

LISTA DE EXERCÍCIOS II - Parte I

Instruções:

- A resolução do exercício deve ser feita **individualmente**. Cópias evidentes entre trabalhos não serão aceitas.
- A entrega deve ser online via Moodle (exclusivamente), somente até a data especificada.
- Para cada uma das tarefas deve-se entregar o código fonte. O nome do arquivo deve identificar a tarefa, exemplo "e2-1a.py" referente ao item "1a" da tarefa. Arquivos corrompidos serão desconsiderados.
- Além do código fonte deve-se entregar um único arquivo PDF com o nome "e1-il.pdf" apresentando o pseudocódigo do algoritmo desenvolvido e os resultados encontrados.
- Não serão aceitos trabalhos atrasados.
- Data de entrega: 17.09.2019 (terça-feira) até as 23:55 via Moodle (<https://moodle.ufrgs.br/login/index.php>).

NOME: CARTÃO:

1. *A hemoglobina dos vertebrados, o transportador de oxigênio nas hemácias, é constituída por quatro cadeias polipeptídicas, duas de um tipo, e duas de outro. As quatro são mantidas juntas por ligações não covalentes. Cada uma contém um grupo heme e um só centro de ligação ao oxigênio. A hemoglobina A, a principal dos adultos, é constituída por duas cadeias alfa (α) e duas beta (β). A seguir está listada a sequência de aminoácidos da hemoglobina α (apenas uma cadeia) em diferentes espécies. Nota-se que ocorrem variações de uma sequência quando comparada a outra, entretanto a hemoglobina exerce a mesma função independente da espécie: é responsável pelo transporte de oxigênio e gás carbônico.*

Human (Homo sapiens) [TaxId: 9606]

2DN3:A|PDBID|CHAIN|SEQUENCE

VLSPADKTNVKAAGKVGAGHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNAL
SALSDELHAKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTISKY

Horse (Equus caballus) [TaxId: 9796]

2D5X:A|PDBID|CHAIN|SEQUENCE

VLSAADKTNVKAAGKVGAGHAGEYGAEALERMFLGFPTTKTYFPHFDLSHGSAQVKAHGKKVGDALTAVGHLDLPGAL
SNLSDELHAKLRVDPVNFKLLSHCLLVTLAVHLPNDFTPAVHASLDKFLSSVSTVLTISKYR

Deer (Odocoileus virginianus) [TaxId: 9874]

1HDS:A|PDBID|CHAIN|SEQUENCE

VLSAANKSNVKAAGKVGGNAPAYGAQALQRMFLSFPTTKTYFPHFDLSHGSAQVKAHGQKVANALTKAQHGLNDLPGTL
SNLSNLHAKLRVNPVNFKLLSHSLLVTLASHLPTNFTPAVHANLNKFLANDSTVLTISKYR

Cow (Bos taurus) [TaxId: 9913]

2QSS:A|PDBID|CHAIN|SEQUENCE

VLSAADKGNVKAAGKVGGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHGAKVAAALTKAVEHLDLPGAL

SELSDLHAHKL RVPVNFKLLSHSLLVTLASHLP SDFTPAVHASL DKFLANVSTVLTSKYR

Pig (*Sus scrofa*) [TaxId: 9823]

1QPW:A|PDBID|CHAIN|SEQUENCE

VLSAADKANVKA AWGKVGQAGAHGAEALERMFLGFPTTKTYFPHFNLSHGSDQVKAHGQKVADALTKAVGHLDDLPGAL
SALS DLHAHKL RVPVNFKLLSHCLLVTLAAHHPDDFNPSVHASL DKFLANVSTVLTSKYR

Maned wolf (*Chrysocyon brachyurus*) [TaxId: 68728]

1FHJ:A|PDBID|CHAIN|SEQUENCE

VLSPADKTNIKSTWDKIGGHAGDYGG EALDRTFQSFPTTKTYFPHFDLSPGSAQVKAHGKKVADALTTAVAHLDLPGAL
SALS DLHAYKL RVPVNFKLLSHCLLVTLACHHPTEFTP AVHASL DKFFTAVSTVLTSKYR

Chicken (*Gallus gallus*) [TaxId: 9031]

1HBR:A|PDBID|CHAIN|SEQUENCE

MLTAEDKKLIQQAWEKAASHQEEFGAEALTRMFTTYPQT KYFPHFDLSPGSDQVRGHGKKVLGALGNAVKNVNLSQAM
AELSNLHAYNL RVPVNFKLLSQCIQVVLAVHMGKDYTPEVHA AFDKFLSAVSAVLA EKYR

Trout (*Oncorhynchus mykiss*) [TaxId: 8022]

