

Dynamic Model of Heavy Metal Detection Biosensor

Minghui Yin, Sherry Dongqi Bao
Tianjin University

October 15, 2018

1 Introduction

Modeling is a powerful tool in synthetic biology. It provides us with a necessary engineering approach to characterize our pathways quantitatively and predict their performance, thus help us test and modify our design. Through the dynamic model of heavy-metal detection biosensor, we hope to gain insights into the characteristics of our whole circuit's dynamics.

2 Methods

2.1 Analysis of metabolic pathways

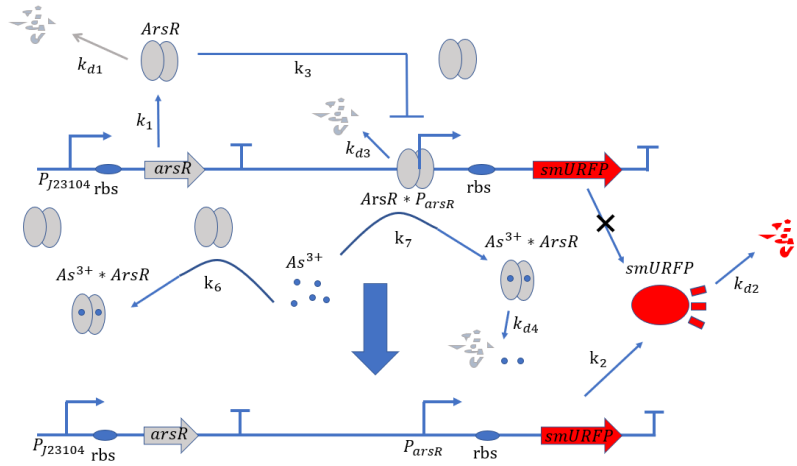
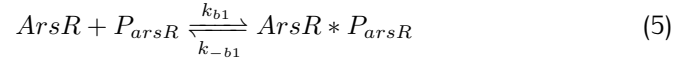
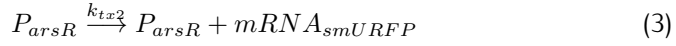
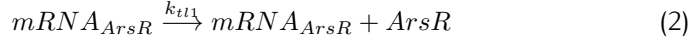


Figure 1: Metabolic pathways related to plasmid#1

At the beginning, on the plasmid#1, the promoter P_{arsR} isn't bound with *ArsR*, thus it is active. *ArsR* and *smURFP* are transcribed and translated under the control of the promoters P_{arsR_u} and P_{arsR_d} , with subscript u and d representing

upstream and downstream separately. The subscript l of smURFP in the equation means leaky expression without the expression of As^{3+} . As $ArsR$ is expressed gradually, it will bind with the promoter P_{arsR} and make it inactive. [1]



On the plasmid#2, the fusion protein of dCas9 and RNAP(RNA polymerase) are produced after transcription and translation, and $sgRNA$ is produced after transcription.

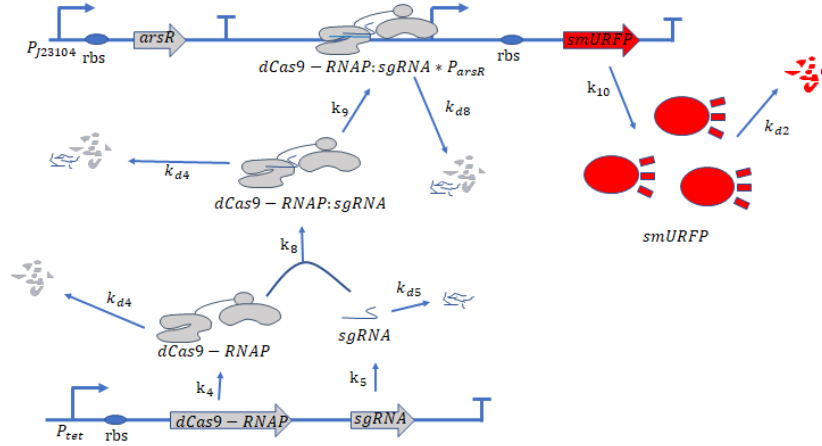
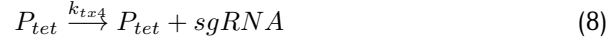
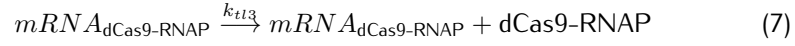
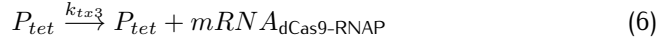


