Development and Characterization of Fluorescent and Luminescent Biosensors for Estrogenic Activity

Ruchi Asthana[1], Donna Lee[1], Michelle Yu[1], Maxwell R. Telmer[3], Niteesh Sundaram[5], Dominique MacCalla[3,4], Wei Mon Lu[2,4], Jordan Tick[5,6], William Casazza[1], Kenneth Li[1], Diana Marculescu[5], Marcel P. Bruchez[1,7], Natasa Miskov-Zivanov[5] and Cheryl A. Telmer[1]

[1]Department of Biological Sciences, [2]Department of Chemical Engineering, [3]Department of Materials Science, [4]Department of Biomedical Engineering, [5] Department of Electrical and Computer Engineering, [6]School of Computer Science, [7]Department of Chemistry, Carnegie Mellon University

Abstract

The scientific community has become increasingly aware of the negative effects compounds with endocrine disrupting activity have, for example fish feminization. The EPA has a sophisticated program to analyze chemicals for estrogenic activity, however it is neither portable nor inexpensive enough to enable widespread testing of water supplies. The goal of this study was to create a rapid, sensitive, low cost, and portable biosensor for estrogenic activity. Fluorescent proteins including BFP, GFP, YFP, OFP and RFP were purified and evaluated in an effort to standardize measurements that are useful for modeling and sharing this data across the synthetic biology community. Luciferase proteins from *Photinus* (firefly), *Renilla* and *Gaussia* were also characterized and tested as reporters. The fluorescent proteins, luciferases, and a simplified protocol for protein purification were then shared with a local community lab. An estrogen sensor bacterial strain containing a two plasmid biosensor was designed. The sensor plasmid contains a phage T7 RNA polymerase (T7 RNAP), which is conditionally activated by estrogen binding to the human estrogen receptor ligand binding domain (ER LBD). The sensor plasmid expresses YFP as an indicator of the amount of sensor synthesized. The reporter plasmid encodes a reporter protein (RFP or *Gaussia* luciferase) under control of T7 RNAP. Upon binding of estrogen and related compounds, the ER LBD alters conformation and reconstitutes the activity of the T7 RNAP, resulting in transcription and translation of the reporter. The RFP biosensor detected a range of 10-100 μM estrogen and ethinylestradiol. The *Gaussia* biosensor detected the presence of these two compounds at nM concentrations, but was not sensitive enough to differentiate between the varying concentrations. Modeling of the system helped predict and verify laboratory results. The code along with instructions, software, and parts for creating an inexpensive 3D printed luminometer/fluorimeter were made publically available to promote Do It Youself (DIY) projects.

I. Introduction

In 2014, the CMU iGEM team initiated the development of STREAM (Sensor That Reports Endocrine Activating Molecules), an inexpensive, fast, and simple sensor to detect molecules in water that bind to an estrogen receptor ($http://2014.igem.org/Team:Carnegie\_Mellon$). Their approach was (1) to construct a biosensor using biological parts and (2) to take advantage of rapid bacterial growth and simple machinery for signal amplification. The sensor was designed such that the binding of estrogen to a receptor ligand binding domain would result in an intein splicing (Liang et al. 2010); this would cause the production of functional T7 RNAP (Liang et al. 2007). T7 RNAP would then bind to a T7 promoter causing subsequent transcription and translation of red fluorescent protein (RFP), resulting in significant signal amplification. It should be noted that the level of estrogen is reflected by the production of functional T7 RNAP, and reported using the expression of RFP. The susbequent inclusion of yellow fluorescence protein (YFP) on the estrogen responsive sensor plasmid later provided a sensor reference level, making it extremely useful in confirming expression levels.A BioNetGen Model (Faeder et al. 2009) was developed and run in Rulebender (Smith et al. 2012) to quantitatively model the performance of the sensor with the addition of estrogen.

This BioNetGen model predicted that the intein-based sensor would fail due to slow splicing kinetics. This was confirmed by biological data when the sensor did not function in the presence of E2. So inteins were removed and the ER-LBD residues 281-595 (McLachlan et al. 2011) were tested at three positions within T7 RNAP (Appendix). From these experiments, it was determined that 179/180 gave some response. Pieces from McLachlan were subsequently tested and the 312/595 fragment, showed an improved response, and thus it was used in our study. BioNetGen models were modified for these new sensors and fluorescent proteins were quantified to improve accuracy of the model.

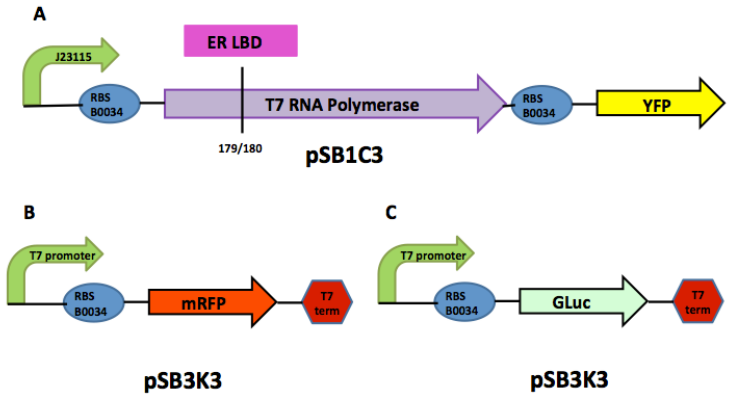


Figure 1: Diagram of (A) Sensor plasmid (B) RFP reporter plasmid and (C) *Gaussia* luciferase reporter plasmid.

II. Materials and Methods

Restriction enzymes, ligase, phusion polymerase, Ni-NTA agarose beads, and chemicals were supplied by Thermo-Fisher. Oligos and gBLOCKS came from IDT. Octyl-thioglucoside was manufactured by Amresco (CAS# 85618-21-9). Table 1 shows the plasmids for expressing fluorescent proteins and luciferases used in this study. Additional cloning details can be found in the Appendix.

A. Cloning, bacterial strains, and vectors

We used the *Escherichia coli (E. coli)* bacterial strain MACH1-T1 (Invitrogen) as the host for cloning. A pSB1C3 (`www.igem.org`) was first modified so that it could include additional restriction sites. Using overlap PCR, the sequence was introduced between the EcoRI and PstI sites in the following restriction site: EcoRI-NotI-XbaI-SphI-HindIII-BshTI-MunI-SpeI-NotI-PstI. The final product was: 5' G/AATTC GCGGCCGC T T/CTAGA G GCATG/C CTT A/AGCTT GGCGGGTCA A/CCGGT GGAGGTTCT CAC/AATTGT TA/CTAGTAGCGGCCGCTGCA/G.

