VICTORIA JUNIOR COLLEGE
BIOLOGY DEPARTMENT
JC2 PRELIMINARY EXAMINATIONS 2017
Higher 3

CANDIDATE NAME

CLASS | | | | | | INDEX NUMBER | | |

BIOLOGY 9815/01

25 September 2017

Additional Materials: Answer Paper **2 hours 30 minutes**

READ THESE INSTRUCTIONS FIRST

Write your CLASS/ INDEX no. and name on all the work you hand in.
Write in dark blue or blue pen.
You may use a soft pencil for any diagrams, graphs or rough working.
Do not use any staples, paper clips, highlighters, glue or correction fluid.

Section A
Answer **all** questions.

Section B
Answer **three** out of **four** questions.

Section C
Answer the question.

At the end of the examinations,
1. Fasten all your work securely;
2. Circle the number of the section B question you have answered in the grid opposite.

For Examiner's Use	
Section A	
1	
2	
3	
4	
5	
Section B	
6	
7	
8	
9	
Section C	
10	
Total	

The number of marks is given in brackets [ ] at the end of each question or part question.

Section A

Answer all questions in this section.

1 **(a)** A genetic study of coat type in purebred dogs has identified a mutation in the keratin gene KRT71, associated with the curly fur phenotype in some breeds.

Keratin is known to be abundant in cysteine. Suggest how a single mutation in KRT71 can lead to this curly fur phenotype. [2]

- Mutation is likely to be a missense mutation that leads to a frameshift in the nucleotide sequence;
- Leading to an unusually high number of codons specifying for the amino acid cysteine in the primary structure of altered keratin;
- Cysteine has a sulfur-containing functional group/ thiol group;
- The abundant number of cysteines (in altered hair keratin) allows numerous disulphide bridges to form between cysteines, resulting in curly hair;

(b) Mandy decided that she wanted to loosen the curls in her poodle's fur. She first hosed her poodle down with water and then applied an ammonium thioglycolate solution (a reducing agent) as she rolled her poodle's fur around larger-diameter rollers. Next, she applied hydrogen peroxide (an oxidising agent) on the fur to set the curls. She was pleased with the results.

Explain how this works at the molecular level. [2]

- When fur keratin is exposed to the reducing agent, hydrogen lead to breaking of pre-existing disulfide bonds between cysteine amino acids;
- New disulfide bonds form when the fur set in rollers is exposed to the oxidising hydrogen peroxide;
- Diameter of curlers determine the number of new disulphide bonds that will form;
- Fur wound around larger-diameter curlers allow only a few disulfide bonds to form; / if the curling rod has a small diameter, lots of disulfide bonds will form for curlier fur; (A! reverse statement)

(c) Mandy was warned by the vet never to feed her poodle chocolates. Dogs cannot metabolize theobromine found in chocolates as well as humans can. For example, it will take a dog 17 ½ hours to metabolize and excrete just half of the theobromine it has ingested. It can thus accumulate in the animal's body to a level that can be toxic or even fatal.

Theobromine is a natural compound found in the cocoa bean and other plants. Genes of the cytochrome P450 family code for enzymes that are involved in detoxification of drugs and compounds such as theobromine. Human enzymes work twice as fast as canine enzymes.

Suggest the advantage of possessing this enzyme and discuss where this discovery places humans and non-canine mammals on the phylogenetic tree with dogs. [4]

- Mammals have the enzyme to break down theobromine;
- suggesting that there must have been a **common ancestor** of mammals that had the ability to break down theobromine;

- This ability to digest theobromine, found naturally in cocoa plants helped our human ancestors to feed on another food source considered toxic to other mammals (ancestors of modern day dogs such as wolves) competing for food in the forest;;
- Human enzymes are more efficient at detoxifying theobromine than canine enzymes possibly resulting from altered aa sequence that confers a configuration that has a higher affinity and hence higher reaction rate of theobromine detoxification / break down / removal;;
- Suggests human mammals and non-canine mammals share a **derived trait** that groups them into a **clade (**through descent with modification) with dogs as an **outgroup**;;

(d) Mandy discovered that her poodle is suffering from biotin deficiency. This is a rare nutritional disorder which can become serious, even fatal, if allowed to progress untreated. Biotin is part of the B vitamin family. Mandy realised that her poodle's biotin deficiency can be caused by consuming raw egg whites over a period of time. Egg whites contain high levels of avidin, a protein that binds biotin strongly.

In fact, this strong avidin-biotin interaction has been exploited in antigen detection in medical research. State the advantage and outline how such an immunoassay works. [2]

- Primary antibody recognises specific antigen through antigen-binding site;
- Secondary antibody, which binds primary antibody, is biotinylated;
- Avidin molecule can be attached to a fluorophore or enzyme for detection step;
- Attached biotin will form complex with avidin-conjugated enzyme or fluorophore to amplify signal / increase sensitivity of detection (enzymatic conversion of chromogenic substrate to a coloured product or detection by fluorescence);

[Total: 10]

2 Human tau is encoded by the MAPT gene, located on chromosome 17. Tau is the major microtubule associated protein (MAP) of a normal mature neuron. It is found naturally as six molecular isoforms in the human brain. An established function of MAP is their interaction with tubulin and promotion of its assembly into microtubules and stabilization of the microtubule network.

