

Tutorial para predição de estrutura de proteínas

VIII Escola de Modelagem Molecular em Sistemas Biológicos

Priscila V. Z. Capriles Goliatt - priscila.capriles@ufjf.edu.br

Fábio Lima Custódio - flc@Incc.br, Gregório Kappaun Rocha - gregorio@Incc.br, Karina Baptista dos Santos - karinabs@Incc.br

Comandos Básicos do Linux:

cd diretório/	Entra na pasta diretório/
ls diretório/	Lista as subpastas e arquivos existentes em diretório/
cd ..	Retorna para a pasta anterior
mkdir diretório/	Cria a pasta diretório/
rm arquivo.txt	Remove arquivo.txt (.txt ou outro formato de arquivo)
rmdir diretório/	Remove a pasta diretório/
cp diretório1/arquivo.txt diretório2/	Copia arquivo.txt no diretório1/ para diretório2/

Atenção: veja também o tutorial básico de linux em <http://www.emmsb.Incc.br/index.php?pagina=26>

Parte 4: Predição via método híbrido: modelagem comparativa + *ab initio*

Passo 1: Construindo o modelo via MHOLline

1. Acessar o site do MHOLline: <http://www.mholline2.Incc.br>
2. Logar na conta “emmsb”.
3. Selecionar todos os programas
4. Fazer o upload do arquivo FASTA (**Atenção: não submeter o job!!!!**)
5. Acesse o resultado na caixa lateral “My Results → All” ou “My Results → Finished”

Atenção: Os resultados do MHOLline já estão sSODEF17s no diretório:

`./modelagem/sodef/jobSODEF17`

Passo 2: Analisar o modelo construído via MHOLline

1. Visualizar o resultado de cada módulo no respectivo ícone.
2. Visualiza o resumo do resultado de modelagem clicando no botão “View”
3. Clique no botão “Refine” e analise o alinhamento global e conseração de estrutura secundária
4. Visualizar a estrutura do molde (*template*) encontrado pelo MHOLline (acesse o site: <http://www.pdb.org/pdb/home/home.do>)

5. Em “Search” digite o nome da estrutura 2KSK.
6. Analisar os dados da estrutura (e.g. Espécie, Método Experimental, Resolução, Ligantes).
7. Clicar em “Sequence”: Analisar a estrutura secundária e verificar a existência ou não de ligações particulares (e.g. ligações dissulfídicas).
8. Visualizar a estrutura do modelo gerado via MHOLline: abrir com um visualizador 3D (vmd) o arquivo ./modelagem/sodef_new/JOBSSODEF17/Modeller/1_SODEF17/1_SODEF17.B99990001.pdb:

Digite na linha de comando de um terminal:

```
vmd -m 1_SODEF17.B99990001.pdb 2KSK_A.atm
```

Atenção: veja o tutorial básico de VMD em
<http://www.emmsb.lncc.br/index.php?pagina=26>

Para mudar a visualização no vmd: Clique em “Graphics → Representations”, em “Coloring Method” use preferencialmente a opção “Secondary Structure” e em “Drawing Method” use preferencialmente a opção “New Cartoon”.

Alinhar as estruturas 3D:

Clique em: “Extensions” → “Analysis” → “MultiSeq”.

Selecionar: “VMD Protein Structures”.

Clicar em: “Tools” → “Stamp Structural Alignment” → “Marked Structures” → “OK”.

Use também outros visualizadores como rasmol ou pymol.

9. Visualizar a estrutura do modelo gerado via MHOLline: abrir com um visualizador 3D (pymol) o arquivo ./modelagem/sodef_new/JOBSSODEF17/Modeller/1_SODEF17/1_SODEF17.B99990001.pdb:

Digite na linha de comando de um terminal:

```
pymol 1_SODEF17.B99990001.pdb 2KSK_A.atm
```

Para mudar a visualização no pymol: No quadrante superior direito use o “Show (S) → as → cartoon” e em “Coloring (C)” use preferencialmente diferentes cores para cada estrutura.

Alinhar as estruturas 3D:

Na linha do template 2KSK_A Clique em: “Action (A)” → “align” → “all to this”.

