

MÉTODOS DE SIMULACIÓN MOLECULAR: UNA REVISIÓN DE LAS HERRAMIENTAS MÁS ACTUALES

Johnny Daniel Alfaro Pérez

Resumen

El siguiente trabajo presenta una revisión de los métodos de simulación molecular más utilizados por científicos, ingenieros e informáticos; de esta manera, se hace una breve descripción del fundamento matemático de cada método. Además, se profundiza en los métodos utilizados para la simulación molecular de la estructura del ADN, los detalles necesarios para su simulación y una reseña de algunas de las herramientas (software) disponibles actualmente para la simulación de esta macromolécula y el análisis de resultados.

Palabras clave: simulación molecular; modelos; software; estructura de ADN.

Abstract

Molecular simulation methods, which are mostly used by scientists, engineers and informatics, are presented with a brief description of the mathematical fundamentals of each method. Furthermore, DNA structure's molecular simulation methods are discussed extensively, including several necessary details for DNA simulation, and a review is provided for some of the most updated software for this macromolecule's simulation and data analysis.

Keywords: molecular simulation; models; software; DNA structure.

Recibido: 22 de abril de 2013 • **Aprobado:** 10 de abril de 2014

1. INTRODUCCIÓN

Los métodos de simulación molecular nacieron como una necesidad en distintas áreas de la ciencia con el fin de estudiar aquellas moléculas que no se pueden observar por medio de microscopía electrónica o que presentan tiempos de equilibrios en escalas tan pequeñas del orden de nanosegundos que no se pueden medir en el laboratorio, o que simplemente su dinámica molecular puede ser entendida con mayor claridad por medio de métodos numéricos y herramientas computacionales.

Entre algunas de las aplicaciones en que los métodos de simulación molecular han sido de gran ayuda, se pueden mencionar el estudio de macromoléculas como ADN, proteínas (desde

su estructura primaria hasta la cuaternaria), modelización de nuevos materiales e inclusive cristales líquidos.

A pesar de que tradicionalmente se ha dependido de sus resultados para poder parametrizar y validar los modelos de simulación molecular, lo que implica que la simulación ha seguido a la experimentación, se espera que en un corto tiempo dichos roles sean intercambiados, es decir, que la simulación ayude a predecir resultados experimentales (con ello reduciendo la cantidad de recursos que se necesitan para la experimentación y el tratamiento de desechos) de manera confiable, y que dichos esfuerzos fructifiquen en otras áreas del conocimiento y que resuelvan problemas tecnológicos.

Por ejemplo, se sabe a partir de la ley de Moore, que la cantidad de transistores de silicio que se pueden colocar en un microprocesador crece de manera exponencial con el tiempo (Cotter, 2011); sin embargo, existe una limitación física con respecto a la cantidad de transistores que se pueden colocar y al proceso de miniaturización, por lo que, se espera a mediano plazo haber alcanzado el límite superior.

Es por ello, que se ha trabajado en resolver este problema a partir de la nanotecnología. Una solución que se genera a partir del conocimiento de la física de los electrones y de las simulaciones moleculares, es el uso de las computadoras cuánticas, que utilizan los principios de la espintrónica (IME, 2013), de manera que las computadoras cuánticas están más cerca de convertirse en una realidad. Se están empezando a realizar los primeros esfuerzos para el desarrollo de microprocesadores utilizando estos conocimientos, y en este proceso los modelos de simulación molecular han tenido un papel importante.

Otra posibilidad que se abre en este procesamiento de datos consiste en el uso de la tecnología del ADN, en el cual se utiliza el modelo de doble hélice de Watson-Crick para la construcción de las series de datos que se puedan almacenar, aprovechando que esta macromolécula se puede autoensamblar (Nelson y Cox, 2008). Sin embargo, para lograr esto se debe tener un conocimiento básico de la biofísica involucrada en la hibridación de ADN, y un mejor desarrollo de técnicas para su estudio, como el uso de multirreglos, espectroscopía de fuerza unimolecular y empacamiento viral de ADN (Knotts, et al, 2007), por lo que, dichos estudios son complementados y reforzados utilizando las técnicas de la simulación molecular.

De aquí nace la importancia de revisar aquellas herramientas que están siendo más utilizadas para el estudio de estas moléculas, así como también la revisión de otros softwares que estarán disponibles en un futuro cercano que simplificarán las simulaciones. Además, es importante comprender los fundamentos por los cuales funcionan los distintos métodos de simulación molecular, ya que, de lo contrario, se estaría realizando un proceso como si fuera una caja negra, donde por “arte de magia” aparecen los resultados y no se conocen las restricciones del modelo.

Es por ello, que en este trabajo se realiza una revisión de los dos métodos de simulación molecular más importantes (que se basan en principios matemáticos distintos), así como un algoritmo relativamente nuevo que ha permitido una mejora significativa de los programas de simulación, y posteriormente, se profundiza en aspectos importantes para la simulación molecular de ADN como limitaciones actuales, modelos utilizados y herramientas para el análisis de datos.

