INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

MID SEMESTER EXAMINATION - BT 305: Computational Biology

Maximum Marks: 40. Time: 2 hours

Questions 1-6 has 5 marks each and question 7 has 10 marks

Q 1. Examine the structure of methane molecule given in the following diagram (Figure 1A). Assume that you define force field with a potential energy (u) Vs angle (ɵ). Draw and spot the position of Potential Energy in the graph (Figure 1B) in your answer book after giving appropriate scales for x and y axis as per your calculations. Attempt it for all possible angle combinations, while there is a bend of 3^{o} and 2^{o} in angle between **H_1, C, H_3** and **H_2, C, H_3** of methane molecule respectively.

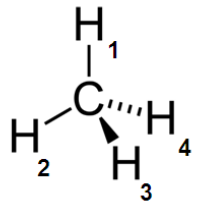


Figure 1A

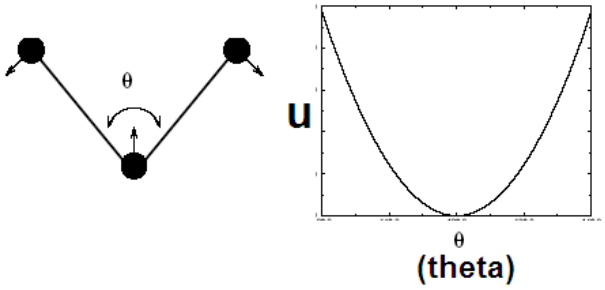
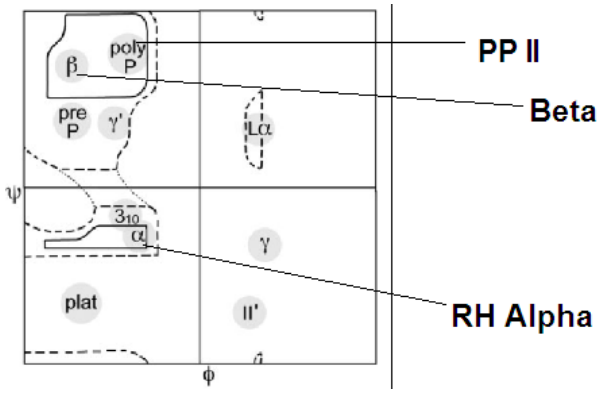


Figure 1B

Q 2. In the following Ramachandran map, β (beta), Polyproline (PP II) and α (alpha) helical region are shown. Write approximate dihedral angles for a nine residue poly-glycine peptide to get the structure of A) Right handed alpha helix; B) Beta sheet; C) Beta strand and D) Random coil.



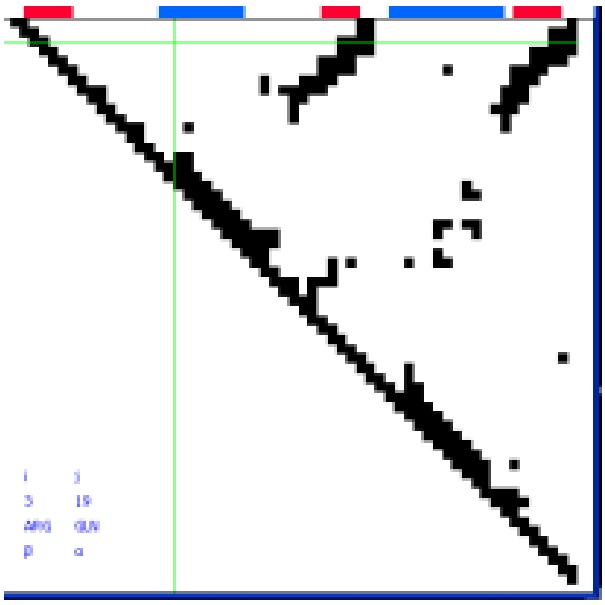
Q 3. Write your name and surname in the format of a continuous string; for e.g. SACHIN KUMAR as SACHINKUMAR. Omit letters that do not represent amino acids. Assuming this as a protein sequence, rationally design another sequence with the same length (and score them), that can likely to have same function as that of the first (your name) sequence. Please note that both sequences should not have any sequence identity. You may use the following substitution matrix for your calculations.

	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

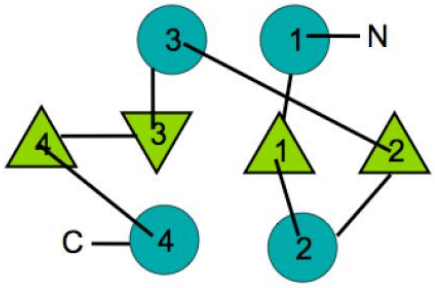
Q 4. What is relevance of united atom approximation while defining a force field? What could be the possible consequence if we switch off improper dihedral angle term while we perform MD simulations of an alanine molecule in water at 298 K.

Q 5.

