Dynamic Model of Heavy Metal Detection Biosensor

Minghui Yin, Sherry Dongqi Bao
Tianjin University

October 14, 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

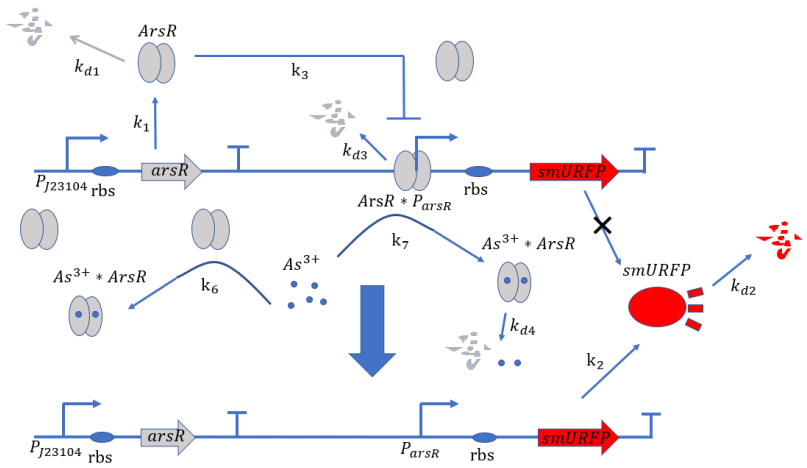


Figure 1: Metabolic pathways related to plasmid#1

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. [**pola2018novel**]

$$P_{J23104} \xrightarrow{k_{tx1}} P_{J23104} + mRNA_{ArsR} \tag{1}$$

$$mRNA_{ArsR} \xrightarrow{k_{tl1}} mRNA_{ArsR} + ArsR \tag{2}$$

$$P_{arsR} \xrightarrow{k_{tx2}} P_{arsR} + mRNA_{smURFP} \tag{3}$$

$$mRNA_{smURFP} \xrightarrow{k_{tl2}} mRNA_{smURFP} + smURFP \tag{4}$$

$$ArsR + P_{arsR} \underset{k_{-b1}}{\overset{k_{b1}}{\rightleftharpoons}} ArsR * P_{arsR} \tag{5}$$

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.

$$P_{tet} \xrightarrow{k_{tx3}} P_{tet} + mRNA_{\text{dCas9-RNAP}} \tag{6}$$

$$mRNA_{\text{dCas9-RNAP}} \xrightarrow{k_{tl3}} mRNA_{\text{dCas9-RNAP}} + \text{dCas9-RNAP} \tag{7}$$

$$P_{tet} \xrightarrow{k_{tx4}} P_{tet} + sgRNA \tag{8}$$

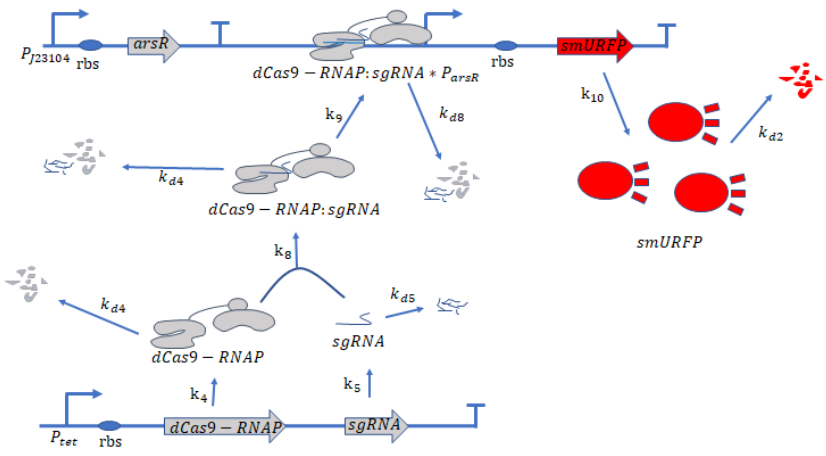


Figure 2: Metabolic pathways related to dCas9/RNAP

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} [**bikard2013programmable**].

