Version SB/final

EGT2
ENGINEERING TRIPOS PART IIA

Friday 5 May 2023 2 to 3.40

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.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed, Supplementary sheet for Q3a

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

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 Your synthetic biology project team is designing a genetic circuit that needs to rapidly respond to the presence of a toxic chemical in an environment. You collected some sample from the target environment and tested the growth of your chassis bacteria in it. The chassis bacterial cells doubled every two hours in this media.

(a) If you use a transcription repressor, which undergoes an allosteric change in activity when bound to this chemical, to control the expression of a fluorescent protein reporter, what is the expected response time of your designed system (time to reach 50 % of the steady-state level). [15%]

(b) After some initial testing you realised that the design is not sensitive to low levels of the toxic chemical. Your team lead suggested you try out three more repressors that are sensitive to this chemical. These new repressor candidates have different cooperativities (n_1, n_2, and n_3) and repression coefficients (K_1, K_2, and K_3). If $n_1 > n_3 > n > n_2$ and $K_2 > K > K_1 > K_3$, where n and K are the corresponding values for the original repressor, which of the three candidates would you pick for your new design and why? [15%]

(c) Since the response time of the above-mentioned design seemed too slow for your application, you decided to use a degradation tag on the reporter protein to speed up the response time.

(i) What other changes do you need in your circuit design to achieve the same steady-state response level? [10%]

(ii) If the degradation tag makes the fluorescent reporter degrade at 0.2 min^{-1} rate, what is the new response time of the designed system? [10%]

(iii) Which of these two versions is expected to be more energetically costly for the chassis cell and why? [10%]

(iv) Could you think of an alternative design for speeding up the response time, which might be energetically more efficient compared to the previous two designs? [10%]

(d) After testing the various designs, you realised that all the circuits are very sensitive to brief pulses of the toxic chemical. You want to avoid that. So you decided to use a feedforward design to ensure that the circuit detects persistent signals only. Draw the detailed schematics of such a feedforward circuit and the corresponding SBOL diagram. [30%]

2 You are using a bacterial plasmid-based expression system for your bioproduction project. Your plan is to produce a fluorescent protein, which is known to have no effect on the growth of the host bacterial cell. You picked a low copy plasmid as the vector for this work, where it is known that on average every newborn cell has 5 copies of the plasmid.

(a) What is the main source of variance of plasmid numbers across the population, if the replication is tightly controlled? [10%]

(b) What is the expected variance of plasmid copy number distribution in the population and what is the expected fraction of plasmid-free cells in the population? [20%]

(c) If each plasmid imposes a 5% demand on the cellular ribosome for the protein production and causes the cell to grow slower proportionally, how would that affect the distribution of plasmid-copy numbers in the population? [20%]

(d) To eliminate any unnecessary burden from the protein production, you have used a transcription activator to turn the protein production on and off as needed. Explain with reason whether or not you expect a plasmid-free and plasmid-containing cell to have the same growth rate, if the protein production is kept off. [15%]

(e) During experiments, you realised that the chemical inducer for the transcription activator seems to affect the growth of the chassis cell. As a result, you sought to avoid keeping the inducer in the growth-medium throughout the bioproduction experiment. Describe a genetic circuit design in which the inducer is added only briefly but transcription is maintained in the activated state for an extended period of time. [10%]

(f) To avoid the takeover by plasmid-free cells, you have also introduced an antibiotic-resistant cassette in the plasmid, and the cells are always grown in the presence of the corresponding antibiotic. Explain with reason if you expect this new system (bacteria containing the newly designed plasmid) to grow faster than the original design. [10%]

(g) After running an extended bioproduction experiment in a culture containing antibiotics, you decided to image the cells under a microscope. Since the cells were grown in the presence of the antibiotic, it is expected that all cells carry the plasmid. However, in your images, you found that a significant number of cells don't show any detectable signal from the fluorescent protein. How would you explain this observation? [15%]