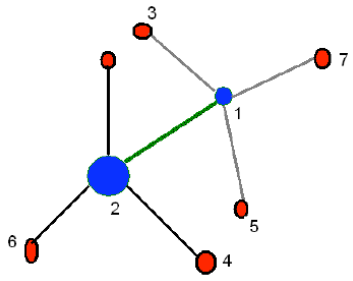
A) Contact map of a protein molecule is shown in the following figure. Draw the same diagram in your answer sheet and identify secondary structure segments.



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



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



Q6. 'The Verlet algorithm uses positions and accelerations at time t and the positions from time t-dt to calculate new positions at time t+dt'. Prove this statement from first principles.

Q7. Protein data bank file of a typical structural segment is shown below.

```
ATOM   186  N    ALA A  26     8.203  11.601  26.881  1.00 14.97           N
ATOM   187  CA   ALA A  26     8.146  12.442  25.683  1.00 13.53           C
ATOM   188  C    ALA A  26     7.714  11.629  24.416  1.00 14.75           C
ATOM   189  O    ALA A  26     6.938  12.111  23.579  1.00 13.31           O
ATOM   190  CB   ALA A  26     9.461  13.156  25.466  1.00 14.89           C
ATOM   191  N    GLU A  27     8.154  10.378  24.341  1.00 15.13           N
ATOM   192  CA   GLU A  27     7.820   9.450  23.250  1.00 17.82           C
ATOM   193  C    GLU A  27     6.311   9.168  23.271  1.00 16.72           C
ATOM   194  O    GLU A  27     5.674   9.075  22.251  1.00 16.01           O
ATOM   195  CB   GLU A  27     8.568   8.142  23.436  1.00 20.13           C
ATOM   196  CG   GLU A  27     8.520   7.306  22.200  1.00 24.41           C
ATOM   197  CD   GLU A  27     8.632   5.841  22.496  1.00 27.22           C
ATOM   198  OE1  GLU A  27     9.616   5.404  23.084  1.00 28.16           O
ATOM   199  OE2  GLU A  27     7.736   5.086  22.132  1.00 29.79           O
ATOM   200  N    LYS A  28     5.737   9.141  24.464  1.00 17.07           N
ATOM   201  CA   LYS A  28     4.309   8.916  24.682  1.00 16.77           C
ATOM   202  C    LYS A  28     3.512  10.055  24.083  1.00 16.33           C
ATOM   203  O    LYS A  28     2.562   9.839  23.323  1.00 15.88           O
ATOM   204  CB   LYS A  28     4.030   8.930  26.195  1.00 19.04           C
ATOM   205  CG   LYS A  28     3.240   7.794  26.717  1.00 22.58           C
ATOM   206  CD   LYS A  28     3.372   7.848  28.264  1.00 24.65           C
ATOM   207  CE   LYS A  28     4.842   7.593  28.786  1.00 24.97           C
ATOM   208  NZ   LYS A  28     4.999   7.896  30.267  1.00 26.72           N
ATOM   209  N    VAL A  29     3.810  11.267  24.514  1.00 14.87           N
ATOM   210  CA   VAL A  29     3.067  12.375  23.977  1.00 15.81           C
ATOM   211  C    VAL A  29     3.321  12.571  22.458  1.00 14.71           C
ATOM   212  O    VAL A  29     2.396  12.980  21.753  1.00 14.07           O
ATOM   213  CB   VAL A  29     3.280  13.702  24.765  1.00 16.97           C
ATOM   214  CG1  VAL A  29     2.657  14.908  24.038  1.00 14.55           C
ATOM   215  CG2  VAL A  29     2.728  13.594  26.066  1.00 17.42           C
ATOM   216  N    PHE A  30     4.545  12.325  21.972  1.00 11.59           N
ATOM   217  CA   PHE A  30     4.843  12.494  20.533  1.00 11.44           C
ATOM   218  C    PHE A  30     4.080  11.466  19.686  1.00 10.70           C
ATOM   219  O    PHE A  30     3.477  11.802  18.675  1.00 11.11           O
ATOM   220  CB   PHE A  30     6.350  12.498  20.265  1.00 10.85           C
ATOM   221  CG   PHE A  30     7.043  13.804  20.666  1.00 10.38           C
ATOM   222  CD1  PHE A  30     6.396  15.038  20.577  1.00  9.64           C
ATOM   223  CD2  PHE A  30     8.380  13.821  21.061  1.00  8.62           C
ATOM   224  CE1  PHE A  30     7.108  16.238  20.873  1.00  9.89           C
ATOM   225  CE2  PHE A  30     9.031  15.017  21.334  1.00  9.58           C
```

ATOM 226 CZ PHE A 30 8.419 16.197 21.243 1.00 6.81 C

A) Calculate Electrostatic interaction energies between all atoms in residue 26 (ala) and residue 30 (phe). Use the following data. (5 marks)

name	type	charge
N	N	-0.280
H	H	0.280
CA	CH2	0.000
C	C	0.380
O	O	-0.380

B) Predict from distance and angle measurements, the possibility of hydrogen bond formation between backbone atoms of residue 26 (ALA) and residue 30 (PHE). Does this give any clue about the possible structure of this protein segment? If yes explain How? (5 marks).

---- END OF QUESTIONS ----