

# SYSTEM IDENTIFICATION FOR IFFL

---

SUMIT MUKHERJEE

mukhes3@uw.edu

## 1 Introduction

The aim of this work is to identify the model parameters for the different cell lines implementing the Incoherent Feed Forward Loop (IFFL). The parameters have then been used in gillespie simulations to generate stochastic trajectories.

### 1.1 Identification Data Set

Provided by Alex Rosenberg. Comprises of several cell lines :-

- Open Loop
- Containing 4 sites complimentary to mir-124a (pv)
- Containing 3 sites (d3)
- Containing 2 sites (d23)
- Containing 1 site (d123)

The data set has different time evolving values of mRNA, miRNA (both from FISH and qPCR) and protein (mCherry) levels for each of the cell lines.

## 2 Model

Initial assumed model :-

$$\begin{aligned}\frac{dm}{dt} &= \alpha_m - \beta_m m - \gamma_s m s \\ \frac{ds}{dt} &= \alpha_s - \beta_s s \\ \frac{dp}{dt} &= \alpha_p m - \beta_p p\end{aligned}$$

Where,  $m$  is the mRNA concentration,  $s$  is the miR-124a concentration,  $p$  is the protein concentration,  $\alpha_m$  is mRNA production rate,  $\alpha_s$  is the miRNA production rate,  $\alpha_p$  is the protein production rate (not constitutive),  $\beta_m$  is the mRNA degradation rate,  $\beta_s$  is the miRNA degradation rate,  $\beta_p$  is the protein degradation rate and  $\gamma_s$  is the miRNA degradation rate for mRNA.

## 3 Parameter Identification

### 3.1 Cost Function

We observe that the time evolution of the miRNA is independent of the mRNA and protein concentration, so we identify it separately and club the mRNA and protein into one identification routine. The cost function used is :-

$$J = \sum_{j=1}^{j=l} a_j \sum_{i=1}^{i=k} \frac{(x_j(i) - x_{act,j}(i))^2}{x_{act,j}(k)^2}$$

Where,  $x_j(i)$  represents the simulated concentration of  $j$ th specie at  $i$ th time point,  $x_{act,j}(i)$  represents the actual concentration of  $j$ th specie at  $i$ th time point,  $k$  is the number of time points,  $l$  is the number of species and  $a_j$  is the weight on the  $j$ th specie. The term in the denominator ensures that the species get normalized (as mRNA, protein and miRNA concentrations are often vastly different). The parameters being optimized here are  $\alpha_m, \alpha_s, \alpha_p, \beta_m, \beta_s, \beta_p, \gamma_s$ .

### 3.2 Parameter Limits

The decay rate limits are chosen based on practical values of half lifes and are chosen to be within one order of magnitude each way. The maximal limit on the production limit is chosen to be a realistic value based on the plots. The limits used for the various parameters are given in Table 1 :-

Parameter	Min	Max
$\beta_m$ (1/h)	.01	.5
$\beta_s$ (1/h)	.02	.04
$\beta_p$ (1/h)	.001	1
$\alpha_m$ (conc./h)	100	3000
$\alpha_s$ (conc./h)	0	500
$\alpha_p$ (1/h)	0	5000
$\gamma_s$ (1/h conc.)	0	.00001

Table 1: Minimum and Maximum values of parameters

### 3.3 Identified Parameters

The identified parameters are given in Table 2.

Parameter	PV	d3	d23	d123	in paper
$\beta_m$ (1/h)	0.26645	0.13174	0.5	0.21695	.1
$\beta_s$ (1/h)	0.029091	0.028633	0.02363	0.04	.03
$\beta_p$ (1/h)	0.054131	0.03673	0.030614	0.018773	.08
$\alpha_m$ (conc./h)	524.9764	517.0953	208.5493	270.1188	0-150
$\alpha_s$ (conc./h)	50.0004	70.7653	22.036	49.0797	0-60
$\alpha_p$ (1/h)	1.5638	1.1415	3.3885	1.6738	10
$\gamma_s$ (1/h conc.)	0.001	0.00051711	0.00097197	0.00028637	.00075

Table 2: Identified values for different parameters

### 3.4 Plots

The plots for each of the classes are shown below :-

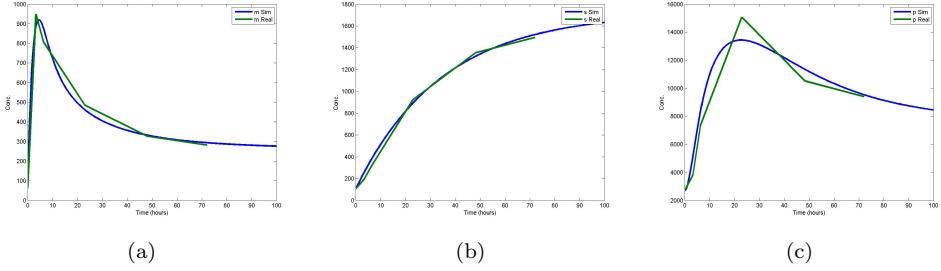


Figure 1: Simulation vs Real plots for PV : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

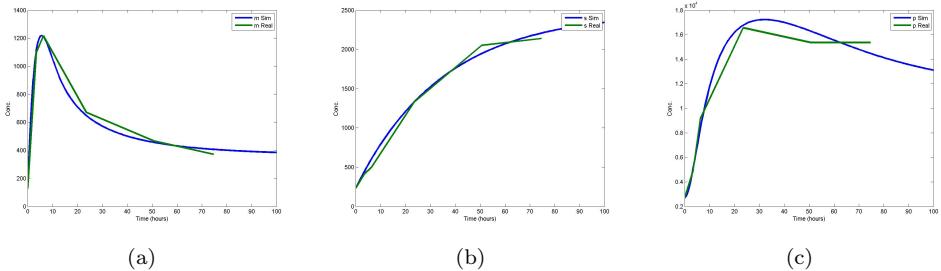


Figure 2: Simulation vs Real plots for d3 : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