3 (a) The restriction enzyme BamHI cuts the DNA motif shown below, with the cuts on each strand being shown with an apostrophe:

```
5'  G'GATC C  3'
3'  C CTAG'G  5'
```

In the list of restriction enzymes below, only the top strand is shown.

```
BamHI   G'GATC C
BglII   A'GATC T
BsrGI   T'GTAC A
KpnI    G GTAC'C
Sau3AI   'GATC
```

Fill out the empty cells of the table on the attached supplementary sheet to indicate which enzyme combinations produce compatible sticky ends. [20%]

(b) Which of the above restriction enzymes will in general be less useful for building DNA constructs and why? [15%]

(c) We would like to make a protein A/protein B fusion. The sequence of DNA that codes for the C-terminus of protein A is shown below, followed by three versions of the DNA sequence of Protein B. The DNA bases are grouped into triplets corresponding to the reading frame. The EcoRI restriction enzyme cuts `G'AATTC` and is to be used to make the fusion.

```
Protein A C-terminus:  ... TCC TGC TAT GTG AAT TCA TAA

Protein B N-terminus version 1:   GAATTC ATG AAA CCC TTT GGG CAC ...

Protein B N-terminus version 2:  GAATTCA ATG AAA CCC TTT GGG CAC ...

Protein B N-terminus version 3: GAATTCAA ATG AAA CCC TTT GGG CAC ...
```

Which of the three alternative versions for the Protein B N-terminus will yield the desired protein fusion? [15%]

(cont.

(d) For the two incorrect versions of the Protein B N-terminus from part (c), describe in detail the consequence of the fusion with reference to the codon table given at the end of the question and comment on the relative lengths of the products of translation. [20%]

(e) If only the protein A gene and version 3 of the Protein B sequence are available as clones, explain how PCR could be used to create the correct fusion. [10%]

(f) There is insufficient information supplied to allow the design of all the primers needed. Give the sequence of those that can be designed with the information given. [20%]

Codon Table

Each codon is followed by the corresponding
single-letter amino-acid code.
* indicates a stop codon.

TTT F	TCT S	TAT Y	TGT C
TTC F	TCC S	TAC Y	TGC C
TTA L	TCA S	TAA *	TGA *
TTG L	TCG S	TAG *	TGG W
CTT L	CCT P	CAT H	CGT R
CTC L	CCC P	CAC H	CGC R
CTA L	CCA P	CAA Q	CGA R
CTG L	CCG P	CAG Q	CGG R
ATT I	ACT T	AAT N	AGT S
ATC I	ACC T	AAC N	AGC S
ATA I	ACA T	AAA K	AGA R
ATG M	ACG T	AAG K	AGG R
GTT V	GCT A	GAT D	GGT G
GTC V	GCC A	GAC D	GGC G
GTA V	GCA A	GAA E	GGA G
GTG V	GCG A	GAG E	GGG G

4 Write short answers to the following questions.

(a) In broad terms, how does a gene drive system work? [10%]

(b) Why does exchange of genetic information enable evolution to proceed faster? [5%]

(c) What is codon optimisation and what problem does it seek to avoid? [10%]

(d) Describe the concept of a library in molecular bioengineering giving a specific example. [10%]

(e) Why are hybridomas important? [5%]

(f) Why is it necessary to humanise therapeutic antibodies? [5%]

(g) What is a substitution matrix as used in protein sequence alignment? [10%]

(h) Why do affine gap penalties give better protein sequence alignments? [10%]

(i) What issues need to be taken into consideration when using DNA as a data store? [10%]

(j) The three stop codons are TGA, TAG and TAA. Design a 12 base sequence that has a stop codon in all three reading frames on both strands. [5%]

(k) Describe the relationship between genome size and complexity of organisms. [5%]

(l) Sketch the various steps in an irreversible strand exchange reaction. [10%]

(m) What change has to be made to the system in part (l) above to enable the reaction, once completed, to be reversed? [5%]

END OF PAPER