10. Visualizar a qualidade da estrutura avaliada com o programa Procheck: abrir a figura ./modelagem/sodef_new/JOBSSODEF17/Procheck/1_SODEF17.B99990001_01.ps

Para abrir o gráfico de Ramachandran gerado, digite na linha de comando de um terminal:

```
okular 1_SODEF17.B99990001_01.ps&
```

11. Visualizar a qualidade da estrutura avaliada com o programa Molprobity: abrir a figura ./modelagem/sodef_new/JOBSSODEF17/Molprobity/1_SODEF17.B99990001.pdf

Para abrir o gráfico de Ramachandran gerado, digite na linha de comando de um terminal:

```
okular SODEF17.B99990001.pdf&
```

Passo 3: Construindo o modelo via Modeller

5.1. Modelagem baseada em mais de um molde (template):

Preparando os arquivos de entrada:

- 1 Realizar o alinhamento múltiplo entre os possíveis templates obtidos com o MHOline e a sequência alvo através do site: <http://www.ebi.ac.uk/Tools/msa/clustalo/>
- 2 Na opção "OUTPUT FORMAT", selecionar o formato "PIR" (formato usado pelo MODELLER) e em "OUTPUT ORDER", selecionar input

- 3 Abaixo, colar as seqüências FASTA dos templates e da sequência alvo contidas no arquivo:

```
/home/vemmsb/modelagem/sodef/sequences/multi.fasta
```

- 4 Clicar em "RUN", visualizar o resultado e salvar o arquivo de resultado no formato ALN.

Atenção: Os resultados do ClustalO já estão salvos no diretório:

```
./modelagem/sodef/sequences
```

Gerando o modelo da proteína alvo com o Modeller:

- 1 Digite na linha de comando do terminal: `mod9.17 model-multi.py`

Atenção: irá aparecer a mensagem: 'import site' failed; use -v for traceback. Não se preocupe!

- 2 O *script* model-multi.py lê o arquivo SODEF17.ali, gera três modelos para a proteína alvo através da função automodel(), lista dentro do arquivo model-multi.log a pontuação dos modelos gerados de acordo com a função de energia molpdf e com os métodos DOPE (*Discrete Optimized Protein Energy*) e GA341, e retorna o nome da melhor estrutura gerada de acordo com a pontuação do DOPE. Através da função cluster(cluster_cut=1.00), clusteriza os modelos gerados e gera uma estrutura representativa já otimizada via Gradiente Conjugado (esta função é interessante para quando se deseja gerar uma grande quantidade de modelos). O resumo dos resultados é gerado dentro do arquivo model-multi.out.

Atenção: Os modelos gerados (SODEF17.B99990001.pdb, SODEF17.B99990002.pdb e SODEF17.B99990003.pdb) já encontram-se no diretório ./sequences/multiopt/

Atenção: A estrutura representativa inicial (cluster.ini) e otimizada (cluster.opt) encontram-se no diretório ./sequences/multiopt/cluster/

Avaliando os modelos 3D gerados:

- 3 Abra com um editor de texto o arquivo ./modelagem/basico/modeller/model-multi.out
- 4 Observe os resultados do DOPE, DOPE normalizado e GA341 para os modelos e para a estrutura representativa (cluster.opt).

Observação: o DOPE normalizado é um Z-score, ou seja, scores positivos estão associados a modelos de baixa qualidade, por outro lado, scores abaixo de -1 estão associados a modelos considerados *native-like*.

- 5 Via terminal, entre no diretório ./sequences/multiplt/
- 6 Digite na linha de comando do terminal:

```
vmd -m SODEF17.B9999000*.pdb cluster/cluster.opt ../*.pdb
```
- 7 Alinhe todas as estruturas.
- 8 Compare todos os modelos com a estrutura dos templates

Validando os modelos gerados via Molprobit:

- 9 Avaliar os modelos gerados no diretório ./sequences/multiplt/ e no diretório ./sequences/multiplt/
- 10 Entre no diretório ./sequences/
- 11 Digite na linha de comando do terminal:
evince ./multiplt/molprobitresults/charts/*-rama.pdf &