2. MÉTODO MONTECARLO

Este es un método estocástico, es decir, que se guía por pasos al azar. Uno de los métodos Montecarlo más utilizados es el muestreo de Metrópolis, donde se hace un recorrido aleatorio de acuerdo a valores de importancia predeterminados (como el criterio del factor de Boltzmann) para obtener un muestreo más rápido y eficiente.

Una de las condiciones del método es la hipótesis ergódica, donde se parte del supuesto que el promedio de todas las coordenadas iniciales del espacio fase (espacio de todos los estados de un sistema) es equivalente al promedio de las coordenadas del espacio fase en un tiempo evolucionado, por lo que, cada punto accesible en el espacio de configuración puede ser alcanzado en un número finito de pasos Montecarlo a partir de otro punto, lo cual implica que el método se basa en hacer un promedio de dichas condiciones para todo el ensamble (Frenkel y Smit, 2002).

Los ensambles que se utilizan en este método son los utilizados para realizar las descripciones de la mecánica estadística, como el ensamble canónico (NVT), el microcanónico (NVE) o el gran canónico (μVT) (Atkins y de Paula, 2010). El ensamble más utilizado en el método Montecarlo se basa en el número fijo de partículas (N), con un volumen (V) a una temperatura (T) dada, aunque se puede simular en otros ensambles utilizando distintas herramientas, por ejemplo, el termostato de Berendsen o el de Nosé-Hoover (Frenkel y Smit, 2002).

Además, otra de las condiciones que se debe cumplir en este método es el balance detallado, en el cual la posibilidad de moverse

a un nuevo estado debe ser igual a la posibilidad que se tiene de regresar al estado actual a partir del nuevo estado que se puede alcanzar (Frenkel y Smit, 2002).

Este método se puede considerar como un proceso de Markov, es decir, un proceso aleatorio que carece de memoria, donde las predicciones de un evento futuro (siguiente paso) sólo están basadas en las condiciones actuales y no en condiciones pasadas.

Una de las limitaciones del método, es que con mucha frecuencia la energía libre de Gibbs ha sido calculada a partir de promedios de propiedades indirectas que pueden ser determinadas durante la simulación; sin embargo, un nuevo enfoque es que la energía libre sea el parámetro que guíe el muestreo del espacio de configuración, esto ha traído como consecuencia un muestreo más eficiente y estimaciones directas de la energía libre (Singh, et al, 2012).

Asimismo, se han desarrollado otros métodos Montecarlo basados en la densidad de estados, que describen el número de estados por intervalo de energía en el cual cada nivel de energía se encuentra disponible (Frenkel y Smit, 2002), de manera que conforme aumenta la densidad de estados, aumenta el número de niveles disponibles.

Algunos de estos nuevos métodos son: sesgado configuracionalmente, con temperatura configuracional, el uso de ventanas múltiples o el uso de algoritmos de viajes redondos. Una descripción de estos nuevos métodos se puede consultar en Singh, et al (2012), mientras que los métodos más tradicionales con algoritmos computacionales se pueden consultar en Frenkel y Smit (2002).

3. MÉTODO DINÁMICA MOLECULAR

Es una técnica que permite calcular las propiedades de transporte y equilibrio de sistemas clásicos de muchas partículas, la cual enfatiza que el movimiento de las partículas obedece las leyes de la mecánica clásica; por ello, se sigue la evolución del sistema con el tiempo para cada partícula conforme se resuelven las ecuaciones de movimiento de Newton (Singh, et al, 2012). Los tiempos de simulación actuales varían desde

los femtosegundos hasta los milisegundos (Salomon, et al, 2012).

Los pasos en este método son los siguientes:

Minimización de energía: normalmente las estructuras iniciales que se crean en los simuladores poseen energías mucho mayores a las que tendría un objeto real, por esta razón, se utilizan algoritmos para calcular las posiciones y fuerzas originales, con el objetivo de minimizarlas y que sean más realistas.

Equilibrio: se resuelven las ecuaciones de movimiento del sistema hasta que las propiedades del sistema no cambian más con el tiempo, como la energía total.

Producción: son las mediciones que se utilizan para los cálculos computacionales, puesto que el sistema ya se encuentra en el equilibrio.

El método de dinámica molecular también se basa en el balance detallado, la hipótesis ergódica y que el promedio en el tiempo no depende de las condiciones iniciales de la simulación, lo cual generalmente puede no ser cierto (Frenkel y Smit, 2002).

Debido a que los cálculos del método de Dinámica Molecular se realizan en los conjuntos clásicos de la mecánica estadística como NPT y NVT, no se puede predecir el comportamiento exacto que tendrá el sistema en cualquier momento, sino más bien se obtienen predicciones estadísticas (Frenkel y Smit, 2002).

