BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI (RAJ.)
FIRST SEMESTER 2016-2017
BIO F417, BIOMOLECULAR MODELING
COMPREHENSIVE EXAMINATION
TOTAL WEITAGE 35% Date: 01.12.2016 DURATION: 3Hrs. (Part A & Part B)
Total Marks (70+35) =105

- **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)

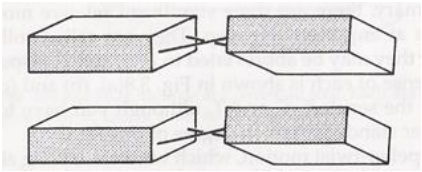
1.a) In nature, it is frequently observed that B-DNA conformation is converted to A-DNA conformation with change of environmental condition. In this context sugar pucker plays an important role. Draw the preferred sugar puckering in B and A-DNA structure respectively. Mention the preferred path through which this conversion takes place. Give the justification. **[1+2+2]**

b) A typical B-DNA structure generally takes up following backbone conformation. With the help of Newman projection diagram show that each of these torsion angles is preferred sterically. **[6]**

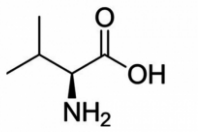
α	β	γ	δ	ε	ζ
-41°	**136**	**38**	**139**	**-133**	**157**

c) Explain why α-helix have particular significance in DNA binding motifs, including helix-turn-helix motifs, leucine zipper motifs and zinc finger motifs. **[2]**

d) Mention approximate value of roll and propeller twist of following DNA dinucleotide step. Justify your answer. **[2]**



2. a) Side chain conformation during model generation is frequently taken from rotamer library. How many possible rotamers exist for following amino acid? Which will be favored among all possible rotamers? Explain with the proper diagram. **[3]**



b) With the help of proper diagram show that a tri-peptide with repetitive ϕ and ψ values of 0° (with trans peptide bond) is highly unfavorable. **[4]**

c) Compactness of protein structure is one of the important parameter to judge protein stability. Give justification. **[2]**

d) With a crude approximation one can estimate that a protein with 300 residues can take up to 10^{300} conformations. How does protine arrive a unique conformation within very short time? **[3]**

e) Why we do not observe any γ-turn without proline residue in second turn position? Explain with proper diagram. **[3]**

3. a) Define the following terms **[1X5=5]**

i) Connectivity matrix **ii)** re-entrant surface **iii)** cooperativity of protein folding **iv)** Molden globule state **v)** Levinthal's paradox

b) In protein databank you came across a molecule with the following atom names. Draw the line diagram of this entity (mention the atom name). Identify this entity. Identify the potential hydrogen bonding sites of this entity which is observed in nature. **[2+1+2]**

P, O1P, O2P, O5', C5', C4', O4', C3', O3', C2', C1', N9, C8, N7, C5, C6, N6, N1, C2, N3, C4

c) Compare knowledge based and *ab initio* technique of protein modeling. **[5]**

4.a) What are the major assumptions of forcefield calculation? Explain. **[3]**

b) Justify why in conjugate gradient method (in comparison with steepest descent method) minima is reached in fewer steps. **[3]**

c) Arrive to the working equation of typical molecular dynamics method from the basic principles. **[4]**

5. Write notes on: **i)** First order minimization methods **ii)** Sequence dependent DNA structure **iii)** Threading technique **iv)** Importance of Glycine in protein structure **v)** Template selection in comparative modeling **[5X3=15]**

****************************Good Luck**

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI (RAJ.)
FIRST SEMESTER 2016-2017
BIO F417, BIOMOLECULAR MODELING
COMPREHENSIVE EXAMINATION
TOTAL WEITAGE 35% Date: 01.12.2016 DURATION: 3Hrs. (Part A & Part B)
Total Marks (70+35) =105

- **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 35)

1. Read the attached discussion section of a review article related to protein folding and answer the following questions (according to the discussion). [**15**]
a) How do proteins fold and why do they fold in that way?
b) What are the features of protein folding that one tries to understand?
c) How does protein reduce conformational search space?

2. a) "Existence of many loops in a modeled protein structure is considered as bad model". Justify the statement **[2]**
b) In a new planet **M**, the life is similar as earth. However, all amino acids in protein are in "D" configuration. What would be Ramachandran plot of planet **M**. **[3]**
c) A majority of promoter region comprised of AT rich sequence. From your knowledge of sequence dependent DNA structure, discuss the advantage and disadvantage of AT rich sequence at promoter region. **[4]**
d) We want to develop a new method for homology modeling with special thrust on correct modeling of loops. Our plan is to generate a loop library and then use that for modeling of loop regions. Write down the basic steps of the algorithm required for generating a loop library. **[4]**

3. a) How does one mimic the experimental condition in a molecular dynamics simulation? **[4]**
b) Explain why cut-off based non-bonded interaction calculation incur larger error in electrostatic energy calculation (with respect to van der Waal's interaction). **[3]**

****************************Good Luck************************************

Discussion

This article considers the fundamental questions of protein folding, previously answered so differently by the classical and new view models. How do proteins fold, and why do they fold in that way? Extensive experience with the folding problem over a 50-y period has shown that clear structural information on the intermediate states that bridge between the unfolded and native states will be required. Experimentation has developed three useful approaches. Folding intermediates can be studied as significantly populated forms during kinetic folding, or as conformationally excited forms present at equilibrium under native conditions, or as equilibrium molten globule forms. Structural results from these different approaches converge on the same conclusions.

