

Modelado de grano grueso de la proteína ubiquitina

C. A. Estrada

Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León

Resumen

La ubiquitina es una proteína que regula el reciclaje de otras proteínas dentro de la célula y que tiene relevancia en el ámbito de la nanotecnología, principalmente en el estudio de las interacciones entre proteínas y nanopartículas para estimar sus posibles efectos tóxicos. El conocer la estructura de las proteínas es vital para comprender su función biológica, y la correcta implementación de herramientas bioinformáticas con técnicas de modelado molecular, tales como el modelado de grano grueso (“coarse-grained modeling”, en inglés), permiten la predicción teórica de las estructuras tridimensionales de las proteínas. En el presente trabajo se realizó la simulación con modelado de grano grueso empleando el modelo de campo de fuerza Martini en la dinámica molecular de la ubiquitina humana, para su posterior análisis estructural y de calidad. La ubiquitina modelada por la simulación no alcanzó una calidad satisfactoria, con un RMSD de 0.7 nm, pero obtuvo un patrón de fluctuación RMSF que corresponde con el de la ubiquitina, por lo que las características estructurales del esqueleto de la ubiquitina se conservaron.

Palabras clave: Proteína, Simulación, Ubiquitina, GROMACS, Martini

1. Introducción

Las proteínas son algunas de las biomoléculas más importantes dentro de la compleja maquinaria que compone a los seres vivos. Un ejemplo de ello es la ubiquitina, una proteína pequeña y de forma globular, compuesta de 76 aminoácidos, cuya función es regular la degradación de otras proteínas de la célula para su reciclaje mediante la unión covalente con el aminoácido lisina de las proteínas blanco, para su posterior degradación intracelular en el proteosoma [1]. La ubiquitina tiene relevancia en el campo de la nanotecnología aplicada a la medicina, ya que, en el caso de las nanopartículas, al interactuar con los sistemas biológicos las nanopartículas pueden alterar la conformación de las proteínas, conllevando efectos tóxicos [2]; con el estudio de la interacción de la ubiquitina con las nanopartículas se pretende comprender los mecanismos con los que las proteínas interactúan

con las nanopartículas aplicadas para la liberación de fármacos y conocer mejor los posibles efectos tóxicos de éstas [3, 4].

El conocer la estructura de las proteínas es de gran importancia para comprender su función biológica, ya que cada proteína posee una estructura tridimensional específica porque su plegamiento se relaciona directamente con la función que desempeña [5]. Para ello, la adecuada adaptación y combinación de herramientas bioinformáticas y técnicas de modelado molecular permiten la predicción teórica de las estructuras tridimensionales de las proteínas [6].

La mayoría de las proteínas son demasiado grandes para ser estudiadas con herramientas clásicas de simulación, como el modelado molecular a nivel atómico, por lo que el modelado de grano grueso (“coarse-grained modeling”, en inglés) se convierte en una buena alternativa para ello. En el modelo de grano grueso se reducen los grados de libertad al ignorar algunos

de los átomos o mediante la formación de grupos de átomos que son tratados como una sola unidad, en lo que se denomina como “pseudoátomos” o granos gruesos [7], pero a la vez permite mantener las principales características fisicoquímicas. Los modelos de grano grueso han sido utilizados exitosamente para el estudio de los mecanismos de plegamiento y para la predicción estructural de proteínas [8], obteniendo resultados favorables.

En la actualidad, se disponen de diversas aproximaciones de modelado de grano grueso, las cuales varían desde aquellas con una aproximación cuantitativa, como los modelos libres de solventes, hasta aquellos modelos más realistas que incluyen especificidad química [9]. Otra aproximación es la de el campo de fuerza Martini (“Martini force field”, en inglés), el cual es un campo de fuerza empleado en modelado de grano grueso adaptado para simulaciones de dinámica molecular de sistemas biológicos [10]. El campo de fuerza Martini combina dos estrategias: primero, las interacciones no vinculadas se basan en la reproducción de las energías libres de partición experimentales entre las fases polares y apolares, y, segundo, las interacciones vinculadas se derivan de simulaciones de referencia de todos los átomos. Este modelo emplea un mapeo de cuatro a uno, lo que significa que se promedian cuatro átomos pesados con sus átomos de hidrógeno asociados y se representan como un solo centro de interacción o pseudoátomo; además, solo se definen cuatro tipos de sitios de interacción: polares, apolares, no polares y con carga, y, a su vez, a cada partícula se le asigna cierto número de subtipos de estas cuatro interacciones, lo que permite representar de manera más precisa la naturaleza química de la molécula [10]. El modelo Martini ha sido adaptado para su empleo en proteínas [11], en donde cada residuo (es decir, cada aminoácido que forma parte de la cadena polipeptídica) posee un pseudoátomo en el esqueleto de la proteína y desde cero a cuatro pseudoátomos de las cadenas laterales, dependiendo del tipo de residuo.

