



Análisis de Modos Normales en una Biomolécula

John Erick Cabrera Ramirez

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Física
Bogotá, Colombia
2016

Análisis de de Modos Normales en una Biomolécula

John Erick Cabrera Ramirez

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Físico

Director(a):
PhD, ,Yuly Edith Sánchez Mendoza

Línea de Investigación:
Biofísica Molecular
Grupo de Investigación:
Biofísica Molecular

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Física
Bogotá, Colombia
2016

(Dedicatoria o un lema)

Su uso es opcional y cada autor podrá determinar la distribución del texto en la página, se sugiere esta presentación. En ella el autor dedica su trabajo en forma especial a personas y/o entidades.

Por ejemplo:

A mis padres

o

La preocupación por el hombre y su destino siempre debe ser el interés primordial de todo esfuerzo técnico. Nunca olvides esto entre tus diagramas y ecuaciones.

Albert Einstein

Agradecimientos

Esta sección es opcional, en ella el autor agradece a las personas o instituciones que colaboraron en la realización de la tesis o trabajo de investigación. Si se incluye esta sección, deben aparecer los nombres completos, los cargos y su aporte al documento.

Resumen

El resumen es una presentación abreviada y precisa (la NTC 1486 de 2008 recomienda revisar la norma ISO 214 de 1976). Se debe usar una extensión máxima de 12 renglones. Se recomienda que este resumen sea analítico, es decir, que sea completo, con información cuantitativa y cualitativa, generalmente incluyendo los siguientes aspectos: objetivos, diseño, lugar y circunstancias, pacientes (u objetivo del estudio), intervención, mediciones y principales resultados, y conclusiones. Al final del resumen se deben usar palabras claves tomadas del texto (mínimo 3 y máximo 7 palabras), las cuales permiten la recuperación de la información.

Palabras clave: (máximo 10 palabras, preferiblemente seleccionadas de las listas internacionales que permitan el indizado cruzado).

A continuación se presentan algunos ejemplos de tesauros que se pueden consultar para asignar las palabras clave, según el área temática:

Artes: AAT: Art y Architecture Thesaurus.

Ciencias agropecuarias: 1) Agrovoc: Multilingual Agricultural Thesaurus - F.A.O. y 2) GEMET: General Multilingual Environmental Thesaurus.

Ciencias sociales y humanas: 1) Tesauro de la UNESCO y 2) Population Multilingual Thesaurus.

Ciencia y tecnología: 1) Astronomy Thesaurus Index. 2) Life Sciences Thesaurus, 3) Subject Vocabulary, Chemical Abstracts Service y 4) InterWATER: Tesauro de IRC - Centro Internacional de Agua Potable y Saneamiento.

Tecnologías y ciencias médicas: 1) MeSH: Medical Subject Headings (National Library of Medicine's USA) y 2) DECS: Descriptores en ciencias de la Salud (Biblioteca Regional de Medicina BIREME-OPS).

Multidisciplinarias: 1) LEMB - Listas de Encabezamientos de Materia y 2) LCSH- Library of Congress Subject Headings.

También se pueden encontrar listas de temas y palabras claves, consultando las distintas bases de datos disponibles a través del Portal del Sistema Nacional de Bibliotecas¹, en la sección Recursos bibliográficos. opción "Bases de datos".

Abstract

Es el mismo resumen pero traducido al inglés. Se debe usar una extensión máxima de 12

¹ver: www.sinab.unal.edu.co

renglones. Al final del Abstract se deben traducir las anteriores palabras claves tomadas del texto (mínimo 3 y máximo 7 palabras), llamadas keywords. Es posible incluir el resumen en otro idioma diferente al español o al inglés, si se considera como importante dentro del tema tratado en la investigación, por ejemplo: un trabajo dedicado a problemas lingüísticos del mandarín seguramente estaría mejor con un resumen en mandarín.