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]; $[P_{arsR_d}] = [P_{arsR_u}] = 0.5[P_{arsR}]$)

$$ArsR + As^{3+} \underset{k_{-b2}}{\overset{k_{b2}}{\rightleftharpoons}} As^{3+} * ArsR \tag{9}$$

$$ArsR * P_{arsR} + As^{3+} \underset{k_{-b3}}{\overset{k_{b3}}{\rightleftharpoons}} P_{arsR} + As^{3+} * ArsR \tag{10}$$

$$\text{dCas9-RNAP} + sgRNA \underset{k_{-b4}}{\overset{k_{b4}}{\rightleftharpoons}} \text{dCas9-RNAP:sgRNA} \tag{11}$$

$$\text{dCas9-RNAP:sgRNA} + P_{arsR} \underset{k_{-b5}}{\overset{k_{b5}}{\rightleftharpoons}} \text{dCas9-RNAP:sgRNA} * P_{arsR} \tag{12}$$

$$\text{dCas9-RNAP:sgRNA} * P_{arsR} \xrightarrow{k_{b6}} \text{dCas9-RNAP:sgRNA} * P_{arsR} + smURFP \tag{13}$$

We then take degradation into account:

$$\text{ArsR} \xrightarrow{k_{d1}} \emptyset \tag{14}$$

$$\text{smURFP} \xrightarrow{k_{d2}} \emptyset \tag{15}$$

$$\text{ArsR*P}_{\text{arsR}} \xrightarrow{k_{d3}} P_{\text{arsR}} \tag{16}$$

$$\text{As}^{3+}\text{*ArsR} \xrightarrow{k_{d4}} \text{As}^{3+} \tag{17}$$

$$\text{dCas9*RNAP} \xrightarrow{k_{d5}} \emptyset \tag{18}$$

$$\text{sgRNA} \xrightarrow{k_{d6}} \emptyset \tag{19}$$

$$\text{dCas9*RNAP:sgRNA} \xrightarrow{k_{d7}} \emptyset \tag{20}$$

$$\text{dCas9*RNAP:sgRNA*P}_{\text{arsR}} \xrightarrow{k_{d8}} P_{\text{arsR}} \tag{21}$$

$$mRNA_{ArsR} \xrightarrow{k_{d9}} \emptyset \tag{22}$$

$$mRNA_{smURFP} \xrightarrow{k_{d10}} \emptyset \tag{23}$$

$$mRNA_{\text{dCas9-RNAP}} \xrightarrow{k_{d11}} \emptyset \tag{24}$$

2.2 Analysis of ODEs

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

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

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

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

$$\frac{d[cplx_3]}{dt} = k_{tl3}[mRNA_{dcplx1}] - k_{b4}[cplx_3][sgRNA] - k_{d5}[cplx_3] \tag{4}$$

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

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

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

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

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

$$\frac{d[P_{J23104}]}{dt} = 0 \tag{10}$$

$$\frac{d[P_{ArsR}]}{dt} = 0 \tag{11}$$

$$\frac{d[P_{tet}]}{dt} = 0 \tag{12}$$

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

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

$$\frac{d[mRNA_{dcplx1}]}{dt} = k_{tx3}[P_{tet}] - k_{d11}[mRNA_{dcplx1}] \tag{15}$$

Table 1: Parameters

Rate constants	Value	units	source
k_{tx1}	1.5e-2	s^{-1}	Berset et al.
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.
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 k2
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 kd1
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 kd5
k_{d8}	1e-1	s^{-1}	Estimated to be slow in comparison to kd7
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.

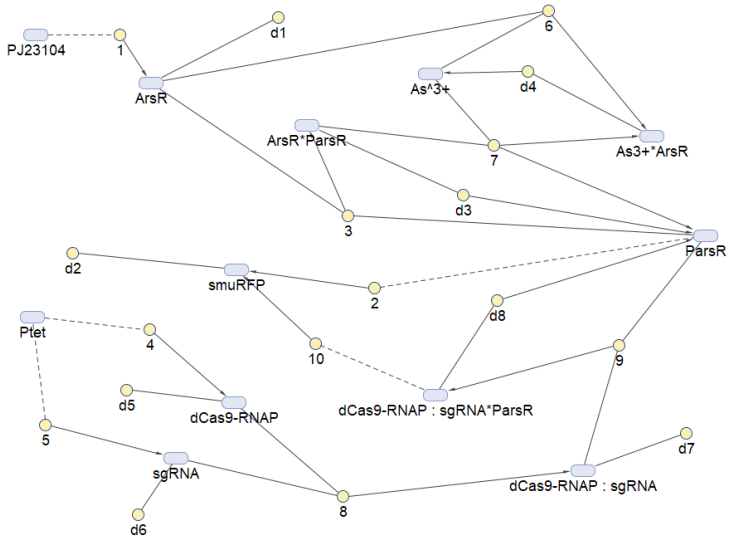


Figure 3: Reaction map generated from the reaction sets above by SimBiology Toolbox

COPASI is freeware developed with the collaboration of VBI and EMLR. It provides almost the same functions as SimBiology, though not quite powerful. However, compared with SimBiology, it provides a friendly user interface for model analysis, such as parameter estimation, and parameter scan.

