



Memoria del XLVIII Taller de Actualización Bioquímica, Facultad de Medicina; UNAM

Aplicaciones de simulaciones de dinámica molecular en proteínas.

Applications of molecular dynamics simulations on proteins.

González Andrade, Martín^{1*}.

1. Laboratorio de Biosensores y Modelaje Molecular, Departamento de Bioquímica, Facultad de Medicina, UNAM.

*Correspondencia. Torre de Investigación, Facultad de Medicina, Circuito Escolar s/n, Ciudad Universitaria, Coyoacan, CDMX, México.
CP. 04510 Tel. +52 (55) 56 23 22 54, martin@bq.unam.mx

Resumen

Las Simulaciones de Dinámica Molecular de proteínas y sus complejos representan actualmente una herramienta eficaz para describir a nivel atómico sistemas biológicos complejos y poder responder preguntas de investigación asociadas a estos sistemas. Esta herramienta computacional es un complemento idóneo a la parte experimental, con diversas aplicaciones para ser utilizada tales como el refinamiento de modelos estructurales, cálculos de parámetros energéticos de unión (proteína-ligando), estudio de plegamiento de proteínas, mecanismos de catálisis enzimática entre los más importantes. Las Simulaciones de Dinámica Molecular sea han visto beneficiadas del avance tecnológico de los sistemas informáticos, tanto a nivel *hardware* como a nivel *software*. Lo anterior, ha permitido tener una gran oferta de software y disminuir el tiempo de cómputo. En este protocolo, proporcionamos conceptos básicos, fundamentos de la técnica, aplicaciones de las Simulaciones de Dinámica Molecular, *softwares* disponibles para realizarlas y analizarlas; así como un protocolo general para un sistema proteína-ligando.

Palabras claves: Simulaciones de Dinámica Molecular; herramienta computacional; AMBER; AMBERTOOLS.

Abstract

Molecular Dynamics Simulations of proteins and their complexes currently represent an efficient tool to describe complex biological systems at the atomic level and answer research questions associated with these systems. This computational tool is an ideal complement to the experimental part, with various applications to be used such as the refinement of structural models, calculations of binding energy parameters (protein-ligand), the study of protein folding, mechanisms of enzymatic catalysis, among the more important. Molecular Dynamics Simulations have benefited from the technological advancement of computer systems, both at the hardware and at the software level; this has allowed us to have a great offer of software and reduce computing time. In this protocol, we provide basic concepts, fundamentals of the technique, applications of Molecular Dynamics Simulations, Software available to perform and analyze them, and a general protocol for a protein-ligand system.

Keywords: Molecular Dynamics Simulations; computational tool; AMBER; AMBERTOOLS.

Introducción

El estudio de la estructura de macromoléculas (proteínas) es clave en la comprensión de la biología estructural. La función biológica se basa en interacciones moleculares entre proteínas, carbohidratos, aminoácidos, lípidos, iones y agua en un sistema dinámico. En el caso de las proteínas, desde las primeras determinaciones de la estructura terciaria en los años 50's, se comenzó a entender la relevancia del conocimiento de cómo las proteínas se comunican e interaccionan con otras moléculas para ejercer su función de manera óptima.

Las reglas de reconocimiento molecular definidas por dicho conocimiento estructural potencian la comprensión de los fenómenos biológicos básicos, como: los mecanismos y la regulación enzimáticos, el transporte a través de membranas, la construcción de grandes estructuras como los ribosomas o las cápsides virales, o cómo se lee el DNA y se controla la transcripción, entre los más importantes.

Actualmente, el estudio y la predicción de las interacciones proteína-proteína o proteína-ligando es uno de los campos en crecimiento en la biología de sistemas. Por ejemplo, una aplicación práctica de este conocimiento es el diseño de fármacos, ya que esta basado en las estructuras de proteínas tridimensionales (3D). El simple análisis visual de estructuras tridimensionales de proteínas ha impulsado un gran número de estudios exitosos en bioquímica. Sin embargo, a pesar de su enorme utilidad, las estructuras sólo proporcionan un conocimiento limitado de la estructura 3D, ya que las proteínas son entidades flexibles y su dinámica puede desempeñar un papel clave en su funcionalidad.

Los procesos de simulación por computadora se pueden clasificar en 4 categorías:

- 1) *ab-initio*, el cual calcula propiedades de sistemas como proteínas, utilizando la ecuación de Schrödinger's para resolver el problema numéricamente. En este método las interacciones entre los átomos son determinada por sus configuraciones electrónicas y la posición en el sistema;
- 2) *semi-empírico*, este simplifica los cálculos del método *ab-initio* y puede ser usado para estudiar grandes sistemas;
- 3) *mecánica molecular*, este método es ampliamente utilizado para el estudio de sistemas grandes usando aproximaciones de

Born-Oppenheimer en el cual los estados del sistema pueden ser determinados usando la posición y velocidad de cada átomo. Este método es el más utilizado para estudiar sistemas que tiene miles de átomos como son las proteínas, y nos permite calcular diversos parámetros teóricos como lo son parámetros termodinámicos, constantes de afinidad, e información de interacciones proteína-proteína o proteína-ligando; por mencionar algunos [1,2]. En este método las interacciones entre cada átomo son estimadas utilizando derivados potenciales interatómicos a partir de los métodos *ab-initio*;

- 4) *semiempíricos de la mecánica cuántica (semiempíricos-QM)* estos métodos utilizan el de cálculo de la estructura electrónica química cuántica que normalmente se basan en un formalismo para la mecánica cuántica *ab initio* (por ejemplo, el orbital molecular) o la teoría funcional de la densidad. Aplican más aproximaciones y utilizan algunos parámetros de datos empíricos [3-5].