A BBa_J23115 promoter, and a BBa_B0034 RBS were also included in our design. The N and C terminus of T7 RNAP (BBa_K145001) were cloned with the ERLBD and inserted between residue 179 and 180 of T7 RNAP. Lastly, the BBa_B0034 RBS and BBa_K1491004 YFP were added after the stop codon of T7 RNAP. To construct the reporter plasmids, pSB3K3 (`www.igem.org`) was used. A BBa_S04423 T7 promoter, BBa_B0034 ribosome binding site (RBS), BBa_E1010 mRFP1, and BBa_K1732003 *Gaussia* luciferase were followed by a BBa_B0016 T7 termination sequence.

Fluorescent proteins and luciferases were ordered as codon-optimized gBLOCKS with the BBa_J23100 promoter, BBa_B0034 RBS, and BBa_B0015 termination sequences. They were then cloned into pSB1C3 to give the following devices (Table 1).

Table 1. BioBrick numbers of parts used to construct the constitutive codon optimized fluorescent protein and luciferase expressing plasmids that were deposited into the iGEM registry.

Constitutively Expressing Plasmid Name	iGEM BioBrick Number
pSB1C3-J23100-B0034-BFPCO-His-B0015	BBa_K1732007
pSB1C3-J23100-B0034-EGFPCO-His-B0015	BBa_K1732008
pSB1C3-J23100-B0034-GFPCO-His-B0015	BBa_K1732009
pSB1C3-J23100-B0034-E0040-His-B0015	BBa_K1732010
pSB1C3-J23100-B0034-YFPCO-His-B0015	BBa_K1732012
pSB1C3-J23100-B0034-OFPCO-His-B0015	BBa_K1732011
pSB1C3-J23100-B0034-RFPCO-His-B0015	BBa_K1732013
pSB1C3-J23115-B0034-RFPCO-His-B0015	BBa_K1732014
pSB1C3-J23100-B0034-PelB-GlucCO-B0015	BBa_K1732000
pSB1C3-J23100-B0034-CDcel-*Gaussia*CO-His-B0015	BBa_K1732004
pSB1C3-J23100-B0034-*Renilla*CO-B0015	BBa_K1732005
pSB1C3-J23100-B0034-FireflyCO-B0015	BBa_K1732006

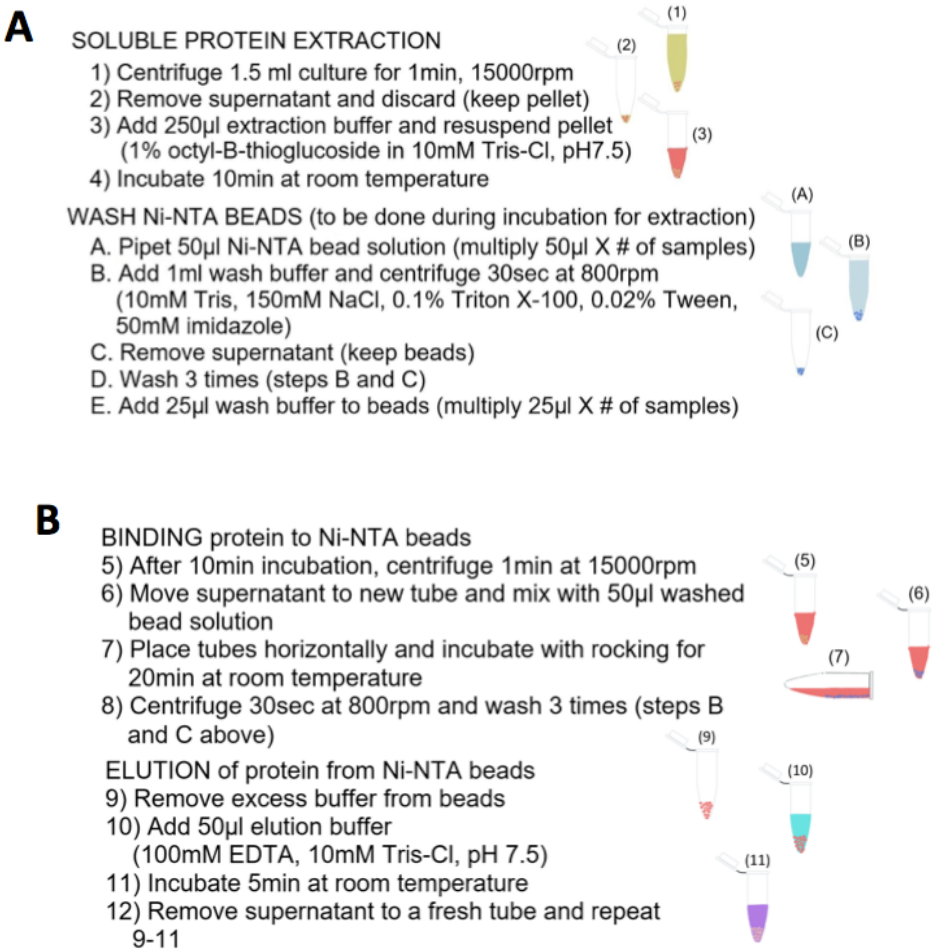


Figure 2: These cards show a simple protocol that was developed for protein purification

B. Fluorescent protein purification

Fluorescent proteins were isolated to quantify and measure fluorescence. The protocol to extract soluble proteins was optimized for 1.5mL of culture. It was then simplified and placed onto cards for easy access to the methods (Figure 2). These instructions were then tested on naive students to troubleshoot the language and clarity of the steps. Fluorescence was measured with a TECAN M1000 using wavelengths for excitation (nm)/ emission (nm) (ex/em). Gain was set at 80 and bandwidth 5nm (RFP 10 nm). The results were: 399/456 (BFP), 488/509 (GFP), 514/527 (YFP), 548/562 (OFP) and 584/607 (RFP).

C. Fluorescent protein quantitation

The Microassay Procedure for Microtiter Plates protocol from the supplier Bio-Rad was followed (`http://www.bio-rad.com/LifeScience/pdf/Bulletin_9004.pdf`). A standard curve was generated using Bradford Standard Assay (BSA), and proteins were analyzed by adding reagent to dilutions of the purified proteins. Absorbance was measured using a TECAN M1000 at 595nm.

D. Luciferase localization

To determine where the luciferases were located, cells expressing luciferase were first centrifuged. The supernatant of these samples was subsequently collected and pellets were tested with coelenterazine (50μM). This mixture was then analyzed in the Bioluminescence mode of the TECAN M1000, using a 100ms integration time.

