

Guia prático de Modeller auxiliar ao guia usado em aula

1. Vamos fazer um BLASTp com uma proteína do vírus da hepatite C para encontrar um depósito de estrutura resolvida depositada no Protein Data Bank, PDB.
 - a. Abra um navegador e procure pelo site do BLAST do NCBI e clique em **Protein BLAST**
 - b. Copie e cole esta sequência em **Enter Query Sequence** e usando as Standard Databases mesmo, mude **Database** de **Non-redundant protein sequences (nr)** para **Protein Data Bank Proteins (pdb)** para buscar em registros de estruturas:

```
>proteína_do_virus_hepatite_C
SLSYSWTGALVTATRREERRHPIGPLSNLTIKHNLVYQTTASASARMTKVTIDREQILDKHY
FDITVTAVKKKASEVTADLLTWDEVARLTPKNTARSKSGLSGSDVRQLTRAARELNSMWQDLLS
DSEELIPTVMKNEVFVSSPTARKPARLIVYDLPVLRACEKRAMYDLFQKLPYAVMGKAYGFQ
YTPRQRVNRLLDMWRHFKNPMGFSYDTKCFDSTVTPHDIDTERDIFLSATLPDEAKTVIKNLTS
RLYRGSPMYNSRGDLVGKRECRASGVFPPTSMGNTLTNFIKATAAAKAAGLSDPQFLICGDDLVC
ITSSKGVEEDEQALREFTSAMTKYSAIPGDLPKPYDLEQITSCSSNVTVAQDRNGRPYYFLTR
DPTTPLARASWETISHSPVNSWLGNIIAFAPT VWVRLVFLTHFFGLLLQODAVDRNYEFEMYGS
TYSVNPLDLPALIIYKLGPEAFDLTNYSPEYVQRVAAALQKLGSPPLRAWKRRAKLDRSKLKVR
GGRYAVVADYLFGFASAYRPKRPAPPGVNSIDVSGWFSIGDDSIGDIYRQ
```

- c. O resultado apresentou vários resultados com similaridade de 99% e valor 0.0 de E-value certo? Vamos abrir com o botão do meio o Accession (última coluna) dos doze primeiros para abrirem em outra aba e verificar do lado direito, abaixo da pequena estrutura, o valor de “Resolution” e repare que a décima segunda tem o menor valor, a 3HKM_A (cadeia A de 3HKM), 1,55 angstroms. Vamos usar ela.
2. Logar no servidor bioufmg@bioinfo.icb.ufmg.br (taxiscarrara) e entrar em sua pasta [cd eusoujacu] e criar uma pasta [mkdir modeller_pratica] e após entrar nela [cd modeller_pratica], dê um pwd para confirmar que está nela.
 3. Copie tudo da conta treinamento o material para a pasta modeller_pratica:

```
cp /home/treinamento/modeller_pratica/* . [não esqueça o ponto de presente diretório]
```

4. Olhe os arquivos com [ls] e veja que vc tem um arquivo pdb ali, mas para este tutorial vamos baixar a 3HKM, com wget do site do rcsb, assim:

```
wget https://files.rcsb.org/view/3HKW.pdb
```

5. Vamos usar três arquivos, dê um less neles
6. [less 3HKW.pdb] e pra sair aperte “q” de quit
7. [less model.ali] que é a sequência a ser modelada em formato PIR que vc não conhecia
8. [less align2d.py] que é um script python preparado para usar como molde a estrutura 3gsz mas vamos mudar para usar a 3HKW com programa vi assim:
 - a. vi align2d.py
 - b. para entrar no modo de inserção aperte a letra “i”
 - c. com as setinhas do teclado pare depois das cinco ocorrências de 3gsz, apague para trás os caracteres 3gsz e substitua por 3HKW
 - d. para sair de INSERÇÃO aperte Esc (lado esquerdo superior do teclado)

