

# Bioinformática estructural

## Tarea 4-Algoritmos 3D

Carina Paola Cornejo Páramo

Claudia Sarai Reyes Ávila

Marzo-2016

### Introducción

La estructura terciaria de una proteína está determinada por su secuencia, sin embargo el número de conformaciones posibles que puede adquirir son demasiadas, ya que se consideran los enlaces del *backbone* de la proteína así como las interacciones que puede establecer (puentes de hidrógeno, interacciones electrostáticas, de Van Der-Waals, hidrofóbicas etc.) Dado el gran número de conformaciones no es fácil predecir la conformación que es funcional, por ello para obtener una conformación aproximada de una proteína se usan moldes o templates que son proteínas relacionadas en secuencia con nuestra proteína de interés, debido a que son homólogas. En el presente trabajo seleccionamos una proteína y evaluamos su estructura tridimensional usando templates como guía.

### Desarrollo

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

En la tarea 3 seleccionamos 3 proteínas de la superfamilia de los citocromos P450, para esta tarea seleccionamos una de éstas, la proteína d2j0da1 de *Homo sapiens*.

- 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.

En la siguiente imagen se muestra el template que seleccionamos, tiene un 22% de identidad.

[illegible]

- 2

[https://github.com/carinapaola/Bioinformatica\\_estructural/blob/master/ModellerScript.py](https://github.com/carinapaola/Bioinformatica_estructural/blob/master/ModellerScript.py)

Calidad para el modelo1:

```
<<< end of ENERGY.  
DOPE score                      : -141886.689193  
Total CPU time [seconds]
```

Calidad para el modelo2:

```
<<< end of ENERGY.  
DOPE score                      : -129091.742344  
Total CPU time [seconds]
```

- 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.

modelo1 vs d2j0da1



PSI(ini)= 94.95 NALI= 432 NORM= 455 RMS= 4.27 NSS= 377  
 PSI(end)= 86.81 NALI= 395 NORM= 455 RMS= 3.18  
 Sstr(LG)= 6178.25 NALI= 395 NORM= 455 RMS= 3.18

E-value= 0.10299921E-17

Z-score= 43.806233 -ln(E)= 41.416981

# Final Structural Alignment

```

*****
Prediction  ....GKLPP GPLPLPGLGN LLHVDFQNTF YCFDQLRRRF GDVFSLQLAW
Prediction  ----SSSSS -SSS---SS SSSSSS--HH HHHHHHHH-- ---SSSS---
          ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Experiment  SSSSS----- -SSS---HH HHH-HHHHHH HHHHHHHH-- ---SSSS---
Experiment  HGLFKKLGIP GPTPLPFLGN ILS.YHKGFC MFDMECHKKY GKVWGFYDQG
          *****

```

```

*****
Prediction  TPVVVLNGLA AVREALVT.H GEDTADRPV PITQILGFGP RSQGVFLARY
Prediction  --SSSS---- HHHHHHHH-H HHHHHHHHHH HHH---SSS- -SSS-----
          ||||| ||||| ||| ||||| ||| ||||| ||
Experiment  -SSSSSS--- HHHHHHHHHH ---SSSSSS- ----S----- SSSS--S-SS
Experiment  QPVLAITDPD MIKTVLVKEC YSVFTNRRP. ....FGPVGK MKSAISI.AE
          *****

```

```

*****
Prediction  GPAWREQRRF ....SVSTLR NLGLGKKSLE QWVTEEAACL CAAFANHSGR
Prediction  --HHHHHHHH ---HHHHHH HHHH--HHHH HHHHHHHHHH HHHHHHH---
          ||||| ||||| ||| ||||| ||| ||||| ||
Experiment  --HHHHHHHH HHHHHHHHHH H---HHHHHH HHHHHHHHHH HHHH-H----
Experiment  DEEWKRLRSL LSPTFTSGKL K...EMVPII AQYGDVLVRN LRRE.AETGK
          *****

