

UNIVERSIDAD NACIONAL DE LA PLATA FACULTAD DE CIENCIAS EXACTAS DEPARTAMENTO DE CIENCIAS BIOLÓGICAS

TÉSIS DOCTORAL

"PROPIEDADES DINÁMICAS Y ESTRUCTURALES DE LAS PROTEÍNAS: EFECTOS DE LA PRESIÓN"

Andrés N. McCarthy

Tesis propuesta para optar por el Título de Doctor de la Facultad de Ciencias Exactas Universidad Nacional de La Plata

(043.2) TESIS 01126 Universidad Nacional de La Plata Facultad de Clenclas Exactas Biblioteca 50 y 115 1º subsuelo biblioteca@exactas.un(p.edu.ar Tel 0221 422-6977/79 int. 129



rzo de 2006

Dedicatoria

AGRADECIMIENTOS

SINAPSIS

TABLA DE CONTENIDO

Capítulo 1

Introducción:

- 1. Definiciones Generales
- 2. Proteínas: organización espacial y problemas relacionados
- 3. Predicción de la estructura a partir de la secuencia: una necesidad
- 4. Estructura de las proteínas: un problema complejo
- 5. Estado nativo ¿Estable o metaestable?
- 6. Breve reseña histórica sobre las fuerzas postuladas como determinantes del plegamiento
 - a. Interacciones de largo alcance: Electrostática
 - b. Puentes de hidrógeno e interacciones de van der Walls
 - c. Efecto hidrofóbico
 - d. Interacciones locales: Propensiones intrínsecas
- 7. Observaciones
- 8. Objetivos de la presente Tesis

Capítulo 2

Técnica: Simulación por Dinámica Molecular:

- 1) Introducción
 - a) Las simulaciones son Clásicas
 - b) Los electrones se encuentran en el estado energético fundamental
 - c) Los Campos de Fuerza (CF) son aproximados
 - d) El Campo de Fuerzas es aditivo y analiza sólo interacciones de a pares
 - e) Las interacciones de largo alcance son tratadas con radios de corte
 - f) Las condiciones de contorno no resultan naturales
- 2) Algoritmo general utilizado en Dinámica Molecular
- 3) El Campo de Fuerzas
 - a) Potenciales no-enlazantes
 - i) Interacción de Lennard-Jones
 - ii) Interacción Coulómbica
 - iii) Interacción Coulómbica con Campo de Reacción
 - iv) Interacciones no-enlazantes modificadas
 - b) Potenciales enlazantes
 - i) Potencial de estiramiento de enlace
 - (1) Potencial armónico
 - (2) Potencial de Morse
 - ii) Potencial armónico de ángulo de enlace

- iii) Potencial para diedros impropios
- iv) Potencial para diedros propios
 - (1) Diedros propios: Potencial periódico
 - (2) Diedros propios: Potencial de Ryckaert-Belleman
- c) Interacciones Especiales
 - i) Restricciones de posición
 - ii) Restricciones de ángulo
 - iii) Restricciones de distancia
- 4) Generalidades del cómputo de fuerzas
 - a) Condiciones iniciales
 - i) La topología y el campo de fuerzas
 - ii) Coordenadas y velocidades
 - iii) Movimiento del centro de masas
 - b) Búsqueda de vecinos
 - i) Generación de listas de pares
 - c) Cómputo de fuerzas
 - i) Energía potencial
 - ii) Energía cinética y temperatura
 - iii) La presión y el virial
 - d) Actualización de la configuración
 - e) Acoplamiento térmico
 - i) Acoplamiento térmico de Berendsen
 - ii) Acoplamiento térmico de Nosé-Hoover
 - iii) Acoplamiento térmico por grupos
 - f) Acoplamiento de presión
 - i) Baróstato de Berendsen
 - ii) Baróstato de Parinello-Raman
- 5) Algoritmos de restricciones topológicas
 - a) SHAKE
 - b) LINCS
 - i) El algoritmo LINCS
 - ii) Las fórmulas de LINCS
- 6) Algoritmo completo de actualización

Capítulo 3

Lisozima en H2O a 3kbar:

- 1) Introducción
- 2) Materiales y métodos
- 3) Resultados y discusión
 - a) Comparación de las estructuras promedio de alta y baja presión
 - b) Screening estructural mediante el cálculo de la raíz de la desviación cuadrática media para alta y baja presión
 - c) Análisis de la evolución de la estructura secundaria
 - d) Análisis de movilidad
 - e) Análisis de la superficie accesible al solvente (SAS)

4) Conclusiones del capítulo

Capítulo 4

Apomioglobina en H2O a 3kbar:

- 1) Introducción
- 2) Materiales y métodos
- 3) Resultados y discusión
 - a) Estructuras promedio de referencia y raíz de la fluctuación cuadrática media (RMSF)
 - b) Análisis de la raíz de la desviación cuadrática media (RMSD) de los carbonos-α a lo largo de todas las trayectorias
 - c) Análisis comparativo de la Estructura Secundaria (SS) y su evolución
 - d) Análisis comparativo de la Superficie Accesible al Solvente (SAS) y su evolución
- 4) Conclusiones del capítulo

Capítulo 5

Agua pura a 1bar y 3kbar:

- 1) Introducción
- 2) Materiales y métodos
- 3) Resultados y discusión
 - a) Función distribución radial (RDF) para los oxígenos
 - b) Distribución de los puentes de hidrógeno por molécula de agua
 - c) Distribución de distancias de los puentes de hidrógeno a 1bar y 3kbar
 - d) Distribución de ángulos de los puentes de hidrógeno a 1bar y 3kbar
 - e) Evolución temporal del número total de puentes de hidrógeno a 1bar y 3kbar
- 4) Conclusiones del capítulo

Capítulo 6

Conclusiones:

- 1) Validación de los sistemas estudiados
- 2) Diferencias de comportamiento entre los modelos
- 3) Estructura secundaria
- 4) El factor común entre ambos modelos
- 5) Cambios en la interacción hidrofóbica
- 6) Cambios de volumen
- 7) Estabilización térmica por presión
- 8) Palabras Finales

Capítulo 7

Perspectivas:

- 1) Estudio detallado de la hidratación de las proteínas en solución.
- 2) Construcción de un modelo de hidratación de proteínas abarcativo, que solucione las deficiencias de los disponibles en la actualidad.

Anexo 1

Técnicas de minimización de energía

Anexo 2

Herramientas de análisis de datos y presentación de resultados

Capítulo 1

INTRODUCCIÓN:

Definiciones generales:

Las proteínas son macromoléculas que poseen varias características únicas.

Desde el punto de vista biológico son las entidades en las cuales se traduce toda la información contenida en un genoma, a la vez que constituyen, directa o indirectamente, las herramientas ejecutoras de todas y cada una de las funciones necesarias para la existencia de la vida tal cual la conocemos.

No está demás recordar que, luego de ser sintetizadas, las proteínas sólo son capaces de desempeñar dichas funciones luego de haber alcanzado su estructura tridimensional definitiva.

El proceso a través del cual las proteínas alcanzan su estructura definitiva es conocido con el nombre de plegamiento y, a pesar de haber sido postulado por primera vez hace algo más de 65 años [1], su elucidación continúa siendo uno de los más grandes desafíos a los cuales los científicos se enfrentan en la actualidad.

Por un lado, se verifica experimentalmente que las proteínas presentes en la naturaleza logran plegarse rápida y confiablemente, a pesar de contar, en principio, con un número de conformaciones posibles de proporciones astronómicas [2].

A pesar de esta última observación, aún los métodos más actuales y sofisticados de predicción ofrecen resultados que no superan los valores de un 65% de acierto para la estructura secundaria [3], o de un 55% (en el mejor de los casos) para la estructura terciaria [4].

Corresponde preguntarse entonces ¿porqué, aún hoy, la calidad de las predicciones estructurales para las proteínas en solución es tan baja?

La respuesta resulta en verdad tan simple como desafiante: no conocemos aún en detalle las leyes fundamentales que gobiernan el proceso de plegamiento.

A pesar de esto, cabe destacar que enormes avances se han hecho en el campo de la biología estructural. En la actualidad contamos con un conocimiento cada vez más detallado de la gran mayoría de las fuerzas que intervienen en el proceso de estabilización estructural de las proteínas, y su comportamiento frente a las variables termodinámicas, como la temperatura y, más recientemente, también la presión.

En particular, estudios recientes han logrado grandes avances en la elucidación de la naturaleza y el comportamiento de muchas fuerzas inter e intramoleculares, tales como los puentes salinos (electrorestricción) [5, 6] o los puentes de hidrógeno [7,8,9,10].

A pesar de esto, y como en ningún otro aspecto de la biofísica molecular, queda aún una fuerte controversia en torno a las interacciones hidrofóbicas y cómo son éstas afectadas con el aumento de la presión.

La literatura actual expone la existencia de teorías antagónicas en torno a este punto [11,12,13,14,15,55]. Por un lado, algunos autores proponen una reducción general del efecto hidrofóbico con el aumento de la presión, debido principalmente a los efectos de ésta sobre el sistema solventeproteína [14,15]. Sin embargo, otros afirman que la formación de clatratos (o jaulas de agua) alrededor de los residuos hidrofóbicos, al involucrar un ΔV positivo para el sistema, resultaría negativamente seleccionado, dado que el sistema proteína-solvente debe moverse hacia estados de menor volumen con el aumento de la presión. Consecuentemente, según esta última teoría, las interacciones hidrofóbicas deberían verse reforzadas, no disminuidas, con el aumento de la presión [11,12,13].

Tales interpretaciones contrapuestas no resultan caprichosas, sino que por el contrario son una consecuencia del comportamiento experimental "anómalo" que presentan las proteínas en solución frente a los cambios en la temperatura o la presión. Más específicamente, se sabe desde hace ya más de medio siglo que la estabilidad óptima de la mayoría de las proteínas se produce alrededor de los 37 °C, y que las mismas pueden ser desplegadas ya sea aumentando o disminuyendo su temperatura [16]. Últimamente se ha podido verificar también un comportamiento "anómalo" con la presión; concretamente, mientras presiones del orden de 10⁴bar producen un desplegamiento completo de las proteínas, Bresiones moderadas del orden de 10²bar o menores logran estabilizar perturbaciones térmicas leves [17].

En la presente tesis se abordan los efectos de la presión sobre la estructura y movilidad de las proteínas, poniendo especial atención sobre todos aquellos parámetros relacionados directa o indirectamente con el efecto hidrofóbico y su comportamiento frente a dicha variable termodinámica.

Proteínas: organización espacial y problemas relacionados

Cuando consideramos el modo en el cual las proteínas se sintetizan en el interior de la célula, podemos observar que una cadena peptídica cualesquiera que crece a una velocidad de aproximadamente 50 aminoácidos [3] por segundo deberá organizase progresivamente para alcanzar una estructura bien definida para el momento en el cual deje el ribosoma.

Más aún, el gran número de modificaciones químicas específicas sufridas por la proteína, (acetilación, fosforilación, glicosilación, clivaje específico de un péptido para la activación de una enzima, etc.) genéricamente denominadas modificaciones post-traduccionales, solo ocurrirán sobre una proteína que haya alcanzado una estructura tridimensional determinada.

Por lo tanto, en general debe existir suficiente información estructural, determinada por la secuencia, para otorgarle a la proteína una estructura tridimensional única y reproducible en el medio en el cual ocurre el plegamiento.

Predicción de la estructura a partir de la secuencia: una necesidad

La primera determinación de la estructura tridimensional de una proteína fue la de la mioglobina en 1961, obtenida por cristalografía de rayos X. Desde entonces, varios miles de estructuras proteicas han sido determinadas a diferentes resoluciones, desde 3,5 hasta 0,69Å.

Más recientemente, merced a los avances tecnológicos y metodológicos en Resonancia Magnética Nuclear (RMN) se han determinado estructuras completas de proteínas pequeñas (PM<15000) en solución. De cualquier manera, resulta importante recordar que hoy sabemos que el número de proteínas diferentes codificadas dentro de un genoma, inclusive el de un organismo simple, resulta del orden de (o incluso superior a) todas las estructuras proteicas diferentes resueltas hasta la fecha.

Al mismo tiempo, las dificultades para cristalizar al utilizar difracción de rayos X, o la solubilidad y el tamaño de las proteínas cuando se usan técnicas de RMN, resultan ser frecuentemente obstáculos serios para el uso de dichos métodos. También resultan muy costosos en términos de tiempo y recursos económicos cuando se debe resolver una estructura tridimensional completa con resolución de escala atómica.

Haemophilus Influenzae fue le primer organismo en ser secuenciado, hace poco más de una década [18]; el primer eucariota unicelular a ser secuenciado, la levadura Sacharomyces Cerevisiae, le siguió sólo un año más tarde [19]; dos años después apareció el primer genoma animal en ser completamente secuenciado, el de *Caenorabditis elegans* [20] y, hace cinco años, la presentación del genoma humano agregó a esta lista la primera secuencia completa de un vertebrado [21,22].

En marzo de 2005 ya se contaba con una lista de secuencias genómicas completas para más de 250 organismos, y se estiman en más de 1000 los organismos que están siendo completamente secuenciados en este momento [4]. Además, nuevas técnicas de secuenciamiento están proveyendo de secuencias a una velocidad sin precedentes [23].

Esta explosión de información genómica disponible ha comenzado a cambiar completamente el rostro de la biología molecular, dado que dispara el desarrollo de nuevas técnicas experimentales de alta eficiencia. Por ejemplo, en lugar de enfocarse arbitrariamente en un receptor particular (porque es el único conocido), los investigadores pueden ahora utilizar como blanco a todos los receptores de la misma familia (que posean una determinada firma y/o fenotipo) del genoma humano completo [4].

Cada genoma enteramente secuenciado revela entre cientos y miles de secuencias proteicas para las cuales la única anotación posible es la de "proteína hipotética". De este modo, tanto en el genoma humano como en los genomas de los organismos patógenos podría haber miles de potenciales blancos terapéuticos aún desconocidos. A partir de aquí resulta claro que la predicción computacional, tanto de la estructura como de la función de estas proteínas, para las cuales sólo conocemos la secuencia, podría jugar un rol preponderante.

Por otro lado, en los últimos años se ha podido mostrar que diversos desórdenes degenerativos, anteriormente inconexos, pueden ser agrupados bajo la definición de desórdenes conformacionales de las proteínas (PCD - Protein Conformational Diseases) [24,25,26,27,28]. Este grupo incluye varios desórdenes muy conocidos como la enfermedad de Alzheimer, la encefalopatía espongiforme transmisible (más conocida como la enfermedad de la vaca loca), algunos tipos de anemia hemolítica, la enfermedad de Huntington, la fibrosis quística, la esclerosis lateral amiotrófica, la enfermedad de Parkinson, la amiloidosis relacionada a diálisis, y otras más de 15 enfermedades menos conocidas [29].

La marca registrada de todas estas enfermedades, agrupadas bajo la denominación de PCD, es el cambio en la estructura tridimensional nativa de una proteína "normal" sin alteración de la estructura primaria, por otra estructura diferente, disfuncional o "mal-plegada".

Es posible que la aparición de la enfermedad promovida por el cambio conformacional sufrido por la proteína en cuestión sea debido, ya sea a la aparición de una actividad "tóxica" por parte del nuevo confórmero, o bien a la simple desaparición de su actividad biológica a causa de la desaparición del estado nativo de dicha proteína [26,27,28].

No existe ninguna homología secuencial ni estructural evidente entre las proteínas implicadas en los PCD.

Sin embargo, la sorprendente característica compartida por todas estas proteínas es su inherente habilidad para adoptar al menos dos conformaciones estables diferentes.

Otra característica importante que varias de estas proteínas "malplegadas" tienen en común es que tienden a formar auto-asociaciones oligoméricas, generando así depósitos de agregados de tipo amiloideos en diversos órganos, induciendo el daño tisular y la consecuente disfunción en los mismos.

En la actualidad se están desarrollando terapéuticas sobre la base de péptidos cortos desarrollados por bioingeniería, diseñados para interactuar específicamente con la porción del confórmero "mal-plegado" que se supone responsable de la oligomerización [27,30].

Resulta claro sin embargo que, hasta que las bases moleculares del plegamiento no sean cabalmente resueltas, dicha terapéutica estará regida necesariamente por una proporción importante e incluso hasta limitante de prueba y error.

Por lo tanto resulta hoy una prioridad encontrar otros métodos, especialmente de alta confiabilidad, que permitan predecir la estructura, esto es: poder especificar la distribución espacial de todos lo átomos de la proteína a partir de la secuencia de los aminoácidos.

Y de hecho, esto sólo será posible una vez que hayamos sido capaces de descifrar el código necesario para predecir inequívocamente la estructura tridimensional que se corresponde con una determinada secuencia, en el medio en el cual esta se expresa.

Estructura de las proteínas: un problema complejo

Los siguientes puntos generales deben ser tenidos en cuenta en relación a la búsqueda de las estructuras de las proteínas:

• Tomemos una proteína con una secuencia de N aminoácidos. La selección de la secuencia no puede haber sido llevada a cabo por un proceso de prueba y error (utilizando solamente mutaciones puntuales). Aún con ritmos de mutación extremadamente elevados, la edad misma del universo (5 x 10¹⁷s) resultaría por lejos demasiado corta para explorar las 20^N posibilidades (Ej. con $N=300 \Rightarrow 20^{300} = 10^{400}$). Debemos asumir que la evolución sólo puede "editar" una secuencia seleccionada al azar [31], siempre y cuando la misma sea capaz de desempeñar una determinada función. Por lo tanto ninguna relación unívoca existe entre la secuencia y la función de una proteína: la estructura no está determinada por la función, mientras que la inversa resulta generalmente verdad. De existir un "código estructural", éste debe ser altamente degenerado dado que diferentes patrones secuenciales de aminoácidos pueden llevar a la misma estructura espacial. Esto nos lleva a que la verificación de la homología estructural, tan frecuentemente utilizada, resulte probablemente una condición necesaria, aunque no suficiente, de homología funcional.

• Una proteína con una nueva función puede ser el resultado del ensamblado, a nivel del RNA mensajero, de varias secuencias de nucleótidos, siendo cada una de estas el resultado de una selección evolutiva previa para el desempeño de funciones más simples. Tal fusión de dominios funcionales acorta el tiempo de evolución requerido, y resulta común encontrar enzimas en eucariotas superiores que resultan del ensamblado de enzimas procariotas más primitivas.

• Desde un punto de vista termodinámico, encontramos (principalmente) las siguientes contribuciones a la energía libre total de una proteína:

- Entropía conformacional por la pérdida de grados de libertad debido a las restricciones topológicas de los aminoácidos y la restricción de movimiento de las cadenas laterales.
- 2) Energía de interacción de los puentes de hidrógeno dentro de la proteína, y entre la proteína y las moléculas de agua externas.
- 3) Energía de los enlaces de Van der Walls o, en términos más generales, de los "enlaces hidrofóbicos".
- Energía Coulómbica de los enlaces electrostáticos (puentes salinos, quelatos internos, etc.) y el acoplamiento entre los dipolos formados por las α-hélices.
- 5) Energía de enlace de valencia y energía de estrés topológico en los puentes disulfuro.
- 6) Entropía del solvente al formar la interfase sobre los residuos hidrofóbicos (apolares) expuestos.

A partir de esta lista (que no es exhaustiva) se puede intentar determinar la energía libre conformacional total de la proteína, y buscar así la conformación que minimice dicha energía libre. Sin embargo, dado que existe un número tan grande de variables, existen también un gran número de mínimos, y resulta muy difícil encontrar cuál de todos estos corresponde al de mínima energía libre total, y por lo tanto al de mayor estabilidad de la proteína. Más aún, resulta posible también que la proteína no se encuentre en un estado estable sino en un estado metaestable con un tiempo de vida media lo suficientemente largo como para otorgarle las mismas características que las de un estado estable.

Estado nativo ¿Estable o metaestable?

Levinthal, ya en 1969 estableció su famosa "paradoja" [2]. Observaba allí que para una proteína de 150 residuos (2000 átomos aproximadamente) existiría en principio un total de 6000 grados de libertad que describirían el sistema. Tomando en cuenta las restricciones topológicas de una cadena polipeptídica, dichos grados de libertad podían verse drásticamente reducidos a 450, siendo 300 los responsables de describir a la cadena principal y 150 debido a los movimientos de las cadenas laterales. Suponiendo que bastaría con determinar los primeros 300, y que es suficiente con poder diferenciar los distintos estados dentro de la décima del radián, habría 10³⁰⁰ configuraciones posibles para esta proteína hipotética.

Dado que en la naturaleza las proteínas tardan del orden de los segundos para alcanzar su estado plegado, y aún asumiendo un tiempo mínimo para moverse de una conformación a otra, las proteínas tendrían tiempo para explorar alrededor de 10^{s} conformaciones diferentes antes de alcanzar su estado final.

Concluía en primer lugar, a partir de las observaciones previas, que el proceso de plegamiento debería estar acelerado y dirigido por la rápida formación de interacciones locales, las cuales luego resultarían determinantes para la continuación de dicho plegamiento.

Por último concluía que la conformación final de una proteína no necesariamente era la de mínima energía. Sin embargo aclaraba que, de encontrarse en un estado metaestable, ésta debería estar atrapada dentro de un pozo energético lo suficientemente profundo como para sobrevivir a las posibles perturbaciones presentes en un sistema biológico.

Tal afirmación parecería estar en contradicción con la "hipótesis termodinámica" fuertemente establecida por Anfinsen, a partir de su exhaustivo trabajo experimental sobre la renaturalización de la ribonucleasa pancreática bovina, en 1973 [32].

Sin embargo, esta aparente contradicción desaparece cuando recordamos que el estado estable será, por definición, el de mínima energía

global respecto de todos los otros estados accesibles dentro de dicha escala de tiempo.

Así, incorporando una (solapada) restricción cinética dentro de un concepto puramente termodinámico logramos resolver este aparente desacuerdo.

Breve reseña histórica sobre las fuerzas postuladas como determinantes del plegamiento:

En este acotado resumen sobre la naturaleza de las fuerzas que intervienen en el plegamiento de las proteínas, y cómo fue cambiando la visión sobre las mismas en el transcurso del siglo pasado, resultará útil establecer la diferencia entre las fuerzas de corto alcance y las de largo alcance, por un lado, y las interacciones locales y no-locales (o a distancia), por el otro.

El alcance de una fuerza resulta de su dependencia con la distancia: las energías que dependen de la distancia r, según r^{-n} serán de largo alcance si $r \leq 3$ (Ej. interacciones ión-ión o ión-dipolo) o serán de corto alcance si r > 3 (Ej. interacciones de Lennard-Jones). Esta dependencia inversa con la tercera potencia es la división natural, dado que para medios simples puros, la integral que da la energía total de un sistema diverge, según esta definición, para fuerzas de largo alcance y converge para las de corto alcance [33].

Para cadenas poliméricas tales como las proteínas, también resulta importante la posición de un segmento de la cadena. Las interacciones locales son aquellas que se producen entre segmentos contiguos de la cadena (i, i+1), o vecinos cercanos de la cadena. Las interacciones no-locales (o a distancia) se refieren a aquellas que se producen entre segmentos significativamente distantes de la cadena. Las interacciones locales pueden aparecer a causa de las fuerzas de corto o largo alcance, al igual que las interacciones a distancia.