(a) **(i)** Suggest how these different six isoforms of the Tau protein can be formed from a single gene. [1]

- Alternative splicing of pre-mRNA to form different types of mature mRNA, which can then be translated into the various isoforms of Tau protein;;

(ii) Explain the biological significance of the mechanism in (i) in normal physiology. [1]

Any 1:

- Developmentally regulated expression of certain isoforms to perform specific functions at specific times of embryonic development;;
- Different isoforms are expressed in different cells and tissues to serve different functions;;
- Different isoforms are expressed in the same cell at different times in appropriate response to different types of stimuli;;

Alzheimer's disease and related neurodegenerative diseases are collectively known as tauopathies as they are characterized by the presence of aggregates of Tau. Tau protein is abnormally hyperphosphorylated and aggregated into bundles of filaments.

(b) Explain how hyperphosphorylation may play a role in the aggregation of Tau protein. [2]

- Addition of phosphate groups to Tau protein altered charge configuration on the protein;
- And result in conformational change to the binding interface of the protein;
- That enable complementary binding of the Tau proteins with one another;
- Reinforced by the formation of weak bonds e.g. hydrogen, ionic attractions etc. between the charged R groups and the added phosphates;

Tau protein consists of unusually high levels of hydrophilic amino acids. Because of its hydrophilic character, it does not adopt the compact folded structure typical of most cytosolic proteins. Instead the entire Tau molecule is shown to be 'natively unfolded' or intrinsically disordered. This means that the polypeptide is highly flexible and mobile.

(c) **(i)** Explain the principles that regulate the compact folding of typical cytosolic proteins. [3]

- Ref to hydrophobic effect;
- When protein place in water / cytoplasm / aqueous medium;
- Folding of the protein to enable the enclosure of hydrophobic R groups in the interior core;
- While hydrophilic groups distributed around this core and presented on surface;
- For interaction with water via H bonds;
- Thus even though protein folding creates order / results in loss in entropy;
- Which is unfavourable;
- Folding results in favourable system as it reduces no. of water molecules interacting with hydrophobic molecules in the polypeptide chain;

(ii) Suggest the biological significance of these characteristics of Tau protein. [1]

- Tau is still able to fulfil its physiological function even after harsh treatment / changes in environmental conditions, as it does not typically fold into a ordered configuration;;

(d) Scientists conducting research on Tau protein were able to extract much information of the protein function from its gene sequence. Explain the benefits of such applications with reference to biological and medical communities. [2]

Any 2:

- Ref modelling the function of protein to carry out effective drug design to target defective proteins in diseases;;
- Able to differentiate between functional and defective proteins that caused diseases;;
- Understand significance of proteins acting in metabolic pathways within the organism;;
- AVP;;

[Total: 10]

3 Six percent of the proteins in the yeast genome show no homology in their sequences when compared with proteins of known function in other organisms (Fig. 3).

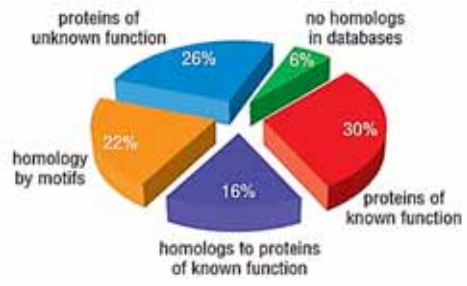


Fig. 3

(a) Protein **Y** is one such protein. Rationalise how two named methods can be used to predict its function. [2]

Any two methods (@ 1 mark per method – ½ m for name of method, ½ m for rationale)

- Use of DNA or protein microarrays / chips;;
- to identify known genes / proteins of similar function to gene / protein **Y** through similar patterns of expression;;

- Inactivation of gene **Y** at the DNA level by insertional mutagenesis;;
- or knockdown of gene expression by RNA interference (RNAi)
- so that the mutant phenotype is an outcome of disrupted biological process impacted by the loss or decrease of protein **Y**, respectively;;

- Attachment of a tag sequence such as GFP to the gene **Y** to monitor localisation of the protein is in the cell via fluorescent GFP;;
- location of protein **Y**-GFP in the cell can provide a valuable clue to its functions through the pathways in which protein **Y** participates in;;

- Yeast two hybrid methods to identify interacting partners of protein **Y**;;
- As proteins tend to function as part of a complex (of two or more gene products), if the function of one of the interacting proteins is known, its binding to protein **Y** will help reveal protein **Y**'s function;;

(b) Outline a named one-step method that can be used to purify yeast proteins that have been expressed in *E. coli*. [4]

- Affinity chromatography of histidine-tagged recombinant proteins;
- Fuse DNA coding for multiple copies of histidine at the N- or C-terminals of yeast ORF in an expression vector;
- Transform into *E coli* and prepare cell lysate from bacterial culture;
- Purify histidine-tagged proteins using IMAC, also known as metal chelate affinity chromatography;
- The medium is first charged with Ni^{2+}, which is immobilized to the medium via a chelating ligand;
- The complex protein sample is applied;
- His-tagged proteins bind the immobilized metal ions;

- Bound protein can be eluted by competitive elution with imidazole, or by lowering the pH;

(c) Chronic myelogenous leukemia (CML) and variants of acute lymphoblastic leukemia (ALL) are blood cancers that arise from a chromosomal translocation resulting in a constitutively active Bcr-Abl tyrosine kinase.

