

TÓPICO: MODELADO MOLECULAR DE PROTEÍNAS

DOCENTES A CARGO: Dr. Ramón Garduño Juárez (ICF) y Dr. Sergio Mares Samano (Catedrático CONACyT)

INTRODUCCIÓN

Los métodos de modelado molecular se han convertido en un complemento indispensable de las técnicas experimentales. En la actualidad son usados rutinariamente para investigar la estructura, dinámica y termodinámica de sistemas biológicos, inorgánicos, y poliméricos. Algunos tipos de actividad biológica que han sido investigados usando el modelado molecular incluyen el plegamiento proteico, la catálisis de enzimas, la estabilidad de proteínas, y el reconocimiento molecular. Este curso va dirigido a estudiantes de licenciatura o de posgrado no especializados en la biología estructural, ni en química computacional, pero que estén interesados en conocer técnicas modernas de modelado molecular y aplicarlas al descubrimiento de nuevos fármacos. El curso proporcionará a los estudiantes las herramientas teóricas y prácticas para que, a partir de la secuencia de aminoácidos de una proteína, sean capaces de generar modelos estructurales de alta calidad y utilizarlos como dianas en campañas de diseño de nuevos fármacos. Específicamente el alumno adquirirá habilidades para i) identificar moldes estructurales utilizando alineamientos de secuencias de aminoácidos, ii) construir modelos tridimensionales de proteínas por medio de modelado por homología, iii) visualizar y examinar modelos tridimensionales de proteínas, iv) minimizar energéticamente los modelos construidos por medio de dinámica molecular, vi) implementar la técnica de acoplamiento molecular para identificar moléculas con potencial actividad moduladora y vii) utilizar métodos de QSAR para modificar o mejorar moléculas con actividad biológica ya existentes. Dado el carácter multidisciplinario de los temas incluidos, el estudiante adquirirá las bases para establecer colaboraciones fructíferas con biólogos, químicos y biofísicos computacionales.

PRERREQUISITOS

- Conceptos básicos de bioquímica, de fisicoquímica y de cálculo.
- Conocimiento básico de programación en Shell y manejo del sistema operativo Linux
- Computadora de escritorio o personal con sistema operativo Linux (se sugiere utilizar Ubuntu) con procesador de al menos dos núcleos, 3 GB en RAM y 100 GB de espacio libre en disco. Disponibilidad de cámara y micrófono; el curso se impartirá a través de videoconferencia.
- El alumno deberá instalar previamente los programas que se utilizarán durante el curso. Se instruirá con oportunidad sobre los sitios de descarga de estos programas y el procedimiento de instalación.
- Contar con cuenta de correo-e en Gmail.

METODOLOGÍA

Durante este curso nos reuniremos una vez por semana durante tres horas. El desarrollo del curso girará en torno a la resolución de un caso práctico. Se presentará a los estudiantes el caso de una proteína cuya actividad es esencial para la aparición de una enfermedad, cuya estructura tridimensional no ha sido dilucidada y sólo está disponible la secuencia de aminoácidos. La meta es entonces que con las herramientas proporcionadas, al final del curso el alumno proponga un conjunto de moléculas que potencialmente puedan inhibir la actividad de esa proteína. El temario ha sido diseñado para que los alumnos vayan adquiriendo secuencialmente las habilidades que les permitirán alcanzar el objetivo. Inicialmente se instruirá sobre los aspectos esenciales de la programación en Shell, fundamentales para utilizar los programas de modelado molecular a lo largo del curso. El modelado por homología les permitirá generar y refinar estructuralmente modelos tridimensionales de la proteína en estudio. En seguida, el acoplamiento molecular les permitirá analizar bases de datos que contienen grandes cantidades de moléculas pequeñas para identificar aquellas que potencialmente podrían inhibir la actividad de la proteína. Los de dinámica molecular ayudará a refinar la selección de las mejores moléculas. Se espera que los estudiantes repasen de antemano, con el apoyo de los libros en formato E-Book, los temas que se presentarán en cada clase. Asimismo, será obligación del estudiante el trabajar de manera independiente los tutoriales de los varios programas que usaremos

LIBRO DE TEXTO

No hay un solo libro de texto que contenga todo el material que se pretende cubrir durante el curso. Dado que muchos libros ya no han tenido ediciones más recientes, algunos libros están disponibles como E-books en las páginas de las editoriales y serán usados como material de consulta.