E. Estrogen Sensor Response

A single colony of bacterial sensor strain containing the sensor plasmid (BBa_K1732015) and either the RFP (BBa_E1010) or the *Gaussia* luciferase

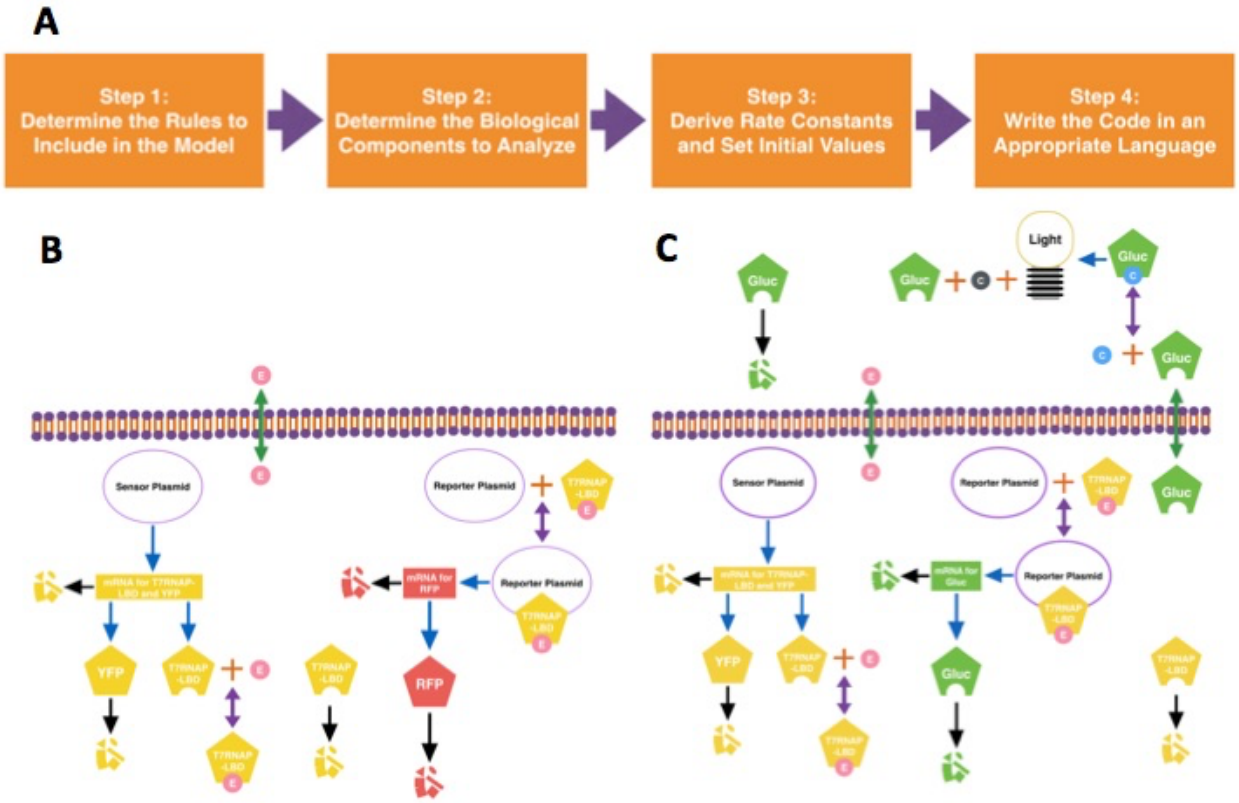


Figure 3: (A) The general outline for constructing a rule-based model (B) Contact map generated from RFP reporter model (C) Contact map generated from *Gaussia* reporter model.

(BBa_K1732003) reporter was cultured overnight in kanamycin 50μg/mL and chloramphenicol 34μ/mL. This mixture was then diluted 1/20 and compounds were added. The 17-β-ethinylestradiol (E2) and 17-α-ethinylestradiol (EE2) were made as stocks in ethanol and added 1:1000 into the cultures for concentrations ranging from 100μM to 10pm. Control cultures received ethanol only to eliminate the carrier effects of other compounds. Cells were incubated at 37°C with shaking at 300rpm. After 24 hours, 100μL of culture was analyzed for fluorescence using a TECAN M1000. To detect YFP, an ex/em of 514/527 with a gain of 80 (5nm bandwidth) was used. To detect RFP, and ex/em of 584/607 with a gain of 100 (10nm bandwidth) was used. Triplicate cultures were analyzed; ratios were calculated and graphed using GraphPad Prism software (GraphPad Software). The procedure used to prepare the sample was the same for the *Gaussia* luciferase biosensor as it was for the RFP biosensor. However, 10μl of 750μM coelenterazine (75μM final) in water was added to the sample 1 minute prior to reading luminescence.

F. Modeling

Two models were created to characterize the estrogen sensor. The first model is based on using red fluorescence protein (RFP) as an intracellular reporter, while the second is based on using *Gaussia* as an extracellular reporter. The biosensor models were written in BioNetGen Language (BNGL), a rule-based modeling language. Rule-based modeling is a type of modeling in which differential equations are generated from a description of how various biological components and systems interact with one another (Figure 3). For more details about the modeling, please see the Appendix.

The models were run in RuleBender version 2.0.382 (http://visualizlab.org/rulebender/), an interactive design environment which is dedicated to analyzing, visualizing, and debugging BNGL models. A useful feature of rule bender is its ability to graphically display the interactions between each of the components in the form of a contact map. The contact map helps the user visualize how each interaction fits together in the overall scheme of the system (Figure 3).

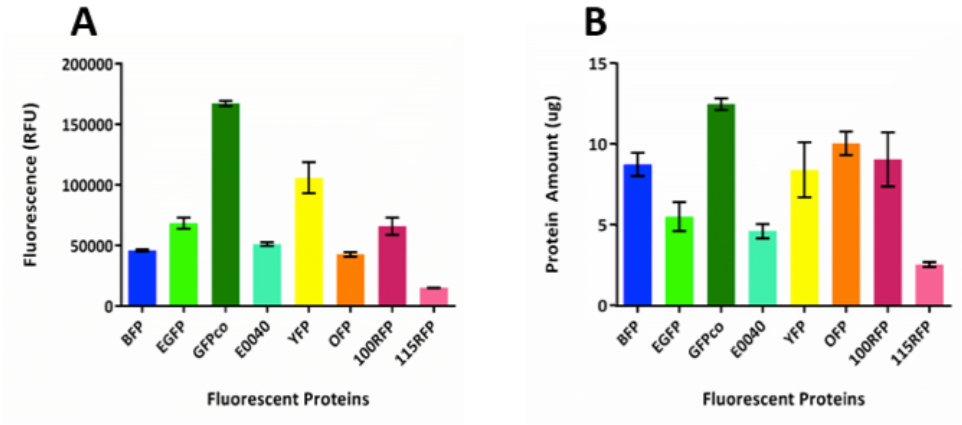


Figure 4: (A) Graph of relative fluorescence calculated for 1mL of culture using TECAN M1000 set for a gain of 80 and bandwidth of 5nm (RFP10nm). Error bards represent standard deviation from triplicate cultures. (B) Graph shows amount of fluorescent protein (ug) present in 1mL of culture for each of the fluorescent proteins. The ratio amount of purified protein/10000 RFU of fluorescence of purified protein (Appendix Figure C) was used to determine total amount of fluorescent protein in 1mL of culture.

