Name : ....

Roll No. : ....

Invigilator's Signature : ....

CS/M.Tech(MBT/PHMB/PHMC)/SEM-2/ MBT/PHMB/PHMC-205/2011

2011

APPLIED BIOINFORMATICS

*Time Allotted* : 3 Hours   *Full Marks* : 70

The figures in the margin indicate full marks.

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

Answer Question No. 1 and any *six* from the rest.

1. Fill in the blanks :

    i)   Low complexity regions of the query sequence are filtered by .... for nucleic acid sequences and by .... for amino acid sequences.   2

    ii) .... is used to identify distantly related sequences through iterative search.   1

    iii) .... is used to identify sequence patterns in protein sequences.   1

    iv) In .... file format the header file starts with a ">" sign.   1

30467 (M.Tech.)   [Turn over
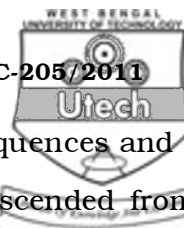
v)   .... may be understood as sequences and the organism in which they occur have descended from a common ancestor.   1

vi) .... algorithm is for local alignment while .... is for global alignment.   2

vii) When two different genes have undergone convergent evolution and given rise to proteins that have a similar active site (but with a different backbone, say), they are called   1

viii) When there is a transfer of genetic material between species that are separated through a long evolutionary distances (as in case of horizontal gene transfer), it is called   1

2.   Write on the following :   5 ∞ 2

a)   Dot plot

b)   Ramachandran plot

c)   Affine gap penalty

d)   PHI-BLAST

e)   PSSM.
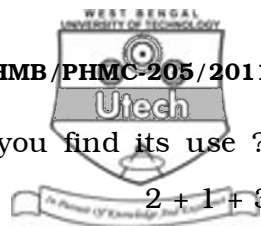
3. a) What is lods ratio ? Where would you find its use ? How would you justify its use there ?   2 + 1 + 3

   b) State how the GOR method differs from the Chou-Fasman method.   4

4. Name the different ways of predicting tertiary structure of a protein. Which one out of these requires structural knowledge of a related protein ? Describe schematically the steps of the method. Is energy minimization of the model thus obtained, a necessary step ? Justify your answer.

   3 + 1 + 3 + 3

5. What is molecular docking ? Mention at least two of its applications. What is SASA ? What is its importance in docking ?   2 + 2 + 2 + 4

6. Write down the basic criterion for a database-search algorithm. Name a very popular tool that uses heuristic method for database search. Describe briefly the algorithm followed by the tool. Define E-value for this tool and mention its significance.   1 + 1 + 5 + 3

7. a) Write down the different applications of multiple sequence alignment.   3

   b) Schematically explain the sequential steps of CLUSTALW.   3

   c) Describe the sum of pairs method.   4

8. a) Use UPGMA to draw a phylogenetic tree using the following distance matrix : 6

	Human	Chimpanzee	Gorilla	Orang-utan	Gibbon
Human	—	0·015	0·045	0·143	0·198
Chimpanzee	—	—	0·030	0·126	0·179
Gorilla	—	—	—	0·092	0·179
Orang-utan	—	—	—	—	0·179
Gibbon	—	—	—	—	—

b) How is this method different from that of the Neighbour-joining method ? 2

c) Write in brief how a phylogenetic tree using the maximum parsimony method, is drawn. 2

———————