Las simulaciones de dinámica molecular (SDM) de sistemas biológicos complejos representan modelos estructurales en computadoras con gran capacidad de procesamiento, para tratar de simular teóricamente lo que sucede experimentalmente a escalas de tiempo [6]. En los últimos años, esta herramienta computacional ha cobrado relevancia con fines de investigación, ya que se puede utilizar para refinar modelos estructurales, proponer mecanismos de acción, calcular energías de unión, estudiar el plegamiento de proteínas, entre los más importantes. Otro aspecto importante que está ligado a las SDM es el avance tecnológico de los sistemas informáticos, tanto a nivel *hardware* como a nivel *software*, ya que el primero permite mayor velocidad de procesamiento y el segundo mejor gestión de datos.

La Figura 1, muestra un esquema general para preparar, realizar y analizar SDM de proteínas. Las SDM son de gran utilidad donde las preguntas de investigación requieren llevar a cabo experimentos difíciles o que éstos únicamente puedan proporcionarnos datos parciales. Otro escenario donde son muy apreciadas las SDM son en los centros de investigación que no cuentan con la infraestructura experimental suficiente para poder responder inquietudes científicas de sistemas biológicos complejos [7]. El concepto de simulaciones moleculares incluye los métodos teórico-computacionales que describen un sistema a nivel

atómico. Estos enfoques computacionales presentan algunas ventajas, tales como: *i)* reducción de tiempo experimental, *ii)* reducción de experimentos costosos, *iii)* proveen evaluaciones preliminares, *iv)* complementan datos experimentales, *v)* permiten realizar hipótesis de manera dirigidas, y *vi)* nos permiten representar sistemas biológicos más cercanos a la realidad. Por lo tanto, las SDM

representan una herramienta poderosa para describir a nivel atómico sistemas biológicos complejos como son las proteínas y sus diferentes ligandos; así como poder responder preguntas de investigación. En este protocolo, proporcionamos conceptos básicos, fundamentos de la técnica, aplicaciones de SDM y un ejemplo de estudio en un sistema proteína-ligando.