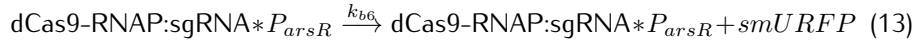
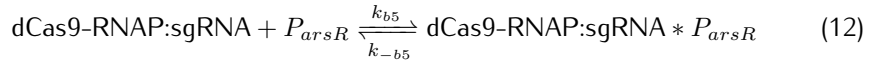
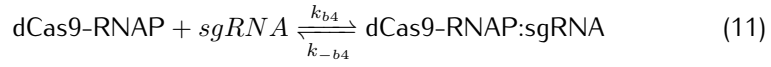
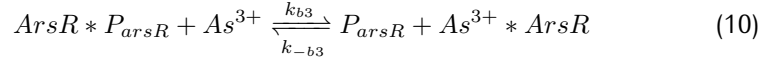
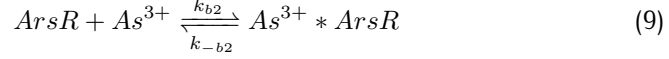
Figure 2: Metabolic pathways related to dCas9/RNAP

$dCas9(*RNAP)$ can bind with its target DNA sequence without cutting, which is at the upstream of the promoter P_{arsR_d} . Simultaneously, $dCas9$ can lead $RNAP$ to bind with the promoter P_{arsR_d} and enhance the transcription of $smURFP$. However, because the promoter P_{arsR_d} has already bound with $ArsR$, as a result, $RNAP$

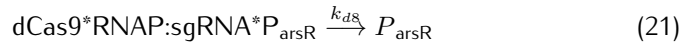
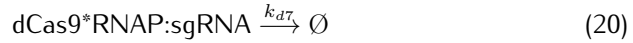
can't bind with the promoter P_{arsR_d} [2].

However, at the presence of As^{3+} , it can bind with ArsR, then dissociate ArsR and P_{arsR_d} , which makes the combination of RNAP and P_{arsR_d} possible.

(Declaration: $[dCas9/RNAP] = [dCas9] = [RNAP]$; $[P_{arsR_d}] = [P_{arsR_u}] = 0.5[P_{arsR}]$)



We then take degradation into account:



2.2 Analysis of ODEs

Applying mass action kinetic laws, we obtain the following set of differential equations. The several complexes involved: $ArsR \cdot P_{arsR}$, $As^{3+} \cdot ArsR$, $dCas9 \cdot RNAP$, $dCas9 \cdot RNAP : sgRNA$, $dCas9 \cdot RNAP : sgRNA \cdot P_{arsR}$, are respectively abbreviated as $cplx_1$, $cplx_2$, $cplx_3$, $cplx_4$, $cplx_5$.

$$\frac{d[ArsR]}{dt} = k_{tl1}[mRNA_{ArsR}] - k_{d1}[ArsR] \quad (1)$$

$$\frac{d[smURFP]}{dt} = k_{tl2}[mRNA_{smURFP}] + k_{b6}[cplx_5] - k_{d2}[smURFP] \quad (2)$$

$$\frac{d[cplx_1]}{dt} = k_{b1}[ArsR][P_{arsR}] - k_{b3}[As^{3+}][cplx_1] - k_{d3}[cplx_1] \quad (3)$$

$$\frac{d[cplx_3]}{dt} = k_{tl3}[mRNA_{dcplx_1}] - k_{b4}[cplx_3][sgRNA] - k_{d5}[cplx_3] \quad (4)$$

$$\frac{d[sgRNA]}{dt} = k_{tx4}[P_{tet}] - k_{b4}[cplx_3][sgRNA] - k_{d6}[sgRNA] \quad (5)$$