Más detalles acerca de simulaciones en otros ensambles menos comunes, así como algoritmos para su implementación y ejemplos, se pueden encontrar en Frenkel y Smit (2002).

4. ALGORITMO DE METADINÁMICA

Este algoritmo originalmente propuesto por Laio y Parinello (2002), se basa en la reducción dimensional y consiste en un muestreo uniforme del espacio de configuración por medio de cierto número finito de variables colectivas, con la finalidad de reconstruir la energía libre de una forma más exacta (Laio y Gervasio, 2008). Este método fue propuesto para acelerar la simulación de eventos que se consideran raros, ya que usando

los algoritmos usuales de Montecarlo y Dinámica Molecular se encuentran problemas, como por ejemplo, pasar de un mínimo local profundo a otro estado (Singh, et al, 2012).

El algoritmo básico es el siguiente (Laio y Gervasio, 2008):

- Selección apropiada de un conjunto de variables colectivas, que pueden ser un ángulo, una distancia, un número de coordinación o la energía potencial.
- Selección óptima de los parámetros.
- Control del error y aseguración de la convergencia.

Dichas variables colectivas deben cumplir ciertas propiedades; por ejemplo: deben ser claramente distinguibles entre el estado inicial, intermedio y final, deben describir todos los eventos lentos que son relevantes para el proceso y sus valores numéricos no deben ser muy grandes (Laio y Gervasio, 2008).

Uno de los problemas que se han reportado con este método es que la cantidad óptima de

variables colectivas ronda entre 2 y 3; además, que si una variable de importancia no es tomada en cuenta durante la simulación, los resultados presentan histéresis y no convergen.

Una visualización del método se puede hacer de la siguiente manera: en una simulación de Dinámica Molecular tradicional, ésta puede quedar atrapada en un mínimo local de energía, por lo que, como solución se utiliza el método porque añade cada cierto tiempo un pequeño potencial Gaussiano repulsivo que ayude al sistema a escapar de este mínimo (Singh, et al, 2012). La altura y espesor del Gaussiano se define de acuerdo a las variables colectivas, de forma tal que, se van superando las “montañas” como las que aparecen en la Figura 1.

Sin embargo, durante los últimos cinco años se han producido avances en este método, sobre todo con la finalidad de asegurar la convergencia, los cuales incluyen: histogramas planos, método bien temperado (para simulaciones a altas temperaturas), flux temperado, múltiples caminantes, entre otros. Una descripción más detallada de estos avances se encuentra en Singh, et al (2012).

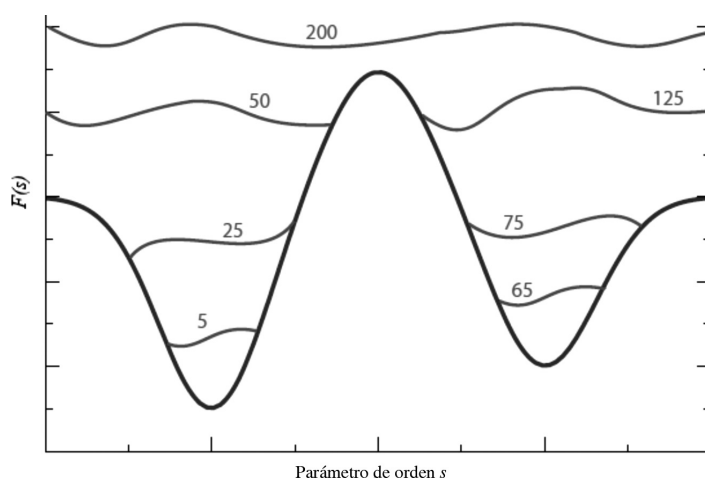


Figura 1. Uso de Gaussianos para superar mínimos locales de energía.

Nota: Tomado de “Density of States-Based Molecular Simulations” por S. Singh, M. Chopra, J. de Pablo, 2012. Annual Review of Chemical and Biomolecular Engineering, 3, 364-394.

Copyright 2012.

5. MODELIZACIÓN DE ADN

El ADN es una macromolécula que posee distintas formas; sin embargo, en condiciones fisiológicas normales, se sabe que la molécula presenta una estructura tipo B, que se encuentra formada por un grupo fosfato, un azúcar de cinco carbonos (desoxiribosa) y una base nitrogenada (adenina, citosina, guanina y timina), donde por la unión entre el grupo fosfato se une cada tripleta, y la doble cadena se forma por puentes de hidrógeno entre cada par de bases nitrogenadas, donde estas se unen por medio de patrones definidos entre adenina con timina, y citosina con guanina, de manera tal que forman una hélice de dos cadenas (Nelson, Cox, 2008; Bruice, 2008).