Keywords: palabras clave en inglés(máximo 10 palabras, preferiblemente seleccionadas de las listas internacionales que permitan el indizado cruzado)

Contenido

Agradecimientos	VII
Resumen	IX
Lista de símbolos	XII
1. Modelos Teóricos	2
1.1. Relevancia	2
1.2. Dinámica de una Biomolécula	2
1.3. Movimientos Globales	3
1.3.1. Modelos de Redes Elásticas (ENM)	3
1.3.2. Análisis de Modos Normales Estándar (NMA)	7
1.3.3. Análisis por Componentes Principales (PCA)	7
1.4. Movimientos Locales	7
2. Estudios Actuales del Cotransportador vSGLT	8
2.1. Transportadores	8
2.2. Algunas Familias Proteicas y enrollamiento LeuT (LeuT fold)	8
2.3. Co-transportador vSGLT	8
3. Simulación Computacional	9
4. Resultados y Análisis	10
5. Conclusiones y recomendaciones	11
5.1. Conclusiones	11
5.2. Recomendaciones	11
A. Anexo:	12
B. Anexo:	13
Bibliografía	14

Lista de símbolos

Esta sección es opcional, dado que existen disciplinas que no manejan símbolos y/o abreviaturas.

Se incluyen símbolos generales (con letras latinas y griegas), subíndices, superíndices y abreviaturas (incluir sólo las clases de símbolos que se utilicen). Cada una de estas listas debe estar ubicada en orden alfabético de acuerdo con la primera letra del símbolo.

Símbolos con letras latinas

Símbolo	Término	Unidad SI	Definición
A	Área	m^2	$\int \int dx dy$
A_{BET}	Área interna del sólido	$\frac{m^2}{g}$	ver DIN ISO 9277
A_g	Área transversal de la fase gaseosa	m^2	Ec...
A_s	Área transversal de la carga a granel	m^2	Ec...
a	Coeficiente	1	Ec...
a	Contenido de ceniza	1	$\frac{m_{ceniza}}{m_{bm,0}}$
c	Contenido de carbono	1	$\frac{m_C}{m}$
c	Longitud de la cuerda	m	Figura...
c	Concentración de la cantidad de materia	$\frac{mol}{m^3}$	$\frac{n}{V}$
D	Diámetro	m	
E_A	Energía de activación	$\frac{kJ}{mol}$	Ec....
F	Fracción de materia volátil	1	ver DIN 51720
Fr	Número de Froude	1	$\frac{\omega^2 R}{g_0}$
\vec{g}	Aceleración de la gravedad	$\frac{m}{s^2}$	$\frac{d^2 \vec{r}}{dt^2}$
H	Entalpía	J	$U + PV$

Símbolo	Término	Unidad SI	Definición
H_o	Poder calorífico superior	$\frac{\text{MJ}}{\text{kg}}$	ver DIN 51857
h	Contenido de hidrógeno	1	$\frac{m_H}{m}$
K	Coeficiente de equilibrio	1	Ec...
L	Longitud	m	DF
L	Longitud del reactor	m	Figura...
m	Masa	kg	DF
\dot{m}	Flujo de masa	$\frac{\text{kg}}{\text{s}}$	$\frac{m}{t}$
n	Velocidad de rotación	$\frac{1}{\text{s}}$	$\frac{\omega}{2\pi}$
n	Cantidad de materia	mol	DF
P	Presión	Pa	$\frac{\vec{F} \cdot \vec{n}}{A}$
Q	Calor	kJ	1. LT
T	Temperatura	K	DF
t	Tiempo	s	DF
x_i	Fracción de la cantidad de materia	1	$\frac{n_i}{n}$
V	Volumen	m^3	$\int dr^3$
\vec{u}	Velocidad	$\frac{\text{m}}{\text{s}}$	$(\frac{dr}{dt}, r \frac{dv}{dt}, \frac{dz}{dt})$
w_i	Fracción en masa del componente i	1	$\frac{m_i}{m_0}$
$w_{w,i}$	Contenido de humedad de la sustancia i	1	$\frac{m_{H_2O}}{m_{i,0}}$
Z	Factor de gases reales	1	$\frac{pv}{RT}$