EVALUACIÓN

- Exámenes - parciales (20%)
- Participación y desarrollo de ejercicios en clase (25%).
- Tareas (25%).
- Presentación de los resultados del problema (30%)

PROTOCOLO

1) Tareas:

De 4 a 8 Tareas se dejarán a lo largo del curso y fundamentalmente versarán sobre la resolución de tutoriales.

2) Presentación de los resultados del problema:

Los resultados deben estructurarse en el formato empleado en la revista científica "Bioinformatics" (https://academic.oup.com/bioinformatics/pages/submission_online). Se espera que el estudiante se reúna con el instructor al menos tres veces durante el semestre para revisar la agenda de la redacción, para revisar el avance y para discutir la presentación y

la versión final. Estas reuniones tendrán que ser agendadas con el instructor ya que se harán por medio de videoconferencia.

BIBLIOGRAFÍA

- Cobbaut, P. (2015). Linux Fundamentals. GNU Free Documentation License. <http://linux-training.be/linuxfun.pdf>
- Fiser, A. (2010). Template-Based Protein Structure Modeling (pp. 73–94). https://doi.org/10.1007/978-1-60761-842-3_6
- Eswar, N., Webb, B., Marti-renom, M. A., Eramian, D., Shen, M., Pieper, U., & Sali, A. (2007). Comparative Protein Structure Modeling Using MODELLER, (November), 1–31. <https://doi.org/10.1002/0471140864.ps0209s50>
- Modi, V., Xu, Q., Adhikari, S., & Dunbrack, R. L. (2016). Assessment of template-based modeling of protein structure in CASP11. *Proteins*. <https://doi.org/10.1002/prot.25049>
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, 11(5), 905–19. <https://doi.org/10.1038/nprot.2016.051>
- Chen, Y. C. (2015). Beware of docking! *Trends in Pharmacological Sciences*, 36(2), 78–95. <https://doi.org/10.1016/j.tips.2014.12.001>
- Jacob, R. B., Andersen, T., & McDougal, O. M. (2012). Accessible high-throughput virtual screening molecular docking software for students and educators. *PLoS Computational Biology*, 8(5), e1002499. <https://doi.org/10.1371/journal.pcbi.1002499>

TEMARIO

1. Introducción a la Programación en Shell (7 horas) – Dr. Mares

- a. Introducción al sistema operativo Linux
- b. Comandos básicos en Linux
- c. Manejo de un editor de texto (VIM)
- d. *Scripts* utilizando awk, grep y sed
- e. Conectarse a servidores remotos usando el protocolo SSH
- f. Transferencia de archivos entre servidores

2. Modelado Comparativo de Proteínas (homología) (11 horas) – Dr. Mares

- a. Introducción al modelado comparativo de proteínas
- b. Bases de datos de secuencias y estructuras de proteínas
 - i. Uniprot
 - ii. *Protein Data Bank*
 - iii. *The National Center for Biotechnology Information*
- c. Visualización de estructuras de proteínas
 - i. Pymol
 - ii. VMD
 - iii. Chimera
- d. Alineamientos de secuencias de proteínas para la identificación de moldes estructurales
 - i. Algoritmos globales
 - Needleman-Wunsch
 - ii. Algoritmos locales
 - Smith-Waterman
 - iii. Herramientas
 - BLAST (Basic Local Alignment Search)
 - T-Coffee
- e. Construcción 3D del modelo
 - i. Modeller
 - ii. Swiss-Model
- f. Validación estructural de modelos 3D
 - i. Estereoquímica
 - ii. Energética
 - iii. Herramientas
 - Molprobit
 - Whatcheck

3. Acoplamiento Molecular (proteína-ligando) (11 horas) – Dr. Mares

- a. Introducción al acoplamiento molecular
- b. Obtención y preparación del receptor (proteína)

4. Métodos de Modelado *ab initio* (12 horas) – Dr. Garduño

- a. La aproximación del Campo de Fuerzas