En el presente trabajo se realizó una simulación con modelado de grano grueso de la ubiquitina humana empleando el modelo de campo de fuerza Martini, para su posterior análisis estructural y de calidad, los cuales se describen en mayor detalle en la sección

de resultados y discusión.

2. Experimentación

2.1. Herramientas

La simulación se realizó en una laptop HP Pavilion con un procesador AMD Ryzen 3. Para la simulación de la proteína se utilizó el paquete de software de libre uso GROMACS 2020.1 [12], así como el software Python 2.7 [13]. La visualización de las estructuras de las proteínas se realizó con el software de libre uso Jmol [14], y el análisis de resultados con el paquete estadístico R versión 4.0.2 [15].

2.2. Simulación

La simulación se llevó a cabo utilizando el archivo proveniente de Protein Data Bank para la ubiquitina humana con el código de indentificación 1UBQ [16], el cual contiene la estructura atomística de la proteína requerida para la simulación, y cuyas características generales se presentan en el cuadro 1. Utilizando el código de uso libre para el modelado de grano grueso con el campo de fuerza Martini [17], se convirtió la estructura atomística de la ubiquitina en una estructura de grano grueso en Python 2.7, especificando la generación de la lista de los átomos involucrados, así como la determinación de la clasificación de la estructura secundaria del esqueleto de la proteína. Posteriormente, utilizando la estructura de grano grueso de la ubiquitina en GROMACS se especificó las condiciones límite periódicas (características de la caja en que se produce la simulación) utilizando una caja de forma dodecaédrica (óptima para las proteínas de conformación globular) con una distancia mínima de 1 nm entre la proteína y cualquier borde de la caja, y se procedió a realizar una minimización en vacío de 100 pasos. Después, se realizó la solvatación del sistema utilizando una caja de agua cuyo sistema está equilibrado (temperatura de 300 K, 1 bar de presión y 2,000 pseudoátomos de agua) y se procedió a una minimización de energía de 100 pasos, con la subsecuente equilibración del sistema con una duración de 10,000 pasos. A partir de ello se realizó la simulación de dinámica molecular del sistema completo con una

duración de 3,000,000 pasos. Una vez finalizada la simulación, se convirtió a las trayectorias generadas de tal manera que la traslación y rotación de la ubiquitina sean removidas, y con ello realizar los análisis estándar que se describen en la siguiente sección.

Cuadro 1: Contenido de la macromolécula 1UBQ.

Peso total de la estructura	8.58 kDa
Cantidad de átomos	660
Cantidad de residuos	76
Cadenas de proteína únicas	1

2.3. Análisis de resultados

Utilizando GROMACS se calcula la desviación cuadrática media de las posiciones atómicas (RMSD, por sus siglas en inglés) y la fluctuación cuadrática media (RMSF, por sus siglas en inglés), mientras que los datos generados son graficados utilizando el paquete estadístico R. La visualización de la nueva estructura de la ubiquitina generada, así como la estructura original, se realizó utilizando Jmol.

