Ref. No. : Ex/PG/BME/T/1210B/2018

MASTER OF BIO-MEDICAL ENGINEERING FIRST YEAR SECOND SEMESTER - 2018

Subject: BIO-INFORMATICS Time: Full Marks : 50

Part I

Instructions: Answer separate answer-script for each Part.

1. Answer all questions: 1 X 10=10

i. The first bioinformatics database was created by
a) Richard Durbin b) Dayhoff
c) Steven Altschul d) David Lipman.

ii. Which of the following database can be used to access protein domain information?
a) Prosite b) DDBJ
c) SANGER d) KEGG.

iii. If you want literature information, what is the best website to visit?
a) OMIM b) Entrez
c) PubMed d) PROSITE.

iv. Which of the following is wrong about European Molecular Biology Laboratory Data Library Format?
a) EMBL maintains DNA and protein sequence databases
b) As with GenBank entries, a large amount of information describing each sequence entry is given
c) Sequence entry includes literature references and information about the function of the sequence, but not locations of mRNAs and coding regions
d) Information is organized into fields, each with an identifier, shown as the first text on each line

v. The format of an entry in the SwissProt protein sequence database is very similar to the EMBL format.
a) True
b) False

vi. According to standard amino acid code letters which of the given pair is not right?
a) K- lysine
b) Y- tyrosine
c) Q- glutamine
d) R- serine

vii. For computer analysis of proteins, it is more convenient to use single-letter than three letter amino acid codes.
a) True
b) False

viii. Investigators are encouraged to submit their newly obtained sequences directly to a member of the International Nucleotide Sequence Database Collaboration, such as the NCBI, DDBJ, and EMBL.
a) True
b) False

ix. What is the source of protein structures in SCOP and CATH?
a) Uniprot
b) Protein Data Bank
c) Ensemble
d) InterPro

x. Which of the following is untrue about homology modeling?

a) Homology modeling predicts protein structures based on sequence homology with known structures

b) It is also known as comparative modeling

c) The principle behind it is that if two proteins share a high enough sequence similarity, they are likely to have very similar three-dimensional structures

d) It doesn't involve the evolutionary distances anywhere

2. Answer any ten questions 2 X10 =20

a) What are the major features of Uniprot?
b) What is the rationale behind protein structure prediction?
c) Define pharmacophore.
d) What are the major strategies for target identification?
e) Name any five docking programs.
f) What do you mean by double blind approach in CASP?
g) What is high throughput screening?
h) Name the four classification systems in CATH.
i) Define QSAR.
j) What is secondary database? Give an example.
k) Define bioinformatics.
l) What is Ramachandran plot ?

3. Answer any four questions: 5 X 4=20

a) What is virtual screening? How does it differ from high throughput screening?
b) Discuss the various structure based drug design approaches.
c) What are the major interatomic forces that determine protein structure?
d) Discuss the major features of Protein Data Bank.
e) Explain the Chou-Fasman method for protein secondary structure prediction.
f) Outline the levels of organization of protein structure.

MASTER OF BIO-MEDICAL ENGINEERING EXAMINATION, 2018

FIRST YEAR SECOND SEMESTER

BIO-INFORMATICS

Time : Three hours Full Marks : 100

Part – II **Full Marks – 50**

Instructions: Use separate Answer scripts for each Part.

Answer any *two* questions in Part-II.

1. What are homologs and orthologs? Why local similarity is preferred over global similarity between any two protein strings? Explain how dot plot is useful for finding global and local similarities. 8+5+12 = 25

2. Compare PAM and BLOSSUM matrices. Given one sequence with 6R and 1M amino acids. Considering all RR, RM, MR and MM pairs, compute the expected probabilities E_{RR} and E_{RM}. Write a short note on Hidden Markov Model. 8+ 10+ 7=25

3. Describe Needleman-Wunsch algorithm. Write the table for two sequences – TACTTA and ATCT following Needleman-Wunsch algorithm and show one alignment with net score for these sequences. Describe the difference between Needleman-Wunsch and Smith-Waterman approaches. 8+12+5=25

— o —