```

```

*****
Prediction  PFRPNGLLDK AVSNVIASLT C...GRRFE YDDPRFLRL DLAQEGLKEE
Prediction  -HHHHHHHHH HHHHHHHH-- -----SSS- ---HHHHHHH HHHHHHHHHH
          ||||| ||||| ||| ||||| ||| ||||| ||
Experiment  HHHHHHHHHH HHHHHHHH-- --SSSSS--- -HHHHHH-- HHHHHHHHH-
Experiment  PVTLDKDVFGA YSMDVITSTS FGVNIDSLNN PQDPFVE... NTKKLLRFDF
          *****

```

```
# total residuos: pdb1 = 445 pdb2 = 472
# total residuos alineados = 413

# coordenadas originales = original.pdb
# superposicion optima:
# archivo PDB = align_fit.pdb
# RMSD = 10.46 Angstrom
```

modelo2 vs d2j0da1



PSI(ini)= 96.85 NALI= 431 NORM= 445 RMS= 4.67 NSS= 370  
 PSI(end)= 87.87 NALI= 391 NORM= 445 RMS= 3.35  
 Sstr(LG)= 5975.17 NALI= 391 NORM= 445 RMS= 3.35

E-value= 0.11926224E-17

Z-score= 43.646737 -ln(E)= 41.270377

# Final Structural Alignment

```

*****
Prediction  ....GKLPP GPLPLPGLGN LLHVDQFQNTF YCFDQLRRRF GDVFSLLQLAW
Prediction  -----SSSSS -SSS----SS SSSSSS--HH HHHHHHHH-- ---SSSS---
          |||| |||||||||          |||| ||||||||| |||||||||
Experiment  SSSSS----- -SSS----HH HHH-HHHHHH HHHHHHHH-- ---SSSS---
Experiment  HGLFKKLGIP GPTPLPFLGN ILS.YHKGFC MFDMECHKKY GKVWGFYDGQ
          *****

```

```

*****
Prediction  TPVVVLNGLA AVREALVT.H GEDTADRPPV PITQILGFGP RSQGVFLARY
Prediction  --SSSS---- HHHHHHHH-H HHHHHHHHHH HHH---SSS- -SSS-----
          ||||||||| ||||||||| ||||||||| |||| || || ||
Experiment  -SSSSSS--- HHHHHHHHHH --SSSSSSSS -----SSSS- -S-----SS
Experiment  QPVLAITDPD MIKTVLVKEC YSVFTNRRPF GPVGFMKSAI SI.....AE
          *****

```

```

*****
Prediction  GPAWREQRRF ....SVSTLR NLGLGKKSLE QWTEEAACL CAAFANHSGR
Prediction  --HHHHHHHH --HHHHHH HHHH--HHHH HHHHHHHHHH HHHHHHH--
          ||||||||| ||||||| ||||| ||||||||| |||| |||||
Experiment  --HHHHHHHH HHHHHHHHHH H--HHHHHH HHHHHHHHHH HHHHH-----
Experiment  DEEWKRLRSL LSPTFTSGKL K...EMVPII AQYGDVLVRN LRREA.ETGK
          *****

```

```

*****
Prediction  PFRPNGLLDK AVSNVIASLT CGRRFEYDDP RFLRLDLAQ EGLKEESGFL
Prediction  -HHHHHHHHH HHHHHHHH-- --SSS----H HHHHHHHHHH HHHHHHHHHH
          ||||||||| ||||||||| ||||||||| ||||||||| ||||
Experiment  HHHHHHHHHH HHHHHHHHHH-- --SSSS---H HHHHHHHHHH- -----HHHHH
Experiment  PVTCLKDVFGA YSMDVITSTS FGVNIDSNPQ DPFVENTKK. ....LLRFF
          *****