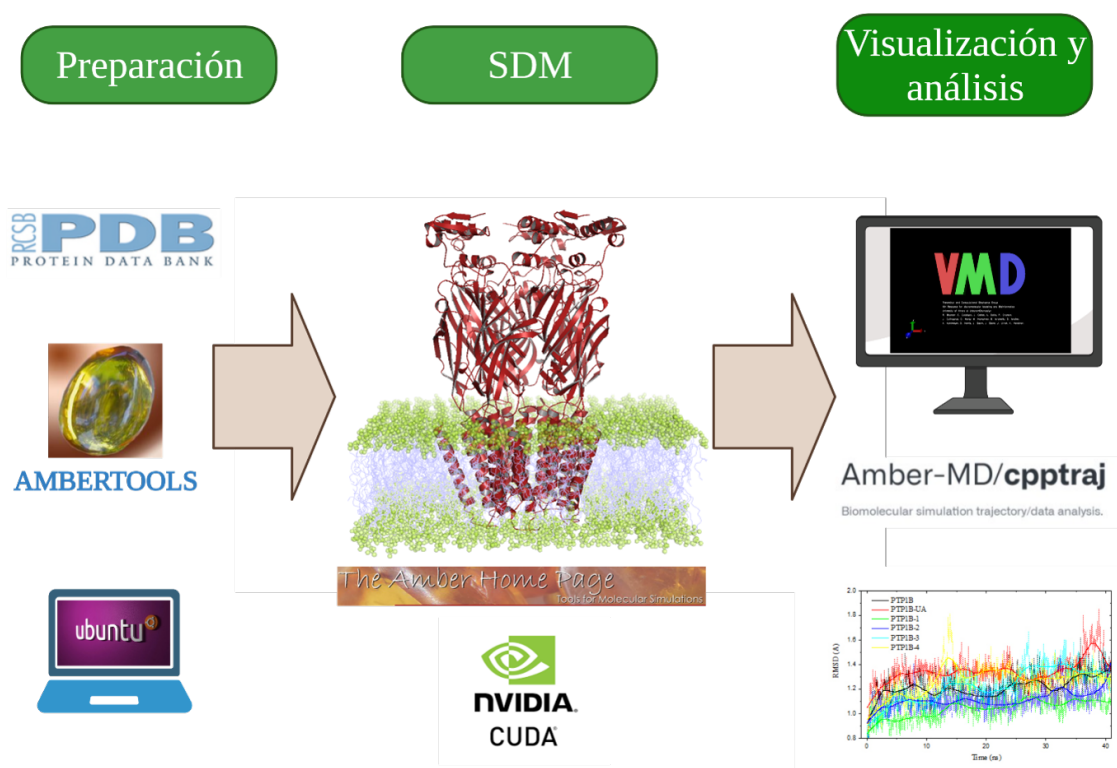


Figura 1. Esquema general para realizar una simulación de dinámica molecular. La etapa de preparación consiste en obtener una estructura 3D a partir del Protein Data Bank (PDB); ajustarla, limpiarla y editarla para ser utilizada por las herramientas de AMBERTOOLS, todo lo anterior en un ambiente Unix (Ubuntu®) para obtener los archivos de topología y coordenadas xyz de nuestro sistema para la SDM. En la segunda etapa se realiza un pre-proceso antes de la SDM, el cual consta de tres etapas fundamentales; 1) minimización, 2) despresurización y 3) equilibrio del sistema. Finalmente, la SDM es realizada en equipo de super computo (tarjetas NVIDIA CUDA® con miles de GPUs) con el programa AMBER [8,9], con la obtención de trayectorias de dinámica molecular del tiempo programado. La última etapa es la visualización y el análisis de la trayectoria de dinámica molecular, lo cual se realiza con los programas VMD y CPPTRAJ [10], respectivamente.

Fundamento de las simulaciones de dinámica molecular

Las bases teóricas de las SDM se concentran en los métodos de mecánica molecular, los cuales usan la aproximación de Born–Oppenheimer. Esta teoría de aproximación sólo contempla la posición y velocidad de los átomos requeridos para describir los estados microscópicos de un sistema (proteína, complejo proteína-ligandos).

La ecuación 1 se utiliza para expresar el Hamiltoniano del sistema como una función de las variables nucleares.

$$\text{Ecuación 1} \quad H(q, p) = K(p) + V(q)$$

donde, H es el Hamiltoniano, $K(p)$ es la energía cinética como una función de los momentos de cada átomo, y $V(q)$ es la energía potencial en función de las coordenadas generalizadas, es decir, contiene los detalles de las interacciones interatómicas. La mecánica molecular considera que la composición atómica de una molécula, como una proteína, es una

colección de masas que interactúan entre sí a través de fuerzas armónicas. Una simplificación adicional sugiere que en la mecánica molecular los átomos se consideran esferas y los enlaces entre ellos se consideran resortes. Ahora, para visualizar y estudiar la dinámica de estos sistemas, necesitamos aplicar una fuerza. Esto es posible generando una ecuación de movimiento en forma newtoniana usando el Hamiltoniano [11].

$$\text{Ecuación 2} \quad \frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i}$$

donde F_i , es definida como la fuerza total experimentada por un átomo i en la dirección r , m_i es la masa de los átomos y r_i es la posición de los átomos i . Por lo tanto, si nosotros consideramos que conocemos la posición y la velocidad de los átomos i a un tiempo t , entonces la posición y la velocidad en un tiempo $t + dt$, donde dt es el intervalo de tiempo entre dos pasos de simulación, podemos estimar solucionando la ecuación 2 usando un método de integración. Uno de los métodos más utilizados para resolver este tipo de ecuaciones es el algoritmo de Verlet [12]. El algoritmo de Verlet requiere los valores de las posiciones actuales, $r(t)$; aceleración $a(t)$; y la posición de los átomos del paso anterior, $r(t-dt)$. Las posiciones y velocidades de los átomos i para el siguiente paso los podemos obtener usando las ecuaciones 3 y 4, respectivamente.

Ecuación 3

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t)$$

Ecuación 4

$$v_t = \frac{r(r + \delta t) - r(t - \delta t)}{2\delta t}$$

Las ecuaciones 3 y 4, representan la forma básica del algoritmo de Verlet y son propensas a errores de procesamiento del orden de dt^4 para la posición y dt^2 para las velocidades. La estimación de las velocidades es importante ya que se utiliza para calcular la energía cinética del sistema; computacionalmente se requiere el cálculo de la energía cinética para probar la conservación de la energía total. Esta prueba es una de las más importantes durante las SDM, ya que verifica la estabilidad y exactitud de la SDM.

En la mecánica molecular, las fuerzas que actúan sobre los átomos i pueden surgir tanto de fuentes internas como externas. Las fuentes internas son las fuerzas de interacción que actúan entre átomos unidos y no unidos. Las fuentes externas pueden ser tensiones

ambientales, como el campo eléctrico, el calor y la presión que se imponen al sistema. Para determinar las fuerzas, es importante obtener la energía potencial del sistema, la cual se puede expresar como una suma de interacciones covalente, no covalentes y entrecruzada. Las derivadas de la función de energía potencial se conocen como campos de fuerza (Ecuación 5).

Ecuación 5

$$E_{total} = E_{covalente} + E_{no-covalente} + E_{entrecruzadas}$$

Las interacciones covalentes también se denominan interacciones de valencia e incorporan términos diagonales como estiramiento de enlace, flexión de ángulo, torsión de ángulo diedro, inversión o interacciones fuera del plano. Dado que en la mecánica molecular se supone que las interacciones se producen a través de fuerzas armónicas, la energía potencial asociada con el estiramiento de los enlaces se puede representar como:

Ecuación 6

$$E_{estiramiento} = \frac{1}{2} k_b (b - b_0)^2$$

donde, k_b es la constante de fuerza para la longitud del enlace, b_0 es la longitud del enlace al equilibrio y b es la longitud actual del enlace. De manera similar para la flexión de ángulos, la expresión armónica simple sería:

Ecuación 7

$$E_{flexion\ de\ angulo} = \frac{1}{2} k_\theta (\theta - \theta_0)^2$$

donde, K_θ es la constante de fuerza para los ángulos de enlace, θ_0 es el ángulo de enlace al equilibrio, y θ es el ángulo de enlace actual. La energía potencial del ángulo de torsión diédrico es representada como una expresión del coseno (Ecuación 8).