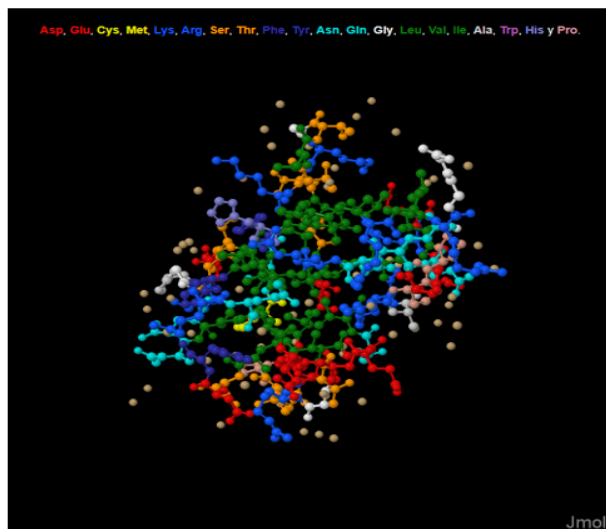
3. Resultados y discusión

En la figura 1 se presentan las estructuras tridimensionales en un modelo de bolas y varillas generados en Jmol de la ubiquitina original proveniente de Protein Data Bank, la misma estructura después de haber sido convertida a granos gruesos (pseudoátomos) y la proteína generada al finalizar la simulación. El código de colores indica los aminoácidos de acuerdo con sus propiedades, por lo que se puede ubicar a simple vista la localización aproximada de los residuos que componen tanto a la proteína original como a la proteína generada por la simulación; se observa que la ubiquitina original posee una gran cantidad de esferas, ya que se modeló su estructura atomística, mientras que la estructura convertida a granos gruesos posee menor cantidad, lo que indica que la conversión fue exitosa. La proteína generada por la simulación posee los granos gruesos de la conversión con sus respectivas interacciones, representando de manera simplista la estructura tridimensional original de la ubiquitina,

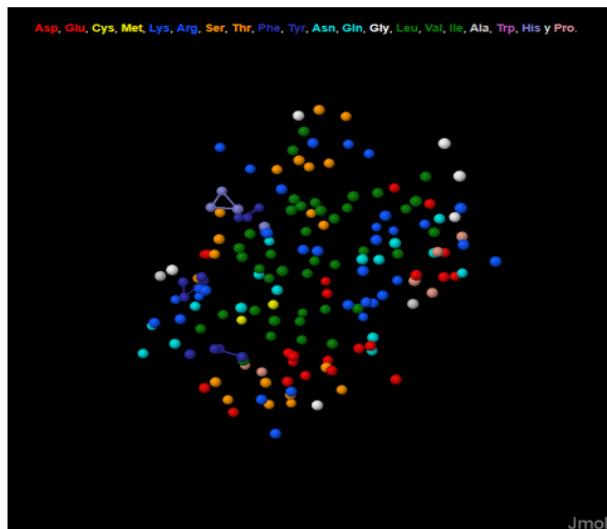
en donde se observa que la posición aproximada de las regiones de cada aminoácido se conservó después de la simulación.

En la figura 2 se presenta la desviación promedio de cada residuo que compone a la proteína simulada a través del tiempo (RMSF); esta medida es un indicador de las porciones de la estructura que fluctúan más a partir de su estructura promedio [18], lo que representa a las regiones flexibles de la proteína, que en este caso serían los bucles. Los bucles se encuentran en particular en las superficies de proteínas globulares [19], como la ubiquitina, lo que explica que se haya presentado un RMSF alto en muchos de los residuos; además, este patrón coincide con el reportado previamente para la ubiquitina [20], ya que los patrones de fluctuación del esqueleto de las proteínas tienden a conservarse [21], mientras que las leves variaciones se deben a la dinámica molecular en particular de la simulación.

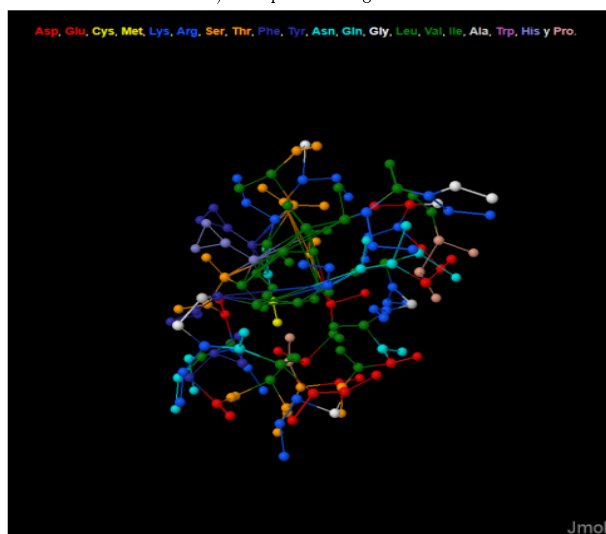
Finalmente, en la figura 3 se presenta la desviación cuadrática media de las posiciones atómicas (RMSD) a partir de la estructura cristalográfica a través del tiempo, la cual es una medida acerca de la exactitud del modelo predicho. Para la ubiquitina simulada, se observa que el RMSD aumenta su valor a través del tiempo hacia una especie de meseta, lo que indica que la estructura tiende a buscar estabilidad a lo largo del tiempo con respecto a la estructura de referencia, pero al tratarse de valores que llegan hasta aproximadamente 0.7 nm, la conformación de la proteína no es estable. El valor de RMSD al finalizar la simulación es de aproximadamente 0.7 nm, el cual se encuentra muy por encima del valor RMSD reportado como control de calidad para simulaciones de proteínas, que es de menos de 0.2 nm [22]; esto pudo deberse a que la simulación, por simplicidad, se realizó utilizando una caja de agua pura, evitando la adición de iones, y es un hecho que las proteínas requieren de cierta cantidad de iones para mantener su estabilidad [23]. En otros reportes con los mismos parámetros utilizados en este trabajo, el RMSD de la ubiquitina mantuvo valores por debajo de los 0.2 nm, con la diferencia que se tomó en cuenta la adición de algunos iones [20], lo que corrobora que fue la omisión de este parámetro lo que causó los valores tan altos obtenidos.