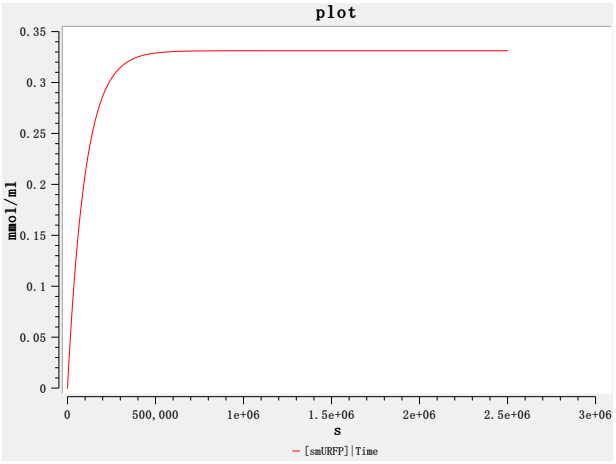


Figure 4: Simulation of smURFP production as a function of time by COPASI

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.

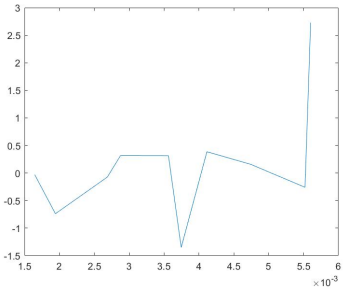
2.4 Sensitivity

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.

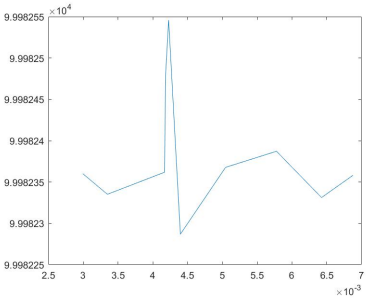
Robustness analysis can also pinpoint which reactions/parameters that are important for obtaining a specific biological behavior. A simple measure for sensitivity is to measure the relative change of a system feature due to a change in a parameter. As for our model, the feature can be the equilibrium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

the smURFP(C) for which the sensitivity (S) to a parameter k is:

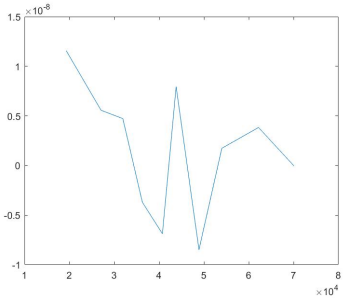
$$S = \frac{\frac{dC}{C}}{\frac{dk}{k}} = \frac{dC}{dk}\frac{k}{c} \approx \frac{\Delta C}{\Delta k}\frac{k}{c} \tag{25}$$



