PROTEOMICS AND PROTEIN ENGINEERING
(BIOT 4121)

Time Allotted : 2½ hrs **Full Marks : 60**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 4 (four) from Group B to E, taking one from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A

1. Answer any twelve: **12 × 1 = 12**

Choose the correct alternative for the following

(i) Which of the following statement holds true for the GAL4 reporter gene in Yeast Two hybrid system?
(a) GAL4 can be fused to the activation domain of a transcription factor
(b) GAL4 can be expressed only if the tested protein interaction occurs
(c) GAL4 fused with DNA binding domain of a transcription factor
(d) GAL4 enquires the presence of Histidine in the growth medium for its expression.

(ii) In order to attain a native 3-dimensional confirmation, polypeptide chains in protein undergo protein folding. Which of the following statements is correct for the process of protein folding?
(a) The protein folding process is governed by an overall decrease in entropy
(b) Folding of proteins favours the position of hydrophilic amino acids towards the outer side
(c) Protein Folding is a nonspontaneous process
(d) Protein folding is majorly driven by various covalent interactions.

(iii) A common experimentation in proteomics for conformational mapping of a protein is
(a) deuterium exchange (b) circular dichroism
(c) protein splicing (d) affinity purification-mass spectrometry.

(iv) A multimeric protein when run on an SDS-PAGE, showed two bands at 20 kd and 40 kd. However, when the protein was run on a native PAGE, it showed a single band at 120 kd. The native form of the protein would be
(a) homotrimer (b) heterotetramer
(c) heterodimer (d) heterotrimer.

(v) Which one of the following activities is not involved in protein folding in the endoplasmic reticulum?
(a) Peptidyl-prolyl-isomerase (b) Protein disulphide isomerase
(c) Protein glycosylation (d) Protein ubiquitination.

(vi) A small fraction of clear cellular lysate was run on an isoelctric focusing gel (IEF) to purify a particular protein, which showed a number of sharp bands corresponding to different pI value. The protein of interest has a pI of 5.2. Therefore, the band corresponding to pI 5.2 was cut, eluted with appropiriate buffer and subjected to SDS-PAGE, which showed 3 distinct bands. One of the following inferences can not be drawn from above observations?
(a) Several different proteins having same pI may be at the single band on IEF gel
(b) SD-PAGE showed 3 distinct bands which may represent molecular mass of a different protein
(c) The protein of interest may be composed of 3 subunits
(d) As the IEF gel showed a distinct band corresponding to pI 5.2, which is the pI of protein of interest, the protein is composed of a single subunit.

(vii) Drug targets with high level of polymorphism within the population are usually unsuitable for which one of the following reasons?
(a) Different variants show similar responses to candidate drugs
(b) Different variants show different responses to candidate drugs
(c) Different genetic and biochemical interaction studies are not required
(d) None of the above.

(viii) A protein undergoes post translational modification. In an experiment to identify the nature of modifications, the following experimental results were obtained.
P. Protein moved more slowly in the SDS-PAGE.
Q. IEF showed that there was no change in the pI.
R. Mass spectrometric analysis showed that the modification was on serine.
The modification that the protein undergoes is likely to be
(a) phosphorylation (b) glycosylation (c) ubiquitinization (d) ADP-ribosylation.

(ix) Which of the following diseases is NOT caused by protein misfolding in humans?
(a) Type 2 Diabetes (b) Cystic fibrosis
(c) Neurofibromatosis type I (d) Parkinson's disease.

(x) Consider a polypeptide with 50 amino acids (corresponding to a very small protein). If there are 5 conformations that each amino acid can adopt, what is total number of conformations for the peptide?
(a) 50^5 (b) 5^{50} (c) 505 (d) 55.

Fill in the blanks with the correct word

(xi) The main application for protein CD spectroscopy is _____.

(xii) Example of homo-oligomeric protein is _____.

(xiii) One *in vivo* technique to study PPI is _____.

(xiv) One technique to study invitro protein folding is _____.

(xv) Inside the cell, formation of disulphide bonds of proteins occurs at _____.

