Real-Time Optogenetics System for Controlling Gene Expression

Guy Soffer

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science (Electrical and Computer Engineering) at

Concordia University

Montreal, Quebec, Canada

March 2020

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Guy Soffer

Entitled: Real-time Optogenetics System for Controlling Gene Expression

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Electrical and Computer Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Chair

Dr. S. Hashtrudi-Zad

External Examiner

Dr. L. Potvin-Trottier (BIOL)

Internal Examiner

Dr. S. Hashtrudi-Zad

Supervisor

Dr. S. Shih

Approved by:

Dr. Y.R. Shayan, Chair

Department of Electrical and Computer Engineering

31/March/2020

Dr. Amir Asif, Dean,

Faculty of Engineering and Computer Science

Abstract

Real-Time Optogenetics System for Controlling Gene Expression

Guy Soffer

Synthetic biology, which combines tools from engineering, molecular biology and computation, aims to explore the biological systems available in nature and engineer new ones that can better serve our needs. Yet, there are many challenges in designing and in building a biological system that will perform the desired functions. Optimization of these newly engineered systems requires careful control over the levels and timing of gene expression. Optogenetic tools are ideal to dynamically control gene expression since only light is required to enable precise control and it provides a convenient interface between digital control software and microbial culture. In this thesis, I present a fully automated closed-loop system which controls gene expression and growth rates in *Escherichia coli* by using the engineered CcaR-CcaS two component system that is activated or repressed by illumination of green or red light respectively. Using an in-house designed computer-controlled light system, we periodically record culture density and fluorescence measurements related to open- and closed-loop control of green fluorescent protein expression. We also verify our experimental results with simulation models - with and without feedback. These results presented here highlight the efficacy of our system, illustrating how it will potentially be used for other applications that require gene expression control such as metabolic pathway engineering and designing biofuel-producing microbes.

Acknowledgments

Many thanks to Dr. Steve Shih for giving me the opportunity to join his research group. I feel privileged to have become part of an interdisciplinary scientific group that creatively bridges electronics, software and biology. I also thank Dr. Shih for giving me the academic freedom to develop my system, and for his useful and interesting comments and insight. Special thanks to James Perry, for helping me with the biological aspects of this project, and for teaching me how to hold a pipette… I learned a lot from him and truly enjoyed our work together and our engaging conversations. I feel I've earned a true friend and an excellent colleague. I would also like to thank all lab members, Ehsan, Kenza, Alaa, Angela, Laura, Fatemeh, Rob and Sam for their pleasant company and collaborative atmosphere. Thanks to them I enjoyed coming to lab every day. I thank Dr. Shahin Hashtrudi Zad for his excellent course on real time system control, and Dr. Nawwaf Kharma for his excellent course on computational synthetic biology. I enjoyed these courses very much and found them extremely useful for my work.

On a personal note, traveling overseas to complete a graduate degree would not have been possible without the support of my dearest family. I thank my beloved parents, Ran and Ayelet, for raising me in a loving and stimulating environment, in a family that encourages curiosity and enjoys technology. And my brother, Elad, and my uncle, Dan, for their close friendship and inspiring conversations and for sharing common fields of interests. My personal and professional growth would not have been the same without them. A huge thanks to my wonderful aunt Rosie who, with endless generosity, hosted me in her house during my studies in Canada and guided me through the freezing Canadian winter. A huge and special thanks to my beloved and wonderful wife – Einav, and children – Tzur, Nave and Tal. My wife, who carried, single handedly, the family's burden while I was away, and still found the time give me her professional advice. I would not have been able to write this thesis without her invaluable comments and input. Her encouragement and support made it all possible. And my brave children, who remind me of the important things in life. They bring much joy and inspiration to my life. I love them very much and feel extremely fortunate to have such a terrific family.

Overview of Chapters

This thesis describes a project I conducted under the supervision of Dr. Steve Shih at Concordia University. It is submitted in partial fulfillment of the requirements for my Master of Applied Science. This thesis includes a detailed description of a fully automated system that can measure, control, and regulate gene expression rate of a specific gene using a photo-regulated biological system. Below I describe the breakdown of my thesis.

Chapter 1 describes the scientific and historical background of synthetic biology, as well as a detailed description of some of the tools implemented in this project, such as the CcaR-CcaS Two Component System. I also describe the main objectives of this project.