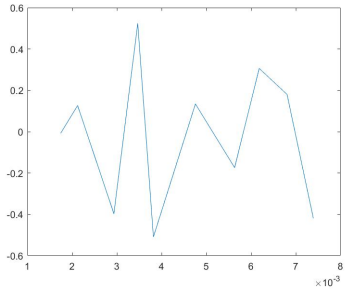
(a) sensitivity of k1



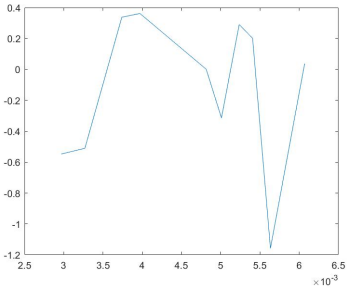
(b) sensitivity of k2



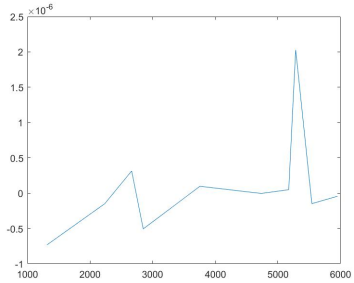
(c) sensitivity of k3



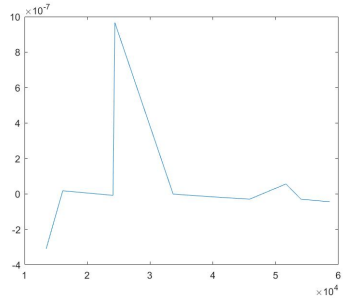
(d) sensitivity of k4



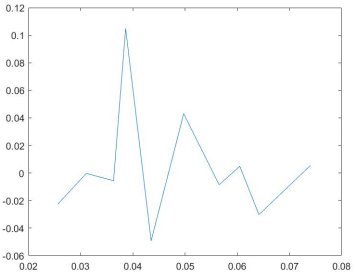
(e) sensitivity of k5



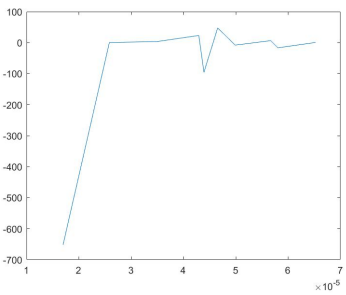
(f) sensitivity of k6



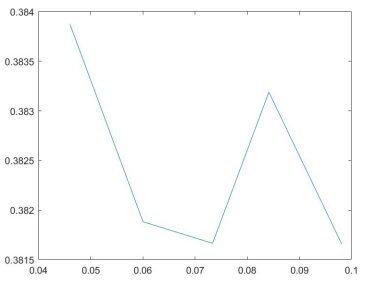
(g) sensitivity of k7



(h) sensitivity of k8

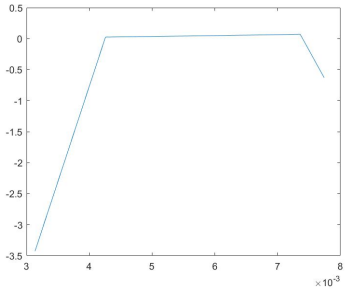


(i) sensitivity of k9

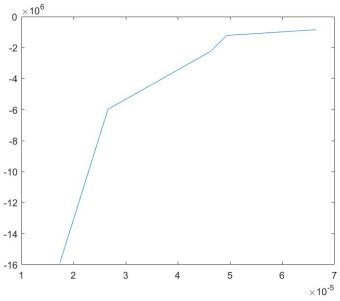


(j) sensitivity of k10

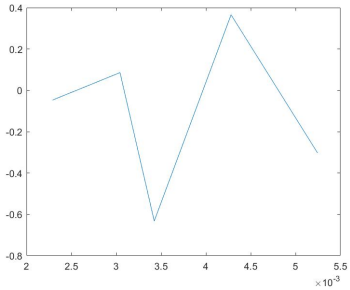
Figure 5: Sensitivity analysis of k1-k10



