

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)**

CLASS: BTECH/BSc.
BRANCH: BIOTECHNOLOGY/CHEMISTRY

SEMESTER : III/ADD
SESSION : MO/2025

SUBJECT: BE24244/BE205 BASICS OF BIOINFORMATICS

TIME: 3 Hours

FULL MARKS: 50

INSTRUCTIONS:

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
 2. Attempt all questions.
 3. The missing data, if any, may be assumed suitably.
 4. Before attempting the question paper, be sure that you have got the correct question paper.
 5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
-

- | | | CO | BL |
|---|-----|----|-----|
| Q.1(a) Select the most appropriate options for the paragraph given below from the list of databases or related terms.
The _____ (1) database provides comprehensive information on protein sequences and their functions, whereas _____ (2) serves as a major repository for storing nucleotide sequences. Experimental gene expression data can be obtained from _____ (3), while protein structural classifications are organized in _____ (4). Data related to biochemical reactions and _____ (5) can be explored through KEGG.
Options: [GenBank, UniProt, PDB, GEO, SCOP, KEGG, nucleotide, protein, metabolic pathway] | [5] | 1 | 1 |
| Q.1(b) List and explain five important properties or features that a good biological database should possess. Discuss the significance of each property. | [5] | 1 | 2 |
| Q.2(a) Explain the basic working principle of the BLAST algorithm. Discuss how the choice of word size and threshold values affects the speed, significance, and sensitivity of the search results. | [5] | 2 | 2,4 |
| Q.2(b) Perform the pairwise alignment using the smith-waterman algorithm. Use the following scoring scheme: (Match +10, Mismatch -5, Gap Penalty -7). Also, report the final optimal alignment(s).
Seq-1: AGCCTAG ; Seq-2: CTCGTC | [5] | 2 | 3 |
| Q.3(a) Explain the following terms in the context of a phylogenetic tree: taxon, internal node, root node, clade, and branch. Describe the significance of each term, and illustrate them clearly using a hypothetical phylogenetic tree diagram. | [5] | 3 | 1 |
| Q.3(b) For the sequences given below, identify the most likely phylogenetic tree using the Maximum Parsimony Method. Also, demonstrate by considering one non-informative site why it does not affect the topology of the most likely tree. | [5] | 3 | 3 |

Sequence	1	2	3	4	5	6	7	8	9
1	A	A	G	A	G	T	G	C	A
2	A	G	C	T	G	T	G	C	G
3	A	G	A	T	A	T	C	C	A
4	A	G	A	A	A	T	C	C	G

- | | | | |
|---|-----|---|-----|
| Q.4(a) Explain the basic steps involved in homology modeling of a protein structure. Enlist at least three important factors to identify a suitable template when multiple templates are obtained after a BLAST search. | [5] | 4 | 2 |
| Q.4(b) Explain the significance of the Ramachandran plot in protein structure analysis. Define the backbone dihedral angles ϕ (phi) and ψ (psi) with the help of an illustration. Identify and label the regions corresponding to α -helices, β -sheets, and left-handed helices on the Ramachandran plot. | [5] | 4 | 1,2 |

PTO

- Q.5(a) Discuss how QSAR modeling and molecular docking complement each other in virtual screening. How can integrating these approaches improve the prediction of ligand binding affinity and biological activity? [5] 5 4
- Q.5(b) Describe the principles of Reverse Vaccinology and explain how immunoinformatics databases aid in identifying potential vaccine targets. Provide examples of such databases and their typical applications. [5] 5 3

.....25/11/2025.....E