- e. para poder dar comando coloque ":" usando shift dois pontos, os dois pontos aparecerão na parte inferior da tela esperando o comando
 - f. digite o comando de salvar que pode ser wq de write and quit ou x!
 - g. pode agora dar um [**less align2d.py**] e repare que ele pede com aln.write para o script escrever as saídas PIR e PAP. Aperte "q" pra sair de less.
9. Vamos rodar o script para alinhar em 2D a sequência a ser modelada com a do molde com o comando abaixo mas vai demorar pra ter o cursos de volta, aguarde:

python align2d.py

10. Como pedido no script com aln.write nas linhas finais, dois arquivos serão escrito, dê um less no alinhamento das duas sequências [**less model-3HKW.ali**] e [**less model-3HKW.pap**] que mostra os resíduos conservados no alinhamento marcados com um asterisco, que serão usados como base para criar a estrutura.
11. Agora vamos gerar o modelo usando os arquivos 3HKW.pdb que vc tem, o alinhamento que vc criou model-3HKW.ali e um script que vamos editar como acima, trocando as citações de 3gsz por 3HKW, abrindo ele com
- a. **vi genmodel.py**
 - b. apertar "i" para abrir modo de INSERÇÃO
 - c. parar depois de 3gsz, apagar, substituir por 3HKW, são apenas 2 ocorrências
 - d. caso queira gerar mais modelos, pois geralmente são usados 100, parar depois de 10 na linha de "a.ending_mode" e adicionar um zero, mas faremos somente 10 modelos propostos pelo Modeller para o objetivo do treinamento.
 - e. Aperte "Esc" pra sair de INSERÇÃO, digite ":" e salve com wq
 - f. Dê um [**less genmodel.py**] para verificar e saia com "q"
12. Agora vamos rodar o script e aguardar até o cursor voltar:

python genmodel.py >&1 | tee genmodel.lo

13. O arquivo resultante genmodel.log tem no final uma tabela com vários valores para os modelos, e vamos escolher aquele modelo que tem o valor mais negativo de "DOPE score" que seria o de menor energia e possivelmente o melhor. Anote o identificador do modelo (concentre-se no último algarismo), abrindo o arquivo de log genmodel.lo com comando para ver a cauda dele, no caso as última 13 linhas [**tail -n 13 genmodel.lo**]. E você pode ordenar pela coluna 3 (-k 3) numericamente (-n), a coluna do DOPE score, assim:

tail -n 13 genmodel.lo | sort -k 3 -n

14. Para capturar o modelo vamos usar o acesso pela web à conta bioufmg acessando em um navegador o link <http://bioinfo.icb.ufmg.br/bioufmg> digitando para acessar usuário bioufmg e senha carrarataxis. A pasta "olhe" dá acesso às nossas pastas, entre na sua, entre em modeller_pratica e com o botão da direita salve na sua área de trabalho o arquivo do modelo que tem o menor DOPE score.
15. Abra o modelo com dois cliques no arquivo ou abra seu Pymol e use na janela branca File... Open... e dê na linha do modelo um hide [H] everything e um show [S] cartoon, vc está vendo a estrutura que vc modelou a partir da estrutura primária.

16. Na janela branca ou mesmo na parte de baixo da preta dê um fetch 3HKW, note que vêm três estruturas cristalizadas. Vamos alinhar agora o modelo com a 3HKW acessando o menu Action [A] do modelo e indicar [A] → align → to molecule → 3HKW
17. Apague, clicando nos nomes, o modelo e a 3HKW, e vc verá pequenas linhas que são as que conectam os resíduos no modelo com o molde, ou seja, mostram como o modelo ficou parecido com a estrutura molde. Verifique que na janela branca foi mostrado o valor de RMS (raiz do valor quadrático médio), e anote o valor, que deve ser parecido com este:

Executive: RMS = 0.108 (489 to 489 atoms)

18. Vc pode colorir usando a letra [C] para cada estrutura, por exemplo para o modelo [C] → reds → red e para a 3HKW [C] → greens → green ou outras cores. E vc pode tb dar um hide [H] everything no modelo e show surface e na 3HKW um hide [H] everything e show [S] mesh. Use o botão da direita do mouse para zoon e mova as estruturas clicando e movendo. Vc obteve um modelo 3D fabricado por modelagem molecular contra a estrutura de melhor resolução, agora vamos checar sua qualidade.
19. Acessar pelo navegador o servidor:

<https://swissmodel.expasy.org/assess>

No campo **Structure file** clique em **+Upload Coordinate File...** e aponte o arquivo do modelo seu modelo. Clique então no botão **Start Assessment**. Avalie o plot de *Ramachandran* que analisa se os ângulos de giro, de um resíduo para os seus vizinhos, está dentro dos limites ótimos (verde escuro) ou bons (verde claro), com ênfase para glicinas e prolinas.