BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI (RAJ.)

FIRST SEMESTER 2010-2011

BIO C417, BIOMOLECULAR MODELING

COMPREHENSIVE EXAMINATION

TOTAL WEITAGE 30% Date: 14.12.2010 DURATION: 3Hrs. (Part A & Part B)

Total Marks (70+30) =100

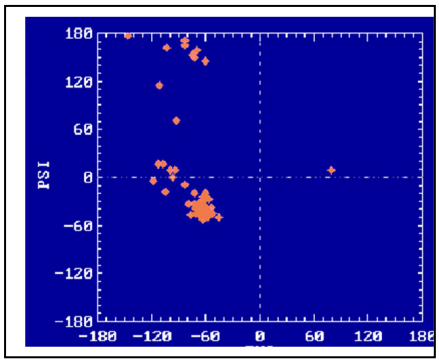
- **Answer Part A and Part B in separate answer sheets.**
- **Irrelevant answer may attract penalty.**

PART – A (CLOSED BOOK) (Max. duration: 2 Hrs., Max. Marks 70)

1. Write notes on **(i)** Collagen Triple helix **(ii)** SIMPLEX method of minimization **(iii)** Weak forces in biomolecules **(iv)** Ramchandran plot **(v)** Grooves of DNA **[5X4=20]**

2. a) What will be the consequences of "D" amino acids in the protein structure? Explain the correlation between side chain orientation (χ) and main chain conformation (ϕ and ψ) of protein with proper example. **[2+2]**

b) What are the utilities of dividing protein structure into different classes? It is claimed that the following Ramachandran plot is for a α-class protein. Is this claim true? Justify your answer. **[2+2]**



c) Why do (0,0) phi (ϕ) and psi (ψ) is disallowed? Explain with the help of figure. **[2]**

3. a) Write down all unique dinuleotide steps and classify them according to their structural feature. Mention all structural features of each class. **[6]**

b) During DNA-protein interaction, which one is which one (between DNA and protein) is responsible for specificity and why? **[2]**

c) Draw a C3'-endo sugar pucker with syn configuration of guanine. **[2]**

4. a) One of the common strategy of comparative modeling is to take multiple templates. What would be the potential problem of choosing multiple templates? **[2]**

b) Among the three possible strategies of Homology modeling, which one is adopted by modeler? Justify your answer. **[2]**

c) How do you integrate Chou-Fasman algorithm into Homology modeling to improve the efficiency of homology modeling? **[3]**

d) What is the crucial step in sequence to structure alignment? Why treading is not as popular as homology modeling? **[3]**

5. a) What are the potential problems in molecular mechanical forcefield? Why do current energy minimization methods cannot find global minima? **[2+2]**

b) Explain critically L-J potential curve of van der Waals interaction and explain why some researcher uses distance based cut-off of van der Waals interaction. **[2+2]**

c) Why it is believed that during simulation of biomolecules, handling non-bonded interaction is more difficult than handling bonded interaction? **[2]**

6. a) Explain Verlet leap-frog algorithm**.** **[3]**

b) Explain the logic behind choosing a time step. **[3]**

c) Explain the advantages and draw backs of choosing implicit solvent model. **[4]**

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI (RAJ.)

FIRST SEMESTER 2010-2011

BIO C417, BIOMOLECULAR MODELING

COMPREHENSIVE EXAMINATION

TOTAL WEITAGE 30% Date: 14.12.2010 DURATION: 3Hrs. (Part A & Part B)

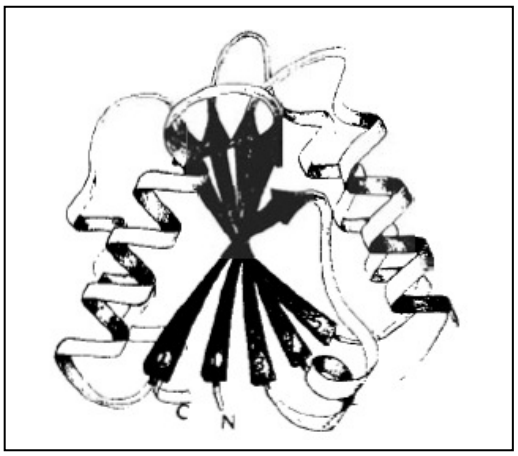
Total Marks (70+30) =100

- **Answer Part A and Part B in separate answer sheets.**
- **Irrelevant answer may attract penalty.**

PART – B (OPEN BOOK) (Max. duration: 1 Hr., Max. Marks 30)

1. a) What type of β turns should be preferred as the direct connection between two antiparallel β strands i.e. in the β hairpin motif. Explain your answer in structural terms. **[3]**

b) Draw TOPS diagram of following protein and identify the protein class it belongs. **[5]**

