



**UNIVERSIDAD AUTONOMA METROPOLITANA**  
UNIDAD IZTAPALAPA

**El Modelo Extendido de Zwanzig y la Teoría de  
Adam-Gibbs en el Plegamiento y  
Desnaturalización de Proteínas.**

TESIS QUE PRESENTA EL

**M. EN C. LUIS OLIVARES QUIROZ**

PARA LA OBTENCIÓN DEL GRADO DE  
DOCTOR EN CIENCIAS.

**ASESOR.**

**DR LEOPOLDO GARCÍA-COLÍN SCHERER.**

*Profesor Emérito Universidad Autónoma Metropolitana-Iztapalapa.*

*Miembro de El Colegio Nacional.*

*Doctor Honoris Causa UNAM 2007.*

**CO-ASESOR**

**DR. LEONARDO DAGDUG LIMA.**

*Profesor Universidad Autónoma Metropolitana-Iztapalapa.*

México Distrito Federal, Julio 2007.



**UNIVERSIDAD AUTONOMA METROPOLITANA**  
UNIDAD IZTAPALAPA

**El Modelo Extendido de Zwanzig y la Teoría de  
Adam-Gibbs en el Plegamiento y  
Desnaturalización de Proteínas.**

Tesis Doctoral.

**COMITÉ SINODAL.**

**DR. LEOPOLDO GARCÍA COLIN SCHERER.**  
DEPARTAMENTO DE FISICA. UAM-IZTAPALAPA

**DR. EUSEBIO JUARISTI Y COSSÍO.**  
DEPARTAMENTO DE QUIMICA. CINVESTAV-MÉXICO.

**DR. MICHEL PICQUART.**  
DEPARTAMENTO DE FÍSICA. UAM-IZTAPALAPA.

**DR. EDGAR VAZQUEZ CONTRERAS.**  
INSTITUTO DE QUIMICA. UNAM.

**DR. ORLANDO GUZMAN LOPEZ.**  
DEPARTAMENTO DE FISICA. UAM-IZTAPALAPA.

## **Agradecimientos.**

*Si bien escribir los resultados de una investigación que duró tres años es un proceso altamente complicado, no es lo menos expresar el agradecimiento a todos aquellos que participaron en forma entusiasta y decidida en el inicio, desarrollo y conclusión de esta etapa tan importante para mí.*

*No sabría bien por dónde empezar, pero en algún punto hay que hacerlo y lo haré diciendo entonces que esto no sería de ninguna forma posible si mis padres, Prof. Fausto Olivares Lara y Profa. Gloria Quiroz Arellanes, no hubiesen creído en un sueño. Gracias a ellos por apoyarme en la decisión de partir hacia un rumbo desconocido y gracias por estar siempre ahí, aún en los momentos en que yo mismo no sabía dónde estaba.*

*A mi hermano Carlos Fernando también debo agradecer su confianza y compañía a lo largo de toda la vida, que si bien a veces físicamente lejos, nunca demasiado.*

*A mis profesores tanto de la licenciatura como del posgrado debo en gran parte estar hoy aquí. En especial, gracias al Dr. Jorge Velasco Hernandez (IMP, México), Dra. Rosa Ma. Velasco Belmont (UAM-I, México) y al Dr. Rafael Barrio Paredes (IFUNAM, México) quienes fueron los primeros que me abrieron las puertas de la ciencia y la investigación. El Dr. Leopoldo García-Colín Scherer merece un agradecimiento especial. La inspiración que me genera conocer su brillante carrera sumado a la oportunidad que me ha dado de conocerlo personalmente, serán elementos que me acompañarán hoy y siempre.*

*De mis amigos y compañeros tendría mucho que decir y poco espacio para hacer justicia a su invaluable compañía y afecto a lo largo de toda la vida. Están mis amigos de siempre, los fisiquitos, Ramón, Fabián, Eva y Marco. Gracias también a Sergio Arguijo por mostrarme que sí, la Vida es Bella, a Daniel Moreno por las largas tertulias, a Omar Vicenzio por demostrar que la distancia más corta entre dos puntos no siempre es la línea recta, a Alfonso Aréchiga por recordarme con el ejemplo que la sencillez está entre nuestros mejores valores, a Carlos Meza y las tardes en Buenos Aires, a Víctor Juárez y a Jorge Estrada por las discusiones sobre la paz mundial... A todos los que están, pero por sobre todo a aquellos que ya no. Sin ustedes, no podría haberlo hecho.*

*Gracias , from the bottom of my heart.*

*Luis Olivares Quiroz  
México D.F., Julio 2007.*

## Índice general

Capítulo 1. Actualidad y Relevancia del Proyecto.	5
1. Perspectiva General del Plegamiento de las Proteínas.	6
2. Implicaciones Biomédicas del Plegamiento de las Proteínas.	14
Capítulo 2. Estructura e Interacciones en las Proteínas.	19
1. Síntesis de una proteína.	19
2. Los aminoácidos: Unidades básicas de la cadena polipeptídica.	22
3. Estructura secundaria: Hélices $\alpha$ y Hojas $\beta$ .	27
4. Estructura Terciaria y el Estado Nativo: Integración en una unidad biológicamente activa.	36
5. ADN, ARN, Proteínas y Evolución.	39
Capítulo 3. Desnaturalización de una proteína inducida químicamente.	47
1. Estabilidad del estado nativo. El efecto hidrofóbico.	47
2. Desnaturalización de una proteína.	50
3. Desnaturalización química de una proteína: Base experimental de la propuesta.	56
4. Teoría del Paisaje Energético: Base Teórica de la Propuesta.	61
5. Modelo Extendido de Zwanzig I. Desnaturalización lineal	68
6. Modelo Extendido de Zwanzig II. Perfil sigmoidal de desnaturalización.	75
Capítulo 4. Cinética del Plegamiento y la Transición Vítreas	93
1. La Transición Vítreas y la Desnaturalización de Proteínas.	93
2. Teoría Básica de la Transición Vítreas.	100
3. Desnaturalización de la Mioglobina I. Funciones Termodinámicas.	104
4. Desnaturalización de la Mioglobina II. Entropía de las Regiones Cooperativas.	111
Capítulo 5. Perspectivas y Conclusiones.	123
1. Primer Eje de Investigación: Termodinámica de la desnaturalización.	123
2. Segundo eje de investigación: Cinética de la desnaturalización.	127
3. Impacto del trabajo.	131
Bibliografía	133

## CAPÍTULO 1

### Actualidad y Relevancia del Proyecto.

En las últimas décadas hemos presenciado un avance vertiginoso en la integración de la Física con otras Ciencias Naturales como la Biología y la Química. Al reconocer que el objeto central de estudio de la Física es en último término el análisis de las interacciones entre la materia y la energía, se han abierto las puertas para su aplicación en campos distintos a aquellos de donde surgió. En este sentido, uno de los éxitos más contundentes de esta amalgama entre Física y otras disciplinas se ha gestado desde hace varias décadas en las ciencias biomédicas, al ponerse de manifiesto que muchos de los mecanismos y procesos que ocurren en los sistemas biológicos tienen una base física y pueden ser descritos en el marco de ésta. Es en esta intersección del conocimiento, llamada por algunos Biofísica o Física Biológica, en donde se ubica el objetivo de este proyecto doctoral. Sin pretender de alguna manera afirmar que la Física tiene por sí misma las respuestas a las preguntas más fundamentales que atañen a la Biología, en este proyecto de investigación pretenderemos mostrar cómo algunos de los procesos bioquímicos en el campo del plegamiento de las proteínas, tienen su origen en la interacción de un conjunto enorme de partículas interactuando entre sí mediante fuerzas de origen físico y que, por tanto, son sistemas susceptibles a ser descritos, con un grado de precisión razonable, por el aparato conceptual y las herramientas de la Mecánica Estadística y la Termodinámica.

Como un ejemplo de la integración entre Física, Química y Biología podemos citar el recientemente descubierto mecanismo de interferencia de RNA-i (*RNA interference mechanism*), el cual consiste en la inhibición de secuencias específicas de nucleótidos en la cadena ADN y que, al ser suprimidas por la cadena de ARN-i no son traducidos por los ribosomas evitando así la expresión de un gen particular. Tal como se ha anunciado, el mecanismo de interferencia de ARN-i tiene implicaciones espectaculares en el control y supresión de genes nocivos y en el tratamiento de enfermedades de origen genético, como algunos tipos de cáncer [Fire et al. , 1998]. Nada inmerecidamente, el mecanismo de supresión genética vía la cadena de ARN-i descubierto por Andrew Fire y Craig Mello ha sido merecedor del Premio Nobel de Medicina y Fisiología 2006.

Éste y otros avances que sería largo de enumerar en este momento, conducen actualmente a las ciencias puras como la Física, la Química, la Biología y las Matemáticas hacia una integración e interdependencia cada vez mayor. Es difícil hallar actualmente un conocimiento de impacto que pueda clasificarse como exclusivo de algunas de estas ciencias. En mi opinión, asistimos hoy a las primeras etapas en la gestación de la ciencia verdaderamente interdisciplinaria. Su desarrollo pleno y madurez seguramente traerá algunas de las respuestas a las preguntas que han inquietado a la Humanidad a lo largo de su existencia.

## 1. Perspectiva General del Plegamiento de las Proteínas.

Se conoce como proteínas a un conjunto muy amplio de macromoléculas biológicas que regulan la inmensa mayoría de los procesos fisicoquímicos de los seres vivos. En esencia, es posible clasificar a las proteínas en cuatro grupos dependiendo de la función que realizan: enzimas (proteínas que catalizan reacciones bioquímicas), proteínas de soporte estructural (colágeno, queratina), proteínas de almacenamiento y transporte (hemoglobina, mioglobina, albúmina) e inmunoglobulinas (anticuerpos) [Guo *et al* , 2005]. Pese a la ubicuidad de las proteínas en la mayoría de las reacciones bioquímicas de los seres vivos, ha sido sólo en las últimas cinco décadas cuando su estudio comenzó a atraer en forma significativa la atención de biólogos moleculares, físicos y químicos [Olivares-Quiroz and García-Colín L., 2004].

En la actualidad se sabe que las proteínas son ensambladas por los ribosomas en el interior de la célula, mediante la traducción de tripletas (codones) de nucleótidos en la cadena de ADN. Una vez sintetizadas, las proteínas son generadas como una cadena unidimensional y secuenciada de subestructuras conocidas como aminoácidos. La proteína así generada tiene una estructura que asemeja a una hebra unidimensional suspendida dentro de un fluido. Esta configuración, conocida en la literatura como **estado desplegado** (*unfolded state*), no es única y las posibles conformaciones que puede adquirir son similares a las diversas conformaciones posibles para una cadena heteropolimérica suspendida dentro de un fluido, como ha sido mostrado en los trabajos clásicos de Flory [Flory P.J. , 1953], [Flory P.G. , 1969], y De Gennes [de Gennes P.G. , 1976] en referencia a los polímeros. En términos generales, una vez definida la secuencia específica de aminoácidos que integra la proteína, la estructura primaria transita por una serie de estados en los cuales su configuración espacial se modifica, hasta alcanzar un estado termodinámicamente estable con una estructura tridimensional particular, la cual depende de la secuencia de aminoácidos inicial y por consecuencia es distinta para cada proteína. Esta estructura, conocida como **estado nativo** (*native state*), define los sitios de unión de la macromolécula<sup>1</sup> y permite a la proteína realizar la función para la cual fue sintetizada. El hecho anterior pone de manifiesto entonces, la conexión intrínseca entre forma y funcionalidad y representa uno de los paradigmas del así llamado plegamiento de las proteínas.

El proceso anterior, mediante el cual la macromolécula transita por diversas configuraciones hasta adquirir su estructura biológicamente activa, se conoce como plegamiento de una proteína (*protein folding*) y su estudio se refiere entonces a la elucidación de los mecanismos fisicoquímicos que permiten a la cadena lineal de aminoácidos evolucionar hacia la estructura biológicamente activa. En principio, no es evidente que un proceso de carácter estrictamente biológico como el plegamiento de las proteínas pueda ser susceptible de ser abordado por las herramientas y métodos de la Física; sin embargo, alrededor de 1961, C. Anfinsen en una serie de experimentos cruciales realizados en el *National Institute of Health* demostró que efectivamente éste es el caso para algunas proteínas [Anfinsen et al. , 1961].

En el medio celular, esto es, *in vivo*, la formación de la estructura nativa de una proteína ocurre casi inmediatamente después de su síntesis por los ribosomas. Una vez que los aminoácidos han sido ensamblados entre sí por los ribosomas, las interacciones

---

<sup>1</sup>Los sitios de unión de una proteína son regiones espaciales en donde se ensamblan otras proteínas o ácidos nucleicos a fin de realizar una función bioquímica particular.

físicas y químicas que ocurren entre la cadena lineal y el medio celular actúan como catalizadores para favorecer (o inhibir en algunos casos) la formación de la estructura nativa correcta. En 1961, C.B. Anfinsen mostró en una serie de experimentos con la proteína de origen pancreático bovino *Ribonucleasa-A* que el proceso de plegamiento podía ocurrir *in vitro*, esto es sin la presencia del medio celular. Asimismo, mostró también que bajo ciertas condiciones la desnaturalización de la proteína y el re-plegamiento (*renaturation*) hacia la estructura nativa es un proceso reversible [Anfinsen et al. , 1961]. Los experimentos realizados por Anfinsen muestran que en las condiciones adecuadas, los ocho aminoácidos *Cis* presentes en la secuencia primaria de la ribonucleasa son capaces de reagruparse adecuadamente en cuatro pares de enlaces disulfuro. A *grosso modo*, la idea esencial de estos experimentos consistió en monitorear la presencia de los ocho grupos sulfidrilos SH mediante técnicas de espectroscopía. El razonamiento seguido por Anfinsen le permitió relacionar la formación de los enlaces disulfuro correspondientes con la aparición de las propiedades espectroscópicas asociadas a la actividad enzimática de la Ribonucleasa-A. En ausencia de los puentes disulfuro, enlaces clave en la formación de la estructura nativa, la actividad enzimática de la proteína se reducía a cero.

Los resultados anteriores, que llevarían a C.B. Anfinsen a obtener el Premio Nobel en 1972, permitieron establecer las bases de lo que hoy en día se conoce como *hipótesis termodinámica*. En esta hipótesis, se considera que la estructura nativa de una proteína inmersa en las condiciones fisicoquímicas adecuadas de su entorno, (pH, temperatura, presión), corresponde a un estado mínimo global en la energía libre  $F$  del sistema. Los argumentos que se derivan de los experimentos de Anfinsen y que dan base a la Hipótesis Termodinámica, podrían agruparse de acuerdo a los siguientes puntos,

- Dado que el plegamiento puede ocurrir en ausencia del medio celular, la información necesaria para que la proteína adquiera su configuración nativa está codificada exclusivamente en la secuencia de aminoácidos. Los componentes del medio celular, ayudan a que este proceso se realice en forma adecuada y eficientemente, evitando por ejemplo la formación de enlaces no favorables entre cadenas laterales.
- El hecho que el mecanismo de plegamiento y desnaturalización sean procesos reversibles<sup>2</sup>, es una característica crucial para permitir una descripción basada en leyes y principios físicos que también tienen la característica de reversibilidad. Si el proceso fuese irreversible, considero que se tendría que admitir la presencia de elementos fuera del campo de la Física que rigen el proceso.<sup>3</sup>

Es importante hacer notar que la reversibilidad en la desnaturalización o plegamiento de la proteína depende no sólo de su masa molecular, sino también del tipo de aminoácidos que participan en la estabilización de la estructura nativa. Por ejemplo,

---

<sup>2</sup>La reversibilidad del plegamiento es un hecho experimental verificado en proteínas globulares pequeñas, cuya masa molecular es menor a 20kDa y que están integradas por < 100 aminoácidos. Para proteínas de masa molecular mayor y con un número mayor de aminoácidos, el proceso de desnaturalización no es necesariamente reversible debido a la cantidad de enlaces que estabilizan esta macroestructura.

<sup>3</sup>La afirmación anterior corresponde a la línea de pensamiento original de Anfinsen. En la actualidad se considera que aún en los casos de plegamiento irreversible es posible describir a los mecanismos en términos físicos mediante el uso de formalismos extendidos a la Termodinámica de Equilibrio.

al aminoácido *Cis* tiene una cadena lateral del tipo  $R = -\text{CH}_2 - \text{SH}$ , la cual en presencia de condiciones fisicoquímicas apropiadas, es capaz de formar un enlace químico con otra cadena similar, generando lo que se conoce como un enlace disulfuro S-S (*disulphide bridge*), el cual es uno de los enlaces de mayor intensidad en la estabilización de la estructura nativa. El papel de los enlaces disulfuro en la estabilización de la estructura nativa es tal que la sustitución de dos enlaces  $S - S$  en la ribonucleasa pancreática-A por enlaces con aminoácidos *alanina*, induce un decremento de  $40^\circ\text{C}$  en la temperatura  $T_m$  de desnaturalización de la proteína [Klink, T.A., *et al*, 2000]. Asimismo, la introducción de un enlace disulfuro en la variante bacteriófaga T4 de la lisosima incrementa su temperatura  $T_m$  en  $4.8^\circ\text{C}$ ,  $6.4^\circ\text{C}$  u  $11^\circ\text{C}$ , dependiendo la posición de la cadena en donde se ubique al enlace [Matthews., 1995].

La relevancia de los experimentos realizados por Anfinsen radica entonces en mostrar que, bajo ciertas condiciones de presión, temperatura, pH, etc., el plegamiento de una proteína es un proceso que puede ocurrir *in vitro*, esto es, sin la presencia del medio celular. Consecuencia inmediata de lo anterior es que un proceso de naturaleza estrictamente biológica como lo es el plegamiento puede ser llevado al laboratorio y por tanto ser sujeto de escrutinio por las herramientas de la Física y La Química. Asimismo, dado que el conjunto de estados desplegados  $D$  y el estado nativo  $N$  son conformaciones de equilibrio termodinámico, son susceptibles a ser abordados por la Termodinámica Clásica, cuyo postulado fundamental es la existencia de estados de equilibrio conectados mediante trayectorias reversibles en el espacio fase. Aún en aquellos casos de macromoléculas que presenten un alto grado de irreversibilidad entre las estructuras primarias y nativas, los estados iniciales y finales son, en última instancia, estados de equilibrio termodinámico y, por tanto, abordables mediante una descripción termodinámica estándar. Conectado intrínsecamente con lo anterior, se halla la posibilidad de una descripción en términos mecánico-estadísticos, dada la relación existente entre la Termodinámica de Equilibrio y la Mecánica Estadística desarrollada por L. Boltzmann y W. Gibbs a mediados del siglo pasado [García-Colin, L.S., 2005], [Gibbs, J.W., 1960]. El postulado principal de la Mecánica Estadística afirma que es posible establecer una conexión unívoca entre los promedios de las propiedades microscópicas de un sistema con sus características termodinámicas. Así, de un sólo golpe Anfinsen no sólo nos muestra que las proteínas son sistemas susceptibles de ser estudiados por los métodos y herramientas de la Física, nos muestra también la forma particular de hacerlo y las condiciones requeridas para tal descripción.

A la luz de lo anterior, ¿cuál es entonces el problema central a analizar en el plegamiento de las proteínas? En forma general, podríamos decir que el análisis presenta básicamente dos vertientes. Por un lado, uno de los aspectos centrales y que atrajeron mucho la atención en las primeras etapas se refiere a la predicción de la estructura nativa que adquirirá la proteína una vez conocida la secuencia inicial de aminoácidos que integra la estructura primaria. Por decirlo en términos físicos, el problema de la predicción de la estructura biológicamente activa es un problema de condiciones iniciales. Dada la secuencia ordenada de  $K$  aminoácidos para una proteína en particular, las condiciones fisicoquímicas del medio celular al que se hallará expuesta (presión, temperatura, agentes químicos, pH), se desea determinar su conformación de equilibrio. Este aspecto, conocido en la literatura como **predicción del estado nativo** implica necesariamente tener una idea general de los mecanismos fisicoquímicos involucrados en tal proceso, y por tanto conlleva en sí mismo la propuesta de



mecanismos de interacción entre la estructura primaria y el medio celular como de interacción entre la estructura primaria consigo misma. A este respecto, podemos decir que en las últimas décadas no hemos estado faltos de propuestas y resultados. Dos de las teorías más importantes y de mayor impacto en este sentido son la *Teoría del Paisaje Energético (Energy Landscape Theory)* propuesta por P. Wolynes y J. Bryngelson a mediados de los años ochenta [Bryngelson and Wolynes., 1989] y la *Teoría del Congelamiento de Estados en Heteropolímeros (Freezing heteropolymer transition theory)* propuesta por V. Pandé *et al* a finales de los noventa [Pande, Grosberg and Tanaka., 2000], de las cuales nos ocuparemos con mayor detalle en el Capítulo III, cuando discutamos los elementos técnicos de esta investigación doctoral.

Como segundo eje de investigación en el plegamiento, se halla otra de las características más sorprendentes del fenómeno. Numerosos experimentos con distintas proteínas han mostrado que el tiempo necesario para que la estructura primaria transite por estados conformacionales intermedios hasta llegar al estado nativo, es increíblemente más corto de lo esperado para un sistema de tal complejidad [Grosberg and Kutnezov., 1992], [Dill and Chan., 1997]. Los experimentos realizados en esta dirección muestran que el tiempo de plegamiento  $\tau$  de una proteína depende de su masa molecular, de su composición en aminoácidos, de las características propias de éstos y de las condiciones fisicoquímicas del solvente en el cual se halla inmersa. Pese a esta complejidad, en términos generales el tiempo de plegamiento  $\tau$  se halla ubicado en un intervalo muy preciso [Vendruscolo, M. *et al* , 2001], que va desde  $10^{-3}$  segundos para las proteínas más pequeñas ( $K < 100$  aminoácidos), hasta  $10^1$  segundos para macromoléculas de masa molecular mayor ( $K > 100$  aminoácidos)<sup>4</sup>. El hecho anterior no deja de ser bastante peculiar y notable, particularmente si tomamos en cuenta que en una proteína típica formada por, digamos  $K = 100$  aminoácidos en donde cada uno pudiese adquirir sólo dos posibles configuraciones, el número de posibles estructuras es del orden de  $C \sim 2^K$ , el cual es número astronómicamente grande. Si la proteína tuviese que hacer una búsqueda aleatoria de su configuración más estable pasando por el conjunto  $C$  de todos los posibles estados, entonces el tiempo de plegamiento  $\tau$  sería del orden de la edad del Universo, lo cual evidentemente no ocurre y plantea pues en definitiva, la existencia de mecanismos que impidan la búsqueda aleatoria del estado conformacional nativo.

En el fondo del planteamiento anterior, conocido como la Paradoja de Levinthal [Levinthal, C.J., 1968], se halla entonces un elemento extra que deberá añadirse para una descripción correcta del tiempo  $\tau$  y que se refiere a la presencia inherente de una tendencia (*bias*) en el diseño de la estructura primaria que codifica de alguna manera una o varias rutas de plegamiento, pero que inhibe la enorme mayoría de posibilidades restantes. Es quizás la solución al problema planteado por Levinthal uno de los mayores retos planteado por el plegamiento de las proteínas para la ciencia moderna. Quizás para algunos el hecho de que las proteínas alcancen un estado conformacional de equilibrio caracterizado por un mínimo energético no sea del todo sorprendente, pero difícilmente alguien podría negar que el hecho de que éste ocurra en tiempos suficientemente cortos deja aún mucho por entender.

---

<sup>4</sup>Existen no obstante, subunidades de las proteínas como *loops hairpins* conformados por pequeños segmentos de la cadena principal, que pueden plegarse en tiempos aún más cortos, del orden de  $10^{-6}$  segundos [Creighton, 1984]

En el párrafo anterior hemos descrito en forma general cuáles son los ejes principales en el estudio del plegamiento. A continuación procederemos a un nivel de especificidad mayor para describir exactamente cuáles es la fenomenología observada en esta transición y que cualquier teoría de la dinámica de proteínas en solución debería poder explicar. Para ello, agruparemos los resultados experimentales observados en dos categorías principales. Por un lado, listaremos aquellos resultados asociados directamente con los aspectos termodinámicos y por otro, los aspectos relacionados intrínsecamente con la cinética del proceso, es decir, con lo que ocurre en las fases intermedias del plegamiento y que afecta directamente al tiempo de plegamiento  $\tau$ .

### Aspectos Termodinámicos del Plegamiento.

Tal como hemos mencionado anteriormente, una de las vertientes en el análisis del plegamiento se refiere a la búsqueda de un modelo mesoscópico sustentado en principios físico-químicos, que permita calcular las diversas propiedades termodinámicas observadas en el estado nativo. En particular, existen tres características de crucial importancia relacionadas con este aspecto y que cualquier modelo que represente adecuadamente la termodinámica del estado nativo debería reproducir.

1. **Existencia de un mínimo para la Energía libre  $F$ .** Diversos experimentos muestran que el cambio de energía libre  $\Delta G$  o  $\Delta F$  del sistema exhibe un máximo para una temperatura  $T_f$ , la cual se halla generalmente en el intervalo de  $0 - 100^\circ C$ . Lo anterior implica, desde un punto de vista termodinámico, que a la temperatura  $T_f$ , la proteína ha alcanzado un mínimo global en su energía libre correspondiente a un estado de equilibrio termodinámico. Asimismo, si notamos que  $G = H - TS$  (o  $F = E - TS$ ), nos percataremos también de que la existencia de este mínimo global para las energías  $G$  y  $F$  implica también que el sistema ha alcanzado un estado de equilibrio entre interacciones termodinámicas. Por un lado, la tendencia natural de los sistemas a disminuir su energía interna  $H$  o  $E$  y por otro lado, de incrementar su entropía  $S$ . El estado nativo surge entonces como aquel que equilibra las dos tendencias anteriores. Se ha observado también que para la mayoría de las proteínas de un sólo dominio<sup>5</sup>, el  $\Delta F$  se halla en el intervalo  $40 - 60$  kJoules/mol [Pfeil and Privalov., 1976a] [Pfeil and Privalov., 1976b].
2. **Entalpía  $H$  y entropía  $S$  en función de  $T$ .** Para un número enorme de proteínas, entre las cuales podríamos citar sólo como unos cuantos ejemplos a la ribonucleasa A, la quimotripsina, la mioglobina, el citocromo c [Privalov, P.L., 1979], [Privalov and Gill., 1988] y una clase amplia de meta-mioglobinas [Kelly and Holladay., 1990], el cambio en la entalpía  $\Delta H$  es una función lineal creciente de  $T$  en tanto que el cambio de entropía  $\Delta S$  muestra un comportamiento también monótonicamente creciente con  $T$  aunque no necesariamente escala linealmente con  $T$ .
3.  **$\Delta C_p$  versus  $T$ .** Las técnicas de micro y nanocalorimetría de barrido muestran que la diferencia  $\Delta C_p$  entre el estado nativo  $N$  y el estado desplegado  $D$  en función de  $T$  exhibe un máximo para una temperatura dada en el intervalo antes mencionado ( $0 - 100^\circ C$ ) [Privalov *et al.*, 1989], [Michnik *et al.*, 2005], [Chateau *et al.*, 2005], [Privalov, P.L., 2007] (Figura 1). Una de las

---

<sup>5</sup>Pese a no haber un consenso en la definición del término *dominio*, generalmente se considera a éste como un subconjunto de la estructura primaria capaz de plegarse en forma independiente al resto de la cadena.

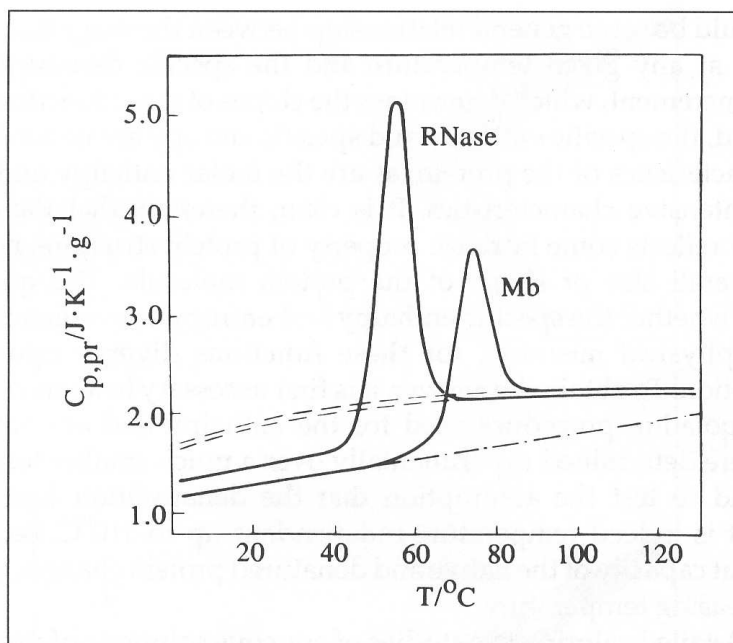


FIGURA 1. Forma funcional del calor específico  $C_p$  en función de la temperatura  $T$  para la proteína *Ribonucleasa pancreática A* y la *Metamioglobina*. La línea sólida corresponde a la estructura terciaria, en tanto que la línea punteada a la cadena sin estructura. Reproducido de Privalov *et al.* [Privalov *et al.*, 1989].

primeras explicaciones para la presencia de este pico de absorción, se halla en la propuesta de Kauzmann [Kauzmann, W., 1959], quien sugiere que en analogía con lo que ocurre en la hidratación de solutos orgánicos no polares en agua, el incremento sustancial del  $\Delta C_p$  se debe a la exposición de los grupos no polares al agua, característica fundamental del efecto hidrofóbico. En el Capítulo III discutiremos con mayor amplitud la base física de este proceso.

### Aspectos Cinéticos del Plegamiento

En párrafos anteriores hemos presentado el esquema general que sigue el plegamiento de las proteínas. Los ribosomas traducen la información codificada por tripletas de nucleótidos en el ADN a fin de generar cada aminoácido que integrará la secuencia primaria de la proteína. Asistidos por las cadenas de ARN-t, los ribosomas ensamblan la cadena polipeptídica que se plegará sobre si misma para formar a la estructura nativa. Un panorama tan simple como el descrito, que ignora por ejemplo los mecanismos de plegamiento asistidos por las proteínas chaperonas, es aplicable en términos generales para proteínas de masa molecular pequeña, en donde sólo existen dos estados termodinámicamente estables, el estado desplegado  $D$  y el estado nativo  $N$ . Cuando las proteínas tienen una masa molecular mayor a 20 kDa<sup>6</sup>, pueden existir un conjunto de estructuras intermedias metaestables, correspondientes a estados activados en la transición  $D \rightleftharpoons N$ . Estas estructuras, conocidas como estados intermedios, tienen un

<sup>6</sup>1 Da equivale a  $1.65 \times 10^{-24}$  gramos

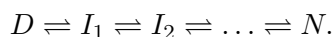
papel fundamental en la cinética del plegamiento, dado que su presencia o ausencia puede incrementar o disminuir en forma considerable el tiempo  $\tau$  necesario para que la proteína alcance el estado de mínima energía.

Se ha discutido a lo largo de mucho tiempo, si el mecanismo de plegamiento es un proceso dirigido termodinámicamente o cinéticamente. Esto es, si el estado nativo es el estado de menor energía o el estado que se forma más rápidamente [Wetlaufer and Ristow., 1973]. Los argumentos que favorecen la hipótesis termodinámica están sustentados básicamente en experimentos de desnaturalización para proteínas pequeñas, en donde en efecto, el estado nativo es el resultado de un proceso de minimización energética. Por otro lado, la idea de que el plegamiento tiene una componente importante de origen cinético tiene su fundamentación en la idea de que la existencia de estados intermedios bien definidos y codificados en la secuencia de aminoácidos, hace posible que el número de conformaciones posibles decrezca rápidamente y por tanto se forme el estado estructuralmente más estable en tiempos relativamente cortos [Go., 1983]. De acuerdo a este esquema, conocido como principio de plegamiento cinético, la secuencia de aminoácidos contiene en forma codificada un número pequeño de trayectorias posibles que minimizan el tiempo de plegamiento. Estas trayectorias de plegamiento (*folding pathways*) conducen a la proteína desde su estado desplegado  $D$  hasta la estructura nativa  $N$  en forma tal que minimizan el tiempo  $\tau$  requerido. En un lenguaje propio de la Física, diríamos que estas trayectorias corresponden a curvas geodésicas temporales en el espacio fase del sistema.

De la misma forma que el enfoque termodinámico busca la caracterización de los mecanismos fisicoquímicos que originan la estabilidad del estado nativo, el enfoque cinético pretende caracterizar con toda precisión las trayectorias de plegamiento que sigue una proteína en función de su masa molecular, secuencia de aminoácidos, etc y la presencia ó ausencia de barreras energéticas y de estados intermediarios que puedan influir en forma directa al tiempo  $\tau$  de plegamiento. Al igual que en el caso termodinámico, dentro del enfoque cinético existe una cantidad enorme de observaciones y resultados experimentales a ser explicados mediante una teoría consistente a nivel microscópico. A continuación enlistamos sólo algunos de los resultados más notables en esta dirección,

1. **Papel fundamental de la estructura primaria.** Las proteínas adquieren su conformación nativa en tiempos muchísimo más cortos que los planteados por la Paradoja de Levinthal [Levinthal, C.J., 1968]. Este hecho plantea la posibilidad de que existan procesos evolutivos en el diseño de las proteínas que eviten la búsqueda aleatoria de la estructura más estable. Si esta idea fuese correcta, entonces la secuencia primaria de aminoácidos sería el elemento crucial en donde se hallaría codificada esta información.
2. **Escalas de tiempo superpuestas.** Para ciertas proteínas, la estructura nativa se forma mediante un proceso secuencial. Diversos estudios muestran que las hélices  $\alpha$  y las hojas  $\beta$  que forman subestructuras estables dentro de la conformación nativa, se pliegan en escalas de tiempo distintas [Baldwin, R., 1975]. Las hélices  $\alpha$  se forman en las primeras etapas del proceso de plegamiento, con tiempos de formación de alrededor de  $t \sim 10^{-3}$  seg., en tanto que las hojas  $\beta$  requieren en general tiempos mayores de formación, aproximadamente  $t \sim 10^{-1}$  seg.

3. **Estados Intermedios.** Para proteínas de masa molecular pequeña ( $M < 20\text{kDa}$ ), el modelo cinético más aceptado para la transición  $D \rightleftharpoons N$  es básicamente un modelo de dos estados [Tanford., 1968]. Sin embargo, para proteínas de masa molecular mayor, puede existir un número importante de estados intermedios  $I_j$  de acuerdo a la reacción siguiente



Para este tipo de macromoléculas, se ha observado que la distribución de tiempos de transición  $\tau_j$  entre las diversas especies  $I_j$  presenta características propias. La transición entre la estructura nativa y los primeros estados intermedios  $I_1, I_2$ , etc.. es mucho más rápida que las transiciones cerca del estado nativo  $N$ . La presencia de fases de plegamiento rápido y lento (*fast and slow folding phases*) está íntimamente ligada con la trayectoria que sigue la proteína dentro del espacio fase de conformaciones y por tanto, puede ser analizada en conjunto con la Teoría del Paisaje Energético [Bryngelson and Wolynes., 1989].

4. **Conexión del plegamiento con la Transición Vítreo (*Glass Transition*).** Se sabe que algunos líquidos, como el glicerol o el ácido láctico, presentan la característica de poder ser sobreenfriados por debajo de su temperatura de cristalización sin formar la estructura cristalina correspondiente, como ocurre en la mayoría de los líquidos [Angell, C., 1995]. Esta transición, conocida como transición vítreo (*glass transition*), presenta rasgos termodinámicos y cinéticos similares al caso de la desnaturalización de las proteínas y permanece como uno de los retos más importantes a resolver en el campo de la física de la materia condensada, [García-Colin, L.S., 1999]. Durante la transición vítreo, los tiempos de relajación estructurales del líquido se incrementan abruptamente en alrededor de 12 órdenes de magnitud, haciendo que el líquido sobreenfriado no alcance un estado de verdadero equilibrio termodinámico y que su viscosidad sea tan grande, que permanezca en cierto sentido ‘congelado’ [Stillinger F.H., 1995]. Asimismo, la transición vítreo va acompañada siempre de un decremento significativo en el calor específico  $\Delta C_p$  entre el estado líquido y el estado vítreo. Estas dos características van íntimamente relacionadas, dado que al incrementarse los tiempos de relajación, el sistema se asemeja a un sólido, para el cual el calor específico  $C_p$  siempre tiene un valor menor que el correspondiente al estado líquido. Simultáneamente al decremento del calor específico  $\Delta C_p < 0$ , se observa un decremento en la entropía configuracional del estado vítreo  $\Delta S < 0$ . Dado que estos elementos se observan experimentalmente en la desnaturalización de las proteínas, es posible en principio considerar que ambos mecanismos tienen elementos en común. Si atendemos al hecho de que la transición vítreo es esencialmente un fenómeno dirigido por mecanismos cinéticos, ¿podríamos decir también que estos mismos rasgos presentes en la desnaturalización del estado nativo son un reflejo de la presencia de mecanismos cinéticos? No sería sorprendente hallar que la teoría completa del plegamiento y desnaturalización de proteínas o bien aquella para la transición vítreo, pudiese explicar también las similitudes termodinámicas observadas en ambos fenómenos.

## 2. Implicaciones Biomédicas del Plegamiento de las Proteínas.

No sería adecuado finalizar una descripción general del problema del plegamiento y desnaturalización de las proteínas sin abordar el aspecto complementario: el plegamiento parcial o completamente incorrecto de estas macromoléculas biológicas. La generación de la estructura primaria de una proteína se inicia con el desdoblamiento de la doble hélice de ADN por la enzima ARN polimerasa. Este proceso genera una cadena de ARN-m y se conoce como transcripción. Posterior a ello, la cadena de ARN-m se acopla con los ribosomas, los cuales en conjunción con cadenas de ARN-t, sintetizan la cadena polipeptídica en un proceso conocido como traducción. Dada la complejidad de este proceso, es posible que algunos de los pasos en los mecanismos de replicación, transcripción y traducción no se realicen en forma correcta, generando que las secuencias de aminoácidos que integrarán a la proteína puedan tener elementos faltantes o ubicados en posiciones diferentes en la cadena a como estaban diseñados a ubicarse. Este proceso, conocido como plegamiento incorrecto (*protein misfolding*), generará una estructura primaria que al plegarse en el ambiente fisicoquímico de la célula producirá una proteína con una estructura distinta a la estructura nativa. Salvo en algunos casos de robustez ante mutaciones, la proteína así generada será incapaz de realizar las funciones propias que le corresponde y tenderá a producir una patología relacionada a la actividad que la proteína debía realizar y que en estas condiciones está impedida a hacer. Estudios recientes han demostrado que para ciertos padecimientos como la *fibrosis quística* y el *enfisema pulmonar familiar*, existe una conexión clara entre el plegamiento incorrecto de una proteína y una disfunción celular específica [Dobson, C.M., 1999]. Asociado también a un plegamiento incorrecto, se encuentra toda una categoría de disfunciones celulares conocidas con el nombre genérico de *amiloidosis*, en las que los plegamientos incorrectos conducen a la formación irreversible de agregados fibrilares insolubles llamados *amiloides*, los cuales interfieren con el desarrollo de las diversas funciones celulares. Dentro de este grupo de patologías se encuentran las diversas variantes de la enfermedad de Alzheimer y Creutzfeld-Jakob [Dobson, C.M., 2001], [Chanez-Cardenas et al., 2002], entre otras. Dados estos hechos, no es de sorprender entonces que el plegamiento de las proteínas sea también un campo de profundo interés para las ciencias médicas. En esta última parte abordaremos con un poco de mayor detalle algunas de las principales disfunciones celulares asociadas con el plegamiento incorrecto de proteínas.

En esencia, el plegamiento incorrecto de una proteína y la formación de agregados fibrilares son procesos precedidos por un mecanismo de agregación. Las proteínas son sintetizadas por los ribosomas en el interior celular partiendo de la terminal N hasta la terminal C. En tanto que algunas proteínas comienzan un plegamiento parcial antes de que los ribosomas hayan sintetizado la cadena completa [Hardesty and Kramer., 2001], algunas otras lo hacen dentro del citoplasma una vez que la síntesis de los ribosomas ha sido totalmente concluída [Deuerling and Bukau., 2004]. En el transcurso del proceso de plegamiento, algunos de los estados parcialmente plegados de la proteína exhiben una tendencia natural a agregarse en el citoplasma, de forma entonces que el proceso de plegamiento *in vivo* requiere generalmente de mecanismos auxiliares de control para evitar el agregado de proteínas en estados parcialmente plegados [Young et al., 2004]. Este proceso de control es realizado en parte por las moléculas chaperonas, las cuales son capaces de reconocer estados conformacionales no nativos mediante la detección en las proteínas de una superficie hidrofóbica en lugar de una

superficie hidrofílica [Young *et al.*, 2004]. Pese a que se conocen las líneas generales de operación de las moléculas chaperonas, los detalles específicos del reconocimiento, acoplamiento y degradación de los estados no nativos por las moléculas chaperonas no ha sido comprendido en su totalidad, con excepción de algunos casos específicos como el complejo GroEL/GroEs [Guo *et al.*, 2005].

Pese a que la mayoría de los resultados conocidos sobre el mecanismo de acción del complejo GroEL/GroES se refieren a situaciones *in vitro*, este es uno de los sistemas de regulación del plegamiento mejor conocidos en la actualidad [Ellis., 2001]. Se sabe por ejemplo que esta macromolécula está constituida por 14 subunidades dispuestas en dos anillos apilados en una geometría cilíndrica. En una primera fase ocurre un acoplamiento entre las unidades de GroEL y la superficie hidrofóbica de la proteína. Una vez que la proteína ha sido encapsulada por GroEL, ocurre una segunda fase de sellamiento de la cavidad de GroEL por la subunidad GroES. La proteína permanece dentro del complejo GroEL/GroES por aproximadamente diez segundos, durante los cuales ocurre un reacción de hidrólisis por ATP seguido de cambios conformacionales en la disposición de los aminoácidos hidrofóbicos de la superficie [Ellis., 2001]. Una vez concluidos, la sección GroES se separa de GroEL para liberar a la proteína al citoplasma. Pese a la eficiencia de este sistema de regulación del plegamiento por las moléculas chaperonas, la presencia de diversos factores puede impedir que la totalidad de las proteínas en estados no nativos sean degradadas por las moléculas chaperonas. Cuando esto ocurre, las proteínas plegadas en forma incorrecta pueden alcanzar el exterior celular y proceder a la fase de agregación molecular, dado que las moléculas chaperonas son prácticamente inexistentes en el exterior de la célula [Guo *et al.*, 2005].

Una de las primeras patologías descubiertas debido a la agregación molecular es la llamada anemia falsiforme (*sickle-cell anemia*). En este tipo de padecimiento, que afecta básicamente a un sexto de la población de afroamericanos en USA [Rucknagel., 1975], los glóbulos rojos presentan un alto grado de rigidez y una forma típica en forma de hoz, que da el nombre al padecimiento y que contrasta con la forma usual de dona de las mismas [Louderback *et al.*, 1999]. Esta patología tiene su origen a nivel molecular en la mutación de un sólo aminoácido en la secuencia primaria de la hemoglobina. Tal mutación se ubica en la sexta posición de la cadena globina  $\beta$ , y corresponde a la sustitución del aminoácido *Gly* por *Val* [Ingram., 1956]. A la forma mutada de la hemoglobina normal (HbA) se le conoce como HbS. Este padecimiento resulta interesante de mencionar no sólo como un ejemplo de las consecuencias generadas por un plegamiento incorrecto, sino también porque es posible identificar perfectamente que la secuencia de aminoácidos en la estructura primaria tiene, efectivamente un papel significativo. El aminoácido *Gly* ocupa una posición estratégica dentro de la variante normal de la hemoglobina HbA y tiene un carácter polar. Cuando es sustituido por el aminoácido *Val* se induce una mutación que afecta directamente a la polaridad de la macromolécula, dado que *Val* es un aminoácido básicamente hidrofóbico. La sustitución  $Gly \rightarrow Val$  induce la presencia de una zona hidrofóbica extra en la superficie de los glóbulos rojos. Como resultado de esto, la forma HbS de la hemoglobina formará agregados que precipitarán en el interior de los glóbulos rojos. Dado que éstos deben tener una cierta flexibilidad para deformarse y transitar por los vasos capilares, la presencia de los agregados en su interior tendrá como resultado una rigidez que impide que puedan deformarse adecuadamente, con la consecuente interrupción del

proceso de oxigenación en algunos tejidos.

De acuerdo a la clasificación médica estándar, el caso de la hemoglobina HbS es un ejemplo de una patología inducida por la formación de agregados de proteínas en el interior de la celular que la alberga. Pese a que en cuestiones biológicas las reglas tienden a ser la excepción, las observaciones muestran que existe una relación entre la formación de agregados de proteínas y la mutación de las estructuras primarias. En forma general se observa que las proteínas que sufren mutaciones tienen una tendencia mayor a la formación de agregados fibrilares, dado que la mutación de la estructura primaria implica una alteración del carácter hidrofóbico-polar de la macromolécula. El Cuadro (1) muestra algunas de las principales patologías asociadas a la presencia de conglomerados de distintas proteínas.

<b>Enfermedad</b>	<b>Proteína Responsable</b>
Sickle-Cell Anemia	Deoxyhemoglobina-S
Huntington	Hungtintina
Parkinson	$\alpha$ -sinucleína
Alzheimer	A- $\beta$ , Tau
Encefalopatías Espongiformes	Proteínas Prión

CUADRO 1. Algunas de las patologías inducidas por agregados proteicos más conocidas. Tomado de [Lesk., 2004].

Dentro del amplio espectro de posibilidades para el plegamiento incorrecto de una proteína, se halla también la posibilidad de formación de agregados en el *exterior* de la célula. Cuando esto ocurre, las proteínas en estados conformacionales no nativos han escapado de los sistemas de regulación y control celular del plegamiento, que en esencia son las moléculas chaperonas y el sistema UPS (*ubiquitin proteasome system*), los cuales asisten el plegamiento y degradan a las proteínas incorrectamente plegadas, respectivamente [Lansbury., 1999]. Ya en el exterior celular, se inicia un mecanismo de agregación para las proteínas parcialmente plegadas que no fueron detectadas por los sistemas de control y regulación del plegamiento. La característica principal de este proceso es la geometría de los agregados fibrilares: las proteínas en estados conformacionales no nativos se asocian entre sí formando una estructura de hojas  $\beta$  altamente insoluble [Walsh *et al.*, 1997] (Figura 2). La formación de estos agregados insolubles y su conexión con diversas patologías (principalmente de tipo neurodegenerativo) se ha convertido en las dos últimas décadas en uno de los temas de mayor relevancia en las ciencias biomédicas. Algunas estimaciones indican que hasta un 50% de todos los padecimientos en nuestra especie tienen un origen en un mecanismo de plegamiento incorrecto y agregación [Bradbury., 2003].

Las patologías resultado de un proceso de agregación extra o intra celular se conocen como amiloidosis y entre ellas se hallan las principales patologías de índole neurodegenerativa, como es el caso de los padecimientos de Alzheimer, de Huntington, de Parkinson, de Gaucher y las encefalopatías espongiformes [Tapan and Subhankar., 2006]. Aunque cada uno de estos padecimientos tiene su sintomatología propia, una de las características generales de todos ellos es la pérdida simultánea de neuronas en los sistemas sensorial, cognoscitivo y motor. En el caso del padecimiento de Alzheimer,



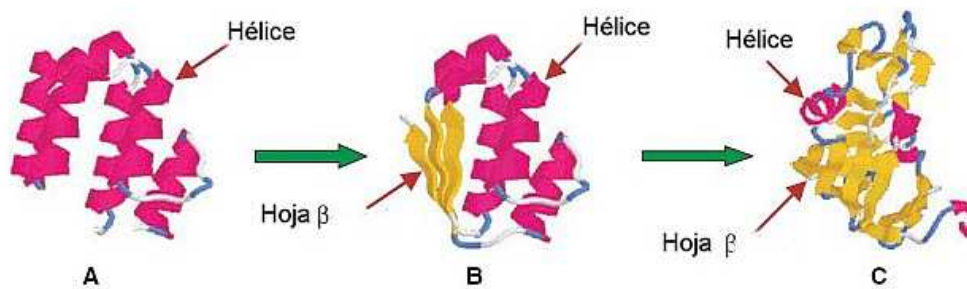


FIGURA 2. El proceso de formación de agregados fibrilares insolubles (amiloides) consiste esencialmente en la transformación de la estructura helicoidal  $\alpha$  de una proteína en una estructura de hoja  $\beta$ . En (A) una proteína con estructura nativa de hélices  $\alpha$ , en (B) los procesos de plegamiento incorrecto generan la conversión de parte de la estructura  $\alpha$  en una topología  $\beta$ . En (C) la proteína ha sido transformada en su estructura nativa hacia una conformación de tipo  $\beta$ . Adaptado de [Tapan and Subhankar., 2006].

se observa el depósito de agregados de la proteína  $A\beta$  en la masa encefálica y la generación de cadenas neurofibrilares anudadas de la proteína Tau (*Tau-neurofibrillary tangled chains*) [Games *et al.*, 1995]. En contraposición al padecimiento de Alzheimer, estudios recientes indican que en el padecimiento de Parkinson los agregados fibrilares se ubican en el espacio intracelular [Forloni *et al.*, 2002]. En referencia a la sintomatología, la enfermedad de Parkinson se caracteriza por una rigidez muscular e inestabilidad en el control de los movimientos musculares más finos. En un nivel de mayor especificidad, el diagnóstico definitivo de un padecimiento de Parkinson se establece al detectar cúmulos de la proteína  $\alpha$  sinucleína en el interior del citoplasma de las neuronas [Agorogiannis *et al.*, 2004].

Finalmente se denomina con el término de enfermedades tipo prión a una variedad amplia de patologías inducidas por la mutación conformacional de las proteínas conocidas como priones,  $PrP^C$ . En su forma normal, esta proteína se encuentra ubicada en la superficie celular de las neuronas y se piensa que juega un papel importante en la recepción de ciertos neurotransmisores. Aún cuando el mecanismo que induce la mutación conformacional no es claro, una vez que éste ha ocurrido, la proteína prión transforma las hélices  $\alpha$  de su estructura nativa en hojas  $\beta$ . Esta variante, conocida como  $PrP^{Sc}$ , es capaz de formar agregados insolubles similares a los descritos anteriormente. Asimismo, las proteínas prión que han mutado a una configuración predominante de hojas  $\beta$  actúan como catalizadores en la producción de más proteína  $PrP^{Sc}$  a partir de la variante normal, actuando entonces como un agente infeccioso [Cohen *et al.*, 1994]. Dentro del grupo principal de patologías conformacionales descritas por el mecanismo anterior se hallan las encefalopatías espongiiformes. El nombre proviene de la estructura que adquieren los cerebros de los individuos infectados, el cual se asemeja a una esponja en donde los espacios vacíos corresponden a las neuronas muertas [Lesk., 2004].

A lo largo de este capítulo hemos presentado un panorama general del tema central de esta tesis doctoral: el plegamiento de las proteínas. Sin pretender desde

luego hacer una revisión exhaustiva, que además no correspondería a la intención primordial, hemos querido destacar la relevancia y actualidad del problema, tanto en sus aspectos fundamentales como en sus consecuencias de mayor impacto en áreas biomédicas. En el siguiente capítulo reduciremos considerablemente el espectro de temas a abordar y al mismo tiempo incrementaremos el nivel de detalle. El Capítulo II estará dedicado a establecer y caracterizar a las proteínas partiendo desde el nivel más básico. Iniciaremos describiendo en términos generales el proceso de síntesis de la cadena secuenciada de aminoácidos a partir de la información codificada y transcrita por el ADN y el ARN. Posteriormente abordaremos el análisis de la estabilidad de la estructura secundaria (hélices  $\alpha$  y hojas  $\beta$ ) y la agrupación de conjuntos de hélices  $\alpha$  y hojas  $\beta$  en unidades funcionales por sí mismas, llamadas dominios. Finalmente concluiremos describiendo las principales características del estado nativo en términos de la estructura primaria, secundaria y terciaria.

## CAPÍTULO 2

# Estructura e Interacciones en las Proteínas.

### 1. Síntesis de una proteína.

La primera etapa en la síntesis de una proteína es la transcripción de la información codificada por los pares de nucleótidos de la molécula de ADN a una cadena de ARN-m la cual se acoplará posteriormente a los ribosomas para iniciar la síntesis de la cadena polipeptídica [Saenger., 1988]. Desde un punto de vista evolutivo, la existencia de un ácido nucleico intermediario entre la molécula de ADN y los ribosomas representa una ventaja, dado que en estas condiciones la doble hélice de ADN permanece expuesta ('abierta') el menor tiempo posible, reduciendo así las posibilidades de un daño a la secuencia de nucleótidos [Gaeta., 1999]. El proceso de transcripción que da origen a las proteínas se lleva a cabo mediante el acoplamiento de la enzima ARN-polimerasa II (ARNP II) al ADN. El ARNP II desdobra a la molécula de ADN abriendo un región de transcripción (*transcription bubble*) que abarca entre 15 – 20 pares de nucleótidos y se desplaza a lo largo de una de las hélices de ADN de forma tal que el tamaño de la región que está siendo transcrita permanece aproximadamente constante. En la fase activa de la transcripción, la enzima ARNP II se desplaza con una rapidez entre  $10 - 10^2$  pares de nucleótidos por segundo y dado que cada par de nucleótidos está enlazado por dos o tres puentes de hidrogeno, la energía involucrada en el mecanismo de transcripción es del orden de  $10^3$  puentes de hidrógeno [Gaeta., 1999] (Figura 1). Conforme la molécula de ARNP-II se desplaza a lo largo de una de las hélices de ADN durante el proceso de transcripción, se genera una cadena del ácido ribonucleico ARN conocida como ARN mensajero (ARN-m). Esta cadena contiene una 'copia' de la información contenida en el ADN y se acoplará en una siguiente fase con los ribosomas y el ácido ribonucleico de transferencia (ARN-t) para ensamblar la cadena polipeptídica que constituye la estructura primaria de una proteína [Watson., 1987].

Al acoplarse la cadena de ANR-m con los ribosomas para iniciar la traducción del código se requiere un mecanismo de diferencie cada uno de los aminoácidos que integran la estructura primaria. Para comprender cómo ocurre este proceso conviene revisar brevemente el almacenamiento de la información genética en la doble hélice de ADN. Cada una de las hélices del ADN está formada por cuatro tipos distintos de nucleótidos: adenina (A), citosina (C), guanina (G) y timina (T). La formación de la doble hélice de ADN ocurre debido a la formación de puentes de hidrógeno entre pares de nucleótidos,  $A - T$ ,  $G - C$ , llamados en forma más específica pares de Watson-Crick [Cantor and Schimmel. , 1980]. Durante el proceso de transcripción, la secuencia de nucleótidos contenida en una sección del ADN es transcrita a la cadena de ARN-m sustituyendo el nucleótido timina  $T$  por el nucleótido uracil  $U$ . De esta forma, la composición del ARN-m es una secuencia de nucleótidos  $A, U, G, C$ .

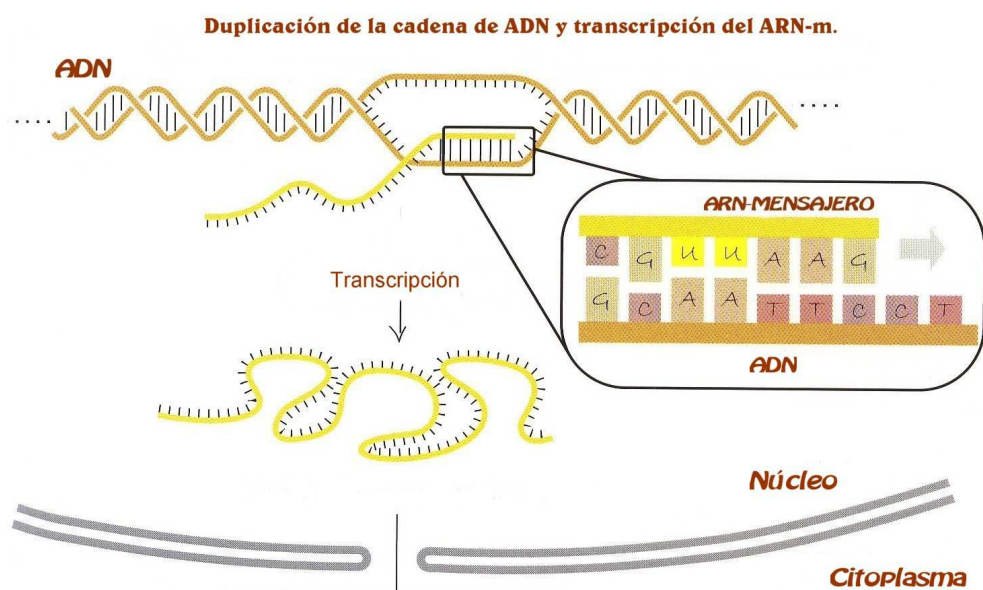


FIGURA 1. Esquematación del proceso de transcripción del ADN por la enzima ARNP-II, el cual genera una cadena de ARN-m. El ARN-m se contiene la información necesaria para la síntesis de la estructura primaria de una proteína, lo cual ocurre durante el proceso de traducción realizado por los ribosomas. Adaptado de [Voet., 1995].

Dado que tenemos solamente cuatro bases (nucleótidos) diferentes y solamente veinte aminoácidos distintos<sup>1</sup>, la primera consecuencia inmediata de este hecho es que se requiere grupos de nucleótidos para poder codificar a un aminoácido en particular. Estos grupos de bases se denominan **codones**. Dado que los codones conformados por sólo dos bases (dupletas) resultarían insuficientes para generar los veinte aminoácidos ( $4^2 = 16 \Rightarrow$  dobletes distintas), se requiere al menos de codones integrados por tripletas de nucleótidos. En este caso, tenemos  $4^3 = 64$  tripletas distintas, lo cual resulta en que código *degenerado* en un sentido matemático; existen tripletas que no codificarán para ningún aminoácido y aminoácidos codificados por más de una tripleta.

Alrededor de 1960, los trabajos de Francis Crick y Sydney Brenner [Crick F.H. et al., 1961] establecieron mediante una serie de experimentos con cadenas de ADN mutantes que el código genético para la síntesis de proteínas está basado efectivamente, en codones formados por tres bases de nucleótidos. El siguiente paso fundamental en el desciframiento del código genético fue realizado en 1961 por Marshall Nirenberg y Heinrich Matthaei [Nirenberg and Matthaei., 1961], al demostrar que la tripleta formada por tres bases uracil (*U, U, U*) codifica al aminoácido *Fenilalanina (Phe)* [Nirenberg and Matthaei., 1961]. Asimismo, experimentos realizados con ARN mutantes constituidos por un sólo nucleótido (poly(A) y poly(C)), condujeron a dos elementos más en el código genético. La cadena poly(A) genera cadenas del aminoácido *Lysina*,

<sup>1</sup>La afirmación anterior se refiere a aminoácidos presentes en proteínas. Es posible de hecho, generar una molécula con un grupo carboxilo y un grupo amino en forma sintética y que no forme parte de una proteína natural.