$$\frac{d[cplx_2]}{dt} = k_{b2}[As^{3+}][ArsR] + k_{b3}[As^{3+}][cplx_1] - k_{d4}[cplx_2] \quad (6)$$

$$\frac{d[As^{3+}]}{dt} = -k_2[As^{3+}][ArsR] - k_{b3}[As^{3+}][cplx_1] \quad (7)$$

$$\frac{d[cplx_4]}{dt} = k_{b4}[cplx_3][sgRNA] - k_{b5}[cplx_4][P_{arsR}] - k_{d7}[cplx_4] \quad (8)$$

$$\frac{d[cplx_5]}{dt} = k_{b5}[cplx_4][P_{arsR}] - k_{d8}[cplx_5] \quad (9)$$

$$\frac{d[P_{J23104}]}{dt} = 0 \quad (10)$$

$$\frac{d[P_{ArsR}]}{dt} = 0 \quad (11)$$

$$\frac{d[P_{tet}]}{dt} = 0 \quad (12)$$

$$\frac{d[mRNA_{ArsR}]}{dt} = k_{tx1}[P_{J12304}] - k_{d9}[mRNA_{ArsR}] \quad (13)$$

$$\frac{d[mRNA_{smURFP}]}{dt} = k_{tx2}[P_{arsR}] - k_{d10}[mRNA_{smURFP}] \quad (14)$$

$$\frac{d[mRNA_{cplx_1}]}{dt} = k_{tx3}[P_{tet}] - k_{d11}[mRNA_{dcplx_1}] \quad (15)$$

Table 1: Parameters

Rate constants	Value	Units	Reference
k_{tx1}	1.5e-2	s^{-1}	Berset et al.[3]
k_{tl1}	7.33e-2	s^{-1}	Berset et al.
k_{tx2}	1.5e-2	s^{-1}	Berset et al.
k_{tl2}	1.84e-13	s^{-1}	Berset et al.
k_{b1}	1e4	$nM^{-1}s^{-1}$	2006 iGEM Edinburgh
k_{-b1}	0.65	$nM^{-1}s^{-1}$	2006 iGEM Edinburgh
k_{tx3}	5e-4	s^{-1}	Estimated to be slow in comparison to k_{tx4}
k_{tl3}	0.072	s^{-1}	Calculated from Eyal Karzbrun et al.[4]
k_{tx4}	1.33e-3	s^{-1}	2007 iGEM Imperial College London
k_{b2}	1e3	$nM^{-1}s^{-1}$	2006 iGEM Edinburgh
k_{-b2}	0.65	$nM^{-1}s^{-1}$	2006 iGEM Edinburgh
k_{b3}	1.26e4	$nM^{-1}s^{-1}$	Berset et al.
k_{b4}	1.6e-2	$nM^{-1}s^{-1}$	2017 iGEM Munich
k_{b5}	1.66e-5	$nM^{-1}s^{-1}$	2017 iGEM Munich
k_{b6}	4e-5	s^{-1}	Estimated to be slow in comparison to k_2
k_{d1}	3.07e-3	s^{-1}	Berset et al.
k_{d2}	1e-5	s^{-1}	Berset et al.
k_{d3}	1e-3	s^{-1}	Berset et al.
k_{d4}	1.53e-3	s^{-1}	Berset et al.
k_{d5}	2e-2	s^{-1}	Estimated to be fast in comparison to k_{d1}
k_{d6}	7.62e-3	s^{-1}	Estimated according to Berset et al.
k_{d7}	1e-2	s^{-1}	Estimated to be slow in comparison to k_{d5}
k_{d8}	1e-1	s^{-1}	Estimated to be slow in comparison to k_{d7}
k_{d9}	2.81e-3	ns^{-1}	Berset et al.
k_{d10}	7.62e-3	s^{-1}	Berset et al.
k_{d11}	8e-4	s^{-1}	Estimated to be slow in comparison to k_{d9}

2.3 Simulation

Our simulation is based on two softwares: MATLAB (SimBiology Toolbox) and COPASI.