Interacciones de largo alcance: Electrostática

Dado que los ácidos y las bases fueron de los primeros agentes desnaturalizantes de las proteínas en ser descubiertos, inicialmente se suponía que las fuerzas determinantes en el proceso de plegamiento debían ser de naturaleza electrostática. El signo inequívoco de un proceso gobernado electrostáticamente es su dependencia del pH y/o de la fuerza iónica del medio. Aplicado a las proteínas, el pH controla la carga total sobre la proteína, mientras que la concentración salina regula el grado interacción entre las mismas, debido a su efecto de apantallamiento. Durante la década de 1930, los pares iónicos eran considerados como las fuerzas principales en la estabilidad conformacional de las proteínas [34,35, 36].

Sin embargo, en 1949 Jacobsen & Liderstrom-Lang establecieron la primer falla insalvable en esta teoría, analizando el cambio de volumen que surgiría de tales interacciones y corroborando que resultan de signo inverso a lo observado experimentalmente durante el proceso de plegamiento [37]. A partir de aquí muchos otros experimentos fueron realizados, produciendo resultados en este mismo sentido [38,39,40].

Puentes de hidrógeno e interacciones de van der Walls

Las interacciones de van der Waals tienen dos orígenes diferentes. La repulsión, debida al principio de exclusión de Pauli, y la atracción (fuerzas de London- o fuerzas de dispersión) que se producen entre átomos o moléculas neutras por la atracción de dipolos instantáneos.

Un puente de hidrógeno ocurre cuando un átomo de hidrógeno es compartido generalmente por dos átomos electronegativos. La energía del enlace puente hidrógeno se encuentra entre los valores de 2 a 10 kcal/mol [41], y depende de la orientación y la electronegatividad de los átomos que lo conforman.

Además, un enlace puente hidrógeno resulta primordialmente de un arreglo lineal de donor, hidrógeno, y aceptor y está compuesto por interacciones de tipo electrostática, de dispersión, de transferencia de carga, e interacciones de repulsión de tipo estéricas [42]. Sin embargo, el componente principal del puente de hidrógeno es electrostático [41,42].

Por lo tanto, aunque el mecanismo microscópico de ambas interacciones es bastante diferente, a los fines prácticos ambas interacciones han sido agrupadas en una misma sección.

En realidad fueron Mirsky & Pauling, en 1936, los primeros en sugerir que los puentes de hidrógeno eran la fuerza principal en el plegamiento de las proteínas [35]. Si bien su enfoque parece haber estado orientado hacia aquellos puentes de hidrógeno producidos por los puentes salinos entre las cadenas laterales, también sugirieron que podían formarse puentes de hidrógeno entre los oxígenos de los carbonilos y los hidrógenos de los aminos de la cadena principal. Su propuesta llevó al descubrimiento de las α -hélices y hojas- β antiparalelas [43] en 1951.

A partir de aquí, se desarrollaron varios modelos teóricos de las transiciones hélice-ovillo que lograron predecir exitosamente su comportamiento frente a la temperatura y la longitud de la cadena [44]. Dichos modelos también pudieron predecir con éxito el comportamiento frente a los cambios de pH [45].

Sin embargo, ya en 1954 Kauzman concluía que, si bien estas fuerzas de corto alcance resultaban de una relevancia indudable, las mismas no podían constituir la fuerza dominante en el plegamiento de las proteínas [46].

Un criterio fundamental para definir la fuerza dominante de un proceso es que ésta debe explicar claramente porqué el estado final resulta indudablemente más favorable que el inicial.

Kauzman argumentaba que los puentes de hidrógeno no podrían satisfacer este criterio, dado que no existía ningún fundamento para suponer que los puentes de hidrógeno intracatenarios formados en el estado plegado eran de menor energía que los formados entre la cadena y el agua en el estado desplegado.

De esta última aseveración se puede concluir, sin embargo, que una proteína plegada deberá contener un gran número de puentes de hidrógeno, ya que de lo contrario la misma se desnaturalizaría espontáneamente.

Efecto hidrofóbico

A partir de dos notables trabajos [46,47], Kauzman construyó la primera posición sólida en defensa de la importancia de las interacciones hidrofóbicas en el plegamiento de las proteínas.

Razonaba allí que la formación de un "enlace" hidrofóbico (el cual llamó "enlace anti-hidrógeno") al plegarse una proteína, involucra la ganancia de un enlace puente hidrógeno *completo* entre moléculas de agua, el cual debería resultar superior en un orden de magnitud a un simple *cambio de fuerza* de un puente de hidrógeno, si fuese el plegamiento un proceso conducido por éste último.

La observación más notable realizada por Kauzman posiblemente la constituya el haber señalado la similitud entre el comportamiento "anómalo" de la energía libre de hidratación de solutos no-polares, y los resultados experimentales [48] que mostraban un comportamiento análogo para el desplegamiento de las proteínas. Esto es, que los solutos no polares muestran un aumento de solubilidad en agua, tanto a bajas como a altas temperaturas, mientras que las proteínas muestran una mayor inestabilidad para los mismos extremos térmicos.

Nació así un muy prometedor modelo para explicar la termodinámica del plegamiento: El modelo de transferencia de hidrocarburos simples, desde un medio acuoso hacia un solvente no-polar. Dicho modelo logró reproducir, a partir de una minuciosa selección entre distintos hidrocarburos y solventes, incluso cuantitativamente el cambio en la estabilidad de las proteínas con la temperatura [49,50,51,52].

Sin embargo, la dependencia de la estabilidad de las proteínas con la presión no se asemeja en absoluto al comportamiento observado para el modelo de transferencia de hidrocarburos simples. Luego de varios trabajos señalando este aspecto [38,40], fue el propio Kauzman [53] quien en 1987 reconoció que: "El modelo del hidrocarburo-líquido falla casi completamente cuando uno intenta extenderlo a los efectos de la presión sobre el plegamiento de las proteínas."

Recientemente se ha intentado abordar este problema utilizando estrategias tanto teóricas como de simulación [54,55], para tratar de explicar esta aparente falla del modelo hidrofóbico para explicar el desplegamiento de las proteínas bajo presión.

A pesar de la ausencia de un modelo definitivo que sea capaz de explicar todos los aspectos del efecto hidrofóbico, se ha ido acumulando un considerable cuerpo de evidencias experimentales en respaldo de que la fuerza dominante en el proceso de plegamiento de las proteínas es la interacción hidrofóbica, y se han desarrollado también algunos modelos que intentan unificar dichas observaciones.

Cabe destacar sin embargo que, a pesar del extremo interés que posee el desarrollo de modelos teóricos, se debe remarcar que lamentablemente los existentes hasta la actualidad han sido basados en el uso de objetos esféricos no-polares simples.

Hummer y colaboradores [55] desarrollaron uno de tales modelos. En el mismo sugirieron como solución para el rompecabezas de la desnaturalización inducida por presión, el considerar la transferencia de agua hacia el interior del "core" hidrofóbico, en contraposición a la transferencia de residuos no polares hacia el medio acuoso. Así, plantearon que la desnaturalización por presión estaría acompañada por una penetración de moléculas de agua hacia el interior de la proteína, más que por una transferencia del interior hidrofóbico hacia el medio exterior.

En contraposición con esta visión del problema aparecen los resultados de reciente publicación aquí presentados [56,57]. Los mismos se realizaron mediante la técnica de dinámica molecular utilizando modelos con detalle atómico. Estos muestran, por un lado, una penetración colectiva de las moléculas de agua hacia el interior de la proteína, producto de la marcada separación de dos regiones específicas de la misma, con el consecuente aumento en la exposición de zonas hidrofóbicas normalmente alejadas del solvente [57]. Pero por otro lado, también muestran importantes cambios en las superficies hidrofóbica e hidrofílica expuestas, las cuales no resultan de

penetración alguna, ni singular ni colectiva, por parte de las moléculas de agua hacia el interior proteico [56].

Por lo tanto la aplicación directa de las conclusiones derivadas de modelos sobre-simplificados de sistemas reales y complejos tales como las proteínas, podría resultar en una peligrosa subestimación de efectos tales como la conectividad, las interacciones polares, los puentes de hidrógeno, y el empaquetamiento tridimensional específico, efectos todos que juegan un rol central en la determinación del estado nativo de una proteína [58]. Parece entonces necesario el uso de modelos que se aproximen de modo menos reduccionista a este problema.

Interacciones locales: Propensiones intrínsecas

El término "propensión intrínseca" no describe ningún tipo de fuerza en particular. Más bien pretende transmitir el concepto de que existen ciertas preferencias conformacionales por parte de los di- o tripéptidos, dependiendo de la secuencia, que resultan de la sumatoria de fuerzas de corto y largo alcance entre residuos vecinos.

Las propensiones intrínsecas han sido estudiadas a través de la medición del equilibrio de las transiciones hélice/ovillo de péptidos en solución[59,60,61], como así también mediante el equilibrio vuelta/ovillo [62,63,64]. Hoy sabemos que la estabilidad de las largas hélices polipeptídicas en solución puede explicarse mediante sus propensiones intrínsecas [44,65,66].

Sin embargo, en ausencia de otras fuerzas, las propensiones intrínsecas resultan aparentemente insuficientes para explicar acabadamente la estabilidad de las hélices dentro de la estructura terciaria de las proteínas [67,68,69].

Kabsch y Sander encontraron que en 6 de 25 casos un determinado pentámero puede encontrarse en la conformación de hélice, mientras que una secuencia idéntica puede estar en conformación de hoja plegada en una proteína diferente, lo cual implica que usar información local únicamente no resulta suficiente para explicar la conformación adoptada por una determinada secuencia en una proteína.

Como se ha dicho previamente, los métodos de este tipo tienen un grado de éxito de aproximadamente 65%.

¿Pero que significa un 65 % de éxito en relación a la magnitud de los factores no-locales que no son tomados en cuenta por estos métodos predictivos?

Si suponemos que la conformación de un residuo cualesquiera puede ser predicha con una probabilidad de éxito $p_0 = 1/3$ [70], siendo éste un estimativo grosero para la clasificación entre hélice, hoja plegada, u otra conformación cualquiera (no necesariamente ovillo), entonces un grado de éxito p conlleva una cantidad de información, $\langle I \rangle$:

$$\langle I \rangle = p \ln\left(\frac{p}{p_0}\right) + (1-p) \ln\left[\frac{(1-p)}{(1-p_0)}\right]$$

Por lo tanto, un grado de éxito p = 0.60 - 0.70 implica que los factores locales dan cuenta únicamente de alrededor de 15-30% de la información total requerida para realizar una predicción perfecta [71].

Observaciones

A partir de este apretado resumen se pretenden resaltar algunos aspectos centrales del problema del plegamiento de las proteínas, a saber:

1. El grado de complejidad del problema explica porqué la resolución del mismo nos ha eludido por más de medio siglo.

2. Su comprensión acabada resulta una necesidad desde el punto de vista de todas las ciencias ligadas a la biología, con especial interés por parte de las biomédicas.

3. Resulta claro que las aproximaciones desde la estadística o la teoría de la información, si bien pueden resultar útiles, no son capaces de ofrecer una solución acabada para este problema. Se precisa entonces de la elaboración de un modelo más abarcativo para la hidratación de proteínas.

4. Para elaborar dicho modelo debe ofrecerse una solución definitiva para el antagonismo teórico existente en torno a las interacciones hidrofóbicas.

5. Las interacciones hidrofóbicas constituyen hasta el presente la fuerza dominante en el proceso de plegamiento de las proteínas

Objetivos de la presente Tesis

En el presente trabajo se aborda el complejo problema de los cambios conformacionales sufridos por las proteínas sometidas a presiones elevadas.

La herramienta elegida para dicho estudio es la simulación por dinámica molecular, debido a que ofrece una perspectiva molecular con resolución a escala atómica, además de brindar información sobre la movilidad estructural de las proteínas. Dicha resolución se plantea hoy como determinante para lograr una perspectiva microscópica válida que permita el desarrollo de un nuevo modelo abarcativo para la hidratación de proteínas en solución.

Las simulaciones por Dinámica Molecular constituyen una herramienta ampliamente utilizada en el estudio de las proteínas en solución. Merced al uso de esta herramienta se han analizado los posibles intermediarios en el plegamiento de un sinnúmero de proteínas sometidas a distintos tipos de perturbaciones [72], se han podido corroborar o reformular hipótesis sobre los mecanismos de funcionamiento de diversas proteínas globulares [73,74,75], y más recientemente han sido utilizadas para estudiar los complejos mecanismos involucrados en el transporte a través de membranas celulares mediado por proteínas [76].

Si bien la Dinámica molecular ha sido utilizada previamente para obtener información detallada del efecto de la alta presión sobre las proteínas [77,78,79,80,81], la escala de tiempo abarcada en estas simulaciones se encontraba en el rango de los cientos de picosegundos a unos pocos nanosegundos totales.

Aunque la escala de tiempos explorada en los referidos trabajos imposibilitó el registro de cualquier fenómeno que tuviese lugar en una escala de tiempo mayor, los resultados allí discutidos bastaron para instalar la idea que:

"Dado que la alta presión no produce una disminución suficiente de las barreras energéticas y que, más aún, disminuye la cinética de tales procesos, ésta no constituye una adecuada perturbación para lograr el desplegamiento de las proteínas en los tiempos accesibles a la simulación." [58]

Y más aún, que:

"Lo que las simulaciones a presión elevada permiten es el estudio de la relajación elástica producida luego de un salto de presión" pero que, dado que dicha relajación elástica constituye una reversible y leve desviación del estado original, la simulación no resulta capaz de mostrar "...ningún incremento relativo sistemático de la superficie accesible al solvente de los residuos hidrofóbicos durante la relajación elástica, ni existe una penetración notable de moléculas de agua hacia el interior del "core" hidrofóbico." [58]

En contradicción con estas últimas afirmaciones, en el presente trabajo se muestra que, llevando la escala de tiempo de las simulaciones al rango de las decenas a centenas de nanosegundos, se observan aumentos sostenidos de la superficie hidrofóbica accesible al solvente, se observa la penetración de moléculas de agua al interior de la proteína, y se observan cambios estructurales no reversibles, lo cual implica cambios conformacionales plásticos (no elásticos).

Adicionalmente, la aplicación de esta herramienta a sistemas que presentan comportamientos experimentales aparentemente antagónicos permite encontrar un factor común entre los mismos, a la vez que ofrece una explicación para las diferencias observadas entre los mismos.

Por último, a parir del presente trabajo surgen elementos que parecerían permitir, luego de un minucioso análisis, aportar a la construcción de un modelo más completo para la hidratación de las proteínas en solución.

Capítulo 2

TÉCNICA: SIMULACIÓN POR DINÁMICA MOLECULAR

1. Introducción:

A modo de resumen introductorio sobre la técnica de Dinámica Molecular se presentará a continuación un breve pero general repaso sobre sus fundamentos, como así también las ventajas y limitaciones que presenta esta técnica en la actualidad.

Las simulaciones por Dinámica Molecular (DM) resuelven las ecuaciones de Newton para un sistema de N partículas que interaccionan entre sí:

$$m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial^2 t_i} = \boldsymbol{F}_i , \quad i = 1...N$$

Las fuerzas son las derivadas negativas de una función potencial $V(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$:

$$\boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i}$$

Las ecuaciones se resuelven simultáneamente en pequeños pasos. El sistema evoluciona durante algún tiempo, guardando el recaudo de que la temperatura y la presión permanezcan dentro de un rango estipulado alrededor de los valores requeridos, registrando también las coordenadas atómicas a intervalos regulares. Las coordenadas en función del tiempo representan la *trayectoria* del sistema.

Luego de los cambios iniciales, el sistema normalmente tenderá a alcanzar un *estado de equilibrio*. Muchas propiedades macroscópicas pueden entonces extraerse de los datos de salida, cuando se promedia sobre una *trayectoria* de *equilibrio*.

En este punto resulta importante considerar las limitaciones de las simulaciones realizadas con esta técnica.

a. Las simulaciones son Clásicas:

El uso de las ecuaciones de Newton para el movimiento implica automáticamente el uso de la "Mecánica Clásica" en la descripción del movimiento de los átomos. Esto es correcto para la mayoría de los átomos a temperaturas normales, pero existen excepciones. Los átomos de hidrógeno, al ser de masa tan pequeña, poseen movimiento de carácter esencialmente mecánico-cuántico. Por ejemplo, un protón puede moverse a través de una barrera de potencial por efecto túnel, lo cual constituye una transferencia de protón a través de un puente de hidrógeno.

Tal proceso no puede ser descripto correctamente por la mecánica clásica.

En la simulación por Dinámica molecular se soluciona este problema tratando a los enlaces covalentes y los respectivos ángulos que éstos forman entre sí como restricciones topológicas o "constraints". Un motivo práctico para dicho tratamiento resulta del eventual aumento en el paso de integración, lo cual resulta esencial a la hora de permitir movimientos realistas, además de un recorrido completo del espacio configuracional [82].

b. Los electrones se encuentran en el estado energético fundamental:

En la Dinámica Molecular utilizamos campos de fuerzas conservativos, los cuales son función de las posiciones de los átomos solamente. Esto significa que los movimientos electrónicos no son tomados en cuenta: se supone que los electrones ajustan su dinámica instantáneamente al cambiar las posiciones atómicas (aproximación de *Born-Oppenheimer*), además de permanecer en el un estado cuántico no excitado. Esto constituye una buena aproximación, casi siempre. Aunque, por supuesto, no pueden ser tomados en cuenta los procesos de transferencia electrónica como así tampoco aquellos con estados electrónicos excitados. Por lo tanto las reacciones químicas en general se encuentran fuera del alcance de ésta técnica.

c. Los Campos de Fuerza (CF) son aproximados:

Las fuerzas son proporcionadas por los Campos de Fuerzas. Estos no son verdaderamente parte del método de Dinámica Molecular y son susceptibles de modificaciones según se necesite, o bien avance el estado del conocimiento en la materia. Si bien el campo de fuerzas utilizado en el presente trabajo posee ciertas limitaciones (no se pueden incorporar polarizaciones y no posee ajuste fino para los parámetros de enlace) resulta de gran utilidad para los sistemas estudiados ya que se encuentra entre los más validados en la actualidad para biomacromoléculas en solución.

d. El Campo de Fuerzas es aditivo y analiza sólo interacciones de a pares:

Esto significa que todas las fuerzas no-enlazantes resultan de la suma lineal de pares de interacciones no-enlazantes.

Las interacciones no-enlazantes no descriptibles por interacciones de a pares, cuyo ejemplo más importante lo constituye la interacción a través de la polarización atómica, son representados a través de los potenciales efectivos de pares. Éstos incorporan sólo contribuciones promedio a las interacciones no-enlazantes no describibles por interacciones de a pares. Esto también significa que las interacciones de a pares no son puras; en otras palabras que no resultan válidas para pares aislados o para situaciones que difieran demasiado de aquellas para las cuales dichos sistemas fueron parametrizados.

De hecho, en la práctica los potenciales efectivos de pares resultan ser una muy buena aproximación, siempre que esta última precaución sea debidamente considerada.

e. Las interacciones de largo alcance son tratadas con radios de corte:

En el presente trabajo se han utilizado radios de corte o "cutoff", tanto para las interacciones de Lennard-Jones como para las interacciones Coulómbicas. Debido a la convención de la mínima imagen (sólo una imagen de cada partícula en las condiciones periódicas de contorno es considerada para la interacción de a pares), el radio de corte utilizado no puede exceder la mitad de la dimensión más pequeña del sistema. Esto puede presentar problemas para el tratamiento de las interacciones Coulómbicas, tales como la acumulación de cargas en la frontera del radio de corte y/o errores muy groseros en los valores de energía.

Para el presente trabajo se utilizó el método del campo de reacción, el cual considera explícitamente a todos los átomos que se encuentran dentro del radio de corte, pero trata al sistema como un medio continuo, de permitividad dieléctrica constante, por fuera de dicho radio.



f. Las condiciones de contorno no resultan naturales:

Dado que los sistemas simulados son pequeños (incluso un sistema de 40.000 partículas resulta pequeño), un cluster de partículas tendrá una gran cantidad de frontera no deseada con su entorno (vacío). Esto debe ser evitado si queremos simular el seno o "bulk" de un sistema. Es así que se utilizan las condiciones periódicas de contorno para evitar las verdaderas interfases. Pero dado que los líquidos no son cristales este tratamiento introduce un artefacto que tiende a exagerar la regularidad.

Si bien para sistemas grandes los errores cometidos son pequeños, para sistemas pequeños con gran correlación espacial interna las condiciones periódicas pueden incrementar dicha correlación. Este efecto se torna especialmente importante cuando se utiliza algún método de sumas continuas para el tratamiento de las interacciones electrostáticas de largo alcance (como sumas de Ewald), ya que éstos son conocidos por introducir un ordenamiento extra en el sistema. Por el contrario, el método elegido para el presente trabajo, Campo de Reacción, tiende a anular dicho artefacto de regularidad.

En la figura 1 puede observarse un esquema gráfico de las condiciones periódicas de contorno para un sistema de dos dimensiones.

2. Algoritmo general utilizado en Dinámica Molecular:

ALGORITMO GLOBAL PARA DM

1. Ingreso de condiciones iniciales

Potencial de interacción V como función de las posiciones atómicas Posiciones r de todos los átomos en el sistema Velocidades v de todos los átomos en el sistema

∜

Repetir 2, 3 y 4 el número requerido de veces

2. Cálculo de fuerzas

La fuerza sobre cualquier átomo

$$\boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i}$$

Se computa a través del cálculo de la fuerza entre los pares de átomos noenlazados:

$$F_i = \sum_j F_{ij}$$

Más las fuerzas debidas a interacciones enlazantes (las cuales pueden depender de 1, 2, 3 o 4 átomos), más las fuerzas de restricción y/o

externas.

Se computan las energías cinética y potencial y el tensor de presión ${\color{black}\Downarrow}$

3. Actualización de la configuración

Se simula el movimiento de los átomos resolviendo numéricamente las ecuaciones de movimiento de Newton

$$\frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \frac{\mathbf{F}_i}{m_i}$$

O bien
$$\frac{\partial \mathbf{r}_i}{\partial t} = \mathbf{v}_i \; ; \; \frac{\partial \mathbf{v}_i}{\partial t} = \frac{\mathbf{F}_i}{m_i}$$

$$\bigcup$$

4. Si se requiere: Paso de Salida

Escribir posiciones, velocidades, energías, temperatura, presión, etc.

3. El Campo de Fuerzas:

Un campo de fuerzas se compone de dos elementos claramente distinguibles:

- El conjunto de ecuaciones (las funciones potenciales) utilizadas para generar las energías potenciales y sus derivadas, las fuerzas.
- Los parámetros utilizados en este conjunto de ecuaciones.

Dentro de un conjunto de ecuaciones pueden utilizarse varios conjuntos distintos de parámetros. Debe tenerse cuidado de que la combinación de ecuaciones y parámetros constituyan un conjunto auto consistente.

En general debe tenerse mucho cuidado al hacer modificaciones *adhoc* de algún subconjunto de parámetros, ya que las distintas contribuciones a la fuerza total son normalmente interdependientes.

Las funciones potenciales pueden subdividirse en tres categorías:

- Potenciales no-enlazantes: de repulsión-dispersión (Lennard-Jones o Buckingham) y electrostáticas (Coulómbicas o Coulómbicas modificadas)
- Potenciales enlazantes: estiramiento del enlace covalente, ángulo de enlace, diedros impropios, y diedros propios.
- Potenciales especiales: restricciones de posición, restricciones de distancia, y restricciones de orientación.

