

**COMPUTATIONAL BIOLOGY**  
**(BIOT 4221)**

**Time Allotted : 3 hrs**

**Full Marks : 70**

*Figures out of the right margin indicate full marks.*

*Candidates are required to answer Group A and  
any 5 (five) from Group B to E, taking at least one from each group.*

*Candidates are required to give answer in their own words as far as practicable.*

**Group – A**  
**(Multiple Choice Type Questions)**

1. Choose the correct alternative for the following: **10 × 1 = 10**
- (i) Number of amino acid residues in a protein of molecular weight 58300 is  
(a) 110 (b) 253  
(c) 530 (d) 330.
  - (ii) The first secondary database developed was  
(a) PRINT (b) PROSITE  
(c) PDB (d) PIR.
  - (iii) Henderson-Hasselbalch Equation is  
(a)  $pK_a = pH + \log_{10} ([HA]/[A^-])$  (b)  $pH = pK_a + \log_{10} ([HA]/[A^-])$   
(c)  $pK_a = pH + \log_{10} ([A^-]/[HA])$  (d)  $pH = pK_a + \log_{10} ([A^-]/[HA])$ .
  - (iv) In Computational Biology Margaret Dayhoff developed first database named  
(a) SWISSPROT  
(b) PDB  
(c) Atlas of protein sequence and structure  
(d) Protein sequence databank.
  - (v) Which variable can give the concrete form to the representation of the transition model?  
(a) Single variable  
(b) Discrete state variable  
(c) Random variable  
(d) Both Single and Discrete state variable.
  - (vi) Proteins often have regions that show specific, coherent patterns of folding or function. These regions are called  
(a) Domains (b) Oligomers  
(c) Peptides (d) Sites.

- (vii) Amongst which of the following is / are the application areas of Python programming?  
(a) Web Development  
(b) Game Development  
(c) Artificial Intelligence and Machine Learning  
(d) All of the mentioned above.
- (viii) Where does the Hidden Markov Model is used?  
(a) Speech recognition  
(b) Understanding of real world  
(c) Both Speech recognition & Understanding of real world  
(d) None of the mentioned.
- (ix) \_\_\_\_\_ is the process of recognizing patterns by using machine learning algorithm.  
(a) Processed Data  
(b) Literate Statistical Programming  
(c) Pattern Recognition  
(d) Likelihood
- (x) TrEMBL is  
(a) primary database  
(b) nucleotide sequence database  
(c) based on architectural classification of protein  
(d) is a computer annotated supplement of SWISS-PROT containing all the translational of EMBL.

### **Group - B**

2. (a) Comment on  $\alpha$ -helix of a protein. *[[CO1](Comment/IOCQ)]*  
(b) Enumerate the mechanism of DNA translation. *[[CO2](Enumerate/HOCQ)]*  
**5 + 7 = 12**
3. (a) Calculate the ATP produced per mole of glucose consumption. *[[CO1](Calculate/IOCQ)]*  
(b) Classify the polysaccharides with examples. *[[CO1](Classify/IOCQ)]*  
(c) State the functions of lipids. *[[CO1](Remember/LOCQ)]*  
**5 + 4 + 3 = 12**

### **Group - C**

4. (a) Define primary database and secondary database. *[[CO3](Understand/LOCQ)]*  
(b) Discuss the two main functions of biological database? *[[CO3](Understand-knowledge/LOCQ)]*  
(c) What type of data are available in biological databases? Write down the names of two primary sequence databases and secondary databases. *[[CO4](Understand-apply/IOCQ)]*  
**(2 + 2) + 2 + 2 + (2 + 2) = 12**

5. (a) Mention based on availability how will you characterise biological databases. [[CO3](Understand/IOCQ)]  
(b) Describe two characteristics of the following: SWISSPROT, PIR. [[CO4](Understand/LOCQ)]  
(c) Write briefly about super secondary structures of proteins. [[CO1](Analyze/IOCQ)]  
**3 + (3 + 3) + 3 = 12**

### Group - D

6. (a) Mention the characteristics of Python. [[CO6](Understand/LOCQ)]  
(b) Mention how Python can be used to for parsing of biological data. [[CO6](Analyze/IOCQ)]  
(c) Mention with example how representation of sequences can be done using Biopython. [[CO6](Remember/LOCQ)]  
**3 + 5 + 4 = 12**
7. (a) What is Hidden Markov Model? [[CO5](Remember/LOCQ)]  
(b) Discuss the application of HMM in computational biology. [[CO5](Remember/LOCQ)]  
(c) What are the advantages and disadvantages of using Artificial Neural Network in computational biology? Mention one important application of ANN in computational biology. [[CO5](Analyze/IOCQ)]  
**2 + 4 + (4 + 2) = 12**

### Group – E

8. (a) Describe the basic steps of drug discovery. [[CO6](Describe/HOCQ)]  
(b) Mention the trials which are needed to be followed in drug discovery process. [[CO6](Examine/IOCQ)]  
(c) Briefly describe any one method used in computational drug design. [[CO6](Analyze/IOCQ)]  
**5 + 3 + 4 = 12**
9. Some mathematical models are used for studying bacterial growth kinetics. These models are divided into structured, unstructured and segregated models. Mention the characteristics of one each microbial growth kinetics model based on the above mentioned models. [[CO6](Remember-evaluate/HOCQ)]  
**(3 × 4) = 12**

---

| <i>Cognition Level</i>         | <i>LOCQ</i> | <i>IOCQ</i> | <i>HOCQ</i> |
|--------------------------------|-------------|-------------|-------------|
| <i>Percentage distribution</i> | 31.2        | 43.8        | 25          |

**Course Outcome (CO):**

After completing the course, the students will be able to:

1. Acquire basic understanding of structures and functions of different biomolecules.
2. Obtain knowledge about the different metabolic pathways.
3. Explain different biological data and biological databases.
4. Understand classification of databases and how the biological data are stored in those databases.
5. Obtain the know ledge of different algorithms and programming languages to manage biological data.
6. Apply different tools and software for analysis of biological data.

*\*LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question.*