<b>Aminoácido</b>	<b>Codón de ARN-m</b>
Alanina (Al)	GCA GCC GCG GCU
Arginina (Ar)	AGA AGG CGA CGC CGG CGU
Asparagina (Asn)	AAC AAU
Acido Aspartico (Asp)	GAC GAU
Cisteína (Cis)	UGC UGU
Acido Glutámico (Glu)	GAA GAG
Glutamina (Gln)	CAA CAG
Glicina (Gly)	GGA GGC GGG GGU
Histidina (His)	CAC CAU
Isoleucina (Ile)	AUA AUC AUU
Leucina (Leu)	UUA UUG CUA CUC CUG CUU
Lisina (Lys)	AAA AAG
Metionina (Met)	AUG
Fenilalanina (Phe)	UUC UUU
Prolina (Pro)	CCA CCC CCG CCU
Serina (Ser)	AGC AGU UCA UCC UCG UCU
Treonina (Thr)	ACA ACC ACG ACU
Triptofán (Trp)	UGG
Tirosina (Tyr)	UAC UAU
Valina (Val)	GUA GUC GUG GUU
Codones que no codifican	UAA UAG UGA

CUADRO 1. Tripletas de nucleótidos de ARN-m que codifican cada uno e los veinte aminoácidos proteinógenos conocidos durante el proceso de traducción. [Watson, J.D., 2003]

en tanto que en presencia de un ARN-m formado por cadenas de poly(C), los ribosomas sintetizan solamente cadenas del aminoácido *Pro*.

Los experimentos anteriores y otros más que los siguieron establecieron el código completo de codones que codifica a cada uno de los veinte aminoácidos conocidos, así como los codones que codifican el inicio y terminación de una síntesis en particular. El Cuadro (1) muestra las tripletas de nucleótidos que codifica a cada uno de los veinte aminoácidos conocidos en proteínas. En términos generales, podemos decir que el código representado en el Cuadro (1) tiene las siguientes características.

- El código genético es un código que presenta un alto nivel de degeneración en términos matemáticos. Existen tres aminoácidos que son codificados por seis codones cada uno y la mayoría restante tiene más de un codón asociado. A los codones que representan al mismo aminoácido se les llama *sinónimos* por lo que entonces por ejemplo los codones  $(C, G, U)$  y  $(C, G, G)$ ,  $(C, G, C)$  son sinónimos.
- Existen solamente tres codones que indican la terminación de un cadena polipeptídica particular. Estos codones corresponden a las tripletas  $(U, A, G)$ ,

$(U, A, A)$  y  $(U, G, A)$ . De la misma forma, los codones  $(A, U, G), (G, U, G)$  indican el inicio de una cadena en particular [Cantor and Schimmel. , 1980].

- Aunque no es en sí una regla, por lo general ocurre que cuando el segundo nucleótido de un codón consiste en una base pirimidina, el aminoácido sintetizado tiene un carácter hidrofóbico. Para aquellos codones cuyo segundo elemento es una base purina, el aminoácido resultante tiene un perfil polar.

A través de la lectura secuencial del código anterior en el acoplamiento de la cadena de ARN-m con los ribosomas, se inicia propiamente la síntesis de la estructura primaria de todas las proteínas, conformada por una secuencia de  $K$  aminoácidos unidos covalentemente entre sí. La estructura primaria de una proteína es entonces una cadena heteropolimérica lineal, la cual en principio podría adquirir un número considerable de configuraciones espaciales mediante flexiones y rotaciones de las monómeros que lo integran. Una vez que la proteína ha sido sintetizada por los ribosomas, entra en contacto con un medio multicomponente (enzimas, coenzimas, otras proteínas) y en condiciones fisicoquímicas particulares (presión, temperatura,  $pH$ , etc) las cuales dirigirán una serie de cambios conformacionales hacia el estado nativo. De los detalles de este proceso nos ocuparemos en los siguientes párrafos, no sin antes discutir brevemente algunas de las propiedades físicas y químicas de los monómeros que integran la estructura primaria: los aminoácidos.

## 2. Los aminoácidos: Unidades básicas de la cadena polipeptídica.

De acuerdo a la información presentada en el Cuadro (1) se conocen en la Naturaleza 20 aminoácidos distintos. A pesar de las diferencias que existen entre ellos, todos comparten una misma estructura química genérica [Finkelstein and Pitysin, 2002]. Cada aminoácido está conformado por un átomo central de carbono, denotado por  $C_\alpha$ , ligado covalentemente a un grupo amino  $NH_2$ , a un grupo carboxilo  $COOH$ , a un átomo de hidrógeno  $H$  y a una molécula llamada cadena lateral  $R$ , la cual es distinta para cada aminoácido y que puede variar enormemente en cuanto a su composición química. A fin de integrar la cadena polipeptídica, los aminoácidos se

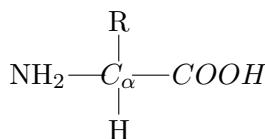


FIGURA 2. Estructura general de un aminoácido. El átomo central de carbono se conoce como  $C_\alpha$  y está ligado covalentemente a los grupos funcionales amino  $NH_2$ , carboxilo  $COOH$ , un átomo de hidrógeno y una cadena lateral  $R$  cuya composición es distinta para cada aminoácido.

acoplan entre sí mediante enlaces peptídicos que unen al grupo amino del aminoácido  $n - 1$  con el grupo carboxilo del aminoácido  $n$ . Este proceso de polimerización de la estructura primaria implica también la expulsión de moléculas de agua en la formación de cada enlace peptídico (Figura 3). El enlace  $COOH - NH_2$  así formado se conoce como *enlace peptídico* (*peptide bond*) y es de naturaleza completamente covalente. La

mayoría de las proteínas consisten de una o más cadenas polipeptídicas, cada una de las cuales está formada entre 40 ~ 4000 aminoácidos. Tomando en cuenta que la masa molecular promedio de un aminoácido es de 110 Da, la masa molecular de una proteína oscila entre 4 ~ 440 kDa.

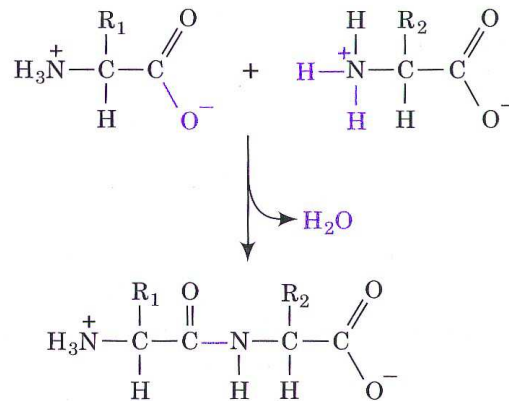


FIGURA 3. La síntesis de la cadena polipeptídica es una reacción de polimerización de los aminoácidos. El enlace entre los grupos carboxilo  $\text{COOH}$  del aminoácido  $n - 1$  y el grupo amino  $\text{NH}_2$  del aminoácido  $n$  implica la expulsión de una molécula de agua. Tomado de [Voet., 1995]

Dado que existen veinte variantes para los monómeros que integran a la estructura primaria, existe un número considerable de posibilidades para el diseño de una proteína en particular. Por ejemplo, si consideramos solamente a las cadenas dipeptídicas, cada una de las veinte variantes de aminoácidos tiene veinte posibilidades a su vez para el segundo monómero, por lo que el número de posibilidades distintas para proteínas dipéptidas es de  $20^2 = 400$ . Dado que las proteínas más pequeñas están formadas por cadenas de 100 aminoácidos, existen  $20^{100}$  distintas posibilidades para la formación de una estructura primaria en particular, un número mucho mayor que el número estimado de átomos en el Universo. Evidentemente no existe tal variedad de proteínas y la existencia de un número limitado de ellas es, de acuerdo a la opinión generalizada en el campo, uno de los resultados más palpables del proceso de selección natural darwiniano [Vickery., 1972].

La naturaleza de la cadena lateral  $R$  (conocido también en inglés como *side chain*) es distinta para cada aminoácido y puede variar en sus características electrostáticas y fisicoquímicas, dando así características distintivas para cada uno de los 20 aminoácidos conocidos en la Naturaleza. De hecho, las propiedades electrostáticas y de polaridad de las cadenas laterales constituyen la base para una de las clasificaciones más comunes de los aminoácidos. Con base en esta clasificación, se tienen tres grupos principales: los aminoácidos cuya cadena lateral es no polar, los que tienen una cadena lateral polar con carga electrostática y aquellos con una cadena lateral polar sin carga electrostática. El aminoácido con la cadena lateral más simple corresponde a aquel en donde  $R$  es un átomo de hidrógeno, este aminoácido se como *Glicina*

(*Gly*). Dos de los aminoácidos con cadenas laterales más complejas corresponden a los aminoácidos *Trp* y *Arg*, los cuales tienen grupos *R* formados por 18 átomos, entre los que se incluyen un anillo aromático en el caso del aminoácido *Trp*. La Figura (4) muestra la composición de las cadenas laterales *R* para cada uno de los aminoácidos conocidos.

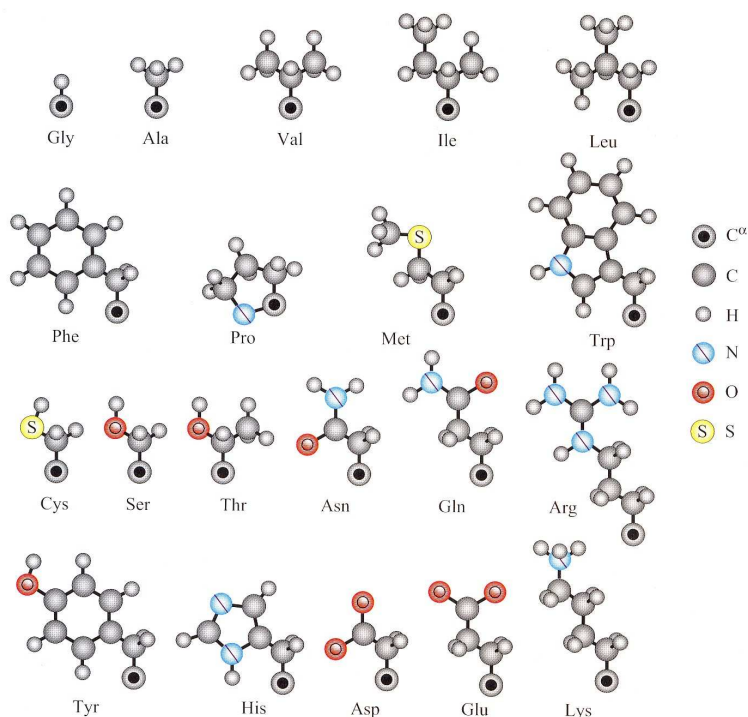


FIGURA 4. Estructura atómica de las cadenas laterales *R* para los 20 aminoácidos presentes en proteínas. El gráfico muestra la composición química de las cadenas *R* en orden creciente de complejidad. Tomado de [Finkelstein and Ptytsin, 2002].

De acuerdo a una de las clasificaciones más utilizadas [Voet., 1995], se tienen 9 aminoácidos no polares, 6 aminoácidos polares sin carga eléctrica y 5 aminoácidos pertenecen a la categoría de polares con carga eléctrica. El Cuadro (2) presenta cuáles aminoácidos pertenecen a cada categoría. Es importante mencionar que la asignación de un aminoácido a una categoría particular no es criterio unívocamente establecido. El aminoácido *Trp*, el cual posee un anillo aromático, es uno de los mejores ejemplos en este sentido. En ocasiones es clasificado como no polar y en otras como polar sin carga eléctrica.

Uno de los aspectos más sorprendentes de la polimerización de la estructura primaria se refiere a la naturaleza bidimensional del enlace peptídico. Este enlace, como discutimos anteriormente, ocurre entre los grupos amino y carboxilo de aminoácidos adyacentes y confiere a la estructura primaria un aspecto similar a una serie de planos unidos entre sí [Branden and Tooze., 1991]. Esta característica fue observada inicialmente por Pauling *et al* [Pauling and Corey., 1951], [Pauling *et al.*, 1951b] y tiene un



Clasificación	Código del Aminoácido
No polar	Gly, Al, Val, Phe, Pro, Met, Ile, Leu, Trp
Polar sin carga	Ser, Thr, Tyr, Cis, Asn, Gln,
Cargados	Asp, Glu, Lys, Arg, Hys

CUADRO 2. Clasificación estándar para los veinte aminoácidos conocidos, de acuerdo a las propiedades electrostáticas de la cadena lateral  $R$  respecto del solvente  $H_2O$ . [Voet., 1995]

origen completamente cuántico. El fenómeno que da origen a la geometría plana del enlace peptídico es la hibridación de los orbitales  $s$  y  $p$  de los átomos  $C$  y  $N$  para generar el orbital híbrido  $sp^2$ . En este proceso, el orbital  $s$  cuya geometría es esférica, se une con dos orbitales  $p$  con geometría en forma de ‘pera’, para generar tres orbitales  $sp^2$  que se extienden desde el núcleo de un átomo  $C$  hasta el núcleo del átomo  $N$ . Debido a que las nubes electrónicas del orbital híbrido  $sp^2$  están deslocalizadas, es decir cubren a ambos átomos, se forma un enlace covalente entre los átomos  $C$  y  $N$  por cada orbital. Para el caso de la hibridación  $sp^2$  entonces, los átomos  $C$  y  $N$  quedan ligados a través de tres enlaces covalentes [Finkelstein and Ptytsin, 2002]. Además de la presencia de los tres orbitales  $sp^2$ , se tiene un orbital más de tipo  $p$  deslocalizado entre los átomos  $C$  y  $N$ . Si bien este orbital no participa directamente en la unión covalente porque no se halla deslocalizado, sí restringe las posibilidades de orientación de los grupos amino y carboxilo. Puesto que éste último orbital es perpendicular al plano de los orbitales  $sp^2$ , el efecto conjunto de los orbitales  $sp^2 + p$  tiene como resultado fijar las posiciones de los átomos  $C$ ,  $O$ ,  $N$  y  $H$  de los grupos carboxilo y amino respectivamente, confiriendo entonces un estructura rígida a cada péptido (Figura 5).

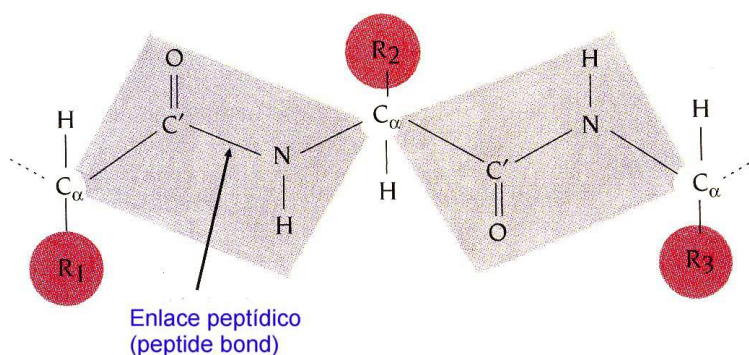


FIGURA 5. La hibridación de un orbital  $s$  y dos orbitales  $p$  genera tres orbitales  $sp^2$ , los cuales confieren un carácter bidimensional a la unidad peptídica. En particular bloquean la rotación del grupo  $NH$  alrededor del eje  $C - N$  y la rotación de  $C - O$  alrededor del mismo eje, resultando en una estructura plana para el péptido (mostrada en gris). Adaptado de [Branden and Tooze., 1991]

Mencionamos al principio del párrafo anterior que la bidimensional del enlace peptídico sólo podría comprenderse adecuadamente dentro del marco de la mecánica cuántica. La razón de esto tiene dos partes. La primera la hemos expuesto anteriormente y está directamente relacionada con la geometría de los orbitales atómicos predichos por la teoría cuántica: la función de onda  $\Psi$  tiene simetría esférica para un estado energético cuyo número cuántico es  $l = 0$ , en tanto que la geometría en forma de *pera* para los orbitales  $p$  es un reflejo de la asimetría de la función de onda  $\Psi$  cuando el sistema tiene un número cuántico  $l = 1$ . El esquema de hibridación de los orbitales para formar un orbital mezclado es también parte de la teoría cuántica y como vimos anteriormente es el punto central en la comprensión de la estructura del enlace peptídico [Finkelstein and Pitytsin, 2002]. La segunda parte del argumento está asociada con el concepto de *deslocalización* de la nube electrónica. Este concepto, también circunscrito a la teoría cuántica, implica que un electrón tiene cierta probabilidad distinta de cero de hallarse en la región de influencia de un átomo o de otro. De hecho, la deslocalización de un electrón es una consecuencia del Principio de incertidumbre de Heisenberg. Para verlo, consideremos el Principio de Incertidumbre de Heisenberg escrito de la forma siguiente [Finkelstein and Pitytsin, 2002]

$$(2.1) \quad \Delta p \Delta x \sim \hbar,$$

en donde  $\Delta p$  representa la incertidumbre en el momento de la partícula,  $\Delta x$  la incertidumbre en la posición de la misma y  $\hbar$  la constante de Planck dividida por  $2\pi$ . Si consideramos que  $\Delta p \sim mv$  y que la energía cinética de una partícula está dada por  $E = (mv^2)/2$ , entonces  $E$  puede expresarse de la siguiente forma,

$$(2.2) \quad E \sim \frac{\hbar^2}{m\Delta x^2}.$$

La ecuación (2.2) muestra entonces una relación inversamente proporcional entre la energía  $E$  de una partícula y la incertidumbre en su posición  $\Delta x$ . Cuando una partícula cuántica, como un electrón, se halla deslocalizado, la incertidumbre en su posición  $\Delta x$  se incrementa, minimizando su energía  $E$ . Por tanto, la conformación más estable para la partícula corresponde al estado cuántico deslocalizado, es decir a la formación de un enlace covalente.

Ahora bien, los enlaces peptídicos que unen a la estructura primaria son susceptibles de exhibir ciertas vibraciones, inducidas ya sea por fluctuaciones térmicas del medio circundante o bien por factores externos. ¿De qué orden de magnitud es la energía requerida para excitar estos modos de vibración? Para un enlace covalente típico en una proteína, la frecuencia de vibración presenta una distribución ubicada en la región infrarroja del espectro (IR) con un máximo ubicado generalmente alrededor de  $\nu \sim 7 \times 10^{13}$  Hz [Ma *et al.*, 2000]. Por ejemplo, para el caso específico del enlace  $C - H$ , la frecuencia de oscilación del modo principal está ubicada alrededor de  $\nu = 6 \times 10^{13}$  Hz; en tanto que para enlaces entre átomos más pesados o radicales completos, como en el caso de  $CH_3-CH_3$ , la frecuencia de oscilación tiene un máximo centrado en  $\nu = 2 \times 10^{13}$  Hz [Finkelstein and Pitytsin, 2002]. En cualquiera de estos casos y a temperaturas ubicadas entre  $20^\circ C$  y  $40^\circ C$ , la energía inducida por las fluctuaciones térmicas del solvente es un orden de magnitud menor que la energía necesaria para excitar los modos de oscilación de los enlaces covalentes, por lo que en estas condiciones los enlaces peptídicos de una proteína permanecen estructuralmente rígidos. Al ser liberada al medio celular, la estructura primaria cambia de conformación, pero no debido a la vibración de los enlaces covalentes, que como hemos visto

permanecen rígidos en estas condiciones. ¿Cuál es entonces el origen de las múltiples conformaciones que puede adquirir una proteína? La respuesta se halla en la rotación tanto de las cadenas laterales  $R$  como de las unidades peptídicas en sí. Recordemos que el análisis de los párrafos anteriores muestra que la unidad peptídica es plana y rígida, lo cual no impide que pueda rotar como una entidad completa respecto de un eje, ya sea del eje  $C_\alpha - N$  o bien del eje  $C_\alpha - C$ . El análisis de cómo ocurre este proceso, los mecanismos que lo originan y las estructuras resultantes constituyen la siguiente parte de esta disgresión.

### 3. Estructura secundaria: Hélices $\alpha$ y Hojas $\beta$ .

Existen diversas acepciones para el término **estructura secundaria** de una proteína. En este trabajo llamaremos estructura secundaria de una proteína a la ubicación geométrica del esqueleto (*backbone*) de la cadena polipeptídica. En general, la descripción anterior se refiere a la estructura general adoptada por los átomos centrales  $C_\alpha$  de la proteína [Cantor and Schimmel, , 1980].

El análisis por métodos físicos de difracción de rayos X de la estructura de cadenas dipéptidas o tripéptidas se remonta a los trabajos pioneros de L. Pauling y R. Corey de 1930 [Pauling and Corey., 1951], [Pauling *et al.*, 1951b]. Uno de los resultados más importantes de estos trabajos ha sido mencionado anteriormente y se refería a la bidimensionalidad de las unidades peptídicas. Derivado de estos trabajos se halla también el hecho de que, con unas pocas excepciones, la mayoría de las cadenas polipeptídicas de las proteínas asumen la llamada configuración *trans* en donde los átomos centrales de carbono  $C_\alpha$  de dos aminoácidos consecutivos se hallan opuestos entre sí. La conformación *cis* de la cadena implica que los átomos  $C_\alpha$  adyacentes se hallen del mismo lado respecto del enlace peptídico que los une y esto necesariamente requiere de una mayor energía puesto que las nubes electrónicas de los átomos  $C_\alpha$  tenderán a repelerse en esta configuración debido a su proximidad. La configuración *trans* es entonces mucho más estable por una diferencia energética de  $\Delta E = 8$  kJ/mol. La figura (6) muestra la ubicación de los átomos  $C_\alpha$ , de los grupos amino y carboxilo en una cadena formada por solamente dos péptidos en una configuración *trans*. Asimismo, muestra los ángulos calculados y observados experimentalmente [Marsh and Donohue., 1967] para la geometría de esta cadena.

La repetición de la Figura (6) un número  $K - 1$  de veces mostraría el aspecto general de una estructura primaria conformada por  $K$  péptidos. El enlace peptídico es plano, lo que no implica que no pueda rotar alrededor de dos ejes: el eje que forma el enlace  $C_\alpha^{(1)} - C$  y el eje que forma el enlace  $C_\alpha^{(2)} - N$ , en donde  $C_\alpha^{(1)}$  indica el átomo central del aminoácido 1 en tanto que  $C_\alpha^{(2)}$  indica el átomo central de carbono del segundo. Son justamente estas rotaciones las que conferirán la flexibilidad necesaria para que los grupos residuales entren en contacto y puedan establecerse interacciones de mayor orden en la cadena, lo que finalmente traerá como consecuencia la estabilidad de la estructura secundaria. Para seguir la notación estándar en la literatura, llamaremos  $\phi$  al ángulo de rotación alrededor del eje  $C_\alpha - N$  y  $\psi$  al ángulo de rotación respecto del eje  $C_\alpha - C$ . Para cada enlace peptídico es posible entonces, especificar un par de ángulos  $(\phi, \psi)$  que definen la orientación exacta del grupo peptídico en el espacio. Al par de ángulos  $(\phi, \psi)$  se les conocen en la literatura como **ángulos**

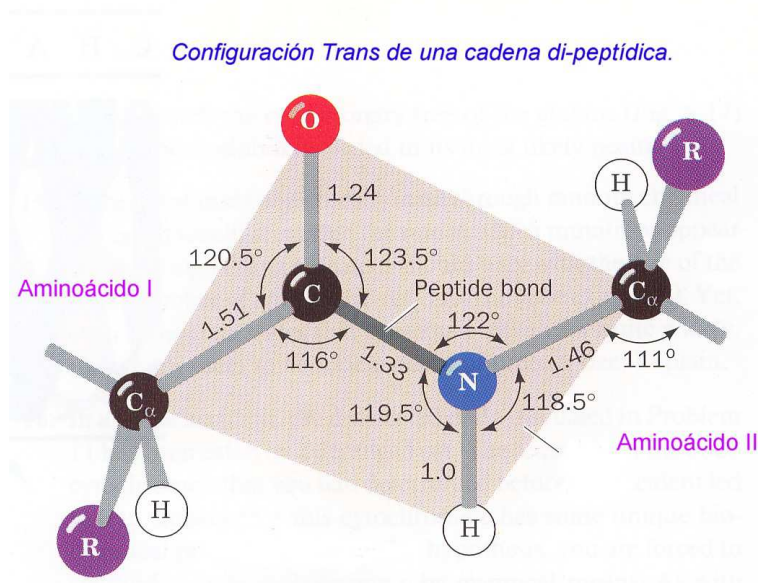


FIGURA 6. Esquema de una cadena di-peptídica de acuerdo a los resultados experimentales de [Marsh and Donohue., 1967]. Se muestran la posición de cada uno de los átomos para cada aminoácido y los ángulos entre los diversos enlaces covalentes. La estructura plana del enlace peptídico se muestra en color gris.

**diédricos.** Dado que para cada uno de los  $K - 1$  enlaces peptídicos correspondientes a una cadena de  $K$  aminoácidos existe un par de ángulos diédricos  $(\phi_j, \psi_j)$ , es posible especificar en forma precisa la geometría de la cadena a través del conjunto  $\Upsilon = \{(\phi_j, \psi_j), j = 1, \dots, K\}$ . Por definición, se asigna el valor  $(\phi = 180^\circ, \psi = 180^\circ)$  a la conformación completamente extendida, en donde todas las unidades peptídicas yacen en el mismo plano [Creighton, 1984].

La configuración completamente extendida de la cadena polipeptídica corresponde a la configuración más simple dentro del conjunto de ángulos diédricos, dado que está asociada con el valor constante  $(\phi = 180^\circ, \psi = 180^\circ)$  para todos los pares de ángulos. Sin embargo, existen evidentemente otras posibilidades para los ángulos diédricos y, consecuentemente, para la configuración de la cadena. Si bien como veremos más adelante, existen numerosas posibilidades para el conjunto  $\Upsilon$  de ángulos diédricos, algunas otras están restringidas desde un principio por las restricciones estéricas provenientes de las fuerzas de Van der Waals entre pares de átomos [Finkelstein and Ptysin, 2002]. Recordemos que un par de átomos no pueden acercarse más allá de una distancia crítica determinado por el potencial de interacción entre ellas. Visto desde esta perspectiva, no todas las configuraciones geométricas de la cadena tendrán igual probabilidad de ocurrencia. Debido a las restricciones del tipo van der Waals, muchas de las configuraciones quedan bloqueadas desde un principio. A través del cálculo teórico de las distancias mínimas de acercamiento entre pares de átomos es posible calcular también teóricamente el conjunto de ángulos diédricos para una cadena polipeptídica en particular. En forma paralela a lo anterior, G. Ramachandran desarrolló una representación gráfica de las posibles conformaciones de una cadena en

términos de los ángulos diédricos en un espacio bidimensional, conocido ahora como mapa de Ramachandran [Ramachandran et al., 1968].

La Figura (7) muestra el espacio configuracional de una cadena de poly(L)-alanina en términos de los ángulos  $\phi$  y  $\psi$  calculado a partir de las distancias de aproximación máximas de Van der Waals entre pares de átomos [Flory, P.G., 1969]. Las regiones sombreadas representan las zonas accesibles en tanto que las zonas blancas son subespacios configuracionales prohibidos debido a restricciones estéricas. Pese a que la Figura (7) fue calculada específicamente para una cadena homopeptídica de Alanina, numerosos cálculos y mediciones experimentales para otras cadenas hetero y homopoliméricas muestran básicamente la misma estructura. En este mapa de Ramachandran, que podríamos considerar entonces como un caso típico, se muestra la existencia de tres zonas configuracionales accesibles a la proteína. De hecho, *todas* las estructuras secundarias conocidas en proteínas caen en alguna de las tres zonas descritas anteriormente. La identificación y caracterización de cada una de las zonas presentes en el mapa de Ramachandran da origen a los diferentes tipos de estructura secundaria hasta hoy conocidas: la hélice  $\alpha$  derecha, que se halla ubicada en la región ( $-30^\circ < \phi < -60^\circ, \psi \sim -45^\circ$ ), la hélice  $\alpha$  izquierda localizada alrededor de ( $30^\circ < \phi < 60^\circ, \psi \sim 45^\circ$ ), la hojas  $\beta$  paralela y antiparalela, las cuales yacen en la región ( $90^\circ < \phi < 150^\circ, -60^\circ < \psi < -150^\circ$ ) [Voet., 1995].

Basados en cálculos teóricos y respaldados por mediciones experimentales, los mapas de Ramachandran nos proveen de información acerca de la posibilidades conformacionales para la estructura secundaria. Tal como mencionamos anteriormente, la gran mayoría de las cadenas polipeptídicas presentan tres zonas conformacionales, a las cuales se ha identificado con el nombre de hélice  $\alpha$  izquierda, hélice  $\alpha$  derecha y hojas  $\beta$ , las cuales tienen dos variantes posibles: la hoja paralela y la hoja antiparalela. A continuación analizaremos brevemente las características de cada una de ellas.

La hélice  $\alpha$  tanto en su variante izquierda como derecha, es la primera estructura secundaria descubierta y analizada. Alrededor de 1951, L. Pauling y H.B. Corey extendiendo argumentos de minimización energética previamente utilizados en el análisis de la estructura de sistemas cristalinos, sugirieron la existencia de estructuras helicoidales estables en macromoléculas biológicas [Pauling and Corey., 1951]. [Pauling et al., 1951b]. Los resultados de esta investigación, publicados en el artículo *The structure of proteins: Two helical configurations of the polypeptide chain*, utilizan como eje principal de argumentación el así llamado *Principio de Equivalencia de Aminocidos*. Este principio afirma que, dado que todos los aminoácidos comparten una misma estructura química, exceptuando a su cadena lateral, existe entonces un conjunto de operaciones formales de rotación y traslación que preservan la simetría de la cadena. Dado que la hélice es una curva matemática que preserva la simetría bajo una rotación y traslación conjunta, Pauling y Corey concluyen que la hélice debía formar parte entonces de las posibles conformaciones estables para cadenas de aminoácidos. La relevancia del trabajo de Pauling y Corey al predecir la existencia de al menos dos configuraciones helicoidales posibles para cadenas polipeptídicas se pondrá de manifiesto unos pocos años más tarde. En 1953, Watson y Crick [Watson and Crick., 1953] demuestran basados en experimentos de difracción de rayos X, que la estructura de la molécula del ADN es, efectivamente, helicoidal. Unos años más tarde, en

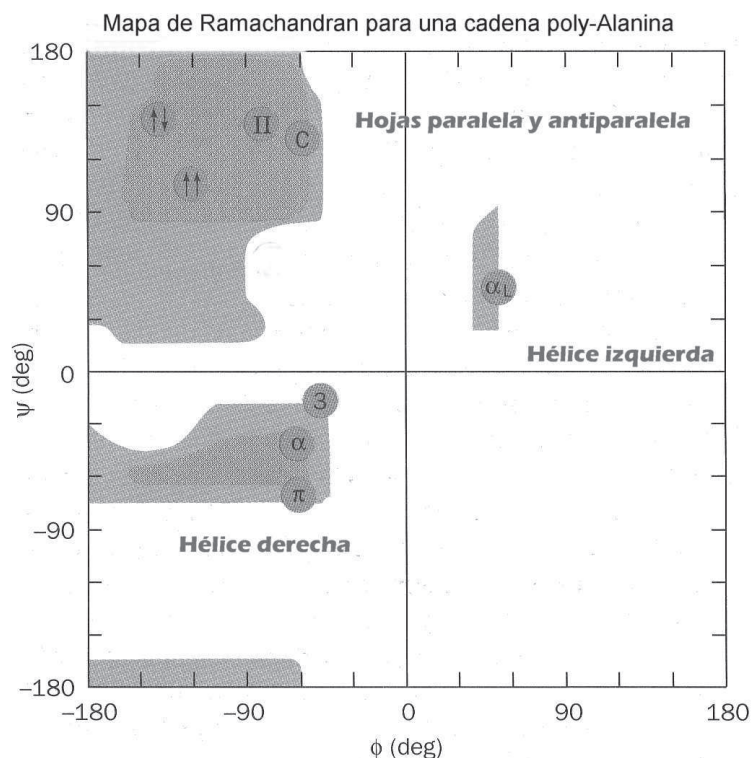


FIGURA 7. Espacio configuracional en términos de los ángulos diédricos  $\phi$  y  $\psi$ , para una cadena poly-L-Alanina. Las zonas sombreadas muestran las regiones físicamente accesibles para la proteína, en donde se ubican los cuatro tipos de estructuras secundarias conocidas. Adaptado de [Flory, P.G., 1969]

1958, Kendrew mostrará también con técnicas experimentales similares que la proteína mioglobina contiene subunidades helicoidales en su estado nativo [Kendrew., 1958], [Kendrew., 1960].

Si bien es cierto que los argumentos presentados por Pauling y Corey afirman que la estructura helicoidal es una de las conformaciones más favorables para la estructura secundaria, se requiere también un elemento adicional al argumento de simetría para garantizar su existencia. No sólo se requiere que la hélice sea una estructura energéticamente favorable para la disposición de los aminoácidos, se requiere también que la estructura sea estable ante fluctuaciones del medio. En el caso de la hélice  $\alpha$ , la disposición de los aminoácidos es tal que se generan puentes de hidrógeno entre los grupos carboxilo y amino de péptidos no adyacentes en la estructura primaria, debido a la torsión de la cadena primaria. Si consideramos que la energía de un puente de hidrógeno es aproximadamente 5 kcal/mol (20.5 kJ/mol) [Finkelstein and Pitysin, 2002] y la energía térmica  $E_{ter} = k_B T$  inducida por fluctuaciones térmicas para temperaturas fisiológicas es  $E_{ter} \sim 0.6$  kcal/mol (2.46 kJ/mol), es posible entender entonces que en estas condiciones, la estabilidad de la hélice  $\alpha$  proviene fundamentalmente de los puentes de hidrógeno entre aminoácidos.

Para caracterizar los diferentes tipos de estructuras helicoidales que pueden existir, hay al menos dos posibilidades. La primera opción contemplaría la especificación de los ángulos diédricos  $\phi$  y  $\psi$  para cada átomo central de carbono  $C_\alpha$  dentro de la cadena. Este tipo de caracterización en realidad no es exclusiva de la estructura helicoidal dado que podría aplicarse a cualquier otra y presenta desventajas en la identificación de los parámetros importantes de la hélice, tales como el número  $n$  de aminoácidos por vuelta y el grado  $p$  de helicidad de la cadena. La segunda opción, que es la estándar en la literatura del campo, contempla la especificación de dos parámetros: el número  $n$  de péptidos por vuelta helicoidal y el grado de helicidad  $p$  (*pitch*, en inglés), el cual se define como la distancia que recorre la hélice a lo largo de su eje en cada vuelta. La especificación de estos dos parámetros es suficiente para caracterizar cualquier tipo de hélice, izquierda o derecha y con cualquier grado de helicidad.

Mediante la especificación de los parámetros  $n$  y  $p$  es posible especificar completamente cualquier tipo de hélice. En principio, los argumentos presentados anteriormente respecto de la estabilidad de esta estructura son aplicables para cualquier valor de  $n$  y  $p$ , con lo que tendríamos en principio toda una gama de posibilidades para la aparición de estructuras helicoidales en proteínas. Sin embargo, el enunciado anterior no es de todo correcto. Existe un valor de  $n$  y  $p$  que maximiza la estabilidad de la estructura. Este tipo particular de hélice, que es justamente de la cual hemos hablado implícitamente a lo largo de esta disgresión, es la llamada **hélice  $\alpha$** , la cual tiene en promedio 3.6 aminoácidos por vuelta y un traslación vertical de 5.4 Å. En términos de los parámetros  $n$  y  $p$ , la hélice  $\alpha$  corresponde al caso  $n = 3.6$ ,  $p = 5.4$  nm. Para estos valores de  $n$  y  $p$ , los ángulos diédricos  $\psi$  y  $\phi$  adquieren sus valores más favorables en términos energéticos y los átomos de la cadena se empaquetan de forma tal que también maximizan las interacciones de Van der Waals. Asimismo, dado que en la hélice  $\alpha$  se tienen 3.6 residuos por vuelta helicoidal, los puentes de hidrógeno que estabilizan a la estructura se establecen entre los átomos de oxígeno del grupos carboxilo del residuo  $j$  y entre el átomo de nitrógeno del grupo amino del residuo  $j + 4$ , aproximadamente.

Uno de los rasgos más sorprendentes de la geometría de este tipo de hélice es que debido a la posición de los átomos centrales de carbono  $C_\alpha$ , los puentes de hidrógeno tienen una longitud óptima para este tipo de interacciones, de aproximadamente 2.86 Å y, más interesante aún, es que su alineación es prácticamente paralela el eje de la hélice, actuando entonces como soportes estructurales de la misma [Brandt., 1968]. La razón por la cual hélices  $\alpha$  dextrógiras son más favorables energéticamente en relación a las hélices levógiras, pese a que ambas tienen los mismos parámetros  $n$  y  $p$ , está relacionado con la orientación de los grupos residuales. En la hélice derecha, los grupos residuales están proyectados hacia *fuera* del esqueleto de la estructura y no interfieren con la formación de los puentes de hidrógeno. Debido a la configuración de la hélice izquierda, los grupos  $R$  en esta estructura se hallan en contacto más directo con el esqueleto de la proteína, dificultando los enlaces de hidrógeno e interfiriendo estéricamente entre sí [Arnott and Dover., 1967]. Como un ejemplo de una estructura secundaria conformada en su gran mayoría por hélices  $\alpha$ , la Figura (8) muestra la conformación nativa de una variante bobina de la proteína *citocromo-c*. El gráfico de la izquierda muestra la localización de los distintos átomos que forman la estructura nativa, en tanto que la figura de la derecha utiliza una representación más esquemática para las hélices  $\alpha$  (listones en rojo), las hojas  $\beta$  (flechas en amarillo) y pequeñas

cadenas que funcionan como estructuras de enlace (gris).

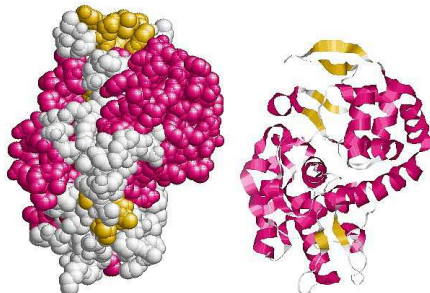


FIGURA 8. Variante bobina de la proteína *cictocroma-c* (PDB: 1EBE), cuyo patrón de estructura secundaria está conformado básicamente por hélices  $\alpha$  representadas en color magenta. En color amarillo se representan las hojas  $\beta$  y en gris las cadenas que enlazan las hélices  $\alpha$  con las hojas  $\beta$ . Gráfico obtenido mediante el programa *Rasmol*<sup>©</sup> basado en los datos obtenidos por difracción de Rayos X.

Pese a que la mayoría de las estructuras secundarias en proteínas corresponden a la hélice  $\alpha$ , se ha observado en algunos casos la presencia de hélices de otro tipo. La hélice derecha  $3_{10}$ , en la cual se tienen 3 aminoácidos por vuelta y 10 átomos en el anillo formado por dos puentes de hidrógeno consecutivos<sup>2</sup>, tiene un parámetro  $p = 6.0 \text{ \AA}$  y  $n = 3$ . En este tipo de hélice, el grado de empaquetamiento de los átomos es menor y consecuentemente, para un mismo número de residuos tiene una longitud mayor que la hélice  $\alpha$ . En términos de sus ángulos diédricos, la hélice  $3_{10}$  se ubica en una región cercana a la de la hélice  $\alpha$  en el mapa de Ramachandran, por lo que se presenta con cierta probabilidad en proteínas, particularmente como un elemento de conexión entre la terminación de una hélice  $\alpha$  y otro elemento de la estructura secundaria [Toniolo and Benedetti., 1991], [Lancelot *et al.*, 2003]. La hélice  $\pi$ , también denotada como hélice  $4.4_{16}$ , es una de las estructuras secundarias menos observadas y su presencia se ha detectado principalmente como anexos a hélices  $\alpha$ . Debido a que el número de aminoácidos por vuelta y el número de átomos comprendidos en el anillo correspondiente es mayor que en el caso de las hélices  $\alpha$  y  $3_{10}$ , la hélice  $4.4_{16}$  presenta una geometría mucho más robusta (en un sentido geométrico) y aplanada que sus similares.

Esta geometría resulta en la formación de un espacio central en la hélice demasiado pequeña para admitir moléculas de agua que a través de puentes de hidrógeno estabilicen a la estructura, y a la vez demasiado grande para una interacción de Van der Waals óptima entre aminoácidos. Estos factores reducen entonces la estabilidad del sistema y consecuentemente, su probabilidad de ocurrencia [Schultz and Schirmer., 1979]. La Figura (9) muestra en forma esquemática los distintos tipos de hélices descritas anteriormente a través de la unión de los aminoácidos correspondientes por

<sup>2</sup>Como referencia, la hélice  $\alpha$  tiene 3.6 aminoácidos por vuelta y 13 átomos formando el anillo entre dos puentes de hidrógeno consecutivos, por lo que en la notación anterior se escribiría como  $3.6_{13}$ .



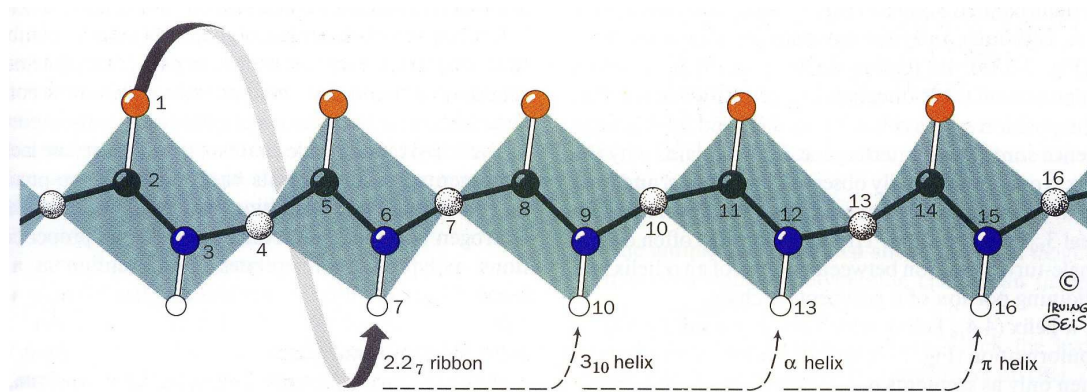


FIGURA 9. Esquematación de los distintos tipos de estructuras helicoidales que pueden presentarse en la estructura secundaria. La diferencia fundamental entre cada tipo de hélice es el número de aminoácidos por vuelta,  $n$ , el cual determina también la ubicación del puente de hidrógeno correspondiente, marcado como una flecha en la Figura. Tomado de *Biochemistry* [Voet., 1995].

puentes de hidrógeno.

El año 1951 parece haber sido excepcional en lo que se refiere a la caracterización y descripción de estructuras secundarias en proteínas. En ese mismo año, L. Pauling y R.B. Corey no sólo propusieron la existencia de la hélice  $\alpha$  sino también sugirieron que las cadenas polipeptídicas podrían agruparse en una conformación distinta [Pauling and Corey., 1951]. Esta nueva configuración, llamada **hoja  $\beta$**  ( $\beta$ -sheet), es en realidad una agrupación de cadenas polipeptídicas individuales enlazadas a través de puentes de hidrógeno, más que una configuración adoptada por una sola cadena. Dependiendo de la orientación de las cadenas que la integran, es posible distinguir entre dos tipos principales de estructuras  $\beta$ : las hojas  $\beta$  paralelas, denotadas por  $\beta \uparrow\uparrow$  y las hojas  $\beta$  antiparalelas, denotadas por  $\beta \uparrow\downarrow$ . Ambos tipos de hojas  $\beta$  contienen puentes de hidrógeno entre las cadenas individuales que estabilizan a la estructura, la diferencia entre una y otra yace en que en la estructura  $\beta \uparrow\uparrow$ , las cadenas individuales están orientadas en la misma dirección: los extremos  $N$  y  $C$  coinciden entre sí. En el caso de la hoja  $\beta \uparrow\downarrow$ , el extremo  $C$  de una cadena coincide con el extremo  $N$  de la otra y viceversa. La Figura (10) muestra la orientación de dos cadenas que forman una hoja  $\beta$  paralela y antiparalela y la ubicación precisa de los puentes de hidrógeno. De igual forma que ocurre en la hélice  $\alpha$ , la configuración de los ángulos  $\psi$  y  $\phi$  de la hoja  $\beta$  corresponde a una zona favorable en el mapa de Ramachandran, por lo que junto con la primera, la hoja  $\beta$  tanto en su versión paralela como antiparalela constituye uno de los patrones más comunes para la estructura secundaria.

En términos generales, las hojas  $\beta$  están constituidas por un promedio de seis hebras individuales, pudiendo existir hojas  $\beta$  conformadas desde sólo dos hebras individuales hasta un máximo de 15. Con un promedio de 15 aminoácidos y 4 Å de anchura por hebra, una estructura  $\beta$  típica tiene aproximadamente 90 aminoácidos y una anchura

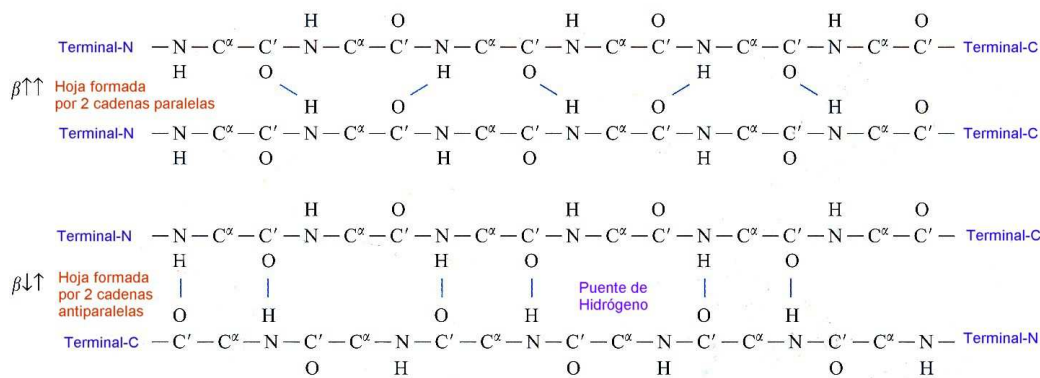


FIGURA 10. En la gráfica se muestran las dos posibilidades de enlace entre dos cadenas polipeptídicas para formar una hoja  $\beta$ . En la parte superior se muestra la conformación  $\beta \uparrow \uparrow$ , la cual corresponde a la formación de puentes de hidrógeno entre los residuos  $n$  de la primer cadena y  $n + 1$  de la segunda. Para el caso de la conformación  $\beta \uparrow \downarrow$ , los puentes de hidrógeno se forman entre los residuos  $n$  de la primer cadena y  $n$  de la segunda. Adaptado de *Protein Physics* [Finkelstein and Ptysin, 2002]

de 6 Å. Un ejemplo de estructura  $\beta \uparrow \downarrow$  formada por dieciséis hojas  $\beta$  que agrupan a 237 aminoácidos es la proteína conocida como *concanavalina-A* cuya estructura atómica se muestra en la Figura (11). En el gráfico de la derecha se ha utilizado la representación de flechas para las hojas  $\beta$  y de listones para las hélices  $\alpha$ , en este caso inexistentes, dado que la *concanavilina-A* es una proteína conformada exclusivamente por hojas  $\beta$ . Los datos experimentales para la determinación de la estructura tridimensional fueron obtenidos del *Protein Data Bank* [Berman *et al.*, 2000], [Westbrook *et al.*, 2003].

Si bien las estructuras secundarias regulares como la hélice  $\alpha$  y la hoja  $\beta$  representan aproximadamente el 50 % de la conformación del estado nativo, los aminoácidos restantes se agrupan en configuraciones menos regulares, llamadas configuraciones de vuelta o configuraciones de enlace (en inglés, *loop conformations*) [Branden and Tooze., 1991]. Una de las configuraciones de enlace más comunes es el giro  $\beta$  ( $\beta$ -turn), el cual generalmente conecta agrupamientos de hojas  $\beta$  antiparalelas. La función principal del giro  $\beta$  y otras conformaciones de enlace que conectan, por ejemplo hélices  $\alpha$  con hojas  $\beta$  o hélices  $\alpha$  entre sí consiste básicamente en agrupar distintos tipos de estructuras secundarias en el menor espacio posible, permitiendo el reacomodo espacial de las mismas. Por ejemplo, el giro  $\beta$  enlaza dos estructuras  $\beta$  antiparalelas que, sin la presencia de esta cadena auxiliar, tendrían que enlazarse ocupando un espacio mayor [Branden and Tooze., 1991]. En la representación tridimensional de la proteína *concanavilina-A* (Figura 11), se observa justamente esta característica: las configuraciones de enlace (dibujadas en colores gris y rojo) permiten conectar y agrupar a las hojas  $\beta$  en la forma más eficiente en términos de espacio ocupado. Es de esperarse entonces que conforme la estructura primaria de una proteína esté integrada por un número mayor de aminoácidos, las conformaciones de enlace se presentarán en mayor

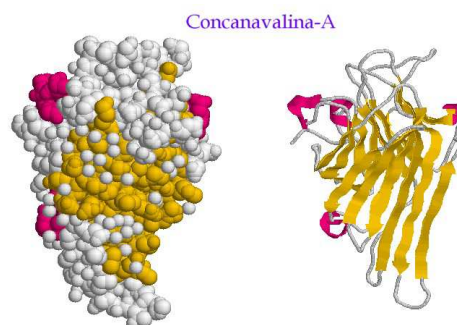


FIGURA 11. Estructura tridimensional de la *concanavalina-A*, cuya estructura nativa está conformada por 16 hojas  $\beta$  las cuales agrupan a 237 aminoácidos. El gráfico de la izquierda representa su estructura molecular en tanto que el gráfico de la derecha es una representación en donde se utilizan flechas para representar a cada estructura  $\beta$ . Ambos datos generados por computadora de acuerdo a los datos de difracción de rayos X depositados en el PDB: 1QNY [Berman *et al.*, 2000]

número dado que se requerirán más elementos adicionales para estabilizar y agrupar a las hélices  $\alpha$  y las hojas  $\beta$ . Un hecho notable a este respecto es que las conformaciones de enlace generalmente se ubican en la superficie del estado nativo, en tanto que en el núcleo principal se hallan los elementos estructurales de mayor regularidad. Además de los giros  $\beta$  existen otro tipo de configuración de enlace llamada **vuelta  $\Omega$** , la cual también se halla generalmente en la superficie del estado nativo. En promedio, una vuelta  $\Omega$  tiene una longitud promedio de 10 Å y agrupa entre 6 – 15 residuos. Dada su ubicación en la superficie del estado nativo, se piensa que tanto el giro  $\beta$ , la vuelta  $\Omega$  y otras configuraciones aún más irregulares podrían tener un papel importante en la actividad biológica del proteína, específicamente como elementos de reconocimiento de otras macromoléculas [Rose *et al.*, 1985].

Pese a que la mayoría de las cadenas heteropoliméricas pueden adquirir alguna de las conformaciones ya descritas, existen al menos dos homopolímeros conformados por aminoácidos incapaces de adoptar una estructura helicoidal o una estructura de hoja  $\beta$ . Tanto en el caso de la cadena de poly-Prolina [Berisio *et al.*, 2006] como de Poly-Glicina [Chung-Phillips., 2005], [Ohnishi *et al.*, 2006] la restricción para adoptar una estructura helicoidal proviene de la naturaleza cíclica de su cadena lateral  $R$ , los cuales interfieren estéricamente en este tipo de configuración. Además de lo anterior, la ausencia de átomos de hidrógeno en su esqueleto impide la formación de puentes de hidrógeno que estabilicen la estructura helicoidal. En condiciones apropiadas de  $pH$  (el cual tiene un papel fundamental en restringir o incrementar las interacciones electrostáticas de los aminoácidos), ambos homopolímeros precipitan como estructuras helicoidales en donde los aminoácidos adoptan la conformación *trans*, que como mencionamos anteriormente, es la configuración de la cadena primaria que evita el contacto cercano entre átomos centrales de carbono. Pese a lo que se podría pensar, las cadenas de poly-prolina y poly-glicina tienen una importancia biológica. La proteína *colágeno*, una de las más abundantes en los vertebrados cuya función principal es proveer tensión y soporte a tejidos como huesos y cartílago, está conformada en

una tercera parte por cadenas de poly-glicina y en otra tercera parte por cadenas de poly-prolina [Engel and Prockop., 1991].

#### 4. Estructura Terciaria y el Estado Nativo: Integración en una unidad biológicamente activa.

En forma implícita, es un acuerdo general el fraccionar el análisis y descripción de las conformaciones que puede adoptar una proteína, en su ruta hacia el estado nativo, en cuatro etapas o estados conformacionales. La estructura primaria, la cual se define como la secuencia ordenada de aminoácidos, y la estructura secundaria, la que se concibe propiamente como un arreglo geométrico *local* de ciertas subsecciones de la cadena principal (generalmente hélices  $\alpha$ , hojas  $\beta$  o conformaciones de enlace), son las dos primeras etapas y han sido el tema fundamental de la disgresión anterior. A continuación procederemos a la descripción de algunas de las propiedades de la **estructura terciaria**, antesala del objetivo último del plegamiento de una proteína, y del **estado nativo**: la más relevante desde el punto de vista biológico y la de mayor complejidad desde el punto de vista estructural y de estabilidad. La descripción de estas dos últimas etapas del proceso de plegamiento, nos conducirá simultáneamente hacia una disgresión sobre la naturaleza física de las fuerzas que dan cohesión y estabilidad a estos conglomerados de miles de átomos. Finalmente podremos constatar entonces, que si bien la dinámica del plegamiento está dirigida básicamente por interacciones electrostáticas, no todo está dicho en la descripción física del plegamiento de estas macromoléculas.

De acuerdo a las definiciones de mayor consenso, la estructura terciaria corresponde a la descripción de la topología general de la proteína en su estado plegado; en tanto que la mención del estado nativo incluye no sólo la descripción de la geometría del estado plegado, sino también la descripción de las interacciones entre las diferentes cadenas polipeptídicas que puedan integrar la estructura terciaria [Hopfinger., 1973]. En el caso de proteínas globulares de masa molecular pequeña<sup>3</sup>, el estado terciario corresponde generalmente a estructuras elípticas u esferoidales dentro de las cuales se agrupan las ya discutidas hélices  $\alpha$ , hojas  $\beta$  y conformaciones de enlace (vueltas- $\beta$  básicamente). Cuando la proteína está conformada por  $K > 200$  aminoácidos, lo cual corresponde generalmente a una masa molecular mayor a los 20kDa, la estructura terciaria comprende dos o más estructuras elipsoidales conocidas como *dominios*, las cuales, aunque están ligadas entre sí, exhiben un grado de interacción menor entre ellas que el grado de interacción entre las unidades que las conforman [Crippen., 1974]. Una de las características más distintivas de los dominios dentro de la estructura terciaria es su relativa independencia del resto de la proteína: una modificación genética a la secuencia primaria que codifica a un dominio generalmente no tiene un efecto significativo en la conformación y/o estabilidad de otro dominio. Tal como ocurre con la definición misma de la estructura terciaria, la asignación del término dominio para una cierta subestructura es criterio del autor y puede variar drásticamente de un caso a otro. Sin embargo, en algunas otras ocasiones la presencia de un dominio en la estructura terciaria es indudable. Éste es el caso de las proteínas conocidas como inmunoglobulinas en donde la estructura terciaria está compuesta de dos estructuras

---

<sup>3</sup>A manera de nota recordatoria diremos que una proteína se considera de masa molecular pequeña cuando ésta es menor a 20 kDa.

terciarias, la primera de las cuales agrupa a cuatro dominios y la segunda a dos. Se sabe también, que los dominios están asociados generalmente a funciones específicas de ligamiento (*binding*) entre la proteína y otras macromoléculas. Por ejemplo, el complejo  $NAD^+$  se asocia al primer dominio de la dehidrogenasa gliceraldehído-3-fosfato. De acuerdo a la definición previamente establecida, la estructura terciaria se define como la conglomeración de las estructuras secundarias en una región elipsoidal. Cada conjunto de estructuras secundarias puede agruparse en distintas formas, lo cual da origen a una clasificación de los patrones posibles. A continuación discutiremos algunos de los patrones de agrupación más comunes de las hélices  $\alpha$  o las hojas  $\beta$ , los cuales dan origen en su conjunto al término estructura terciaria.

El patrón de agrupamiento más simple con una función bioquímica conocida, es el de dos hélices  $\alpha$  conectadas por un segmento polipeptídico de enlace (*loop conformation*) [Hopfinger., 1973]. Este patrón, al cual denotaremos como  $\alpha \leftrightarrow \alpha$ , ha sido observado en dos variantes que dependen de la orientación de las hélices  $\alpha$  entre sí. Cada una de las dos posibilidades para el patrón de agrupamiento  $\alpha \leftrightarrow \alpha$  está asociado con una función bioquímica específica; enlace con secciones de la molécula de ADN en un caso y ligamiento de átomos de calcio en el otro. Las proteínas cuya estructura terciaria está integrada por el patrón  $\alpha \leftrightarrow \alpha$  son entonces, proteínas que se ligan a la cadena de ADN en procesos de transcripción o inhibición de secuencias de nucleótidos (*DNA binding proteins*) o bien proteínas cuya función es unir átomos de calcio del medio celular con el propósito de regular la actividad celular (*calcium binding proteins*). Dos ejemplos bastante conocidos de proteínas con esta estructura terciaria son la *calmodulina* y la *parvalbúmina*, las cuales participan en el intercambio de iones de calcio entre el medio extracelular y el citoplasma [Branden and Tooze., 1991].

La proteína *Parvalbúmina* es una proteína clasificada desde el punto de vista biológico como una proteína muscular, dado que se considera que el ligamiento de átomos de calcio realizado por esta macromolécula tiene un papel relevante en la relajación muscular [Moews and Kretsinger., 1975]. Con una secuencia primaria integrada por  $K = 109$  aminoácidos, presenta una estructura terciaria en donde se observan 3 patrones  $\alpha \leftrightarrow \alpha$ . Quizás el lector atento habrá notado que hemos asociado la presencia del patrón  $\alpha \leftrightarrow \alpha$  a dos funciones biológicas específicas: la participación en la unión a cadenas de ADN y átomos de calcio. El caso anterior es un ejemplo típico de una situación mucho más general: la relevancia biológica del estado nativo de una proteína yace justamente en la conformación de sus estructuras secundarias y terciarias, las cuales por su diseño molecular son capaces de realizar funciones bastante específicas. La presencia de tres agrupamientos  $\alpha \leftrightarrow \alpha$  en la *Parvalbúmina*, en donde cada uno de ellos enlaza a un átomo de  $Ca$ , es sólo un ejemplo más de la relación directa que existe entre estructura y función en las proteínas y en el caso específico de la *Parvalbúmina* del nivel de especificidad bioquímica que es posible alcanzar mediante la selección adecuada de la secuencia de aminoácidos (estructura primaria), de la conformación local de éstos (estructura secundaria) y de la manera de agruparles (estructura terciaria).

