INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

END SEMESTER EXAMINATION- BT 305: Computational Biology

May 4, 2022. Maximum Marks: 100. Time: 3 hours

All questions carry 10 marks each; this question paper has 2 pages, 10 questions

1. In a Monte carlo simulation, energy is calculated at each step and compared with the previous step. What is the probability of accepting the step? How are the difference in energy between subsequent steps and temperature affects the likelihood of acceptance?

2. In MD simulations, we allow the system to evolve in time indefinitely, that system will eventually pass through all possible states with a hope of generating enough representative conformations. What is the statistical relevance of this exercise? What do we achieve from this?

3. Explain DOCK algorithm.

4. The SMILES representation of four molecules are as follows. Attempt to draw an approximate 2D structure of these molecules.

```
CCO
CC(=O)O
C1CCCCC1
c1cnccc1
```

5. Bit strings for query and test molecules in a typical fingerprinting exercise is given below.

Mol 1	0	1	0	0	0	1	1	1	0
Mol 2	1	1	0	1	1	0	0	0	0
Mol 3	1	1	1	1	1	1	1	1	1
Mol 4	1	1	1	1	1	1	1	0	0

Calculate Tanimato coefficients for structure comparison and attempt the following question:

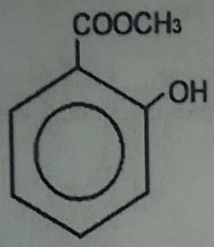
If receptors of the four candidate structures Mol 1 Mol 2 Mol 3 and Mol 4 are respectively Rec 1, Rec 2, Rec 3 and Rec 4. Which molecule is more likely to bind with maximum number of Receptors? Rationalize your answer.

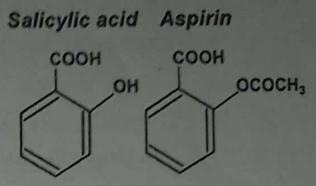
6. Show how Euclidean and city block distances vary while measuring dissimilarity between two individual patterns.

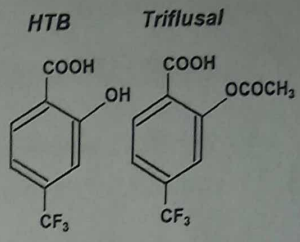
7. Consider the following situation that I have five amino acids A F E G K and D. I have arranged these amino acids in a peptide sequence AFEGKD. i) What is the probability of getting the same sequence, if I repeat this exercise? ii) Suppose, this exercise is manipulated such that polar residues are accurately repeated at their respective positions, calculate log odds ratio between exercise i and exercise ii.

8.

A) Suppose we are performing an experiment of structure based molecular profiling for methyl salicylate (below structure) in a structure pool of four members on the right hand side, how do we determine the closest structure by calculating Tanimato coefficient?

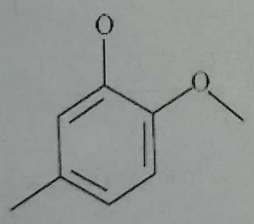
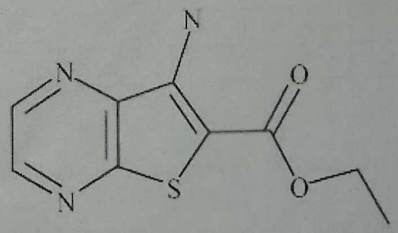






9. Explain briefly the following concepts.
A) Pharmacophore B) Receptor C) Lipinski rule of five D) Molecular descriptor.

10. Write the 1D string representation (SMILES) for the following molecules.

A)	B)
	

C) Attempt to draw a Reduced Graph representation of structures A and B, with nodes corresponding to aromatic rings, aliphatic rings, functional groups and linking groups.

END OF QUESTIONS