CHEME 5440/7770: Take Home Prelim 1 S2019

---

1. Take Home Prelim 1 has three questions which are collectively worth 100 points.
2. Take Home Prelim 1 is due at the beginning of class on T March 26, 2019.
3. You may use your course notes, literature, the internet, or any other course materials to formulate your solutions.
4. You *cannot* consult with any other person regarding the prelim (except the TA, JV or MP). You *cannot* use any form of electronic communication to discuss the prelim questions with any other person (except the TA, JV or MP via a direct message in Slack). Violation of this policy will result in a ZERO for the prelim, and an honor code violation.
5. Mistakes/corrections/clarifications to the prelim will be made on the #general Slack channel by the TA, JV or MP.
6. In all problems, show your work and state all assumptions or simplifications. Start from the general, and work your way to the specific.
7. **Submission**: Submit a link to the #prelim-1 channel on Slack that points to your solutions. Your solutions should include all written material, source code, and instructions to reproduce your calculations/figures.

---

1. (20 points). Derive an expression for the *kinetic limit* of transcription for gene j ($r_{X,j}$) in a set of $\mathcal{N}$ genes. Assume the same four elementary steps proposed in class:

$$\begin{aligned} \mathcal{G}_j + R_X^\circ &\rightleftharpoons (\mathcal{G}_j : \mathrm{R_X})_C \\ (\mathcal{G}_j : \mathrm{R_X})_C &\longrightarrow (\mathcal{G}_j : \mathrm{R_X})_O \\ (\mathcal{G}_j : \mathrm{R_X})_O &\longrightarrow R_X + \mathcal{G}_j \\ (\mathcal{G}_j : \mathrm{R_X})_O &\longrightarrow m_j + R_X + \mathcal{G}_j \end{aligned}$$

where $\mathcal{G}_j$, R_X° denote the gene and *free* RNAP concentration, and $(\mathcal{G}_j : \mathrm{R_X})_O$, $(\mathcal{G}_j : \mathrm{R_X})_C$ denote the open and closed complex concentrations, respectively. Let the kinetic limit of transcription be directly proportional to the concentration of the open complex:

$$r_{X,j} = k_{E,j} \left(\mathcal{G}_j : \mathrm{R_X}\right)_O$$

where $k_{E,j}$ is the elongation rate constant for gene j.

a) (15 points). Starting from the proposed elementary steps and the RNAP balance:

$$R_{X,T} = R_X^\circ + (\mathcal{G}_j : \mathrm{R_X})_C + (\mathcal{G}_j : \mathrm{R_X})_O + \sum_{i=1,j}^{\mathcal{N}} \Big\{ (\mathcal{G}_i : \mathrm{R_X})_C + (\mathcal{G}_i : \mathrm{R_X})_O \Big\}$$

show that:

$$r_{X,j} = k_{E,j} R_{X,T} \left(\frac{\mathcal{G}_j}{\tau_{X,j} K_{X,j} + (1 + \tau_{X,j}) \mathcal{G}_j + \mathcal{E}_j} \right)$$

where:

$$\mathcal{E}_j = \sum_{i=1,j}^{\mathcal{N}} \frac{K_{X,j} \tau_{X,j}}{K_{X,i} \tau_{X,i}} (1 + \tau_{X,i}) \mathcal{G}_i$$

The saturation and time constants are defined as $K_{X,j}^{-1} \equiv k_{+,j} / (k_{-,j} + k_{I,j})$ and $\tau_{X,j}^{-1} \equiv k_{I,j} / (k_{A,j} + k_{E,j})$, respectively.

b) (5 points). Under what circumstances would an $\mathcal{N}$-gene system ($\mathcal{N} \gg 1$) be approximately equivalent to the 1-gene system we derived in class?

