Quantitative principles in biological systems
Problem Set #2
Due by noon on 2025/04/16

1. Growth: Let's analyze omics data to explore how bacteria balance the competing demands of maximizing metabolic fluxes while minimizing proteome costs. The file `growth_data.xlsx`[1] contains data for key reactions in *E. coli* central metabolism, including reaction fluxes, enzyme concentrations, catalytic rates, and equilibrium constants. Data were obtained in similar experimental conditions.
    a. The flux through a reaction depends on enzyme and metabolite concentrations. Write down an equation to describe some of the possible relationships among these variables.
    b. A given set of fluxes can be realized by different combinations of enzyme and metabolite concentrations. Given the fluxes in the datasheet, make a parameter-free prediction for the corresponding enzyme levels. Use similar assumptions as in the work on multi-objective optimality of fluxes by Schuetz et al.
    c. Make a more refined prediction following the MetE example in Gene-Wei Li et al. Did your predictions improve after using more parameters? Why or why not? Interpret your predictions with respect to the metabolic pathways map.
    d. Describe how you might incorporate thus far unused parameters in your predictions. What hypotheses will you be testing by doing so?
    e. An important and learnable research skill is the ability to identify and appreciate open questions, even in fields unfamiliar to you. Consider the following two papers on growth rate shifts: [#1, https://doi.org/10.1038/nature24299] and [#2, https://doi.org/10.1038/s41586-020-2505-4]. What are some connections between these two papers? What do you think about the models and their similarities or differences?

2. Lac repressor: [Problem 19.2 in *PBoC*.] Data for this problem can be found at the *PBoC* website (Figs. 1.11, 19.22, and 19.27 at https://www.rpgroup.caltech.edu/pboc2/#/Figures). These data from Oehler et al describe the repression of *lacZ* expression in *E. coli* mutants differing in the number, sequence, and position of the Lac operator sites.
    a. Using the data for Fig. 19.22, determine the in vivo binding energy of the Lac repressor to each of its operators and reproduce Fig. 19.23.
    b. Using the data for Fig. 19.27 – in which repression was measured in cells with two operators present, which leads to DNA looping – determine the looping energy and reproduce Fig. 19.27.
    c. Using the data for Fig. 1.11, in which the distance between operators was varied, determine the looping energy as a function of the distance between operators. Interpret your results.

3. Persisters: Growing fast is not always the best survival strategy: *Persisters* are non-growing cells that survive exposure to antibiotics otherwise lethal to genetically identical normal cells. Cells appear to stochastically switch in and out of persistence via still unclear mechanisms.
    a. We saw in class that gene expression is stochastic, so perhaps one hypothesis is that noise in metabolic enzymes can propagate to cause fluctuations in growth rate.
        i. Test this hypothesis using data from Taniguchi et al. Based on Table S6, how noisy are enzymes in central metabolism? Compared to what? Consider the Pdh module as an example.
        ii. How much growth rate variation can be generated from this amount of noise? Try to ground your answers using data or propose an experiment.
    b. Another commonly proposed mechanism is the toxin-antitoxin motif, where a cell expresses two genes, one for a toxic protein inhibiting growth and another for an antitoxin that binds to and sequesters the toxic protein.
        i. Suppose that toxin T and antitoxin A are expressed constitutively with rate r_T and r_A, respectively. Further assume that T and A immediately and irreversibly bind together, and that all reactants are degraded at the same rate $d \ll r_T, r_A$ to model dilution by growth. Implement this model using the Gillespie algorithm for a range of r_T/r_A. Plot simulated trajectories of T and A over time.

---

[1] Fluxes from van Rijsewijk et al. *Mol Syst Biol* (2011). Concentrations from Schmidt et al. *Nat Biotechnol* (2016). Rate constants from BRENDA. Equilibrium constants from eQuilibrator.

ii. Explore the parameter space. Define non-inhibited periods as when there are zero T molecules. When do you get short non-inhibited periods separated by long periods of inhibition? eg Fig. 1B in Lord et al. *Science* (2019)[2].

iii. Simulate and analyze the deterministic version of this model and compare to the stochastic version. How many steady states are there in either case?

---

[2] DOI: 10.1126/science.aaw4506