Suggest how you will design a RNAi-based strategy to treat this type of cancers. [2]

- Use siRNAs that recognize the breakpoint of the fusion protein to see if they are effective at specifically silencing Bcr-Abl tyrosine kinase expression / suppression of Bcr-Abl tyrosine kinase activity;;
- in both cell lines that express this hyperactive form or isolated CML cells from patients vs normal cells as siRNA should not interfere with the expression of wild type *c-abl* or *c-bcr*;;

(d) Describe one advantage and one limitation each, of fluorescent and chromogenic detection labels, used in immunohistochemistry. [4]

Fluorescent labels emit light while chromogenic labels are converted from soluble substrates to insoluble, chromogenic end products by antibody-conjugated enzymes.

Advantages of fluorescent labels (any one):

- Easier multiplexing as more colours for multiple targets due to their ability to emit light at unique wavelengths are available compared to chromogenic dyes;;
- Confers better target colocalization as fluorescent dyes allow separate identification of co-localized targets;;
- Easier to visualize rare and high abundant targets on the same slide;;
- Fewer steps since there is no step for substrate addition;;

Disadvantages of fluorescent label:

- Susceptible to photobleaching as exposure to light may diminish fluorescent signal over time;;

Advantages of chromogenic labels (any one):

- Greater sensitivity due to signal amplification via indirect chromogenic detection that increases signal strength;;
- Signal are longer-lasting as chromogenic stains are more resistant to photobleaching compared to fluorochromes;;

Disadvantages (any one):

- Difficult to study target colocalization due to difficulty in distinguishing mixed colours from when targets co-localize;;
- Difficult to visualize rare and high abundant targets on the same slide;;
- Difficult multiplexing as there are fewer colours available compared to the broader range of fluorescent dyes which offers a wider range of emission spectra;;

[Total:12]

4 A protein from bacteria *Thermococcus* species that lives in a thermal vent has been isolated. You obtained the gene sequence and synthesized the protein in *Escherichia coli* bacteria that grow at room temperature.

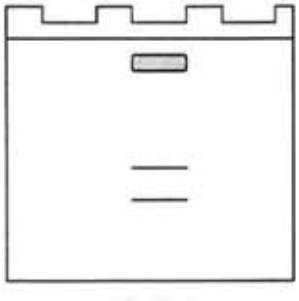
(a) Explain briefly why and how the gene sequence of the desired protein and *E.coli* were used for the study of the protein structure [3]

- Protein extracted from the bacteria *Thermococcus* is in **too minute amount** / insufficient for the study of the protein structure;

- Moreover, it is **difficult to purify** the desired protein from the rest of the proteins from the same bacteria;
- Difficult to create the environment for the culture of the original bacteria;
- Cloning the gene sequence of the desired protein into an expression vector followed by transformation into the *E.coli* facilitates **expression of a large amount of the proteins** for further studies;
- *E.coli* cells are also **safe** and can **multiply quickly** at 37^{o}C;

- Recombinant expression allows the protein to be **tagged** (e.g. His-tags) to **facilitate purification;**

During the purification process, you heated the *E.coli* cells to 90^{o}C and then subjected the soluble portion of the heated mixture to native polyacrylamide gel electrophoresis (PAGE). After applying Coomassie stain to the gel, you obtained the gel as shown below.



(b) Suggest why the purification process involved the heating of the *E.coli* cells to 90^{o}C. [1]

- At 90^{o}C, many of the proteins from the *E.coli* will be **denatured** by heating / unfold from their 3D configuration and **expose their hydrophobic R groups** which are in the interior of the molecule;
- Hence they become **insoluble in water** and precipitate out of solution, leaving behind the desired heat stable protein in the soluble portion of the heated mixture;

(c) Give reasons why your procedure will not allow you to accurately determine the molecular weight of your protein. [3]

- **Lack of protein ladder**/ marker;
- **Native gel / lacks SDS**;
- protein chain / polypeptide chain is **not coated by –ve charges**, hence **not denatured**;

- The protein from the *Thermococcus* living in the thermal vent should be able to **withstand high temperature;**
- and **remained folded in its original specific shape** unlike the other *E. coli* proteins;

- The folding pattern / specific shape of the protein will **affect its migration through the gel matrix;**
- Hence, the molecular weight cannot be accurately determined

(d) Nevertheless, you want to isolate the protein so that you can further elucidate its structure and function. State one method which allows you to do so without affecting its structure or function. [1]

- Size exclusion chromatography;;

(e) Your lab assistant helps you isolate the protein based on your method in **(d).** Upon isolation of the protein, he incubates half of your protein sample at room temperature (25^{o}C) overnight and then applies it to an SDS-PAGE gel. He finds the protein no longer migrates as one band, but many. However, the other half of the protein sample, which he had left overnight at 80^{o}C, still migrates as one single band in the non-SDS gel. Explain this observation. [2]

- SDS is an **anionic detergent** that **denatures the protein**, which is made up of **more than one subunit polypeptide chain**, by applying **many negative charges** to the protein;
- It **disrupts H bonds, ionic bonds and hydrophobic interaction** within and **between the chains**;
- The **subunit chains are separated** and hence show up as multiple bands;

- However, for the proteins heated at 80^{o}C overnight, the protein which is derived from thermophilic bacteria is able to **remain in its specific shape / not denatured / quaternary structure not affected** and hence only one band is shown;

[Total:10]

5 Proteins rarely act alone as they are built from a large number of protein components organized by their protein–protein interactions.