Las funciones potenciales utilizadas en el presente trabajo son las siguientes:

a. Potenciales no-enlazantes:

Las interacciones no-enlazantes son aditivas de a pares y centrosimetricas:

$$V(\mathbf{r}_1, \dots \mathbf{r}_N) = \sum_{i < j} V_{ij}(\mathbf{r}_{ij}) ;$$

$$F_i = -\sum \frac{dV_{ij}(\mathbf{r}_{ij})}{dr_{ij}} \frac{\mathbf{r}_{ij}}{r_{ij}} = -F_j$$

Las interacciones no-enlazantes contienen un término de repulsión, uno de dispersión, y un término Coulómbico. Los términos de repulsión y dispersión se encuentran combinados, ya sea en el potencial de Lennard-Jones (o interacción 12-6) o en el potencial de Buckingham (o interacción exp-6). Adicionalmente, átomos cargados (parcialmente) actúan a través del término Coulómbico.

i. Interacción de Lennard-Jones:

El potencial de Lennard-Jones entre un par de átomos puede escribirse de la siguiente manera:

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}}$$

Ver la figura 2 para una representación gráfica del potencial.

Los parámetros $C_{ij}^{(12)}$ y $C_{ij}^{(6)}$ dependen del par de átomos considerado para la interacción; consecuentemente los mismos se toman de una matriz de parámetros de Lennard-Jones.

La fuerza derivada de este potencial resulta por lo tanto:

$$\boldsymbol{F}_{i}(\boldsymbol{r}_{ij}) = \left(12\frac{C_{ij}^{(12)}}{r_{ij}^{12}} - 6\frac{C_{ij}^{(6)}}{r_{ij}^{6}}\right) \boldsymbol{r}_{ij}$$

El potencial de Lennard-Jones puede también ser escrito de la siguiente forma:

$$V_{LJ}(r_{ij}) = 4 \epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right)$$

Consecuentemente, para construir la matriz de parámetros para las interacciones no-enlazantes de Lennard-Jones pueden utilizarse dos reglas de combinación:

$$C_{ij}^{(6)} = \left(C_{ii}^{(6)} * C_{jj}^{(6)}\right)^{1/2}$$
$$C_{ij}^{(12)} = \left(C_{ii}^{(12)} * C_{jj}^{(12)}\right)^{1/2}$$

O bien, alternativamente,

$$\sigma_{ij} = \frac{1}{2} (\sigma_{ii} + \sigma_{jj})$$
$$\epsilon_{ij} = (\epsilon_{ii} \epsilon_{jj})^{1/2}$$



ii. Interacción Coulómbica:

La interacción Coulómbica entre dos partículas cargadas viene dada por:

$$V_C\left(r_{ij}\right) = f \frac{q_i q_j}{\varepsilon_r r_{ij}}$$

donde $f = \frac{1}{4\pi\varepsilon_0}$, dicho potencial se encuentra graficado en la figura 3.

La fuerza derivada de este potencial resulta por lo tanto:

$$\boldsymbol{F}_{i}\left(\boldsymbol{r}_{ij}\right) = f \frac{q_{i}q_{j}}{\varepsilon_{r}r_{ij}^{2}} \frac{\boldsymbol{r}_{ij}}{r_{ij}}$$

iii. Interacción Coulómbica con Campo de Reacción:

La interacción Coulómbica puede ser modificada para sistemas relativamente homogéneos, suponiendo la existencia de un medio dieléctrico continuo por fuera del radio de corte r_c con constante dieléctrica ε_{rf} . Dicha interacción resulta por lo tanto:

$$V_{crf}\left(r_{ij}\right) = f \frac{q_i q_j}{r_{ij}} \left[1 + \frac{\varepsilon_{rf} - 1}{2\varepsilon_{rf} + 1} \frac{r_{ij}^3}{r_c^3} \right] - f \frac{q_i q_j}{r_c} \frac{\varepsilon_{rf}}{2\varepsilon_{rf} + 1}$$

En la cual el término constante de la derecha lleva el potencial a cero a la distancia r_c . Por simplicidad podemos reescribirla de la siguiente manera:

$$V_{crf}\left(r_{ij}\right) = fq_{i}q_{j}\left[\frac{1}{r_{ij}} + k_{rf}r_{ij}^{2} - c_{rf}\right]$$

donde:

$$k_{rf} = \frac{1}{r_c^3} \frac{\varepsilon_{rf} - 1}{\left(2\varepsilon_{rf} + 1\right)}$$

y

$$c_{rf} = \frac{1}{r_c} + k_{rf} r_c^2 = \frac{1}{r_c} \frac{3\varepsilon_{rf}}{(2\varepsilon_{rf} + 1)}$$

Para valores grandes de \mathcal{E}_{rf} el valor de k_{rf} tiende a $\frac{1}{2}r_c^{-3}$, mientras que para $\mathcal{E}_{rf} = 1$ esta corrección desaparece. Esto hace posible utilizar la misma expresión con y sin campo de reacción, empero con algún costo computacional extra.

En la figura 3 se grafica la interacción modificada, de donde resulta evidente que la derivada de dicho potencial con respecto a r_{ij} (=-fuerza) tiende a cero dentro de la distancia del radio de corte. La fuerza derivada de este potencial resulta entonces:

$$\boldsymbol{F}_{ij}\left(\boldsymbol{r}_{ij}\right) = fq_{i}q_{j}\left[\frac{1}{r_{ij}^{2}} - 2k_{rf}r_{ij}\right]\frac{\boldsymbol{r}_{ij}}{r_{ij}}$$



Figura 3: La interacción Coulómbica (para partículas cargadas de igual signo) con y sin campo de reacción (RF). Para el último caso se tomó $\varepsilon_{rf} = 78$ y $r_c = 0.9$ nm. Ambas líneas punteadas difieren sólo en el valor de una constante C, y pasamos de una a otra mediante la operación RF-C.

iv. Interacciones no-enlazantes modificadas:

Los potenciales no-enlazantes del campo de fuerzas pueden modificarse utilizando una función de "shift" o de corte suave. El objeto de esto es el de reemplazar las fuerzas truncadas por fuerzas continuas con derivadas continuas alrededor del radio de corte. Con dichas fuerzas los pasos de integración producen errores mucho más pequeños y complicaciones tales como la creación de cargas netas a partir de dipolos por el proceso de truncado dejan de existir. De cualquier manera, la función de corte suave produce una modificación importante de del potencial Coulómbico. A no ser que el término de largo alcance "perdido" sea calculado correctamente y agregado (a través del uso de algún método de sumas continuas como Ewald, PME o PPPM), el efecto de dichas modificaciones debe ser evaluado cuidadosamente. La modificación del potencial de repulsión-dispersión resulta mucho, menor pero logra eliminar el ruido causado por los efectos del truncado. No existe ninguna diferencia fundamental entre la función de corte suave (o "shift", la cual agrega una función a la fuerza o potencial) y la función "switch" o de llave (la cual multiplica el potencial con una función). De hecho, la función "switch" es un caso particular de la función "shift", la cual es aplicada sobre la función de la fuerza F(r), relacionada con las fuerzas electrostática o de Van der Walls que actúan sobre la partícula *i* debido a la partícula *j*, de manera que:

$$\boldsymbol{F}_{i} = cF\left(r_{ij}\right)\frac{\boldsymbol{r}_{ij}}{r_{ij}}$$

Para las interacciones Coulombicas o de Lennard-Jones puras:

$$F(r) = F_{\alpha}(r) = r^{-(\alpha+1)}$$

La fuerza de corte suave $F_s(r)$ puede en general escribirse como:

$$F_{s}(r) = F_{\alpha}(r) \qquad r < r_{1}$$

$$F_{s}(r) = F_{\alpha}(r) + S(r) \qquad r_{1} \le r < r_{c}$$

$$F_{s}(r) = 0 \qquad r_{c} \le r$$

Cuando $r_1 = 0$ esto resulta en una función tradicional de corte suave ("shift") de lo contrario actúa como una función de llave ("switch"). El potencial Coulómbico modificado resulta entonces:

$$V_{s}\left(r_{ij}\right) = f \mathbf{\Phi}_{s}\left(r_{ij}\right) q_{i} q_{j}$$

Donde $\Phi_{i}(r)$ es la función potencial:

$$\Phi_{s}(r) = \int_{r}^{\infty} F_{s}(x) dx$$

Dado que la función de corte deberá ser suave en las fronteras, las siguientes condiciones de frontera deben satisfacerse:

$$S(r_1) = 0$$

$$S'(r_1) = 0$$

$$S(r_c) = -F_{\alpha}(r_c)$$

$$S'(r_c) = -F'_{\alpha}(r_c)$$

Un polinomio de 3º grado de la forma

$$S(r) = A(r-r_1)^2 + B(r-r_1)^3$$

satisface correctamente estos requerimientos. Las constantes A y B vienen dadas por la condición de frontera en r_c :

$$A = -\frac{(\alpha + 4)r_{c} - (\alpha + 1)r_{1}}{r_{c}^{\alpha + 2}(r_{c} - r_{1})^{2}}$$
$$B = \frac{(\alpha + 3)r_{c} - (\alpha + 1)r_{1}}{r_{c}^{\alpha + 2}(r_{c} - r_{1})^{3}}$$

Así, la función de fuerza total resulta:

$$F_{s}(r) = \frac{1}{r^{\alpha+1}} + A(r-r_{1})^{2} + B(r-r_{1})^{3}$$

Y la función potencial será:

$$\Phi_{s}(r) = \frac{1}{r^{\alpha}} - \frac{A}{3}(r - r_{1})^{3} - \frac{B}{4}(r - r_{1})^{4} - C$$

Donde:

$$C = \frac{1}{r^{\alpha}} - \frac{A}{3} (r_c - r_1)^3 - \frac{B}{4} (r_c - r_1)^4$$

Cuando $r_1 = 0$, la función Coulómbica modificada de fuerza resulta:

$$F_{s}(r) = \frac{1}{r^{2}} - \frac{5r^{2}}{r_{c}^{4}} + \frac{4r^{3}}{r_{c}^{5}}$$

Ecuación idéntica a la función de fuerza parabólica que se recomienda para ser utilizada como función potencial de corto alcance en conjunción con el método de Poisson para la parte de largo alcance [83]. La función de potencial Coulómbico modificada es:

$$\Phi_{s}(r) = \frac{1}{r^{2}} - \frac{5}{3r_{c}} + \frac{5r^{3}}{3r_{c}^{4}} - \frac{4r^{4}}{r_{c}^{5}}$$


b. Potenciales enlazantes:

Las interacciones enlazantes se basan en una lista fija de átomos. Estas no están compuestas exclusivamente por interacciones de a pares, sino que incluyen también interacciones de 3 y 4 cuerpos. Están compuestas por las interacciones de estiramiento de enlace (2 cuerpos), de ángulo de enlace (3 cuerpos) y de ángulo diedro (4 cuerpos). Se utiliza un tipo especial de diedro (llamado diedro impropio) para que los átomos permanezcan en un determinado plano, o bien para prevenir la transición a una configuración de quiralidad opuesta (imagen especular).

i. Potencial de estiramiento de enlace:

Potencial armónico:

El estiramiento (o compresión) entre dos átomos i y j que se encuentran enlazados covalentemente puede representarse a través de un potencial armónico:

$$V_{b} = \frac{1}{2} k_{ij}^{b} \left(r_{ij} - b_{ij} \right)^{2}$$

Donde la fuerza viene dada por:

$$\boldsymbol{F}_{b}\left(\boldsymbol{r}_{ij}\right) = k_{ij}^{b}\left(\boldsymbol{r}_{ij} - \boldsymbol{b}_{ij}\right) \frac{\boldsymbol{r}_{ij}}{\boldsymbol{r}_{ij}}$$



Potencial de Morse:

Para ciertos sistemas donde se requiere un potencial no-armónico de enlace entre los átomos i y j, el potencial de Morse resulta una alternativa muy aceptable [84]. Este potencial difiere del potencial armónico en que posee un pozo potencial asimétrico además de tener fuerza de enlace cero a distancia infinita. Su forma funcional resulta:

$$V_{morse}\left(r_{ij}\right) = D_{ij}\left[1 - \exp\left(-\beta_{ij}\left(r_{ij} - b_{ij}\right)\right)\right]^{2}$$

Cuya fuerza correspondiente es:

$$\boldsymbol{F}_{morse}\left(\boldsymbol{r}_{ij}\right) = 2D_{ij}\beta_{ij}\boldsymbol{r}_{ij}\exp\left(-\beta_{ij}\left(\boldsymbol{r}_{ij}-\boldsymbol{b}_{ij}\right)\right) * \left[1-\exp\left(-\beta_{ij}\left(\boldsymbol{r}_{ij}-\boldsymbol{b}_{ij}\right)\right)\right]\frac{\boldsymbol{r}_{ij}}{\boldsymbol{r}_{ij}}$$

Donde D_{ij} representa la profundidad del pozo potencial en kJ/mol, β_{ij} define que tan agudo es el pozo potencial (en nm⁻¹) y b_{ij} es la distancia de equilibrio en nm.



ii. Potencial armónico de ángulo de enlace:

La vibración del ángulo de enlace entre un triplete de átomos i - j - kpuede representarse también a través de un potencial armónico sobre el ángulo θ_{iik} :

$$V_{a}\left(\theta_{ijk}\right) = \frac{1}{2} k_{ijk}^{\theta} \left(\theta_{ijk} - \theta_{ijk}^{0}\right)^{2}$$

Dado que la vibración del ángulo de enlace se representa con un potencial armónico, su forma es la misma que para el potencial armónico de estiramiento de enlace (Figura 7).



Las ecuaciones para la fuerza vienen dadas por la regla de la cadena:

$$F_{i} = -\frac{dV_{a}(\theta_{ijk})}{dr_{i}}$$

$$F_{k} = -\frac{dV_{a}(\theta_{ijk})}{dr_{k}} \quad \text{donde} \quad \theta_{ijk} = \arccos\frac{(r_{ij}, r_{kj})}{r_{ij}r_{kj}}$$

$$F_{i} = -F_{i} - F_{k}$$

La nomenclatura *i*, *j*, *k* representa la secuencia de enlace de los átomos *i*, *j*, *k* covalentemente enlazados. El átomo *j* se encuentra en el medio, los átomos *i* y *k* están en los extremos (ver Figura 7).



iii. Potencial para diedros impropios:

Los diedros impropios son utilizados para mantener la planaridad en aquellos grupos que así lo requieran (anillos aromáticos) además de mantener la quiralidad de aquellos grupos que cuenten con algún carbono asimétrico (casi todos los aminoácidos). En la Figura 8 se representan dichos casos esquemáticamente.

$$V_{di}\left(\xi_{ijkl}\right) = k_{\xi}\left(\xi_{ijkl} - \xi_0\right)^2$$

Esto también resulta ser un potencial armónico y se encuentra representado en la Figura 9. Cabe destacar que, dado que es un potencial armónico, la periodicidad no es tomada en cuenta, por lo tanto resulta recomendable definir un valor de ξ_0 lo más alejado posible de $\pm 180^{\circ}$.



iv. Potencial para diedros propios:

La interacción normal para los diedros propios puede ser abordada de dos maneras diferentes; la *función periódica* clásica o bien una función basada en una expansión de potencias del $\cos \phi$ (el *potencial de Ryckaert-Belleman*). Esta elección tiene consecuencias ya que, si se elige la función periódica, debe incluirse una interacción especial entre el 1^{er} y 4^{to} átomo de la tétrada que define al diedro.

Diedros propios: Potencial periódico

Los ángulos diedros propios de definen de acuerdo a la convención IUPAC/IUB, donde ϕ es el ángulo entre los planos *ijk* y *jkl*, con el cero correspondiente a la configuración *cis* (átomos *i* y *l* del mismo lado).

$$V_d\left(\phi_{ijkl}\right) = k_{\phi}\left(1 + \cos\left(n\phi - \phi_0\right)\right)$$



Diedros propios: Potencial de Ryckaert-Belleman

Para alcanos suele utilizarse el siguiente potencial (ver Figura 11)

$$V_{rb}\left(\phi_{ijkl}\right) = \sum_{n=0}^{5} C_n \left(\cos\left(\psi\right)\right)^n$$

Donde $\psi = \phi - 180^{\circ}$

Nota: puede lograrse la conversión de una convención a la otra multiplicando cada coeficiente C_n por $(-1)^n$.



(Nota: el uso de este potencial implica la exclusión de las interacciones 1-4 de Lennard-Jones, y ψ está definido de acuerdo a la "convención de polímeros" ($\psi_{trans} = 0$).)

c. Interacciones Especiales:

Las interacciones especiales son utilizadas para imponer restricciones sobre el sistema, ya sea para evitar desviaciones desastrosas (por Ej. para equilibrar el sistema una vez solvatado), o para incluir el conocimiento de datos experimentales (por Ej. datos provenientes de experimentos de RMN). En cualquier caso éstas no forman parte del campo de fuerzas y la confiabilidad de sus parámetros no resulta de relevancia.

i. Restricciones de posición

Estas se utilizan para restringir partículas a una posición fija de referencia $R_{\rm r}$. Pueden utilizarse durante le equilibrado del sistema de manera

de evitar el reacomodamiento drástico de zonas críticas (para restringir el movimiento de una proteína sujeta a fuerzas importantes debidas a un solvente que aún no se encuentra equilibrado). Otra aplicación es la de restringir partículas en una cáscara alrededor de una región que se desea simular en detalle, mientras que la cáscara solo es aproximada ya que carece de las interacciones apropiadas por parte de las partículas faltantes que están por fuera de la misma. Las restricciones mantendrán la integridad del interior. Para casquetes esféricos es recomendable utilizar una constante de fuerza para la restricción que dependa del radio, creciendo desde cero para el interior del casquete hasta un valor elevado en el límite exterior del mismo.

Se utiliza la siguiente forma general:

$$V_{rp}\left(\boldsymbol{r}_{i}\right) = \frac{1}{2}k_{rp}\left|\boldsymbol{r}_{i}-\boldsymbol{R}_{i}\right|^{2}$$

Este potencial se grafica en la Figura 12.

El potencial puede ser escrito sin pérdida de generalidad también como:

$$V_{rp}(\mathbf{r}_{i}) = \frac{1}{2} \left[k_{rp}^{x} (x_{i} - X_{i})^{2} \, \hat{\mathbf{x}} + k_{rp}^{y} (y_{i} - Y_{i})^{2} \, \hat{\mathbf{y}} + k_{rp}^{z} (z_{i} - Z_{i})^{2} \, \hat{\mathbf{z}} \right]$$

Donde las fuerzas son:

$$F_{i}^{x} = k_{rp}^{x} (x_{i} - X_{i})^{2}$$
$$F_{i}^{y} = k_{rp}^{y} (y_{i} - Y_{i})^{2}$$
$$F_{i}^{z} = k_{rp}^{z} (z_{i} - Z_{i})^{2}$$

Las restricciones de posición pueden "apagarse" o "prenderse" para cada dimensión espacial utilizando tres constantes de fuerza diferentes; esto significa que los átomos pueden restringirse armónicamente a un punto, un plano o una recta.



ii. Restricciones de ángulo

Estas se utilizan para restringir el ángulo entre dos pares de partículas o entre un par de partículas y el eje Z. La forma funcional es similar a la de un diedro propio. Para dos pares de átomos:

$$V_{ra}(\mathbf{r}_{i}, \mathbf{r}_{j}, \mathbf{r}_{k}, \mathbf{r}_{l}) = k_{ra} \left(1 - \cos \left(n \left(\phi - \phi_{0} \right) \right) \right)$$

Donde $\phi = \arccos \left(\frac{\mathbf{r}_{j} - \mathbf{r}_{i}}{\|\mathbf{r}_{j} - \mathbf{r}_{i}\|} \cdot \frac{\mathbf{r}_{l} - \mathbf{r}_{k}}{\|\mathbf{r}_{l} - \mathbf{r}_{k}\|} \right)$

Para un par de átomos y el eje Z:

$$V_{ra}\left(\boldsymbol{r}_{i},\boldsymbol{r}_{j},\boldsymbol{r}_{k},\boldsymbol{r}_{l}\right) = k_{ra}\left(1 - \cos\left(n\left(\phi - \phi_{0}\right)\right)\right)$$

Donde $\phi = \arccos\left(\frac{\boldsymbol{r}_{j} - \boldsymbol{r}_{i}}{\|\boldsymbol{r}_{j} - \boldsymbol{r}_{i}\|} \cdot \begin{pmatrix} 0\\0\\1 \end{pmatrix}\right)$

iii. Restricciones de distancia

Las restricciones de distancia agregan una penalidad al potencial entre pares previamente especificados de átomos, cuando la distancia entre ellos excede un valor límite. Esto se utiliza normalmente para imponer restricciones experimentales sobre el movimiento de un sistema, como en el caso de experimentos de resonancia magnética nuclear (RMN). De este modo la Dinámica Molecular puede utilizarse para el proceso de refinamiento estructural de los datos provenientes de experimentos de RMN.

La forma del potencial para las restricciones de distancia es cuadrática por debajo de un valor especificado, como así también entre los dos valores superiores especificados, y lineal por encima del valor del límite superior mayor.

$$V_{rd}(r_{ij}) = \begin{cases} \frac{1}{2} k_{rd} (r_{ij} - r_0)^2 & para & r_{ij} < r_0 \\ 0 & para & r_0 \le r_{ij} < r_1 \\ \frac{1}{2} k_{rd} (r_{ij} - r_1)^2 & para & r_1 \le r_{ij} < r_2 \\ \frac{1}{2} k_{rd} (r_2 - r_1) (2r_{ij} - r_2 - r_1) & para & r_2 \le r_{ij} \end{cases}$$

Dicho potencial está graficado en la Figura 13.

Las fuerzas son

$$F_{i}(\mathbf{r}_{ij}) = \begin{cases} -k_{rd}(r_{ij} - r_{0})\frac{\mathbf{r}_{ij}}{r_{ij}} & para & r_{ij} < r_{0} \\ 0 & para & r_{0} \leq r_{ij} < r_{1} \\ -k_{rd}(r_{ij} - r_{1})\frac{\mathbf{r}_{ij}}{r_{ij}} & para & r_{1} \leq r_{ij} < r_{2} \\ -k_{rd}(r_{2} - r_{1})\frac{\mathbf{r}_{ij}}{r_{ij}} & para & r_{2} \leq r_{ij} \end{cases}$$



4. Generalidades del cómputo de fuerzas:

En la sección 2 del presente capítulo se presenta un esquema global de flujo para la Dinámica Molecular. Toda simulación de DM requiere como entrada un conjunto de coordenadas iniciales y, opcionalmente, velocidades iniciales para todas las partículas involucradas. Las primeras suelen provenir de datos experimentales (cristalografía de rayos X, Resonancia Magnética Nuclear, etc.), mientras que las últimas suelen ser generadas inicialmente a partir de una distribución de Maxwell.

a. Condiciones iniciales:

La topología y el campo de fuerzas

La topología, incluyendo la descripción completa del campo de fuerzas, debe ser provista. Estos elementos fueron descriptos en detalle en la sección anterior. Toda esta información es estática y nunca resulta modificada durante la corrida.

Coordenadas y velocidades

Así, antes de comenzar una corrida se requieren las coordenadas y velocidades de las partículas, como así también las dimensiones de la caja de simulación. El tamaño de la caja queda determinado por tres vectores (nueve números) b_1 , b_2 , b_3 , los cuales representan los tres vectores que definen la caja periódica.

Si la corrida comienza en $t = t_0$, las coordenadas en $t = t_0$ deben ser conocidas. El algoritmo "leap-frog" o de salto de rana, utilizado para actualizar el paso de integración Δt (ver sección X), requiere que se conozcan las velocidades en $t = t_0 + \frac{\Delta t}{2}$. Si las velocidades $v_i, i = 1...3N$ no son conocidas, estas pueden generarse a partir de una distribución de Maxwell para una dada temperatura T:

$$p(v_i) = \sqrt{\frac{m_i}{2\pi kT}} \exp\left(-\frac{m_i v_i^2}{2kT}\right)$$

Donde k es la constante de Boltzmann.