Ecuación 8

$$E_{torsion} = \frac{1}{2} k_\theta (1 + \cos(n\theta - \theta_0))$$

donde, K_θ es la constante de fuerza para el ángulo diédrico, θ_0 es la referencia del ángulo de torsión y θ es el ángulo de torsión actual.

Todas las funciones de energía potencial mencionadas anteriormente pertenecen a campos de fuerza de Clase I, que únicamente incorporan términos diagonales e interacciones no-covalentes. Esta clase de campos de fuerza son ampliamente utilizadas para

simular sistemas complejos como son las proteínas. Algunos de los campos de fuerza de Clase I más utilizados son AMBER [8,9], CHARMM [13] y GROMOS [14,15]. Las interacciones no-covalentes definidas en esta clase de campos de fuerza incluyen interacciones electrostáticas y Van der Waals. Las interacciones electrostáticas pueden ser descritas como una interacción entre dos partículas cargadas y es expresada con una función potencial de Coulomb:

Ecuación 9

$$E_{electrostatica} = \sum_{ij} \frac{q_i q_j}{4\pi D r_{ij}}$$

donde, q es la carga de la partícula, D es la constante dieléctrica y r_{ij} es la distancia entre los centros de las partículas i y j .

Las interacciones de Van der Waals entre átomos no-covalentes pueden ser incorporadas usando el potencial de Lennard-Jones. En mecánica molecular, la función de energía potencial más simple es el potencial de esfera dura. Este asume que las partículas viajan en línea recta hasta chocar con otra partícula, lo que provocaría una dispersión elástica. Este potencial no incluye los componentes atractivos y repulsivos, que lo hacen inadecuado para el estudio de grandes sistemas como las proteínas. El potencial de Lennard-Jones incluye componentes tanto atractivos como repulsivos y se puede expresar mediante la Ecuación 10.

Ecuación 10

$$E = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$

donde, ε es la profundidad del potencial, r es la distancia entre las partículas y σ es la distancia a la que el potencial entre partículas es 0. Esta función también se denomina potencial 12-6, la cual representa los términos del exponente para los componentes repulsivos y atractivos, respectivamente. Usando la Ecuación 10, las interacciones de Van der Waals entre partículas no-covalentes se puede presentar como:

Ecuación 11

$$E_{VDW} = \sum_{ij} \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right]$$

donde, r_{ij} es la distancia entre las partículas, ε_{ij} es el valor mínimo para el término Van der Waals, y R_{min} es el radio donde el término Van der Waals es mínimo.

Softwares disponibles para realizar simulación de dinámica molecular y su análisis

El desarrollo de programas informáticos adecuados para realizar SDM es un proceso que requiere la inversión de mucho tiempo, habilidades de personas especializadas en diversas áreas y financiamiento público y privado. Afortunadamente, un esfuerzo concertado de científicos de todo el mundo ha contribuido con una variedad de aplicaciones comerciales y de código abierto adecuadas para este fin (Tabla 1).

Adicionalmente, existen varios software y plataformas en línea que se utilizan para construir moléculas (Avogadro, Hyperchem <http://www.softwareactivo.com.mx/portfolio/hyperchem/>, Materials Studio <http://accelrys.com/products/collaborative-science/biovia-materials-studio/>, Marvin <https://chemaxon.com/products/marvin/>), para determinar los parámetros del campo de fuerza (PRODRG <http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg>, Swiss Param <https://www.swissparam.ch/>, ATB <https://atb.uq.edu.au/>, Gaussian suite <https://gaussian.com/>) y otros para analizar y visualizar las SDM (VMD <https://www.ks.uiuc.edu/Research/vmd/>, PyMOL <https://pymol.org/2/>, CHIMERA <https://www.cgl.ucsf.edu/chimera/>, OriginLab <https://www.originlab.com/>, R Proyecto para estadística <https://www.r-project.org/>, Grace <https://plasma-gate.weizmann.ac.il/Grace/>).

Estos softwares nos proporcionan resultados dinámicos del sistema simulado y proveen información necesaria sobre los atributos de calidad críticos del propio sistema. Sin embargo, estos softwares tienen algunas desventajas que limitan su uso. Primero, es necesario mejorar los campos de fuerza actuales para mejorar la precisión de las predicciones de energía libre. Los efectos de la polarización, la transferencia de carga y otras interacciones electrónicas se ignoran total o parcialmente. En segundo lugar, no es posible ejecutar simulaciones de rutina de más de un microsegundo debido a la alta demanda computacional. Dado que los eventos reales, como el plegamiento de proteínas, toman segundos, los sistemas informáticos actuales son incapaces de manejar tales cálculos. Por tanto, la selección apropiada del paquete de software para las simulaciones depende del sistema biológico a simular,

de las variables de entrada, el tiempo de simulación y la disponibilidad computacional.