Símbolos con letras griegas

Símbolo	Término	Unidad SI	Definición
α_{BET}	Factor de superficie	$\frac{\text{m}^2}{\text{g}}$	$(w_{\text{F,waf}})(A_{\text{BET}})$
β_i	Grado de formación del componente i	1	$\frac{m_i}{m_{\text{bm},0}}$
γ	Wandhaftreibwinkel (Stahlblech)	1	Sección...
ϵ	Porosidad de la partícula	1	$1 - \frac{\rho_s}{\rho_w}$
η	mittlere Bettneigungswinkel (Stürzen)	1	Figura...

Símbolo	Término	Unidad SI	Definición
θ	Ángulo de inclinación de la cama	1	Figura...
θ_O	Ángulo superior de avalancha	1	Figura...
θ_U	Ángulo inferior de avalancha	1	Figura...
κ	Velocidad de calentamiento	$\frac{K}{s}$	$\frac{dT}{dt}$
ν	Coeficiente estequiométrico	1	ver DIN 13345
ρ_b	Densidad a granel	$\frac{kg}{m^3}$	$\frac{m_S}{V_S}$ (Sección...)
ρ_s	Densidad aparente	$\frac{kg}{m^3}$	$\frac{m_F}{V_P}$ (Sección...)
ρ_w	Densidad verdadera	$\frac{kg}{m^3}$	$\frac{m_F}{V_F}$ (Sección...)
τ	Tiempo adimensional	1	Ec....
Φ_V	Flujo volumétrico	$\frac{m^3}{s}$	$\frac{\Delta V}{\Delta t}$
ω	Velocidad angular	$\frac{1}{s}$	$\frac{dv}{dt}$

Subíndices

Subíndice	Término
bm	materia orgánica
DR	Dubinin-Radushkevich
E	Experimental
g	Fase gaseosa
k	Condensado
Ma	Macroporos
P	Partícula
p	Poros
p	Pirolizado
R	Reacción
t	Total
wf	Libre de agua
waf	Libre de agua y de ceniza
0	Estado de referencia

Superíndices

Superíndice	Término
n	Coefficiente x

Abreviaturas

Abreviatura	Término
1. <i>LT</i>	Primera ley de la termodinámica
<i>DF</i>	Dimensión fundamental
<i>RFF</i>	Racimos de fruta fresca

Introducción

EDITAR El estudio de los cotransportadores de azúcar, en particular, transportadores de glucosa SGLT dependientes de sodio son esenciales en la producción de metabolismo y la energía celular. Los cotransportadores SGLT son miembros de la familia de portadores de soluto (SLC5) y algunos de estos transportadores de interés tienen una secuencia y estructura similar tridimensional similar. En este caso se examinó el co-transportador dependiente sodio galactosa del *Vibrio parahaemolyticus* (vSGLT), que media el transporte de galactosa en el citoplasma de las bacterias *Vibrio parahaemolyticus*. Según la literatura, la cinética del co-transportador tiene entre 5 y 6 estados o conformaciones, pero en este caso de que la desvinculación de los sustratos se estudia la conformación, también conocido como modelo de liberación estado de co-transportador que mira hacia dentro. se realizó un estudio computacional para analizar los movimientos globales de un transportador vSGLT, y comparamos nuestros resultados computacionales con los que se encuentran en los anteriores informes experimentales. análisis de modos normales con un modelo elástico de red (ENM) fue utilizado para explorar los cambios en los movimientos globales entre vSGLT en la presencia o ausencia de los iones que transportan (Na^+ , galactosa). ENM se ha demostrado que es un cálculo útil herramienta para predecir la dinámica de las proteínas de membrana en muchas aplicaciones. los modos normales más bajas generadas por la ENM proporcionar información valiosa sobre la dinámica global de las biomoléculas que son relevantes para su función.

