Preliminary Exam I

PROBLEM I

Four elementary steps:

$$G_j + R_X^\circ \rightleftharpoons (G_j + R_X)_C$$
$$(G_j + R_X)_C \longrightarrow (G_j + R_X)_O$$
$$(G_j + R_X)_O \longrightarrow R_X + G_j$$
$$(G_j + R_X)_O \longrightarrow m_j + R_X + G_j$$

Kinetic limit of transcription:

$$r_{X,j} = k_{E,j}(G_j : R_X)_O$$

Proposed elementary steps and the RNAP balance:

$$R_{X,T} = R_X^\circ + (G_j : R_X)_C + (G_j : R_X)_O + \sum_{i=1,j}^{\mathcal{N}} \{(G_i : R_X)_C + (G_i : R_X)_O\}$$

a. Show that:

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

where,

$$\varepsilon_j = \sum_{i=1,j}^{\mathcal{N}} \frac{K_{X,j}\tau_{X,j}}{K_{X,i}\tau_{X,i}} (1 + \tau_{X,i}) G_i$$

Saturation constant:

$$K_{X,j}^{-1} \equiv \frac{k_{+,j}}{(k_{-,j} + k_{I,j})}$$

Time constant:

$$\tau_{X,j}^{-1} \equiv \frac{k_{I,j}}{(k_{A,j} + k_{E,j})}$$

Material balances around the closed and open complex for gene *j* are given by:

$$\frac{d}{dt}(G_j : R_X)_C = k_+\left(G_j\right)(R_X) - k_-(G_j : R_X)_C - k_I(G_j : R_X)_C$$
$$\frac{d}{dt}(G_j : R_X)_O = k_I(G_j : R_X)_C - k_A(G_j : R_X)_O - k_{E,j}(G_j : R_X)_O$$

At steady state, revised balance equations:

$$(G_j{:}\,R_X)_C \backsimeq \left(\frac{k_+}{k_- + k_I}\right)(G_j)(R_X)$$

$$(G_j{:}\,R_X)_O \backsimeq \left(\frac{k_I}{k_A + k_E}\right)(G_j{:}\,R_X)_C$$

$$(G_{j,i}{:}\,R_{X,i})_C \backsimeq \left(\frac{k_{+,i}}{k_{-,i} + k_{I,i}}\right)(G_{j,i})(R_{X,i})$$

$$(G_{j,i}{:}\,R_{X,i})_O \backsimeq \left(\frac{k_{I,i}}{k_{A,i} + k_{E,i}}\right)(G_{j,i}{:}\,R_{X,i})_C$$

In regards to the steady-state case, only the open complex will be taken into consideration:

$$(G_j{:}\,R_X)_O \backsimeq (K_{X,j}^{-1})(\tau_{X,j}^{-1})(G_j)(R_X)$$

Solve for the free RNAP concentration, R_X:

$$R_{X,T} = R_X^\circ + (G_j{:}\,R_X)_C + (G_j{:}\,R_X)_O + \sum_{i=1,j}^{\mathcal{N}}\{(G_i{:}\,R_X)_C + (G_i{:}\,R_X)_O\}$$

$$R_{X,T} = R_X^\circ + (K_{X,j}^{-1})(G_j)(R_X) + (K_{X,j}^{-1})(\tau_{X,j}^{-1})(G_j)(R_X) + \sum_{i=1,j}^{\mathcal{N}}\left\{(K_{X,i}^{-1})(G_i)(R_X) + (K_{X,i}^{-1})(\tau_{X,i}^{-1})(G_i)(R_X)\right\}$$

$$R_{X,T} = R_X + (K_{X,j}^{-1})(G_j)(R_X) + (K_{X,j}^{-1})(\tau_{X,j}^{-1})(G_j)(R_X) + (R_X)\sum_{i=1,j}^{\mathcal{N}}(1+\tau_{X,i}^{-1})(K_{X,i}^{-1})(G_i)$$

$$R_X = \frac{R_X}{1 + (1+\tau_{X,j}^{-1})(K_{X,j}^{-1})(G_j) + \sum_{i=1,j}^{\mathcal{N}}(1+\tau_{X,i}^{-1})(K_{X,i}^{-1})(G_i)}$$

$$R_X = \frac{R_{X,T}(\tau_{X,j}K_{X,j})}{\tau_{X,j}K_{X,j} + (\tau_{X,j}+1)G_j + \sum_{i=1,\,j}^{\mathcal{N}}\frac{K_{X,j}\tau_{X,j}}{K_{X,i}\tau_{X,i}}(1+\tau_{X,i})(G_i)}$$