a) Ubiquitina original.



b) Ubiquitina convertida a granos gruesos.



c) Ubiquitina simulada.

Figura 1: Estructuras tridimensionales de la ubiquitina en modelo de bolas y varillas generadas con Jmol.

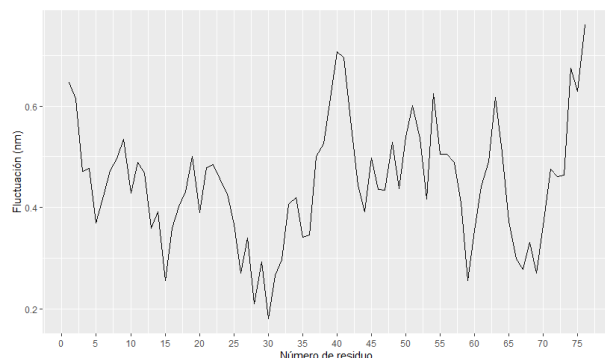


Figura 2: Fluctuación cuadrática media (RMSF) en función de cada número de residuo de la secuencia a través del tiempo de la ubiquitina simulada.

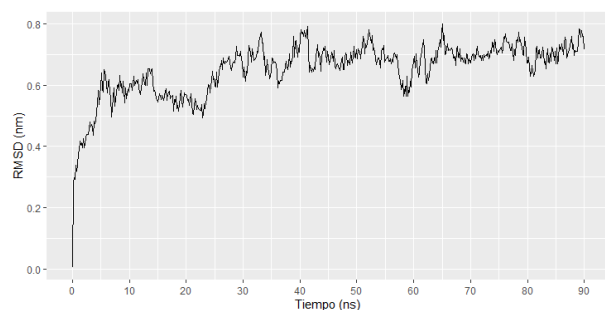


Figura 3: Desviación cuadrática media de las posiciones atómicas (RMSD) a partir de la estructura cristalográfica en función del tiempo.

4. Conclusiones

La ubiquitina modelada por grano grueso obtuvo una estructura tridimensional similar a la ubiquitina original, lo cual se ve corroborado por el RMSF obtenido, que coincide con el patrón de fluctuación del esqueleto de la ubiquitina reportado en la literatura. No obstante, el modelo producido obtuvo un RMSD de aproximadamente 0.7 nm al finalizar la simulación, por lo que no alcanzó una calidad satisfactoria, ya que esto indica que su conformación es inestable.

Se requeriría realizar la simulación nuevamente en el futuro, incluyendo iones en la fase de solvatación del sistema para incrementar la estabilidad de la ubiquitina a lo largo del tiempo. De igual forma, se podrían incorporar nuevos parámetros que permitan el estudio de su interacción con otras proteínas, o incluso nanopartículas.

5. Agradecimientos

Se agradece encarecidamente al Lic. Juan Garza por su asesoría para la instalación de GROMACS, al igual que a la Dra. Elisa Schaeffer por sus correcciones en la estructura original del trabajo.

Referencias

- [1] P. J. Bertics, C. J. Koziol-White, M. L. Gavala, G. J. Wiepz, 11 - signal transduction, in: N. F. Adkinson, B. S. Bochner, A. W. Burks, W. W. Busse, S. T. Holgate, R. F. Lemanske, R. E. O’Hehir (Eds.), Middleton’s Allergy (Eighth Edition), eighth edition Edition, W.B. Saunders, London, 2014, pp. 184–202. doi:<https://doi.org/10.1016/B978-0-323-08593-9.00012-7>. URL <http://www.sciencedirect.com/science/article/pii/B9780323085939000127>
- [2] F. Ding, S. Radic, R. Chen, P. Chen, N. K. Geitner, J. M. Brown, P. C. Ke, Direct observation of a single nanoparticle–ubiquitin corona formation, *Nanoscale* 5 (19) (2013) 9162–9169.