```

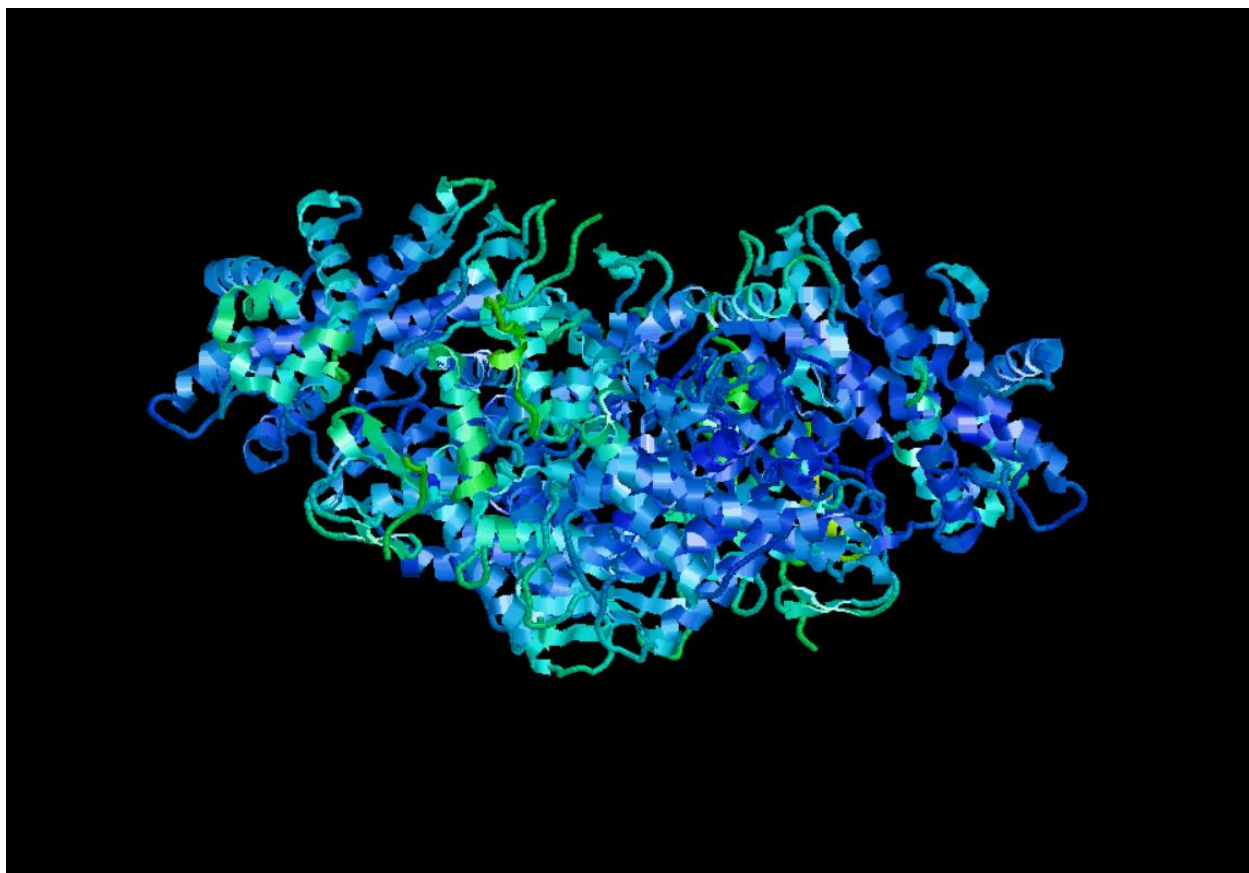
```
# total residuos: pdb1 = 445 pdb2 = 472  
# total residuos alineados = 413  
  
# coordenadas originales = original.pdb  
# superposicion optima:  
# archivo PDB = align_fit.pdb  
# RMSD = 8.76 Angstrom
```