Con respecto a la proteína *Parvalbúmina*, Kretsinger *et al* [Moews and Kretsinger., 1975] observaron que el mecanismo de unión de átomos de  $Ca$  por el agrupamiento  $\alpha \leftrightarrow \alpha$  se realiza en particular en la región donde se ubica la conformación

de enlace, esto es, en la región intermedia entre las hélices  $\alpha$ , la cual hemos denotado como  $\rightsquigarrow$ . A fin de enlazar al átomo de  $Ca$ , se requiere la participación de los grupos carboxilo  $COOH$  de los aminoácidos *Asp* y *Glu*, ubicados estratégicamente en una zona en donde se forma un receptáculo de naturaleza electrostática que atrapa a la nube electrónica del átomo de  $Ca$ . Partiendo de un análisis detallado de la geometría de esta proteína, en 1973 Kretsinger *et al* dedujeron un conjunto de condiciones que cualquier secuencia de aminoácidos debería satisfacer a fin de poder generar un patrón de agrupamiento  $\alpha \rightsquigarrow \alpha$ , uno de los primeros en ser observados y caracterizados experimentalmente. Entre las restricciones observadas se encuentra el número de aminoácidos necesarios para integrar la cadena de enlace entre las dos hélices, la ubicación en esta cadena de los aminoácidos hidrofóbicos e hidrofílicos y el aminoácido específico que debía ubicarse en la posición seis, posición central de la conformación de enlace, la cual debía estar formada por doce aminoácidos. Basado en estas consideraciones, Kretsinger *et al* efectuaron la predicción teórica de que las proteínas conocidas hasta el momento que tenían una participación directa en el enlace de átomos de  $Ca$ , deberían exhibir este tipo de estructura terciaria. La determinación experimental de la estructura terciaria y nativa de las proteínas *calmodulina* y *Troponina-c* por Herzberg y James en 1985 [Herzberg., 1985] demostraron que la predicción de Kretsinger era correcta.

El análogo en cuanto al nivel de complejidad para el patrón de agrupamiento  $\alpha \rightsquigarrow \alpha$  para las hojas  $\beta$  es la llamada horquilla  $\beta$  ( *$\beta$ -hairpin* por su nombre en inglés) y consiste en dos hojas  $\beta$  antiparalelas enlazadas mediante una conformación de enlace irregular. En la notación que hemos utilizado anteriormente, esta estructura terciaria correspondería a  $\beta \uparrow \downarrow \rightsquigarrow \beta \uparrow \downarrow$ . Pese a que este tipo de estructura es bastante frecuente en proteínas que tienen cadenas  $\beta$  dado que es una forma económica en términos energéticos de unir dos o más cadenas  $\beta$  en una sola estructura, el patrón de agrupamiento  $\beta \uparrow \downarrow \rightsquigarrow \beta \uparrow \downarrow$  no ha sido asociado hasta el momento a ninguna función biológica particular. Para el caso de la proteína de 54 aminoácidos conocida como *Inhibidor de tripsina bovina*<sup>4</sup>, el patrón de agrupamiento  $\beta \uparrow \downarrow \rightsquigarrow \beta \uparrow \downarrow$  se presenta como un elemento aislado dentro de su estructura terciaria.

Aunque el motivo  $\beta \uparrow \downarrow \rightsquigarrow \beta \uparrow \downarrow$  es una forma económica de enlazar a dos cadenas  $\beta$  antiparalelas, en ocasiones se requiere agrupar a dos cadenas  $\beta$  paralelas. Para efectuar este agrupamiento, los aminoácidos de la secuencia primaria conforman un patrón de agrupamiento diferente a los anteriores, conocido como  $\beta\alpha\beta$ , el cual consiste en dos hojas  $\beta \uparrow \uparrow$  conectadas entre sí mediante una hélice  $\alpha$  y una conformación de enlace (*loop*); esto es  $\beta \uparrow \uparrow \rightsquigarrow \alpha \rightsquigarrow \beta \uparrow \uparrow$ . Dado que este tipo de agrupamiento es uno de los más simples que permite enlazar a dos cadenas  $\beta$  paralelas, se halla presente en prácticamente todas las proteínas que poseen hojas  $\beta$  paralelas, como es el caso de la proteína *triosafosfato isomerasa*, la cual en su estado nativo está formada por repeticiones consecutivas de este patrón de agrupamiento. Dado que las hojas  $\beta$  en este caso son paralelas, la hélice del patrón  $\beta\alpha\beta$  conecta la terminal carboxilo de una de ellas con la terminal amino de la otra y en términos generales, se halla que el eje de la hélice  $\alpha$  está orientado en forma paralela al de las cadenas  $\beta$ . El tamaño y función de las conformaciones de enlace merece un par de comentarios más. Se ha determinado que en este tipo de patrón de agrupamiento, las conformaciones de enlace  $\rightsquigarrow$  pueden formarse desde dos aminoácidos hasta una centena de ellos. Asimismo, en lo que se

<sup>4</sup>Esta proteína tiene como función principal inhibir la acción de la proteasa digestiva.

refiere a su funcionalidad más allá de su evidente función de conectar dos hojas  $\beta \uparrow\uparrow$ , se sabe con certeza que al menos una de ellas, aquella que enlaza la terminal amino de la hélice con la terminal carboxilo de la hoja  $\beta$ , tiene una participación importante en la formación del sitio activo de este arreglo. El papel o función bioquímica específica de la segunda conformación de enlace  $\leftrightarrow$  permanece indeterminado.

Aunque no existe un acuerdo general para el uso del término **dominio**, se considera a éste como la unidad mínima en cuanto a tamaño que es capaz de plegarse en forma independiente del resto de la cadena y que exhibe cierta funcionalidad biológica. De hecho ocurre con bastante frecuencia que dominios diferentes en el estado nativo de una proteína están asociados con funciones distintas de la misma. Un caso típico en este sentido es la proteína *lambda-repressor*, en la cual uno de los dominios se enlaza con cierta sección de la molécula de ADN en tanto que el segundo participa en la dimerización de la molécula misma [Branden and Tooze., 1991]. Asimismo, el estado nativo de una proteína puede estar formado por un sólo dominio, como es el caso de *EGF: epidermal growth factor protein*, o bien de varios docenas de ellos, como ocurre por ejemplo en los *plasminógenos*, los cuales participan activamente en los procesos de coagulación sanguínea [Branden and Tooze., 1991]. De acuerdo a la clasificación propuesta por Levitt y Crothia, se puede proceder a una clasificación de los dominios existentes en tres clases principales: dominios  $\alpha$ , dominios  $\beta$ , y dominios  $\alpha/\beta$ . La denominación anterior refiere al contenido principal del núcleo del estado nativo. En las proteínas con dominios tipo  $\alpha$ , el núcleo de su estado nativo está conformado exclusivamente por hélices  $\alpha$ . En el caso de dominios tipo  $\beta$ , el núcleo aglutina primordialmente a estructuras  $\beta$  en tanto que para las proteínas con dominio tipo  $\alpha/\beta$ , el núcleo del estado nativo corresponde básicamente a hojas  $\beta$  rodeadas por una coraza de hélices  $\alpha$ .

La intención a lo largo de los párrafos anteriores ha sido proveer al lector no familiarizado con el campo de la ciencia de proteínas, de un conjunto de definiciones, conceptos y notación estándar utilizados en el campo. Si bien no todos los elementos discutidos aquí serán utilizados en secciones posteriores, considero de interés general presentarlos como una introducción general. De ninguna manera se pretende hacer una revisión exhaustiva y por lo tanto referimos al lector interesado a los textos de bioquímica aquí mismo citados para una ampliación de los temas discutidos. En este mismo sentido general, a continuación dedicamos la última sección de este capítulo a la discusión que existe en la actualidad sobre el sustento bioquímico para las teorías de la evolución. En este sentido, el conocimiento de la secuencia primaria de aminoácidos de diversas proteínas y del tipo de plegamiento que se observa, han generado una discusión muy activa e interesante sobre las posibles rutas de evolución que han seguido los sistemas procariontes y eucariontes a lo largo de la historia de la vida misma. Considero pues que estos últimos párrafos podrán complementar al lector de una visión del papel tan importante que ha jugado la determinación de los distintos niveles de organización en las proteínas, no sólo en las teorías modernas del plegamiento, sino también en aquellas que intentan trazar el curso de la historia de los seres vivos.

## 5. ADN, ARN, Proteínas y Evolución.

El hecho de que todas las proteínas estén conformadas por exclusivamente veinte tipos distintos de aminoácidos y que a su vez éstos estén codificados en un conjunto

de sólo cuatro nucleótidos, es sin duda un hecho notable. Más notable aún si pensamos en que no estamos considerando solamente a las proteínas correspondientes al ser humano, sino a todas las especies que coexisten con él. Este hecho ha inspirado y sostenido durante mucho tiempo la idea de que la vida en el planeta Tierra necesariamente tuvo un origen común. El escenario más probable para la generación de estructuras autoreplicables y que conducirían a la generación de sistemas biológicos auto sustentables es, de acuerdo a la opinión de los expertos, el propuesto por Oparin y Haldane alrededor de 1965 [Oparin., 1964]. En este escenario, la composición química del planeta hace aproximadamente 3 o 4 billones de años era básicamente una mezcla de compuestos de carbono ( $CO_2$ ,  $CO$ ), nitrógeno ( $N_2$ ) y agua ( $H_2O$ ). Debido a la presencia de la radiación UV (entre otros factores), estos compuestos pudieron reaccionar para formar compuestos más sofisticados como los aminoácidos y nucleótidos y a partir de aquí generar al primer organismo capaz de auto replicarse. El cómo se dio la generación de esta primera estructura auto replicables no es de ninguna forma claro, ni en el esquema de Oparin y Haldane ni en las teorías modernas de la evolución. En realidad el problema es el siguiente. Actualmente sabemos que para un organismo de auto replique es necesario que tenga una cadena de nucleótidos que codifique para algunas proteínas, las cuales a su vez transcriban a la cadena de nucleotidos para posterior replicación. Esto es, se requiere ácidos nucleicos que codifiquen proteínas y proteínas que transcriban a los ácidos nucleicos. En cierto sentido es una problema de ordenamiento temporal, ¿cúal de las dos estructuras se generó en primer lugar? Hasta el momento no se conoce un mecanismo bioquímico para transcribir la información de la estructura primaria y generar ácidos nucleicos. Lo que sí ocurre es el proceso inverso: los ácidos nucleicos codifican la generación de proteínas, por lo que si durante el proceso evolutivo no se alteraron los mecanismos anteriores, lo más plausible es que los ácidos nucleicos se hayan formado en primera instancia [Woese, 1979], [Woese, 2004].

El panorama descrito por Oparin y Haldane contempla entonces, poniendo aparte la pregunta del cómo surgió aquel primer organismo auto replicable, un proceso bastante claro en esencia. En algún punto, surgió este organismo capaz de usar los veinte aminoácidos presentes en el ambiente fisicoquímico terrestre para producir proteínas como las que conocemos en la actualidad. La elección precisa de estos veinte aminoácidos y no otros que muy probablemente estuvieron a disposición es un hecho que se considera fue azaroso en esencia, dado que no se tiene evidencia en otra dirección. Así, una vez que el organismo tuvo la capacidad de utilizar y ensamblar los aminoácidos presentes en el ambiente para poder generar estructuras bioquímicas más complejas, se considera que dio inicio un nuevo paso evolutivo. La razón de ello se halla en el hecho de que el suministro natural de aminoácidos debió haber llegado en algún punto a un estado crítico, induciendo así una presión natural para que aquellos organismos que pudieran sintetizar sus propios aminoácidos partiendo de estructuras más básicas, es decir generar sus propias reacciones bioquímicas, se hallaran en una posición mucho más favorable en comparación que aquellos que dependían del suministro natural. Este proceso a la vez eliminaría al aminoácido correspondiente en el medio natural y habría seleccionado al mismo tiempo a aquellos organismos con la habilidad necesaria para producirlo por su propia cuenta. La repetición de este proceso en forma gradual para los demás constituyentes del ambiente original habría dado como resultado la creación de las diversas reacciones bioquímicas que se conocen en la actualidad, las cuales forman parte sustancial en la supervivencia y reproducción



de cualquier ser vivo [Weber and Miller, 1981].

La utilización de los veinte aminoácidos por los primeros organismos para la producción de complejos moleculares necesarios para la sustentación y replicación ha sido analizada minuciosamente por diversos investigadores. Entre ellos, Weber y Miller [Weber and Miller, 1981] y Cloud [Cloud, 1982] han sugerido que en las primeras etapas del mecanismo de Oparin y Haldane al menos uno de los veinte aminoácidos que actualmente se conocen era distinto. Este aminoácido, conocido como *ornitidina*, fue sustituido a lo largo de un proceso evolutivo por el actual *Arginina*. A fin de alcanzar esta conclusión, Weber, Miller y Cloud observaron que la distribución de los aminoácidos en las proteínas actuales no es una distribución uniforme para todos los aminoácidos, eso es, hay algunos que tienen mayor probabilidad de aparecer en una secuencia primaria y que, por otro lado, esta no uniformidad en la distribución de los aminoácidos no depende de la especie analizada. La Figura (12) muestra el porcentaje en el que aparece cada uno de los veinte aminoácidos para una base de datos de 207 proteínas no correlacionadas entre sí en un estudio de secuenciación llevado a cabo por Kappler en 1977. [Kappler, 1977] El hecho de que la distribución de la gráfica

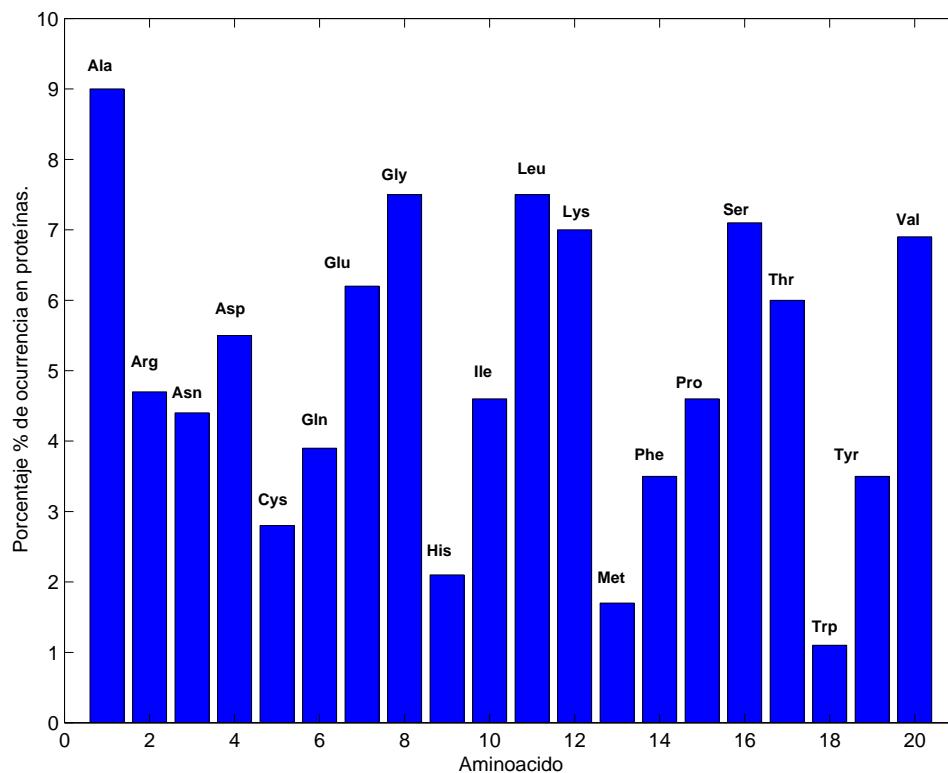


FIGURA 12. Cada uno de los veinte aminoácidos conocidos aparece en la estructura primaria de las proteínas en una proporción distinta. El presente gráfico está generado a partir de los datos obtenidos para la secuencia primaria de 207 proteínas no correlacionadas entre sí obtenido por M.H Kappler [Kappler, 1977]

(12) no sea uniforme está íntimamente ligado con el hecho de que algunos aminoácidos son codificados por un número mayor de codones que otros (Cuadro 1). Por ejemplo,

el aminoácido *Trp* es codificado solamente por el codón ( $U, G, G$ ) y tiene la menor frecuencia de aparición, aproximadamente el 1.1 %, en tanto que el aminoácido *Ser* tiene seis codones asociados y ocurre con una frecuencia mucho mayor. Sin embargo, Weber, Miller y Cloud se percataron de que esta asociación entre la frecuencia de aparición de un aminoácido en particular y el número de codones asociados a él presenta una excepción importante para el aminoácido *Arg*, el cual ocurre con la mitad del porcentaje del que debería ocurrir con la regla anterior. Uno de los argumentos propuestos para explicar esta discrepancia sugiere que el aminoácido *Orn*, que actualmente ha dejado de formar parte de las secuencias primarias de las proteínas actuales, formaba parte de los veinte aminoácidos fundamentales, dado que su sustituto actual, *Arg*, no se hallaba en el medio en cantidades suficientes. A medida que los organismos fueron capaces de sintetizar sus propios aminoácidos en lugar de absorberlos del medio fisicoquímico circundante, el aminoácido *Arg* comenzó a ser producido por los organismos y a competir con el aminoácido *Orn* por su incorporación a las secuencias primarias.<sup>5</sup> La razón por la que *Arg* permaneció dentro del conjunto de veinte aminoácidos básicos y *Orn* desapareció por completo es un tema de debate para los biólogos moleculares de la evolución. Sin embargo, algunas propuestas indican que la razón más probable para la desaparición del aminoácido *Orn* se debió a la cantidad de energía requerida por los organismos para sintetizar su cadena lateral, el cual es más complejo que el de aminoácidos similares como *Arg* o *Lys*. Lo anterior pudo haber inducido una presión evolutiva para que los organismos ‘seleccionaran’, en el sentido más amplio de este término, a aquellas estructuras que implicasen un menor gasto energético [Voet., 1995].

La idea central subyacente en la teoría de Evolución plantea que todos los organismos se originaron a partir de un conjunto reducido a ancestros comunes, los cuales al reproducirse, introdujeron un alto grado de variabilidad en la composición genética de sus descendientes, induciendo así la diversidad de especies conocidas. Si la hipótesis anterior es correcta, deberíamos observar ciertas similitudes en la composición genética de especies relacionadas y diferencias importantes en la estructura molecular de aquellas no tan cercanas y esto es justamente lo que ocurre. Se sabe en la actualidad que las estructuras primarias de las proteínas de especies cercanas presentan un alto grado de similaridad. Por ejemplo, las proteínas *citocromo-c*, las hemoglobinas  $\alpha$ ,  $\beta$  y  $\gamma$  y los fibrinopéptidos de los chimpancés y de los humanos son prácticamente indistinguibles, con sólo un aminoácido de diferencia en la cadena  $\delta$  de sus hemoglobinas y mioglobinas. Paralelamente, el número de diferencias de aminoácidos entre especies poco relacionadas crece considerablemente conforme se separan del ancestro común. El Cuadro (3) presenta una matriz del número de diferencias de aminoácidos en la secuencia primaria de la proteína citocromo-c para varias especies. Si bien es correcto afirmar que el grado de similaridad entre la misma proteína correspondiente a dos especies es mayor cuanto más cercanas sean entre sí las especies, diversos estudios muestran también que distintas proteínas han evolucionado en formas diferentes. Para la *citocromo-c*, la diferencia entre la variante humana y la del mono *Rhesus* es sólo del 1 %. Sin embargo alcanza hasta un 30 % en la estructura primaria de los fibrinopéptidos correspondientes. Distintos grupos de proteínas evolucionan y mutan, entonces, a ritmos diferentes. La comparación hecha

<sup>5</sup>El aminoácido *Arg* puede ser sintetizado por los seres vivos mediante el ciclo de la generación de urea en el metabolismo. Dada la complejidad de esta reacción, se considera que ésta apareció en etapas posteriores a las primeras reacciones metabólicas [Creighton, 1984].

–	Chimapancé	Oveja	Víbora	Caracol	Coliflor
Humano	0	10	14	29	44
Oveja	–	0	20	24	46
Víbora	–	–	0	28	45

CUADRO 3. Número de diferencias de aminoácidos en las estructuras primarias de la proteína *citocromo-c* para diversas especies. Todas las proteínas están integradas por 104 aminoácidos. Adaptado de [Creighton, 1984] a partir de la Base de Datos publicada por Dayhoff M.O. (Atlas of Protein Sequences and Structure, 1972)[Dayhoff, 1972].

anteriormente entre el *homo sapiens* y el mono *Rhesus* mostraría entonces que la tasa de variación de la citocromas-c es mucho más lenta que en el caso de los fibrinopéptidos. Este hecho es de fundamental importancia, dado que ha servido como un parámetro para determinar en forma aproximada, los tiempos evolutivos de cada especie y reconstruir, bajo una base molecular, el árbol filogenético de las especies. ¿Cómo se puede explicar esta diferencia entre las tasas de mutación de las secuencias primarias de las diversas proteínas? El escenario más plausible contempla una idea que es bien conocida en la actualidad. Para muchísimas proteínas, una mutación que altere la posición o composición de unos pocos aminoácidos en la estructura primaria puede no alterar significativamente su funcionalidad. Si la proteína continúa realizando adecuadamente su función con una estructura primaria ligeramente distinta, podrán entonces haber un número considerable de mutaciones permitidas, haciendo que la tasa de variación de una secuencia primaria sea mucho más rápida. En esencia, el planteamiento anterior sugiere que las proteínas con mucha mayor estabilidad frente a mutaciones serán aquellas con mayor tasa de variabilidad evolutiva [Kimura, 1968].

Los fibrinopéptidos y los citocromos-c parecen ser excelentes ejemplo de este esquema [Kimura, 1968]. Para el caso de los primeros, formados por sólo veinte aminoácidos, se sabe que su función principal consiste en la inhibición de la agregación de la proteína fibrina en los tejidos vasculares [Vindigni *et al.*, 1996]. Cuando ocurre un sangrado, los fibrinopéptidos son removidos de su ubicación usual para dejar libres ciertos segmentos que pueden unirse y reparar el daño. Para realizar esta función, los fibrinopéptidos requieren de la presencia de un aminoácido *Arg* en una posición específica en su estructura primaria el cual es responsable en su totalidad de su actividad inhibitoria [Pechick *et al.*, 2006]. Si éste aminoácido específico es eliminado de la secuencia inicial, la proteína es biológicamente inactiva. Dado este nivel de especificidad en la relación aminoácido-función, la estructura primaria de los fibrinopéptidos admite una variabilidad importante con respecto de los demás aminoácidos de la cadena sin afectar su funcionalidad. De hecho, se tiene claramente establecido que los demás aminoácidos pueden ser sustituidos con cierta libertad, a condición de que el posicionamiento del *Arg* no se altere [Fitch and Langley, 1979]. Dado este requisito mínimo para su funcionalidad, es posible entonces que coexistan múltiples variaciones de esta proteína con plenas capacidades funcionales y que esto indudablemente, induzca una tasa de evolución mucho más rápida en comparación con otras proteínas mucho más sensibles a su composición en aminoácidos. En el otro extremo de la correlación composición específica-funcionalidad, se hallan las citocromas-c. A diferencia de los fibrinopéptidos, que difieren hasta en un 85 % en su

secuencia primaria entre el caballo y el ser humano, los citocromos-c humanos y de los de los caballos difieren sólo en un 12 %. Dado que su función principal se halla en el proceso de oxidación de compuestos moleculares en el interior de la célula para obtener energía, se considera que el diseño de los citocromos-c es tal que permite muy pocas mutaciones que dejen invariante sus sitios activos y su funcionalidad. En este caso, no es sólo la presencia de un conjunto pequeño de aminoácidos los que definirán los sitios activos de la enzima, sino subconjuntos de ellos los que actuarán acoplada-mente para generar la actividad requerida. La Figura 13 muestra la tasa de variación de los fibrinopéptidos  $A$ ,  $B$ , las hemoglobinas  $\alpha$ ,  $\beta$  y la citocromo-c en función del tiempo desde el cual se separaron de su ancestro común (Adaptado de [Creighton, 1984]).

Al presentar al lector las ideas anteriores, hemos querido dejar la idea de que el entendimiento de la evolución de las primeras cadenas polipeptídicas y de algunas proteínas precursoras, es también el trazado de la historia de las primeras formas orgánicas sobre la superficie terrestre. En los siguientes dos capítulos presentaremos los resultados de esta investigación doctoral.

En primera instancia el Capítulo III presenta los resultados obtenidos de la aplicación de la mecánica estadística de equilibrio a la generalización del modelo de Zwanzig para calcular el cambio energético  $\Delta F$  entre el estado nativo  $N$  y el estado desplegado  $D$  cuando la proteína se halla en presencia de un agente químico desnaturante. Al concluir este capítulo estaremos en posibilidad de proveer de la forma funcional para el cambio  $\Delta F$  en función tanto de la concentración  $\zeta$  de agente desnaturante como del número promedio  $\nu$  de grados de libertad rotacionales de los grupos residuales. Este resultado nos permitirá derivar dos de las conclusiones centrales de esta primera fase. Posterior a ello, en el Capítulo IV utilizaremos la analogía termodinámica entre el proceso conocido como transición vítrea en líquidos orgánicos y la desnaturación de proteínas, para construir la distribución de tiempos  $\tau_r$  en los cuales ocurre la reacción  $D \rightleftharpoons N$  en función de la temperatura  $T$ . En el proceso para construir este resultado, generaremos uno de los resultados centrales de esta segunda fase, a saber, la entropía configuracional  $s$  para una cadena polipeptídica que transita entre un estado desplegado y una estructura helicoidal. Este resultado, generado mediante un formalismo de geometría diferencial, merece una discusión en sí mismo, dado que es un resultado completamente original de este trabajo. Al integrar el conocimiento de la entropía configuracional  $s$  para la estructura helicoidal con el esquema de Adam y Gibbs podremos mostrar y justificar una de las suposiciones estándares en la transición vítrea, la cual establece que la contribución mayor a la distribución de tiempos de relajación estructurales proviene del cambio de entropía conformacional global del sistema y no de la transición de las regiones cooperativas.

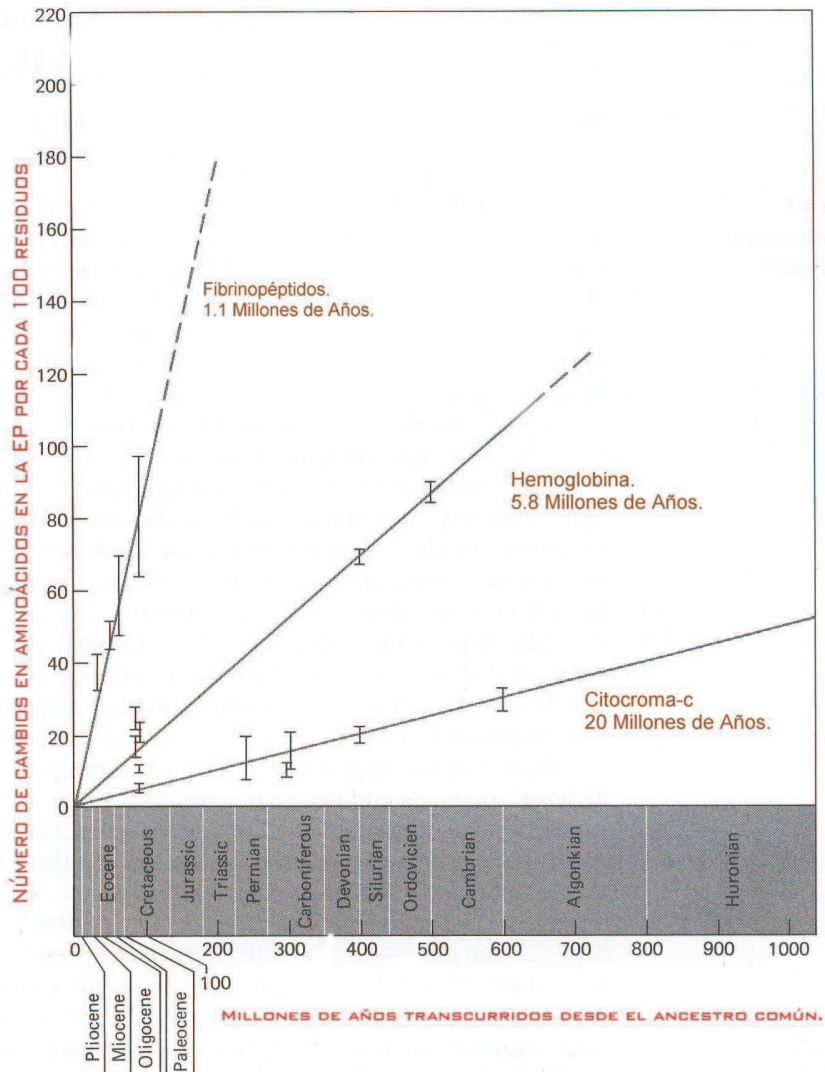


FIGURA 13. El gráfico muestra la tasa de evolución para tres clases distintas de proteínas, fibrinopéptidos, hemoglobinas y citocromas-c. El eje vertical corresponde al número de mutaciones en aminoácidos por cada 100 residuos en tanto que el eje horizontal muestra una escala temporal en millones de años. Adaptado de [Creighton, 1984].

## CAPÍTULO 3

# Desnaturalización de una proteína inducida químicamente.

En este capítulo presentamos una generalización al modelo de Zwanzig a fin de analizar el papel de los grados de libertad conformacionales de la cadena polipeptídica en la estructura del Paisaje Energético cuando la reacción  $N \rightleftharpoons D$  se realiza en presencia de un agente químico, uno de los casos más comunes en la desnaturalización de la estructura nativa. Asimismo, utilizaremos esta generalización para validar uno de los resultados experimentales más conocidos en el campo de la termodinámica de proteínas, el cual establece que el cambio  $\Delta F$  en presencia de un agente químico es una funcional lineal, y en ocasiones cuadrática, de la concentración  $\zeta$  de agente desnaturalizante.

### 1. Estabilidad del estado nativo. El efecto hidrofóbico.

A lo largo de los capítulos anteriores nos hemos ocupado de la descripción de los distintos niveles de organización por los cuales transita la estructura primaria en su ruta hacia el estado nativo. Hemos observado que a partir de la síntesis de la cadena polipeptídica por parte de los ribosomas, la proteína presenta diversos niveles de organización estructural que parten de la conformación de la hélices  $\alpha$  y las hojas  $\beta$ , a la agrupación de conjuntos de estas estructuras para la formación de *dominios*, los cuales representan el primer paso en integración de estructuras con actividad biológica. Una vez conformado el estado nativo en sí, la proteína es capaz de realizar la función bioquímica correspondiente.

A fin de que una proteína pueda concebirse como un agente primordial en diversas funciones bioquímicas debe tener poseer estabilidad ante fluctuaciones fisicoquímicas del medio en el cual está diseñada a operar. De otra forma, cualquier fluctuación térmica, de presión o de agentes químicos podría desdoblarse la estructura nativa e impedir la realización la reacción bioquímica correspondiente. En lo referente a la estabilidad del estado nativo, una de las explicaciones más aceptadas en la actualidad hace referencia a la interacción entre los grupos residuales de los aminoácidos y el solvente en el cual se halla inmersa, generalmente agua. En esta propuesta, sugerida inicialmente por Kauzmann [Kauzmann, W., 1959], la interacción primordial que dirige el proceso de plegamiento y que mantiene a la estructura nativa estable es la interacción hidrofóbica entre los grupos residuales de los aminoácidos y las moléculas de agua. En esta descripción, los aminoácidos cuyos grupos residuales presentan un carácter apolar tienden a aglutinarse en el núcleo del estado nativo en tanto que los grupos residuales hidrofílicos tienden a posicionarse en la superficie exterior de la proteína, en donde pueden formar puentes de hidrógeno con las moléculas de agua circundantes.

A fin de comprender con mayor claridad en papel del efecto hidrofóbico en la conformación y estabilización del estado nativo, conviene revisar con mayor detalle las interacciones entre los aminoácidos, el esqueleto de la proteína (*backbone*) y las moléculas de agua en el que se halla inmersa. Es un hecho bien establecido que las moléculas de  $H_2O$  son moléculas polares, es decir, presentan un momento dipolar distinto de cero, debido a la atracción electrostática que ejerce el átomo de oxígeno sobre las nubes electrónicas de los átomos de hidrógeno. Esta interacción electrostática genera un desplazamiento no simétrico de las nubes electrónicas de los átomos de hidrógeno hacia el átomo de oxígeno generando un déficit de carga electrostática en los primeros y un exceso de carga electrostática en el último (Figura 1). De esta ma-

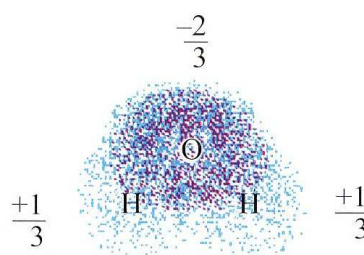


FIGURA 1. El momento dipolar neto de la molécula de  $H_2O$  proviene de la electronegatividad del átomo de oxígeno, la cual genera un desplazamiento no simétrico de la nube electrónica de los átomos de hidrógeno en la dirección del enlace O-H. . Tomado de [Finkelstein and Pitytsin, 2002]

nera, cuando una molécula de  $H_2O$  se aproxima a otra, la polaridad de ambas induce una atracción electrostática entre los átomos de  $H$  de una molécula y el átomo  $O$  de la otra, generando lo que se conoce como un *punte de hidrógeno*, que no es más que una interacción de origen electrostático debido a la electronegatividad de los átomos participantes.

Es importante mencionar que la formación de un puente de hidrógeno depende, evidentemente, de la distancia a la cual se hallen las moléculas correspondientes. Existe una distancia óptima entre los átomos de oxígeno e hidrógeno que favorece la formación de este enlace, para distancias más cortas la cercanía con los demás átomos tendría el efecto contrario, en tanto que una distancia mayor tendría como resultado una interacción demasiado débil para formar el enlace correspondiente. La distancia  $r$  a la cual se observa la presencia de puentes de hidrógeno en el caso de estructuras cristalinas corresponde aproximadamente al intervalo entre  $1.8 - 2.1 \text{ \AA}$ , en tanto que para sistemas con un grado menor de empaquetamiento molecular, la distancia  $r$  se ubica generalmente entre  $2.35 - 2.75 \text{ \AA}$ . Basados en información sobre la distancia en donde se ubican los átomos participantes y suponiendo una interacción de van der Waals, Finkelstein y Pitytsin [Finkelstein and Pitytsin, 2002] estiman que la energía de formación de un puente de hidrógeno es de  $20 \text{ kJ/mol}$ , en tanto que otros autores como Pace *et al* sugieren el intervalo  $4 - 8 \text{ kJ/mol}$  como una aproximación más moderada [Pace *et al*, 1996].

Dado que la composición del esqueleto de una proteína contiene átomos con diferentes electronegatividades, el fenómeno de desplazamiento de las nubes electrónicas e inducción de un momento dipolar neto también se presenta en las cadenas polipeptídicas de las proteínas. Si consideramos que la energía típica de un puente de hidrógeno se halla en el intervalo 4 – 20 kJ/mol, la energía asociada a las fluctuaciones térmicas del medio ( $E_{ter} = k_B T = 2.5$  kJ/mol a  $T = 298$  K) es un orden de magnitud menor, lo cual indica que los puentes de hidrógeno son estructuras estables a temperatura ambiente.

En presencia de un solvente polar, cada elemento del esqueleto de la proteína que es capaz de participar en un puente de hidrógeno se enlaza con una de las moléculas de  $H_2O$  presentes en un proceso de minimización de energía. Sin embargo, la tendencia natural de las cadenas laterales hidrofóbicas a aglutinarse en el núcleo de la proteína induce una torsión en la estructura primaria y la sustitución de dos puentes de hidrógeno entre dos elementos del esqueleto y dos moléculas de agua, por un puente de hidrógeno entre los elementos del esqueleto y un puente de hidrógeno entre dos nuevas moléculas de agua (Figura 2). Este mecanismo, que es en esencia el efecto hidrofóbico, es en términos energéticos equivalente entre sí: los dos puentes de hidrógeno iniciales formados entre el grupo A de la proteína y una molécula de agua y entre el grupo B de la proteína con otra molécula de agua son equivalentes energéticamente hablando, a los dos puentes de hidrógeno finales formados entre los grupos A y B y dos moléculas de agua (Figura 2). Sin embargo, la diferencia sustancial entre ambos procesos y que de hecho hace que exista una preferencia del sistema proteína+solvente hacia la formación de puentes de hidrógeno entre la cadena polipeptídica y la expulsión de moléculas de agua del solvente, está asociada con la entropía inherente al proceso, tal como veremos a continuación.

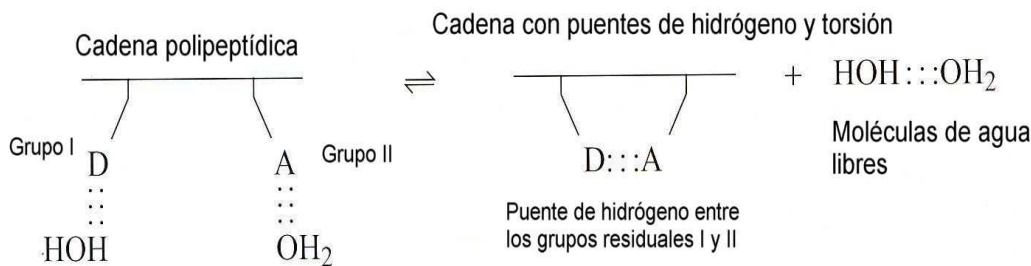


FIGURA 2. El plegamiento y torsión de la estructura primaria de la proteína ocurre dado que la sustitución de dos puentes de hidrógeno entre grupos residuales y el solvente, por dos enlaces de hidrógeno entre los grupos residuales entre sí es una reacción más favorable en términos entrópicos [Finkelstein and Pitytsin, 2002].

Al momento de que distintos segmentos de la cadena polipeptídica se enlazan entre sí mediante puentes de hidrógeno y expulsan al exterior moléculas de agua que a su vez se unirán por este tipo de enlace, la entropía configuracional del solvente



se incrementa. La razón de esto se halla en el hecho de que las moléculas de agua ya no se encuentran unidas a la cadena polipeptídica y que por tanto pueden moverse con mayor libertad dentro de líquido. Debido a este incremento en la entropía configuracional del líquido  $S_{liq}$ , la energía libre  $F$

$$(1.1) \quad F = E - TS$$

del sistema proteína+solvente disminuye en aproximadamente 6 kJ/mol por cada puente de hidrógeno formado. Es justamente este argumento el que nos permite comprender la razón por la cual el solvente juega un papel fundamental en el plegamiento de una proteína. Simultáneamente nos indica que el plegamiento dirigido por mecanismos hidrofóbicos (de formación de puentes de hidrógeno) es un proceso de naturaleza netamente entrópica, más que energética. La ruptura de los enlaces de hidrógeno entre las moléculas de agua y la cadena polipeptídica y su sustitución por enlaces entre moléculas de agua entre sí provee de grados de libertad extras a las moléculas de agua y conlleva un incremento en la entropía configuracional del líquido. El decremento de la energía libre  $F$  del sistema no proviene de los enlaces de hidrógeno entre grupos residuales, proviene de un incremento en las posibilidades conformacionales de las moléculas del líquido [Finkelstein and Pitytsin, 2002].

Si bien el efecto hidrofóbico representa uno de los mecanismos más significativos para el plegamiento de una proteína, existen dos procesos más que promueven (o inhiben) la formación del estado nativo. El primero de ellos se refiere a la presencia de moléculas de agua *dentro* del núcleo hidrofóbico de la proteína. Estas moléculas, que no son expulsadas durante la formación del estado nativo, se conoce como moléculas de agua estructural (*structural water*) y es un descubrimiento reciente en el campo de las proteínas [Rydberg *et al*, 2007]. Las moléculas de agua estructural en el núcleo de la proteína funcionan como elementos estabilizadores adicionales, dado que forman nuevos puentes de hidrógeno con cadenas laterales y residuos en donde inicialmente no estaban presentes. El segundo de los mecanismos adicionales de estabilización de la estructura nativa se refiere a la formación de un enlace covalente entre átomos de azufre presente en grupos residuales como los del aminoácido *Cis*. Este tipo de enlaces se conocen como enlaces disulfuro (*disulfide bonds*) y provee, cuando se halla presente de un nivel de estabilidad mayor al estado nativo, dada su naturaleza covalente. De hecho, dependiendo de las condiciones fisicoquímicas, la presencia de enlaces disulfuro puede ser la diferencia entre un proceso de desnaturalización reversible e irreversible.

## 2. Desnaturalización de una proteína.

Al discutir en los párrafos anteriores el origen de la estabilidad termodinámica del estado nativo, hemos introducido implícitamente el papel fundamental que en este proceso juega el solvente en el cual ocurre el plegamiento. Difícilmente podría ocurrir la formación de los puentes de hidrógeno, claves en la estabilidad estructural de la macromolécula, en una solución en la cual el solvente no tuviese las propiedades de polaridad y electronegatividad que poseen las moléculas de agua. A partir de este momento dirigiremos la atención hacia el proceso inverso del plegamiento, es decir, nos interesará describir las condiciones y mecanismos que generan el desdoblamiento de la estructura nativa de la proteína, conocido como *desdoblamiento* o *desnaturalización* (*protein unfolding*). Justamente la descripción de mecanismo de desnaturalización y las propiedades termodinámicas de la transición entre el estado nativo  $N$  y el estado

desnaturalizado  $D$  constituyen el núcleo central de esta tesis doctoral.

Desde los primeros experimentos de desnaturalización reversible realizados por C. Anfinsen [Anfinsen et al. , 1961] para el caso de la ribonucleasa-A, se ha observado que la transición  $N \rightarrow D$  puede inducirse *in vitro* a través de una modificación de las condiciones fisicoquímicas del medio, por ejemplo temperatura, presión, agentes químicos, pH.. En el caso del pH, por ejemplo, se sabe que la desnaturalización ocurre en soluciones cuyo pH se halla en los extremos de la escala logarítmica que lo define. Tanto en el caso de soluciones extremadamente ácidas ( $pH \rightarrow 1$ ), como en el caso de soluciones muy básicas, ( $pH \rightarrow 14$ ), la desnaturalización ocurre debido a la ionización de los grupos residuales en la superficie de la proteína expuesta al solvente. En términos generales, podría afirmarse que en este caso la transición  $N \rightarrow D$  obedece a mecanismos de origen puramente electrostático: el exceso o la carencia de iones de hidrógeno en la solución induce un desequilibrio electrostático en el sistema [Ghelis and Yon, 1982]. Sin embargo, el hecho de que en algunas proteínas la desnaturalización en soluciones ácidas presenta algunas diferencias con la de soluciones básicas pone de manifiesto que el proceso es aún más complejo. En general se observa que cuando  $N \rightarrow D$  ocurre en presencia de valores grandes de pH, la proteína tiende a conservar un fracción mucho menor de su estructura nativa que en el caso de desnaturalización para pH con valores pequeños.

Existen también numerosos ejemplos en donde la desnaturalización inducida por  $pH$  está asociada directamente con la ionización de cadenas laterales específicas, ubicadas generalmente en el núcleo hidrofóbico del estado nativo y cuya ionización induce la transición conformacional. Un ejemplo de ello es la transición conformacional de las serinas proteasas asociada a la ionización de la cadena lateral Asp-194. La Figura (3) muestra el perfil de desnaturalización para la  $\delta$ -quimotripsina en función del pH. En forma sintetizada podemos decir entonces que la desnaturalización inducida por pH tiene básicamente dos orígenes. Por un lado puede deberse a desequilibrios electrostáticos que ionizan a los grupos residuales en la superficie hidrofílica de la molécula o bien, a la ionización de grupos residuales hidrofóbicos específicos en el núcleo del estado nativo. En cualquier caso, el estado desnaturalizado en uno y otro caso es indistinguible: una hebra unidimensional.

El caso de desnaturalización de una proteína inducido por temperatura, esto es, mediante la adición de energía térmica al sistema solvente+proteína, ha recibido un atención considerable debido a su conexión intrínseca con las variables naturales de una descripción termodinámica: temperatura  $T$ , entalpía  $H$ , entropía  $S$ . Asimismo, el desarrollo de técnicas cada vez más precisas de microcalorimetría que permiten determinar con mayor exactitud la cantidad de calor absorbido por el sistema en un espectro amplio de condiciones de temperatura, ha permitido a los grupos de investigación acumular un gran cantidad de información experimental acerca de las características termodinámicas de la transición  $N \rightleftharpoons D$  inducida por temperatura.

En este sentido, uno de los primeros resultados experimentales arrojados por la microcalorimetría de barrido se refiere a la existencia de un máximo de absorción de energía térmica por parte del sistema proteína+solvente, cuando es desnaturalizado por un incremento de temperatura  $T$  [Privalov and Gill., 1988], [Privalov, 1982]. La gráfica (3) muestra la curva del calor específico a presión constante  $C_p$  en función

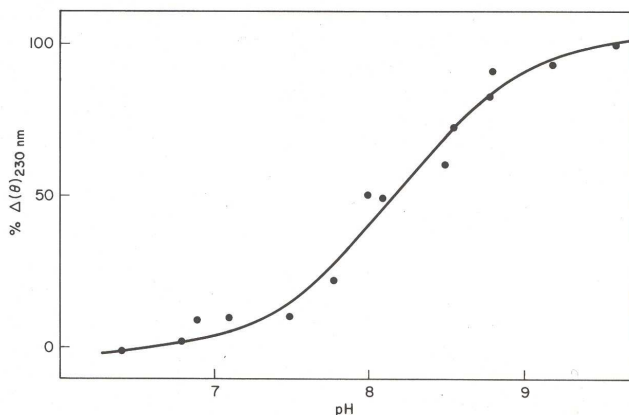


FIGURA 3. Desnaturalización de  $\delta$ -quimotripsina en función de la basicidad del solvente, monitoreada a través del cambio en la absorbancia  $\Delta\theta$  a 230 nm. La concentración de proteína en la solución es de  $10 \mu\text{M}$  a  $T = 25 \text{ }^\circ\text{C}$ . Tomado de Ghelis, C and Yon J [Ghelis and Yon, 1982].

de la temperatura  $T$  para dos proteínas de masa molecular pequeña en distintas condiciones de  $pH$  [Privalov and Dragan, 2007].

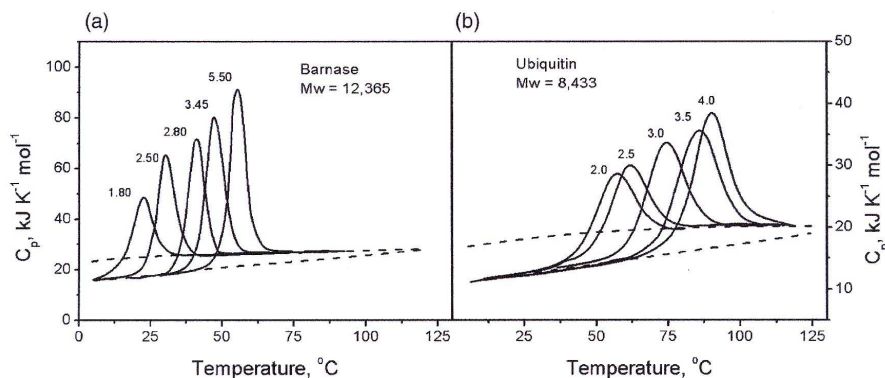


FIGURA 4. Diferencia del calor específico molar  $\Delta C_p$  en función de la temperatura  $T$  para dos proteínas a) Barnasa con una masa molecular de  $M_w = 12.4 \text{ kDa}$  y para b) Ubiquitina de masa molecular  $M_w = 8.4 \text{ kDa}$ . De acuerdo a Privalov P.L. y Dragan A.I. [Privalov and Dragan, 2007]

La interpretación más aceptada para el gráfico anterior propone que la existencia de este pico resonante en la absorción de energía térmica está directamente ligado a la exposición de los grupos no polares del núcleo de la proteína a las moléculas del solvente, dado que un fenómeno similar de incremento del  $C_p$  se observa en la adición de solutos no polares inorgánicos a un solvente polar como el agua. Además de lo anterior, la gráfica (3) sugiere la existencia de un mecanismo de dos estados bien definidos en la desnaturalización de la proteína: por un lado el estado nativo  $N$  y por el otro, el estado desnaturalizado  $D$ . Si bien la existencia de un máximo significativo en la curva  $C_p$  versus  $T$  es un indicador importante de la existencia de dos estados exclusivos en una proteína, no es el único. Experimentos ópticos de absorción y emisión de

luz muestran un cambio abrupto en la actividad biológica de la macromolécula en un intervalo pequeño de temperatura, sugiriendo también que la fractura de la estructura nativa es un proceso cooperativo. Por debajo de un valor crítico  $T_m$ , la macromolécula conserva su estructura tridimensional ante fluctuaciones del medio, sin embargo conforme  $T \rightarrow T_m$ , pequeñas perturbaciones térmicas inducen la fractura del estado nativo. El esquema anterior se conoce como una transición cooperativa y es una de las características de la transición  $N \rightarrow D$  para proteínas de masa molecular pequeña.

La descripción de la transición  $N \rightarrow D$  a través de un mecanismo de dos estados ha ganado aceptación en las últimas décadas debido a que permite también una descripción cuantitativa de las variables termodinámicas. Utilizando un esquema del tipo,



es posible escribir las relaciones termodinámicas de equilibrio del proceso, en la forma siguiente,

$$(2.2) \quad \begin{aligned} \Delta_N^D G &= -RT \log K_{eq} \\ \Delta_N^D H &= R \frac{\partial \log K_{eq}}{\partial (1/T)}, \end{aligned}$$

en donde  $K_{eq}$  es la constante de equilibrio definida como  $K_{eq} = [D]/[N]$ .

La aplicación correcta de las ecuaciones anteriores depende de que tanto la transición entre el estado  $N$  y  $D$  pueda representarse efectivamente por un mecanismo de dos estados. Si bien la evidencia experimental de cambios abruptos en alguna propiedad monitoreada durante la transición es buen criterio, no es definitivo. Existen situaciones en las cuales una proteína formada por varios dominios puede exhibir una dinámica de este tipo, pero la desnaturalización podría ocurrir a través de la ruptura de las estructuras terciarias de cada uno de los dominios. De hecho, la desnaturalización por temperatura es prácticamente el único proceso a través del cual se puede decidir si la transición es cooperativa o no, debido a que en este tipo de experimentos es posible medir con suficiente la curva de absorción de calor  $C_p$  y, consecuentemente el comportamiento del cambio de entalpía  $\Delta_N^D H$  del sistema [Privalov and Photekhin, 1986], [Dill and Bromberg, 2003].

El criterio calorimétrico utilizado se refiere al cálculo del cambio de entalpía  $\Delta_N^D$  en la temperatura  $T_m$  a la cual ocurre la transición. Tal como mencionamos anteriormente, esta temperatura representa el equilibrio del sistema entre los estados  $N$  y  $D$ , por lo que el cambio de energía libre  $\Delta_N^D G(T_m) = 0$ . Si  $T_m$  se halla ubicada en un intervalo de temperaturas muy pequeño, el cambio de entalpía entre los estados  $N$  y  $D$  puede aproximarse por la entalpía de van 't Hoff  $\Delta_N^D H_{vh}$  [Dill and Bromberg, 2003]. Esto es,

$$(2.3) \quad \Delta_N^D H \rightarrow \Delta_N^D H_{vh}(T_m) = 4RT_m^2 \frac{\Delta C_p(T_m)}{Q},$$

en donde  $\Delta C_p(T_m)$  es la altura de la curva  $C_p$  versus  $T$  en el punto medio de la transición y  $Q$  representa el área bajo la curva  $C_p$ . Con estas consideraciones, estamos en posición de describir en forma precisa una desnaturalización de dos estados. Se dice que la reacción  $N \rightleftharpoons D$  de desnaturalización o plegamiento puede representarse mediante un reacción de dos estados, si el valor de la entalpía medida caloriméricamente

$pH$	Solvente	$T_m(^{\circ}C)$	$\Delta H_{cal}(T_m)$	$\Delta H_{vh}(T_m)$	$\Delta C_p(T_m)$
0.36	HCl	31.5	61.2	62	1.90
1.05	NaCl	29.9	59.0	59	2.02
2.02	Amortiguador Glicina	31.2	65.6	67	1.98
2.80	Amortiguador Glicina	31.2	65.6	67	2.11
3.28	Amortiguador Glicina	45.8	105.0	67	2.0
5.0	Amortiguador Acetato	57.8	151.0	60	2.25
6.23	Amortiguador Acetato	60.8	155.5	66	2.05
7.80	Amortiguador acetato	61.2	178.0	72	2.0

CUADRO 1. Datos termodinámicos para la desnaturalización térmica de la Ribonucleasa-A para diferentes valores de  $pH$ . La temperatura  $T_m$  está dada en unidades de grados Celsius, en tanto que los cambios de entalpía  $\Delta_N^D H_{cal}$ ,  $\Delta_N^D H_{vh}(T_m)$  en kcal/mol. El cambio del calor específico  $\Delta C_p$  en kcal/grado Celsius mol. [Tsong *et al*, 1970].

$\Delta_N^D H_{cal}$ , dada por

$$(2.4) \quad \int_N^D \Delta_N^D C_p(T) dT = \Delta_N^D H_{cal}$$

coincide *exactamente* con la entalpía de van 't Hoff  $\Delta_N^D H_{vh}(T_m)$ , de forma entonces que

$$(2.5) \quad \Delta_N^D H = \Delta_N^D H_{vh}(T_m),$$

para un mecanismo ideal de dos estados. En referencia a la coincidencia entre las entalpías de van 't Hoff y la entalpía calorimétrica, se ha encontrado que para proteínas globulares de masa molecular pequeña, ambos resultados difieren en no más del 5%, apoyando entonces la hipótesis de un mecanismo de dos estados para estos casos [Freire and Biltonen, 1978]. Para proteínas de masa molecular mayor, el perfil de desnaturalización tiene un aspecto mucho más complejo [Privalov, 1982], por lo que en estos casos

$$(2.6) \quad \Delta_N^D H > \Delta_N^D H_{vh}(T_m).$$

El Cuadro (1) muestra un estudio de microcalorimetría realizado por Tsong *et al* [Tsong *et al*, 1970] para la desnaturalización térmica de la Ribonucleasa-A. Observando los valores obtenidos para  $\Delta_N^D H_{cal}$  y  $\Delta_N^D H_{vh}$  es posible afirmar que el mecanismo de desnaturalización en este caso se aproxima con bastante precisión a un sistema de dos estados en el caso de soluciones con  $pH \rightarrow 1$ , pero se desvía de este comportamiento en el caso de soluciones a  $pH$  mayor. Privalov y Khechinavishii [Privalov and Khechinavshvili, 1974] han analizado también, mediante microcalorimetría de barrido, la desnaturalización térmica de la lisosima, quimotripsina, citocromo-c y mioglobina, hallando que el cociente

$$(2.7) \quad \Delta_N^D H_{cal} / \Delta_N^D H_{vh}(T_m) \sim 1.05,$$

un indicador de que estamos en presencia de sistemas de dos estados. La desviación observada que es del orden del 5% se atribuye generalmente a la presencia de estados intermedios con poca estabilidad y que rápidamente son degradados hacia cualquiera de los dos estados de equilibrio del sistema.

El hecho de que la desnaturalización  $N \rightarrow D$  pueda representarse mediante un mecanismo de dos estados tiene también una connotación importante: implica que no existen estados intermedios estables o al menos con un tiempo de vida suficientemente largo para ser detectados por alguna técnica experimental y que, por tanto, la transición entre el estado  $N$  y  $D$  exhibe un alto grado de cooperatividad. Cuando éste es el caso, es posible escribir las funciones termodinámicas del proceso observando que, para la temperatura  $T_m$  a la cual sucede la fractura de la estructura nativa, el sistema se halla en equilibrio termodinámico entre los estados  $N$  y  $D$  y entonces podemos escribir que

$$(2.8) \quad \Delta G_N^D(T_m) = \Delta H_N^D(T_m) - T_m \Delta S_N^D(T_m) = 0.$$

Dado que  $\partial \Delta H(T)/\partial T = \Delta C_p$  y que  $\partial \Delta S(T)/\partial T = \Delta C_p/T$ , es posible en principio calcular el cambio de entalpía y de entropía en función de la temperatura, si se conoce con suficiente precisión el  $\Delta C_p$  para un intervalo de temperaturas. Por tanto, uno de los aspectos centrales en el estudio de la desnaturalización térmica se refiere a la medición, con el mayor grado de precisión posible, del calor específico  $C_p$  en función de la temperatura para los estados  $N$  y  $U$ , dado que  $\Delta_N^D C_p(T) = C_p^N(T) - C_p^D(T)$ .

En los albores del análisis por calorimetría de la desnaturalización de proteínas, cuando la sensibilidad de los instrumentos no permitía apreciar variaciones pequeñas en la absorción de calor, se pensó que la diferencia  $\Delta_N^D C_p$  era una función constante de la temperatura [Privalov and Khechinashvili, 1974]. Si se asume que el  $\Delta_N^D C_p$  es constante y se utiliza su relación con el cambio de entalpía

$$(2.9) \quad \Delta H(T) = \Delta H(T_m) + \int_{T_m}^T \Delta C_p(T) dT,$$

y con el cambio de entropía

$$(2.10) \quad \Delta S(T) = \Delta S(T_m) + \int_{T_m}^T \frac{\Delta C_p(T)}{T} dT,$$

podemos escribir las formas funcionales para  $\Delta H(T)$  y  $\Delta S(T)$  de acuerdo a,

$$(2.11) \quad \Delta H(T) = \Delta H(T_m) + (T - T_m) \Delta C_p,$$

y

$$(2.12) \quad \Delta S(T) = \frac{\Delta H(T_m)}{T_m} + \Delta C_p \log \frac{T}{T_m}.$$

Consecuentemente, el cambio de energía libre  $\Delta_N^D G$  en este caso se escribe de la siguiente manera,

$$(2.13) \quad \Delta_N^D G(T) = \left( \frac{T_m - T}{T_m} \right) \Delta_N^D H(T_m) - T \Delta_N^D C_p \log \frac{T}{T_m} + (T - T_m) \Delta_N^D C_p.$$

Los cálculos anteriores muestran que el perfil del cambio de la entalpía es una función lineal de la temperatura  $T$  y que el cambio de energía libre  $\Delta G$  exhibe un máximo para una temperatura  $T_g$ , la cual está dada por

$$(2.14) \quad T_g = T_m \exp \left( - \frac{\Delta_N^D H(T_m)}{T_m \Delta_N^D C_p} \right).$$

Una discusión más detallada de las implicaciones y detalles en el cálculo de las relaciones anteriores puede hallarse en el texto de Creighton T. (1999) [Creighton, T., 1999] o bien al artículo de revisión [Olivares-Quiroz, L. and Garcia-Colin, L.S., 2004].

