

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

COLEGIO DE CIENCIAS Y HUMANIDADES

MAESTRÍA EN DINÁMICA NO LINEAL Y SISTEMAS
COMPLEJOS

Física Estadística de Cadenas Heteropoliméricas
ideales. Modelo de grano grueso basado en la
aproximación de Zwanzig

TESIS QUE PARA OBTENER EL GRADO DE MAESTRO EN
DINÁMICA NO LINEAL Y SISTEMAS COMPLEJOS

PRESENTA:

Arturo Rodríguez Chacón

Director de tesis:

Dr. Luis Agustín Olivares Quiroz

México, D.F. Agosto, 2015.

SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN



UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA

RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

DERECHOS RESERVADOS[©]

La presente obra y cada uno de sus elementos está protegido por la Ley Federal del Derecho de Autor; por la Ley de la Universidad Autónoma de la Ciudad de México, así como lo dispuesto por el Estatuto General Orgánico de la Universidad Autónoma de la Ciudad de México; del mismo modo por lo establecido en el Acuerdo por el cual se aprueba la Norma mediante la que se Modifican, Adicionan y Derogan Diversas Disposiciones del Estatuto Orgánico de la Universidad de la Ciudad de México, aprobado por el Consejo de Gobierno el 29 de enero de 2002, con el objeto de definir las atribuciones de las diferentes unidades que forman la estructura de la Universidad Autónoma de la Ciudad de México como organismo público autónomo y lo establecido en el Reglamento de Titulación de la Universidad Autónoma de la Ciudad de México.

Por lo que el uso de su contenido, así como cada una de las partes que lo integran y que están bajo la tutela de la Ley Federal de Derecho de Autor, obliga a quien haga uso de la presente obra a considerar que solo lo realizará si es para fines educativos, académicos, de investigación o informativos y se compromete a citar esta fuente, así como a su autor ó autores. Por lo tanto, queda prohibida su reproducción total o parcial y cualquier uso diferente a los ya mencionados, los cuales serán reclamados por el titular de los derechos y sancionados conforme a la legislación aplicable.

Índice general

1. INTRODUCCIÓN	7
1.1. Planteamiento	7
1.2. Objetivos generales y particulares.	10
1.2.1. Objetivos Generales	10
1.2.2. Objetivos Particulares.	10
1.3. Marco teórico	10
1.3.1. Estructura de las proteínas.	13
1.3.2. Enlace peptídico.	20
2. METODOLOGÍA	26
2.1. Modelos de grano grueso (<i>Coarsed-Grained Models</i>)	26
2.2. Termodinámica y Física estadística	28
2.2.1. Conceptos básicos de termodinámica	28
2.2.2. Física Estadística.	30
2.3. Función de partición Z	32
2.3.1. Transiciones de fase configuracionales	35
2.3.2. Modelo de Zwanzig para homopolímeros ideales	39
3. MODELO DE HETEROPOLIMERO IDEAL	47
3.1. Modelo para un heteropolímero basado en el Modelo de Zwanzig	47
3.2. Comparación del MHI con el modelo de Zwanzig	53
3.2.1. Caso de estudio I: Mutaciones en péptidos	55
3.2.2. Caso de estudio II: Heteropolímeros aleatorios	60
3.2.3. Caso de estudio III: Péptidos opiáceos	66
3.2.4. Antecedentes e importancia	68
3.2.5. $[met^5]$ -Encefalina, $[leu^5]$ -Encefalina y sus mutaciones	71
3.2.6. Degradaciones de la $[met^5]$ -encefalina y $[leu^5]$ -encefalina	76
3.2.7. Caso de estudio IV: comparación del Cp experimental con el Cv del MHI para proteínas reales	79
4. CONCLUSIONES Y PERSPECTIVAS	87
4.1. Conclusiones	87
4.2. Perspectivas	91

4.2.1. Espaciamiento y acotamiento del espectro de energías E_α	91
4.2.2. Expresión analítica de la función de degeneracidad G_α para heteropolímeros	93
4.2.3. Interacción a primeros y segundos vecinos en homopolímeros	94
4.2.4. Simulaciones numéricas	100
A. Función de degeneracidad g_α y función de partición Z	102
A.1. Función de degeneración g_α del modelo de Zwanzig	102
A.2. Función de partición del modelo de Zwanzig	103
B. Códigos en <i>Mathematica</i>	105
B.1. Función de degeneracidad G	105
B.2. Función de degeneracidad G con elección aleatoria	106
C. Partícula Cuántica en una caja.	108

DATOS DEL JURADO

- Presidente: Dr. Luis Agustín Olivares Quiroz. UACM.
- Secretario: Dr. Fernando Ramirez Alatraste. UACM.
- Vocal: Dr. Damian Hernández Herrán. UACM.
- Vocal: Dr. Rafael Barrio Paredes. UNAM.
- Vocal: Dr. Orlando Guzmán López. UAM-I.

AGRADECIMIENTOS

- *A mi asesor de tesis, Dr. Luis A. Olivares Quiroz por todo el apoyo que me ha brindado en varios temas y la paciencia que tuvo al explicármelos. Muchas gracias Luis.*
- *A mis sinodales, Dr. Fernando Ramirez Alatrister. Dr. Damian Hernández Herrán. Dr. Rafael Barrio Paredes. Dr. Orlando Guzman Lopez por aceptar con gusto ser parte de mi jurado, corregir mis errores en la redacción de la tesis y enseñarme en sus clases..*
- *Al posgrado de Dinámica no Lineal y Sistemas Complejos. UACM.*
- *Proyecto SECITI No. P2011-14 (Financiamiento)*
- *A todos mis amigos que me han apoyado: Rodrigo, Julio, Ivonne, Carlos, Oro, Isaac, Ivan, Delia, Nohemi, Itzel, Ricardo, David, Alain, Miguel, Victor, Kiyoshi, Mario, etc.*

DEDICATORIAS

- *A mi hija, Sofía Rodríguez Ventre, por inspirar a seguir adelante. Desde el primer día que supe que venías.*
- *A Kenya, por el gran apoyo que me ha brindado, pese a todo. Por estar en el camino conmigo y comprender que todo logro requiere sacrificios.*
- *A mis Padres, Catalina Chacón Uribe y José Rogelio Rodríguez Talavera, los cuales son mi fuente de inspiración. Gracias por el apoyo incondicional y amor que siempre me brindaron a lo largo de este camino. Sin su ejemplo, este trabajo no lo hubiese podido realizar.*
- *A mis hermanos, Rogelio, Alejandra y Gabriela. Parte fundamental de mi trayecto y de mi vida. Nuestro pasado ayudo a forjar nuestro futuro.*
- *A mis cuñados, sobrinos y toda mi familia en general por todo el apoyo que hemos recibido de su parte.*
- *A mi amigo, el señor Mario Raul Camacho Medina (QEPD) por el apoyo y palabras que me brindo en momentos clave. Donde quiera que se encuentre, gracias.*
- *A toda la gente que ha dejado huella en mi. Gracias.*

Capítulo 1

INTRODUCCIÓN

1.1. Planteamiento

Las proteínas son macromoléculas biológicas que los organismos vivos utilizan para realizar una gran diversidad de funciones bioquímicas. Están formadas por aminoácidos, moléculas cuya estructura esta bien definida, habiendo 20 diferentes en la naturaleza[1] y donde se diferencian por tener en su estructura una cadena llamada *lateral* distinta entre si. Las proteínas constituyen gran parte de las unidades estructurales con las que se forman las células y sus múltiples formas dan como resultado una gran cantidad de funciones, las cuales pueden ser agrupadas en las siguientes categorías:

- Estructural. Dan soporte mecánico, por ejemplo, el colágeno y la queratina.
- Transportadoras. Transportan moléculas dentro de los organismos, por ejemplo, la hemoglobina, la cual es capaz de transportar oxígeno en la sangre [2].
- Catalíticas. Tienen la capacidad de acelerar procesos bioquímicos, como es el caso de las enzimas digestivas, las cuales pueden descomponer los alimentos, por ejemplo, las enzimas digestivas [3].

Todas las funciones bioquímicas que realizan las proteínas están relacionadas con la secuencia que tienen, llamada: estructura primaria, la cual consiste únicamente en la secuencia específica de aminoácidos que la forman, sin hacer referencia de los enlaces covalentes, peptídicos disulfuro que se crean en dicha estructura. La estructura secundaria consiste en la formación de hélices α y hojas β cuya formación depende de la secuencia

de aminoácidos involucrados.

El siguiente nivel de organización es la estructura tridimensional que adopta la cadena polipeptídica cuya forma es única y en la cual intervienen todos los residuos en la cadena. Esta última forma de organización estructural da lugar a que toda la proteína se doble sobre sí misma, es decir, se enrolle con el fin de realizar alguna actividad biológica específica[1]. Dicho plegamiento puede ser realizado de forma espontánea o en ocasiones es necesaria la asistencia de otra proteína distinta, la cual es capaz de ayudar a que este proceso de plegamiento se realice por completo. Este estado se le conoce como el *estado nativo*. Puede ocurrir que alguna cadena polipeptídica tenga una secuencia de aminoácido erróneo, es decir, que algún elemento de la misma no corresponda a la secuencia que conduce al estado nativo. Esto puede originar que el plegamiento se realice de forma distinta y la forma tridimensional lo sea también, adoptando un estado no nativo, cuyas funciones bioquímicas puedan cambiar, generando procesos anómalos dentro de los organismos, lo que se traduce como una enfermedad o padecimiento en muchas ocasiones.

Existen varias enfermedades que están relacionadas con el plegamiento incorrecto de las proteínas [6], del cual más adelante se hablará con más detalle, es responsable de patologías como son las llamadas *amiloidosis* las cuales conllevan a la formación irreversible de agregados fibrilares insolubles llamados amiloides. En este grupo hay diversas enfermedades como son el Alzheimer, la diabetes tipo II y la encefalopatía espongiiforme bovina entre otras[11]. Las amiloidosis están relacionadas con anomalías en la estructura nativa de las proteínas, cuya causa puede ser hereditaria o adquirida. Se caracterizan por el depósito extracelular amiloide[47], en órganos como cerebro, bazo, riñón e hígado, de proteínas precursoras particulares, o sus fragmentos. Toda esta problemática origina un gran interés para el estudio del comportamiento de este tipo de procesos como lo es el plegamiento de proteínas [7].

Por otro lado, la física estadística es una rama de la física, la cual, a partir de conceptos de la teoría atómica, estudia las propiedades macroscópicas fundamentales de sistemas formados de muchos átomos o moléculas. En dichos sistemas se incluyen los gases, líquidos, sólidos y recientemente se ha extendido su uso para el estudio de macromoléculas biológicas[8]. El comportamiento dinámico de los átomos está bien definido a partir, tanto de las leyes de la mecánica cuántica y de la mecánica clásica, así como de las interacciones entre partículas debido a fuerzas electromagnéticas. Estos últimos puntos acerca del conocimiento de las leyes de la física microscópica nos permiten deducir, en principio, las propiedades de cualquier sistema macroscópico en equilibrio, en particular, de sistemas biológicos [9].

Un comportamiento importante de los sistemas macroscópicos en la física estadística son las transiciones de fase, las cuales son cambios en la estructura homogénea de un sistema, la cual se conoce como *fase*. Estos estados dependen de las condiciones experimentales a las cuales este sometido el sistema como pueden ser cambios en la temperatura, presión, campo magnético o eléctrico. Dichos cambios se ven reflejados en el comportamiento de la energía libre como función de variables termodinámicas que competen al sistema, es decir, presentan alguna discontinuidad precisamente en dicha transición.

Las diversas transiciones sólido-líquido-gas pueden ser clasificadas como transiciones de primer orden, ya que implican un cambio discontinuo en la densidad, que es la primera derivada de la energía libre con respecto al potencial químico. Transiciones de fase de segundo orden son continuas en la primera derivada, pero exhiben una discontinuidad en la segunda derivada de la energía libre. Esta incluyen la transición de fase en materiales ferromagnéticos tales como el hierro, en donde la magnetización, que es la primera derivada de la energía libre con respecto a la intensidad del campo magnético aplicado, aumenta continuamente desde cero a medida que se baja la temperatura por debajo de la temperatura de Curie¹. Esta clasificación se le denomina: clasificación de Ehrenfest [12].

Modelos estadísticos se han hecho para describir el comportamiento termodinámico del plegamiento de las proteínas [84], considerando efectos de interacción con agua. Dichos modelos intentan reproducir resultados, como el calor específico, similares a los obtenidos experimentalmente, dado que consideran las posibles interacciones a las que están sujetas las proteínas. Algunas de las cuestiones de interés se pueden abordar a partir de estos modelos están relacionadas con las transiciones de fase que tienen estas cadenas polipeptídicas y los factores que afectan dicha transición.

A partir de herramientas de la Física estadística es posible calcular los indicadores termodinámicos como es la energía interna a partir de la función de partición Z donde esta última variable puede ser medida a partir del Cv , con el fin de poder caracterizar el tipo de transición de fase que puede presentar cualquier cadena heteropolimérica al variar su temperatura.

¹Temperatura correspondiente a la transición entre ferromagnetismo y paramagnetismo, llamada también punto de Curie.

1.2. Objetivos generales y particulares.

1.2.1. Objetivos Generales

En el presente trabajo desarrollaremos un modelo de grano grueso basado en la aproximación de Zwanzig para cadenas heteropoliméricas y calcular sus propiedades termodinámicas, particularmente los potenciales termodinámicos y así obtener las temperaturas de transición en función de la secuencia de aminoácidos de una cadena heteropoliméricas.

1.2.2. Objetivos Particulares.

- Calcular la generalización de la función de degeneración para el modelo de Zwanzig G_{α}^K que presentan los heteropolímeros ideales. Posteriormente obtener una expresión para la función de partición de heteropolímeros Z_{het} para el cálculo de los indicadores termodinámicos a partir de nuevas variables no antes definidas en el modelo original.
- Estudiar el comportamiento de las variables termodinámicas de heteropolímeros aleatorios a partir de las posibles mutaciones con el aminoácido prolina que puedan tener en algún residuo de su secuencia. Posteriormente observar el comportamiento de las temperaturas de desnaturalización que presenta cada una de ellas.
- Analizar el comportamiento de péptidos pequeños como son los neuropéptidos *[met⁵]-encefalina* y *[leu⁵]-encefalina* mediante la función de partición Z_{het} , los cuales son de gran importancia en los mecanismos de regulación de dolor, entre otras funciones. y calcular las temperaturas de transición que presentan dichas cadenas.
- Comparar los resultados calculados del modelo de heteropolímero ideal con resultados reales obtenidos mediante mediciones experimentales del C_p en vacío y con algún resultado de dinámica molecular.

1.3. Marco teórico

La física es una ciencia que estudia las propiedades de la materia, así como sus interacciones. La aplicación de sus métodos en diferentes áreas como química, biología, economía, sociología, etc. ha hecho que esta ciencia sea de amplio interés para las ramas multidisciplinarias. En particular,

la física biológica es una rama de la biología la cual, con principios de la física, pretende describir los comportamientos de sistemas vivos a escalas microscópicas y obtener información de sus propiedades macroscópicas. Ejemplos en ese sentido son la biomecánica[13], la cual estudia el movimiento y equilibrio de los seres vivos a través del análisis de las fuerzas internas y externas que actúan sobre algún cuerpo. Esta rama tiene una gran variedad de aplicaciones incorporadas a la práctica médica. Un ejemplo de aplicación son las sofisticadas prótesis en miembros ausentes en el cuerpo o incluso las válvulas cardíacas, las cuales tienen la función de controlar el flujo de sangre en el cuerpo, entre otros campos de la biología abordada por la física.

Una aplicación reciente de la física a la biología es en el tema del plegamiento de las proteínas, las cuales son polímeros que no son lo suficientemente grandes como para poderlos tratar como un sistema mecánico, a la vez que no son lo suficientemente pequeños como para tratarlos como moléculas simples en solución. Las proteínas son imprescindibles para diversas funciones en los sistemas biológicos, por ejemplo, el crecimiento humano. Realizan una enorme cantidad de funciones bioquímicas diversas, entre las que destacan la función enzimática[14], la cual ayuda a realizar una reacción química en un tiempo menor en comparación del mismo proceso sin presencia de dicha enzima, sin alterar sus propiedades. Estas enzimas de naturaleza proteica, funcionan disminuyendo la energía de activación ΔG de una reacción ², de forma que se acelera sustancialmente la tasa de reacción.

Las enzimas no alteran el balance energético de las reacciones en que intervienen, ni modifican, por lo tanto, el equilibrio de la reacción, pero consiguen acelerar el proceso incluso millones de veces. Una reacción que se produce bajo el control de una enzima, o de un catalizador en general, alcanza el equilibrio mucho más deprisa que la correspondiente reacción no catalizada. En la figura [1.1] se muestra un ejemplo de la enzima Triosa fosfato isomerasa [6], la cual está involucrada en la transformación de azúcares en energía dentro de las células.

La función estructural [4] es de vital importancia para los organismos vivos debido a que este tipo de proteínas se agregan a otras moléculas de la misma proteína para originar una estructura mayor. Un ejemplo de este tipo de proteínas es el colágeno[15], el cual es capaz de formar estructuras que resisten las fuerzas de tracción. Su diámetro en los diferentes tejidos es muy variable y su organización también; en la piel de los mamíferos están

²La energía de activación en química y biología es la energía que necesita un sistema antes de poder iniciar un determinado proceso. La energía de activación suele utilizarse para denominar la energía mínima necesaria para que se produzca una reacción química dada.

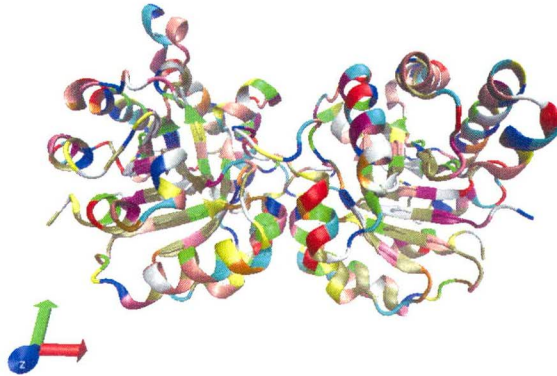


Figura 1.1: Estructura molecular de la Triosa fosfato isomerasa, enzima implicada en la glucólisis, proceso por el cual las azúcares son transformadas en energía. Imagen creada a partir del archivo pdb1ney.ent y realizada por el programa Visual Molecular Dynamics (VMD) 1.9.2).

organizadas como cestos o sillas de mimbre, lo que permite la oposición a las tracciones ejercidas desde múltiples direcciones. En los tendones lo están en haces paralelos que se alinean a lo largo del eje principal de tracción [16]. En el tejido óseo adulto y en la córnea se disponen en láminas delgadas y superpuestas, paralelas entre sí, mientras las fibras forman ángulo recto con las de las capas adyacentes.

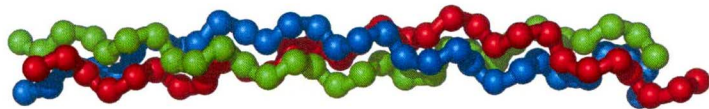


Figura 1.2: Estructura molecular de la triple hélice del colágeno. (Collagentriplehelix. Publicado bajo la licencia Creative Commons Attribution-Share Alike 3.0 vía Wikimedia Commons).

Existen proteínas que tienen la capacidad de transportar átomos o moléculas a través de canales de comunicación como es el caso de la hemoglobina. La hemoglobina es capaz de transportar el oxígeno a los

tejidos y el dióxido de carbono desde los tejidos hasta los pulmones que lo eliminan por medio de la sangre, entre otras funciones, la razón de esta función es debido a que posee un átomo de hierro (Fe) el cual es capaz de unir de forma reversible a una molécula de oxígeno[17].

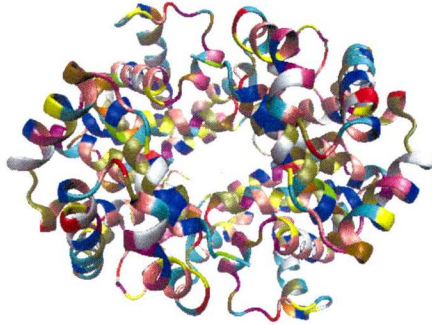


Figura 1.3: Estructura molecular de la hemoglobina. Esta proteína tiene una función de transporte la cual lleva una molécula de oxígeno de los órganos respiratorios hasta los tejidos. Figura creada a partir del archivo pdb1c7d.ent, obtenido del Protein Data Bank y visualizado por el programa VMD 1.9.2.

La figura [1.2] muestra la estructura de la molécula de colágeno, la cual está formada por tres cadenas, que se mantienen unidas por los puentes de hidrógeno al menos en tres cuartas partes de la longitud de la cadena.

En las siguientes subsecciones hablaremos más estructuradamente de estas moléculas que son las proteínas para tener un panorama amplio del tema, mencionando los puntos más relevantes que describen a estas moléculas.

1.3.1. Estructura de las proteínas.

Las proteínas están formadas por moléculas más pequeñas, llamadas aminoácidos. Los aminoácidos, también llamados residuos, son moléculas cuya estructura consiste en un carbono central llamado *carbono* C_{α} , cuyo número de valencia es 4, por lo que solamente es posible tener 4 enlaces covalentes[18]. En primer lugar, realiza un enlace covalente con un grupo carboxilo (-COOH), en el enlace contrario se encuentra un grupo amino

($-NH_2$), el tercer enlace lo realiza con un átomo de hidrógeno. El cuarto enlace esta hecho con la cadena lateral R la cual hace la diferencia entre los 20 aminoácidos que existen. Cuando los residuos se unen entre si toman el nombre de péptidos, siendo llamado enlace peptídico a la unión de dos aminoácidos. En la figura 1.4 es presentada la estructura de un aminoácido junto con sus cuatro enlaces.

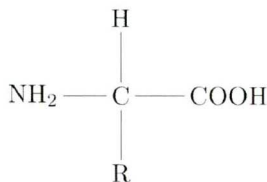


Figura 1.4: Representación de una molécula de un aminoácido. A la izquierda y derecha del carbono central se encuentran los grupos amino y carboxilo respectivamente. En el enlace de arriba se encuentra un átomo de hidrógeno y abajo la cadena lateral R, la cual es característica de cada aminoácido..

Una propiedad que tienen los aminoácidos es la quiralidad, la cual consiste en que ningún residuo puede superponerse con su imagen especular. Para un mejor entendimiento de este concepto, en la imagen 1.5 se muestra un aminoácido con su imagen especular, donde se nota que no es posible superponer a los residuos.

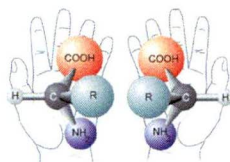


Figura 1.5: Imágenes especulares de un aminoácido. Se observa que al superponer las dos imágenes no pueden coincidir entre ellas, al igual que las manos del cuerpo humano.

Dentro de un medio acuoso, las proteínas se pliegan, es decir, adquieren una estructura tridimensional debido a las interacciones que existen entre la molécula en si misma y las moléculas de agua. Algunos aminoácidos con carga buscan estar en contacto con el agua, llamados hidrofílicos. Otros aminoácidos repelen las moléculas de agua, llamados hidrofóbicos. Estas interacciones hacen que los aminoácidos estén más cercanos unos con los otros

de tal forma que también interaccionan entre sí: se atraen, se repelen a otros, incluso forman enlaces covalentes entre ellos y también enlaces disulfuros o del tipo puente de hidrógeno, afectando aún más la forma de la cadena ya incluso plegada. Estas interacciones originan estructuras tridimensionales que son conocidas como secundarias [20] si son a nivel local o terciarias cuando se considera la forma global de la cadena plegada. La función de una proteína está determinada por su estructura debido a que la forma que adopta después de haberse plegado, genera lugares espaciales de su misma estructura los cuales pueden interaccionar con otras moléculas o inclusive con otras proteínas. A estas regiones se les conoce como sitios activos. Por ejemplo, las enzimas tienen una forma tridimensional que les permite unirse a las sustancias químicas que hacen reaccionar; las proteínas que se unen a los genes, como las polimerasas o factores de transcripción, tienen una forma tal que les permite unirse a los fosfatos dispuestos en hélice del DNA [20].

El plegamiento de las proteínas fue abordado por primera vez por Christian Anfinsen [21], el cual estudió la relación entre la secuencia de aminoácidos de la Ribonucleasa A y su estructura tridimensional biológicamente activa. La RNasa-A consta de un sólo polipéptido de 124 aminoácidos; entre ellos, hay cuatro puentes disulfuro en total. En condiciones óptimas de temperatura, salinidad y pH, con su estructura tridimensional adecuada, tiene la función enzimática de degradar moléculas de RNA.

Si la estructura terciaria depende únicamente de la secuencia de aminoácidos, entonces, alterar la estructura tridimensional sin cambiar la secuencia y posteriormente regresar a las mismas condiciones donde pueda recuperar su capacidad catalítica, sería indicio de que recuperó su estructura [6].

Anfinsen realizó dos experimentos con RNasa-A en donde agregó mercaptoetanol con el objetivo de eliminar sus puentes disulfuro, y posteriormente lo colocó en una solución concentrada de urea, de tal modo que la proteína se encontró en un estado no nativo. Para probar que se encontraba en un estado desnaturalizado, agregó RNA al medio y observó que las proteínas habían perdido prácticamente por completo su actividad, quedando sólo un remanente de un 2%. Después, por medio de diálisis, eliminó la urea, restableciendo las condiciones óptimas para su actividad, consiguiendo recuperar hasta un 80% de la actividad normal de la ribonucleasa, es decir, la proteína había recuperado su estructura nativa. Como no había nada más que el polipéptido, sales y agua, Anfinsen concluyó que, en las condiciones ambientales en las que ocurre el plegamiento normal de las proteínas, al menos para proteínas globulares pequeñas, la estructura biológicamente activa es única, estable y está determinada por la secuencia de aminoácidos de la proteína misma. A esto se le conoce como el Paradigma de Anfinsen y le valió el Premio Nobel a la Química en 1972 [19].

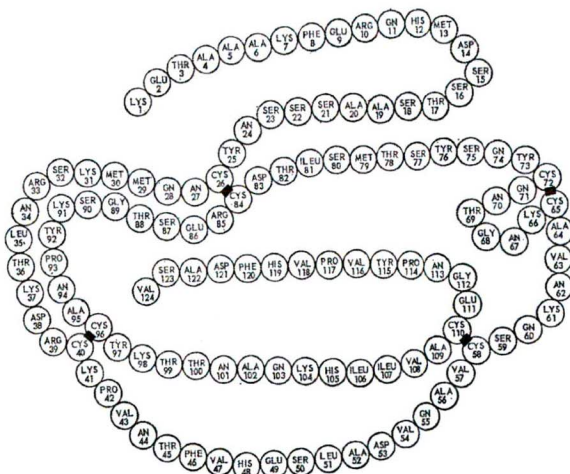


Figura 1.6: Estructura primaria de la secuencia de aminoácidos de la ribonucleasa pancreática bovina. Figura tomada del artículo de Christian B. Anfinsen, 1973 [21].

En los aminoácidos que forman a las proteínas, se encuentran habitualmente 20 tipos de cadenas laterales que varían en tamaño, forma, carga eléctrica, capacidad de formar enlaces de hidrógeno³, carácter hidrofóbico y reactividad química. Cabe señalar que todas las proteínas de todas las especies (bacterias, arqueas, células eucariotas⁴) se construyen con exactamente los mismos aminoácidos con unas pocas excepciones. La gran cantidad de funciones que pueden hacer las proteínas es el resultado de la diversidad de los 20 pilares de construcción (Tabla 1.1).

Las hélices α son estructuras secundarias de las proteínas muy estables debido a los enlaces de puentes de hidrógeno que se forman cuando se va generando dicha estructura. Cada secuencia de aminoácidos puede ser participe de dicha estructura la cual depende directamente de los aminoácidos que la compongan. En las hélices α se encuentran en promedio 3.6 aminoácidos por vuelta, tal y como se muestra en la figura 1.8; a su vez dicha estructura se logra gracias al enlace de hidrógeno que se forma entre

³Un enlace de hidrógeno es la fuerza atractiva entre un átomo electronegativo y un átomo de hidrógeno unido covalentemente a otro átomo electronegativo. La energía de este tipo de enlace está entre los valores de 5 a 30 kJ/mol.

⁴Se llama célula eucariota a todas las células con un núcleo celular delimitado dentro de una doble capa lipídica: la envoltura nuclear, la cual es porosa y contiene su material hereditario, fundamentalmente su información genética.

Aminoácido	Abreviación	Símbolo	Formula química
Alifáticos.			
Glicina	Gly	G	$C_2H_5NO_2$
Alanina	Ala	A	$C_3H_7NO_2$
Valina	Val	V	$C_5H_{11}NO_2$
Leucina	Leu	L	$C_6H_{13}NO_2$
Isoleucina	Ile	I	$C_6H_{13}NO_2$
Aromáticos.			
Fenilalanina	Phe	F	$C_9H_{11}NO_2$
Tirosina	Tyr	Y	$C_9H_{11}N_1O_3$
Triptófano	Trp	W	$C_{11}H_{12}N_2O_2$
Alcoholes.			
Serina	Ser	S	$C_3H_7NO_3$
Treonina	Thr	T	$C_4H_9NO_3$
Con Azufre.			
Cisteína	Cys	C	$C_3H_7NO_2S$
Metionína	Met	M	$C_5H_{11}NO_2S$
Ácidos.			
Ácido Aspártico	Asp	D	$C_4H_7NO_4$

Tabla 1.1: Tabla de aminoácidos estándares enlistados respecto a la cadena lateral R.

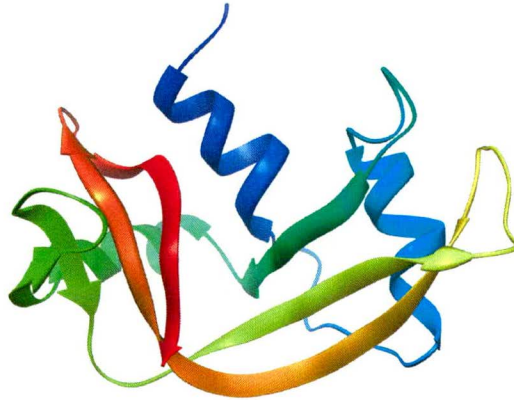


Figura 1.7: Estructura terciaria de la Ribonucleasa pancreática bovina. Esta estructura toma lugar en un espacio tridimensional. Una proteína con determinada estructura puede estar constituida por uno o varios dominios que pueden tener funciones específicas y separadas.

residuos al realizar un giro. Que un aminoácido pueda participar en una hélice α , depende de la estructura de cada uno de ellos, ya que los enlaces del carbono pueden o no girar dependiendo de la cadena lateral que tiene cada residuo.

Por otro lado, la propensidad helicoidal es una cantidad que indica la posibilidad que tiene un aminoácido en participar en una hélice α , representada con la letra ω [27]. Cada aminoácido tiene diferente tendencia a formar parte de una estructura helicoidal siendo esta última una estructura muy estable. En la tabla 1.2 se muestran los valores de las propensidades helicoidales de los 20 aminoácidos esenciales. Dichos valores son obtenidos experimentalmente y dependen de las condiciones específicas en las que fue realizado el experimento [27].

De acuerdo a la escala propuesta por Chakrabartty *et al* [25], el aminoácido alanina (Ala) tiene la propensidad helicoidal mas alta, por lo que es altamente probable que se encuentre en una hélice α . Contrario al aminoácido prolina (Pro), cuya propensidad helicoidal es 3 ordenes de magnitud menor que *Ala*, siendo este un aminoácido que difícilmente participa en una estructura hélice α . En la referencia [25] se denominan a los aminoácidos con baja propensidad helicoidal como disruptores de hélices (*helix-breakers*), refiriéndose a que si en una secuencia de péptidos se encuentra alguno de ellos, no habrá alguna hélice que se forme en la

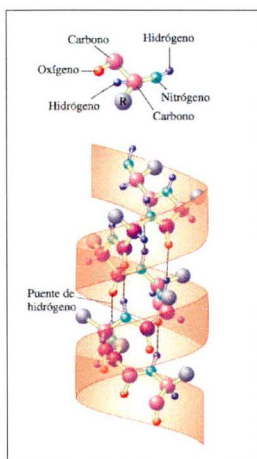


Figura 1.8: Estructura secundaria de la hélice α . En esta figura se muestran los aminoácidos que forman una hélice α , donde se forma un enlace de hidrógeno entre aminoácidos por cada vuelta que se realiza en la estructura, dándole soporte estructural. Imagen tomada de la referencia [2].

estructura terciaria, y esto tienen que ver con el hecho de que los enlaces del carbono α no tienen libertad de giro. La teoría de Lifson-Roig permite el cálculo de las propensidades helicoidales, las cuales se mencionarán más adelante, en donde existen restricciones para los valores de dichos ángulos, proponiendo que existe una tendencia de los aminoácidos a participar en una hélice α , poniendo la transición hélice-ovillo ⁵ en términos de un conjunto de probabilidades condicionales (u, v, ω) y una función de

⁵La transición hélice-ovillo se refiere al cambio estructural de una proteína que se enrolla en sí misma, dando origen al plegamiento de la misma.