SimBiology Toolbox provides functions for modeling, simulating, and analyzing biochemical pathways by the powerful computing engine of MATLAB.

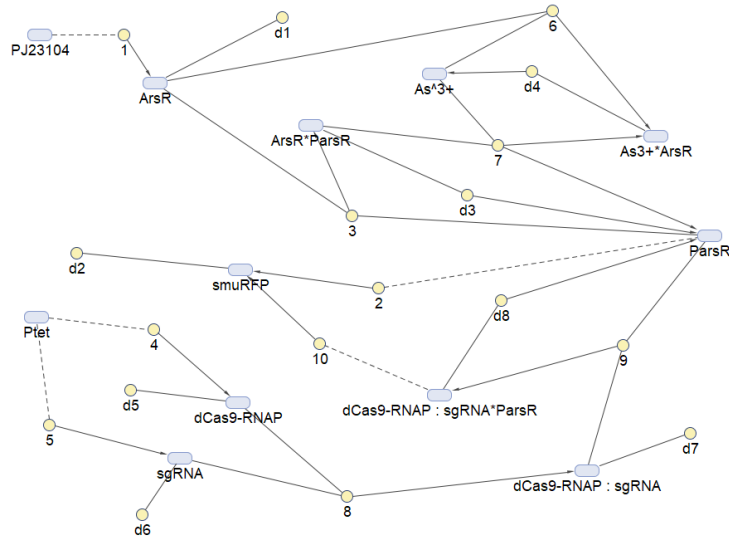


Figure 3: Reaction map generated from the reaction sets above by SimBiology Toolbox

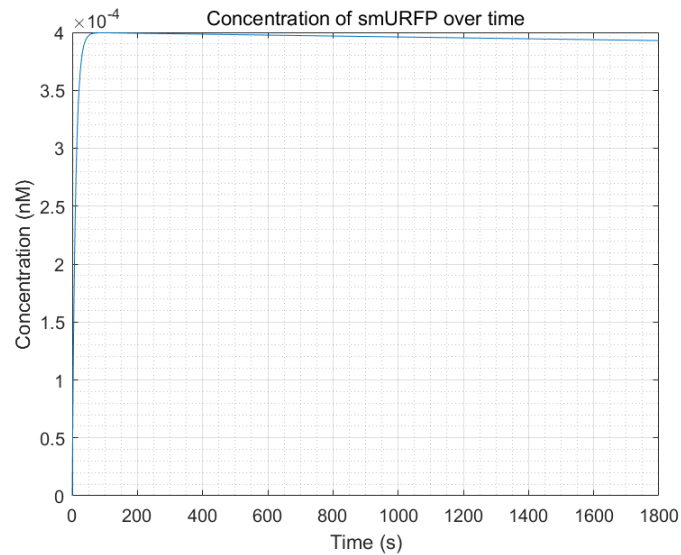


Figure 4: Simulation of smURFP production as a function of time by MATLAB

Through the figure, we can see that the smURFP can gradually increase and

reach a steady state after a period in the presence of arsenic ions.

2.4 Sensitivity

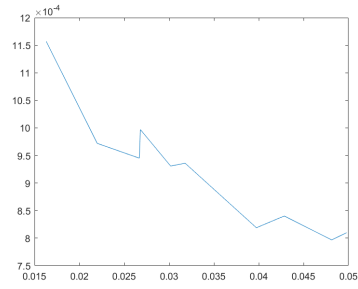
A good biosystem should have certain stability towards fluctuations in parameters. A good model should reflect this, and hence a test for robustness can be essential to the model.