Si bien el esquema anterior nos permite reproducir algunas de las características de la desnaturalización térmica inducida por temperatura, estudios de microcalorimetría mucho más recientes han mostrado que la dependencia en la temperatura del  $\Delta C_p$  no es en realidad constante para muchos casos, por lo que el formalismo anterior constituye sólo una primera aproximación al problema. En el capítulo IV, cuando discutamos una propuesta para el cálculo de la distribución de tiempos de relajación para una proteína, abordaremos el caso de la mioglobina, en donde el  $\Delta C_p$  es de hecho una función cuadrática en  $T$ .

### 3. Desnaturalización química de una proteína: Base experimental de la propuesta.

A continuación procederemos a establecer la fenomenología y características principales en uno de los mecanismos de desnaturalización más estudiados en la actualidad y el cual constituye uno de los ejes centrales en este proyecto doctoral. Nos referimos a la desnaturalización inducida químicamente (*chemically driven unfolding*), el cual agrupa todos aquellos procesos en donde la desnaturalización de la estructura nativa de una proteína ocurre debido a la adición de solutos orgánicos o inorgánicos al sistema proteína+solvente.

Se sabe que la adición de solutos orgánicos como la urea o el clorhidrato de guanidina (GdnHCl) al sistema proteína+solvente es un catalizador que induce el desdoblamiento del estado nativo. Con excepción de algunos proteínas extremadamente estables e insensibles a la urea o al GdnHCl, para la mayoría de las proteínas el proceso de desnaturalización se inicia para concentraciones molares pequeñas de soluto,  $[\zeta] \sim 0.5$  M y concluye con la ruptura del estado nativo para concentraciones que oscilan entre  $\zeta \sim 6 - 8$  M. [Ghelis and Yon, 1982]. Como resultado de este proceso particular de desnaturalización, se han identificado dos tipos principales de estructuras residuales. Por un lado, si el estado nativo de la proteína contiene enlaces disulfuro  $S - S$  entre aminoácidos *Cis*, el estado desnaturalizado resultante es un cadena lineal unidimensional en donde se han eliminado los puentes de hidrógeno que estabilizaban al núcleo hidrofóbico, pero con la permanencia de los enlaces disulfuro  $S - S$  los cuales tiene un margen de estabilidad mayor y no son afectados significativamente por los compuestos urea o GdnHCl. Esta estructura se conoce como *cross-linked random coil*, en contraposición a la estructura resultante cuando el estado nativo no tenía enlaces disulfuro. En este caso, los solutos orgánicos mencionados desdoblan por completo la configuración tridimensional, obteniéndose un cadena lineal completamente desplegada (*linear random coil*) [Ghelis and Yon, 1982].

Aún cuando ambos compuestos, urea y GdnHCl, son potentes agentes desnaturalizantes, se ha observado que la transición entre el estado nativo y una estructura completamente desnaturalizada se alcanza preferentemente en presencia de GdnHCl cuando la temperatura del sistema se halla entre  $0 - 50$  °C y  $pH \sim 7$  [Ghelis and Yon, 1982]. Estos resultados motivaron a Hibbard y Tulinsky [Hibbard and Tulinsky, 1978] a sugerir un mecanismo de acción para GdnHCl y urea distinto. Basados en el análisis de mapas de densidad electrónica para cristales de  $\alpha$ -quimotripsina en urea (3M) y GdnHCl (2M), Hibbard y Tulinsky han propuesto que en tanto que el soluto urea tiene una acción preferentemente sobre los grupos polares de la proteína ubicados en su superficie, el soluto GdnHCl actúa sobre éstos y también sobre los

grupos residuales hidrofóbicos ubicados en el interior del estado nativo, de forma tal que el mecanismo de desnaturalización en este caso procede simultáneamente desde el interior y exterior del estado nativo [Roseman and Jenks., 1975].

Además de los elementos descritos anteriormente, existe un cúmulo de información experimental muy amplio y bien documentado acerca de las curvas de desnaturalización que sigue una proteína en función no sólo de la concentración  $\zeta$  de soluto presente, sino también en función de las otras variables discutidas previamente, presión  $P$ , o valor del  $pH$  [Paliwal *et al.*, 2004]. Las curvas de desnaturalización son, en esencia, gráficos de la actividad biológica, espectro de emisión o absorbancia, contenido helicoidal, etc, versus la concentración de soluto orgánico presente, medidos en condiciones de equilibrio termodinámico. En principio, existen diversos métodos para detectar y seguir los cambios conformacionales entre el estado  $N$  y el estado  $D$ , la elección de cuál es más conveniente en cada caso depende del tipo de información que se desea extraer y del tipo de estructura nativa de cada proteína. Las técnicas experimentales de la hidrodinámica de fluidos en solución, que permiten medir el coeficiente de viscosidad volumétrica  $\eta$  son utilizados para detectar cambios globales en la transición  $N \rightarrow D$ , dado que la viscosidad  $\eta$  en el caso de una estructura globular compacta como es el estado nativo es significativamente diferente al del caso esplegado, considerado como una hebra unidimensional. En contraposición, los métodos experimentales basados en espectroscopía óptica, ya sea de emisión o absorción de radiación EM, permiten observar cambios conformacionales debido a que se basan en el cambio del índice de refracción del sistema cuando aminoácidos que se hallan en el núcleo hidrofóbico, particularmente *Trp*, *Tyr* o *Phe*, son expuestos al solvente en el proceso de desnaturalización [Creighton, 1984].

Las técnicas experimentales basadas en absorción-emisión de radiación EM en la región ultravioleta (UV) son uno de los métodos más utilizados para monitorear los cambios conformacionales. Durante la transición  $N \rightarrow D$ , los aminoácidos que contienen grupos aromáticos como *Phe*, *Try*, *Trp* y que generalmente se encuentran en el interior hidrofóbico de la proteína, son expuestos al solvente induciendo una modificación del espectro de absorción del sistema, caracterizado por un cambio  $\Delta\epsilon$  en la intensidad de la radiación asociada a un cambio  $\Delta\lambda$  en la longitud de onda [Cantor and Schimmel., 1980a]. Puesto que cada aminoácido responde de manera distinta a la radiación EM, el conocimiento previo de esta información es un requisito esencial para identificar en el perfil completo de desnaturalización, las contribuciones realizadas por la exposición de cada grupo cromóforo. Por ejemplo, la cadena lateral aromática de *Tyr* exhibe un cambio máximo  $\Delta\epsilon$  alrededor de 286 nm y un segundo máximo en 278 nm. La superposición de esta información para cada tipo de aminoácido y la filtración de elementos atribuibles a otros procesos, permite generar el perfil de desnaturalización de una proteína en función de distintos agentes promotores o inhibidores.

La Figura (5) muestra la curva de transición  $N \rightarrow D$  obtenida mediante espectroscopía UV en el caso de la proteína quimotripsinógeno en presencia de urea y de diversas concentraciones de  $pH$  para una emisión de 293 nm. Además de que la transición entre los estados  $N$  y  $D$  ocurre en forma más pronunciada conforme el  $pH$  se incrementa, la curva  $\Delta\epsilon$  versus  $\zeta$  exhibe un perfil tipo sigmoidal cuya curvatura depende del  $pH$  de la solución, el cual permite estimar el número de grupos cromóforos expuestos al solvente, de acuerdo al esquema propuesto por Herskovitts [Herskovitts,



1965]. Tal como veremos a continuación, el perfil mostrado en la gráfica (5) no es un caso aislado en ningún sentido. De hecho, corresponde en realidad al prototipo de perfil de desnaturalización en sistemas cuya transición  $N \rightarrow D$  está dirigido por un mecanismo de dos estados.

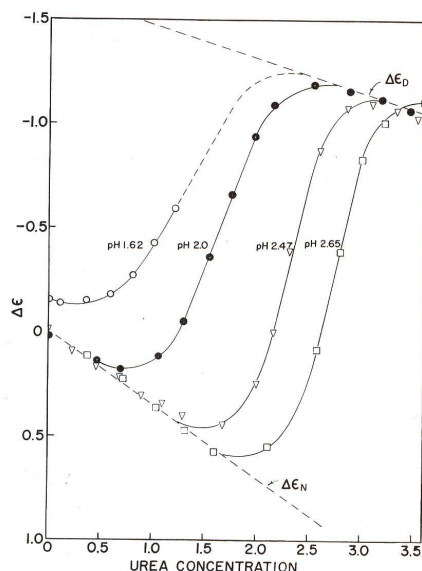


FIGURA 5. Variación  $\Delta\epsilon$  del espectro de absorción de la proteína quimotripsinógeno en función de la concentración de urea presente y del  $pH$  de la solución a 293 nm. Las líneas rectas representan  $\Delta\epsilon$  para los estados  $N$  y  $D$ . De acuerdo a Brandts [Brandts, 1965]

Uno de los principales problemas prácticos en la implementación de cualquier técnica de espectroscopía está relacionado con la concentración de proteína en el solvente, dado que por encima de un valor crítico la mayoría de ellas tiende a formar agregados insolubles [Ghelis and Yon, 1982]. En este sentido, la técnica experimental de fluorescencia de cromóforos (cadenas laterales específicas que emiten/absorben radiación EM) tiene ventajas sustanciales sobre la de espectrometría de absorbancia presentada anteriormente, dado que requiere una menor concentración de proteína dentro del solvente. Tanto las técnicas de absorbancia como de fluorescencia basan su funcionamiento en la misma idea: la dispersión y reemisión de radiación EM por aminoácidos con grupos aromáticos particularmente sensibles a un tipo específico de radiación durante su exposición al solvente en el proceso de desnaturalización. La Figura (6) muestra un espectro de emisión de fluorescencia para la proteína Fosfatasa alcalina humana (*Human Alkaline Phosphatase*) en presencia de la sal NaCl y utilizando urea como agente desnaturalizante. En este caso en particular, Hung *et al* [Hung and Chang, 2001] monitorearon la fluorescencia específica del aminoácido *Trp* durante el desplegamiento.

En forma alternativa a las técnicas anteriores, es posible también seguir los cambios conformacionales de la transición  $N \rightarrow D$  mediante la técnica de dicroísmo circular (CD, por sus siglas en inglés). En esencia, y tal como discutimos brevemente

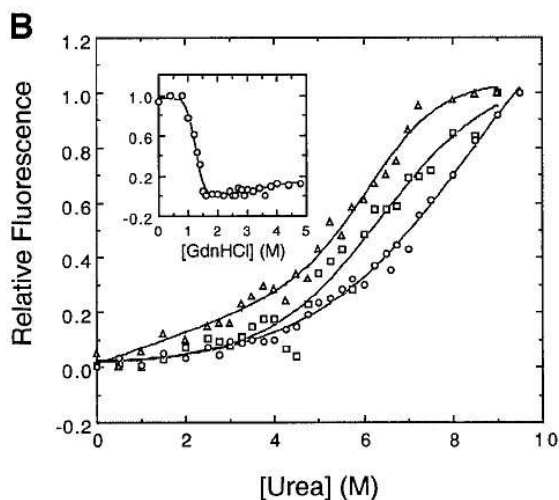


FIGURA 6. Espectro de fluorescencia durante la transición  $N \rightarrow D$  de la proteína *Human Alkaline Phosphatase* excitada con una radiación incidente a 280 nm y con un barrido entre 300 – 400 nm. La concentración de la proteína es de  $12.9 \mu\text{g}/\text{mL}$  en un amortiguador de triacetato a 0.1 M y  $\text{pH} = 7.5$ . El cuadro pequeño muestra las condiciones de desnaturalización en ausencia de la sal NaCl y en el cuadro mayor se muestra a distintas concentraciones de la misma sal. De acuerdo a los resultados de Hung *et al* [Hung and Chang, 2001].

en el Capítulo II, la técnica CD basa su funcionamiento en la interacción de radiación EM en el espectro UV con el esqueleto de la proteína y en la respuesta de éste en función de la polarización de la luz incidente [Jirgersson *et al*, 1973], [Adler *et al*, 1973]. Como resultado de este tipo de experimentos es posible medir el cambio en la elipticidad molar  $\Delta\theta$  en función de la concentración  $\zeta$  de soluto desnaturalizante. Dado que es posible estimar el contenido helicoidal de la proteína en función de  $\Delta\theta$  [Cheng and Yang, 1972], los perfiles de desnaturalización obtenidos de esta manera representan la fracción de estructura nativa presente para proteínas con dominios agrupados en estructuras  $\alpha$  principalmente. Los resultados de la aplicación del CD en la detección de cambios conformacionales para la proteína *Ervatamina-B*, la cual presenta una estructura terciaria de tipo  $\alpha + \beta$ , se presentan en la figura (7), de acuerdo a los resultados de Sundd *et al* [Sundd *et al*, 2002]. En este caso, las curvas muestran un perfil sigmoideal, característico de un mecanismo de desnaturalización primordialmente de dos estados.

El perfil de desnaturalización en función de un agente químico tipo sigmoideal mostrado aquí para el caso de *Ervatamina-B* no es un caso particular en ningún sentido. De hecho, múltiples proteínas con distintos patrones de agrupamiento en la estructura terciaria ( $\alpha + \beta$ ,  $\alpha/\beta$ ,  $\alpha\beta\alpha$ , etc.), en presencia de distintos tipos de solutos orgánicos actuando como agentes desnaturalizantes (urea, GdnHCl), en condiciones varias de pH y monitoreados por distintas técnicas de espectroscopía, fluorescencia, CD, etc., muestran características similares a los ejemplos presentados aquí. La observación anterior sugiere entonces que, pese a la gran diversidad que existe entre cada uno de los estados nativos de las proteínas y la complejidad de factores fisicoquímicos

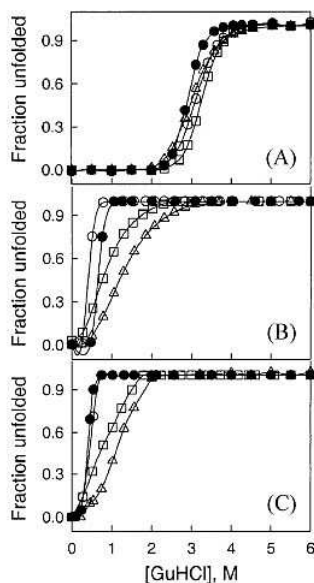


FIGURA 7. Mecanismo de desnaturalización de *Ervatamina-B* en presencia de GuHCl para diferentes valores de pH. (A), pH=7.0, (B), pH=4.0, (C), pH=3.0 mediante la técnica de DC con radiación incidente sintonizada en  $\lambda = 278$  nm ( $\bullet$ ); a  $\lambda = 222$ nm ( $\Delta$ ); actividad proteolítica ( $\circ$ ) y emisión a  $\lambda = 298$  nm ( $\square$ ). Tomado de Sundd *et al* [Sundd *et al*, 2002].

que intervienen en la interacción entre la estructura molecular de los solutos orgánicos urea, GdnHCl, y los aminoácidos de la proteína, existen mecanismos cooperativos generales similares en todos los casos que dirigen la desnaturalización.

Si bien es cierto que una descripción detallada de la estructura atómica-molecular del estado nativo, de las moléculas que integran al solvente y a los solutos orgánicos sumada a una descripción precisa de las interacciones electrostáticas entre cada una de las partes de este sistema, nos llevaría indudablemente al conocimiento preciso de los detalles en el mecanismo de desnaturalización; también es cierto que quizás esta descripción tan detallada terminaría reproduciendo los perfiles de desnaturalización aquí presentados, en donde muchos de los detalles de las interacciones proteína-solvente-soluto se han colapsado para generar un perfil estándar de desnaturalización para un grupo específico de proteínas. En forma similar a como ocurre en la descripción de la dinámica de la materia condensada en general (sólidos, líquidos, gases) en donde los detalles de la dinámica molecular no constituyen un elemento significativo en el comportamiento global del sistema, es altamente probable que en la desnaturalización de muchas proteínas los detalles del o de los mecanismos fisicoquímicos que dirigen el proceso se integren de forma tal que sólo observemos los rasgos macroscópicos más generales del mismo. Es factible pensar también que la razón por la que observamos un mismo perfil de desnaturalización para diversas proteínas, tenga su origen en una limitante de tipo experimental. Probablemente, el nivel de sofisticación que se ha alcanzado en las técnicas experimentales sean aún insuficiente para monitorear los detalles de los mecanismos de desnaturalización y que, en consecuencia, no

podamos discernir diferencias donde en realidad las hay. Cualesquiera que sean las razones de fondo para este fenómeno, la universalidad del proceso de desnaturalización, por decirlo de algún modo, será una característica que utilizaremos en la descripción de los potenciales termodinámicos de la transición  $N \rightarrow D$ .

#### 4. Teoría del Paisaje Energético: Base Teórica de la Propuesta.

Hemos dedicado las secciones anteriores a una descripción de los diversos procesos bajo los cuales es posible inducir el desdoblamiento del estado nativo dado que justamente el estudio termodinámico de este proceso corresponde a uno de los ejes de investigación de esta tesis doctoral. Se sabe, tal como lo discutimos en su oportunidad, que el estado nativo corresponde a un mínimo global en la energía libre  $F$  del proceso  $D \rightarrow N$ . Esto implica desde el punto de vista termodinámico que el cambio en la energía libre del sistema entre ambos estados  $\Delta_N^D F < 0$  en condiciones naturales, dado que el proceso ocurre en forma espontánea [Dill and Bromberg, 2003]. El conocimiento a detalle de la forma funcional de la energía libre  $F$  en función de las coordenadas de reacción adecuadas<sup>1</sup>, nos permitiría no sólo establecer con claridad las trayectorias posibles para los diversos mecanismos de desnaturalización o plegamiento de la proteína, sino también cuál es el papel de estas coordenadas de reacción en la ubicación del estado nativo dentro de un paisaje energético y, por tanto, nos permitiría analizar la estabilidad del mismo. Justo estas ideas dan sustento y motivación a la primera fase de este proyecto: Utilizando la universalidad del mecanismo de desnaturalización apoyado por la evidencia experimental, pretendemos determinar la forma funcional de la energía libre  $F$  en función de la coordenada de reacción que en este caso dirige el desdoblamiento de la estructura nativa, esto es, de la concentración  $\zeta$  de un soluto orgánico presente en el solvente. El cálculo de la energía libre  $F$  y de su variación  $\Delta F$  a lo largo de la transición  $N \rightarrow D$  nos permitirá entonces una disgresión sobre el papel que juega la coordenada reactiva en la forma funcional de  $F$  y, como consecuencia, nos permitirá también establecer cuál es el papel de ésta en la estabilidad del estado nativo.

Para desarrollar este programa, presentaremos a continuación en forma sintética las ideas principales que han llevado a la conceptualización del estado nativo como mínimo global en el espacio fase del sistema, conocido también como Paisaje Energético (*Energy Landscape*). Esta revisión resultará importante puesto que de ella extraeremos uno de los elementos teóricos esenciales para nuestro objetivo central: la idea de que una proteína es un sistema con múltiples grados de libertad conformacionales cuyas interacciones entre sí generan un espacio fase multidimensional con una estructura particular, en la cual existe un único estado de energía mínima global y un número considerable de estados metaestables representados por mínimos. Posterior a esto, integraremos esta idea junto con la universalidad de las curvas de desnaturalización para generar un formalismo que nos permita alcanzar el objetivo central: El cálculo de  $\Delta F$  en función de las coordenadas reactivas apropiadas y, por ende, de una versión reducida del paisaje energético del sistema. Finalmente, aplicaremos el formalismo así generado, que llamaremos *Modelo Extendido de Zwanzig* (MEZ),

<sup>1</sup>Se denomina coordenadas de reacción a las variables relevantes de un proceso físico y en términos de las cuales es posible escribir los potenciales termodinámicos. La definición anterior proviene de la Mecánica Clásica, en donde las coordenadas de reacción corresponden a las coordenadas generalizadas en términos de las cuales se escribe la función de Hamilton.

en dos casos de análisis: el desdoblamiento del estado nativo en función lineal de la concentración  $\zeta$  y el caso de desnaturalización para la proteína  $T70N$ , la cual exhibe un perfil de desdoblamiento sigmoidal como los presentados anteriormente y que por tanto, nos permitirá establecer propiedades generales para  $F$  y  $\Delta F$  válidos para aquel conjunto de proteínas que comparten esta propiedad.

Uno de los primeros pasos en el desarrollo de una teoría microscópica del plegamiento proviene de la analogía observada entre el plegamiento de una proteína y ciertas transiciones de fase observadas para las cadenas homopoliméricas en solución. En los trabajos clásicos de Flory [Flory., 1969], DeGennes [de Gennes P.G. , 1976] y Volkenstein [Volkenstein., 1963], se establece que una cadena homopolimérica unidimensional puede colapsarse para formar una estructura globular compacta, en donde las fluctuaciones de la densidad se han reducido a su mínimo valor. Esta transición, conocida en la literatura como transición hebra-glóbulo (*coil-globule transition*) ha sido discutida analíticamente por Lifshitz *et al* [Lifshitz, Grosberg and Khokhlov, 1978] y, más recientemente, por Grosberg y Kuznetsov [Grosberg and Kutnezov., 1992] para cadenas homo y heteropoliméricas, respectivamente. El resultado más importante que se desprende de los estudios realizados por Lifshitz *et al* [Lifshitz, Grosberg and Khokhlov, 1978], es que la transición hebra-glóbulo constituye una transición de fase en términos estrictos, es decir, en el límite termodinámico, se observa una discontinuidad en las funciones termodinámicas del sistema.

La limitante más importante tanto en el tratamiento de Lifshitz *et al* como en el de Grosberg y Kuznetov para aceptarla como una teoría consistente del plegamiento, radica en que el estado globular compacto resultante es distinto para cada secuencia de monómeros, en tanto que en las proteínas reales se observa que múltiples secuencias o incluso mutaciones en ellas pueden dar origen a una misma estructura globular. Para tomar en cuenta este efecto, Bryngelson y Wolynes [Bryngelson and Wolynes., 1989] han propuesto que una proteína, al estar constituida por aminoácidos de distintas clases y por tanto con interacciones energéticas de distintos tipos, debe poseer propiedades similares a las observadas en los vidrios de espín. Este enfoque, conocido como **Teoría del Paisaje Energético** (*Energy Landscape Theory*, ELT) supone que una proteína posee un espacio conformacional mucho más amplio, dado no sólo por el número de configuraciones espaciales en principio accesibles, sino también por un número equivalente de posibles secuencias de aminoácidos consistentes con una misma estructura nativa. Dado que muchas de estas secuencias son similares a la secuencia específica que producirá el plegamiento, la energía de estas secuencias es similar a la energía de la estructura nativa, generando que el paisaje energético de la macromolécula, es decir, la gráfica multidimensional de la energía libre  $F$  del sistema en función de la configuración, exhiba un aspecto rugoso, con múltiples mínimos locales separados por barreras energéticas y un mínimo global correspondiente al estado nativo (Figura 8).

En la teoría de los vidrios de espín, análogo propuesto por Bryngelson y Wolynes para el plegamiento de una proteína, cada átomo individual tiene un momento magnético que puede orientarse, en principio, en múltiples conformaciones. A medida que el sistema se enfría, los espines se reacomodan tratando de encontrar su configuración de mínima energía. Dado que existen múltiples interacciones tanto de corto como de largo alcance, se produce una situación de conflicto para el sistema al tratar

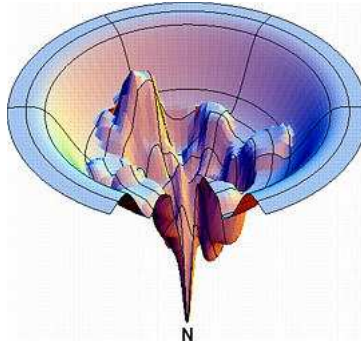


FIGURA 8. Gráfico genérico de la energía libre  $F$  de una macromolécula, en donde las diversas interacciones energéticas entre los aminoácidos produce una estructura rugosa, con múltiples mínimos locales asociados a configuraciones intermediarias y un estado mínimo global correspondiente a la estructura terciaria. (Tomada de la referencia [Dill and Chan., 1997]).

de encontrar la configuración energética más favorable. La razón para esto se halla en que dos configuraciones muy similares tendrán energía similares y por tanto el sistema tendrá en esencia, la misma probabilidad de ocupar cualquiera de estos estados. Este fenómeno, conocido en la teoría de vidrios de espín como *frustración*, constituye también la base de la Teoría del Paisaje Energético [Bryngelson and Wolynes., 1989]. En un sistema con un alto grado de frustración, el estado global de mínima energía no difiere significativamente de los estados metaestables circundantes, por lo que la probabilidad de ocurrencia de este estado es esencialmente la misma a la de las conformaciones ubicadas en mínimos locales. Siguiendo entonces esta línea de pensamiento, la Teoría del Paisaje Energético implica que para que una proteína tenga un único estado de mínima energía con un alto grado de estabilidad (que es lo que de hecho se observa en la Naturaleza), las proteínas deben ser sistemas con frustración mínima, a fin de evitar el congelamiento de la transición  $N \rightarrow D$  en un estado intermedio. Dado que el grado de frustración de una proteína está asociado en proporción directa con la aleatoriedad de las secuencias de aminoácidos, una proteína tendría que ser de acuerdo a la ELT un sistema con mínima frustración y cuyas secuencias de aminoácidos no exhiban una distribución aleatoria.

La propuesta del ELT de que las proteínas son entonces sistemas con frustración mínima y por tanto, ser el resultado de un proceso evolutivo en donde se han seleccionado las secuencias de aminoácidos que mejor resuelven este problema, introduce una componente netamente biológica al proceso de plegamiento al mismo tiempo que induce una estructura particular al paisaje energético. A fin de resolver el problema de la frustración, las secuencias de aminoácidos de las proteínas existentes deben codificar un mapa de interacciones tal que exista un estado de mínima energía global separado de los demás estados conformacionales posibles por una barrera energética significativa. En este esquema, el paisaje energético podría tener la forma de un embudo, en donde la parte superior representa los diversos estados metaestables y la parte inferior al estado nativo. (Figura 9).

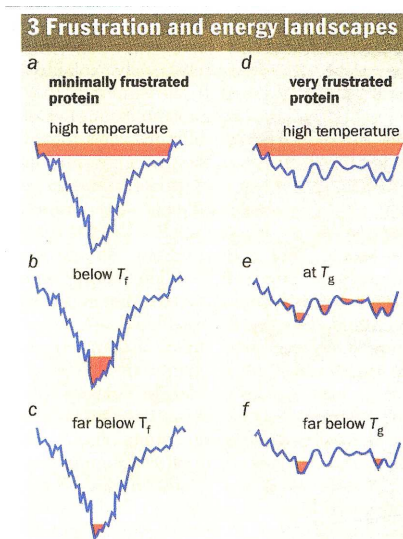


FIGURA 9. Una proteína cuya secuencia de aminoácidos es aleatoria, exhibe el fenómeno de frustración con mayor intensidad debido a la enorme cantidad de interacciones presentes, en forma similar a los vidrios de espín. En la actualidad se considera que las secuencias de aminoácidos seleccionadas mediante procesos evolutivos son aquellas con frustración mínima, las cuales presentan un paisaje energético en forma de embudo.

Además de lo anterior, el ELT tiene una predicción importante en referencia al número de posibles conformaciones en función de la temperatura, conocida como *transición por congelamiento de estados (heteropolymer freezing transition)* [Shakhnovich and Gutin, 1989]. En esta transición, el número  $O$  de estados accesibles al sistema cambia de orden de magnitud, de  $O \sim N$ , a  $O \sim \exp(N)$  cuando el sistema cruza un umbral de temperatura en forma similar a lo que ocurre en los vidrios de espín. Cuando esta transición fue predicha, primero fenomenológicamente por Bryngelson y Wolynes [Bryngelson and Wolynes., 1989] y después dentro de un esquema microscópico computacional por Shakhnovich y Gutin [Shakhnovich and Gutin, 1989], parecía ser una respuesta satisfactoria para el plegamiento de las proteínas, dado que en este modelo, existe la posibilidad de un estado termodinámico *congelado*, el cual podría identificarse, en principio, con el estado nativo de las proteínas. Sin embargo, de acuerdo a análisis posteriores realizados por Pandé *et al* [Pande, Grosberg and Tanaka., 2000], se ha mostrado que el esquema de Bryngelson-Wolynes-Shakhnovich presenta los siguientes inconvenientes

- Sólo existe una secuencia de aminoácidos que conduce al estado termodinámicamente más estable. Asimismo, la configuración tridimensional asociada a este estado tiene un carácter aleatorio y por tanto no es única. Ambos hechos son inaceptables desde el punto de vista biológico, en donde diversas secuencias de aminoácidos pueden generar el mismo estado nativo y éste, además es único.
- Aunque el estado base de las cadenas heteropoliméricas es termodinámicamente estable, los estados conformacionales próximos a él en el espacio fase

sólo difieren energéticamente por un factor  $(\sqrt{N})^{-1}$  pese a tener configuraciones nativas completamente distintas. Dada esta similitud, estos estados metaestables funcionan como trampas energéticas, haciendo que el tiempo de plegamiento se incremente en varios órdenes de magnitud a lo observado experimentalmente.

Para resolver este problema, Pande, Grosberg y Tanaka [Pande, Grosberg and Tanaka., 2000] han propuesto extender el formalismo anterior utilizando secuencias de aminoácidos mínimamente frustradas, las cuales, como se mencionó anteriormente, están asociadas con paisajes energéticos de tipo embudo y distan de ser aleatorias. En términos técnicos, el enfoque utilizado por Pande *et al.*, consiste en calcular la energía libre de  $F$  del sistema mediante un promedio estadístico sobre un espacio de secuencias  $E_S$ , el cual corresponde a todas las posibles cadenas lineales que pueden formarse con  $N$  elementos de  $q$  clases distintas y sobre un espacio de configuraciones  $E_C$ , el cual corresponde a las posibles estructuras nativas que pueden formarse con las secuencias mencionadas. La energía libre  $F$  del sistema puede calcularse entonces de acuerdo a,

$$(4.1) \quad F_{\text{secuencias+conformaciones}} = \langle \langle F \rangle_{E_S} \rangle_{E_C},$$

en donde la notación  $\langle x \rangle_y$  denota un promedio estadístico de la variable  $x$  sobre un ensamble en el espacio  $y$ . La ecuación (4.1) indica que el promedio se realiza en primer lugar sobre las posibles secuencias a fin de encontrar las más óptimas desde el punto de vista de la frustración y, en segundo lugar, se realiza el promedio sobre las posibles configuraciones tridimensionales accesibles para dicha secuencia. Para realizar tales promedios, se utiliza la función de partición  $Z$  definida sobre ambos espacios y su relación con la energía libre  $F$ . Esto es,

$$(4.2) \quad F = -k_B T \langle \langle \log Z \rangle_{E_S} \rangle_{E_C}$$

en donde  $k_B$  denota la constante de Boltzmann y  $T$  es la temperatura termodinámica del sistema. La aparente simplicidad de la ecuación (4.2) es engañosa. Para estar en posibilidades de calcular la energía libre  $F$  del sistema bajo este esquema, se requiere conocer la función de partición completa del sistema, la cual depende a su vez de la función de Hamilton  $H$ , requiriendo entonces de un modelo para las interacciones entre los monómeros de la cadena.

Abordando el problema desde esta perspectiva, diversos grupos de investigación han dedicado esfuerzos considerables a proponer, bajo diversas y muy distintas argumentaciones, funciones hamiltonianas  $H$ . Entre las más destacadas se halla la propuesta de Wolynes, Bryngelson y Derrida [Derrida, 1981], conocida como *Random Energy Model (REM)* el cual propone en esencia que, dado que no es posible conocer con exactitud la totalidad y la forma específica de las interacciones entre los aminoácidos de la cadena polipeptídica primaria, podemos optar por una descripción de naturaleza estocástica, en donde el sistema exhibe una distribución de energías  $P(E)$  con perfil gaussiano dada por,

$$(4.3) \quad P(E) = \frac{1}{2\pi\Delta^2} \exp(-(E - \bar{E})/2\Delta^2).$$

en donde  $\bar{E} = Nz\epsilon$  es la energía media del sistema,  $\epsilon$  es la energía de contacto,  $N$  es el número de residuos de la cadena y  $\Delta^2 = zN$  es la varianza de la distribución.



El modelo anterior es la propuesta más simple para un sistema, que en analogía con los vidrios de espín, exhibe el fenómeno de la frustración. El resultado más importante del REM es que existe una energía crítica  $E_c$ , tal que para  $E > E_c$ , la densidad de estados  $\eta(\bar{E}) \gg 1$ , en tanto que si  $E < E_c$ ,  $\eta(\bar{E}) \ll 1$ . En el lenguaje de las proteínas, el modelo REM predice la existencia de un estado energéticamente favorable en un sistema con frustración, el cual podría asociarse con el estado nativo. Pese a este éxito, el estado energéticamente más favorable en este esquema no necesariamente es único, dado que estamos hablando de un modelo estocástico y en realidad nos estamos refiriendo a una distribución de estados energéticamente favorables, cada uno de los cuales tiene una cierta probabilidad  $p$  de ocurrencia.

Más allá de la idea de REM, Wolynes y Takada [Takada and Wolynes, 1997] han propuesto un modelo determinista para la función de energía  $H$  de una macromolécula del tipo desarrollo del virial, en donde se incluyen interacciones de corto alcance entre los diversos aminoácidos. El hamiltoniano  $H$  propuesto tiene una estructura del tipo,

$$(4.4) \quad H = k_B T \sum_i \frac{(\vec{r}_{i+1} - \vec{r}_i)^2}{2a^2} + \frac{v}{2} \sum_{i \neq j} b_{ij} \delta(\vec{r}_i - \vec{r}_j) + c \frac{v^2}{6} \sum_{i \neq j \neq k} \delta(\vec{r}_i - \vec{r}_j) \delta(\vec{r}_j - \vec{r}_k),$$

en donde  $\vec{r}_j$  representa la ubicación del átomo de carbono  $C_\alpha$  de cada aminoácido,  $a$  es la longitud de Kuhn de las cadenas homopoliméricas y  $b_{ij}$  y  $c$  son coeficientes viriales [Grosberg and Khokhlov, 1994]. Dependiendo del tipo de aminoácidos presente, cada una de las energías de interacción  $b_{ij}$  puede obtenerse de una distribución gaussiana, por ejemplo.

Con esta propuesta y utilizando el método de réplicas, se puede calcular la energía libre  $F$  del sistema, a través de la relación

$$(4.5) \quad [F]_{av} = -k_B T [\log Z]_{av}$$

en donde la notación  $[x]$  indica un promedio estadístico sobre la distribución de las energías de interacción  $b_{ij}$ . A fin evitar calcular directamente el promedio del logaritmo de  $Z$ , con frecuencia se utiliza la identidad matemática,

$$(4.6) \quad \log Z = \lim_{n \rightarrow 0} \frac{x^n - 1}{n},$$

con la cual, la ecuación (4.5) puede expresarse de acuerdo a

$$(4.7) \quad -\beta [F]_{av} = [\log Z]_{av} = \lim_{n \rightarrow 0} \frac{[Z^n]_{av} - 1}{n}.$$

La función de partición  $Z^n$  promediada sobre la distribución probabilística de las energías individuales de interacción puede determinarse a través de la relación [Takada and Wolynes, 1997],

$$(4.8) \quad [Z^n]_{av} = \int \Pi_\alpha \mathcal{D}\vec{r}_i^\alpha e^{-\beta H}.$$

en donde

$$(4.9) \quad \mathcal{D}\vec{r}_i \equiv \Pi_j d\vec{r}_j \delta \left( \sum_j \vec{r}_j \right).$$

La integración de la ecuación (4.8) es prácticamente imposible de hacerse en manera exacta, de forma entonces que Takada y Wolynes han propuesto extender la desigualdad de Bogoliubov-Feynmann [Feynmann, 1965], al espacio de las réplicas  $Z^n$ . La desigualdad de Bogoliubov-Feynmann establece que si un sistema tiene una función de Hamilton  $H$  y una energía libre  $F$ , entonces para cualquier otro hamiltoniano  $H_0$ , la energía  $F$  es una cota inferior para la energía libre  $F_0$  del nuevo sistema. Esto es,

$$(4.10) \quad F \leq \langle H \rangle_0 - TS_0$$

en donde el promedio  $\langle H \rangle_0$  se realiza sobre la matriz de densidad del hamiltoniano original  $H$ . Es decir,

$$(4.11) \quad \langle H \rangle_0 = \frac{\text{Tr}[He^{-\beta H_0}]}{\text{Tr}[e^{-\beta H_0}]}.$$

Por tanto, podemos escribir la siguiente relación, que es la bien conocida **Desigualdad de Bogoliubov-Feynmann**,

$$(4.12) \quad F \leq F_0 + \frac{\text{Tr}[H - H_0 e^{-\beta H_0}]}{\text{Tr}[e^{-\beta H_0}]},$$

en donde  $F_0 = \langle H_0 \rangle_0 - TS_0$ . Una demostración de la desigualdad anterior y de diversas aplicaciones a sistemas de osciladores armónicos y anarmónicos cuánticos se halla en el texto de R.P. Feynmann, *Statistical Mechanics: a set of lectures*. [Feynmann, 1972].

El esquema que consiste en proponer una función  $H$  y después utilizar diversos procedimientos y técnicas para calcular la función de partición  $Z$  (en algunos casos sólo cotas mínimas a la energía libre  $F$  como en el caso de la desigualdad BF), ha sido explorada por diversos grupos de investigación [Pande, Grosberg and Tanaka., 2000]. Los resultados principales que ha arrojado este procedimiento, en sus diversas modalidades, son los siguientes,

- La energía libre  $F$  del sistema exhibe propiedades de rugosidad, como las esperadas en un sistema con múltiples interacciones que compiten entre sí.
- Existen dos temperaturas tipo transición vítrea,  $T_A$  y  $T_K$ , con  $T_A > T_K$ . Para  $T > T_A$ , el paisaje energético es monotónico y prácticamente liso.
- Para temperaturas  $T_K < T < T_A$ , el paisaje energético presenta numerosos mínimos locales metaestables, los cuales corresponden a solución de rompimiento de simetrías en el espacio de réplicas.
- Finalmente, para  $T < T_K$ , existe un número pequeño de estados energéticos estables. Se considera que el plegamiento completo de la macromolécula ocurre a una temperatura cercana a  $T_K$ , en donde comienza a manifestarse el escenario previsto de escape de trampas energéticas para alcanzar un estado único.

Los resultados mencionados aquí forman parte de los diversos enfoques para analizar la transición al estado biológicamente activo de una proteína. Como se mencionó, cada uno de los enfoques ha permitido avanzar en ciertas direcciones y obtener resultados importantes en la comprensión global de este fenómeno. Sin embargo, aún se considera que el problema central del plegamiento, esto es, la predicción completa del estado nativo a partir de la secuencia de aminoácidos continúa como un problema abierto.

En términos generales, las propuestas que han producido resultados más apropiados en el contexto de las proteínas están asociados con los enfoques basados en mecánica estadística de cadenas heteropoliméricas [Pande, Grosberg and Tanaka., 2000]. En este tipo de modelos, la construcción de funciones de Hamilton  $H$  es un elemento crucial, dado que a partir de este nivel se integra en forma numérica a la función de partición del sistema y la energía libre  $F$ . Si bien es factible que cada proteína puede seguir una ruta distinta en el proceso de construcción del estado nativo, el proceso de desnaturalización no parece compartir esta característica. Tal como hemos revisado en la sección anterior, diferentes tipos de proteínas desnaturalizadas por distintos factores fisicoquímicos muestran comportamientos similares, por lo que para abordar el mecanismo de desnaturalización es altamente probable que podamos prescindir de un nivel de detalle tan sofisticado como podría ser la construcción de funciones de Hamilton. La presencia de elementos cooperativos en la transición  $N \rightarrow D$ , que colapsan los detalles de la interacción a nivel molecular para generar un comportamiento global típico, nos permitirá construir una aproximación mesoscópica para establecer las propiedades termodinámicas de la desnaturalización. En la siguiente sección, procederemos entonces a la construcción y aplicación de esta idea.

### 5. Modelo Extendido de Zwanzig I. Desnaturalización lineal

Como punto de partida para analizar la desnaturalización de una proteína en función de las condiciones fisicoquímicas del solvente en el cual se halla inmersa, partiremos del esquema propuesto por Zwanzig [Zwanzig, 1995] para describir la función de partición  $Z$  de una cadena que puede exhibir múltiples estados conformacionales. Este modelo no pretende responder la pregunta planteada anteriormente acerca del papel de las secuencias de aminoácidos que mejor resuelven el problema de la frustración. Mas bien, asume que existen tales secuencias para la cadena y busca construir una función de partición adecuada que reproduzca algunas de las características fenomenológicas observadas, incorporando la idea central del paisaje energético, en la cual el estado nativo corresponde a un mínimo global en la energía del sistema.

En el esquema de Zwanzig (EZ) [Zwanzig, 1995] se considera a una proteína como un sistema integrado por un conjunto de  $K$  variables cuyos valores definen en forma unívoca al estado del sistema. Cada una de estas  $K$  variables puede asociarse con un variable física, por ejemplo con las coordenadas espaciales de los átomos centrales de carbono  $C_\alpha$  de la cadena, con el par de ángulo diédricos  $(\phi, \psi)$  que definen la orientación de cada enlace peptídico o con alguna otra variable que permita especificar un estado conformacional de la proteína. Por otro lado, si notamos que tanto las coordenadas espaciales  $(x, y, z)$  de cada átomo central de carbono  $C_\alpha$  o bien cada par de ángulos diédricos puede adquirir un cierto número de valores dependiendo del estado conformacional en el que se halle el sistema, podemos ver con claridad que cada una de estas  $K$  variables puede exhibir un número distinto de valores<sup>2</sup>, al cual definiremos como  $\nu + 1$ . De acuerdo a lo anterior, el número de configuraciones  $\omega$

---

<sup>2</sup>En forma general, cada una de las coordenadas generalizadas del conjunto  $K$  podría tomar un número distinto de valores  $\nu_j + 1, j = 1, ..K$ . Este punto lo retomaremos más adelante, sin embargo para los propósitos de esta discusión consideraremos que todas las variables en el conjunto  $K$  tienen el mismo número de estados accesibles.

accesibles en principio a la macromolécula, esta dado por,

$$(5.1) \quad \omega = (\nu + 1)^N.$$

Si asociamos a la variable  $K$  con el número de aminoácidos de una proteína, entonces la variable  $\nu + 1$  describe el número de estados conformacionales al cual cada uno de ellos puede acceder. Dado que el estado nativo de una proteína es único, éste corresponde a una ubicación particular de cada uno de las cadenas laterales  $R$  de la macromolécula, por tanto de las  $\nu + 1$  configuraciones accesibles para cada grupo  $R$ , existe una que corresponde a su ubicación cuando se genera la estructura nativa y las  $\nu$  restantes corresponden a estados no-nativos de carácter metaestables. Es posible definir entonces un parámetro de orden  $\alpha$  el cual representará el número de grupos residuales ubicados en el estado nativo para una configuración dada. En términos de  $\alpha$ , la configuración  $\omega_\alpha$  que corresponde al estado nativo, es  $\omega_K$  en tanto que la estructura primaria, en donde no se han formado los puentes de hidrógeno, enlaces disulfuro, etc que generan al estado nativo, corresponde a la configuración  $\omega_0$ . El parámetro de orden  $\alpha$  resulta un elemento esencial en esta descripción puesto que nos permite caracterizar los distintos estados conformacionales de la macromolécula y simultáneamente, definir en forma operativa donde se halla el estado nativo  $N$  y la cadena sin estructura  $D$ .

A fin de construir una función de partición adecuada, Zwanzig propone utilizar uno de los resultados centrales de la teoría del paisaje energético [Bryngelson and Wolynes., 1989]. Como mencionamos anteriormente, a fin de que una proteína pueda adquirir su estructura nativa en tiempo relativamente cortos, debe poseer un grado mínimo de frustración y por tanto el paisaje energético correspondiente deberá poseer una estructura particular con un mínimo global separado por barreras energéticas significativas de los demás estados. Esta idea permite construir *ex profeso* una forma genérica para la energía interna  $E$  de la macromolécula en donde la energía  $E^*$  del estado base se halle en la región negativa del espectro y las energías  $U$  de los estados conformacionales metaestables en la región positiva. Asimismo, siguiendo esta idea, supondremos que el espectro de energías  $E$  del sistema es función exclusivamente del número de aminoácidos en una conformación particular, de modo tal que podemos escribir a  $E$  en función del parámetro de orden  $\alpha$  de acuerdo a,

$$(5.2) \quad E_\alpha = (K - \alpha)U - E^* \delta_{K-\alpha,0},$$

en donde  $K$  es el número de grupos residuales de la macromolécula,  $U > 0$  es la contribución energética de cada aminoácido cuando se halla en una conformación no nativa,  $E^* > 0$  es la energía del estado nativo y  $\delta_{K-\alpha,0}$  representa la función delta de Kronecker, la cual  $\delta_{a,b} = 0$  si  $a \neq b$ . La estructura de la ecuación (6.1) implica también, tal como ocurre en un sistema de dos estados, que el estado nativo corresponde a un efecto cooperativo de transición de todos los aminoácidos puesto que sólo en la configuración  $\omega_K$  correspondiente a  $\alpha = K$ ,  $E(K) = -E^*$ .

Tal como se ha definido la estructura de los niveles de energía  $E$ , éstos sólo dependen del número  $\alpha$  de grupos residuales ubicados en el estado nativo y no de su ubicación específica dentro de la cadena, Esto es, cualquier secuencia de aminoácidos que tenga el mismo valor de  $\alpha$  tendrá la misma energía  $E$ , no importando en donde se hallen ubicados estos aminoácidos dentro de la cadena. Esta consideración introduce un grado de degeneración en los niveles de energía  $E$  puesto que existen múltiples ordenamientos de cadenas tamaño  $K$  con  $\alpha$  aminoácidos ubicados en el estado nativo

consistentes con un mismo valor de  $E$ . Si llamamos  $g_\alpha$  a la degeneración del nivel energético  $E_\alpha$ , entonces podemos escribir que

$$(5.3) \quad g_\alpha = \nu^{K-\alpha} \binom{K}{\alpha}.$$

en donde  $\binom{K}{\alpha}$  representa el coeficiente binomial en las variables  $K$  y  $\alpha$ .

A partir de esta suposición para la forma funcional del espectro de energías y suponiendo que el sistema se halla en equilibrio termodinámico con un baño térmico a temperatura  $T$ , la función de partición  $Z$  del sistema está dada por

$$(5.4) \quad Z = \sum_{\alpha=0}^N \nu^{K-\alpha} \binom{K}{\alpha} e^{-\beta((K-\alpha)U - E^* \delta_{K-\alpha,0})}.$$

en donde hemos sustituido las ecuaciones (Eq 5.3) y (Eq 5.2) en

$$(5.5) \quad Z = \sum_{\alpha=0}^N g_\alpha e^{-\beta E_\alpha},$$

que es la definición de la función de partición en el conjunto canónico.

El tratamiento anterior para el espectro de energías y la función de partición correspondiente considera a una proteína como un sistema con  $K(\nu + 1)$  grados de libertad conformacionales que definen cada uno de los estados accesibles. Asimismo, partiendo de la teoría del paisaje energético construye una hipótesis sobre la forma funcional del espectro de energías  $E$  de la macromolécula y lo incorpora dentro del formalismo de la mecánica estadística de equilibrio para determinar entonces la energía libre  $F$  del sistema a través de su conexión con  $Z$ ,  $F = -K_B T \log(Z)$ . Tal como ha sido diseñado, el esquema de Zwanzig así planteado considera que el proceso de desnaturalización o plegamiento de una proteína está dirigido en esencia por la temperatura, dado que al estar inmerso en un baño térmico con un reservorio de calor a temperatura  $T$ , las fluctuaciones térmicas de éste último inducen los cambios conformacionales en el primero. Estas ideas forman nuestro punto de partida para construir una generalización al EZ que permita incluir en esta descripción la influencia de otro tipo de variables fisicoquímicas que puedan provocar la fractura del estado nativo en el caso de la desnaturalización. Tal como ya revisamos, existen diversos factores de este tipo que generan la ruptura de los enlaces estabilizadores de esta estructura, el factor temperatura es uno de ellos, pero también y es en este caso en el que nos centraremos, la presencia de solutos orgánicos como urea o GdnHCl, entre otros.

La base para la generalización del esquema de Zwanzig, al cual llamaremos Modelo Extendido de Zwanzig (MEZ) tiene su fundamentación de acuerdo a las siguientes consideraciones. Como primer punto, es importante notar que los esquemas teóricos discutidos en la sección anterior están basados en la construcción de funciones de Hamilton de las interacciones entre aminoácidos y las moléculas del solvente para después proceder a construir la función de partición correspondiente. Como ya vimos, el nivel de detalle en las funciones  $H$  diseñadas no garantiza necesariamente una descripción completa dado que esto implicaría establecer y formular a detalle el conjunto completo de interacciones entre la macromolécula y el solvente, lo cual es prácticamente imposible. Por otro lado, si estamos interesados en describir propiedades globales en la desnaturalización o plegamiento de una proteína, bastará con una

descripción mucho más modesta, pero que capture los rasgos esenciales del proceso. En este sentido, el MEZ pretende buscar una opción mesoscópica para el diseño de la función  $H$ , o lo que es equivalente, para el espectro de energías del sistema. Basados en estas consideraciones, propondremos el espectro de energías  $E$  del sistema para el caso en el cual se halla presente un agente desnaturizante que pueda inducir el desplegamiento. La forma funcional que proponemos para  $E$  no está deducida de primeros principios porque eso implicaría regresar al punto anterior, sino propuesta en base a información experimental que podamos extraer de los distintos experimentos de desnaturalización que se han llevado a cabo con numerosas proteínas.

En segundo lugar, se construye la forma funcional genérica que debe tener el espectro de energías  $E$  tomando en consideración los argumentos y resultados de la Teoría del Paisaje Energético. En esencia, el hecho de que las proteínas tengan tiempos de plegamiento relativamente cortos es una confirmación de que el proceso de plegamiento está dirigido y codificado de antemano por secuencias de aminoácidos con un grado de frustración mínimo. A fin de que esta condición se cumpla, la estructura del espacio fase debe tener una forma tal que dirija el proceso hacia un estado base de mínima energía global, siendo la estructura tipo embudo presentada en su oportunidad el ejemplo más característico. Asumiremos pues, que cualesquiera que sea el espectro de energías  $E$  propuesto, deberá satisfacer este tipo de topología. La forma funcional para  $E$  propuesta por Zwanzig es un caso discreto de este tipo de forma funcional. En el MEZ discutiremos cómo generalizar este tipo de función al caso continuo.

Finalmente procedemos a establecer el tercer punto y piedra angular del MEZ. Tal como revisamos anteriormente, existen diversos mecanismos para inducir la desnaturalización del estado nativo. Pese a que los principios físico-químicos mediante los cuales ocurre cada uno de estos procesos son diferentes, existen comportamientos generales comunes a toda una clase de proteínas, independientemente de su composición en aminoácidos y del tipo de agente inductor de la desnaturalización. Esto nos permite considerar entonces que si  $\alpha$  representa el número de aminoácidos en el estado nativo, esta cantidad debe ser necesariamente una función de todas las variables fisicoquímicas que influyen en la proteína, estableciendo así una conexión entre las variables que caracterizan al solvente y la conformación que puede adquirir una macromolécula. En términos matemáticos, podemos escribir lo anterior en forma directa de acuerdo a la siguiente hipótesis,

$$(5.6) \quad \alpha = \alpha(T, pH, \zeta, ..),$$

en donde  $T$ ,  $pH$ , y  $\zeta$  representan las distintas variables que pueden inducir la desnaturalización de la estructura nativa. Si a lo anterior sumamos la idea de que el espectro de energías  $E$  del sistema depende a su vez del parámetro de orden  $\alpha$ , entonces tenemos un esquema completo que, una vez determinada la forma específica para la ecuación (5.6), nos pone en condiciones para proceder a un cálculo de la función de partición  $Z$  a través de las ecuaciones estándares de mecánica estadística de equilibrio.

Ahora bien, ¿cómo obtener la dependencia funcional específica del número de aminoácidos en el estado nativo respecto de las condiciones fisicoquímicas del solvente? Bien, para ello recordemos que las curvas de desnaturalización obtenidas mediante técnicas experimentales diversas miden, en esencia, los cambios conformacionales de

la macromolécula en función del pH, concentración de soluto orgánico  $\zeta$ , etc. Dicho en otros términos, lo que los experimentos reportan es el cambio en la fracción de estructura nativa en función de las variables anteriores. Si asociamos la variable  $\alpha/K$ , que es la fracción de aminoácidos que exhiben la estructura nativa, con las curvas de desnaturalización, estaremos entonces en posición de construir y calcular la función de partición del sistema en estas condiciones a partir de un enfoque sustentado en observaciones y datos experimentales.

Antes de proceder con la aplicación de las ideas anteriores a un caso particular, sólo unas breves líneas sobre el significado y uso bajo el cual se debe entender a la Ec (5.6). Dado que esta ecuación establece que la conformación de una proteína depende de las variables fisicoquímicas del medio circundante, debe considerarse como un tipo de ecuación constitutiva para el sistema, entendiendo éste termino en el mismo sentido que se le da en Hidrodinámica clásica, en donde la ecuación constitutiva contiene la respuesta del fluido a perturbaciones externas. Para ciertos casos, es posible deducir de primeros principios ecuaciones constitutivas para fluidos, pero la gran mayoría de ellas proviene ciertamente de información experimental, ajena al formalismo de la mecánica estadística o la termodinámica. En ese mismo contexto, la ecuación (5.6) contiene la respuesta de la macromolécula a la presencia del soluto orgánico. En la medida de que podamos deducir de primeros principios o proponer fenomenológicamente una ecuación que represente la mayor cantidad de detalles microscópicos, será la medida en la cual podremos extraer la mayor cantidad de información sobre la dinámica del sistema.

Aplicaremos el formalismo del MEZ para el caso en que la fracción de estructura nativa, denotada por  $\alpha' = \alpha/K$  depende sólo de la concentración de soluto orgánico desnaturalizante. Esto es, consideraremos un caso particular de la Ec (5.6), en donde  $\alpha = \alpha(\zeta)$ . Consideraremos en primera instancia el caso de una proteína que se desnaturaliza siguiendo un perfil lineal en términos de la concentración  $\zeta$ , dado que es la relación matemática más simple que puede existir. Si bien, hemos discutido ampliamente que la mayoría de los perfiles de desplegamiento observados corresponden a curvas sigmoidales, el análisis de este caso nos permitirá desarrollar cierta intuición para el caso más complicado que analizaremos posteriormente. Este último caso, que corresponde a una desnaturalización de dos estados con perfil sigmoidal, será discutido ampliamente en la siguiente sección.

A fin de construir un perfil de desnaturalización consistente, debemos establecer dos condiciones sobre la función  $\alpha'$ ,

- En ausencia del agente desnaturalizante, la proteína permanece en su estructura nativa. Por tanto  $\alpha'(0) = 1$ .
- Cuando se ha agregado una concentración crítica  $\zeta^*$ , la proteína se ha desnaturalizado completamente, por lo que  $\alpha'(\zeta^*) = 0$ .

Con estos requerimientos,  $\alpha(\zeta)$  es una recta con pendiente negativa, definida por la ecuación,  $\alpha = -A\zeta + K$ , en donde  $A = \frac{K}{\zeta^*}$ . Suponiendo entonces una forma funcional del tipo (Ec 6.1) para el espectro de energías y sustituyendo la relación  $\alpha = -A\zeta + K$ , obtenemos una expresión para  $E$  en términos de la concentración  $\zeta$  de soluto orgánico dada por,

$$(5.7) \quad E(\zeta) = A\zeta U - E^* \delta(A\zeta),$$

en donde  $A, \zeta, U$  y  $E^*$  representan las cantidades ya mencionadas y hemos sustituido a la función delta de Kronecker por una función delta de Dirac centrada en  $\zeta = 0$ . El cambio de la delta de Kronecker, la cual es una función finita acotada por la función delta de Dirac, la cual es infinita para  $\zeta = 0$  se debe a que en este caso la variable de la cual depende el espectro de energías  $E$  es una variable continua, a diferencia del caso discreto de Zwanzig en donde  $E$  depende exclusivamente de  $\alpha$ , una variable de naturaleza discreta.

En la ecuación (5.7) así obtenida, observamos que el espectro fenomenológico para este tipo de desnaturalización exhibe las propiedades esperadas de acuerdo al paisaje energético. En  $\zeta = 0$  se tiene un pozo energético negativo, en tanto que para valores  $\zeta > 0$ , la energía interna  $E$  es positiva creciente en forma lineal. Ciertamente la sustitución de la función delta de Kronecker por la función delta de Dirac introduce una divergencia infinita en el espectro  $E$ , la cual de momento mantendremos por ser éste una primera aproximación al mecanismo del MEZ. Cuando discutamos el caso real de un perfil sigmoidal procederemos en forma diferente para modificar este término.