Aminoácido	ω	Aminoácido	ω	Aminoácido	ω
Ala	1.61	Ile	0.44	Trp	0.30
Arg	1.2	Tyr	0.39	Phe	0.29
Leu	0.96	His	0.38	Val	0.23
Lys	0.82	Ser	0.38	Thr	0.14
Glu	0.66	Cys	0.35	His	0.06
Met	0.63	Asn	0.31	Gly	0.05
Gln	0.56	Asp	0.31	Pro	0.001

Tabla 1.2: Tabla de valores de las propensidades helicoidales de cada aminoácido, obtenidos de la referencia [27].

partición que dependen de dichas probabilidades $Z(u, v, \omega)$. Este conjunto puede ser escrito en terminos de integrales de una función potencial $V(\phi, \psi)$ de toda la cadena evaluada en regiones particulares de los mapas de Ramachandran, donde en líneas posteriores se mencionarán, y las variables ϕ y ψ son los ángulos diedros ⁶.

1.3.2. Enlace peptídico.

Los átomos tienden a formar enlaces químicos a través de sus orbitales. Por ejemplo, en el caso de la molécula de agua, el oxígeno, al ser más electronegativo que el hidrógeno, atrae más, hacia éste, los electrones compartidos en los enlaces covalentes con el hidrógeno, cargándose negativamente, mientras los átomos de hidrógeno se cargan positivamente, estableciéndose así dipolos eléctricos.

En algunas moléculas se presenta el problema de explicar cómo a partir de orbitales atómicos diferentes se pueden formar enlaces idénticos. Esta descripción la realizó el químico Linus Pauling, quien observó que un orbital atómico está dado por la solución de la ecuación de Schroedinger, que es un conjunto de funciones de onda Ψ con un cierto valor de la energía. Cuando dos o mas átomos se unen entre si, también se combinan linealmente las funciones de onda correspondientes, apareciendo nuevas funciones de onda las cuales se les denominan *orbitales híbridos*[31]. La hibridación combina los orbitales de átomos que se unen para formar un nuevo orbital llamado híbrido el cual, dependiendo del enlace, se generara una geometría distinta.

Los orbitales atómicos híbridos son los que se superponen en la formación de los enlaces, dentro de la teoría del enlace de valencia⁷, y justifican la geometría molecular. Existen varios tipos de hibridación, en particular, la hibridación sp^2 , la cual consiste en la combinación de un orbital s y $2p$, para formar 3 orbitales híbridos, que se disponen en un plano formando ángulos de 120° . Los átomos que forman hibridaciones sp^2 pueden formar compuestos con enlaces dobles. Forman un ángulo de 120° y su molécula es de forma plana. A los enlaces simples se les conoce como enlaces sigma σ y los enlaces dobles están compuestos por un enlace sigma y un enlace pi (π).

⁶Un ángulo diedro es cada una de las dos partes del espacio delimitadas por dos semiplanos que parten de una arista común. En el contexto de péptidos, son aquellos que se forman a partir de los planos de los grupos que se encuentran en los lados del carbono α .

⁷En química, la teoría del enlace de valencia explica la naturaleza de un enlace químico en una molécula, en términos de las valencias atómicas. La teoría del enlace de valencia resume la regla que el átomo central en una molécula tiende a formar pares de electrones, en concordancia con restricciones geométricas, según está definido por la regla del octeto. La teoría del enlace de valencia está cercanamente relacionada con la teoría del orbital molecular.

Las reglas de ubicación de los electrones en estos casos, como el alqueno etileno obligan a una hibridación llamada sp^2 , en la cual un electrón del orbital $2s$ se mezcla sólo con dos de los orbitales $2p$: surge a partir o al unirse el orbital s con dos orbitales p ; por consiguiente, se producen tres nuevos orbitales sp^2 , cada orbital nuevo produce enlaces covalentes[31].

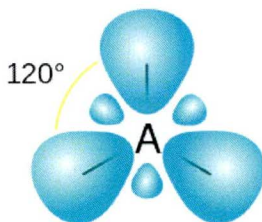


Figura 1.9: Configuración de orbitales sp^2 . Notese que a partir del átomo central están distribuidos los demás átomos con un ángulo de 120° , dándole una estructura triangular plana.

Dependiendo del número de aminoácidos que se tengan en la molécula se le denomina, según el prefijo, dipéptido, tripéptido, etc. Pero cuando el número se hace muy grande se le denomina *polipéptido*. En la siguiente figura se muestra a dos aminoácidos formando un enlace péptido y por ende un péptido.

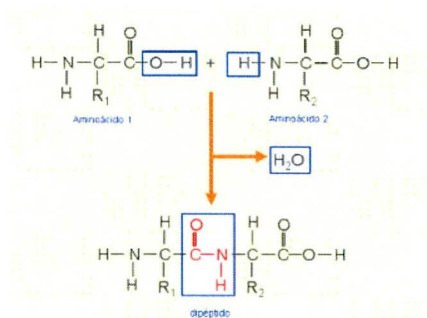


Figura 1.10: Dos aminoácidos. El grupo amino del primer aminoácido se une con el grupo carboxilo del segundo, liberando una molécula de agua y formando un enlace péptido entre los dos.

Una vez que se forme la cadena polipeptídica, los aminoácidos reciben el nombre de *residuos*. Los extremos de la cadena tienen cierta denominación, por ejemplo, el N-terminal, contiene un aminoácido con un grupo α -amino

libre que no está enlazado, mientras que el otro extremo $C_{terminal}$ tiene un grupo α -carboxilo libre. El enlace peptídico -CO-NH- se representa normalmente como un enlace sencillo [1].

Sin embargo, el enlace peptídico posee una serie de características que lo aproximan más a un doble enlace. Los átomos de C, N y O que participan en el enlace amida presentan una hibridación sp^2 , de forma que los 5 orbitales enlazantes de tipo s adoptan una disposición triangular plana. Los electrones p de los átomos de C, N y O no implicados en el enlace s pueden formar orbitales híbridos p . Como el N es menos electronegativo que el O, el enlace C-O tiene un 60% de carácter de doble enlace mientras que el enlace C-N tiene un 40%. Por lo tanto, se puede considerar que los enlaces C-O y N-C que participan en el enlace peptídico tienen estructuras intermedias entre el enlace sencillo y el enlace doble. Esta teoría se confirma por el hecho de que las distancias interatómicas medidas en el enlace C-O y en el enlace C-N son intermedias entre el enlace sencillo y el doble enlace. Esta disposición atómica está estabilizada por resonancia⁸, de forma que los seis átomos implicados en la formación del enlace peptídico están contenidos en el mismo plano. El carácter parcial de doble enlace impide la libre rotación de los átomos de C y N, al ser éste un enlace rígido. Por lo tanto, las posibilidades conformacionales de los péptidos son limitadas. Los péptidos sólo podrán cambiar de conformación por giro de los carbonos α tetraédricos (con hibridación sp^3), pero no por rotación de los átomos con hibridación sp^2 . El carbono α (C_α) de cada aminoácido implicado puede establecer dos ángulos de torsión con los planos de los enlaces peptídicos contiguos: los llamados ángulos ϕ y ψ . El ángulo ϕ es la rotación del enlace N- C_α , y el ángulo ψ la del enlace C- C_α . En la estructura de un péptido o de una proteína, cada aminoácido presenta un valor de ϕ y otro de ψ determinados. Como estos son los únicos grados de libertad que presenta la estructura, la conformación de la cadena polipeptídica queda completamente determinada cuando se conocen los valores de dichos ángulos para cada aminoácido[26]. Los giros alrededor del carbono α se concretan en los llamados *ángulos de torsión* [1].

Un giro en la dirección de las agujas del reloj sobre cualquiera de estos enlaces visto desde el átomo de nitrógeno hacia el carbono α o desde el carbonilo hacia el carbono α corresponde a un valor positivo. Los ángulos ϕ y ψ determinan la arquitectura de la cadena polipeptídica. La figura 1.11 se muestran los ángulos de torsión en un péptido.

⁸La resonancia en química se refiere a un sistema de enlaces entre átomos de una molécula que, debido a la compleja distribución de sus electrones, obtiene una mayor estabilidad que con un enlace simple[23].

Estructura	ϕ	ψ
Hélice α dextrógira	-57	-47
Hélice 3_{10}	49	-27
Hoja β paralela	-119	113
Hoja β antiparalela	-139	135
Poliprolina tipo II^b	-78	149

Tabla 1.3: En esta tabla se muestran los ángulos de rotación de los enlaces ϕ y ψ para estructuras secundarias regulares comunes. Dentro de estas estructuras, la hélice α y la hoja β son las conformaciones de las cadenas polipeptídicas generalmente mas estables termodinámicamente.

Si consideramos las diferentes combinaciones que se podrían hacer en la arquitectura de las cadenas polipeptídicas variando los ángulos de rotación, tendríamos una infinidad de configuraciones para la misma. Estas posibles configuraciones de ángulos que puede tener la cadena polipeptídica se pueden mapear en una gráfica llamada Mapa de Ramachandran, el cual son las representaciones del espacio conformacional accesible para una cadena polipeptídica y cada coordenada representa cada par de los ángulos diedros. La rigidez de la unidad peptídica y los límites de los ángulos de enlace ϕ y ψ limitan el suficiente número de estructuras posibles para que la cadena desplegada se pliegue [26].

En algunos segmentos de un péptido, existen configuraciones de las estructuras secundarias cuando todos los ángulos ϕ son iguales entre si en dicho segmento y todos los ángulos ψ también son iguales. Sin embargo, también existen regiones significativas de estructura al azar en la que los ángulos ϕ y ψ no son iguales.

Un péptido esta formado, como se mencionó anteriormente, por una secuencia de aminoácidos, los cuales solamente difieren por la cadena lateral R, a su vez, los carbonos α y los enlaces peptídicos crean lo se le conoce como cadena principal (*backbone*). La cadena lateral R tiene una gran capacidad de formar puentes de hidrógeno. Cada residuo tiene un grupo carbonilo C=O que acepta muy bien los puentes de hidrógeno y, con excepción de la prolina, un grupo NH que cede, de igual forma los puentes de hidrógeno. La mayoría de las cadenas polipeptídicas naturales tienen entre 50 y 2000 residuos de aminoácidos y se conocen como proteínas.

Un enlace peptídico plano tiene dos configuraciones posibles: *trans* y *cis*. En la configuración *trans* pasa que los dos átomos de carbono α están en lados opuestos del enlace peptídico. Casi todos los enlaces peptídicos

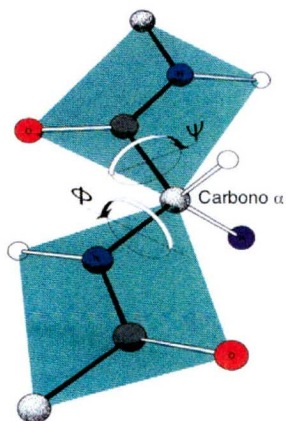
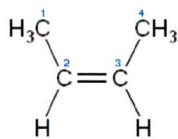
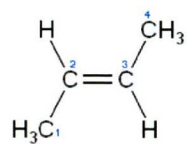


Figura 1.11: Representación gráfica de los ángulos de torsión ψ y ϕ en un aminoácido los cuales están a los costados del carbono α .



Configuración *Cis*



Configuración *Trans*

Figura 1.12: Configuraciones *Cis* y *Trans* para la molécula de buteno.

son *trans*. En la configuración *cis*, los dos grupos de las cadenas laterales se encuentran en el mismo lado del enlace doble entre el carbono α y un nitrógeno, lo que da lugar a fuerzas estéricas de repulsión desfavorables entre las cadenas laterales. Esta preferencia del *trans sobre el cis* se puede explicar por el hecho de que los choques estéricos entre los grupos unidos a los átomos de carbono α impiden la adopción de la forma *cis* mientras que no se producen en la forma *trans*.^[1]

Los enlaces peptídicos *cis* mas comunes son las uniones X-Pro, donde X es la cantidad de residuos en la cadena. Estos enlaces muestran menos preferencia por la configuración *trans* porque el nitrógeno de la *prolina* esta unido a dos átomos de carbono tetraédricos, lo que limita las diferencias estéricas entre las formas *trans y cis*.

En el siguiente capítulo se mencionarán las herramientas que se utilizaran para realizar el estudio de cadenas que no tienen los mismos aminoácidos en su secuencia, llamadas cadenas heteropoliméricas. En particular centraremos la atención en los modelos de grano grueso, y el modelo de Zwanzig para cadenas homopoliméricas.

Capítulo 2

METODOLOGÍA

Existen varios métodos para estudiar las propiedades físicas de cadenas homo y heteropoliméricas proteinogéneas. Una de estas técnicas es la dinámica molecular (DM), la cual permite a partículas atómicas como átomos o moléculas interactuar entre sí y poder observar el movimiento que realizan. El uso de computadoras es vital en esta técnica debido a que la DM se basa en la solución numérica de las ecuaciones de Newton con el fin de obtener las ecuaciones de movimiento del sistema. La dinámica molecular resulta ser una herramienta muy útil cuando se necesita modelar un conjunto de muchas moléculas interactuando entre sí.

Por otro lado, existe otro enfoque para describir a sistemas de muchas moléculas el cual se conoce como modelos de grano grueso (*Coarsed-Grained Models*). En estos modelos se reduce considerablemente la cantidad de grados de libertad de todas las partículas en el sistema. Debido a la reducción en los grados de libertad y la eliminación de detalles de interacciones finas, la simulación de un sistema de grano grueso (CG) requiere menos recursos y se realiza en tiempos menores que para el mismo sistema en la representación de todos los átomos. En este trabajo nos centraremos en modelos de grano grueso basados en física estadística.

2.1. Modelos de grano grueso (*Coarsed-Grained Models*)

Para describir el comportamiento de un sistema complejo es necesario implementar un modelo, el cual involucre la menor cantidad posible de variables, cuya información sea relevante para el sistema que se encuentre en estudio.

Como ejemplo consideremos una proteína sometida a varios tipos de

interacción, como son las atracciones o repulsiones estéricas, fuerzas de Van Der Waals e interacciones electrostáticas. El incluir todas las interacciones en un modelo que pretenda describir la dinámica de esta molécula sería un problema complicado. Esto nos lleva a crear un modelo que no considere los detalles dinámicos, atómicos y estructurales, el cual denominaremos *modelo de grano-grueso (Coarsed-Grained Model, CGM)*.

Muchos trabajos de dinámica molecular se basan en este tipo de consideraciones ya que facilita, hasta cierto punto, los cálculos pertinentes a dicho problema. Los equipos de cómputo son la herramienta necesaria para dichos cálculos; existen varios programas que sirven para estos fines como son: GROMACS [32], MARTINI [41], entre otros, los cuales simulan a las moléculas que intervienen en el problema y sus interacciones. Existen trabajos en los cuales se puede simular moléculas de colesterol en una bicapa lipida, el cual ayuda a disminuir su permeabilidad y observar el comportamiento de la condensación sobre dichas capas [41]. Este ejemplo de simulación puede verse gráficamente en la figura 2.1 donde se presenta a las moléculas de colesterol en el ambiente mencionado.

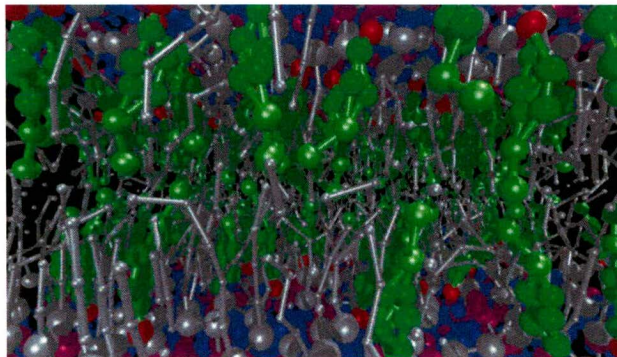


Figura 2.1: Estructura de una bicapa DPPC / colesterol en 0,3 moles fracción de colesterol después de 1 microsegundo simulación con un modelo de CG. Las moléculas del colesterol se muestran en verde. El sitio de colesterol en representación del grupo hidroxilo se muestra en rojo. Las colas de lípidos DPPC se muestran en plata, con las esferas más grandes que denotan el resto de glicerol. Imagen obtenida de la referencia [41].

Los modelos de grano grueso son de gran utilidad debido a que simplifican los cálculos donde la solución únicamente pueda realizarse con métodos numéricos. A continuación se presentará un modelo de grano grueso con el fin de dar una explicación de las bases del trabajo realizado en este trabajo de tesis. A su vez, se aplicaran los criterios pertinentes a

los cuales estará sujeto dicho modelo con el objetivo de obtener las propiedades termodinámicas a partir de una expresión para la función de partición.

2.2. Termodinámica y Física estadística

La termodinámica es una rama de la física que estudia las propiedades macroscópicas de sistemas físicos tales como calores específicos, energía interna, etc. Esta rama de la física no realiza ninguna hipótesis acerca de la estructura microscópica y se basa en principios generalizados a partir de la experiencia [36].

Por otro lado, la física estadística pretende, a partir de un comportamiento estadístico de los átomos y moléculas de un sistema, encontrar las propiedades termodinámicas macroscópicas y determinar la distribución de grandes conjuntos de moléculas que constituyen una porción macroscópica de la materia[9].

En este capítulo nos enfocaremos en las herramientas que se utilizarán en este trabajo para realizar la descripción del comportamiento termodinámico de cadenas heteropoliméricas, enfocandonos en la cantidad de conformaciones que puede adoptar una cadena heteropolimérica con respecto a su estado nativo y a sus estados no nativos.

2.2.1. Conceptos básicos de termodinámica

Desde el punto de vista atómico, una muestra macroscópica de materia es un conglomerado de un número enorme de electrones y núcleos atómicos. Una descripción matemática completa de la muestra consiste en especificar las coordenadas adecuadas para cada electrón y cada núcleo. De este modo, se requieren del orden de 10^{24} o 10^{25} coordenadas para describir el estado, por ejemplo, de un litro de alcohol. Para describir macroscópicamente un sistema es conveniente tener un número reducido de parámetros, en contraste con el enorme número requerido para una descripción atómica.

En la transición desde el nivel de descripción atómico al macroscópico se produce una simplificación enorme, así como una drástica reducción en el número de variables necesarias. Esta simplificación es debida a la lentitud de las medidas macroscópicas en comparación con la rapidez de los movimientos atómicos. Las medidas macroscópicas detectan solamente valores promedio de las coordenadas atómicas [9].

La termodinámica es el estudio de las propiedades macroscópicas de un sistema, las cuales pueden ser la gran cantidad de coordenadas atómicas que, en virtud del promedio estadístico, no aparecen explícitamente en la descripción macroscópica de un sistema dado. Es de gran generalidad, aplicable a sistemas de estructura complicada con todo tipo de propiedades mecánicas complejas, eléctricas y térmicas. A su vez, la termodinámica también se aplica a una amplia variedad de temas de la ciencia y la ingeniería, como motores, transiciones de fase, reacciones químicas, fenómenos de transporte, y los agujeros negros aún ya que todos estos ejemplos pueden ser función de la temperatura. Los resultados de los cálculos termodinámicos son esenciales para otros campos de la física y la química, ingeniería química, ingeniería aeroespacial, ingeniería mecánica, biología celular, ingeniería biomédica, y la ciencia y los materiales útiles en otros campos como la economía.

El punto de partida conveniente son las ecuaciones fundamentales de un sistema PVT monofásico, donde P representa a la presión, V el volumen y T la temperatura del sistema. No se consideran campos eléctricos, magnéticos o gravitacionales. Para un sistema de componentes múltiples, donde la cantidad total de sustancia está dado por $n = \sum_i n_i$, la ecuación característica fundamental de la energía es representada por la ecuación siguiente:

$$d(nU) = Td(nS) - Pd(nV) + \sum_i \mu_i dn_i, \quad (2.1)$$

donde U es la energía interna molar y S es la entropía molar del fluido. El parámetro intensivo, μ_i , proporcionado por las derivadas parciales de primer orden de la energía interna con respecto a la cantidad de sustancia de componente i, es el potencial químico. Su introducción se extiende para sistemas de una sola fase en la que el n_i puede variar, ya sea mediante el intercambio de materia con su entorno (sistema abierto) o por cambios en la composición que se producen como resultado de reacciones químicas (sistema reactivo) o ambos[37].

La representación de la energía de las cantidades extensivas son las variables matemáticamente independientes, mientras que los parámetros intensivos se derivan. La elección de S y V como variables independientes en la ecuación fundamental en la representación de la energía no es conveniente, y la ecuación para la energía interna U sugiere la definición de funciones alternativas útiles a base de energía primaria. El método adecuado para la generación de ellas sin pérdida de información son las transformaciones de Legendre.

Aplicando la transformada de Legendre para la relación fundamental $U = U(S, V, N_1, \dots, N_r)$ y teniendo en cuenta que las derivadas $\left(\frac{\partial U}{\partial S}\right)_{V, N_i}$,

$(\frac{\partial U}{\partial V})_{S,N}$, $(\frac{\partial U}{\partial N})_{S,V}$ corresponde a los parámetros intensivos T , $-P$, μ_i respectivamente. Estas funciones transformadas de Legendre obtenidas se denominan *potenciales termodinámicos*, de los cuales nos enfocaremos en uno en particular que se llama *potencial de Helmholtz* F . Este potencial F es la transformada de Legendre parcial de U que reemplaza como variable independiente la entropía por la temperatura. Las variables naturales de F son T, V, N , esto es que $F = F(T, V, N)$ y constituye una relación fundamental.

$$U = U(S, V, N_i) \quad (2.2)$$

$$T = \frac{\partial U}{\partial S} \quad (2.3)$$

$$F = U - TS \quad (2.4)$$

Eliminando U y S se obtiene:

$$F = F(T, V, N_i) \quad (2.5)$$

La diferencial de completa dF es:

$$dF = -SdT - PdV + \mu_1dN_1 + \mu_2dN_2 + \dots \quad (2.6)$$

La condición de equilibrio del potencial de Helmholtz dice que: *El valor de equilibrio de cualquier parámetro interno sin ligadura de un sistema que se halla en contacto diatérmico con una fuente de calor, minimiza el potencial de Helmholtz a temperatura constante (igual a la de la fuente de calor)*. Es claro que U puede ser obtenido a partir de F de la siguiente forma:

$$U = F - TS \quad (2.7)$$

$$U = F - T \left(\frac{\partial F}{\partial T} \right)_V \quad (2.8)$$

Las variables canónicas del potencial de Helmholtz F son V y T . Nos enfocamos en el potencial F debido a que el sistema que deseamos estudiar en el capítulo tres, depende de las variables T y V y cuyo manejo se puede utilizar para manipular las propiedades termodinámicas del mismo.

2.2.2. Física Estadística.

La física estadística proporciona un marco para relacionar las propiedades microscópicas de los átomos y moléculas individuales con las propiedades macroscópicas de los materiales que se pueden observar en la vida cotidiana como resultado de las descripciones de la mecánica clásica y

cuántica, de la estadística y la mecánica del nivel microscópico. La física estadística proporciona una interpretación a nivel molecular de las cantidades termodinámicas macroscópicas tales como el trabajo, el calor, la energía libre y entropía. Esta capacidad de hacer predicciones macroscópicas sobre la base de propiedades microscópicas es de gran ayuda para la termodinámica clásica. Ambas teorías se rigen por la segunda ley de la termodinámica a través del medio de la entropía. Sin embargo, la entropía en termodinámica sólo puede ser conocida empíricamente, mientras que en la física estadística, es una función de la distribución del sistema de sus microestados.

Los microestados de un sistema se definen como una configuración en la que es posible encontrar al sistema y a su vez, le corresponde un cierto valor de energía. Cada microestado es caracterizado por alguna configuración del sistema en particular restringido a ciertas condiciones las cuales dependen de la energía. Por otro lado, un macroestado de un sistema es aquel que puede describir sus propiedades macroscópicas, independientemente de las posibles configuraciones que pueda adoptar. Un macroestado determinado puede tener muchos microestados diferentes. Por ejemplo, si se tienen dos átomos con los posibles valores de espín s , donde puede tener los valores $s = -1$ y $s = 1$, y definimos a la energía del sistema como el producto de dichos espines, o sea $E = s_j s_i$, los microestados del sistema son las posibles configuraciones que puede adoptar el sistema de dos átomos, las cuales son: $(s=1)(s=1)$, $(s=1)(s=-1)$, $(s=-1)(s=1)$ y $(s=-1)(s=-1)$; y los macroestados son aquellos que tienen energía E igual a 1 y -1 .

El postulado fundamental en la física estadística es el siguiente:

Dado un sistema aislado en equilibrio, todos los microestados compatibles con un estado macroscópico son igualmente probables.

Este postulado es una premisa fundamental en la física estadística que establece que un sistema en equilibrio no tiene ninguna preferencia por alguno de sus microestados disponibles.

Dados Ω microestados en una energía particular, la probabilidad de encontrar el sistema en un microestado particular es $p = 1/\Omega$. Este postulado es necesario, ya que permite a la conclusión de que, para un sistema en equilibrio, el estado termodinámico (macroestado) que podría resultar del mayor número de microestados es también el más probable macroestado del sistema. El postulado se justifica, en parte, por los sistemas clásicos, por el teorema de Liouville (Hamilton), lo que demuestra que si la distribución de los puntos del sistema a través del espacio de fase accesible es uniforme en algún momento, lo sigue siendo en momentos posteriores.

La formulación moderna de la física estadística se basa en la descripción del sistema físico por un conjunto que representa todas las posibles configuraciones del sistema y la probabilidad de la realización de cada configuración. Cada conjunto está asociado con una función de partición que, con la manipulación matemática, se puede utilizar para extraer valores de las propiedades termodinámicas del sistema.

El ensamble canónico describe un sistema en equilibrio térmico con su entorno. Sólo se puede intercambiar energía en forma de calor con el exterior. En el conjunto canónico, N , V y T son fijos. Invocando el concepto del conjunto canónico, es posible derivar la probabilidad P_i de que un sistema macroscópico en equilibrio térmico con su entorno, será en un microestado dado con energía E_i de acuerdo con la distribución de Boltzmann:

$$P_i = \frac{e^{-\beta E_i}}{\sum_{j=1}^{j_{max}} e^{-\beta E_j}} \quad (2.9)$$

La temperatura T se deriva del hecho de que el sistema está en equilibrio térmico con su entorno. Las probabilidades de los diferentes microestados deben sumar uno, y el factor de normalización en el denominador es la función de partición canónica, definida de la siguiente forma:

$$Z = \sum_{j=1}^{j_{max}} e^{-\beta E_j} \quad (2.10)$$

donde β esta definida como $\beta = \frac{1}{k_B T}$ y k_B es la contante de Boltzmann.

En este trabajo sólo tomaremos el ensamble canónico en consideración debido a que vamos a considerar intercambio de energía con el exterior por alguna fuente de calor, lo que no consideraremos es que exista un intercambio de masa, es decir, cuando tengamos una cadena polimérica con K residuos, estos permanecerán constantes ($K = cte$).

2.3. Función de partición Z

La función de partición describe las propiedades estadísticas de un sistema en equilibrio termodinámico con un baño térmico a temperatura T . Es funcion de la temperatura y otros parámetros que determinan los niveles de energía. La mayoría de las variables termodinámicas del sistema, tales como la energía total, energía libre, la entropía, y la presión, se puede expresar en términos de la función de partición. En realidad, hay varios tipos diferentes de funciones de partición, cada uno correspondiente a diferentes tipos de conjunto estadístico (o, de manera equivalente, diferentes

tipos de energía libre) [9].

La función de partición canónica se aplica a un conjunto canónico, en el que se permite que el sistema intercambie calor con el medio ambiente a temperatura, volumen, y el número de partículas fijos.

Ahora, suponemos que las variables de estado T, V, N permanecen constantes. Entonces la condición de equilibrio del sistema esta dado por la ecuación:

$$dF = dU - TdS = 0 \quad (2.11)$$

Si consideramos a la entropía en función de las probabilidades p_j :

$\frac{S}{k} = -\sum_{j=1}^{j_{max}} p_j \ln(p_j)$ y diferenciando de ambos lados nos queda:

$$dS = -k \sum_{j=1}^{j_{max}} [1 + \ln(p_j)] dp_j \quad (2.12)$$

Sabemos que dU puede ser obtenida teniendo en cuenta que es el promedio de los niveles de energía microscópicos,

$$U = \langle E \rangle = \sum_{j=1}^{j_{max}} p_j E_j \quad (2.13)$$

tomando la derivada de U tenemos

$$dU = \sum_{j=1}^{j_{max}} (p_j dE_j + E_j dp_j) \quad (2.14)$$

Sabemos que la energía E_j es función únicamente del volumen y del número de partículas del sistema, es decir, $E_j = E_j(V, N)$, por lo que $dE_j = (\frac{\partial E_j}{\partial V})dV + (\frac{\partial E_j}{\partial N})dN = 0$ ya que consideramos que V y N permanecerán constantes, por lo que la ecuación de la energía interna queda de la forma:

$$U = \langle E \rangle = \sum_{j=1}^{j_{max}} E_j dp_j \quad (2.15)$$

Ahora bien, consideramos que la suma de las probabilidades p_j es igual a 1 ($\sum_{j=1}^{j_{max}} p_j = 1$) como una restricción a la condición de equilibrio $dF = dU - TdS = 0$ y sustituyendo los términos anteriores en la misma, escribiendo a la restricción como un multiplicador de Lagrange.

$$dF = d\langle E \rangle - TdS \quad (2.16)$$

$$dF = \sum_{j=1}^{j_{max}} E_j dp_j - T \left(-k \sum_{j=1}^{j_{max}} (1 + \ln(p_j)) dp_j \right) \quad (2.17)$$

$$dF = \sum_{j=1}^{j_{max}} [E_j + kT(1 + \ln(p_j)) + \omega] dp_j = 0 \quad (2.18)$$

Para que esta ecuación se cumpla, el termino $E_j + kT[1 + \ln(p_j)] + \omega$ tiene que ser igual a cero para cada valor de j .

$$E_j + kT(1 + \ln(p_j)) + \omega = 0 \quad (2.19)$$

Despejando el termino $\ln(p_j)$:

$$\ln(p_j) = -\frac{E_j}{kT} - \frac{\omega}{kT} - 1 \quad (2.20)$$

Aplicando de los dos lados de la ecuación la función exponencial:

$$p_j = e^{\left(\frac{E_j}{kT}\right)} e^{\left(\frac{-\omega}{kT}\right)-1} \quad (2.21)$$

$$Z = \sum_{j=1}^{j_{max}} e^{\left(\frac{-E_j}{kT}\right)} \quad (2.22)$$

La función de partición nos dice como van a estar las partículas distribuidas en los estados accesibles de energía del sistema. Esta función es de gran ayuda para el calculo de variables termodinámicas, por ejemplo, para un sistema de N partículas, la energía interna puede ser expresada en términos de Z :

$$U = - \left(\frac{\partial \ln(Z)}{\partial \beta} \right) \quad (2.23)$$

De aquí mismo podemos calcular el calor específico de la forma siguiente:

$$C_v = \frac{\partial}{\partial T} \left(\frac{\partial \ln(Z)}{\partial \beta} \right) \quad (2.24)$$

$$C_v = \frac{1}{kT^2} \left(\frac{\partial^2 \ln(Z)}{\partial \beta^2} \right), \quad (2.25)$$

y como $C_v = \frac{\partial E}{\partial T}$, entonces podemos decir que el calor específico C_v nos dice como va a cambiar la energía promedio del sistema con respecto a la temperatura que tenga en el ensamble canónico. Esta expresión es muy importante debido a que es posible observar cambios abruptos en el sistema conforme la temperatura cambia, lo que implicaría una reconfiguración de

la estructura interna, lo que se conoce como una transición de fase, tema que se discutirá brevemente en la sección siguiente.