Robustness analysis can also pinpoint which reactions/parameters that are important for obtaining a specific biological behavior. A simple measure for sensitivity is to measure the relative change of a system feature due to a change in a parameter. As for our model, the feature can be the equilibrium concentration of the smURFP(C) for which the sensitivity (S) to a parameter k is:

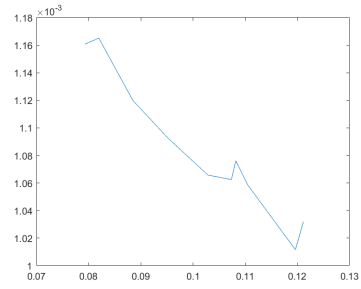
$$S = \frac{\frac{dC}{C}}{\frac{dk}{k}} = \frac{dC}{dk} \frac{k}{C} \approx \frac{\Delta C}{\Delta k} \frac{k}{C} \quad (25)$$

After analysis, we found that the concentration of smURFP is relatively sensitive to parameters such as ktx3, ktl3, ktx4, kb4, kb6, kd2, kd5, kd6, kd7, kd8, kd11, etc. Among these parameters, except for the parameters that directly affect the production and degradation of smURFP, the rest of them are all related to dCas9-RNAP:sgRNA. It shows that our model reflects the critical role of dCas9-RNAP:sgRNA, which initially confirms our hypothesis: dCas9-RNAP can enhance transcription to increase the concentration of smURFP. However, due to the lack of previous modeling studies on dCas9-RNAP, some kinetic parameters may not be very accurate, and due to time limitation, we have not implemented experiments to measure related parameters, which may lead to some deviations in our model.

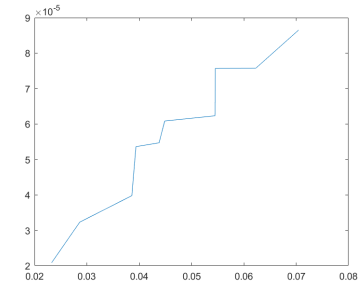
The sensitivity of each parameter is shown in the figures below.