III. Results

A. Codon Optimized Fluorescent Proteins

In order to provide quantitative data for modeling it is important to be able to relate fluorescent readings with amounts of protein. Fluorescence is a measurement of light that is detected and changed to an electrical signal that is then amplified; values are obtained in relative fluorescent units (RFU). RFU are relative to some background; data collected with respect to RFU are then compared to a control sample to obtain meaningful data. These measurements are not always useful for modeling therefore fluorescent proteins were purified and fluorescence was related to quantity. Figure 4 shows the fluorescence of bacterial cultures expressing fluorescent proteins. As you will notice codon optimized GFP cultures showed 3-4 times more fluorescence than the BBa˙E0040 GFP cultures. YFP, EGFP, BFP and OFP showed decreasing signal, and RFP expressed from the strong J23100 promoter showed more fluorescence than RFP expressed from the weak J23115 promoter (Figure 4A). Figure 4B shows that the amount of protein per culture varied 2-3 fold across fluorescent proteins. Codon optimization produced at least twice as much protein (GFPco vs. E0040). A 3-4 fold change was observed with promoter differences. These numbers can be used in models involving fluorescent proteins. Note that differences in fluorescence are not the same as differences in amount due to spectral properties of the specific fluorescent proteins. If labs were to calibrate their fluorescence readings to protein amount it would allow for standardization of fluorescence data.

B. Luciferases

Fluorescent proteins are very convenient reporters because they do not require a substrate; however, there is significant autofluorescence at the wavelengths used for detection. Therefore we investigated the use of luciferases for reporters. Luciferases that remained in the cytoplasm such as firefly and *Renilla* were compared to that from *Gaussia* which is an extracellular luciferase. Additionally, two extracellular domains were tested, one predicted to target the media (CDcel) and the other predicted to target the periplasmic space (PelB).

Renilla and firefly luciferase did not contain a targeting sequence and were localized intracellularly as expected. *Gaussia* luciferase with the targeting domain derived from the catalytic domain of cellulase (Gao et al. 2015) , called CDcel-*Gaussia*, was found mainly in the media as expected. *Gaussia* luciferase with the PelB leader sequence (Lei et al. 1988) , called PelB-*Gaussia*, was expected to remain in the periplasmic space however 95 % of it was found in the media (Table 2).

Table 2. This table shows the luciferases that were produced in bacteria and their subcellular locations.

Transformed Cells	Expected Location	Observed Location
CD-cel Gaussia	Extracellular	Extracellular
PelB Gaussia	Periplasmic Space	Extracellular
Renilla	Intracellular	Intracellular
Firefly	Intracellular	Intracellular

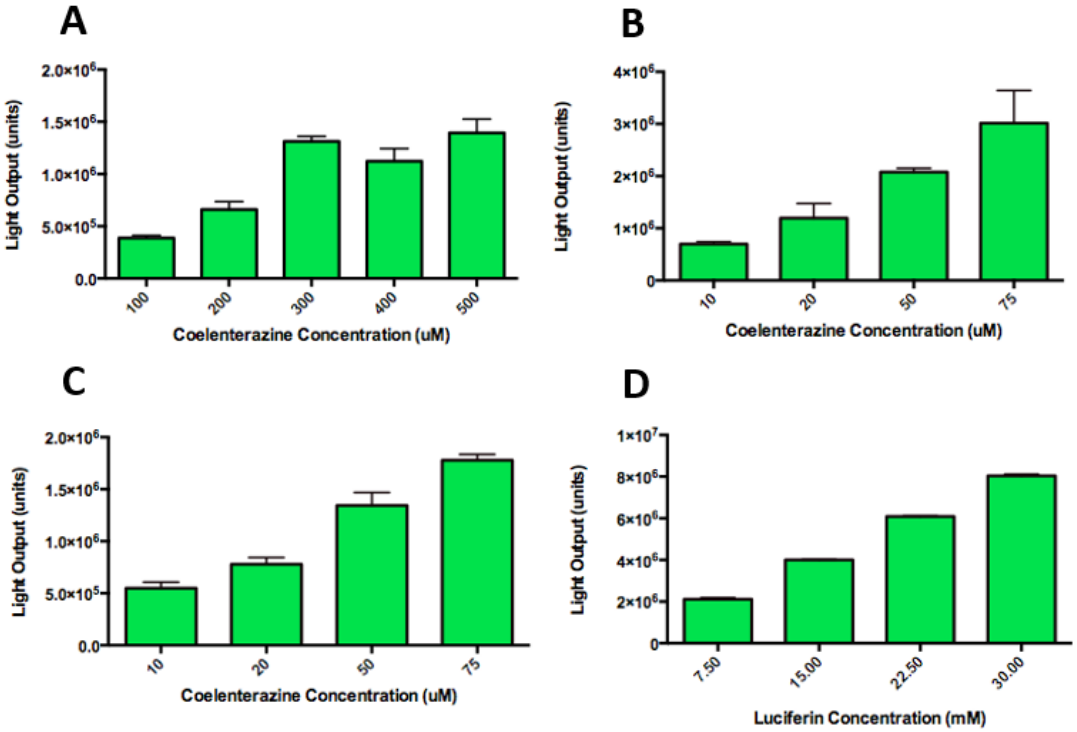


Figure 5: (A) This graph shows the relative luminescence of cultures of CDcel-*Gaussia* expressing bacteria with various concentrations of coelenterazine added and measured after 1 min. (B) This graph shows the relative luminescence of cultures of PelB-*Gaussia* expressing bacteria with various concentrations of coelenterazine added and measured after 1 min. (C) This graph shows the relative luminescence of cultures of *Renilla* expressing bacteria with various concentrations of coelenterazine added and measured after 1 min. (D) This graph shows the relative luminescence of cultures of Firefly expressing bacteria with various concentrations of luciferin added and measured after 1 min.

Of the coelenterazine using luciferases, PelB-*Gaussia* expressing cells produced the highest level of light output. The firefly luciferase produced the highest light output but uses luciferin, a different substrate, and the concentration required mM instead of uM as for the coelenterazine utilizing luciferases. It is important to note that we were unable to determine the concentration that saturates the firefly culture because our stock solution of luciferin is 30mM and we could not test any higher concentrations. There is a possibility that a higher concentration of luciferin would produce even greater light output. We also had problems with CDcel-*Gaussia*, which did not produce much light (Figures 5).