En la teoría de perturbaciones, el calor específico juega un papel importante ya que es posible acceder a variables microscópicas a partir de una variable que se puede medir en el laboratorio, tal y como lo es el calor específico C_v y establecer relaciones fundamentales, entre las fluctuaciones (generalmente, difíciles de medir) de magnitudes extensivas, como E alrededor de las medias que caracterizan el estado de equilibrio y otras magnitudes que, como es el C_v . En la referencia [36] se calcula la desviación cuadrática media de la energía, la cual esta regida por el calor específico a volumen constante, considerando las fluctuaciones de energía de un sistema simple en contacto diatermico con una fuente de calor pero confinado, por paredes impermeables rígidas.

$$\frac{\partial U}{\partial \beta} = - \langle E^2 \rangle + \langle E \rangle^2 \quad (2.26)$$

pero $-\langle (\Delta E)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2$, por lo que podemos acceder a la siguiente ecuación tomando en cuenta la definición del calor específico:

$$\langle E^2 \rangle - \langle E \rangle^2 = k_B T^2 C_v \quad (2.27)$$

La fluctuación de cada parámetro depende de las fuentes con las cuales esta en contacto el sistema. Esta ecuación es otra forma de relacionar al calor específico con las fluctuaciones que puede tener algún sistema en estudio.

Tener una expresión analítica para la función de partición facilita la obtención de las variables termodinámicas para un sistema del cual podemos conocer su energía.

2.3.1. Transiciones de fase configuracionales

Una transición de fase es el cambio de un estado homogéneo a otro estado también homogéneo, como resultado de un cambio en las condiciones externas, tales como temperatura, presión, o un campo magnético o eléctrico externo [35]. El ejemplo más conocido de la vida cotidiana es el agua, la cual tiene tres estados de agregación de la materia: sólido, líquido y gaseoso. Para cambiar de una fase a otra, decimos que ocurre una transición de fase (TF).

Las transiciones de fase generalmente se encuentran asociadas con arreglos de diferente simetría, la transición que va de una fase más simétrica

a otra menos simétrica, es un proceso de rompimiento de simetría. La transición ferromagnética es un ejemplo, la simetría es representada como espines hacia arriba o hacia abajo, y se rompe con la formación de dominios magnéticos. Este rompimiento de simetría es muy importante en el comportamiento de las transiciones de fase.

Cuando se rompe la simetría, es necesario introducir una o más variables extras para poder definir el estado del sistema. Por ejemplo, en la fase ferromagnética es necesario dar la magnetización total. Estas variables se pueden considerar como parámetros de orden.

En el caso de las proteínas un estado nativo se puede considerar como una fase conformacional en la cual se encuentra dicho péptido con respecto a todas las posibles conformaciones en las que se encuentren cada uno de los residuos que la forman. Por otro lado, el estado no nativo o desplegado de la proteína también es otra fase en la cual alguno o todos los residuos se encuentran en un estado desnaturalizado, donde la actividad biológica no se realiza¹. El cambio del estado nativo al estado no nativo se considera una transición de fase configuracional, por lo que se llega a la necesidad de estudiar las propiedades termodinámicas de las cadenas polipeptídicas y conocer en que valores de la temperatura ocurren estos cambios configuracionales.

La descripción y el análisis de las transiciones de fase requiere el uso de la termodinámica y la física estadística. En termodinámica cada estado de un sistema se define por alguna energía característica. Si el estado del sistema se define por su temperatura T y su presión P o el volumen V , esta energía se llama la energía libre. Una parte de esta energía es la energía E a $T = 0$ del sistema, mientras que la otra parte depende de la temperatura y la entropía S del sistema. Si las variables independientes son la temperatura y la presión, entonces el potencial termodinámico relevante es la energía libre de Gibbs ($G = E + PV - TS$), mientras que si son la temperatura y el volumen es la energía libre de Helmholtz ($F = E - TS$). Las diferenciales de estas energías libres para un sistema simple son:

$$dG = -SdT + VdP, \quad dF = -SdT - PdV. \quad (2.28)$$

Las primeras derivadas de la energía libre 2.28 son precisamente variables del sistema V y S en el caso del potencial de Gibbs 2.29, S y P para el caso del potencial de Helmholtz 2.30.

¹En realidad existe todo un conjunto de posibles conformaciones no nativas posibles.

$$\left(\frac{\partial G}{\partial P}\right) = V, \quad \left(\frac{\partial G}{\partial T}\right) = -S, \quad (2.29)$$

y

$$\left(\frac{\partial F}{\partial T}\right) = -S, \quad \left(\frac{\partial F}{\partial V}\right) = -P. \quad (2.30)$$

Ahora, obteniendo las segundas derivadas se puede calcular el calor específico C_p y C_v de la siguiente forma:

$$C_p = \left(\frac{\partial^2 G}{\partial T^2}\right)_P, \quad C_v = -T \left(\frac{\partial^2 F}{\partial T^2}\right)_V. \quad (2.31)$$

La relación que existe entre los potenciales termodinámicos esta dada por la ecuación $G = F + PV$, por lo que dado uno de los dos potenciales, se puede obtener el otro. Las transiciones de fase de primer orden, por definición, son cambios estructurales que llevan a un sistema en una fase inicial a otra fase distinta y están caracterizadas por una discontinuidad finita en la primera derivada del potencial de Gibbs $G(T, P)$.

En la física estadística se catalogan diferentes transiciones de fase de distinto orden. La clasificación se basa en la cercanías o vecindades de puntos críticos sobre las curvas (en algún diagrama de fase) correspondientes a una ecuación de estado termodinámico propio del sistema. Como ejemplo se tiene la ecuación de Van der Waals [22], que históricamente fue la primera ecuación de estado que exhibía una transición de fase en las isotermas del diagrama PV; siendo mas específicos, cuando la derivada parcial $\frac{\partial P}{\partial V}_T > 0$ el coeficiente de compresibilidad $\kappa_T = -\frac{1}{v} \frac{\partial v}{\partial P}_T$ (siendo v el volumen y P la presión) es negativo, lo que está en contradicción con las condiciones de estabilidad de los sistemas termodinámicos, los cuales exigen que este coeficiente sea siempre positivo. Por tanto se debe tratar de una zona no física, donde no se puede encontrar al sistema homogéneo y en equilibrio. Cuando la transición de solido a líquido o de líquido a gaseoso toma lugar, hay un cambio discontinuo en la entropía.

Un ejemplo mas ilustrativo consiste en graficar la entropía molar $S_m = -\left(\frac{\partial G_m}{\partial T}\right)_P$, donde G_m es el potencial molar de Gibbs, como función de la temperatura T a P y N constantes. La figura 2.2 muestra tres gráficas de X (Entropía o volumen), del potencial químico y del calor específico (C_p). El potencial químico cambia continuamente pero su derivada es discontinua en $T = T_{trans}$. A la temperatura de transición, debido a la existencia de calor latente, el calor específico ($= \frac{\Delta Q}{\Delta T}$) tienen una singularidad ya que se

vuelve infinito, es decir, el calor absorbido ΔQ no causa ningún cambio en la temperatura (ΔT). Este tipo de cambios son clasificados como *transiciones de fase de primer orden* [40].

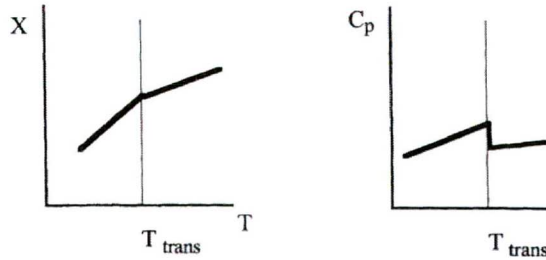


Figura 2.2: En esta figura se muestra el cambio de las variables termodinámicas en una transición de fase TF de primer orden que ocurre a una temperatura $T = T_{trans}$ llamada *temperatura de transición*. X es una cantidad extensiva tal como S_m o V_m las cuales se someten a cambios discontinuos. (Imagen tomada de Kondepudi, D. Prigogine, I. Modern thermodynamics (Wiley, 1998))

Un ejemplo en donde ocurre una transición de fase es el modelo de Ising en 2 dimensiones [42], el cual tiene como objetivo estudiar el comportamiento de materiales ferromagnéticos a partir de definir un arreglo lineal de espines ² y definiendo la energía de interacción a primeros vecinos, la cual es clave fundamental para el calculo de las propiedades termodinámicas y por ende, calcular la temperatura en la cual dicha retícula tendrá una transición de fase[39]. El plantear y resolver analíticamente este modelo en 2 dimensiones, le valió el premio nobel a Lars Onsager (1968). Este modelo es una herramienta muy útil debido a que se puede aplicar a diferentes sistemas en los cuales se pueda lograr hacer una analogía en donde se tenga una secuencia de elementos y estos tengan interacción a primeros vecinos tal y como sucede en las proteínas las cuales tienen ese tipo de interacción debida a los enlaces que crean entre si.

En el capítulo 3, se mostrarán las transiciones de fase que ocurren en una cadena heteropolimérica, donde las fluctuaciones de la energía promedio nos indicarán en que valores de la temperatura ocurre la transición entre el estado nativo y el estado no nativo de dichas secuencias.

²El modelo de Ising también esta en 1 y 3 dimensiones, definiendo retículas de espines en dichas dimensiones.

2.3.2. Modelo de Zwanzig para homopolímeros ideales

El modelo de Zwanzig (MZ) [33] es un modelo que utiliza herramientas de la física estadística, cuyo objetivo es determinar las propiedades termodinámicas de una cadena polipeptídica. Mediante la aproximación de grano grueso, el modelo de Zwanzig no considera interacciones de largo alcance o efectos estéricos los cuales pueda presentar el sistema que se este estudiando.

A continuación, describiremos el modelo y sus principales suposiciones. En primer lugar, se considera una cadena de K aminoácidos idénticos, los cuales tienen cada uno $\nu + 1$ conformaciones, donde ν es la cantidad de estados *no nativos* que puede tomar un aminoácido y solo uno es un estado *nativo*, por lo que la cadena completa presenta un total de $(\nu + 1)^K$ conformaciones.

Dada una configuración dada, la cantidad de aminoácidos que se encuentran en el estado nativo la denotaremos por α , por lo que $(K - \alpha)$ es la cantidad de aminoácidos que están en un estado no nativo. Una de las suposiciones básicas del modelo es que la energía E depende directamente de la cantidad de aminoácidos en un estado nativo α que se encuentren en la cadena homopolimérica debido a que la energía en un estado nativo o plegado es menor que cualquier estado no nativo cualquiera. El calculo de la energía E_α de la cadena polipeptídica es una tarea muy complicada, ya que se tendría que dar una expresión analítica para el potencial de interacción y resolver la ecuación de Schrodinger, por lo que Zwanzig[33] propone una expresión para el espectro de energías E_α de la siguiente forma:

$$E_\alpha = (K - \alpha)U - \epsilon\delta_{K,\alpha} \quad (2.32)$$

donde

- U es la energía de cada aminoácido en el estado no nativo la cual se supone igual para todos los residuos.
- ϵ es la energía de toda la cadena polipeptídica en el estado nativo.

Como se muestra en la figura 4.1a, el espectro de energía tiene niveles igualmente espaciados. a diferencia del espectro de energías de un átomo de hidrógeno, que aumentan como el cuadrado del nivel energético n de la siguiente forma $E_n = \frac{n^2 h^2}{8mL^2}$. El modelo de Zwanzig, en el caso de cadenas homopoliméricas, es un modelo de grano grueso meramente configuracional, lo cual indica que solamente hace uso de los estados *nativos* y *no-nativos* de cada aminoácido en la cadena homopolimérica residual en cuestión, por lo que se excluyen las posiciones espaciales de los mismos residuos en la cadena que podrían estar dados, por ejemplo, por

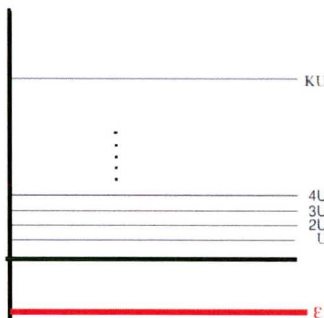


Figura 2.3: Espectro de energía del modelo de Zwanzig. La línea con el valor ϵ corresponde a la energía del estado base del modelo de Zwanzig para el valor de $\alpha = K$. Los demás niveles corresponden a estados meta estables en los cuales la cadena homopolimérica puede encontrarse. El estado con el valor de $\epsilon = KU$ es el estado donde la cadena tiene a todos los residuos en un estado *no nativo*, con $\alpha = 0$.

un vector de posición \vec{r}_j el cual determinaría las posiciones de cada molécula.

Para ejemplificar el contenido físico del modelo de Zwanzig, consideramos el caso de una cadena de 3 aminoácidos $K = 3$ con un aminoácido en una conformación nativa ($\alpha = 1$) y dos en estado no nativo ($K - \alpha = 2$), tendríamos que la energía para esa conformación es $E_{\alpha=1} = 2U$. Para estos mismos valores de α , existen tres configuraciones, las cuales comparten la misma energía, por lo cual, el sistema en general presenta degeneración. Para mostrar dicha degeneración en la figura 2.4 se expone un ejemplo de cadena homopolimérica, donde cada residuo únicamente tiene dos conformaciones: la conformación *nativa* y la *no nativa*.

Para calcular la degeneración asociada a cada nivel de energía, es necesario contar todos los estados que poseen la misma energía y para hacerlo solamente necesitamos considerar las formas en las cuales se puede acomodar α objetos en K lugares $\binom{K}{\alpha}$ multiplicada por el número de conformaciones que se tienen $\nu^{K-\alpha}$, por lo que la función de degeneración g_α tendrá la siguiente forma:

$$g_\alpha = \nu^{K-\alpha} \binom{K}{\alpha}. \quad (2.33)$$

La expresión de la energía E_α planteada por Zwanzig es un punto crucial en el modelo debido a que, como se mencionó en líneas anteriores, calcular

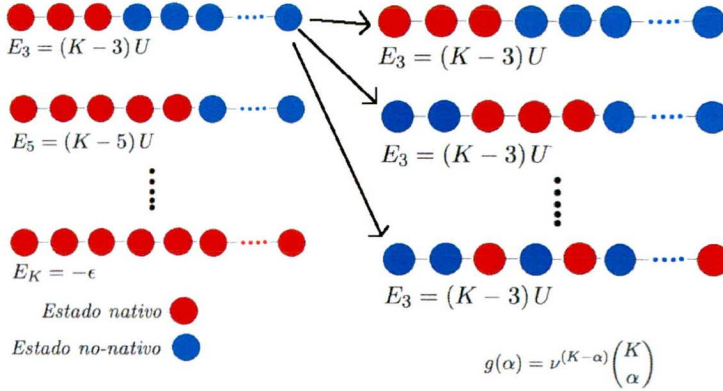


Figura 2.4: Cadena homopolimérica con K aminoácidos. De lado izquierdo se muestra la cadena para distinto número de aminoácidos en el estado nativo. Cada cadena de la derecha representa un microestado del sistema. Para una cadena con α conformaciones *nativas* existen distintas configuraciones. Del lado derecho se encuentran cadenas con el mismo número de aminoácidos y la misma cantidad de residuos en el estado nativo pero tienen distinta distribución de los mismos.

dicha expresión para E_α podría ser una tarea extremadamente complicada debido a que se tendría que definir un potencial de interacción para la cadena polipeptídica con K aminoácidos que la componen y posteriormente resolver la ecuación de Schrodinger para dicho potencial³; en lugar de todo ese formalismo, se propone E_α con ciertas hipótesis que la sustentan para poder obtener, en base a la mecánica estadística, la función de partición Z_{hom} , la cual su cálculo es parte importante del modelo planteado para poder obtener las variables termodinámicas de nuestro interés.

Una vez que se tiene una expresión analítica para la energía E_α podemos calcular a partir de la función de partición Z_{hom} (capítulo 2) las propiedades termodinámicas inherentes a la cadena homopolimérica como son la energía promedio $\langle E \rangle$, el calor específico Cv y la entropía configuracional S .

$$Z = \sum_{\alpha=0}^K g_\alpha e^{\beta E_\alpha} \quad (2.34)$$

donde $\beta = \frac{1}{k_B T}$ es el factor de Boltzmann. Sustituyendo las expresiones

³En el apéndice de este trabajo se hace el cálculo para la energía (E) de una partícula libre sometida a un cierto potencial para ejemplificar el procedimiento para obtener E . Ahora para el caso de una cadena polimérica el procedimiento se vuelve una tarea casi imposible de resolver usando la ecuación de Schrodinger $H\Psi_j = \epsilon_j\Psi_j$.

de g_α y E_α en Z_{hom} , tenemos que:

$$Z_{hom} = \sum_{\alpha=0}^K \nu^{K-\alpha} \binom{K}{\alpha} e^{\beta(K-\alpha)U - \epsilon\delta_{K,\alpha}} \quad (2.35)$$

Desarrollando esta expresión⁴, la función de partición finalmente queda como:

$$Z_{hom} = e^{\beta\epsilon} + \left(1 + \nu e^{-\beta U}\right)^K - 1 \quad (2.36)$$

Teniendo la función de partición para una cadena homopolimérica del modelo de Zwanzig podemos obtener toda la información termodinámica de dicho sistema. Como se ha mencionado anteriormente, la función de partición Z_{hom} en el modelo de Zwanzig cuenta los estados que son accesibles al sistema, que corresponden a la expresión E_α propuesta en el modelo. Dado que ya tenemos una expresión para la función de partición, podemos obtener variables termodinámicas como son el calor específico Cv , la energía libre F , la entropía S , entre otras a partir de cantidades microscópicas como ϵ , U , y ν .

La siguiente gráfica corresponde al Cv obtenido a partir de la función de partición del modelo de Zwanzig para diferentes valores de la cantidad de residuos en la cadena homopolimérica K y como función de la temperatura T .

En la figura 2.6 se muestra la gráfica de la entropía configuracional como función de la temperatura T , donde se aprecia un cambio de un estado de baja entropía a otro de mayor entropía. Este cambio se vuelve mas abrupto cuando la cantidad de aminoácidos en la cadena polipeptídica aumenta. Como se mencionó en el capítulo 2, la entropía puede ser expresada en términos de la energía libre de Gibbs por la expresión (2.37) y a su vez, podemos caracterizar la transición de fase en la temperatura crítica $T = T_c$ observando si existe una discontinuidad en la entropía dando lugar a una transición de fase de primer orden; en el caso contrario, que no existiera dicha discontinuidad, la transición seria de segundo orden. En la figura (2.5) podemos ver las distintas curvas de la entropía para varias cantidades de residuos K . La entropía para $K = 3$ es una curva suave; pero al ir aumentado hasta $k = 50$ se puede observar que esa curva tiende a tener una discontinuidad. Aunado a esto, el límite termodinámico considera que $K \rightarrow \infty$, por lo que la entropía para una cadena muy larga tendrá una discontinuidad, llevando a la conclusión que el sistema en cuestión presenta una transición de fase de primer orden.

⁴El desarrollo completo de la expresión esta en el apéndice A.

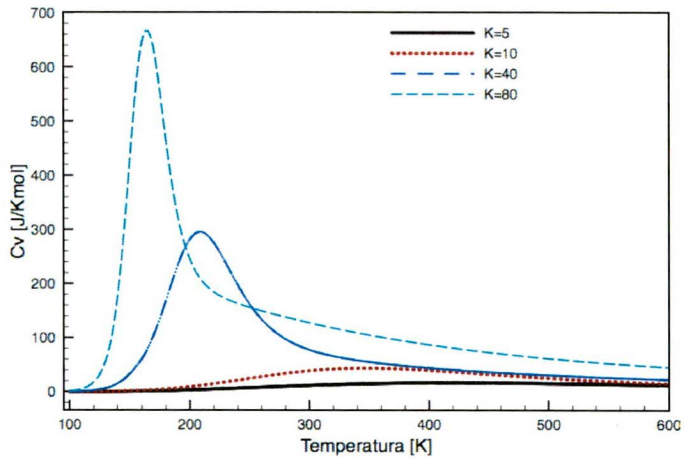


Figura 2.5: Curvas del calor específico del modelo de Zwanzig con diferentes valores de K en las secuencias. $U = 3 \times 10^3 J/mol$, $\epsilon = 5 \times 10^4 J/mol$.

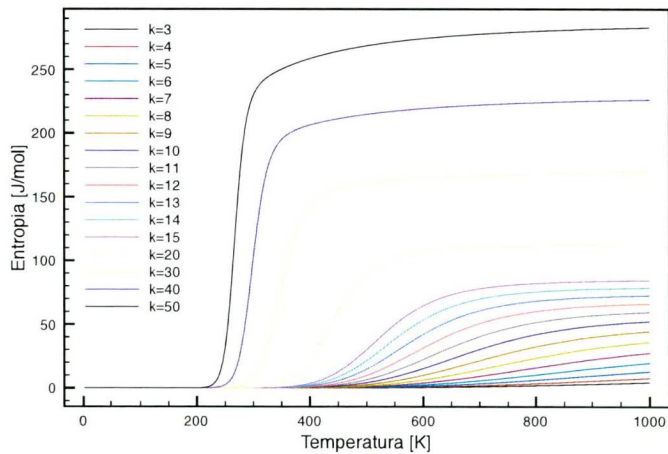


Figura 2.6: Gráfica de la entropía del modelo de Zwanzig con diferentes valores de K residuos. $U = 3 \times 10^3 J/mol$, $\epsilon = 5 \times 10^4 J/mol$.

El modelo de Zwanzig permite el estudio analítico para el estudio de las propiedades termodinámicas a partir del planteamiento de la energía E_α , pero al depender directamente de las variables microscópicas ϵ y U lo vuelven limitado. Esto se debe a que en los experimentos con proteínas en los cuales se accede a dichas variables a través de cambios en variables macroscópicas, no hay un control sobre de ellas experimentalmente.

$$S = - \left(\frac{\partial G}{\partial T} \right)_P \quad (2.37)$$

Debido a esto se hace uso de nuevas variables [38] que se pueden medir experimentalmente definiendo a f y Δ de la siguiente forma:

$$f = \left| \frac{\epsilon}{KU} \right| \quad (2.38)$$

f se define como la razón positiva entre la energía del estado fundamental $E_{\alpha=K}$ y la energía de la estructura totalmente desplegada $E_{\alpha=0}$ y

$$\Delta = -(\epsilon + KU) \quad (2.39)$$

donde Δ es la diferencia entre las energías del estado plegado y la energía del estado no plegado.

A partir de estas variables se podrá verificar las propiedades termodinámicas para observar las transiciones de fase como implicación de los detalles del espectro de energía del modelo de Zwanzig cambiando los valores de f y Δ .

Como se ha mencionado en capítulos anteriores, la propensidad helicoidal, la cual da información acerca de la posibilidad que un aminoácido tiene para participar en una hélice α y esta relacionada directamente con el número de conformaciones no nativas ν . A partir de la primera ley de la termodinámica, de un estado no nativo al estado nativo, tenemos que el cambio de la energía libre de Helmholtz por residuo esta dada por:

$$\Delta F_i = \Delta H_i - T \Delta S_i \quad (2.40)$$

donde ΔH_i y ΔS_i son la entalpía y la entropía por residuo respectivamente. Si únicamente nos enfocamos en la transición hélice-ovillo [43] de la cadena polipeptídica, la propensidad helicoidal ω del modelo de Lifson-Roig [27] puede ser la función de partición de los estados helicoidales, por lo que el cambio de la energía libre puede ser expresada en términos de dicha función de partición como $\Delta F_i = -RT \log(\omega_i)$. Por otro lado, cada residuo puede estar en una transición de un estado no nativo ν al estado

nativo, entonces el cambio en la entropía por residuo ΔS_i esta dada por $\Delta S_i = S_f - S_i = R \log(1) - R \log(\nu)$, tenemos que $\Delta S_i = -R \log(\nu)$.

Teniendo estas expresiones para ΔF_i y ΔS_i , la ecuación 2.40 toma la siguiente forma:

$$-RT \log(\omega) = \Delta H_i - T(-R \log(\nu)) \quad (2.41)$$

$$\log(\omega) = -\beta \Delta H_i - \log(\nu) \quad (2.42)$$

$$\log(\omega) - \log(\nu) = -\beta \Delta H_i \quad (2.43)$$

$$\log\left(\frac{\omega}{\nu}\right) = -\beta \Delta H_i \quad (2.44)$$

$$(2.45)$$

Finalmente, tenemos ahora una expresión que relaciona a la propensidad helicoidal ω con la cantidad de conformaciones no nativas ν , por lo que es posible implementar en el modelo de Zwanzig cualquier aminoácido ya que cada uno de los mismos tiene una propensidad helicoidal propia, siendo Ala con el valor mas alto y Pro con el valor mas bajo.

$$\omega = \frac{1}{\nu} \exp(-\beta \Delta H_i) \quad (2.46)$$

donde ΔH_i es la contribución de la entalpía por residuo de pasar de algún estado no plegado al estado plegado. Los detalles del calculo de Q están dados en [38].

Una vez que se sustituyen las variables f y U en la expresión de la energía E_α , la nueva función de partición Z_{hom} del modelo de Zwanzig tiene la siguiente forma:

$$Z_{hom} = \exp\left(\frac{-\beta f \Delta}{1+f}\right) + \left[1 + \omega^{-1} \exp\left(\beta \left[\frac{\Delta}{K(1+f)} - \Delta H\right]\right)\right]^K - 1 \quad (2.47)$$

A partir de Z_{hom} podemos obtener las funciones termodinámicas como el C_v , la energía promedio $\langle E \rangle$ y la entropía.

Existen regiones para valores de f , las cuales indicarán por medio del C_v que tan favorable es que ocurra una transición de fase para fluctuaciones de la energía. Para observar este comportamiento nos basamos en un ejemplo de una cadena de $K = 20, 30, 40$ aminoácidos donde los residuos son todos alaninas (poly-Ala). En la figura 2.7 se muestra el comportamiento del calor específico C_v y, en la subfigura, el comportamiento de la entropía configuracional S . La temperatura de desnaturalización T_d disminuye

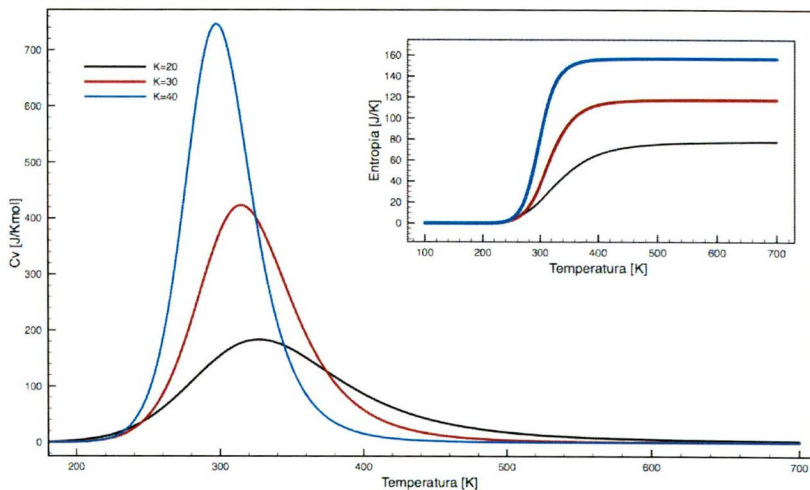


Figura 2.7: En esta gráfica se muestra el comportamiento del calor específico C_v de tres cadenas poly-Ala con 20, 30 y 40 aminoácidos en su secuencia. Es claro observar que las temperaturas de desnaturalización T_d cambian con respecto a la longitud K de la cadena en cuestión, habiendo una disminución en la misma conforme K incrementa. $f = 1$, $\Delta = -5 \times 10^4 \text{ J/mol}$, $\Delta H = 0$, valores utilizados para el calculo. También se muestra en la subfigura el comportamiento de la entropía configuracional S de las tres cadenas.

conforme el valor de K aumenta, concluyendo que la transición de fase que ocurre en las cadenas homopoliméricas dependen del tamaño o longitud de la secuencia en el vacío [38]. Esta tendencia es debida a que los grados de libertad aumentan conforme K aumenta. Dicha transición es también comprobada en el comportamiento de S y a su vez se realiza de manera abrupta conforme K aumenta.

En el siguiente capítulo se desarrollará la generalización del modelo de Zwanzig a partir del cambio de la función de degeneración g_α en función de los valores de las conformaciones no nativas ν de cada residuo, que a su vez dependen del valor de las propensidades helicoidales de los aminoácidos que forman a alguna cadena heteropolimérica.

Capítulo 3

MODELO DE HETEROPOLIMERO IDEAL

En el capítulo 2 hemos planteado el modelo de Zwanzig para cadenas homopoliméricas, el cual plantea una expresión para la energía E_α que depende de las variables $K, f, \Delta, \omega, \Delta H$ y T . La función de degeneración juega un rol importante al calcular la función de partición Z_{hom} debido a que dicha función calcula la cantidad de microestados que una cadena homopolimérica pueda tener. El carácter homopolimérico viene precisamente de que todos los K aminoácidos tienen el mismo valor de ν , por lo que la cadena tiene en total $(\nu+1)^K$ posibles configuraciones.

Tomaremos ahora el caso en el cual cada aminoácido tenga una cantidad de conformaciones *no nativas* ν_j distinta, en general, a los demás residuos donde $0 \leq j \leq K$. Por ejemplo, teniendo una cadena con $K = 5$ aminoácidos, la cantidad de configuraciones totales que tiene la cadena es $(\nu_1 + 1)(\nu_2 + 1)(\nu_3 + 1)(\nu_4 + 1)(\nu_5 + 1)$, donde $(\nu_1 + 1)$ corresponde al primer residuo en la secuencia de aminoácidos y así sucesivamente. A continuación se plantea un modelo de grano grueso basado en el modelo de Zwanzig para cadenas heteropoliméricas, definiendo una nueva función de degeneración G_α^K la cual es parte fundamental en este trabajo.

3.1. Modelo para un heteropolímero basado en el Modelo de Zwanzig

En la teoría de los paisajes de energía (Energy Landscape Theory)[6] se propone que el plegamiento de una proteína esta fuertemente relacionado

con los mínimos en la energía libre y al decremento de la entropía. Esto lleva a que la energía libre $F = E - ST$ se mantenga siempre en un mínimo global y se alcance el estado nativo en el cual puede realizar la función biológica específica que caracteriza a cada una de ellas.

Definimos a Υ como el vector cuyas entradas son los valores de las conformaciones no nativas que tienen cada residuo en la secuencia heteropolimérica, es decir:

$$\Upsilon = (\nu_1, \nu_2, \nu_3, \nu_4, \dots, \nu_k) \quad (3.1)$$

La degeneración que presenta un heteropolímero es una función que calcula la cantidad de microestados con la misma energía, y la representaremos como G_α^K . Ahora, a diferencia de las cadenas homopoliméricas, las cadenas heteropoliméricas difieren de esta hipótesis, donde todos los residuos tienen la misma cantidad de conformaciones, asignando a cada residuo de la cadena polipeptídica un conjunto de conformaciones ν_j , donde j es el aminoácido j -ésimo, por lo que el primer residuo $K = 1$ tiene ν_1 conformaciones no nativas, el segundo residuo $K = 2$ tiene ν_2 conformaciones no nativas y así sucesivamente hasta el residuo K que tiene ν_K conformaciones no nativas.

En la tabla 3.1 se muestra un ejemplo para una cadena heteropolimérica con tres residuos, cada uno de los cuales tiene $\nu_j + 1$ conformaciones, donde ν_j son las conformaciones no nativas y únicamente hay 1 conformación nativa. Cabe mencionar que generalmente la cantidad de conformaciones no nativas de cada aminoácido son distintas entre sí.

$\nu_1 + 1$	$\nu_2 + 1$	$\nu_3 + 1$
-------------	-------------	-------------

Tabla 3.1: Cadena heteropolimérica con 3 residuos distintos donde cada bloque representa un aminoácido y cada uno de ellos tiene $\nu_j + 1$ posibles conformaciones donde ν_j son las no nativas y solamente 1 es nativa.

Analicemos cada uno de los casos donde una cadena con tres aminoácidos puede tener distintos valores de la función G_α^K para todos los valores de α , donde $0 \leq \alpha \leq K$. En esta cadena, para $\alpha = 0$, la degeneración de la cadena sería $G_{\alpha=0}^{K=3} = \nu_1 \nu_2 \nu_3$, ya que ningún residuo se encuentra en la conformación nativa y cada residuo tiene un conjunto de conformaciones ν_j no nativas.

