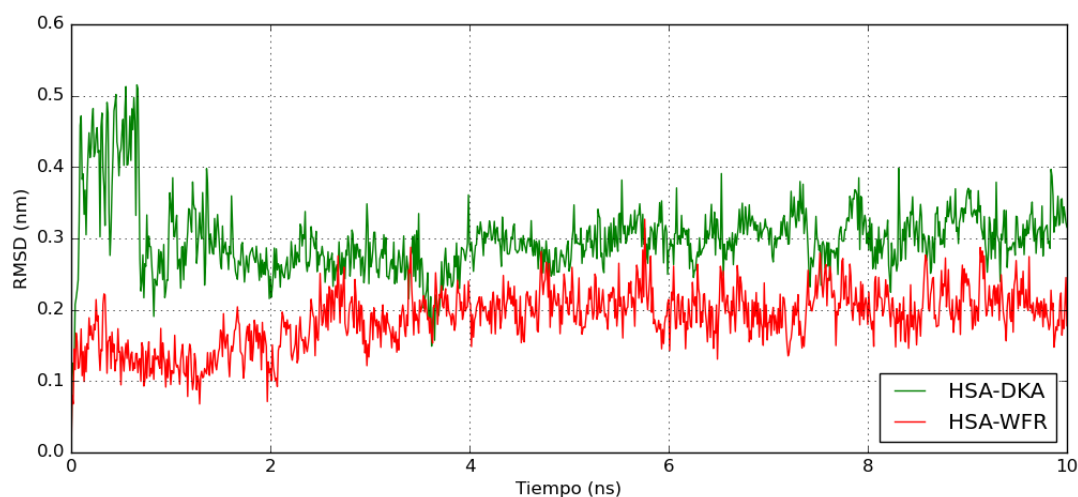


Ilustración 16-3: Evolución del RMSD de los complejos proteína-ligando, con 10 ns de simulación. La HSA es la referencia y la estructura de estudio el ligando.

Realizado por: Chávez, I., 2022.

En la ilustración 15-3, se puede notar que durante el primer nanosegundo de simulación el ligando DKA tiene un comportamiento extraño con respecto de la proteína, ya que inicialmente sus átomos tienen un promedio de 0.1 nm con respecto a los átomos de la HSA, luego existe un salto abrupto de hasta 0.5 nm en el rmsd. Aproximadamente en 1.5 ns de simulación el ligando se estabiliza en un valor de 0.3 nm y sus fluctuaciones no son tan elevadas como en los momentos iniciales.

Un caso muy diferente se observa con el complejo HSA-WFR ya que desde el inicio de la simulación hasta el final de esta no existen picos elevados de rmsd, aunque se puede notar que utilizó el mismo tiempo que el DKA en establecer su posición. Las fluctuaciones en sus valores de rmsd nos son elevadas, alrededor de 0.2 nm de distancia entre la proteína y el ligando, lo que en la práctica se traduce en que el WFR se compacta mejor con la HSA.

En segunda instancia, luego de las observaciones del RMSD, se analizan los valores del radio de giro de la proteína en cada uno de los sistemas complejos. Se puede observar en la ilustración 16-3 que los valores de la proteína HSA y el complejo HSA-WFR no tienen mayor diferencia, por lo que no existe un mayor cambio en la flexibilidad de la estructura. Lo contrario ocurre cuando se analizan los cambios del Rg en el tiempo con el complejo HSA-DKA, donde se puede notar que el radio de giro inicial es mayor y durante toda la simulación mantiene valores mayores al de los otros sistemas. En los 10 ns que dura la simulación el radio de giro incrementa su valor hasta llegar a niveles de 2.95 nanómetros.

Una vez más se puede observar que la Warfarina no afecta en el comportamiento ni en la estructura de la proteína HSA en mayor medida, puesto que se puede inferir que la flexibilidad se mantiene luego de haberse formado el complejo HSA-WFR.

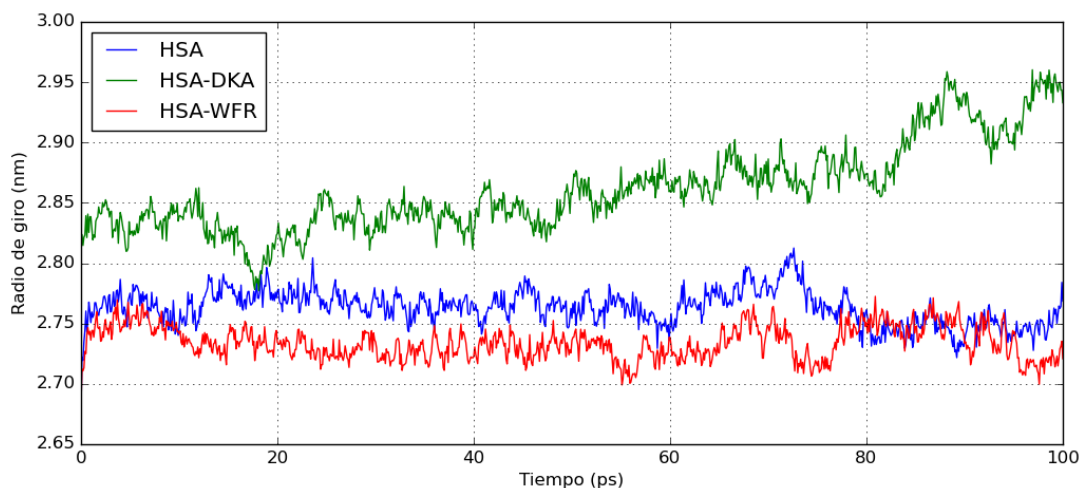


Ilustración 17-3: Evolución temporal del Rg de la HSA en los tres sistemas estudiados durante 10 ns de simulación.

Realizado por: Chávez, I., 2022.

En este punto, se requiere observar la flexibilidad del ligando durante la DM, para observar si el cambio de flexibilidad del complejo se debe a características intrínsecas del ligando o depende de cómo la HSA se acopló con cada ligando. Por esta razón, en la siguiente etapa se analiza el radio de giro de cada ligando. En la ilustración 17-3 se muestra el resultado del radio de giro de la DKA y el WFR en cada sistema complejo. Se logra observar que la flexibilidad del ligando ácido graso tiene mayor variación durante la simulación que el WFR, el cual parte de un radio de giro de 0.35 nm sin grandes fluctuaciones. Por otro lado, el ácido decanoico tiene una variación relativamente alta con respecto al otro ligando, donde su nivel más bajo fue de 0.32 nm y el más alto de hasta 0.41 nm, lo que implica que este ligando es mucho más flexible.