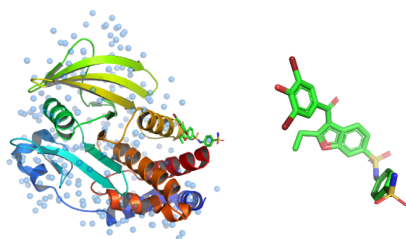
Tabla 1. Programas computacionales disponibles para realizar Simulación de Dinámica Molecular (SDM).

Programa	Versión	Tipo de licencia	Copilado para CPU/GPU	Campo de fuerza	Página web	Referencia
AMBER	Amber20	Comercial	Ambos	AMBER (ff19SB)	https://ambermd.org/	[8,9]
GROMACS	4.6	Libre	Ambos	AMBER, CHARMM, GROMOS y OPLS	https://www.gromacs.org/	[14,15]
NAMD	2.14	Comercial	Ambos	AMBER	https://www.ks.uiuc.edu/Research/namd/	[16]
CHARMM	C45b2	Libre (Academicos)	Ambos	CHARMM	https://www.charmm.org/	[13]
OpenMM	7.4	Libre	Ambos	CHARMM	https://openmm.org/	[17]

Procedimiento para realizar una Simulación de Dinámica Molecular usando las herramientas de AMBERTOOLS y AMBER

Preparación

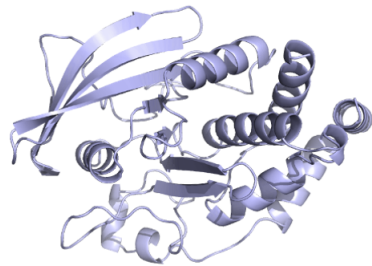
1. Selección de las coordenadas iniciales de la estructura de la biomolécula (Protein Data Bank), y ligandos (obtenidos o generados).



2. Eliminar moléculas de aguas y co-solvente

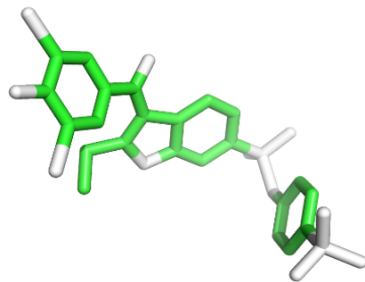


Este paso se puede realizar con cualquier editor de texto o con PyMOL

3. Corrección y formato de las coordenadas xyz a amber (*pdb4amber*)

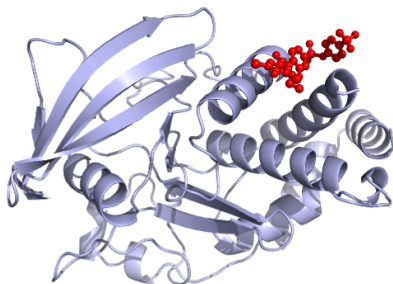
Se renumera el archivo pdb y se asigna la protonación de algunos residuos.

4. Parametrización de ligandos (antechamber y parmchk2)

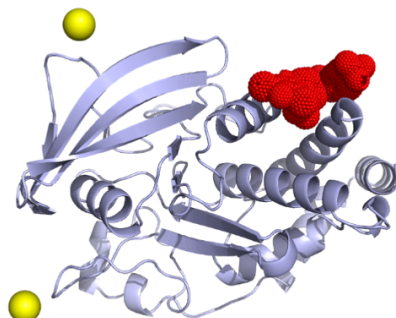


```
@TRIPOS=MOLECULE
llg
36 39 1 0 0
SMALL
bcc
@TRIPOS=ATOM
1 C01 61.8250 31.5140 21.2400 c1 301 llg -0.071500
2 C1 56.0390 32.2780 19.5100 c1 301 llg -0.019000
3 N01 59.8900 31.3950 19.1730 n2 301 llg -0.757600
4 O01 58.1540 39.5970 17.2000 o 301 llg -0.892000
5 O01 58.3950 31.5720 16.3290 o6 301 llg -2.852000
6 C02 60.1670 32.6780 20.3060 c2 301 llg 0.328000
7 C2 37.8040 31.3220 19.3990 c2 301 llg -0.007000
8 N02 59.9120 35.3260 24.5510 n1 301 llg -1.153000
9 O02 58.2710 32.9100 17.7430 o 301 llg -0.892000
10 S02 61.1870 34.2870 24.1580 s6 301 llg 2.221000
11 C03 52.7590 31.0090 22.4530 c2 301 llg 0.362000
12 C3 57.0290 30.1380 20.1280 c1 301 llg 0.096000
```

5. Construir el complejo proteína-ligando

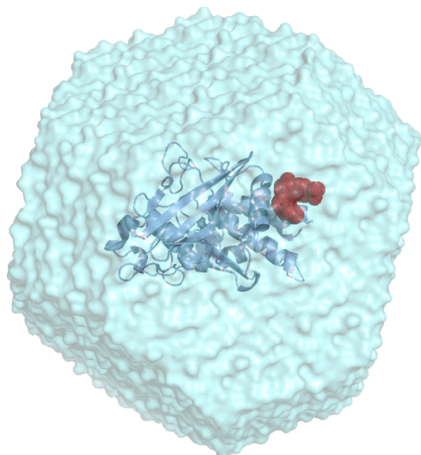


Es importante asignar al ligando un número de residuo diferente al de la macromolécula.