Las uniones de bases nitrogenadas determinan los parámetros físicos del ADN y puede ser tantas hasta alcanzar varios miles de uniones de bases nitrogenadas por cadena. Sin embargo, la simulación de cadenas tan largas resulta virtualmente imposible utilizando las herramientas computacionales existentes, por lo que, se realizan ciertas simplificaciones o se hacen estudios en oligómeros, que son pequeñas secuencias de ADN, que normalmente no superan las 30 bases nitrogenadas.

Asimismo, las uniones de bases nitrogenadas determina la formación de dos surcos alternados que se distinguen con facilidad en la hélice de ADN, un surco mayor de aproximadamente 22 Å y un surco menor de 11 Å (Wing, et al, 1980). Sin embargo, estos valores pueden variar dependiendo del orden de las uniones de bases nitrogenadas; por ejemplo, se sabe que secuencias de adenina y timina prolongadas provocan surcos menores muy angostos, los cuales se ha comprobado que son beneficiosos para el acople con ciertos tipos de proteínas (Rohs, et al, 2009). También de aquí nace la importancia de predecir el tamaño de estos surcos, puesto que también cambian durante el acople de ADN-proteína, y juegan un papel importante en el comportamiento físico y químico del ADN.

5.1. Escalas de simulación

Los modelos de simulación se basan en el tamaño relativo de los oligómeros de ADN, de manera que se tiene la siguiente escala de trabajo:

- **Atomístico:** se presenta en el tamaño de los nanómetros, donde se resuelven las ecuaciones para cada uno de los átomos que lo forman. Presenta la ventaja de obtener resultados muy exactos; sin embargo, su limitación principal es el tiempo de cálculo computacional. Esta restricción hace que en la década anterior sólo se haya podido utilizar en oligómeros de menos de 50 bases nitrogenadas por varios picosegundos, o en oligómeros sencillos de 10 bases nitrogenadas con simulaciones extendidas hasta el rango de los nanosegundos (Knotts, et al, 2007).
- **Mesoescala:** se da entre los dos nanómetros hasta los dos micrómetros. En esta escala, los cálculos de los modelos atomísticos son demasiados y se obtienen problemas con la resolución, en especial si se quiere mantener la continuidad que prevé la mecánica cuántica. Por ello, se requieren de modelos más simplificados, como los modelos “coarse-grained”.
- **Todo el genoma:** es mayor a los dos micrómetros, de manera que, como dice su nombre trata de simular una secuencia completa de ADN. Para ello son necesarios modelos “coarse-grained” o de bolas-resortes (Knotts, et al, 2007).

En la Figura 2 se muestra un ejemplo gráfico de la diferencia entre las dimensiones de cada escala de simulación.

5.2. Modelos de agua e iones

Uno podría realizar una simulación del ADN en condiciones de vacío perfecto, es decir, donde no se encuentre en contacto con otras moléculas; sin embargo, estas simulaciones no generan resultados significativos, ya que en las condiciones fisiológicas normales, esta molécula se encuentra influenciada por el medio solvente (que en este caso se trata de agua), el cual tiene sales disueltas.

El ADN por sí solo es una molécula cargada negativamente (Bruice, 2008), lo que implica que en forma natural se encuentra eléctricamente unida a cationes, lo cual, estabiliza a la molécula; por ello, es necesario realizar una parametrización adecuada de la fuerza eléctrica relativa de los iones que se agregan en la simulación.

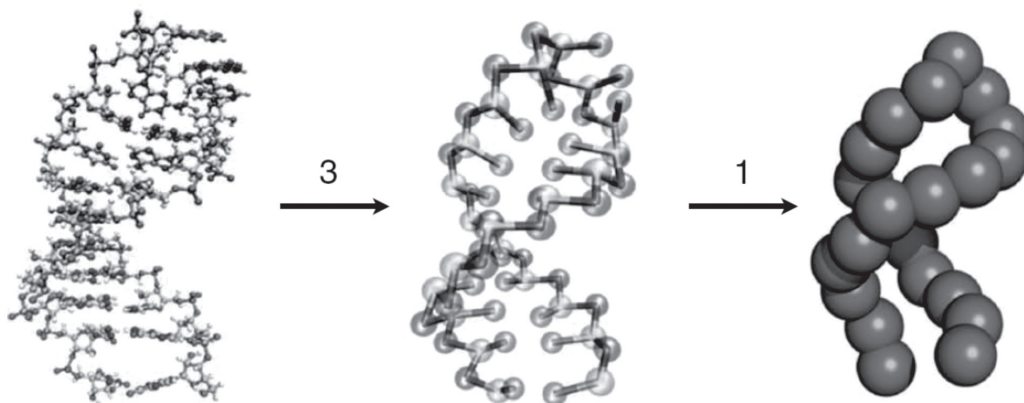


Figura 2. Escalas de simulación de ADN.