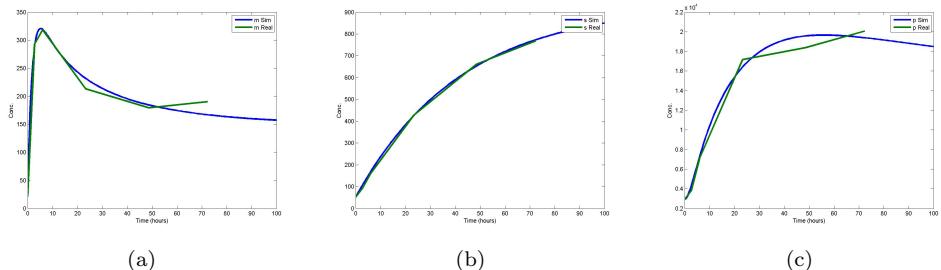


Figure 3: Simulation vs Real plots for d23 : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

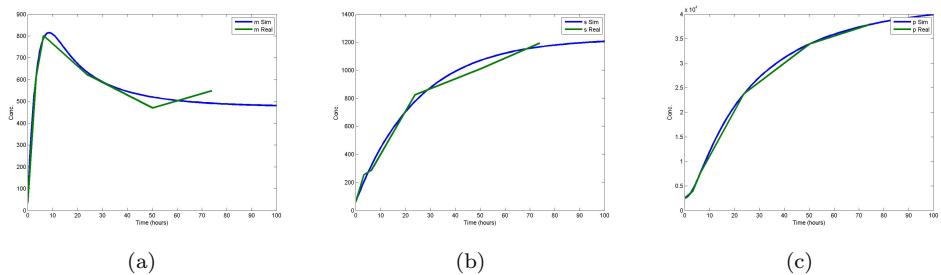


Figure 4: Simulation vs Real plots for d123 : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

### 3.5 Open Loop Plots and Parameters

Table 4 shows the values of the parameters identified with the open loop data (the mRNA data is used from the FISH data for open loop). The plots of the open loop identification are shown below in Figure 5.

Parameter	P
$\beta_m$ (1/h)	0.3169
$\beta_s$ (1/h)	0.0247
$\beta_p$ (1/h)	0.0201
$\alpha_m$ (conc./h)	140.7183
$\alpha_s$ (conc./h)	51.0592
$\alpha_p$ (1/h)	6.5936
$\gamma_s$ (1/h conc.)	0

Table 3: Identified values for different parameters

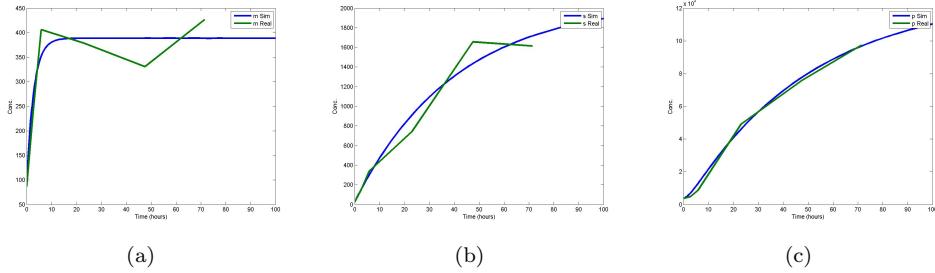


Figure 5: Simulation vs Real plots for P (open loop) : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

### 3.6 Re-identification with FISH data

We have mRNA data from FISH (Fluorescence in situ hybridization) for both P (open loop) and PV cell lines. The re-identification of the PV cell line yields the parameter values as seen in table and the plots can be seen in figure

Parameter	P
$\beta_m$ (1/h)	0.2615
$\beta_s$ (1/h)	0.0290
$\beta_p$ (1/h)	0.0514
$\alpha_m$ (conc./h)	194.9786
$\alpha_s$ (conc./h)	49.967
$\alpha_p$ (1/h)	3.9719
$\gamma_s$ (1/h conc.)	0

Table 4: Identified values for different parameters

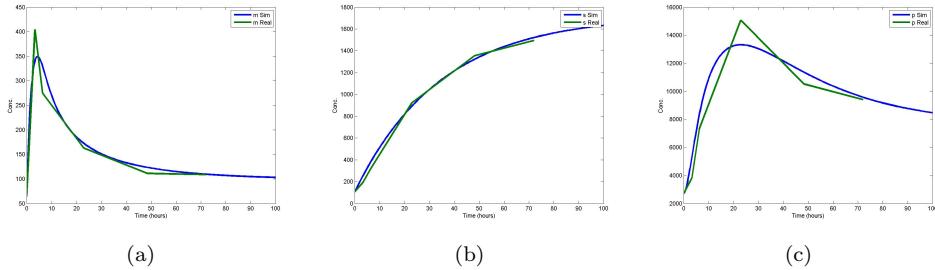


Figure 6: Simulation vs Real plots for PV : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

### 3.7 Open loop plots with PV-FISH parameters

Theoretically, the open loop plot can be obtained by setting  $gs = 0$  and making the initial conditions to be the same as the initial conditions for the open loop case. This yields the plots seen in Figure 7.