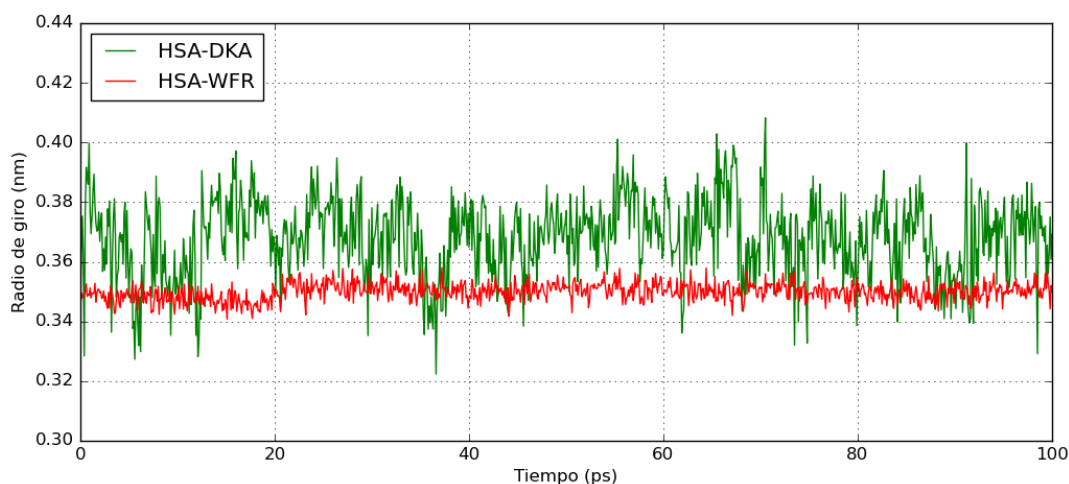


Ilustración 18-3: Evolución temporal del Rg del ligando en los complejos proteína-ligando estudiados durante 10 ns de simulación.

Realizado por: Chávez, I., 2022.

Usando los datos con los que se realizaron las gráficas de esta sección se calcula un promedio para cada medida física en los sistemas de estudio. Los cálculos fueron hechos utilizando programación en Python, lo que permitió obtener valores reales junto a sus errores de medición. En la siguiente tabla se resumen los resultados obtenidos.

Tabla 5-3: Promedios de las medidas analizadas en luego de la simulación de DM, para detectar cambios conformacionales.

	HSA	HSA-DKA	HSA-WFR
RMSD (nm)			
Backbone-Sistema	5.1168 ± 0.56	4.9957 ± 0.59	4.8731 ± 0.54
Backbone-Proteína	0.3884 ± 0.23	0.3318 ± 0.23	0.2754 ± 0.17
Proteína-Ligando		0.2993 ± 0.26	0.1878 ± 0.16
Rg (nm)			
Proteína	2.7631 ± 0.07	2.8610 ± 0.09	2.7341 ± 0.04
Ligando		0.3679 ± 0.04	0.3499 ± 0.01

Realizado por: Chávez, I., 2022.

La Tabla 5-3 permite reafirmar algunas observaciones antes expuestas sobre la interacción y cambios conformacionales de la proteína HSA en presencia de los ligandos. Los valores del RMSD y el radio de giro para los átomos que conforman el sistema de la HSA sin ligandos tienen un valor promedio muy similar al obtenido por Fujiwara y Amisaki (2006b: pp. 732-733), los cuales fueron de $0.3690 \pm 0.035nm$ y $2.720 \pm 0.023nm$. De igual manera, el radio de giro de la HSA en una solución acuosa fue medido experimentalmente por Kiselev y et al. (2001, p. 426) reportando un valor de $2.740 \pm 0.035nm$. La similitud de los resultados

obtenidos con los datos de trabajos anteriores indica que la simulación de DM presentado en este trabajo ha reproducido de buena manera el comportamiento de la HSA en una solución acuosa.

Los valores del RMSD en los sistemas con ligando son menores comparando con el sistema sin ligando, esto puede ser ocasionado debido a que la proteína tiende a adaptarse para que sus sitios de enlace se ubiquen en la posición adecuada al ligando, lo que indica esta variación de distancia promedio es que el complejo se está volviendo más compacto. Lo que a su vez concuerda con los datos obtenidos en la sección 3.1. sobre el volumen de las distintas celdas unitarias, donde se mostró un valor mayor cuando la proteína no tenía ningún ligando.

Se puede observar que rmsd de la proteína al ligando es menor en el WFR, esto indica que el enlace intermolecular entre la proteína y este ligando tiene una menor distancia, esto tiene sentido con lo que hemos mencionado antes sobre como la HSA se compacta mejor con el WFR, mientras que valor con el DKA es mayor indicando una menor conformación relativamente menor compacta.

Con lo que respecta al radio de giro se puede observar que los valores del HSA sin ligando y el HSA-WFR con prácticamente iguales, mientras que el HSA-DKA tiene niveles más altos y la fluctuación de su valor es superior al resto; esto se puede notar en el error. Lo que, en la práctica, indica que el DKA cuando interacciona con la HSA le brinda una mayor flexibilidad conformacional, mientras que el WFR no cambia esta propiedad de la proteína.

