CHEME 5440/7770: Take Home Prelim 1 S2022

---

1. Take Home Prelim 1 has three questions which are collectively worth 60 points.
2. Take Home Prelim 1 is due on M March 28, 2021 by 4:59 PM Ithaca time
3. You may use your course notes, literature, the internet, or other course materials to formulate your solutions.
4. You *cannot* consult with any other person regarding the prelim (except the teaching team). You *cannot* use any form of electronic communication to discuss the prelim questions with any other person (except the teaching team 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 document will be made on the #general Slack channel by the teaching team.
6. In all problems, show your work and state all assumptions or simplifications and **clearly state your answers**.
7. You can use any computing tools, languages, etc to formulate your solutions.
8. **Submission**: Submit your solution to the teaching team email by the deadline. Your solution should include all written material, links to source code if posted on GitHub or the source code itself, and instructions to reproduce your calculations/figures.

---

1. (20 points). Derive the *kinetic limit* of transcription for gene j ($r_{X,j}$) in a set of $\mathcal{N}$ genes, where transcription occurs via the four elementary steps:

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

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. The second elementary step describes initiation (closed to open complex). The third elementary step accounts for the abort mechanism, where transcription produces an incomplete truncated message $m_j^\star$ that is not translated. Finally, the fourth elementary step accounts for elongation.

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.

Questions:

a) (16 points) Starting from the proposed elementary steps and the RNAP balance:

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

where the $\sum_{i=1,j}^{\star}$ notation denotes a summation excluding the index j, 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:

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

The saturation and time constants are defined as $K_{X,j}^{-1} \equiv k_{+,j} / \left(k_{-,j} + k_{I,j}\right)$ and

$\tau_{X,j}^{-1} \equiv k_{I,j}/\left(k_{A,j}+k_{E,j}\right)$, respectively, where $k_{\pm,j}$ denote the rate constant(s) governing the on- and off-rate of RNAP for gene j, $k_{I,j}$ denotes the rate constant for initiation (elementary reaction 2), $k_{A,j}$ denotes the rate constant for RNAP abort (elementary reaction 3) and $k_{E,j}$ denotes the rate constant for elongation (elementary reaction 4).

b) (4 points) For the single gene case, with a negligible RNAP abort rate, if $\tau_{X,j} \ll 1$, is transcription initiation or elongation limited?

Table 1: PFK rate measurements as a function of AMP concentration.

3′-5′-AMP (mM)	mean rate (μM/h)	standard deviation (μM/h)
0.0	3.0	0.59
0.055	6.3	1.20
0.093	29.8	5.7
0.181	52.0	10.2
0.405	60.3	11.8
0.990	68.7	13.3
1.0	68.9	10.0

2. (20 points). Develop an expression for the allosteric regulation of Phosphofructokinase (PFK) using a discrete state regulatory model. PFK catalyzes the conversion of D-fructose 6-phosphate (F6P) to D-fructose 1,6-bisphosphate (F16BP):

$$\mathrm{ATP} + \mathrm{D-fructose\ 6-phosphate} \longrightarrow \mathrm{ADP} + \mathrm{D-fructose\ 1,6-bisphosphate} \quad (1)$$

PFK activity is strongly activated in the presence of 3′-5′-AMP, a signalling molecule produced when glucose is transported into cells.

Assume: (i) the concentration of F6P in the assay equals 0.1 mM and is constant; (ii) the concentration of ATP in the assay equals 2.3 mM and is constant; (iii) the concentration of PFK in the assay equals 0.12 μM and is constant; (iv) K_{F6P} = 0.11 mM, K_{ATP} = 0.42 mM, and k_{cat} = 0.4 s^{-1}.

Questions:

Let the model take the form $\hat{r}_j = r_j v\left(\ldots\right)_j$ where $\hat{r}_j$, the overall rate of the PFK reaction (μM h^{-1}), is the product of a kinetic limit r_j (μM h^{-1}) and a control variable $0 \leq v\left(\ldots\right)_j \leq 1$ (dimensionless) that describes the influence of effector molecules.

a) (5 points) Formulate a three micro-state model for PFK activity; State 0 (base): no effector+no substrate (no activity), State 1: no effector+substrate (what does

the data say?) and State 2: effector+substrate (activity). The probability of each microstate p_i is given by:

$$p_i = \frac{W_i f_i}{Z} \tag{2}$$

where Z denotes the partition function, W_i denotes the weight of configuration i, and f_i denotes the accessibility of state i. Let's assume the accessibility factor $f_i \in [0, 1]$ is either set to 1 or is given by hill-type binding function $f_i = (x/K_i)^{n_i}/(1 + (x/K_i)^{n_i})$ which describes the fraction of bound activator/inhibitor (x) for configuration i; K_i denotes a binding constant (mM), and n_j denotes an order parameter (dimensionless). Finally, let the kinetic limit for the PFK reaction be given by:

$$r_1 = k_{cat} E_1 \left(\frac{F6P}{K_{F6P} + F6P} \right) \left(\frac{ATP}{K_{ATP} + ATP} \right) \tag{3}$$

where the concentration of the PFK enzyme is denoted by E_1 (units: μM).