Chapter 2 provides a description of methods and materials used in this project. Specifically, all the biological protocols for sample preparation, modeling requirements (e.g., software and equations), and procedures for controlling gene expression under open- or closed-loop control.

Chapter 3 provides a complete description of my system including the mechanical and electronical design, mathematical modeling and controller design. In addition, it contains a detailed description of 'wet lab' experimental results in open- and closed- loop experiments, using 96-well plates and microfluidics device.

Chapter 4 describes future directions and improvements. It ends with concluding remarks.

Appendix A describes the controllers stability using the root-locus method.

Appendix B describes a side project done in digital microfluidics. The software infrastructure and the two electrical boards developed as part of this project were used in published work.

Overview of Author Contributions

This work described in my thesis was made possible with the help of colleagues from the Shih Lab. Here, I outline the contributions that each author made to the work.

The research was designed by Dr. Steve Shih, James Perry, and myself. The research article relevant to this work was written and edited by Dr. Steve Shih, James Perry, and myself. All experiments and analysis were conducted by Dr. Steve Shih, James Perry, and myself. The resulting figures were revised and approved by Dr. Steve Shih. Bacterial culture work (cell passaging, maintenance) and biological methods were performed by James Perry. The mathematical modeling, simulations and controller design were done by myself. The mechanical, electrical and software design were developed by myself.

Table of Contents

List of Figures

List of Equations

List of Abbreviations

API – Application-Protocol-Interface

ADC – Analog-To-Digital

CAD – Computer-Aided-Design

CcaSR – CcaS-CcaR system

CNC – Computer-Numerical-Control

CV – Computer-Vision

CMOS – Complementary Metal Oxide Semiconductor

DAC – Digital-To-Analog

DNA –Deoxyribonucleic Acid

EWOD – Electro-Wetting-On-Dielectric

FOH – First-Order-Hold

FPGA – Field-Programmable-Gate-Array

GRN – Genetic-Regulatory-Network

GFP – Green-Fluorescence-Protein

sfGFP – superfolded Green-Fluorescence-Protein

IP – Internet-Protocol

IR – Infra-Red

LED – Light-Emitting-Diode

LSB – Least-Significant-Bit

LQR – Linear-Quadratic-Regulator

MBD – Model Based Design

MCU – Micro-Controller-Unit

MPC – Model-Predict-Controller

mL – Milliliter

mRNA – Messenger-Ribonucleic-Acid (Macromolecules Messenger Ribonucleic Acids)

ms – Millisecond

OD – Optical-Density

ODE – Ordinary-Differential-Equation

OS – Operating-System

PC – Personal-Computer

PCB – Printed-Circuit-Board

PDMS – Polydimethylsiloxane, also known as dimethylpolysiloxane or dimethicone

PID – Proportional-Integral-Derivative

PLA – Polylactic-Acid or Polylactide

PSU – Power-Supply-Unit

PWM – Pulse-Width-Modulation

RAM – Random-Access-Memory

RNA – Ribonucleic-Acid

RTOS – Real-Time-Operating-System

TCP – Transmission-Control-Protocol

TCS – Two-Component-System

TEC – Thermo-Electric-Coupler.

UDP – User-Datagram-Protocol

µL – Microliter

µS – Microsecond

USB – Universal-Serial-Bus

ZOH – Zero-Order-Hold

List of Foundations and Funding Sources

I would like to thank the following funding sources for providing financial support to this project:

- Natural Sciences and Engineering Council (NSERC)
- Fonds de Recherche Nature et Technologies (FRQNT)
- Canadian Foundation of Innovation (CFI)

I would also like to thank the following institutions and organisations for their contribution for funding and resources:

- Concordia University Department of Electrical and Computer Engineering for FRS funding and academic resources.
- Centre for Applied Synthetic Biology (CASB) for equipment and technical support.
- Concordia University Department of Biology for academic resources and for their tissue culturing facility.

List of Co-Authored Publications

Integration of World-to-Chip Interface with Digital Microfluidics.