CONCLUSIONES

- En este proyecto se modelaron estructuras de proteína - ligando. La bibliografía investigada determinó que el campo de fuerza más aplicable para modelar la proteína es CHARMM36m y para los ligandos se utilizó la herramienta CGenFF, luego para cada sistema se analizaron diversos parámetros desde el punto de vista estructural y de igual manera se obtuvieron medidas para observar los cambios conformacionales tales como el RMSD y Radio de Giro.
- Los archivos necesarios para realizar el proceso de dinámica molecular fueron preprocesados a través de un algoritmo creado con bibliotecas de python especializadas. El algoritmo permitió automatizar el proceso de complementación de aminoácidos en la cadena a través de mutagénesis y demostró que existen varias posibles disposiciones espaciales para cada sistema. El mejor modelo fue escogido a través de la puntuación molpdf y se pudo comprobar la estabilidad del mismo con la visualización del complejo.
- En el presente trabajo se ha creado y analizado tres muestras con la proteína HSA y dos distintos ligandos; un ácido graso y un fármaco, mediante herramientas de dinámica molecular. Bajo esta metodología se pudo analizar la conformación de los tres sistemas a través de mediciones físicas y la visualización de los complejos formados. Dando como resultado que la celda unitaria de la proteína sin ligandos tiende a ser menos compacta, por lo que su volumen es mayor al de los otros sistemas. Este valor obedece la teoría de plegamiento de la proteína para compactarse al ligando con el que está actuando.
- Después de tener los tres sistemas formados se realizó varios procesos de dinámica molecular que generaron distintos comportamientos en cada fase. El complejo HSA-DKA durante la minimización de energía y los ensambles fue el menos estable, mostrando oscilaciones con mayor amplitud durante la ejecución de las simulaciones. El WFR no mostró un cambio sustancial en el comportamiento de la molécula HSA durante estas fases, lo que en resumen permite aseverar que la proteína tiene una mejor afinidad para interactuar con ligandos de tipo fármaco durante procesos con temperatura y presión constante.
- Las ilustraciones con los resultados del RMSD y radio de giro con respecto al tiempo de la proteína HSA sin ligandos dio valores similares a estudios preexistentes los cuales fueron citados en la sección de resultados, esto nos permite concluir que la metodología utilizada en el presente trabajo de integración logró simular el comportamiento de dinámica molecular del sistema proteína-ligando.

- Se representaron los datos de RMSD en función del tiempo de simulación durante 10 nanosegundos. Lo que nos permitió observar el comportamiento de las trayectorias de cada ligando con respecto a la HSA durante la dinámica molecular. EL HSA sin ligandos tiende a cambiar más sus posiciones iniciales llegando a valores cercanos a 0.4 nanómetros de diferencia, mientras que el WFR hace que los valores disminuyan, además las variaciones de este complejo no son muy grandes. Por otro lado, cuando el HSA está interactuando con el ligando DKA se observan oscilaciones pronunciadas y a su vez el radio de giro es mayor, esto permite concluir que el ligando ácido graso le da más movilidad y elasticidad a la proteína.

RECOMENDACIONES

- La metodología utilizada para este estudio puede ser extendida para otros ligandos de diferente tipo que tienen interacción con la proteína HSA. Lo que permitiría dar un mejor conocimiento sobre el transporte de sustancias a través del torrente sanguíneo, puesto que la HSA es la proteína más abundante en el plasma sanguíneo.
- Una extensión del presente estudio se puede aplicar a través de docking molecular, lo que permite realizar una simulación con la proteína HSA y ambos ligandos estudiados en el presente trabajo de integración curricular. Este nuevo medio de dinámica molecular permitiría analizar la competencia que existe entre estos ligandos cuando ocupan los sitios de enlace.
- Aunque la metodología del trabajo fue revisada por la bibliografía investigada, aun se puede realizar una validación del campo de fuerza más extenso para estructuras no convencionales por ejemplo ADN o ARN lo que implica un análisis más amplio se recomienda evaluar: Parámetros de la rotación por residuo, las longitudes de los enlaces de hidrógeno o la inclinación por par de bases.

GLOSARIO

Algoritmo: Secuencia finita de pasos lógicos y ordenados que se deben seguir para resolver un problema o realizar una actividad. Pueden representarse a través de pseudocódigo, diagramas de flujo, entre otros (Mejía, 2015, párr. 3).

Canónico: En Física o matemática se refiere en el sentido de objetivo, universal, no dependiente del sujeto, e intrínseco al objeto. También describe los estados de un sistema macroscópico que tiene determinado su número de partículas, volumen y temperatura (Británica the editors of encyclopaedia, 2007: párr. 1).

Cristalografía: Rama de la ciencia que se ocupa de discernir la disposición y el enlace de los átomos en sólidos cristalinos y con la estructura geométrica de las redes cristalinas. Clásicamente, las propiedades ópticas de los cristales tenían valor en mineralogía y química para la identificación de sustancias (Britannica the editors of encyclopaedia, 2018: párr. 1).

Entropía: Es un concepto matemático que representa de manera precisa si la segunda ley de la termodinámica es violada por un proceso en particular. Por definición, la entropía depende de la cantidad de calor que fluye en un sistema con una temperatura constante (Gordon, 2022a: párr. 1).

Entalpía: Es la suma de la energía interna y el producto de la presión y el volumen del sistema termodinámico estudiado. Es una función de estado determinado enteramente por la temperatura, presión y la composición; no por la evolución del sistema. Si solamente el trabajo realizado produce un cambio de volumen a presión constante, entonces la variación de la entalpía es exactamente igual al calor transferido al sistema (Britannica the editors of encyclopaedia, 2020, párr. 1-3).

Hamiltoniano: La función hamiltoniana es una definición matemática para expresar el cambio de las condiciones físicas del sistema con respecto al tiempo. El hamiltoniano de un sistema especifica la energía total del mismo (suma de la energía cinética y potencial) en términos de la función lagrangiana derivada en estudios previos, la posición y el momento de cada una de las partículas (Britannica the editors of encyclopaedia, 2012, párr. 1).

Lagrangiano: es una función que caracteriza el estado de un sistema físico. En mecánica, el lagrangiano es solamente la energía cinética menos la energía potencial (Britannica the editors of encyclopaedia, 2016, párr. 1).

Mutagénesis: Se le puede considerar una mutación, lo cual se refiere a una alteración en el material genético. El genoma está compuesto por una larga cadena de moléculas de ADN y la mutación puede ocurrir en cualquiera de ellas. Un gen está compuesto por una región regulatoria, la cual es responsable de la transcripción apropiada del código genético de una molécula funcional, generalmente una proteína (Griffiths, 2022, párr. 1-2).

Nucleación: El proceso inicial que ocurre en la formación de un cristal a partir de una solución, un líquido o un vapor, en el que una pequeña cantidad de iones, átomos o moléculas se organizan

en un patrón característico de un sólido cristalino, formando un sitio sobre el cual se depositan partículas adicionales a medida que crece el cristal (Britannica the editors of encyclopaedia, 2020, párr. 1).

Plasma sanguíneo: Es la sustancia donde están suspendidas las células, oxígeno y nutrientes que son transportados por el torrente sanguíneo (Schwartz & Conley, 2022, párr. 1-2).

Sistema aislado: Se refiere cuando el conjunto de propiedades del sistema físico estudiado no tiene intercambio de energía ni de materia con el exterior, por eso también se le conoce como sistema cerrado (Gordon, 2022b: párr. 8).