Dado entonces  $E(\zeta)$ , procedemos a escribir la función de partición correspondiente a un conjunto canónico, la cual debe escribirse en términos de una integral sobre el espacio configuracional de la macromolécula. Esto es, la expresión original

$$(5.8) \quad Z = \sum_{\alpha=0}^K \nu^{K-\alpha} \binom{K}{\alpha} e^{-\beta((K-\alpha)U - E^* \delta_{K-\alpha,0})}.$$

debe transformarse en un integral sobre la variable  $\zeta$ , dado que ésta es continua sobre el intervalo en el cual ocurre la desnaturalización. Es decir

$$(5.9) \quad \sum_{\alpha} \rightarrow \int d\zeta$$

Para proceder con esta transformación, no sólo se requiere sustituir al espectro de energías  $E(\alpha)$  por  $E(\zeta)$  (Ecuación 5.7), también se requiere transformar el coeficiente binomial que aparece en (5.8) a fin de escribirlo en términos de una variable continua. Para escribir el término binomial, utilizaremos la aproximación de Stirling [Arfken., 1985], la cual nos permite escribir la siguiente ecuación

$$(5.10) \quad \binom{K}{\alpha} = \frac{K!}{\alpha!(K-\alpha)!} \simeq \frac{1}{\sqrt{2\pi}} \frac{K^{K+1/2}}{(K-\alpha)^{(K-\alpha+1/2)}(\alpha)^{(\alpha+1/2)}}.$$

Debido a la presencia de los términos factoriales, el coeficiente binomial  $\binom{K}{\alpha}$  exhibe un máximo en el valor  $\alpha = K/2$  y tiene un perfil de tipo gaussiano [Butkov., 1968], de forma entonces que podemos representar al coeficiente binomial como una función gaussiana centrada en  $\alpha = K/2$  con un coeficiente dado por el valor máximo de la ecuación (5.10) evaluado en  $\alpha = K/2$ . De esta manera, podemos escribir la siguiente aproximación,

$$(5.11) \quad \binom{K}{\alpha} \simeq \frac{2^{K-b}}{\sqrt{2\pi K}} \exp\left(-\frac{2}{K} \left(\alpha - \frac{K}{2}\right)^2\right),$$



en donde  $b$  es un parámetro ajustable que depende de  $K$ . Para los efectos de la integración de la función de partición  $Z$  no es relevante la forma funcional en que depende  $b$  de  $K$ , dado que esta última variable es una constante respecto de la variable de integración  $\zeta$ .

Sustituyendo la relación  $\alpha = -A\zeta + K$  en (5.11), obtenemos entonces la expresión para el coeficiente binomial en términos de la variable continua  $\zeta$ . Esto es,

$$(5.12) \quad \binom{K}{\alpha(\zeta)} \simeq \frac{2^{K-b}}{\sqrt{2\pi K}} \exp\left(-\frac{2}{K} \left(\frac{K}{2} - A\zeta\right)^2\right),$$

en donde  $A$  tiene el valor definido previamente  $A = \frac{K}{\zeta^*}$ .

La función de partición se escribe entonces de la siguiente manera,

$$(5.13) \quad Z = \int_{\zeta=0}^{\zeta=\zeta^*} d\zeta \nu^{A\zeta} \frac{2^{K-b}}{\sqrt{2\pi K}} \exp\left(-\frac{2}{K} \left(\frac{K}{2} - A\zeta\right)^2\right) \exp[-\beta(UA\zeta - E^* \delta_{A\zeta,0})]$$

La función de partición anterior puede calcularse en forma exacta si separamos la integral en dos contribuciones. Para ello, observemos que el término que corresponde a la energía del estado nativo sólo contribuye a la integral cuando  $\zeta = 0$ . Por tanto, separaremos la integral en dos partes evaluando el integrando en  $\zeta = 0$  y posteriormente efectuando la integración sobre el intervalo semiabierto  $(0, \zeta^*]$ . Esto conduce a la siguiente expresión,

$$(5.14) \quad Z = \frac{2^{K-b}}{\sqrt{2\pi K}} \left\{ \exp\left(-\frac{K}{2} + \beta E^*\right) + \lim_{j \rightarrow 0} \int_j^{\zeta^*} \nu^{A\zeta} \exp\left[-\frac{2}{K} \left(\frac{K}{2} - A\zeta\right)^2 - \beta U A \zeta\right] \right\}.$$

La integral que aparece en el segundo término de la Ec (5.14) puede calcularse explícitamente, obteniéndose como resultado para la función  $Z$  la expresión siguiente,

$$(5.15) \quad Z = \frac{2^{K-b}}{\sqrt{2\pi K}} \left\{ \exp\left(-\frac{K}{2} + \beta E^*\right) + \frac{2^{K-b}}{\sqrt{2\pi K}} \left\{ \lim_{j \rightarrow 0} \left[ \frac{1}{2A} \nu^{-\frac{K}{4}(U\beta-2)} [Erf(X) - Erf(Y)] \exp\left(\frac{K}{8} (U\beta(U\beta-4) + \log \nu)\right) \right] \right\} \right\},$$

en donde las funciones de error  $Erf(X)$  y  $Erf(Y)$  están definidas por,

$$(5.16) \quad Erf(X) = Erf\left[\frac{4A\zeta^* - 2K + KU\beta - K \log \nu}{\sqrt{8K}}\right],$$

y

$$(5.17) \quad Erf(Y) = Erf\left[\frac{4Aj - 2K + KU\beta - K \log \nu}{\sqrt{8K}}\right].$$

Finalmente, la relación  $F = -RT \log Z$  nos provee con una fórmula explícita para la energía libre de este sistema en función del número promedio de grados de libertad  $\nu$ , de la temperatura  $T$  y del número de residuos  $K$ , por mencionar las variables de mayor relevancia.

### 6. Modelo Extendido de Zwanzig II. Perfil sigmoidal de desnaturalización.

En la sección anterior presentamos los detalles operativos del MEZ en el caso de un perfil de desnaturalización lineal. La simplicidad de la relación propuesta entre la fracción de contactos nativos  $\alpha$  y la concentración de agente desnaturalizante  $\zeta$  nos permitió calcular, con un mínimo de aproximaciones, la forma analítica de la función de partición  $Z$  y de la energía libre  $F$  en función de las variables discutidas anteriormente. En esta sección procederemos a analizar las consecuencias e implicaciones del MEZ para una proteína que exhibe un perfil de desnaturalización de tipo sigmoidal.

Como caso de análisis consideraremos la desnaturalización inducida por GdnHCl de la variante T70N de la lisosima de origen humano. La lisosima humana es una proteína de tipo enzima cuya actividad hidrolizante sobre el segmento  $\beta-(1-4)$  de la cadena homopolimérica en la enzima N-acetilquitina presente en las paredes celulares de las bacterias de tipo Gram positivas, permite clasificarla como un enzima de tipo agente antibacterial [Salton, 1960]. Integrada por  $K = 130$  aminoácidos distribuidos en una estructura de tipo  $\alpha + \beta$ , con 4 hélices  $\alpha$ , 3 hojas  $\beta$   $\uparrow\downarrow$ , 4 puentes  $S-S$  que estabilizan la estructura terciaria y una masa molecular de  $M_w = 14,300$  Da, es considerada como una proteína pequeña [Kumeta et al., 2003]. La Figura (10) muestra un gráfico generado mediante el programa Rasmol de la estructura nativa de T70N basado en los datos de difracción de rayos X reportados por Johnson *et al* [Johnson et al., 2005] y depositados en el Protein Data Bank con el código 1W08.

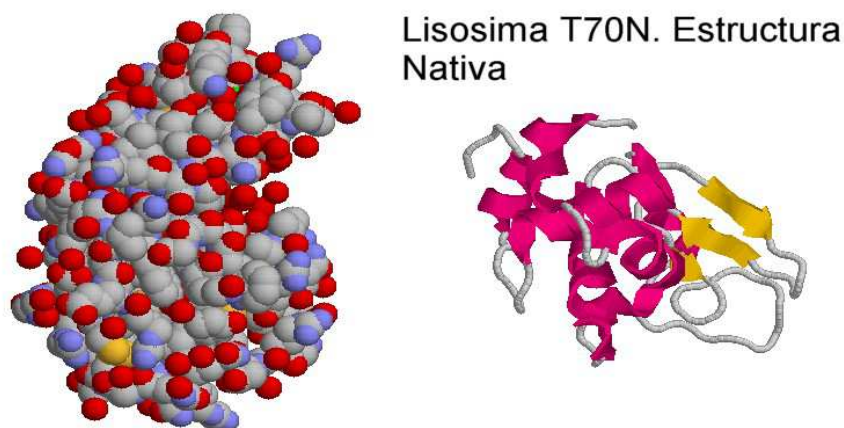


FIGURA 10. Gráfico de la estructura nativa de la variante T70N de la enzima lisosima de origen humano de acuerdo a los datos de difracción de rayos X reportados por [Johnson et al., 2005] en el PDB. El lado izquierdo muestra una representación atómica de los átomos de carbono (gris), oxígeno (rojo), Nitrógeno (azul) y azufre (amarillo). El lado derecho utiliza una representación más general, en términos de hélices  $\alpha$  (rojo) y hojas  $\beta$  (amarillo).

La proteína lisosima y sus cinco variantes mutagénicas conocidas e identificadas con los códigos I56T, F57I, W64R, D67H y T70N del *Protein Data Bank*, han recibido

en la última década una enorme atención debido a su conexión con el estudio de las enfermedades moleculares, particularmente de la clase de patologías conocidas como amiloidosis [Pepys et al., 2003]. En este tipo de patologías inducidas por mutaciones de la secuencia principal de aminoácidos, la estructura helicoidal del estado nativo es transformada a una conformación de tipo  $\beta$  con una distribución de carga electrostática diferente a la original. Cuando esto ocurre, la proteína mutada es incapaz de solubilizarse en forma adecuada y se aglutina ya sea en el interior o exterior celular formando agregados, que eventualmente impedirán otras funciones bioquímicas al interior o exterior de la célula.

En particular en el caso de la proteína lisosima, se sabe que tanto la variante natural como cuatro de sus mutaciones están asociadas con padecimientos amiloidogénicos presentes en hígado y riñones, principalmente [Yazaki *et al.*, 2003]. De las cinco mutaciones conocidas, sólo T70N, que corresponde a la sustitución del aminoácido Trp en la posición 70 por el aminoácido Asn, no se ha detectado que participe en la formación de agregados fibrilares *in vivo*, esto es, en pacientes que han desarrollado alguna amiloidosis de riñón o hígado [Booth *et al.*, 2000]. Si tomamos en cuenta que la diferencia entre las variantes amiloidogénicas de la lisosima y la variante T70N es solamente un aminoácido ubicado en un *loop* de la superficie hidrofílica del estado nativo, resulta natural comprender que el estudio de la desnaturalización, plegamiento y dinámica de la variante T70N de la lisosima ocupe un lugar preponderante en los estudios de proteínas amiloidogénicas.

Además de lo anterior, T70N tiene un comportamiento anormal en referente a la cinética del proceso  $D \rightarrow N$ . En general, las hélices  $\alpha$  son estructuras que se forman en etapas tempranas del plegamiento, con tiempos de formación en la escala de los microsegundos, a diferencia de la estructuras  $\beta$ , las cuales se forman en una etapa posterior y además con tiempos mucho más largos de formación. Para el caso de T70N, sin embargo, el monitoreo del plegamiento ha mostrado un mecanismo diferente. Canet *et al* han mostrado que existen dos escalas de tiempo superpuestas en la formación del estado nativo de T70N. En primer lugar, ocurre la formación de dos hélices  $\alpha$  para posteriormente ocurrir, en forma simultánea la formación de las estructuras  $\beta$  y de las dos hélices  $\alpha$  restantes. La escala de tiempos es además distinta. El primer proceso ocurre en tiempos del orden microsegundos, en tanto que el segundo ocurre en forma más lenta, del orden de milisegundos [Canet *et al.*, 1999].

Los experimentos de desnaturalización de T70N, de la variante natural y de los mutantes D67H, I56T, reportados por Esposito *et al.*, muestra un perfil sigmoideal para la fracción de estructura nativa en función de GdnHCl (Figura 11) [Esposito *et al.*, 2003]. El monitoreo del desdoblamiento del estado nativo se realizó mediante la técnica de emisión de fluorescencia con una longitud de onda incidente  $\lambda = 340$  nm, en una solución con pH=6.5 y en equilibrio térmico a una temperatura  $T = 20$  grados Celsius, mostrando una transición del tipo de dos estados para todas las variantes, incluida T70N. Para el caso específico de esta última, Esposito *et al* reportan que el punto medio de la transición ocurre en  $\zeta_m = 3.4$  Molar con un error experimental de  $\Delta\zeta_m = \pm 0.1$  Molar y que la desnaturalización del estado nativo ocurre en  $\zeta^* = 6$  Molar de GdnHCl. El perfil sigmoideal de desnaturalización de T70N, junto con los datos experimentales anteriores nos permitirán aplicar el formalismo del MEZ para

describir el espectro de energías  $E(\zeta)$ .

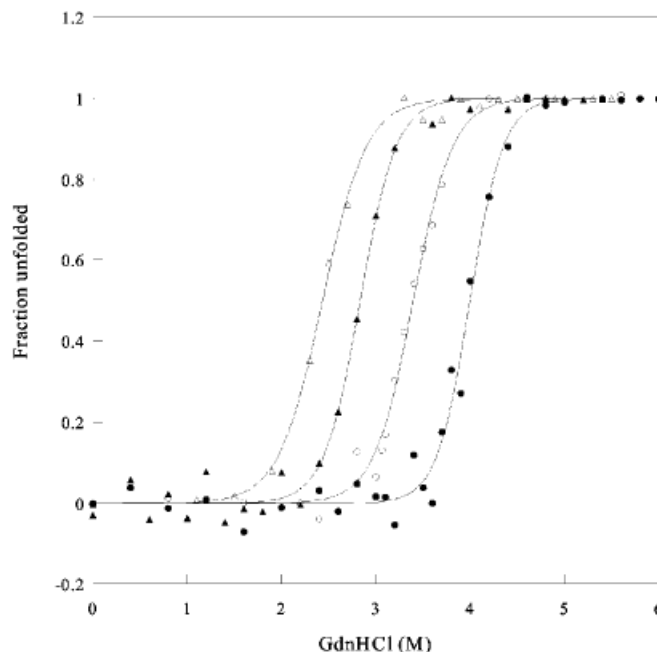


FIGURA 11. Desnaturalización química de T70N inducida por GdnHCl monitoreada por emisión de fluorescencia a  $\lambda = 340$  nm en un amortiguador de pH=6.5 a temperatura constante  $T = 20$  grados Celsius. Variante silvestre (●); T70N (○); D67H (▲) y I56T (△)

De acuerdo al MEZ, el espectro de energías  $E$  de la macromolécula está dado por la ecuación

$$(6.1) \quad E_{\alpha} = (K - \alpha)U - E^* \delta_{K-\alpha,0},$$

en donde  $\alpha$  representa el número de aminoácidos en el estado nativo y  $E^*$  representa la energía del estado base. A fin de poder escribir la Ec (6.1) para el caso de la desnaturalización de T70N requerimos conocer la dependencia funcional de la variable  $\alpha$  en términos de la concentración  $\zeta$  de GdnHCl. Para este fin consideraremos la curva experimental de desnaturalización reportada por Esposito *et al* (Figura 11) y definiremos la variable  $\alpha' = \alpha/K$  como la fracción de aminoácidos existentes en una configuración dada. Si aceptamos que la curva de desnaturalización representa a la variable  $\alpha'$  en función de  $\zeta$ , entonces podemos proponer una forma funcional general que reproduzca un comportamiento de tipo sigmoidal como el observado en la Figura (11). La función que propondremos es la siguiente,

$$(6.2) \quad \frac{\alpha}{K} \equiv \alpha' = p - q \tanh(\zeta - r),$$

la cual corresponde a la función tangente hiperbólica más general que puede construirse. Las cantidades  $p, q$  y  $r$  representan en términos matemáticos el escalamiento de la función a lo largo de los ejes correspondientes y, tal como veremos a continuación, pueden escribirse en términos de las cantidades experimentales que caracterizan

la desnaturalización de T70N.

A fin de escribir a la tripleta  $(p, q, r)$  en términos de cantidades físicas, debemos observar que para que la ecuación (6.2) represente adecuadamente a la curva en la Figura (11), ésta debe cumplir con los siguientes requisitos,

- ◇ En forma estricta, en ausencia del solvente GdnHCl la fracción de estructura nativa  $\alpha'$  debe ser la unidad. Sin embargo, debido a que la función tangente hiperbólica está definida a través de funciones exponenciales las cuales exhiben un comportamiento asintótico conforme  $\zeta$  se aproxima a las cotas  $+\infty$  y  $-\infty$ , utilizaremos esta propiedad para sustituir esta condición por

$$(6.3) \quad \lim_{\zeta \rightarrow -\infty} \alpha'(\zeta) = 1.$$

La sustitución anterior no representa mayor problema dado que la función  $\tanh(x)$  converge rápidamente a su asíntota superior, por lo que  $\lim \alpha' \rightarrow 1$  para valores muy cercanos a cero.

- ◇ Diversos estudios experimentales indican que en el caso de desnaturalización inducida químicamente, la mayoría de las proteínas no se desnaturalizan completamente aún en presencia de concentraciones grandes de soluto orgánico [Ghelis and Yon, 1982]. Para incorporar este hecho matemáticamente, diremos que la estructura desplegada de la proteína se alcanza sólo en el límite de concentraciones infinitas. Esto es,

$$(6.4) \quad \lim_{\zeta \rightarrow +\infty} \alpha'(\zeta) = 0.$$

- ◇ Finalmente, consideraremos que si bien no es posible alcanzar la desnaturalización completa del estado  $N$  para una concentración finita, existe (tal como lo demuestran los experimentos) un valor crítico  $\zeta^*$  para el cual la actividad biológica de la proteína se ha reducido a cero. Este valor, para el caso de T70N corresponde a  $\zeta^* = 6$  Molar. La cantidad  $\zeta^*$  es un indicador que marca la frontera entre nuestra capacidad para distinguir experimentalmente el estado desplegado de algún remanente del estado nativo. Si existe tal remanente del estado  $N$  por arriba de la concentración  $\zeta^*$  no es detectable mediante las técnicas experimentales correspondientes. Así pues, diremos que cuando se ha alcanzado este valor, la fracción de estructura nativa corresponde a la cantidad  $\epsilon$ , la cual evidentemente es un número adimensional extremadamente pequeño. En la notación que hemos establecido, lo anterior se escribe como,

$$(6.5) \quad \alpha'(\zeta^*) = \epsilon, \quad \text{para } \epsilon \rightarrow 0$$

Las condiciones anteriores definen un sistema consistente para calcular los valores de la tripleta  $(p, q, r)$  en términos de  $(K, \zeta^*, \epsilon)$ , por lo que como resultado de aplicar las ecuaciones (6.3), (6.4) y (6.5) a la ecuación (6.2), obtenemos que

$$(6.6) \quad p = q = \frac{1}{2}, \quad \text{y que}$$

$$(6.7) \quad r = \zeta^* + \frac{1}{2} \log \left( \frac{1 - \frac{p-\epsilon}{q}}{1 + \frac{p-\epsilon}{q}} \right)$$

por lo que entonces,

$$(6.8) \quad r = \zeta^* + \frac{1}{2} \log \left( \frac{\epsilon}{1 - \epsilon} \right).$$

La ecuación (6.8) muestra la naturaleza física de la variable  $r$ . Dado que está escrita en términos de  $\zeta^*$ , representa entonces una concentración y dado que  $r$  no puede ser más grande que  $\zeta^*$ , este mismo resultado nos da una nueva cota superior para  $\epsilon < 1/2$ , en consistencia con la relación (6.5). Además de lo anterior, es posible ver a partir de las ecuaciones (6.2) y (6.6) que la variable  $r$  no sólo es una concentración sino que de hecho corresponde al punto medio de la transición  $N \rightarrow D$ , esto es, al valor de  $\zeta$  tal que  $\alpha = K/2$ . Esta es una cantidad que se mide experimentalmente, por ejemplo en el caso de T70N,  $r$  corresponde al valor  $3.4 \pm 0.1$  Molar. Una vez identificada esta variable, la escribiremos en la forma estándar de la literatura, en donde se denota por  $C_{1/2}$ . Dado que  $\epsilon$  representa la fracción de estructura nativa remanente una vez que se ha alcanzado la concentración crítica  $\zeta^*$ , podemos ajustar este valor para representar cualquier curva de desnaturalización dada su conexión con la variable  $r$  definida por (6.8). Si elegimos  $\epsilon = 2/3K$ , lo cual implica con  $K = 130$  que  $\epsilon \sim 5 \times 10^{-3}$ , la concentración  $C_{1/2}$  para la desnaturalización de T70N en este esquema teórico queda como  $C_{1/2} = 3.36$ , que es un valor dentro del margen de error experimental reportado. Sustituyendo entonces las ecuaciones (6.6) y (6.8) en (6.5) podemos escribir que la variable  $\alpha$  se escribe en términos de  $\zeta$  de acuerdo a,

$$(6.9) \quad \alpha = \frac{K}{2} [1 - \tanh(\zeta - C_{1/2})],$$

la cual no depende ya de los parámetros  $(p, q, r)$  y se escribe solamente en función de cantidades observadas experimentalmente. La Figura (12) muestra un gráfico de la Ecuación (6.9) en el caso de T70N, en donde  $K = 130$  y  $C_{1/2} = 3.36$  Molar.

Para escribir la expresión correspondiente al espectro de energías  $E$  en función de la concentración, debemos observar que en la expresión original de Zwanzig, el estado nativo está representado mediante un pozo energético finito de magnitud  $E^*$ . En el caso de la desnaturalización lineal discutida anteriormente, utilizamos una función delta de Dirac para representar la energía de este estado, sin embargo en este análisis, sustituiremos la función delta de Dirac por una aproximación finita. Para ello, utilizaremos a las funciones secuencia- $\delta$ , las cuales constituyen una aproximación finita a la función delta de Dirac. De hecho, es posible demostrar que una secuencia- $\delta$ , denotada aquí como  $S_n(x)$ , tiene propiedades similares a la función  $\delta(x)$  respecto de un operador integral de Riemann [Butkov., 1968]. Esto es,

$$(6.10) \quad \lim_{n \rightarrow \infty} \int S_n(x) f(x) dx = f(0).$$

Existen diversas funciones  $S_n(x)$  que satisfacen esta propiedad, además de exhibir un máximo pronunciado para un valor en el dominio de la función, por lo que para este caso utilizaremos a la función

$$(6.11) \quad S_n(x) = \frac{n}{\sqrt{\pi}} e^{-x^2 n^2},$$

como una representación finita e integrable de la energía del estado nativo. En la ecuación (6.11), el variable  $n$  es un parámetro ajustable que define la anchura del máximo de  $S_n(x)$ . La Figura (13) muestra el perfil de la función  $S_n$  definida por la ecuación (6.11) para distintos valores de  $n$  en el dominio  $[-1, 1]$ . Sustituyendo

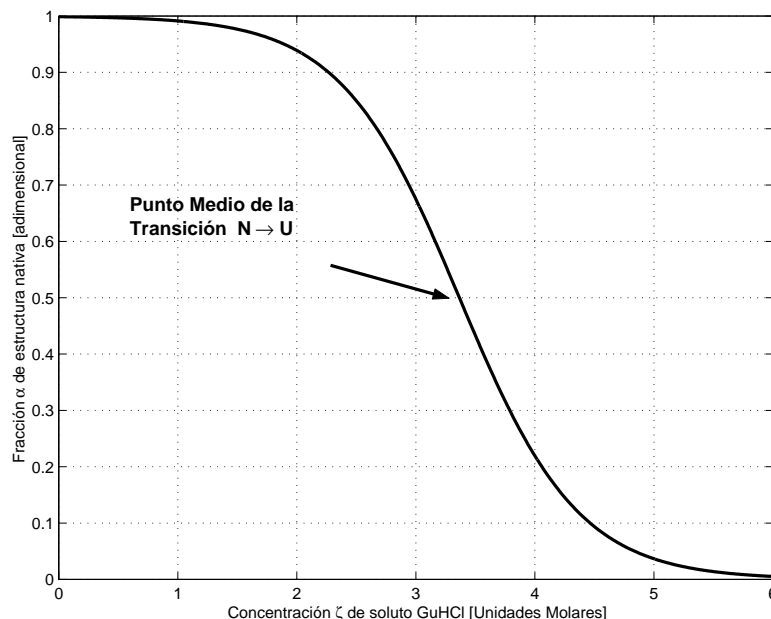


FIGURA 12. Gráfico de la ecuación (6.9) que describe la desnaturalización de T70N en función de la concentración  $\zeta$  de soluto orgánico. La concentración crítica  $\zeta^* = 6$  Molar, en tanto que  $K = 130$  y  $C_{1/2} = 3.36$ . Para  $\zeta = 0$ , la fracción  $\alpha' = 0.998$  en tanto que para  $\zeta = \zeta^*$ , la fracción  $\alpha' = 0.005$ .

entonces el término delta de Kronecker del modelo original de Zwanzig (Ecuación 6.1) por la función secuencia- $\delta$ , (Ecuación 6.11) obtenemos la forma del espectro de energías para el caso de una variable continua,

$$(6.12) \quad E_\alpha = (K - \alpha)U - \frac{nE^*}{\sqrt{\pi}} e^{-(K-\alpha)^2 n^2}.$$

en donde los símbolos  $K$ ,  $E^*$ ,  $U$  y  $n$  tienen el significado físico discutido anteriormente. Es importante recalcar que en la ecuación (6.12) la variable  $\alpha$  es una función continua de  $\zeta$ , por tanto  $E = E(\zeta)$ . Insertando entonces la forma funcional que hemos deducido para representar a la desnaturalización química de T70N (Ecuación 6.9) en (6.12) se obtiene la forma final para el espectro *continuo* de energías de la macromolécula en función de la concentración de soluto orgánico, a saber,

$$(6.13) \quad E(\zeta) = \frac{KU}{2} [1 + \tanh(\zeta - C_{1/2})] - \tilde{E}^* \exp \left[ \frac{-K^2 n^2}{4} (1 + \tanh(\zeta - C_{1/2}))^2 \right],$$

donde hemos definido la cantidad  $E^*$  como,

$$(6.14) \quad \tilde{E}^* \equiv \frac{nE^*}{\sqrt{\pi}}.$$

Para efectuar un análisis de la ecuación (6.13) consideraremos que el valor absoluto de la energía del estado nativo,  $E^*$  puede relacionarse con la energía de los estados no nativos  $U$  y que de hecho es un múltiplo de ella. De forma tal, que a partir de este momento consideraremos entonces que  $E^* = KU$ . Con esta suposición, para una proteína de tamaño fijo  $K$ , los niveles de energía  $E$  sólo dependen de la concentración  $\zeta$  de soluto y de la energía  $U$  de cada aminoácido en un estado no nativo. Observemos

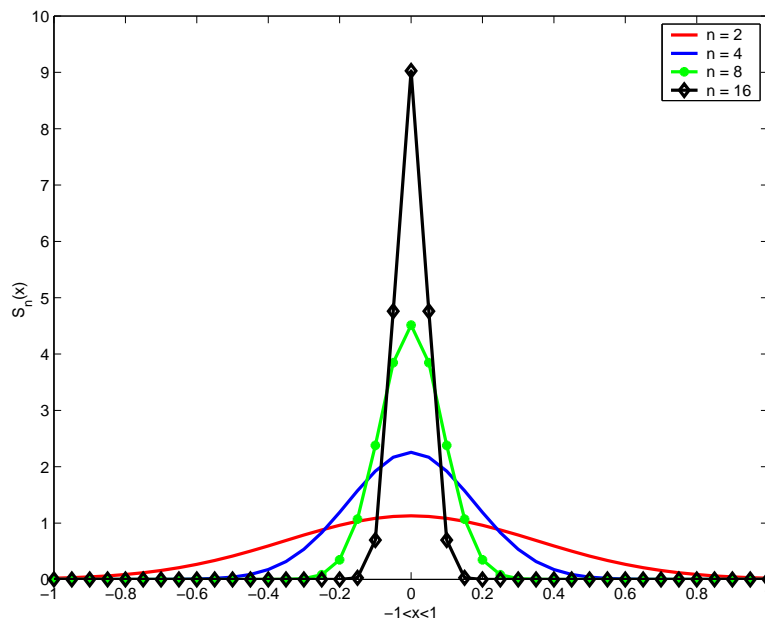


FIGURA 13. Aproximación de la función delta de Dirac a través de las funciones secuencia- $\delta$ ,  $S_n(x)$ . En la gráfica se muestra el comportamiento de  $S_n(x) = \frac{n}{\sqrt{\pi}} e^{-x^2 n^2}$  en función del parámetro  $n$ . Se observa que la función  $S_n$  muestra un máximo significativo para  $x = 0$  el cual nos permitirá utilizarla como una representación del pozo energético en donde se halla el estado nativo.

también que dado que la energía del estado base ha sido seleccionada como  $-KU$  y cada aminoácido en un estado no nativo contribuye con una energía  $+U$ , la diferencia energética entre el estado base y cualquier otro es proporcional a  $U$ , por lo que esta variable representa la diferencia energética entre los estados nativo y desplegado.

La Figura (14) muestra un gráfico tridimensional de la energía  $E$  del sistema en función de  $U$  y de  $\zeta$ , de acuerdo a la ecuación (6.13). Tal como mencionamos en el párrafo anterior, la variable  $U$  está asociada directamente con diferencias energéticas entre el estado nativo y los demás estados metaestables. Una de las consecuencias de la ecuación (6.13) es que conforme esta diferencia se incrementa, la transición entre los estados de mayor energía, para  $\zeta \rightarrow \zeta^*$  y los estados de menor energía, caracterizados por  $\zeta \rightarrow 0$ , tiende a ser más pronunciada, indicando entonces que en el formalismo que estamos desarrollando la existencia de transiciones cooperativas (*all-or-none transitions*) es una consecuencia de la forma estructural del paisaje energético y de la magnitud de la variable  $U$  que separa al estado base del sistema de los demás. De acuerdo a Finkelstein y Pitytsin, justo esta es una característica que diferencia a las transiciones conformacionales entre las cadenas heteropoliméricas inorgánicas y las macromoléculas biológicas. En éstas últimas, la presencia de separaciones energéticas significativas y la cooperatividad de la transición impiden que la proteína quede atrapada en un estado intermedio, como sucede en la mayoría de las cadenas heteropoliméricas inorgánicas [Finkelstein and Pitytsin, 2002].



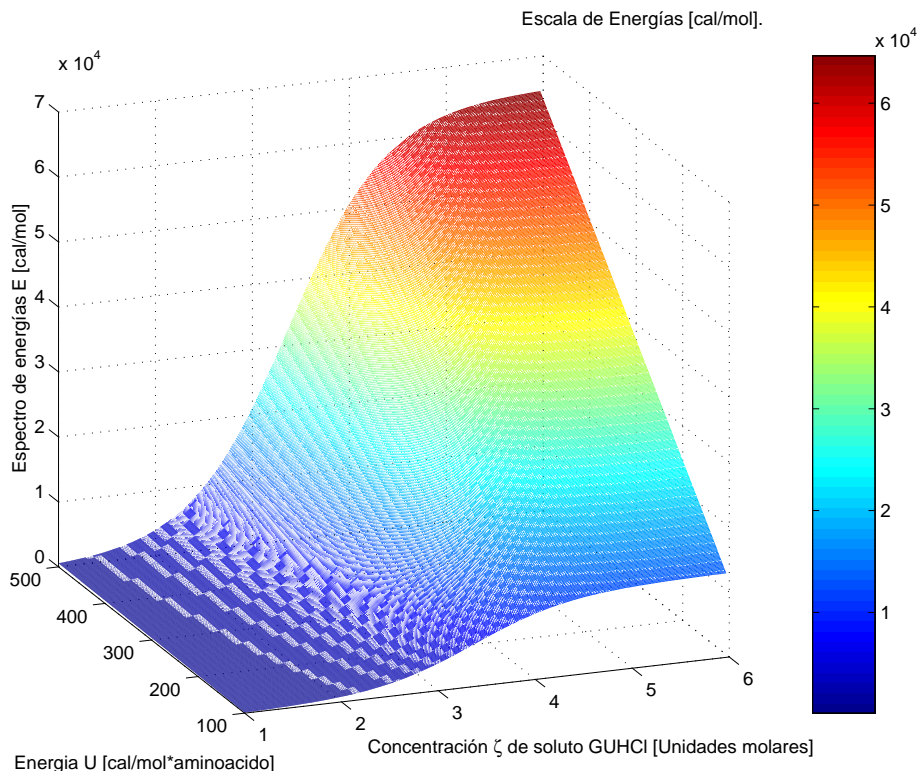


FIGURA 14. Gráfico en tres dimensiones del espectro de energía accesible a la macromolécula de acuerdo a la Ecuación (6.13). Es importante notar que conforme  $U$  se incrementa, la transición entre los estados nativo y desnaturalizado tiende a ser más pronunciada, indicando entonces que en el MEZ, el mecanismo de desnaturalización está intrínsecamente asociado a la existencia de separaciones energéticas significativas entre los estados nativo y desplegado. La gráfica anterior se realizó para  $K = 130$ ,  $C_{1/2} = 3.36$  y  $n = 8$  en la ecuación (6.13).

Una vez que hemos determinado el espectro de energías en función de la concentración, procedemos a escribir la generalización correspondiente a la función de partición  $Z$ , la cual en el esquema original de Zwanzig está escrita en términos de la variable discreta  $\alpha$  como una suma sobre el espacio configuracional de la macromolécula (Ecuación 5.5. En esta expresión, la degeneración  $g_\alpha$  asociada a cada nivel de energía está representada por el coeficiente binomial,

$$(6.15) \quad g_\alpha = \nu^{K-\alpha} \binom{K}{\alpha}.$$

en donde  $\alpha$  es una variable discreta. Dado que en el MEZ,  $\alpha$  es una función continua de la concentración, la traducción correcta de la ecuación (6.15) implica dos consideraciones. Por un lado, la sustitución obvia del operador  $\sum$  por un operador integral  $\int$  y, por otro lado la sustitución del coeficiente binomial por una expresión continua en  $\zeta$ . Para hacer esto último, utilizamos la fórmula de Stirling, la cual establece que

[Arfken., 1985]

$$(6.16) \quad K! \sim \sqrt{(2\pi)} K^{K+1/2} \exp(-K).$$

Utilizando la ecuación (6.16) es posible escribir al coeficiente binomial de (6.15) como

$$(6.17) \quad \binom{K}{\alpha} = \frac{K!}{\alpha!(K-\alpha)!} \simeq \frac{1}{\sqrt{2\pi}} \left( \frac{K^{K+\frac{1}{2}}}{(K-\alpha)^{(K-\alpha+\frac{1}{2})} \alpha^{(\alpha+\frac{1}{2})}} \right).$$

la cual es una fórmula válida incluso para valores de  $K$  del orden de  $K \sim 10^1$ . Para mostrar la validez de la aproximación (6.17), la Figura (15) muestra en el lado izquierdo una gráfica de la superposición del coeficiente binomial calculado en forma exacta para  $0 < \alpha < K$  con  $K = 130$  y de la aproximación hecha a través de la ecuación (6.17), en tanto que en el lado derecho se tiene una gráfica del error relativo que existe entre ambas expresiones. El error relativo  $\Xi$ , calculado como,

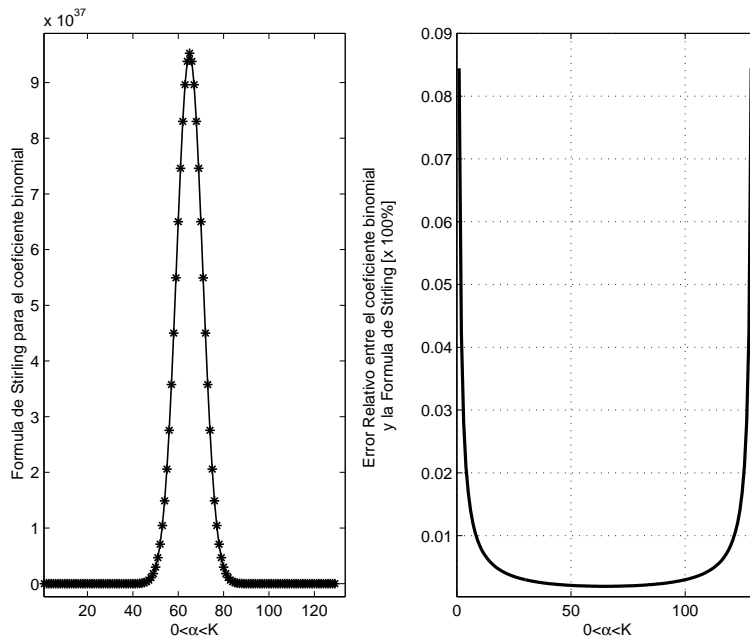


FIGURA 15. Aproximación de Stirling para el coeficiente binomial con  $0 < \alpha < K = 130$ . El lado izquierdo muestra una superposición entre el coeficiente binomial  $\binom{K}{\alpha}$  (línea sólida) y la aproximación de Stirling hecha a través de la Ecuación (6.17) (puntos). El lado derecho es una gráfica del error relativo entre ambas expresiones.

$$(6.18) \quad \Xi = \frac{\frac{1}{\sqrt{2\pi}} \left( \frac{K^{K+\frac{1}{2}}}{(K-\alpha)^{(K-\alpha+\frac{1}{2})} \alpha^{(\alpha+\frac{1}{2})}} \right) - \binom{K}{\alpha}}{\binom{K}{\alpha}}$$

muestra que la aproximación hecha es perfectamente válida para el valor de  $K$  utilizado<sup>3</sup>. El error relativo  $\Xi$  se aproxima a cero para valores de  $\alpha \rightarrow K/2$ , esto es hacia el punto medio de la distribución binomial, que es donde el factor de degeneración

<sup>3</sup>La deducción formal de (6.17) implica que es válido incluso para  $K$  menores a la seleccionada aquí.

$g_\alpha$  alcanza su máximo valor, en tanto que en los extremos de la distribución, el error  $\Xi$  es del orden del 8 %. Dado que el máximo margen de error en la aproximación de Stirling se observa en los extremos de la distribución y es justamente ahí donde  $g_\alpha$  se aproxima a cero, consideraremos que este argumento nos permite utilizar la aproximación para describir adecuadamente el comportamiento de  $g_\alpha$ , particularmente en la región central de la distribución, en donde  $\Xi \rightarrow 0$ .

En acuerdo con los argumentos y análisis presentados, podemos escribir la función de partición  $Z$  del sistema como,

$$(6.19) \quad Z = \frac{K^{K+\frac{1}{2}}}{\sqrt{2\pi}} \int_0^{\zeta^*} \frac{\nu^{\frac{K}{2}(1+\tanh(\zeta-C_{1/2}))} \exp\left[\frac{-\beta KU}{2}(1+\tanh(\zeta-C_{1/2}))\right]}{\left[\frac{K}{2}(1-\tanh(\zeta-C_{1/2}))\right]^{\frac{1}{2}[1+K(1-\tanh(\zeta-C_{1/2}))]} \times} \\ \times \frac{\exp\left[\beta \tilde{E}^* \exp\left(-\frac{K^2 n^2}{4}[1+\tanh(\zeta-C_{1/2})]^2\right)\right]}{\left[\frac{K}{2}(1+\tanh(\zeta-C_{1/2}))\right]^{\frac{1}{2}+\frac{K}{2}(1+\tanh(\zeta-C_{1/2}))}} d\zeta,$$

en donde se han sustituido las expresiones para el espectro de energías  $E(\zeta)$  (Ecuación 6.13) y la fórmula de Stirling para el factor  $g_\zeta$  (Ecuaciones 6.15 y 6.17) respectivamente, en la ecuación (5.8).

Debido a la complejidad manifiesta en el integrando de la función de partición  $Z$  se procedió en primera instancia a un análisis computacional del integrando  $I$ . Un conjunto extenso de simulaciones numéricas que exploran el comportamiento de  $I$  versus  $\zeta$  muestran que para distintos valores de  $U$ ,  $\nu$  y temperaturas entre 273 – 373 Kelvin, el integrando  $I$  muestra una contribución máxima para algún valor de  $\zeta$ , que evidentemente depende de las variables anteriores. Este valor  $\zeta^\dagger$ , obtenido como solución numérica de la ecuación

$$(6.20) \quad \left(\frac{\partial I}{\partial \zeta}\right)_{\zeta=\zeta^\dagger} = 0,$$

satisface la ecuación,

$$(6.21) \quad \frac{K}{2} \log \nu - \\ + \frac{\beta KU}{2} - \frac{\beta \tilde{E}^* K^2 n^2}{2} \left[1 + \tanh(\zeta^\dagger - C_{1/2})\right] \exp\left(\frac{K^2 n^2}{4}[1 + \tanh(\zeta^\dagger - C_{1/2})]^2\right) + \\ - \frac{K}{2} \left(\log \left[\frac{1 + \tanh(\zeta^\dagger - C_{1/2})}{1 - \tanh(\zeta^\dagger - C_{1/2})}\right] - \frac{2 \tanh(\zeta^\dagger - C_{1/2})}{K(1 - \tanh^2(\zeta^\dagger - C_{1/2}))}\right) = 0.$$

La Figura (16) muestra la curva  $I$  versus  $\zeta$  para el caso de T70N, en donde  $K = 130$ ,  $C_{1/2} = 3.36$  M y  $\zeta^* = 6$  Molar, una temperatura  $T = 293$  K, un valor  $U = 360$  cal/mol y  $\nu = 2$ . Para los valores anteriores, la solución numérica de la ecuación (6.21) arroja que el máximo del integrando ocurre para el valor  $\zeta^\dagger = 3.38$  M. No está demás insistir que este resultado es aplicable sólo para los valores descritos anteriormente; distintos valores de las variables  $U$ ,  $\nu$ ,  $C_{1/2}$ ,  $U$  y  $T$  generarán un valor distinto de  $\zeta^\dagger$ . Lo que sí es común a todos estos casos es la existencia de un valor resonante para el integrando en función de la concentración. El origen de este pico resonante en el integrando  $I$ , que como consecuencia implica una contribución máxima también en la función de partición  $Z$  puede comprenderse directamente a partir de los postulados

de la mecánica estadística de equilibrio. En ésta, uno de los postulados fundamentales garantiza que, en condiciones de equilibrio termodinámico, existe un macroestado cuya multiplicidad de microestados es máxima [Garcia-Colin, L.S., 2005b].

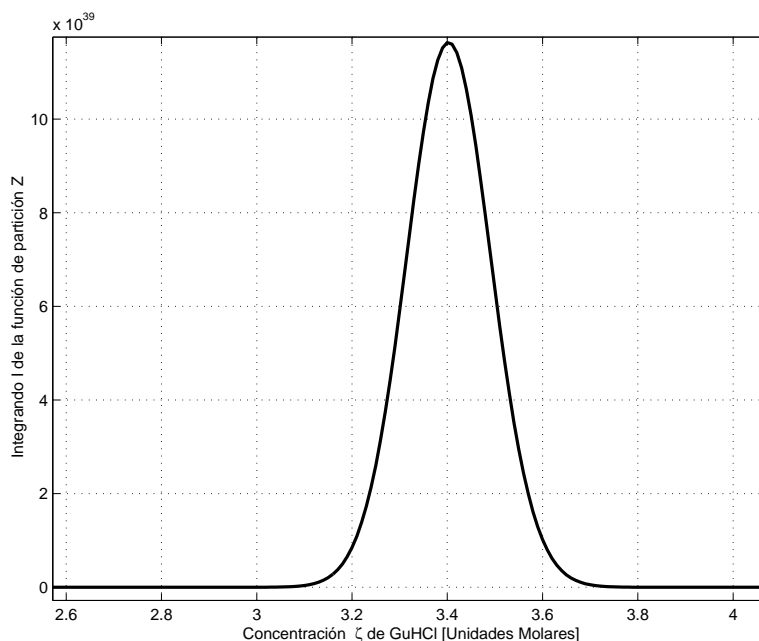


FIGURA 16. Una exploración numérica extensiva del comportamiento de  $I$  en función de  $\zeta$  muestra que en todos los casos analizados, éste exhibe un perfil como el mostrado en la gráfica, en donde existe un valor  $\zeta^\dagger$  para el cual  $I$  es máximo. Para el caso  $K = 130$ ,  $C_{1/2} = 3.36$ ,  $\zeta^* = 6$  M,  $U = 360$  cal/mol y  $\nu = 2$ , el máximo de  $I$  ocurre en  $\zeta^\dagger = 3.38$  Molar

La existencia de este máximo en  $I$  nos permitirá hacer una aproximación a primer orden para calcular la integral indicada en (6.19). Utilizando el método de cuadratura Gaussiana, que en esencia divide el intervalo de integración en subintervalos iguales y construye rectángulos cuya altura está determinada por el valor de la función, podemos aproximar la integral anterior como el área bajo un rectángulo cuyo altura corresponde a  $I(\zeta^\dagger)$  y base  $\Delta_\zeta$ . Entonces, a primer orden de aproximación,

$$(6.22) \quad Z = \frac{K^{K+1/2}}{\sqrt{2\pi}} \mathbb{Z},$$

en donde  $\mathbb{Z}$  se escribe como,

$$(6.23) \quad \mathbb{Z} = \Delta_\zeta \frac{\nu^{\frac{K}{2}} [1 + \tanh(\zeta^\dagger - C_{1/2})] \exp\left[-\frac{\beta K U}{2} (1 + \tanh(\zeta^\dagger - C_{1/2}))\right]}{\left[\frac{K}{2} (1 - \tanh(\zeta^\dagger - C_{1/2}))\right]^{\frac{1}{2}} [1 + K(1 - \tanh(\zeta^\dagger - r))]} \\ \times \frac{\exp\left[\beta \tilde{E}^* \exp\left[-\frac{K^2 n^2}{4} (1 + \tanh(\zeta^\dagger - C_{1/2}))^2\right]\right]}{\left[\frac{K}{2} (1 + \tanh(\zeta^\dagger - C_{1/2}))\right]^{\frac{1}{2} + \frac{K}{2}} [1 + \tanh(\zeta^\dagger - C_{1/2})]}$$

A través de la relación  $F = -k_B T \log Z$ , la energía libre del sistema  $F$  se escribe entonces como

$$(6.24) \quad F = -RT \left( \frac{K}{2} \log \nu - \frac{\beta KU}{2} + \left(K + \frac{1}{2}\right) \log K - \log \sqrt{2\pi} + \log \Delta_\zeta \right) + \\ - RT \beta \tilde{E}^* \exp \left( \frac{-n^2 K^2}{4} \left[ 1 + \tanh(\zeta^\dagger - C_{1/2}) \right]^2 \right) + \\ - RT \tanh(\zeta^\dagger - C_{1/2}) \left\{ \frac{K}{2} \log \nu - \frac{\beta KU}{2} + \frac{K}{2} \log \left[ \frac{1 + \tanh(\zeta^\dagger - C_{1/2})}{1 - \tanh(\zeta^\dagger - C_{1/2})} \right] \right\} + \\ + \frac{RT}{2} \left\{ \log \left[ \frac{K^2}{4} \left( 1 - \tanh^2(\zeta^\dagger - C_{1/2}) \right) \right] + K \log \left[ \frac{K^2}{4} \left( 1 - \tanh^2(\zeta^\dagger - C_{1/2}) \right) \right] \right\}.$$

La ecuación (6.24), que representa el paisaje energético de una proteína que sigue un perfil de desnaturalización de tipo sigmoidal como es el caso de T70N, constituye el resultado central del primer eje de investigación en este trabajo doctoral [Olivares-Quiroz L., and Garcia-Colin L.S., 2007]. Su deducción se ha sustentado a lo largo de esta sección y en los siguientes párrafos centraremos la atención en las implicaciones que pueden derivarse de él.

En primera instancia discutiremos el papel de los grados de libertad de la macromolécula, representados en el MEZ por la variable  $\nu$ . De acuerdo a lo que mencionamos al inicio de este capítulo, el MEZ, al igual que el modelo original de Zwanzig considera que cada aminoácido puede adquirir en promedio un número  $\nu + 1$  de posibles conformaciones,  $\nu$  de las cuales están asociadas con estados metaestables y la configuración restante con la posición espacial que ocupa el aminoácido en cuestión en el estado nativo. El primer resultado de interés en este sentido que puede extraerse de la ecuación (6.24) se refiere al papel que juega esta variable en la estabilidad del estado nativo. Observemos que el valor de la diferencia de energías  $\Delta_N^D F$  entre los estados nativo y desnaturalizado es un indicador de la ubicación del estado  $N$  dentro del paisaje energético y por consecuencia de su estabilidad energética. A medida que la diferencia  $\Delta_N^D F$  se incrementa, el estado  $N$  se halla necesariamente en estratos de mayor profundidad dentro del paisaje energético, dado que se requiere una mayor cantidad de energía para inducir la desnaturalización. Paralelo a lo anterior es posible afirmar también que en este caso, el estado  $N$  tiene un grado de estabilidad mayor que aquel caso en donde la diferencia  $\Delta_N^D F$  tuviese un valor menor. En el resultado que hemos derivado, la diferencia  $\Delta_N^D F$  tiene dos contribuciones asociadas con los grados de libertad de la macromolécula. Estos términos, representados por las expresiones,

$$(6.25) \quad F_{\nu,1} = \frac{RTK}{2} \log \nu,$$

y

$$(6.26) \quad F_{\nu,2} = \frac{RTK}{2} \tanh(\zeta^\dagger - C_{1/2}) \log \nu.$$

implican que la variable  $\nu$  juega un papel relevante en la ubicación y estabilidad del estado  $N$ . Dado que el primer término  $-F_{\nu,1}$  es siempre negativo para  $\nu > 1$ , su contribución a  $\Delta_N^D F$  tiende a ubicar al estado  $N$  en regiones de energía más negativa conforme  $\nu$  se incrementa. Por otro lado, el término  $F_{\nu,2}$ , que si bien puede ser negativo o positivo dependiendo de la diferencia de concentraciones  $(\zeta^\dagger - C_{1/2})$ , está acotado

debido a la función tangente hiperbólica, entre los valores  $[-F_{\nu,1}, F_{\nu,1}]$ , haciendo que aún en el caso de que la concentración  $\zeta^\dagger$  se desplazase en forma extrema hacia los valores  $\zeta = 0$  o bien  $\zeta = \zeta^*$ , sólo tuviese como efecto anular la contribución del primer término. Aunque no tenemos argumentos para afirmar que este caso no pueda ocurrir, el conjunto de simulaciones realizadas para intervalos de temperaturas fisiológicas y valores de la separación energética en el intervalo  $[0.3, 30]$  kcal/mol, que corresponden a las cotas energéticas establecidas experimentalmente para la estabilidad del estado nativo, no muestran evidencias en esta dirección. Tal resultado nos motiva a pensar entonces que para las condiciones mencionadas, el término  $F_{\nu,2}$  es en general una contribución menor a  $\Delta_N^D F$ . Si aceptamos como válidas las consideraciones anteriores, el MEZ está en condiciones entonces de hacer una predicción teórica sobre la estabilidad del estado nativo en función de los grados de libertad del sistema, esto es, en función del espacio conformacional disponible para la cadena primaria de aminoácidos, la cual enunciamos a continuación:

*La estabilidad del estado nativo depende logarítmicamente del número de grados de libertad conformacionales de la secuencia de aminoácidos. Esto implica que el estado nativo de proteínas cuya secuencia de aminoácidos tenga un espacio conformacional de dimensión mayor, tenderá a ubicarse en regiones más profundas del paisaje energético y por tanto, a exhibir un mayor nivel de estabilidad ante fluctuaciones del medio.*

Esta afirmación indica entonces que debemos buscar el origen de la estabilidad del estado nativo de una proteína en las propiedades físicas de la secuencia de aminoácidos. En particular, en el número promedio de conformaciones que puede adquirir cada uno de ellos. El punto de vista anterior es también consistente con el hecho que la entropía configuracional  $S_{config}$  del sistema está directamente asociado con el logaritmo del número de microestados accesibles al sistema y que por tanto, un sistema cuya entropía  $S$  se incrementa, decrece necesariamente su energía libre  $F$ . Dado que los mapas de Ramachandran de una proteína, ya sean calculados teóricamente mediante potenciales de interacción de tipo van der Waals o bien obtenidos mediante mediciones experimentales, nos proveen información sobre el espacio conformacional de una proteína en su estado nativo, la confrontación entre los mapas de Ramachandran y los datos termodinámicos para la desnaturalización de una macromolécula podrían arrojar indicios sobre la veracidad del resultado que aquí hemos formulado. Por otro lado, el caso  $\nu = 1$ , el cual corresponde a una proteína cuya secuencia de aminoácidos solo puede adquirir en promedio dos configuraciones, el estado  $N$  y el estado  $D$ , generaría que una contribución idénticamente cero de los términos  $F_{\nu,1}$  y  $F_{\nu,2}$ . Sin embargo, éste caso correspondería a una proteína con un espacio conformacional altamente restringido en donde la macromolécula no exhibe tal libertad conformacional en sus componentes. El MEZ está diseñado para explorar las implicaciones de esta propiedad, por lo que este caso no resulta de interés.

La ecuación (6.24) también nos permite una digresión sobre el comportamiento de  $\Delta_N^D F$  en función de la concentración  $C_{1/2}$  la cual representa el punto medio de la transición  $N \rightarrow D$ . Este análisis resulta interesante dado que se ha publicado en la literatura una cantidad considerable de información a este respecto. Para proceder a este análisis debemos observar que al considerar que la integral para la función de partición  $Z$  puede representarse mediante su contribución máxima, de acuerdo a la

ecuación (6.22), hemos introducido dos variables adicionales en la expresión para la energía libre  $F$ . Estas variables son la concentración máxima  $\zeta^\dagger$  donde ocurre el pico resonante de  $Z$  y el ancho  $\Delta_\zeta$  de éste. Estas dos variables dependen de la concentración  $C_{1/2}$ , dado que la ubicación del pico resonante es la solución de la ecuación (6.21), la cual contiene a la cantidad  $C_{1/2}$ .

Para determinar el perfil de la variable  $\zeta^\dagger$  en función de la concentración  $C_{1/2}$ , se resolvió numéricamente la ecuación (6.21) para distintos valores de la variable  $C_{1/2}$ , manteniendo constantes los demás parámetros. Los resultados de este procedimiento se muestran en forma explícita en el Cuadro (2) y en forma gráfica en la figura (17).

$C_{1/2}$	$\zeta^\dagger$
1	1.024
1.5	1.524
2	1.875
2.5	2.524
3	3.024
3.5	3.524
4	4.025
4.5	4.526
5	5.029
6	5.31

CUADRO 2. Valores para la concentración  $\zeta^\dagger$  en donde ocurre el máximo de la función de partición  $Z$  en función de la concentración  $C_{1/2}$ . Estos valores se obtuvieron a través de la solución numérica de la ecuación (6.21) la cual define el valor de  $\zeta^\dagger$  en función de  $C_{1/2}$ .

Realizando un ajuste a un polinomio de grado uno para la variable  $\zeta^\dagger$  en términos de  $C_{1/2}$  del tipo

$$(6.27) \quad \zeta^\dagger = a_0 + a_1 C_{1/2}$$

se obtiene que las constantes  $a_0 = 0.14$  y  $a_1 = 0.9$ . Este análisis nos permite entonces representar a la concentración máxima  $\zeta^\dagger$  como una función lineal de la concentración  $C_{1/2}$  de acuerdo a la ecuación (6.27), con los valores para las constantes  $a_0$  y  $a_1$  determinadas previamente.

Para calcular la integral de la función de partición  $Z$  hemos supuesto que, debido a la existencia de un máximo significativo en el intervalo de concentraciones considerado, el integral puede representarse mediante el área de un rectángulo cuya altura es  $I(\zeta^\dagger)$  y cuya base es  $\Delta_\zeta$ . Si la variable  $\zeta^\dagger$  se modifica en función de la concentración  $C_{1/2}$  tal como se demostró en los párrafos anteriores, es de esperarse que el ancho  $\Delta_\zeta$  también exhiba una dependencia en la concentración  $C_{1/2}$ . Para verlo con mayor claridad, definimos la base  $\Delta_\zeta$  en términos de los puntos de inflexión de la curva  $I(\zeta)$ . Esto es, si llamamos  $\zeta_{i,1}$  y  $\zeta_{i,2}$  a los puntos donde la segunda derivada de  $I(\zeta)$  es cero, entonces los puntos  $\zeta_{i,1}$  y  $\zeta_{i,2}$  satisfacen la ecuación

$$(6.28) \quad \left( \frac{\partial^2 I}{\partial \zeta^2} \right)_{\zeta=\zeta_i} = 0.$$

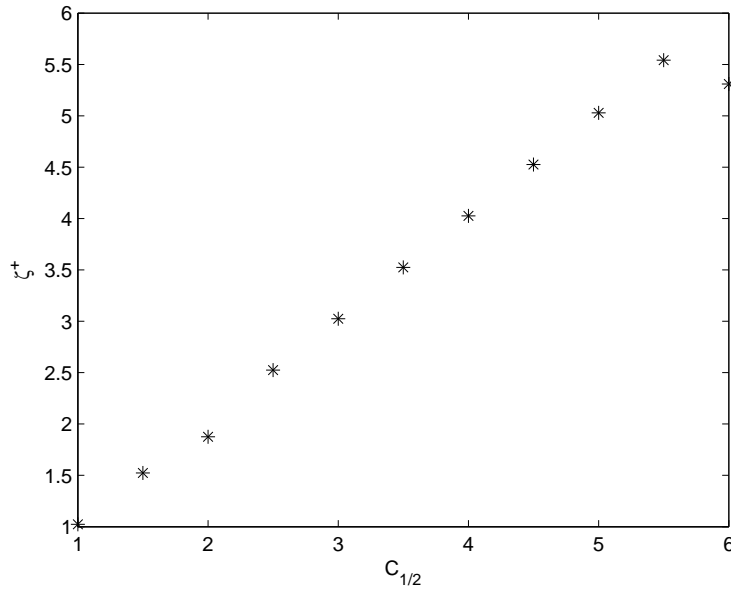


FIGURA 17. Gráfico de los datos para  $\zeta^\dagger$  en función de  $C_{1/2}$  provenientes de Cuadro (2). El ajuste de los datos a un polinomio de grado uno en la variable  $C_{1/2}$  del tipo  $\zeta^\dagger = a_0 + a_1 C_{1/2}$  genera los valores  $a_0 = 0.14$  y  $a_1 = 0.9$ .

El ancho  $\Delta_\zeta$  entonces se define a través de la ecuación,

$$(6.29) \quad \Delta_\zeta = \zeta_{i,1} - \zeta_{i,2}.$$

Puesto que  $\zeta_{i,1}$  y  $\zeta_{i,2}$  son soluciones de la ecuación (6.28), la cual depende de la variable  $C_{1/2}$ , el análisis anterior muestra que, efectivamente el ancho  $\Delta_\zeta$  del máximo de la función de partición depende de la concentración  $C_{1/2}$ .

La solución numérica de la ecuación (6.28) para diversos valores de  $C_{1/2}$  se muestran a continuación en el Cuadro (3). El algoritmo utilizado para calcular cada uno de los valores de  $\Delta_\zeta$  en función de  $C_{1/2}$  es el siguiente. En primera instancia se determina la segunda derivada de  $I$  respecto de  $\zeta$ , posteriormente se sustituyen los valores para los parámetros  $U$ ,  $K$ ,  $\nu$  y  $T$  considerados previamente. Se procede a una solución numérica de esta ecuación para determinar los valores de  $\zeta_{i,1}$  y  $\zeta_{i,2}$  y finalmente se utiliza la relación  $\Delta_\zeta = \zeta_{i,1} - \zeta_{i,2}$ .