(a) Define protein-protein interactions. [1]

- the physical contacts of high specificity established between two or more protein molecules;;

(b) Describe the protein-protein interactions between

(i) the subunits of a named protein. [2]

- Ref. to two α helices bind together to form a coiled-coil / helix-helix binding between leucine zipper in a transcription factor;;
- Ref. to hydrophobic interactions between leucine side chains;;

OR

- Ref to matching of complementary surface between subunits of haemoglobin;;
- Ref to bonds between these surfaces such as hydrophobic interaction, ionic bonds and hydrogen bonds;;

(ii) named proteins in a complex. [2]

- Surface-surface interaction between the enzymes of pyruvate dehydrogenase multi-enzyme complex;; (Ref. to non-covalent bonds at the surfaces)
- Ref to name of enzymes: pyruvate dehydrogenase; dihydrolipoyl transacetylase E2; dihydrolipoyl dehydrogenase; (named E1, E2 and E3 – 1m)

(c) Aberrant protein-protein interactions are the basis of multiple aggregation-related diseases such as Alzheimer's disease. Suggest how multiple aggregation of proteins in aberrant protein-protein binding can result in diseases such as Alzheimer's. [3]

- To function, most proteins fold into a strictly defined three-dimensional structure, their native conformation;
- Proteins in a non-native conformation may aggregate and / or engage in aberrant interactions with other proteins;
- These disease-related proteins can drive disease processes by altering individual protein complexes and protein network dynamics;
- Ref. to more protein-protein interactions gained than lost in disease state;
- Ref. to aggregated proteins being cytotoxic leading to disease state;

[Total: 8]

Section B

Answer 3 out of the 4 questions in this section.

6 **(a)** p53 is a well-known tumour suppressor protein with a proline-rich region implicated in initiating cell apoptosis via interaction with different promoters. Suggest the structural and functional significance of having an abundance of proline in these regions. [4]

- When proline is involved in forming a peptide bond, it does not have a hydrogen on the α amino group;
- cannot donate a hydrogen bond to stabilize an α helix or a β sheet;
- proline is known as a helix-breaker;
- causes a slight bend to the secondary structure due to the lack of the hydrogen bond;.
- Proline is often found at the end of α helix or in turns or loops;
- Unlike other amino acids which exist almost exclusively in the *trans*- form in polypeptides, proline can exist in the *cis*-configuration in peptides;
- The *cis* and *trans* forms are nearly isoenergetic / can alternate between isomers readily;
- The *cis/trans* isomerization can play an important role in the folding of proteins which allows flexibility for interaction with different promoters;

The N-terminal proline-rich domain of human p53 has been shown to be important for the induction of apoptosis. However, the corresponding region in mouse and other species is not highly conserved. Research efforts to confirm this domain's role in eliciting cell death have been inconclusive. While one mutant missing a large proportion of this proline-rich domain is selectively defective in causing cell death in the mouse, other mutants display wildtype phenotype in all assays performed.

(b) What conclusions can currently be drawn here about the importance of this region for the purpose of tumour suppression? [2]

- These findings are not consistent with the interpretation that the proline-rich region of mouse p53 is an 'apoptotic domain';
- P53 is a transcription factor that binds the promoter of target genes;
- N-terminal proline-rich domain may play a role in promoter discrimination;

- The low sequence conservation with human p53 suggests it is subjected to frequent mutation and the protein probably served its function of promoting apoptosis via some other mechanisms and not the proline rich domain;

(c) The p53 DNA-binding domain recognizes specific regulatory sites on the DNA. Suggest how this is achieved. [2]

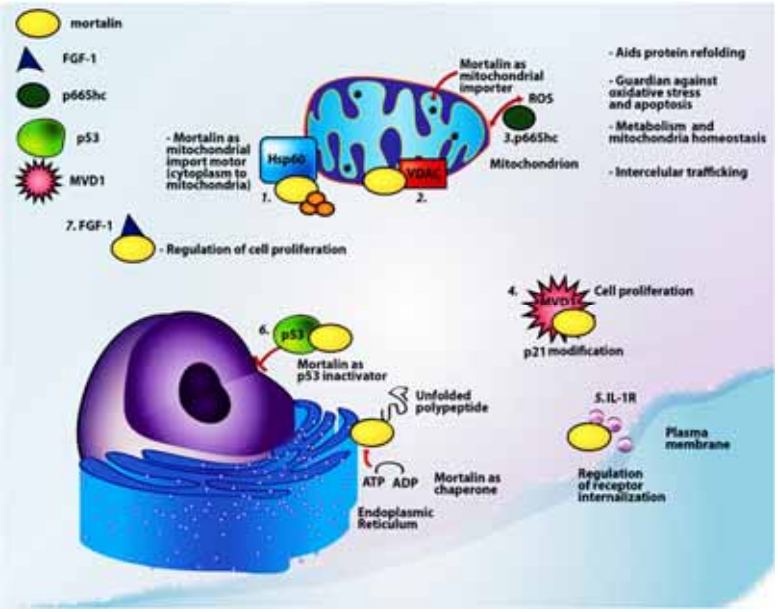
- The DNA binding portion must be made up of many amino acids with positively charged residues; (*FYI - it is rich in positively charged arginine amino acids that interact with DNA)
- because DNA is negatively charged due to the abundant phosphate groups in the backbone;
- This domain needs to have a **complementary shape** to bind to the DNA helix;
- so that interacting R groups fit well into the minor and major grooves of the DNA helix to ensure a specific binding;

(d) Mortalin, a member of the heat shock protein (HSP) 70 family, is overexpressed in different tumour types. It is thought to contribute to the process of carcinogenesis by multiple ways.