Ehsan Moazami[1,2],James Perry[2,3],**Guy Soffer[1,2]**,Mathieu C. Husser[2,3],Steve C.C. Shih[1-3*]

[1]Department of Electrical and Computer Engineering, Concordia University, Montréal, Québec, Canada

[2]Centre for Applied Synthetic Biology, Concordia University, Montréal, Québec, Canada

[3]Department of Biology, Concordia University, Montréal, Québec, Canada

Abstract

Digital microfluidics (DMF) represents an alternative to the conventional microfluidic paradigm of transporting fluids in enclosed channels. One of the major benefits of DMF is that fluid motion and control is achieved without external pumps. The automation component of DMF have pushed the barriers of this "lab-on-chip" technology. However, integration with external components (i.e., "world-to-chip") interfaces have been a challenge. Two common "world-to-chip" challenges are (1) delivering biological samples to DMF devices and (2) accurately controlling temperatures on device. To address these challenges, this work describes two "world-to-chip" interface features that have been integrated on a DMF platform: a reagent delivery system and a thermal control apparatus. This platform enables a variety of biological or chemical experiments to be conducted on-chip while reducing manual intervention. Specifically, our platform increases reagent volumes available to device reservoirs volume by at least 50-fold eliminating the need to manually refill reservoirs while improving droplet dispensing reproducibility. In addition, we have integrated a closed-loop temperature control system that offers precise temperature control on-chip. To validate our "world-to-chip" interface, we have automated bacterial transformation and enzymatic assay protocols, showing that such a system enhances DMF performance. Overall, we propose that this system will improve biological experimentation which requires fluidic and temperature control integrated on DMF platforms.

Moazami, E., Perry, J.M., Soffer, G., Husser, M.C. and Shih, S.C.C., 2019. ***Integration of World-to-Chip Interfaces with Digital Microfluidics for Bacterial Transformation and Enzymatic Assays.*** *Anal. Chem., 91(8) 5159-5168.*

A fucosyltransferase inhibition assay using image-analysis and digital microfluidic.

Laura M. Leclerc[2,3], **Guy Soffer[1, 2]**, David H. Kwan[1-3], Steve C.C. Shih[1-3]*.

A fucosyltransferase inhibition assay using image-analysis and digital microfluidics.

1 Department of Electrical and Computer Engineering, Concordia University, Montréal, Québec, Canada
2 Centre for Applied Synthetic Biology, Concordia University, Montréal, Québec, Canada
3 Department of Biology, Concordia University, Montréal, Québec, Canada

Abstract

Sialyl-LewisX and LewisX are cell-surface glycans that influence cell-cell adhesion behaviors. These glycans are assembled by α(1,3)-fucosyltransferase enzymes. Their increased expression plays a role in inflammatory disease, viral and microbial infections, and cancer. Efficient screens for specific glycan modifications such as those catalyzed by fucosyltransferases are tended toward costly materials and large instrumentation. We demonstrate for the first time a fucosylation inhibition assay on a digital microfluidic system with the integration of image-based techniques. Specifically, we report a novel lab-on-a-chip approach to perform a fluorescence-based inhibition assay for the fucosylation of a labeled synthetic disaccharide, 4-methylumbelliferyl β-N-acetyllactosaminide. As a proof-of-concept, guanosine 5′-diphosphate has been used to inhibit Helicobacter pylori α(1,3)-fucosyltransferase. An electrode shape (termed "skewed wave") is designed to minimize electrode density and improve droplet movement compared to conventional square-based electrodes. The device is used to generate a 10 000-fold serial dilution of the inhibitor and to perform fucosylation reactions in aqueous droplets surrounded by an oil shell. Using an image-based method of calculating dilutions, referred to as "pixel count," inhibition curves along with IC50 values are obtained on-device. We propose the combination of integrating image analysis and digital microfluidics is suitable for automating a wide range of enzymatic assays.

*Leclerc, L.M., **Soffer, G**., Kwan, D.H. and Shih, S.C. 2019. **A fucosyltransferase inhibition assay using image-analysis and digital microfluidics**. Biomicrofluidics 13(3), p.034106.*

One cell, one drop, one click: hybrid microfluidic mammalian single-cell isolation

Kenza Samlali[1, 2], Fatemeh Ahmadi[1, 2], Angela B.V. Quach[2,3], **Guy Soffer[1, 2]**, Steve C.C. Shih[1-3]*

1 Department of Electrical and Computer Engineering, Concordia University, Montréal, Québec, Canada

2 Centre for Applied Synthetic Biology, Concordia University, Montréal, Québec, Canada

3 Department of Biology, Concordia University, Montréal, Québec, Canada

Abstract

The process of generating a stable knockout cell line is a complex process that can take several months to complete. In this work, we introduce a microfluidic method that is capable of isolating single cells, electing successful edited clones, and expansion of these isoclones. Using a hybrid microfluidics method, droplets in channels can be individually addressed using a co-planar electrode system. In our hybrid microfluidic device, we show that we can trap single cells and subsequently encapsulate them on demand into pL-sized droplets. Furthermore, individual cells inside the droplet can be released from the traps or merged with other droplets by simple application of an electric potential to the electrodes that is actuated trough a user interface. We use this high precision control to successfully sort and to recover single isoclones to establish monoclonal cell lines, which is demonstrated with a heterozygous NCI-H1299 lung squamous cell population resulting from loss-of-function eGFP and RAF1 gene knock-out transfections.

