Indraprastha Institute of Information Technology Delhi (IIITD) Department of Computational Biotechnology

BIO213 – Introduction to Quantitative Biology

ASSIGNMENT-2 (April 13, 2022)

Instructions:

- 1. You are required to submit the assignments by next Wednesday, 20 April 2022.
- 2. Use Python to write your program.
- 3. Use your Roll no. and name to save the file.
- 4. You have to submit well commented scripts, and a PDF file with all the properly labelled answers.
- 5. Requests for extension of submission deadline will not be entertained.

Question 1. Write a program to implement Chou and Fasman method of secondary structure prediction (Helix and Beta strand) for the following protein sequence:

SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNPNYEDL LIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQTFSVLA CYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGF

a) Provide a well commented code for the same.

(20 marks)

b) Display the output using the notation – H: Helix, S: Beta strand.

(10 marks)

Answer:

If the output is correct give full 30 marks. If less than 5 residues are marked wrong give 28. If less than 10 and more than 5 residues are wrong give 25. If many errors give 15.

Question 2. Use the given protein.pdb file (structural coordinates for the protein from question 1) to predict the secondary structural elements using the STRIDE webserver (http://webclu.bio.wzw.tum.de/cgi-bin/stride/stridecgi.py). When prompted to choose from the PDB entries, do not use the same and proceed with the secondary structure prediction with the

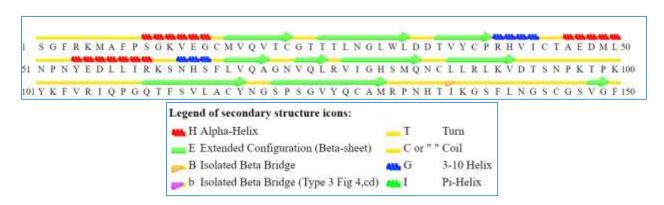
server.

a) Display the output and compare with the result obtained in question 1. Mention the regions differing in both in terms of assigned secondary structure. (10 marks)

Answer:

56	SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDML	1	
	TTTT HHHHHH EEEEEETTEEEEEEETTEEEEEGGGG HHHHH		
100		51	
100	HHHHHHH GGG EEEETTEEE EEEEEETTEEEEEE TTTT		
	98 98 989 89 89		
150	YKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGF	101	

Or



If this output has been shown and a comparison of residues is there, give 10 marks. If only output has been shown, give 6 marks.

b) Comment on the differences observed with proper reasoning. (10 marks)

Answer: Give marks if they have talked about the use of structural information (like hydrogen bonds, torsional angles) in Stride, and that there are many different types of secondary structural elements and not only helices, strands and turns in general.

Chou and Fasman parameters to be used for the prediction are as following:

Residue	Ρα	Residue	Рβ
Glu	1.53	Met	1.67
Ala	1.45	Val	1.65
Leu	1.34	lle	1.60
His	1.24	Cys	1.30
Met	1.20	Tyr	1.29
Gln	1.17	Phe	1.28
Trp	1.14	Gln	1.23
Val	1.14	Leu	1.22
Phe	1.12	Thr	1.20
Lys	1.07	Trp	1.19
lle	1.00	Ala	0.97
Asp	0.98	Arg	0.90
Thr	0.82	Gly	0.81
Ser	0.79	Asp	0.80
Arg	0.79	Lys	0.74
Cys	0.77	Ser	0.72
Asn	0.73	His	0.71
Tyr	0.61	Asn	0.65
Pro	0.59	Pro	0.62
Gly	0.53	Glu	0.26