Soluto: En química, el soluto es la sustancia que esta suelta a una solución. Si ambos componentes son líquidos se le puede distinguir por el que tiene menos concentración (Britannica the editors of encyclopaedia, 2022, párr. 4).

Solvente: En química, el solvente es el líquido que conforma la solución. Si ambos componentes son líquidos se le puede distinguir por el que tiene mayor concentración (Britannica the editors of encyclopaedia, 2022, párr. 4).

Viscosidad: Propiedad física de un fluido que mide su resistencia a la deformación, la relación entre el esfuerzo cortante y la velocidad del flujo de cada fluido determina un coeficiente de viscosidad, los fluidos que cumplen esta relación se conocen como fluidos newtonianos (Cengel & Cimbala, 2012, p. 50-52).

BIBLIOGRAFÍA

ABRAHAM, J., MURTOLA, T., SCHULZ, R., PÁLL, S., SMITH, J., et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. En línea. *SoftwareX*, vol. 1-2 (septiembre de 2015), pp. 19–25. ISSN 2352-7110. Disponible en: <https://doi.org/10.1016/j.softx.2015.06.001>. [consultado el 19/07/2022].

ALONSO, H., BLIZNYUK A. & GREADY J. Combining docking and molecular dynamic simulations in drug design. *ChemInform*, vol. 37 (noviembre de 2006), n. ° 45. ISSN 1522-2667.

ANDERSEN, C. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, vol. 72 (febrero de 1980), n. ° 4, pp. 2384–2393. ISSN 1089-7690.

ARTALI, R., BOMBIERI G., CALABI, L. & DEL PRA, A. A molecular dynamics study of human serum albumin binding sites. *Il Farmaco*, vol. 60 (junio de 2005), n.º 6-7, pp. 485–495. ISSN 0014-827X.

AVOGADRO CHEMISTRY. *Avogadro*. En línea. avogadro.cc. [s. f.]. Disponible en: <https://avogadro.cc/>. [consultado el 29/05/2022].

BARRAGAN, D. Evaluación del campo de fuerza CHARMM27 implementado en NAMD para la simulación de adn g- cuádruple tetramoleculares y paralelos. En línea. Trabajo de titulación. Riobamba: ESPOCH, 2015. Disponible en: <http://dspace.esPOCH.edu.ec/handle/123456789/4808>. [consultado el 05/06/2022].

BERENDSEN, H., POSTMA, J., VAN GUNSTEREN, W., DINOLA, A. & HAAK, J. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, vol. 81 (octubre de 1984), n. ° 8, pp. 3684–3690. ISSN 1089-7690.

BERENDSEN, H. & VAN GUNSTEREN, W. Practical algorithms for dynamic simulations. En: *Molecular-dynamic simulation of statistical mechanical systems*, pp. 43–64. Amsterdam: North-Holland Physics, 1986.

BERMAN, H., WESTBROOK, J., FENG, Z., GILLILAND, Z., BHAT, T. et al. *A structural view of biology*. En línea. RCSB PDB: Homepage. Enero de 2000. Disponible en: <https://www.rcsb.org/>. [consultado el 04/06/2022].

BEST, B., ZHU, X., SHIM, J., LOPES, P., MITTAL, J. et al. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Sidechain χ_1 and χ_2 Dihedral Angles. *Journal of Chemical Theory and Computation*, vol. 8 (agosto de 2012), n. ° 9, pp. 3257–3273. ISSN 1549-9626.

BHATTACHARYA, A., GRÜNE, T. & CURRY, S. Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin 1 edited by R. huber. *Journal of Molecular Biology*, vol. 303 (noviembre de 2000), n. ° 5, pp. 721–732. ISSN 0022-2836.

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Canonical ensemble. En línea. Encyclopedia Britannica. 16/07/2007. Disponible en: <https://www.britannica.com/science/canonical-ensemble>. [consultado el 14/09/2022].

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Crystallography. En línea. Encyclopedia Britannica. 09/05/2018. Disponible en: <https://www.britannica.com/science/crystallography>. [consultado el 14/09/2022].

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Enthalpy. En línea. Encyclopedia Britannica. 09/04/2020. Disponible en: <https://www.britannica.com/science/enthalpy>. [consultado el 14/09/2022].

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Hamiltonian function. En línea. Encyclopedia Britannica. 01/02/2012. Disponible en: <https://www.britannica.com/science/Hamiltonian-function>. [consultado el 14/09/2022].

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Lagrangian function. En línea. Encyclopedia Britannica. 25/10/2016. Disponible en: <https://www.britannica.com/science/Lagrangian-function>. [consultado el 14/09/2022].

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Nucleation. En línea. Encyclopedia Britannica. 11/02/2020. Disponible en: <https://www.britannica.com/science/nucleation>. [consultado el 14/09/2022].

BRITANNICA, THE EDITORS OF ENCYCLOPAEDIA. Solution. En línea. Encyclopedia Britannica. 17/08/2022. Disponible en: <https://www.britannica.com/science/solution-chemistry>. [consultado el 14/09/2022].

BROOKS, B., BRUCCOLERI, R., OLAFSON, B., STATES, D., SWAMINATHAN, S. et al. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, vol. 4 (1983), n. ° 2, pp. 187–217. ISSN 1096-987X.

CEDIA. *Supercomputador.* En línea. cedia.edu.ec. 2022. Disponible en: <https://cedia.edu.ec/infraestructura/supercomputador#descripcion/>. [consultado el 30/05/2022].

CENGEL, Y. y CIMBALA, J. *Mecánica de fluidos. fundamentos y aplicaciones.* 2a ed. Mexico D.F.: McGraw-Hill, 2012. ISBN 978-607-15-0779-2.

DA SILVA, T. *Efeito de interações de longo alcance no sistema macromolecular Ibuprofeno-Albumina do soro humano.* Tesis doctoral. Maceió: UNIVERSIDADE FEDERAL DE ALAGOAS, 2017.

EARLS, A. *What is container image? - definition from whatis.com.* En línea. SearchITOperations. 28/02/2018. Disponible en: [https://www.techtarget.com/searchitoperations/definition/container-image#:~:text=A%20container%20image%20is%20an,information%20technology%20\(IT\)%20infrastructure.](https://www.techtarget.com/searchitoperations/definition/container-image#:~:text=A%20container%20image%20is%20an,information%20technology%20(IT)%20infrastructure.) [consultado el 16/06/2022].

ERCOLESSI, F. A molecular dynamics primer. *Spring college in computational physics, ICTP, Trieste*, 1997, vol. 19.