■ 2 a 4 páginas

Capítulo 1

Modelos Teóricos

1.1. Relevancia

Es importante estudiar los movimientos de una biomolécula ya que como es señalado en [Lezon et al., 2009] y en [Rader et al., 2006], la dinámica de la molécula vincula la estructura con la función de la biomolécula. La función es el papel que desempeña la biomolécula y que está intimamente relacionado con las interacciones de la biomolécula a un ligando. La estructura de una biomolécula tiene 4 niveles de organización, denominadas *estructuras primaria, secundaria, terciaria y cuaternaria*.

Por ejemplo, en una proteína la estructura primaria es la secuencia u orden de los monómeros (aminoácidos) que la constituyen. La estructura secundaria está determinada por los puentes de hidrógeno existentes entre los grupos amino y carboxilo de dos aminoácidos diferentes; predice la estructura tridimensional en forma local. Mientras que la estructura terciaria da el arreglo tridimensional de cada uno de sus átomos, específicamente se determina con las coordenadas de cada uno de los átomos. La estructura cuaternaria es el arreglo de unidades proteicas o cadenas peptídicas formando lo que se conoce como complejo multiproteico.

Algunas proteínas están en la clase de proteínas de transporte que se encuentran en la membrana celular, la función que cumplen es actuar como mediadoras para el transporte de iones y sustratos, convirtiéndose este en un paso previo al metabolismo y la energética celular al interior de la célula.

1.2. Dinámica de una Biomolécula

La dinámica de una biomolécula se determina por las ecuaciones de movimiento para cada uno de los átomos que la constituyen. Usualmente en una biomolécula el número de monómeros es mayor a 20, que al multiplicarlo por el número de átomos en cada monómero incrementa considerablemente el número de ecuaciones de movimiento a resolver, de ahí que

sea necesario realizar *dinámica molecular* (Molecular Dynamics que por sus siglas en inglés es MD) la cual estudia mediante simulaciones computacionales el movimiento de los átomos, de acuerdo a las interacciones que presenten.

Las ecuaciones de movimiento se pueden conocer a partir de los formalismos lagrangiano o hamiltoniano [Goldstein et al., 2001], en los cuales es necesario conocer los potenciales con los que interactúan los átomos. Las soluciones a las ecuaciones de movimiento se encuentran mediante los métodos de la dinámica molecular o los análisis de modos normales (Normal Mode Analysis que por sus siglas en inglés es NMA) en los cuales se escogen los modelos de potencial.

Los diversos modelos de potencial pueden ser tomados según la naturaleza del polímero a analizar, ver [Lezon et al., 2009]. Sin embargo, al escoger el potencial para hacer un análisis *in silico* de la dinámica de una biomolécula, debe tenerse en cuenta el costo computacional requerido, esto es, el tiempo de simulación de la molécula y la exactitud requerida en el movimiento de cada uno de los constituyentes de la molécula.

De acuerdo a los parámetros de costo y tiempo, las simulaciones de biomoléculas se pueden hacer analizando los *movimientos locales* y los *movimientos globales*.

1.3. Movimientos Globales

Son aquéllas simulaciones en las que se desean conocer los *cambios globales* o el aspecto general que exhibe el movimiento de una biomolécula haciendo simplificaciones, ya sea en los potenciales presentes o en el número de átomos interconectados. Este tipo de simulaciones pueden ser realizadas a un orden de magnitud de los microsegundos, lo cual facilita su uso en computadores personales, al respecto ver [Gur et al., 2013].

Un conjunto de modelos que permite calcular los movimientos globales de una molécula son los *Modelos de Redes Elásticas* (Elastic Network Models o ENM por sus siglas en inglés). Otros modelos que describen los movimientos globales son los análisis por componentes principales (Principal Component Analysis o PCA por sus siglas en inglés) y el análisis por modos normales estándar (Normal Mode Analysis o NMA por sus siglas en inglés).