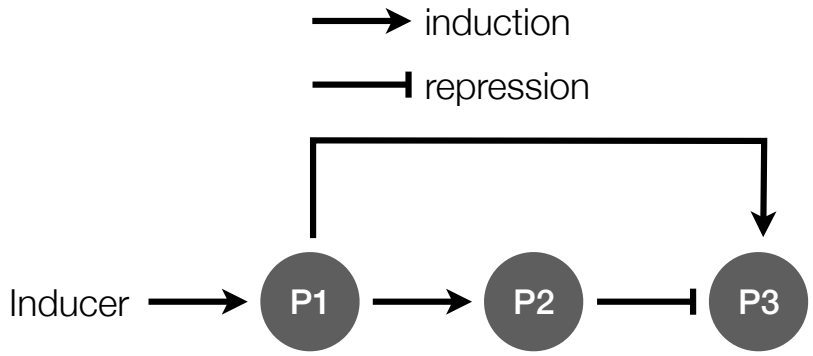

Fig. 1: Schematic of the type 1 incoherent feed forward loop. Inducer induces the expression of protein 1 (P1) which in turn induces P2 and P3. However, P2 represses the expression of P3.

2. (50 points). Analyze a Type 1 incoherent feed forward loop (Fig. 1). Alon and coworkers analyzed the structure of gene regulatory networks in *E. coli* (1). Let's analyze a common motif, the Type 1 incoherent feed forward loop, in a growing population of *E. coli* cells with a doubling time of τ_d = 40 min and the promoter/binding parameters given in Table(s) 1 and 2. Use nmol/gDW as the concentration system for the mRNA and proteins in the motif.

   Assume: (i) the motif is encoded on a plasmid present at 200 copies per cell (constant); (ii) characteristic lengths of $\mathcal{L}_X$ = 1000 nt and $\mathcal{L}_T$ = 333 AA; (iii) the promoter control models follow the Moon/Voigt formulation; (iv) translation operates at the kinetic limit; (v) The copy numbers of RNAP (4600 copies/cell) and Ribosome (50000/cell) are constant; (vi) $\mathcal{L}_{X,1}$ = 1200 nt, $\mathcal{L}_{X,2}$ = 2400 nt and $\mathcal{L}_{X,3}$ = 600 nt. Let $\mathcal{L}_{L,i} \simeq (1/3)\mathcal{L}_{X,i}$; (vii) an *E.coli* cell weighs 2.8×10^{-13} g/cell and is 70% water; (viii) the half-life of mRNA equals 2.1 min, while the protein half-life is 24 hrs; (ix) the RNAP elongation rate $\dot{v}_X$ = 60 nt/s, while the Ribosome elongation rate $\dot{v}_L$ = 16.5 aa/s; (x) K_X = 0.24 nmol/gDW and K_L = 2.4 nmol/gDW.

   a) (10 points). Simulate the response to 10 mM Inducer using the discrete modeling approach from `PS2` with a time step of 1.0 min using the protocol: i) run the model to steady-state without inducer; ii) from the steady-state, run the model without inducer for an additional 60 min (Phase 1); ii) add inducer and run the model for 300 min (Phase 2). Plot P1, P2 and P3 versus time.

   b) (20 points). Compute the scaled state sensitivity coefficients as a function of

time:

$$s_{ij}(t) = \left(\frac{p_i}{x_i}\right)_\star \frac{\partial x_i}{\partial p_j}\Big|_t \tag{1}$$

by approximating the partial derivative using a *central difference* for all states and model parameters (p_j) for a window in Phase 1 and early/late Phase 2. These time windows should be approximately 20 min in length. **Note**: the $(\cdot)_\star$ scaling term is evaluated along the unperturbed trajectory at the same time point as derivative.

c) (20 points). Rank-order the importance of model species using Singular Value Decomposition (SVD) of the time-averaged sensitivity array(s) for Phase 1, and early/late Phase 2. Use the absolute magnitude of the first column of the **U** matrix as the basis for the ranking in each of the three time windows. Approximate the integrals in the time-averaged sensitivity array using the trapezoidal rule. Both `Julia` and `MATLAB` have `svd` implementations (in `Julia`, the `svd` is in the `LinearAlgebra` package). Explain the rankings and any shifts in the rankings during the time windows.