Movimiento del centro de masas

La velocidad del centro de masas normalmente se lleva a cero en cada paso. Normalmente no existe ninguna fuerza neta externa actuando sobre el sistema y le velocidad del centro de masas debería permanecer constante. En la práctica, de cualquier manera, el algoritmo de actualización desarrolla un cambio lento en la velocidad del centro de masas, y por lo tanto en la energía cinética total del sistema, especialmente cuando se acopla un baño térmico al mismo. Si estos cambios no se amortiguan, eventualmente se desarrolla un movimiento apreciable del centro de masas para corridas largas, lo cual resultará en una significativa mala interpretación de la temperatura. Lo mismo podría suceder con el movimiento de rotación global, pero solamente cuando se simula un cluster de partículas aislado. En sistemas periódicos con cajas llenas de partículas, el movimiento rotacional global se encuentra acoplado a otros grados de libertad y por lo tanto no ocasiona problemas.

b. Búsqueda de vecinos:

Como se mencionó en la sección 3, las fuerzas son generadas, ya sea por listas fijas (estáticas) de átomos, o bien por listas dinámicas. Estas últimas son siempre debidas a las interacciones no-enlazantes entre un par cualesquiera de partículas.

Generación de listas de pares

Las fuerzas entre pares no-enlazantes necesitan calcularse solamente para aquellos pares i,j para los cuales la distancia r_i entre la partícula i y la imagen más cercana de la partícula j sea menor que un dado radio de corte R_c . Algunos de los pares de partículas que satisfacen este criterio deben ser excluidos de la lista, dado que su interacción ha sido totalmente tomada en cuenta por las interacciones enlazantes.

Para realizar dicha lista de vecinos deben hallarse todas las partículas que se encuentran cerca (dentro del radio de corte) de una dada partícula. Esta búsqueda involucra condiciones periódicas de contorno y por lo tanto la determinación de la imagen más cercana. Sin condiciones periódicas de contorno debe usarse un algoritmo simple $O(N^2)$. Con condiciones periódicas de contorno puede aplicarse una búsqueda tipo grilla, la cual resulta O(N), y por lo tanto menos costosa en términos computacionales.

c. Cómputo de fuerzas:

Energía potencial

Cuando se computan las fuerzas, también se computa la energía potencial para cada término de interacción. Por lo tanto la energía potencial total resulta de la suma de varias contribuciones, tales como Lennard-Jones, Coulomb, y las interacciones enlazantes.

Energía cinética y temperatura

La temperatura viene dada por la energía cinética total del sistema de N partículas:

$$E_{cin} = \frac{1}{2} \sum_{i=1}^{N} m_i v_i^2$$

De aquí, la temperatura absoluta T puede ser computada utilizando:

$$\frac{1}{2}N_{gl}kT = E_{cin}$$

Donde k es la constante de Boltzmann y N_{gl} es el número total de grados de libertad, los cuales pueden calcularse a partir de:

$$N_{gl} = 3N - N_c - N_{com}$$

Aquí N_c es el número de restricciones impuestas sobre el sistema. Cuando se realiza una simulación por Dinámica Molecular deben ser restados adicionalmente $N_{com} = 3$ grados de libertad, dado que las tres velocidades del centro de masas son constantes del movimiento, y normalmente se las hace valer cero. Cuando la simulación es realizada en vacío, puede restarse también la rotación alrededor del centro de masas, en este caso $N_{com} = 6$.

Cuando se utiliza más de un grupo de partículas acoplados independientemente al baño de temperatura, el número de grados de libertad para el grupo *i* es:

$$N_{gl}^{i} = (3N^{i} - N_{c}^{i})\frac{3N - N_{c} - N_{com}}{3N - N_{c}}$$

La energía cinética puede ser escrita también como un tensor, lo cual resulta necesario frecuentemente para el cálculo de la presión, o bien para sistemas en donde se imponen fuerzas de corte:

$$\mathbf{E}_{cin} = \frac{1}{2} \sum_{i=1}^{N} m_i \mathbf{v}_i \otimes \mathbf{v}_i$$

La presión y el virial

El tensor de la presión P se calcula como la diferencia entre la energía cinética \mathbf{E}_{cin} y el virial Ξ :

$$\mathbf{P} = \frac{2}{V} \left(\mathbf{E}_{cin} - \Xi \right)$$

Donde V es el volumen de la caja de simulación. La presión escalar P, la cual puede utilizarse para el acoplamiento al baño de presión para sistemas isotrópicos, puede calcularse según:

$$P = \text{traza}(\mathbf{P})/3$$

El tensor Ξ del virial se define como:

$$\Xi = -\frac{1}{2} \sum_{i < j} \boldsymbol{r}_{ij} \otimes \boldsymbol{F}_{ij}$$

d. Actualización de la configuración:

El algoritmo utilizado para la integración del las ecuaciones de movimiento es el denominado "leap-frog" o de salto de rana [85]. Este algoritmo utiliza posiciones r a tiempo t y velocidades v a tiempo $t - \frac{\Delta t}{2}$; actualiza tanto posiciones como velocidades utilizando las fuerzas F(t) determinadas por las posiciones al tiempo t:

$$\mathbf{v}\left(t+\frac{\Delta t}{2}\right) = \mathbf{v}\left(t-\frac{\Delta t}{2}\right) + \frac{\mathbf{F}(t)}{m}\Delta t$$
$$\mathbf{r}\left(t+\Delta t\right) = \mathbf{r}\left(t\right) + \mathbf{v}\left(t+\frac{\Delta t}{2}\right)\Delta t$$

El algoritmo puede visualizarse en la Figura 15. Es equivalente al algoritmo de Verlet [86]:

$$\boldsymbol{r}(t+\Delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t-\Delta t) + \frac{\boldsymbol{F}(t)}{m}\Delta t^{2} + O(\Delta t^{4})$$

Este algoritmo es de 3^{er} orden en r y es reversible en el tiempo. Las ventajas de este algoritmo y su comparación con otros algoritmos de integración temporal pueden encontrarse en la referencia [87].

Las ecuaciones de movimiento se modifican para incorporar el acoplamiento a los baños de presión y temperatura, además de ser extendidas para incluir la conservación de las restricciones, todo lo cual se describe a continuación.



e. Acoplamiento térmico:

Por varios motivos (deriva durante el equilibrado del sistema, deriva como resultado del truncado de fuerzas y errores de integración, calentamiento debido a fuerzas externas o de fricción), es necesario controlar la temperatura del sistema. Para esto puede utilizarse el esquema de acoplamiento débil de Berendsen [88] o el esquema de ensamble extendido de Nosé-Hoover [89,90].

Acoplamiento térmico de Berendsen

El algoritmo de Berendsen reproduce un acoplamiento débil a un baño térmico externo de una dada temperatura T_0 , con cinética de primer orden. El efecto de la aplicación de este algoritmo es el de corregir una desviación térmica de T_0 del sistema según:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}$$

Lo cual significa que una dada desviación térmica decae exponencialmente con una constante de tiempo τ . Este método de acoplamiento tiene la ventaja que la fuerza con la cual se encuentra acoplado el baño puede variarse, adaptándose a las necesidades del sistema: para llevar un sistema al equilibrio puede utilizarse una constante pequeña (0.01ps), mientras que para lograr corridas confiables de un sistema en equilibrio puede utilizarse una constante mucho mayor (0.5ps) en cuyo caso prácticamente no se afecta la dinámica conservativa [91].

El flujo de calor desde o hacia el sistema es realizado escalando a cada paso las velocidades de cada partícula mediante un factor $\lambda(t)$, dado por:

$$\lambda(t) = \left[1 + \frac{\Delta t}{\tau_T} \left\{\frac{T_0}{T(t - \Delta t/2)} - 1\right\}\right]^{1/2}$$

El valor del parámetro τ_{τ} es cercano al de la constante de acoplamiento τ , aunque no exactamente igual:

$$\tau = \frac{2C_V \tau_T}{N_{gl}k}$$

Donde C_{ν} es la capacidad calorífica total del sistema, k es la constante de Boltzmann y N_{gl} es el número total de grados de libertad. La razón por la cual $\tau \neq \tau_T$ es que el cambio en la energía cinética causado por el escalado de las velocidades es parcialmente redistribuido entre las energías cinética y potencial, y por lo tanto el cambio en la temperatura resulta menos que el escalado de la energía. En la práctica, el valor de la relación τ/τ_T va desde 1 (gas) a 2 (sólido armónico) a 3 (agua). El término "constante temporal de acoplamiento térmico" se utiliza para el parámetro τ_T .

El algoritmo de Berendsen resulta estable hasta valores de $\tau_T \approx \Delta t$

Acoplamiento térmico de Nosé-Hoover

El algoritmo de Berendsen resulta extremadamente eficiente para relajar un sistema hacia la temperatura de equilibrio perseguida, pero una vez que el mismo haya alcanzado el equilibrio puede resultar importante lograr cubrir un ensamble canónico. Esto lamentablemente no resulta ser el caso del esquema de acoplamiento débil, aunque la diferencia resulta despreciable para la mayoría de los casos [91].

La aproximación de ensamble extendido fue originalmente propuesta por Nosé [89] y luego modificada por Hoover [90]. El Hamiltoniano del sistema es extendido merced a la introducción de un reservorio térmico y un término de fricción en las ecuaciones de movimiento. La fuerza de fricción es proporcional al producto entre la velocidad de cada partícula y un parámetro de fricción ξ . Este parámetro de fricción (o variable de "baño térmico") es una cantidad completamente dinámica, la cual posee su propia ecuación de movimiento; la derivada respecto del tiempo se calcula a partir de la diferencia entre la energía cinética del sistema y la temperatura de referencia.

En la formulación de Hoover, las ecuaciones de movimiento de las partículas presentadas en la sección 2 del presente capítulo son reemplazadas por:

$$\frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \frac{\mathbf{F}_i}{m_i} - \xi \frac{\partial \mathbf{r}_i}{\partial t}$$

Donde la ecuación de movimiento para el parámetro ξ del baño térmico es:

$$\frac{d\xi}{dt} = \frac{1}{Q} \left(T - T_0 \right)$$

La temperatura de referencia se denota por T_0 mientras que *T* representa la temperatura instantánea del sistema. La fuerza del acoplamiento viene determinada por la constante Q (normalmente llamada "parámetro de masa" del reservorio) en combinación con la temperatura de referencia.

De cualquier manera resulta importante tener presente la diferencia entre el esquema de acoplamiento débil y el algoritmo de Nosé-Hoover: Al utilizar el acoplamiento débil se obtiene un sistema de *relajación exponencial* fuertemente suavizado, mientras que el uso del algoritmo de Nosé-Hoover produce una relajación oscilatoria.

Acoplamiento térmico por grupos

El motivo por el cual se introduce este esquema se debe a que el intercambio de energía entre los distintos componentes del sistema no es perfecto debido a diferentes efectos incluyendo los radios de corte, etc. Si todo un sistema, compuesto (por ejemplo) de una proteína hidratada, es acoplado al mismo baño, el agua (la cual experimenta el mayor ruido debido al radio de corte) tenderá a calentarse, mientras que la proteína tenderá a enfriarse. Diferencias de hasta 100K pueden frecuentemente obtenerse. Con el uso de algún método (PME) para las fuerzas electrostáticas de largo alcance, esta diferencia puede reducirse mucho, pero sigue sin ser despreciable.

f. Acoplamiento de presión:

Análogamente al acoplamiento térmico, el sistema puede ser acoplado también a un "baño" de presión.

Para ajustar el valor de la presión del sistema puede utilizarse el algoritmo de Berendsen [88] o el esquema de ensamble extendido de Parinello-Raman. Ambos pueden combinarse con cualquiera de los algoritmos para el acoplamiento térmico descriptos con anterioridad.

Baróstato de Berendsen

El algoritmo de Berendsen re-escala las coordenadas y los vectores de la caja a cada paso de la integración con la matriz μ , lo cual tiene un efecto de una cinética de relajación de primer orden sobre la presión, hacia un valor dado de referencia \mathbf{P}_0 de presión:

$$\frac{d\mathbf{P}}{dt} = \frac{\mathbf{P}_0 - \mathbf{P}}{\tau_p}$$

La matriz μ de escalado viene dada por:

$$\mu_{ij} = \delta_{ij} - \frac{\Delta t}{3\tau_p} \beta_{ij} \left\{ P_{0ij} - P_{ij}(t) \right\}$$

Aquí β es la compresibilidad isotérmica del sistema. En la mayoría de los casos esto será una matriz diagonal, con elementos iguales sobre la diagonal, cuyo valor resulta generalmente desconocido. Alcanza con tomar un valor aproximado del mismo dado que el valor de β solamente tiene influencia sobre la constante temporal no crítica de la relajación de la presión sin afectar el valor promedio de la presión.

Cuando el escalado es completamente anisotrópico, el sistema debe ser rotado de manera de cumplimentar con las restricciones de la caja. Esta rotación puede aproximarse en el primer orden del escalado, el cual es normalmente menor que 10^{-4} . La matriz de escalado μ^{*} es:

$$\mu' = \begin{pmatrix} \mu_{xx} & \mu_{xy} + \mu_{yx} & \mu_{xz} + \mu_{zx} \\ 0 & \mu_{yy} & \mu_{yz} + \mu_{zy} \\ 0 & 0 & \mu_{zz} \end{pmatrix}$$

Las velocidades no se rotan ni se escalan.

El escalado de Berendsen puede también realizarse isotrópicamente, lo cual significa que puede utilizarse una matriz diagonal de tamaño traza $(\mathbf{P})/3$ en lugar de la matriz \mathbf{P} .

Baróstato de Parinello-Raman

Para los casos donde las fluctuaciones de la presión o el volumen son importantes *per se* (en el cálculo de propiedades termodinámicas) puede resultar problemático, al menos teóricamente, que el ensamble exacto no esté correctamente definido por el esquema del acoplamiento débil.

Para resolver esto pueden realizarse simulaciones utilizando la proximación de Parinello-Raman [92, 93], el cual resulta similar al acoplamiento térmico de Nosé-Hoover.

Con el baróstato de Parinello-Raman los vectores de la caja se representan a través de la matriz \boldsymbol{b} y obedecen la siguiente ecuación matricial de movimiento:

$$\frac{d\boldsymbol{b}^2}{dt^2} = V \boldsymbol{W}^{-1} \boldsymbol{b}^{r-1} \left(\boldsymbol{P} - \boldsymbol{P}_{ref} \right)$$

El volumen de la caja está dado por V, y W es el parámetro matricial que determina la fuerza del acoplamiento. Las matrices P y P_{ref} representan las presiones instantánea y de referencia respectivamente.

Análogamente al caso del acoplamiento de Nosé-Hoover, las ecuaciones de movimiento de las partículas deben también cambiarse.

En la mayoría de los casos resultaría razonable combinar el baróstato de Parinello-Raman con el termóstato de Nosé-Hoover.

El parámetro matricial (inverso) de masa W^{-1} determina la fuerza del acoplamiento, además del modo en el cual puede ser deformada la caja.

Se debe contar además con los valores aproximados para la compresibilidad isotérmica β y la constante temporal de acoplamiento del baróstato $\tau_p(L$ es el mayor elemento de la matriz de la caja):

$$\left(\boldsymbol{W}^{-1}\right)_{ij} = \frac{4\pi^2\beta_{ij}}{3\tau_p L}$$

5. Algoritmos de restricciones topológicas:

Las restricciones topológicas pueden resolverse utilizando diversos algoritmos desarrollados para este fin. Haremos una breve reseña de los más utilizados y los más eficientes en términos de costo computacional.

a. SHAKE

El algoritmo SHAKE [94] cambia un grupo de coordenadas norestringidas r' por otro grupo de coordenadas r'' que cumplimentan con una serie de restricciones de distancia, utilizando de referencia el conjunto de coordenadas r:

SHAKE
$$(r' \rightarrow r''; r)$$

Esta acción es consistente con la resolución de un conjunto de multiplicadores de Lagrange en las ecuaciones de movimiento restringidas. En el algoritmo SHAKE se necesita definir un parámetro de *tolerancia* (TOL); este continuará con las iteraciones hasta que se satisfagan todas las restricciones dentro del valor de *tolerancia* dada.

Asumiendo que las ecuaciones de movimiento deben cumplir con un conjunto de restricciones holonómicas K, expresadas por:

$$\sigma_k(\mathbf{r}_1...\mathbf{r}_N) = 0; \quad k = 1...K$$

Entonces las fuerzas quedan definidas por:

$$-\frac{\partial}{\partial \boldsymbol{r}_{i}}\left(V+\sum_{k=1}^{K}\lambda_{k}\boldsymbol{\sigma}_{k}\right)$$

Donde λ_k son los multiplicadores de Lagrange que deben ser resueltos para satisfacer las ecuaciones de restricción. La segunda parte de esta suma determina las fuerzas de restricción G_i , definidas por:

$$\boldsymbol{G}_{i} = -\sum_{k=1}^{K} \lambda_{k} \frac{\partial \boldsymbol{\sigma}_{k}}{\partial \boldsymbol{r}_{i}}$$

El desplazamiento debido a las fuerzas de restricción en los algoritmos de integración "leap-frog" o verlet es igual a $(G_i/m_i)(\Delta t)^2$. La resolución de los multiplicadores de Lagrange (y por lo tanto de los desplazamientos) requiere de resolver un conjunto de ecuaciones acopladas de segundo grado. Este conjunto de ecuaciones es resuelto iterativamente en el algoritmo SHAKE.

Dado que en simulaciones de sistemas hidratados el solvente suele representar más del 80% del mismo, existen ciertas alternativas para el uso de este algoritmo.

Para el caso de modelos de agua rígidos puede utilizarse una versión optimizada (y simplificada) de este algoritmo, denominado SETTLE [95], el cual consiste en una aplicación de una solución analítica de SHAKE.

b. LINCS

El algoritmo LINCS

LINCS es un algoritmo que reconstituye enlaces a su longitud correcta luego de una actualización no-restringida [96]. El método es noiterativo, dado que siempre se resuelve en dos pasos. Aunque LINCS está basado en matrice, no precisa de la resolución de multiplicaciones entre matrices. El método es más rápido y estable que SHAKE, pero sólo puede utilizarse con restricciones de longitud de enlace y restricciones de ángulos de enlace aislados, tales como el ángulo del protón en el enlace OH.



Las fórmulas de LINCS

Consideremos un sistema de N partículas, con posiciones dadas por un vector 3N, denominado r(t). Para la Dinámica Molecular las ecuaciones de movimiento vienen dadas por las ecuaciones de Newton:

$$\frac{\partial^2 \boldsymbol{r}}{\partial t^2} = \boldsymbol{M}^{-1} \boldsymbol{F}$$

Donde F es el vector 3N de la fuerza y M es la matriz diagonal $3N \times 3N$, conteniendo las masas de las partículas. El sistema se encuentra restringido topológicamente por K ecuaciones de restricción, dependientes del tiempo

$$g_i(\mathbf{r}) = |\mathbf{r}_{i_1} - \mathbf{r}_{i_2}| - d_i = 0$$
 donde $i = 1, ..., K$

En el esquema de integración numérica LINCS (al igual que SHAKE) es aplicado luego de una actualización no-restringida. El algoritmo trabaja en dos pasos (ver Figura 16). Durante el primer paso se llevan a cero las proyecciones del nuevo enlace sobre el viejo. En el segundo paso se aplica una corrección por estiramiento de enlace debido a la rotación. Puede verse una derivación completa del algoritmo en la referencia [96].

6. Algoritmo completo de actualización:

EL ALGORITMO DE ACTUALIZACION

Dadas: Posiciones r de todos los átomos a tiempo tVelocidades v a tiempo $t - \frac{\Delta t}{2}$ Aceleraciones F/m sobre todos los átomos a tiempo t. (Las fuerzas se computan sin tomar en cuenta ninguna restricción) La energía cinética total y el virial. 11 1. Cómputo de los factores λ y μ de escalado (Ecuaciones X e Y) 11 2. Actualización y escalado de velocidades: $v' = \lambda (v + a\Delta t)$ 11 3. Cálculo de las nuevas coordenadas no-restringidas: $r' = r + v' \Delta t$ 4. Aplicación de algoritmo de restricción a las coordenadas: restringir $(\mathbf{r'} \rightarrow \mathbf{r''}; \mathbf{r})$ 11 5. Corrección de velocidades según restricciones: $v = (r'' - r) / \Delta t$][6. Escalado de coordenadas y caja: $r = \mu r''$; $b = \mu b$

Capítulo 3

LISOZIMA EN H2O A 3KBAR:

1. Introducción:

En el presente capítulo se presentan los resultados del estudio por Dinámica Molecular de los efectos de la presión sobre la lisozima en solución, además del correspondiente análisis de los mismos. Dicho sistema fue estudiado a las presiones de 1bar (baja presión) y 3kbar (alta presión).

Los resultados guardan un importante acuerdo con los datos experimentales disponibles, lo cual permite el análisis de otros aspectos del efecto de la presión sobre esta proteína en solución.

Los estudios de movilidad muestran que aunque la misma se ve en general reducida para la mayoría de los residuos bajo presión, esto no resulta cierto para algunos residuos en particular.

Del análisis de la estructura secundaria a lo largo de las trayectorias se observa que la conformación de la proteína bajo presión resulta más estable, lo cual sugiere que la alta presión actúa como "selector de confórmeros" sobre esta proteína.

El comportamiento de la superficie accesible al solvente muestra una clara inversión de la relación de superficies accesibles hidrofílica a hidrofóbica, lo cual indica por lo tanto que la interacción hidrofóbica resulta considerablemente menor con el aumento de la presión.

2. Materiales y métodos:

Las simulaciones de Dinámica Molecular se realizaron utilizando el paquete GROMACS 3.2.1 [97,98]. Se utilizó el campo de fuerzas para todos los átomos explícitos (ff2gmx) [99,100,101,102,103,104], tanto para los procedimientos de minimización de energía y equilibrado del sistema, como para todos los pasos de simulación por Dinámica Molecular.

Se mantuvieron las restricciones topológicas para las longitudes de enlace utilizando el algoritmo LINCS [96]. Las restricciones topológicas de las moléculas de agua se resolvieron utilizando el algoritmo SETTLE [95]. Para el cálculo de las fuerzas electrostáticas se utilizó el método de Campo de Reacción [105]. Las interacciones de Lennard-Jones fueron calculadas explícitamente dentro de un radio de corte de 1,4nm. Para todas las simulaciones de este sistema se utilizó un cluster de 4 procesadores Pentium III, corriendo bajo GNU/Linux, utilizando para las visualizaciones de trayectorias, estructuras y representaciones gráficas el software de distribución libre Visual Molecular Dynamics (VMD), Swiss PDB Viewer (SPDBV) y XMGrace respectivamente.

Como estructuras de partida para la lisozima se utilizaron las resueltas por RMN-2D (Resonancia Magnética Nuclear bidimensional) de alta y baja presión [106] (códigos de PDB 1GXX y 1GXV respectivamente) publicadas y disponibles en el "Protein Data Bank".

La topología para la lisozima se generó utilizando la herramienta PDB2GMX, con estados de protonación estándar para todos los aminoácidos (pH 7).

Se utilizó el modelo de agua SPC/E [107], tanto para el sistema de alta presión como para el de baja presión.

El sistema inicial consistió de una caja cúbica de dimensiones X=Y=Z=4,9588nm, con un volumen de 121,9354nm, conteniendo una molécula de lisozima y 3497 moléculas de agua.

