Estudio de la evolución dinámica de las proteínas intrínsecamente desordenadas usando cuaterniones duales y aprendizaje automático profundo.

Introducción

Para poder comprender con detalle la estructura, la función y los mecanismos de reacción de sistemas biológicos y químicos es necesario emplear métodos teóricos de modelado molecular y de simulación que complementen los resultados obtenidos mediante técnicas experimentales. A pesar de que la cristalografía de rayos X permite estudiar las propiedades físicas y estructurales de las moléculas, la mera visualización de una única estructura molecular no es capaz de resolver problemas más complejos como, por ejemplo, el movimiento de una proteína o su unión a un ligando. En estos casos es necesario acudir a la simulación de tales procesos. El principal objetivo del modelado molecular aplicado en distintas áreas como la bioquímica o la farmacología, consiste en describir las interacciones biomoleculares en base a las leyes generales de la química y de la física.

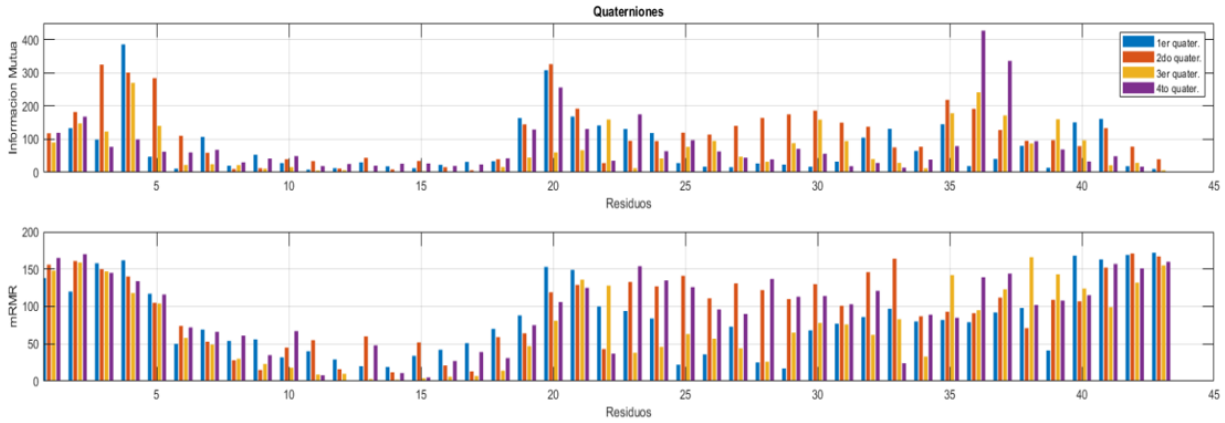
Hoy en día, el modelado molecular está íntimamente ligado al uso de computadoras y gráficos interactivos. El espectacular desarrollo de los ordenadores y de las técnicas computacionales ha permitido el uso de modelos teóricos para el estudio de todos aquellos problemas relacionados con la geometría y la energía de las moléculas. De hecho, se podría llegar a modelar con exactitud cualquier sistema biológico con suficiente poder de cálculo y un alto nivel de teoría. Sin embargo, en la práctica, esto no es siempre posible debido a limitaciones tanto de tiempo como de recursos materiales. Por este motivo, es necesario introducir aproximaciones para mantener la viabilidad del experimento aplicando distintos niveles de teoría en función del sistema que va a ser analizado. Como norma general, a medida que aumenta el nivel teórico aumenta la calidad de los resultados obtenidos, pero también se incrementa notablemente el coste computacional.