---

d2j0da1



modelo1

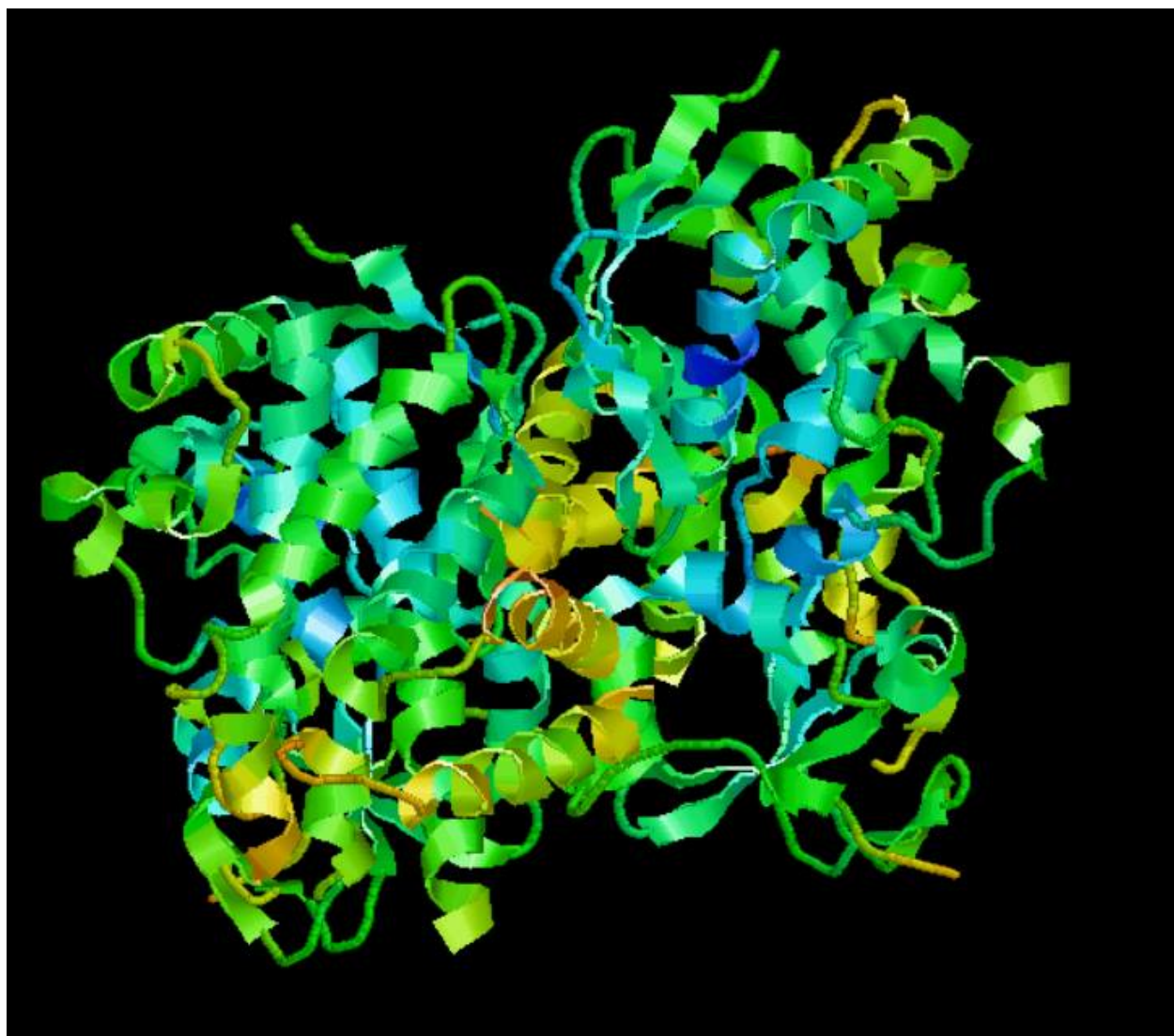


d2j0da1 vs modelo1

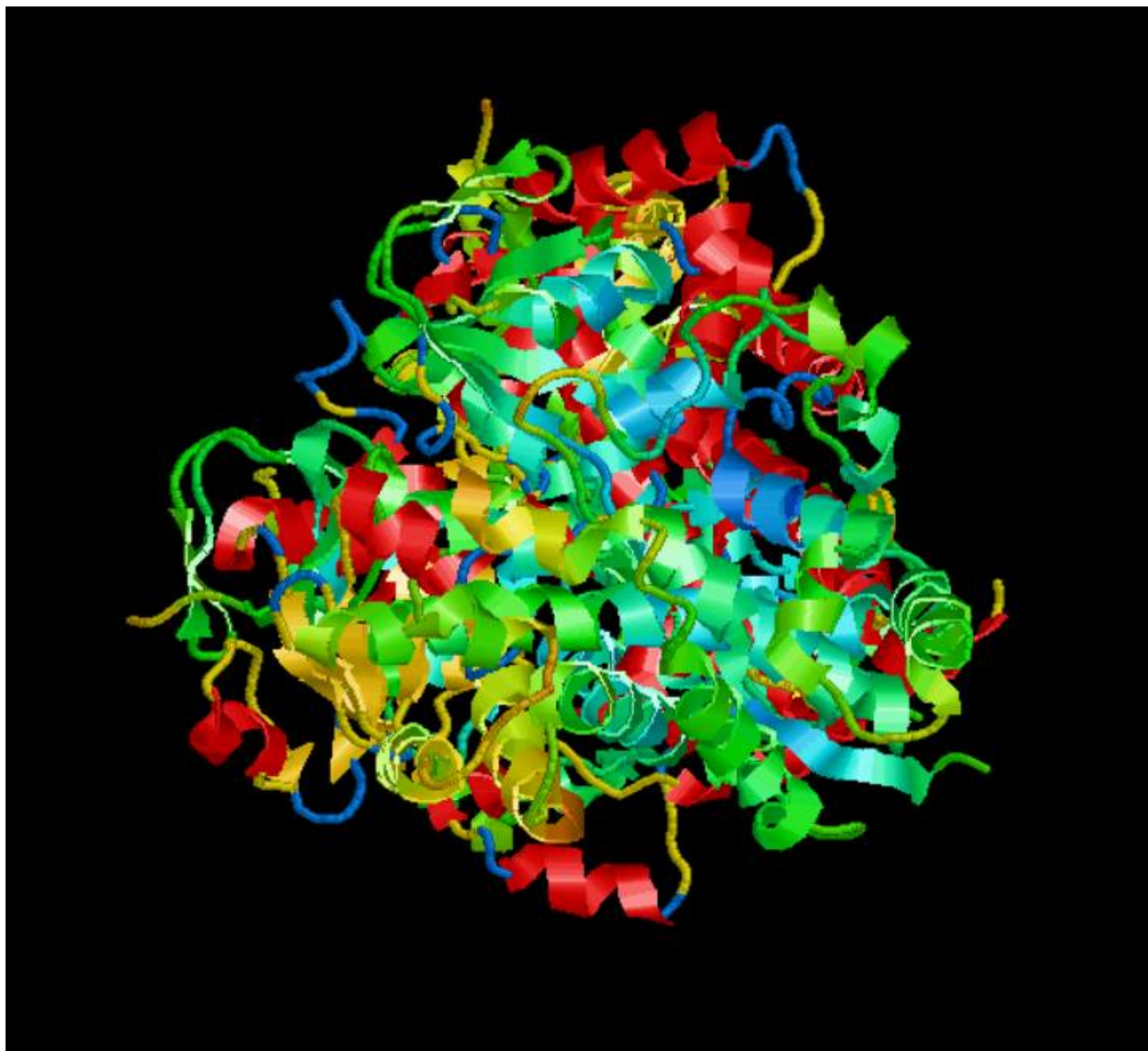




modelo2



d2j0da1 vs modelo2



Se aprecia que los modelos si son diferentes del citocromo P450 d2j0da1, suponemos que se debe a que nuestro template tiene solo 22% de identidad con la proteína query, escogimos dicho template porque el resto de los posibles templates tenían valores muy cercanos, todos alrededor del 20% de identidad, o tenían el 100% de identidad, pudiendo ser la misma proteína. A simple vista los modelos no parecen exactamente iguales a la proteína d2j0da1, sin embargo los valores DOPE, que nos indican la calidad del modelo es negativa, lo cual indica que son buenos modelos.

El DOPE *score* del modelo1 es más negativo, además su RMSD es ligeramente mayor que la del modelo2, por lo que asumimos que el modelo 1 es un mejor modelo de nuestra proteína.