- [3] G. Brancolini, D. B. Kokh, L. Calzolari, R. C. Wade, S. Corni, Docking of ubiquitin to gold nanoparticles, *ACS nano* 6 (11) (2012) 9863–9878.
- [4] L. Calzolari, F. Franchini, D. Gilliland, F. Rossi, Protein- nanoparticle interaction: identification of the ubiquitin- gold nanoparticle interaction site, *Nano letters* 10 (8) (2010) 3101–3105.
- [5] D. De Lucrezia, D. Slanzi, I. Poli, F. Polticelli, G. Minervini, Do natural proteins differ from random sequences polypeptides? natural vs. random proteins classification using an evolutionary neural network, *PLoS One* 7 (5) (2012) e36634.
- [6] J. Haas, S. Roth, K. Arnold, F. Kiefer, T. Schmidt, L. Bordoli, T. Schwede, The protein model portal—a comprehensive resource for protein structure and model information, *Database* 2013.
- [7] M. Blaszczyk, D. Gront, S. Kmiecik, K. Ziolkowska, M. Panek, A. Kolinski, Coarse-grained protein models in structure prediction, in: *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes*, Springer, 2014, pp. 25–53.
- [8] S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, A. Kolinski, Coarse-grained protein models and their applications, *Chemical reviews* 116 (14) (2016) 7898–7936.
- [9] G. A. Voth, *Coarse-graining of condensed phase and biomolecular systems*, CRC press, 2008.
- [10] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, The martini force field: coarse grained model for biomolecular simulations, *The journal of physical chemistry B* 111 (27) (2007) 7812–7824.
- [11] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, S.-J. Marrink, The martini coarse-grained force field: extension to proteins, *Journal of chemical theory and computation* 4 (5) (2008) 819–834.
- [12] Lindahl, Abraham, Hess, van der Spoel, *Gromacs 2020.1 manual* (Mar. 2020). doi:10.5281/zenodo.3685920. URL <https://doi.org/10.5281/zenodo.3685920>
- [13] P. Software Foundation, *Python Language Reference*, version 2.7 (2020). URL <https://www.python.org/>
- [14] J. development team, *Jmol: an open-source java viewer for chemical structures in 3d* (2020). URL <http://www.jmol.org/>
- [15] T. R. Foundation, *The R Project for Statistical Computing* (2020). URL <https://www.r-project.org/>
- [16] S. Vijay-Kumar, C. E. Bugg, W. J. Cook, PDB ID: 1UBQ, *Journal of molecular biology* 194 (3) (1987) 531–544. URL <https://www.rcsb.org/structure/1UBQ>
- [17] D. H. de Jong, G. Singh, W. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schafer, X. Periole, D. P. Tieleman, S. J. Marrink, Improved parameters for the martini coarse-grained protein force field, *Journal of chemical theory and computation* 9 (1) (2013) 687–697. URL <http://www.cgmartini.nl/images/stories/tutorial/2017/protein-tutorial-v2016.3/soluble-protein/ubiquitin/martini/martinize.py>
- [18] Y. Kiyoshi, *Root-mean-square fluctuation (rmsf)* (2016). URL <https://www.r-ccs.riken.jp/labs/cbrt/tutorial/analysis-tutorials/tutorial-4-3/>
- [19] J. F. Leszczynski, G. D. Rose, Loops in globular proteins: a novel category of secondary structure, *Science* 234 (4778) (1986) 849–855.
- [20] D. B. Kony, P. H. Hünenberger, W. F. Van Gunsteren, Molecular dynamics simulations of the native and partially folded states of ubiquitin: Influence of methanol cosolvent, ph, and tempe-

- rature on the protein structure and dynamics, Protein science 16 (6) (2007) 1101–1118.
- [21] E. Fuglebakk, J. Echave, N. Reuter, Measuring and comparing structural fluctuation patterns in large protein datasets, Bioinformatics 28 (19) (2012) 2431–2440.
- [22] D. van der Spoel, Structure and dynamics of peptides: theoretical aspects of protein folding.
- [23] M. T. Record Jr, W. Zhang, C. F. Anderson, Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, hofmeister effects, and osmotic effects of salts., Advances in protein chemistry 51 (1998) 281.