Michael Minehan

ChemE 7770 Problem Set 1

1a)

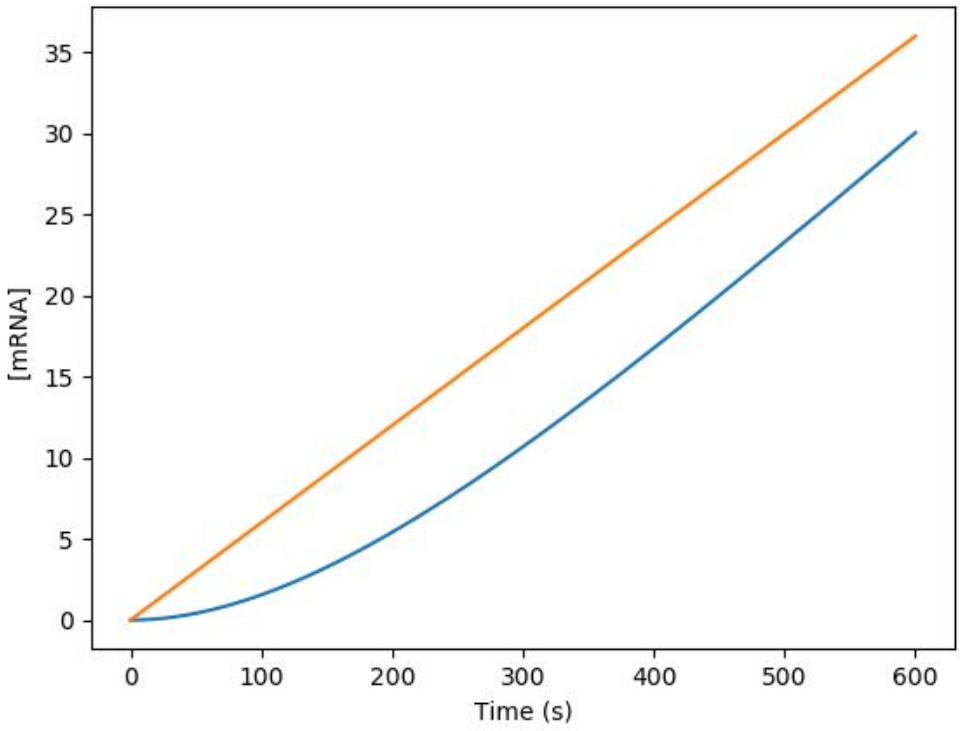
$r_{T,j} = \bar{r}_{T,j} u_j$,

$r_{T,j} = \bar{r}_{T,j}(1 - e^{-k_{ocf}*t})$, where k_{ocf} is the rate constant associated with open complex formation

$$r_{T,j} = [e_T L_T (\frac{[RNAP]}{[RNAP]_T})(\frac{L_T}{L_{T,j}})(\frac{G_j}{K_{T,j}+G_j})][(1 - e^{-\frac{k_1 k_2 [P]}{k_1[P]+k_{-1}+k_2}*t})]$$