La gráfica (18) muestra el comportamiento del logaritmo del ancho  $\Delta_\zeta$  en función de la concentración  $C_{1/2}$ . Dado el carácter lineal observado para  $\log \Delta_\zeta$  en función de  $C_{1/2}$ , se procedió a realizar un ajuste por mínimos cuadrados para una dependencia lineal entre  $\log \Delta_\zeta$  y  $C_{1/2}$  de acuerdo a la siguiente relación

$$(6.30) \quad \log \Delta_\zeta = \log A_0 + A_1 C_{1/2}$$

en donde  $\log A_0$  representa la ordenada al origen de la recta mostrada en la Figura (18) y  $A_1$  es la pendiente de la recta. El ajuste por mínimos cuadrados indica que  $A_0 = 0.006$  y que  $A_1 = 0.96$ . La ecuación (6.30) implica también que el ancho de la



$C_{1/2}$	$\Delta_\zeta$
2.5	0.08
3.0	0.12
3.32	0.16
3.49	0.20
3.71	0.24
4.0	0.32
4.31	0.44
4.5	0.54

CUADRO 3. Valores para el ancho  $\Delta_\zeta$  en donde ocurre el máximo de la función de partición  $Z$  en función de la concentración  $C_{1/2}$ .

curva sigue un crecimiento exponencial en términos de  $C_{1/2}$  de acuerdo a la relación,

$$(6.31) \quad \Delta_\zeta = A_0 \exp(A_1 C_{1/2}).$$

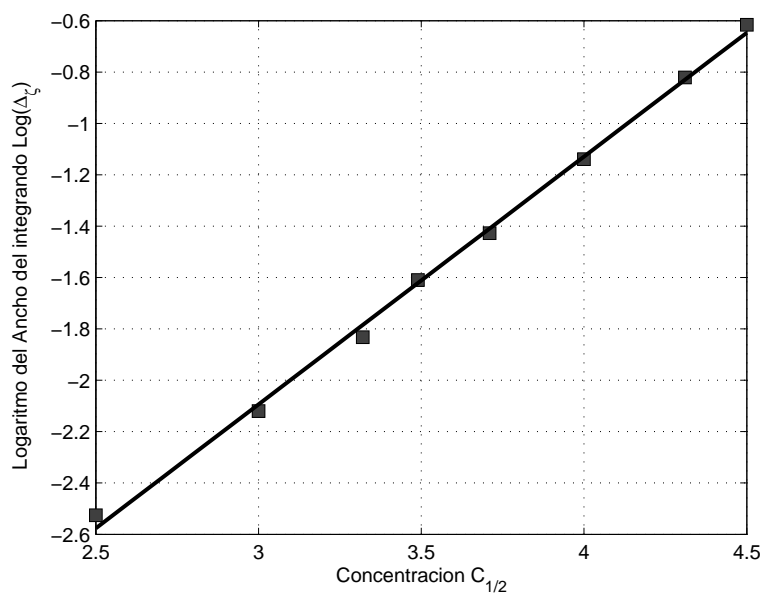


FIGURA 18. Logaritmo del ancho  $\Delta_\zeta$  en función de la concentración  $C_{1/2}$  obtenidos como una solución a la ecuación (6.28) y (6.29). Los puntos (■) representan los valores numéricos obtenidos en tanto que la línea sólida es un ajuste por mínimos cuadrados a un polinomio de grado uno de la forma  $\log \Delta_\zeta = \log A_0 + A_1 C_{1/2}$ . El ajuste por mínimos cuadrados nos permite determinar los valores  $A_0 = 0.006$  y que  $A_1 = 0.96$

Los resultados anteriores indican que tanto la ubicación del máximo  $\zeta^\dagger$  de la función de partición, como el ancho  $\Delta_\zeta$  del mismo son funciones de la concentración

$C_{1/2}$ . Las expresiones deducidas para estas relaciones funcionales son las siguientes,

$$(6.32) \quad \begin{aligned} \zeta^\dagger &= a_0 + a_1 C_{1/2} \\ \Delta_\zeta &= A_0 \exp(A_1 C_{1/2}), \end{aligned}$$

con los valores numéricos para  $a_0, a_1, A_0$  y  $A_1$  descritos anteriormente. Estos dos elementos nos permiten analizar a continuación la dependencia funcional del cambio  $\Delta_N^D F$  en términos de la concentración  $C_{1/2}$ . No está demás insistir en la relevancia de este punto, dado que los experimentos de desnaturalización de proteínas sugieren que la cantidad  $\Delta_N^D F$  depende efectivamente, de la concentración  $C_{1/2}$ .

Al sustituir las relaciones (6.32) en la ecuación para la energía libre  $F$  (Ecuación 6.24), obtenemos la siguiente expresión

$$(6.33) \quad \begin{aligned} F &= -RT \left( \frac{K}{2} \log \nu - \frac{\beta KU}{2} + \left(K + \frac{1}{2}\right) \log K - \log \sqrt{2\pi} + \log A_0 + A_1 C_{1/2} \right) + \\ &\quad - RT \beta \tilde{E}^* \exp \left( \frac{-n^2 K^2}{4} [1 + \tanh(a_0 - a'_1 C_{1/2})]^2 \right) + \\ &\quad - RT \tanh(a_0 - a'_1 C_{1/2}) \left\{ \frac{K}{2} \log \nu - \frac{\beta KU}{2} + \frac{K}{2} \log \left[ \frac{1 + \tanh(a_0 - a'_1 C_{1/2})}{1 - \tanh(a_0 - a'_1 C_{1/2})} \right] \right\} + \\ &\quad + \frac{RT}{2} \left\{ \log \left[ \frac{K^2}{4} (1 - \tanh^2(a_0 - a'_1 C_{1/2})) \right] + K \log \left[ \frac{K^2}{4} (1 - \tanh^2(a_0 - a'_1 C_{1/2})) \right] \right\}. \end{aligned}$$

en donde  $a'_1 = |a_1 - 1|$ .

Para el intervalo de concentraciones de urea o GdnHCl necesarios para desnaturalizar a una proteína en condiciones estándar de presión y temperatura, [0 – 8] M [Ghelis and Yon, 1982], la cantidad  $|a_0 - a'_1 C_{1/2}| < \pi/2$ , por lo que podemos utilizar esta propiedad para hacer un desarrollo en serie de Taylor de la ecuación (6.33) que haga mucho más explícito el papel de la variable  $C_{1/2}$  en la energía  $\Delta_N^D F$ . A primer orden de aproximación en la variable  $|a_0 - a'_1 C_{1/2}|$  se obtiene que,

$$(6.34) \quad \begin{aligned} \Delta_N^D F &= -RTA(K, U, \nu) - RTA_1 C_{1/2} + \\ &\quad - RT(a_0 - a'_1 C_{1/2}) \left[ \frac{K}{2} \log \nu - \frac{\beta KU}{2} + K(a_0 - a'_1 C_{1/2}) \right] + \\ &\quad + \frac{RT(K+1)}{2} \log \left[ \frac{K^2}{4} (1 - (a_0 - a'_1 C_{1/2})^2) \right], \end{aligned}$$

en donde el coeficiente  $A(K, U, \nu)$  está dado por la expresión,

$$(6.35) \quad A(K, U, \nu) = \left( \frac{K}{2} \log \nu - \frac{\beta KU}{2} + \left(K + \frac{1}{2}\right) \log K - \log \sqrt{2\pi} + \log A_0 \right).$$

La expresión (6.34) implica que a primer orden en la aproximación hecha, el cambio  $\Delta_N^D F$  tiene tres contribuciones en función de  $C_{1/2}$ , un término lineal, un término cuadrático y un término logarítmico, el cual contiene como argumento al cuadrado de  $C_{1/2}$ . La dependencia funcional del cambio energético en una desnaturalización inducida químicamente es en la actualidad un tema de discusión abierto [Asgeirsson *et al*, 2006], [Sharma *et al*, 2006]. Para proteínas como la lizosima en su variante natural, la ribonucleasa-A y ciertos tipos de citocromo-c, los resultados experimentales para  $\Delta_N^D F$  muestran una dependencia lineal en función de  $C_{1/2}$  [Gupta and

Ahmad, 1999], en tanto que para otras proteínas, como la *hisactofilina*, los resultados experimentales están en mejor correspondencia con un perfil cuadrático [Wong, 2004]. El resultado que hemos derivado del MEZ (Ecuación 6.34) contiene ambos comportamientos. El coeficiente  $m_1$  del término lineal depende en proporción directa del tamaño del espacio conformacional de la macromolécula a través de la variable  $\nu$ , de la separación energética  $U$  entre los estados no nativos y de la temperatura  $T$  del sistema, de acuerdo a,

$$(6.36) \quad m_1 = \frac{RTK}{2} \log \nu - \frac{RKU}{2k_B} - RTA_1.$$

Tal como se puede observar de la Ecuación (6.36), tanto los grados de libertad de la macromolécula como el tamaño de la separación energética entre los estados no nativos juegan un papel sustancial en el valor de  $m_1$ . De forma tal, que siendo consistente con los resultados anteriores podríamos decir que, para aquellas proteínas con espacios conformacionales de mayor dimensión, y por tanto de mayor estabilidad, el coeficiente  $m_1$  podría representar la contribución dominante para el cambio de energía libre y que por tanto en este caso,  $\Delta_N^D F$  se escribiría como

$$(6.37) \quad \Delta_N^D F \sim \Delta_N^D F^0 - m_1 C_{1/2},$$

resultado observado experimentalmente para las proteínas mencionadas anteriormente. Para el caso del coeficiente cuadrático  $m_2$ , éste depende sólo del tamaño de la macromolécula y de la temperatura, de acuerdo a

$$(6.38) \quad m_2 = RTK.$$

Dado que la diferencia sustancial entre los coeficientes  $m_1$  y  $m_2$  radica en la presencia de los grados de libertad en el primero, podríamos sugerir entonces que el comportamiento lineal observado en algunos casos de desnaturalización tiene su origen en la multiplicidad de estados de la estructura desplegada. Modelos teóricos que calculan el coeficiente  $m_1$  partiendo de modelos microscópicos de la proteína, asocian al coeficiente  $m_1$  con el número de grupos residuales expuestos al solvente durante el mecanismo de desnaturalización. La idea que presentamos aquí puede resultar en una descripción complementaria a la anterior.

Los resultados anteriores constituyen el núcleo central del primer eje de investigación de este trabajo, el cual se refiere a la estructura del paisaje energético y la estabilidad del estado nativo de una proteína en presencia de agentes químicos [Olivares-Quiroz L., and Garcia-Colin L.S., 2007]. Las perspectivas del MEZ son diversas y serán discutidas en el Capítulo V. A continuación procederemos con el programa planteado al inicio enfocando la atención a la distribución de tiempos  $\tau_r$  en los cuales ocurre la desnaturalización de una proteína en función de la temperatura  $T$ . Para abordar esta problemática, estableceremos en primera instancia una conexión entre la transición vítrea en líquidos orgánicos y la desnaturalización térmica de proteínas. Posterior a ello, deduciremos una expresión analítica para la entropía configuracional de un polímero que puede adquirir una estructura helicoidal, la cual nos servirá como punto de partida para establecer la forma funcional de los tiempos de relajación estructural de una macromolécula en función de la temperatura  $T$ .

## CAPÍTULO 4

### Cinética del Plegamiento y la Transición Vítreas

En el Capítulo III propusimos un formalismo fenomenológico para calcular el cambio de energía libre  $\Delta_N^D F$  entre los estados nativo y desnaturalizado de una proteína, los cuales de acuerdo a la Hipótesis Termodinámica de Anfinsen (Capítulo I), son estados de equilibrio termodinámico [Anfinsen et al. , 1961]. En este capítulo centramos la atención en el segundo eje de investigación de este proyecto doctoral, el cual se refiere a aspectos cinéticos del mecanismo  $N \rightarrow D$  y su conexión con el fenómeno de la transición vítrea (*glass transition*) en líquidos sobreenfriados, esto es, con la generación de estructuras cuasi-cristalinas fuera de equilibrio por debajo de la temperatura normal de cristalización del líquido [Angell, C., 1995], [Angell, C.A. , 1995b]. Propondremos que dadas las similitudes termodinámicas observadas entre la transición vítrea en líquidos y la desnaturalización de proteínas, es posible utilizar el formalismo propuesto por Adam y Gibbs [Adam and Gibbs, 1965] para calcular el tiempo  $\tau$  de relajación estructural entre el estado  $N$  y  $D$  en una proteína.

El programa del capítulo es el siguiente. En primera instancia mostraremos algunas de las líneas generales sobre las cuales se desarrolla la investigación en la cinética del plegamiento. Posterior a ello, presentaremos la evidencia experimental que sustenta una analogía entre la transición vítrea y la desnaturalización de proteínas, al menos en un nivel termodinámico. En la sección subsecuente, calcularemos las funciones termodinámicas  $\Delta_N^D H$ ,  $\Delta_N^D S$  y  $\Delta_N^D G$  para la desnaturalización térmica de la mioglobina, proteína que tomaremos como prototipo de análisis en este Capítulo. Una vez calculadas, sugeriremos un formalismo basado en geometría diferencial de curvas en el espacio, para calcular la entropía configuracional  $\Delta s$  de un polímero en una conformación helicoidal. Así, el cálculo del cambio de entropía global  $\Delta_N^D S$  del sistema y del cambio de entropía configuracional de una subestructura del mismo  $\Delta s$  nos permitirán construir los elementos necesarios para la aplicación del formalismo de Adam y Gibbs a la desnaturalización de proteínas. En la última parte mostraremos la forma funcional para  $\tau_r$  y discutiremos algunas implicaciones de estos resultados.

#### 1. La Transición Vítreas y la Desnaturalización de Proteínas.

Una de las clasificaciones más simples y estándar en cursos elementales de Física organiza a la materia de acuerdo a su habilidad para fluir. Se dice que los líquidos, gases y plasmas son estados de la materia en los cuales su organización molecular es tal que les permite fluir, esto es, mostrar desplazamientos relativos entre sí. Por otro lado se define a los sólidos, ya sea cristalinos o amorfos como sistemas para los cuales su grado de fluidez se ha reducido casi a cero, para condiciones estándares de presión y temperatura. La definición anterior, que es esencialmente correcta y funciona bastante bien para la mayoría de los líquidos y sólidos estándar, utiliza en el fondo el concepto de difusión molecular. En términos de esta propiedad, se diría entonces que los gases exhiben un grado mayor de difusión molecular que los líquidos, los cuales a

su vez un grado mayor que lo sólidos, y así sucesivamente. El uso del término difusión molecular nos parece más apropiado para definir a los estados de la materia, dado que es en este sentido en el cual sucede la definición estándar del estado vítreo.

Para estar en posición de dar una definición más precisa del estado vítreo, discutiremos brevemente la fenomenología observada durante la transición del estado líquido al estado vitrificado. Una vez que se inicia el proceso de enfriamiento gradual del líquido formador de vidrio, se observa que conforme la temperatura  $T$  se aproxima al valor crítico  $T_g$ , en donde normalmente se observaría la transición al estado cristalino, el sistema toma un ruta diferente y accesa a un estado distinto, en donde el tiempo necesario para que el sistema alcance un estado de equilibrio termodinámico se incrementa en forma significativa [Angell, C., 1995]. Asociado a este proceso, mediciones calorimétricas del calor específico  $C_p$  tanto en la fase líquida como en este nuevo estado muestran un decremento significativo en el cambio del calor específico entre estos dos estados. Esto es,  $\Delta_l C_p \equiv C_p^g - C_p^l < 0$ , en donde  $C_p^l$  se refiere al calor específico en el estado líquido y  $C_p^g$  se refiere al calor específico en el estado vítreo. Diversos estudios realizados con dinámica molecular confrontados con esquemas teóricos deducidos de la teoría de acoplamiento de modos (*Mode Coupling Theory*) han permitido elucidar un panorama general de lo que ocurre a nivel molecular durante la transición vítrea [Bengtzelius, 1986],[Gotze., 1989].

De acuerdo a esto, a temperaturas estándares para el estado líquido, el sistema puede considerarse como un conjunto de partículas en constante movimiento caótico debido a las fluctuaciones térmicas del medio, las cuales colisionan entre sí de acuerdo a lo predicho por la Mecánica Estadística de Equilibrio. A esta temperatura y densidad, las colisiones son lo suficientemente rápidas y ocurren en distancias intermoleculares suficientemente amplias comparadas con las distancias intermoleculares que en promedio, pueden difundirse dentro del líquido permitiendo alcanzar al sistema un estado de equilibrio termodinámico a nivel macroscópico. Conforme la temperatura  $T$  y el volumen comienzan a disminuir, la densidad del sistema se incrementa de forma que las partículas ahora tienen un espacio menor para la difusión y las colisiones se hacen más frecuentes dentro del mismo volumen. Esto es, el grado de difusión molecular del sistema decrece significativamente en comparación con el estado líquido. En términos del llamado tiempo de relajación estructural  $\tau_r$ , que no es más que el tiempo requerido para que las moléculas puedan recorrer una distancia interatómica, el estado vítreo corresponde entonces a un estado en donde el tiempo  $\tau_r$  requerido para que las partículas se difundan, se ha incrementado enormemente en comparación con el del estado líquido, impidiendo así que el sistema alcance un estado de equilibrio termodinámico. [Avramov, 2005].

La idea anterior permite entonces definir al estado vítreo en términos operacionales ya sea mediante la medición del cambio  $\Delta C_p$  o bien mediante experimentos de difracción de neutrones, los cuales permiten la medición indirecta del tiempo de relajación estructural  $\tau_r$ . La Figura (1) muestra el comportamiento del calor específico  $C_p$  en función de la temperatura  $T$  en la vecindad de la transición vítrea para diversos sistemas líquidos como el Tolueno (A), para aleaciones metálicas como Au-Si (B), sistemas unidos covalentemente como la mezcla  $As_2Se_3$  (C) y compuestos estructurados en redes abiertas como el caso de  $GeO_2$  (D). En todos los casos, se observa un cambio abrupto del  $C_p$  en la vecindad de la temperatura  $T_g$ . El análisis de estas

curvas provee de una definición operacional para la temperatura  $T_g$  a la cual ocurre la transición vítrea y que corresponde a la región en donde finaliza el decremento de  $\Delta C_p$ .

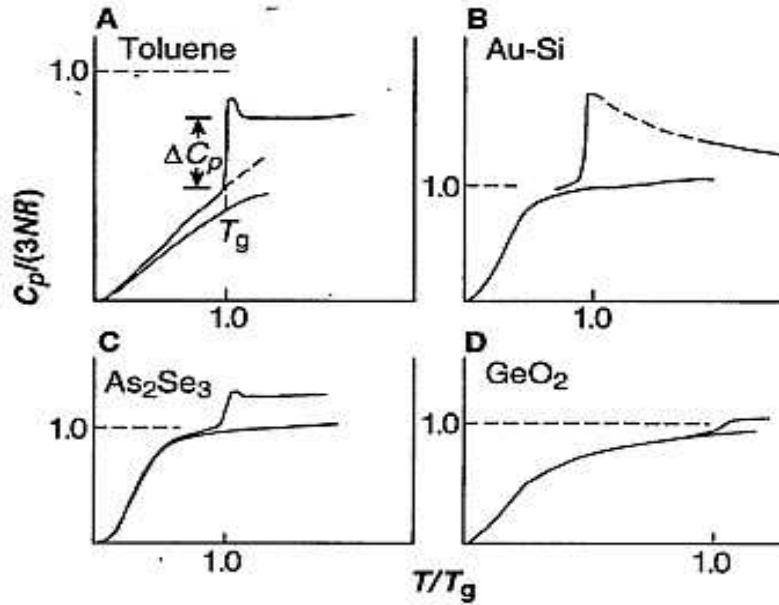


FIGURA 1. Cambio del calor específico  $C_p$  en función de la temperatura  $T$  para el tolueno (A), Au-Si (B),  $As_2Se_3$  (C) y  $GeO_2$  (D). Aunque la intensidad del cambio  $\Delta C_p$  es distinta para cada tipo de material, se observa como rasgo general que conforme la temperatura del sistema ha disminuido por debajo de la temperatura de transición vítrea  $T_g$ , la forma funcional del  $C_p$  se aproxima al comportamiento lineal típico del estado cristalino. Tomado de [Angell, C., 1995].

El decremento abrupto del calor específico  $\Delta C_p$  en la vecindad de la temperatura  $T_g$  es reconocido como uno de los signos más característicos de que el sistema ha entrado en una conformación estructural distinta a la de un líquido. Sumado a lo anterior, la viscosidad  $\eta$  del sistema sufre también modificaciones sustanciales, dado que ésta es una propiedad macroscópica del sistema asociada con una propiedad microscópica, que es el tiempo de relajación estructural  $\tau_r$ , de acuerdo a la ecuación propuesta por Maxwell

$$(1.1) \quad \eta = G_\infty \tau_r,$$

en donde  $G_\infty$  es el tensor de esfuerzos y  $\eta$  la viscosidad del sistema. En vista entonces de la conexión entre el tiempo de relajación estructural  $\tau_r$  y la viscosidad, es posible añadir un elemento más a la caracterización de la transición vítrea mediante mediciones experimentales de la viscosidad. Si se elige una viscosidad de referencia  $\eta_g = 10^{13}$  poise para la temperatura  $T_g$ , es posible establecer una escala universal para el comportamiento de  $\eta$  en función de  $T$  y agrupar entonces las gráficas de viscosidad versus temperatura para diversos sistemas. La Figura (2) muestra el comportamiento de la variable  $\log \eta$  en función de la variable adimensional  $T_g/T$  para diversos sistemas.

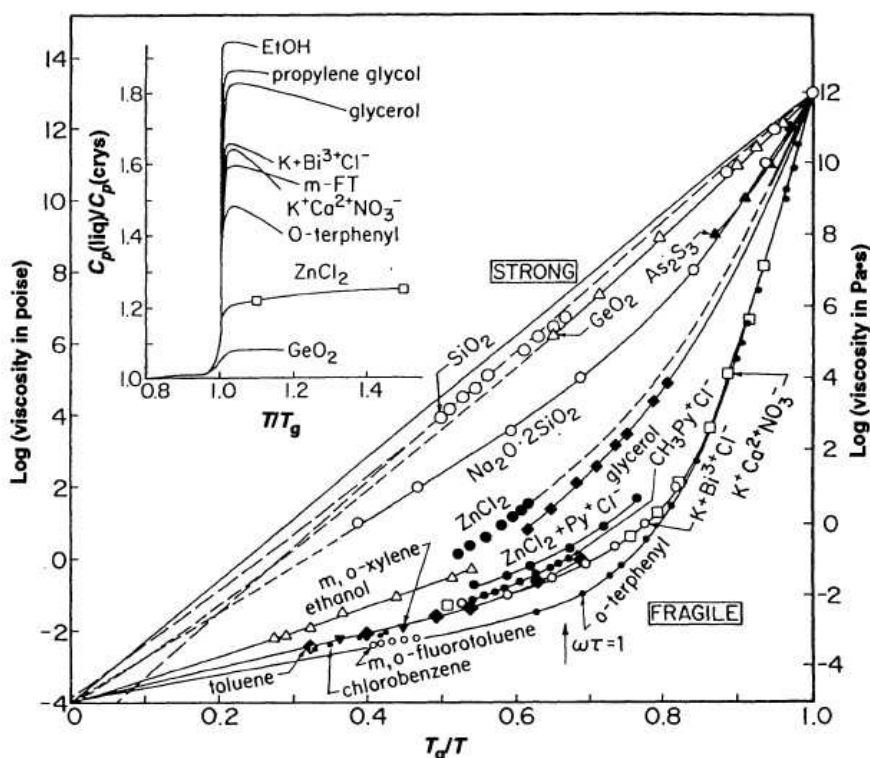


FIGURA 2. Gráfico de la viscosidad  $\eta$  en función de la temperatura  $T$  para diversos materiales de acuerdo a [Angell, C., 1995]. Los puntos corresponden a los datos experimentales en tanto que las curvas sólidas son ajustes de los datos para la ecuación 1.2. El comportamiento observado para  $\eta(T)$  permite clasificar a los sistemas formadores de vidrios en frágiles, para aquellos que no siguen una dinámica de tipo Arrhenius; y fuertes, para aquellos que sí la siguen.

La Figura(2), que apareció por primera vez a mediados de los noventa muestra que existen en esencia dos tipos de comportamientos diferentes para la viscosidad de líquidos formadores de vidrios. Por un lado, para sistemas como  $SiO_2$  y  $GeO_3$ , se observa que  $\log(\eta)$  tiene un comportamiento lineal en función de la variable adimensional  $T_g/T$ , en tanto que otro tipo de sistemas, como el glicerol o el tolueno exhiben una viscosidad con una dinámica diferente. Para ajustar los datos experimentales observados en la gráfica (2), Fuchel-Vogel-Tamman propusieron una ecuación fenomenológica de tipo Arrhenius para la viscosidad en función de la temperatura. Esta ecuación contiene sólo un parámetro ajustable, el cual indica el grado de desviación del sistema de un comportamiento de tipo Arrhenius ideal. Esta relación está dada por,

$$(1.2) \quad \eta = \eta_0 \exp\left(\frac{DT_0}{T - T_0}\right),$$

en donde  $T_0$  es una temperatura de referencia,  $\eta_0$  es la viscosidad del sistema a esta temperatura y  $D$  es el parámetro ajustable que mencionamos anteriormente. Uno de los hechos más interesantes en referencia a la ecuación (1.2), es que reproduce los

perfiles de la viscosidad mostrados en la gráfica mediante la elección adecuada del parámetro  $D$ , el cual generalmente oscila entre  $5 < D < 100$  [García-Colin, L.S., 1999b]. El hecho de que existan en esencia dos tipos de comportamiento para la viscosidad  $\eta$  en función de  $T$ , indica también la presencia de dos dinámicas distintas para el comportamiento de los tiempos de relajación estructural del sistema, en vista de la ecuación (1.1). Esta característica ha sugerido una clasificación para los tipos de estados vítreos en función de su viscosidad. Dentro de este esquema, se conoce como líquidos fuertes a aquellos, como el  $SiO_2$ , cuya viscosidad sigue una dinámica de tipo Arrhenius y vidrios frágiles a aquellos cuyos líquidos que los forman, muestran desviaciones significativas de este comportamiento. En términos de la ecuación (1.2) y del parámetro  $D$ , los vidrios fuertes corresponden a valores mayores del parámetro  $D$ , en tanto que los vidrios frágiles tienden a ser representados por valores menores de  $D$ . Si bien la ecuación (1.2) ha mostrado ser un excelente ajuste para los datos experimentales provenientes de líquidos inorgánicos, no reproduce correctamente la dinámica del estado vítreo para líquidos poliméricos, dado que en este caso la viscosidad también depende del peso molecular del líquido. Para resolver esta cuestión, William, Landel y Ferry [Angell, C., 1995] introdujeron un enfoque distinto, basado en una propuesta fenomenológica directa para el tiempo de relajación estructural  $\tau_r$ , la cual funciona bastante bien para vidrios poliméricos si se eligen las constantes adecuadas. De momento no abundaremos más en este tópico dado que lo retomaremos con detalle en párrafos subsecuentes cuando discutamos el formalismo de Adam y Gibbs, el cual será nuestro eje base para analizar la desnaturalización de proteínas.

Hemos discutido aquí que la transición vítrea es un proceso que va acompañado de una disminución del calor específico  $\Delta C_p$  cuando el sistema cruza la temperatura de transición  $T_g$ . Esto es, conforme el sistema pasa del estado líquido estándar al estado vitrificado, el  $\Delta C_p < 0$ , cuya intensidad depende del comportamiento en la viscosidad  $\eta$  en función de la temperatura, esto es, del grado de 'fragilidad' del líquido en cuestión. Para reconocer la importancia de esta característica termodinámica en la desnaturalización de proteínas, recordemos que tal como lo discutimos en el Capítulo III, los experimentos calorimétricos que monitorean el desdoblamiento del estado nativo  $N$  y su conversión en la estructura desplegada  $D$ , indican que este proceso se realiza con una absorción de calor significativa, lo cual se refleja en un cambio positivo para el  $\Delta C_p$  de este proceso. Tal como lo muestra la Figura (3) y que reproducimos aquí para facilidad de comparación, el incremento del pH tiende a incrementar la estabilidad de la proteína y por tanto a intensificar el pico de absorción de calor. Las líneas punteadas muestran el comportamiento de los calores específicos para el estado nativo  $C_p^N$  y para el estado desnaturalizado  $C_p^D$  en función de  $T$ , los cuales muestran un comportamiento creciente con una ligera curvatura. Esta característica induce a su vez la presencia de una curvatura positiva para la diferencia  $\Delta_N^D C_p = C_p^D - C_p^N$ , de forma tal que en los experimentos de calorimetría de barrido se observa que la desnaturalización de proteínas va acompañada por un incremento en el calor específico. Esto, es  $\Delta C_p > 0$ .

Para explicar el incremento en el calor específico, Kauzmann propuso el ahora llamado *efecto hidrofóbico* que tal como revisamos en el Capítulo III, propone que dado que el estado nativo está conformado por un núcleo de grupos residuales hidrofóbicos y una superficie hidrofílica, la transición entre este estado y la estructura desplegada



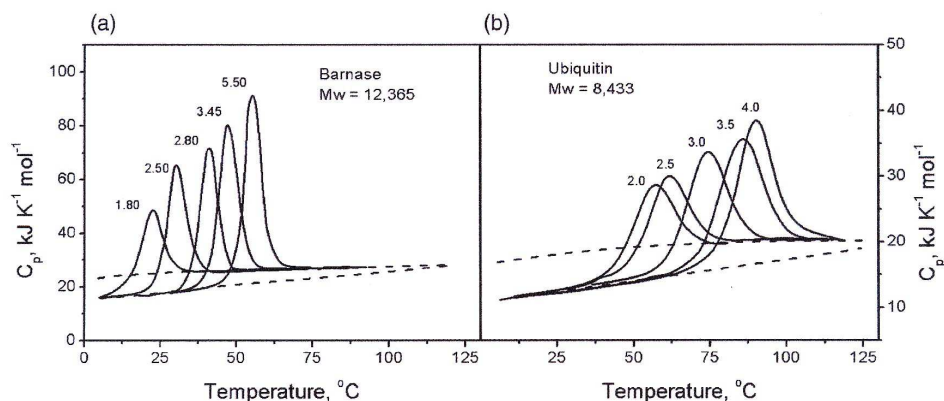


FIGURA 3. El calor específico molar  $C_p$  para dos proteínas a) Barnasa con una masa molecular de  $M_w = 12.4$  kDa y para b) Ubiquitina de masa molecular  $M_w = 8.4$  kDa en presencia de diferentes valores de pH. De acuerdo a Privalov P.L. y Dragan A.I. [Privalov and Dragan, 2007]

implica entonces la exposición de los grupos hidrofóbicos a las moléculas del solvente. Esta situación implica entonces dos efectos en el sistema proteína+solvente,

- Por un lado la desorganización de la proteína en sí genera un incremento en la entropía  $\Delta S > 0$ , dado que el estado desnaturalizado tiene un número mayor de microestados accesibles. Este efecto produce una contribución positiva en el  $\Delta C_p$  de la transición.
- Por otro lado, la exposición de los grupos no polares de la proteína al solvente implica también una reorganización estructural de las moléculas de agua. Como se sabe, en el estado líquido las moléculas de agua se hallan unidas en su mayoría por puentes de hidrógeno, con una energía promedio 5 kcal/mol por puente de hidrógeno. Al ser liberados los grupos no polares hacia la superficie de la proteína, las moléculas de agua circundantes deben modificar su orientación a fin de evitar el contacto directo con ellos. Este proceso implica un decremento en la entropía del sistema  $\Delta S < 0$ , dado que existirá entonces un número menor de conformaciones accesibles para las moléculas de agua.

Estos dos efectos, por un lado un incremento en la entropía proveniente de la ruptura de la estructura nativa en sí y por otro lado un decremento en la entropía como resultado de la organización molecular del solvente, generan una contribución total  $\Delta S < 0$ , dado que el segundo efecto es de mayor intensidad. Si observamos entonces que el cambio de la energía libre  $\Delta F$  está dado por

$$(1.3) \quad \Delta G = \Delta H - T\Delta S,$$

veremos que la contribución entrópica negativa  $-T\Delta S$  genera un término positivo para  $\Delta G$ , haciendo entonces que en forma global, el proceso de exposición de los grupos polares de la proteína al solvente sea desfavorable en términos termodinámicos [Finkelstein and Ptysin, 2002]. Por otro lado, dado que el cambio  $\Delta C_p$  está asociado

Table 1: Data for 13 Unique Single Domains in Our Study<sup>a</sup>

protein	$N_{res}$	PDB entry	method	resolution (Å)	$T_G$ (°C)	$\Delta H_G$ (kcal/mol)	$\Delta C_p$ (kcal mol <sup>-1</sup> K <sup>-1</sup> )
Sso7d	62	1SSO	NMR		97.8	63.4	0.62 ± 0.04
TmCsp	66				82	62.6	1.1 ± 0.1
BsHpr	87	2HID	NMR		73.4 ± 0.2	58.1 ± 1.7	1.17 ± 0.05
$\lambda$ repressor <sub>6-85</sub>	80	1LMB	X-ray	1.7	57.2 ± 0.1	68.0 ± 1.0	1.44 ± 0.03
Barstar	90	1A19	X-ray	2.76	72.7 ± 0.3	72.4 ± 3.6	1.27 ± 0.24
Snase	149	1EY0	X-ray	1.6	52.8	96 ± 2	2.2
FeCyt $b_{562}$	106	1QPU	NMR		67.2 ± 0.5	94 ± 5	2.4 ± 0.4
GDH domain II	150	1B26	X-ray	3.0	69.6	70.2 ± 4.0	1.4 ± 0.3
C <sub>L</sub> of Ig $\lambda$	104	1A8J	X-ray	2.7	61 ± 0.5	66.5 ± 0.9	1.8
K-STI	181	1AVU	X-ray	2.3	59.0	102.5 ± 1.4	2.6 ± 0.1
apoflavodoxin	168	1FTG	X-ray	2.0	57.3 ± 0.1	63.1 ± 0.7	1.34 ± 0.02
thioredoxin	108	2TRX	X-ray	1.68	87.0	106.9 ± 1.1	1.66 ± 0.05
ADA2h	80	1AYE	X-ray	1.8	77.0	47.6	0.86 ± 0.33

FIGURA 4. Datos termodinámicos de la desnaturalización  $N \rightleftharpoons D$  para trece proteínas conformadas por un sólo dominio analizadas por Kumar *et al.* [Kumar, Tsai and Nussinov, 2003]. Estas proteínas exhiben su máxima estabilidad a temperatura ambiente y tiene una temperatura de desnaturalización  $T_G$ . Se observa que el cambio  $\Delta C_p$  es positivo en todos los casos y oscila entre 0.5 – 2.5 kcal/Kelvin mol

con el cambio de energía libre  $\Delta F$  a través de la relación

$$(1.4) \quad \Delta C_p = -T \frac{\partial^2 \Delta F}{\partial T^2}$$

es posible ver en forma directa que una curvatura negativa en el cambio de energía libre  $\Delta F$  y por tanto la existencia de un máximo en la estabilidad del sistema, inducirán un  $\Delta C_p$  positivo. En la actualidad, se considera que el mecanismo anterior representa la contribución más importante para el cambio positivo en el calor específico de las proteínas en un proceso de desnaturalización estándar. Uno de los cálculos más conservadores en este sentido, es el propuesto por Schellman *et al.*, el cual sugiere que el 75% del  $\Delta C_p$  proviene justamente del efecto hidrofóbico [Schellmann, 1997]. Pese a que otros autores proponen que tanto el efecto hidrofóbico como las interacciones electrostáticas contribuyen de igual forma a la estabilidad del estado nativo [Pace *et al.*, 1996], las simulaciones computacionales y los análisis termodinámicos de los experimentos de desnaturalización indican, que al menos para proteínas globulares de dos estados, el efecto hidrofóbico representa la contribución dominante [?]. La Figura (4) muestra los datos experimentales reportados por Kumar *et al.* [Kumar, Tsai and Nussinov, 2003], para trece proteínas conformadas por un sólo dominio y que exhiben una transición  $N \rightleftharpoons D$  altamente cooperativa con un grado de reversibilidad de aproximadamente el 95%. Tal como se muestra, en todos los casos la transición  $N \rightarrow D$  está acompañada por un incremento sustancial del calor específico  $C_p$  el cual oscila entre 0.5 – 2.5 kcal/Kelvin mol.

Para el caso estándar de la desnaturalización de proteínas, en donde los grupos no polares se hallan en el núcleo hidrofóbico, el  $\Delta C_p > 0$ . Sin embargo, estudios de cadenas polipeptídicas conformadas por grupos residuales exclusivamente hidrofílicos

muestran que en este caso el  $\Delta C_p < 0$ , como sucede en la transición vítrea. De forma entonces que el signo de  $\Delta C_p$  permite distinguir la solvatación de grupos no polares (+), de la solvatación de grupos hidrofílicos (-), en contraposición con el signo del cambio de entalpía  $\Delta H$  y de entropía  $\Delta S$  que son ambos negativos tanto para los grupos polares como para los hidrofóbicos [Murphy *et al*, 1990], [Baldwin, 1986]. Asimismo, la conexión entre el cambio de energía libre  $\Delta F$  y el calor específico  $\Delta C_p$  (Ecuación 1.4) implica que para un  $\Delta C_p > 0$ , la curvatura de  $\Delta F < 0$  en función de  $T$  es negativa y consecuentemente que  $\Delta F$  versus  $T$  tiene un máximo de estabilidad para una temperatura dada. Esta característica coincide con lo observado experimentalmente y también implica que la estabilidad de las proteínas disminuye conforme la temperatura desciende dando origen al fenómeno conocido como desnaturalización en frío (*cold denaturation*) [Robertson and Murphy, 1997], [Prabhu and Sharp, 2005].

En este nivel de la discusión hemos presentado dos características comunes entre la transición vítrea y la desnaturalización de proteínas. Por un lado el comportamiento anómalo del calor específico durante la transición,  $\Delta C_p > 0$  para la desnaturalización de la estructura nativa y  $\Delta C_p < 0$  para el sobreenfriamiento de un líquido formador de vidrios y por otro, el decremento en ambos casos de la entropía configuracional del sistema. En el caso de la transición vítrea, el cambio  $\Delta S < 0$  es una consecuencia directa del incremento en el tiempo de relajación estructural  $\tau_r$  y del consecuente congelamiento de grados de libertad colectivos del sistema. En el caso de la desnaturalización de proteínas, el decremento  $\Delta S < 0$  observado proviene de la interacción entre la cadena polipeptídica y las moléculas del solvente. Como ya establecimos, la exposición de los grupos no polares del núcleo hidrofóbico induce un rearrreglo sustancial en la orientación de los puentes de hidrógeno en el solvente, restringiendo así el número de posibles conformaciones accesibles a las moléculas del solvente. Este efecto, de mayor intensidad que el cambio entrópico positivo debido a la desorganización de la cadena polipeptídica, representa la mayor contribución energética  $-T\Delta S$  para el cambio de energía  $\Delta F$  a temperatura ambiente.

La similitud entre los procesos de desnaturalización de proteínas y la transición vítrea va más allá de las características termodinámicas discutidas aquí. Se extienden también a procesos de tipo cinético, como es el caso de la enorme similitud existente entre los espectros Raman y de difracción de neutrones de las proteínas hidratadas y de ciertos tipos de estructuras vítreas lo cual indica la existencia en ambos casos de un pico bosónico en el factor de estructura  $S(q, \omega)$  [Orecchini, 2001], [Leyser *et al*, 1999]. Por otro lado, la presencia de distintas escalas de tiempo para los procesos de relajación hacia el equilibrio tanto en proteínas como en vidrios [Piazza *et al*, 2005], ha sugerido que la analogía entre ambos sistemas va más allá de simples coincidencias. En este sentido, la mayoría de los autores sugieren que la presencia de múltiples interacciones entre los componentes de cada uno de estos sistemas, aunados a la presencia del fenómeno de la frustración y de un paisaje energético con una geometría de múltiples mínimos locales, puede conducir a la explicación más plausible para las coincidencias observadas.

## 2. Teoría Básica de la Transición Vítrea.

De acuerdo entonces a lo discutido anteriormente, es posible en principio construir una analogía entre la desnaturalización de proteínas y la transición vítrea a distintos

niveles de descripción. Desde una teoría microscópica basada en la construcción de un Paisaje Energético hasta un enfoque fenomenológico que tome en cuenta las similitudes termodinámicas observadas en ambos casos. En la presente sección partiremos de esta última opción para establecer una conexión entre la transición vítrea y la desnaturalización de proteínas. Esto es, dado que ambos fenómenos están caracterizados por un cambio de entropía configuracional negativo  $\Delta S < 0$  y un incremento sustancial del calor específico  $\Delta C_p$  (positivo para la desnaturalización y negativo para la transición vítrea), consideraremos entonces que ambos procesos pueden describirse, al menos en un nivel termodinámico, mediante formalismos similares. Si aceptamos como verdadera esta premisa, entonces podremos utilizar algunas de las herramientas formales desarrolladas en el campo de la materia condensada para calcular cantidades de relevancia en la desnaturalización de proteínas. En este caso, estaremos interesados en particular en una variable dinámica del proceso, el así llamado tiempo de relajación estructural  $\tau_r$ . En el caso de la transición vítrea, la variable  $\tau_r$  es un indicador macroscópico de los procesos microscópicos de congelamiento de los grados de libertad del sistema, dado que depende explícitamente del cambio entrópico entre el estado líquido y el estado vítreo del sistema. En el caso de la desnaturalización de proteínas, la interpretación del tiempo  $\tau_r$  se refiere al tiempo requerido para cambiar la conformación estructural de la macromolécula y, dado que justo esta característica es la que define la transición entre el estado nativo  $N$  y el estado desplegado  $D$ , podemos ver entonces que en forma natural el equivalente del tiempo de relajación estructural en vidrios corresponde al tiempo de plegamiento y/o desnaturalización de una proteína cuando el proceso es reversible.

El programa que desarrollaremos en esta sección entonces puede desglosarse de acuerdo a lo siguiente. En una primera parte, describiremos el formalismo propuesto por Adam y Gibbs para calcular el tiempo de relajación estructural en vidrios de naturaleza polimérica, cuya viscosidad  $\eta$  como vimos anteriormente, no sigue una curva de Arrhenius en forma estricta. La propuesta de Adam y Gibbs toma en consideración esta propiedad y, basándose en argumentos de mecánica estadística, calcula el tiempo  $\tau_r$  en función de la temperatura, del cambio de entropía configuracional  $\Delta S$  del sistema y de la entropía conformacional de subregiones del sistema capaces de reorganizarse estructuralmente en forma independiente del resto del mismo. Para estar en condiciones de aplicar estas ideas al caso de la desnaturalización de proteínas, discutiremos posteriormente la manera de determinar tanto el cambio de entropía configuracional  $\Delta S_c$  entre los estados  $N$  y  $D$  como el cambio de entropía conformacional  $\Delta s$  de una subregión de la estructura nativa de una proteína. Finalmente aplicaremos las ideas desarrolladas para el caso de la desnaturalización inducida por temperatura de la mioglobina, una proteína esencial para el transporte de oxígeno a los tejidos del organismo y cuya estructura nativa está conformada esencialmente por hélices  $\alpha$ .

**2.1. Formalismo de Adam y Gibbs para la transición vítrea en líquidos orgánicos.** Tal como mencionamos brevemente en la sección anterior, para el caso de líquidos poliméricos formadores de vidrios la viscosidad  $\eta$  y el tiempo de relajación  $\tau_r$  no siguen un comportamiento de tipo Arrhenius en función de la temperatura  $T$  como el descrito por la ecuación (1.2). Para tratar de reproducir el comportamiento observado experimentalmente para la variable  $\tau_r$ , Williams, Landel y Ferry [William

*et al*, 1955] propusieron empíricamente una ecuación de la forma

$$(2.1) \quad -\log a_T = \frac{C_1(T - T_s)}{C_2 + (T - T_s)},$$

en donde  $a = \tau_r(T)/\tau_r(T_s)$  es el cociente de los tiempos  $\tau_r$  en la temperatura  $T$  y una temperatura de referencia  $T_s$  y  $C_1$  y  $C_2$  son dos constantes para ajustar a los datos experimentales. La ecuación (2.1) escrita en términos del tiempo  $\tau_r(T)$  tiene entonces la forma siguiente

$$(2.2) \quad \tau_r(T) = \tau_r(T_s) \exp\left(-\frac{C_1(T - T_s)}{C_2 + (T - T_s)}\right),$$

en donde se observa en forma más clara que el tiempo  $\tau_r$  tiene en este caso un comportamiento de tipo no-Arrhenius.

La relevancia de la ecuación (2.2), conocida como Ecuación WLF, radica en su capacidad para reproducir el comportamiento de la variable  $\tau_r$  cerca de la temperatura de transición vítrea  $T_g$  para un número considerable de líquidos formadores de vidrio eligiendo un único par de constantes  $C_1 = 8.86$  y  $C_2 = 101.6^\circ C$  para todos los sistemas. En cierto sentido, las constantes  $C_1$  y  $C_2$  tienen entonces un carácter universal para el intervalo de temperaturas analizadas [Williams., 1955]. Asimismo, otra de las características interesantes de la ecuación WLF se refiere a la observación hecha por sus mismos autores, en referencia a que la temperatura de referencia  $T_s$  elegida para cada sistema consistente con los valores de las constantes  $C_1$  y  $C_2$  está relacionada en todos los casos con la temperatura  $T_g$  a la cual ocurre la transición vítrea para cada uno de ellos. La relación exacta entre  $T_s$  y  $T_g$  implica también un comportamiento 'universal' para este tipo de sistemas, dado que la diferencia

$$(2.3) \quad T_s - T_g \sim 50^\circ C,$$

es una constante para todos los casos. La ecuación (2.3) implica que la temperatura de referencia  $T_s$  está ubicada  $50^\circ C$  por encima de la temperatura de la transición vítrea  $T_g$  en todos los casos discutidos.

Así planteado, la intención entonces del formalismo propuesto por Adam y Gibbs [Adam and Gibbs, 1965] es proveer de una fundamentación para la ecuación WLF en términos de argumentos mecánico-estadísticos. Para ello, se toma como punto de partida la observación experimental que conforme  $T \rightarrow T_g$ , la entropía configuracional  $S_c \rightarrow 0$  y el tiempo de relajación estructural  $\tau_r$  se incrementa significativamente. En referencia justamente a la entropía configuracional, Kauzmann observó que si se extrapolan los gráficos de la entropía  $S_c$  en función de la temperatura por debajo de la temperatura  $T_g$ , la entropía adquiriría valores por debajo de la entropía del cristal correspondiente, el cual corresponde al estado de mayor ordenamiento molecular del sistema y por ende, el correspondiente a la menor entropía conformacional posible. Este hecho, conocido como la Paradoja de Kauzmann [Kauzmann, W., 1959], ha recibido una atención considerable desde su planteamiento y forma parte de las motivaciones de Adam y Gibbs para construir una fundamentación para la ecuación WLF. A este respecto y a manera de explicación, Gibbs y Di Marzio propusieron que la extrapolación de los gráficos  $\Delta S_c$  versus  $T$  para  $T < T_g$  es esencialmente incorrecto, dado que en este intervalo de temperaturas el sistema exhibe un transición de fase de segundo orden a una temperatura  $T_2$ , en donde la entropía conformacional se hace idénticamente cero y permanece constante conforme  $T \rightarrow 0$ , en concordancia

con el Principio de Nerst. Tal como lo discuten Adam y Gibbs en su artículo, las temperaturas  $T_g$  y  $T_2$  tienen una categoría distinta en cuanto a su significado físico. Mientras que la temperatura  $T_g$  en la cual se manifiesta la transición vítrea depende del proceso de sobreenfriamiento del líquido y por tanto no es única, la temperatura  $T_2$  a la cual ocurre la transición de fase que anula la entropía del sistema está definida sin ambigüedad a través de la relación  $S_c(T_2) = 0$ .

El primer paso en el esquema de Adam y Gibbs supone que el líquido formador de vidrio puede subdividirse en  $n$  regiones espaciales que pueden modificar su conformación estructural en forma independiente y  $N - n$  regiones que no presentan esta propiedad. Estas  $n$  regiones las denomina *regiones cooperativas* (*Cooperative Rearrangements*) y están definidas en esencia como subregiones del sistema que interactúan débilmente con las demás, de forma entonces que pueden reestructurarse en forma independiente del resto del sistema. Si llamamos  $z$  al número de moléculas o segmentos en una cadena polimérica que integran a cada una de las regiones cooperativas del líquido, entonces el sistema está compuesto por  $n$  regiones cooperativas con una distribución de tamaños  $\{z_i\}$ . Asumiendo un estado de equilibrio termodinámico a temperatura  $T$ , la probabilidad  $W$  de que alguna de las  $n$  regiones cooperativas exhiba un cambio estructural está dado por

$$(2.4) \quad W(T) = A \exp\left(-\frac{z\Delta\mu}{k_b T}\right),$$

en donde  $\Delta\mu = (G' - G)/z$  y  $G'$  y  $G$  representa las energías libres de las regiones cooperativas y no cooperativas, respectivamente. En términos físicos, la cantidad  $\Delta\mu$  representa la barrera energética entre las subregiones cooperativas y las que no lo son [Adam and Gibbs, 1965]. Para determinar la probabilidad media de transición  $\bar{W}(T)$  se debe sumar sobre todas los tamaños posibles para las regiones cooperativas. Esto es,

$$(2.5) \quad \bar{W}(T) = \sum_{j=z^*}^{\infty} A \left[ \exp\left(-\frac{z\Delta\mu}{k_b T}\right) \right]^j,$$

cuyo resultado es

$$(2.6) \quad \bar{W}(T) = \frac{A}{1 - \exp\left(-\frac{z^*\Delta\mu}{k_b T}\right)} \exp\left(-\frac{z^*\Delta\mu}{k_b T}\right),$$

en donde  $z^*$  representa el tamaño mínimo necesario que debe tener una región cooperativa para que exhiba una transición conformacional en forma independiente de los demás. El resultado mostrado en la ecuación (2.6) indica entonces que la máxima probabilidad para que suceda un cambio estructural ocurre básicamente para aquellas regiones cooperativas de tamaño crítico  $z^*$ . Asumiendo que el Hamiltoniano del sistema puede separarse en una parte cinética y una parte potencial, el tamaño crítico  $z^*$  puede escribirse en términos de la entropía conformacional  $s^*$  correspondiente a la región cooperativa de menor tamaño que es capaz de exhibir una transición estructural y de la entropía conformacional  $S_c$  del sistema de acuerdo a  $z^* = N_A s^*/S_c$  y entonces la probabilidad media de transición  $\bar{W}$  queda escrita como

$$(2.7) \quad \bar{W}(T) = \bar{A} \exp\left(-\frac{s^*\Delta\mu}{k_b T S_c}\right),$$

en donde hemos definido

$$(2.8) \quad \bar{A} \equiv \frac{A}{1 - \exp\left(-\frac{z\Delta\mu}{k_b T}\right)},$$

el cual es en esencia un factor constante en la vecindad de la temperatura de la transición vítrea.

Finalmente, dado que el tiempo  $\tau_r$  es proporcional al recíproco de la probabilidad media de transición  $\bar{W}$ ,

$$(2.9) \quad \tau_r \sim \frac{1}{\bar{W}(T)},$$

entonces es posible escribir al factor de corrimiento  $a$  de la ecuación de WLF (Ecuación 2.1) de acuerdo a,

$$(2.10) \quad -\log a_T \equiv -\log\left(\frac{\tau_r(T)}{\tau_r(T_s)}\right) = -\log\left[\frac{\bar{W}(T_s)}{\bar{W}(T)}\right].$$

La sustitución de la ecuación (2.7) en (2.10) conduce entonces a la forma final para el tiempo  $\tau_r$  en función de la temperatura  $T$ , de la entropía configuracional  $s^*$  de la región cooperativa de tamaño crítico  $z^*$  y de la entropía configuracional  $S_c$  del sistema, la cual se escribe como

$$(2.11) \quad \tau_r(T) = \tau_r(T_s) \exp\left(-\frac{C}{RS_c} \left[\frac{1}{T_s} - \frac{1}{T}\right]\right),$$

donde  $C = \Delta\mu s^*$  contiene la dependencia en la entropía conformacional de la región cooperativa de tamaño  $z^*$  y  $R$  es la constante de los gases. Como caso de análisis, Adam y Gibbs observan que si se considera que el  $\Delta C_p$  de la transición líquido-estructura vítrea no depende de la temperatura  $T$ , entonces es posible calcular el cambio en la entropía configuracional  $\Delta S_c$  obteniéndose en ese caso una forma funcional para el tiempo de relajación estructural del tipo ecuación WLF [Adam and Gibbs, 1965]. La obtención de una ecuación del tipo WLF a partir del formalismo de Adam y Gibbs suponiendo que el cambio del calor específico  $\Delta C_p$  es constante, sólo puede considerarse como una primera aproximación al comportamiento real del calor específico y válida sólo en un intervalo de temperaturas muy cercanas a la transición vítrea.

Es justo aquí entonces donde nuestro enfoque para analizar la desnaturalización de proteínas diverge de lo planteado por Adam y Gibbs. Si bien pudiera considerarse que el  $\Delta C_p$  tiene un comportamiento constante en la vecindad de la temperatura  $T_g$  en el caso de la transición vítrea, no es el caso para la desnaturalización de proteínas. Tal como lo expusimos en secciones anteriores, los primeros experimentos de microcalorimetría de barrido para macromoléculas biológicas fueron analizados con esta suposición, la cual refinamientos posteriores en las técnicas experimentales mostraron como equivocada. Procederemos pues a continuación a analizar la fractura del estado nativo para la mioglobina, una proteína prototipo en donde el cambio del calor específico  $\Delta C_p$  no es constante respecto de la temperatura.

### 3. Desnaturalización de la Mioglobina I. Funciones Termodinámicas.

Para determinar el tiempo de desnaturalización  $\tau$  de una macromolécula biológica desde la perspectiva propuesta por Adam y Gibbs, requerimos conocer dos cantidades

termodinámicas que caracterizan a la transición  $D \rightleftharpoons N$ . Por un lado, es necesario conocer el cambio de la entropía configuracional  $\Delta_N^D S$  en función de la temperatura para la macromolécula en conjunto y por otro, debemos determinar el cambio de entropía conformacional  $\Delta s$  de la región cooperativa de tamaño mínimo  $z^*$  que es capaz de exhibir una transición conformacional. La observación anterior define entonces las líneas a seguir. En primera instancia, abordaremos el cálculo del cambio de entropía configuracional  $\Delta S$  a partir de su conexión con el calor específico  $\Delta C_p$  y de la información experimental que tenemos sobre éste. En segunda instancia, presentaremos un enfoque teórico basado en geometría diferencial para calcular el cambio en la entropía conformacional  $\Delta s$  de las regiones cooperativas. El enfoque de geometría diferencial presentado aquí resulta un elemento novedoso en el campo de la teorías de la transición hebra-hélice, las cuales generalmente utilizan el formalismo de Zimm-Bragg y de variaciones sobre éste. Una de las ventajas de este enfoque sobre las teorías estándares de la transición hebra-hélice consistirá en la eliminación del elemento estocástico presente en ellas a través del parámetro  $\sigma$ , que representa la probabilidad de iniciar el mecanismo de nucleación en la cadena polipeptídica. Finalmente integraremos los elementos anteriores para determinar el tiempo de desnaturalización  $\tau$  (o plegamiento dado que estamos considerando reacciones reversibles) incorporando esta información en la ecuación (2.11).

Para enfocar las disgresiones subsecuentes en el marco del plegamiento de las proteínas con relevancia biológica, consideraremos como caso de análisis a la proteína *mioglobina*, cuya desnaturalización ha sido analizada extensivamente [Privalov and Makhatadze., 1990], [Abruzzetti *et al*, 2006], [Miksovska *et al*, 2006]. Pese a que los estudios de desnaturalización térmica de la mioglobina son relativamente recientes, la determinación de su estructura nativa corresponde a la década de los sesenta y es considerado como una de las primeras aplicaciones de esta técnica al campo de la biología molecular. De hecho, el descubrimiento de la conformación tridimensional del estado nativo de la mioglobina fue una de las primeras confirmaciones experimentales de la existencia de estructuras helicoidales estables en macromoléculas biológicas [Kendrew., 1958].

La mioglobina y la hemoglobina pertenecen a la categoría de proteínas conocidas como *hemo-proteínas*, cuya relevancia desde el punto de vista biológico se halla en el hecho de su capacidad para la unión, transporte y almacenamiento del oxígeno molecular presente en el torrente sanguíneo [Branden and Tooze., 1991]. A diferencia de la hemoglobina, que es en realidad un tetrámero conformado por cuatro cadenas polipeptídicas que pueden plegarse en forma independiente, la mioglobina está integrada por una sola cadena polipeptídica que se pliega en ocho hélices  $\alpha$ , denominadas con letras, de la A a la H. Dada esta característica, el contenido helicoidal del estado nativo de la mioglobina corresponde aproximadamente al 77 % al total de la macromolécula, en tanto que el 23 % restante corresponde a estructuras no helicoidales, principalmente  $\beta$ -turns que enlazan a las distintas hélices  $\alpha$  [Voet., 1995]. Además de lo anterior, la inexistencia de hojas  $\beta$  y de enlaces disulfuro entre aminoácidos *Cis* permite que los mecanismos de formación para la estructura nativa presenten un alto grado de reversibilidad, dado que como hemos discutido anteriormente los procesos de formación y ruptura de enlaces disulfuro generalmente impiden la recuperación de la estructura nativa en ausencia de las enzimas que catalizan esta reacción. La estructura primaria de la mioglobina está integrada por  $K = 153$  aminoácidos que en



conjunto proporcionan una masa molecular de  $M_w = 16.7$  kDa, lo cual la ubica dentro de la categoría de proteínas pequeñas que exhiben un alto grado de cooperatividad durante la transición  $D \rightleftharpoons N$ .

En referencia a la capacidad que exhibe esta macromolécula para ligar y almacenar átomos de oxígeno, ésta se debe a la existencia de un átomo central de Fe dentro del grupo hemo, el cual dado que generalmente se halla en un estado ionizado del tipo  $Fe^{+2}$ , exhibe una afinidad significativa para el ligamiento con el oxígeno molecular. Asimismo, la presencia del átomo de nitrógeno del aminoácido *His* ubicado justo por encima del plano del grupo hemo, genera un incremento adicional en la estabilidad de la reacción correspondiente, haciendo así que el  $\Delta G$  para esta reacción sea negativo [Reeder *et al*, 2004]. La Figura (5) muestra la estructura nativa de la molécula de mioglobina obtenida mediante los datos de difracción de rayos X depositados en el *Protein Data Bank*: (1MBN).