With reference to the figure below and your knowledge of how cancer arises, suggest two ways in which mortalin may work to promote the development of cancerous cells. [2]

- Mortalin sequesters of p53 in the cytoplasm so that it cannot enter the nucleus to function as a TF to activate expression of genes involved in tumor suppression;;
- Alters p21 shape so that it cannot function in cell cycle arrest for repair of DNA damage;;

- In mitochondria, mortalin promotes metabolism and facilitates cell growth and eventual proliferation;;
- In ER, mortalin prevents movement of polypeptides into ER lumen to fold into their specific shape, that contributes to tumour suppression;;



(Source: http://www.mdpi.com/biomolecules/biomolecules-02-00143/article_deploy/html/images/biomolecules-02-00143-g002.png)

[Total: 10]

7 In order to investigate the subcellular locations of proteins **1**, **2** and **3**, rat liver cells were homogenised and the cellular components separated into cytosol and lysosomes by centrifugation. The temperature during these steps was maintained between 0-4°C.

The components were then subjected to 2D gel electrophoresis and the results are shown in Fig. 7.1 and Fig. 7.2.

Fig. 7.1

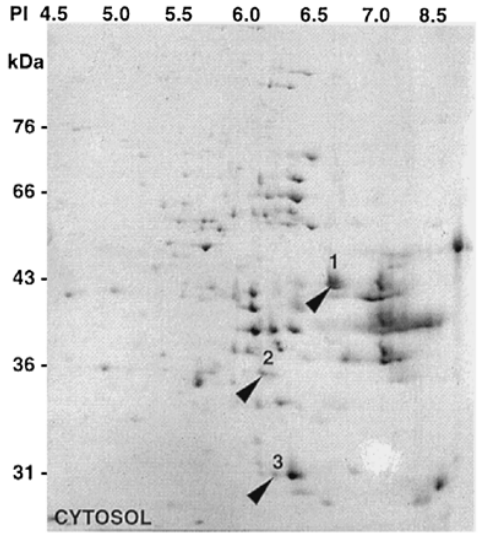
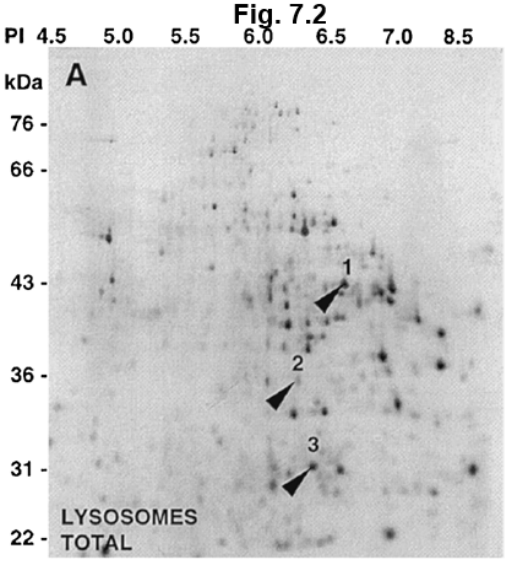


Fig. 7.2



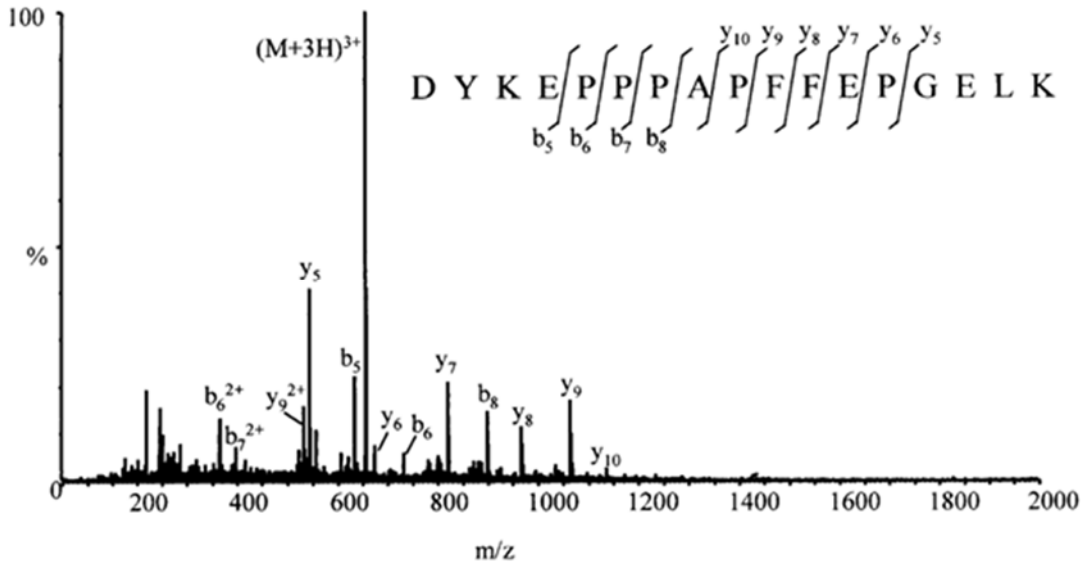
(a) Explain the principles behind 2D gel electrophoresis. [3]

- Separation by isoelectric point in 1st dimension;
- Proteins move under influence of electric current through a gel with pH gradient until at pH at which they have no net charge;
- Gel is rotated 90º;
- Separation by size in the 2nd dimension;
- By SDS-PAGE;
- SDS coats each polypeptide such that all have equal charge to mass ratio;
- smaller proteins move through gel faster;

(b) The pH of lysosomes is usually 4.5, and yet Fig. 7.2 shows that proteins are not clustered around pH 4.5 but are instead spread across the entire pH gradient of the gel. Comment. [2]

- Fig. 7.2 shows proteins at their isoelectric points;;
- Optimal pH is not one where enzymes are at isoelectric points;;

The spots which are labelled 1, 2 and 3 were identified as proteins 1, 2 and 2 when the spots were excised, subjected to trypsin digestion and then fed into tandem mass spectrometry (MS/MS). The mass spectrum results of one of the peptides are shown in Fig. 7.3.