C. Response of estrogen sensor

The biosensor is a bacterial cell containing two-plasmids. The sensor plasmid is a high-copy plasmid, which has the ligand binding domain of the human estrogen receptor alpha (ER-LBD) inserted into T7 RNAP and YFP is transcribed from the same message and used for normalization. When the ER-LBD binds estrogen, it causes a conformational change (McLachlan et al. 2011) that brings together the separated domains of T7 RNAP and the activity of the T7 RNAP is reconstituted (Shis and Bennet et al. 2012). The second plasmid is the reporter plasmid and is a low-copy plasmid, which has the T7 promoter driving expression of RFP. When the T7 RNAP is reconstituted upon binding to estrogen, it allows for binding to the T7 promoter on the reporter plasmid, transcription of the RFP mRNA and then translation to produce RFP. Several luciferases were then characterized and the RFP was substituted for a *Gaussia* luciferase reporter. Two EDCs with high activities are natural estrogen hormones known as 17-β-ethinylestradiol (E2) and a synthetically produced estrogen 17-α-ethinylestradiol (EE2). E2 is a major estrogen used to naturally regulate the female reproductive system and to maintain sexual characteristics. EE2 is derived from E2 and is used mainly as a component of contraceptives. It is important to note that the negative control (labeled as μM estrogen) contained only 100% ethanol as it was the solvent in which the estrogenic compounds were resuspended in. The relative flu-

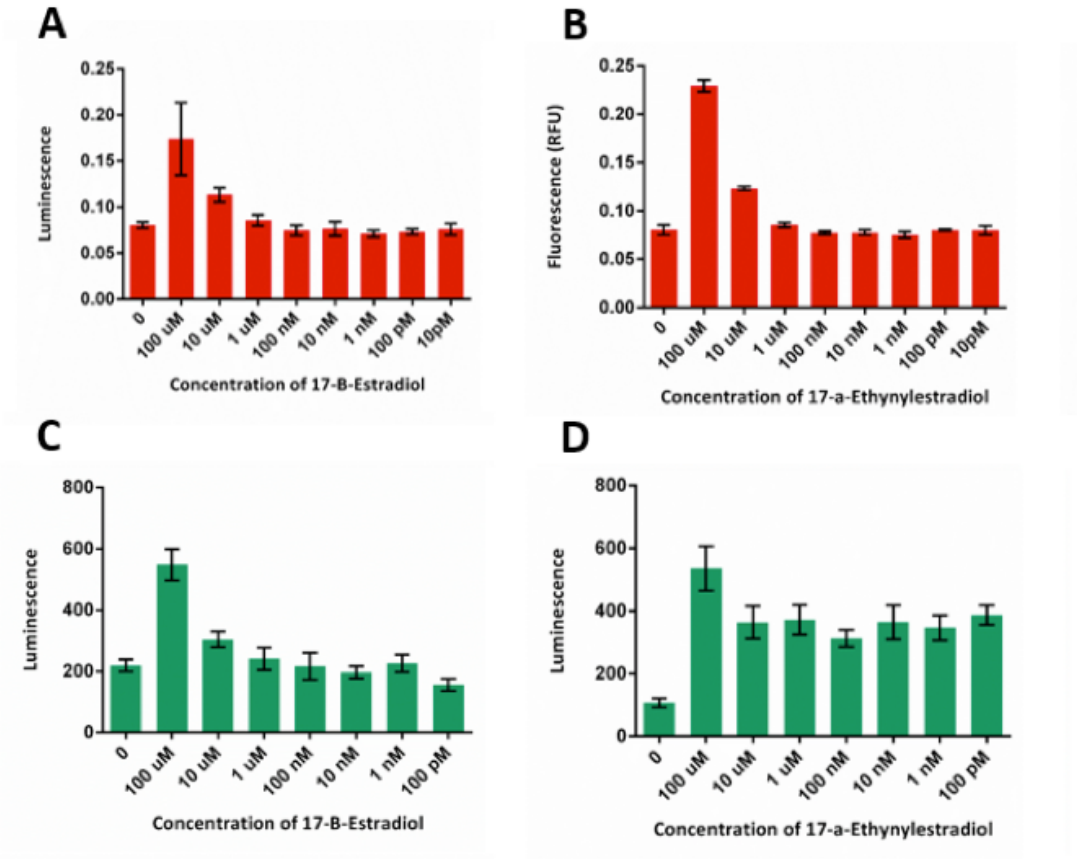


Figure 6: RFP/YFP fluorescence ratio at concentrations of (A) E2 or (B) EE2. Lower concentrations of E2 or EE2 produce lower RFU, as RFU is proportional to the concentration of estrogen that can be detected by the biological assay. At concentrations lower than 1μM, the biosensor reports similar RFU. *Gaussia* luciferase bioluminescence at concentrations of (C) E2 or (D) EE2. Lower concentrations of E2 or EE2 produce lower outputs, as light is proportional to the concentration of estrogen that can be detected by the biological assay. At concentrations lower than 14μM, the biosensor reports similar luminescence in response to E2 and at lower than 10uM in response to EE2

orescent units (RFU) of RFP was ratioed to the RFU of YFP (Figure 6).

The *Gaussia* luciferase biosensor also shows a response to the addition of two different types of estrogenic compounds, E2 and EE2 (Figure 6). Higher concentrations (100 uM) are detected more strongly whereas concentrations lower than 10uM do not seem to be easily distinguishable. It is important to note that while the readings from the RFP biosensor were very close to zero, making it more difficult to distinguish between the varying concentrations. The *Gaussia* Luciferase biosensor readings remain above 100, allowing for more sensitive detection among the varying concentrations of EE2.

IV. DISCUSSION

This report describes the development of a fast and inexpensive estrogenic responding biosensor. Fluorescent proteins and luciferases were characterized for use as reporters and an estrogen responsive T7 RNAP was developed, tested and modeled. Additionally, software, a user interface and parts were 3D printed in an effort to construct a portable, low cost luminometer/fluorimeter detection device (for a greater description see `igem2015.org/Team:Carnegie_Mellon`).

The simple protein purification protocol that was developed can have widespread application as a method to produce standards for fluorescent proteins. This protocol is also very useful as a teaching tool to familiarize scientists of all ages with the concepts and meth-

ods involved in protein purification. These cards are in use at the Citizen Science Lab in Pittsburgh, PA, USA (`http://www.thecitizensciencelab.org/`).