Table 1: Promoter weights for the Type 1 incoherent feed forward loop problem. The terms W_{ij} denote the influence of protein i on the expression of protein j (from $i \rightarrow j$).

Protein	Weight	Value
Inducer	W_{I1}	100
P1	W_{11}	0.01
	W_{12}	10.0
	W_{13}	5.0
P2	W_{22}	0.01
	W_{23}	25.0
P3	W_{33}	0.01

Table 2: Promoter binding parameters for the Type 1 incoherent feed forward loop problem. The binding function: f = (effector^n)/(k^n+effector^n).

Effector	Target	Symbol	Value	Units	Description
Inducer	P1	k	0.30	mM	Binding parameter
Inducer	P1	n	1.5	-	Binding order
P1	P2	k	1.0	nmol/gDW	Binding parameter
P1	P2	n	1.5	-	Binding order
P1	P3	k	1.0	nmol/gDW	Binding parameter
P1	P3	n	1.5	-	Binding order
P2	P3	k_b	1.0	nmol/gDW	Binding parameter
P2	P3	n	10.0	-	Binding order

3. (30 points). Estimate the resource requirements for cell free protein synthesis (CFPS) using flux balance analysis (FBA). Simulate the expression of a peptide of length 308 aa encoded by a gene of length 924 nt in an *E. coli* cell free extract under the conditions shown in Table 3.

   Assume: (i) the peptide is encoded on a plasmid present at 5nM in the cell free reaction mixture (constant); (ii) the cell free reaction volume is 15μL; (iii) protein expression is induced by inducer I (see Table 3 for parameters); (iv) translation operates at the kinetic limit; (v) assume the exchange reactions $(\star)[e] \longleftrightarrow (\star)$ (where $(\star)[e]$ denotes an external species) are reversible with bounds $-100000.0 \leq b_\star \leq 100000.0$ μM/hr.

   a) (10 points). Construct a stoichiometry matrix **S** from the reaction scheme shown in Table 1 of Allen and Palsson (2). Formulate expressions for the rate of transcription $\hat{r}_X$ and translation $\hat{r}_L$ in the cell free reaction. These rates will act as bounds in the FBA calculation; let `R2` in the Allen paper be equal to $\hat{r}_X$, while `R5` be bounded by $0 \leq v_5 \leq \hat{r}_L$.

   b) (10 points). Maximize the translation rate (v_5) for I = 0.0001 mM to I = 10.0 mM. Let W_1 = 0.26, W_2 = 300.0, K = 0.30 mM and n = 1.5. Plot the steady-state protein level versus inducer concentration I on a `semilogx` scale.

   c) (10 points). Which exchange flux bounds is the translation rate most sensitive to? (**Hint**: Look up the concept of a shadow price).

Table 3: Parameters for sequence specific flux balance analysis

Description	Parameter	Value	Units
RNA polymerase concentration	R_X	0.15	μM
Ribosome concentration	R_L	1.6	μM
Transcription elongation rate	$\dot{v}_X$	60	nt s^{-1}
Translation elongation rate	$\dot{v}_L$	16.5	aa s^{-1}
Saturation constant transcription	K_X	0.3	μM
Saturation constant translation	K_L	57.0	μM
Time constant transcription	τ_X	2.7	dimensionless
Time constant translation	τ_X	0.8	dimensionless
degradation constant mRNA	$k_{d,X}$	8.35	h^{-1}
degradation constant protein	$k_{d,L}$	9.9×10^{-3}	h^{-1}
characteristic gene length	$\mathcal{L}_X$	1000	nt
characteristic protein length	$\mathcal{L}_T$	330	aa

References

1. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet. 2002;31(1):64–8. doi:10.1038/ng881.
2. Allen TE, Palsson BØ. Sequence-based analysis of metabolic demands for protein synthesis in prokaryotes. J Theor Biol. 2003;220(1):1–18.