3.1. MODELO PARA UN HETEROPOLÍMERO BASADO EN EL MODELO DE ZWANZIG49

Ahora, si $\alpha = 1$, habrá un residuo en la conformación nativa y los demás tendrán sus respectivos valores de ν , por lo que los posibles estados en los que se puede encontrar la cadena heteropolimérica $K = 3$ están representados en la tabla 3.2

1	ν_2	ν_3
ν_1	1	ν_3
ν_1	ν_2	1

Tabla 3.2: Conformaciones de una cadena heteropolimérica de tres residuos ($k = 3$) con un estado nativo $\alpha = 1$.

Ya que hay tres configuraciones distintas como se muestra en la tabla 3.2, la función de degeneracidad estará dada por la siguiente expresión, contando todos los estados posibles en los que se encuentre la cadena con $\alpha = 1$:

$$G_{\alpha=1}^{K=3} = \nu_1\nu_2(1) + \nu_1(1)\nu_3 + (1)\nu_2\nu_3. \quad (3.2)$$

El caso en el cual $\alpha = 2$, tendremos que la cadena tendrá dos residuos en un estado nativo y sólo un residuo con un conjunto de conformaciones ν_j . En la tabla siguiente se muestran los posibles estados:

ν_1	1	1
1	ν_2	1
1	1	ν_3

Tabla 3.3: Conformaciones de una cadena heteropolimérica de tres residuos ($k = 3$) con dos estados nativos $\alpha = 2$.

Para el primer residuo se tiene que las conformaciones totales son $\nu_1(1)(1) = \nu_1$ mas las conformaciones que puede adoptar el segundo residuo cuando los otros dos se encuentran en un estado nativo, teniendo $(1)\nu_2(1) = \nu_2$ mas las conformaciones del tercer residuo manteniendo a los dos restantes con una conformación nativa $(1)(1)\nu_3 = \nu_3$, por lo que la función de degeneracidad de la cadena está dada por la suma de cada conformación no nativa como se muestra en la ecuación siguiente:

$$\begin{aligned} G_{\alpha=2} &= \nu_1(1)(1) + (1)\nu_2(1) + (1)(1)\nu_3 \\ &= \nu_1 + \nu_2 + \nu_3 \end{aligned} \quad (3.3)$$

Finalmente, el caso en donde la cadena heteropolimérica se encuentra en un estado totalmente nativo ($\alpha = 3$), solamente existe una conformación como se muestra en la tabla 3.4 donde cada residuo tiene una única

conformación que es la nativa.

1	1	1
---	---	---

Tabla 3.4: Conformaciones de una cadena heteropolimérica de tres residuos ($k = 3$) con dos estados nativos $\alpha = 3$.

Evidentemente la función de degeneración es:

$$G_{\alpha=3}^{K=3} = 1 \quad (3.4)$$

En la tabla 3.5 se muestran los resultados anteriores, considerando que α toma los valores de cero hasta el número total de aminoácidos k ($0 \leq \alpha \leq K$). En la columna *Cadena* se aprecian las posibles configuraciones que adopta la cadena con $K = 3$ para los diferentes valores de α . Por ejemplo, en el caso de $\alpha = 1$ se tienen tres posibles microestados, los cuales cada uno de ellos tiene un residuo en el estado nativo, tal y como están escritos de manera explícita en dicha tabla.

Valor de α	Cadena	Función $G_{\alpha}^{K=3}$
$\alpha = 0$	$\nu_1 \nu_2 \nu_3$	$G_{\alpha=0}^{K=3} = \nu_1 \nu_2 \nu_3$
$\alpha = 1$	$\nu_1 \nu_2 (1)$ ó $\nu_1(1) \nu_3$ ó $(1)\nu_2 \nu_3$	$G_{\alpha=1}^{K=3} = \nu_1 \nu_2 + \nu_1 \nu_3 + \nu_2 \nu_3$
$\alpha = 2$	$(1)(1)\nu_3$ ó $(1)\nu_2(1)$ ó $\nu_1(1)(1)$	$G_{\alpha=2}^{K=3} = \nu_3 + \nu_2 + \nu_1$
$\alpha = 3$	$(1)(1)(1)$	$G_{\alpha=3}^{K=3} = 1$

Tabla 3.5: En esta tabla se presentan los posibles valores de la función de degeneración G_{α}^K en función de los valores de ν_i para el caso de una cadena con $K = 3$ residuos.

Al aumentar el número de residuos en el heteropolímero la función de degeneración se vuelve muy complicada debido al gran número de estados posibles que se pueden adoptar, ya que depende de los valores de cada conformación de las ν_k . Para calcular la función de degeneración, se construyó un script¹ en el programa de computo *Mathematica* que se presenta en el apéndice de este trabajo.

La idea general del algoritmo es generar vectores con ceros y unos, donde cero se refiere a un estado nativo y un uno se refiere a un estado no nativo. Posteriormente se cuentan las posibles permutaciones que se pueden formar teniendo diferentes cantidades de estados nativos. Al tener dichas

¹Es un programa que por lo regular se almacena en un archivo y sigue una serie de instrucciones para un cierto fin.

permutaciones se asocian a cada valor de la energía.

La descripción del programa es la siguiente:

- Se define una cadena con un número de aminoácidos fijos k .
- Definimos el vector ph el cual tiene todos los valores de las propensidades helicoidales ω de cada aminoácido ordenada de mayor a menor valor.
- Definimos el vector de conformaciones Υ antes mencionado como el vector v cuya dimensión es igual al número de aminoácidos en la cadena residual.
- Definimos un vector a el cual tendrá la misma dimensión que el vector Υ y en sus entradas habrá solamente elementos de valor 1.
- Ahora, se generan una cantidad de K vectores B_j con K entradas (con $0 \leq j \leq k$) los cuales tendrán j -ceros en sus primeras entradas y en el resto de las entradas mantendrá solamente valores de 1, por ejemplo, si $j = 2$, el vector B tendrá dos ceros; con B_5 , el vector B tendrá 5 ceros hasta que el vector B_k tenga en todas sus entradas igual a cero. Estos ejemplos quedan representados de la siguiente forma:

$$B_2 = (0, 0, 1, 1, 1, 1, 1, \dots, 1) \quad (3.5)$$

$$B_5 = (0, 0, 0, 0, 0, 1, 1, \dots, 1) \quad (3.6)$$

$$B_k = (0, 0, 0, 0, 0, 0, \dots, 0, 0) \quad (3.7)$$

En cada vector B_j tiene tantos ceros como número de conformaciones *nativas* hay, es decir, si tenemos el vector B_6 con $j = 6$, entonces hay 6 aminoácidos en la conformación *nativa*. Tomando el ejemplo anterior, donde se tiene una cadena de tres aminoácidos ($k = 3$), considerando el caso donde $\alpha = 1$, la función de degeneracidad G_α es:

$$G_\alpha = \nu_1\nu_2 + \nu_1\nu_3 + \nu_2\nu_3 \quad (3.8)$$

Para este caso, el vector B_1 está escrito de la siguiente forma:

$$B_1 = (0, 1, 1) \quad (3.9)$$

Ahora, los números 1 en el vector B_j representan las conformaciones no nativas ν_j y el cero la conformación *nativa*, ya que al contar las

conformaciones totales para la función G , únicamente nos interesan las que consideran a las *no nativas*, lo único que faltaría es calcular las permutaciones² del mismo B_1 , que serían: $(1, 1, 0)$, $(1, 0, 1)$, $(0, 1, 1)$, teniendo tres configuraciones diferentes para un mismo valor de $\alpha = 1$ y dan lugar a los valores de la función de degeneracidad mencionados en el ejemplo en el que nos basamos.

- Se realiza un ciclo llamado *Do* el cual creará un vector P_j con K entradas en las cuales se guardará la información de todas las posibles permutaciones que puede tener algún vector B_i respecto a sus entradas, por ejemplo, si se tiene el vector B_2 con la configuración $B = 0, 1, 0$, la cantidad de permutaciones que se tiene es 3, $\{\{0, 1, 1\}, \{1, 0, 1\}, \{1, 1, 0\}\}$. El objetivo de este punto es crear las diferentes posibilidades de conformaciones que se puedan tener en una cadena residual.
- Se crea otro ciclo *Do* para generar los vectores q_i con K entradas cuyos valores se obtendrán a partir del producto de las entradas de los vectores v y el vector d que contiene a la partición de la tabla de valores de los estados nativo y no nativo. El primer elemento del primer vector se multiplica por el primer elemento del segundo vector y dicho resultado será el valor de la primera entrada de q_i y así sucesivamente con las demás entradas.
- Por último, se crea otro ciclo *Do* para generar los elementos de la función de degeneracidad g_α a partir de la suma de los elementos del vector q_i definiendo los límites de las sumas y los productos de dichos elementos.

La cantidad de procesos que realiza este programa aumenta conforme la cantidad de residuos K crece.

Por otro lado, se formuló una expresión matemática con la intención de obtener la función de degeneracidad G_α de forma analítica, pero no se llegó a dicho objetivo debido a la complejidad de dicho desarrollo y el tiempo para obtenerlo, por lo que solamente se llegó a una expresión en la cual podemos sustituir los valores de α para obtener la cantidad de estados degenerados.

²Una permutación es la variación del orden o de la disposición de los elementos de un conjunto. Por ejemplo, si se tiene el siguiente conjunto $\{1, 2, 3\}$, cada ordenación posible de sus elementos, sin repetirlos, es una permutación, es decir, existen seis permutaciones para el ejemplo anterior, teniendo las siguientes ordenamientos: "1, 2, 3", "1, 3, 2", "2, 1, 3", "2, 3, 1", "3, 1, 2" y "3, 2, 1".

$$G_{\alpha}^K = \left(\prod_{i=1}^K \nu_i \right) \left(\sum_{\xi_1=1}^{K-\alpha+1} \frac{1}{\nu_{\xi_1}} \sum_{\xi_2=1}^{K-\alpha+2} \frac{1}{\nu_{\xi_2}} \cdots \sum_{\xi_{\alpha}=1}^K \frac{1}{\nu_{\xi_{\alpha}}} \right) \quad (3.10)$$

donde $\prod_{i=1}^K \nu_i$ muestra la cantidad total de estados en el estado *no nativos* y $\sum_{\xi_{\alpha}=1}^{K-\alpha+1} \frac{1}{\nu_{\xi_{\alpha}}}$ es la suma de términos alternados los cuales tienen términos extras los cuales se eliminarán con la expresión anterior.

Por ejemplo, si tenemos una cadena con $K = 3$ residuos, podemos sustituir en la expresión anterior dichos valores para $\alpha = 2$ teniendo el siguiente resultado:

$$G_{\alpha=0}^{K=3} = \prod_{i=1}^3 \nu_i \sum_{j=1}^2 \left(\frac{1}{\nu_j} \sum_{j+1}^3 \frac{1}{\nu_{j+1}} \right) \quad (3.11)$$

desarrollando esta expresión, tenemos:

$$G_{\alpha=0}^{K=3} = (\nu_1 \nu_2 \nu_3) \left[\frac{1}{\nu_1} \sum_{j=2}^3 \frac{1}{\nu_1} + \frac{1}{\nu_2} \sum_3^3 \frac{1}{\nu_1} \right] \quad (3.12)$$

reduciendo esta expresión nos queda:

$$G_{\alpha=0}^{K=3} = (\nu_1 \nu_2 \nu_3) \left[\frac{1}{\nu_1 \nu_2} + \frac{1}{\nu_1 \nu_3} + \frac{1}{\nu_2 \nu_3} \right] \quad (3.13)$$

$$G_{\alpha=2}^{K=3} = \nu_3 + \nu_2 + \nu_1 \quad (3.14)$$

Este resultado es congruente con lo esperado en el caso de un heteropolímero, con $K = 3$ residuos y $\alpha = 2$, únicamente que los términos de las sumas no son una forma analítica de tal forma que se pueda programar de manera directa y fácil como en el caso del modelo de Zwanzig, en donde la expresión de la degeneracidad es una expresión totalmente analítica.

3.2. Comparación del MHI con el modelo de Zwanzig

Los valores de las conformaciones ν_j en general son diferentes entre sí, pero si fuesen iguales, estaríamos en el caso de un homopolímero, regresando al modelo de Zwanzig. La manera de comprobar que los resultados del modelo de heteropolímeros ideales son correctos, es comparar las funciones de partición Z_{hom} y Z_{het} del modelo de Zwanzig y modelo de heteropolímeros respectivamente en el caso en que $\nu_i = \nu_j$. Para ello se

utilizaran los programas de computo hechos para cada modelo.

En la figura ??, se muestran las dos gráficas del calor específico Cv para el caso de dos cadenas con $K = 5$, las cuales fueron estudiadas por el modelo de Zwanzig (a) y por el modelo de heteropolímero ideal (b), respectivamente. En (a) se consideró que el valor de las conformaciones no nativas para cada residuo fueran iguales, es decir, que $\nu_j = \nu$ para cualquier valor de j , en particular, se escogió el valor de ν promedio para el caso del aminoácido alanina ($\omega = 1.61$, $\nu = 1/1.61$). Al obtener sus respectivas funciones de partición, las cuales son iguales entre si por la función de degeneracidad g_α^K y G_α^K respectivamente, se calculó el calor específico para obtener la temperatura de desnaturalización de cada cadena. Los resultados se muestran en la figura ??, donde se muestran las dos gráficas del calor específico de los dos modelos, observando que son exactamente las mismas entre si, teniendo un valor para la temperatura de transición de $T = 304K$ en los dos casos, concluyendo que el modelo de heteropolímero ideal es una generalización del modelo de Zwanzig.

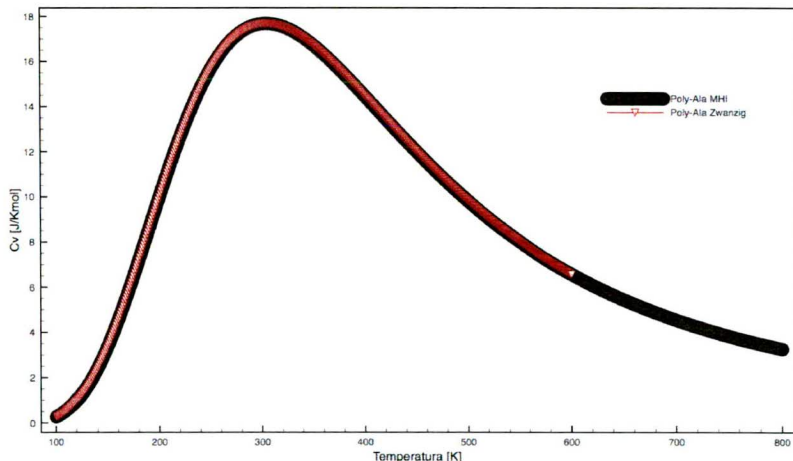


Figura 3.1: Gráficas del calor específico de la cadena Poly-Ala calculadas con los programas para cadenas homopoliméricas; y cadenas heteropoliméricas. Los valores utilizados en las dos gráficas fueron $K = 5$, $\Delta = -1.3 \times 10^4 J/mol$, $f = 1$, $\epsilon = 6500 J/mol$, $U = 1300 J/mol$. Es evidente que las curvas son iguales independientemente de los dos programas utilizados, concluyendo que el programa que calcula las propiedades termodinámicas de cadenas heteropoliméricas se reduce al caso de los homopolímeros.

Una forma de comprobar que efectivamente la función G_α^K es una ge-

Función de degeneracidad Zwanzig	Función de degeneracidad heteropolímeros
$g_0 = \nu^5$	$G_0^{K=5} = \nu_1\nu_2\nu_3\nu_4\nu_5 = \nu^5$
$g_1 = 5\nu^4$	$G_1^{K=5} = \nu + \nu + \nu + \nu + \nu$
$g_2 = 10\nu^3$	$G_2^{K=5} = 10\nu^3$
$g_3 = 10\nu^2$	$G_3^{K=5} = 10\nu^2$
$g_4 = 5\nu^1$	$G_4^{K=5} = 5\nu^1$
$g_5 = 1$	$G_5^{K=5} = 1$

Tabla 3.6: En la primera columna se muestra la función de degeneracidad de un homopolímero con $K=5$ residuos los cuales tienen ν conformaciones no nativas en cada aminoácido. En la columna derecha se muestra una cadena heteropolimérica la cual ha tenido el caso que sus conformaciones son todas iguales, reduciendo al modelo de Zwanzig.

neralización de g_α sería, en primera instancia, obtener una forma analítica para G_α^K y reducirla a la función g_α para el caso en que todas las conformaciones no nativas sean iguales, o sea, $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \dots = \nu_K$. Al no contar con dicha expresión analítica, probamos para ciertos casos. Como un ejemplo, dos cadenas las cuales tengan el mismo número de residuos K . A partir de las funciones de degeneracidad se presenta la tabla [3.7] con dos cadenas con $K = 5$ residuos donde en la primer secuencia se utiliza la función g_α y en la segunda G_α^K . Cuando los valores $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \nu_5$ entonces $G_0^{K=5} = \nu^5$, $G_1^{K=5} = 5\nu^4$, $G_2^{K=5} = 10\nu^3$, $G_3^{K=5} = 10\nu^2$, $G_4^{K=5} = 5\nu^1$ y $G_5^{K=5} = 1$ lo cual indica que el algoritmo de la función de degeneracidad para cadenas heteropoliméricas G_α reproduce los resultados del modelo de Zwanzig con la misma expresión de la energía E_α .

3.2.1. Caso de estudio I: Mutaciones en péptidos

Se sabe que el ADN en conjunto con el ARN realizan la tarea de crear las secuencias específicas de las proteínas a través de un proceso llamado transcripción en el cual la información genética contenida en la molécula de ADN es copiada a una molécula de ARN, normalmente conocido como ARN-mensajero[29]. Los tres procesos que establecen la conservación y transmisión de la información genética son los siguientes:

- Replicación o copiado del ADN para formar moléculas idénticas
- Transcripción, proceso por el que el mensaje genético contenido en el ADN se transcribe al ARN mensajero

- Traducción, proceso por el cual, el mensaje genético es llevado a los ribosomas donde el ARN mensajero, actuando como molde, dirige la secuencia específica de aminoácidos durante la biosíntesis de proteínas.

Esta síntesis de proteínas da lugar a la estructura primaria de las proteínas la cual posteriormente obtiene su estructura terciaria con el objetivo de realizar una función biológica específica.

La replicación de la molécula de ADN es primordial en la creación de proteínas. Una célula solo tiene generalmente una o dos copias de las secuencias de ADN. Las proteínas y las moléculas de ARN dañadas pueden ser reemplazadas rápidamente utilizando información codificada en el ADN; pero en el caso del ADN, las moléculas son irremplazables [30]. Este hecho hace de la información contenida en el ADN sea de las prioridades más altas dentro de la célula para que se mantenga intacta, por lo que se tiene un sistema muy elaborado para la reparación de ADN. Debido a varios factores ambientales, el ADN puede resultar dañado, incluso en la propia replicación llega a suceder que la información contenida se vea afectada al dejar las bases mal apareadas. En el caso de que una lesión en el ADN no sea reparada, existirá un cambio en la secuencia de bases de ADN que si se replica y transmite a las generaciones futuras se hace permanente. Un cambio permanente en la secuencia de nucleótida del ADN hace que la secuencia de aminoácidos creada por el ARN sea diferente a la secuencia original, para algún péptido dado.

En general, se esperaría que cuando un péptido tiene una mutación de alguno de sus residuos, la estructura terciaria cambie. Este hecho no siempre se cumple. Mencionaremos el caso de la proteína *lysozima*, la cual tiene una función enzimática en organismos vivos, ya que su función consiste en partir o fraccionar otras proteínas con el fin de dañar a las células bacterianas. Existen varios tipos de lysozimas, las cuales dependen del organismo en la que se encuentren. Por ejemplo, la HEW-lysozima³ [82] es un péptido que se encuentra en la clara de huevo en abundancia. La lisozima humana se encuentra, por ejemplo, en la saliva, o en las lagrimas, donde la función catalítica es básicamente para proteger de agentes bacterianos que entran en el organismo.

Las secuencias de cada una de las lysozimas mencionadas no son iguales entre si respecto a la secuencia inherente a cada proteína. En la figura 3.2 se muestran las secuencias de la HEW-lysozima y la lysozima humana, donde

³La lisozima de clara de huevo HEWL fue la primera proteína que fue secuenciada y que tiene los veinte aminoácidos usuales, además de ser también la primera de la que se obtuvo un análisis cristalográfico completo y de la que se propuso un mecanismo de reacción detallado.

evidentemente no son iguales entre si. Pero las funciones biológicas son muy similares en cada organismo, donde la estructura terciaria de cada una de ellas también son muy similares entre si, tal y como se muestra en la figura 3.3.

Secuencia de la HEW-Lysozima	Secuencia de la Lysozima humana
KVFGRCELAA	KVFERCELAR
AMKRHGLDNY	TLKRLGMDGY
RGYSLGNWVC	RGISLANWMCL
AAKFESNFNTQ	AKWESGYNTRA
ATNRNTDGSTD	TNYNAGDRSTD
YGILQINSRWW	YGIFQINSRYWC
CNDGRTPGSRN	NDGKTPGAVNA
LCNIPCSALLSS	CHLSCSALLQD
DITASVNC AKKI	NIADAVACAKRV
VSDGNGMNAW	VRDPQGIRAWV
VAWRNRCKGTD	AWRNRQNRD
VQAWIRGCRLL	VRQYVQGCGV

(a)

(b)

Figura 3.2: En esta figura se muestran las secuencias de las dos proteínas: (a)HEW-lysozima y (b) lysozima humana. En este caso es utilizada la abreviación IUB/IUPAC, donde cada residuo le corresponde una letra del alfabeto, por ejemplo, Kcorresponde al aminoácido Lys.

Los cambios en la secuencia de una cadena heteropolimérica no necesariamente cambian las funciones biológicas del péptido. El plegamiento de dichas secuencias es el punto crucial en las funciones que realiza cada péptido en los organismos.

A continuación, se realizará el estudio del comportamiento termodinámico de cadenas polipeptídicas cuando se realiza una mutación en alguno de sus residuos utilizando el modelo desarrollado aquí. Uno de los objetivos es observar si la temperatura de desnaturalización que denotaremos como Td_j , donde j representa el lugar de la secuencia donde se realiza la mutación. Para diferenciar entre cada residuo simplemente relacionamos el valor de sus conformaciones no nativas ν con el valor de su propensidad helicoidal ω la cual, considerando una barrera entalpica $\Delta H = 0$ (ver sección anterior), tenemos que $\nu = \frac{1}{\omega}$.

Tomando una cadena Poly-Ala como en el ejemplo anterior, mutaremos a cada residuo de la cadena, donde elegiremos al aminoácido Pro para realizar dicho cambio, debido a la diferencia entre sus valores de propensidad helicoidal ($\omega = 0.001$) y Alanina ($\omega = 1.61$). La figura 3.4 muestra las temperaturas de transición Td_j que existen entre la cadena Poly-Ala con $K = 5$ y las mutaciones que presenta en cada lugar de su secuencia, donde

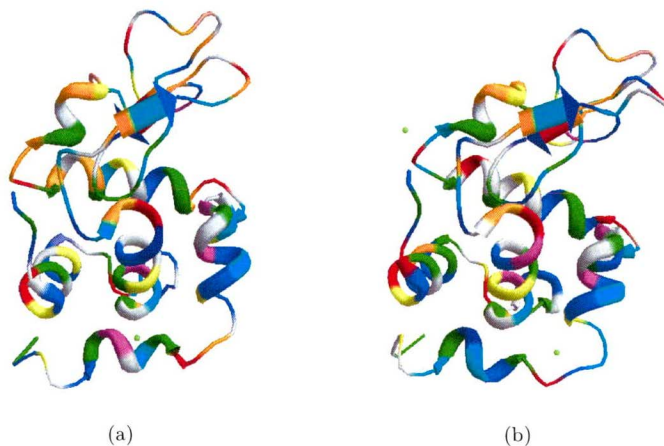


Figura 3.3: En esta figura se muestran a las proteínas (a)HEW-Lysozima y (b)H-Lysozima, cuyas funciones son muy similares entre sí, difiriendo en cuanto al organismo que lo contiene. Notese que, pese a la diferencia entre las secuencias residuales, la estructura terciaria que adoptan es muy similar una con respecto a otra. Las imágenes fueron creadas con el programa de computo *Mathematica* a partir de los archivos del Protein Data Bank 1DPX.pdb y 1IWT.pdb respectivamente.

el subíndice j indica el lugar de la mutación en la cadena, por ejemplo, si $j = 3$ indica que la mutación se realizó en el tercer aminoácido en la secuencia, visto en la gráfica. El eje horizontal de la figura se refiere al lugar de la posición j , teniendo como 0 a la cadena sin mutar, a 1 como la mutación en el primer residuo y así sucesivamente; el eje vertical indica el valor de la temperatura de transición en grados Kelvin.

Los resultados muestran que existe una disminución en la Td_j al mutar la cadena con prolina debido a que la función de degeneración depende de los valores de ω los cuales son muy distintos entre sí, siendo tres ordenes de magnitud menor el de la prolina que el de la alanina, teniendo cambios en Z_{het} y por ende en los indicadores termodinámicos como el Cv . Esto se ve reflejado en la función de partición, ya que es proporcional a ω . Ahora, es natural preguntarse la dependencia que existe entre las diferencias que existen entre la Td_j de la cadena Poly-Ala y las Td_j de las mutaciones con algún péptido X y la propensidad helicoidal ω , por lo que ahora mutamos a la cadena Poly-Ala con distintos residuos.

La figura 3.5 muestra los valores de estas diferencias, las cuales representaremos como $\Delta\tau$, cuyo valor dependen de ω , o sea que $\Delta\tau = \Delta\tau(\omega)$.

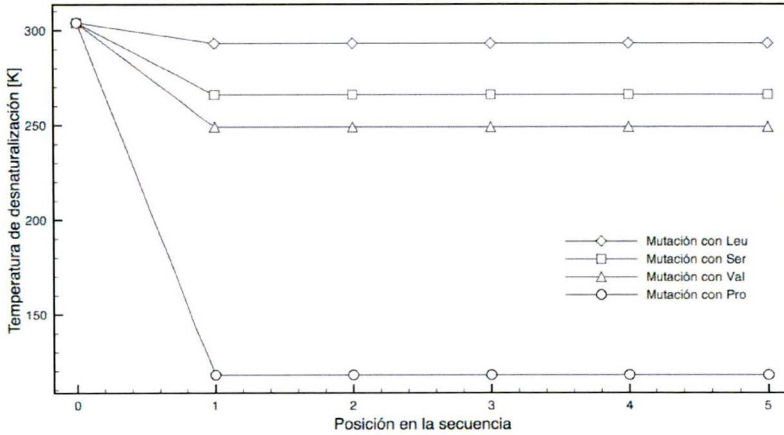


Figura 3.4: Temperaturas de transición de una cadena Poly-Ala con 4 residuos y sus mutaciones con el aminoácido Pro. Se observa que la temperatura de transición disminuye al ser mutada la cadena disminuye. Este decremento de la Td es proporcional al valor de la propensidad helicoidal del aminoácido con el que se este realizando la mutación.

Es claro ver la disminución de temperatura de desnaturalización de una cadena homopolimérica con sus posibles mutaciones la cual depende de ω . Con esta dependencia de ω podemos controlar, hasta cierto punto, la temperatura de desnaturalización de un péptido Poly-X, donde X se refiere a cualquiera de los 20 aminoácidos que existen, ya sea elevarla o disminuirla. Esta manipulación esta restringida al tipo de cadena que se maneje ya que, por ejemplo, en el caso de tener una cadena homopolimérica del tipo Poly-Pro, la cual tiene sus valores de ω muy pequeños, la única posibilidad que se tendría en un cambio de Td_j sería elevarla con alguna mutación de cualquiera de los otros 19 residuos ya que éstos tienen propensidades helicoidales mas altas que la misma prolina.

El punto importante de los resultados de la figura 3.4 es el hecho de que los valores de la temperatura de desnaturalización Td_j de una cadena mutada son los mismos entre si, lo cual podemos decir en este caso que los valores de Td_j para $j \neq 0$ ⁴ no dependen del lugar donde se realice la mutación en un homopolímero con algún aminoácido en específico. Esto es debido a que en la función de degeneracidad, al realizar una mutación en la secuencia por un termino ν_i , los sumandos de dicha función quedarán

⁴Esta condición de $j \neq 0$ es debido a que este caso es cuando no existe alguna mutación en la cadena homopolimérica.

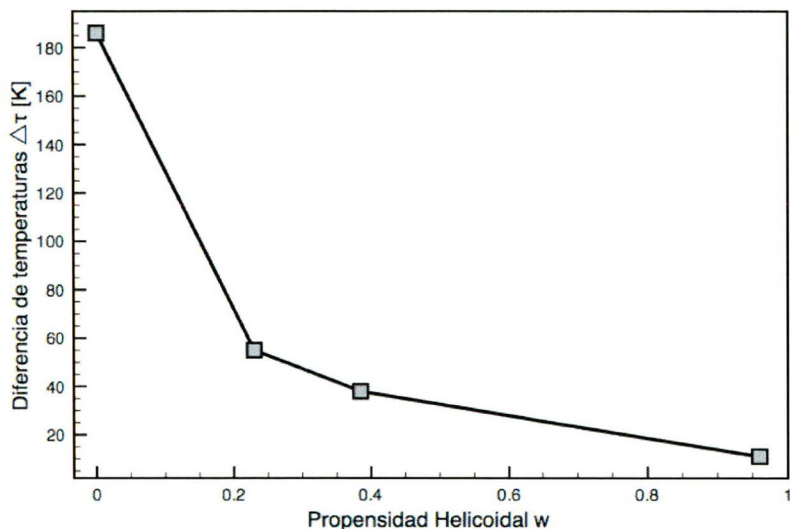


Figura 3.5: Diferencias de las temperaturas de transición de una cadena Poly-Ala con distintos residuos. Los cuatro puntos muestran el valor de la diferencia entre la Td_j de Poly-Ala y sus mutaciones con: Pro, Val, Ser y Leu respectivamente. Las variables utilizadas fueron: $\Delta = -13000J/mol$, $f = 1$, $\epsilon = 6500J/mol$ y $U = 1300$.

iguales, siempre y cuando se realice el cambio con un mismo residuo. En la siguiente sección se realizarán las mutaciones en cadenas heteropoliméricas y se calcularán las Td_j con el objetivo de observar el comportamiento de dichas temperaturas con respecto a la posición en el que se realice la mutación.

3.2.2. Caso de estudio II: Heteropolímeros aleatorios

Consideraremos ahora cadenas heteropoliméricas aleatorias donde las conformaciones totales para cada aminoácido son diferentes entre sí, es decir, un aminoácido en la posición j tendrá $\nu_j + 1$ conformaciones, donde ν_j representa al número de conformaciones no nativas del aminoácido en la secuencia heteropolimérica.

Consideraremos el caso de un heteropolímero con K aminoácidos y calcularemos las temperaturas de transición Td_j para una cadena heteropolimérica con $K = 10$ residuos y la secuencia sea elegida aleatoriamente con una probabilidad uniforme de $\frac{1}{19K}$ con el objetivo de calcular las Td_j

de las mutaciones y observar su comportamiento. Para ello realizamos un programa en *Mathematica* el cual pueda calcular las temperaturas de desnaturalización de cadenas elegidas aleatoriamente. Dicho código se encuentra escrito en el apéndice de este trabajo.

El procedimiento que realiza el script para secuencias aleatorias se comenta a continuación.

- Se define un vector con $K = 10$ de la forma $v = \nu_1, \nu_2, \nu_3, \dots, \nu_{10}$ donde cada valor de ν_j sea elegido aleatoriamente de una lista donde se tengan los valores de las propensidades helicoidales de los 19 aminoácidos sin contar a la prolina debido a que las mutaciones las realizamos con ese péptido. El comando *Random* en el comando que realiza la elección de la lista de los 19 aminoácidos que se encuentran en un vector denominado *ph* donde la distribución de probabilidad es uniforme, es decir, que cada aminoácido tiene la misma probabilidad de ser elegido por el programa.
- Después, esta elección es colocada en otro vector llamado *S* el cual contiene a los inversos de *ph* que son los valores de las conformaciones no nativas ν ya que la relación que existe entre estas dos variables es $\nu = \frac{1}{\omega}$.
- Posteriormente se realizan los cálculos para la función de partición de la misma manera que en el MHI.