The RFP biosensor and *Gaussia* luciferase biosensor were both successful in detecting two different estrogenic compounds, E2 and EE2. The intracellular RFP detection of both compounds was similar, with the strongest response at 100μM and indistinguishable responses at concentrations below 10μM. The extracellular luciferase reporter detected E2 above 10μM but not below and detected the presence of EE2 at pM concentrations but could not distinguish between concentrations.

While the two biosensors responded to the presence of estrogen, there is a need to improve sensitivity and test additional estrogenic compounds. This will expand upon the functionality of the sensor as the environment is influenced by many endocrine disrupting chemicals aside from estrogen. When used in industrial field work to detect concentrations in the environment, amounts of EDCs are very low due to the large bodies of water they are found in (ppt or fM concentrations). Increased sensitivity will allow for more accurate readings and inference of potential effects they may have. To further increase the efficiency of the biosensor for industrial purposes, a device should also be designed to host the biosensor to more conveniently detect EDCs and generate real-time measured.Taking advantage of the robustness of T7 RNAP it would be of interest to utilize an *in vitro* RNA output system to have a more rapid test.

V. Contributions

Lab Experiments:Ruchi Asthana, Donna Lee, Michelle Yu, Maxwell R. Telmer, Dominique MacCalla, Wei Mon Lu, William Casazza, Kenneth Li, Cheryl A. Telmer
Modeling and Device Design: Niteesh Sundaram, Jordan Tick, Maxwell R. Telmer,Natasa Miskov-Zivanov
Manuscript Preparation: Ruchi Asthana, Donna Lee, Michelle Yu, Maxwell R. Telmer, Niteesh Sundaram, Cheryl A. Telmer
Reviewing: Cheryl A. Telmer, Natasa Miskov-Zivanov, Marcel P. Bruchez
Editing: Diana Marculescu, Dominique MacCalla, Wei Mon Lu, Jordan Tick
Funding Acquisition: Diana Marculescu, Marcel P. Bruchez, Natasa Miskov-Zivanov ,Cheryl A. Telmer

VI. Acknowledgements

We would like to acknowledge the Carnegie Mellon 2014 Team, especially Danielle Peters and Lena Wang for cloning of the original intein sensor and the RFP reporter plasmids. Funding was generously provided by Carnegie Mellon University, College of Engineering, and Mellon College of Science, Departments of Electrical and Computer Engineering and Biological Sciences, the Molecular Biosensor and Imaging Center, and the Council of Electronic Design Automation and the National Sciences Foundation. Reagents were donated by IDT, RegisTechnologies Inc. and Thermo Fisher.

References

[1] Bistan M, Podgorelec M, Marinek Logar R, Tiler T. Yeast Estrogen Screen Assay as a Tool for Detecting Estrogenic Activity in Water Bodies. Food Technol Biotechnol. 2012;50(4):427-433.

[2] Brenzel S, Kurpiers T, Mootz H. Engineering Artificially Split Inteins for Applications in Protein Chemistry: Biochemical Characterization of the Split Ssp DnaB Intein and Comparison to the Split Sce VMA Intein . Biochemistry. 2006;45(6):1571-1578.

[3] BioNetWiki [Internet]. Bionetgen.org. 2016. Available from: http://bionetgen.org/index.php/Citation˙guidelines

[4] Endocrine Disruption — US EPA [Internet]. Epa.gov. 2016. Available from: https://www.epa.gov/endocrine-disruption

[5] Faeder JR, Blinov ML, Hlavacek, WS. Rule-based modeling of biochemical systems with BioNetGen, Methods Mol. Biol. 2009; 500, 11367.

[6] Gaido K, Leonard L, Lovell S, Gould J, Baba D, Portier C et al. Evaluation of Chemicals with Endocrine Modulating Activity in a Yeast-Based Steroid Hormone Receptor Gene Transcription Assay. Toxicology and Applied Pharmacology. 1997;143(1):205-212.

[7] Gao D, Wang S, Li H, Yu H, Qi Q. Identification of a heterologous cellulase and its N-terminus that can guide recombinant proteins out of Escherichia coli. Microb Cell Fact. 2015;14(1):1-8.

[8] Hampl R, Kubtov J, Strka L. Steroids and Endocrine Disruptors History, Recent State of Art and Open Questions. The Journal of Steroid Biochemistry and Molecular Biology. 2016;155:217-223.

[9] Lei S, Lin H, Wang S, Wilcox G. Characterization of the Erwinia carotovora pelA gene and its product pectate lyase A. Gene. 1988;62(1):159-164.

[10] Liang R, Liu X, Liu J, Ren Q, Liang P, Lin Z et al. A T7-expression system under temperature control could create temperature-sensitive phenotype of target gene in Escherichia coli. Journal of Microbiological Methods. 2007;68(3):497-506.

[11] Liang R, Zhou J, Liu J. Construction of a Bacterial Assay for Estrogen Detection Based on an Estrogen-Sensitive Intein. Applied and Environmental Microbiology. 2011;77(7):2488-2495.

[12] McLachlan M, Katzenellenbogen J, Zhao H. A new fluorescence complementation biosensor for detection of estrogenic compounds. Biotechnol Bioeng. 2011;108(12):2794-2803.

[13] niteeshsun/iGEM2015. GitHub. 2016. Available from: https://github.com/niteeshsun/iGEM2015

[14] Rich R, Hoth L, Geoghegan K, Brown T, LeMotte P, Simons S et al. Kinetic analysis of estrogen receptor/ligand interactions. Proceedings of the National Academy of Sciences. 2002;99(13):8562-8567.

[15] Routledge E, Sumpter J. Estrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen. Environmental Toxicology and Chemistry. 1996;15(3):241-248.

[16] Schaerli Y, Gili M, Isalan M. A split intein T7 RNA polymerase for transcriptional AND-logic. Nucleic Acids Research. 2014;42(19):12322-12328.

[17] Shis DBennett M. Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. Proceedings of the National Academy of Sciences. 2013;110(13):5028-5033.

[18] Smith A, Xu W, Sun Y, Faeder J, Marai G. RuleBender: integrated modeling, simulation and visualization for rule-based intracellular biochemistry. BMC Bioinformatics. 2012;13(Suppl 8):S3.

