Práctica 4

Anastasia Hernández y Alán Muñoz

February 15, 2016

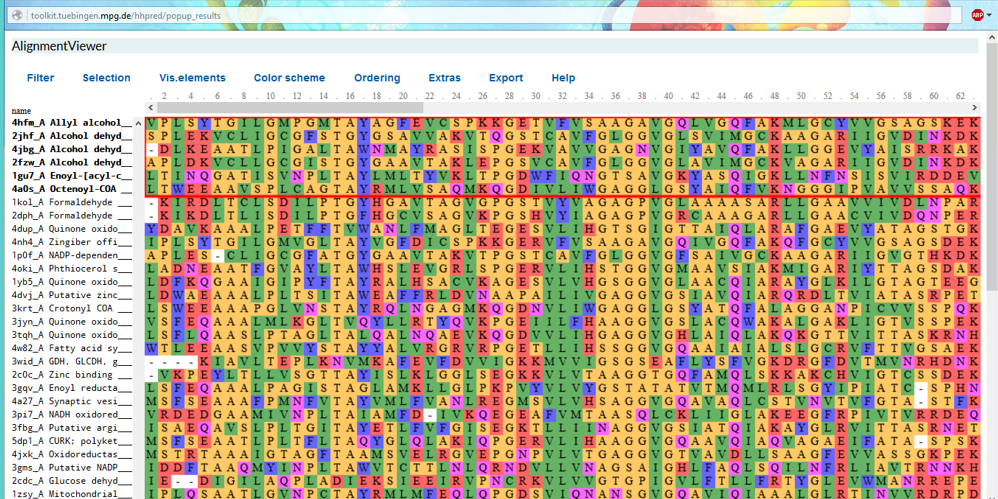
# Práctica 4

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

= De la superfamilia NAD(P)-binding Rossmann-fold domains, se elige la secuencia de d1llua2, obtenida de: <http://scop.berkeley.edu/sunid=91065>.

1. Usando HHpred (<http://toolkit.tuebingen.mpg.de/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.

= Con base en esta secuencia usando los parámetros por default, haciendo énfasis en que el alineamiento fuera local y la base de datos fuera pdb70; se usó HHpred para obtener los alineamientos perfil-perfil obteniendo los siguientes resultados <http://toolkit.tuebingen.mpg.de/hhpred/results/2408346>.



Tras esto, se seleccionaron 2 de los resultados para refinar el alineamiento local de los segmentos alineados y obtener las coordenadas peptídicas en formato .pdb

2jhf\_A Alcohol dehyd\_\_\_\_\_\_\_\_\_\_\_\_\_\_ SPLEKVCLIGCGFSTGYGSAVVAKVTQGSTCAVFGLGGVGLSVIMGCKAAGARIIGVDINKDKFAKAKEVGATECVNPQD KPIQEVLTEMSNGVDFSFEVIGRLDTMVTALSCCQEAGVSVIVGVrstudrsPPDSQNLSMPMLLLSGRTWKGAIFGSKDSVPKLVA DFMAKK

4jbg\_A Alcohol dehyd\_\_\_\_\_\_\_\_\_\_\_\_\_\_ -DLKEAATLPIGALTAWNMAYRASISPGEKVAVVGAGNVGIYAVQFAKLLGGEVYAISRRKAKVSILKSAGADAVLTP-- DEAKSAAP-----FDVVLDPTGS-ASWDLSFGVLGRGGRYVTAGALTAEVRLDLRRLYGMQILVIGATGGRRADFNTVVR LLEAGR

Asegurando así, que el porcentaje de identidad fuera el correcto.

1. De acuerdo con el ejemplo de <http://eead-csic-compbio.github.io/bioinformatica_estructural/node34.html> [ <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.

Se toma la secuencia de d1llua2 en formato fasta para realizar el alineamiento usando ClustalW con las secuencias query. Esto, para obtener los alineamientos necesarios para hacer uso del script 3.3 para la construcción de modelos en formato .pdb.

