

Homework 4

1. You have designed an advanced biofuel production pathway in *E. coli* MG1655 using a recombinant plasmid. This plasmid includes:

- A high copy pUC origin
- A streptomycin resistance cassette flanked by FRT sites oriented to the right.
- Regulatory genes for expression via aTc (promoter, TetR, terminator, RBS, etc)
- And your full biofuel production pathway in a single operon, whose individual parts have been validated to work

While you are able to produce biofuels, you notice that the construct is unstable, and quickly accumulates mutations that inactivate the pathway. Thus, you move your plasmid to DH10B as its genotype is more amenable to plasmid stability.

- a) After transformation, you isolate a large number of DH10B colonies. However, none produce biofuel. Make a hypothesis for the problem and propose an experiment to test it (5 points).

The plasmid was lost because it lacked selection pressure as DH10B plasmids already have streptomycin resistance. The new plasmid did not provide a novel form of selection pressure and thus was not produced.

- b) Your test successfully diagnoses the issue with the plasmid. You quickly generate a replacement part via PCR that is formatted to resolve the problem when combined with your defective plasmid and a single enzyme. What specific enzyme would you use and how would you format the replacement part? (5 points)

I would use recombinase to perform an insertion between specific sites by putting FRTs on each side of the part. With PCR, you can create hanging ends that match the FRT sites to then insert the replacement part.

2. While implementing a biofuel pathway in *E. coli*, you decide to use alcohol dehydrogenase from yeast as it may have superior kinetics.

- a) What issues must you consider when preparing your gene sequences for *E. coli* expression (2 points)

You must consider the codon frequency differences between *E. coli* and the yeast to ensure the best codon availability as well as ensuring that the sequence does not have restriction enzyme recognition sites.

- b) Given the first 7 yeast codons below, indicate any modifications you would make to the sequence. Justify your choices by showing your work (8 points)

5' ATG TCT ATC CCA GAA ACT CAA 3'

5' START S I P E T Q 3'

The start codon is the same for all organisms so that will not change.

5' ATG TCT ATC CCA GAA ACT CAA 3'

In *E. coli*, the most common codon for S is AGC (0.25), so TCT will be replaced with this.

5' ATG AGC ATC CCA GAA ACT CAA 3'

In *E. coli*, the most common codon for I is ATT (0.49), so ATC will be replaced with this.

5' ATG AGC ATT CCA GAA ACT CAA 3'

In *E. coli*, the most common codon for P is CCG (0.49), so CCA will be replaced with this.

5' ATG AGC ATT CCG GAA ACT CAA 3'

In *E. coli*, the most common codon for E is GAA, so this will not change.

5' ATG AGC ATT CCG GAA ACT CAA 3'

In *E. coli*, the most common codon for T is ACC (0.40), so ACT will be replaced with this.

5' ATG AGC ATT CCG GAA ACC CAA 3'

In *E. coli*, the most common codon for Q is CAG (0.66), so CAA will be replaced with this.

5' ATG AGC ATT CCG GAA ACC CAG 3'

When the sequence is run through a restriction enzyme recognition site finder, none of the common restriction enzymes came up as having recognition sequences within this sequence, so this sequence is acceptable. (Bpu10I and BsmI were identified).

5' ATG AGC ATT CCG GAA ACC CAG 3'

3. You introduce a series of mutations into a chromosome via λ -RED mediated recombineering. For each mutation, you use a marker flanked by FRT sites for selection and subsequently remove it before the next integration. After successive rounds, you notice that mutations that are near to each other (< 30kb) on the chromosome appear to be neither wildtype nor the introduced mutation.
 - a) Propose a hypothesis for what has gone wrong (2 points).

The λ -RED mediated recombineering is only efficient for insertions with an upper bound of 10-20 kb. The insertion must be greater than 20 kb, thus being vulnerable to extra unwanted mutations.
 - b) Propose a strategy to introduce these mutations that does not suffer from this issue (2 points)

The CRISPR/Cas9 method is more efficient, assuming that the chromosome has PAMs near the desired insertion area.

4. You are engineering a pathway that branches from glycolysis. To optimize flux, you decide to rewire glycolysis by deregulating *glk*, the first committed step of glycolysis*.
- a) Draw a map of *glk* and its native regulators as presented on EcoCyc. (2 points)



- b) To deregulate this gene, what features must be altered? (2 points)
Either the two promoters or the coding sequence must be altered/deleted.
- c) Prepare a step-by-step protocol of how you would alter this area via recombineering without leaving a selection marker behind at the end. Describe sequences of any oligos that you would order (5 points)
As we do not want to leave an antibiotic selection marker behind, we must select two regions of homology before and after where we want to use site-specific recombinases. In this case, these regions of homology are immediately before and after the promoters for the *glk* gene (highlighted). In our donor sequence, we must have two FRT sequences facing the same direction between our regions of homology to create an insertion. In between those, we must insert a nonsense sequence to prevent transcription of the gene.
 *first step under selected conditions
5. DESIGN PROJECT (2 points)
- a) Describe the specific parts needed to execute your project. What are they? How many of them are they? What circuits will they form?

We will have four coding regions, but three promoters, RBSs, and terminators.

The first gene cassette will have a promoter sensitive to polio and will produce *lacI* if polio is not present.

If *lacI* is produced, it will inhibit OFP (orange fluorescent protein) and another operator (TBD) production by the second cassette.

The final cassette will produce OFP (blue fluorescent protein) due to its promoter's sensitivity to concentrations of a small molecule such as calcium. If the small molecule is present, the cassette is activated. However, it is also inhibited by the second operator (TBD) production by the second cassette.

- b) Discuss your assembly strategy for constructing the circuits discussed above.

We plan to use Golden Gate assembly to build the circuits with four parts for each coding domain. If we put the promoters and RBSs together for each circuit, we have exactly ten parts, enough to perform only one Golden Gate assembly. We plan to check for restriction enzyme recognition sites within each part prior to performing the assembly.