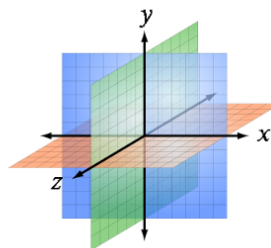
6. Neutralización del sistema con un contraión (K^+ o Cl^-)

Primero hay que identificar la carga total del sistema, para saber con qué contraión neutralizaremos.

7. Solvatación del sistema (cuadrada u octaédrica)



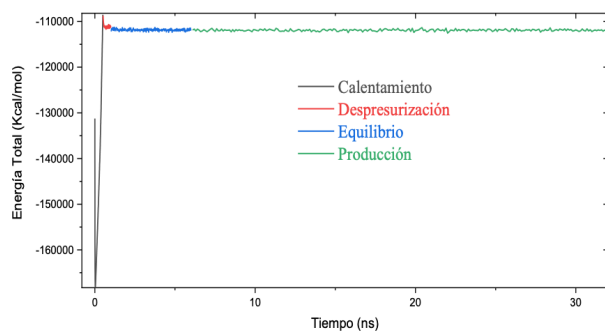
Se recomienda 12 Å de capa de solvatación.

8. Generación de archivos de topología y coordenadas xyz (*tLeap*)*Pre-Proceso*

9. Minimización (elimina impedimentos estéricos y geometrías inapropiadas).
10. Calentamiento (calienta el sistema de 0 K a la temperatura elegida, utilizando una función termostato)
11. Despresurización (mantiene la presión a 1Atmosfera)
12. Equilibrio (equilibra el sistema antes de empezar la SDM)

Proceso

13. Producción de la SDM (*sander o pmemd.cuda*)



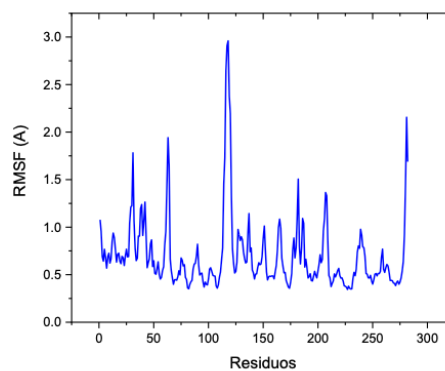
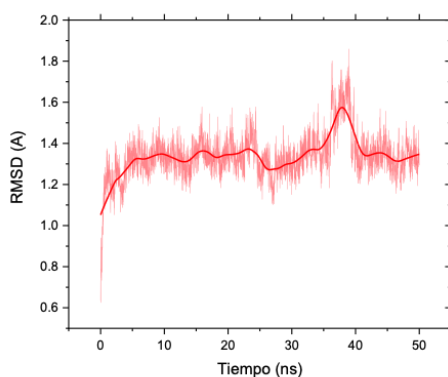
Grafica de la Energía total del sistema en función del tiempo, donde se muestran las etapas de una SDM.

Post-Proceso

14. Visualización de las SDM (VMD)



15. Ejemplos de análisis de las trayectorias de dinámica molecular con CPPTRAJ

**Ejemplo de un protocolo de Simulación de Dinámica Molecular para un complejo Proteína-Ligando**

Para este protocolo se requiere de los siguientes programas:

- I. Ambiente Unix
- II. AmbersTools2021
- III. Amber construido para GPUs
- IV. VMD y PyMOL

1. Obtenemos la estructura de la macromolécula XXX.pdb y eliminamos ligandos, aguas y protones (en PyMOL o editor de texto).
2. Preparamos a la macromolécula, ejecutando el programa pdb4amber (para corregir y ordena el pdb para ser leído por ambertools)

```
pdb4amber -i XXX.pdb -o XXX-amber.pdb
```

3. Parametrizamos el ligando colocamos el siguiente comando

```
antechamber -i XXX.pdb -fi pdb -o XXX.mol2 -fo mol2 -rn LIG -c bcc -nc 0
```

4. Generamos la topología del ligando con:

```
parmchk2 -i XXX.mol2 -f mol2 -o XXX.frcmod
```

5. Generamos los archivos de entrada para la dinámica molecular utilizando tleap

```
tleap -s -f leaprc.protein.ff19SB  
> source leaprc.water.tip3p  
> source leaprc.gaff2
```

6. Cargamos el ligando en tleap

```
> LIG=loadmol2 LIG.mol2
```

7. Cargamos el archivo frcmod

```
> loadamberparams LIG.frcmod
```

8. Verificamos que el ligando está cargado adecuadamente en tleap

> *check LIG*

9. Escribimos un nuevo pdb del ligando en tleap

> *savepdb LIG LIG2.pdb*

10. En otra terminal construimos un archivo pdb del complejo, utilizando la macromolécula (XXX-amber.pdb) y el ligando generado en el paso anterior (LIG2.pdb); nombrándolo XXX-LIG-amber.pdb