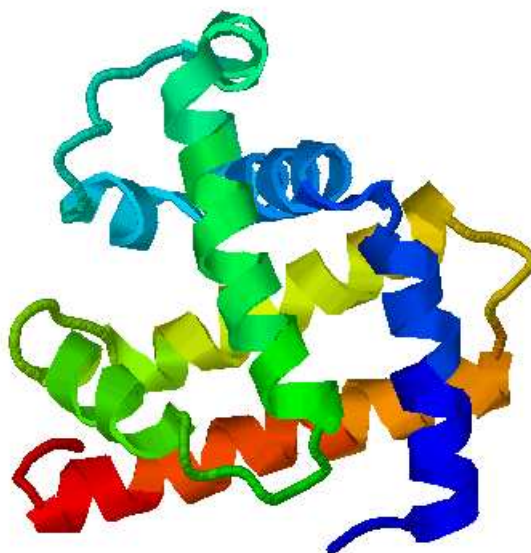


FIGURA 5. Estructura nativa de la mioglobina (PDB: 1MBN) obtenida mediante los datos de difracción de rayos X reportados por Watson y Kendrew [Watson and Kendrew, 1969] depositados en el *Protein Data Bank*. La estructura nativa está conformada por ocho hélices  $\alpha$  y  $K = 153$  aminoácidos. El contenido helicoidal corresponde al 77 % del total de la macromolécula y el 23 % restante corresponde a estructuras no helicoidales de enlace.

El esquema de Adam-Gibbs sugiere que la forma funcional del tiempo de relajación estructural  $\tau$  está asociado directamente con el cambio de entropía configuracional  $\Delta S$ . A fin de determinar esta cantidad procederemos a través de la relación que

conecta al cambio en el calor específico  $\Delta C_p$  con el cambio de entropía  $\Delta S$ . Esto es,

$$(3.1) \quad \frac{\partial \Delta S}{\partial T} = \frac{\Delta C_p}{T}.$$

La ecuación (3.1) es una relación completamente general válida para cualquier sistema termodinámico. Para calcular el cambio de entropía configuracional requerida por el esquema de Adam-Gibbs es necesario considerar entonces la contribución configuracional al calor específico. Aún en el caso de líquidos orgánicos o inorgánicos de composición molecular simple, la determinación de la contribución exclusivamente configuracional al  $\Delta C_p$  es un problema que plantea retos en el campo experimental y teórico [Cangialosi *et al*, 2006]. En el caso de los sistemas vítreos estándares, generalmente se considera que la diferencia entre el calor específico del estado líquido  $C_p^l$  y el estado cristalino  $C_p^c$  es proporcional a la contribución configuracional, de forma entonces que se considera que  $\Delta C_p^{config} = C_p^c - C_p^l$  [Cangialosi *et al*, 2006]. En el caso de la desnaturalización de las proteínas, la información experimental obtenida directamente de experimentos de microcalorimetría se refiere a la contribución proveniente de la solvatación de los grupos no polares del núcleo hidrofóbico al solvente [Privalov and Dragan, 2007], de forma entonces que consideraremos que el calor específico  $\Delta C_p$  puede escribirse de acuerdo a lo siguiente,

$$(3.2) \quad \Delta C_p \simeq \Delta C_p^{hyd},$$

en donde  $\Delta C_p^{hyd}$  se refiere a la contribución de solvatación antes mencionada. En términos generales, la solvatación de los grupos no polares tiene dos contribuciones, una parte vibracional  $\Delta C_p^v$  y una parte configuracional  $\Delta C_p^{config}$ , las cuales provienen de la excitación de modos vibracionales y rotacionales de la proteína en el estado nativo.

Para el caso de la mioglobina, un análisis en términos de modos normales de oscilación y su comparación con resultados experimentales mediante espectroscopía Mossbauer [Achterhold *et al*, 2002], muestra que la contribución vibracional al calor específico  $\Delta C_p^v$  es del orden de  $10^{-3}$  veces la contribución configuracional, de forma entonces que para esta molécula podemos escribir que

$$(3.3) \quad \Delta C_p^{hyd} \simeq \Delta C_p^{config}.$$

Para determinar la forma funcional del  $\Delta C_p^{hyd}$  en función de la temperatura, utilizamos los datos experimentales reportados por Privalov y Makhatadze [Privalov and Makhatadze., 1990] para la desnaturalización de la mioglobina en el intervalo de temperaturas  $5^\circ C < T < 125^\circ C$ , los cuales se presentan en el Cuadro (1).

Temperatura (Grados Celsius)	5 °C	25 °C	50 °C	75 °C	100 °C	125 °C
$\Delta C_p^{hyd}$ (kJ / Kelvin * mol)	10.2	11.5	12.1	11.5	10.8	9.5

CUADRO 1. Variación del  $\Delta C_p^{hyd}$  para la desnaturalización de la mioglobina en función de la temperatura  $T$  de acuerdo a [Privalov and Makhatadze., 1990].

Los datos anteriores muestran que el  $\Delta C_p^{hyd}$  no es constante respecto de la temperatura en el intervalo analizado y que, tal como se espera en una desnaturalización dirigida fundamentalmente por el efecto hidrofóbico, el  $\Delta C_p > 0$ . Si bien es posible en principio ajustar diversas formas funcionales consistentes con el comportamiento

observado para  $\Delta C_p$  versus  $T$ , en este trabajo propondremos una representación para  $\Delta C_p^{hyd}$  en términos de un polinomio de grado dos en  $T$ . La razón para esta elección se basa en dos argumentos. Por un lado, un polinomio cuadrático es la función con el menor número de parámetros ajustables consistente con el comportamiento observado y por el otro, se sabe que la desnaturalización de varias proteínas, entre ellas la *quimotripsina*, es consistente con un modelo cuadrático para el  $\Delta C_p^{hyd}$  en función de  $T$  [Brandts et al., 1992]. Así dicho, propondremos que en el caso de la mioglobina, el  $\Delta C_p^{hyd}$  en función de  $T$  está dado por la ecuación

$$(3.4) \quad \Delta C_p^{hyd}(T) = \Delta C_p^0 + M_1(T - T_0) + M_2(T - T_0)^2,$$

en donde  $\Delta C_p^0$ ,  $M_1$  y  $M_2$  son constantes a ajustar a partir de los datos experimentales y  $T_0$  es una temperatura de referencia, la cual como veremos en párrafos subsecuentes corresponde a la temperatura en donde el cambio de energía libre correspondiente  $\Delta G(T_0) = 0$  y por tanto es la temperatura de equilibrio del sistema. En términos de la reacción  $N \rightleftharpoons D$ ,  $T_0$  es la temperatura a la cual la fracción de moléculas en el estado nativo es idénticamente igual a la concentración de moléculas en el estado desplegado.

El ajuste por mínimos cuadrados de la ecuación (3.4) a los datos experimentales mostrados en el Cuadro (1) se muestra en la Figura (6) y conduce a los valores  $\Delta C_p^0 = 1.13 \times 10^4$  J/Kelvin\* $\text{mol}$ ,  $M_1 = 36.36$  J/Kelvin<sup>2</sup>\* $\text{mol}$  y  $M_2 = -0.55$  J/Kelvin<sup>3</sup>\* $\text{mol}$ , en donde se ha elegido la temperatura de equilibrio de la desnaturalización como  $T_0 = 298$  K. Es importante notar que el coeficiente  $M_2$  necesariamente debe ser negativo a fin de que la ecuación (3.4) exhiba un máximo como lo indican los resultados experimentales. Considerando entonces que la solvatación de los grupos hidrofóbicos es el efecto dominante en la desnaturalización de proteínas y que en el caso específico de la mioglobina la contribución vibracional es mucho menor que la contribución rotacional, procedemos a la integración de la ecuación

$$(3.5) \quad \frac{\partial \Delta S^{config}}{\partial T} = \frac{\Delta C_p^{hyd}}{T}.$$

para determinar el cambio de entropía configuracional  $\Delta S^{config}$  asociado a esta desnaturalización.

Integrando las relaciones termodinámicas [Pathria, 1972],

$$(3.6) \quad \begin{aligned} \frac{\partial \Delta H}{\partial T} &= \Delta C_p \\ \frac{\partial \Delta S}{\partial T} &= \frac{\Delta C_p}{T} \\ \Delta G(T) &= \Delta H(T) - T\Delta S(T), \end{aligned}$$

(3.7)

se tienen entonces los siguientes resultados para el cambio de entalpía  $\Delta H$ , para el cambio de entropía  $\Delta S$  y para el cambio de la energía libre  $\Delta G$ . Esto es,

$$(3.8) \quad \Delta H^{config}(T) = \Delta H^0 + \Delta C_p^0(T - T_0) + \frac{1}{2}M_1(T - T_0)^2 + \frac{1}{3}M_2(T - T_0)^3,$$

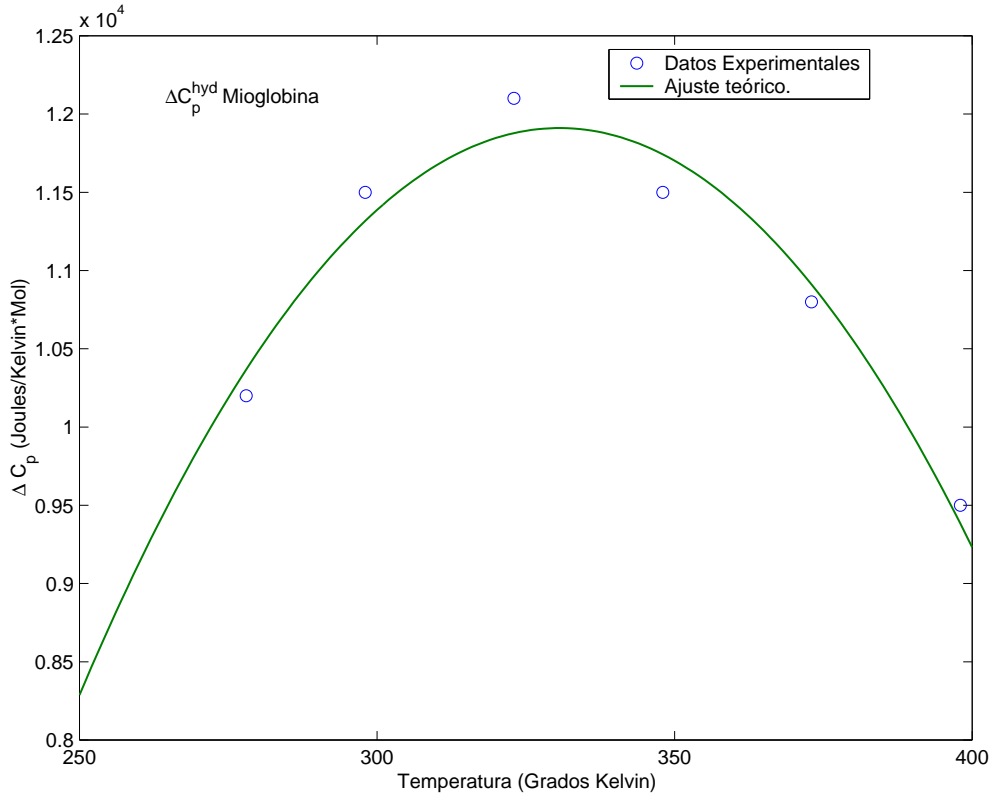


FIGURA 6. Comportamiento del  $\Delta C_p^{hyd}$  en función de  $T$  para la mioglobina (PDB: 1MBN). Los puntos representan los datos experimentales de microcalorimetría de barrido reportados por [Privalov and Makhatadze., 1990] en tanto que la línea sólida representa el ajuste por mínimos cuadrados a estos datos de la ecuación (3.4). La temperatura de equilibrio de la reacción  $N \rightleftharpoons D$  se eligió como  $T_0 = 298$  K. Los valores obtenidos para las constantes son  $\Delta C_p^0 = 1.13 \times 10^4$  J/Kelvin\*mol,  $M_1 = 36.36$  J/Kelvin<sup>2</sup>\*mol y  $M_2 = -0.55$  J/Kelvin<sup>3</sup>\*mol

(3.9)

$$\Delta S^{config}(T) = \Delta S^0 + \Delta C_p^0 \log(T/T_0) + M_1(T - T_0) - DT_0 \log(T/T_0) + \frac{M_2}{2}(T^2 - T_0^2) + 2M_2T_0(T - T_0) + M_2T_0^2 \log(T/T_0),$$

y

$$(3.10) \quad \Delta G^{config} = \Delta H^0 + \Delta C_p^0(T - T_0) + \frac{M_1}{2}(T - T_0)^2 + \frac{M_2}{3}(T - T_0)^3 + T\Delta S^0 - T\Delta C_p^0 \log(T/T_0) - TM_1(T - T_0) + M_1T_0T \log(T/T_0) - \frac{TM_2}{2}(T^2 - T_0^2) + 2M_2T_0T(T - T_0) - TD_1T_0^2 \log(T/T_0).$$

Las ecuaciones (3.8), (2.7) y (3.10) representan entonces las funciones termodinámicas para la desnaturalización de la mioglobina, en donde  $M_1$ ,  $M_2$  y  $\Delta C_p^0$  tiene los valores numéricos reportados previamente y las cantidades  $\Delta H^0$ ,  $\Delta S^0$  y  $\Delta G^0$

están referidas a la temperatura de equilibrio  $T_0 = 298$  K. A partir de este momento, omitiremos la notación que hace referencia explícita a la contribución configuracional, por lo que a menos que se indique lo contrario, los potenciales termodinámicos que se describen a continuación provienen de contribuciones rotacionales.

La Figura (7) muestra el comportamiento de  $\Delta H(T)$  y  $\Delta G(T)$  en función de  $T$ , en tanto que la Figura (8) muestra la dinámica del cambio de entropía  $\Delta S(T)$ .

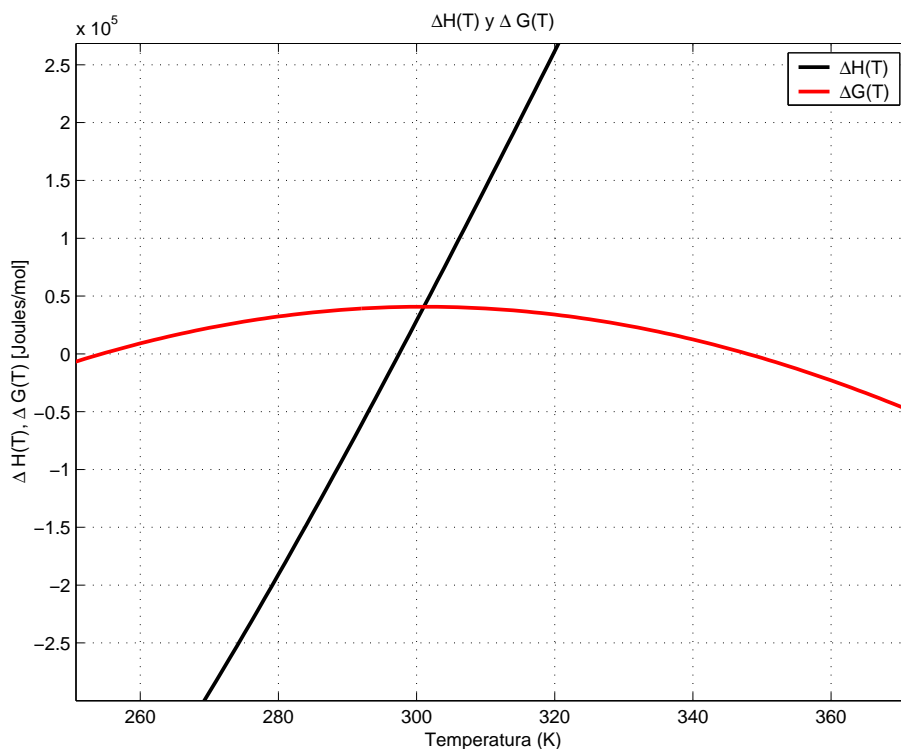


FIGURA 7. Comportamiento de  $\Delta H(T)$  (negro) y  $\Delta G(T)$  (rojo) en función de la temperatura de acuerdo a las ecuaciones (3.8) y (3.10). Se observa un comportamiento esencialmente lineal para  $\Delta H$  en función de  $T$ , en tanto que la curva  $\Delta G$  muestra un máximo de estabilidad en  $T_0 = 298$  K.

Pese a que la ecuación (3.8) muestra un comportamiento logarítmico para el cambio de entalpía en función de la temperatura  $T$ , el valor de los coeficientes que rigen este comportamiento es mucho menor que su contraparte lineal. De forma entonces que  $\Delta H$  muestra un perfil lineal en función de  $T$ , con un cambio de signo a la temperatura  $T_0$ . Para  $T < T_0$ , la entalpía  $\Delta H < 0$  indicando entonces que el proceso de desnaturalización está dirigido básicamente por el término entrópico  $-T\Delta S > 0$ , dado que  $\Delta G > 0$  en esta región. Para  $T > T_0$  y hasta una temperatura aproximada de  $T = 350$  K, la entalpía  $\Delta H > 0$  y  $\Delta S > 0$  de forma que el proceso comienza a ser dominado por la contribución entálpica, hasta alcanzar la temperatura aproximada de  $T = 350$  Kelvin, a partir de donde la contribución entrópica vuelve a ser dominante, induciendo el cambio de signo  $\Delta G < 0$ .

#### 4. DESNATURALIZACIÓN DE LA MIOGLOBINA II. ENTROPÍA DE LAS REGIONES COOPERATIVAS

La Figura (8) que se refiere al cambio de entropía  $\Delta S$  en función de  $T$  muestra que para temperaturas  $T < T_0$ , la contribución entrópica a la desnaturalización es positiva, dado que  $\Delta S < 0$ .

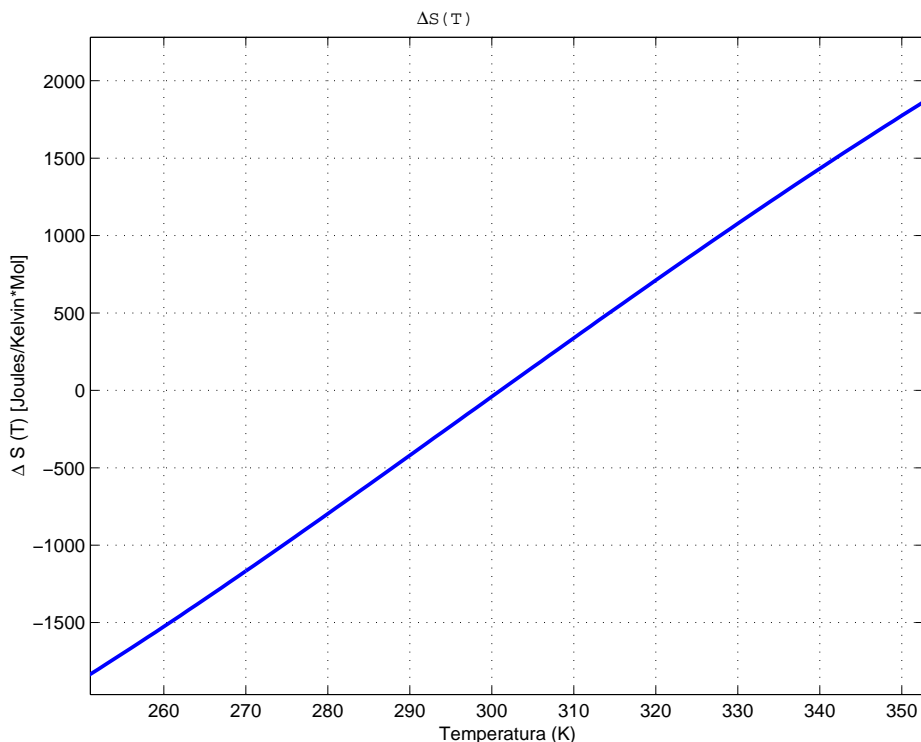


FIGURA 8. Comportamiento de  $\Delta S(T)$  en función de  $T$  de acuerdo a la ecuación (2.7). Para  $T < T_0$ , el cambio  $\Delta S < 0$  de forma entonces que la contribución entrópica a la energía libre  $\Delta G$  es positiva. Para  $T > T_0$  y hasta un valor límite 350 Kelvin aproximadamente, el efecto entrópico comienza a disminuir la estabilidad del sistema, dado que en este intervalo  $\Delta S > 0$  y por tanto el término  $-T\Delta S < 0$

El resultado anterior puede interpretarse de la siguiente manera. Conforme la temperatura desciende por debajo de  $T_0$ , el cambio de entropía entre el estado nativo  $N$  y el estado desnaturalizado se incrementa negativamente, esto es, la entropía del estado nativo es mayor que la del estado desnaturalizado, en acuerdo con lo esperado. Sin embargo, para  $T > T_0$ , el cambio de entropía  $\Delta S > 0$  sugiriendo entonces que conforme la temperatura se incrementa, se inicia propiamente la marca del efecto hidrofóbico; los grupos no polares son expuestos al solvente debido a la fractura del estado nativo induciendo un rearrreglo conformacional en las moléculas polares del solvente.

#### 4. Desnaturalización de la Mioglobina II. Entropía de las Regiones Cooperativas.

De acuerdo al programa planteado por la teoría de Adam-Gibbs, una vez establecido el cambio de entropía configuracional  $\Delta S$  para el sistema en conjunto, requerimos

determinar el cambio configuracional  $\Delta s$  de la región cooperativa de menor tamaño que es capaz de exhibir una transición conformacional. En el caso de las proteínas, las regiones cooperativas pueden identificarse en forma natural con las regiones conocidas como **dominios**, las cuales corresponden a subestructuras del estado nativo que son capaces de plegarse o desnaturalizarse en forma independiente del resto de la cadena polipeptídica y que generalmente tienen una función biológica específica. Para nuestro caso de análisis, consideraremos que las regiones cooperativas del formalismo de Adam y Gibbs pueden identificarse con las hélices  $\alpha$  de la estructura nativa de la mioglobina. De forma entonces que la determinación del cambio de entropía configuracional  $\Delta s$  mencionado al principio, se referirá al cambio de la entropía configuracional entre el estado helicoidal de un conjunto de aminoácidos y su estado desplegado. La descripción de las características termodinámicas y cinéticas de transición anterior se agrupan generalmente dentro de la llamada *transición hebra-hélice*, la cual ha sido discutida ampliamente desde los trabajos pioneros de Lifshitz *et al* [Lifshitz, Grosberg and Khokhlov, 1978] y más recientemente por Grosberg y Kuznetsov [Grosberg and Khokhlov, 1994]. Junto con las teorías anteriores, la propuesta de Zimm-Bragg para describir la formación y estabilidad de la estructura helicoidal en términos del parámetro de nucleación  $\sigma$  representan el panorama más conocido en esta dirección.

En este trabajo abordaremos, sin embargo, un enfoque distinto. A diferencia del formalismo de Zimm-Bragg que contiene elementos estocásticos en esencia, el esquema que proponemos es completamente determinista y está basado en el cálculo de la energía de formación de una estructura helicoidal considerándola como un polímero elástico. Partimos de la idea entonces que para deformar un polímero lineal se requiere una energía de deformación, la cual asociaremos con la energía interna del sistema. A partir de este punto, utilizaremos las herramientas estándares de la mecánica estadística de equilibrio para calcular la función de partición a través de su conexión con la energía interna del sistema  $Z = \int_{\Omega} \exp(-\beta E_c) dC$ , en donde  $\Omega$  representa el espacio de conformaciones accesibles al sistema y  $E_c$  la energía elástica del polímero en función de su conformación. Para determinar la energía necesaria para construir la estructura helicoidal del polímero elástico utilizaremos el marco de la geometría diferencial de curvas en el espacio. Este enfoque ha sido utilizado en trabajos anteriores para analizar la deformación de estructuras helicoidales de ADN y otro tipo de polímeros no orgánicos [Marko and Siggia, 1994], [Kamien *et al*, 1997].

Consideremos entonces un polímero elástico de longitud  $L$ , el cual puede deformarse para generar una estructura helicoidal. Para representarlo, definimos una curva  $\mathbf{R}$  en el espacio parametrizada por su longitud de arco  $\gamma$ , de acuerdo a  $\mathbf{R}(\gamma) = [x(\gamma), y(\gamma), z(\gamma)]$ , en donde  $\mathbf{R}(\gamma)$  es entonces una función vectorial en la variable real  $\gamma$ . Asociado a esta curva, construimos el vector tangente  $\mathbf{T}$ , el cual es función también de la longitud de arco  $\gamma$  y está dado por la ecuación,

$$(4.1) \quad \mathbf{T}(\gamma) = \frac{d\mathbf{R}(\gamma)}{d\gamma}.$$

De la definición misma del vector tangente (4.1) observamos que para el caso en que éste es constante, la curva  $\mathbf{R}$  corresponde a una línea recta, por lo que en este caso tendríamos una conformación lineal para el polímero. Esta observación es crucial puesto que entonces sugiere que cualquier variación conformacional de la estructura

lineal debe estar asociada con la derivada del vector  $\mathbf{T}$ . Así, la energía  $E_c$  de deformación del polímero puede escribirse en una primera aproximación como directamente proporcional al cuadrado de la primera derivada del vector tangente [Kamien, R.D., 2002]. Esto es,

$$(4.2) \quad E_c = A \int_0^L \left[ \frac{d\mathbf{T}(\gamma)}{d\gamma} \right]^2 d\gamma,$$

en donde  $A$  es una constante relacionada con la rigidez de los monómeros que conforman al polímero. Dentro del formalismo de la geometría diferencial de curvas en el espacio, la magnitud de la derivada del vector tangente  $\|\mathbf{T}'(\gamma)\|$  es justamente la *curvatura*  $\kappa(\gamma)$ , de forma entonces que la ecuación (4.2) puede escribirse de acuerdo a,

$$(4.3) \quad E_c = A \int_0^L \kappa^2(\gamma) d\gamma.$$

La ecuación (4.3) representa nuestro punto de partida e indica entonces que la energía requerida para deformar una curva  $\mathbf{R}$  es proporcional al cuadrado de la curvatura  $\kappa$  de la misma. Las ecuaciones de Frenet-Serret en geometría diferencial muestran que de hecho, sólo se requieren dos parámetros para definir unívocamente a una curva en el espacio [Stoker., 1969], la curvatura  $\kappa$  previamente definida y la torsión  $\varsigma$ , la cual representa la rotación del vector tangente  $\mathbf{T}$  respecto del vector normal  $\mathbf{N}$ , definido por  $\mathbf{T} \cdot \mathbf{N} = 0$  para toda  $\gamma$  en el intervalo de existencia de la curva.

La parametrización canónica de una curva helicoidal en términos de la longitud de arco  $\gamma$  está dada por la ecuación,

$$(4.4) \quad \mathbf{R}(u) = \left[ a \cos \left( \frac{\gamma}{\sqrt{a^2 + b^2}} \right), a \sin \left( \frac{\gamma}{\sqrt{a^2 + b^2}} \right), \frac{b\gamma}{\sqrt{a^2 + b^2}} \right].$$

en donde  $a$  representa el radio de la hélice y  $b$  es un parámetro directamente proporcional al desplazamiento vertical de la curva (*pitch*) (Figura 9).

Utilizando las ecuaciones de Frenet-Serret [Stoker., 1969] que definen a la curvatura  $\kappa$  y la torsión  $\varsigma$  en términos del vector  $\mathbf{R}$  y sus derivadas, se puede demostrar que para una curva helicoidal del tipo (Ecuación 4.4), tanto  $\kappa$  como  $\varsigma$  son independientes de la longitud de arco y se escriben en términos de  $a$  y  $b$  de acuerdo a,

$$(4.5) \quad \begin{aligned} \kappa &= \frac{a}{a^2 + b^2} \\ \varsigma &= \frac{b}{a^2 + b^2}. \end{aligned}$$

Es posible también escribir una representación distinta para la hélice en términos de su radio  $r$  y del desplazamiento vertical  $p$  (*pitch*). En este nuevo conjunto de parámetros, el parámetro  $a = r$  y el parámetro  $b = p/2\pi$ . La ventaja al utilizar el dupleta  $(r, p)$  para describir a la hélice radica en que estos parámetros están bien determinados para distintas estructuras helicoidales en proteínas. En el caso de una hélice  $\alpha$ , los parámetros  $r = 0.30$  nm y  $p = 0.54$  nm [Branden and Tooze., 1991]. En términos de  $(r, p)$ , la curvatura  $\kappa$  queda escrita como

$$(4.6) \quad \kappa = \frac{4\pi^2 r}{p^2 + 4\pi^2 r^2}.$$

Sustituyendo la ecuación (4.6) en (4.3), obtenemos una expresión analítica para la energía requerida para deformar un polímero de longitud  $L$  desde una conformación



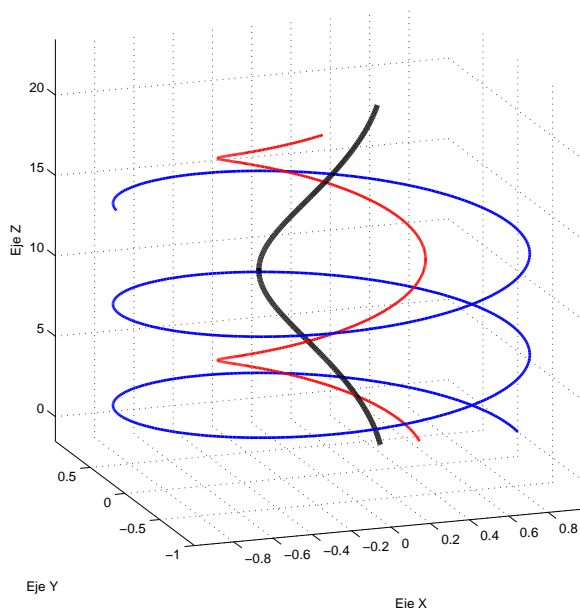


FIGURA 9. Representación de una curva helicoidal en 3D utilizando la parametrización canónica en términos de la longitud de arco  $\gamma$  y los parámetros  $a$  y  $b$  (Ecuación 4.4) En azul ( $a = 1, b = 1$ ), en rojo ( $a = 0.5, b = 2$ ), en negro ( $a = 0.3, b = 4$ ). La longitud de arco  $\gamma$  recorre el intervalo  $[0, 7\pi]$ .

lineal a una conformación helicoidal de radio  $r$  y desplazamiento vertical  $p$ . Esta expresión es,

$$(4.7) \quad E_c = A \int_0^L \left( \frac{4\pi^2 r}{p^2 + 4\pi^2 r^2} \right)^2 d\gamma.$$

Puesto que la curvatura  $\kappa$  es constante respecto de la longitud de arco  $\gamma$  [Guggenheimer., 1963], la integración de (4.7) es directa, obteniéndose que la energía de formación de la hélice está dada por

$$(4.8) \quad E_c = AL \left( \frac{4\pi^2 r}{p^2 + 4\pi^2 r^2} \right)^2 = \frac{16AL\pi^4 r^2}{(4\pi^2 r^2 + p^2)^2},$$

en donde la constante de rigidez del polímero  $A$  tiene unidades de [Joule· metro/mol].

La ecuación (4.8) indica que la energía  $E_c$  tiene una relación inversamente proporcional tanto con el radio  $r$  de la hélice, como con el desplazamiento vertical  $p$  (*pitch*). Más precisamente,  $E_r \sim 1/r^2$  y  $E_r \sim 1/p^4$ . Lo anterior es razonable si pensamos que conforme  $r \rightarrow 0$ , requerimos un mayor gasto energético para conformar la estructura helicoidal, dado que la constante de rigidez  $A$  del polímero tiene un valor fijo y por tanto se debe invertir una mayor cantidad de energía para doblar el polímero en un espacio menor. Por otro lado, la dependencia en el parámetro  $p$  es fundamental y muestra un acuerdo con lo esperado. Recordemos que la cantidad  $p$  indica la distancia que se desplaza la hélice a lo largo del eje vertical en cada vuelta. El valor  $p = 0$

se refiere al caso límite de una hélice restringida a una curva bidimensional, esto es a un círculo de radio  $r$ , como puede verse de la ecuación (4.4). Fuera de este caso trivial, conforme  $p$  se incrementa, la hélice se halla en conformaciones más alargadas, de forma entonces que deberíamos esperar que la energía requerida para esta conformación disminuya a medida que  $p$  se incrementa, en acuerdo con lo observado en la forma funcional de  $E_r$  respecto de  $p$  (Ecuación 4.8). Finalmente observemos que para el caso  $p = 0$ , el polímero ha adquirido una conformación circular plana de radio  $r$  y entonces su energía  $E_r$  debería estar dada por el cuadrado de su curvatura y que en el caso de un círculo corresponde a  $\kappa = 1/r$ , resultado que se obtiene al sustituir  $p = 0$  en (4.8).

Considerando entonces como una primera aproximación para la energía del sistema a la energía de deformación  $E_c$ , la función de partición se escribe de acuerdo a,

$$(4.9) \quad Z = \int_{\Omega} \exp[-\beta E_c(T)] d\Omega$$

en donde  $\Omega$  denota el espacio configuracional donde se realiza la integración de la energía  $E_r(T)$  y  $\beta = (k_B T)^{-1}$  representa el factor de Boltzmann. En forma estricta, la integral definida en (4.9) está definida sobre un espacio configuracional semi-infinito en las variables  $(r, p)$ , dado que una transición hebra-hélice puede ocurrir de infinitas formas dentro de este espacio conformacional. Sin embargo, haremos la suposición de que para un dominio helicoidal en una cadena polipeptídica biológica, como es el caso de una proteína, sólo es posible que la transición hebra-hélice ocurra entre dos estados: el estado totalmente desnaturalizado y un estado helicoidal específico, caracterizado por un dupleta  $(r_0, p_0)$ . La suposición anterior se basa en el hecho de que la mayoría de las transiciones hebra-hélice en cadenas polipeptídicas biológicas ocurren entre estados desplegados y tipos específicos de estructuras helicoidales (hélices  $\alpha$ , hélices  $\pi_{10}$ , etc). Es muy improbable que una estructura en proceso de plegamiento quede atrapada en un conformación helicoidal distinta a las mencionadas anteriormente.

Con esta suposición, consideraremos entonces que la función de partición puede representarse como una suma de dos estados. Esto es,

$$(4.10) \quad Z = \sum_{j=0}^1 g_j \exp(-\beta E_j) = g_0 \exp(-\beta E_0) + g_1 \exp(-\beta E_1),$$

en donde  $E_0$  es la energía de la estructura helicoidal,  $E_1$  es la energía del estado desplegado y  $g_0$  y  $g_1$  son las degeneraciones correspondientes. El desarrollo presentado en la sección anterior nos indica que el valor absoluto de la energía necesaria para deformar un polímero está dada por la ecuación (4.8). Si consideramos por tanto que la estructura desplegada corresponde a la conformación lineal del polímero, la energía  $E_1 = 0$ , puesto que en este estado la curvatura  $\kappa$  es idénticamente cero. Para determinar la energía  $E_0$  del estado helicoidal óptimo, asumiremos que ésta corresponde a la energía  $-E_c(r_0, p_0)$ , en donde los valores  $r_0$  y  $p_0$  corresponden a una estructura helicoidal de tipo hélice  $\alpha$ . De esta forma tenemos entonces un sistema con dos niveles de energía accesibles, en el cual existe un estado de menor energía identificable con la energía de deformación necesaria para construir la estructura helicoidal. Es importante mencionar que en esta argumentación, no estamos deduciendo que el estado

$E_0 = -E_c(r_0, p_0)$  es el estado que corresponde al mínimo de energía. Estamos en realidad proponiendo que *existe* ese valor mínimo y lo identificamos con la cantidad  $-E_c(r_0, p_0)$ . De esta forma, la función  $Z$  se escribe como,

$$(4.11) \quad Z = g_1 + g_0 \exp [\beta E_c(r_0, p_0)]$$

El estado con energía  $E_0$  es único, por lo que  $g_0 = 1$ . Sin embargo, la conformación lineal con energía  $E_1 = 0$  tiene una dos subestados en el espacio conformacional  $\Omega$ . El primero corresponde a una elongación infinita ( $p \rightarrow \infty$ ) y una contracción del radio ( $r \rightarrow 0$ ), en tanto que la segunda se obtiene estirando el radio de un círculo en forma infinita ( $p = 0, r \rightarrow \infty$ ). Ambas situaciones límite tiene energía de deformación cero, como puede verse de la ecuación (4.8). Esto implica entonces que  $g_1 = 2$ , por lo que la función  $Z$  se escribe finalmente como,

$$(4.12) \quad Z = 2 + \exp [\beta E_c(r_0, p_0)].$$

Utilizando la ecuación (4.12), y la relación que conecta el cambio de entropía con la función de partición  $Z$  [Binney *et al*, 1992],

$$(4.13) \quad S(T) = -\frac{1}{T} \left( \frac{\partial \log Z}{\partial \beta} \right) + R \log Z,$$

podemos calcular el cambio de entropía configuracional  $\Delta s$  para una transición entre una conformación lineal y una conformación helicoidal, obteniéndose

$$(4.14) \quad \Delta s = - \left[ \frac{E_c(r_0, p_0)/T}{2 \exp(-E_c(r_0, p_0)/RT) + 1} \right] + R \log [2 + \exp(E_c(r_0, p_0)/RT)],$$

en donde  $E_c(r_0, p_0)$  esta dado por,

$$(4.15) \quad E_c(r_0, p_0) = \frac{16AL\pi^4 r_0^2}{(4\pi^2 r_0^2 + p_0^2)^2}$$

Con respecto de la temperatura  $T$ , el cambio de entropía  $\Delta s$  en la ecuación (4.17) muestra un comportamiento monótonico creciente para distintos valores de la energía  $E_0$  de la estructura helicoidal. La Figura (10) muestra este hecho para un valor  $E_0 = 2.5$  kJ/mol, que es la energía inducida por las fluctuaciones térmicas del medio a temperatura  $T = 298$  Kelvin. En este escenario, la formación de la estructura helicoidal es un proceso marginalmente estable, dado que la energía  $E_0$  de formación de la hélice es del mismo orden de magnitud que el de las fluctuaciones térmicas. Asimismo, la Figura (10) muestra que en esta situación, el cambio de entropía  $\Delta s$  se aproxima a un valor límite conforme  $T \rightarrow \infty$ . En el intervalo de temperaturas [273 – 300] Kelvin, el  $\Delta s$  es esencialmente constante, dado que exhibe una variación muy pequeña en el intervalo [7 – 9] Joules/mol Kelvin.

Utilizando entonces las expresiones para el cambio de entropía configuracional  $\Delta S^{config}$  proveniente del ajuste teórico a los datos experimentales de microcalorimetría de barrido,

$$(4.16) \quad \Delta S^{config}(T) = \Delta S^0 + \Delta C_p^0 \log(T/T_0) + M_1(T - T_0) - DT_0 \log(T/T_0) + \frac{M_2}{2}(T^2 - T_0^2) + \\ - 2M_2T_0(T - T_0) + M_2T_0^2 \log(T/T_0),$$

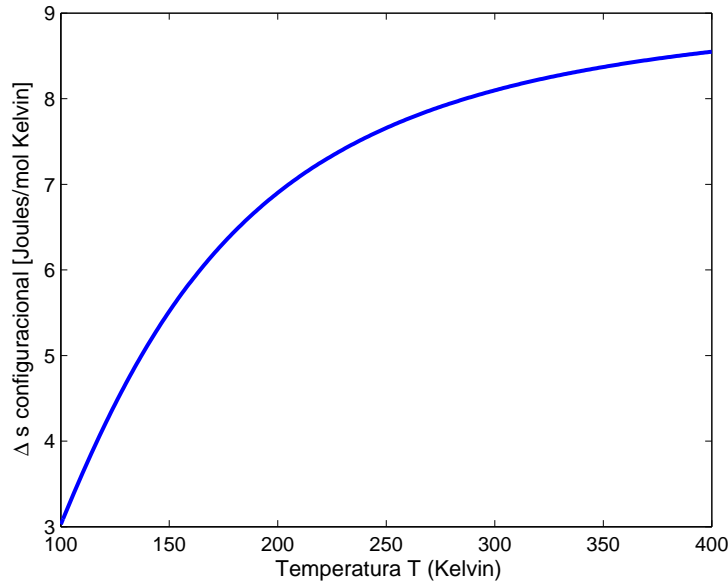


FIGURA 10.  $\Delta s$  en función de  $T$  de acuerdo a la ecuación (4.17). En este caso, la curva corresponde a un valor para la energía de formación de la estructura helicoidal  $E_0 = 2.5$  kJ/mol, el cual corresponde al orden de magnitud de la energía térmica del sistema. En esta situación, la hélice es metaestable y el proceso de formación y ruptura de la estructura helicoidal está dirigido por las fluctuaciones térmicas del medio.

y la expresión analítica deducida para el cambio de entropía configuracional  $\Delta s$  de las regiones cooperativas del sistema,

$$(4.17) \quad \Delta s = - \left[ \frac{E_c(r_0, p_0)/T}{2 \exp(-E_c(r_0, p_0)/RT) + 1} \right] + R \log [2 + \exp(E_c(r_0, p_0)/RT)],$$

podemos escribir la expresión completa para el tiempo de relajación  $\tau_r(T)$  sustituyendo las ecuaciones anteriores en,

$$(4.18) \quad \tau_r(T) = \tau_r(T_s) \exp \left( - \frac{\Delta \mu \Delta s}{R \Delta S_c} \left[ \frac{1}{T_s} - \frac{1}{T} \right] \right).$$

en donde como discutimos anteriormente,  $\Delta \mu$  es la barrera energética necesaria para que ocurra una transición conformacional en las regiones cooperativas y  $T_s$  es una temperatura de referencia en la cual el tiempo  $\tau_r$ , adquiere el valor  $\tau_r(T_s)$ .

En la teoría estándar de la transición vítrea de Adam-Gibbs, el término  $\Delta s$  se asume como un término constante para describir los procesos de relajación en un sistema formador de vidrio. Los resultados que presentamos aquí apoyan esta idea para el caso de una proteína que tiene una transición conformacional de dos estados. Tal como la muestra la Figura (10), la ecuación (4.17) indica que  $\Delta s$  proveniente del cambio de conformación en las regiones cooperativas tiene una contribución menor en comparación con la contribución  $\Delta S$  del cambio configuracional global del

sistema cuando la energía de formación de la estructura helicoidal se halla en el orden de magnitud de las fluctuaciones térmicas. En otras palabras, si las transiciones conformacionales entre las estructuras desplegadas y las estructuras helicoidales son dirigidas esencialmente por las fluctuaciones térmicas, la contribución  $\Delta s$  puede considerarse constante. Considerando el caso de una energía de formación  $E_0 = 10kT$ , equivalente a diez veces la energía de las fluctuaciones térmicas, la ecuación (4.17) predice una contribución dos órdenes de magnitud menor que en el primer caso. En este caso, la energía de formación de la estructura helicoidal se halla en un intervalo superior al de la energía de las fluctuaciones térmicas y por tanto, el baño térmico donde se halla inmersa a temperatura  $T$  no podría proveer de la energía necesaria para iniciar la transición conformacional. La ecuación (4.17) también indica que si la energía de formación  $E_0$  es diez veces menor que la energía térmica  $E = kT$  entonces,  $\Delta s$  contribuye con una cantidad esencialmente constante, que en ese caso corresponde a 9 J/mol Kelvin, que de acuerdo a la información experimental calorimétrica para la mioglobina, es tres órdenes de magnitud menor que  $\Delta S$ .

El término  $\Delta S$  es entonces la contribución dominante para el tiempo de relajación estructural en este sistema. Para los sistemas líquidos formadores de vidrios se asume en general que la variación del  $\Delta C_p$  es constante en la región cercana a la transición vítrea. Esta suposición implica una variación logarítmica para el cambio  $\Delta S$  en función de  $T$  [Adam and Gibbs, 1965], [García-Colin, L.S., 1999], la cual genera ecuaciones del tipo William-Landel y Ferry para el tiempo de relajación estructural. En el caso que estamos analizando, el  $\Delta C_p$  es una función cuadrática en la temperatura  $T$  y genera un perfil más complejo para el tiempo  $\tau_r$  que en el caso de la transición vítrea estándar. La Figura (11) muestra el perfil de la variable  $\tau_r$ ,

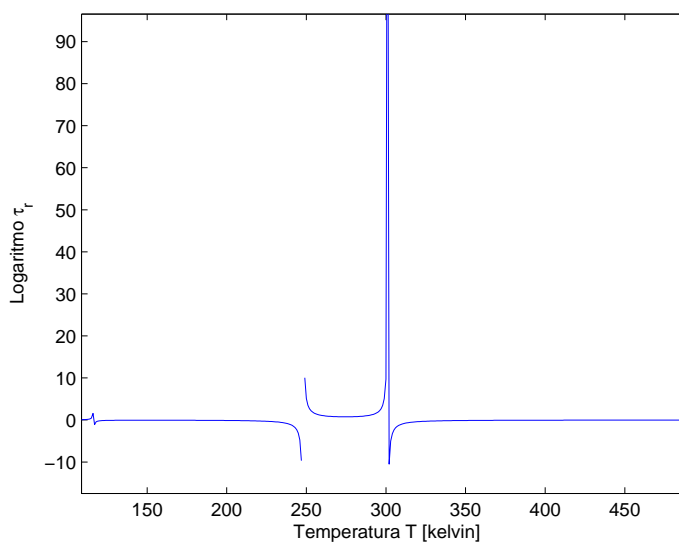


FIGURA 11. Comportamiento del tiempo de relajación estructural  $\tau_r$  en función de  $T$ , de acuerdo a la ecuación (4.23) para  $T_s = 248$  Kelvin,  $\Delta\mu = 2.5$  kJ/mol. La dinámica para  $\tau_r$  muestra una divergencia ubicada en  $T_s$  y un pico resonante ubicada a la temperatura  $T = 300$  Kelvin.

de acuerdo a la ecuación (4.23) cuando se han sustituido las expresiones para  $\Delta s$  y  $\Delta S$  correspondientes y se ha tomado como valor para la barrera energética  $\Delta\mu = 2.5$  kJ/mol y una temperatura de referencia  $T_s = 248$  Kelvin en acuerdo con el criterio propuesto por Adam-Gibbs ( $T_s = T_g - 50$ ). La especificidad de los valores elegidos para la temperatura de referencia  $T_s$  y la barrera energética  $\Delta\mu$  no tiene mayor relevancia en el perfil del tiempo de relajación  $\tau_r$  en función de  $T$ , dado que son valores independientes de la temperatura. De forma entonces que el comportamiento para el tiempo  $\tau_r$  en función de  $T$  mostrado en la Figura (11) es general; las variables  $T_s$  y  $\Delta\mu$  juegan en este caso un papel de escalamiento y corrimiento del perfil.

La dinámica para  $\tau_r$  mostrada en la Figura (11) indica la existencia de una divergencia esencial, esto es, la presencia de un valor *infinito* para el tiempo de relajación en términos estrictos para la temperatura  $T = T_s$  y de un pico resonante *finito* para  $\tau_r$  a la temperatura  $T = T_0$ . La divergencia que ocurre en  $T = T_s$  es una singularidad esencial, esto significa que es inherente a la estructura de la ecuación (4.23). Por el contrario, el pico resonante que exhibe la variable  $\tau_r$  en  $T = T_0$  es finito y proviene del hecho que a la temperatura  $T_0$  la macromolécula tiene su máxima estabilidad (Ver sección anterior). A esta temperatura, el cambio de entropía configuracional del sistema tiene un valor finito  $\Delta S^0$  (Ecuación 2.7). Para demostrar que el pico resonante en  $T = T_0$  es finito, basta sustituir  $T = T_0$  en (4.23), obteniéndose la expresión para  $\tau_r(T_0)$ ,

$$(4.19) \quad \tau_r(T_0) = \tau_r(T_s) \exp\left(-\frac{\Delta\mu\Delta s(T_0)}{R\Delta S^0} \left[\frac{1}{T_s} - \frac{1}{T_0}\right]\right),$$

que en términos del factor de corrimiento  $\log a_T$ , definido por

$$(4.20) \quad \log a_T = \log \left[ \frac{\tau_r(T)}{\tau_r(T_s)} \right],$$

puede escribirse como,

$$(4.21) \quad \log a_{T_0} = -\frac{\Delta\mu\Delta s(T_0)}{R\Delta S^0} \left[ \frac{1}{T_s} - \frac{1}{T_0} \right].$$

La ecuación (4.23) establece entonces la existencia de una singularidad esencial y un pico resonante para el tiempo de relajación estructural. La interpretación más directa de este fenómeno es que, a estas temperaturas los modos de relajación del sistema son extremadamente lentos. Para establecer una discusión en esta dirección, analizaremos la dinámica de la variable  $\tau_r$  en función de  $T$  en regiones de temperatura en la proximidad de la singularidad y el pico resonante. Es justo en esta región en donde podremos establecer una comparación entre lo que ocurre en la transición vítrea estándar y la desnaturalización de proteínas.

En el caso de los líquidos formadores de vidrios, el  $\Delta C_p$  es una constante respecto de la temperatura  $T$  [Angell, C.A. , 1995b], en tanto que para la desnaturalización de proteínas,  $\Delta C_p$  generalmente es una función cuadrática de  $T$  [Privalov and Dragan, 2007]. La suposición de que el  $\Delta C_p$  es una constante respecto de  $T$  permite escribir

a la fórmula de Adam-Gibbs como ,

$$(4.22) \quad \tau_r(T) = \tau_r(T_s) \exp \left( - \frac{\Delta\mu\Delta s}{R\Delta S^0 + R\Delta C_p^0 \log(T_0/T)} \left[ \frac{1}{T_s} - \frac{1}{T} \right] \right).$$

Al expresar la ecuación (4) como un desarrollo en serie de Taylor, es posible deducir una fórmula del tipo de William-Landel y Ferry [Adam and Gibbs, 1965], la cual como ya discutimos en su oportunidad, representa un ajuste teórico muy aceptable para los datos experimentales de viscosidad en líquidos inorgánicos.

En las primeras etapas de los estudios de calorimetría en la desnaturalización de proteínas, se consideró que el cambio en el calor específico  $\Delta C_p$  era una función lineal de la temperatura  $T$ . Desde el formalismo de Adam-Gibbs esto implica que el tiempo de relajación estructural tiene una forma funcional del tipo,

$$(4.23) \quad \tau_r(T) = \tau_r(T_s) \exp \left( - \frac{\Delta\mu\Delta s}{R\Delta S^0 + R \log(T_0/T)(\Delta C_p^0 - M_1 T_0) + R M_1 (T - T_0)} \left[ \frac{1}{T_s} - \frac{1}{T} \right] \right).$$

que indica un decaimiento más rápido de la variable  $\tau_r$  en función de  $T$  debido a la presencia del término  $R M_1 (T - T_0)$ . Para el caso de perfiles de desnaturalización cuadráticos, como es el caso de la mioglobina, la magnitud del coeficiente  $M_2$  es una medida del grado de concavidad de la función, de forma entonces que se debería esperar un decremento aún más abrupto que en el caso lineal conforme el cambio del calor específico  $\Delta C_p$  sea más pronunciada.

Más allá de estas consideraciones, los resultados presentados en este Capítulo son los siguientes. En primera instancia, hemos propuesto un procedimiento basado en elementos de geometría diferencial para calcular la energía de formación de una estructura helicoidal. Como resultado de esto, hemos podido escribir a la energía  $E_c$  del sistema en función del radio  $r$  de la hélice y del desplazamiento vertical  $p$ . Aunque en este momento no hemos explotado completamente las ventajas de este enfoque, considero que es un punto de partida para la extensión de este trabajo, no sólo en el caso de estructuras helicoidales sino también de conformaciones de tipo hoja  $\beta$ . Asumiendo que la transición hebra-hélice sólo puede ocurrir entre dos estados, se ha calculado la función de partición  $Z$  y la entropía  $\Delta s$  asociada a este cambio. Con esta aproximación los resultados muestran que, en comparación con el cambio de entropía  $\Delta S$  proveniente de todo el sistema (proteína, moléculas de agua del solvente, moléculas de agua estructural, etc), el cambio conformacional  $\Delta s$  es significativamente menor que su contraparte  $\Delta S$  y que por tanto puede ser considerado como una constante cuyo valor oscila entre [7–9] Joules/mol Kelvin. Es importante hacer notar que estos resultados son válidos solamente en el régimen considerado y a la luz de las aproximaciones hechas. Un cálculo más detallado de la función de partición  $Z$  podría establecer características distintas a las discutidas aquí.

En segunda instancia, tomando como base el esquema de Adam-Gibbs, hemos establecido una fórmula analítica para el tiempo de relajación estructural  $\tau_r$  no sólo en función de la temperatura  $T$ , como es el caso estándar en las descripciones estándar de la transición vítrea utilizando las ecuaciones de Vogel-Fulcher-Tamman y William-Landel y Ferry, sino también en términos de propiedades geométricas de la estructura helicoidal a través de las variables  $r$  y  $p$ , y de las variables  $M_1$  y  $M_2$ , las cuales pueden

medirse experimentalmente y que están relacionadas con los perfiles de desnaturalización de proteínas. Estos resultados en su conjunto indican entonces que el tipo de relajación esperado para el caso de la desnaturalización de proteínas no es, en general, de tipo Arrhenius. Es fundamental subrayar que esta conclusión ha de evaluarse con sumo cuidado, y siempre a la luz de las consideraciones, aproximaciones e hipótesis hechas a lo largo de este Capítulo.



## Perspectivas y Conclusiones.

### 1. Primer Eje de Investigación: Termodinámica de la desnaturalización.

**1.1. Filosofía del trabajo.** En la actualidad, el plegamiento y desnaturalización de proteínas es una de las líneas de investigación más activas y de mayor relevancia en las ciencias biomédicas. Tal como discutimos en los Capítulos I y II, el conocimiento de los mecanismos de interacción entre las cadenas polipeptídicas y las condiciones fisicoquímicas del medio en el cual se hallan inmersas no sólo resulta de fundamental importancia desde un punto de vista científico *per se*, sino que también constituye la base teórica necesaria para su implementación futura en el diseño molecular de fármacos y procedimientos terapéuticos para diversos padecimientos.

En este sentido, en las últimas décadas se han presentado, discutido y analizado diversos formalismos teóricos para caracterizar los procesos de plegamiento y desnaturalización de proteínas. En particular, los esquemas teóricos que han alcanzado mayor aceptación, se han centrado en la modelación molecular de las interacciones entre las cadenas polipeptídicas y el solvente estándar, que es agua. Estos enfoques proceden, en primera instancia, con la construcción de modelos para las diversas interacciones presentes entre los residuos de una proteína y las moléculas del solvente a fin de generar funciones de Hamilton  $H$  que abarquen la mayor cantidad de interacciones posibles: electrostáticas, de van der Waals, formación de puentes de hidrógeno, etc. En estas teorías se pretende establecer desde primeros principios, las características termodinámicas y cinéticas del plegamiento y desnaturalización con el mayor nivel de detalle posible. A diferencia de este enfoque, en este trabajo hemos presentado una discusión del plegamiento y desnaturalización desde una perspectiva mesoscópica. En este tipo de aproximación estamos interesados en caracterizar los rasgos generales bajo los cuales ocurre el plegamiento y/o desnaturalización tomando como punto de partida un modelo relativamente simple que permita un análisis directo de las variables y mecanismos involucrados en el proceso. Una de las argumentaciones que validan el uso de descripciones mesoscópicas, se halla en la observación experimental de los perfiles de desnaturalización de proteínas. Para muchos casos, particularmente aquellos referentes a proteínas de masa molecular pequeña y sin la presencia de enlaces disulfuro, el mecanismo de desnaturalización es un proceso reversible que sigue un perfil característico, independiente de la secuencia específica de aminoácidos y la geometría particular del estado nativo. Esta observación es crucial, dado que indica que aunque los detalles particulares de la interacción térmica o de los solutos urea, GdnHCl, etc con las cadenas laterales de cada proteína, pueda ser diferente en cada caso, los perfiles mostrados por las diversas técnicas experimentales sugieren mecanismos que operan a un nivel mesoscópico y que prescinden del detalle molecular. Así planteado, en este trabajo nos hemos centrado en la construcción de un formalismo mesoscópico que nos permita una disgresión sobre los rasgos generales del proceso de plegamiento y desnaturalización en términos de características observadas experimentalmente. La

generalidad del formalismo propuesto es, simultáneamente, su fortaleza y punto débil. Si bien la construcción del MEZ no permite hasta el momento, su aplicación y comparación directa con los resultados numéricos obtenidos en el campo experimental, si permite establecer características generales para el comportamiento de los potenciales termodinámicos, en particular para el cambio de energía libre  $\Delta_N^D F$ , el cual es una de las variables termodinámicas más analizadas en la literatura.

**1.2. Modelo Extendido de Zwanzig.** Para alcanzar este objetivo, hemos considerado como punto de partida el esquema teórico propuesto por Zwanzig [Zwanzig, 1995] y hemos establecido las bases sobre las cuales puede generalizarse para incluir información de tipo experimental respecto del desdoblamiento del estado nativo en función de las condiciones fisicoquímicas del medio. En este sentido, el MEZ considera a una proteína como un sistema capaz de adquirir múltiples estados energéticos metaestables que describen al estado desplegado de una proteína y un estado energético más favorable, asociado con la conformación nativa de la macromolécula, en acuerdo con la Teoría del Paisaje Energético [Bryngelson and Wolynes., 1989]. Asimismo, el MEZ toma en cuenta el efecto del solvente sobre la energía de las conformaciones no nativas y desnaturalizadas a través de la inclusión de información proveniente del campo experimental.

En este trabajo hemos centrado la atención en las implicaciones del MEZ para un tipo particular de interacción entre el solvente y la proteína: aquella que genera un perfil sigmoideal de desnaturalización en función de un agente químico. Sin embargo, el MEZ puede incluir y analizar otros perfiles de desnaturalización, como de hecho se presentó en el Capítulo III cuando se discutió el caso lineal de desdoblamiento de la estructura nativa. En este nivel, el MEZ no contempla la composición específica de aminoácidos en una cadena particular. Lo que sí considera es el número promedio  $\nu$  de grados de libertad de la cadena y diferencia dos estructuras primarias solamente en ese nivel. Asimismo, contempla también las diferencias existentes entre cadenas heteropoliméricas no biológicas y las proteínas a través de la diferencia energética  $U$  entre los estados no nativos. Dado que la transición entre estados conformacionales de polímeros no biológicos tiende a ser menos cooperativa que en el caso de las proteínas [Finkelstein and Ptysin, 2002], la variación de  $U$  permite establecer esta diferencia. Partiendo entonces de una descripción de la proteína en términos del número de aminoácidos  $K$  y el número promedio  $\nu$  de grados de libertad conformacionales de la cadena, de una descripción de los niveles de energía del sistema en términos de la separación energética  $U$  entre estados no nativos y la energía  $E^*$  del estado nativo y de una caracterización del perfil de desnaturalización en términos de la concentración media  $C_{1/2}$ , el MEZ puede establecer el perfil esperado para el cambio de energía libre  $\Delta F$  entre los estados desplegados  $D$  y el estado nativo  $N$ .