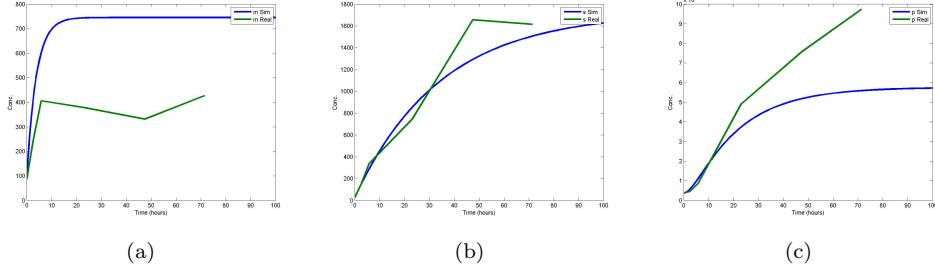


Figure 7: Simulation vs Real plots for PV : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

Clearly this does not lead to very well fitted plots. We assume that this could be a result of a translation inhibition. If only that was the case, the mRNA levels would roughly remain the same but the steady state value seems of the open loop case seems to be around half the steady state value of the PV cell line. Hence we fix all other parameters and just vary  $\alpha_p$  and  $\alpha_m$  to see if we can better fits. The plots obtained are shown in figure

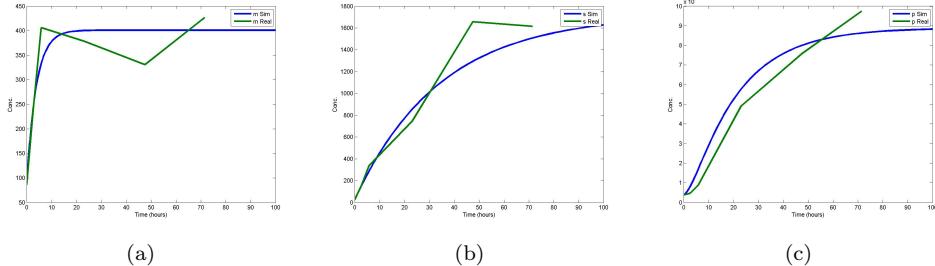


Figure 8: Simulation vs Real plots for PV : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

The identified values of parameters are shown in table

Parameter	P
$\alpha_m$ (conc./h)	104.91
$\alpha_p$ (1/h)	11.3891

Table 5: Identified values for different parameters

This yields slightly better fits. However, we are interested to see whether this is just due to cell line variability or whether translation inhibition has a role to play here. So, now we just vary the  $\alpha_m$  (change it to re-identified value) and keep everything else fixed (to the first case) and observe the results in Figure

So clearly, there is translation inhibition happening which causes the  $\alpha_p$  value to be higher in the re-identified case.

## 4 Identification at multiple Dox levels

Protein fluorescence data from at different Dox induction levels was available for various cell lines. Unfortunately the miRNA and mRNA time course data aren't available. Thus the problem reduces to finding the initial conditions along with the parameters. However, it is expected that the other parameters remain the

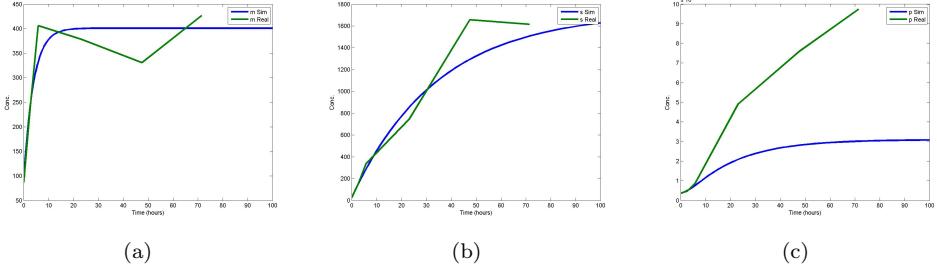


Figure 9: Simulation vs Real plots for PV : a) mRNA b) miRNA (mir-124a) c) protein (mCherry)

same and only the  $a_m$  and  $b_m$  varies. However, their ratio is assumed to be constant, so we need to only identify  $a_m$  for different Dox level. So, the optimization variables are  $m_0$ ,  $s_0$  and  $a_m$ . The plots for the protein fits is seen in Figure 10 for the Full Vamp3 cell line. The predicted plot for fluorescence and mRNA levels at different Dox induction levels is shown in the Figure 11.

