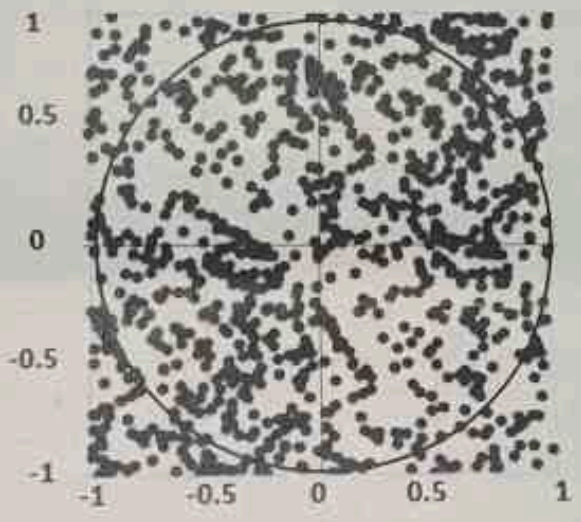
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

MID SEMESTER EXAMINATION - BT 305 Computational Biology

Maximum Marks: 80. Time: 2 hours

All questions (1-8) has 10 marks each

Q1. The dots in the following figure are a result of a series of attempts to throw a dart on to a square board with an imaginary circle of diameter 2 units. Explain the principle of convergence by calculating the value of Pi (π). Write an algorithm to simulate random (x, y) points in a 2-D plane (like the below one) centered on (0,0).

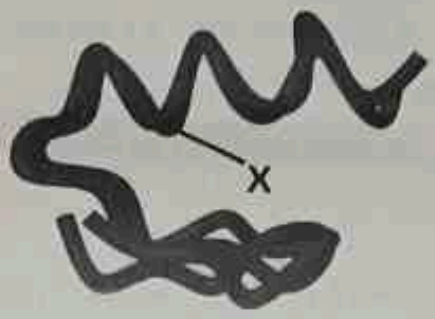


Q2. Define forcefield. What are the components of a molecular forcefield used for protein simulations.

Q3. What are integration algorithms. The Verlet algorithm uses no explicit velocities. Prove this statement from first principles.

Q4. In MD simulations, we allow the system to evolve in time indefinitely, such that, the 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? How does it conceptually mimic a lab experiment, in light of ergodic hypothesis?

Q5. In the following protein backbone template, a rotamer is chosen at the site X in the core of the protein.



A. Suppose the residue X is tryptophan, how do you estimate the solvation energy as a result of choosing tryptophan at position X ?

B. Three amino acid residues (W, F and L) are competing for the position X. The energetic advantage, ΔE of the three rotamers of each amino acids as a result of chi1 rotation at position X is in the following table. With respect to the trans rotamer of phenylalanine, which all rotamers can be eliminated by adopting simple DEE (Dead end elimination algorithm) and Goldstein DEE.

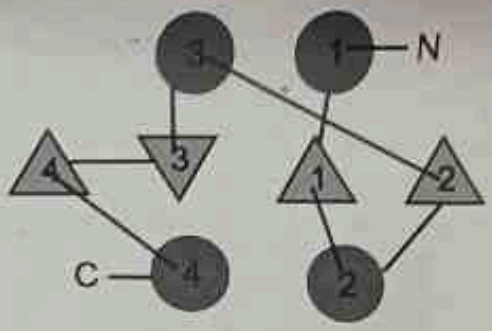
Amino acid	g+	g-	trans
Tryptophan	-10 kcal/mol	-20 kcal/mol	-16 kcal/mol
Phenylalanine	-16 kcal/mol	-17 kcal/mol	-18 kcal/mol
Leucine	-15 kcal/mol	-7 kcal/mol	-12 kcal/mol

Q6. Explain the following concepts with necessary illustrations:

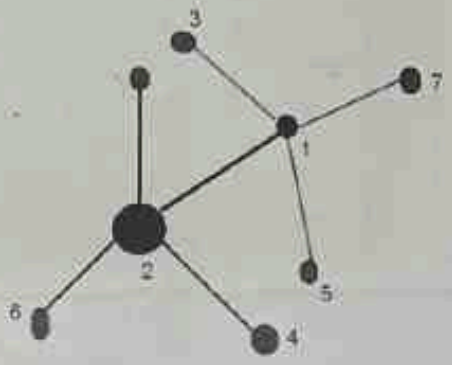
1. Ensemble
2. Contact map
3. United atom approximation
4. Solvent accessible surface area
5. Hydrogen bond

Q7.

A) Sketch an approximate structure from the TOPS diagram shown below.



B) Assuming bigger atom as carbons and smaller ones hydrogen in this ethane molecule, write an approximate internal coordinate representation from the following figure.



Q 8. Write your name and surname in the format of a continuous string; for e.g. SACHIN TENDULKAR as SACHINTENDULKAR. Omit letters that do not represent amino acids. Assume your name sequence as the sequence of a hypothetical template (structure) of a protein.

Which of the following sequence can best substitute your name sequence, show alignment with decreasing order of similarity.

1. ROHITSHARMA
2. SMRITIMANDANA
3. SURYAKUMARYADAV

You may use the following substitution matrix for your calculations. What changes can you expect if you reverse the sign of the entries in this substitution matrix.

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4	-1	-2	-2	0	-1	-1	0	-2	-1	-1	-1	-1	-2	-1	1	0	-3	-2	0
R	-1	5	0	-2	-3	1	0	-2	0	-3	-2	2	-1	-3	-2	-1	-1	-3	-2	-3
N	-2	0	6	1	-3	0	0	0	1	-3	-3	0	-2	-3	-2	1	0	-4	-2	-3
D	-2	-2	1	6	-3	0	2	-1	-1	-3	-4	-1	-3	-3	-1	0	-1	-4	-3	-3
C	0	-3	-3	-3	9	-3	-4	-3	-3	-1	-1	-3	-1	-2	-3	-1	-1	-2	-2	-1
Q	-1	1	0	0	-3	5	2	-2	0	-3	-2	1	0	-3	-1	0	-1	-2	-1	-2
E	-1	0	0	2	-4	2	5	-2	0	-3	-3	1	-2	-3	-1	0	-1	-3	-2	-2
G	0	-2	0	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-2	-3	-3
H	-2	0	1	-1	-3	0	0	-2	8	-3	-3	-1	-2	-1	-2	-1	-2	-2	2	-3
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4	-2	2	0	-3	-2	-1	-2	-1	1
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5	-1	-3	-1	0	-1	-3	-2	-2
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5	0	-2	-1	-1	-1	-1	1
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6	-4	-2	-2	1	3	-1
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7	-1	-1	-4	-3	-2
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4	1	-3	-2	-2
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5	-2	-2	0
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	2	-3
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	-1
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

---- END OF QUESTIONS ----