[19] Team:Carnegie Mellon 2015.igem.org. Available from: http://2015.igem.org/Team:Carnegie˙Mellon

VII. APPENDIX

A. Construction of Biosensor

The US Environmental Protection Agency has established an Endocrine Disruptor Screening Program that utilizes methods requiring sophisticated reagents and equipment for detecting estrogenic activity of chemicals and one of these tests the yeast estrogen-screening assay, YES assay is used to test water for the presence of estrogenic compounds. This method employs S. cerevisiae strains containing the estrogen-binding domain of the human estrogen receptor alpha to bind to estrogen responsive elements initiating production of reporters (Routledge and Sumpter 1996; Gaido et al. 1997; Bistan et al. 2012). However, this yeast estrogen-screening assay (YES assay) is slow in detecting estrogen. It usually takes several days to incubate the reporter cells with the water samples in order to accumulate enough reporter protein and produce a measurable signal, and this is not really suitable for large-scale sample screening. There is a need for an inexpensive, fast, simple and portable test.

A more rapid assay utilized bacterial beta-galactosidase and the E. coli strain DIER to detect estrogenic compounds (Liang et al. 2010). This strain was engineered to contain a conditionally splicing intein. An intein is a splicing protein, sometimes called a protein intron. When bound to its specific molecule, conditional inteins will splice out and produce a peptide bond between the two parts of protein. This strain was engineered to contain the ligand binding domain of the human estrogen receptor within the yeast VMA intein to form an estrogen responsive intein. This intein was tested at two sites (between the Gly122 and Cys123 residues and Ala328 and Cys329 residues) in the essential region of the constitutively expressed lacZ gene (Liang et al. 2010). In the presence of estrogenic compounds like 17-β estradiol (E2), the intein would bind those, and splice out, to produce a functional beta-galactosidase enzyme that catalyzes the color change of the ONPG substrate. This assay required a two hour incubation with E2 for efficient splicing and was not very sensitive and unable to detect certain compounds such as benz[a]anthracene and pyrene. This may be due to the E. coli cell wall and transport system selectively decreasing a particular chemicals potency or remaining fully impermeable to it (Liang et al. 2010). This assay also required a substrate to produce a color change indicating the presence of estrogen.

In 2014, the CMU iGEM team initiated development of an inexpensive, fast, and simple sensor to detect molecules in water that bind to the estrogen receptor named STREAM (Sensor That Reports Endocrine Activating Molecules). To increase sensitivity, the approach was to use the robust T7 RNA Polymerase (T7 RNAP) and a fluorescent reporter. T7 RNAP is a strong viral polymerase requiring no additional factors, making its expression straightforward. T7 RNAP had been used previously with a temperature sensitive intein (Liang et al. 2007). At the permissive temperature the intein was spliced out to form functional T7 RNAP resulting in transcription from the T7 promoter to the terminator only at the permissive temperature of 18 C, but not at the restrictive temperature of 37 C. The S. cerevisiae VMA intein was inserted in between the Ala491 and Cys492 residues of the T7 RNAP. The T7 promoter was placed upstream of the lacZ gene, and was transcribed and translated resulting in blue colonies upon the production of functional T7 RNAP.

The 2014 sensor had the estrogen responsive intein inside T7 RNAP between residues 491 and 492. In the presence of estrogen, the estrogen receptor ligand binding domain would bind estrogen causing the intein to splice, functional T7 RNAP would be produced, readily bind to the T7 promoter and the transcription and translation of red fluorescent protein, would result in signal amplification in the presence of estrogen. The level

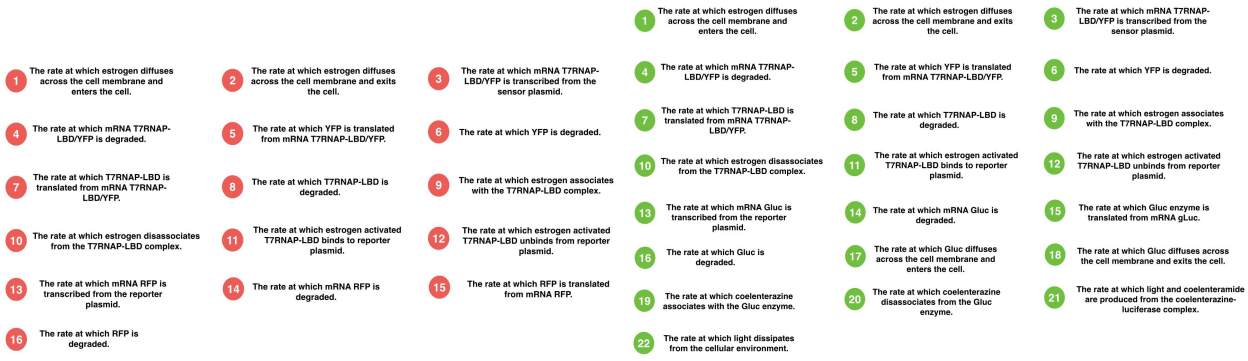


Figure 7: (Left) 16 reactions captured by the RFP model (Right) 22 reactions captured by our *Gaussia* reporter estrogen model. These images can be seen in more detail at http://2015.igem.org/Team:Carnegie˙Mellon/Modeling

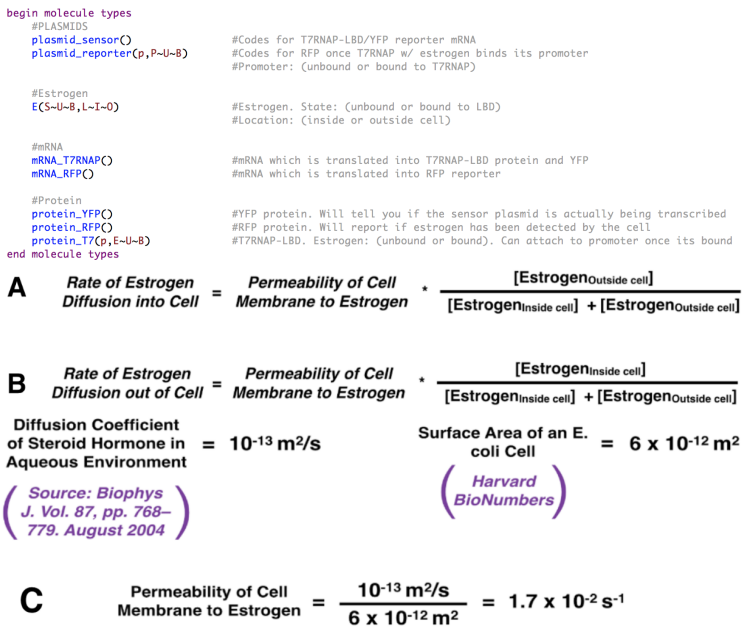


Figure 8: Class definitions were used for the components of the RFP reporter estrogen model (top). Reaction rules were then used to determine rate constants of each reaction.Equation A shows rate of diffusion of estrogen into the cell. Equation B shows rate of diffusion of estrogen out of the cell. Equation C shows the derivation of the cells permeability to estrogen.

of the estrogen-responsive intein sensor corresponds to the amount of yellow fluorescent protein and the level of estrogen, measured by the production of functional T7 RNAP, would be reported using the red fluorescent protein.