11. Regresamos a la terminal donde se está ejecutando tleap

12. Cargamos el complejo

> *complex=loadpdb XXX-amber.pdb*

13. Checamos el complejo

> *check complex*

14. Identificamos la carga del complejo

> *charge complex*

15. Neutralizamos el sistema con algún contraion según sea el caso (K⁺ o Cl⁻)

> *addions complex Cl- 0*

16. Solvamos el sistema

> *solvateoct complex TIP3PBOX 10*

17. Salvamos las topologías y coordenadas del complejo neutralizado y solvado

> *saveamberparm complex XXX-LIG-WAT.top XXX-LIG-WAT.xyz*

18. Opcionalmente, podemos salvar en formato pdb el complejo neutralizado y solvado para visualizarlos en PyMOL

> *savepdb complex XXX-LIG-WAT.pdb*

19. Salimos de tleap

> *quit*

20. Obtenemos y editamos los archivos de entrada (.in) y los archivos de ejecución (.bsub) para realizar la dinámica molecular.

min.in

```
&cntrl
imin=1,maxcyc=10000,ncyc=500,
cut=8.0,ntb=1,
ntc=2,ntf=2,
ntpr=100,
ntr=1, restraintmask=':1-282',
restraint_wt=2.0
/
```

heat.in

```
&cntrl
imin=0,irest=0,ntx=1,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,ntb=1,
ntpr=100,ntwx=100,
ntt=3,gamma_ln=2.0,
temp0=298.15,
ntr=1, restraintmask=':1-282',
restraint_wt=2.0,
nmropt=1
/
&wt TYPE='TEMP0', istep1=0, istep2=25000,
value1=0.1, value2=298.15, /
```

density.in

```
&cntrl
imin=0,irest=1,ntx=5,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,ntb=2, ntp=1, taup=1.0,
ntpr=500,ntwx=500,
ntt=3,gamma_ln=2.0,
temp0=298.15,
ntr=1, restraintmask=':1-282',
restraint_wt=2.0,
/
```

equil.in

```
&cntrl
imin=0,irest=1,ntx=5,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,ntb=2, ntp=1, taup=2.0,
ntpr=1000,ntwx=1000,
ntt=3,gamma_ln=2.0,
temp0=298.15,
/
```

prod.in

```
&cntrl
imin=0,irest=1,ntx=5,
nstlim=25000000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,ntb=2,ntp=1,taup=2.0,
ntpr=5000,ntwx=5000,
ntt=3,gamma_ln=2.0,
temp0=298.15,
/
```

min-q_gpu.bsub

```
pmemd.cuda -O -i min.in -o XXX-LIG-WAT-min.out -p XXX-LIG-WAT.top -c XXX-LIG-WAT.xyz -r
XXX-LIG-WAT-min.rst -ref XXX-LIG-WAT.xyz
```

heat-q_gpu.bsub

```
pmemd.cuda -O -i heat.in -o XXX-LIG-WAT-heat.out -p XXX-LIG-WAT.top -c XXX-LIG-WAT-min.rst -r
XXX-LIG-WAT-heat.rst -x XXX-LIG-WAT-heat.mdcrd -ref XXX-LIG-WAT-min.rst
```

density-q_gpu.bsub

```
pmemd.cuda -O -i density.in -o XXX-LIG-WAT-density.out -p XXX-LIG-WAT.top -c XXX-LIG-WAT-
heat.rst -r XXX-LIG-WAT-density.rst -x XXX-LIG-WAT-density.mdcrd -ref XXX-LIG-WAT-heat.rst
```

equil-q_gpu.bsub

```
pmemd.cuda -O -i equil.in -o XXX-LIG-WAT-equil.out -p XXX-LIG-WAT.top -c XXX-LIG-WAT-
density.rst -r XXX-LIG-WAT-equil.rst -x XXX-LIG-WAT-equil.mdcrd -ref XXX-LIG-WAT-density.rst
```

prod-q_gpu.bsub

```
pmemd.cuda -O -i prod.in -o XXX-LIG-WAT-prod.out -p XXX-LIG-WAT.top -c XXX-LIG-WAT-equil.rst
-r XXX-LIG-WAT-prod.rst -x XXX-LIG-WAT-prod.mdcrd -ref XXX-LIG-WAT-equil.rst
```

21. Ejecutamos en orden concatenando los scrips min-q_gpu.bsub > heat-q_gpu.bsub > densityq_gpu.bsub > equil-q_gpu.bsub > prod-q_gpu.bsub

22. Analizamos la DM con CPPTRAJ

23. Creamos una carpeta llamada análisis dentro de nuestra carpeta de trabajo y nos movemos a ésta.

24. Cargamos el módulo cpptraj

```
cpptraj
>parm ../XXX-LIG-WAT.top
>trajin ../XXX_LIG-WAT-prod.mdcrd
>reference ../XXX-LIG-WAT-equil.rst
>rms reference out XXX-LIG-WAT.rms @C,CA,N
>atomicfluct out XXX-LIG-WAT.rmsf :1-282@CA
byres
>trajout XXX-LIG-fragmes.pdb onlyframes
1,100,200,300,400,500
>run
```

25. Los archivos de salida se pueden procesar en algún programa de análisis estadísticos (OriginLab, gnuplot, R, Matlab, etc)

26. Para visualizar la trayectoria de dinámica molecular abrimos VMD en la ventana donde estan nuestros resultados.

27. En VMD

[file--New Molecule--Browse--XXX-LIG-WAT.top Determine file type--AMBER7 Parm Load](#)
[file--New Molecule--Browse-- XXX-LIG -WAT-heat.mdcrd Determine file type--AMBER Coordinates with Periodic Box Load](#)
[file--New Molecule--Browse-- XXX-LIG -WAT-density.mdcrd Determine file type--AMBER Coordinates with Periodic Box Load](#)