Let $\varepsilon_j = \sum_{i=1,j}^{\mathcal{N}}\frac{K_{X,j}\tau_{X,j}}{K_{X,i}\tau_{X,i}}(1+\tau_{X,i})G_i$ and solve it for the open complex equation:

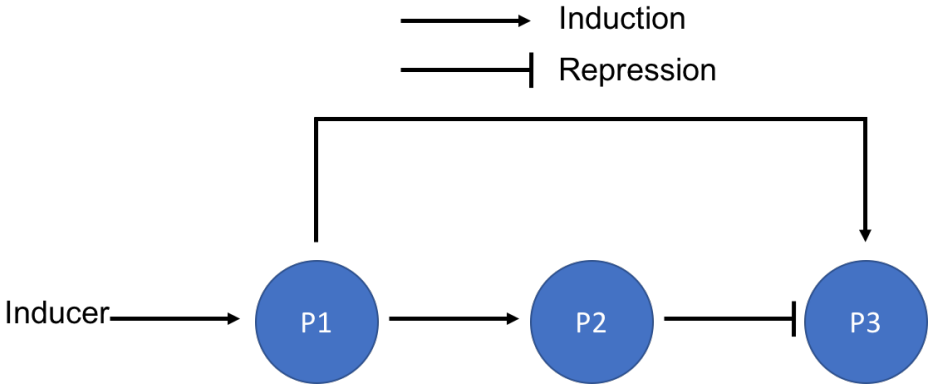
$$(G_j{:}\,R_X)_O \backsimeq \frac{R_{X,T}G_j}{\tau_{X,j}K_{X,j} + (\tau_{X,j}+1)G_j + \varepsilon_j}$$

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

b. Conditions under which N $\gg 1$ would be approximately equivalent to the 1-gene system:

The N gene system would be equivalent under elongation limited circumstances as $\tau_{X,j} \ll 1$, ε_j can degrade down to 1 eliminating the summation, thus resulting into the same 1-gene system. However, the values of i and j are relative therefore it can depend on those values.

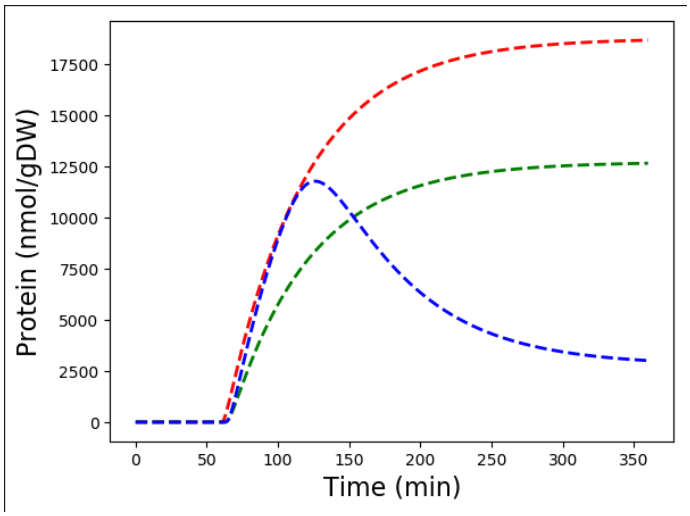
PROBLEM II



Assumptions:

- Motif is encoded on a plasmid present at 200 copies per cell (constant)
- Characteristic lengths of L_X = 1000 nt and L_T = 333 AA.
- Promoter control models follow the Moon/Voigt formulation
- Translation operates at the kinetic limit
- Copy numbers of RNAP (4600 copies/cell) and ribsome (50000/cell) are constant
- $L_{X,1}$ = 1200 nt, $L_{X,2}$ = 2400 nt,a nd $L_{X,3}$ = 600 nt
- *E. coli* weighs 2.8e-13 g/cell and is 70% water
- Half-life of mRNA = 2.1 min; protein half-life is 24 hrs
- RNAP elongation rate = 60 nt/s; ribosome elongation rate = 16.5
- K_X = 0.24 nmol/gDW and K_L = 2.4 nmol/gDW

a. Stimulate the response 10 mM inducer using the discrete modeling approach with a time step of 1.0 min using the protocol.



Run PrelimQ2.jl to generate the figure above. ICT1.json is the secondary code that the Julia file calls for.

b. Compute the scaled state sensitivity coefficients as a function of time by approximating the partial derivative using a central difference for all states and model parameters for a window in Phase 1 and early/late Phase 2.

I am not entirely sure how to code a central difference or alter a .json file to do the sensitivity test. However, if I did know how to code for it, there would be a code where I would have changed the weights of the three proteins. Then, I would plot the difference between the maximized values and original values to analyze the scaled state sensitivity coefficients

c. Rank-order the importance of model species using Singular Value Decomposition of the time-averaged sensitivity arrays for Phase 1 and early/late Phase 2. Explain the rankings and any shifts in the rankings during the time windows.

This question is dependent on the results I obtain from part b. I was unable to generate a plot or code, but if there was a working code so I would have compared the changes from part a and part b and determined which protein had the most to least impact in the forward loop.

PROBLEM III

Estimate the resource requirements for cell free protein synthesis using FBA. Stimulate 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.

$$G + RNAP \xrightarrow{v1} G^*$$
$$G^* + nNTP \xrightarrow{v2} mRNA + G + RNAP + 2nP_i$$
$$mRNA \xrightarrow{v3} nNMP$$
$$mRNA + rib \xrightarrow{v4} rib^*$$
$$rib^* + aAAtRNA + 2aGTP \xrightarrow{v5} atRNA + 2aGDP + 2aP_i + rib + mRNA + protein$$
$$AA + tRNA + ATP \xrightarrow{v6} AMP + 2P_i + AAtRNA$$
$$AA_{ext} \xrightarrow{b1} AA$$
$$NTP_{ext} \xrightarrow{b2} NTP$$
$$protein \xrightarrow{b3} protein_{ext}$$
$$NMP \xrightarrow{b4} NMP_{ext}$$
$$ATP_{ext} \xrightarrow{b5} ATP$$
$$AMP \xrightarrow{b6} AMP_{ext}$$
$$GTP_{ext} \xrightarrow{b7} GTP$$
$$GDP \xrightarrow{b8} GDP_{ext}$$
$$P_i \xrightarrow{b9} P_{i,ext}$$

Assumptions:

- Peptide is encoded on a plasmid present at 5 nM in the cell free reaction mixture (constant)
- Cell free reaction volume is 15 uL
- Protein expression is induced by inducer *I*
- Translation operates at the kinetic limit
- Assume exchange reactions are reversible with bounds

a. Construct a stoichiometric matrix. Formulate expressions for the rate of transcription and translation in the cell free reaction.

Refer to PrelimQ3_matrices.jl for the stoichiometric matrices.

b. Maximize the translation rate for I = 0.0001 mM to I = 10.0 mM. Let W1 = 0.26; W2 = 300.0, K = 0.30 mM and n = 1.5.

Refer to PrelimQ3.jl for matrices. My code kept on generating an error where it says the default bounds array is not defined even though I have defined it in PrelimQ3_matrices.jl. However, if it

were to have worked it would have generated a matrix where you can identify the maximized species in the reactions.

c. Which exchange flux bounds is the translation rate most sensitive to?

It is the most sensitive to the time constants from translation and transcription.