

Bioinformática

Exercícios - Previsão da estrutura de proteínas

1. A proteína CspA (major cold shock protein) é produzida por certas bactérias em resposta a uma descida brusca de temperatura (cold shock). Esta proteína liga-se aos ácidos nucleicos em cadeia simples e acredita-se que contribua para a sua estabilização em condições extremas. Existem estruturas de vários membros desta família, pretendendo-se aqui construir um modelo para a proteína CspA de *Staphylococcus aureus*, da qual se conhece a sequência, mas cuja estrutura não foi ainda obtida experimentalmente.
 - a) Faça uma pesquisa da sequência da CspA contra as sequências das proteínas de estrutura conhecida. Para tal, obtenha a sequência no site Uniprot (www.uniprot.org) e usando o "tab" Blast, faça uma pesquisa contra a base de dados PDB. Seleccione o resultado na base do E-value e comprimento do alinhamento (pretendemos um "template" que cubra a totalidade da sequência).
 - b) Analise a entrada Uniprot do template e escolha um ficheiro PDB com uma resolução inferior a 2.0 Å.
 - c) Para construir o modelo, usamos o site <http://swissmodel.expasy.org>, com a opção "Automated Mode". Nas "Advanced Options" insira o código do ficheiro PDB escolhida na alínea anterior, bem com da cadeia (chain). Deverá introduzir um email válido, sendo que a página do site é actualizada após alguns minutos com o resultado da modelação.
 - d) Faça download do ficheiro pdb do modelo ("download model as pdb"), e carregue-o no PyMOL, juntamente com ficheiro PDB do template que escolheu. Compare as duas estruturas.
2. Vamos considerar agora um caso bastante mais difícil, a construção de um modelo para proteína dihidrofolato redutase (dehydrofolate reductase) de *Vibrio Cholerae*, mais uma vez com sequência, mas não estrutura conhecida. Neste caso os "templates" que irá encontrar terão uma percentagem de identidade da ordem dos 35%. Repita os passos exactos do exemplo anterior, e observa as diferenças entre o modelo final obtido e o template.