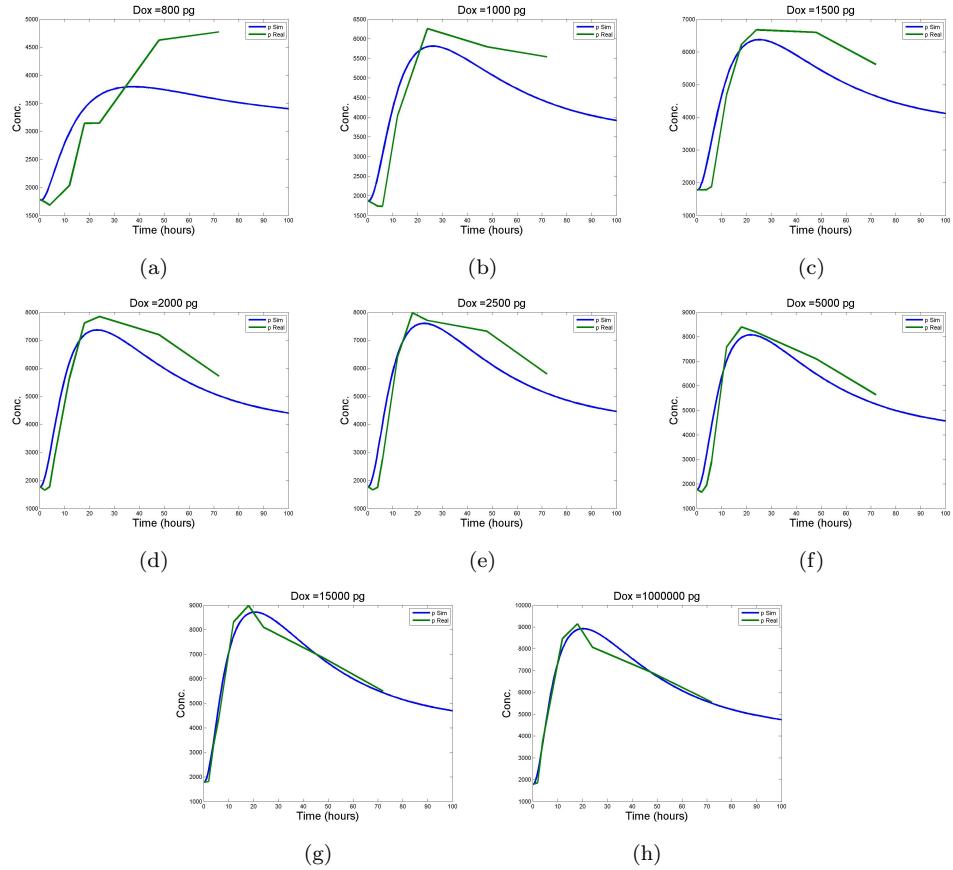


Figure 10: Protein conc. fits at different Dox induction level a)800 ng b) $1\mu\text{g}$  c) $1.5\mu\text{g}$  d) $2\mu\text{g}$  e)  $2.5\mu\text{g}$  f) $5\mu\text{g}$  g) $15\mu\text{g}$  h) $1\text{mg}$

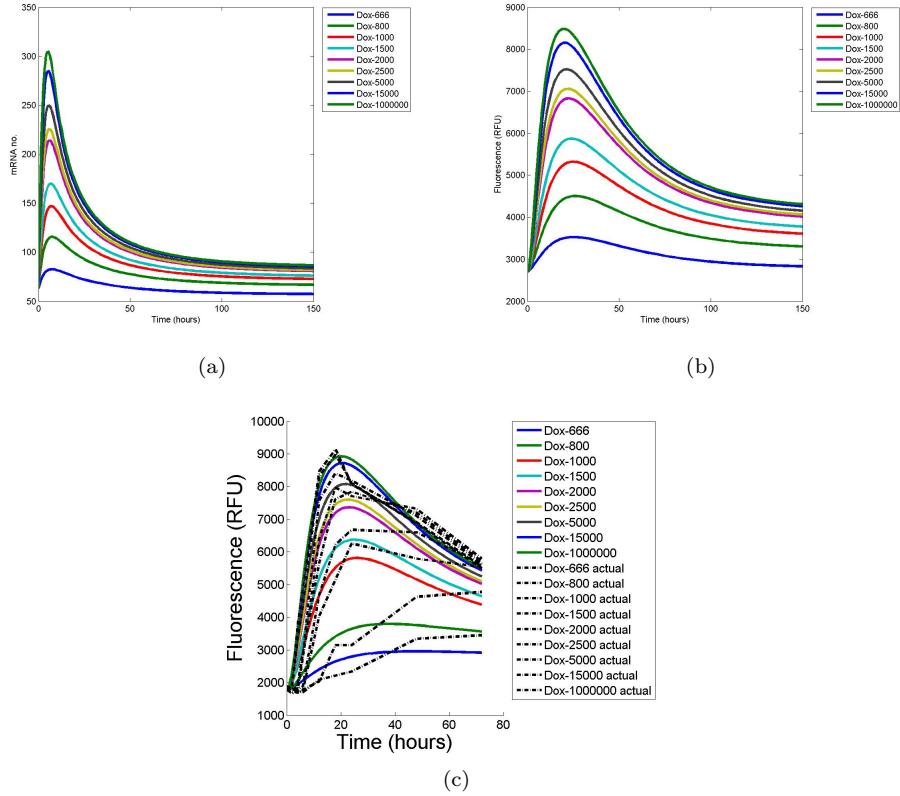


Figure 11: Predicted plots at different Dox induction levels for a) mRNA b) Protein c) Protein with observed data

The same is performed for the open loop data. The plots for the protein fits is seen in Figure 12 for the control cell line. The predicted plot for fluorescence and mRNA levels at different Dox induction levels is shown in the Figure 13.