The Foldon Hypothesis: In all of these observations, cooperative foldon units play a pivotal role. Foldon units were first discovered and characterized in the initial native state HX experiment. The experiment showed that native Cyt c at equilibrium under native conditions repeatedly unfolds and refolds. A series of experiments showed that the foldon unfolding reactions occur in a sequential pathway-like manner rather than independently. That chain of research was rather complex; it developed over a period of years and has evidently been difficult for most investigators to follow. However, reversible partial unfolding and refolding steps have now been seen in various ways for many proteins, and they have often been connected to the protein folding process. Most pointedly, a recently advanced HX MS capability made it possible to observe matching behavior as it occurs during kinetic folding for MBP, RNase H, and Cytc, as just described. In all cases one sees that unfolding and refolding proceed in steps that subtract or add one more native-like cooperative foldon unit. The detailed foldon construction of Cyt c and RNase H is illustrated in Fig. 6. Both fold by first forming their blue foldon, then an immediately adjacent foldon to form the blue + green PUF, and so on.

The centrally important point is this: contrary to previous belief, proteins are multistate objects built from separately cooperative foldon units. This fundamental insight leads to a foldon-based hypothesis that suggests the “how” and the “why” of protein folding. The cooperative foldon construction of proteins predisposes them to unfold and refold through foldon-determined steps. The discrete steps produce an ordered repeatable macroscopic folding pathway because pre- viously formed foldons tend to guide and stabilize the formation of incoming foldons that they are designed to interact with in the native protein.

Time and Energy: A successful folding model must resolve major questions concerning folding time and energy. Levinthal pointedout that the vast array of protein conformations in unfolded space cannot simply reequilibrate and reach the unique native state by an undirected random search in any reasonable time. Early theoretical work therefore focused on the downhill energetic drive and the many independent routes that heterogeneity and microscopic thermal searching alone seemed to require. The new view answer to the “why” question is that, from the microscopic point of view, there seems to be no other viable choice.

Experimental work recounted here reveals an emergent macroscopic behavior that provides a previously unrecognized mechanism. Random search does not have to carry the protein all of the way to the native state. It only needs to accomplish the formation of a first native-like foldon. This process is thermodynamically downhill and is guidedby the inherent cooperativity of native foldon units. Present information indicates that the

first-formed foldon tends to be stable in the context of the rest of the protein. The still-unfolded regions can shield and energy minimize unfavorably exposed groups, as in the molten globule situation described before.

The time scale for forming a first foldon unit by an unguided search, perhaps two segments ~20 residues in length, is shorter by far than for a reference 100-residue protein [3^{100} /(2×3^{20}) ~ 10^{40}]. The formation of subsequent foldons must proceed by way of similar microscopic searching but in a more guided way analogous to the process of "folding upon binding." The concept that proteins start folding by forming a native-like structural nucleus has been widely accepted. This minimal structure can be sufficient to seed subsequent foldon–foldon interaction steps in a sequence of more guided searches that follow through, rapidly, to the native target.

Does this process have the energetic bias necessary to select specific folding steps and drive folding to completion in a short time? Zwanzig et al. calculated that a free energy bias of 2 kT toward correct interactions is necessary for a folding sequence to complete on a time scale of seconds. It should be appreciated that this degree of bias, more than 1 kcal/mol, is unreasonable at the individual residue level. A single residue has very low probability for finding its correct native partners in a sea of non-native alternatives. Certainly, microscopic thermal searching must underlie any structure formation process. However, given the required energy bias computed by Zwanzig et al., it seems that microscopic-level searching alone cannot swiftly reach the native state.

By contrast, in a more macroscopic foldon-based scenario each correct native-like choice is driven by the collective energy of many interaction sites held stereochemically in a native-like geometry in partner foldons. This mechanism has been described be-fore as sequential stabilization. It is analogous to the well-known folding upon binding process, except that here the incoming disordered segment is advantageously tethered to its already structured partner. The macroscopic foldon-level factors provide both the qualitative structural basis and the quantitative energetic bias required to rapidly and repeatably select discrete determinate pathway steps in competition with all of the other possible alternatives.

****