Analysis of a peptide fragment by MS/MS. The C-terminal (y-type) or N-terminal (b-type) fragments produced are due to the way that the peptide is fragmented.

Fig. 7.3

(c) Describe briefly the principles of mass spectrometry, and state exactly what is measured. [3]

- A sample undergoes fragmentation forming ions in gaseous form;
- The ions are accelerated/travel through magnetic fields in a mass spectrometer;
- Magnetic field deflects ions according to mass & charge;
- smaller mass & higher charge lead to greater deflection;
- Greater deflection leads to shorter time of flight;
- Ions detected by detector;
- The ratio of mass to charge (m/z) of the particles;

(d) Explain why Fig. 7.3 shows two different peaks for fragment y_9. [1]

- The 2 fragments have different charges;
- The one with +2 charge travels faster than the fragment with +1 charge;

(e) Name another method which can be used to determine amino acid sequence, and state one advantage of this method over using MS/MS. [1]

- Edman degradation;
- Can easily sequence novel proteins which are not in database;

[Total: 10]

8 **(a)** Multi-protein complexes are critical to biological processes inside cells. Proteins that have not been functionally annotated can be determined based on their role as components of complexes. Multiprotein complexes can be purified by affinity purification and then analysed by mass spectrometry.

Describe how tandem affinity can be used to examine large scale *in vivo* protein-protein interactions of a protein named **FUNK**. [4]

- TAP uses two tags sequentially to isolate and purify complexes under native conditions;
- Fuse the cDNA coding for the TAP tag to the cDNA for the target protein FUNK;
- Introduce this into the host cell or organism as a bait to trap endogenous interacting partners;
- Tagged FUNK is isolated along with its associated partners via interaction of the Protein A tag with immunoglobulins immobilized on agarose beads;
- Proteins that are non-specifically bound to the column are washed off,
- The retrieved FUNK complex is released by protease cleavage using the TEV (Tobacco etch virus) protease that cleaves a TEV recognition sequence located between the first (Protein A tag) and the second tag (calmodulin binding peptide, CBD).
- In the second affinity step, the complex is immobilized to calmodulin coated beads via the CBP tag.
- This step removes the TEV protease and further contaminants present. The CBP–calmodulin interaction is calcium dependent;
- Removal of calcium ions with chelating agents (e.g. EGTA) is the second specific elution step that yields the final protein complex;

Simplified WTTE:

- The TAP tag consists of two affinity modules, so that the target protein FUNK and its binding partners can be isolated in two consecutive purification steps;
- The tag consists of a calmodulin binding domain (CBD), a cleavage site for the tobacco etch virus (TEV), and the IgG binding units of the protein A of *S aureus*;;
- In the first purification step, the complex of TAP tagged FUNK and its interacting partners is bound to an IgG column and then eluted by cleaving off the protein complex from the column using the TEV cleavage site of the TAP tag;
- In the second purification step, the TAP-tagged complex is bound to calmodulin beads and then eluted using EGTA;
- Interaction partners of the TAP-tagged bait protein FUNK are then subjected to downstream analysis for identification;

(b) Identify one advantage and one disadvantage of using epitope tagging in the isolation of protein complexes. [2]

Advantages (any 1):

- Bait proteins can be fused to an epitope-tag and an antibody can be used that is directed against the tag instead of the bait protein is used for complex retrieval;;
- Many different cDNAs can be fused to the same tag in parallel and complexes retrieved using the same antibody;;
- Antibodies against these tags are commercially available;;

Disadvantages (any 1):

- Overexpression may adversely affect the assembly of a protein complex;;
- Overexpression may cause cytotoxicity;;
- Tag may interfere with protein folding, protein function, or the ability to interact with other proteins;;

(c) How does DNA transposon strategy differ from RNAi in characterising gene function? [2]

Any two:

- Transposon inserts into genes to disrupt gene and its function at the DNA level whereas RNAi uses RNA-based gene-silencing;;
- Transposon makes use of homologous sequences to replace endogenous genes in a targeted manner whereas RNAi makes use of specific small double-stranded RNA (dsRNA) molecules to direct homology-dependent control of gene activity;;
- Transposon uses transposase to cause excision and insertion into host genome whereas RNAi uses two ribonuclease (RNase) III enzymes, Drosha and Dicer to generate siRNA from dsRNA;;

(d) To generate the stable expression of siRNAs in mammalian systems, mammalian expression vectors must be able to direct the intracellular synthesis of target-specific siRNAs.