Nota: Tomado de “Capturing the essence of folding and functions of biomolecules using coarse-grained models” por C. Hyeon, D. Thirumalai, 2011, *Nature Communications*, 2:487. Reproducido con permiso de Macmillan Publishers Ltd: Nature Communications.

Además, se requiere realizar una modelización del agua, de forma que pueda ser incluida en las simulaciones con solvente implícito y explícito, ya que, las interacciones del agua y las sales con el ADN son dependientes de las ecuaciones de movimiento, lo que se convierte en un problema adicional en las simulaciones a escala atómica.

Normalmente, los campos de fuerza (que son el conjunto de parametrizaciones realizadas a las moléculas de agua, iones y ADN) se encuentran formulados a partir de observaciones experimentales y se encuentran disponibles en los paquetes computacionales.

Para el caso del agua, se han realizado principalmente dos tipos de modelos: de tres sitios (SPC y TIP3P) y de cuatro sitios (TIP4P). Algunos estudios realizados presentan conclusiones interesantes:

- Según Mark y Nilsson (2001), SPC/E presenta una mejor descripción de la dinámica y estructura del agua pura, mientras que el TIP3P presenta una dinámica más rápida que la obtenida en datos experimentales.
- Van der Spoel et al (1998), considera que SPC/E trabaja mejor con otras moléculas no polares, y que TIP4P presenta mejores resultados.

- Jorgensen y Jenson (1997) concluyen que TIP4P presenta mejor ajuste con los datos termodinámicos experimentales que TIP3P y SPC.
- Zielkiewicz (2005) recomienda SPC y TIP4P para el método de dinámica molecular en soluciones acuosas de biomoléculas.

Por lo tanto, de acuerdo con los estudios mencionados, para el caso de ADN las moléculas modelizadas TIP4P son las más recomendadas para simulaciones moleculares. Además, una revisión de todos los tipos de modelos utilizados para la modelización de agua con sus características se puede encontrar en Guillot (2002).

5.3. Tipos de cajas de simulación

En los paquetes computacionales más importantes, se incluye la posibilidad de realizar las simulaciones en distintos arreglos tridimensionales que contienen el ADN, las moléculas de agua e iones. Los dos tipos más importantes son:

- Cúbica y prismática: presentan la desventaja de que en cadenas largas de ADN, la rotación

natural que presenta la macromolécula, hace que ciertos átomos o elementos de la molécula empiecen a interactuar con sus contrapartes, debido a las condiciones periódicas de frontera que presentan las restricciones de la mayoría de los modelos utilizados.

- Octaédrica truncada: es una caja más cercana a una esfera, por lo que, permite una rotación del ADN sin que se presenten los problemas que se encuentran con las cajas cúbicas y prismáticas (Amber, 2000). El problema que presentan es que se requieren mayor cantidad de moléculas de solvente para realizar la descripción del sistema, aumentando significativamente el tiempo requerido para realizar la simulación, y que la caja puede “explotar” con mucha facilidad si la energía total del sistema es suficiente.

5.4. Paquetes computacionales

Existen varios paquetes computacionales de simulación molecular, que se encuentran basados principalmente en el método de dinámica molecular, los cuales se pueden utilizar para la simulación de ADN. Algunos de estos paquetes se describen a continuación.

AMBER: es un paquete con licencia, desarrollado por más de 40 investigadores de distintas universidades de Estados Unidos (Salomon, et al, 2012), el cual se utiliza para realizar simulaciones a escala atómica. Existen muchos campos de fuerza incorporados en AMBER, que actualmente corresponden a la versión 12; sin embargo, debido a que muchos de ellos son para uso general, los resultados para ácidos nucleicos no son muy exactos, por ello, se recomienda para simulaciones de ADN utilizar el campo de fuerza propuesto por Pérez et al (2007), el cual no aparece en todas las versiones de AMBER.

Este paquete posee numerosos programas para la instalación, desarrollo y análisis de Dinámica Molecular (este último se consigue por separado como AmberTools). Posee tres motores de simulación: sander (más tradicional, permite simulaciones utilizando la transformada rápida de Fourier), pmemd (una versión mejorada

de sander, que permite hacer simulaciones en paralelo) y pnmmd.cuda (una versión mejorada que permite el uso de GPUs por medio del lenguaje de programación CUDA desarrollado por NVIDIA). Una descripción más detallada de todas las herramientas se puede encontrar en Salomon, et al (2012) y Amber, (2013).

GROMACS: es un paquete de licencia libre para realizar simulaciones en escala atómica, desarrollado inicialmente por investigadores de Suecia y Alemania. Este paquete está orientado a la realización de simulaciones de Dinámica Molecular para moléculas bioquímicas como el caso de proteínas, lípidos y ácidos nucleicos (GROMACS, 2013).