FEUNANG, Y., WISHART, D., GUO, A., LO, E., MARC, A. et al. *(R)-warfarin uses, interactions, mechanism of action.* En línea. DrugBank Online: Database for Drug and Drug Target Info. Julio de 2020. Disponible en: <https://go.drugbank.com/drugs/DB08496>. [consultado el 19/07/2022].

FU, Y., ZHAO, J. & CHEN, Z. Insights into the molecular mechanisms of protein-ligand interactions by molecular docking and molecular dynamics simulation: a case of oligopeptide binding protein. *Computational and Mathematical Methods in Medicine*, vol. 2018 (diciembre de 2018), pp. 1–12. ISSN 1748-6718.

FUJIWARA, S. & AMISAKI, T. Molecular dynamics study of conformational changes in human serum albumin by binding of fatty acids. *Proteins: Structure, Function, and Bioinformatics*, vol. 64 (junio de 2006), n. ° 3, pp. 730–739. ISSN 0887-3585.

GORDON, D. Entropy. En línea. Encyclopedia Britannica. 25/08/2022. Disponible en: <https://www.britannica.com/science/thermodynamics/Entropy>. [consultado el 14/09/2022].

GRIFFITHS, A. Mutation. En línea. Encyclopedia Britannica. 06/09/2022. Disponible en: <https://www.britannica.com/science/thermodynamics/mutation-genetics>. [consultado el 14/09/2022].

GROMACS DEVELOPMENT TEAM. Long range electrostatics. En línea. GROMACS 2022.2 documentation Reference Manual. 2022. Disponible en: <https://manual.gromacs.org/current/reference-manual/functions/long-range-electrostatics.html>. [consultado el 07/07/2022].

GROMACS DEVELOPMENT TEAM. Long range van der waals interactions. En línea. GROMACS 2019.1 documentation Reference Manual. 2019. Disponible en: <https://manual.gromacs.org/2019.1/reference-manual/functions/long-range-vdw.html>. [consultado el 06/07/2022].

GUGLIELMELLI, A., RIZZUTI, B. & GUZZI, R. Stereoselective and domain-specific effects of ibuprofen on the thermal stability of human serum albumin. *European Journal of Pharmaceutical Sciences*, vol. 112 (enero de 2018), pp. 122–131. ISSN 0928-0987.

GUTMAN, I. & POLANSKY, E. Mathematical concepts in organic chemistry. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986. ISBN 9783642709845.

HANSEN, J. Molecular dynamic simulation of coulomb systems in two and three dimention. En: *Molecular dynamic simulation of statistical mechanical systems*, pp. 89–129. Amsterdam: North-Holland Physics, 1986.

HERNANDEZ, E. Molecular Dynamics: from basic techniques to applications (A Molecular Dynamics Primer). AIP Conference Proceedings, vol. 1077 (2008), n. ° 1, pp. 95–123.

HOLLINGSWORTH, S. & DROR R. Molecular dynamics simulation for all. *Neuron*, vol. 99 (septiembre de 2018), n. ° 6, pp. 1129–1143. ISSN 0896-6273.

HUANG, J., RAUSCHER, S., NAWROCKI, G., RAN, T., FEIG, M., et al. CHARMM36: an improved force field for folded and intrinsically disordered proteins. *Biophysical Journal*, vol. 112 (febrero de 2017), n. ° 3, pp. 175–176. ISSN 0006-3495.

HÜNENBERG, P. Thermostat algorithms for molecular dynamics simulations. *Advances in polymer sciences*, 2005, n. ° 173, pp. 105–149.

IBM. *What is HPC? Introduction to high-performance computing | IBM.* En línea. IBM - Deutschland. [s. f.]. Disponible en: <https://www.ibm.com/topics/hpc>. [consultado el 01/07/2022].

JORGENSEN, W., MAXWELL, S., & TIRADO-RIVES J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, vol. 118 (enero de 1996), n. ° 45, pp. 11225–11236. ISSN 1520-5126.

JORGENSEN, W., CHANDRASEKHAR, J., MADURA, J., IMPEY J. & KLEIN, M. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, vol. 79 (julio de 1983), n. ° 2, pp. 926–935. ISSN 1089-7690.

JUPYTER ORG. *The Jupyter Notebook 7.0.0a5 documentation.* En línea. Jupyter. 2015. Disponible en: <https://jupyter-notebook.readthedocs.io/en/latest/notebook.html>. [consultado el 15/06/2022].

KAMINSKI, G. & JORGENSEN, W. Performance of the AMBER94, MMFF94, and OPLS-AA force fields for modeling organic liquids. *The Journal of Physical Chemistry*, vol. 100 (enero de 1996), n. ° 46, pp. 18010–18013. ISSN 1541-5740.

KISELEV, M., GRYZUNOV, Y., DOBRETSOV, G. & KOMAROVA, M. Size of a human serum albumin molecule in solution. *Biofizika*, vol. 46 (2001), n. ° 3, pp. 423–427.

KLAUDA, J., VENABLE J., FREITES J., O'CONNOR, J., TOBIAS D., et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *The Journal of Physical Chemistry B*, vol. 114 (junio de 2010), n. ° 23, pp. 7830–7843. ISSN 1520-5207.

KLEIN, M. Structure and dynamics of molecular crystals. En: G. CICCOTTI y W. HOOVER (eds.). *Molecular-dynamic simulation of statistical mechanical systems*, pp. 424–476. Amsterdam: North-Holland Physics, 1986.

LI, J. Basic molecular dynamics. En línea. En: *Handbook of materials modeling*, pp. 565–588. Dordrecht: Springer Netherlands, 2005. ISBN 9781402032875. Disponible en: https://doi.org/10.1007/978-1-4020-3286-8_29. [consultado el 12/06/2022].

LUMITOS A. *Ácido graso*. En línea. quimica.es – el portal informativo químico, desde el laboratorio hasta los procesos. 2008. Disponible en: https://www.quimica.es/enciclopedia/Ácido_graso.html#_note-Dev/. [consultado el 12/07/2022].

MACKERELL Lab. *CHARMM force field files*. En línea. MacKerell Lab: gromacs charmm_ff. [s. f.]. Disponible en: http://mackerell.umaryland.edu/charmm_ff.shtml#gromacs. [consultado el 01/07/2022].

MACKERELL Lab. *Welcome to CGenFF*. En línea. CGenFF. 17/08/2018. Disponible en: <https://cgenff.umaryland.edu/>. [consultado el 03/07/2022].