*Submitted to **bioRxiv** on January 24, 2020*

1 Introduction to Synthetic Biology and Optogenetics

In this section, I will introduce general concepts in synthetic biology and optogenetics as a method to control gene expression. I will also describe basic concepts regarding system control theory and model-based design methodology.

1.1 Synthetic biology

Synthetic biology is the process of engineering biological systems. The field of synthetic biology describes a multidisciplinary, engineering-based approach using a design-build-test-learn cycle to study, model, and re-design biological systems[1]–[3]. One approach of synthetic biology is to think of a living cell as a circuit of connected biological parts with the ability to process logical operations that will translate to a biological function. Combining engineering and biology has inspired creation of biological regulation networks which models biological function by using electrical circuit theory [3], [4]. Advances in molecular biology techniques such as DNA sequencing and DNA synthesis sparked by large scale endeavors such as The Human Genome Project [5], which have enabled major progress in the past 20 years. These techniques have set the foundations of a new field in biology, namely systems biology, which is aimed at analyzing and modeling complex biological systems to better understand cellular signaling and metabolic networks [6], [7].

Yet, in contrast to classical genetics, which includes forward genetic approaches (studying function by gain- or loss- of a gene), or reverse genetics (looking for an unknown gene, based on a given phenotype), synthetic biology studies biological functions and networks outside of their 'natural' biological context. A common approach is to introduce new biological functions to an organism by modifying the DNA of the genome to provide the metabolism to generate valuable bio-chemicals such as pharmaceuticals[8] and biofuels[9] or to break-down harmful chemicals in the environment [10]. Another example is editing human cells to possibly eradicate growing tumors related to cancer [11]. An important step in the synthetic biology process is to search for biological parts to broaden our "synbio" toolbox, as well as contribute to our understanding of the biological system[1], [12]–[14].

In addition to modifying existing species, efforts have been made to generate completely new synthetic life forms[15]–[17]. For example, synthetic biologists are often engaged in the engineering of novel biological circuits and regulatory networks. These circuits can be designed to generate certain

functions, which can also be tested against model prediction. The most successful and well-known design and construction of a synthetic gene networks are the genetic toggle switch[18], [19] and the repressilator[20]. These networks have shown to be equivalents of electronic memory storage and timekeeping respectively, to mimic natural behaviors of the cellular system. These networks have been used to model several natural processes. For example, bi-stability is modeled by mutual inhibition of the lysogenic cycle in bacteriophage[21] while a ring oscillator models the periodic circadian clocks in photosynthetic bacteria[22] and other types of eukaryotic cells[23], [24]. Within this framework, these fundamental components allow synthetic biologists to build more complex systems by obtaining well-defined DNA parts from other natural models to construct additional genetic switches, oscillators[25], [26], as well as other "electronic"-inspired genetic devices and elements[27]–[29] that can be used to expedite our understanding of biological functions.

1.2 Why do we need control over of gene expression?

Gene expression is first controlled at the level of transcription. Regions of DNA upstream from a target gene, called promoters, initiate transcription by providing a binding site for RNA polymerase. In conjunction with the promoter sequences, adjacent regions of DNA, called operators, selectively bind proteins called transcription factors. The binding of transcription factors changes the secondary structure of DNA, either enhancing or restricting the ability of RNA polymerase to bind to the promoter. Chemicals called inducers bind and alter the transcription factor's DNA binding affinity. The availability of these chemical inducers thus have control over the timing and rate of transcription of a target gene and subsequent protein production. Chemically regulated or *inducible* promoters are commonly used to turn ON or OFF expression of a target protein[30]. Controlling gene expression plays a critical role in maintaining a precise biochemical balance for cell function and survival [4], [31].

In the context of synthetic biology experiments, exogenous DNA is introduced into a cell which provides new metabolic capabilities such as waste-chemical degradation [32], production of biofuels[33], nutrients [34] or drugs [8], [35]. However, if production of these target proteins is not well controlled, they can become a metabolic burden or toxic, resulting in bottlenecks in the target pathway or lead to a buildup of toxic precursors. Therefore, control over exogenous gene expression is essential for maintaining healthy cells and efficient production.