De acuerdo a las ventajas que presenta frente a otros programas de simulación atómica como AMBER o CHARMM, sus desarrolladores aseguran que utilizan algoritmos que hacen que GROMACS sea más rápido en el cálculo de las interacciones no enlazantes (que usualmente son las que dominan los cálculos en este tipo de simulaciones), que es fácil de utilizar y que es mucho más rápido computacionalmente, debido a optimizaciones algorítmicas (GROMACS, 2013).

Entre las ventajas que presenta la cuarta versión (más reciente) es que utiliza la electrostática de Separación de Partículas de Ewald (PME en inglés), así como el uso de dos transformadas rápidas de Fourier (FTT) en tres dimensiones (Hess, et al, 2008).

3SPN.2: este es un modelo “coarse-grained”, desarrollado por investigadores de la Universidad de Wisconsin-Madison (Knotts, et al, 2007, Sambriski, et al, 2009, Freeman, et al, 2011, Hinckley, et al, 2013), que hace simulaciones en la mesoescala, de manera que permite simular cadenas de ADN más largas por mayor tiempo. El modelo se basa en la simplificación de cada parte del ADN (azúcar, grupo fosfato, base nitrogenada) puede ser descrito a través de un sitio, de ahí su nombre “Modelo tres sitios por nucleótido” y se encuentra implementado sobre la plataforma LAMPPS. En la Figura 3 se muestra cómo se realiza la modelización del ADN usando este programa.

El modelo describe los efectos salinos, el fenómeno de hibridación y derretimiento,

propiedades mecánicas y estructurales del ADN, así como atracciones débiles del tipo solvente inducidas. El modelo anterior (3SPN.1) es del tipo G σ , donde se integran todas las fuerzas que actúan sobre las moléculas en un único campo potencial; sin embargo, la nueva versión (3SPN.2) ha reemplazado este campo por potenciales que son dependientes de los ángulos (usando el potencial armónico de Morse) y torsiones (usando pozos de Gaussianos, uso del algoritmo de Metadinámica), y contempla las interacciones entre el apilamiento de bases nitrogenadas, el acoplamiento de dichas bases, así como la interacción debido al apilamiento entre hebras de ADN. Las ecuaciones por las cuales se calculan las interacciones del campo de fuerzas se encuentran bien documentadas en Hinckley, et al (2013). El mayor problema que presenta este modelo es que no utiliza iones en forma explícita.

Algunos otros modelos del tipo “coarse-grained” que se han desarrollado se pueden encontrar en Dans, et al (2010), Ouldrige, et al (2011) y Linak, et al (2011).

5.5. Análisis de datos

En esta sección se presenta una pequeña descripción de algunas herramientas computacionales que se han desarrollado para poder analizar los resultados obtenidos de simulaciones moleculares, principalmente de ADN.

3DNA: es un software libre diseñado para el análisis, reconstrucción y visualización de estructuras en tres dimensiones de ácidos nucleicos. El método se basa en tomar el archivo de coordenadas en formato PDB (Protein Data Bank) o de trayectorias de simulaciones moleculares, de manera que se identifican y categorizan todas las interacciones entre las bases nitrogenadas, con lo que se genera una estructura en términos de parámetros relacionados con bases nitrogenadas sucesivas.

Cada base nitrogenada es caracterizada a partir de un conjunto de seis parámetros de cuerpos rígidos, así como seis parámetros

complementarios de bases nitrogenadas. El problema que se ha encontrado con este programa es que no es capaz de describir correctamente la curvatura del ADN (Lavery, et al, 2009). Una descripción más detallada del funcionamiento del programa se presenta en Lu y Olson (2003).

Curves+: es un software libre que permite realizar un análisis conformacional de ácidos nucleicos. La idea de este programa es relacionar la noción de un eje de hélice del ácido nucleico para obtener estructuras curvilíneas. A partir de archivos PDB o archivos de trayectoria de simulaciones moleculares, puede calcular los parámetros de la hélice y del esqueleto de fosfato-azúcar, así como relacionar la posición de cada base nitrogenada con el eje de la hélice. En la Figura 4 se muestra una descripción de la hélice y de los parámetros calculados para un oligómero de 20 bases nitrogenadas. Una descripción más detallada del funcionamiento del programa se presenta en Lavery, et al (2009).

La hemicelulosa es un polímero heterogéneo que representa, en general, 15-35% de la biomasa de la planta. Consiste en cadenas cortas y ramificadas de azúcares. Contiene azúcares de cinco carbonos, generalmente xilosa y arabinosa y azúcares de seis carbonos (Girío et al, 2010).

La lignina tiene una estructura amorfa que consiste en unidades diferentes de fenilpropano. La función principal de la lignina es dar soporte estructural a la planta, impermeabilidad, resistencia contra ataques microbianos y al estrés oxidativo. Es insoluble en agua y ópticamente inactiva lo que la hace difícil de degradar (Hendricks & Zeeman, 2009).