Dada una secuencia elegida aleatoriamente, se sustituye el primer residuo $j = 1$ por una prolina para posteriormente calcular la temperatura de transición $Td_{j=1}$ a partir de obtener el máximo del C_v . Este proceso se realiza para todos los demás aminoácidos que forman a la cadena residual, por lo que tendremos un conjunto de $K + 1$ temperaturas de transición, donde K sera la cantidad de temperaturas de transición de cadenas mutadas y solamente 1 temperatura de transición no mutada.

Al calcular las temperaturas de desnaturalización Td_j de una cadena heteropolimérica y sus mutaciones se observa su comportamiento respecto a la posición donde se realiza la mutación. Este valor depende del aminoácido que se este mutando. La tabla ?? presenta a los valores de Td_j y son mostrados en la figura 3.6 dicho comportamiento con los valores utilizados para realizar el calculo. En este caso, las fluctuaciones que se observan en la gráfica de las Td_j difieren entre si, no mas de $5K$ entre ellas.

Mutación	0	1	2	3	4	5	6	7	8	9	10
Td_j [K]	310	197	200	201	204	197	203	200	204	200	201

Tabla 3.7: Valores de las temperaturas de desnaturalización Td_j de una cadena heteropolimérica aleatoria con secuencia Ala-Glu-Gnl-Trp-Ala-Cys-Met-Trp-Met-Gnl y sus posibles mutaciones con prolina. $\Delta = -35000J/mol$, $f = 1$, $\epsilon = 17500J/mol$ y $U = 1750J/mol$.

Por otro lado, se pretende estudiar el comportamiento colectivo de varias cadenas heteropoliméricas aleatorias con el fin de observar las variaciones de los valores de sus temperaturas de desnaturalización con respecto de otras cadenas aleatorias con la misma cantidad de aminoácidos en su secuencia. Se realizaran promedios de las Td con respecto al lugar donde se haga la mutación. Las pequeñas variaciones de las temperaturas de desnaturalización sugiere la existencia de una T_d promedio que pueden tener cada cadena polipeptídica. Para ejemplificar el proceso que se realizará, se eligieron cinco cadenas, donde su secuencias han sido elegidas aleatoriamente como en el caso de la cadena del ejemplo anterior.

Las secuencias son presentadas en la tabla [3.9] en la columna de la izquierda y en la columna de la derecha están sus respectivas temperaturas de desnaturalización; estos datos se muestran gráficamente en la figura 3.7. Esta gráfica muestra a las cuatro cadenas aleatorias junto con la cadena del primer ejemplo con sus respectivas mutaciones. Dentro de la misma imagen hay un subgráfico el cual hace énfasis de las Td_j de cada una de las cadenas en las cuales la diferencia que hay entre el máximo y el mínimo de la secuencia Val-Asn-Trp-Tyr-Met-Ile-Ile-Asn-Lys-Phe es de $5K$, la ecuación 3.15 muestra como realizar el calculo. En la tabla [3.9], la columna de la derecha se encuentran los valores de estas diferencias las cuales son muy pequeñas en comparación a la disminución de la Td_j al ser mutada por prolina. **Estas diferencias tan pequeñas motivan a pensar que, en promedio, no hay una dependencia significativa del lugar en la secuencia donde se realice la mutación tal como se ve en el subgráfico de la figura 3.7 donde todas las cinco curvas son muy planas.**

$$\Delta Td_j = Td_{max} - Td_{min}, \quad (3.15)$$

donde Td_{max} es el valor de la temperatura de desnaturalización mas grande de una cadena mutada y Td_{min} la temperatura de desnaturalización mas chica. Por lo que ΔTd_j calcula el intervalo donde estarán todas las Td_j de las mutaciones de la cadena polipeptídica.

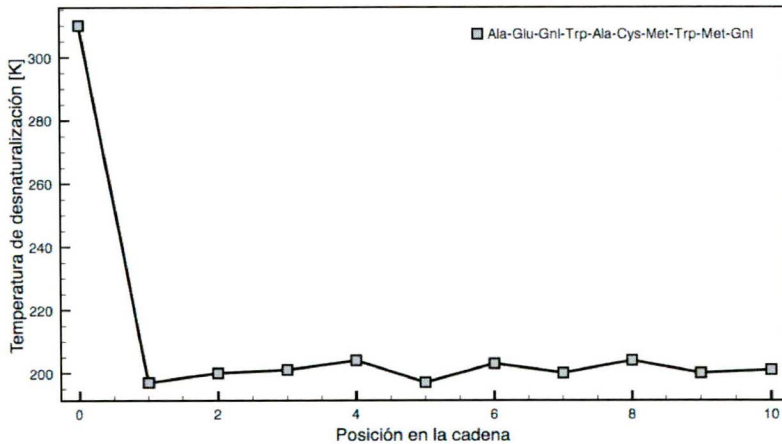


Figura 3.6: Esta gráfica muestra el comportamiento de las temperaturas de desnaturalización de la cadena elegida aleatoriamente con secuencia Ala-Glu-Gnl-Trp-Ala-Cys-Met-Trp-Met-Gnl. Los datos numéricos están presentados en la tabla [3.9]. Existe una variación pequeña que existe entre las Td_j de las mutaciones, a diferencia de las mutaciones que existen en una cadena homopolimérica en las cuales sus Td_j mutadas son las mismas por lo que no depende en que lugar se realice la mutación.

El comportamiento de las Td_j promedio puede observarse en una gráfica en la cual se promedie los valores de la Td_j en la j -ésima posición de cada cadena, por ejemplo, si tomamos las temperaturas Td_j de las cinco cadenas de la mutación en el primer residuo y realizamos el promedio de estas cinco temperaturas, el resultado es de $194K$, para la segunda posición tenemos como temperatura promedio de desnaturalización un valor de $193K$ y así sucesivamente para los restantes ocho residuos, por lo que la gráfica 3.7 puede ser expresada como otra gráfica donde los valores que contenga sean promedios de las Td_j .

La figura 3.8 en la imagen de la derecha muestra el comportamiento promedio de las cinco cadenas heteropoliméricas. Es claro observar que el tamaño de la franja ΔTd_j es muy pequeña, por lo que nos lleva a la siguiente pregunta: ¿En promedio las temperaturas Td_j dependen de la posición en la cadena? Para responder esta pregunta es necesario calcular Td_j de todas las posibles secuencias que existan para una cadena de K aminoácidos, cantidad que es extraordinariamente grande como se mencionó en párrafos anteriores 19^K , por lo que recurrimos realizar un muestreo de cadenas

Secuencia aleatoria	Td [K]	ΔTd [K]
Ala-Glu-Gnl-Trp-Ala-Cys-Met-Trp-Met-Gnl	310	7
Val-Asn-Trp-Tyr-Met-Ile-Ile-Asn-Lys-Phe	260	5
Asp-Thr-Phe-His-Ser-Asn-Leu-Trp-Ile-Ala	258	10
Thr-Gnl-Asp-Arg-Ala-His-Lys-Gnl-Arg-Val	288	12
Lys-Ala-Ile-Ser-Arg-Ile-Arg-Ser-His-Glu	321	6

Tabla 3.8: Valores de las temperaturas de desnaturalización Td_j de cuatro cadenas heteropoliméricas aleatorias con su respectiva secuencia del lado izquierdo. $\Delta = -35000J/mol$, $f = 1$, $\epsilon = 17500J/mol$ y $U = 1750J/mol$.

aleatorias a través de un script el cual realice rutinas para una cantidad considerable de posibles secuencias elegidas aleatoriamente.

A partir del script creado en *Mathematica* para una sola cadena, creamos una rutina mas grande que realice el calculo de las temperaturas de desnaturalización Td_j para una cantidad J de cadenas aleatorias. Debido al proceso de enumeración extensivo y el creciente uso de memoria, se opto por realizar este proceso con tamaños de i del orden de: 1, 10, 30, 50, 100, 200, 300, 500 y 1000 con $K = 8$ residuos en total. Este primer programa calcula una serie de procesos para obtener la temperatura de desnaturalización Td_j para una sola cadena y sus posibles K mutaciones obteniendo $K + 1$ temperaturas las cuales han sido mostradas en secciones anteriores y denotándolas como $Td_{j,i}$, donde el subíndice i indica el número de cadena elegida. Los valores obtenidos de $Td_{j,i}$ se almacenan en un vector, el cual denominaremos n_1 . En el caso de una cadena con $K = 8$ residuos, el vector n seria un vector con la $Td_{j,\mu}$ de la cadena propuesta con $K + 1$ elementos: $n_1 = (Td_{j=0}, Td_{j=1}, Td_{j=2}, Td_{j=3}, Td_{j=4}, Td_{j=5}, Td_{j=6}, Td_{j=7}, Td_{j=8})$.

Posteriormente requerimos otra rutina que repita este proceso para m cadenas aleatorias las cuales son elegidas aleatoriamente como se realizó en el programa de heteropolímeros aleatorios. La cantidad de cadenas que se calcularán dependen de la capacidad de computo que se tenga. En este caso, la máxima cantidad de procesos realizados en este trabajo fue para la cantidad de $J = 2000$ cadenas aleatorias.

La figura 3.10 muestra el comportamiento del promedio de las franjas de temperaturas de desnaturalización $\Delta T_{max-min}$ para el caso de la elección aleatoria con probabilidad uniforme de elección de cada aminoácido, en las cuales es claro observar que entre mas cadenas aleatorias se promedien, estas variaciones tenderán a ser cero. Este punto es de suma importancia

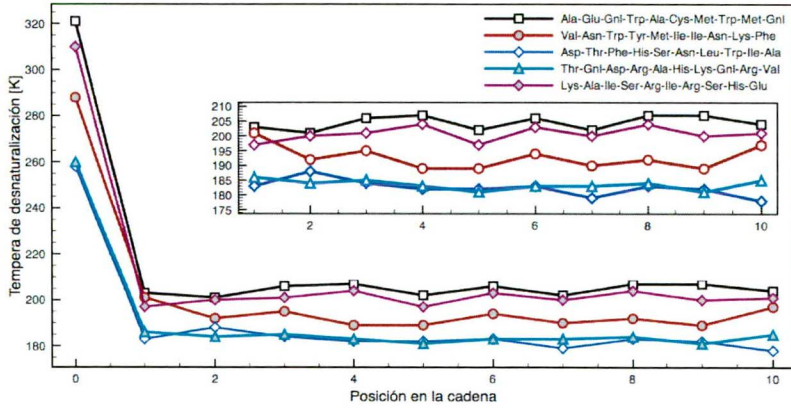


Figura 3.7: En esta imagen se muestran a cinco cadenas heteropoliméricas las cuales han sido elegidas aleatoriamente respecto a su secuencia. Las T_d^* mutadas son mas bajas respecto a la cadena original sin mutar y es claro ver que la diferencia entre T_d_j respecto a su posición disminuye.

J	30	50	80	100	150	200	500	1000	2000
$\Delta T_{max-min}$	3.13	3.4	2.3	1.79	1.80	1.12	0.71	0.611	0.2275

Tabla 3.9: Valores de las temperaturas de desnaturalización $Td_{j,\mu}$ promedio para $\mu=30, 50, 80, 100, 150, 200, 500, 1000$ y 2000 cadenas. $K = 8$. $\Delta = -50000J/mol$, $f = 1$, $\epsilon = 25000J/mol$ y $U = 25000J/mol$.

debido a que esto indica que en promedio, una mutación colocada en alguna posición en la cadena, no alterará significativamente la $Td_{j,i}$, es decir que en promedio si colocamos una mutación en la posición $j = 2$, por ejemplo, la temperatura de desnaturalización en cualquier otra posición será la misma en una sola cadena aleatoria. Este comportamiento estadístico también sugiere que existe un valor de la temperatura de transición de cualquier cadena con distribución uniforme y con K residuos específica, teniendo en general una temperatura $Td_{j,i}(K)$ para cada cadena con distintas cantidades de aminoácidos en su secuencia, lo cual se puede decir que existe una temperatura Td_j promedio para cualquier secuencia con K residuos, debido al comportamiento colectivo de todas las posibles temperaturas de desnaturalización, en las cuales podemos realizar promedios debido a la gran cantidad de Td_j que se obtienen de las 20^K cadenas aleatorias posibles.

Esta diferencia de temperaturas de transición entre la cadena aleatoria

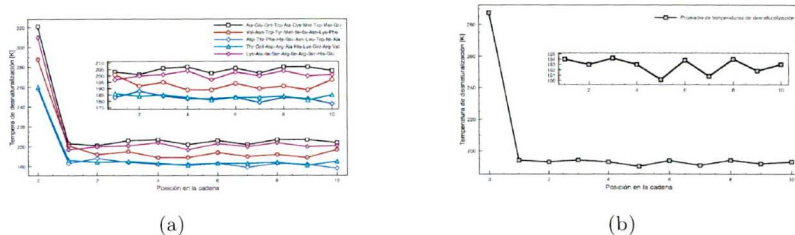


Figura 3.8: Estas dos gráficas muestran las temperaturas de desnaturalización de cinco cadenas heteropoliméricas aleatorias. La diferencia entre ellas es que la imagen (b) es un promedio de la imagen (a), haciendo los promedios con respecto a la posición de la mutación con prolina.

y sus posibles mutaciones dependen directamente del aminoácido con el que se realice la mutación debido a que la estabilidad depende del valor de la propensidad helicoidal del residuo que participa en el cambio en la secuencia, siendo la prolina un rompedor de hélices α ; en cambio, si la mutación se realiza con una alanina, cuya propensidad helicoidal es mas grande por tres ordenes de magnitud que la prolina, la cadena será mas estable estructuralmente, provocando que la transición de fase en alguna cadena se realice en un valor de temperatura mas alto.

1. Las cadenas elegidas aleatoriamente con probabilidad uniforme, fueron mutadas con el aminoácido prolina, disminuyendo su temperatura de desnaturalización T_d .
2. Encontramos que al promediar las temperaturas T_d de un conjunto de 50, 100, 200, 500, 1000 y 2000 cadenas mutadas elegidas aleatoriamente con la misma probabilidad, se tiene una temperatura de desnaturalización, cuyo valor no difiere en demasía con las T_d de cada secuencia.

3.2.3. Caso de estudio III: Péptidos opiáceos

A diferencia de la sección anterior donde se estudiaron las propiedades termodinámicas de cadenas aleatorias (random heteropolymer) donde se tiene una gran cantidad de posibles secuencias específicas, en esta sección estudiaremos cadenas heteropoliméricas, cuya función son de gran relevancia para el ser humano, y se les conoce como *péptidos opiáceos* ya que tienen efectos similares al opio el cual es un regulador de dolor, incluso existen estudios donde se emplean para el tratamiento de tumores cancerígenos[57]. Los péptidos en los cuales centraremos nuestro estudio

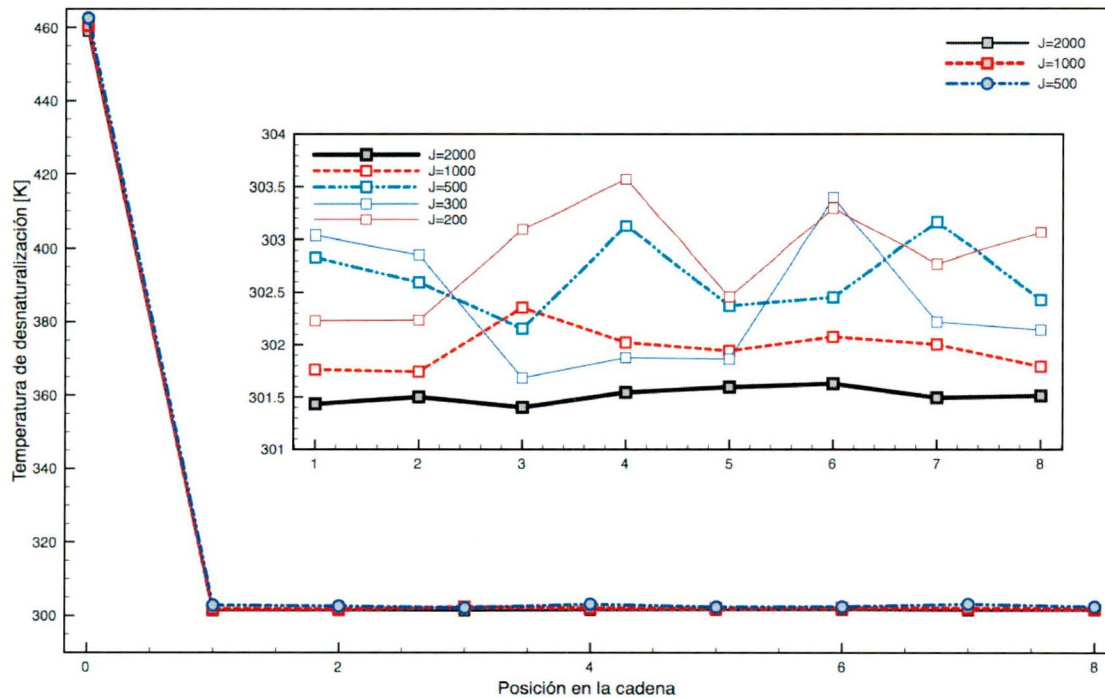


Figura 3.9: Esta figura muestra el comportamiento promedio de j cadenas aleatorias mutadas. Se observa en la gráfica que las temperaturas de desnaturalización o de transición, en promedio tienden a un valor de $Td_j \sim 300K$. En la subfigura se muestra en detalle los valores de las temperaturas de desnaturalización de todas las mutaciones en cada lugar de la cadena heteropolimérica en promedio, observando que existen fluctuaciones entre ellas.

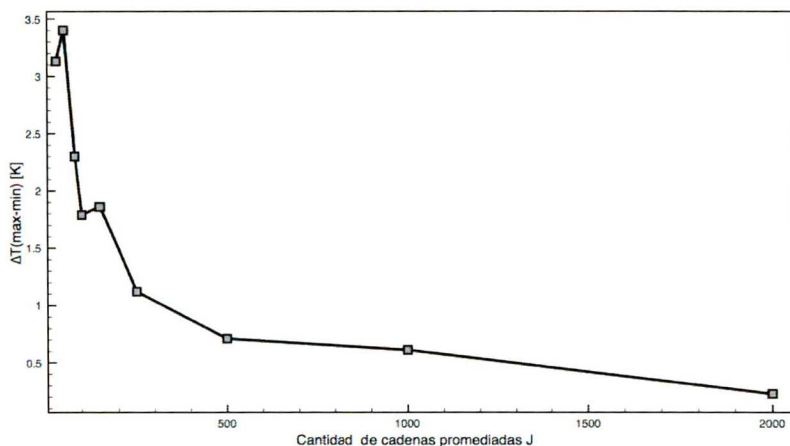


Figura 3.10: Esta gráfica muestra la tendencia de las temperaturas de transición $Td_{j,i}$ de las cadenas mutadas con $K = 8$ a tener un valor promedio ya que la franja de las diferencia de temperatura tienden a ser cero. El eje horizontal indica la cantidad J de cadenas que fueron calculadas en el programa hecho en *Mathematica*. El eje vertical indica la diferencia de temperaturas ΔT entre el mayor valor de temperatura de una cadena mutada en algún residuo con el menor valor de la misma el cualquier otro residuo. $\Delta = -2.2 \times 10^4$, $f = 1$, $\epsilon = 1.1 \times 10^4$, $U = 2.2 \times 10^3$

son Met y Leu Encefalina y otros que son derivadas de estos mencionados.

3.2.4. Antecedentes e importancia

Una clase de péptidos, cuya importancia es relevante para el ser humano son los llamados *péptidos opiáceos* (po), los cuales tienen diversas funciones en los organismos vivos; la regulación del crecimiento de tumores ó la función de neuromoduladores y neurotransmisores[50] son algunas de las muchas funciones que tienen estos péptidos.

A estos péptidos se les llama opiáceos endógenos. El término tiene origen en el opio la cual es una mezcla de sustancias que se extrae de una planta llamada *Papaver somniferum*. Contiene sustancias narcóticas y analgésicas como la morfina, entre otros alcaloides⁵. El descubrimiento de los opiáceos

⁵Se llaman alcaloides a aquellos metabolitos secundarios de las plantas sintetizados, generalmente, a partir de aminoácidos, que tienen en común su hidrosolubilidad a pH ácido y su solubilidad en solventes orgánicos a pH alcalino. La mayoría de los alcaloides

Péptido opiáceo	Secuencia de aminoácidos
Met-enkefalina	Tyr-Gly-Gly-Phe-Met
Leu-enkefalina	Tyr-Gly-Gly-Phe-Leu
[Des- <i>Met</i> ⁵]-enkefalina	Tyr-Gly-Gly-Phe
[Des- <i>Tyr</i> ¹ - <i>Met</i> ⁵]-enkefalina	Gly-Gly-Phe-Met

Tabla 3.10: Tabla con las secuencias de aminoácidos que conforman a algunos péptidos opiáceos de los que se estudiarán en este trabajo. En los últimos dos péptidos, la notación de los superíndices indican la posición en la cadena donde se omitió el residuo que lleva dicho número, por ejemplo, *Met*⁵ indica que no hay un residuo *Met* en el lugar cinco. También notese que la [Des-*Met*⁵]-enkefalina es muy similar a la met-enkefalina, menos el residuo *Met*.

endógenos y el reconocimiento de su importancia en la función cerebral, en la homeostasis y en regulación neuroendocrina ha sido uno de los grandes acontecimientos de la biología moderna. En 1978, Kosterlitz [66] señaló que los efectos antidolorosos de la estimulación eléctrica de ciertas regiones del mesencéfalo o cerebro medio eran revertidos por la administración de un antagonista morfínico ⁶ como la naloxona ⁷, este fenómeno se explica mejor por la suposición de que la estimulación eléctrica libera una sustancia endógena con acciones similares a la morfina, efecto que era revertido por el antagonista naloxona.

Los primeros opiáceos endógenos en ser aislados en laboratorio a partir de extractos celulares fueron llamados *enkefalinas*, por sus descubridores, Hughes y Kosterlitz *et al*[51], en 1975. Entre los péptidos endógenos más importantes están la β -*endorfina*, la *met-enkefalina*, la *leu-Enkefalina* y la *dinorfina*. Estos péptidos difieren entre si por el aminoácido met y leu en uno de los extremos de la secuencia; en la tabla [3.11] se muestran las secuencias de dichos péptidos. En la figura 3.11 se muestra una imagen del neuropéptido met-enkefalina la cual fue obtenida a partir del PDB y generada con el programa VMD.

Es importante conocer las propiedades termodinámicas que tienen estos péptidos opiáceos a partir de los indicadores termodinámicos,

poseen acción fisiológica intensa en los animales aun a bajas dosis con efectos psicoactivos, por lo que son muy usados en medicina para tratar problemas de la mente y calmar el dolor. Ejemplos conocidos son la cocaína, la morfina, la atropina, la colchicina, la quinina, cafeína, la estricnina, la nicotina y la marihuana.

⁶Un antagonista es un tipo de fármaco que al unirse a un receptor celular inhibe una respuesta biológica, pero bloquea o detiene respuestas mediadas por agonistas.

⁷La naloxona es un antagonista de los receptores opiáceos, muy usado en el tratamiento de la intoxicación aguda por opiáceos.

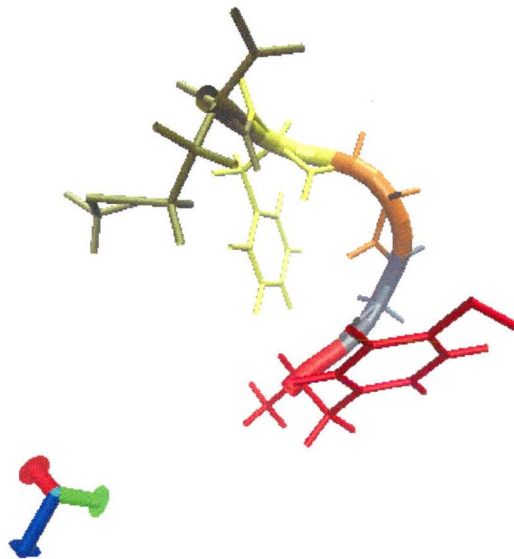


Figura 3.11: Molécula de $[met^5]$ -encefalina. En la imagen se aprecia un tubo (llamado *backbone*) el cual representa a todo el péptido. Por otro lado, también se muestra la representación de la forma explícita de las moléculas de la met-encefalina, donde aparecen los anillos de benceno, por ejemplo. Imagen creada a partir del archivo 1PLW.pdb de la base de datos Protein Data Bank y creada con el visor de archivos VMD 1.9.2.

con los cuales calculamos las temperaturas de desnaturalización de dichos péptidos, donde la estructura cambia y las funciones biológicas cambian. Por ejemplo, pudiera ser el caso en el cual un péptido opiáceo no se junte con algún receptor o cambien su actividad eléctrica neuronal[28].

3.2.5. $[met^5]$ -Encefalina, $[leu^5]$ -Encefalina y sus mutaciones

En particular, en este apartado analizaremos las propiedades termodinámicas en el vacío del neuropéptido $[met^5]$ -encefalina. Dicho péptido tiene gran importancia debido a las funciones que realiza, como la regulación de la transmisión de señales dolorosas las cuales se activan en las neuronas como respuesta a estímulos en el organismo[5].

La secuencia de la $[met^5]$ -encefalina y sus posibles mutaciones con prolina son presentadas en la tabla 3.13. Calculamos las temperaturas de desnaturalización T_d a partir de los indicadores termodinámicos $\langle E \rangle$, C_v y S a partir de la función de partición Z_{het} con $K = 5$ con el objetivo de calcular la temperatura de desnaturalización T_d y a partir de ahí obtener las demás temperaturas de transición de las cadenas mutadas que puedan ocurrir en el péptido $[met^5]$ -encefalina. El calor específico esta presentado en la figura 3.14, donde se muestra el comportamiento del C_v de la cadena $[met^5]$ -encefalina, a su vez, se incluyen los comportamientos de los calores específicos de las cadenas mutadas de la $[met^5]$ -encefalina donde dicha mutación se hizo con el aminoácido prolina. Dicho comportamiento depende de las perturbaciones de la temperatura en la energía promedio; cuando la temperatura aumenta, la cantidad de moléculas de alguna proteína cambia desde el estado nativo hasta el estado desnaturalizado. Este proceso de desnaturalización de proteínas se acompaña de cambios en el calor específico, el valor de C_v en el estado no nativo es mayor siempre que en el estado nativo[77].

Hasta el momento, no se tienen referencias de resultados experimentales de los valores del C_v para los péptidos opiáceos para poder comprar nuestros resultados y verificar si el MHI nos entrega resultados coherentes con la realidad.

Teniendo expresiones del C_v podemos encontrar las T_d de cada secuencia, presentadas en la tabla 3.13. También se muestran los valores de las temperaturas de desnaturalización de las mutaciones de la Met-Encefalina con Prolina y con Glutamina, péptidos que tienen una propensión helicoidal muy distinta entre sí, ($\omega = 0.001$ y $\omega = 0.66$). Se muestra este comportamiento En la figura 3.12 se muestra el comportamiento de Td_j , donde el promedio de los valores de las temperaturas Td_j de las secuen-

Posición en la secuencia	Secuencia de aminoácidos	T_d [K] Pro	T_d [K] Glu
0	Tyr-Gly-Gly-Phe-Met	223	223
1	Pro-Gly-Gly-Phe-Met	148	227
2	Tyr-Pro-Gly-Phe-Met	160	267
3	Tyr-Gly-Pro-Phe-Met	160	267
4	Tyr-Gly-Gly-Pro-Met	149	230
5	Tyr-Gly-Gly-Phe-Pro	147	223

Tabla 3.11: Lista del neopéptido $[Met^5]$ -Encefalina y sus temperaturas de desnaturalización con las mutaciones en cada residuo con los aminoácidos Pro y Glu, las cuales tienen cada una los siguientes valores de los parámetros: $\Delta = -2.2 \times 10^4$, $f = 1$, $\epsilon = 1.1 \times 10^4$, $U = 2.2 \times 10^3$.

cias mutadas difiere con la temperatura T_d de la cadena original por $72.2K$.

En la secuencia de la $[met^5]$ -encefalina tenemos dos aminoácidos que son iguales los cuales son las glicinas (Gly) y están en las posiciones 2 y 3. Es claro observar que las T_d de estas mutaciones tienen el mismo valor de 160 K, lo cual está relacionado con la función de degeneración G_α^K ya que al intercambiar un residuo con otro, la forma de G_α^K no se altera. Esto sugiere que cuando se tienen dos péptidos iguales en una cadena heteropolimérica y estos son sustituidos por una prolina, las temperaturas de desnaturalización serán las mismas. La tabla 3.13 mostrará que efectivamente sucede este hecho.

Así como la Met-Encefalina es un neuropéptido que se produce con el fin de unirse a los receptores opiáceos los cuales realizan la sensación de bienestar entre otras más funciones, la Leu-Encefalina también realiza funciones relacionadas con el aprendizaje y la memoria, el control emocional, la respuesta al estrés y el dolor [58]. También existen enfermedades que están relacionadas con mutaciones genéticas las cuales se han descubierto recientemente en los seres humanos y casualmente vinculado a trastornos neurodegenerativos [59] como lo es la enfermedad de Huntington. Por otro lado, las endorfinas están ligadas a la inhibición del crecimiento de células cancerosas [57] [50]. Existen mecanismos fisiopatológicos que incluyen la epilepsia, adicción, depresión y esquizofrenia.

Las propiedades termodinámicas son obtenidas a partir de la función de partición Z_{het} , la cual depende de los valores de la propensión helicoidal de la secuencia específica. El calor específico C_v es calculado a partir de la derivada de la energía promedio $\langle E \rangle$, y como se mencionó en las secciones pasadas, esto es con el fin de encontrar las temperaturas de desnaturaliza-

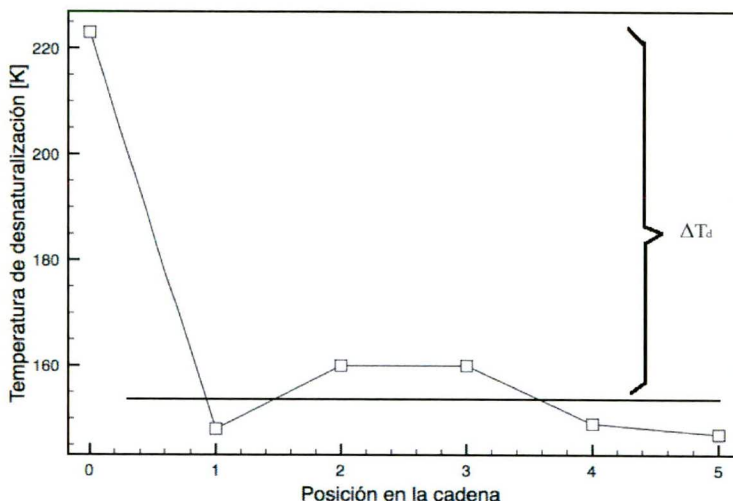


Figura 3.12: En esta figura se muestran las temperaturas T_d de la $[met^5]$ -encefalina y sus mutaciones con el aminoácido prolina en cada posición en la cadena. Existe una Franja de temperaturas ΔT_d la cual es de $72.2K$.

ción T_d de la secuencia original y sus posibles mutaciones las cuales tienen un efecto directamente sobre dichas temperaturas, debido al valor de ω que tiene la Prolina y se ve reflejado en la función de degeneración G_α^K . La franja ΔT_d entre el promedio de las T_d de las mutaciones y la T_d de la cadena original es de $167K$ presentando el mismo comportamiento en las mutaciones 2 y 3 como se mostró en el caso de la met-encefalina.

En la figura 3.16 se muestra la gráfica de las temperaturas de desnaturalización T_d de los péptidos $[met^5]$ -encefalina y $[leu^5]$ -encefalina y de las cadenas mutadas de cada uno de estos dos péptidos. En este caso se realizó dicha mutación con dos péptidos diferentes, los cuales son Glu y Pro cuyos valores de su propensión helicoidal son de $\omega = 0.66$ y $\omega = 0.001$ respectivamente. Es claro observar que el cambio en la temperatura de desnaturalización T_d depende del aminoácido con el que se realice la mutación en algún heteropolímero.