MARTOŇÁK, R., LAIO, A. & PARRINELLO, M. Predicting crystal structures: the parrinello-rahman method revisited. En línea. *Physical Review Letters*, vol. 90 (febrero de 2003), n.º 7. ISSN 1079-7114. Disponible en: <https://doi.org/10.1103/physrevlett.90.075503>. [consultado el 22/07/2022].

MEJÍA, A. Para iniciar debes saber. En línea. Algoritmos, primeros pasos. 24/06/2015. Disponible en: <https://todoalgoritmos.weebly.com/#:~:text=Algoritmo:,%20diagramas%20de%20flujo,%20etc>. [consultado el 14/09/2022].

MEZA, C. Steepest descent. En línea. *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2 (septiembre de 2010), n.º 6, pp. 719–722. ISSN 1939-0068. Disponible en: <https://doi.org/10.1002/wics.117>. [consultado el 02/07/2022].

MISHRA, V. & HEATH, R. Structural and biochemical features of human serum albumin essential for eukaryotic cell culture. En línea. *International Journal of Molecular Sciences*,

vol. 22 (agosto de 2021), n. ° 16, p. 8411. ISSN 1422-0067. Disponible en: <https://doi.org/10.3390/ijms22168411>. [consultado el 19/07/2022].

MONTICELLI, L. & TIELEMAN, P. Force field for classical molecular dynamics. *Methods in molecular biology*, 2013, n. ° 924, pp. 197–213.

NOSÉ, S. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, vol. 52 (junio de 1984), n.° 2, pp. 255–268. ISSN 1362-3028.

NVIDIA CORPORATION. *A simple yet powerful tool to turn traditional container/os images into unprivileged sandboxes.* En línea. GitHub. 2022. Disponible en: <https://github.com/NVIDIA/enroot>. [consultado el 25/06/2022].

PETERS T. Serum albumin. *Advances in protein chemistry*, 1985, vol. 37, p. 161-245.

PETITPAS, I., BHATTACHARYA, A., TWINE, S., EAST, M. & CURRY, S. Crystal structure analysis of warfarin binding to human serum albumin. *Journal of Biological Chemistry*, vol. 276 (abril de 2001), n. ° 25, pp. 22804–22809. ISSN 1083-351X.

RETTIE, A. & GUOYING, T. The pharmacogenomics of warfarin: closing in on personalized medicine. En línea. *Molecular Interventions*, vol. 6 (agosto de 2006), n.° 4, pp. 223–227. ISSN 1534-0384. Disponible en: <https://doi.org/10.1124/mi.6.4.8>. [consultado el 19/07/2022].

RICKAERT, J. The method of constraints in molecular dynamics. General aspects and application to chain molecules. En: *Molecular dynamic simulation of statistical mechanical systems*, pp. 329–340. Amsterdam: North-Holland Physics, 1986.

RONDÓN, I. *¿Qué es anaconda?* En línea. Escuela Internacional de Posgrados. 04/02/2022. Disponible en: <https://eiposgrados.com/blog-python/que-es-anaconda/>. [consultado el 29/05/2022].

SARGSYAN, K., GRAUFFEL, C. & LIM, C. How molecular size impacts RMSD applications in molecular dynamics simulations. *Journal of Chemical Theory and Computation*, vol. 13 (marzo de 2017), n. ° 4, pp. 1518–1524. ISSN 1549-9626.

SCHWARTZ, D. & LOCKARD, C. Blood. En línea. Encyclopedia Britannica. 04/04/2022. Disponible en: <https://www.britannica.com/science/thermodynamics/blood-biochemistry>. [consultado el 14/09/2022].

SUGIO, S., KASHIMA, A., MOCHIZUKI, S., NODA, M. & KOBAYASHI, K. Crystal structure of human serum albumin at 2.5 Å resolution. *Protein Engineering, Design and Selection*, vol. 12 (junio de 1999), n.º 6, pp. 439–446. ISSN 1741-0134.

TECHPOWERUP. *NVIDIA A100 SXM4 40 GB specs*. En línea. [techpowerup.com](https://www.techpowerup.com). [s. f.]. Disponible en: <https://www.techpowerup.com/gpu-specs/a100-sxm4-40-gb.c3506>. [consultado el 03/06/2022].

THEORETICAL AND COMPUTATIONAL BIOPHYSICS GROUP. *What is VMD?* En línea. University of Illinois Urbana-Champaign. Marzo de 2008. Disponible en: https://www.ks.uiuc.edu/Research/vmd/allversions/what_is_vmd.html. [consultado el 28/05/2022].

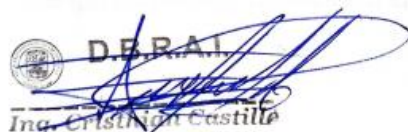
VANDERPLAS, J. *Python data science handbook: essential tools for working with data*. Dawn SCHANAFELT (ed.). O'Reilly Media, 2016. ISBN 9781491912058.

VANOMMESLAEGHE, K., HATCHER, E., ACHARYA, C., KUNDU, S., ZHONG, S., et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, 2009, pp. 671–690. ISSN 1096-987X.

VANOMMESLAEGHE, K. & MACKERELL, A. Automation of the CHARMM general force field (cgenff) I: bond perception and atom typing. *Journal of Chemical Information and Modeling*, vol. 52 (noviembre de 2012), n.º 12, pp. 3144–3154. ISSN 1549-960X.

WISHART, D., KNOX, C., GUO, A., SHRIVASTAVA, S., HASSANALI, M. et al. *Capric acid: uses, interactions, mechanism of action*. En línea. DrugBank Online: Database for Drug and Drug Target Info. Junio de 2022. Disponible en: <https://go.drugbank.com/drugs/DB03600>. [consultado el 13/07/2022].

ZHU, X., LOPES, P. & MACKERELL A. Recent developments and applications of the CHARMM force fields. *WIREs Computational Molecular Science*, vol. 2 (junio de 2011), n.º 1, pp. 167–185. ISSN 1759-0884.