Unfortunately this sensor did not function in the presence of E2 so it was modified. The VMA and Mtu inteins (Skretas and Wood) were codon optimized and tested, in three different positions within the T7 RNAP between residues 491/492 (Liang et al. 2007), 514/515 (Schaerli et al. 2014) and 179/180 (Shis and Bennet 2007). None of these worked so the inteins were removed and the ER-LBD residues 281-595 (McLachlan et al. 2011) were tested at the 3 positions and the 179/180 gave some response. The pieces from McLachlan were tested and the 312/595 fragment showed the best response and was used in this study.

For two plasmid biosensor system consists of a reporter plasmid and a sensor component. The sensor component is a high-copy plasmid with a weak promoter that constitutively expresses the sensor with the ligand binding domain of the human estrogen receptor alpha (ER-LBD) inserted into T7 RNA polymerase (T7 RNAP) [3] and YFP for normalization. T7 RNA polymerase is a strong phage RNA polymerase that requires no additional factor. YFP is found on the same transcript as the sensor to serve as a baseline to ratio the RFP. The RFP reporter plasmid for the RFP biosensor consists of a T7 promoter driving expression of RFP.

When estrogen binds to the ER-LBD of the sensor protein, it causes a conformational change. Separated domains of T7 RNAP reconstitute and T7 RNAP activity is restored to bind the T7 promoter on reporter plasmid, allowing transcription of RFP mRNA that is then translated to produce RFP.

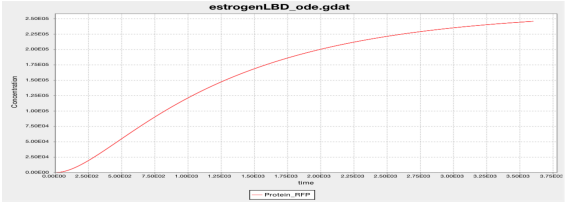
B. Modeling Methods

In order to translate these written definitions to code, we instantiated class definitions. These class definitions are useful in informing the model as to which cellular components need to be included, and which components are not being directly analyzed. We created class definitions to specify the names of the components as they appear in the code, all possible binding sites of the components as well as to which molecules they can bind, and the possible states of the components (e.g. phosphorylation state, methylation state, etc.). Estrogen is represented as E (), while the full class definition for estrogen is E(S∼U∼B,L∼I∼O). The S∼U∼B, indicates that the state (S) of estrogen can either be unbound (∼U) or bound (∼B) to the ligand binding domain of T7 RNA polymerase, and the L∼I∼O indicates that the location (L) of estrogen can either be inside (∼I) or outside (∼O) of the cell.

After class definitions were created, the components that represented the concentration of estrogen outside the cell and concentration of reporter and sensor plasmid inside the cell were initialized to predetermined values in order to begin simulation. For components whose class definition includes multiple states, such as estrogen (E(∼S∼U∼B,L∼I∼O)) and reporter plasmid ((p,P∼U∼B)), each particular instance must be initialized.

Before assembling the reaction rules, it is important to determine the rate constants of each reaction. Rate constants for the rate estrogen associates/disassociates to the LBD were obtained from R. L. Rich et al. Rate constants for the rate T7 RNA Polymerase associates/disassociates to the RFP plasmid were obtained from S. Brenzel et al. Finally rate constants concerning the rate of RFP and YFP transcription, translation, and degradation were obtained from experimental data from the 2009 PKU-Beijing iGEM team (http://2009.igem.org/Team:PKU˙Beijing). Rate constants for the diffusion of estrogen to cell and from the cell were calculated for by equations A and B (Figure 8).

A final step in constructing the rule-based model is to compile the aforementioned information into single lines of executable code, which correspond to a reaction rule. It should be noted that each line of code will be turned into a differential equation or set of differential equations that can be run for a specified number of trials. Due to the fact that our models were based on experimental data, we were able to predict the outcome of experimental wet-lab trials under a variety of different conditions. The predicted outcomes were also useful in guiding wet-lab experiments and giving insight into some of the biological underpinnings of experimentation. For example, the RFP reporter model demonstrated that the reason the signal produced by the 2014 version of the sensor was due to the fact that the old sensor simply did not possess a high enough concentration of sensor plasmids per E. coli.

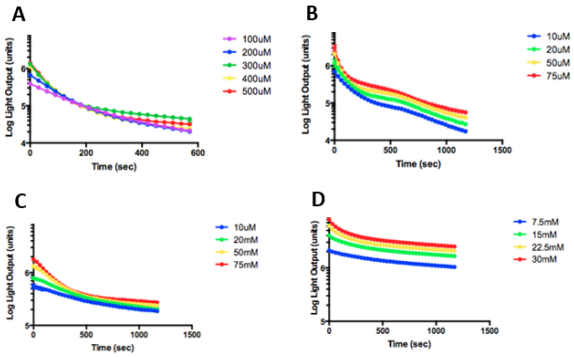


Above is graph that demonstrates the biological results of this model.

The modified sensor contains an estrogen responsive T7 RNAP allowingt it to produce a detectable RFP signal. The units for the x-axis are time (minutes), and the units of the y-axis are concentration (nM). Unlike 2014s sensor, the amount of RFP produced within a few hours is significantly greater than the minimum threshold detection of 100 nM.

C. Luciferase and Fluorescent Protein Characterization

Figure 8 shows details from the fluorescent protein quantification protocol.



Above are time courses for each luciferase and their respective substrates. Higher concentrations of substrate resulted in higher light output. Higher concentrations of substrate plateaued at higher output levels (A) CDcel-*Gaussia*, (B) PelB-*Gaussia*, (C) *Renilla* and (D) Firefly.

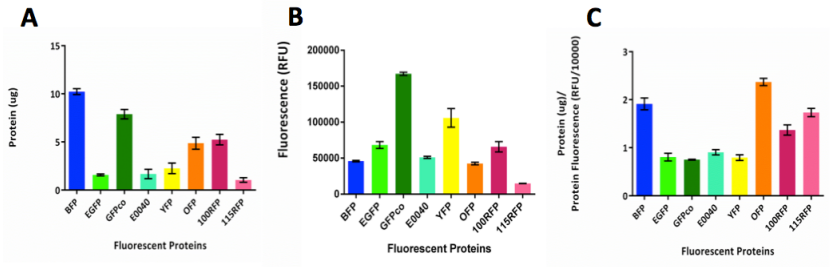


Figure 9: (A) Total protein isolated from 1mL of culture expressing various fluorescent proteins (B) Total fluorescence of proteins isolated from 1 mL of culture (C) Ration of protein amount per 10,000 RFU.