1.3.1. Modelos de Redes Elásticas (ENM)

Los ENM, como la palabra *elástico* lo indica, se basan en una simplificación de la energía potencial a una energía potencial elástica, es decir de tipo Hooke. Un requisito para que sea posible hacer dicha simplificación, es la de minimizar la energía potencial.

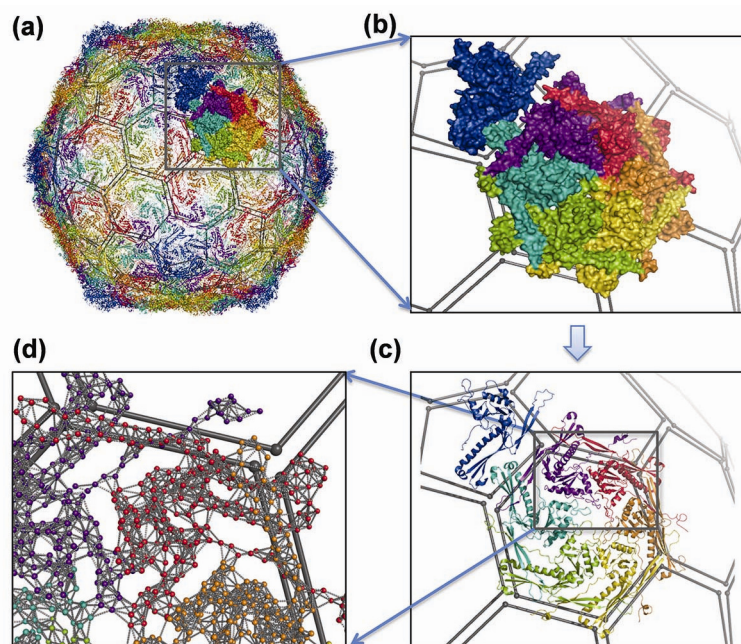


Figura 1-1.: (a) Vista exterior de un cápside vírico HK97 coloreado por cada cadena, todas las proteínas son idénticas. (b) Vista del arreglo proteínico en una cara del cápside. (c) Vista de la estructura secundaria de las proteínas (d) Esquema de cada proteína mostrando cada uno de sus átomos, las aristas de cada cara son carbonos α unidos por lados (ligaduras elásticas). Tomado de [Lezon et al., 2009].

Al simplificar el potencial, la biomolécula original se convierte en una red cuyos nodos están sometidos al potencial elástico, ver figura 1-1. Los nodos se consideran como bloques constituyentes de la biomolécula y no siempre los nodos coincidirán con cada uno de los átomos en la biomolécula. La elección del bloque constituyente depende de la compatibilidad del modelo con los datos experimentales, que se encuentra reflejado en la estabilidad de los enlaces con respecto a su posición de equilibrio, de tal manera que cada bloque constituyente pueda considerarse como una partícula puntual o incluso como un cuerpo rígido.

Descripción Mecánica del Modelo

Considérese una biomolécula con N partículas constituyentes, el tipo de constituyente depende del modelo apropiado para la biomolécula, por ejemplo en las proteínas como la BPTI, ver [Gur et al., 2013], los constituyentes son los carbonos α de los aminoácidos.

La energía potencial V que representa las interacciones entre los constituyentes de la biomolécula, se puede expresar alrededor de las posiciones de equilibrio $\mathbf{q}_0 = \mathbf{0}$ tal como describe