1.3 Optogenetics and the CcaS-CcaR (CcaSR) Two-Component-System

Most genetic networks (e.g., the genetic toggle switch) are either activated or repressed by chemical-based inducers or other environmental signals to enable gene expression. For example, in *E.coli* bacteria, the Lac operon consists of a set of genes that are regulated by a combination of glucose and lactose inducers, which leads to efficient expression and activation of relevant genes and transporters required in lactose metabolism [36]. Although induction by chemicals have been frequently shown in literature[31], their high costs and scalability make them unsuitable for downstream processes like fermentation. Furthermore, their controllability and programmability are very difficult to fine tune which will produce highly variable outputs. Recently, in cyanobacteria, a regulatory network contains light-responsive elements that is capable to induce gene expression without the administration of chemical-based inducers [36]. Using light as an inducer can eliminate the drawbacks of using chemical-based inducers. For example, using light as an inducer is easily integrated via automation – i.e. light can be controlled via switches and a computer. Using an automation scheme, light can easily control the behavior of biological circuits inside cell by a computer-generated sequence of light of different wavelengths. The ease of controllability can tremendously fine tune the programmability of the cells simply by a 'click' of a button. Other key advantages are that we avoid the possible toxic effects of chemical inducers on genes and light enables precise modulation of the cells without changing the media concentration in which cells are growing in.

One system that uses light-sensitive transcription factors is the CcaS-CcaR-v2 (CcaSR), which is a two-component gene regulatory network. In the CcaSR system, the CcaS protein undergoes distinct conformational changes in the presence of different wavelengths of light. Under 535 nm light (green), also known as the "Pg state" the protein undergoes phosphorylation. Under 670 nm light (red), also known as "Pr state" the protein will de-phosphorylate. This is described in Hirose et al. [37] and is previously shown to be a binary system [38], [39], [40]. In the green-light "Pg state", a phosphate group is transferred from the CcaS protein to the CcaR protein which enables CcaR to bind to the promoter, initiating gene expression **(Figure 1.1)**. Transcription activation leads to the expression of a reporter gene – sfGFP (superfolder Green-Fluorescence-Protein), which is a variant of the GFP protein, with improved folding kinetics. However, under exposure of red light, phosphorylated CcaR in not maintained, and gene expression is arrested [41].

