CHEMENG 355 Spring 2018

Problem Set 1 Due April 19, 2018

Question 1 – Rates of transcription and translation

In class we discussed the importance of manufacturing the human peptide hormone insulin using a recombinant host. One way to make insulin in *Escherichia coli* mentioned in class is to produce the two subunits separately and then combine them *in vitro* (*i.e.* outside the cell). Another method involves production of insulin in *E. coli* as a single peptide and then digestion of proinsulin after it is isolated to yield the final form of the peptide hormone with the two subunits linked by disulfide bonds. The heterologous production of proteins in *E. coli* commonly features a promoter borrowed from the T7 bacteriophage (a virus that infects *E. coli*), which when turned on can cause a protein of interest to account for up to half of all the mRNA transcripts and translation products in a cell.

Let's examine insulin production in an *E. coli* batch fermentation (using the method where insulin is produced as a single peptide) and assume the culture is saturated and cellular growth rate is negligible (OD_{600} = 10, 5 liter culture volume; OD_{600} of 1.0 for *E. coli* corresponds to 5 x 10^8 cells per ml). Further estimate that ATP "usefulness" is only 50% (BNID 102605) - that is, half of all ATP made in the cell is lost in membrane potential buildup and leakage - and that this ATP is made by 3000 ATP synthases per cell (von Meyenburg, 1984).

A. Using estimates for transcription and translation energetics in bacteria discussed in lecture, calculate the minimum time required to produce a day's worth of insulin for a single patient (68 units of insulin per day, where 1 unit of insulin is defined as 34.7 ug of pure crystalline insulin) if ATP availability limits the total amount of insulin that can be made. Assume that each ATP synthase can produce 100 ATP/second.

B. From a biotechnological standpoint, does this calculated time make sense? Do you think it is reasonable to assume ATP availability alone limits the rate of protein synthesis? Why?

Let us instead analyze the production on a transcriptionally-limited basis. Assume that transcription of insulin is induced at stationary phase (such that every cell in the culture is making protein); mRNA transcripts have an average lifetime of 8 minutes; and each recombinant *E. coli* cell contains 60 copies of the insulin-gene-harboring expression plasmid. Further assume that translation and transcription initiation rates are maximal: Translation occurs at 48 nucleotides per second, with 120 nucleotides between each ribosome on the mRNA on average. Transcription occurs at 50 nucleotides per gene per second.

C. Keeping in mind the original culture parameters above, how many transcripts per cell are required to synthesize a day's worth of insulin for a single patient? How does this compare to the average number of mRNA transcripts per gene in the *E. coli* genome?

D. In the transcriptionally-limited regime, what is the length of time required for the culture to produce a day's worth of insulin for a single patient? Is insulin production likely limited by ATP availability or transcription?

Question 2 – Engineering translation

In recent years, several research groups have undertaken efforts to engineer *E. coli* and yeast to produce proteins that contain unnatural amino acids, reasoning that these additional amino acids could endow proteins with new and useful functions (e.g., in catalysis and imaging). The most commonly used strategy is to 'expand the genetic code' by using one of the three stop codons to encode the unnatural amino acid.

A) What are the two biomolecules that need to be engineered for such an effort? *Hint: one is an enzyme and the other is a nucleic acid.*

B) What are two key challenges of engineering the enzyme from part A starting from a native variant of the same enzyme? One should address its ability to function as desired, and the other should address the potential problems of introducing a new enzyme of this sort into a cell.

C) Bonus: Add a third key challenge that is a more general concern about the potential effects on the translation of other cellular proteins.

Question 3 – Making and using ATP

One of the most fascinating aspects of using cells as biocatalysts is that they can self-replicate. For example, ATP synthase produces ATP, which in turn is required for its own production. However, cellular metabolism of growth and protein production comes at an energetic cost that competes with product formation in an engineered strain.

a) Using data provided in the Weissman paper (Cell 157, 624–635, April 24, 2014), estimate how long it takes a cell to produce 1000 new ATP synthases. Assume that a generation time for an *E. coli* cell is 20 minutes.

b) How long does an ATP synthase need to operate to generate enough ATP to pay for the cost of the protein translation that generated it? (Assume an ATP synthase turnover rate of 100 ATP/second)

c) In E. coli, ATP Synthase is composed of several subunits (AtpA-AtpH). The subunits are all produced from a single mRNA transcript but are not needed in a 1:1 ratio. From Figure 2 of the paper, which subunit is the most abundant in ATP synthase? Are the subunits synthesized in a 1:1 ratio, or does protein synthesis more accurately reflect the stoichiometry of the subunits in the ATP synthase complex? What mechanism is likely to account for any differences in synthesis rate?