**Proyecto de Curso**

Identificación de potenciales inhibidores específicos de la enzima PI3K en tres especies de hongos patogénicos:

*Histoplasma capsulatum, Fusarium oxysporum y Aspergillus fumigatus*

**Tareas Generales del Proyecto**

1. Analizar y comparar un conjunto de estructuras de complejos proteína-ligando para la enzima PI3K, disponibles en la base Protein Data Bank (PBD).
2. Construir un modelo computacional de la estructura 3D de la enzima PI3K del hongo, con la calidad requerida para su utilización en simulaciones de acoplamiento molecular (docking).
3. Realizar simulaciones de docking con inhibidores conocidos de la enzima PI3K humana, y comparar los resultados con datos experimentales de la base PDB.
4. *Realizar simulaciones de cribado virtual con el modelo 3D construido de la enzima PI3K, utilizando un subconjunto de estructuras de moléculas pequeñas de una base de datos comercial.*

TAREA EVALUATIVA: Desarrollar la Tarea 4 del proyecto.

Para esta tarea utilizaremos una base de datos de 1000 moléculas en formato PDBQT, proporcionada por el profesor y que se descargará de la Uvirtual (BD-1000m.zip). La base incluye unos pocos inhibidores conocidos de la enzima PI3K.

Cada estudiante utilizará el modelo de PI3K construido para la especie de hongo asignada.

La tarea constará de varias fases:

1. Preparar un script en Python o Bash para correr automáticamente las simulaciones de cribado virtual con el programa Autodock Vina para toda la base de datos.

Notas:

- Crear un archivo de configuración de Vina con los parámetros comunes para todas los moléculas (receptor (proteína), parámetros de la caja, exhaustiveness, número de modelos de salida (2) ).

- Incluir los parámetros asociados al ligando en la línea de comando

2. Ejecutar la simulación de cribado virtual

3. Análisis general de los resultados

3.1. Programar script (función) para leer archivos PDBQT y guardar la información en arreglos en memoria.

3.2. Programar, o incluir en el script anterior, una función para extraer energías de los archivos de resultados de las simulaciones (.pdbqt)

3.3. Construir un histograma de energías (scores) resultantes de las simulaciones (un gráfico para cada pose del ligando)

4. Análisis molecular de los resultados (aplicar solo para la 1ra pose del ligando)

4.1. Identificar las interacciones polares en cada complejo proteína-ligando (entre átomos de oxígeno o nitrógeno en la proteína, y átomos de los elementos químicos N, O, S, F en el ligando). Utilizar un criterio de distancia interatómica (d < 3.5 A) para identificar estas interacciones.

En otras palabras: ¿Qué átomos de N, O, S, F del ligando se encuentran a menos de 3.5 A de algún átomo de oxígeno o nitrógeno de la proteína?

Realizar el análisis solo para el modelo 1 del resultado de docking.

Resultado: Un único archivo que incluya, para cada ligando de la base de datos, la lista de las parejas de átomos identificados y su distancia de interacción,

4.2. Construir un histograma de la frecuencia de interaciones polares por ligando.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ejercicios de prueba de la función de lectura de archivo .pdbqt

1. Utilizar la función construida en un script que lea el archivo pdbqt del modelo de Vps34 construido, e imprima la siguiente información en archivos independientes:

1.1. Carbonos alfa de la proteína (Atom name: CA: Aminoácido (Residue name, con código de 3 letras), Número del aminoácido (Residue sequence number), y coordenadas X, Y, Z

1.2. Carbonos beta (CB) dentro de una distancia de 12 A del carbono alfa (CA) del aminoácido No. 10 (residue sequence number = 10). Imprimir los mismos datos que en 1.1, más el valor de la distancia (con 3 dígitos después del punto).