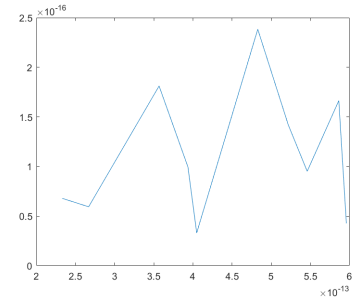
(a) sensitivity of ktx1



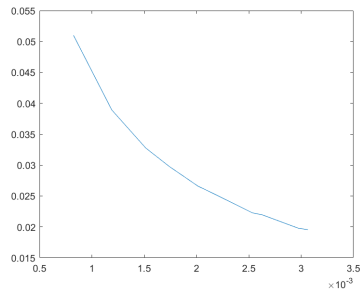
(b) sensitivity of ktl1



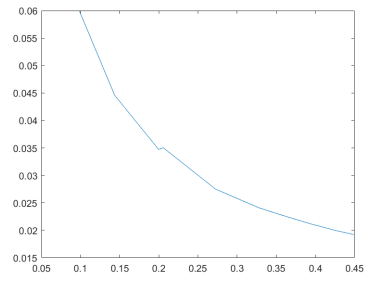
(c) sensitivity of ktx2



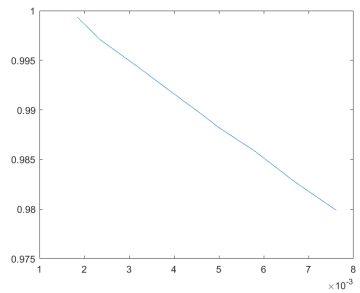
(d) sensitivity of ktl2



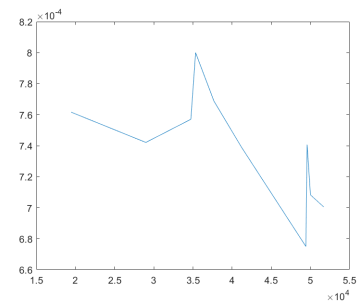
(e) sensitivity of ktx3



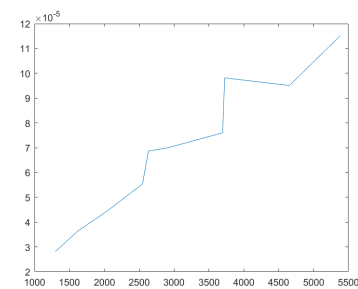
(f) sensitivity of ktl3



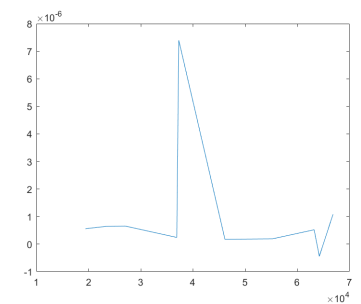
(g) sensitivity of ktx4



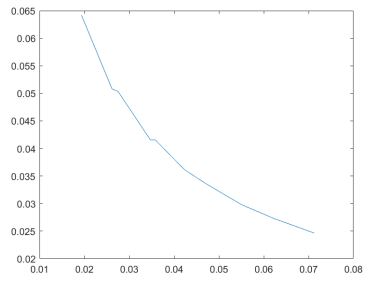
(h) sensitivity of kb1



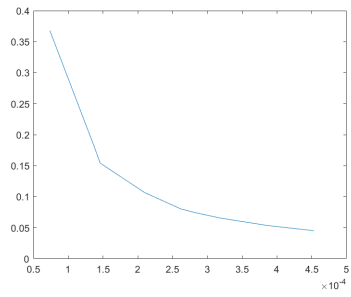
(i) sensitivity of kb2



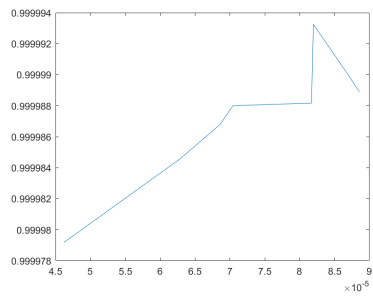
(j) sensitivity of kb3



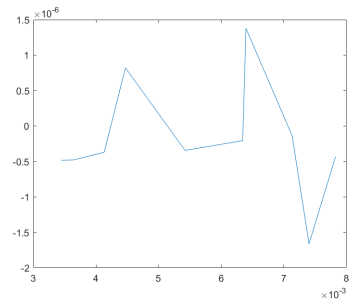
(a) sensitivity of kb4



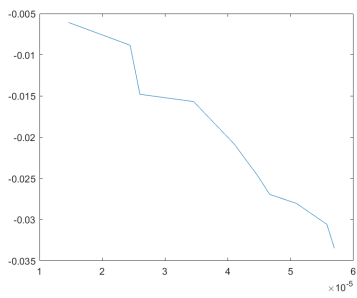
(b) sensitivity of kb5



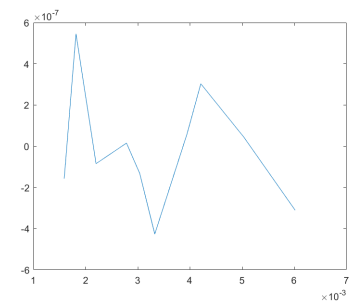
(c) sensitivity of kb6



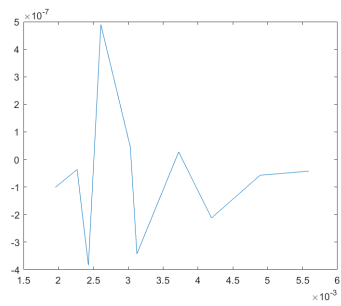
(d) sensitivity of kd1



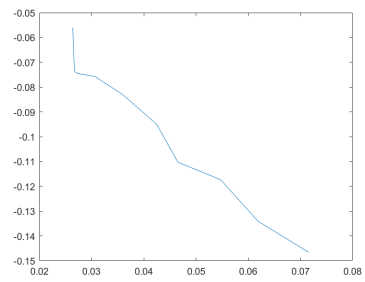
(e) sensitivity of kd2



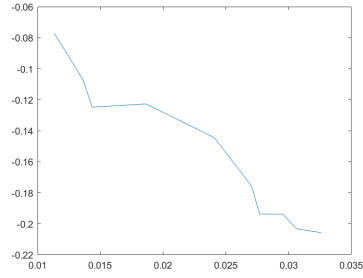
(f) sensitivity of kd3



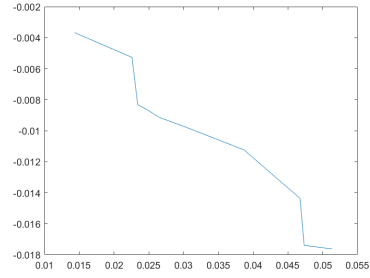
(g) sensitivity of kd4



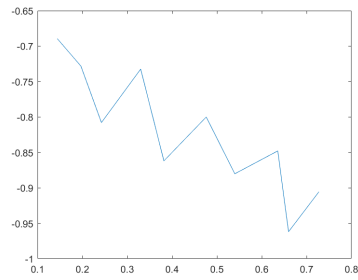
(h) sensitivity of kd5



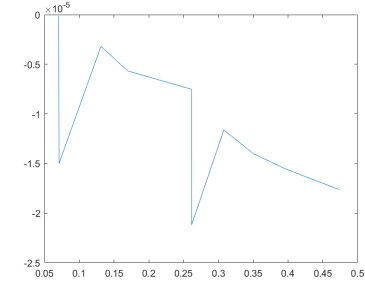
(a) sensitivity of kd6



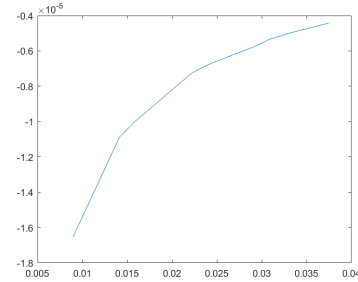
(b) sensitivity of kd7



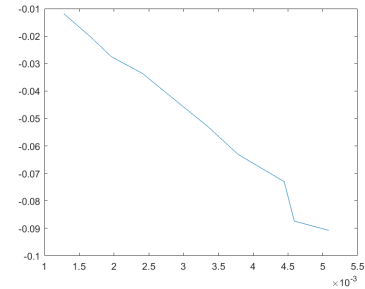
(c) sensitivity of kd8



(d) sensitivity of kd9



(e) sensitivity of kd10