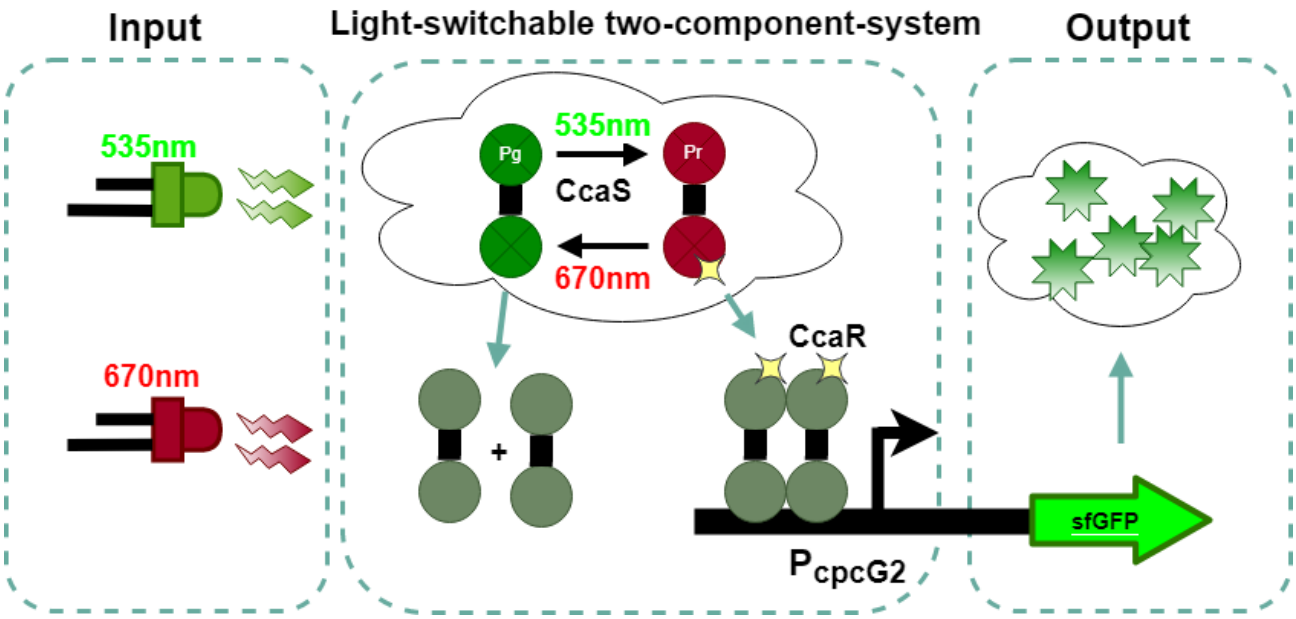

Figure *1*.1: The CcaS-CcaR gene regulatory network model.

A schematic showing the mechanism for CcaS-CcaR-v2 light switchable two component system. The green light increases the transfer rate of the phosphate between the CcaS to CcaR molecules. When the phosphate binds to the CcaR molecule it can bind to the cpcG2 promoter and enable the transcription process. Following exposure to red light, de-phosphorylation of CcaR occurs and as a result the gene transcription is arrested[41].

1.4 System-control theory: open- vs. closed- loop systems

An open-loop system does not monitor its output nor correct for disturbances. Typically, it consists of an input signal that drives the process, or plant, while external disturbances are usually added into the system. The distinguishing characteristic of an open-loop system is that it cannot monitor the condition of its output signal nor does it compensate for any errors between the two (**Figure 1.2A**). Additionally, open-loop systems are vulnerable to external disturbances since there is no feedback. However, a closed-loop system works by having the plant's input (G(s)) be a function of the plant's output response through a feedback path and a controller (**Figure 1.2B**). The output signal – which is the comparison between the controller's input to the feedback signal (i.e. the error signal – E(s)) - is processed by the controller (K(s)) to generate a new output signal to the plant to reduce the error.

A special case of a closed-loop system is a system that already has an internal closed-loop topology. In fact, most closed-loop systems have a complex multi-loop system(s) nested inside them. Even in these cases, an additional controller can be designed and implemented as an additional loop over the subsystem.

If properly designed, the additional controller can improve the overall performance of the system in criteria such as, reaching the target value quicker, making it stable or more accurate. (**Figure 1.2C**).

