

Bioinformática Estructural

Amaranta Manrique de Lara y Ramirez, Valeria Erendira Mateo Estrada

4 de marzo de 2016

Tarea 4: Modelaje de proteínas

En esta práctica se dispone de la secuencia de una proteína A y se quiere conocer su estructura tridimensional.

4.1) Elegir una secuencia S de la superfamilia de la tarea 3.

Elegimos la secuencia de la acuaporina de humano que utilizamos en la práctica anterior (**d1h6ia**). Obtuvimos las estructuras template para hacer el modelo empleando la página (<http://toolkit.tuebingen.mpg.de/hhpred>).

4.2) Seleccionar al menos una estructura molde usando HHpred.

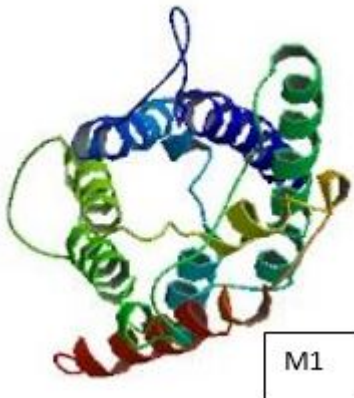
Se escogieron las siguientes estructuras:

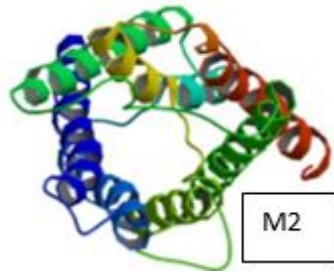
Identificador	Identidad
1j4n_.pdb	80%
3d9s_.pdb	49%
4nef_.pdb	46%

4.3) Utilizar MODELLER para construir dos modelos M1 y M2 de S y comprobar su estima de calidad con DOPE.

Los valores menores en DOPE indican que el modelo es más óptimo.

Modelo	DOPE score
M1	-18567.000000
M2	-25923.000000





4.4) Evaluar la calidad de los modelos obtenidos comparándolos con la estructura conocida.

Para esto se utilizó MAMMOTH, como en la práctica anterior.

Se indican el alineamiento obtenido, el RMSD y se muestra la superposición entre modelo y estructura experimental.

M1

```
mammoth -p d1h6ia_.pdb -e modbase-model_7430d04623f7d90369572c262d50e9c8.pdb -o log.out
```