(f) sensitivity of kd11

Note: The ordinate axis represents the sensitivity S , and the abscissa axis is the parameter k for which we want to evaluate the sensitivity.

2.5 Application of the model

Since the goal of our project is to increase the sensitivity of biosensors by introducing a complex of dCas9-RNAP and sgRNA, and one of the purposes of our model is to explore whether this complex is effective. So we assume a reasonable and large enough concentration value for this complex. We use the concentration of glyceraldehyde-3-phosphate dehydrogenase A as the assumed concentration. Glyceraldehyde-3-phosphate dehydrogenase A (gapA) is a crucial enzyme in the glycolytic pathway, and the gene encoding this enzyme is a housekeeping gene in *E. coli* cells with high expression levels. We find in the literature that the protein mass of gapA is 48645 fg/cell, and its molecular weight is 35492 Da.[5] The amount of abundance of Glyceraldehyde-3-phosphate dehydrogenase A protein per cell can be calculated as follows:

$$n = \frac{m}{M} = \frac{48645 * 10^{-15} g}{35492 g/mol} = 1.37 * 10^{-15} mol$$

As for the size of E. coli, we found relevant data from the literature, as the figure below shows.[6]

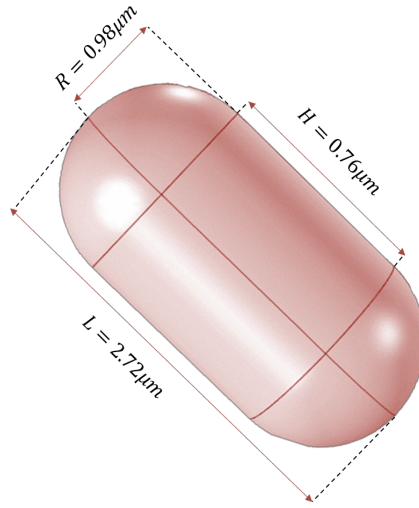


Figure 8: Size of E. coli

The volume of E. coli can be calculated as follows:

$$V_{E.coli} = \frac{4}{3}\pi R^3 + \pi R^2 H = \frac{4}{3}\pi (0.98 \mu m)^3 + \pi (0.98 \mu m)^2 (0.76 \mu m) = 6.24 \mu m^3 = 6.24 * 10^{-15} L$$