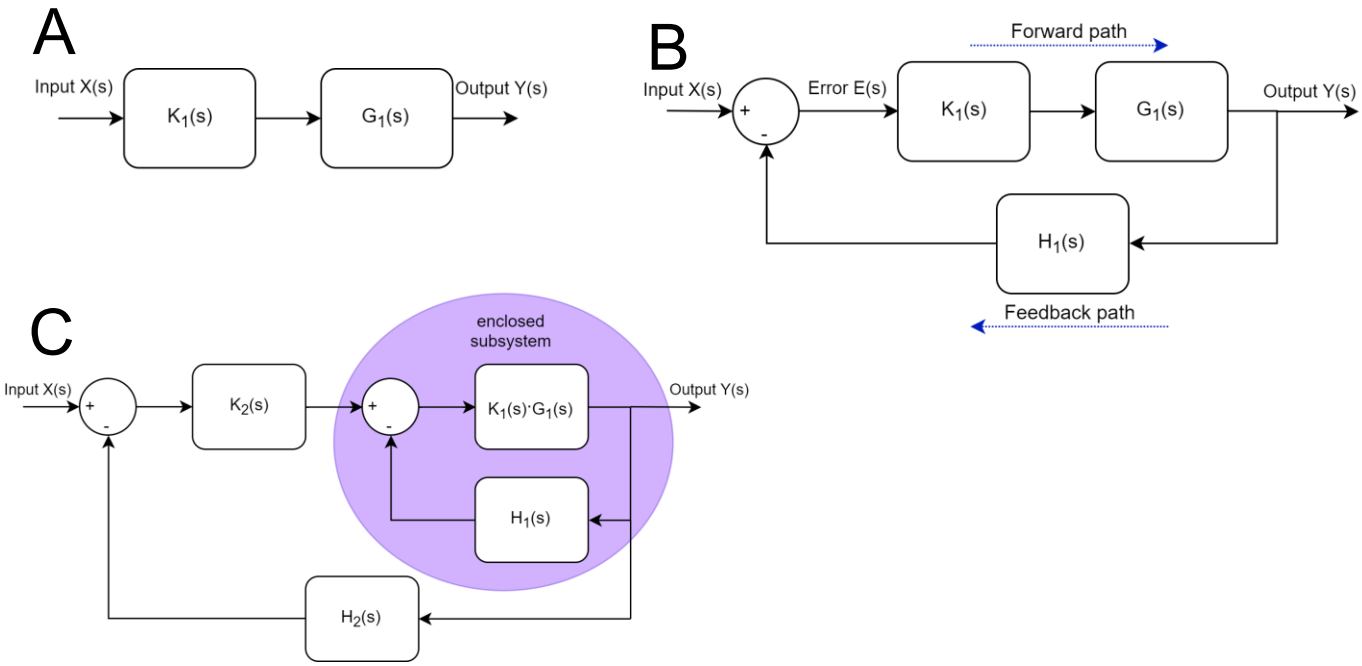

Figure 1.2: Open-loop vs closed-loop topologies.

(A) Open-loop topology – the controller K_1 is driven by an input signal X and the output is driving the plant G_1. **(B)** Closed-loop topology – the controller K_1 is driven by the error signal E and its output is driving the plant G_1. The plant output Y is converted to a feedback signal by H_1 and subtracted from the input signal X to generate the error signal E. **(C)** Closing a loop on a closed-loop subsystem – a closed-loop controller K_2 is added over an enclosed closed-loop subsystem.

1.5 Closed-loop Controller: The proportional–integral–derivative (PID) controller

A commonly used closed-loop controller is the proportional, integral, and derivative (PID) controller. The controller receives the error input which is the difference between the setpoint and the feedback output. The PID controller is typically given by the equation 1.1 and shown in **Figure 1.3**.

$$\boldsymbol{y(t) = K_p \cdot e(t) + K_i \cdot \int_0^t e(t) \cdot dt + K_d \cdot \frac{d}{dt} e(t)}$$

Equation (1.1) Basic PID controller

This formula describes how the control signal is the sum of three terms: (1) a term that is proportional to the error, (2) a derivative term that is proportional to the derivative of the error, and (3) an integral term that is proportional to the integral of the error. Each of these terms have gain parameters that can be tuned to reach our desired output. After summing these terms, the signal is then fed-forward to the Plant to control the desired output.

In synthetic biology, there have been previous works that have shown the use of PID controllers to control the gene expression in a biological system[42], [43], [44]. Recent studies have demonstrated quantitative control of gene expression using an external real-time controller which monitors the biological system and periodically induces the promoter to minimize the error between the desired and measured gene expression [42], [44]. Furthermore, some of the engineered regulatory genetic networks use closed-loop topologies - the toggle switch[19], represillator[24], and the JK-flipflop[27], [28]. Although these studies use mathematical tools to predict stability (or instability) of the system, the majority of their research abstains from working in the dynamic region of the feedback signal. This means the output feedback signal is used to saturate the input to achieve full activation or deactivation of a specific a gene. Currently none of these regulatory genetic networks show a continuous nature of operation, i.e. 'close the loop' to track a quantitative signal with constant or time varying behavior.

