## Bioinformática

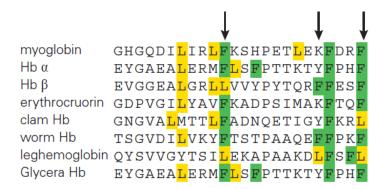
## Exercícios TP4: Algoritmos de alinhamento e matrizes de score

1. Considere as duas sequências:

HEAGAWGHEE

**PAWHEAE** 

- a) Construa a matriz de alinhamento para estas sequências e calcule o alinhamento global óptimo pelo algoritmo de Needleman-Wunsch, considerando uma gap penalty de -8 e uma matriz de score BLOSUM50.
- b) Construa um alinhamento semelhante, mas usando o método de alinhamento local de Smith-Waterman.
- 2. Obtenha as sequências da tripisina humana e de ratazana. Produza o seu alinhamento global com o progama "needle" (<a href="www.ebi.ac.uk/emboss/Tools/psa">www.ebi.ac.uk/emboss/Tools/psa</a>). O software usa por defeito uma matriz BLOSUM62 para fazer este alinhamento, mas qual lhe parece que deveria ser a matriz apropriada para esta situação. Volte a correr "needl" com esta matriz e compare com o resultado anterior. Obtenha a sequência da tripsina de Litopenaeus vannamei (white leg shrimp) e faça o alinhamento desta com a humana. Qual seria a matriz de alinhamento mais apropriada nesta situação?... Repita o alinhamento com esta última e compare.
- **3.** Utilize o site <u>DALI</u> para fazer o alinhamento estrutural das duas proteínas tripsina 1 humana e tripsina da bactéria *Streptomyces griseus*. Compare este alinhamento com o obtido com a ferrramenta EMBOSS (<u>water</u>). Modifique os valores por defeito das penalidades de abertura e extenão de *gaps* e obseve o efeito sobre o socre o alinhamento, procurando encontrar os valores que melhore fazem corresponder o alinhamento observado ao alinhado estrutural.
  - **4.** Considere o seguinte fragmento de um bloco da base de dados BLOCKS, para a família das globulinas:



com base nas frequências de ocorrência dos aminoácidos e frequência de ocorrência dos pares fenilalanina-leucina (FL), obtenha um valor aproximado para o score da substituição Phe-Leu na matriz BLOSUM.