D. B. R. A. I.
Ing. Cristóbal Castilla



ANEXOS

ANEXO A: PROGRAMA DE PYTHON PARA HACER LA RECONSTRUCCIÓN DE LA SECUENCIA DE AMINOÁCIDOS DE LA PROTEÍNA.

```
def generarAligment(file1,file2,resp):
    alig=open('alignment.ali','w')
    fastaRead=file1.read().splitlines()
    fasta_cad=fastaRead[1]
    if (resp):
        fasta_cad=fasta_cad+'/'+fasta_cad
    seqRead=file2.read().splitlines()
    lista=seqRead[3:]
    cad=""
    for elem in lista:cad=cad+elem
    seq_cad=cad[:-1]
    #seq_cad=seq_cad.replace('/',"")
    for i in range(len(fasta_cad)):
        if i<len(seq_cad):
            if seq_cad[i]!=fasta_cad[i]:
                seq_cad=seq_cad[:i]+'-'+seq_cad[i:]
    guions=len(fasta_cad)-len(seq_cad)
    seq_cad=seq_cad+'-'*guions
    alig.write(seqRead[1]+\n')
    alig.write(seqRead[2]+\n')
    alig.write(seq_cad+'*\n')
    alig.write(seqRead[1]+'_fill'+\n')
    alig.write('sequence:.....'+\n')
    alig.write(fastaRead[1]+'*')
    alig.close()
```

ANEXO B: PROGRAMA DE PYTHON QUE REALIZA LA MUTAGÉNESIS DE LA HSA CREANDO DISTINTOS MODELOS CON SU RESPECTIVA PUNTUACIÓN

```
from modeller import *
from modeller.automodel import *
def HSA():
    ruta='simulacionHSA/Files/HSA.pdb'
    fasta='simulacionHSA/Files/HSA.fasta'
    e = Environ()
    m = Model(e, file=ruta)
    aln = Alignment(e)
    aln.append_model(m, align_codes=ruta)
    aln.write(file=ruta+'.seq')
    fasta=open(fasta,'r')
    seq=open(ruta+'.seq','r')
    generarAligment(fasta,seq,False)
    env= Environ()
    env.io.atom_files_directory = ['.', './atom_files']
    a = LoopModel(env, alnfile = 'alignment.ali',knowns = ruta, sequence = ruta+'_fill')
    a.starting_model= 1
    a.ending_model = 1
    a.loop.starting_model = 1
    a.loop.ending_model = 2
    a.loop.md_level = refine.fast
    a.make()
    seq.close()
    fasta.close()
    return
```

ANEXO C: PROGRAMA DE PYTHON QUE REALIZA LA MUTAGÉNESIS DEL SISTEMA HSA - DKA CREANDO DISTINTOS MODELOS.

```
from modeller import *
from modeller.automodel import *
def HSADKA():
    ruta='simulacionHSADKA/Files/HSADKA.pdb'
    fasta='simulacionHSADKA/Files/HSADKA.fasta'
    e = Environ()
    m = Model(e, file=ruta)
    aln = Alignment(e)
    aln.append_model(m, align_codes=ruta)
    aln.write(file=ruta+'.seq')
    fasta=open(fasta,'r')
    seq=open(ruta+'.seq','r')
    generarAligment(fasta,seq,True)
    log.verbose()
    env= Environ()
    env.io.atom_files_directory = ['.', './atom_files']
    class MyModel(AutoModel):
        def select_atoms(self):
            return Selection(self.residue_range('1:A', '2:A'),self.residues['585:A'])
    a = MyModel(env, alnfile = 'alignment.ali',knowns = ruta, sequence = ruta+'_fill')
    a.starting_model= 1
    a.ending_model = 1
    a.make()
    return
```

ANEXO D: PROGRAMA DE PYTHON QUE REALIZA LA MUTAGÉNESIS DEL SISTEMA HSA - WFR CREANDO DISTINTOS MODELOS.

```
from modeller import *
from modeller.automodel import *
def HSAWFR():
    ruta='simulacionHSAWFR/Files/HSAWFR.pdb'
    fasta='simulacionHSAWFR/Files/HSAWFR.fasta'
    e = Environ()
    m = Model(e, file=ruta)
    aln = Alignment(e)
    aln.append_model(m, align_codes=ruta)
    aln.write(file=ruta+'.seq')
    fasta=open(fasta,'r')
    seq=open(ruta+'.seq','r')
    generarAligment(fasta,seq,True)
    log.verbose()
    env= Environ()
    env.io.atom_files_directory = ['.', './atom_files']
    a = LoopModel(env, alnfile = 'alignment.ali',knowns = ruta, sequence = ruta+'_fill')
    a.starting_model= 1
    a.ending_model = 1
    a.make()
    return
```

ANEXO E: ARCHIVO MDP BASE PARA LA EJECUCIÓN DE LA FASE DE MINIMIZACIÓN DE ENERGÍA

```
; LINES STARTING WITH ';' ARE COMMENTS
title          = HSA_Minimization ; Title of run

; Parameters describing what to do, when to stop and what to save
integrator     = steep           ; Algorithm (steep = steepest descent
minimization)
emtol         = 1000.0          ; Stop minimization when the maximum force < 10.0
kJ/mol
emstep        = 0.01           ; Energy step size
nsteps        = 50000          ; Maximum number of (minimization) steps to
perform

; Parameters describing how to find the neighbors of each atom and how to calculate the
interactions
nstlist       = 1              ; Frequency to update the neighbor list and long
range forces
cutoff-scheme = Verlet
ns_type       = grid           ; Method to determine neighbor list (simple, grid)
rlist         = 1.2            ; Cut-off for making neighbor list (short range forces)
coulombtype   = PME            ; Treatment of long range electrostatic
interactions
rcoulomb      = 1.2            ; long range electrostatic cut-off
vdwtype       = cutoff
vdw-modifier  = force-switch
rvdw-switch   = 1.0
rvdw         = 1.2             ; long range Van der Waals cut-off
pbc           = xyz            ; Periodic Boundary Conditions
DispCorr     = no
```