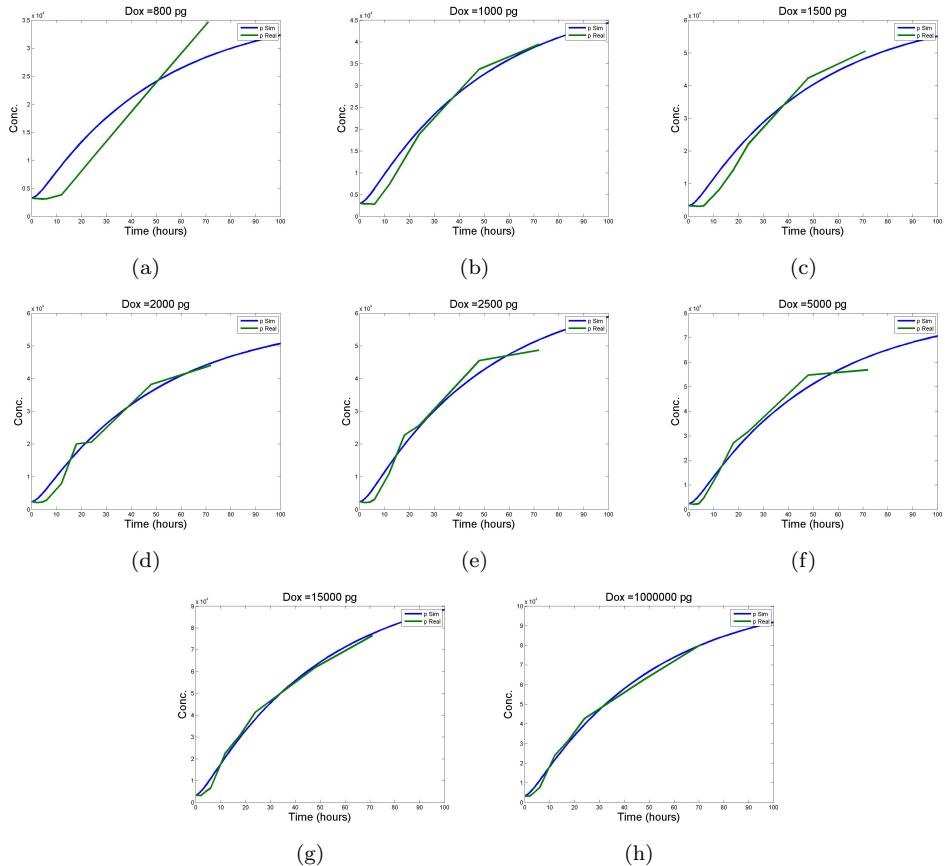


Figure 12: Protein conc. fits at different Dox induction level a)800 ng b) $1\mu\text{g}$  c) $1.5\mu\text{g}$  d) $2\mu\text{g}$  e)  $2.5\mu\text{g}$  f) $5\mu\text{g}$  g) $15\mu\text{g}$  h) $1\text{mg}$

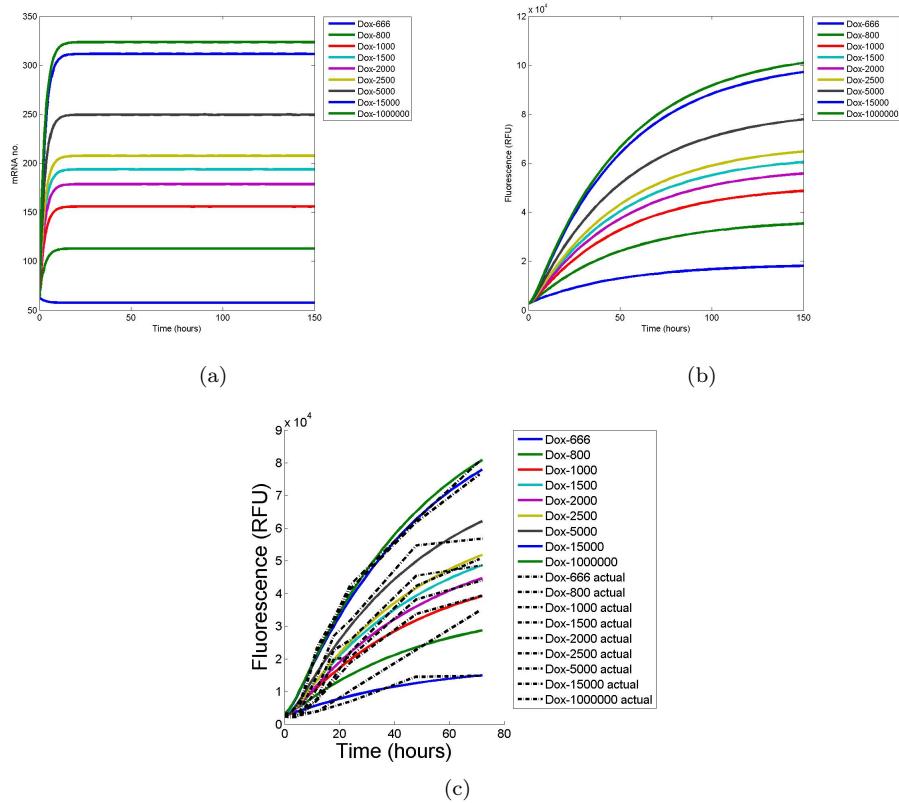


Figure 13: Predicted plots at different Dox induction levels for a) mRNA b) Protein c) Protein with observed data

The same analysis is now performed on two other cell lines (d3 and d23) and can be seen in Figures 14, 15, 16, 17.