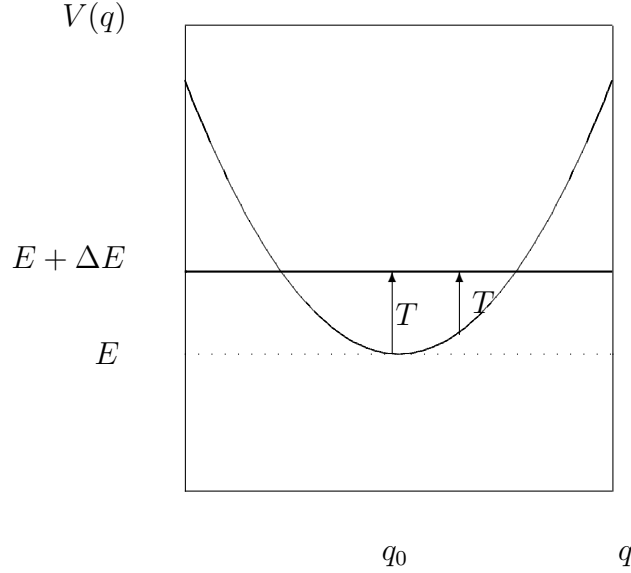


Figura 1-2.: Potencial en función de la posición y energía total E

la teoría de pequeñas oscilaciones, ver [Goldstein et al., 2001]:

$$V(q) = V(\mathbf{0}) + \sum_{i=1}^n \frac{\partial V}{\partial q_i} q_i + \sum_{ij} \frac{\partial^2 V}{\partial q_i \partial q_j} q_i q_j + \dots \quad (1-1)$$

Donde q_i son los desplazamientos con respecto a las posiciones de equilibrio, n es el número de posibles desplazamientos en la biomolécula. $V(\mathbf{0})$ es el potencial en equilibrio que por conveniencia puede ser calibrado a cero: $V(\mathbf{0}) = 0$.

El sistema se encuentra alrededor del equilibrio cuando las fuerzas generalizadas se anulan, esto es:

$$\frac{\partial V}{\partial q_i} = 0 \quad (1-2)$$

En este tipo de casos como la energía se minimiza, se dice que hay un equilibrio estable. Para entender esto, supóngase que la energía total, que en este caso es $E = T + V$ corresponde al punto de equilibrio donde los nodos no se mueven, es decir, $E = V_{min}$ (línea punteada de la figura 1-2). Ahora suponga que hay un incremento de la energía total en una cantidad ΔE , este incremento generará un aumento en la energía potencial, línea continua. Si el sistema se desvía de la posición de equilibrio, por conservación de la energía y como $E = T + V$, la energía cinética disminuye. En un equilibrio inestable la energía cinética aumenta, alejando al sistema del equilibrio. [Goldstein et al., 2001]

Considerando la condición de equilibrio (1.3.1) y despreciando desplazamientos de orden

superior se tiene que:

$$V(\mathbf{q}) = \sum_{i,j=1}^n \frac{\partial^2 V}{\partial q_i \partial q_j} q_i q_j \quad (1-3)$$

Donde se identifican las constantes elásticas como:

$$U_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j} \quad (1-4)$$

En términos matriciales el potencial se puede escribir como:

$$V(\mathbf{q}) = \mathbf{q}^t \mathbf{U} \mathbf{q} \quad (1-5)$$

En la ecuación (1.3.1) \mathbf{q} es el vector columna formado por los desplazamientos de las posiciones de equilibrio para cada constituyente, i. e., $\mathbf{q}^t = (q_1, q_2, \dots, q_n)$. Por otro lado la matriz de constantes elásticas \mathbf{U} , es una matriz simétrica debido a que la fuerza generalizada se considera conservativa, lo cual permite intercambiar el orden de las derivadas parciales.

Para las pequeñas oscilaciones, no existen ligaduras dependientes explícitamente de el tiempo (holonómicas) luego, la energía cinética de los constituyentes sólo dependerá de los cuadrados de las velocidades generalizadas:

$$T = \frac{1}{2} \mathbf{q}^t \mathbf{M} \mathbf{q} \quad (1-6)$$

Donde \mathbf{M} es la masa generalizada, la cual se expresa en términos de los factores de escala entre sistemas coordenados:

$$M_{jk} = \sum_{i=1}^N m_i \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} \quad (1-7)$$

En este caso:

$$\frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = \delta_{ij} \delta_{jk} \quad (1-8)$$

Ya que las coordenadas de posición son lineales a los desplazamientos de las posiciones de equilibrio. Esto permite llamar a \mathbf{M} simplemente la masa del sistema.