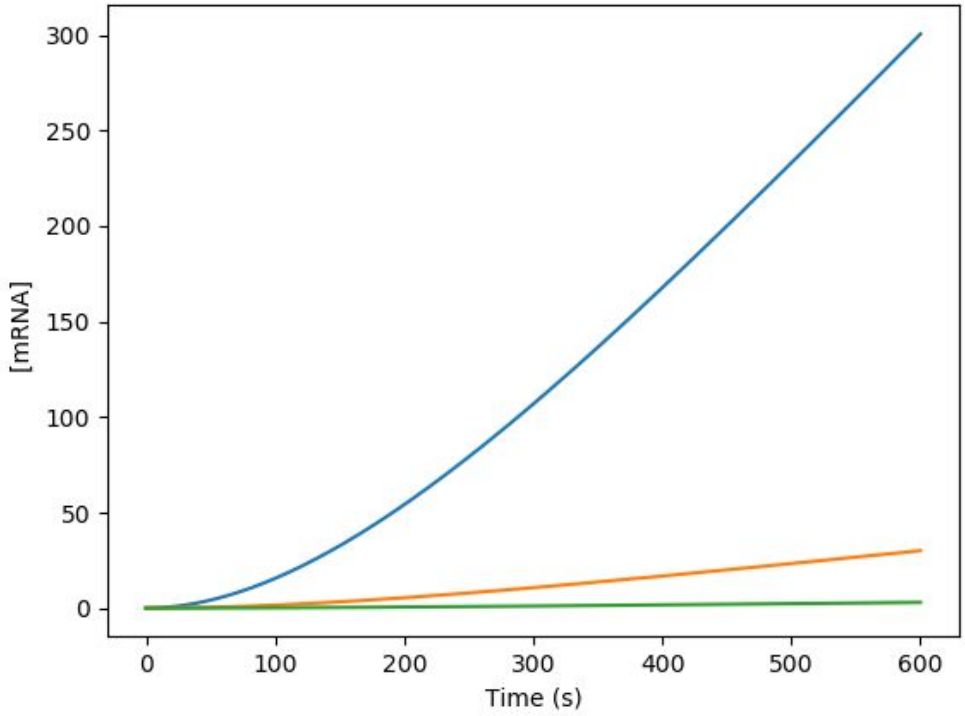
1b)

Qualitatively, the pre-incubated RNAP + DNA template + promoter had a much lower lag time in mRNA production than the unincubated condition prior to transcription initiation. This is most likely primarily due to the opportunity for open complex formation that the pre-incubation condition had. In the figure below, the orange line represents the case of RNAP pre-incubated with DNA+ 1 of 2 necessary promoter molecules, and the blue line represents the case of un-incubated condition. $r_{T,j} = \bar{r}_{T,j}(1 - e^{-k_{ocf}*t})$ was used with the default value for transcription rate, and k_ocf was set to 10^3.



1c)

The model for rate of transcription ($r_{T,j}$) incorporates gene length as a simple multiplicative factor. With a characteristic gene read length of 100bp, systems with genes of length 100bp, 10bp, and 1000bp would have relative transcription rates of 1, 10, and 0.1, respectively, assuming all other factors are equal. These changes are represented in the graph below with blue - 10bp gene, orange - 100bp, and green 1000bp. The transcription model including promoter lag was used.



2a)

According to the included rank-ordered row decomposition of the parameter sensitivity matrix, the most important set of parameters was RNAPII concentration, ribosome concentration, mRNA degradation constant, kcat_transcription, kcat_translation, maximum specific growth rate, saturation constant_transctiption, and saturation constant_translation.

2b)

The parameters estimable from only [mRNA]_1 and [protein]_3 data using an epsilon of 0.01 were RNAPII concentration, and kcat_translation. The parameters estimable from all of the species using the same epsilon were the control parameter associated with gene_1 and RNAP, the control parameter associated with gene_1 and gene_2, kcat_transcription, and kcat_translation, respectively.

3a)

In order to determine the biochemical connections that most closely approximate the true model, the following strategy will be implemented:

1. Gene connectivity models (including inducer) will be generated from default parameters based on intuition.
2. Sensitivity analysis will be performed on each model, and the key parameters identified in order to save computation time in Step 3. (NOT IMPLEMENTED)
3. The key parameters identified will be optimized via perturbed randomly and an error function finding the sum of the squared error of the three protein concentrations will be minimized to find the optimal parameter combinations.
4. The error of the parameter-optimized model will quantified and stored in a model_error.dat file.
5. Each model_error.dat will be read and used to rank the models based on accuracy.

3b)

Out of the models I tested, I found the following logic to minimize error with the provided protein levels following parameter optimization:

Inducer activates Gene 1
Gene 1 induces Gene 2
Gene 1 inhibits Gene 3
Gene 2 inhibits Gene 3

The second-best was

Inducer activates Gene 1
Gene 1 induces Gene 2
Gene 1 inhibits Gene 3
Gene 2 inhibits Gene 3
Gene 3 inhibits Gene 2