Z-score	E-value	RMSD
29.204150	0.87496677E-12	4.38 Angstrom

```

Final Structural Alignment
-----

*****
Prediction  ...LFWRRAV VAEFLATTLF VFISIGSALG F.....KYP VGNNQTAVQD
Prediction  ---HHHHHH HHHHHHHHHH HHHHHHHHHH H-----HH- ----HHHHHH
          ||||| ||||| ||||| ||||| || || |||||
Experiment  SSS-HHHHHH HHHHHHHHHH HHHHHHHHHH HHHHHHHHH- ----HHHHHH
Experiment  MASEFKKKLF WRRAVVAEFLA MILFIFISIG SALGFHYPIK SNQITGAVQD
          *****

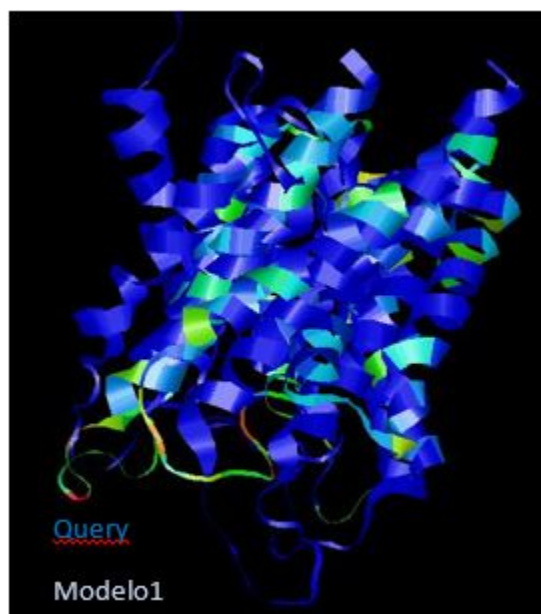
*****
Prediction  NVKVSLAFLG SIATLAQSVG HISGAHLNPA VTLGLLSQC ISIFRALMYI
Prediction  HHHHHHHHHH HHHHHHHHHH ---SSSS--H HHHHHH--H HHHHHHHHHH
          ||||| ||||| ||||| ||||| || || |||||
Experiment  HHHHHHHHHH HHHHHHHHHH ---SSSS--H HHHHHH--H HHHHHHHHHH
Experiment  NVKVSLAFLG SIATLAQSVG HISGAHLNPA VTLGLLSQC ISVLRIMYI
          *****

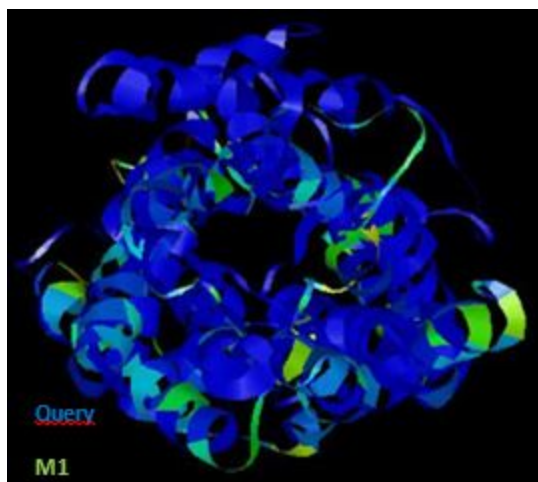
*****
Prediction  IAQCVGAIYA TAILSGITSS LTGNSLGRND LADGVNSGQG LGIEIIGTLQ
Prediction  HHHHHHHHHH HHHHHHH--- ---SSS---S SS--HHHHHH HHHHHHHHHH
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Experiment  HHHHHHHHHH HHHHHHHHHH ---SSS--SS SS---SS-HH HHHHHHHHHH
Experiment  IAQCVGAIYA TAILSGITSS LPDNSLGLNA LAPGVNSGQG LGIEIIGTLQ
          *****

*****
Prediction  LVLVLAATD RRRDLGGSA PLAIGLSVAL GHLLAIDYTG CGINPARSFG
Prediction  HHHHHHHHHH H---HHHHH HHHHHHHHHH HHHHHH--S SSS--HHHHH
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Experiment  HHHHHHHHHH -----HHHH HHHHHHHHHH HHHHHH--S SSS--HHHHH
Experiment  LVLVLAATD RRRDLGGSG PLAIGFSVAL GHLLAIDYTG CGINPARSFG
          *****

*****
Prediction  SAVITHNFSN HWIFWVGPF I GGALAVLIYD FILA.....
Prediction  HHH--SSS-H HHHHHHHHHH HHHHHHHHHH HHHH-----
          ||| || |||| ||||| ||||| |||||
Experiment  HHH--HHHH HHHHHHHHHH HHHHHHHHHH H---HHHHHH HHHHHHHH
Experiment  SSVITHNFQD HWIFWVGPF I GAALAVLIYD FILAPRSSDL TDRVKVWT
          *****

```





M2

```
mammoth -p dlh6ia_.pdb -e modbase-model_6420d06613f7d90369372c262750e4c6.pdb -o log.out
```