La similitud de los comportamientos por pares radica en el hecho de que las secuencias $[met^5]$ y $[leu^5]$ -encefalina son muy similares entre sí respecto a la secuencia que tienen, difiriendo entre ellas un sólo aminoácido, Met y Leu, cuyos valores de propensiones helicoidales $\omega = 0.63$ y $\omega = 0.96$ respectivamente, tienen solamente una diferencia de 0.33 , por lo que la

Mutación	T_d [K]
0	272
1	165
2	171
3	171
4	166
5	162

Tabla 3.12: Tabla con las temperaturas de desnaturalización T_d del neuropéptido leu-encefalina con las mutaciones hechas en cada residuo con el aminoácido prolina. Es claro observar que la leu-encefalina tienen una $T_d = 272K$, mientras que las temperaturas de desnaturalización T_d de cada cadena mutada son mas bajas. La diferencia de temperaturas entre la cadena leu-encefalina junto con la cadena mutada con temperatura $T_d = 162$ mas baja es de 110K, con los valores $\Delta = -2.2x10^4$, $f = 1$, $\epsilon = 1.1x10^4$, y $U = 2.2x10^3 J/mol$.

Mutación	Cv_{max} [K]	Td_j [K]
0	187.354	303
1	442.192	202
2	343.174	219
3	343.174	219
4	430.099	203
5	458.763	200

Tabla 3.13: En esta tabla se presentan los valores del calor específico Cv_{max} como función de la temperatura, de la cadena $[met^5]$ -encefalina y sus posibles mutaciones con el aminoácido prolina. Los máximos en el calor específico cambian respecto al lugar de la mutación y a su vez dependen del aminoácido que fue sustituido, ya que en el caso de la mutación en el aminoácido Met, coincide con el Cv_{max} mas grande que en los demás casos. $\Delta = -2.2x10^4$, $f = 1$, $\epsilon = 1.1x10^4$, y $U = 2.2x10^3 J/mol$.

Mutación 2 $\nu_1\nu_{P_{ro}}\nu_{Gly}\nu_4\nu_5$	Mutación 3 $\nu_1\nu_{Gly}\nu_{P_{ro}}\nu_4\nu_5$
$G_0 = \nu_1\nu_{P_{ro}}\nu_{Gly}\nu_4\nu_5$	$G_0 = \nu_1\nu_{Gly}\nu_{P_{ro}}\nu_4\nu_5$
$G_1 = \nu_1\nu_{P_{ro}}\nu_{Gly}\nu_4 + \nu_1\nu_{P_{ro}}\nu_{Gly}\nu_5 +$ $\nu_1\nu_{P_{ro}}\nu_4\nu_5 +$ $\nu_1\nu_{Gly}\nu_4\nu_5 +$ $\nu_{P_{ro}}\nu_{Gly}\nu_4\nu_5$	$G_1 = \nu_1\nu_{Gly}\nu_{P_{ro}}\nu_4 + \nu_1\nu_{Gly}\nu_{P_{ro}}\nu_5 +$ $\nu_1\nu_{Gly}\nu_4\nu_5 +$ $\nu_1\nu_{P_{ro}}\nu_4\nu_5 +$ $\nu_{Gly}\nu_{P_{ro}}\nu_4\nu_5$
$G_2 = \nu_1\nu_{P_{ro}}\nu_{Gly} + \nu_1\nu_{P_{ro}}\nu_4 +$ $\nu_1\nu_{P_{ro}}\nu_5 + \nu_{P_{ro}}\nu_{Gly}\nu_4 +$ $\nu_{P_{ro}}\nu_{Gly}\nu_5 + \nu_{Gly}\nu_4\nu_5 +$ $\nu_1\nu_{Gly}\nu_4 + \nu_1\nu_{Gly}\nu_5 +$ $\nu_1\nu_4\nu_5 + \nu_{P_{ro}}\nu_4\nu_5$	$G_2 = \nu_1\nu_{Gly}\nu_{P_{ro}} + \nu_1\nu_{Gly}\nu_4 +$ $\nu_1\nu_{Gly}\nu_5 + \nu_{Gly}\nu_{P_{ro}}\nu_4 +$ $\nu_{Gly}\nu_{P_{ro}}\nu_5 + \nu_{P_{ro}}\nu_4\nu_5 +$ $\nu_1\nu_{P_{ro}}\nu_4 + \nu_1\nu_{P_{ro}}\nu_5 +$ $\nu_1\nu_4\nu_5 + \nu_{Gly}\nu_4\nu_5$
$G_3 = \nu_1\nu_{P_{ro}} + \nu_1\nu_{Gly} +$ $\nu_1\nu_4 + \nu_1\nu_5 +$ $\nu_{P_{ro}}\nu_{Gly} + \nu_{P_{ro}}\nu_4 +$ $\nu_{P_{ro}}\nu_5 + \nu_{Gly}\nu_4 +$ $\nu_{Gly}\nu_5 + \nu_4\nu_5$	$G_3 = \nu_1\nu_{Gly} + \nu_1\nu_{P_{ro}} +$ $\nu_1\nu_4 + \nu_1\nu_5 +$ $\nu_{Gly}\nu_{P_{ro}} + \nu_{Gly}\nu_4 +$ $\nu_{Gly}\nu_5 + \nu_{P_{ro}}\nu_4 +$ $\nu_{P_{ro}}\nu_5 + \nu_4\nu_5$
$G_4 = \nu_1 + \nu_{P_{ro}} + \nu_{Gly} + \nu_4\nu_5$	$G_4 = \nu_1 + \nu_{Gly} + \nu_{P_{ro}} + \nu_4\nu_5$
$G_5 = 1$	$G_5 = 1$

(a)
(b)

Figura 3.13: En la tabla (a) se muestra el desarrollo de la función de degeneración G_α para una cadena con $K = 5$ aminoácidos, donde el segundo y tercer residuo son prolina y glicina respectivamente. En la tabla (b) se muestra de igual forma a la función de degeneración para una cadena polipeptídica con la misma cantidad de aminoácidos pero en el segundo y tercer lugar se encuentran la glicina y la prolina, respectivamente. Notese que los valores de G_α que se encuentran de ambos lados de la tabla son los mismos, para cada valor de α , por lo que podemos concluir, que el cambio de lugar de dos residuos no altera a la función de degeneración ni a la función de partición y por ende, no cambia la temperatura de desnaturalización T_d de la cadena heteropolimérica.

función de degeneración G_α no se ve afectada de manera significativa y por ende tampoco las funciones termodinámicas que se derivan de la misma. A su vez, en el caso de las cadenas $[Des - Met^5]$ -encefalina y la $[Des - Tyr^1 - Met^5]$ -encefalina sucede algo similar en el comportamiento termodinámico respecto a los valores de las propensidades helicoidales de cada uno de los cuatro aminoácidos que conforman las dos secuencias de las degradaciones de la $[met^5]$ y $[leu^5]$ -encefalina. La longitud de los péptidos K también tiene influencia en la función G_α debido a los términos que se agregan en la expresión, cambiando así las propiedades termodinámicas de cada cadena.

Los resultados del calor específico Cv_{max} son mostrados en la tabla 3.13. El comportamiento del Cv se ve alterado al realizar mutaciones en

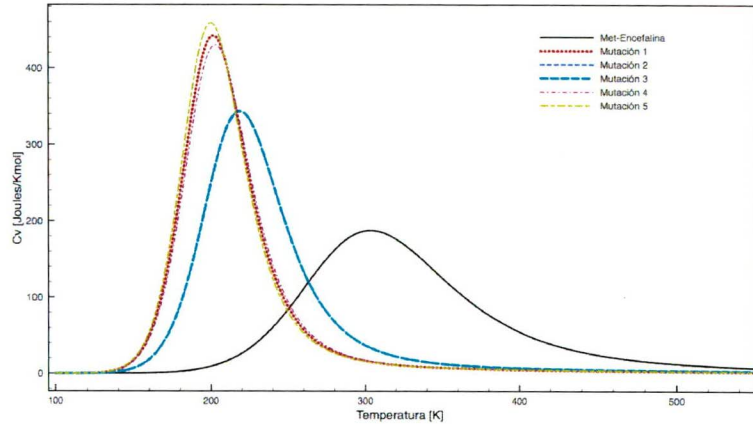


Figura 3.14: El punto máximo de la curva del C_v indica un cambio abrupto en la energía promedio de la estructura polimérica, indicando la transición de fase que presenta el péptido $[met^5]$ -encefalina y sus respectivas mutaciones con el aminoácido prolina.

sus residuos cambiando sus punto máximos, observando que en el cuando se realizó la mutación con prolina en el quinto aminoácido Met, el $C_{v_{max}}$ fue el mayor que en los otros casos, siendo Met el residuo con mayor propensidad helicoidal que los demás. A su vez, la transición de fase en dicho péptido se realiza con una temperatura de $200K$, siendo la menor respecto a las otras temperaturas de desnaturalización T_{d_j} de las mutaciones

3.2.6. Degradaciones de la $[met^5]$ -encefalina y $[leu^5]$ -encefalina

En la sección anterior se calcularon las T_{d_j} de los péptidos opiáceos con el objetivo de caracterizar la transición de fase que ocurre en la cadena configuracionalmente. Ahora, en la referencia [50] se menciona que existen comportamientos análogos entre los péptidos $[met^5]$ -encefalina y $[leu^5]$ -encefalina con respecto a la función que tienen como factores de crecimiento. Existen cadenas polipéptidas, las cuales son degradaciones de otras cadenas ya definidas, como es el caso de $[Des - Met^5]$ -encefalina y la $[Des - Tyr^1 - Met^5]$ las cuales son cadenas que corresponden a las secuencias *Tyr Gly Gly Phe* y *Gly Gly Phe Met* respectivamente. Es claro notar que estas secuencias con $K = 4$ residuos surgen a partir de remover un aminoácido de la $[met^5]$ y $[leu^5]$ encefalina. Ahora, cuando estas secuencias son degradadas en un aminoácido de su secuencia, la actividad biológica que originalmente tenían se pierden, realizando otras funciones biológicas como neoplasias o crecimiento de tejido anormal[50]. Utilizando

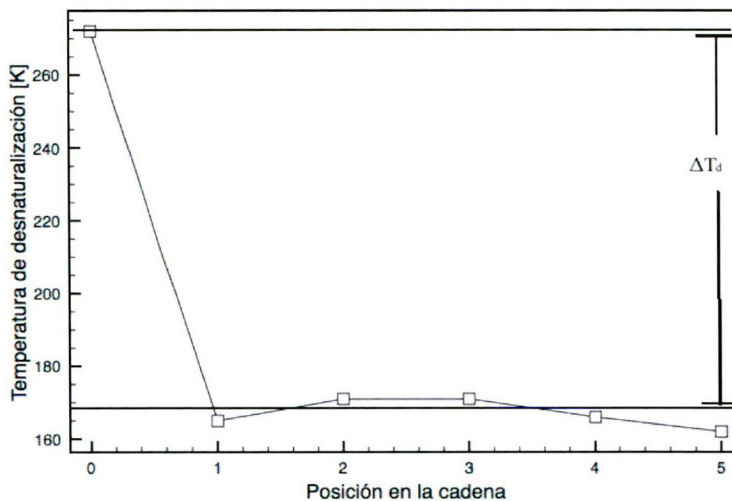


Figura 3.15: Temperaturas de desnaturalización T_d del neuropéptido $[leu^5]$ -encefalina y las posibles secuencias mutadas con prolina. Es claro observar que la $[leu^5]$ -encefalina tienen una $T_d = 272K$, mientras que las mutaciones tienen sus T_d más bajas, con una diferencia entre la más baja de $110K$, con los valores $\Delta = -2.2 \times 10^4$, $f=1$, $\epsilon = 1.1 \times 10^4$, y $U = 2.2 \times 10^3 J/mol$. La franja $\Delta T_d = 105K$.

el formalismo que hemos desarrollado, podemos calcular sus propiedades termodinámicas a partir de la función de partición para cadenas heteropoliméricas Z_{het} para obtener información del comportamiento de las cadenas en relación a la transición de fase configuracional del mismo modo que se realizó en capítulos anteriores. En la figura 3.17 se muestra el comportamiento de la energía promedio de las cadenas $[Des - Met^5]$ -encefalina y $[Des - Tyr^1 - Met^5]$. El comportamiento de las curvas de las Met y Leu-encefalinas es muy similar termodinámicamente hablando, ya que los cambios de la energía promedio $\langle E \rangle$ unicamente difieren entre si en un 2%.

Por otro lado, el comportamiento de las dos cadenas $[Des - Met^5]$ -encefalina y $[Des - Tyr^1 - Met^5]$ también son muy similares en sus variables termodinámicas, por ejemplo, en el caso del cambio en la entropía ΔS varían entre el 2% y el 3% respecto a las otras secuencias $[met^5]$ y $[leu^5]$ encefalina. Este hecho es debido a que las secuencias solo difieren entre si por un solo aminoácido en su secuencia, donde Met y Tyr son los residuos que cambian entre si, por ende, al calcular su función de partición Z_{het} , esta no se alterará cuantitativamente y por consecuencia las variables termodinámicas tendrán una formas similar, reflejándose en las temperaturas de desnaturalización de las dos cadenas, ya que dichos valores solamente difieren entre 7 grados Kelvin.

En la grafica 3.18 se muestran los calores específicos de $[met^5]$ y $[leu^5]$ -encefalina, $[Des - Met^5]$ y $[Des - Tyr^1 - Met^5]$ - encefalina. Observamos a partir de los máximos de esta función podemos obtener las temperaturas de desnaturalización Td_j , donde se realiza la transición de fase de un estado nativo a otro no nativo. Estos valores de Td_j pueden ser manipulados mutando alguno de los residuos que conforman a cualquiera de las dos cadenas que estudiamos en esta sección.

Las cadenas $[Des - Met^5]$ y $[Des - Tyr^1 - Met^5]$ - encefalina tienen un comportamiento termodinámico similar debido a que los máximos de sus correspondientes calores específicos Cv también son muy similares debido a que las diferencias de entropía en en caso de $[met^5]$ -encefalina y $[leu^5]$ -encefalina son: $\Delta S = 80.721J/K$ y $\Delta S = 78.741J/K$ valores que son muy similares entre si, salvo una diferencia de $1.979J/K$ por ende un 2.45% de diferencia entre los dos; por otro lado las diferencias de entropía en los casos de $[Des - Met^5]$ -encefalina y la $[Des - Tyr^1 - Met^5]$ -encefalina son de $\Delta S = 72.609J/Kmol$ y $\Delta S = 69.870J/Kmol$ teniendo una diferencia de $2.739J/K$ y un 3.77% de diferencia entre las dos cadenas. $\Delta = -3 \times 10^4 J/mol$, $f = 1$, $\epsilon = 1.5 \times 10^4 J/mol$ y $U = 3 \times 10^3 J/mol$.

LA similitud de los comportamientos por pares radica en el hecho de que las secuencias de $[met^5]$ y $[leu^5]$ encefalina son muy similares entre

Cadena	Td_j [K]	C_v [$Jmol^{-1} K^{-1}$]
Met-Encefalina	303	187.35
Leu-Encefalina	307	179.35
[<i>Des-Met</i> ⁵]-Encefalina	343	154.678
[<i>Des-Tyr</i> ¹ - <i>Met</i> ⁵]-encefalina	350	144.523

Tabla 3.14: Temperaturas de desnaturalización T_d del neuropéptido met-enkefalina y leu-enkefalina con sus respectivas degradaciones. $\Delta = -3 \times 10^4$, $f = 1$, $\epsilon = 1.1 \times 10^4$, $yU = 2.2 \times 10^3 J/mol$.

si, teniendo de diferencia un aminoácido solamente. Los valores de las propensidades helicoidales de [*met*⁵] y [*leu*⁵] encefalina son 0.63 y 0.96 respectivamente. Al sustituir en la función de degeneración G_α^K en cada caso, la función de partición no se vera afectada cuantitativamente. Así, las funciones termodinámicas serán muy diferentes entre si.

A continuación se realizará una comparación entre los resultados obtenidos del C_v con el MHI y resultados experimentales del C_p con el fin de conocer la precisión del MHI.

3.2.7. Caso de estudio IV: comparación del C_p experimental con el C_v del MHI para proteínas reales

Los modelos matemáticos que se aplican en el campo de la física tienen relevancia una vez que la información que brindan son compatibles, a partir de criterios específicos, con los resultados empíricos obtenidos a partir de experimentos realizados en algún laboratorio. El modelo de heteropolímero ideal calcula la función de partición de una cadena heteropolimérica en el vacío (sin interacción) para obtener las variables termodinámicas E , C_v y S , con el objetivo de conocer el valor de la temperatura en la cual dicha cadena tenga una transición de fase.

Es conveniente comparar los valores del calor específico C_v de los péptidos opioides obtenidos del MHI con experimentos que se hayan realizado en el laboratorio. En la referencia [83] se presentan las gráficas con los valores del calor específico C_p para los péptidos HEW-Lysozima y RNasa-A, cuyos valores se muestran en la tabla 3.15, los cuales se encuentran en una solución acuosa. Hay que recordar que en nuestros cálculos de las propiedades termodinámicas se considera que los péptidos están en el vacío, o en fase gaseosa, lo que implica que no tienen interacción alguna con algún solvente, por ende, en el calor específico no se consideran

contribuciones de hidratación.

Hay varios factores que pueden contribuir a la medida del calor específico, los cuales están relacionados con las propiedades estructurales globales de las proteínas en los estados nativos y no nativos y cómo estos estados interactúan con el solvente[37]. En nuestro estudio consideramos que los péptidos se encuentran en el vacío (fase gaseosa), y no consideramos la interacción con las moléculas del solvente en donde esté inmerso el péptido.

La comparación que realizamos en la tabla 3.15 hace referencia al calor específico a presión constante C_p el cual esta relacionado con el C_v de la siguiente manera:

$$C_p = C_v + \frac{TV\vartheta_p^2}{\kappa_T} \quad (3.16)$$

donde ϑ_p es el coeficiente de expansión térmica y κ_T es la compresibilidad isotérmica. En general, dicha ecuación cumple que $C_p \geq C_v$. El término $\frac{TV\vartheta_p^2}{\kappa_T}$ es obtenido de las referencias [89] y [88] donde se encuentran los valores de ϑ_p y de κ_T , V respectivamente. Los resultados se presentan en la tabla 3.15.

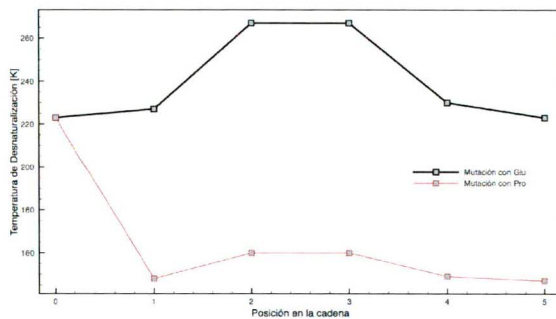
En la gráfica 3.19 se presentan los valores del calor específico C_p de la proteína RNasa-A obtenido de manera experimental [83]. El máximo valor de C_p corresponde a $7.65 \times 10^4 J/Kmol$ a una temperatura de 326 K. Para poder comparar los datos analíticos y los datos experimentales es necesario ajustar los valores de Δ y f del MHI de tal forma que los máximos del C_v coincidan con los de la gráfica 3.19. La comparación de los resultados obtenidos con el MHI y con los resultados experimentales, como se muestra en la tabla 3.15 los cuales difieren, en el caso de la HEW-Lysozima un 31.67% y en el caso de la RNasa A un 26.73%. Podemos decir que el modelo de heteropolímero ideal nos brinda información termodinámica de cadenas polipeptídicas en los ordenes de magnitud de datos experimentales.

El en modelo de heteropolímero ideal, es necesario implementar en la expresión de la energía los términos de interacciones a los que esta sujeta la cadena heteropolimérica del tipo electrostáticas, Van der Waals entre otras.

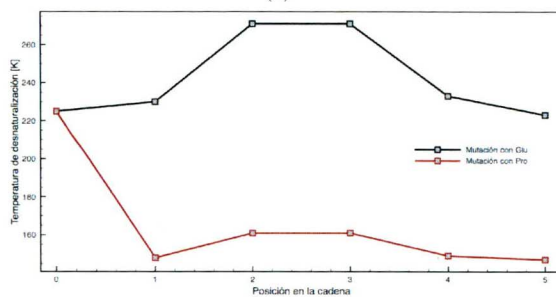
En la sección siguiente se presentan las conclusiones de todo el trabajo realizado y escrito en estas hojas, donde se incluyen las posibles vertientes que se pueden seguir a partir del modelo de heteropolímero ideal. Tal y como se menciona en el párrafo anterior, una perspectiva importante incluye en el MHI términos de interacción con los residuos adyacentes en la misma secuencia.

Proteína	HEW-Lysozyme (pH=2.5)	Ribonucleasa A
Archivo pdb	1DPX.pdb	4OOH.pdb
$C_p [JK^{-1}mol^{-1}]$ (Exp)	8.7×10^4	7.65×10^4
$C_v [JK^{-1}mol^{-1}]$ (MHI)	5.83217×10^4	5.36496×10^4
Temperatura Experimental[K]	337	326
Temperatura MHI [K]	X	Y
f	1×10^2	1×10^2
Δ	-3×10^5	-2.8×10^5
ϵ	297030	277228
U	23.02	22.35
$\frac{TV\vartheta^2}{\kappa}$	1125.4	2395.01
Diferencia entre $C_{p,exp}$ y $C_{p,MHI}$ [%]	31.67	26.73
Diferencia entre $T_{d,exp}$ y $T_{d,MHI}$ [%]	31.67	26.73

Tabla 3.15: Tabla con los valores del calor específico C_p medidos experimentalmente y obtenidos de la referencia [83]. y el C_v calculado con el MHI. *[78]. Los valores utilizados para la compresibilidad isotérmica fueron obtenidos de la referencia [89]. Para los coeficientes de expansión térmica ϑ y el volumen V recurrimos a la referencia [89]. El porcentaje de diferencia entre los valores experimentales y los valores obtenidos analíticamente a partir del MHI de las dos proteínas HEW-Lysozyme y RNasa-A fueron 31.67% y 26.73% respectivamente. $T_{d,1}$ corresponde a la temperatura de desnaturalización calculada experimentalmente y $T_{d,2}$ la temperatura de desnaturalización calculada con el MHI.



(a)



(b)

Figura 3.16: En la gráfica (a) se muestran los valores de las temperaturas de desnaturalización T_d para $[met^5]$ -encefalina y de sus correspondientes secuencias mutadas con los residuos Glu y Pro. Las diferencias de temperaturas T_d entre la $[met^5]$ -encefalina y las secuencias mutadas son menores cuando se realiza la mutación con el aminoácido Glu, a diferencia de que se hiciera con el residuo Pro. La gráfica (b) muestra de igual forma las temperaturas de desnaturalización de la $[leu^5]$ -encefalina, teniendo el mismo tipo de comportamiento respecto al tipo de aminoácido con el que se realice la mutación en la secuencia. $\Delta = -2.2 \times 10^4$, $f=1$, $\epsilon = 1.1 \times 10^4$, y $U = 2.2 \times 10^3 \text{ J/mol}$.

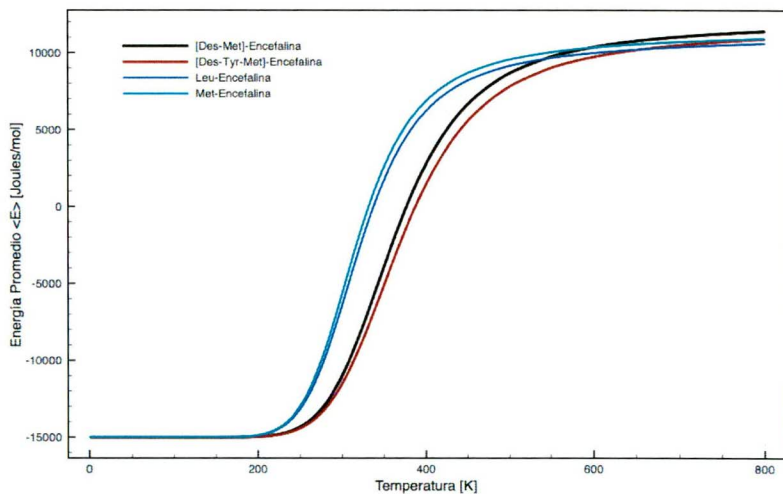


Figura 3.17: Gráfica de la energía promedio $\langle E \rangle$ de los péptidos opiáceos $[met^5]$ -encefalina, $[leu^5]$ -encefalina y sus degradaciones. Los dos primeros al tener propiedades similares en relación a su actividad biológica, también tienen una energía promedio similar. Sus degradaciones a su vez tienen un comportamiento similar en su energía promedio. $\Delta = -3 \times 10^4 J/mol$, $f = 1$, $\epsilon = 1.5 \times 10^4 J/mol$ y $U = 3 \times 10^3 J/mol$.

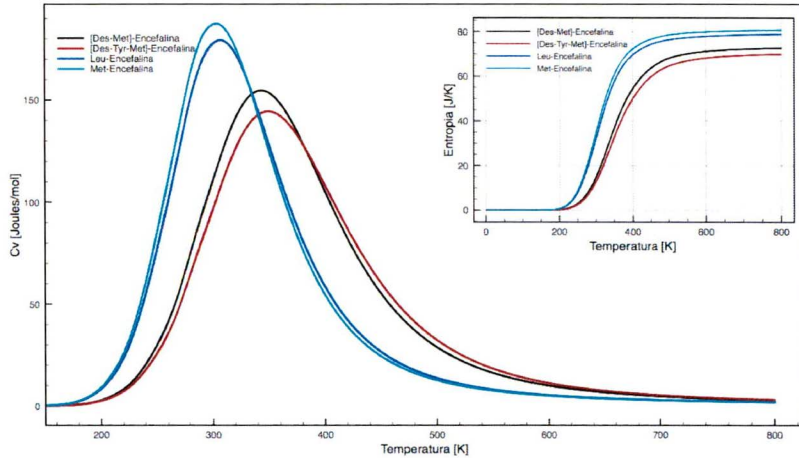


Figura 3.18: Gráfica del calor específico C_v y la entropía configuracional S de los péptidos opiáceos met-enkefalina, leu-enkefalina y sus degradaciones. Sus degradaciones a su vez tienen un comportamiento similar respecto a C_v y S . $\Delta = -3 \times 10^4 \text{ J/mol}$, $f = 1$, $\epsilon = 1.5 \times 10^4 \text{ J/mol}$ y $U = 3 \times 10^3 \text{ J/mol}$. El comportamiento en pares destaca en las dos gráficas mostradas, ya que las cadenas $[\text{met}^5]$ y $[\text{leu}^5]$ -encefalinas presentan comportamientos termodinámicos similares debido a que solamente existe una diferencia de un aminoácido en su secuencia, los cuales son precisamente los residuos Met y Leu, donde las propensidades helicoidales correspondientes son 0.63 y 0.96 respectivamente, donde la diferencia entre estos valores es de 0.33, valor que no es muy grande en comparación con la prolina, donde la diferencia es 3 órdenes de magnitud, por lo que en la función de degeneración no existirá un cambio abrupto.

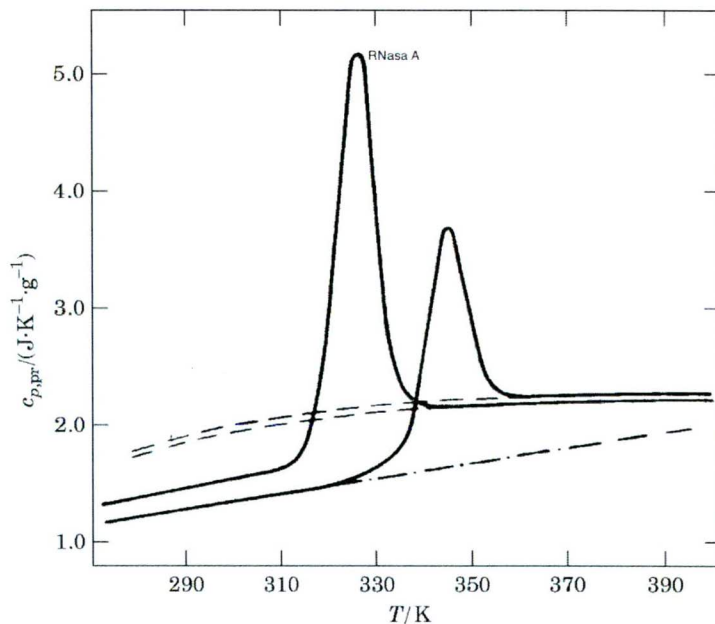


Figura 3.19: Gráfica del calor específico C_p de la ribonucleasa pancreática A (RNasa A). La figura ha sido extraída de la referencia [83]. Las líneas discontinuas corresponden a la función de la capacidad calorífica para una proteína completamente desplegada: RNasa sin disulfuro enlaces cruzados. Las demás, son extrapolaciones lineales de los estados plegados que no utilizaremos en este trabajo.

Capítulo 4

CONCLUSIONES Y PERSPECTIVAS

En este capítulo presentamos un resumen de los resultados obtenidos. A su vez, damos un panorama de los posibles horizontes de estudio a partir de todos los resultados que se obtuvieron.

4.1. Conclusiones

La generalización del modelo de Zwanzig permitió hacer el análisis de las propiedades termodinámicas para cadenas heteropoliméricas, como se sabe, las proteínas que son de interés, no son homopolímeros, donde todos los residuos son iguales entre ellos, si no que en general son heteropolímeros. Tal como lo mencionamos, el modelo de Zwanzig está limitado para cadenas donde cada aminoácido es el mismo que el resto de los residuos, lo cual imposibilita el estudio de cadenas reales. Esta generalización elimina la restricción y amplía el estudio para cadenas reales basándonos en la relación que existe entre la cantidad de conformaciones no nativa y la propensión helicoidal de cada aminoácido.

Para realizar la generalización se consideró la misma expresión para el espectro de energía E_α que en el modelo de homopolímero. Por otro lado, la función de degeneración G_α^K se modificó debido a que la cantidad de microestados en el sistema es diferente. Como se mencionó en el capítulo 3, el encontrar una expresión analítica para G_α^K es complejo debido al conteo de los microestados, por lo que se utilizó un script escrito en *Mathematica* donde se obtuvieron de manera numérica los microestados que puede tener una cadena con K aminoácidos, para cada valor de α , limitando dicho cálculo hasta valores de $K = 20$ residuos.

La estructura del script realizado hace que limitemos nuestro estudio hasta $K=20$ residuos. Posteriormente se comprobó que los resultados obtenidos del script para la G_{α}^K fueran correctos a partir del modelo de Zwanzig solamente considerando que los valores de ν_j de un heteropolímero sean todos iguales, es decir, $\nu_j = \nu_i$, para cualquier valor de i y j menores o iguales a K , concluyendo que podemos obtener todos los microestados para los valores de E_{α} requeridos, a partir del script, de cadenas heteropoliméricas.

Debido a que los recursos de memoria, donde se utilizaron 8 GB, en las computadoras que se utilizaron para realizar los cálculos de la degeneración no fueron suficientes, únicamente analizamos cadenas hasta $K = 20$ aminoácidos como máximo. Es importante extender este estudio para cadenas mayores a 20 residuos debido a que existen cadenas heteropoliméricas las cuales poseen una actividad biológica fundamental dentro de organismos vivos, como es el caso de la hemoglobina con $K = 574$ residuos, la cual es una proteína que se encarga del transporte de oxígeno por la sangre, abasteciendo a los tejidos de dicho elemento. Existen padecimientos que involucran a esta proteína como son las hemoglobinopatías, las cuales consisten en una estructura anormal de una de las cadenas que la conforman, ya sea por un plegamiento incorrecto o por alguna mutación que se presente[62].

La cantidad de cadenas heteropoliméricas con diferente secuencia que se pueden formar en una cadena de K residuos es de 20^K , número que es sumamente grande conforme K se incrementa. El intentar estudiar todas las posibles combinaciones sería casi imposible debido al tiempo que se requeriría para realizar dicho cálculo. Por ejemplo, si analizamos a las cadenas con $K = 8$ residuos y el tiempo que tarda en calcular las temperaturas de desnaturalización es de 1 segundo, entonces requeriríamos un tiempo de 20^8 segundos, equivalente a 811 años aproximadamente, lo cual imposibilita estudiar todo ese conjunto.