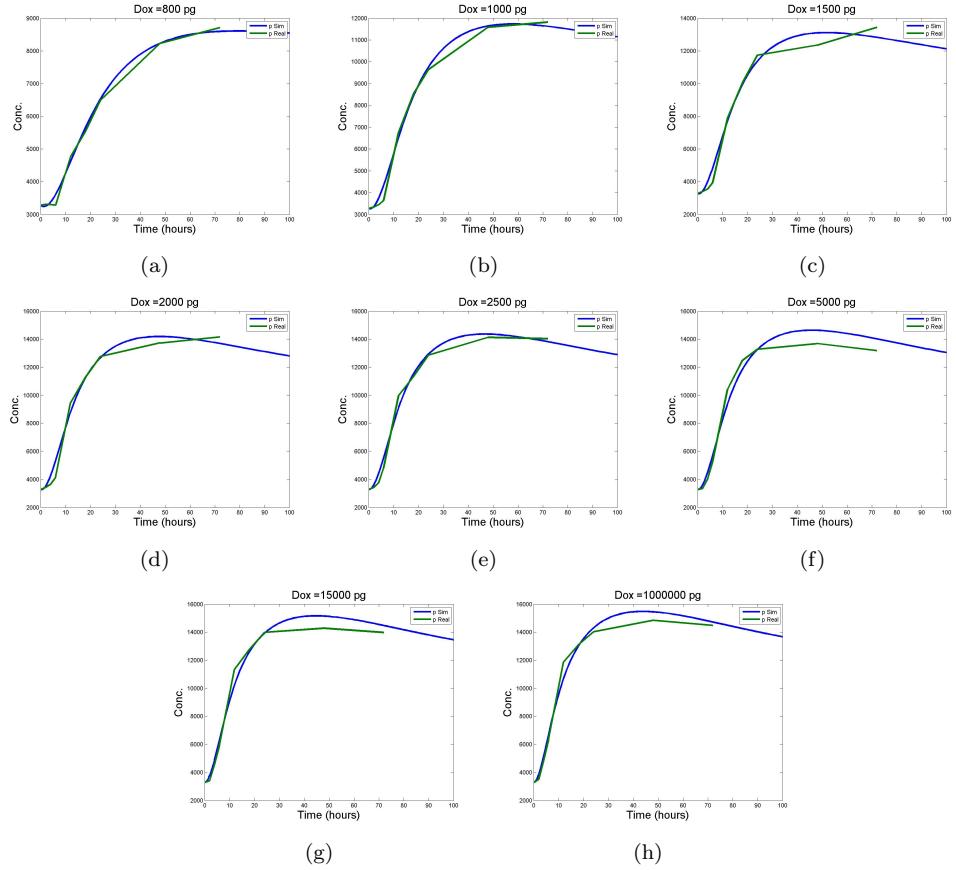


Figure 14: Protein conc. fits at different Dox induction level a)800 ng b) $1\mu\text{g}$  c) $1.5\mu\text{g}$  d) $2\mu\text{g}$  e)  $2.5\mu\text{g}$  f) $5\mu\text{g}$  g) $15\mu\text{g}$  h) $1\text{mg}$

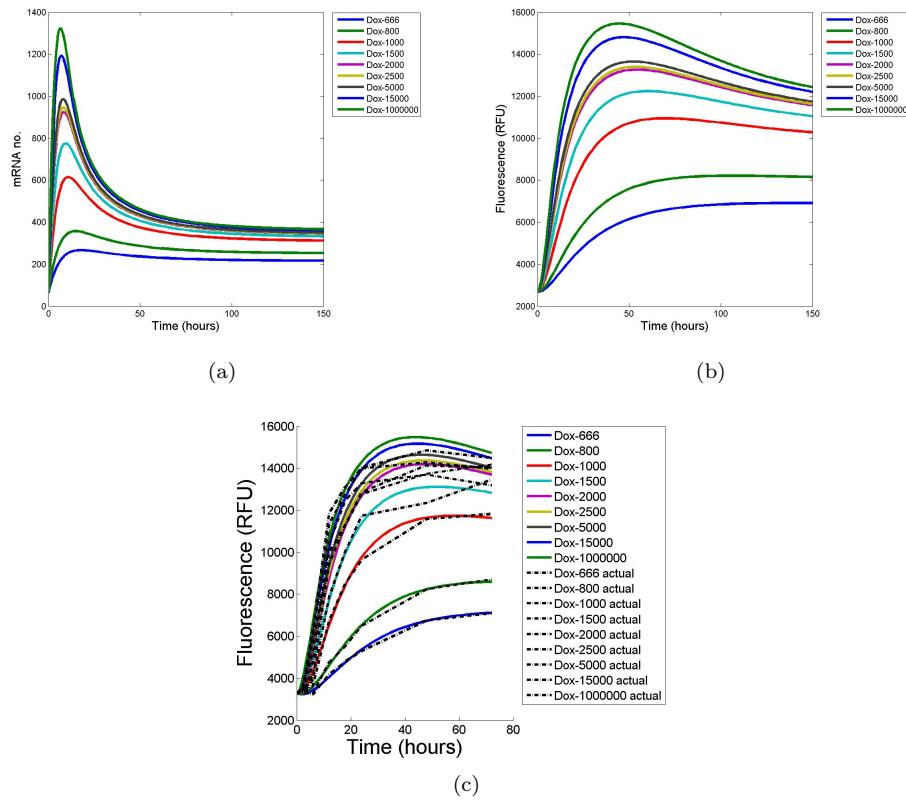


Figure 15: Predicted plots at different Dox induction levels for a) mRNA b) Protein c) Protein with observed data