Sobre este sistema se realizó una primera minimización de energía por el método de "steepest descent" o pendiente pronunciada, la cual convergió al valor de la precisión del cómputo. Luego se realizó una segunda minimización de energía por el método de gradientes conjugados, el cual logró converger dentro de los primeros 20 ciclos.

Luego de la minimización de energía se permitió que el solvente se relajara alrededor de la proteína, realizando una breve simulación por Dinámica Molecular (10ps) sobre el sistema, aplicando restricciones de posición sobre todos los átomos de la proteína.

Desde el comienzo el sistema fue acoplado a un baño débil de temperatura como de presión, de manera de trabajar en el ensamble isotérmico e isobárico [88], a la temperatura T=315K y presión P=1bar.

A continuación se permitió que todo el sistema se relajara (se quitaron las restricciones de posición de todos los átomos) durante 200ps, manteniendo el acoplamiento del sistema al termóstato y al baróstato.

En este punto el sistema fue dividido en dos. La primera copia fue sometida lentamente a presiones hidrostáticas ascendentes, mientras que la segunda se mantuvo a presión atmosférica.

Al primer sistema se le aumentó la presión en 22 pasos concatenados, durante un tiempo total de 4,4ns según se describe a continuación: 1-10bar, durante 200ps; 10-100bar, durante 200ps; 100-1000bar (incrementos de 100bar durante 200ps cada uno); 1000-3000bar (incrementos de 200bar durante 200ps cada uno). Al segundo sistema se le permitió evolucionar durante un tiempo equivalente (4,4ns) pero a presión de 1bar.

Habiendo alcanzado este punto se permitió que ambos sistemas evolucionaran durante un período adicional de 8ns, de modo de producir puntos de partida equilibrados y comparables para ambos sistemas.

Una vez alcanzado este punto se realizó una simulación de 24ns de longitud, tanto para el sistema a 1bar como para el sistema a 3kbar.

Todos los resultados presentados a continuación corresponden a estas simulaciones de 24ns de duración.

3. Resultados y discusión:

a. Comparación de las estructuras promedio de alta y baja presión

El siguiente análisis fue realizado calculando la estructura promedio de la proteína, previo ajuste progresivo por cuadrados mínimos de cada estructura sobre la anterior, utilizando las posiciones atómicas registradas durante los 24ns totales para ambos sistemas.

La Figura 17 muestra la diferencia de distancias para cada carbono- α con el resto, entre las estructuras promedio a baja y alta presión. Las diferencias positivas indican que los átomos se han acercado entre sí en la estructura promedio de alta presión, comparado con la estructura promedio a baja presión.

La primera observación a destacar es que la mayoría de los carbonos- α se acercan al resto de sus pares (tonos rojos), lo cual implica que en general la estructura se vuelve más compacta.

Por otro lado, notablemente, algunos carbonos- α se alejan del resto (tonos azules), lo cual implica que, aunque se observa una compresión general de la estructura, algunos residuos específicos se mueven hacia fuera.

Esta observación resulta cualitativamente concordante con los datos experimentales ofrecidos por RMN-2D [106].



b. Screening estructural mediante el cálculo de la raíz de la desviación cuadrática media para alta y baja presión

De modo de poder encontrar similitudes y/o diferencias estructurales durante las simulaciones, se calculó la matriz para la raíz de la desviación cuadrática media posicional de los carbonos- α , de cada estructura contra todas las otras registradas (luego de realizar un ajuste estructural por cuadrados mínimos para cada par calculado), para las trayectorias completas de baja y alta presión.

Los resultados se presentan en la Figura 18, en formato de código de escala de grises. Blanco corresponde a RMSD=0nm, mientras que negro corresponde a RMSD≥0,248nm.

Como primera observación, puede verse que el sistema de alta presión muestra una homogeneidad estructural mucho mayor, a lo largo de toda la corrida, que la correspondiente para el sistema a baja presión.

Este resultado también se encuentra en muy buena concordancia con las observaciones experimentales disponibles respecto a los efectos de la presión sobre las transiciones estructurales de la lisozima [108].



Cabe destacar también que existen, casi homogéneamente distribuidas a lo largo de toda la corrida de baja presión, algunas estructuras compatibles con los confórmeros más homogéneos del sistema de alta presión.

c. Análisis de la evolución de la estructura secundaria

En esta sección se presentan los resultados del análisis de la evolución de la estructura secundaria de la lisozima para los sistemas de alta y baja presión.

Los resultados presentados fueron calculados utilizando las convenciones establecidas en el "Dictionary of Secondary Structure for Proteins" (DSSP) [69].

En las Figuras 19a y 19b se presenta la evolución temporal del recuento total para los diferentes tipos de estructura secundaria presentes en la lisozima, a baja y alta presión respectivamente.



En las Figuras 20a y 20b se presenta información detallada respecto de cómo participa cada residuo dentro de la estructura secundaria de la proteína, como así también la evolución de dicha propiedad en el tiempo, a baja y alta presión respectivamente.



Una primera comparación de estos gráficos confirma la impresión inicial, obtenida merced al screening realizado a través de la matriz de RMSD, que para el caso de la lisozima solvatada las fluctuaciones en la estructura (secundaria) se ven notablemente reducidas con el aumento de la presión, incrementando por lo tanto la homogeneidad estructural del sistema al someterlo a la presión de 3kbar.

De igual modo puede observarse que el recuento de la estructura total (calculado como la suma helice- α + hoja- β + giro) se incrementa significativamente, además de estabilizarse, con el aumento de la presión.

Por otro lado, un estudio más detallado de las Figuras X e Y revelan el hecho de que la estructura prevalente durante los primeros 2,2ns de la corrida de baja presión resulta muy compatible con la estructura promedio que resulta estabilizada durante casi toda la corrida a 3kbar.

Esto resulta más notorio cuando ponemos el foco sobre las estructuras secundarias β .

Esta última observación resulta altamente compatible con el rol de "Selector de Confórmeros", propuesto recientemente para la presión [109].

En suma, podría decirse que la información complementaria y más detallada presentada en esta sección confirma las conclusiones preliminares de

la sección anterior, en cuanto a la diferencia en la variabilidad estructural entre las corridas de baja y alta presión.



Puede claramente concluirse a partir de un primer análisis de las figuras de esta sección que muchos residuos muestran una mayor variabilidad en el tipo de estructuras que adoptan a lo largo de la corrida a 1bar, comparados con su comportamiento a 3kbar.

Podemos observar también que, aunque la mayoría de los residuos mantienen la estructura que tenían a 1bar, algunos otros por el contrario presentan cambios significativos al ser sometidos a la presión de 3kbar. Buenos ejemplos de este comportamiento lo constituyen los residuos 2-3, 20, 23 y 39, los cuales adoptan una estructura permanente de hoja-β a la presión de 3kbar, cuando esta de ningún modo resulta su estructura prevalente a 1bar.

Sin embargo cabe destacar que, aunque en una proporción a veces muy reducida, para cada residuo la estructura prevalente a 3kbar también se encuentra presente durante la corrida a 1 bar.

Además, al analizar la evolución de la estructura secundaria global mostrada en las Figuras J. K. puede verse que la estructura estabilizada a alta presión corresponde a confórmeros presentes a baja presión, aunque la proporción en la cual se encuentran los mismos en estas condiciones resulta mucho menor.



d. Análisis de movilidad

Para llevar a cabo el análisis de la movilidad de los átomos de la proteína en los sistemas de baja y alta presión, calculamos la Raíz de la Fluctuación Cuadrática Media (RMSF) de los carbonos- α para ambas trayectorias.

La Figura 21 muestra los valores de RMSF de cada carbono- α , tanto para alta como para baja presión, además de la diferencia entre ambos. Puede verse inmediatamente que la movilidad promedio de los residuos se ve significativamente disminuida con el aumento de la presión.

Existen, sin embargo, ciertos residuos particulares cuyas movilidades se ven aumentadas. Tal es el caso de los residuos superficiales Asp18, Asp48, Asp87 y Asp119.

A fin de analizar esta situación se calculó la función distribución radial (RDF) para los oxígenos de las aguas, alrededor de los oxígenos de los carbonilos laterales para estos aminoácidos de movilidad aumentada. Estas RDF se muestran en la Figura 22.



Podemos observar en la Figura 22 (a, b y c) que la segunda capa de hidratación se ve disminuida para los residuos 18, 48, y 87 para el sistema de alta presión en relación al de baja presión. Para estos residuos resultaría razonable suponer que dicha reducción en la capa de hidratación redundaría en un mayor grado de libertad del residuo y por lo tanto en un consecuente aumento de su movilidad.

De cualquier manera, este razonamiento no puede aplicarse a los resultados obtenidos para el Asp119 (Figura 22d), ya que en este caso no se verifica prácticamente ningún cambio en la esfera de hidratación. Debemos recordar que la función distribución radial no describe las características de los enlaces puente de hidrógeno formados por la esfera de hidratación, sino que sólo nos proporciona información sobre el número de aguas que rodean el residuo. Podría suceder que la presión esté produciendo un cambio en esta red de puentes de hidrógeno, lo cual podría resultar en un aumento de movilidad para este residuo. Esta especulación no puede probarse sin un minucioso estudio de la red de puentes de hidrógeno que se establece alrededor del residuo, lo cual resulta una tarea aún pendiente.

La reducción general en la movilidad de ciertas proteínas globulares con el incremento de la presión es un hecho experimental ampliamente observado [110,111,112,113], además de estar de acuerdo con los resultados citados disponibles para la lisozima [108].

Se ha podido, por lo tanto, reproducir no sólo la información experimental en cuanto al comportamiento estructural promedio, sino también en cuanto al comportamiento dinámico que se le ha asignado recientemente a la lisozima al ser sometida a presiones hidrostáticas elevadas.

Estos resultados, junto con el importante grado de acuerdo mencionado con anterioridad, entre los datos experimentales disponibles y los presentes cálculos de Dinámica Molecular permiten establecer fehacientemente la existencia de un importante grado de correlación entre el presente modelo y el sistema real.



Figura 22: Función Distribución Radial (RDF) g(r) de los oxígenos de las aguas alrededor de de los oxígenos de los carbonilos laterales de los residuos Asp, para las presiones de 1bar y 3kbar: (a) Asp18; (b) Asp48; (c) Asp87; (d) Asp119. La función distribución radial g(r) es la densidad de probabilidad de encontrar una partícula a la distancia r de la partícula de referencia, normalizada a una probabilidad igual a 1 para valores de r grandes.

e. Análisis de la superficie accesible al solvente (SAS)

La Figura 23 muestra la superficie accesible al solvente (SAS) para los residuos hidrofóbicos e hidrofílicos de la lisozima a 1bar, mientras que la Figura 24 muestra el gráfico correspondiente para la lisozima a 3kbar. Esta SAS (o superficie de Connelly) se calcula como el área expuesta a una sonda esférica con un radio de 0,14nm, y fue calculada utilizando el programa g_sas [114,117].

Se observa que a baja presión (Figura 23) la SAS hidrofílica es predominante sobre la SAS hidrofóbica, resultados que resultan esperables para una situación normal en la cual los residuos hidrofóbicos evitan el contacto con el agua. Sin embargo este comportamiento se revierte en el sistema de alta presión.

El incremento relativo en el número de residuos no-polares en contacto con el solvente indica la pérdida del comportamiento regular de las sustancias hidrofóbicas en agua con el aumento de la presión.

Dado que la presión no tiene ningún efecto sobre las interacciones directas la respuesta debe ser la de la pérdida, al menos parcial, del efecto hidrofóbico con el aumento de la presión. Esto podría explicarse mediante la reducción del grado de estructura presente en el seno del solvente acuoso con el aumento de la presión [115].



Figura 23: Superficie Accesible al Solvente (SAS) hidrofóbica e hidrofílica vs. el tiempo total de simulación para el sistema a 1bar. Las líneas rectas representan el valor promedio para la corrida completa.
Sin embargo, resulta correcto destacar la importante variabilidad en los resultados mostrados en la Figura 23.

Resulta particularmente notable el hecho de que, durante ciertos períodos, puede observarse una inversión de la relación de SAS hidrofílica/hidrofóbica en el sistema de baja presión. Al comparar este comportamiento con el de la gráfica mostrada en la Figura 24, puede observarse una disminución de alrededor de 1nm² del rango entre los valores mínimo y máximo, tanto para la SAS hidrofóbica como para la SAS hidrofílica.

Esto, al analizarse a la luz de los resultados presentados en las secciones anteriores, refuerza la conclusión de que, para el caso de la lisozima, la presión efectivamente actúa como "Selector de Confórmeros", reduciendo consecuentemente la variabilidad estructural del sistema de alta presión.





El efecto de la presión sobre la estructura del agua puede verse claramente reflejado en el debilitamiento de la red de puentes de hidrógeno [116]. Dicha situación redundaría claramente en una reducción del efecto hidrofóbico. Resulta razonable esperar entonces que resulte más fácil exponer un residuo a un solvente menos estructurado, lo cual se constituiría en la principal fuerza motriz de los cambios conformacionales de las proteínas bajo presión.

4. Conclusiones del capítulo

Este modelo no sólo ha sido exitoso para reproducir el comportamiento global de la lisozima bajo presión, sino que también ha podido dar cuenta del comportamiento aparentemente anómalo de ciertas regiones específicas de la proteína, tal como fuese descripto recientemente a partir de datos experimentales. Explícitamente, el modelo reproduce exitosamente una compactación global de la estructura bajo presión, a la vez que describe correctamente el comportamiento de aquellas pocas regiones que han sido recientemente descriptas como zonas que se alejan de la estructura global bajo los efectos de la presión hidrostática elevada.

También se ha tenido éxito en describir un aumento de la homogeneidad estructural general bajo presión. En este sentido, al enfocarnos sobre elementos estructurales específicos tales como las fluctuaciones en la estructura secundaria regional, surgen nuevos elementos microscópicos que proporcionan argumentos independientes que apuntan en la dirección de confirmar el rol de "Selector de Confórmeros" recientemente asignado a la presión, para proteínas en solución bajo presión.

En cuanto al estudio de la movilidad, una primera observación de los resultados presentados por el presente estudio de Dinámica Molecular confirma las aseveraciones realizadas por estudios previos referidos a que la presión produce una reducción general de los grados de libertad estructurales de esta proteína en solución.

Aunque el punto de ofrecer una explicación completa para el incremento de movilidad, aparentemente anómalo, de ciertos residuos superficiales cargados aún no ha sido plenamente resuelto, una primera aproximación preliminar parece apuntar en la dirección del cambio en la esfera de hidratación de estos residuos como un factor importante que podría explicar, al menos parcialmente, dicho comportamiento.

La evaluación del cambio en la superficie accesible al solvente (SAS) con la presión muestra una clara inversión de la relación de las SAS hidrofílica/hidrofóbica, cambio que es sostenido dinámicamente a la largo de la totalidad del tiempo de simulación. Dicha inversión de relación constituye, por lejos, el cambio más notable en todo el sistema. Esta diferencia debe interpretarse como una disminución del efecto hidrofóbico con el aumento en la presión, cambio que razonablemente explicaría una gran parte de la fuerza motriz de los cambios conformacionales para las proteínas bajo presión.

Capítulo 4

APOMIOGLOBINA EN H2O A 3KBAR:

1. Introducción:

En el presente capítulo se presenta el estudio por Dinámica Molecular del efecto de la presión sobre la estructura y movilidad de la Apomioglobina de Cachalote, llevado a cabo a las presiones de 1bar y 3kbar.

Los resultados presentan un muy buen grado de acuerdo con los datos experimentales de mayor precisión disponibles para alta presión [118]. Esto permite el análisis subsiguiente de otros aspectos del efecto de la presión sobre esta proteína en solución.

Del análisis de la evolución de la estructura secundaria (SS) a la largo de toda la corrida, puede observarse que la estructura de hélice- α se ve favorecida con el aumento de la presión, a expensas de estructuras β , giros y hélices-3'.

El estudio de la movilidad muestra que aunque la movilidad general se ve restringida con el aumento de la presión, ciertos residuos específicos, por el contrario, ven su movilidad aumentada.

Los estudios de la evolución de la estructura terciaria muestran importantes cambios conformacionales con el aumento de la presión.

La evolución de la superficie accesible al solvente (SAS) aumenta notablemente con la presión, debido casi completamente a un aumento sesgado del área hidrofóbica expuesta, lo cual en consecuencia muestra que la interacción hidrofóbica se vuelve considerablemente más débil bajo condiciones de alta presión hidrostática.

Adicionalmente, los hallazgos presentados en el presente capítulo indican que los importantes cambios conformacionales hallados ocurren en la escala de tiempo de 10ns a 30ns, lo cual explica porqué las simulaciones de Dinámica Molecular realizadas con anterioridad sobre proteínas solvatadas bajo presión no lograron observar ningún cambio estructural de envergadura como consecuencia de la perturbación a la presión.

2. Materiales y métodos:

Las simulaciones de Dinámica Molecular se realizaron utilizando el paquete GROMACS 3.2.1 [97,98]. Se utilizó el campo de fuerzas para todos los átomos explícitos (ff2gmx) [99,100,101,102,103,104], tanto para los procedimientos de minimización de energía y equilibrado del sistema, como para todos los pasos de simulación por Dinámica Molecular.

Se mantuvieron las restricciones topológicas para las longitudes de enlace utilizando el algoritmo LINCS [96]. Las restricciones topológicas de las moléculas de agua se resolvieron utilizando el algoritmo SETTLE [95]. Para el cálculo de las fuerzas electrostáticas se utilizó el método de Campo de Reacción [105]. Las interacciones de Lennard-Jones fueron calculadas explícitamente dentro de un radio de corte de 1,4nm.

Para todas las simulaciones de este sistema se utilizó un cluster de 2 procesadores Dual-Xeon, corriendo bajo GNU/Linux, utilizando para las visualizaciones de trayectorias, estructuras y representaciones gráficas el software de distribución libre Visual Molecular Dynamics (VMD), Swiss PDB Viewer (SPDBV) y XMGrace respectivamente.

Como estructuras de partida para la Apomioglobina se utilizaron las resueltas por cristalografía de rayos X de alta y baja presión (códigos de PDB 1JP8 y 1VXD respectivamente) publicadas y disponibles en el "Protein Data Bank".

La topología para la apomioglobina se generó utilizando la herramienta PDB2GMX, con estados de protonación estándar para todos los aminoácidos (pH 6).

Se utilizó el modelo de agua SPC/E [107], tanto para el sistema de alta presión como para el de baja presión.

El sistema inicial consistió de una caja cúbica de dimensiones X=Y=Z=7,19617nm, con un volumen de $372,653nm^3$, conteniendo una molécula de apomioglobina y 11663 moléculas de agua.

Sobre este sistema se realizó una primera minimización de energía por el método de "steepest descent" o pendiente pronunciada, la cual convergió al valor de la precisión del cómputo. Luego se realizó una segunda minimización de energía por el método de gradientes conjugados, el cual logró converger dentro de los primeros 20 ciclos.

Luego de la minimización de energía se permitió que el solvente se relajara alrededor de la proteína, realizando una breve simulación por Dinámica Molecular (10ps) sobre el sistema, aplicando restricciones de posición sobre todos los átomos de la proteína. Desde el comienzo el sistema fue acoplado a un baño débil, tanto de temperatura como de presión, de manera de trabajar en el ensamble isotérmico e isobárico [88], a la temperatura T=300K y presión P=1bar.

A continuación se permitió que todo el sistema se relajara (se quitaron las restricciones de posición de todos los átomos) durante 200ps, manteniendo el acoplamiento del sistema al termóstato y al baróstato.

En este punto el sistema fue dividido en dos. La primera copia fue sometida lentamente a presiones hidrostáticas ascendentes, mientras que la segunda se mantuvo a presión atmosférica.

Al primer sistema se le aumentó la presión lentamente durante un tiempo total de 2ns, utilizando una constante de relajación muy lenta para el baño de presión ($\tau_p = 400 \, ps$). Al segundo sistema se le permitió evolucionar durante un tiempo equivalente (2ns) pero a presión de 1bar.

Habiendo alcanzado este punto se permitió que ambos sistemas evolucionaran durante un período adicional de 8ns, de modo de producir puntos de partida equilibrados y comparables para ambos sistemas.

Una vez alcanzado este punto se realizó una simulación de 16ns de longitud, tanto para el sistema a 1bar como para el sistema a 3kbar. El sistema de alta presión fue continuado por un tiempo adicional de simulación de 34ns debido al hecho de que la estabilidad conformacional no se había logrado en el período previo de 16ns.

Todos los resultados presentados a continuación corresponden a estas simulaciones de 16 y 50ns de duración.

Todos los datos mostrados en los gráficos y matrices presentadas en el presente capítulo fueron calculados utilizando los programas correspondientes del paquete GROMACS. La superficie accesible al solvente (SAS o superficie de Connelly) se calculó utilizando el programa g_sas [114,117] con una sonda esferica de radio 0,14nm.

3. Resultados y discusión:

a. Estructuras promedio de referencia y raíz de la fluctuación cuadrática media (RMSF):

Luego de equilibrar los sistemas tanto de alta como de baja presión durante un tiempo total de 10ns, el último nanosegundo (ns) de dicho período se utilizó para producir estructuras promedio de referencia adecuadas de toda la proteína, como así también para obtener los valores de referencia para la desviación estándar de todos los carbonos- α de la proteína. La Figura 25 muestra una representación tipo cinta de ambas estructuras superpuestas. Dicha superposición se realizó a través de un ajuste de cuadrados mínimos de la estructura promedio (9ns a 10ns) de alta presión (en color) contra la estructura promedio (9ns a 10ns) de baja presión (en blanco).

Puede observarse en esta figura que el efecto general de la presión, en este punto, puede describirse como una contracción, globalmente homogénea, de la estructura terciaria de la proteína. Sin embargo cabe destacar que dichas estructuras mantienen las asociaciones tridimensionales generales (tal como puede observarse por el tono azul preponderante, representando valores bajos de RMS, sobre la estructura de alta presión), como así también casi las mismas regiones para las estructuras secundarias, lo cual está indicando probablemente que estamos observando básicamente pequeñas variaciones de un mismo confórmero.



Figura 25: Representación tipo cinta de las estructuras promedio, calculadas entre los 9ns y 10ns de simulación, de la Apomioglobina, para los sistemas de 1bar (en blanco) y 3kbar (en color). Las estructuras fueron espacialmente ajustadas por cuadrados mínimos una sobre la otra. El código de color sobre la estructura de 3kbar debe leerse: rojo = máxima desviación posicional; azul = mínima desviación posicional.

Los valores para las desviaciones estándar de todos los carbonos- α (RMSF) entre el nanosegundo 9 y el nanosegundo 10 se muestran en la Figura 26, tanto para el sistema a 1bar (negro) como para el sistema a 3kbar (rojo). Las líneas horizontales representan el valor promedio para cada condición. La primera conclusión que puede extraerse de este gráfico es que la movilidad promedio de los carbonos- α a 1 bar resulta aproximadamente 1,5 veces mayor

que la correspondiente movilidad promedio de los carbonos- α a 3kbar. Esto resulta cierto para casi todos los carbonos- α , con la notable excepción de aquellos pertenecientes a los residuos 38, 39, 41, y 42. Por otro lado, la mayor reducción en la movilidad se produce alrededor de los residuos 112 a 126, junto con el residuo terminal 153. Estos hallazgos serán abordados nuevamente cuando discutamos los cambios en la estructura secundaria.



