Practical 7
22 March 2024
Questions

1. Compute the amino acid composition of the following sequences. Provide the output as a table of amino acid percentage values for each sequence and comment on the results.

    1. RATPTRWPVGCFNRPWTKWSYDEALDGIKAAGYAWTGLLTASKPSLHHATATPEYLAALKQKSRHAA
    2. AAAVMMGLAAIGAAIGIGILGGKFLEGAARQPDLIPLLRTQFFIVMGLVDAIPMIAVGLGLYVMFAVA
    3. AADVSAAVGATGQSGMTYRLGLSWDWDKSWWQTSTGRLTGYWDAGYTYWEGGDEGAGKHSLSFAPVFVYEFAGDSIKPFIEAGIGVAAFSGTRVGDQNLGSSLNFEDRIGAGLKFANGQSVGVRAIHYSNAGLKQPNDGIESYSLFYKIPI

2. Assume the molecular weights of the 20 amino acid residues as given below. Compute the molecular weight of the three sequences given in question 1.

Ala: 85	Cys: 115	Asp: 130	Glu: 145	Phe: 160	Gly: 70	Trp: 200
His: 150	Ile: 125	Lys: 145	Leu: 125	Met: 143	Asn: 130	Tyr: 175
Pro: 110	Gln: 140	Arg: 170	Ser: 100	Thr: 115	Val: 110	

3. The amino acid composition of a standard set of Group A (first value) and Group B (second value) proteins are given below. Identify whether the given sequences in Question 1 belong to Group A or Group B and write your answer.

Ala: 8.47, 8.95	Asp: 5.97, 5.91	Cys: 1.39, 0.47	Glu: 6.32, 4.78	Thr: 5.79, 6.54
Phe: 3.91, 3.68	Gly: 7.82, 8.54	His: 2.26, 1.25	Ile: 5.71, 4.77	Val: 7.02, 6.76
Lys: 5.76, 4.93	Leu: 8.48, 8.78	Met: 2.21, 1.56	Asn: 4.54, 5.74	Trp: 1.44, 1.24
Pro: 4.63, 3.74	Gln: 3.82, 4.75	Arg: 4.93, 5.24	Ser: 5.94, 8.05	Tyr: 3.58, 4.13

4. Compute the residue pair preference for the three sequences given in question 1. The required output is a 20x20 table showing the pair preferences (a) $[N_{ij}*100/(N_i+N_j)]$, (b) $[N_{ij}*100/(N-1)]$ and (c) $[N_{ij}*100/(N_{i*}N_j)]$. List the top 10 preferred residues from each of the three pair-preferences.

5. Compute average hydrophobicity (H_{gm}), Helical contact area (Ca) and Total non-bonded energy (Et) for the sequences in Q1 and comment on the results. (Refer www.iitm.ac.in/bioinfo/fold_rate/prop_orig.html for the properties).

Deadline: 30th March 2024