**1.3. Aplicación del MEZ en el caso lineal.** En el capítulo III analizamos el caso de desnaturalización lineal de una proteína obteniéndose que es posible en este caso calcular en forma exacta y con un mínimo de suposiciones, la forma de la función de partición  $Z$ . Aún cuando este caso sea poco realista en el campo de las proteínas, nos permite establecer una línea de investigación posterior sí conectada con la desnaturalización de polímeros biológicos: *el perfil sigmoideal observado en proteínas de masa molecular pequeña, puede aproximarse mediante una parametrización lineal en las regiones en donde ocurre la desnaturalización y dos asíntotas rectas en las zonas extremas del perfil*. Esto implica que podemos utilizar los resultados presentados para

el caso lineal, para calcular en forma exacta  $\Delta_n^D F$  en el caso sigmoidal. Ya en esa dirección, nos interesará discutir el papel de la temperatura  $T$  en la dinámica del calor específico  $C_p$ , característica que no pudimos realizar en el caso sigmoidal debido a las aproximaciones utilizadas.

**1.4. Aplicación del MEZ en el caso sigmoidal.** Para el caso de desnaturalización sigmoidal, una de las primeras hipótesis de trabajo se refiere al número promedio de conformaciones  $\nu + 1$  que puede adquirir cada cadena lateral. En general, cada cadena lateral  $R_i$  es capaz de adquirir un número distinto  $\nu_i$  de conformaciones, dependiendo no sólo de su composición química, sino también de su posición en la cadena. Más aún, las posibles rotaciones de cada cadena lateral dependen de la forma de los orbitales electrónicos de los vecinos inmediatos, pero también incluso podrían observarse correlaciones de mayor alcance, particularmente en estructuras helicoidales  $\alpha$ , en donde los aminoácidos  $i$  e  $i + 4$  se hallan en contacto cercano debido a la curvatura de la hélice. En estas condiciones pues, un tratamiento más realista de la variable  $\nu_i$  debería referirse a una distribución de valores de  $\nu$  en un intervalo dado. En este trabajo hemos considerado solamente el número *promedio* de conformaciones de todos los aminoácidos y en este nivel de descripción sólo podemos describir el comportamiento de cadenas polipeptídicas con números promedio  $\nu$  de conformaciones diferentes. Dos secuencias de aminoácidos con distintas distribuciones de valores  $\nu_j$  pero con el mismo número promedio  $\nu$  son indistinguibles en este nivel.

En referencia a la forma funcional del espectro de energías  $E$  de la macromolécula hemos considerado que la propuesta de Zwanzig contiene uno de los elementos clave de la teoría del paisaje energético, la existencia de un estado único de energía mínima  $E^* < 0$  y diversos estados metaestables con energía positiva  $U > 0$ . En principio, es posible proponer otras formas funcionales para  $E$  en función del parámetro de orden  $\alpha$ , sin embargo consideramos que los elementos clave requeridos para que el proceso de plegamiento sea un proceso dirigido están contenidos en la propuesta de Zwanzig y ésta corresponde a la forma funcional más simple. Es importante notar que en tanto en el esquema original de Zwanzig como en la generalización que hemos propuesto, el espectro de energías  $E$  no depende explícitamente de la temperatura  $T$ . La razón para no considerar la dependencia en  $T$  de  $E$  se halla en que decidimos enfocar la atención a los resultados que pueden derivarse cuando se considera que el proceso de desnaturalización está dirigido esencialmente por la presencia del soluto orgánico, y no de la temperatura. Al hacer esta consideración y la aproximación a primer orden en la función de partición  $Z$ , la dependencia en  $T$  en  $Z$  queda definida exclusivamente a través del factor de Boltzmann  $\beta = (K_B T)^{-1}$  y por tanto, al calcular el  $\Delta F$  se obtiene una expresión que depende exponencialmente en el factor de Boltzmann. Dado que el calor específico  $C_v$  se define en términos de una segunda derivada de la energía libre  $\Delta F$ , al hacer una aproximación de este tipo para  $F$ , estamos imposibilitados para extraer información en referencia a fluctuaciones térmicas de las variables  $E$  y  $S$ . Esta consideración nos lleva a plantear entonces como un proyecto a futuro, por un lado la inclusión de la dependencia en  $T$  de los niveles de energía y por el otro el desarrollo de métodos numéricos de integración más precisos los cuales nos permitan calcular la integral de la función de partición en forma exacta, tal como se hizo en el caso lineal (ver MEZ en el caso lineal). Esta vertiente nos permitirá discutir las propiedades del calor específico  $C_v$  en función de la temperatura  $T$  y confrontarlo con

los perfiles obtenidos experimentalmente.

El primer resultado extraído del MEZ para un perfil sigmoïdal de desnaturalización pone de manifiesto que la estabilidad del estado nativo  $N$ , es decir, su ubicación dentro de los estratos del Paisaje Energético depende significativamente del espacio conformacional de la macromolécula. Tal como discutimos en su oportunidad, el MEZ tiene una predicción teórica a este respecto. Dado que los términos  $F_{\nu,1}$  y  $F_{\nu,2}$  tienen una contribución negativa a  $\Delta F$ , éstos tienden a ubicar al estado nativo en regiones más profundas del paisaje energético conforme  $\nu$  se incrementa. En una primera aproximación de la variable  $\nu$  como constante para todos los aminoácidos, el MEZ muestra que aquellas proteínas con secuencias de aminoácidos con espacios conformacionales de dimensión mayor, tenderán a tener estructuras nativas con mayor estabilidad frente a fluctuaciones del medio. Hasta el conocimiento del autor, no se ha publicado en la literatura un resultado en esta dirección, por lo que en este sentido el MEZ provee de un formalismo para explicar los diferentes márgenes de estabilidad de una proteína en función de su estructura misma. En conjunción con la posibilidad de considerar en un trabajo posterior la distribución de valores  $\nu_i$ , el MEZ podrá dar resultados con un nivel de detalle mayor, indicando muy probablemente cuáles aminoácidos dentro de la secuencia primaria son los que contribuyen en mayor medida al efecto global mostrado en esta primera aproximación. Asimismo, los resultados obtenidos ya en este trabajo podrían ser comparados con los mapas de Ramachandran para cadenas homopoliméricas conformadas por aminoácidos *Gly* y *Pro*, los cuales representan los extremos de complejidad en sus cadenas laterales. Para el caso de cadenas polipeptídicas conformadas exclusivamente por aminoácidos *Gly* se espera que los mapas de Ramachandran muestren un área mayor de accesibilidad conformacional y que los perfiles de desnaturalización experimentales indiquen valores numéricos mayores para  $\Delta_N^D F$  en comparación con cadenas polipeptídicas conformadas por el aminoácido *Pro*, dado que éstos últimos tienen una cadena lateral *R* que presenta mayores restricciones estéricas que restringen las posibilidades conformacionales.

Finalmente, en referencia a la dependencia de  $\Delta F$  de la concentración media  $C_{1/2}$  podemos decir que esta es también una contribución significativa del MEZ al estudio de la desnaturalización de una proteína. Tal como se discutió previamente, la literatura experimental del campo indica que en la mayoría de las proteínas, el cambio  $\Delta_N^D F$  depende lineal o cuadráticamente en términos de  $C_{1/2}$ . Los resultados aquí derivados contienen ambos fenómenos y define también las condiciones bajo los cuales se presentan cada uno de ellos. Tomando como base un perfil de desnaturalización sigmoïdal para representar la respuesta de la proteína al agente desnaturalizante, el MEZ sugiere que la dependencia lineal de  $\Delta_N^D F$  en  $C_{1/2}$  debería observarse en proteínas con secuencias de aminoácidos con mayor libertad conformacional. En contraposición, propone también que los perfiles cuadráticos para  $\Delta_N^D F$  deberían ser los menos, dado que éstos corresponderían a proteínas con espacios conformacionales más restringidos. Es interesante notar que hasta el conocimiento del autor, esto justamente lo que ocurre. La mayor cantidad de resultados experimentales favorecen el comportamiento lineal.

## 2. Segundo eje de investigación: Cinética de la desnaturalización.

**2.1. Fundamentación para el uso de la Teoría Adam-Gibbs en las proteínas.** La teoría de Adam-Gibbs [Adam and Gibbs, 1965] es un formalismo basado en argumentos de mecánica estadística para describir los procesos de relajación que ocurren en la vecindad de la transición vítrea de sistemas líquidos. Es un resultado bien establecido que la transición del estado líquido al estado vítreo está caracterizada por un decremento en el calor específico  $C_v$  y en la entropía configuracional del sistema  $\Delta S$  [Angell, C., 1995], [García-Colin, L.S., 1999]. El esquema propuesto por Adam-Gibbs sugiere que el incremento sustancial en el tiempo de relajación estructural  $\tau_r$  y, consecuentemente en la viscosidad  $\eta$  del sistema, se halla asociado a procesos de carácter entrópico. Una de las razones del uso actual de la teoría de Adam-Gibbs para analizar los procesos de relajación en sistemas vítreos se debe a que, en el caso de  $\Delta C_p$  constante, el formalismo de Adam-Gibbs reproduce una ecuación del tipo William-Landel y Ferry, la cual reproduce los perfiles de relajación de diversos sistemas vítreos [Angell, C.A., 1995b].

Aún cuando la transición entre el estado nativo  $N$  de una proteína y su estado desplegado  $D$  es un fenómeno distinto, las observaciones experimentales discutidas en el Capítulo IV muestran que existen similitudes termodinámicas. El cambio de entropía configuracional del sistema proteína+solvente es negativo  $\Delta S$  debido a que el ordenamiento de la cadena polipeptídica reduce las posibilidades conformacionales de las moléculas de agua circundantes (efecto hidrofóbico) y el incremento del calor específico  $\Delta C_p$  en este caso es positivo, debido a la concavidad negativa de la energía libre  $\Delta F$ . Estas similitudes sugieren entonces que, si la teoría de Adam-Gibbs puede aplicarse en este caso, el tiempo de transición entre la estructura nativa  $N$  y los estados desplegados  $D$  es una variable determinada por el cambio de entropía configuracional  $\Delta s$  de las regiones cooperativas que integran al estado nativo, como del cambio de entropía global  $\Delta S$  del sistema en su conjunto. En el caso de la transición vítrea estándar, el cambio  $\Delta C_p$  es una constante respecto de la temperatura  $T$ , en contraposición al caso de las proteínas, en donde el cambio del calor específico  $\Delta C_p$  no sólo no es constante respecto de  $T$ , sino que exhibe perfiles característicos bien establecidos.

En este trabajo presentamos la aplicación del formalismo de Adam-Gibbs para el caso de una proteína cuyo estado nativo  $N$  está integrado básicamente por estructuras hélice  $\alpha$  a fin de analizar la dependencia del tiempo de relajación estructural  $\tau_r$  entre el estado  $N$  y el estado desplegado  $D$ . Para el caso de sistemas vítreos, la suposición de que el cambio de entropía  $\Delta s$  de las regiones cooperativas es constante respecto de  $T$ , permite reproducir los resultados experimentales a través de las ecuaciones de Vogel-Fulcher-Tamman y William-Landel-Ferry, sin embargo éste no necesariamente tendría que ser el caso en la desnaturalización de proteínas. Una de las contribuciones de este trabajo es la propuesta de un formalismo basado en geometría diferencial para calcular el cambio de entropía  $\Delta s$  entre una conformación helicoidal y una conformación lineal. Para analizar la transición hebra-hélice se utiliza generalmente la teoría de Zimm-Bragg, la cual es en esencia una teoría estocástica. Consideramos que las ideas presentadas aquí sobre cómo analizar el cambio conformacional hebra-hélice desde un punto de vista geométrico pueden aportar elementos nuevos a los esquemas

ya establecidos.

**2.2. Cambio de entropía global  $\Delta S$ .** La teoría de Adam-Gibbs se basa en su operación en dos elementos centrales: el cálculo de la entropía  $\Delta s$  de las regiones cooperativas y el cambio de entropía configuracional  $\Delta S$  global del sistema. Para determinar el variable  $\Delta S$  en la desnaturalización de la mioglobina, tomamos como punto de partida la información experimental de experimentos de microcalorimetría para el comportamiento de  $\Delta C_p$  en función de  $T$  [Privalov, 1982]. El  $\Delta C_p$  así medido contiene tanto las contribuciones vibracionales como las contribuciones rotacionales. Esto es,

$$(2.1) \quad \Delta C_p = \Delta C_p^v + \Delta C_p^r,$$

en donde  $\Delta C_p^v$  denota la contribución vibracional y  $\Delta C_p^r$  la contribución rotacional. Sin embargo, para el caso específico de la mioglobina, la contribución vibracional  $\Delta C_p^v$  es al menos tres órdenes de magnitud menor que la contribución rotacional [Achterhold *et al.*, 2002], de forma entonces que consideraremos que

$$(2.2) \quad \Delta C_p \simeq \Delta C_p^r.$$

La asociación de los datos experimentales reportados por [Privalov, 1982] con la contribución rotacional, nos permite calcular el cambio de entropía configuracional global  $\Delta S$  a través de las relaciones estándar de la termodinámica de equilibrio [García-Colin, L.S., 2005b], [Pathria, 1972]. El cálculo anterior es válido para el caso de proteínas en las cuales pueda omitirse la contribución vibracional al calor específico, como es el caso de la mioglobina. Cuando éste no sea el caso, la ecuación (2.1) indica que la contribución vibracional debe restarse al calor específico medido calorimétricamente. De hecho, la identificación de las contribuciones vibracionales y rotacionales al calor específico  $C_v$  desde un punto de vista teórico y experimental no es un problema menor [Avramov, 2005].

**2.3. Cálculo de  $\Delta s$  de las regiones cooperativas.** El análisis estándar realizado por Adam-Gibbs [Adam and Gibbs, 1965] para la transición del estado líquido al estado vítreo asume que el cambio de entropía configuracional  $\Delta s$  de las regiones cooperativas del sistema es una constante. Como mencionamos, éste no necesariamente tendría que ser el caso y en este trabajo presentamos un enfoque para calcular la posible contribución de  $\Delta s$  al tiempo de relajación. La idea central se basa en la representación de una conformación dada mediante una curva en el espacio a través de una función vectorial de variable real  $\mathbf{R}$  parametrizada por su longitud de arco  $\varsigma$ . Dado que una recta en el espacio tiene, por definición curvatura  $\kappa = 0$  y cualquier otra conformación de la curva tendrá una curvatura  $\kappa \neq 0$ , proponemos que la energía  $E_c$  de una conformación dada es proporcional al cuadrado de la curvatura<sup>1</sup>. Esta suposición implica que la conformación de menor energía es aquella con curvatura cero, la cual corresponde a una conformación lineal. Asimismo también implica que una conformación helicoidal tiene una curvatura  $\kappa$  constante, independiente de la longitud de arco  $\gamma$ . Partiendo entonces de la representación paramétrica de un hélice en el espacio, determinamos la curvatura  $\kappa$  correspondiente en términos del radio  $r$  de

<sup>1</sup>La dependencia de la energía en términos del cuadrado de la curvatura y no de la curvatura en sí, proviene del hecho de que la energía es una cantidad escalar y la curvatura es proporcional a la derivada del vector tangente. De forma entonces que  $\kappa^2$  sí es una cantidad escalar

la hélice y del desplazamiento vertical  $p$ .

El proceso anterior nos permite entonces determinar la energía de cualquier conformación helicoidal de radio  $r$  y desplazamiento vertical  $p$ . Es importante notar que ésta es sólo una primera aproximación a la transición cadena lineal-hélice real, dado que en ésta, la conformación desplegada  $D$  no necesariamente corresponde a una estructura lineal con curvatura cero. En general, el conjunto de conformaciones desplegadas exhiben un cierto grado de curvatura  $\kappa$  y torsión  $\zeta$  puesto que son conformaciones aleatorias. Esta observación sugiere una vertiente futura para este trabajo. A fin de representar con mayor precisión las conformaciones de los estados desplegados, puede considerarse un curva arbitraria en el espacio con curvatura  $\kappa(\gamma)$  y torsión  $\zeta(\gamma)$  e inducir a que el sistema adquiera una conformación preferente con curvatura  $\kappa_0$ , la cual podría asociarse con la curvatura de una hélice específica. De esta manera, la energía  $E$  de una conformación dada, podría escribirse como

$$(2.3) \quad E(\gamma) \sim [\kappa(\gamma) - \kappa_0]^2,$$

la cual tiene un mínimo, como puede verse de la Ecuación (2.3), para la curvatura preferente  $\kappa_0$ . Vale la pena mencionar que esta propuesta no deduce que la conformación de menor energía es aquella con curvatura  $\kappa_0$ , propone que así sea. Aún con esta consideración, el esquema para describir la transición cadena lineal-hélice estaría incompleto. En el Capítulo II discutimos que en el caso de la formación de hélices  $\alpha$ , la estabilidad de la estructura helicoidal en cadenas polipeptídicas ocurre debido a la formación de puentes de hidrógeno entre residuos no necesariamente adyacentes en la estructura primaria. Estos puentes de hidrógeno estabilizan y dan soporte estructural a la hélice  $\alpha$  debido a que la formación de cada uno de ellos disminuye la energía total del sistema en aproximadamente 2 kJ/mol. Esto implica entonces que la preferencia de las cadenas polipeptídicas a formar estructuras helicoidales no proviene de la geometría misma de la hélice, sino de la formación de los puentes de hidrógeno ubicados en las posiciones adecuadas.

Esta observación es crucial cuando se trata de generar una teoría que deduzca que el estado nativo de menor energía corresponde a una estructura helicoidal. Para tales propósitos, la energía  $E$  de la conformación deberá contener términos proporcionales al número de puentes de hidrógeno formados entre residuos. De forma tal que si  $h$  representa el número de puentes de hidrógeno y  $\varepsilon$  la energía de cada uno de ellos,  $E$  podría escribirse como

$$(2.4) \quad E = E_{helice} - h\varepsilon,$$

en donde  $E_{helice}$  representa la contribución energética proveniente de la estructura geométrica en sí. Asimismo, la discusión anterior implica que el planteamiento de que la energía de la conformación depende sólo de la geometría del sistema (ya sea a través de la curvatura  $\kappa$  y la torsión  $\zeta$ , o de otros parametros geométricos) no garantiza la existencia de un estado de mínima energía, se requieren elementos adicionales. Dada esta característica, procedimos entonces asumiendo que existe tal estado de menor energía y lo identificamos con la energía de una conformación helicoidal de radio  $r_0$  y  $p_0$ , en donde  $r_0$  y  $p_0$  corresponden a los valores estimados experimentalmente para un hélice  $\alpha$  [Cantor and Schimmel. , 1980]. Este estado lo denotamos como  $-E_0(r_0, p_0)$ . Asimismo, consideramos que la conformación desplegada puede representarse en una primera aproximación por una conformación lineal.

Así planteado, la energía  $E_0$  determinada mediante la integración de la curvatura  $\kappa$  de una hélice es entonces una primera aproximación a la energía de una estructura helicoidal biológica. Sin embargo, cualesquiera que sea la forma funcional de la energía  $E_0$  de la hélice  $\alpha$ , las estimaciones experimentales indican que  $E_0$  se halla en el intervalo  $[10 - 60]$  kJoules/mol, el cual es el margen de estabilidad del estado nativo en la mayoría de las proteínas no termofílicas. De forma entonces que en este trabajo, procedimos a calcular la función de partición  $Z$  y el cambio de entropía  $\Delta s$  asumiendo que, independientemente de la expresión analítica de  $E_0$  en términos de la geometría y de la inclusión del efecto estabilizador de los puentes de hidrógeno, su valor numérico se halla en el intervalo anteriormente descrito. A fin de analizar cuál es la contribución de  $\Delta s$  al tiempo de relajación procedimos suponiendo que la transición hebra-hélice ocurre solamente entre dos estados: la estructura lineal con energía  $E_1 = 0$  y la conformación helicoidal con energía  $E_0$ . Esta aproximación es realmente la parte más delicada puesto que ignora por completo cualquier otra contribución a la función de partición  $Z$  y al cambio de entropía  $\Delta s$  y conduce a la expresión,

$$(2.5) \quad \Delta s = - \left[ \frac{E_c(r_0, p_0)/T}{2 \exp(-E_c(r_0, p_0)/RT) + 1} \right] + R \log [2 + \exp(E_c(r_0, p_0)/RT)].$$

El resultado más notable que se desprende de este análisis es que, en el régimen de aproximaciones hechas y con valores para la energía de formación  $E_0$  en el intervalo antes discutido, la magnitud de la cantidad  $\Delta s$  es tres órdenes de magnitud menor respecto de la contribución  $\Delta S$ . Ahora bien, si los experimentos confirmasen que efectivamente éste es el caso para la desnaturalización de proteínas, podríamos pensar entonces la transición de los estados desplegados  $D$  al estado helicoidal es, efectivamente, una transición de dos estados. Por el contrario, si los experimentos indicasen que la contribución  $\Delta s$  sí juega un papel central en el tiempo de relajación, deberíamos proceder a una revisión de la aproximación de dos estados realizada, dado que ésta podría haber inducido una sobresimplificación del fenómeno.

Habiendo hechas las consideraciones anteriores, determinamos la forma funcional del tiempo de relajación  $\tau_r$  en función de la temperatura  $T$ . Tal como se espera de un análisis directo de la ecuación

$$(2.6) \quad \tau_r(T_0) = \tau_r(T_s) \exp \left( - \frac{\Delta \mu \Delta s(T)}{R \Delta S(T)} \left[ \frac{1}{T_s} - \frac{1}{T_0} \right] \right),$$

el perfil de  $\tau_r$  respecto de  $T$  muestra una divergencia esencial localizada en la temperatura  $T_s$ . Asimismo, muestra un pico resonante finito en la temperatura  $T = T_0$ , dado que a esta temperatura, el cambio de entropía configuracional  $\Delta S = \Delta S^0$ . La interpretación de que para  $T = T_0$  el tiempo de relajación se incrementa sustancialmente indica que la proteína se halla en un estado de máxima estabilidad y que por tanto, el sistema se halla ‘congelado’ en un estado de equilibrio termodinámico. Para discutir la dinámica del proceso de relajación hacia el estado desplegado  $D$  se requiere analizar la forma funcional de  $\tau_r$  en función de  $T$  en la vecindad del pico resonante  $T = T_0$ .

Para el caso de la transición vítrea, el  $\Delta C_p$  es constante y el cambio de entropía  $\Delta S$  depende logarítmicamente de la temperatura. Al incluir estas características en la ecuación Adam-Gibbs, es posible reproducir las curvas de experimentales de viscosidad  $\eta$  en función de  $T$ . Para el caso de las proteínas,  $\Delta C_p$  no es constante y esto implica una forma funcional para  $\Delta S$  del tipo,



(2.7)

$$\Delta S^{config}(T) = \Delta S^0 + \Delta C_p^0 \log(T/T_0) + M_1(T - T_0) - DT_0 \log(T/T_0) + \frac{M_2}{2}(T^2 - T_0^2) - 2M_2T_0(T - T_0) + M_2T_0^2 \log(T/T_0),$$

la cual contiene los casos particulares  $\Delta C_p$  constante y  $\Delta C_p$  lineal eligiendo los parámetros  $M_1$  y  $M_2$  como cero en forma adecuada. Cada proceso de desnaturalización particular está representado por un conjunto de valores específicos para las cantidades  $T_0$ ,  $\Delta S^0$ ,  $M_1$  y  $M_2$ . En el caso de la mioglobina, por ejemplo,  $\Delta S^0 = -116$  J/mol kelvin [Privalov, 1982] y  $M_1$  y  $M_2$  tienen los valores obtenidos previamente por el ajuste de mínimos cuadrados. Para este caso, en el intervalo [290 – 310] Kelvin, la entropía configuracional  $\Delta S$  oscila entre [–500, 500] Joules/mol Kelvin, en tanto que  $\Delta s \sim 9$  Joules/mol Kelvin. Esto justifica, tal como mencionamos anteriormente que  $\Delta S$  es la contribución principal al tiempo de relajación. Sin embargo, para otras proteínas, la contribución de los términos lineales, logarítmicos y cuadráticos de  $\Delta S$  podría ser distinta. En particular, para perfiles del  $\Delta C_p$  en donde el coeficiente del término cuadrático  $M_2$  domine sobre las contribuciones lineales (coeficiente  $M_1$ ), el tiempo de transición entre las estructuras desplegadas y nativa debería ser mucho más abrupto que en la transición vítrea. Observemos que las cantidades  $\Delta S^0$ ,  $M_1$  y  $M_2$  son medibles experimentalmente, de forma entonces que en principio, es posible establecer una comparación entre los resultados presentados aquí e información experimental para el tiempo en el cual ocurre la transición  $N \rightleftharpoons D$ . Para estar en posibilidades de establecer esta comparación, una línea posterior de trabajo en esta dirección necesariamente implica un análisis del papel de la barrera energética  $\Delta\mu$  necesaria para la transición de las regiones cooperativas suceda. Ya sea que se proceda mediante un análisis teórico o mediante un ajuste de datos experimentales, la discusión del papel de esta variable podría arrojar nueva información sobre el perfil del tiempo de relajación  $\tau_r$ .

### 3. Impacto del trabajo.

Las ideas presentadas en este trabajo han sido discutidas por el autor en diferentes foros nacionales e internacionales bajo las modalidades de cartel y exposición oral. A continuación se enlistan algunas de las participaciones

- ◇ **Protein Folding and Misfolding.** European Molecular Biology Association. Universitá La Sapienza, Roma, Italia, 2003.
- ◇ **Fifth International Conference on Biological Physics.** Ghotenburg, Suecia. Agosto 2004.
- ◇ **49th Annual Meeting of the Biophysical Society USA.** Long Beach, California. Poster: *The free energy profile of a protein in a chemically induced denaturation mechanism.* L. Olivares-Quiroz and L.S. García-Colín. Febrero 2005.
- ◇ **Congreso Nacional de Física 2005.** Universidad de Guadalajara. Guadalajara, Jalisco, 2005. Poster: *Relación Lineal para el cambio de energía libre  $\Delta F$  en un proceso de desnaturalización sigmoidal.* L. Olivares-Quiroz

and L.S. García-Colín.

- ◇ **Congreso Nacional de Bioquímica 2006.** Universidad de Guanajuato. Guanajuato, Gto. Sesión simultánea: *Tiempo de plegamiento de una proteína desde el formalismo de Adam y Gibbs.* L. Olivares-Quiroz and L.S. García-Colín.
- ◇ **Seminario del Area de Líquidos 2007.** Departamento de Física. Area de Líquidos. Universidad Autónoma Metropolitana-Iztapalapa. *Desnaturalización química de un proteína.* L. Olivares-Quiroz. Enero 2007.
- ◇ **Seminario de Estudiantes del Departamento de Física 2007.** Departamento de Física. Universidad Autónoma Metropolitana-Iztapalapa. *La desnaturalización de proteínas y la Transición Vítrea.* L. Olivares-Quiroz. Marzo 2007

En referencia al material escrito, hasta el momento se han publicado dos trabajos los cuales se citan en las referencias de este reporte y a continuación se reproducen,

- ◇ Olivares-Quiroz, L. and García-Colín L.S. *Rev Soc. Quím. Mex.* **48**, 95-105, 2004.
- ◇ Olivares-Quiroz, L. and García-Colín L.S. *J. Theor. Biology.* **246**, 214-224, 2007.

Deseo expresar mi agradecimiento a todos aquellos que han contribuido al desarrollo y conclusión de este trabajo. Particularmente a aquellos que con sus ideas, comentarios y sugerencias contribuyeron a consolidar las ideas aquí presentadas y me permitieron también exponerlas en distintos foros nacionales e internacionales. Sin ningún orden en especial y a todos ellos con mi más sincera estimación y afecto,

- Dr. Leopoldo García-Colín Scherer,
- Dr. Hugo Morales Técotl,
- Dr. Michel Picquart,
- Dr. Rafael Barrio Paredes
- Dr. Jorge Velasco Hernández
- Dr. Orlando Guzmán Lopez,
- Dr. Emmanuel Haro Poniatowski
- Dr. Roberto Olayo González
- Dr. Edgar Vazquez Contreras
- Dr. Eusebio Juaristi.

## Bibliografía

- Fire, A., Xu, S., Kostas MK, Driver SE and Mello. C. Potent and specific genetic interference by double stranded RNA in *C. Elgans*. *Nature*.**391**. 806-811. (1998).
- Guo, W., Shea, J., and Berry, S. The physics of the interactions governing folding and association of proteins. *Ann. N.Y. Acad. Sci.***1066**, 34-53, (2005).
- Olivares-Quiroz, L. and García-Colín LS. *Rev Soc Quim.***48**. 95-105. (2004).
- Flory, P.J. *Principles of Polymer Chemistry*. Cornell University, Ithaca, (1953).
- Flory, P.J. *Statistical mechanics of chain molecules*. Interscience, New York, (1969).
- de Gennes, P.G. *J. Phys (Paris)* **37**, 1445 (1976).
- Anfinsen, C.B., Haber, E., Sela, M., and White F.W. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. USA* **47**, 1309-1314, (1961).
- Klink, T.A., Woycechowski, K.M., and Raynes, R.T. Contributions to disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *Eur.J. Biochem.* **267**, 566-572, (2000).
- Matthews, B.W. Studies on protein stability with T4 lysozyme *Adv. Protein Chem.* **46**, 249-278, (1995).
- García-Colín L.S. *Introducción a la Física Estadística*. Edit. El Colegio Nacional. Primera Edición 2005.
- Gibbs, J.W. *Elementary Principles in Statistical Mechanics*. Dover Publications NY. 1960
- Bryngelson, J.D. and P. G. Wolynes. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem.*, **93**, 6902-6915, (1989).
- Pande, V.S, Grosberg, A.Yu., and Tanaka, T. Heteropolymer freezing and design: Towards physical models of protein folding. *Rev. Mod. Phys.*, **72**, 259, (2000).
- Grosberg, A. Yu and D.V. Kuznetsov. *Macromolecules.*, **25**,1970, (1992).
- Dill, K.A., and Chan, H.S. From Levinthal to pathways to funnels. *Nat. Struct.Biol.*, **4**, 1, (1997).
- Vendruscolo, M.,Paci E., Dobson CM., and Karplus M. The protein folding 'speed limit'.*Curr. Op. Struct. Biol.*, **14**, 76-88, (2001).
- Levinthal, C. Are there pathways to protein folding?. *J. Chim. Phys.*. **65**, 44-45, (1968).
- Kauzmann, W. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.*,**14**, 1-63, (1959).
- Pfeil, W. and Privalov, P.L. Thermodynamic investigations of proteins I: Standard functions for proteins with lysozyme as an example. *Biophys.Chem.*,**4**, 23-32, (1976).
- Pfeil, W. and Privalov, P.L. Thermodynamic investigations of proteins II: Calorimetric study of lysozyme denaturation by guanidine hydrochloride. *Biophys.Chem.*,**4**, 33-40, (1976).

- Privalov, P.L. Stability of proteins: small globular proteins. *Adv. Protein Chem.* **33**, 167-241, (1979).
- Privalov, P.L. and Gill, S.J. Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.*, **39**, 191-234, (1988).
- Kelly L., and Holladay L.A. A comparative study of the unfolding thermodynamics of vertebrate metmyoglobins. *Biochemistry* **29**, 5062-9, (1990).
- Kelly L., and Holladay L.A. Conformational free energy of armadillo metmyoglobin. *Int. J. Pept. Protein Res.* **35**, 235-240, (1990).
- Privalov, P.L., Tiktopulo, E.I., Venyaminou S., Griko Y., Makhatadze G.I., and Keshinashivishii N.N. Heat capacity and conformation of proteins in the denatured state. *J. Mol. Biol.* **205** (4), 737-750, (1989).
- Michnik, A., Drzazga Z., Kluczevska, A., and Michalik K. Differential scanning microcalorimetry study of the thermal denaturation of haemoglobin. *Biophys. Chem.* **118**, 93-101, (2005).
- Chateau, E., Garden, J.L., Bourgeois, O., and Chaussy, J. Physical Kinetics and thermodynamics of phase transitions probed by dynamical nanocalorimetry. *Applied Phys. Lett.* **86**, 151913, (2005).
- Privalov, P.L. Reflections on the origins of microcalorimetry of biopolymers. *Biophysical Chem.* **126**, 13-15, (2007).
- Wetlaufer D.B., and Ristow S. Acquisition of three dimensional structure of proteins. *Annu Rev Biochem.* **42**, 135-158, (1973).
- Go, N. Theoretical studies of protein folding. *Annu Rev Biophys. Bioeng.* **12**, 183-210, (1983).
- Baldwin, R.L. Intermediates in protein folding reactions and the mechanism of protein folding. *Annu. Rev. Biochem.* **44**, 453-475, (1975).
- Dubey, V.K., Pande, M., and Jaganadham, M.V. Snapshots of protein folding problem: Implications of folding and misfolding studies. *Protein Pept. Lett.* **13**, 883-888, (2006). bibitem[Tanford.(1968)]tanford Tanford C. Protein denaturation. *Adv. Protein Chem.* **23**, 121-282, (1968).
- Tanford C. Macromolecules *Protein Sci.* **3** (5), 857-861, (1994).
- Angell C.A. Formation of glasses from liquids and biopolymers. *Science* **267**, 1924-1935, (1995).
- Angell C.A. The old problems of glass and the glass transition, and the many new twists. *Proc. Natl. Acad. Sci. USA*, **92**, 6675-6682, (1995).
- Swallen, S.F., Kearns, K.L., Mapes, M.K., Kim, Y.S. Organic glasses with exceptional thermodynamics and kinetic stability. *Science*, **315**, 353, (2007).
- Yavari, A.R. Metallic glasses: the changing faces of disorder. *Nature Materials*, **6**, 181-182, (2007).
- LS García-Colín *Rev Mex Fis* **45**, 11-17, (1999).
- Stillinger F.H. A topographic view of supercooled liquids and glass formation. *Science* **267**, 1935, (1995).
- Dobson, C.M. Trends Biochem. Sci. **24**, 329-332. (1999).
- Dobson, C.M. The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc. B*, **356**, 133-145, (2001)
- M.E. Chanez-Cárdenas, D. Cisneros-Armas, L. del Pozo-Yauner, B. Becerril., and A.Fernandez-Velasco. Mensaje Bioquímico Vol.XXVI. UNAM, (2002).
- Hardesty, B., and Kramer G. Folding of a nascent peptide on the ribosome. *Prog. Nucl. Acid. Res. Mol. Biol.* **66**, 41-66. (2001).

- Deuerling, E., and Bukau, B. Chaperone-assisted folding of newly synthesized proteins in the cytosol. *Crit. Rev. Biochem. Molec. Biol.* **39**, 261-277. (2004).
- Young, J.C., Agashe, V.R., Siegers K., and Hartl, F.U. Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* **5**, 781-791. (2004).
- Ellis, R.J. Molecular chaperones: inside and outside the Anfinsen cage. *Curr. Biol.* **11**, 114-119. (2001).
- Rucknagel, D.L. In *Sickle-Cell Anemia and Other Hemoglobinopathies*. R.D. Levere. Editor. Academic Press NY, (1975).
- Louderback J.G., Ballas, S.K. and Kim-Shapiro D.B. Sickle hemoglobin polymer melting in high concentration phosphate buffer. *Biophys. J.*, **76**, 2216-2222, (1999).
- Wang, W.C. The pathophysiology, prevention and treatment of stroke in sickle cell anemia. *Curr. Opin. Hematol.* **14**, 191-197, (2007).
- Ingram V.M. A specific chemical difference between the globins of normal human and sickle cell anaemia haemoglobin. *Nature*, **178**, 792-794, (1956).
- Ingram V.M. Sickle-cell anemia hemoglobin: the molecular biology of the first molecular disease. *Genetics*, **167** (1), 1-7, (2004).
- Lesk A. M. *Introduction to protein science*. Oxford University Press, 2004.
- Lansbury, P.T. Jr. Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci. USA*, **96**, 3342-4, (1999).
- Walsh, D.M., Hartley D.M., Kusumoto, Y., Fezoul Y., Condron M.M., Lomakin, A., Benedek G.B., Selkoe D.J., and Teplow D.B. Amyloid  $\beta$ -protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *J. Biol. Chem.*, **274**, 25945-52, (1997).
- Bradbury, J. Chaperones: keeping a close eye on protein folding. *Lancet*, **361**, 1194-1195, (2003).
- Tapan, K.C., and Subhankar, P. Protein folding diseases and chaperone-based therapeutic approaches. *FEBS Journal*, **273**, 1331-1349, (2006).
- Games, D., Adams, D., Alessandrini, R. Alzheimer-type neuropathology in transgenic mice overexpressing V717F  $\beta$ -amyloid precursor protein. *Nature*, **373**, 523-527, (1995).
- Forloni, G., Terreni, L., Bertani, I. Protein misfolding in Alzheimer's and Parkinson's disease: genetics and molecular mechanisms. *Neurobiol. Aging*, **23**, 957-976, (2002).
- Agorogiannis E.I., Agorogiannis G.I., Papadimitrou, A., and Hadjigeorgou G.M. Protein misfolding in neurodegenerative diseases. *Neuropathology and Applied Neurobiology*, **30**, 215-224, (2004).
- Cohen F.E., Pan, K.M., Huang, Z., Baldwin, M., Fletterick R.J., and Prusiner S.B. Structural clues to prion replication. *Science*, **264**, 530-531, (1994).
- Saenger, W. Principles of nucleic acid structure. Springer (Berlin), (1988).
- Gaeta, G. Results and limitations of the soliton theory of DNA transcription. *J. Biol. Phys.*, **24**, 81-96, (1999).
- Watson, J.D., Hopkins N.H., Roberts, J.W., Steitz, J.A., and Weiner A.M. Molecular biology of the gene. *Benjamin Publishing Co.* (1987).
- Voet, D., and Voet, J. Biochemistry. Second Edition. John Wiley and Sons. (1995).
- Cantor C.R., and Schimmel P.R. Biophysical Chemistry. Part I: The conformation of biological macromolecules. W.H. Freeman and Co. (1980).
- Crick FHC, Barnett, L. Brenner S. and Watts-Tobin R.J. General nature of the genetic code for proteins. *Nature*, **192**, 1227-1232, (1961).

- Nirenberg M.W. and Matthaei J.H. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci.*, **47**, 1588-1602, (1961).
- Watson J. D., *DNA. The secret of life*. Alfred Knopp Publications. 2003
- Barret G.C. *Chemistry and Biochemistry of Aminoacids*. Chapman and Hall (1985).
- Vickery H.B. *Adv Protein Chem.* **26**, 81-171 (1972).
- Finkelstein A.V., and Ptitsyn O. *Protein Physics*. Academic Press, (2002).
- Branden C., and Tooze J. *Introduction to Protein Structure*. Garland Publishing Inc. Second Edition, (1991).
- Pauling L. Corey R.B. Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets. *Proc. Natl. Acad. Sci. USA*, **37**. 729-740, (1951).
- Pauling L., Corey R.B., and Branson, H.R. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chains. *Proc. Natl. Acad. Sci. USA*, **37**. 205-211, (1951).
- Ma, B., Tsai, C., Nussinov, R. A systematic study of the vibrational free energies of polypeptides in folded and random states. *Biophys. J.*, **79**. 2739-2753, (2000)
- Adam G and Gibbs J.H. *J Chem Phys.* **43**. 139, 1965.
- Marsh R.E. and Donohue J. Crystal structure studies of aminoacids and peptides. *Adv. Protein Chem.*, **22**. 235-256, 1967.
- Ramachandran G.N., Sasisekharan V. Conformation of polypeptides and proteins. *Adv. Protein Chem.*, **23**, 283-438, 1968.
- Flory P.J. *Statistical mechanics of chain molecules*. Pag 263. InterScience 1969.
- Watson, J.D., and Crick, F.C.H. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738, (1953)
- Kendrew, J.C. A three dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*. **181**. 662-666, 1958.
- Kendrew, J.C., Watson, H.C., Strandberg, B.E., Dickerson, R.E., Phillips, D.C., and Shore, V.C. The aminoacid sequence x-ray methods and its correlation with chemical data. *Nature*. **190**. 666-670, (1961).
- Brant D.A., and Schimmel, P.R. Analysis of the skeletal configuration of crystalline hen egg-white lysozyme. *Proc. Natl. Acad. Sci. USA*. **58**, 428-435, (1967).
- Arnott S., and Dover S.D. Refinement of bond angles of an  $\alpha$  helix. *J. Mol. Biol.*, **30**. 209-212, (1967).
- Toniolo C., and Benedetti E. The polypeptide  $3_{10}$  helix. *Trends Biochem. Sci.*, **16**. 350-353, (1991).
- Lancelot, N., Elbayed, K., Raya, J., Piotta, M., Briand, J.P., Formaggio, F., Toniolo, C., Bianco, A. Characterization of the  $3_{10}$  helix in model peptides by HRMAS NMR spectroscopy. *Chemistry*, **9**, 1317-1323, (2003).
- Schultz, G.E., and Schirmer R.H. *Principles of Protein Structure*. Springer Verlag (1979).
- Berman H.M, Westbrook J., Feng Z., Gilliland G. The Protein Data Bank. *Nucleic Acids Research*. **28**. 235-242, (2000).
- Westbrook, J., Feng, Z., Chen, I., Yang, H., and Berman, H.M. The Protein Data Bank and structural genomics. *Nucleic Acids Res.*, **1**. 489-491, (2003).
- Rose G.D., Gierasch, L.M., and Smith J.A. Turns in proteins and peptides. *Adv. Protein Chem.*, **37**. 1-109, (1985).
- Berisio R., Loguercio S., DeSimone A. and Zagari A. PolyProline helices in protein structure: an statistical survey. *Protein Pept. Lett.*, (**13(8)**), 847-854, (2006).

- Chung-Phillips, A. Polyglycine conformational analysis: calculated vs experimental gas-phase basicities and proton affinities. *J. Phys. Chem-A*, **109(26)**, 5917-5932, (2005)
- Ohnishi S., Kamikubo, H., Onitsuka, M. and Kataoka M. Conformational preference of polyglycine in solution to elongated structure. *J. Am. Chem. Soc.*, **128(50)**. 16338-16344, (2006).
- Engel, J. and Prockop D.J. The zipper-like folding of collagen triple helices and the effect of mutations that disrupt the zipper. *Annu. Rev. Biophys. Chem.*, **20**. 137-152, (1991).
- Hopfinger, A.J. *Conformational properties of macromolecules*. Academic Press N. Y. (1973).
- Teale F.J.W and Weber J. *Biochem. J.* **65**. 476-482, 1957.
- Egan W., Shindo H., and Cohen J.S. *Annu Rev Biophys Bioeng.* **6**. 383-417, 1977.
- Crippen G.M. Topology of globular proteins. *J. Theor. Biol.*, **45**. 327-338, (1974).
- Moews P.C., and Kretsinger R.H. Terbium replacement of calcium in carp-muscle calcium-binding parvalbumin by model building and difference Fourier analysis. *J. Mol. Biol.*, **91**. 229-232, (1975).
- Herzberg O., and James, M.N.G. Structure of the calcium regulatory muscle protein troponin-C and other  $Ca^{+2}$  binding proteins. *Nature*, **313**, 653-659, (1985).
- Oparin A.I. *Life: Its Origin, Nature and Development*. New York Academic Press (1964).
- Woese C.R. A proposal concerning the origin of life on the planet Earth. *J. Mol. Evol.*, **13**, 95-101, (1979).
- Woese C.R. A new biology for a new century. *Microbiol. Mol. Biol. Rev.*, **68**, 173-86, (2004).
- Weber A.L., and Miller S.L. Reasons for the occurrence for the twenty coded protein aminoacids. *J. Mol. Evol.*, **17**, 273-284, (1981).
- Cloud P. How life began. *Nature*, **296**, 198-199, (1982).
- Kappler M.H. The independent distribution of aminoacids near neighbor pairs into polypeptides. *Biochem. Biophys. Res. Commun.*, **78**, 1018-1024, (1977).
- Creighton, Thomas E. *Proteins: Structure and Molecular Properties*. W.H. Freeman and Co. New York. 1984.
- Dayhoff M.O. *Atlas of Protein Sequences and Structure*. Washington DC. National Biomedical Foundation. 1972.
- Kimura M. Evolutionary rate at the molecular level. *Nature*, **217**, 624-626, (1968).
- Vindigni A., Di Cera A. Release of fibrinopeptides by the slow and fast forms of thrombin. *Biochemistry*, **35(14)**, 4417-4426, (1996).
- Pechik I., Yakovlev S., Mosesson M.W, Gilleland G.L., and Medved L. Structural basis for sequential cleavage of fibrinopeptides upon fibrin assembly. *Biochemistry*, **45(11)**, 3588-3597, (2006).
- Gillespie J.H. and Langley C.H. Protein evolution and the molecular clock. *J. Mol. Evol.*, **13**, 27-34, (1979).
- Rydberg, P., Rod, T.H., Olsen, L., and Ryde, U. Dynamics of water molecules in the active-site cavity of human cytochromes P450. *J. Phys. Chem. B*. **111**, 5445-5457, (2007).
- Ghelis Ch. and Yon J. *Protein Folding*. Academic Press NY. 1982.
- Privalov P.L. and Dragan, A.I. Microcalorimetry of biological macromolecules. *Biophys. Chem.*, **126**, 16-24, (2007).

- Privalov P.L., and Pothekin, S.A. Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.*, **131**, 4-51, (1986).
- Dill, K.A., and Bromberg, S. Molecular Driving Forces. Statistical Thermodynamics in chemistry and biology. Garland Science. (2003).
- Freire, E., and Biltonen, R.L. Thermodynamics of transfer ribonucleic acids: the effect of sodium on the thermal unfolding of yeast t-RNA Phe. *Biopolymers*. **17**, 1257-1272, (1978).
- Privalov, P.L. Stability of proteins: proteins which do not represent a single cooperative system. *Adv. Protein Chem.* **35**, 1-104, (1982).
- Tsong T.Y., Hearn R.P. Watrall, D.P. and Sturtevant J.M. A calorimetric study of thermally induced conformational transitions of ribonuclease-A and certain of its derivatives. *Biochemistry*. **2**, 2666-2667, (1970).
- Privalov P.L., and Khechinashvili, N.N. A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J. Mol. Biol.* **86**, 665-684, (1974).
- Creighton T. E. *Protein Folding*. W.H. Freeman and Co. N.Y. (1999)
- Olivares-Quiroz L. and García-Colín L.S. El plegamiento de las proteínas: un problema interdisciplinario. *J. Mex. Chem. Soc.* **48**, 95-105, (2004).
- Tanford C. Protein denaturation. *Adv. Protein Chem.* **23**, 121-282, (1968).
- Hibbard L.S. and Tulinsky A. Expression of functionality of  $\alpha$ -chymotrypsin. Effects of guanidine hydrochloride and urea in the onset of denaturation. *Biochemistry*. **17**, 5460-5468, (1978).
- Roseman D., and Jenks W.P. *J. Am. Chem. Soc.* **97**, 631-640, 1975.
- Paliwal, A., Asthagiri, D., Bossev, D.P., and Paulatis, M.E. Pressure denaturation of staphylococcal nuclease studied by neutron small-angle scattering and molecular simulation. *Biophys. J.* **87**, 3479-3492, (2004).
- Cantor, C.R., and Schimmel, P.R. Biophysical Chemistry II. Techniques for the study of biological function and structure. W. H. Freeman and Co. (1980).
- Brandts J.F. The nature of the complexities in the ribonuclease conformational transition and the implications regarding calthrating. *J. Am. Chem. Soc.* **87**, 2759-2760, (1965).
- Herskovitts J. Jr. *Methods Biochem. Ann.* **13**, 81-111, 1965.
- Hung, Hui-Chi and Chang Gu-Han. Multiple unfolding intermediates of human placental alkaline phosphatase in equilibrium urea denaturation. *Biophys. J.* **81**, 3456-3471, (2001).
- Jirgersons, B. *Optical activity of proteins and other macromolecules*. Springer-Verlag. Berlin and New York, 1973.
- Adler, A., Greenfield N.J., and Fasman, G.D. En *Enzyme Structure. Methods in Enzimology* Vol 27, 675-735. Academic Press New York, 1973.
- Cheng Y.H., Yang J.T. and Martinez H.M. *Biochemistry*. **11**, 4120-4131, 1972.
- Sundd, M., Kundu, M., and Jagganadham M. Acid and chemical induced conformational changes in ervatamin B: sequential unfolding of domains. *J. Biochem. Mol. Biol.* **35**(2), 143-154, (2002).
- Flory, P.J. *Statistical mechanics of chain molecules*. Interscience, New York, (1969).
- Volkenstein, M. V. *Configurational Statistics of Polymer Chains*. Moscow. InterScience, New York, (1963)
- Lifshitz, I. M., A. Y., Grosberg and A.R. Khokhlov. Some problems of the statistical physics of polymer chains with volume interactions. *Rev. Mod. Phys.*, **50**, 683, (1978).



- Shakhnovich, E.I. and A. Gutin. Formation of unique structure in polypeptide chains: theoretical investigations with the aid of a replica approach. *Biophys. Chem.*, **34**, 187-99, (1989).
- Derrida, B. The random energy model, an exactly solvable model of disordered systems. *Phys. Rev. B* **24**, 2613-2626, (1981).
- Takada, S., and Wolynes, P. An elementary mode coupling theory of random heteropolymer dynamics. *Proc. Natl. Acad. Sci. USA*. **94**, 2318-21, (1997).
- Grosberg, A.Y. and A.R. Khokhlov. *Statistical physics of macromolecules*. AIP Press, New York, (1994).
- Feynman, R.P. *Quantum Mechanics and Path Integrals*. McGraw Hill, New York, (1965)
- Feynman, R.P. *Statistical mechanics: a set of lectures*. Addison- Wesley, (1972).
- Zwanzig, R. *Simple model of protein folding kinetics*. *Proc. Nat. Acad. Sci. USA* **92**, 9801, (1995).
- Olivares-Quiroz L. and García-Colín LS. *Calculo de propiedades termodinámicas de una macromolécula: Generalización del Modelo de Zwanzig*. Jornadas del Posgrado Divisional. UAM-I. El Colegio Nacional (2004).
- Salton M.R.J., and Ghuysen M.J. Properties of lysozyme and its action on microorganisms. *Biochim. Biophys. Acta*. **45**, 355, 363, (1960).
- Kumeta H., Miura, A., Kobashigawa, Y. Low-temperature-induced structural changes in human lysozyme elucidated by three dimensional NMR spectroscopy. *Biochemistry*. **42**, 1209-1216, (2003).
- Johnson R., Christodoulou J., Domoulin, M. Rationalising lysozyme amyloidosis: insights from the structure and solution dynamics of T70N lysozyme. *J. Molecular Biol.* **352**, 823-836, (2005).
- Pepys M.B., Hawkins P.N., Booth D.R., Vigushin D.N. Tenneth G.A., and Soutar A.K. *Nature* **362**, 553-557, (2003).
- Yazaki, M., Farrell, S.A. and Benson, M.D. A novel lysozyme mutation Phe57Ile associated with hereditary renal amyloidosis. *Kidney Int.* **63**, 1652-1657, (2003).
- Booth D.R., Pepys, M.B. and Hawkins, P.N. A novel variant of human lysozyme (T70N) is common in the normal population. *Human Mutat.* **16**, 180-181, (2000).
- Canet, D., Sunde M., Last A., Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behaviour in its disease-related variants. *Biochemistry* **38**, 6419-6427, (1999).
- Esposito G., García, J., Mangione, P., Giogetti, S., Corazza, A. Structural and folding dynamic properties of the T70N variant of human lysozyme. *J. Biol. Chem.* **278**, 25910-25918, (2003).
- Butkov, E. *Mathematical Physics*. Addison-Wesley, 1968.
- Arfken, G. *Mathematical Methods for Physicists*. Academic Press Inc. 1985.
- García-Colín, LS. *Introducción a la Termodinámica Estadística*. El Colegio Nacional, 2005.
- Olivares-Quiroz L., and García-Colín L.S. Protein's native state stability in a chemically induced denaturation mechanism. *Journal of Theoretical Biology*. **246(2)**, 214-224, (2007).
- Asgeirsson, B., Guojonsdottir, K. Reversible inactivation of alkaline phosphatase from Atlantic cod in urea. *Biochim. Biophys. Acta*, **1764**, 190-98 (2006).
- Sharma, S. *et al. Protein Pept Lett*, **13** (4), 323-329 (2006).
- Gupta, R., and Ahmad, F. Protein stability: functional dependence of denaturational Gibbs energy on urea concentration. *Biochemistry*, **38**, 2471-2479, (1999).

- Wong H.J., Sthapoulos, P.B., Bonner, J.M. Non-linear effects of temperature and urea on the thermodynamics and kinetics of folding and unfolding of hisactophilin. *J. Mol. Biol.* **344**, 1089-1107, (2004).
- Epstein, H.F., Schechter, A.N., Chen, R.F., and Anfinsen C.B. Folding of the staphylococcal nuclease: kinetic studies of two processes in acid denaturation. *J. Mol. Biol.* **60**,499-508, (1971).
- Ikai, A., and Tanford C. Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins. *Nature* **230**, 100-102, (1971).
- Garel J.R., and Baldwin R.L. Both the fast and slow refolding reactions of ribonuclease A yield native enzyme. *Proc. Natl. Acad. Sci. USA.* **70**, 3347-3351, (1973).
- Brandts, J.F., Halvorson, H.R., Brennan, M. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **14**, 4953-4963, (1975).
- Stewart D.E., Sarkar, A., and Wampler J.E. Occurrence and role of cis peptide bonds in protein structures. *J. Mol. Biol.* **214**,253-260, (1990).
- Fisher G., Bang, H., and Mech., C. *Biomed. Biochim. Acta* **10**, 1101-1111, 1984.
- Bengtzelius, U. Theoretical calculations on liquid-glass transitions in Lennard-Jones systems. *Phys. Rev. A.* **33**, 3433-3439, (1986)
- Götze, W. in *Liquids, Freezing and the Glass Transition..* J.P. Hansen and D Levesque Eds. 1989.
- Avramov, I., Viscosity in disordered media. *J. Non. Cryst. Solids.* **351**, 3163-3173, (2005).
- García-Colín, L. S. Remarks on the glass transition. *Rev. Mex. Fis.* **1**, 11-17, (1999).
- Schellman, J.A. Temperature, stability and the hydrophobic interaction. *Biophysics J.* **73**, 2960-2964, (1997).
- Pace C.N., Shirley B.A., McNutt M., and Gajiwala K. Forces contributing to the conformational stability of proteins. *FASEB J.* **10**, 75-83, (1996).
- Kumar, S., Tsai C.J., and Nussinov R. Maximal stabilities of reversible two state proteins. *Biochemistry* **41**, 5359-5374, (2002).
- Kumar, S., Tsai C.J., and Nussinov R. Temperature range of thermodynamic stability for the native state of reversible two-state proteins. *Biochemistry* **42**, 4864-4873, (2003).
- Murphy K.P., Privalov P.L. and Gill S.J. Common features of protein unfolding and dissolution of hydrophobic compounds. *Science* **247**, 559-561, (1990).
- Baldwin R. Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. USA* **83**, 8069-8072, (1986).
- Robertson A.D., Murphy K.P. Protein structure and the energetics of protein stability. *Chem. Rev.* **97**, 1251-1267, (1997).
- Prabhu N.V., and Sharp K.A. Heat capacity in proteins. *Annu. Rev. Phys. Chem.* **56**, 521-548, (2005).
- Orecchini, A., Paciaroni, A., Bizarri, A.R., and Cannistraro, S. Low-frequency vibrational anomalies in  $\beta$ -lactoglobulin *J. Phys. Chem B* **105**, 12150-12156, (2001).
- Leyser H., Doster, W., and Diehl M. *Phys. Rev. Lett.* **82**, 2987, (1999)
- Piazza F., De los Rios, P. Slow energy relaxation of macromolecules and nanoclusters in solution. *Phys. Rev. Lett.* **94**, 145502, (2005)
- Adam G., and Gibbs, J. On the temperature dependence of cooperative relaxation properties in glass-forming liquids. *J. of Chem. Phys.* **43**, 139, (1965).
- Williams, M.L., Landel, R.F., and Ferry J. *J. Am. Chem. Soc.* **77**, 3701, 1955.
- Williams, M.L. *J. Phys. Chem.* **59**, 95, 1955.

- Privalov P.L., and Makhatadze G.I. Heat capacity of proteins I. Partial molar heat capacity of individual aminoacids residues in aqueous solution. *J. Mol. Biol.* **215**, 585-591, (1990)
- Abruzzetti, S., Sottini, S., Viapanni, C., and Corrie, J.E. Acid induced unfolding of myoglobin triggered by a laser pH jump method. *Photochem. Photobiol. Sci.* **5**, 621-628, (2006).
- Miksovská, J., Gennis R.B., Larsen R.W. Spectroscopy and photothermal study of myoglobin conformational changes in the presence of sodium dodecyl sulphate. *Biomacromolecules.* **7**, 476-482, (2006).
- Reeder B.J., Svistunenko D.A., Cooper C.E., and Wilson M.T. The radical and redox chemistry of myoglobin and hemoglobin: from in vitro studies to human pathology. *Antiox. Redox Signal* **6** (6), 964-966, (2004).
- Watson H.C., and Kendrew J.C. The aminoacid sequence of sperm whale myoglobin. *Nature* **190**, 670-672, (1961).
- Cangialosi, D., Alegría, A., Colmenero, J. A thermodynamic approach to the fragility of glass forming polymers. *J. Chem. Phys.* **124**, 24906, (2006).
- Achterhold, K., Sturhanhn, W., and Alp, E.E. Phonon-assisted Mosbauer effect: the vibrational density of states of myoglobin. *Hyperfine Interact.* **141**, 3-12, (2002).
- Brandts J.F. *Biochemistry.* **31**, 3947-55, 1992.
- Pathria, R.K. *Statistical Mechanics.* Pergamon Press. (1972).
- Marko J.F., and E.D. Siggia. Statistical mechanics of supercoiled DNA. *Phys. Rev. E.* **52**, 2912-2938, (1995).
- Kamien R.D., Lubensky T.C., and Nelson P. Direct determination of DNA twist-streht coupling. *Europhys Lett.* **38**, 237-242, (1997)
- Kamien R.D. The geometry of soft materials: a primer. *Rev. Mod. Phys.* **74**, 953, (2002)
- Stoker J.J. *Differential Geometry.* Wiley-Interscience, 1969
- Guggenheimer H.W. *Differential Geometry.* Dover Publications New York, 1963.
- Van Holde, K., Johnson W., and Shing P. *Physical Biochemistry.* Prentice Hall Inc, (1998).
- Binney, J.J., Newman, M.E.J. *The theory of critical phenomena.* Claredon Press. Oxford (1992).