Tamando en cuenta que

Ensamble Estadístico

Modelo de Redes Gaussianas (GNM)

Modelos de Redes Anisotrópicas (ANM)

1.3.2. Análisis de Modos Normales Estándar (NMA)

Describe un sistema oscilatorio en el que todos los constituyentes del sistema oscilan sinusoidalmente y con la misma frecuencia.

1.3.3. Análisis por Componentes Principales (PCA)

1.4. Movimientos Locales

Hace referencia a las simulaciones en las que se incluyen todos los átomos junto con las interacciones presentes, es decir, en las que se analizan los *cambios locales*. Estas se pueden simular a un orden de magnitud de los nanosegundos en una máquina usual, al respecto ver [Gur et al., 2013].

Como caso particular se pueden tomar los potenciales usados en [Amber, 2016], que siguen el modelo de Amber. El modelo de Amber tiene en cuenta las contribuciones debidas a:

- Interacciones intermoleculares: Son las producidas por los enlaces covalentes entre grupos de átomos, las de valencia y las torsiones.
- Interacciones entre pares: Lennard Jones, electrostático.

Análisis por Componentes Principales

-

Estudios Actuales del Cotransportador vSGLT

El cotransportador de Na^+ /Galactosa presente en la proteobacteria *Vibrio Parparahaemolyticus* denotado como vSGLT, es un simportador presente en la familia de simportadores de sodio-soluto SSS.

2.1. Transportadores

2.2. Algunas Familias Proteicas y enrollamiento LeuT (LeuT fold)

2.3. Co-transportador vSGLT

El co-transportador es similar en la estructura primaria y terciaria a otros transportadores de la familia SLC5 como por ejemplo NIS, SGLT1, SGLT2.

Simulación Computacional

Capítulo 4

Resultados y Análisis

Conclusiones y recomendaciones

5.1. Conclusiones

Las conclusiones constituyen un capítulo independiente y presentan, en forma lógica, los resultados de la tesis o trabajo de investigación. Las conclusiones deben ser la respuesta a los objetivos o propósitos planteados. Se deben titular con la palabra conclusiones en el mismo formato de los títulos de los capítulos anteriores (Títulos primer nivel), precedida por el numeral correspondiente (según la presente plantilla).

5.2. Recomendaciones

Se presentan como una serie de aspectos que se podrían realizar en un futuro para emprender investigaciones similares o fortalecer la investigación realizada. Deben contemplar las perspectivas de la investigación, las cuales son sugerencias, proyecciones o alternativas que se presentan para modificar, cambiar o incidir sobre una situación específica o una problemática encontrada. Pueden presentarse como un texto con características argumentativas, resultado de una reflexión acerca de la tesis o trabajo de investigación.

Apéndice A

Anexo:

Apéndice B

Anexo:

Bibliografía

- [Amber, 2016] Amber, C. (2016). Amber 2016 Reference Manual.
- [Goldstein et al., 2001] Goldstein, H., Poole, C. P., and Safko, J. (2001). *Classical Mechanics*, volume 1. Adison Wesley, Pearson, 3 edition.
- [Gur et al., 2013] Gur, M., Zomot, E., and Bahar, I. (2013). Global motions exhibited by proteins in micro- to milliseconds simulations concur with anisotropic network model predictions. *The Journal of chemical physics*, 139(12):121912.
- [Lezon et al., 2009] Lezon, T. R., Shrivastava, I. H., Yang, Z., and Bahar, I. (2009). Elastic Network Models For Biomolecular Dynamics: Theory and Application to Membrane Proteins and Viruses. *Handbook on Biological Networks*, pages 129–158.
- [Rader et al., 2006] Rader, A. J., Chennubhotla, C., Yang, L.-W., and Bahar, I. (2006). The Gaussian Network Model: theory and applications. *Normal Mode Analysis - theory and applications to biological and chemical systems*, 10(20):41–64.