Estima de calidad con DOPE: -14589.48828 y -14772.64844

Para visualizar directamente los resultados obtenidos: <https://github.com/afermg/bioinfo_lcg/tree/master/Practice%204/modeller>

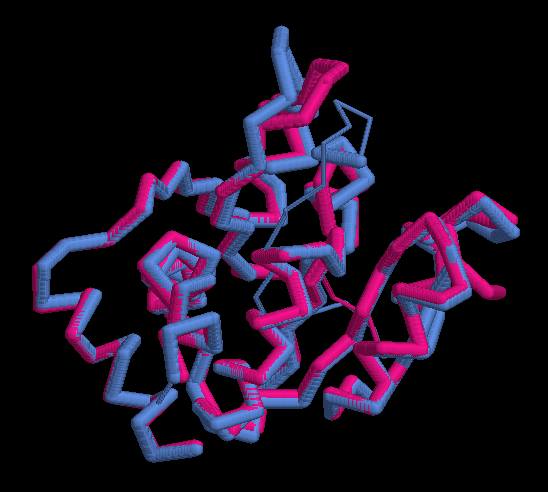
1. Evalúa la calidad de los modelos M obtenidos comparándolos con la estructura conocida, que descargaste de SCOP en la tarea 3. Para ello puedes usar MAMMOTH. En tu informe por favor indica el alineamiento obtenido, el RMSD y al menos una imagen de su superposición para brevemente comentar las diferencias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

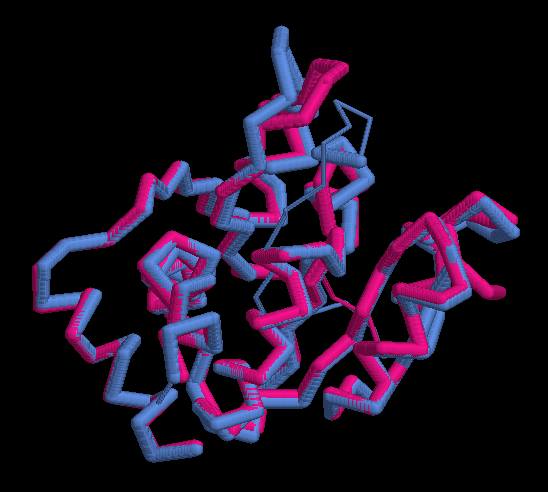
= Ambos modelos se compararon con la secuencia previamente conocida, los alineamientos para M1 y M2 son iguales en secuencia;lo que varía son los valores de z-score (21.2 /20.6)/e-value (.1x10-8/.2x10-8/), aunque los cambios no son significativos.

En el RMSD, los resultados obtenidos para ambos modelos son iguales:

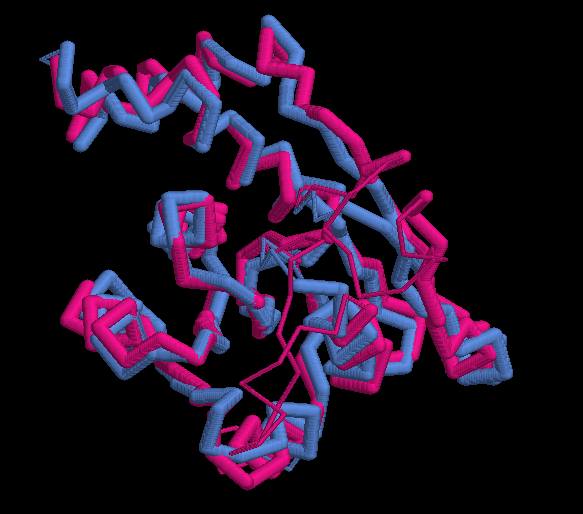
Para M1: total residuos: pdb1 = 157 pdb2 = 166 total residuos alineados = 116 RMSD = 7.28 Angstrom

Para M2: total residuos: pdb1 = 157 pdb2 = 166 total residuos alineados = 115 RMSD = 7.28 Angstrom

Imágenes visualizadas con Rasmol: 



En las imágenes obtenidas puede observarse que la estructura obtenida por cada uno de los modelos (M1 y M2), es la misma.



También, al compararlo con la estructura original de la secuencia d1llua2 elegida en un principio puede observarse que la estcutura original tiene un mayor número de zonas muy plegadas que en cualquiera de los modelos (los cuales únicamente tiene una), pero también tiene zonas conservadas como la zona muy delgada de la estructura que es observable en las 3 estructuras obtenidas.

Para visualizar directamente los resultados obtenidos: <https://github.com/afermg/bioinfo_lcg/tree/master/Practice%204/alns>