Then the concentration of Glyceraldehyde-3-phosphate dehydrogenase A protein in the cell can be determined:

$$c = \frac{n}{V_{E.coli}} = \frac{1.37 * 10^{-15} mol}{6.24 * 10^{-15} L} = 0.22 mol/L$$

With this concentration, we can get very nice results:

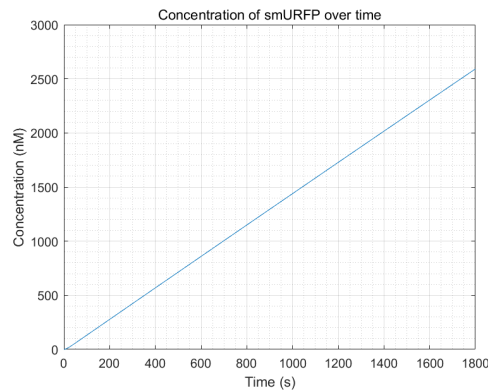


Figure 9: smURFP production with enough dCas9-RNAP:sgRNA

Compared to the diagram without introducing dCas9-RNAP:sgRNA:

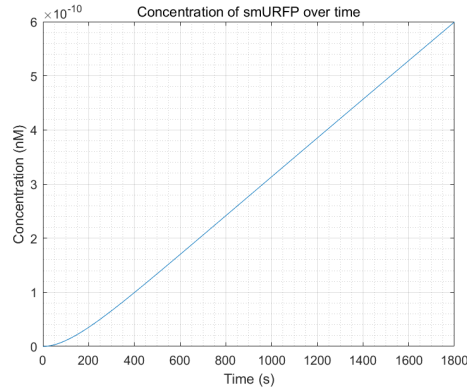


Figure 10: smURFP production within a reasonable time frame

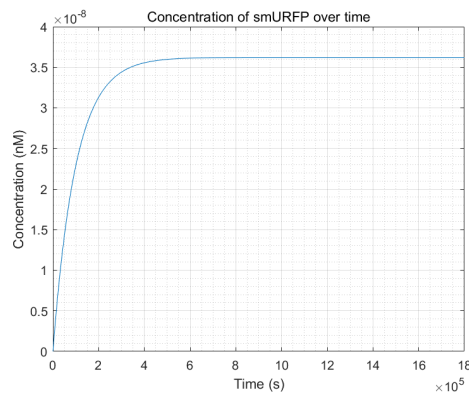


Figure 11: smURFP production reaches equilibrium but it takes a long time

From these three figures, we can conclude that dCas9-RNAP:sgRNA does have the effect of promoting transcription and increasing fluorescence intensity, thereby increasing sensitivity, as long as its concentration is sufficient. This result enhances the confidence of the experimental group, and they need to try to improve the expression of dCas9-RNAP:sgRNA in *E. coli* without having to doubt its role.

References

- [1] LA Pola-López et al. "Novel arsenic biosensor "POLA" obtained by a genetically modified *E. coli* bioreporter cell". In: *Sensors and Actuators B: Chemical* 254 (2018), pp. 1061–1068.
- [2] David Bikard et al. "Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system". In: *Nucleic acids research* 41.15 (2013), pp. 7429–7437.

- [3] Yves Berset et al. "Mechanistic Modeling of Genetic Circuits for ArsR Arsenic Regulation". In: *ACS synthetic biology* 6.5 (2017), pp. 862–874.
- [4] Eyal Karzbrun et al. "Coarse-grained dynamics of protein synthesis in a cell-free system". In: *Physical review letters* 106.4 (2011), p. 048104.
- [5] Yasushi Ishihama et al. "Exponentially modified protein abundance index (em-PAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein". In: *Molecular & Cellular Proteomics* 4.9 (2005), pp. 1265–1272.
- [6] Nili Grossman, Eliora Z Ron, and Conrad L Woldringh. "Changes in cell dimensions during amino acid starvation of Escherichia coli." In: *Journal of bacteriology* 152.1 (1982), pp. 35–41.