Referencias

- Sosa-Peinado, A., Leon-Cruz, E., Velazquez-Lopez, I., Matuz-Mares, D., Cano-Sanchez, P., and Gonzalez-Andrade, M. (2020) Theoretical-experimental studies of calmodulin-peptide interactions at different calcium equivalents. *J Biomol Struct Dyn*, 1-12
- Gonzalez-Andrade, M., Mata, R., Madariaga-Mazon, A., Rodriguez-Sotres, R., Del Pozo-Yauner, L., and Sosa-Peinado, A. (2013) Importance of the interaction protein-protein of the CaM-PDE1A and CaM-MLCK complexes in the development of new anti-CaM drugs. *J Mol Recognit* **26**, 165-174
- Zhang, Y. (2005) Improved pseudobonds for combined ab initio quantum mechanical/molecular mechanical methods. *J Chem Phys* **122**, 024114
- Titmuss, S. J., Cummins, P. L., Rendell, A. P., Bliznyuk, A. A., and Gready, J. E. (2002) Comparison of linear-scaling semiempirical methods and combined quantum mechanical/molecular mechanical methods for enzymic reactions. II. An energy decomposition analysis. *J Comput Chem* **23**, 1314-1322
- Rohs, R., Etchebest, C., and Lavery, R. (1999) Unraveling proteins: a molecular mechanics study. *Biophys J* **76**, 2760-2768
- Karplus, M., and McCammon, J. A. (2002) Molecular dynamics simulations of biomolecules. *Nat Struct Biol* **9**, 646-652
- Singh, D., Mahajan, S., Singh, G., and Razdan, K. (2020) Applications of Molecular Dynamic Simulations in Lipid-Based Drug Delivery System. *Assay Drug Dev Technol* **18**, 269-273
- Case, D. A., Cheatham III, T. E., Darden, T., Gohlke, H., Luo, R., Merz Jr., K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. **26**, 1668-1688
- D.A. Case, H. M. A., K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-

[file--New Molecule--Browse-- XXX-LIG -WAT-equil.mdcrd Determine file type--AMBER Coordinates with Periodic Box Load](#)
[file--New Molecule--Browse-- XXX-LIG -WAT-prod.mdcrd Determine file type--AMBER Coordinates with Periodic Box Load](#)

Agradecimientos

A los apoyos financieros otorgados por la Facultad de Medicina, UNAM, DGAPA-UNAM (PAPIIT-IN203719) y DGTIC-UNAM (LANCAD-UNAM-DGTIC-313), así como al apoyo de la Dirección General de Cómputo y de Tecnologías de Información y Comunicación, UNAM.

- Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman. (2021) Amber 2021. 2020 Ed.
- Roe, D. R., and Cheatham, T. E., 3rd. (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **9**, 3084-3095
- Leimkuhler, B. J., and Skeel, R. D. (1994) Symplectic Numerical Integrators in Constrained Hamiltonian Systems. *Journal of Computational Physics* **112**, 117-125
- Adcock, S. A., and McCammon, J. A. (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* **106**, 1589-1615
- Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoseck, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545-1614
- Kutzner, C., Páll, S., Fechner, M., Esztermann, A., de Groot, B. L., and Grubmüller, H. (2019) More bang for your buck: Improved use of GPU nodes for GROMACS 2018. **40**, 2418-2431
- Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43-56
- Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., Buch, R., Fiorin, G., Hénin, J., Jiang, W., McGreevy, R., Melo, M. C. R., Radak, B. K., Skeel, R. D., Singharoy, A., Wang, Y., Roux, B., Aksimentiev, A., Luthey-Schulten, Z., Kalé, L. V., Schulten, K., Chipot, C., and Tajkhorshid, E. (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. **153**, 044130
- Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L. P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., and Pande, V. S. (2017) OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* **13**, e1005659



DR. MARTÍN GONZÁLEZ ANDRADE

ID ORCID: 0000-0002-8910-3035

Obtuvo la Licenciatura de Químico Farmacéutico Biólogo en la Universidad Nacional Autónoma de México (UNAM). Realizó estudios de Maestría y Doctorado en ciencias bioquímicas en la Facultad de Medicina de la UNAM.

Ha realizado estancias de investigación en el Departamento de química analítica en la Universidad

Complutense de Madrid, España así como en el Departamento de farmacia en la Facultad de Química de la UNAM.

De manera complementaria, realizó estudios de modelaje molecular (dinámica molecular, acoplamiento molecular “*docking*”, relación estructura actividad, etc.) para tratar de explicar a través de modelos estructurales las interacciones de estos sistemas con sus ligandos.

Es profesor titular A de TC en la Facultad de Medicina, UNAM y miembro del Sistema Nacional de Investigadores (SNI), Nivel I.

Su área de investigación es la biotecnología aplicada en el desarrollo de biosensores basados en interacciones proteína-ligando, DNA-proteína para el desarrollo y búsqueda de nuevos compuestos activos de interés farmacológico.

Ha publicado diversos artículos de investigación original, en revistas de circulación internacional; así como capítulos de libros nacionales e internacionales. Cuenta con una patente registrada a nivel nacional (IMPI) e internacional (PCT).