Figura 26: Desviación estandar (Raiz de la fluctuación cuadratica media -RMSF o σ - alrededor de la posición promedio) calculada para cada carbono- α , utilizando todos los cuadros grabados en la trayectoria desde los 9ns hasta los 10ns de simulación. En negro están representados los valores calculados para el sistema a 1bar, mientras que en rojo se representan los correspondientes para el sistema a 3kbar. Las líneas horizontales negra y roja representan el RMSF promediado sobre todos los carbonos- α , para los sistemas de baja y alta presión respectivamente.

b. Análisis de la raíz de la desviación cuadrática media (RMSD) de los carbonos-α a lo largo de todas las trayectorias:

Para el cálculo del RMSD de los carbonos- α se utilizaron las estructuras promedio calculadas y discutidas en la sección anterior. Todas las gráficas correspondientes al RMSD de los carbonos- α fueron normalizadas dividiéndolas por los valores correspondientes de desviación estándar para los carbonos- α calculadas y discutidas en la sección anterior.

Esto se realizó de modo de poder comparar efectivamente los resultados provenientes de estos dos sistemas, ya que ambos presentan movilidades muy disímiles para sus carbonos- α .



Figura 27: Raíz de la desviación cuadrática media normalizada (RMSD) de los carbonos- α , tanto para el sistema a 1bar (negro) como para el sistema a 3kbar (rojo), entre 10ns y 26ns de simulación. El proceso de normalización se llevó a cabo dividiendo valor de RMSD, calculado para cada cuadro registrado (luego de realizar un ajuste por cuadrados mínimos sobre cada estructura contra la estructura de referencia), por el valor correspondiente de RMSF (o σ) promedio de los carbonos- α , representados por las líneas horizontales en la Figura 26 de la sección anterior.

La Figura 27 muestra la evolución del RMSD normalizado tanto para el sistema de alta presión (rojo) como para el de baja presión (negro), desde los 10ns hasta los 26ns de simulación. Al analizar este período de simulación de 16ns puede observarse un comportamiento muy razonable para el sistema a 1bar. De hecho, durante todo este período su valor fluctúa alrededor de 1,3 unidades de σ (desviación estándar). Cabe destacar aquí que los 10ns de estabilización resultaron ser tiempo suficiente para que este sistema evolucionara desde su estructura de holo-proteína hasta la correspondiente estructura apo-proteica. Dicho de otro modo, en 10ns el sistema logró reaccionar completamente a la remoción de su grupo hemo.

Por el contrario, cuando analizamos el comportamiento del sistema a 3kbar, observamos un comportamiento drásticamente diferente. Concretamente, los valores del RMSD normalizado sólo oscilan alrededor de 1,3 unidades de σ desde los 10ns hasta los 19ns, e incluso durante este corto período el comportamiento del sistema de alta presión se presenta como muy errático cuando es comparado con los resultados correspondientes del sistema de baja presión. Entre los 19ns y 20ns el RMSD normalizado muestra un salto muy claro, alcanzando lo que aparentaría ser un nuevo estado estacionario en un valor de alrededor de 2 unidades de σ . Así, podemos concluir que, mientras que el sistema de baja presión ha alcanzado un punto de estabilidad relativa, alrededor del cual fluctúa, el sistema de alta presión muestra un comportamiento mucho más inestable. Consecuentemente se continuó con el estudio de la evolución del sistema a 3kbar, extendiendo el período de simulación de 26ns a 60ns totales.



La Figura 28 muestra la evolución del RMSD normalizado para el sistema a 3kbar, durante los 60ns totales de simulación. Lo que resulta claro de un primer análisis de esta figura es que el RMSD normalizado no converge

horizontales en la Figura 26 de la sección anterior.

hacia ningún valor dado durante la totalidad del período entre los 10ns y los 60ns de simulación. Esto indica que, lejos de estabilizarse en alguna conformación dada, la estructura tridimensional del la proteína sufre una evolución permanente durante la totalidad de la simulación de alta presión.



Figura 29: Matriz de valores de la raíz de la desviación cuadrática media (RMSD) para la totalidad del tiempo de simulación del sistema a 1bar (de 10ns a 26ns). En este caso, el valor de RMSD se calculó de a pares, cubriendo todas la posibles combinaciones entre cuadros registrados durante la simulación. Los resultados se presentan en formato de código de color. La barra muestra la equivalencia color/nanometro. El color puede interpretarse en unidades de σ (o desviación estándar) según la siguiente conversión: amarillo=1,36 σ ; verde=2,72 σ ; celeste=4,1 σ ; azul=5,44 σ .

Las Figuras 29 y 30 permiten un análisis más detallado del presente aspecto. En el caso de la Figura 29, el código de color de la referencia fue

escalado en un factor de 1,44 (Cociente entre las desviaciones estándar: $\frac{\sigma_{lbar}^{\overline{aC}}}{\sigma_{3kbar}^{\overline{aC}}} = 1,44$) de modo de producir resultados que sean gráficamente comparables con aquellos presentados en la Figura 30.

La Figura 29 muestra dos regiones estables (amarillo-naranja) que dominan casi toda la simulación, conectadas a través de una región intermedia de fluctuación levemente mayores (amarillo-verde). con valores Generalizando, a partir de una interpretación gráfica de los datos podemos decir que prácticamente todas las fluctuaciones en la matriz (de la Figura 29) se mantienen por debajo de valores de RMSD normalizado de 1,75 σ (amarillo=1,36 σ ; verde=2,72 σ ; celeste=4,1 σ ; azul=5,44 σ). En este sentido, la Figura 29 parece confirmar la primera impresión obtenida a partir de la Figura 27. Concretamente, que la simulación del sistema de baja presión ha alcanzado una conformación de relativa estabilidad, fluctuando levemente alrededor de dicho punto.

Análogamente, a partir de un estudio gráfico general de la Figura 30, podemos concluir que las estructuras del sistema a 3kbar no tienden a repetirse. De hecho, aún en los períodos de relativa estabilidad (evidenciados por las regiones-verde amarillas) la mayoría de las estructuras difieren entre sí en valores mayores a $1,75 \sigma$. La diferencia alcanza valores mayores que 5σ cuando se comparan las primeras y las últimas estructuras de la simulación (regiones azules). Por otro lado, a medida que avanzamos en el tiempo, podemos observar que los períodos de relativa estabilidad se vuelven cada vez más cortos, lo cual parece indicar, junto con las observaciones previas, en la dirección de una progresiva pérdida de estabilidad debida probablemente a la concomitante pérdida parcial de la conformación nativa original de la proteína.



Figura 30: Matriz de valores de la raíz de la desviación cuadrática media (RMSD) para la totalidad del tiempo de simulación del sistema a 3kbar (de 10ns a 60ns). Al igual que en la Figura 29, en este caso, el valor de RMSD se calculó de a pares, cubriendo todas la posibles combinaciones entre cuadros registrados durante la simulación. Los resultados se presentan en formato de código de color. El código de color se muestra al fondo de la figura. Utilizando la barra, puede traducirse inmediatamente el color en valores de RMSD, en nanometros (nm), como se muestra en la leyenda. La conversión de color a unidades de σ (o desviación estándar) puede realizarse según la siguiente equivalencia: amarillo=1,36 σ ; verde=2,72 σ ; celeste=4,1 σ ; azul=5,44 σ .

c. Análisis comparativo de la Estructura Secundaria (SS) y su evolución:

El análisis de la estructura secundaria se realizó utilizando las definiciones del "Dictionary of Secondary Structure for Proteins" (DSSP) [69], y se calculó utilizando el programa do_dssp, perteneciente al paquete de GROMACS.



para la totalidad del tiempo de simulación (10ns a 26ns) del sistema a 1bar. El código de color puede interpretarse a partir de la información presente en la leyenda.

A continuación se presenta un análisis comparativo completo de las Figuras 31 y 32, presentadas en esta sección:

Residuos 1 a 3: (Coil1) ovillo aleatorio estable a 1bar, sin variaciones a 3kbar. Residuos 4 a 19: (Helix1) α -hélice estable a 1bar, sin variaciones a 3kbar. Residuo 20: oscila entre giro, vuelta y ovillo aleatorio, mayormente sin variaciones a 3kbar.

Residuos 21 a 23: (Turn1) vuelta estable a 1bar, sin variaciones a 3kbar.

Residuos 24 a 35: (Helix2) α-hélice estable a 1bar, sin variaciones a 3kbar.

Residuo 36: oscila entre giro y ovillo aleatorio, mayormente sin variaciones a 3kbar.

Residuos 37 a 38: (pre-Helix3) oscila entre α -hélice y vuelta a 1bar, estabilizándose como α -hélice a 3kbar.

Residuos 39 a 42: (Helix3) α-hélice estable a 1bar, sin variaciones a 3kbar.

Residuos 43 a 48: (Helix3') oscila entre giro, vuelta y ovillo aleatorio (e incluso a veces hélice-3' para los residuos 46 a 48). Toda esta región es estabilizada como α -hélice, con la única excepción del residuo 43, el cual actúa como bisagra conectado la Helix3 con la Helix3' a la presión de 3kbar.

Residuos 49 a 50: (Coil2) ovillo aleatorio estable a 1bar, sin variaciones a 3kbar.

Residuo 51: giro mayormente estable a 1bar, sin variaciones a 3kbar.

Residuos 52 a 57: (Helix4) a-hélice estable a 1bar, sin variaciones a 3kbar.

Residuo 58: oscila entre giro y ovillo aleatorio, sin variaciones a 3kbar.

Residuos 59 a 77: (Helix5) α -hélice mayormente estable a 1bar, notando que los residuos 73 a 76 tienden también a formar vuelta (e incluso a veces hélice-3'). A 3kbar toda esta región es estabilizada como α -hélice, extendiendo con frecuencia la hélice hasta el residuo 77.

Residuos 78 a 80: oscila entre giro, vuelta y ovillo aleatorio, mayormente sin variaciones a 3kbar.

Residuos 81 a 82: (Coil3) ovillo aleatorio estable a 1bar, sin variaciones a 3kbar.

Residuos 83 a 86: (pre-Helix6) oscila entre α -hélice y vuelta a 1bar, estabilizándose como α -hélice a 3kbar.

Residuos 87 a 91: (Helix6) \alpha-hélice estable a 1bar, notablemente extendida desde el residuo 83 hasta el residuo 94 a 3kbar.

Residuos 92 a 94: (post-Helix6) oscila entre α -hélice y vuelta a 1bar, estabilizándose como α -hélice a 3kbar.

Residuo 95: predominantemente α-hélice a 1bar, estabilizándose como vuelta a 3kbar.

Residuo 96: vuelta mayormente estable a 1bar, sin variaciones a 3kbar.

Residuo 97: giro mayormente estable a 1bar, sin variaciones a 3kbar.

Residuos 98 a 100: (Coil4) ovillo aleatorio estable a 1bar, sin variaciones a 3kbar.

Residuos 101 a 118: (Helix7) α -hélice estable a 1bar. Esta región se ve mayormente afectada por la alta presión en el residuo 101, el cual prácticamente deja de participar de la α -hélice, al igual que los residuos 113 a 118, los cuales solo logran volver a la configuración de α -hélice luego de 20ns. Sin embargo a partir de este punto la α -hélice parece constituir una configuración muy estable para estos residuos bajo condiciones de alta presión. Debeos hacer notar en este punto que son estos residuos los que vieron su movilidad más disminuida con la presión.

Residuos 119 a 124: esta región no posee una estructura secundaria definida, ni a baja ni a alta presión. Sin embargo, tanto el residuo 119 como el residuo 124 parecen tener una preferencia por la conformación de ovillo aleatorio, mientras que el resto de los residuos de esta región fluctúan mayormente entre las estructuras de vuelta y giro. También resulta notoria la ausencia total de puente- β entre los residuos 120 y 123 a 3kbar (el cual se presenta esporádicamente a 1bar).

Residuos 125 a 149: (Helix8) α -hélice estable a 1bar, mayormente sin variaciones a 3kbar, con la excepción de los residuos de los extremos de esta

región (residuos 125, 126, 148 y 149) los cuales fluctúan entre las conformaciones de hélice y vuelta a alta presión.

Residuo 150: oscila entre giro y vuelta, mayormente sin variaciones a 3kbar. Residuos 151 a 153: (Coil5) ovillo aleatorio estable a 1bar, sin variaciones a 3kbar.

Nota: los nombres de las regiones para la estructura secundaria de la apomioglobina, escritos entre paréntesis, se dejan en inglés para facilitar la comparación con las publicaciones citadas como referencia experimental en el presente manuscrito.



Figura 32: Matriz de la evolución temporal de la estructura secundaria (SS) para la totalidad del tiempo de simulación (10ns a 60ns) del sistema a 3kbar. El código de color puede interpretarse a partir de la información presente en la leyenda.

Resumen de las observaciones sobre la estructura secundaria (SS) y su comportamiento bajo presión:

La mayoría de las regiones permanecen sin cambios con el aumento de la presión, aunque sí aparecen ciertos cambios de carácter local que deben ser mencionados.

Casi todos los cambios observados en la estructura secundaria con el aumento de la presión obedecen a alguno de los siguientes patrones:

1bar		3kbar
Vuelta	\rightarrow	α-hélice
Giro	\rightarrow	α-hélice
hélice-3'	\rightarrow	α-hélice

Esto es cierto para todos los cambios observados, con la excepción de los residuos 95, 125, 126, 148 y 149, los cuales obedecen a un patrón de cambio de α -hélice a vuelta. Cabe igualmente destacar que ninguno de estos residuos llega a perder completamente su conformación de α -hélice, sino que más bien se observa es algún grado de pérdida de estabilidad de la misma. Por último, si bien los resultados presentados en esta sección parecen apuntar en la dirección de una menor reducción en la movilidad de aquellos residuos que participan de una α -hélice, tal conclusión requiere de un análisis más detallado y sistemático de los efectos observados a través de la simulación por dinámica molecular de sistemas de alta presión.

d. Análisis comparativo de la Superficie Accesible al Solvente (SAS) y su evolución:

A modo de corroboración gráfica de las observaciones y conclusiones extraídas a partir del análisis y discusión detallados de los datos de RMSD y estructura secundaria presentados en las secciones previas, procederemos a visualizar la estructura de la proteína.



Figura 33: Representación tipo cilindro (o "cartoon") de la estructura promedio, de 10ns a 26ns, del sistema a 1bar (cyan) superpuesta por cuadrados mínimos sobre de la estructura promedio, de 58ns a 59ns, del sistema a 3kbar (púrpura).

La Figura 33 muestra una representación tipo cilindro (o "cartoon") de las estructuras promedio, de 10ns a 26ns del sistema a 1bar (cyan) y de 58ns a 59ns del sistema a 3kbar (púrpura), superpuestas una sobre la otra por cuadrados mínimos. Una primer mirada rápida de esta figura podría darnos la

impresión errónea de que, a pesar del análisis previo, parecieran no haber ocurrido cambios conformacionales de relevancia entre los sistemas de baja y alta presión.

Sin embargo, cuando elegimos la representación algo más realista (aunque más incomoda y computacionalmente más pesada) de la superficie proteica accesible al solvente (SAS ó superficie de Connely) vemos completamente revelado el cambio conformacional hacia el cual todos los datos previamente analizados apuntaban. La superficie accesible al solvente para las estructuras promedio de baja y alta presión representadas en la Figura 33, se presentan por separado en las Figuras 34 y 35, respectivamente.



Figura 34: Superficie Accesible al Solvente (SAS o superficie de Connely) calculada para la estructura promedio, de 10ns a 26ns, del sistema a 1bar. Para el cálculo de la superficie se utilizó una sonda esférica de radio 1,4Å. La orientación presentada en esta figura es exactamente la misma que la elegida para la Figura 33, tal cual queda de manifiesto al observar la orientación de los ejes en la esquina inferior izquierda de ambas figuras.

Mientras que en la Figura 34 podemos ver una superficie medianamente regular, lo cual resulta razonable dado el carácter globular de una proteína como la apomioglobina, la Figura 35 muestra, por el contrario, una forma de lo más irregular. Una observación más detallada de las imágenes revela la existencia de una grieta en la estructura promedio del sistema de alta presión, que está ausente de en el sistema a baja presión. Esta grieta corresponde a la notable separación entre la región $\alpha 3$ (α -hélice) y la región $\alpha 6$ (α -hélice), lo cual da cuenta de gran parte del cambio conformacional. Esta grieta facilita el acceso del solvente al interior de la proteína, y puede ser el resultado del descenso en la interacción hidrofóbica con el aumento de la presión, como discutiremos a continuación. Resulta importante destacar en este punto que la separación entre las regiones $\alpha 3$ y $\alpha 6$ ha sido observada experimentalmente por resonancia magnética nuclear [118].



Figura 35: Superficie Accesible al Solvente (SAS o superficie de Connely) calculada para la estructura promedio, de 58ns a 59ns, del sistema a 3kbar. Para el cálculo de la superficie se utilizó una sonda esférica de radio 1,4Å. La orientación presentada en esta figura es exactamente la misma que la elegida para la Figura 33, tal cual queda de manifiesto al observar la orientación de los ejes en la esquina inferior izquierda de ambas figuras.

Con el objetivo de lograr una perspectiva más minuciosa que permita el análisis de la fuerza principal que gobierna el cambio conformacional observado con el aumento de la presión, se han calculado y graficado las SAS para la totalidad de las trayectorias de alta y baja presión. Las Figuras 36a, 36b y 36c muestran la evolución temporal de la SAS para la trayectoria completa del sistema de 1bar, de 10ns a 26ns, presentando separadamente las superficies hidrofóbica, hidrofílica y total, expuestas al solvente. Una vez más puede observarse que el sistema a 1bar fluctúa alrededor de valores aparentemente estables de SAS. Las líneas rectas presentes en las Figuras 36a, 36b y 36c, muestran el valor promedio para cada gráfica.



Análogamente, las Figuras 37a, 37b y 37c muestran la evolución temporal de la SAS para la trayectoria completa del sistema de 3kbar, de 10ns a 60ns, presentando separadamente las superficies hidrofóbica, hidrofílica y total, expuestas al solvente. Contrariamente a lo observado para el sistema de baja presión, el sistema a 3kbar muestra un comportamiento relativamente estable solamente para la superficie hidrofílica expuesta al solvente, mientras que la superficie hidrofóbica expuesta muestra un permanente incremento a lo largo de la totalidad del período de 10ns a 60ns de simulación. Lo mismo puede decirse de la superficie accesible total, lo cual se debe obviamente en su mayoría al aumento en la SAS hidrofóbica.

para cada figura.

Para lograr una comparación efectiva entre los sistemas de alta y baja presión, las líneas rectas presentes en las Figuras 37a, 37b y 37c corresponden a los valores promedio a 1bar, <u>no</u> a los promedios del sistema a 3kbar. De este modo, podemos observar que la SAS total sufre un incremento de alrededor de 8nm², desde el comienzo hasta el final de la simulación, alcanzando valores aún mayores al promedio para la SAS total del sistema a 1bar.



simulación del sistema a 3kbar, mostrando separadamente: a) SAS total (verde); b) SAS hidrofóbica (negro); c) SAS hidrofílica (rojo). A diferencia de la Figura 36, las líneas horizontales en a), b), y c) muestran los correspondientes valores promedio del sistema a 1bar, <u>no</u> a los promedios del sistema a 3kbar.

Al analizar el comportamiento específico de las SAS hidrofóbica e hidrofílica separadamente, deben notarse dos aspectos relevantes. El primero lo constituye el hecho de que, analizando la SAS hidrofílica, el sistema a 3kbar casi nunca alcanza el valor promedio para el sistema a 1bar, y nunca lo supera.

El segundo aspecto lo constituye el hecho de que mientras que la SAS hidrofóbica muestra una clara tendencia de aumento sostenido hacia valores de exposición mayores, no puede decirse lo mismo respecto del comportamiento de la SAS hidrofólica. Por último, volviendo al comportamiento de la SAS hidrofóbica, puede observarse que su evolución a lo largo de la totalidad de la simulación de alta presión la lleva a valores que se encuentran claramente por encima del valor promedio para el sistema de baja presión, durante prácticamente toda la segunda mitad de la simulación de alta presión.

Por lo tanto, a la luz de los datos presentados, puede decirse que la evolución del sistema de alta presión puede seguirse a través del incremento del área proteica total expuesta, el cual se debe casi con exclusividad al aumento de la SAS hidrofóbica.

Una explicación plausible para esta observación podría lograrse prontamente si asumimos que la presión actúa rompiendo la estructura natural del solvente, reduciendo por lo tanto la capacidad del solvente de estabilizar efectivamente la estructura proteica nativa a través del efecto solvofóbico (específicamente hidrofóbico en este caso). Este debilitamiento de la interacción hidrofóbica se debe a la ruptura de la red de puentes de hidrógeno a causa del aumento de la presión, hecho que ha sido demostrado tanto experimentalmente [119] como por simulación [116,120].

Un mayor sustento para la afirmación previa puede encontrarse cuando observamos el lapso de tiempo que debe transcurrir en la simulación para que dichos cambios se vuelvan evidentes. Al revisar ya sea las figuras de la sección correspondiente al análisis del RMSD, o bien los gráficos de la SAS de la presente sección, podemos observar que cualquiera de los parámetros estudiados tarda un mínimo de entre 4ns y 10ns en manifestar una tendencia clara, y por lo tanto volverse evidente. Esto último indicaría que esta es en efecto la escala de tiempo en la cual la proteína es capaz de explorar su repentinamente accesible espacio conformacional (o ensamble), debido a un cambio en un factor entrópico de estabilidad, tal como lo es la reducción del efecto hidrofóbico del agua, causado por un incremento en la presión ejercida sobre el sistema.

4. Conclusiones del capítulo:

El estudio comparativo de la apomioglobina en solución a baja y alta presión, llevado a cabo por medio de la técnica de simulación por dinámica molecular, permite un análisis detallado de los cambios en la conformación de la proteína bajo los efectos de la presión. Sin embargo, resulta claro que no es posible detectar los efectos de la presión en una simulación corta (de una duración del orden de las unidades de nanosegundos. Para que dichos efectos se vuelvan evidentes, los tiempos de simulación deben entrar en el orden de las decenas de nanosegundos. También es de destacar que, en el caso de la apomioglobina, la reacción completa de la proteína a los efectos de la presión requiere de una extensión de tiempo de simulación que no ha sido cubierto en el presente trabajo. De todos modos pueden extraerse datos interesantes a partir de la evolución parcial observada para la estructura de la proteína. El primer efecto causado por la presión es la compresión global de la proteína, con importante preservación de la estructura tridimensional global. A medida que el proceso continúa esta primera afirmación pierde validez, y la proteína comienza a perder su forma compacta, incrementando de manera sostenida su superficie accesible al solvente. El solvente penetra en el interior de la proteína, produciendo así la separación entre ciertas α-hélices específicas.

Este efecto puede explicarse conociendo el hecho de que la presión produce la ruptura de la red de puentes de hidrógeno del agua, la cual resulta esencial para garantizar la existencia del efecto hidrofóbico. Este efecto puede ser visto claramente al observar el aumento de la superficie hidrofóbica expuesta al solvente, lo cual significa que la "repulsión" entre los grupos nopolares y el agua se ha visto disminuida de modo notable. Esta ruptura en la red de puentes de hidrógeno del agua parece resultar de fundamental importancia en relación con la desestabilización de la conformación proteica nativa bajo presión, siempre y cuando la presión exceda los valores debajo de los cuales el agua mantenga un comportamiento regular.

Podría especularse que, de mantenerse la presión durante suficiente tiempo, la proteína se desplegaría completamente, perdiendo todo vestigio de su conformación nativa. Sin embargo, dentro de los lapsos de tiempo explorados en el presente trabajo, la presión parece estabilizar la estructura secundaria de α -hélice, lo cual podría implicar que el estado no-nativo podría no resultar una estructura de ovillo aleatorio bajo condiciones de alta presión. La elucidación de tal planteo requiere, por supuesto, de ulteriores investigaciones.