State how stable and appropriate expression is made possible. [1]

- Inducible promoter and appropriate control elements can be inserted in the expression vector for regulated expression of siRNA after introduction into cells or organisms;;

(e) State the strengths underpinning immunohistochemistry in studying protein function. [1]

- Combines specificity, sensitivity and resolution through combination of immunological, biochemical and histological techniques;;

WTTE:

- To identify specific tissue components via **specific** antigen-antibody interaction that can be visually detected through a visible tag/ label (sensitivity of detection)
- Obtain the spatial and temporal distribution and localisation of specific cellular components or proteins within a cell or tissue (resolution);;

[Total: 10]

9 **(a)** Platelet-derived growth factor is received by a fibroblast cell via its receptor tyrosine kinase (RTK). The ligand-bound RTK activates a SH2 domain-containing protein, which in turn activates the G protein Ras. Activated Ras then activates the first out of three serine-threonine kinases of the mitogen-activated protein (MAP) kinase pathway. All three kinases in this pathway activate multiple substrates at each step. The final enzyme in the pathway causes a change in gene transcription, which results in the proliferation of the fibroblast for blood vessel formation.

Account for the conformational changes that occur to RTK during signal reception. [2]

- Dimerization;
- Activation of tyrosine kinase region of cytoplasmic tail;
- Cross-phosphorylation of tyrosine residues on cytoplasmic tails;
- Allows for recognition and binding by SH2 domain-containing protein;

(b) Explain what a protein domain is and the role of the SH2 domain in its interaction with the activated RTK. [2]

- Definition of domain: specific 3D structure within protein; that can evolve, function and exist independently of the rest of protein;;
- SH2 domain specifically recognizes and binds to the phosphorylated tyrosine side chain of the RTK;
- Allows activation of SH2 domain-containing protein;

(c) Outline the conformational changes and interactions involving the G protein Ras. [2]

- Inactive G protein is bound by GDP;
- Binding by SH2 domain-containing protein;
- GDP replaced by GTP;
- Causes conformational change of G protein, activating it;
- Activated G protein activates 1st kinase of MAPK pathway;

(d) Discuss the significance of the MAP kinase pathway in cell signalling. [2]

- Signal transduction;
- Phosphorylation cascade;
- One activated kinase phosphorylates and activates the next in sequence;
- Signal amplification (multiple substrates activated at each step);

(e) Numerous lines of evidence support a role for RTKs in the growth and progression of human malignancies.

Suggest how selective SH2 inhibitors can be used to treat cancer. [2]

- By disrupting SH2-RTK interactions, prevents transduction of signal from ligand-RTK activation;;
- inhibit Ras activation, thus blocking MAPK signalling that activates genes promoting the growth of malignant cells;;

[Total: 10]

Section C

Answer the question in this section.

10 **(a)** Fig. 10 shows the workflow involved in the identification of Protein **S** (boxed band) from a protein mixture extracted from a tissue.

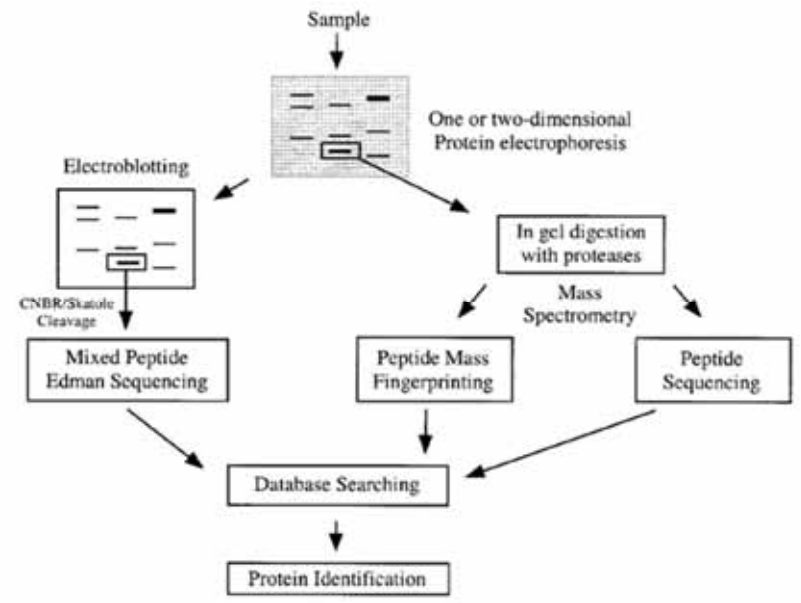


Fig. 10 (Source: Molecular Biologist's Guide to Proteomics.)

After the full sequence of protein **S** has been determined following database search of its peptide mass fingerprint, what can be done to predict its function using an *in silico* approach? [8]

- Make sequence comparison of protein **S,** in **FASTA** format, and using **BLAST** data analysis tool;
- to search for **homologous** sequences against protein databases such as **SwissProt**, **UniProt**;
- Finding related **proteins that belong to the same family**; / homologues of protein **S**, whose functions are known;

Points below at max of 2m:

- Since 3D structure is more conserved than primary structure, related proteins can be identified by comparing their 3D structures using **PDB** database for function prediction of protein S;;
- Refer to homology modelling, threading and *ab initio* techniques for obtaining 3D structures to predict function (max 1 ½ m)
  - *Ab initio* if protein S has no homologues to proteins in PDB database that have a solved structure that and is a small protein (<100 residues);
  - Threading against a library of structural templates/ folds, to produce a list of scores with the aid of a scoring function that assesses the fit of protein S sequence to a given fold;
  - Homology modelling if protein S has homologues with known structures as they are likely to share very similar structures and thus function;