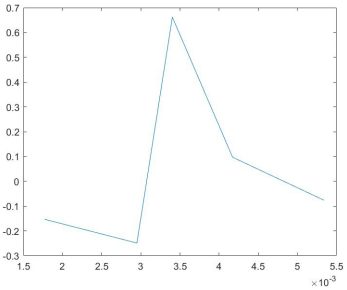
(a) sensitivity of kd1



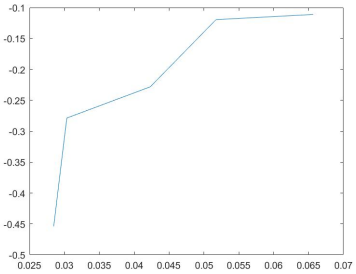
(b) sensitivity of kd2



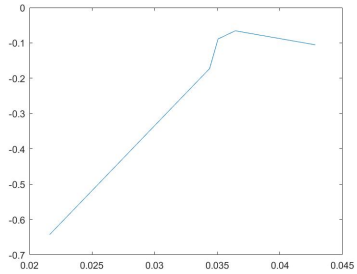
(c) sensitivity of kd3



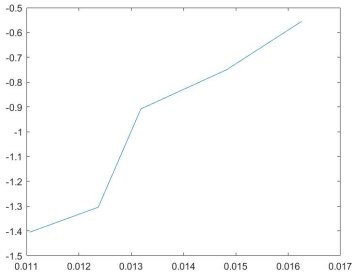
(d) sensitivity of kd4



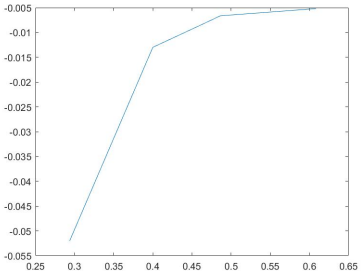
(e) sensitivity of kd5



(f) sensitivity of kd6



(g) sensitivity of kd7



(h) sensitivity of kd8

Figure 6: Sensitivity analysis of kd1-kd8

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 48645fg/cell, and its molecular weight is 35492 Da. 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.

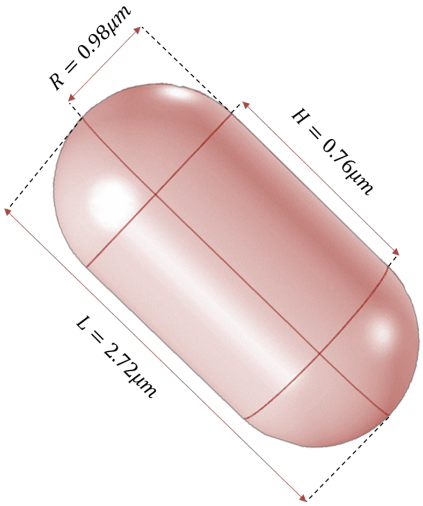


Figure 7: Size of E. coli

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:
Compared to the diagram without introducing dCas9:

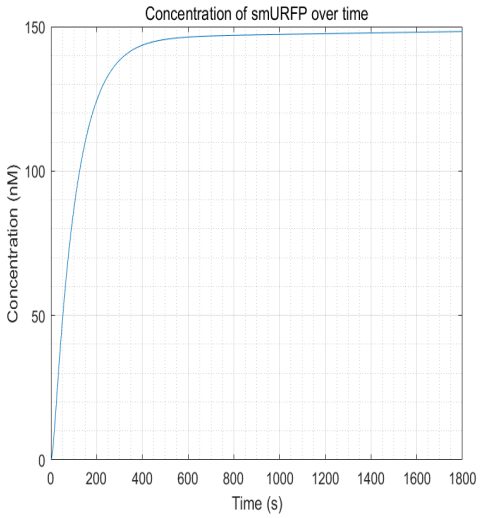


Figure 8: Schematic diagram of smURFP fluorescence

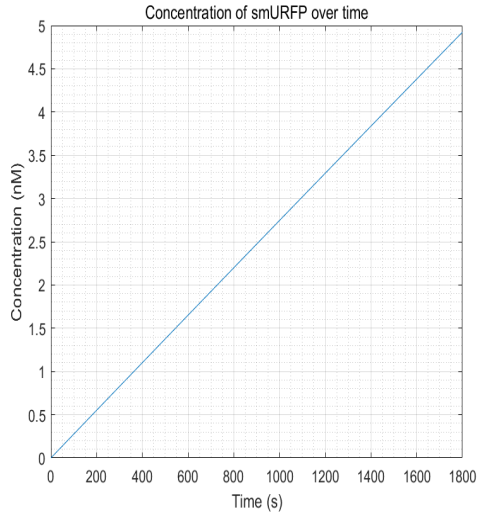


Figure 9: smURFP fluorescence within a reasonable time frame

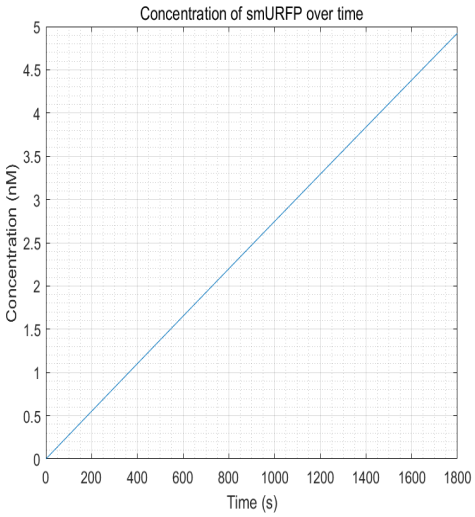


Figure 10: smURFP fluorescence reaches equilibrium but costs too long

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.