Metodología

Analizar las diferentes conformaciones de una proteína obtenida mediante la técnica de resonancia magnética nuclear y depositada en la base de datos de proteínas (PDB). Para estas mismas proteínas se realizará un estudio mediante dinámica molecular clásica tratando de obtener un amplio muestreo de las diferentes conformaciones que adoptan las secciones desordenadas de los PID. Para ello, se utilizará el código informático gromacs (software libre). Para asegurar un amplio muestreo de las diferentes conformaciones [2]. La evolución dinámica de la estructura de la proteína se realizará describiendo las conformaciones de las proteínas mediante cuaterniones [8-11] y midiendo la distancia entre ellos. De esta forma obtendremos series de tiempo de la evolución de los parámetros estructurales que serán investigados con técnicas de aprendizaje automático como las Técnicas No Lineales para determinar la correlación del movimiento entre las diferentes partes de la proteína, estos son el Error de Predicción y la Dimensión de Incrustación [12], Información Mutua, Mínima Redundancia y Máxima Relevancia. La proteína que se pretenden estudiar es la 6ES5.

Resultados

En la figura muestra el análisis con información mutua y MRMR para el campo de fuerza OPLS, donde los residuos 4, 20 y 36 son los que comparten mayor información.



Para comprender con detalle la estructura, función y mecanismos de reacción de sistemas biológicos es necesario emplear métodos teóricos de modelado molecular.

Referencias

[1] F. Bemporad, J. Gsponer, H. I. Hopearuoho, G. Plakoutsi, G. Stati, M. Stefano, N. Taddei, M. Vendruscolo, F. Chiti Embo J.27, 1525-1535 (2008).

[2] V. N. Uversky Front. Phys. 7, Art. No. 10 (2019).

[3] A. K. Dunker, J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K.W. Hipps, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner, Z.Obradovic J. Mol. Graph Model 19, 26-59 (2001).

[4] A. L. Fink Curr. Opin. Struc. Biol. 15, 35-41 (2005).

[5] B. Xue., A. K. Dunker, V. N. Uversky J. Biomol. Struc. Dyn. 30, 137-149 (2012).

[6] J. J. Ward, J. S. Soghi, L. J. McGuffin, B. F. Buxton, D. T. Jones J. Mol. Biol. 337, 635-645 (2004).

[7] V. N. Uversky, A. K. Dunker Biochim. Biophys. Acta 1804, 1231-1264 (2010).

[8] P. Radivojac, L. M. Iakoucheva, C. J. Oldfield, Z. Obradovic, V. N. Uversky, A. K. Dunker Biophys. J. 92, 1439-1456(2007).

[9] P. Romero, Z. Obradovic, X. Li, E. C. Garner, C. J. Brown, A. K. Dunker Proteins 42, 38-48 (2001).

[10] D. Shukla, C. X. Hernández, J. K. Weber, V. S. Pande Acc. Chem. Res. 48, 414-422 (2015).

[11] K. Henzler-Wildman, D. Kern Nature 450, 964-972 (2007).

[12] J. D. Forman-Kay, T. Mittag Structure 21, 1492-1499 (2013).

[13] J.-F. Yu, Z. Cao, Y. Yang, C.-L. Wang, Z.-D. Su, Y.-W. Zhao, J.-H. Wang, Y. Zhou Cell. Mol. Life. Sci. 73, 2949-2957(2016).

[14] Y.-S. Lin, G. R. Bowman, K. A. Beauchamp, V. S. Pande Biophys. J. 102, 315-324 (2012).

[15] Q. Qiao, G. R. Bowman, X. Huang J. Am. Chem. Soc. 135, 16092-16101 (2013).

[16] Lindahl, E., Hess, B., \& Van Der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. Molecular modeling annual, 7(8), 306-317.

[17] G. F. Torres del Castillo. (January 1999) La representación de rotaciones mediante

cuaterniones. ResearchGate. 43-50

[18] Fermani S, Trivelli X, Sparla F, et al. Conformational selection and folding-uponbinding of intrinsically disordered protein CP12 regulate photosynthetic enzymes assembly. The Journal of Biological Chemistry. 2012 Jun;287(25):21372-21383. DOI: 10.1074/jbc.m112.350355.