Group - B

2. (a) "Sample preparation is the vital step for successful 2D-GE result", justify the statement. Also, the explain the two important separation steps of 2D-GE with diagram. *[(CO1)(Evaluate-Remember/HOCQ)]*

(b) What is the purpose of PTM in eukaryotic cells? Explain mechanism of the QC for protein product in ER through Calnexin/calreticulin. *[(CO2)(Remember-Understand/IOCQ)]*

(c) Write different types of molecular chaperon present in cytoplasm of cells. Explain the features and functions of all molecular chaperons. *[(CO2)(Remember/LOCQ)]*

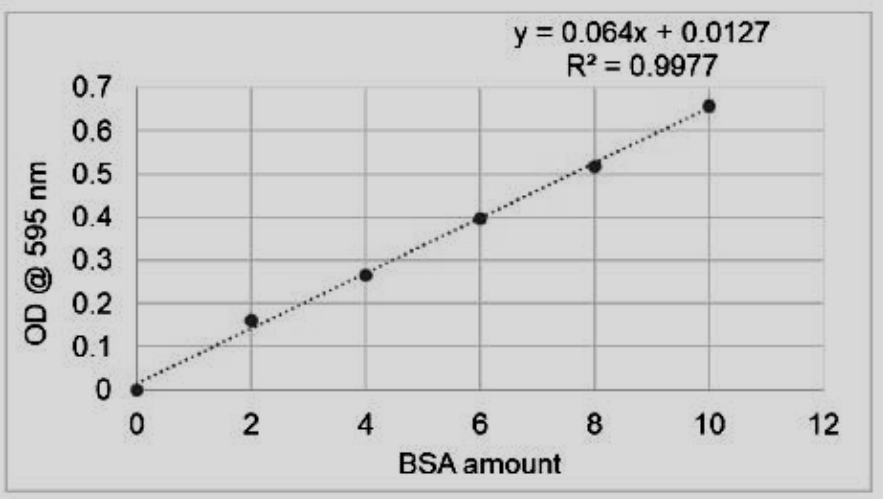
(1 + 3) + (1 + 3) + (1 + 3) = 12

3. (a) Explain how affinity chromatography can used to study protein-protein interaction with a labeled diagram. *[(CO1)(Explain-Remember/HOCQ)]*

(b) Write the names of five common types of HPLC detectors. Explain operation of any two detectors that you have mentioned. *[(CO2)(Remember-Understand/IOCQ)]*

(c) An individual X wants to perform the SDS-PAGE for four protein samples A, B, C and D, for which he first quantified the protein concentration using Bradford method of protein quantification. He had dissolved his protein samples in 50 µl of buffer and for each sample, then 2 µl of each protein sample was taken in duplicates to perform the quantification. Standard protein BSA was used to plot the standard curve taking different amount of BSA as 2 µg, 4 µg, 6 µg, 8 µg and 10 µg. Given below is the standard plot obtained for BSA along with the respective absorbance.

BSA amount	OD1	OD2	Avg OD
0	0	0	0
2	0.163	0.159	0.161
4	0.261	0.271	0.266
6	0.433	0.358	0.396
8	0.511	0.522	0.517
10	0.656	0.657	0.657



The absorbance obtained for samples A to D were 0.505, 0.353, 0.1845 and 0.2125, respectively. If the individual wishes to load 15 µg protein for each sample in SDS-PAGE, what would be the volume required for the samples A and D. *[(CO2) (Analysis/HOCQ)]*

4 + (1 + 4) + 3 = 12

Group - C

4. (a) In x-ray diffraction of proteins, what are the definitions and influence of structure factor and temperature factor on the diffraction pattern? What does high temperature factor for a crystal indicate? *[(CO3)[(understand-analyze-IOCQ)]*

(b) (i) Why has cancer as a disease been a primary target for proteomic analysis? (ii) Use a labelled diagram/flowchart *only* to explain how cancer proteomics has led to the development of novel biomarkers, technologies and diagnostic patterns for treatment. (iii) The protein strathmin has been used as a reliable biomarker for leukemia. What is the distinguishing feature of this protein biomarker? If strathmin were to be developed as target for leukemia, what would be its preferred mode of delivery? Explain your answer. *[(CO6)(Understand-Analyze-IOCQ)]*