ANEXO F: ARCHIVO MDP BASE PARA LA EJECUCIÓN DEL ENSAMBLE NVT

```
title          = Protein-HSA NVT equilibration
define         = -DPOSRES ; position restrain the protein and ligand

; Run parameters
integrator     = md      ; leap-frog integrator
nsteps        = 50000    ; 2 * 50000 = 100 ps
dt            = 0.002    ; 2 fs

; Output control
nstenergy     = 5000     ; save energies every 10 ps
nstlog        = 5000    ; update log file every 10 ps
nstxout-compressed = 5000 ; save coordinates every 10 ps
nstvout       = 5000    ; save velocities every 10 ps

; Bond parameters
continuation   = no      ; first dynamics run
constraint_algorithm = lincs ; holonomic constraints
constraints    = h-bonds ; bonds to H are constrained
lincs_iter     = 1      ; accuracy of LINCS
lincs_order    = 4      ; also related to accuracy

; Neighbor searching and vdW
cutoff-scheme  = Verlet
ns_type        = grid    ; search neighboring grid cells
nstlist        = 20     ; largely irrelevant with Verlet
rlist          = 1.2
vdwtype        = cutoff
vdw-modifier   = force-switch
rvdw-switch    = 1.0
rvdw           = 1.2    ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype    = PME     ; Particle Mesh Ewald for long-range electrostatics
rcoulomb       = 1.2     ; short-range electrostatic cutoff (in nm)
pme_order      = 4      ; cubic interpolation
fourierspacing = 0.16   ; grid spacing for FFT

; Temperature coupling
tcoupl         = V-rescale ; modified Berendsen thermostat
tc-grps       = Protein Non-Protein ; two coupling groups - more accurate
tau_t         = 0.1 0.1   ; time constant, in ps
ref_t         = 300 300   ; reference temperature, one for each group, in K

; Pressure coupling
pcoupl        = no      ; no pressure coupling in NVT

; Periodic boundary conditions
pbc           = xyz     ; 3-D PBC

; Dispersion correction is not used for proteins with the C36 additive FF
DispCorr      = no

; Velocity generation
gen_vel       = yes     ; assign velocities from Maxwell distribution
gen_temp      = 300     ; temperature for Maxwell distribution
gen_seed      = -1     ; generate a random seed
```


ANEXO G: ARCHIVO MDP BASE PARA LA EJECUCIÓN DEL ENSAMBLE NPT

```
title      = Protein-HSA NPT equilibration
define     = -DPOSRES      ; position restrain the protein and ligand

; Run parameters
integrator = md            ; leap-frog integrator
nsteps     = 50000        ; 2 * 50000 = 100 ps
dt         = 0.002        ; 2 fs

; Output control
nstenergy  = 5000         ; save energies every 10 ps
nstlog     = 5000         ; update log file every 10 ps
nstxout-compressed = 5000 ; save coordinates every 10 ps

; Bond parameters
continuation = yes        ; continuing from NVT
constraint_algorithm = lincs ; holonomic constraints
constraints  = h-bonds    ; bonds to H are constrained
lincs_iter   = 1          ; accuracy of LINCS
lincs_order  = 4          ; also related to accuracy

; Neighbor searching and vdW
cutoff-scheme = Verlet
ns_type       = grid      ; search neighboring grid cells
nstlist       = 20        ; largely irrelevant with Verlet
rlist         = 1.2
vdwtype       = cutoff
vdw-modifier  = force-switch
rvdw-switch   = 1.0
rvdw          = 1.2       ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype   = PME       ; Particle Mesh Ewald for long-range electrostatics
rcoulomb      = 1.2
pme_order     = 4         ; cubic interpolation
fourierspacing = 0.16    ; grid spacing for FFT

; Temperature coupling
tcoupl        = V-rescale ; modified Berendsen thermostat
tc-grps       = Protein Non-Protein ; two coupling groups - more accurate
tau_t         = 0.1 0.1   ; time constant, in ps
ref_t         = 300 300   ; reference temperature, one for each group, in K

; Pressure coupling
pcoupl        = Parrinello-Rahman ; pressure coupling is on for NPT
pcoupltype    = isotropic        ; uniform scaling of box vectors
tau_p         = 2.0               ; time constant, in ps
ref_p         = 1.0               ; reference pressure, in bar
compressibility = 4.5e-5          ; isothermal compressibility of water, bar^-1
refcoord_scaling = com

; Periodic boundary conditions
pbc           = xyz      ; 3-D PBC
; Dispersion correction is not used for proteins with the C36 additive FF
DispCorr      = no

; Velocity generation
gen_vel       = no       ; velocity generation off after NVT
```

ANEXO H: ARCHIVO MDP BASE PARA LA EJECUCIÓN DE LA DINÁMICA MOLECULAR

```

title          = HSA_protein MD simulation

; Run parameters
integrator     = md          ; leap-frog integrator
nsteps        = 5000000     ; 2 * 5000000 = 10000 ps (10 ns)
dt            = 0.002      ; 2 fs

; Output control
nstxout       = 0          ; suppress bulky .trr file by specifying
nstvout       = 0          ; 0 for output frequency of nstxout,
nstfout       = 0          ; nstfout, and nstfout
nstenergy     = 5000       ; save energies every 10.0 ps
nstlog        = 5000       ; update log file every 10.0 ps
nstxout-compressed = 5000   ; save compressed coordinates every 10.0 ps
compressed-x-grps = System ; save the whole system

; Bond parameters
continuation   = yes          ; Restarting after NPT
constraint_algorithm = lincs ; holonomic constraints
constraints    = h-bonds     ; bonds involving H are constrained
lincs_iter     = 1          ; accuracy of LINCS
lincs_order    = 4          ; also related to accuracy

; Neighborsearching
cutoff-scheme  = Verlet      ; Buffered neighbor searching
ns_type        = grid        ; search neighboring grid cells
nstlist        = 10         ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb       = 1.0        ; short-range electrostatic cutoff (in nm)
rvdw           = 1.0        ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype    = PME        ; Particle Mesh Ewald for long-range electrostatics
pme_order      = 4          ; cubic interpolation
fourierspacing = 0.16      ; grid spacing for FFT

; Temperature coupling is on
tcoupl        = V-rescale    ; modified Berendsen thermostat
tc-grps       = Protein Non-Protein ; two coupling groups - more accurate
tau_t         = 0.1 0.1     ; time constant, in ps
ref_t         = 300 300     ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl        = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype    = isotropic    ; uniform scaling of box vectors
tau_p         = 2.0         ; time constant, in ps
ref_p         = 1.0         ; reference pressure, in bar
compressibility = 4.5e-5     ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc           = xyz        ; 3-D PBC

; Dispersion correction
DispCorr      = EnerPres ; account for cut-off vdW scheme

; Velocity generation
gen_vel       = no        ; Velocity generation is off

```



epoch

Dirección de Bibliotecas y
Recursos del Aprendizaje

UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y
DOCUMENTAL

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 19 / 01 / 2023

INFORMACIÓN DEL AUTOR/A (S)
Nombres – Apellidos: Iván Fabricio Chávez Velasco
INFORMACIÓN INSTITUCIONAL
Facultad: Facultad de Ciencias
Carrera: Física
Título a optar: Físico
f. responsable: Ing. Cristhian Fernando Castillo Ruiz



0008-DBRA-UTP-2023