b) (10 points) Estimate the parameter(s) W_0 (base), W_1 (no 3′-5′-AMP), W_2 (with 3′-5′-AMP), the binding constants and order parameters appearing in f_i using the experimental dataset in Table 1. **Note**: this can be done analytically, but need not be.

c) (5 points) Plot your estimated overall rate $\hat{r}_1$ (y-axis), along with the measured rate (with errorbars), versus the 3′-5′-AMP concentration (x-axis) on the same axes. How well does the proposed model describe the data?

Table 2: Average messenger RNA (lac_Z) copy number $< n >$ per cell for the P_{lac} promoter as a function of extracellular inducer I = IPTG.

IPTG (mM)	$< n >$ (mRNA/cell)	low (mRNA/cell)	high (mRNA/cell)
0.0	19	18	20
5e-4	21	17	26
0.005	41	37	44
0.012	67	65	69
0.053	86	84	88
0.216	93	91	95
1.0	93	92	94

3. (20 points). Golding and coworkers measured the average mRNA copy number per cell for several promoters using single-molecule fluorescence in situ hybridization (smFISH) in single dividing *E. coli* cells (1).

   Assume: (i) all experiments were conducted in a well-mixed exponentially growing population of *E. coli* cells with a doubling time of $\tau_d \simeq$ 40 min; (ii) 1 molecule per cell is equivalent to an intracellular concentration of 1 nM; (iii) the average mass of an *E. coli* cell $< m_c >$ = 2$\times 10^{-13}$ g; (iv) the average volume of an *E. coli* cell $< V_c >$ = 2.75 μm^3; (v) an *E. coli* cell is 70% water; (vi) write the promoter function in terms of extracellular inducer (ignore inducer transport); (vii) the lacZ gene is present at two copies/cell; (viii) the lacZ mRNA half-life is 5 minutes; (ix) lacZ is 3075 base pairs in length; (x) RNAP polymerase has a transcription rate e_X=35 nt/s and is present at 4,600 copies/cell.

   Questions:

   a) (3 points) Convert the $< n >$ values in Table 2 to cell mass specific concentration units nmol/gDW.

   b) (2 points) Derive the gain function $\mathcal{K}_X$ and formulate $\bar{u}_i$. Starting from the

mRNA balance:

$$\dot{m}_i = r_{X,i}\bar{u}_i - (\mu + \theta_{m,i})m_i \tag{4}$$

derive the steady-state mRNA abundance $m^\star$:

$$m^\star = \mathcal{K}_X\left(G, \dots\right)\bar{u}(I, \kappa) \tag{5}$$

where $r_{X,i}\bar{u}_i$ (units: nmol/gDW-time) denotes the specific rate of transcription of gene i (production rate of mRNA i), $\bar{u}_i$ denotes the promoter activity function describing both regulated and unregulated gene expression activity, μ denotes the specific growth rate (units: 1/time), $\theta_{m,i}$ denotes the mRNA degradation constant (units: 1/time), G denotes the lacZ gene abundance, I denotes the extracellular IPTG inducer abundance, and κ, θ denotes parameter vectors (various kinetic and promoter model constants). Assume the specific kinetic limit of transcription $r_{X,i}$ takes the form:

$$r_{X,i} = V_{max,i}\left(\frac{G_i}{K_{X,i} + G_i}\right) \tag{6}$$

and $\bar{u}_i$ is formulated as a three state discrete promoter model. To make the problem easier, assume P_{lac} is a positively inducible promoter that responds to IPTG with the states:

- State 0: Base state, just gene G_i. No transcription possible.
- State 1: RNAP bound to G_i at I = 0. What does the data say?
- State 2: RNAP + inducer bound to G_i. Transcription possible.

c) (10 points) Use the data in Table 2 (converted to the correct units), to estimate the discrete state promoter model parameters in $\bar{u}_i$, and the gain $\mathcal{K}_X\left(G, \dots\right)$ such that $m^\star$ is consistent with the measured copy number as a function of IPTG concentration. If any additional information is required, estimate any required values and justify your estimate.

d) (5 points) Plot (on the same axes) the converted data, and the estimated average copy number from your model as a function of IPTG, does the model fit?

References

1. So LH, Ghosh A, Zong C, Sepúlveda LA, Segev R, Golding I. General properties of transcriptional time series in Escherichia coli. Nat Genet. 2011;43(6):554–60. doi:10.1038/ng.821.