Version GM/final

EGT2
ENGINEERING TRIPOS PART IIA

Friday 7 May 2021 13.30 to 15.10

Module 3G1

MOLECULAR BIOENGINEERING I

*Answer not more than **three** questions.*

All questions carry the same number of marks.

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet and at the top of each answer sheet.*

STATIONERY REQUIREMENTS
Write on single-sided paper.

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed.
You are allowed access to the electronic version of the Engineering Data Books.

10 minutes reading time is allowed for this paper at the start of the exam.

The time taken for scanning/uploading answers is 15 minutes.

Your script is to be uploaded as a single consolidated pdf containing all answers.

1 Consider a repressilator circuit that has three repressor proteins (A, B, and C) controlling each other's expression in a cyclic manner. Protein C also controls the expression of a red fluorescent reporter protein (RFP, maturation time = 40 min) that can be used to monitor the function of this circuit using fluorescence microscopy.

(a) (i) Sketch the expected changes in the concentration of all of the three repressors over time. On a separate sketch underneath this, using the same time scale, sketch the expected concentration of the RFP reporter and also its signal.

[20%]

(ii) Explain what is shown by the plots you have drawn above.

[10%]

(b) Imagine that all three of the repressors have the same repression coefficient. Imagine also that their peak concentrations are 100 times higher than their repression coefficients. If all of these proteins are only diluted through growth, and the generation time of this kind of bacteria is 20 minutes, what is the expected period of this oscillator?

[25%]

(c) What would be the expected period of this oscillator if you add a degradation tag to each of these repressors and to the RFP gene as well, which reduces their half-lives to 2 min? [10%]

(i) Why can you no longer use the RFP reporter for monitoring the dynamics? [10%]

(ii) What modification is needed to this circuit design to observe the dynamics? [10%]

(d) To achieve a slower periodicity, we can add two more repressors in the cycle named D and E. Draw the conceptual design of the overall circuit as it would be laid out on a plasmid. [15%]

2 (a) Sketch and then design the sequence of a single strand of synthetic DNA that will fold to make a double-stranded molecule in the shape of the capital letter T. Take the lengths of both the horizontal and vertical portions as 2k. Ignore the thickness of the double strand itself and assume that a five base loop is needed for a DNA strand to fold back on itself.

[25%]

(b) Suppose you wish to make the equivalent molecule but from enzymatically generated RNA and without the use of PCR or cloning.

(i) Describe how you would derive a new design based on the first, and what components you would need in the enzymatic reaction.

[40%]

(ii) Considering your answer to part (a), what problem might you have in making the reaction work with the new design?

[15%]

(iii) Without using other enzymes or gel electrophoresis, how can you modify your design to remove the DNA once the RNA synthesis has taken place?

[20%]

3 A frog has been discovered that secretes antibiotic from its skin and it is hoped that, if it can be produced in sufficient quantities, it might be of therapeutic value. Characterisation of the antibiotic indicates that it is in fact a mixture of peptides that have antibiotic function. Therefore we would like to clone the genes encoding these peptides in order to allow them to be expressed for bulk production of the peptides. The peptides are analysed by direct peptide sequencing and it is discovered that they share their first 7 amino acids in common and also their last 6 amino acids in common. Between these common sequences are found variable peptide-specific sequences of 5-10 amino acids.

(a) You are provided with the genome sequence of the frog and with the amino acid sequences of the peptides. As they have different alphabets, outline how you would go about converting them in order that you can search for the corresponding genes.

[10%]

(b) Having found the segments of the genome that correspond to the peptides, you discover that the two common regions are each encoded by a single segment of the genome nearby on the same chromosome. Furthmore, the many variable segments are arrayed in between these two constant segments. You hypothesise that perhaps some process like antibody gene rearrangement leads to the fusing of the terminal common segments with one variable region to yield a mature peptide-producing gene. You consider that perhaps different cells in the frog's skin contain different such rearranged genes, so leading to the observed diversity of antibiotic peptides secreted by the skin.

Based on the above and starting with an mRNA sample derived from the frog skin, outline an approach by which you can generate a library of all the genes that encode these peptides, ready to transform into host cells to generate a clone library.

[50%]

(c) Having constructed a library as above in (b) using an *E. coli* expression vector, the mixture is transformed into *E. coli*. However the transformation efficiency is extremely poor compared to an identical transformation carried out with the empty vector, and the colonies grow very weakly. What has gone wrong and how might the problem be fixed?

[20%]

(d) As an alternative to expression in living *E. coli* cells, the library is added to a cell-free transcription/translation system made from *E. coli* cells. It is possible to detect transcription of the genes, but production of the peptides is still weak. Give two possible explanations for these observations.

[20%]

4 A yeast gene knock-out library has been created that is a collection of strains. In each strain a different yeast gene has been knocked out by replacing it with a selectable marker. Next to the marker is a barcode that identifies the knocked-out gene. Across the library, all the barcodes are flanked by the same PCR primers.

(a) If there are 6000 strains in the collection, justify the minimum length of barcode that would be required to distinguish all strains.

[10%]

(b) Why in practice might one want to use a longer barcode sequence than above, and what features might it include?

[20%]

(c) Suppose that we wish to discover whether any members of the gene knock-out library confer resistance to a toxin. Outline an efficient process by which we could do this.

[20%]

(d) Having carried out the process above, how would the resistant strains be identified?

[20%]

(e) Curiously, it is found that strains that are resistant to the toxin now also require uracil to be supplied in their growth medium in order to be able to grow. Outline how this fact can be exploited to identify human genes that carry out the same function as the knocked-out yeast gene.

[15%]

(f) Having isolated a plasmid carrying a 3000 base candidate human gene, justify which method of DNA sequencing you would use to sequence it.

[15%]

END OF PAPER