VMD: por sus siglas en inglés, es Visual Molecular Dynamics. Es una herramienta computacional que permite visualizar las distintas estructuras que se le introducen, por lo que, permite no sólo visualizar ADN, sino también otras biomoléculas y estructuras de interés. Permite obtener imágenes de cada tiempo en el cual se han registrado las posiciones de las moléculas, así como seguir su evolución con el tiempo en forma de vídeo. El documento original al que se refiere esta herramienta corresponde a Humphrey, et al (1996).

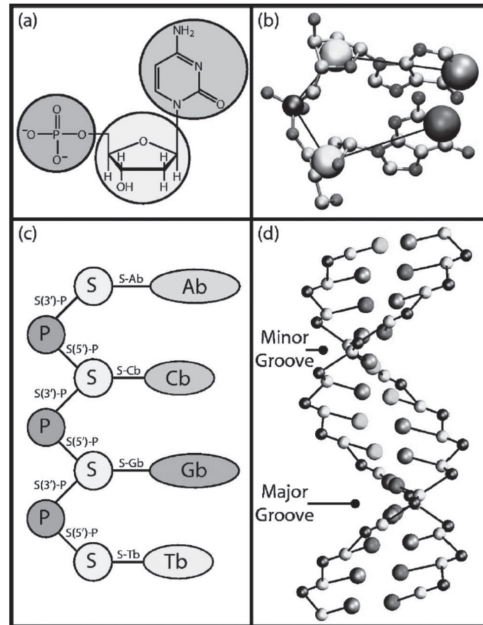


Figura 3. Modelización del ADN en modelo 3SPN.2. (a) Agrupamiento de los átomos para cada sitio para un nucleótido de citosina. (b) Mapeo de sitios atómicos a "coarse-grained" para un dinucleótido monofosfato 5'-GA-3'. (c) Topología de una hebra que ilustra la polaridad de las hélices. (d) Modelo para un oligonucleótido de 13 bases nitrogenadas.

Nota: Tomado de "A coarse grain model for DNA" por T. Knotts, N. Rathore, D. Schwartz, J. de Pablo, 2007, Journal of Chemical Physics, 126, 084901. Copyright 2007.

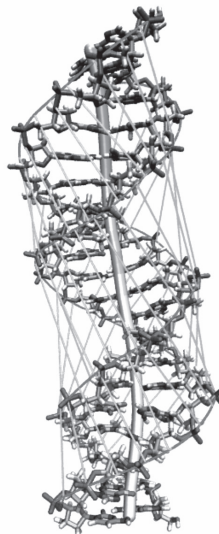


Figura 4. Oligómero de 20 bases nitrogenadas analizado con Curves+ y visualizado con VMD.

6. CONCLUSIONES

- Las simulaciones moleculares se pueden realizar a través de dos métodos: Montecarlo, el cual realiza recorridos aleatorios para la exploración de un ensamble termodinámico y Dinámica Molecular que se basa en la resolución de ecuaciones de movimiento, utilizando los métodos de la mecánica estadística, por medio de la evolución del sistema en el tiempo.
- El algoritmo de Metadinámica se ha desarrollado como una solución para el cálculo de la energía libre en sistemas difíciles de simular por los métodos tradicionales de simulación molecular.
- El ADN es una macromolécula que presenta una gran complejidad para su simulación a escala atómica, por lo que, utilizando programas como AMBER o GROMACS sólo se pueden realizar simulaciones de oligómeros de menos de 50 bases nitrogenadas y/o a tiempos pequeños (máximo unos cuantos nanosegundos).
- Si se utilizan modelos que parametrizan el ADN en la mesoescala, como el modelo 3SPN.2, se pueden realizar simulaciones de secuencias de nucleótidos más largas, así como tiempos de simulaciones más extensos, debido a que se simplifican los cálculos computacionales y el tiempo requerido para realizar estas simulaciones.
- La existencia de herramientas para el procesamiento de datos como VMD, X3DNA y Curves, ayudan a obtener los parámetros físicos de las distintas secuencias simuladas con gran facilidad y permiten a los científicos estudiar la evolución de estos parámetros en escalas de tiempo que no son accesibles por otros métodos.