El comportamiento termodinámico individual de cadenas heteropoliméricas aleatorias y sus posibles mutaciones es importante para el estudio específico de alguna secuencia de interés. Si deseamos conocer las propiedades termodinámicas de todas las cadenas con una cantidad K de residuos fija, donde existen 20^K posibles secuencias, es una tarea exhaustiva, por lo que optamos por tomar muestras aleatorias de cadenas para conseguir que sus propiedades sean extrapolables a todo dicho conjunto, lo cual nos llevó a obtener los indicadores termodinámicos a partir de un script hecho también en *Mathematica*. Como se menciona en la sección de heteropolímero aleatorios, fue un script que realizara el cálculo, en primera instancia de una sola cadena elegida aleatoriamente para así obtener su temperatura de desnaturalización $T_{d,j,i}$, donde el

subíndice j indica la cantidad de cadenas que se analizaran e i indica el lugar en la secuencia de la cadena donde se realiza la mutación con algún residuo. Posteriormente elegirá otra secuencia aleatoria volviendo a calcular sus propiedades termodinámicas y así sucesivamente. El tamaño de las muestras fueron de 5, 100, 500, 1000 y 2000 cadenas aleatorias con $K = 8$, y a su vez cada una de ellas fue mutada en cada uno de sus residuos, a fin de mostrar el comportamiento global de todas las 20^K secuencias.

A partir de aquí podemos concluir que existe una temperatura de transición promedio, característica para cada secuencia con K residuos distribuidos uniformemente, donde dicho valor no depende de los aminoácidos que la componen. A su vez, el mismo programa calculó las mutaciones de todas las cadenas aleatorias en diferentes posiciones tomando como base al aminoácido prolina. Este tiene una diferencia notoria en los valores de las temperaturas de desnaturalización T_d debido al valor de su propensidad helicoidal ω el cual es tres ordenes de magnitud menor que los demás aminoácidos. Posteriormente obteniendo sus respectivas T_d , teniendo $K + 1$ temperaturas de desnaturalización por secuencia aleatoria, y para cada una de ellas también se realizó un promedio de las mismas. Se observó que no existía diferencia significativa de dichas temperaturas con respecto a la posición en la que se encuentra la mutación en el caso de mutar con prolina, tendiendo a una temperatura de transición $T_{d,j,i}$ promedio característica de cadenas mutadas.

El script para el calculo de la temperatura de desnaturalización de todas las posibles secuencias aleatorias junto a sus K cadenas mutadas con algún aminoácido fue el punto crucial en la sección de cadenas aleatorias. Para cada una de ellas es necesario el calculo de la degeneracidad G_α^K , de la función de partición y de sus respectivas funciones termodinámicas. Posteriormente se busca el máximo del calor específico C_v y a partir de obtener su valor máximo, podemos tener la temperatura de desnaturalización. Estos procesos requieren de una gran capacidad de computo y un tiempo grande proporcional a K . Este tiempo de calculo puede disminuir reestructurando el script a una forma mas eficiente.

En referencia a los péptidos opiáceos podemos mencionar la importancia biológica que presentan en los organismos vivos, en los cuales se realizan una serie de funciones donde destaca la regulación de dolor, entre varias otras. Se calcularon las temperaturas de desnaturalización de $[met^5]$ y $[leu^5]$ -encefalinas y de las $[Des-Met^5]$ y $[Des-Tyr^1-Met^5]$ -encefalinas en el vacío, donde estas dos ultimas son de cierta forma una degradación de las primeras mencionadas. Para ello se calcularon las funciones termodinámicas como la energía promedio $\langle E \rangle$, el calor específico C_v y la entropía configuracional S de los neuropéptidos con el fin de encontrar la temperatura de

desnaturalización donde ocurre una transición de fase en la cadena. Dichas temperaturas fueron comparadas hallándose que podemos agruparlas en pares entre la $[met^5]y[leu^5]$ -encefalinas y las cadenas degradadas, encontrando que el comportamiento termodinámico de $[met^5]y[leu^5]$ es muy similar debido a que las transiciones de fase de cada una de las secuencias ocurre en valores similares de la temperatura, con una diferencia de $4K$ tal y como se muestra en la tabla 3.14. De la misma forma, el comportamiento de sus respectivas degradaciones $[Des - Met^5] - encefalina$ y $[Des - Leu^5] - encefalina$ resulto ser también muy similar entre ellas con una diferencia entre estas dos cadenas de $7K$ tal y como se muestra en la tabla 3.14. Esta comparación se realizó con el incremento en la entropía configuracional ΔS , las cuales difieren entre las encefalinas un 2.45% y en el caso de sus degradaciones un 3.77% únicamente. Como se menciona en la literatura[50], las propiedades biológicas de estos péptidos son similares entre si. Una de las consideraciones iniciales en estas cadenas es que se encuentran en el vacío, no considerando interacciones con el solvente que las contiene como es el caso del agua la cual es el fluido en el cual generalmente se encuentran los opiáceos. Se esperaría que los resultados, considerando interacciones con el solvente, aumenten, debido a que el sistema se restringe disminuyendo los grados de libertad.

Hasta estos momentos, no es posible encontrar resultados experimentales de los péptidos opiáceos con los cuales se pueda comparar con nuestros resultados del MHI. Esto no significa que no se puedan estudiar otras cadenas donde si se tengan datos del calor específico C_v con los cuales podamos contrastar con datos calculados teóricamente. Por otro lado existen experimentos en los cuales podemos encontrar a la proteína en las condiciones que este modelo tiene inherentes, como es el estar en el vacío[74], donde observan transiciones estructurales de las proteínas.

En general, las proteínas se encuentran en algún solvente como es el agua, por ejemplo. En el modelo de heteropolímero ideal se considero que las cadenas no se encuentran en dicho solvente, es decir, que no hay interacción con otras moléculas. Es posible tener a las proteínas en un estado en el cual no este inmerso en un solvente, a partir de una técnica experimental llamada *electrospray ionization*[63], la cual libera al conjunto de péptidos en la célula y mantenerlos en una fase gaseosa ¹[69]. Existen experimentos donde se estudia el proceso de plegamiento de las proteínas en el vacío, donde aun se debate si la estructura del péptido se mantiene al pasar a una fase gaseosa (gas-phase) [64] [65] [67] [68]. Por esta razón, podemos decir que nuestro modelo tiene sentido utilizarlo en este tipo de sistemas en fase gaseosa, aun-

¹La fase gaseosa de las proteínas se refiere al estado en el cual esta en ausencia de un solvente, por lo general agua.

que el plegamiento puede ser alterado, no necesariamente por el solvente en el que se encuentre, si no por las cadenas laterales de los demás péptidos[70].

4.2. Perspectivas

A partir del estudio de las cadenas heteropoliméricas surgen posibles vertientes cuya importancia e interés vale la pena mencionar en esta sección. Para alguna de estas ideas, se mostraran algunos de los avances y resultados que se tienen hasta el momento con el fin de mostrar un panorama mas amplio de las posibles lineas que puedan abordarse a partir de resultados aquí presentados.

4.2.1. Espaciamiento y acotamiento del espectro de energías E_α

En el modelo de Zwanzig se pueden realizar modificaciones al espectro de energía propuesto por el autor con el fin de obtener las variables termodinámicas para su estudio. En esta sección se mencionarán dos posibles propuestas en E_α que pueden mejorar la aproximación de este modelo. Dicho modelo se centra en la expresión propuesta del espectro de energía E_α , el cual, como se vio en la figura 4.1a, tiene todos sus niveles de energía igualmente espaciados, lo cual no necesariamente tiene que ser de esta forma.

Por ejemplo, en el caso del átomo de hidrógeno, a partir de la ecuación de Schrodinger podemos tener una expresión para el espectro de energía haciendo las consideraciones pertinentes. Teniendo una partícula que solamente se puede mover sobre el eje x en un rango de 0 a L ($0 < x < L$). En las fronteras consideraremos que el potencial $V(x)$ es infinito, es decir, $V(0) = V(L) = \infty$. Teniendo la ecuación de Schrodinger

$$H\psi = E_j\psi \quad (4.1)$$

donde el operador H aplicado a la función de onda ψ es:

$$H = - \left(\frac{h^2}{8\pi^2m} \right) \frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x) \quad (4.2)$$

Como la partícula es libre, quiere decir que no esta sujeta a ninguna fuerza potencial, por lo que $V(x) = 0$, quedando así la ecuación de Schrodinger como:

$$- \left(\frac{h^2}{8\pi^2m} \right) \frac{\partial^2\psi}{\partial x^2} = E_j\psi \quad (4.3)$$

Para resolver la ecuación anterior, se propone una solución de la siguiente forma, tomando $k = \frac{E_j}{\left(-\frac{\hbar^2}{8n^2m}\right)}$:

$$\psi(x) = A \operatorname{sen}(\sqrt{k}x) + B \operatorname{cos}(\sqrt{k}x) \quad (4.4)$$

Realizando el álgebra pertinente, tenemos la expresión para los niveles de energía, que solamente dependen del cuadrado de n :

$$E_n = \frac{n^2 \hbar^2}{8mL^2} \quad (4.5)$$

Es claro que los niveles de energía de una partícula libre cambia como n^2 , lo que quiere decir que el espaciamiento que hay entre cada nivel no son iguales. Este hecho nos lleva a pensar que el espectro de energía para el modelo de Zwanzig puede ser reformulado de tal forma que sus niveles, los cuales dependen linealmente de α , no sean igualmente espaciados por lo que podemos proponer a E_α como una expresión no lineal de la siguiente forma:

$$E_\alpha = (K - \alpha)^2 - \epsilon \delta_{K,\alpha} \quad (4.6)$$

Es necesario argumentar el hecho del porque se tendrían niveles de esta forma ya que en el caso del modelo de Zwanzig, el argumento es debido a que el termino $(K - \alpha)$ indica la cantidad de residuos en un estado no nativo, teniendo cada uno de ellos una energía U . Basándonos en el espectro de energía de una partícula cuántica en la cual los niveles de energía se encuentran espaciados como n^2 , tomamos el termino $(K - \alpha)$ el cual tiene a todos los niveles energéticos y suponemos que tienen un espaciamiento cuadrático, teniendo $(K - \alpha)^2$ como el termino análogo al nivel de energía de la partícula libre.

Por otro lado, como se menciona en la sección anterior, el espectro de energía E_α propuesto por Zwanzig tienen dos cotas, superior e inferior, las cuales coinciden en los valores $\alpha = 0$ y $\alpha = K$. Este hecho contrasta con los resultados calculados en la subsección anterior donde se considera a una partícula libre en una caja sujeta a un potencial $V = 0$ y donde el espectro de energías, en principio puede adoptar valores desde el estado base $E_1 = \frac{\hbar^2}{8mL^2}$ hasta infinito cuando n tiende a infinito.

Debido a este acotamiento de E_α , es posible llegar a temperaturas negativas ya que esto implica que la entropía no es una función siempre creciente de la energía. En la referencia [60] se discuten las circunstancias en las que pueden producirse temperaturas absolutas negativas, y se desarrollan los principios de la termodinámica y la mecánica estadística a temperaturas negativas. Si la entropía de un sistema termodinámico no es una función monótonamente creciente de su energía interna, posee una temperatura negativa siempre que $\left(\frac{\partial S}{\partial U}\right)$ sea negativo. Este tipo de

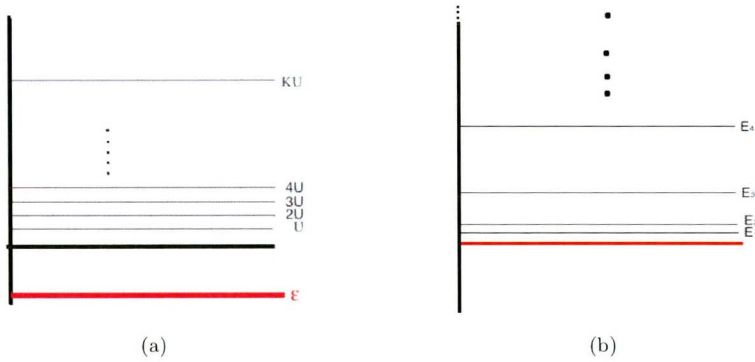


Figura 4.1: Espectros de energía del modelo de Zwanzig (a) y partícula cuántica libre (b). En esta figura se muestran las diferencias entre los dos espectros de energía para el modelo de Zwanzig y para una partícula cuántica la cual no esta sujeta a un potencial de interacción. Notese que en (a) los niveles de energía están igualmente espaciados, mientras que en (b) los niveles están separados como n^2 .

inconvenientes incita a proponer un espectro de energías donde no se tenga una cota superior con las mismas variables que se manejan, las cuales son, α , U , ϵ y K . Dicho espectro, a primera instancia, no tendría que tener el termino $(K - \alpha)$ ya que limita a E_α cuando $\alpha = K$ en su cota superior.

4.2.2. Expresión analítica de la función de degeneracidad G_α para heteropolímeros

Para el calculo de la función de partición en el caso de heteropolímeros fue necesario calcular la función de degeneracidad la cual cuenta el número de microestados para algún valor de α . En el caso de los homopolímeros, g_α^K se calcula de manera analítica con la forma $g_\alpha = \nu^{K-\alpha} \binom{K}{\alpha}$. Esta expresión se obtiene de manera sencilla ya que los valores de la conformación no nativa ν siempre son los mismos para cada residuo en la secuencia. Pero en el caso de heteropolímero, se tiene una distribución de distintos valores de ν_j donde cada residuo en la posición i tiene un valor ν_i el cual esta relacionado con el inverso de la propensidad helicoidal ω . Al intentar escribir una expresión analítica para la degeneracidad G_α nos queda una expresión de la siguiente forma:

$$G_{\alpha, \Upsilon} = \left(\prod_{i=1}^K \nu_i \right) \sum_{\xi_1=1}^{K-\alpha+1} \frac{1}{\nu_{\xi_1}} \sum_{\xi_2=1}^{K-\alpha+2} \frac{1}{\nu_{\xi_2}} \dots \sum_{\xi_\alpha=1}^K \frac{1}{\nu_{\xi_\alpha}} \quad (4.7)$$

El desarrollo de esta ecuación, hasta estos momento de realización

del trabajo presentado, ha sido complicado reducirla a una forma mas sintetizada. Este hecho crea la necesidad de contar los microestados G_{α}^K a partir de rutinas realizadas en un script en el programa *Mathematica* de tal forma que los recursos de memoria utilizados sean mínimos. Los cálculos a los que nos limitamos fueron para cadenas con $K = 20$ residuos. En trabajos recientes realizados en colaboración con el Dr. Damian Hernandez, hemos determinado ya una expresión analítica para G_{α}^K en una representación integral. Esta nueva formula facilita los cálculos para la función de partición Z_{het} , donde se pueden realizar cálculos para cadenas mayores a $K = 20$ residuos.

4.2.3. Interacción a primeros y segundos vecinos en homopolímeros

Hablando en sentido estricto, los átomos que conforman a las moléculas están en constante interacción con sus vecinos, ya sea que esta interacción sea de corto o largo alcance. Estas interacciones pueden ser, por ejemplo, de Van der Waals debido a los enlaces covalentes que conforman al péptido. Pero en la expresión del espectro de energía de Zwanzig, no se considera interacción alguna entre residuos, lo cual limita al modelo. Existen varios modelos en los cuales se consideran interacciones a primeros y segundos vecinos. Incluso existen modelos que consideran a las cadenas como flexibles, abordando el tema como una caminata aleatoria (*random walk*), donde los N pasos que realiza el caminante corresponderían análogamente a los lugares que ocupan los residuos en la secuencia.[61].

El modelo de Ising, llamada así por el físico Ernst Ising, es un modelo matemático de ferromagnetismo en la mecánica estadística el cual considera interacciones a primeros vecinos. El modelo consta de variables discretas que representan momentos dipolares magnéticos de los espines atómicos que pueden estar en uno de dos estados (+1 o -1). Los espines están dispuestos en un gráfico, por lo general, una red, permitiendo que cada giro interactúe con sus vecinos. El modelo permite la identificación de las transiciones de fase, como un modelo simplificado de la realidad.

El modelo de Ising en una red bidimensional [42] es uno de los modelos estadísticos simples para mostrar una transición de fase. En la figura 4.2 se muestra dicha red de espines. El modelo presentado en una dimensión es una cadena lineal de k espines, mostrada en la figura 4.3, donde cada espín se le asigna un valor de s_i en la posición i , donde $s = \pm 1$, dependiendo de la orientación que tenga cada uno y va a depender de la interacción que exista entre sus primeros vecinos y a su vez de fluctuaciones térmicas. El modelo de Ising también puede considerar que el sistema se encuentre bajo

un campo magnético externo B donde los espines interactúan con el con una energía de $H_B = B s_i$; para fines prácticos consideraremos que el campo externo es cero, o lo que es lo mismo, no hay campo externo. Cada espín va a interactuar con sus primeros vecinos, es decir, el espín s_i va a tener interacción con s_{i-1} y s_{i+1} y va a estar representada por el término J , el cual es un parámetro del sistema con unidades de energía.

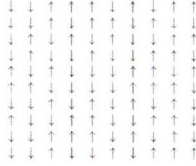


Figura 4.2: Representación gráfica de una cadena lineal de espines para el modelo de Ising en 2D.

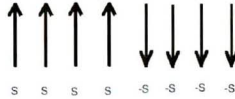


Figura 4.3: Representación gráfica de una cadena lineal de espines para el modelo de Ising en 1D.

El hamiltoniano de este sistema se plantea de la siguiente forma simplemente sumando todas las interacciones que hay entre los espines a primeros vecinos:

$$H = -J_i \sum_{i=1}^{k-1} s_i s_{i+1} \quad (4.8)$$

donde

$$s_i = \begin{cases} 1 \\ -1 \end{cases} \quad (4.9)$$

Ya que tenemos una expresión para la energía de interacción del sistema, calculamos la función de partición para obtener las variables termodinámicas y poder observar si la cadena lineal de espines presenta una transición de fase debido a las fluctuaciones de la temperatura:

$$Z = \sum_{i=1}^k e^{-\beta H} = \sum_{i=1}^k e^{-\beta(-J_i \sum_{i=1}^{k-1} s_i s_{i+1})} \quad (4.10)$$

Desarrollando esta expresión tenemos finalmente que la función de partición, para un caso uniforme en el cual J sea la misma para todas las interacciones ($J_i = J$):

$$Z = 2^k \cosh^{k-1}(\beta J) \quad (4.11)$$

Sin duda, la virtud principal de este modelo, en cierto modo académico, consiste en que, aplicado a una red bidimensional permite un tratamiento exacto en mecánica estadística, y es el único ejemplo de transición de fase (esta transición es de orden a desorden) que puede tratarse de forma matemáticamente exacta y rigurosa, y es el paradigma de las transiciones de fase. Así mismo, tampoco se conocen soluciones exactas del modelo de Ising (bidimensional) cuando el campo externo aplicado es distinto de cero.

Por otro lado, el modelo de Ising puede aplicarse a otros modelos que no sean ferromagnéticos. Con ello queremos decir que podemos reproducir otros sistemas, y la solución del modelo de Ising puede llevarse a ellos. Entre tales sistemas, se encuentra el gas de red (*lattice gas*) donde el modelo consiste en que las partículas están dispuestas en una red y coinciden con los nodos de la misma y es utilizado, por ejemplo, para estudiar la termodinámica de fluidos. En la figura 4.4 se muestra el comportamiento del calor específico de tres cadenas, cada una de ellas tiene 5, 10 y 18 partículas cada uno para el modelo de Ising.

Sin embargo, el modelo de Ising ha encontrado muchas aplicaciones fuera del contexto de la física. En biología molecular ha servido para modelizar macromoléculas y desnaturalización de ADN, y en sociología, para modelizar el aislamiento cultural, entre otros.

Por otro lado el modelo de Zwanzig, a diferencia del modelo de Ising, es un modelo donde la expresión de la energía esta relacionado con la energía que tiene cada residuo estando en una conformación nativa o no nativa individualmente y no se considera la energía de interacción que existe entre cada residuo. Como una perspectiva, proponemos un termino adicional en el modelo de Zwanzig en el cual se considere la energía de interacción entre residuos (Ising).

Esta mezcla entre los dos modelos no puede darse de manera directa debido a la forma en la que esta definido el modelo de Ising en cuestión de las interacciones por pares. Por ejemplo, en el caso de tener 3 residuos en un estado nativo, definiendo a $s_j = 1$ como el estado nativo y $s_j = -1$ como estado no nativo, donde j representa el lugar de ocupación de algún aminoácido en la cadena. Ahora, si los tres aminoácidos se encuentran en una conformación nativa en alguna secuencia tienen interacción s_1, s_2 y s_3

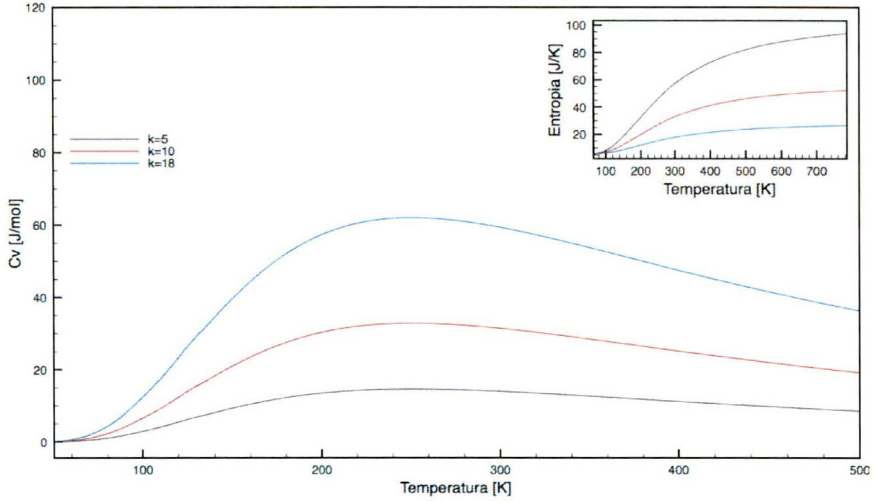


Figura 4.4: En esta gráfica se muestra el comportamiento del calor específico C_v de tres cadenas con $K = 5, 10$ y 18 partículas con espín $s = \pm 1$ para el modelo de Ising. En la subfigura se muestra la entropía del sistema de espines. $J = 1000$

entre ellos, entonces $H = -J_i \sum_{i=1}^{K=3} s_i s_{i+1}$. El desarrollo de H está dado por la siguiente expresión:

$$H = -J_i \sum_{i=1}^{K=3} s_i s_{i+1} \quad (4.12)$$

$$= -J [(s_1 s_2) + (s_2 s_3)] \quad (4.13)$$

$$= -J [((1)(1)) + ((1)(1))] \quad (4.14)$$

$$= -J(1 + 1) \quad (4.15)$$

$$= -2J \quad (4.16)$$

siendo los términos s_1 y s_2 serían iguales a 1, por lo que la suma de estas energías sería $H = -J_i \sum_{i=1}^{K=3} s_i s_{i+1} = -2J$ teniendo una energía en un estado totalmente nativo, pero el inconveniente real se muestra cuando se tiene un estado totalmente desnaturalizado, donde si hacemos la analogía directa tenemos que los valores de las dos interacciones en el caso con $K = 3$ residuos son $s_1 = -1$ y $s_2 = -1$, resultando un valor de la energía que es exactamente el del estado nativo llegando a una inconsistencia. Podría resolverse redefiniendo la interacción de la siguiente manera:

$$\Gamma_{i,i+1}^{(1)} = \begin{cases} 2J & si \begin{cases} i \rightarrow D \\ i+1 \rightarrow D \end{cases} \\ -1J & si \begin{cases} i \rightarrow D \\ i+1 \rightarrow N \end{cases} \\ -1J & si \begin{cases} i \rightarrow N \\ i+1 \rightarrow D \end{cases} \\ -2J & si \begin{cases} i \rightarrow N \\ i+1 \rightarrow N \end{cases} \end{cases}$$

donde N representa al estado nativo y D al estado no nativo.

Este operador $\Gamma_{i,i+1}^{(1)}$ toma valores dependiendo del estado en el que se encuentren los residuos en la cadena heteropolimérica. En el caso de estar los dos residuos en un estado no nativo, su energía de interacción será $2J$. Por otro lado, en el caso de que los dos residuos se encuentren en un estado nativo, su energía de interacción será $-2J$, cuyo valor es congruente con el estado totalmente nativo de toda la cadena, ya que el gasto energético es menor. Un caso intermedio sería el caso en el cual los estados sean contrarios entre si, es decir, un residuo en estado nativo y el otro en estado no nativo. La expresión para el espectro de energías del modelo de Zwanzig con interacción a primeros vecinos quedaría definida por la siguiente expresión:

$$E_\alpha = (k - \alpha)U - c\delta_{k,\alpha} - J \sum_{i=1}^{k-1} (\Gamma_{i,i+1}^{(1)}) \quad (4.17)$$

Teniendo el espectro de energías, podemos calcular las propiedades termodinámicas de las cadenas interactuantes y obtener la temperatura de desnaturalización T_d de la misma. Al calcular la función de partición para una cadena de K aminoácidos considerando, además del espectro de energías propuesto por Zwanzig, la interacción a primeros vecinos que es un termino inspirado en el modelo de Ising, tenemos una expresión no analítica para la entropía y el calor específico debido a la complejidad de la ecuación y los términos que hay en ella por lo que se hicieron los cálculos numéricamente.

Al graficar dicha expresión numérica, tenemos la figura 4.5 del calor específico C_v para dos cadenas polipeptídicas. Los máximos del C_v son menores conforme la cantidad de residuos K en la cadena aumenta, lo que marca una incongruencia en la termodinámica ya que esto indicaría que se necesita menos energía para incrementar la temperatura a la cadena heteropolimérica conforme sus residuos K aumentan. Este comportamiento

inconsistente puede abordarse a través de la consideración de las interacciones a primeros y segundos vecinos, ya que en el modelo de heteropolímero ideal no se consideran este tipo de factores. Escribimos un planteamiento de una expresión en donde si se considera este tipo de interacciones a partir de definir otro operador donde contenga las posibles energías de interacción a segundos vecinos, de forma análoga como se hizo en el caso de interacciones a primeros vecinos.

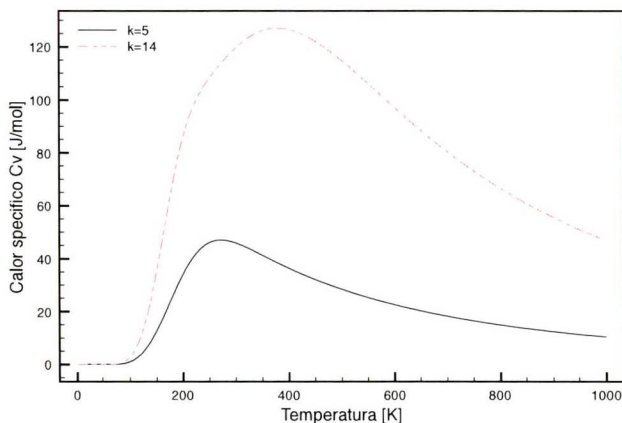


Figura 4.5: Calor específico del modelo Zwanzig con interacción a segundos vecinos. Esta gráfica muestra los comportamientos del calor específico del modelo de Zwanzig con interacción a segundos vecinos. $\Delta = -1.3 \times 10^4 J/mol$, $f = 1$, $\epsilon = 6500 J/mol$, $U = 1300 J/mol$. Se observa que la inversión de curvas se corrige considerando la interacción a segundos vecinos ya que el máximo en el Cv de la cadena con $K = 5$ esta a una temperatura menor que la cadena con $K = 14$ residuos.

El espectro de energías propuesto por Zwanzig considera únicamente energías por residuo en estado nativo y no nativo, excluyendo toda interacción que se pueda tener en el sistema, es decir, como las cadenas polipeptídicas están en un medio con partículas de agua las cuales son polares, existen interacciones de tipo electrostático que influyen en el plegamiento de la secuencia hasta alcanzar su estructura terciaria. El hecho de haber incorporado a los heteropolímeros a través de las propensidades helicoidales que tienen cada residuo nos da un panorama mas amplio del comportamiento de un péptido con diferentes residuos, los cuales pueden ser, como se vio en el capítulo 3, un neuropéptido u otro que presente algún tipo de mutación en su secuencia. Las perspectivas que se tienen a partir de este estudio es

proponer una expresión para la energía E_α la cual considere interacciones de corto y largo alcance, agregando varios términos adicionales que dependan del lugar donde se encuentre cada residuo individualmente.

4.2.4. Simulaciones numéricas

Los modelos analíticos analizados en este trabajo se basan principalmente en la forma funcional del espectro de energías E_α y el cálculo la función de degeneración G_α^K . A partir de esto se calcula directamente la función de partición Z_{het} con el fin de obtener las variables termodinámicas como E , Cv y S . Este modelo considera que los péptidos estudiados se encuentran en el vacío o fase gaseosa, como se le denomina en las referencias [64][65]. En general, los péptidos se encuentran inmersos en un solvente, por ejemplo en agua, donde existen interacciones cadena-solvente como pueden ser de carácter Coulombianas y de Van der Waals. Estos términos de interacción influyen en el espectro de energía. Sin embargo, en el modelo de Zwanzig no se consideran dichos términos.

Las simulaciones de dinámica molecular son técnicas computacionales para la modelación de grandes complejos agregados moleculares donde si se pueden incluir interacciones de distintos tipos, como son la interacción coulombiana y la de Van der Waals, por ejemplo. La dinámica molecular pretende dar una descripción de los sistemas complejos en términos de un modelo atómico realista con el fin de predecir las propiedades termodinámicas macroscópicas del sistema en base al conocimiento que se tiene a nivel atómico.

Un software que es capaz de realizar simulaciones de moléculas propuestas y sus posibles interacciones con otras moléculas es el llamado GROMACS. Dicho programa es capaz de realizar cálculos de dinámica molecular a partir de resolver numéricamente las ecuaciones de Newton para un sistema de N cuerpos interactuantes utilizando técnicas de minimización de energía. Esta herramienta es útil para el estudio de sistemas en donde el número de partículas es suficientemente grande. Basándonos en el tema de los péptidos opiáceos, por ejemplo, en las ecuaciones propuestas se considera al péptido sin interacción alguna[72][73][74][76], hecho que no es necesariamente cierto ya que naturalmente este péptido esta inmerso en un medio acuoso, es decir, a su alrededor existen moléculas de agua que constantemente están interactuando con los péptidos, modificando así las propiedades termodinámicas inherentes a este sistema. En la figura [4.5] se muestra al péptido [*met*⁵]-encefalina rodeado de moléculas de agua. Esta figura es una mejor representación de lo que realmente esta sucediendo a niveles moleculares, por lo que un enfoque será considerar una simulación en dinámica molecular del comportamiento termodinámico de cadenas

heteropoliméricas sometidas a un campo de fuerzas creado por las moléculas de agua[75].

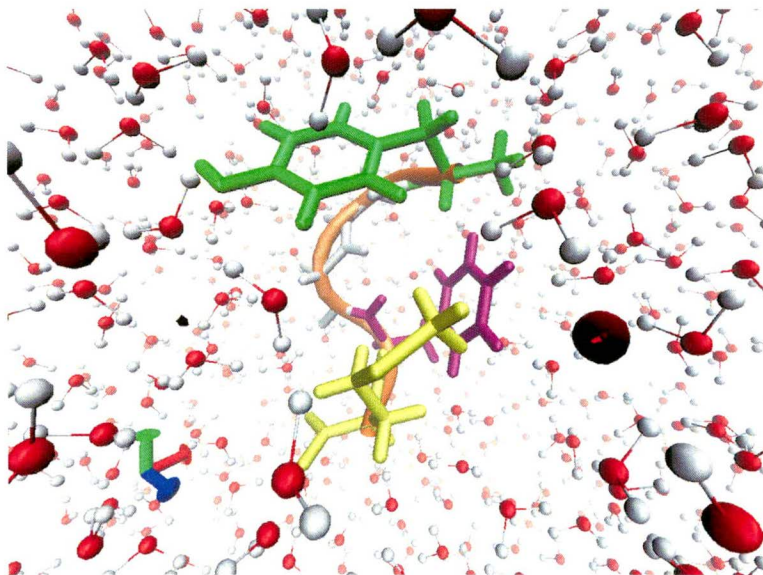


Figura 4.6: Molécula de $[met^5]$ -encefalina inmersa en una caja con 10,000 moléculas de agua. Esta figura fue creada a partir de una simulación hecha en el programa *Gromacs* y exportada a un archivo para su visualización en VMD.

Este trabajo nos ayuda a estudiar el comportamiento termodinámico de cadenas heteropoliméricas a partir de la modificación del modelo de Zwanzig el cual solo es posible calcularlas para cadenas poly-X, donde X representa alguno de los 20 aminoácidos y a su vez, obtener el comportamiento promedio de cadenas con K residuos. A partir de las perspectivas presentadas se pretende abrir nuevas ventanas de investigación donde los modelos sean mas elaborados y puedan compararse con los experimentos realizados en el laboratorio.