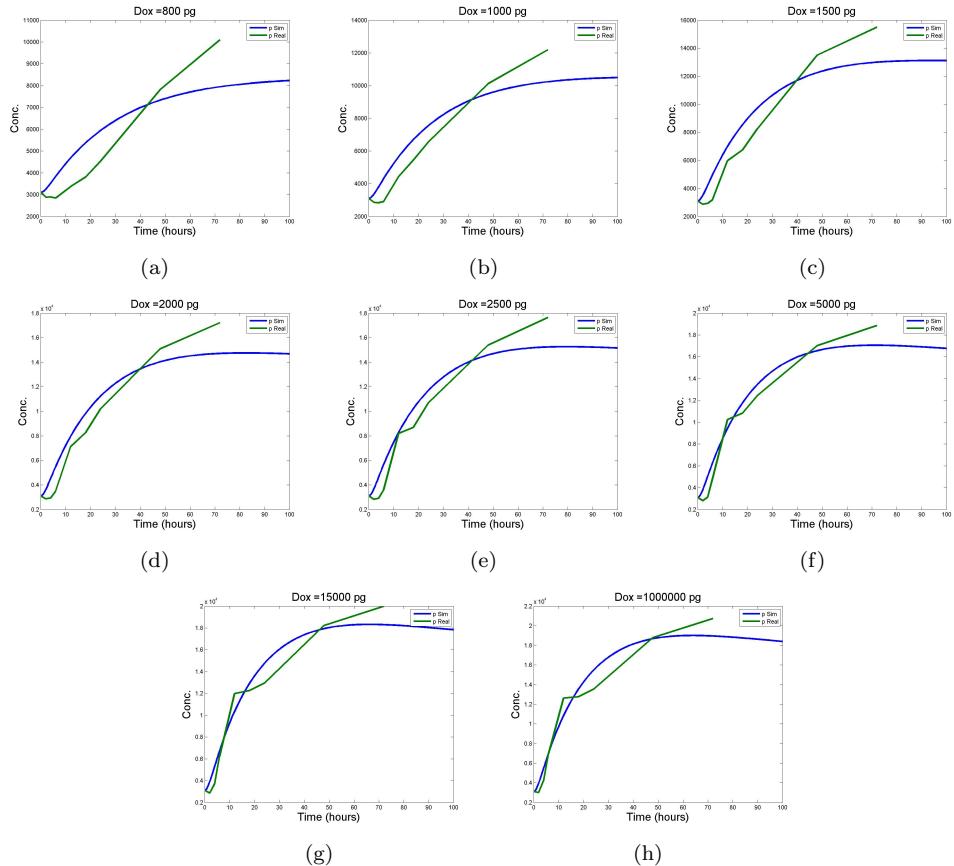


Figure 16: Protein conc. fits at different Dox induction level a)800 ng b) $1\mu\text{g}$  c) $1.5\mu\text{g}$  d) $2\mu\text{g}$  e)  $2.5\mu\text{g}$  f) $5\mu\text{g}$  g) $15\mu\text{g}$  h) $1\text{mg}$

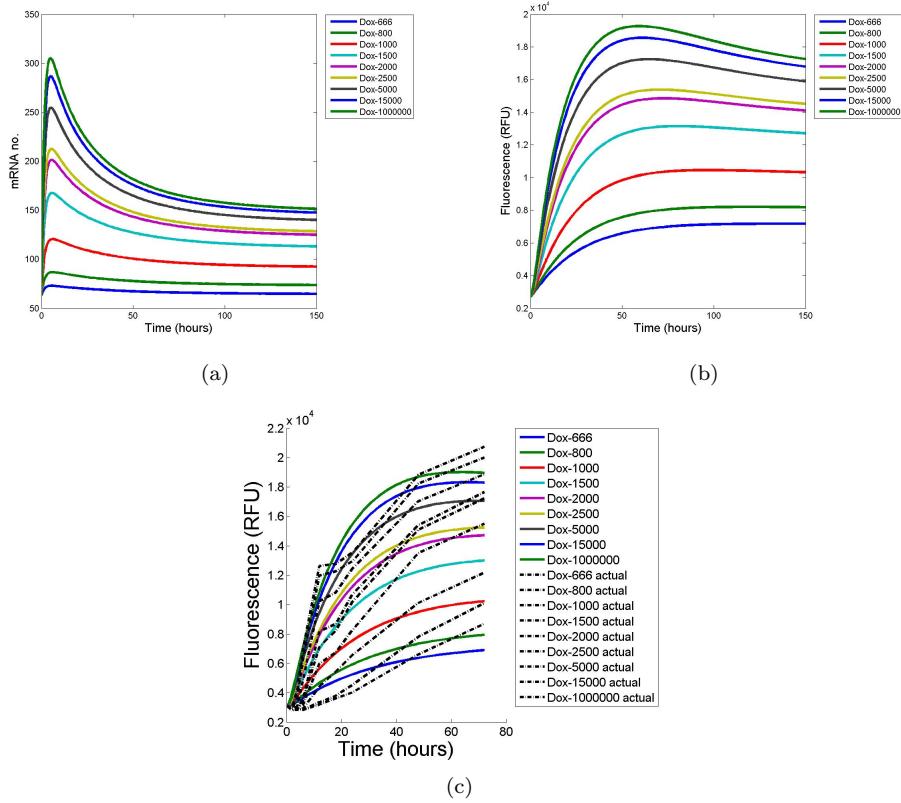


Figure 17: Predicted plots at different Dox induction levels for a) mRNA b) Protein c) Protein with observed data

The comparison of the the steady state protein level with varying levels of Dox induction for the four different cell lines are shown in Fig 18 :-

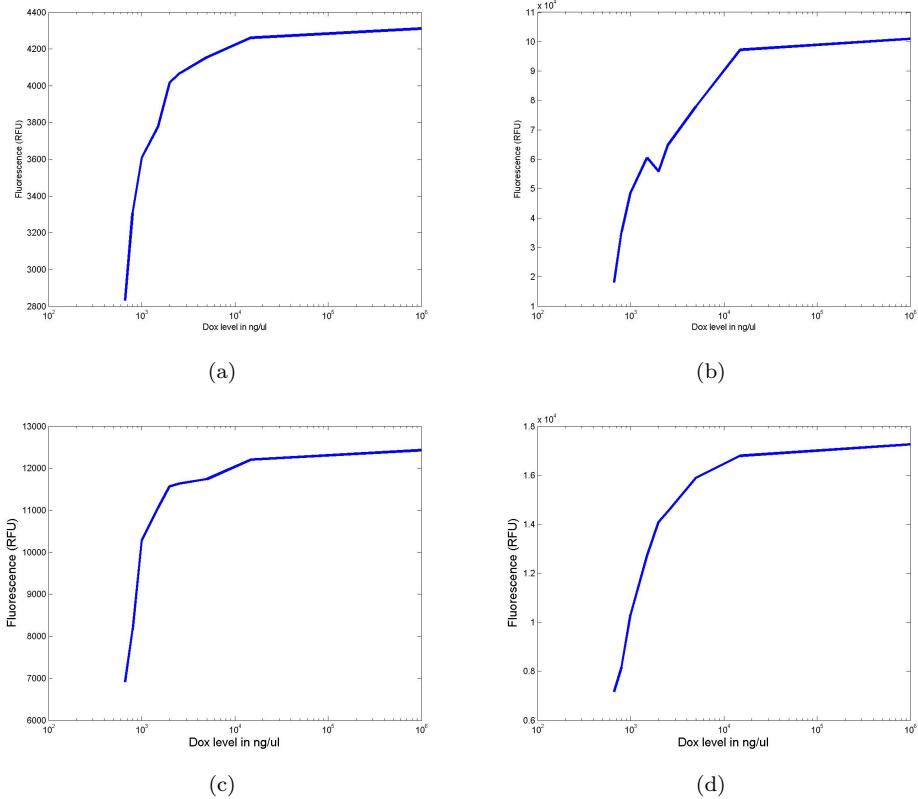


Figure 18: Dox Vs Steady State protein values for a) PV b) P c)d3 d)d23

#### 4.1 Model for $\alpha_m$ variance with induction

We then hypothesize that the  $\alpha_m$  values will have a similar functional relationship with Dox induction levels. For the purpose of the study we assume two models i)  $\alpha_m = \frac{ax}{bx+1}$ , ii)  $\alpha_m = \frac{ax^n}{bx^n+1}$ . The matlab function 'fit' was used to fit these models. The plots are seen in Figure 19. The fits are seen in Figure 19.