Z-score	E-value	RMSD
31.404220	0.83596975E-12	8.72 Angstrom

```

-----
Final Structural Alignment
-----

*****
Prediction LFWRAVVAEF LATTLFVFIS IGSALGFKYP VGNNQTAVQD NVKVSLAFGL
Prediction HHHHHHHHHH HHHHHHHHHH HHHHHHHHHH ----HHHHHH HHHHHHHHHH
          ||||| ||||| ||||| ||||| |||||
Experiment HHHHHHHHHH HHHHHHHHHH HHHHHHHHHH ----HHHH HHHHHHHHHH
Experiment LFWRAVVAEF LATTLFVFIS IGSALGFKYP VGNNQTAVQD NVKVSLAFGL
          *****

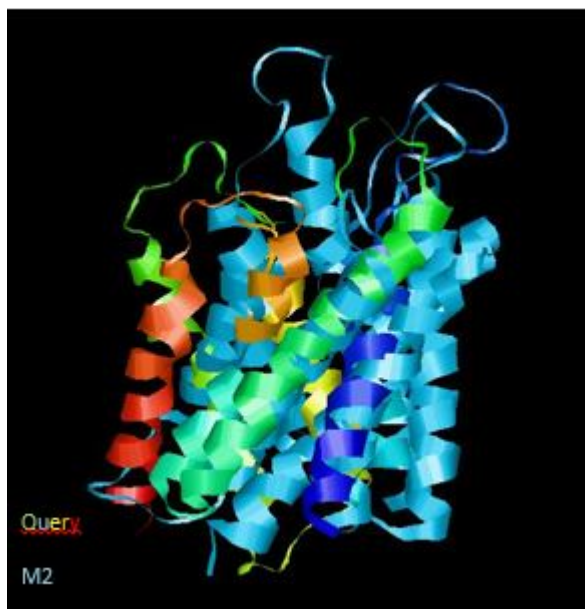
*****
Prediction SIATLAQSVG HISGAHLNPA VTLGLLLSCQ ISIFRALMYI IAQCVGAIVA
Prediction HHHHHHHHHH ---SSSS--H HHHHHH--H HHHHHHHHHH HHHHHHHHHH
          ||||| ||||| || || ||||| |||||
Experiment HHHHHHHHHH ---SSSS-- HHHHHH--H HHHHHHHHHH HHHHHHHHHH
Experiment SIATLAQSVG HISGAHLNPA VTLGLLLSCQ ISIFRALMYI IAQCVGAIVA
          *****

*****
Prediction TAILSGITSS LTGNSLGRND LADGVNSGQG LGIEIIGTLQ LVLCVLATTD
Prediction HHHHHHHH--- ---SSS---S SS--HHHHHH HHHHHHHHHH HHHHHHHHHH
          |||| ||||| ||||| ||||| |||||
Experiment HHHHHHHHHH ---SSS--SS SS--SS-HH HHHHHHHHHH HHHHHHHHHH
Experiment TAILSGITSS LTGNSLGRND LADGVNSGQG LGIEIIGTLQ LVLCVLATTD
          *****

*****
Prediction RRRRDLGGSA PLAIGLSVAL GHLLAIDYTG CGINPARSFG SAVITHNFSN
Prediction H---HHHHH HHHHHHHHHH HHHHHH---S SSS--HHHHH HHH--SSS-H
          ||||| ||||| ||||| ||||| |||||
Experiment ----HHHHH HHHHHHHHHH HHHHHH---S SSS--HHHHH HHH--HHHH
Experiment RRRRDLGGSA PLAIGLSVAL GHLLAIDYTG CGINPARSFG SAVITHNFSN
          *****

*****
Prediction HWIFWVGPF I GGALAVLIYD FILA
Prediction HHHHHHHHHH HHHHHHHHHH HHHH
          ||||| |||||
Experiment HHHHHHHHHH HHHHHHHHHH HHHH
Experiment HWIFWVGPF I GGALAVLIYD FILA
          *****

```





Evaluación

Ya que M1 y la estructura experimental sobrelapan, este modelo es el que mejor se adecua a la acuaporina query. Aunque los loops coinciden en cuanto a las dimensiones, hay problemas en la transposición. Sin embargo, los datos indican que éste es el mejor modelo para la estructura tridimensional de **d1h6ia**. Por ejemplo, el RMSD indica que la estructura tiene un mejor ajuste. En cambio, el modelo M2 tiene un menor ajuste, existe menos superposición entre estructuras, y tanto el Z-score como el RMSD son más altos; todos estos datos indican que M2 está más alejado de la proteína experimental.