REFERENCIAS BIBLIOGRÁFICAS

- Amber. (2013). *Amber Home Page*. Extraído de la página: <http://ambermd.org/>, el día 12 de abril de 2013.
- Amber. (2000). *Example 4: using a truncated octahedral box*. Extraído de la página: <http://ambermd.org/doc6/html/AMBER-sh-4.4.html>, el día 12 de abril de 2013.
- Atkins, P., de Paula, J. (2010). *Physical chemistry*. (9na ed). Nueva York: W.H. Freeman and Company.
- Bruice, P. (2008). *Química orgánica*. 5ta ed. México: Pearson Prentice Hall.
- Cotter, G. (2011). *Administración de la producción*. (2da ed). San Pedro, Costa Rica: Editorial de la Universidad de Costa Rica.
- Dans, P., Zeida, A., Machado, M., Pantano, S. (2010). A coarse grained model for Atomic-Detailed DNA simulations with explicit electrostatics. *J. Chem. Theory Comp*, 6(5), 1711-1725.
- Freeman, G., Hinckley, D., de Pablo, J. (2011). A coarse-grain three-site-per-nucleotide model for DNA with explicit ions. *J. Chem. Phys*, 135, 165104.
- Frenkel, D., Smit, B. (2002). *Understanding molecular simulation: From Algorithms to Applications*. San Diego: Academic Press.
- GROMACS. (2013). *About Gromacs*. Extraído de la página: http://www.gromacs.org/About_Gromacs, el día 12 de abril de 2013.
- Guillot, B. (2002). A reappraisal of what we have learnt during three decades of computer simulation on water. *Journal of Molecular Liquids*, 101(1-3), 219-260.
- Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E. (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3), 435-447.
- Hinckley, D., Freeman, G., Whitmer, J., de Pablo, J. (2013). 3SPN.2: An Experimentally-Informed Coarse-Grained Model of DNA. *J. Chem. Phys.*, 139(14), 144903.
- Humphrey, W., Dalke, A., Schulten, K. (1996). VMD: Visual molecular dynamics. *J. Mol. Graphics*, 14(1), 33-38.
- Hyeon, C., Thirumalai, D. (2011). Capturing the essence of folding and functions of biomolecules using coarse-grained models. *Nature Communications*, 2:487, DOI: 10.1038/ncomms1481.
- Institute of Molecular Engineering (IME). (2013). *Researchers review advances in the exploring of quantum computing*. Extraído

- de la página: http://ime.uchicago.edu/about/news/researchers_review_advances_in_the_exploration_of_quantum_computing/, el 29 de abril de 2014.
- Jorgensen, W., Jenson, C. (1997). Temperature dependence of TIP3P, SPC, and TIP4P Water from NPT Monte Carlo Simulations: Seeking Temperatures of Maximum Density. *Journal of Computational Chemistry*, 19(10), 1179-1186.
- Knotts, T., Rathore, N., Schwartz, D., de Pablo, J. (2007). A coarse grain model for DNA. *J. Chem. Phys.*, 126, 084901.
- Laio, A., Parinello, M. (2002). Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA*, 99, 12562-12566.
- Laio, A., Gervasio, F. (2008). Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics*, 71, 126601.
- Lavery, R., Moakher, M., Maddocks, J., Petkeviciute, D., Zakrzewska, K. (2009). Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Research*, 37(17), 5917-5929.
- Linak, M., Tourdot, R., Dorfman, K. (2011). Moving beyond Watson-Crick models of coarse grained DNA dynamics. *J. Chem. Phys.*, 135, 205102.
- Lu, X., Olson, W. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research*, 31(17), 5108-5121.
- Mark, P., Nilsson, L. (2001). Structure and Dynamics of the TIP3P, SPC and SPC/E Water Models of 298 K. *J. Phys. Chem.*, 105, 9954-9960.
- Nelson, D., Cox, M. (2008). *Lehninger Principles of Biochemistry*. (5ta ed). Nueva York: W.H. Freeman and Company.
- Ouldrige, T., Louis, A., Doye, J. (2011). Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.*, 134, 085101.
- Pérez, A., Marchán, I., Svozil, D., Spöner, J., Cheatham, T., Laughton, C., Orozco, M. (2007). Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophysical Journal*, 92(11), 3817-3829.
- Rohs, R., West, S., Sosinsky, A., Liu, P., Mann, R., Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature*, 461, 1248-1254.
- Salomon, R., Case, D., Walker, R. (2012). An overview of the Amber biomolecular simulation package. *WIREs Comp. Mol. Sci.*, doi: 10.1002/wcms.1121.
- Sambriski, E., Schwartz, D., de Pablo, J. (2009). A mesoscale model of DNA and its renaturation. *Biophysical Journal*, 96(5), 1675-1690.
- Singh, S., Chopra, M., de Pablo, J. (2012). Density of States-Based Molecular Simulations. *Annual Review of Chemical and Biomolecular Engineering*, 3, 364-394.
- van der Spoel, D., van Maaren, P., Berendsen, H. (1998). A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *J. Chem. Phys.*, 108, 10220.
- Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., Dickerson, R. (1980). Crystal structure analysis of a complete turn of B-DNA. *Nature.*, 287 (5784), 755-758.
- Zielkiewicz, J. (2005). Structural properties of water: Comparison of the SPC, SPCE, TIP4P and TIP5P models of water. *J. Chem. Phys.*, 123, 104501.

SOBRE EL AUTOR

Johnny Daniel Alfaro Pérez, The University of Chicago, Institute for Molecular Engineering. Estudiante de doctorado en Ingeniería Molecular, Asistente Graduado de Investigación. Correo electrónico: alfaro@uchicago.edu