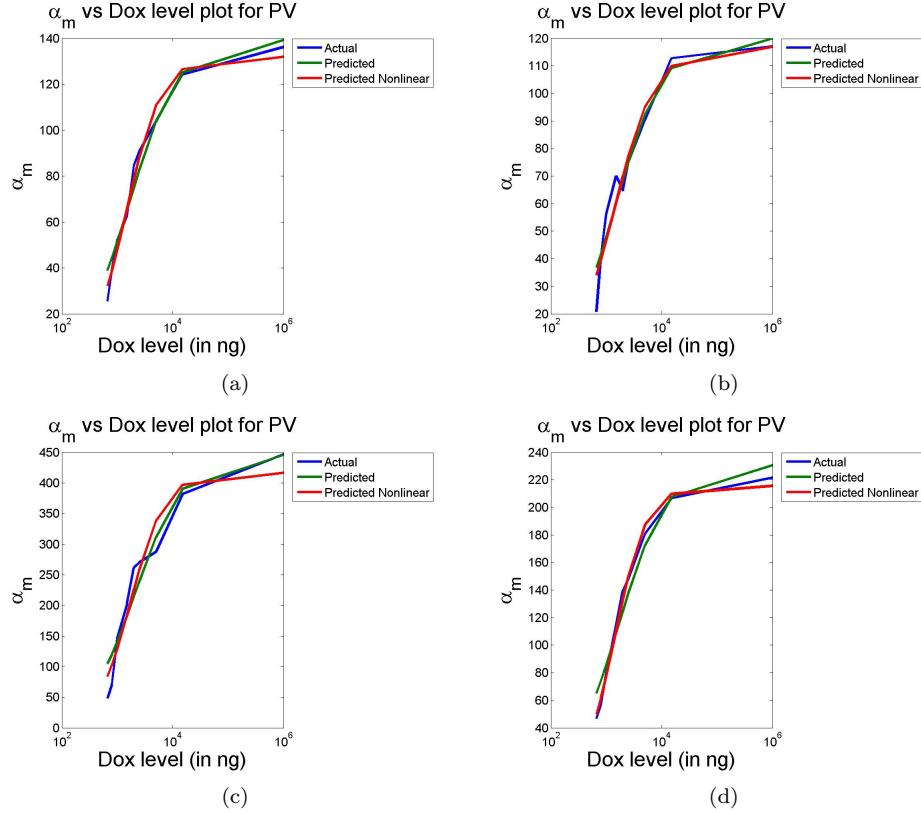


Figure 19: Dox Vs  $\alpha_m$  for a) PV b) P c)d3 d)d23

Cell Line	Model Type	a	b	n	SSE
PV	Linear	0.0812	5.82e-4	1	385.5
PV	Nonlinear	0.0053	4.04e-5	1.3825	186.6
P	Linear	0.0798	6.644e-4	1	471.7
P	Nonlinear	0.0241	2.061e-4	1.17	433.7
d3	Linear	0.205	4.61e-4	1	9913.7
d3	Nonlinear	0.0114	2.73e-5	1.405	8044.7
d23	Linear	0.1362	5.89 e-4	1	1183.6
d23	Nonlinear	0.0031	1.42e-5	1.53	183.3

Table 6: Parameter Values for  $\alpha_m$  Vs Dox plots

## 5 Incorporating Translation Inhibition into Models

It has been demonstrated in several papers that miRNA's not only bind to mRNA's causing them to degrade but also play a role in translational inhibition. This is demonstrated in Figure 20.

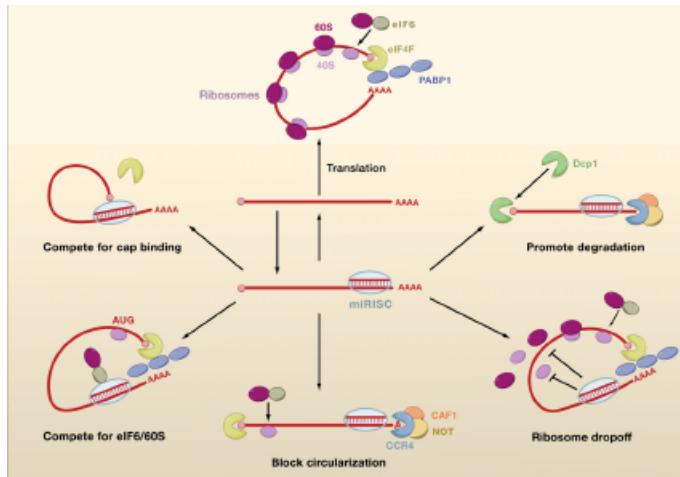


Figure 20: Modes of translation inhibition in miRNAs

### 5.1 Modeling

Based on Figure 20, miRNA not only causes enhanced mRNA degradation by binding to it but also competes for cap binding and causes ribosome dropoff once they are attached to a mRNA. Both these effects cause the translation rate to decrease for the fraction of mRNA's that are attached to the miRNA's.

## 6 Future Direction

- Identify the exact order and mechanism of miRNA based repression
- Come up with a mathematical model to capture the translation inhibition
- Demonstrate how this could work out to our benefit in synthetic control circuits