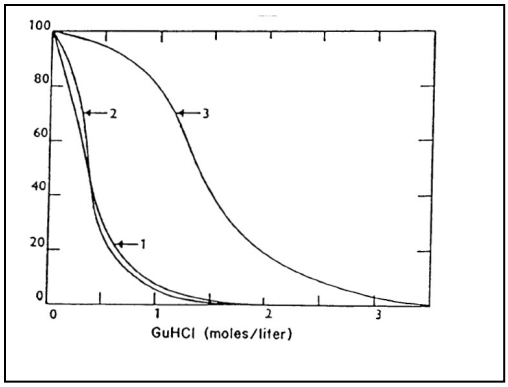


c) Explain why sugar pucker of DNA is C2'endo whereas sugar pucker of RNA is C3'-endo. **[2]**

2. a) The protein penicillinase from *Staphylococcus aureus* can be unfolded by high concentration of standard denaturants. The figure below shows the unfolding of penicillinase as a function of [GuHCl] when monitored by three different methods. Curve 1 is a plot of UV absorbance at 278 nm, near the λ_{max} for aromatic residues. Curve 2 monitors the increase in viscosity of the protein solution as a function of increasing [GuHCl]. Curve 3 is the molar ellipticity at 220 nm as measured by circular dichroism (CD) spectroscopy. All three properties have been normalized to a scale of 0 to 100 for presentation on the same plot.

(i) What does this data imply about the validity, for penicillinase, of the "two-state" assumption often made in the analysis of such data? **[3]**

(ii) Do these unfolding curves provide support for any of the generic model for the protein folding process? **[2]**



[Hints: GuHCl is Guanidinium Hydrochloride, one of the strongest denaturant. CD is used to study the presence of secondary structure in protein and 220 nm ellipticity often indicates the presence of α-helix]

b) Discuss why so called "protein folding problem" is still an open problem. **[5]**

3. Read attached methodology of a research work and explain the purpose of selecting each field of simulation protocol. **[10]**

****

The initial structure for the unfolding simulations was taken from the NMR ensemble (PDB code1FSD). Models 1-10 were chosen for the simulation. The initial structure was solvated using a truncated octahedral box of water molecules represented according to the TIP3P model, ensuring that the edge of the solvent box was at least 9 Å away from the solute. This required a box with sides of length 50 Å and a total of ~11 000 atoms. The system was minimized and equilibrated via a constant pressure, constant-temperature simulation. After the equilibration phase at each trajectory, constant-volume, constant-temperature simulations were performed, and the coordinates were saved every 20 ps. The MD simulations were conducted with the AMBER simulation package, and the protein was represented using the Duan et al. force field (AMBER ff03). Particle Mesh Ewald (PME) was applied to calculate long-range electrostatic interactions; SHAKE was applied to freeze the vibrations of the bonds connecting hydrogen atoms. A 2.0-fs time step was used. The unfolding temperature was set to 500 K, and 10 independent trajectories were run, with each extending to 10.0 ns starting from the native structure. This set of trajectories is labeled UTRAJ. For comparison, 10 trajectories were run at 300 K for 10 ns each starting from the native state, and this set is denoted NTRAJ. Five independent simulations at 300.0 K were performed, with each running to 200.0 ns, to investigate the early folding process. In this set of simulations, the initial structure was the fully extended state. After an initial collapsing process modeled by a short simulation in the Generalized- Born solvent model, the root-mean-square deviation (RMSD) reached ~8 Å. Five extended structures were selected from which the folding simulations continued using the same protocol and solvent models as used in the UTRAJ and NTRAJ sets. This trajectory set is labeled FTRAJ. An initial estimate of the Transition State Ensembles (TSE) was obtained from the unfolding simulations at 500 K by analyses of the free-energy landscape, which allowed identification of an area as defined by the conditions RMSD) 4.0 (0.2 Å and radius of gyration (*R*g)) 9.1 (0.2 Å. There were a total of 42 snapshots in the defined area. Ten conformations were selected from this set of 42 structures for approximately 2 frames per each of the 5 trajectories that were structurally dissimilar from one another by visual inspection. Using these 10 structures as the starting points, 10 different trajectories were run for 10.0 ns at 300 K. This trajectory set is referred to as TSTRAJ. A summary of the simulations is provided in Table 1.

Summary of the Simulation (Table-1)

Set	Starting Point	Temp (K)	Length of Simulation	No. of Independent Trajectories	Description
UTRAJ	Native	500	10	10	Unfolding
NTRAJ	Native	300	10	10	Native
FTRAJ	Unfolding	300	200	5	Early folding
TSTRAJ	Unfolding TS	300	10	10	Folding/unfolding