Max 3m for points below:

- **ClustalW** alignment tool for **multiple sequence alignment** of protein sequences;
- Homologous amino acid sequences from different species that share a common ancestor are called orthologs;
- Have identical or similar functions through the presence of conserved regions that indicates conserved structural and/or functional elements (motifs) within the sequence;;

- **ClustalOmega** can also be used to construct phylogenetic trees for viewing the evolutionary relationships between protein of interest and proteins of known functions in the databases;;
- Perform **phylogenetic profiling** (rationale - proteins that interact are more likely to co-evolve);
- Look for homologous interacting pairs to find related proteins in other species as the known functions of the interacting partner will provide clues to function of protein **S**;

Points below at max 2m:

- Use **ScanProsite** program to scan sequence of protein of interest (POI) against **SWISS-PROT** database to look for the occurrence of patterns and profiles;;
- Compare protein sequences with secondary protein databases such as **Pfam, ProSite, ProDom** databases, which contains information on motifs, signatures and protein domains;;

Note: Further experiments (e.g. using Y2H, BRET, FRET, TAP, Co-IP) to confirm predicted functions and functional associations of shortlisted candidate interacting partners (this is not a mark point as it is not in silico but a nice conclusion about experimental assays to confirm function).

(b) You are the leader of an R&D laboratory. Your task is to investigate how a potential drug, Molecule **D**, interacts with a protein kinase that has been implicated in the signalling pathway in cancer.

Outline a strategy to generate the 3-dimensional structure of this protein-drug interaction that will aid drug design. [6]

In XRC, electrons in the crystal of protein complex diffract the X-rays and the scattered rays are recorded by X-ray detector. Mathematical analysis of the diffraction pattern produces an image of the electron clouds surrounding atoms in the crystal. This electron density map reveals the overall shape of the molecule and the positions of each of the atoms in 3D space. The primary structure of the protein is then fitted into the electron density map obtained in model building stage. To obtain the 3D structure of protein kinase-Molecule D complex, the major steps are as follows:

- Express high quantities of pure POI, the protein kinase and molecule **D in the same eukaryotic cell (> 10 mg/ml);**
- Idea of **TAP-tagged POI** and **isolating complex under native conditions;**
- Obtain ordered/ high quality crystals of the protein kinase-Molecule **D** complex for XRC (e.g. optimising conditions for crystallisation by hanging drop vapour diffusion method);
- Subject crystal to a narrow beam of X rays and the scattered rays from the electrons in the atoms are detected and captured as reflections in a diffraction pattern; (max 1 ½ m)
  - The protein crystal is mounted and positioned in a precise orientation with respect to the x-ray beam and the X ray detector;
  - The crystal is repeatedly exposed to the X-ray beam, while changing its orientation (usually by rotating, 1 degree a time, through at least 30 degrees);
  - The scattered waves reinforce one another at the detector if they are in phase (in step) there, and they cancel one another if they are out of phase;
  - Each exposure provides an image and each spot on the image is a diffracted X-ray beam, which emerged from the crystal and is registered by the X-ray detector;

- The **diffraction pattern** obtained on the X-ray detector appears a series of discrete spots which are known as reflections (max 1 ½ m)
  - Each atom contributes to the intensity of each reflection. Each spot from a diffracted beam is defined by three properties: (i) the amplitude, which is measured from the intensity of the spot; (ii) the wavelength, which is set by the X-ray source; and (iii) the phase;
  - The amplitude of the wave scattered by an atom is proportional to its number of electrons (e.g. a carbon atom scatters six times as strongly as a hydrogen atom does.);
- The phase information is obtained through a variety of methods such as Molecular Replacement (if there exists a starting model, i.e. coordinates of a similar protein in PDB)
  - In the absence of a starting model, Isomorphous Replacement methods such as Multiple isomorphous replacement (MIR) or Multi-wavelength Anomalous Diffraction (MAD) can be used;;
- The intensity and position of each spot is measured by complex mathematical computation called a Fourier transform to calculate the electron density of the molecules within the crystal;
- The **electron density map** generated provides information where the atoms are located, which can be used to build a model of the protein kinase-Molecule D complex;
- To which model building of 3D structure can be applied by fitting the primary structures of the protein kinase and of the interacting Molecule **D**;
- The 3D structure quality of protein kinase-Molecule **D** complex will be indicated by the **resolution (known as the R value)** of the model, especially the Free R;
- Generally, reliable crystallographic models have **resolutions of 2.0 Å or less**, i.e. R values of ≤ 0.20 and Free R values of ≤ 0.25;

(c) With reference to three named examples, explain how protein modification or cleavage is involved in gene expression and its regulation. [6]

- Chromatin remodelling;
- Ref to histone acetylation;
- Reduced positive charge; weaken interaction with negatively-charged DNA;
- Displaces nucleosomes;
- Increases transcription rate;

- Ref to cell signalling pathways that upregulate/downregulate gene expression;
- Binding of G-proteins by GTP for activation;
- Hydrolysis of GTP to GDP to inactivate G protein and return to inactive conformation;

- Ref to kinases in signal transduction/phosphorylation cascades;
- Activation by phosphorylation; conformational changes;
- Deactivation by phosphatases;

- Ref. to post-translational modifications of proteins to make them fully functional;
- E.g. insulin via proteolytic cleavage/ cleavage of signal peptides;

- AVP/any other valid examples

[Total: 20]