(4 + 2) + (1 + 2 + 3) = 12

5. (a) A novel protein is being investigated as a drug target. Determination of an accurate relative molecular mass M_r as part of protein purification is an important parameter to establish in the preclinical phase of drug development. Following are some relevant information and parameters for this process: (i) separation of the novel protein is being done by size exclusion (SE) chromatography on a Sephadex column (ii) the enzymes aldolase, catalase, ferritin, thyroglobulin and blue dextran are being used as standards (iii) M_rS and the retention volumes (V_rS) of the standards are known and (iv) the retention volume of the unknown protein is known. Stepwise explain how you would deduce the relative molecular mass of this unknown protein. Show all calculations. *[(CO6)(calculate-analyze/IOCQ)]*

(b) "In the field of proteomics based drug discovery, the large scale high throughput technologies associated with proteomics have led to increased discovery of potential new targets but has caused a downstream bottle neck at the target validation

stage". Elucidate this statement with a flowchart of the protein based drug discovery process. Explain in quantitative terms why target validation is a low throughput enterprise. Cite four specific methods of proteomics based target validation. Using a specific example explain how protein-protein interactions can be used for target validation.

[(CO3)(Understand-analyze/IOCQ)]

4 + 6 + 2 = 12

Group – D

6. (a) How will you quantitatively parameterize whether proteins change conformation upon protein-ligand or protein-protein complex formation? Briefly define the features and parameters. *[(CO6)(Understand-analyze/IOCQ)]*

   (b) "In protein engineering, manipulation of the gene of interest prior to expression of the protein product is advantageous". Elucidate this statement by citing four specific advantages of this process. How is intracellular degradation of the cloned protein prevented by this process? *[(CO4)(Understand-analyze/IOCQ]*

(3 + 3) + (4 + 2) = 12

7. (a) Biotechnological applications of protein splicing have moved from "proof-of-concept experiments to productive applications. How would the identification of split inteins with large variation tolerances help in such applications?

[(CO4)(Understand-analyze/IOCQ)]

   (b) "Genetically engineered proteins for inorganic materials (GEPIs) are important tools for the self-assembly of molecular systems in nanobio-technology". Using an example, explain how this has been achieved. For what purposes have such GEPIs been used for? Itemize the applications that higher generation function-specific peptides have been utilized for.

[(CO4)(Analyze-HOCQ)]

6 + (3 + 1 + 2) = 12

Group - E

8. (a) How you will determine thermodynamic stability of protein, assuming that *in vitro* protein folding is a reversible reaction between native state (N) and unfolded state (U). *[(CO6)(Evaluate/IOCQ)]*

   (b) For a certain solution of protein RNaseA, in which the total protein concentration is 2.0×10^{-5} M, the concentration of the native and denatured protein at 50 and 100°C are listed below.

Temperature	Protein (denatured)	Protein (native)
50°C	5.1×10^{-6} M	2.0×10^{-3} M
100°C	2.8×10^{-4} M	1.7×10^{-3} M

Determine ΔS^o and ΔH^o for the folding reaction. *[(CO6)(Analyse/HOCQ)]*

   (c) Write names of three experimental techniques for the study of *in vitro* protein folding. Explain the principle and steps to study protein folding through anyone of those techniques you have mentioned. *[(CO3)(Understand/LOCQ)]*

4 + 4 + (1 + 3) = 12

9. (a) Write names three different techniques for the study of protein folding. Describe anyone techniques for the study of protein folding that you have mentioned, with labelled diagram. *[(CO6)(Understand/IOCQ)]*

   (b) (i) Write the four properties by which MG state of proteins can be characterized.

   (ii) Write the steps of misfolded protein degradation by UPS inside the cell cytoplasm, with labelled diagram.

[(CO3)(Understand/LOCQ)]

(1 + 4) + (3 + 4) = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	15.62	56.25	28.13

Course Outcome (CO):

After the completion of the course students will be able to

CO1. Understand different large scale protein separation, estimation, identification and sequencing techniques. Apply the knowledge to solve and analysis of proteome.
CO2. Understand the in vivo and in vitro protein-protein interactions techniques.
CO3. Describe the techniques for structural proteomics and apply knowledge of proteomics in drug discovery.
CO4. Describe the basics and significance of protein engineering; demonstrate the modification and design of protein according to the demand of industry and application.
CO5. Understand the stability of protein structure and mechanism of protein folding; apply this knowledge in study of protein misfolding related diseases.
CO6. Analyze and solve problems related to proteomics and protein engineering technology.

*LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question