AGUA PURA A 1BAR Y 3KBAR:

1. Introducción:

En el presente capítulo se discuten los resultados del estudio de simulaciones por Dinámica Molecular de agua pura tanto a la presión de referencia de 1bar como así también a la presión de 3kbar.

Se ha prestado especial atención al análisis detallado de aquellos parámetros que puedan ofrecer información clara sobre el grado de estructuración de las moléculas en el seno del agua, como así también a la comparación de dichos resultados entre las condiciones de alta y baja presión utilizadas.

Se intenta además establecer en este capítulo una relación más clara entre los cambios estructurales en el seno del agua y el comportamiento del efecto hidrofóbico, una propiedad de naturaleza esencialmente entrópica.

Las propiedades del agua y los efectos causados por la presión sobre la misma han sido largamente estudiadas, especialmente aquellas que presentan un comportamiento frecuentemente denominado "anómalo".

En este sentido, una notable explicación cualitativa para el comportamiento "anómalo" del coeficiente de auto-difusión del agua para diferentes densidades [121,122] había sido propuesta en base al cambio en la distribución de los puentes de hidrógeno [123]. La misma ha sido recientemente corroborada mediante métodos de simulación [116], estableciendo un posible punto de contacto entre dicho comportamiento "anómalo" y el cambio en la solubilidad de ciertas substancias hidrofóbicas de interés biológico en tales condiciones [124,125], aunque aún no se ha podido elaborar ninguna explicación definitiva para tales eventos.

Es posible que, dada la parcialidad de las respuestas halladas hasta el momento, una explicación cabal y acabada precise aún de algunos años de posteriores estudios, además de la elaboración de una teoría más abarcativa.

En este punto surge entonces la necesidad de realizar el análisis específico del efecto de la presión sobre la estructura del seno del agua, bajo las condiciones específicas de perturbación aplicadas en los capítulos anteriores. En este capítulo presentamos los resultados del análisis sobre dos simulaciones de agua pura, a las presiones de 1bar y 3kbar. Se computaron tres aspectos diferentes, aunque complementarios de la distribución de los puentes de hidrógeno, como así también la función de distribución radial (RDF) entre los átomos de oxígeno del agua, además del cálculo de la evolución temporal del número total de puentes de hidrógeno en ambas condiciones referidas.

Los resultados muestran que, para todos los parámetros estudiados, existen variaciones que van desde lo moderado hasta lo notorio e indican una clara pérdida de estructura con el aumento de la presión.

2. Materiales y métodos:

Las simulaciones de Dinámica Molecular se realizaron utilizando el paquete GROMACS 3.2.1 [97,98]. Se utilizó el campo de fuerzas para todos los átomos explícitos (ff2gmx) [99,100,101,102,103,104], tanto para los procedimientos de minimización de energía y equilibrado del sistema, como para todos los pasos de simulación por Dinámica Molecular.

Se mantuvieron las restricciones topológicas para las longitudes de enlace utilizando el algoritmo LINCS [96]. Las restricciones topológicas de las moléculas de agua se resolvieron utilizando el algoritmo SETTLE [95]. Para el cálculo de las fuerzas electrostáticas se utilizó el método de Campo de Reacción [105]. Las interacciones de Lennard-Jones fueron calculadas explícitamente dentro de un radio de corte de 1,4nm.

Para todas las simulaciones de este sistema se utilizó un cluster de 2 procesadores Dual-Xeon, corriendo bajo GNU/Linux, utilizando para las visualizaciones de trayectorias, estructuras y representaciones gráficas el software de distribución libre Visual Molecular Dynamics (VMD), Swiss PDB Viewer (SPDBV) y XMGrace respectivamente.

Como estructuras de partida para la el seno del agua se utilizaron réplicas apiladas de una caja de 216 moléculas de agua, previamente estabilizadas.

Se utilizó el modelo de agua SPC/E [107], tanto para el sistema de alta presión como para el de baja presión.

El sistema inicial consistió de una caja cúbica de dimensiones X=Y=Z=3,72529nm, con un volumen de 51,69878nm³, conteniendo 1728 moléculas de agua.

Sobre este sistema se realizó una primera minimización de energía por el método de "steepest descent" o pendiente pronunciada, la cual convergió al valor de la precisión del cómputo. Luego se realizó una segunda minimización de energía por el método de gradientes conjugados, el cual logró converger dentro de los primeros 20 ciclos.

Luego de la minimización de energía se realizó una breve simulación por Dinámica Molecular (20ps) sobre el sistema, con el objeto de lograr la relajación del solvente.

Desde el comienzo el sistema fue acoplado a un baño débil, tanto de temperatura como de presión, de manera de trabajar en el ensamble isotérmico e isobárico [88], a la temperatura T=300K y presión P=1bar.

A partir de este punto el sistema fue dividido en dos. La primera copia fue sometida lentamente a presiones hidrostáticas ascendentes, mientras que la segunda se mantuvo a presión atmosférica.

Al primer sistema se le aumentó la presión lentamente durante un tiempo total de 0,5ns, utilizando una constante de relajación lenta para el baño de presión ($\tau_p = 100 ps$). Al segundo sistema se le permitió evolucionar durante un tiempo equivalente (0,5ns) pero a la presión de 1bar.

Habiendo alcanzado este punto se permitió que ambos sistemas evolucionaran durante un período adicional de 0,5ns, de modo contar con simulaciones equilibradas y comparables para ambos sistemas.

Todos los resultados presentados a continuación corresponden a estas últimas simulaciones de 0,5ns de duración.

Todos los datos mostrados en los gráficos y curvas presentadas en el presente capítulo fueron calculados utilizando los programas correspondientes del paquete GROMACS, con la excepción de la Figura 40 cuyos datos fueron calculados utilizando un programa desarrollado primariamente por el Dr. Ernesto Caffarena, y luego levemente modificado por mí.

Resultados y discusión: a. Función distribución radial (RDF) para los oxígenos

En esta sección se presentan tanto los resultados del cálculo de la función distribución radial acumulada (CN-RDF), como también los obtenidos para el cálculo de la función distribución radial propiamente dicha (RDF).

En la Figura 38 se observan las gráficas de las funciones distribución radial acumulada (CN-RDF) para el sistema a 1bar (negro) y para el sistema a 3kbar (rojo).



Un análisis directo de esta gráfica muestra un incremento claro en el número de moléculas vecinas, para cualquier radio analizado, con el aumento de la presión. Tal resultado no es más que lo esperable, dado el aumento en la densidad que acompaña al incremento de la presión. Cabe destacar que a partir de dicha curva no es posible extraer ninguna conclusión directa respecto a los cambios estructurales sufridos en el seno del agua, más allá del hecho previamente mencionado respecto de la compactación sufrida por este sistema a presiones elevadas.

En la Figura 39 se observan las gráficas de las funciones distribución radial propiamente dicha (RDF) para el sistema a 1bar (negro) y para el sistema a 3kbar (rojo).



A diferencia de la gráfica anterior, en la Figura 39 pueden observarse los primeros indicios claros de los cambios estructurales dramáticos sufridos por el sistema a 3kbar.

Podemos observar en primer lugar una pequeña disminución en el máximo del primer pico, con el aumento de la presión. Esto se debe primariamente al incremento en la densidad promedio del seno, el cual se ve reflejado solo levemente sobre la primera esfera de hidratación, lo cual redunda en una disminución neta del valor máximo del pico merced al proceso de normalización aplicado sobre la curva. Por el contrario, la posición del máximo no sufre variaciones, lo cual indica que el acercamiento de cualquier molécula a su primera esfera de hidratación, aunque resulta razonable suponer que ocurre, resulta ser de relevancia despreciable.

Sin embargo, el resultado más notable lo constituye la virtual desaparición concertada del primer valle y el segundo pico. Es este resultado el que nos indica con mayor claridad la evidente pérdida de estructura, evidente a partir de estos resultados esencialmente alrededor de la segunda esfera de hidratación.

b. Distribución de los puentes de hidrógeno por molécula de agua

Para el cálculo de los puentes de hidrógeno que se presenta en esta sección se ha utilizado un criterio geométrico en el cual el enlace puente de hidrogeno se considera formado cuando éste cumple simultáneamente con los siguientes dos criterios, a saber:

- La distancia entre el aceptor y el protón es menor o igual a 0,24 nm.
- El ángulo O-H-O pertenece al intervalo comprendido entre 0° y 45°.

El criterio utilizado fue el mismo, tanto para el sistema a 1bar como para el sistema a 3kbar.

En la Figura 40 puede observarse la distribución de los puentes de hidrógeno por molécula de agua para el sistema a 1bar (negro) y el sistema a 3kbar (rojo).



Se puede observar rápidamente a partir de dicha gráfica que ha ocurrido una notable redistribución de los puentes de hidrógeno por molécula de agua con el aumento de la presión.

Concretamente se observa que los estados con 1 y 4 puentes ven disminuida su población relativa en condiciones de alta presión.

En el caso del estado con un solo puente establecido, dicha disminución aparece como razonable dado que, al implicar un estado de bajo número de coordinación, debería verse más desfavorecido que el resto de los estados con el aumento de la densidad.

En cuanto al estado con cuatro puentes, esta disminución puede solamente ser interpretada como una pérdida en el grado de estructuración promedio del seno del agua al aumentar de la presión, producto del aumento del número de coordinación (la presión favorecería una estructura hexagonal compacta contra la regular del agua - tetragonal abierta).

Al analizar los estados con 2, 3, y 5 puentes de hidrógeno establecidos, podemos ver que son éstos cuya población se ve favorecida a 3kbar.

Para el caso de los estados con 2 y 3 puentes establecidos sólo cabe mencionar que su aumento relativo es resultado directo y complementario de las disminuciones en los estados con 1 y 4 puentes establecidos.

Para analizar el caso del incremento poblacional del estado con 5 puentes cabe recordar que éste sólo es posible si existe una importante distorsión en todos los puentes formados, en términos de la longitud pero sobre todo en cuanto a los ángulos de los mismos. Así, este aumento poblacional relativo, lejos de interpretarse como un aumento en el establecimiento de la red de puentes de hidrógeno en el seno del agua, debe interpretarse como una pérdida neta en el grado de estructuración.

c. Distribución de distancias de los puentes de hidrógeno a 1bar y 3kbar

En esta sección se presenta el cálculo de la distribución de distancias de los puentes de hidrógeno. Se utilizó para tal fin un criterio geométrico, aunque más abarcativo que el de la sección anterior, en el cual el enlace puente de hidrogeno se considera formado cuando éste cumple simultáneamente con los siguientes dos criterios, a saber:

- La distancia entre el aceptor y el protón es menor o igual a 0,25 nm.
- El ángulo O-H-O pertenece al intervalo comprendido entre 0° y 60°.

El criterio utilizado fue el mismo, tanto para el sistema a 1bar como para el sistema a 3kbar, y se establecieron criterios más abarcativos que los de la sección anterior con el sólo objeto de presentar los efectos de la presión sobre dichas distribuciones con una mayor claridad.

En la Figura 41 puede observarse la distribución de distancias de los puentes de hidrógeno para el sistema a 1bar (negro) y el sistema a 3kbar (rojo).



De la comparación entre ambas curvas de distribución surgen las siguientes tres observaciones.

Por un lado vemos que la posición del máximo resulta invariante para ambos sistemas. Sin embargo el valor del mismo muestra una pequeña disminución para el sistema a 3kbar. Adicionalmente, puede verse un leve angostamiento del pico de la distribución, acompañado de un corrimiento de la curva hacia distancias menores, lo cual resulta solidario con el efecto general del aumento de la densidad con la presión. Por último se puede observar un incremento neto de la distribución para valores mayores que 0,22nm.

Estas últimas tres observaciones deben ser interpretadas conjuntamente como el resultado de un incremento en la dispersión de la distribución de distancias para el sistema a 3kbar. Esto último debe traducirse como un acortamiento de la distancia de los puentes de hidrógeno formados por el sistema de alta presión.

d. Distribución de ángulos de los puentes de hidrógeno a 1bar y 3kbar

En esta sección, análogamente a lo presentado en la sección anterior, se presenta el cálculo de la distribución de ángulos de los puentes de hidrógeno. Se utilizó para tal fin el mismo criterio geométrico que para la sección anterior, en el cual el enlace puente de hidrogeno se considera formado cuando éste cumple simultáneamente con los siguientes dos criterios, a saber:

- La distancia entre el aceptor y el protón es menor o igual a 0,25 nm.
- El ángulo O-H-O pertenece al intervalo comprendido entre 0° y 60°.

Al igual que en la sección anterior, el criterio utilizado fue el mismo, tanto para el sistema a 1bar como para el sistema a 3kbar. También para esta sección vale la aclaración de que se establecieron criterios más abarcativos con el sólo objeto de presentar los efectos de la presión sobre dichas distribuciones con una mayor claridad.

En la Figura 42 puede observarse la distribución de ángulos de los puentes de hidrógeno para el sistema a 1bar (negro) y el sistema a 3kbar (rojo).



De la comparación entre ambas curvas de distribución surgen varias observaciones, las cuales ratifican y amplían lo discutido a partir de la Figura 41 de la sección anterior.

Por un lado vemos que la posición del máximo, a diferencia de lo descripto en la sección anterior, sufre dos cambios notorios. No sólo se desplaza hacia valores algo mayores, sino que sufre una importante disminución para el sistema a 3kbar. Respecto del ancho del pico, éste no se vuelve más agudo sino que presenta más bien los efectos de un aplastamiento general. Por último, el incremento neto que presenta la distribución para el sistema a 3kbar resulta más notorio, además de abarcar una mayor proporción de la distribución, que lo observado para la distribución de distancias.

Si bien estas últimas tres observaciones difieren algo de las presentadas en la sección anterior, al ser interpretadas conjuntamente también deben ser leídas como el resultado de un incremento en la dispersión de la distribución, en este caso de ángulos, para el sistema a 3kbar. Esto último, aún más notablemente que en la sección anterior, debe traducirse como una disminución de la calidad de los puentes de hidrógeno formados por el sistema de alta presión, en este caso en relación al ángulo.

e. Evolución temporal del número total de puentes de hidrógeno a 1bar y 3kbar

Para el cálculo de los puentes de hidrógeno que se presenta en esta sección, al igual que para el estudio de la distribución del los puentes de hidrógeno por molécula de agua, fue utilizado un criterio geométrico en el cual el enlace puente de hidrogeno se considera formado cuando éste cumple simultáneamente con los siguientes dos criterios, a saber:

- La distancia entre el aceptor y el protón es menor o igual a 0,24 nm.
- El ángulo O-H-O pertenece al intervalo comprendido entre 0° y 45°.

El criterio utilizado fue el mismo, tanto para el sistema a 1bar como para el sistema a 3kbar.

En la Figura 43 puede observarse la evolución temporal del número total de puentes de hidrógeno a 1bar (negro) y 3kbar (rojo). Las líneas horizontales representan el valor promedio correspondiente a cada sistema, calculado para los 0,5ns totales de simulación de equilibrio.



Figura 43: Evolución temporal del número total de puentes de hidrógeno a 1bar (negro) y 3kbar (rojo). Las líneas horizontales representan el valor promedio correspondiente a cada sistema, calculado para los 0,5ns totales de simulación de equilibrio.

En este caso la interpretación de la curva resulta simple y directa. Resulta claro que durante casi todo el lapso simulado el número total de puentes de hidrógeno formados fue menor en el sistema a 3kbar que en el sistema a 1bar.

Ambos sistemas muestran un comportamiento de oscilaciones alrededor de un valor de equilibrio propio, lo cual permite suponer que un mejor análisis puede extraerse al comparar los valores promedio para ambas curvas. Al hacer esto lo que se manifiesta resulta ser una disminución en el número total de puentes de hidrógeno de casi el 1%, con el aumento de la presión.

Podemos decir por lo tanto que el aumento de la presión produce una disminución en el número total de puentes de hidrógeno que un determinado número de moléculas de agua es capaz de formar.

4. Conclusiones del capítulo:

A modo de resumen de lo observado y discutido en las secciones previas podemos decir que todos los parámetros relacionados con la estructura del agua, analizados en el presente capítulo, muestran una disminución, desde moderada hasta notable, con el aumento de la presión. Concretamente se ve que, con el aumento de la presión, el agua sufre varios cambios estructurales. Por un lado pierde una moderada proporción de su característica estructural tetraédrica. Por otro lado resulta claro que disminuye notablemente la calidad de los puentes de hidrógeno formados, analizados ya sea desde el punto de vista de la longitud o el ángulo de dicho enlace. Por último se verifica una disminución en el número total de puentes de hidrógeno formados.

Analizando desde esta perspectiva las consecuencias de la alta presión sobre la red de puentes de hidrógeno formada en el seno del agua deberíamos decir que ésta estará formada por un menor número de enlaces, los cuales a su vez serán de menor calidad (al estar más distorsionados), y un menor número de molécula de agua con coordinación tetraédrica. No obstante el número de coordinación resultaría mayor, según lo observado en la figura 38.

En pocas palabras mayor presión sobre el agua significa: menos estructura, menos enlaces y de menor calidad.

En cuanto a la aspiración presentada al comienzo del presente capítulo respecto a establecer una relación más clara entre los cambios estructurales en el seno del agua y el comportamiento del efecto hidrofóbico, parece razonable plantear que, dado que en general es válido decir que el efecto solvofóbico (hidrofóbico en este caso) depende directamente del grado de estructuración presente en el seno del solvente, al volverse éste menos estructurado disminuye consecuentemente el aporte que el mismo es capaz de realizar hacia la estabilización de tal propiedad.

CONCLUSIONES:

Validación de los sistemas estudiados:

Todos los sistemas estudiados mostraron resultados cualitativa e incluso cuantitativamente compatibles con los resultados experimentales disponibles hasta el momento.

En el caso de la lisozima, el modelo reproduce exitosamente una compactación global de la estructura bajo presión, a la vez que reproduce correctamente el comportamiento de aquellas regiones recientemente descriptas como zonas que se alejan de la estructura global bajo los efectos de la alta presión.

También se ha tenido éxito en describir un aumento de la homogeneidad estructural general bajo presión para dicha proteína.

Los resultados experimentales para la apomioglobina bajo presión no permiten comparaciones estructurales directas con la simulación, debido a que ésta se despliega a partir de los 1-2kbar. No obstante, el modelo ha resultado exitoso en la reproducción de un estado globular homogéneo a la presión de 1bar, a la vez se pudieron observar cambios conformacionales "plásticos" a 3kbar, compatibles con los primeros pasos del desplegamiento registrado experimentalmente para la misma presión.

En cuanto al sistema de agua pura, se verificó que dicho modelo reproduce con éxito la pérdida de estructura de la red de puentes de hidrógeno, hecho que ha sido descripto experimentalmente desde hace ya algún tiempo. Por otro lado, también se obtuvieron datos compatibles con los resultados previos existentes en literatura, provenientes de simulaciones realizadas bajo esquemas levemente diferentes al utilizado en estos trabajos.

Adicionalmente se ha establecido que, cuando se comparan escalas de tiempo similares, los resultados obtenidos a partir de los presentes estudios resultan totalmente compatibles con aquellos presentes en literatura, provenientes de estudios previos de dinámica molecular a presiones elevadas.

Diferencias de comportamiento entre los modelos:

Un aspecto notable que resulta del análisis de los resultados es la marcada diferencia de comportamiento entre ambos sistemas proteicos.

Mientras que la lisozima reduce su variabilidad estructural con el aumento de la presión, la apomioglobina, muy por el contrario, comienza a desplegarse sin llegar a estabilizarse en ninguna conformación determinada durante todo el período simulado.

Esta última observación constituye la validación más valiosa para los modelos estudiados, en cuanto a su capacidad para reproducir el comportamiento experimental. Gracias a estudios recientes de RMN-2D sobre soluciones bajo presión, tanto de lisozima como de apomioglobina, hoy sabemos que la presión actúa sobre la lisozima como "selector de confórmeros" [106] mientras que sobre la apomioglobina produce directamente el desplegamiento [118].

Estas diferencias de comportamiento observadas a la presión de 3kbar, tanto experimentalmente como por simulación, podrían explicarse merced a la existencia de los puentes disulfuro presentes en la lisozima pero ausentes en la apomioglobina.

Por un lado, se sabe que la adición de un único puente disulfuro previamente inexistente en la estructura de una proteína [126] puede lograr la estabilización térmica de dicha estructura en valores de hasta 29 °C.

Por otro lado, desde el punto de vista microscópico, la existencia de puentes disulfuro agrega importantes barreras de estrés topológico que deben superarse para que la proteína pueda sufrir cambios conformacionales de envergadura.

Por último, desde el punto de vista mecánico-estadístico, el agregado de restricciones covalentes no-locales, como los puentes disulfuro, disminuye notablemente el número de conformaciones accesibles al estado desplegado, lo cual resulta en una disminución de la ganancia entrópica (conformacional de la proteína) inherente a la transición plegado/desplegado [127,128,129].

A partir de estas últimas observaciones resultaría razonable que una perturbación en la presión del orden de los 3kbar, aunque resulte suficiente para causar el desplegamiento de la apomioglobina, sería insuficiente para desestabilizar a la lisozima.

Estructura secundaria:

Hoy sabemos que la estructura secundaria que presentan las proteínas no puede ser explicada solamente a partir de la secuencia, sino que es necesario tomar en cuenta las interacciones no-locales que se producen luego de que la proteína se pliega.
En este sentido, no resulta sorprendente que en ninguno de los dos sistemas proteicos se hayan observado cambios de envergadura en torno a las estructuras secundarias presentes en los estados nativos correspondientes.

En el caso de la lisozima, dado que los cambios sufridos por parte de la estructura tridimensional resultan pequeños, no se observan modificaciones sustanciales de las asociaciones no-locales, y por lo tanto resulta razonable la invariabilidad general de la estructura secundaria con la presión.

Sin embargo, en el caso de la apomioglobina, a pesar de que se presentan importantes cambios en las asociaciones no-locales, estos no se traducen en perturbaciones sostenidas y evidentes de la estructura secundaria. Esta observación resultaría satisfactoriamente explicada si suponemos que el cambio estructural observado en el tiempo cubierto por la simulación no resulta suficiente para desestabilizar a la estructura secundaria. De todos modos no puede descartarse la hipótesis de que el estado desplegado bajo presión podría no ser la de un ovillo estadístico.

El factor común entre ambos modelos:

El resultado que surge fuertemente a partir de los resultados presentados en esta tesis como elemento unificador es el comportamiento de la superficie accesible al solvente.

Para el caso de la lisozima se observa una clara y sostenida inversión de la relación entre las superficies accesibles hidrofóbica e hidrofílica.

Concretamente, mientras que a presión atmosférica resulta preponderante la superficie hidrofílica por sobre la hidrofóbica, a 3kbar esta situación se ve revertida, siendo la superficie hidrofóbica la primordialmente expuesta.

Al observar el comportamiento de la apomioglobina, el aumento sostenido de la superficie hidrofóbica accesible al solvente se presenta indiscutiblemente como la variable más correlacionada con los cambios conformacionales en el tiempo. Puede plantearse razonablemente entonces que su incremento gobierna el proceso de desplegamiento.

Así, se puede interpretar que ambos sistemas muestran una disminución del efecto hidrofóbico con el aumento de la presión. También se pueden reinterpretar las diferencias de comportamiento entre ambos sistemas esencialmente sobre la base de la presencia de los cuatro puentes disulfuro en la lisozima, restricciones éstas que están ausentes en la apomioglobina.