Apéndice A

Función de degeneración g_α y función de partición Z

A.1. Función de degeneración g_α del modelo de Zwanzig

Para la función de degeneración g_α , Se menciona el siguiente ejemplo para realizar la deducción de la misma.

Se tiene una cadena con dos aminoácidos, los cuales solo pueden estar en los estados: A, B y C, donde los dos primeros no estados no nativos y el tercero es el estado nativo.

Como se puede apreciar en la tabla anterior existen 4 cadenas en las cuales hay 2 estados incorrectos, 4 cadenas con 1 estado incorrecto y una sola cadena con todos los estados correctos la cual es la que se encuentra en el estado nativo.

Para dar una expresión analítica de la función g_α es necesario contar las

Tabla A.1: Tabla de valores de energía del estado base para varios átomos.

Cadena	α
AA	2
AB	2
AC	1
BA	2
BB	2
BC	1
CA	1
CB	1
CC	0

cadenas con un cierto numero de estados incorrectos. Sabemos que en este ejemplo, los estados A, B son estados incorrectos, pero no hay distinción entre ellos hablando en el sentido de el conteo de α estados no nativos, por lo que si en K casillas se pueden acomodar los estados incorrectos, que en este caso son α , lo que indica que estamos haciendo es encontrar el numero de formas en que se pueden acomodar α elementos en K casillas, por lo que el primer termino que contribuye a la función de degeneración es:

$$\binom{K}{\alpha} \quad (\text{A.1})$$

Ahora, el segundo termino que contribuye en la función de degeneración se calcula tomando en cuenta que ahora fijando uno de los estado incorrectos, por ejemplo el A, los demás tienen ν formas de acomodarse multiplicado por el otro estado en el cual también pueden conformarse de ν formas y así hasta α -veces, ya que es el estado que se quiere, por lo que tenemos el siguiente termino:

$$\nu^\alpha \quad (\text{A.2})$$

Como cada uno de los eventos son independientes, multiplicamos el primer termino por el segundo, obteniendo finalmente la función de degeneración deseada:

$$g(\alpha) = \nu^{(K-\alpha)} \binom{K}{\alpha} \quad (\text{A.3})$$

A.2. Función de partición del modelo de Zwanzig

En la sección anterior se obtuvo la función de degeneración $g(\alpha)$ la cual nos dice cuantas configuraciones hay con un cierto numero de estados incorrectos. Ahora, sabemos que la función de partición quedaría de la siguiente forma contando todos los estados que se puede encontrar el sistema teniendo en cuenta que varios comparten la misma energía E_α :

$$Z = \sum_{\alpha=0}^K g_\alpha e^{-\beta E_\alpha} \quad (\text{A.4})$$

Sustituyendo el valor de la energía $E_\alpha = US - \epsilon \delta_{(\alpha,0)}$ en Z y el valor de la $g(\alpha)$, tenemos lo siguiente:

$$Z = \sum_{\alpha=0}^K \binom{K}{\alpha} \nu^\alpha e^{-\beta(US - \epsilon \delta_{(\alpha,0)})} \quad (\text{A.5})$$

Esta ecuación se puede desarrollar para quedar en una forma mas sencilla de la siguiente manera:

$$Z = \sum_{\alpha=0}^K \binom{K}{\alpha} \nu^{K-\alpha} e^{-\beta(U(K-\alpha)-\epsilon\delta_{\alpha,K})} = \sum_{\alpha=0}^K \binom{K}{\alpha} \nu^\alpha e^{-\beta U \alpha} e^{-\beta \epsilon \delta_{(\alpha,0)}} \quad (\text{A.6})$$

Quitamos un termino de la suma, el cual es $e^{-\beta \epsilon \delta_{(\alpha,0)}}$ ya que en $\alpha = 0$

$$Z = g_0 e^{-\beta \epsilon} + \sum_{\alpha=1}^K e^{-\beta \alpha U} \quad (\text{A.7})$$

$$Z = g_0 e^{-\beta \epsilon} + \sum_{\alpha=1}^K \binom{K}{\alpha} \nu^\alpha e^{-\beta U \alpha} \quad (\text{A.8})$$

Factorizamos el exponente α de ν y de la exponencial, nos queda:

$$Z = g_0 e^{-\beta \epsilon} + \sum_{\alpha=1}^K \binom{K}{\alpha} (\nu e^{-\beta U})^\alpha \quad (\text{A.9})$$

Esta forma es muy similar al teorema del binomio, salvo a que la sumatoria que tenemos aquí empieza en 1 y en el caso del teorema empieza en 0, por lo que solamente sumamos un cero de tal forma que la sumatoria empiece en 0 y podamos factorizar esta expresión para visualizarla de una manera mas sencilla.

$$Z = g_0 e^{-\beta \epsilon} + \sum_{\alpha=0}^K \binom{K}{\alpha} (\nu e^{-\beta U})^\alpha - 1 \quad (\text{A.10})$$

Este -1 sale precisamente del valor que toma la función en $\alpha = 0$, por lo que solamente había que sumarle un 1 dentro de la sumatoria y eliminarlo de esa forma, teniendo así la siguiente expresión:

$$Z = e^{\beta \epsilon} + (1 + \nu e^{-\beta U})^K - 1 \quad (\text{A.11})$$

Apéndice B

Códigos en *Mathematica*

A continuación se describe de manera general y resumida la estructura del algoritmo computacional con el que se realizaron los cálculos pertinentes para la función de degeneracidad del modelo de heteropolímero ideal.

B.1. Función de degeneracidad G

.....
Script (Mathematica)

```
% \begin{doublespace}
\noindent\(\pmb{\text{ClearAll}}[\text{{' '}}Global$\grave{ }$*{' '}]}\
\pmb{\text{Clear}}[\text{{' '}}Subscript{' '}]}\
\pmb{K=1;}\
\pmb{\beta =\frac{1}{R*T};}\
\pmb{R=8.31;}\
\pmb{f=1;\text{{(*)}16744923\text{{(*)}}}}\
\pmb{\Delta =-50000;}\
\pmb{\epsilon =N\left[\frac{-f*\Delta }{1+f}\right]}\
\pmb{U=N\left[\frac{-\Delta }{k*(1+f)}\right]}\
\pmb{\text{ph}}=\{1.61, 1.2, 0.96, 0.82, 0.66, 0.63, 0.56, 0.44,}\
\pmb{0.39, 0.38, 0.385, 0.35, 0.31, 0.315, 0.30, 0.29, 0.23, 0.14, 0.05, 0.001}\
\pmb{\text{aa}}=\{\}\
\pmb{S=1/\text{ph};}\
\pmb{m=1/b}\
\pmb{\text{prom}}=(\text{Sum}[b[[i]],\{i,1,k\}])/k}\
\pmb{1/\text{prom}}}\
\pmb{v=\text{Table}\left[\nu _i,\{i,k\}\right];}\
\pmb{a=\text{Table}[1,\{i,k\}];}
```

```

\pmb{B_0=a;}\
\pmb{\text{For}[i=1,i\leq k,i\text{++},]\}
\pmb{\text{ {      }B_i= \text{ReplacePart}\left[B_{i-1},i\to 0\right]}\}
\pmb{\quad ;}\}
\pmb{\text{Do}\left[P_i=\text{Permutations}\left[B_i\right]\}
, \{i,0,k\}\right] \} \}
\pmb{\text{totalH}=\text{Sum}\left[\text{Length}\}
\left[\text{Permutations}
\left[B_i\right]\right], \{i,0,k\}\right];}\}
\pmb{W=\text{Flatten}\left[\text{Join}\left[\text{Table}\left[P_i, \{i,0,1\}
\right]\right]\right];}\}
\pmb{d=\text{Partition}[W,k];}\}
\pmb{\text{Do}\left[q_i=(v*d[[i]]) \text{/} 0\to 1, \{1, \text{Length}[d]
\right];}\}
\pmb{\}
\pmb{\}
\pmb{\text{Do}\left[g_{\alpha} = \sum_{i=\left(\sum_{j=0}^{\alpha} \}
\text{Length}\left[\text{Permutations}\}
\left[B_j\right]\right)\}
-\text{Length}\left[\text{Permutations}\left[B_{\alpha}
\right]\right]+1\right)}^{\sum_{j=0}^{\alpha} \text{Length}\}
\left[\text{Permutations}\left[B_j\right]\right] \} \}
\left(\prod_{h=1}^k q_i[[h]\right),
\{\alpha, 0, k\}\right];}\}
\pmb{\text{ZHet}=e^{-\beta * \epsilon} + \sum_{\alpha=0}^{k-1}
\left(\left(g_{\alpha}\right)*\left(e^{-\beta * U*(k-\alpha)}\right)\}
\right);}\}
%\end{doublespace}

```

B.2. Función de degeneracidad G con elección aleatoria

.....
 Script (Mathematica)

B.2. FUNCIÓN DE DEGENERACIDAD G CON ELECCIÓN ALEATORIA107

```

\begin{doublespace}
\noindent\(\pmb{\text{ClearAll}}[\text{{'Global$\grave{}$*{'}}}]\\
\pmb{k=8;}\\
\pmb{\beta =\frac{1}{R*T};}\\
\pmb{R=8.31;}\\
\pmb{f=1;\text{{(*)16744923\text{{(*)}}}}\\
\pmb{\Delta =-60000;}\\
\pmb{\epsilon =N\left[\frac{-\Delta}{1+f}\right];}\\
\pmb{U=N\left[\frac{-\Delta}{k*(1+f)}\right];}\\
\pmb{\text{{ph}}=\{1.61, \dots, 0.29, 0.23, 0.14, 0.05, 0.001\};}\\
\pmb{S=1/\text{{ph}};}\\
\noindent\(\pmb{r=\text{{Table}}[S[\text{{RandomInteger}}[\{1,19\}]],\{k\}];}\\
\pmb{\text{{Do}}\left[s_j=\text{{ReplacePart}}[r,j\to 0.001],\{j,1,8\}\right]}\\
\pmb{\text{{}})}\\
\noindent\(\pmb{v=\text{{Table}}\left[\nu_{i,k}\right];}\\
\pmb{a=\text{{Table}}[1,\{i,k\}];}\\
\pmb{B_0=a;}\\
\pmb{\text{{For}}[i=1,i\leq k,\text{{++}],}\\
\pmb{\text{{text{{}}}}B_i=\text{{ReplacePart}}\left[B_{i-1},i\to 0\right]}\\
\pmb{\text{{quad }}];}\\
\pmb{\text{{}})}\\
\pmb{\text{{Do}}\left[P_i=\text{{Permutations}}\left[B_i\right],\right)}

\pmb{v[\{i,0,k\}\right]}\\
\pmb{\text{{totalH}}=\text{{Sum}}\left[\text{{Length}}\right]}
\left[\text{{Permutations}}\left[B_i\right],\{i,0,k\}\right];\\
\pmb{W=\text{{Flatten}}\left[\text{{Join}}\left[\text{{Table}}\right]}
\left[P_i,\{i,0,k\}\right]\right];\\
\pmb{d=\text{{Partition}}[W,k];}\\
\pmb{\text{{Do}}\left[q_i=(v*d[[i]]) \text{{/.}} 0\to 1,\{i,1,\text{{Length}}[d]\right]}
\pmb{\text{{}})}\\
\pmb{\text{{}})}\\
\pmb{\text{{Do}}\left[g_{\alpha} = \sum_{i=\left(\sum_{j=0}^{\alpha}\right)}\right]}
\left[\text{{Length}}\left[\text{{Permutations}}\left[B_j\right]\right]-\right]}
\left[\text{{Length}}\left[\text{{Permutations}}\left[B_{\alpha}\right]\right]}
\right]+1\right)^{\sum_{j=0}^{\alpha}\text{{Length}}\left[\text{{Permutations}}\left[B_j\right]\right]}
\left[\text{{Prod}}_{h=1}^k q_{i[[h]]}\right], \{\alpha,0,k\}\right]}\\
\pmb{\text{{ZHet}}=e^{-\beta * \epsilon} + \sum_{\alpha=0}^{k-1}}
\left[\left(g_{\alpha}\right)*\left(e^{-\beta *U*(k-\alpha)}\right)\right];\\
\end{doublespace}

```

Apéndice C

Partícula Cuántica en una caja.

En este apartado se calculará la función de partición para una partícula libre. Vamos a considerar una partícula que solamente se puede mover sobre el eje x en un rango de 0 a L ($0 < x < L$). En las fronteras consideraremos que el potencial $V(x)$ es infinito, es decir, $V(0) = V(L) = \infty$.

Teniendo la ecuación de Schrodinger:

$$H\psi = E_j\psi, \quad (\text{C.1})$$

donde el operador H aplicado a la función de onda ψ es:

$$H = -\left(\frac{\hbar^2}{8\pi^2m}\right) \frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x). \quad (\text{C.2})$$

Como la partícula es libre, quiere decir que no está sujeta a ninguna fuerza potencial, por lo que $V(x) = 0$, quedando así la ecuación de Schrodinger como:

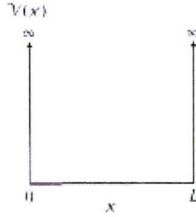
$$-\left(\frac{\hbar^2}{8\pi^2m}\right) \frac{\partial^2\psi}{\partial x^2} = E_j\psi. \quad (\text{C.3})$$

Realizamos el siguiente cambio:

$$k = \frac{E_j}{\left(-\frac{\hbar^2}{8\pi^2m}\right)}. \quad (\text{C.4})$$

Nuestra ecuación se reduce a una ecuación diferencial de segundo orden:

$$\frac{\partial^2\psi}{\partial x^2} - k\psi(x) = 0. \quad (\text{C.5})$$



h

Figura C.1: Partícula libre en una caja.

Para resolver la ecuación anterior, se propone una solución de la siguiente forma:

$$\psi(x) = A \operatorname{sen}(\sqrt{k}x) + B \operatorname{cos}(\sqrt{k}x). \quad (\text{C.6})$$

Como $V(0) = V(L) = \infty$, quiere decir que la partícula no puede estar en las paredes ($x = 0$ y $x = L$), por lo que las condiciones a la frontera estarán dadas por:

$$\psi(0) = 0; \psi(L) = 0. \quad (\text{C.7})$$

Usando la primera condición de frontera ($\psi(0) = 0$), sustituimos en la solución propuesta:

$$\psi(0) = A \operatorname{sen}(\sqrt{k}(0)) + B \operatorname{cos}(\sqrt{k}(0)). \quad (\text{C.8})$$

En la parte de la función *cos*, la única manera de que ese termino sea cero es que B sea igual a cero, por lo que la solución general nos queda como:

$$\psi(x) = A \operatorname{sen}(\sqrt{k}x) \quad (\text{C.9})$$

Ahora, usamos la segunda condición de frontera ($\psi(L) = 0$):

$$\psi(L) = A \operatorname{sen}(\sqrt{k}L) = 0 \quad (\text{C.10})$$

Para que la función *sen* sea cero, su argumento tiene que ser un múltiplo entero de π , o sea:

$$\sqrt{k}L = n\pi; n \in \text{Enteros}. \quad (\text{C.11})$$

$$k = \left(\frac{n\pi}{L}\right)^2 \quad (\text{C.12})$$

$$E_j = \frac{n^2 \pi^2}{L^2} * \left(-\frac{h^2}{8\pi^2 m} \right) \quad (\text{C.13})$$

Realizando el álgebra necesaria, tenemos la expresión para los niveles de energía, que solamente dependen de n :

$$E_n = \frac{n^2 h^2}{8mL^2} \quad (\text{C.14})$$

Esta ecuación nos da la energía para cada n que demos.

Ahora, para que la función de onda quede bien definida, tenemos que calcular el valor de la constante A , y la determinaremos normalizando la función, es decir:

$$\int_0^L (A \text{sen}(\sqrt{k}x))^2 dx \quad (\text{C.15})$$

Resolviendo esta integral con una sustitución trigonométrica ($\text{sen}^2(x) = \frac{1}{2} - \frac{1}{2}\cos(2x)$) obtenemos:

$$\int_0^L (A \text{sen}(\sqrt{k}x))^2 dx = A^2 \left[\frac{L}{2} - \frac{L}{2\pi n} \text{sen}(zx) \right] \quad (\text{C.16})$$

$$= A^2 \left[\frac{L}{2} - \frac{L}{2\pi n} \text{sen}(2\pi n) \right] \quad (\text{C.17})$$

$$= 1 \quad (\text{C.18})$$

Como $\text{sen}(2\pi n) = 0$, entonces:

$$A^2 \left[\frac{L}{2} \right] = 1 \quad (\text{C.19})$$

$$A = \left[\frac{2}{L} \right]^{\frac{1}{2}}. \quad (\text{C.20})$$

La función de onda queda totalmente determinada por la siguiente expresión:

$$\psi_n(x) = \left[\frac{2}{L} \right]^{\frac{1}{2}} \text{sen} \left(\frac{n\pi x}{L} \right) \quad (\text{C.21})$$

Ahora, sabemos que la función de partición translacional esta definida de la siguiente manera:

$$q_{\text{translacional}} = \sum_{n=1}^{\infty} e^{\left(\frac{\epsilon_n}{kT}\right)} \quad (\text{C.22})$$

$$= \sum_{n=1}^{\infty} e^{\left(\frac{n^2 h^2}{8mL^2 kT}\right)} \quad (\text{C.23})$$

En el argumento de la función exponencial tiene que tener una cantidad adimensional, por lo que el término $\frac{h^2}{8mL^2 k}$ lo definimos como una cantidad en la cual se tenga unidades de temperatura, definiendola de la siguiente manera:

$$\theta_{\text{translacional}} = \frac{h^2}{8mL^2 k} \quad (\text{C.24})$$

Ahora, si los términos del argumento de la exponencial son muy grandes, por ejemplo, m , L o T , entonces el término:

$$\frac{\theta_{\text{translacional}}}{T} \ll 1 \quad (\text{C.25})$$

por lo que los niveles de energía están muy cercanos entre si, tanto que podemos aproximar la expresión de la función de partición por una integral, quedando así:

$$\int_0^{\infty} e^{\left(\frac{n^2 h^2}{8mL^2 kT}\right)} dn \quad (\text{C.26})$$

Haciendo uso de formulas de integrales, tenemos que para alguna a dada:

$$\int_0^{\infty} e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \quad (\text{C.27})$$

Por lo que nuestra integral toma la siguiente forma:

$$\int_0^{\infty} e^{\left(\frac{n^2 h^2}{8mL^2 kT}\right)} dn = \frac{1}{2} \sqrt{\frac{\pi}{\frac{h^2}{8mL^2 kT}}} \quad (\text{C.28})$$

Entonces:

$$q_{\text{translacional}} = L \sqrt{\frac{2\pi mkT}{h^2}} \quad (\text{C.29})$$

Bibliografía

- [1] Arboledas Brihuega, David. Jerarquía estructural de las proteínas. Editorial club universitario. 2008.
- [2] Berg, Jeremy M. . Bioquímica. Edit. Reverté.
- [3] Devlin, Thomas M. Bioquímica: libro de texto con aplicaciones clínicas. Ed. Reverte. 2004.
- [4] Bray, Dennis. Introducción a la biología celular. Ed. Medica Panamericana. 2a Ed. 2004.
- [5] Holum, John. Fundamentos de Química general, orgánica y bioquímica para las ciencias de la salud. Limusa -Wiley. 2007.
- [6] Voet, Donald. Bioquímica. Ed. medica panamericana. 2006.
- [7] Chancz-Cárdenas, Ma. Elena, et-al. Enfermedades relacionadas al plegamiento anómalo de las proteínas. UNAM. México, D.F.
- [8] Dill, Ken A. Bromberg, Sarina. Molecular Driving Forces, Statistical Thermodynamics in Chemistry and Biology.
- [9] Reif. F. Física estadística. Editorial Reverte. 1993.
- [10] Bakk, Audun. Heat Capacity of Protein Folding. Biophysical Journal. Volume 81. August, 2001. 710-714.
- [11] IICA-México. Enfermedades Exóticas de los Animales. IICA Biblioteca Venezuela, 2000.
- [12] De Paula, Atkins; Julio de Paula. Química Física. Ed. Médica Panamericana, 2007. Pag. 129.
- [13] Aguilar Gutiérrez, Miguel. Biomecánica: la física y la fisiología. Editorial CSIC Press. 2000.
- [14] Pereto, Juli. Fundamentos de bioquímica. Universitat de València, 2007.

- [15] Garrido, Amando; et al. Fundamentos de bioquímica estructural. Ed. Tebar, 2a edición. 2006.
- [16] Rodríguez Merchan, Carlos. Prótesis de Rodilla Primaria: Estado Actual. Ed. Medica Panamericana. 2008.
- [17] Macarulla, José M. Goñi, Félix M. Bioquímica humana: curso básico. Reverte, 1994. Pag. 100
- [18] Branden, Carl. Introduction to protein structure. 2 Ed. Garland Publishing. 1998. Pag. 3.
- [19] http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1972/anfinsen-bio.html
- [20] Pena, Antonio. Bioquímica. Ed. Limusa. 2a edición. 2004. Pags. 345-348.
- [21] Anfinsen, Christian B. Principles that Govern the Folding of Protein Chains. Science, New Series, Vol. 181, No. 4096 (Jul. 20, 1973), pp. 223-230
- [22] Cussó Pérez, Fernando. Fundamentos físicos de los procesos biológicos, Vol. 2. Editorial club universitario. 2013.
- [23] Kerber, Robert C. If It's Resonance, What Is Resonating? J. Chem. Educ. 2006. 83 223
- [24] Branden, C.; Tooze, J. Introduction to Protein Structure. Second Edition Garland Publishing Co. 1999.
- [25] Chakrabarty A, Kortemme Tand Baldwin RL 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix stabilizing side-chain interactions Protein Sci. 3 843-852
- [26] Lehninger AL, Nelson DL y Cox MM. Principios de Bioquímica. 2ª ed. Barcelona: Ediciones Omega; 1995.p. 118-121.
- [27] Lifson. Roig. Theory of helix-coil transition in polypeptides J. Chem. Phys. 34 1963.
- [28] Reginald H. Garrett, Charles M. Grisham. Biochemistry. Brooks Cole (2004).
- [29] Pacheco Leal, Daniel. Bioquímica, Estructural y aplicada a la medicina. IPN, D.R. 1996.
- [30] Nelson, David L. Principios de Bioquímica. Ediciones Omega. 4ª Ed. 2008.

- [31] M. Angelini. et al. *Temas de Química General*. Eudeba. 1994.
- [32] M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GRO-MACS development team, *GROMACS User Manual version 5.0.4*, www.gromacs.org (2014)
- [33] Zwanzig, Robert. Simple model of protein folding kinetics. (1995). *Proc. Natl. Acad. Sci. USA*. Vol. 92. pp. 9801-9804.
- [34] Finkelstein, Alexei V. *Protein Physics A Course of Lectures*. Oleg B. Ptitsyn
- [35] Gitterman, Moshe. *Phase Transitions. A Brief Account with Modern Applications*. World Scientific Publishing Co. Pte. Ltd.
- [36] Callen, Herbert. *Termodinamica*. Ed. AC.
- [37] Wilhelm, Emmerich. Letcher, Trevor M. *Heat Capacities. Liquids, Solution and Vapours*. RSC Publishing, 2010.
- [38] Olivares Quiroz, Luis A. Thermodynamics of ideal proteinogenic homopolymer chains as a function of the energy spectrum E , helical propensity ω and enthalpic energy barrier. *J. Phys.: Condens. Matter* 25 (2013) 155103.
- [39] Stanley, Eugene. *Introduction to phase transitions and critical phenomena*. Clarendon Press, Oxford. 1971. Pag. 131.
- [40] Kondepudi D., Prigogine I. *Modern thermodynamics* (Wiley, 1998)(ISBN 0471973947).
- [41] Siewert J. Marrink, et al. J. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *Phys. Chem. B* 2007, 111, 7812-7824.
- [42] L. Onsager. *Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition*. *Physical Review*, 65, 117 (1944).
- [43] Skolnick, Jeffrey. Holtzer, Alfred. *Theory of α -Helix-to-Random-Coil Transition of Two-Chain, Coiled Coils. Application of the Augmented Theory to Thermal Denaturation of α -Tropomyosin*. American Chemical Society. 1985.
- [44] Mensaje bioquímico. *Enfermedades relacionadas al plegamiento anómalo de las proteínas*. Vol. XXVI (2002).
- [45] <http://www.nature.com/horizon/proteinfolding/background/disease.html>

- [46] Menendez, Silvia Gra. Padrón Perez, Noel. Llibre Rodríguez, Juan de Jesús. Péptido Beta amiloide, Proteína Tau y Enfermedad de Alzheimer. 2002.
- [47] Blázquez Fernández, Enrique. Fundamentos moleculares de la medicina. Real Academia Nac. Medicina, 2005. Pag. 120.
- [48] Felsenstein KM, Lewis-Higgins LL. Processing of the b-amyloid precursor protein carrying the familial, Dutch-type, and a novel recombinant C-terminal mutation. *Neurosci Lett* 1993; 152:185-9.
- [49] Goto, Satoshi. MD, Asao Hirano, MD and Sadayuki Matsumoto, MD. Met-enkephalin immunoreactivity in the basal ganglia in Parkinson's disease and striatonigral degeneration. *Neurology* July 1990 vol. 40 no. 7 1051
- [50] Zagon, Ian S. et. al. The biology of the opioid growth factor receptor (OGFr). *Brain Research Reviews* 38 (2002) 351-376
- [51] Reichlin S. Neuroendocrinología. En: Williams, Tratado de Endocrinología. t1. La Habana: Científico-técnica; 1997. p. 656-8.
- [52] Guyton-Hall. Sensaciones somáticas: II. Dolor, cefalea y sensaciones de temperatura. En: Tratado de Fisiología Médica. La Habana: Ciencias Médicas; 1999. p. 666.
- [53] Cardellá, R. Hernández F. Hormonas. En: Bioquímica Médica. t3. La Habana: Científico-técnica; 1999. p. 380-2.
- [54] Leihninger AL, Nelson DL y Cox MM. Principios de Bioquímica. 2ª ed. Barcelona: Ediciones Omega; 1995. p. 334-6.
- [55] www.multimedgrm.sld.cu/articulos/2003/v7-2/12.html
- [56] Protein Data Bank. Troubled Hemoglobins. [http : //www.pdb.org/pdb/101/motm.domomID = 41](http://www.pdb.org/pdb/101/motm.domomID=41)
- [57] Emson, P.C. Regional distribution of methionine-enkephalin and substance P-like immunoreactivity in normal human brain and in Huntington's disease. *Brain Research*. Volume 199, Issue 1, 13 October 1980, Pages 147-160
- [58] Schwarzer C. 30 years of dynorphins-new insights on their functions in neuropsychiatric diseases. *Pharmacol Ther*. 2009 Sep; 123(3):353-70.
- [59] Björnerås J1, Gräslund A, Måler L. Membrane interaction of disease-related dynorphin A variants. *Biochemistry*. 2013 Jun 18;52(24):4157-67.

- [60] Ramsey, N.F. *Phys. Rev.* 103 (1956) 20.
- [61] Piñeiro Redondo, Yolanda. Simulación de Monte Carlo de sistemas complejos en red. Univ Santiago de Compostela.
- [62] Steinberg MH. Sickle cell disease and associated hemoglobinopathies. In: Goldman L, Schafer AI, eds. *Cecil Medicine*. 24th ed. Philadelphia, PA: Saunders Elsevier; 2011:chap 166.
- [63] Lentz, Nicholas B. *Electrospray Ionization Mass Spectrometry: From Cluster Ions to Toxic Metal Ions in Biology*. ProQuest 2008.
- [64] Milosavljevic, A. R. et al. Gas phase spectroscopy of a protein. *Journal of Physics: Conference Series* 257 (2010)
- [65] Mackerell, Alexander D. *Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations*. Wiley Periodicals, Inc. 2004.
- [66] Kosterlitz, Hans W. Endogenous opioid peptides: multiple agonists and receptors. *Nature* 267, 495 - 499. 1977
- [67] Gillig, Kent J. and Chen, Chung-Hsuan. Critical Examination of Gas-Phase Protein Conformation/Multimer Ion Formation by Electrospray Ion Mobility-Mass Spectrometry. *Anal. Chem.* 2013, 85, 2177–2182.
- [68] Oomens Jos, et al. Charge-state resolved mid-infrared spectroscopy of a gas-phase protein. *Phys. Chem. Chem. Phys.*, 2005, 7, 1345–1348.
- [69] Jurchen, John C. and Williams, Evan R. Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers. *J. Am. Chem. Soc.* 2003, 125, 2817-2826 9 2817
- [70] Warnke, Stephan. Von Helden, Gert. and Pagel, Kevin. *Protein Structure in the Gas Phase: The Influence of Side-Chain Microsolvation*, *J. Am. Chem. Soc.* 2013, 135, 1177–1180,
- [71] Shelimov, Konstantin B. and Jarrold, Martin F. Conformations, Unfolding, and Refolding of Apomyoglobin in Vacuum: An Activation Barrier for Gas-Phase Protein Folding. *J. Am. Chem. Soc.* 1997, 119, 2987-2994.
- [72] Arteca, Gustavo A. and Tapia, O. Structural transitions in neutral and charged proteins in vacuo. *Journal of Molecular Graphics and Modelling* 19, 102–118, 2001.
- [73] Arteca, Gustavo A. Reimann, C.T. and Tapia, O. Proteins in vacuo: denaturing and folding mechanisms studied with computer- simulated molecular dynamics. *Mass Spectrometry Reviews*, 2001, 20, 402 – 422.

- [74] Patriksson, Alexandra. From Solution into Vacuum - Structural Transitions in Proteins. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 360.
- [75] Reimann, C. T. Velazquez, I. and Tapia, O. Proteins in Vacuo. Denaturation of Highly-Charged Lysozyme Studied by Molecular Dynamics Simulations. *J. Phys. Chem. B* 1998, 102, 9344-9352.
- [76] Rouse, Sarah L. Marcoux, Julien. Robinson, Carol V. and Sansom, Mark S. P. Dodecyl Maltoside Protects Membrane Proteins In Vacuo. *Biophysical Journal* Volume 105 August 2013 648-656.
- [77] I. Makhatadze, G. I. and Privalov, P. L. *Adv. Protein Chem.*, 1995, 47, 307.
- [78] Gomez, Javier. Hilser, Vincent J. The heat Capacity of Proteins. *PROTEINS: Structure, Function and Genetics*. 1995. Wiley. 404-412.
- [79] Privalov, P.L. Makhatadze, G. I. Partial Molar Heat Capacity of Individual Amino Acid Residues in Aqueous Solution: Hydration Effect. *J. Mol. Biol.* (1990) 213, 375-384.
- [80] Privalov, P.L. Makhatadze, G.I. Heat capacity of proteins. 11. Partial molar heat capacity of the unfolded polypeptide chain of proteins: Protein unfolding effects. *J. Mol. Biol.* 213:385-391, 1990.
- [81] Privalov, P. L. & Gill, S. J. (1989). *Advan. Protein Chem.* 39, 191-234.
- [82] Sanz Morales, Jesús Miguel. Estudios estructurales y aplicaciones biotecnológicas de la Lisozima del bacteriofago Cp-I de neumococo. Colección tesis doctorales No. 206/92. Madrid. 1992.
- [83] Privalov, P. L. Thermodynamics of protein folding. *J. Chem. Thermodynamics* 1997, 29, 447-474.
- [84] Bakk, Audun. Heat Capacity of Protein Folding. *Biophysical Journal* Volume 81 August 2001 710-714.
- [85] Makhatadze, G.I., Clore, G.M., Gronenborn, A., Privalov, P.L. Thermodynamics of unfolding of the all β -sheet protein Interleukin 1- β . *Biochemistry* 31:9327-9332, 1994.
- [86] Yu, Y. Makhatadze, G.I., Pace, C.N. Privalov, P.L. Energetics of ribonuclease T1 structure. *Biochemistry* 33:3312-3319. 1994.
- [87] Wintrode, P.L. Makhatadze, G.I. Privalov, P.L. Thermodynamics of ubiquitin unfolding. *Proteins* 18:246-253, 1994.

- [88] Voichita M. Insights into Protein Compressibility from Molecular Dynamics Simulations. *J. Phys. Chem. B* 2001, 105, 715-724.
- [89] Lin, Lung-Nan. Determination of the Volumetric Properties of Proteins and Other Solutes Using Pressure Perturbation Calorimetry. *Analytical Biochemistry* 302, 144-160 (2002).