

Reporte 4: Modelado de proteínas
Bioinformática II: Bioinformática estructural
Licenciatura en Ciencias Genómica
Integrantes del equipo:
Jessica Samantha Cruz Ruiz
Lorena Elizabeth Fajardo Brígido

El modelado de proteínas ha sido un campo de estudio muy importante para la biología y para la bioinformática, debido a que de esta manera se ha intentado predecir cuáles son las posibles maneras en las que se puede plegar una proteína a partir de su secuencia. En este reporte se hace un modelado con proteínas ya reportadas experimentalmente, para concluir que, a partir de la estructura de dos proteínas parecidas en secuencia, se puede predecir su plegamiento.

Las instrucciones que se siguieron en este trabajo y sus respectivos resultados fueron los siguientes:

1) Elige una secuencia S de la superfamilia que elegiste para la tarea 3.

Se utilizó el motivo d1ymga1 de la secuencia de Aquaporina 0 del organismo *Bos taurus*, utilizada también en el ejercicio anterior.

2) Usando HHpred (<http://toolkit.tuebingen.mpg.de/hhpred>) selecciona al menos una estructura molde o template que puedas usar para modelar S, asegurándote que tiene menos del 90% de identidad si fuera posible.

La secuencia que se eligió fue la de 4NEF, la cual es una proteína homóloga de aquaporina en *Homo sapiens* y muestra una identidad de 61% con la proteína de *Bos taurus*.

3) De acuerdo con el ejemplo de http://eead-csic-compbio.github.io/bioinformatica_estructural/node34.html y la documentación de MODELLER construye dos modelos M1 y M2 de S y comprueba su estima de calidad con DOPE.

Los archivos pdb de los dos modelos contruidos están en los archivos query.B999990001.pdb y query.B999990002.pdb, en la carpeta Reporte_3. Su valor de DOPE fue de -30225.38477 y -30289.07617 respectivamente, lo cual indica una buena calidad del modelaje.

4) Evalúa la calidad de los modelos M obtenidos comparándolos con la estructura conocida, que descargaste de SCOP en la tarea 3. Para ello puedes usar MAMMOTH. En tu informe por favor indica el alineamiento obtenido, el RMSD y al menos una imagen de su superposición para brevemente comentar las diferencias que observas entre cada modelo y la estructura experimental.

- Alineamiento con MAMMOTH de la secuencia query.B99990002.pdb (el archivo resultante del alineamiento es log_2.out).



En primer lugar, las imágenes producidas por Rasmol para visualizar la superposición de los modelos con la secuencia original muestra que si tienen un parecido significativo en ambos casos. Ahora, si tomamos en cuenta los valores de RMSD, si bien no son tan pequeños como deberían ser (según la gráfica que se observa en la página de modelado de proteínas), nosotras pensamos que es un buen acercamiento, basando este argumento de igual manera en los alineamientos hechos por Mammoth. Debido a esto, creemos que el modelado se realizó de buena manera, aunque podrían existir parámetros que podríamos mejorar para que el resultado fuera mejor.