Concretamente, mientras que la lisozima sólo puede reaccionar elásticamente a la perturbación de la presión a causa de sus puentes disulfuro, la apomioglobina por el contrario tiene libertad de explorar diferentes conformaciones, las cuales le resultan accesibles a 3kbar debido al mencionado debilitamiento que sufren las interacciones hidrofóbicas a presiones de esta magnitud, mostrando así una reacción conformacional plástica frente a la misma perturbación.

Cambios en la interacción hidrofóbica:

La interacción hidrofóbica es de naturaleza puramente entrópica, y su prevalencia depende del grado de estructuración existente en el seno del agua (el efecto solvofóbico no existe en solventes que carecen de estructura).

Por lo tanto, no basta con observar lo que sucede con la superficie hidrofóbica accesible al solvente, sino que resulta necesario el análisis de los cambios producidos en el seno del solvente.

Al estudiar los efectos de la presión sobre la estructura del agua y la calidad de la red de puentes de hidrógeno formada, aparecen resultados tan significativos como los previamente discutidos para los modelos proteicos.

Concretamente, resulta evidente, a partir del estudio del modelo de agua pura a 3kbar, que ésta presenta no sólo forma menos puentes de hidrógeno, sino que éstos son de menor calidad que los formados a 1bar, tanto desde el punto de vista de sus longitudes como de sus ángulos.

Por otro lado, los resultados de la distribución del número de puentes de hidrógeno por molécula de agua muestran una fuerte disminución de la coordinación tetraédrica, lo cual implica que las redes de puentes de hidrógeno formadas de a 3kbar se encuentran distorsionadas.

Así, podemos concluir razonablemente que dado que el grado de estructuración presente en el seno del agua disminuye con la presión, en estas condiciones el aporte entrópico que la misma puede realizar al efecto hidrofóbico también se ve disminuido.

Cambios de volumen:

Una observación interesante de los trabajos presentados, que no ha sido señalado previamente, es que el volumen de todos los sistemas estudiados responde completamente a las perturbaciones de presión aplicadas, alcanzando el equilibrio, en una escala de tiempo inferior al nanosegundo.

Indudablemente los cambios producidos en cualquier sistema que sea sometido a un aumento de presión deberán obedecer el principio de la búsqueda del menor volumen. Sin embargo, no parecería razonable fundamentar a partir de esta última afirmación la aparición de los cambios conformacionales presentados en esta tesis, dado que éstos son eventos que se producen en una escala de tiempos superior en uno a dos órdenes de magnitud a la necesaria para que el volumen alcance su valor de equilibrio una vez aplicada la presión.

Estabilización térmica por presión:

En este punto cabría aclarar que ninguna de las conclusiones expuestas hasta aquí permiten explicar el hecho de que, a presiones moderadas, pero por encima de la atmosférica (aprox. 10²bar), la mayoría de las proteínas globulares logran estabilizar perturbaciones térmicas leves (40-60 °C) [17].

En este sentido es importante remarcar que los estudios presentados en esta tesis han sido diseñados a fin de trabajar dentro del rango de presiones en el cual la estructura "normal" del agua se ve perturbada.

Es posible sin embargo que a presiones moderadas los efectos observados experimentalmente se deban a factores diferentes que los analizados aquí, especialmente si bajo tales condiciones experimentales no se registran en el seno del agua cambios estructurales de relevancia.De hecho, si la estructura del agua no se perturba (como ocurre a 100bar) las relaciones termodinámicas referentes a la estabilidad de la interacción hidrofóbica con la presión son perfectamente validas. Es decir más presión, menor volumen, y por lo tanto refuerzo de la interacción hidrofóbica debería resultar en este caso un planteo totalmente válido.

Palabras Finales

Por último, correspondería resumir los resultados presentados, su análisis, discusión y conclusiones de la siguiente manera:

El principal efecto ocasionado por la aplicación de altas presiones sobre las proteínas en solución es el de la disminución del grado de estructuración en el seno del agua. Esto genera, a su vez, una disminución del efecto hidrofóbico. Cuando esta disminución ocasiona una pérdida de estabilidad en la estructura nativa lo suficientemente grande como para superar la barrera de activación para el cambio conformacional, ésta se constituye entonces en la fuerza motriz de tales cambios, produciendo como resultado el desplegamiento.

Capítulo 7

PERSPECTIVAS:

Estudio detallado de la hidratación de las proteínas en solución.

Si bien el trabajo presentado en esta tesis ha dejado prácticamente de lado el estudio de la hidratación de las proteínas en el sentido del comportamiento específico de cada residuo, han aparecido muchos resultados que permiten depositar expectativas en la dinámica molecular como herramienta para el estudio de la hidratación de las proteínas, con especial interés en la hidratación hidrofóbica y su relación con la presión.

Contar con resultados con resolución a escala atómica de tales procesos nos permitiría acceder a una gran cantidad de elementos en relación con los mecanismos moleculares que intervienen en la estabilidad de las proteínas, particularmente en relación con aquellos factores de naturaleza esencialmente entrópica.

Construcción de un modelo de hidratación de proteínas abarcativo, que solucione las deficiencias de los disponibles en la actualidad.

Seguramente, a partir de estudios como los referidos en el párrafo anterior, resultaría factible comenzar con la elaboración de un modelo de hidratación que logre interpretar acabadamente el complejo problema de las proteínas en solución.

Está claro que éste resulta un enorme desafío, y que no podrá ser abordado exclusivamente desde modelos teóricos o computacionales, sino que deberá plantearse una relación de retroalimentación dialéctica entre teoría, modelos y experimentación que posibilite la acumulación suficiente de conocimientos que nos permitan la elaboración de una teoría lo suficientemente ajustada como para proveernos de capacidad predictiva, no sólo en cuanto al comportamiento de una determinada estructura proteica frente a una perturbación térmica o de barométrica, sino esencialmente para poder predecir la estructura que una determinada secuencia proteica tenderá a formar bajo ciertas condiciones.

Bibliografía

- ¹ Wu, H. (1929) Am. J. Physiol. 90, 562.
- ² C. Levinthal, How to Fold Graciously. Proc. Mossbauer Spect. in Biol. Systems; (1969) Illinois Univ. Press 22-24
- ³ M. Daune, Molecular Biophysics; (1999) Oxford Univ. Press
- ⁴ Y. Ofran, M. Punta, R. Schneider and B. Rost, (2005) DDT 10:1475-1482.
- ⁵ S.D. Hamann, Rev. Phys. Chem. Jpn. 50 (1980) 147–168.
- ⁶ T.V. Chalikian, A.P. Sarvazyan, K.J. Breslauer, J. Phys. Chem. 97 (1993) 13017–13026.
- ⁷ J.W. Linowski, N.-I. Liu, J. Jonas, J. Chem. Phys. 65 (1976) 3383–3384.
- ⁸ J.W. Linowski, N.-I. Liu, J. Jonas, J. Magn. Reson. 23 (1976) 455–460.
- ⁹ K. Goossens, L. Smeller, K. Heremans, J. Chem. Phys. 99 (1993) 5736– 5741.
- ¹⁰ H. Li, H. Yamada, K. Akasaka, Biochemistry 37 (1998) 1167–1173.
- ¹¹ B.B. Boonyaratanakornkit, C.B. Park, D.S. Clark, Biochim. Biophys. Acta 1595 (2002) 235–249.
- ¹² D.J. Hei, D.S. Clark, Appl. Environ. Microbiol. 60 (1994) 932–939.
- ¹³ P.C. Michels, D. Hei, D.S. Clark, Adv. Protein Chem. 48 (1996) 341–376.
- ¹⁴ W. Kauzmann, Nature 325 (1987) 763–764.
- ¹⁵ K.J. Frye, C.A. Royer, Protein Sci. 7 (1998) 2217–2222.
- ¹⁶ S. Kunugi, N. Tanaka, Cold denaturation of proteins under high pressure. BBA 1595 (2002) 329-344.
- ¹⁷ K. Heremans, L. Smeller, Protein structure and dynamics at high pressure. BBA 1386 (1998) 353-370.

- ¹⁸ Fleischmann, R.D. et al. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269, 496–512.
- ¹⁹ Goffeau, A. et al. (1996) Life with 6000 genes. Science 274, 546–567.
- ²⁰ C. elegans Sequencing Consortium. (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. Science 282, 2012–2018.
- ²¹ Lander, E.S. et al. (2001) Initial sequencing and analysis of the human genome. Nature 409, 860–921.
- ²² Venter, J.C. et al. (2001) The human genome. Science 291, 1304–1351.
- ²³ Venter, J.C. et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304, 66–74.
- ²⁴ Carrell, R.W. and Lomas, D.A. (1997) Lancet 350, 134-138.
- ²⁵ Kelly, J.W. (1996) Curr. Opin. Struct. Biol. 6, 11-17.
- ²⁶ Thomas, P.J., Qu, B.-H. and Pedersen, P.L. (1995) Trends Biochem. Sci. 20, 456-459.
- ²⁷ Soto, C. (1999) J. Mol. Med. 77, 412-418.
- ²⁸ Carrell, R.W. and Gooptu, B. (1998) Curr. Opin. Struct. Biol. 8,799-809.
- ²⁹ Soto, C. Protein misfolding and disease; protein refolding and therapy. FEBS Letters 498 (2001) 204-207.
- ³⁰ Soto, C. (1999) CNS Drugs 12, 347-356.
- ³¹ Ptitsyn, O.B. y Volkenstein, M.V. (1986) Protein structures and neutral theory of evolution. J. Biomol. Struct. Dynamics, 4:137-156.
- ³² Afinsen, C. B. (1973) Science 181, 223.
- ³³ Hill, T. L. (1960) en Introduction to Statistical Thermodynamics, Addison-Wesley, Reading, MA.
- ³⁴ Cohn, E. J., et al. (1933) J. Biol. Chem. 100, 3.
- ³⁵ Mirsky, A. E., & Pauling, L. (1936) Proc. Natl. Acad. Sci. U.S.A. 22, 439.
- ³⁶ Eyring, H., & Stearn, A. E. (1939) Chem. Rev. 24, 253.

- ³⁷ Jacobsen, C. F., & Linderstrom-Lang, K. (1949) Nature 164, 411.
- ³⁸ Zipp, A. & Kauzmann, W. (1973) Biochemistry 12, 4217.
- ³⁹ Edelhoch, H., & Osborne, J. C. (1976) Adv. Protein Chem. 30, 183.
- ⁴⁰ Brandts, J. F., Oliveira, R. J., & Westort, C. (1970) Biochemistry 9, 1038.
- ⁴¹ Pauling, L. (1960) en The Nature of the Chemical Bond, 3rd ed., Cornell University Press, Ithica, NY.
- ⁴² Vinogradov, S. N., & Linnell, R. H. (1971) en Hydrogen Bonding, Van Nostrand Reinhold, New York.
- ⁴³ Pauling, L., Corey, R. B., & Branson, H. R. (1951) Proc. Natl. Acad. Sci. U.S.A. 37, 205.
- ⁴⁴ Poland, D. C., & Scheraga, H. A. (1970) Theory of the Helix-Coil Transition, Academic Press, New York.
- ⁴⁵ Zimm, B. H., et al. (1959) Proc. Nail. Acad. Sci. U.S.A. 45, 1601.
- ⁴⁶ Kauzmann, W. (1954) en The Mechanism of Enzyme Action (McElroy, W. D., & Glass, B., Eds.) p70, Johns Hopkins Press, Baltimore, MD.
- ⁴⁷ Kauzmann, W. (1959) Adv. Protein Chem. 14, 1.
- ⁴⁸ Christensen, J. K. (1952) C. R. Trau. Lab. Carlsberg, Ser. Chim. 28, 37.
- ⁴⁹ Nozaki, Y., & Tanford, C. (1971) J. Biol. Chem. 246, 2211.
- ⁵⁰ Kyte, J., & Doolittle, R. F. (1982) J. Mol. Biol. 157, 105.
- ⁵¹ Fauchere, J.-L., & Pliska, V. E. (1983) Eur. J. Med. Chem.-Chem. Therm. 18, 369.
- ⁵² Rose, G. D., et al. (1985) Science 229, 834.
- ⁵³ Kauzmann, W. (1987) Nature 325, 763.
- ⁵⁴ V.A. Payne, N. Matubayasi, L.R. Murphy, R.M. Levy, Monte Carlo study of the efect of pressure on hydrophobic association, J. Phys. Chem. B 101 (1997) 2054-2060., 18
- ⁵⁵ G. Hummer, S. Garde, A.E. Garcia, M.E. Paulatis, L.R. Pratt, The pressure dependence of hydrophobic interactions is consistent with the

observed pressure denaturations of proteins, Proc. Natl. Acad. Sci. USA 95 (1998) 1552-1555.

- ⁵⁶ A.N. McCarthy and J.R. Grigera, Effect of pressure on the conformation of proteins. A molecular dynamics simulation of lysozyme. Journal of Molecular Graphics and Modelling, 24 (January 2006) 254-261.
- ⁵⁷ A.N. McCarthy and J.R. Grigera, Pressure denaturation of apomyoglobin: A molecular dynamics simulation study. Biochim Biophys Acta. 2005 Dec 27; [Epub ahead of print]
- ⁵⁸ E. Paci, High pressure simulations of biomolecules. BBA 1595 (2002) 185-200.
- ⁵⁹ Marqusee, S., & Baldwin, R. L. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 8898.
- ⁶⁰ Marqusee, S., Robbins, V. H., & Baldwin, R. L. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 5286.
- ⁶¹ Sueki, M., et al. (1984) Macromolecules 17, 148.
- ⁶² Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) Biochemistry 27, 7167.
- ⁶³ Dyson, H. J. et al. (1985) Nature 318, 480.
- ⁶⁴ Dyson, H. J. et al. (1988) J. Mol. Biol. 201, 161.
- ⁶⁵ Goodman M. et al. (1969) Proc. Natl. Acad. Sci. U.S.A. 64, 444.
- ⁶⁶ Zimm B. H. & Bragg J. K. (1959) J. Chem. Phys. 31, 526.
- ⁶⁷ Srinivasan, R. (1976) Indian J. Biochem. Biophys. 13, 192.
- ⁶⁸ Levitt M., & Greer, J. (1977) J. Mol. Biol. 114, 181.
- ⁶⁹ Kabsch W. & Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. (1983) Biopolymers 22, 2577.
- ⁷⁰ Chou P. Y. & Fasman G. D. (1978) Adv. Enzymol. 47, 45.
- ⁷¹ K. A. Dill, Dominant Forces in Protein Folding. Biochemistry, (1990) 29 (31):7133-7155
- ⁷² M. Cieplak, J.I. Sułkowska, J. Chem. Phys. 123, 194908 (2005)

- ⁷³ J. Zeng, H.R. Treutlein, T. Simonson, Proteins-SFG 35:89-100 (1999) BiophJ 77:505-515 (1999)
- ⁷⁴ C.A. Sotriffer, O. Krämer, G. Klebe, Proteins-SFB 56:52-66 (2004)
- ⁷⁵ G.H. Peters, T.M. Frimurer, J.N. Andersen, O.H. Olsen, Bioph. J. 77:505-515 (1999)
- ⁷⁶ B.L. de Groot, Science (2001) 294(5550):2353-2357.
- ⁷⁷ D.B. Kitchen, L.H. Reed, R.M. Levy, Biochemistry 31 (1992) 10083– 10093.
- ⁷⁸ E. Paci, M. Marchi, Constant-pressure molecular dynamics techniques applied to complex molecular systems and solvated proteins. J. Phys. Chem. 100 (1996) 4314–4322.
- ⁷⁹ R.M. Brunne, W.F. van Gunsteren, FEBS Lett. 323 (1993) 215–217.
- ⁸⁰ P.E. Hunenberger, A.E. Mark, W.F. van Gunsteren, Computational approaches to study protein unfolding: hen egg white lyzozyme as a case study. Proteins 21 (1995) 196–213.
- ⁸¹ M. Marchi, K. Akasaka, J. Phys. Chem., B 105 (2001) 711–714.
- ⁸² van Gunsteren, W. F., Berendsen, H. J. C. Algorithms for macromolecular dynamics and constraint dynamics. Mol. Phys. 34:1311– 1327, 1977.
- ⁸³ Berendsen, H. J. C. Electrostatic interactions. In: Computer Simulation of Biomolecular Systems. van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J. eds. . ESCOM Leiden 1993 161–181.
- ⁸⁴ Morse, P. M. Diatomic molecules according to the wave mechanics. II. vibrational levels. Phys. Rev. 34:57–64, 1929.
- ⁸⁵ Hockney, R. W., Goel, S. P. J. Comp. Phys. 14:148, 1974.
- ⁸⁶ Verlet., L. Phys. Rev. 34:1311–1327, 1967.
- ⁸⁷ Berendsen, H. J. C. and van Gunsteren, W. F. (1986). Practical algorithms for dynamic simulations. En: Ciccotti, G. and Hoover, W., editors, Molecular-Dynamics Simulation of Statistical-Mechanical Systems. North-Holland, Amsterdam.

- ⁸⁸ Berendsen, H. J. C., Postma, J. P. M., DiNola, A., Haak, J. R. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81:3684–3690, 1984.
- ⁸⁹ Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. Mol. Phys. 52:255–268, 1984.
- ⁹⁰ Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. Phys. Rev. A 31:1695–1697, 1985.
- ⁹¹ Berendsen, H. J. C. Transport properties computed by linear response through weak coupling to a bath. En: Computer Simulations in Material Science. Meyer, M., Pontikis, V. eds. Kluwer 1991 139–155.
- ⁹² Parrinello, M., Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys. 52:7182–7190, 1981.
- ⁹³ Nosé, S., Klein, M. L. Constant pressure molecular dynamics for molecular systems. Mol. Phys. 50:1055–1076, 1983.
- ⁹⁴ Ryckaert, J. P., Ciccotti, G., Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. J. Comp. Phys. 23:327–341, 1977.
- ⁹⁵ Miyamoto, S., Kollman, P. A. SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. J. Comp. Chem. 13:952–962, 1992.
- ⁹⁶ Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M. LINCS: A linear constraint solver for molecular simulations. J. Comp. Chem. 18:1463–1472, 1997.
- ⁹⁷ E. Lindahl, B. Hess, D. van der Spoel, J. Mol. Mod. 7 (2001) 306–317.
- ⁹⁸ H.J.C. Berendsen, D. van der Spoel, R. van Drunen, Comput. Phys. Commun. 91 (1995) 43–56.
- ⁹⁹ W.F. van Gunsteren, H.J.C. Berendsen, Biomos BV Nijenborgh 4, 9747 AG Groningen, the Netherlands, 1987.
- ¹⁰⁰ A.R. van Buuren, S.J. Marrink, H.J.C. Berendsen, Phys. Chem. 97 (1993) 9206–9212.
- ¹⁰¹ A.E. Mark, S.P. van Helden, P.E. Smith, L.H.M. Janssen, W.F. van Gunsteren, Am. Chem. Soc. 116 (1994) 6293–6302.

- ¹⁰² W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926–935.
- ¹⁰³ A.R. van Buuren, H.J.C. Berendsen, Biopolymers 33 (1993) 1159–1166.
- ¹⁰⁴ H. Liu, F. Müller-Plathe, W.F.J. van Gunsteren, Am. Chem. Soc. 117 (1995) 4363–4366.
- ¹⁰⁵ I.G. Tironi, R. Sperb, P.E. Smith, W.F. van Gunsteren, J. Chem. Phys. 102 (1995) 5451–5459.
- ¹⁰⁶ M. Refaee, T. Tezuka, K. Akasaka, M.P. Williamson, J. Mol. Biol. 327 (2003) 857–865.
- ¹⁰⁷ H.J.C. Berendsen, J.R. Grigera, T.P.J. Straatsma, Phys. Chem 91 (1987) 6269–6271.
- ¹⁰⁸ K. Sasahara, M. Sakurai, K. Nitta, Protein Struct. Funct. Genet. 44 (2001) 180–187.
- ¹⁰⁹ K. Akasaka, Biochemistry 42 (2003) 10875–10885.
- ¹¹⁰ M.M.C. Sun, N. Tolliday, C. Vetriani, F.T. Robb, D.S. Clark, Protein Sci. 8 (1999) 1056–1063.
- ¹¹¹ N. Tanaka, C. Ikeda, K. Kanaori, K. Hiraga, T. Konno, S. Kunugi, Biochemistry 39 (2000) 12063–12068.
- ¹¹² N. Tanaka, D. Mitani, S. Kunugi, Biochemistry 40 (2001) 5914–5920.
- ¹¹³ P. Cioni, G.B. Strambini, J. Mol. Biol. 242 (1994) 291–301.
- ¹¹⁴ Frank Eisenhaber, Philip Lijnzaad, Patrick Argos, Chris Sander, Michael Scharf, J. Comp. Chem. 16 (1995) 273–284.
- ¹¹⁵ M.A. Anisimov, J.V. Sengers, J.M.H. Levelt Sengers, in: D.A. Palmer, R. Fernandez-Prini, A.H. Harvey (Eds.), Physical Chemistry in Water, Steam and Hydrothermal Solutions, Elsevier, Amsterdam, 2004.
- ¹¹⁶ E.R. Caffarena, J.R. Grigera, On the Hydrogen Bond Structure of Water at different densities. Physica A 342 (2004) 34–39.
- ¹¹⁷ D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen, Gromacs User Manual version 3.2, www.gromacs.org (2004).

- ¹¹⁸ Ryo Kitahara, Hiroaki Yamada, Kazuyuki Akasaka and Peter E. Wright, High Pressure NMR Reveals that Apomyoglobin is an Equilibrium Mixture from the Native to the Unfolded, Journal of Molecular Biology, 320 (2002) 311-319.
- ¹¹⁹ Y. Q. Cai, H.-K. Mao, P. C. Chow, J. S. Tse, Y. Ma, S. Patchkovskii, J. F. Shu, V. Struzhkin, R. J. Hemley, H. Ishii, C. C. Chen, I. Jarrige, C. T. Chen, S. R. Shieh, E. P. Huang, and C. C. Kao, Ordering of Hydrogen Bonds in High-Pressure Low-Temperature H2O, Phys. Rev. Lett. 94 (2005) 025502.
- ¹²⁰ M. I. Marques, J. M. Borreguero, H. E. Stanley, and N. V. Dokholyan, A Possible Mechanism for Cold Denaturation of Proteins at High Pressure, Phys. Rev. Lett. 91 (2003) 138103.
- ¹²¹ F.W. Starr, F. Sciortino, H.E. Stanley, Dynamics of simulated water under pressure, Phys. Rev. E 60 (1999) 6757.
- ¹²² F.W. Starr, S.T. Harrington, F. Sciortino, H.E. Stanley, Slow dynamics of water under pressure. Phys. Rev. Lett. 82 (1999) 3629.
- ¹²³ P.A. Netz, F. Starr, H.E. Stanley, M. Barbosa, J. Chem. Phys. 115 (2001) 344.
- ¹²⁴ M. Groß, R. Jaenicke, Pressure-induced dissociation of tight couple ribosomes, FEBS Lett. 267 (1990) 239.
- ¹²⁵ M. Groß, R. Jaenicke, Growth inhibition of lysozyme crystals at high hydrostatic pressure, FEBS Lett. 284 (1991) 87.
- ¹²⁶ Alber, T. (1989) Annu. Rev. Biochem. 58, 765.
- ¹²⁷ Chan, H.S., & Dill, K.A. (1989) J. Chem. Phys. 90, 492.
- ¹²⁸ Pace, C. N., Grimsley, G. R., Thomson, J. A., & Barnett, B. J. (1988) J. Biol. Chem. 263, 11820.
- ¹²⁹ Matsumura, M., Matthews, B.W., Levitt, M., & Becktel, W. J. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 6562.