1OUT:A|PDBID|CHAIN|SEQUENCE

XSLTAKDKSVVKA FWGKISGKADVVGAEALGRMLTAYPQTKTYF SHWADLSPGSGPVKKHGGIIMGAIGKAVGLMDDL VG
GMSALS DLHAFKL RVPDGNFKILSHN ILVTLAIHFPSDFTPEVHIAVDKFLAAVSAALADKYR

- (a) Implemente um conjunto de rotinas computacionais para realizar as tarefas que seguem:
- Desenvolva um algoritmo que leia duas seqüências n e m .
 - Implemente o algoritmo de Needleman-Wunsch para alinhar estas seqüências.
 - O programa deve imprimir na tela a tabela final de alinhamento, as duas cadeias alinhadas e identidade do alinhamento.
 - O programa deve imprimir na tela o score do alinhamento.
 - Regras:
 - Valor para GAP: -4
 - Valor para Match: +5
 - Valor para Mismatch: -3
 - $S_{i,j} = \text{MAX} \begin{bmatrix} S_{i-1,j-1} + s(a_i, b_j)(\text{match/mismatch}) \\ S_{i,j-1} + \text{gap}(\text{seq1}) \\ S_{i-1,j} + \text{gap}(\text{seq2}) \end{bmatrix}$
 - Com base na identidade e no score do alinhamento informe qual das espécies acima apresenta a maior semelhança, em termos de seqüência, com a espécie humana (*homo sapiens*).
2. A matriz BLOSUM (BLOCKs of Amino Acid SUBstitution Matrix) é uma matriz de substituição usada para o alinhamento de seqüências de proteínas. Matrizes BLOSUM são usadas para pontuar alinhamentos entre seqüências de proteínas divergentes.

- (a) Implemente um conjunto de rotinas computacionais para realizar as tarefas que seguem:
- Desenvolva um algoritmo que leia duas seqüências n e m .
 - Implemente o algoritmo de Needleman-Wunsch para alinhar estas seqüências.
 - O programa deve imprimir na tela a tabela final de alinhamento, as duas cadeias alinhadas e identidade do alinhamento.
 - O programa deve imprimir na tela o score do alinhamento.
 - Regras:
 - Valor para GAP: -4
 - Valor para Match: considerar a pontuação da Matriz Blosum 62

- $$\bullet S_{i,j} = \text{MAX} \begin{bmatrix} S_{i-1,j-1} + s(a_i, b_j)(\text{match/mismatch}) \\ S_{i,j-1} + \text{gap}(\text{seq1}) \\ S_{i-1,j} + \text{gap}(\text{seq2}) \end{bmatrix}$$

vi. Com base na identidade e no score do alinhamento informe qual das espécies da questão 1 apresenta a maior semelhança, em termos de sequência, com a espécie humana (*homo sapiens*).

	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
Ala	4																			
Arg	-1	5																		
Asn	-2	0	6																	
Asp	-2	-2	1	6																
Cys	0	-3	-3	-3	9															
Gln	-1	1	0	0	-3	5														
Glu	-1	0	0	2	-4	2	5													
Gly	0	-2	0	-1	-3	-2	-2	6												
His	-2	0	1	-1	-3	0	0	-2	8											
Ile	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
Leu	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
Lys	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
Met	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
Phe	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
Pro	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
Ser	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
Thr	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
Trp	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Tyr	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	1	-2	-1	3	-3	-2	-2	2	7	
Val	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	0	-3	-1	4	

Figure 1: Matriz de Substituição Blosun62.

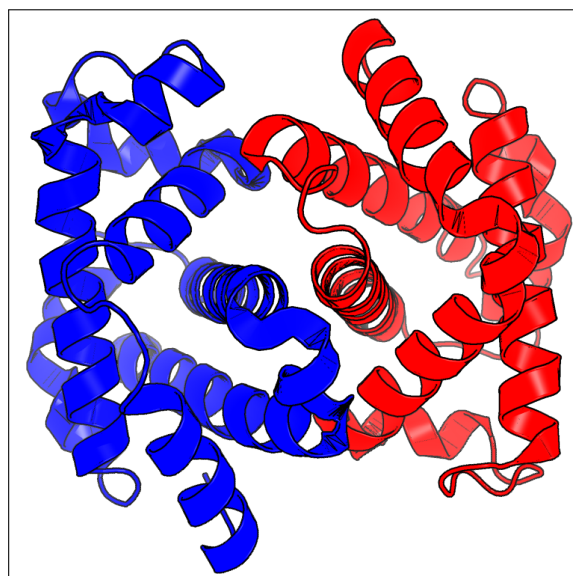


Figure 2: Human (Homo sapiens) Hemoglobin. PDB ID: 2DN3.