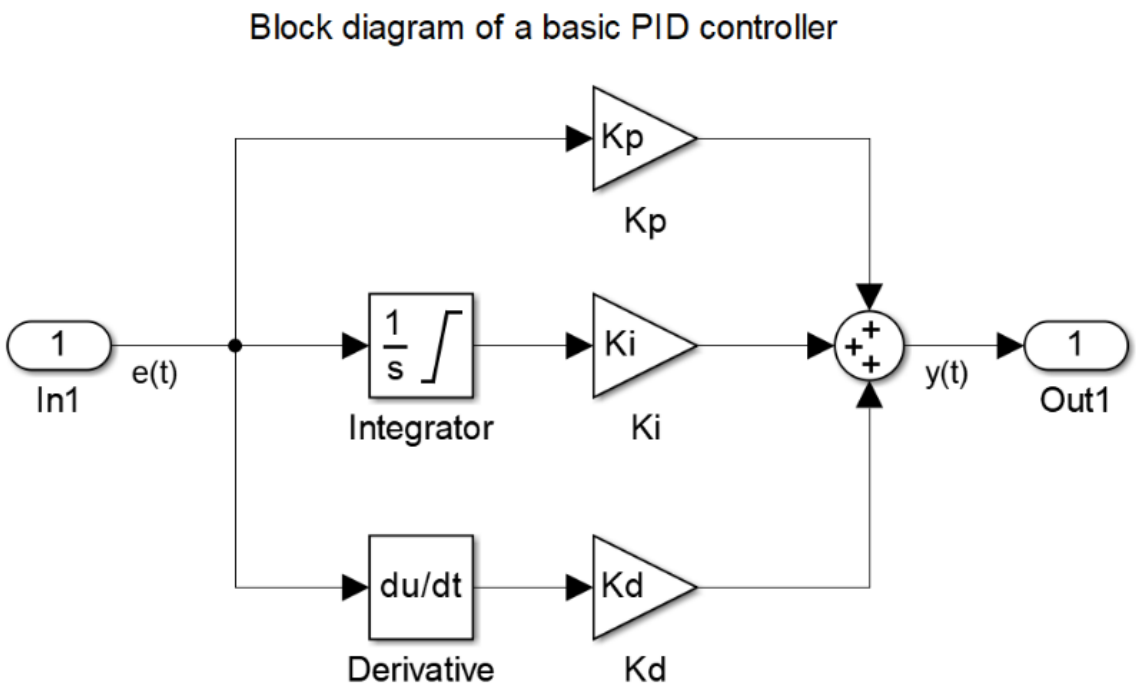

Figure 1.3: Basic PID controller block diagram.

The block diagram of a basic PID controller created by MATLAB's Simulink software. The input to the PID controller is the error signal e(t). The output of the controller is a superposition of three terms. The proportional term is the error value multiplied by the factor K_P. The integral term is the integrated error signal multiplied by the factor K_I. The derivative term is the derivative of the error signal multiplied by the factor K_D. Equation 1.1 is the mathematical representation of Figure 1.3.

One strategy to understand the dynamic behavior from these genetic systems is to use computational tools[31], [45]–[47] to model the cell behavior based on experimental and theoretical knowledge [26], [48], [49]. Despite the benefits of modeling cells, the complexities within a cellular system restricts our ability to build complete and accurate models. The main reason is that for biological systems that have known parts, it is still very difficult to understand precisely how these parts interact to make the desired process work[27], [28]. Since there is genetic variation between individual cells and differences across the micro-environment, there is constant dynamic changes in the physical and chemical surroundings which can cause differences in their cellular activity.[23], [50] However, quantitative models are still useful in determining *a priori* data sets that can be used as predictors to identify target gene or proteins and can serve as a basis of tools to aid in the optimization of existing biological systems and development of new designs.

1.6 Model based design

Model-based design (MBD) methodology simplifies the development of large-scale systems by enabling developers to test and to verify their solution at early stages during the design, rather than prototyping the real system. The main emphasis is placed on the modeling of the plant (i.e. the process of the system). This is done by using continuous-time and discrete-time building blocks which can be quickly and accurately evaluated with analytical or numerical tools. The model can also be used to develop real-time simulation tools that can further accelerate the development of the software and algorithms (i.e. hardware in the loop). In complex or slow-reactive systems, the MBD approach is quicker and much more efficient when compared with traditional design methodology (e.g. prototyping).

1.6.1 The central dogma in molecular biology model

The typical modeling approach in synthetic biology is to view each cell as a collection of modules – which can be represented as individual genetic elements. In a living cell, the DNA contains the genetic information that will be transcribed into messenger RNA (mRNA) which will be used as a template to produce proteins (from 20 different amino acids) by the process of translation. The proteins that are produced will either activate/repress the same or other transcription processes to control gene expression (or other biological functions) in the cell[3], [4], [31]. At the same time, mRNA and proteins will also go through the process of degradation and have played a large role in the design and control of a synthetic network[51], [52]. The block diagram describing this model is shown in **Figure 1.4**.

In synthetic biology, the cell can be modeled as a collection of modules. The typical modeling approach in synthetic biology is to view each cell as a collection of modules – which can be represented as individual genetic elements. In a living cell, the DNA contains the genetic information that will be transcribed into messenger RNA (mRNA) which will be used as a template to produce proteins (from 20 different amino acids) by the process of translation. The proteins that are produced will either activate/repress the same or other transcription processes to control gene expression (or other biological function) in the cell[3], [4], [31]. At the same time, mRNA and proteins will also go through the process of degradation and have played a large role in the design and control of a synthetic network [51], [52]. The block diagram describing this model is shown in **Figure 1.4**.

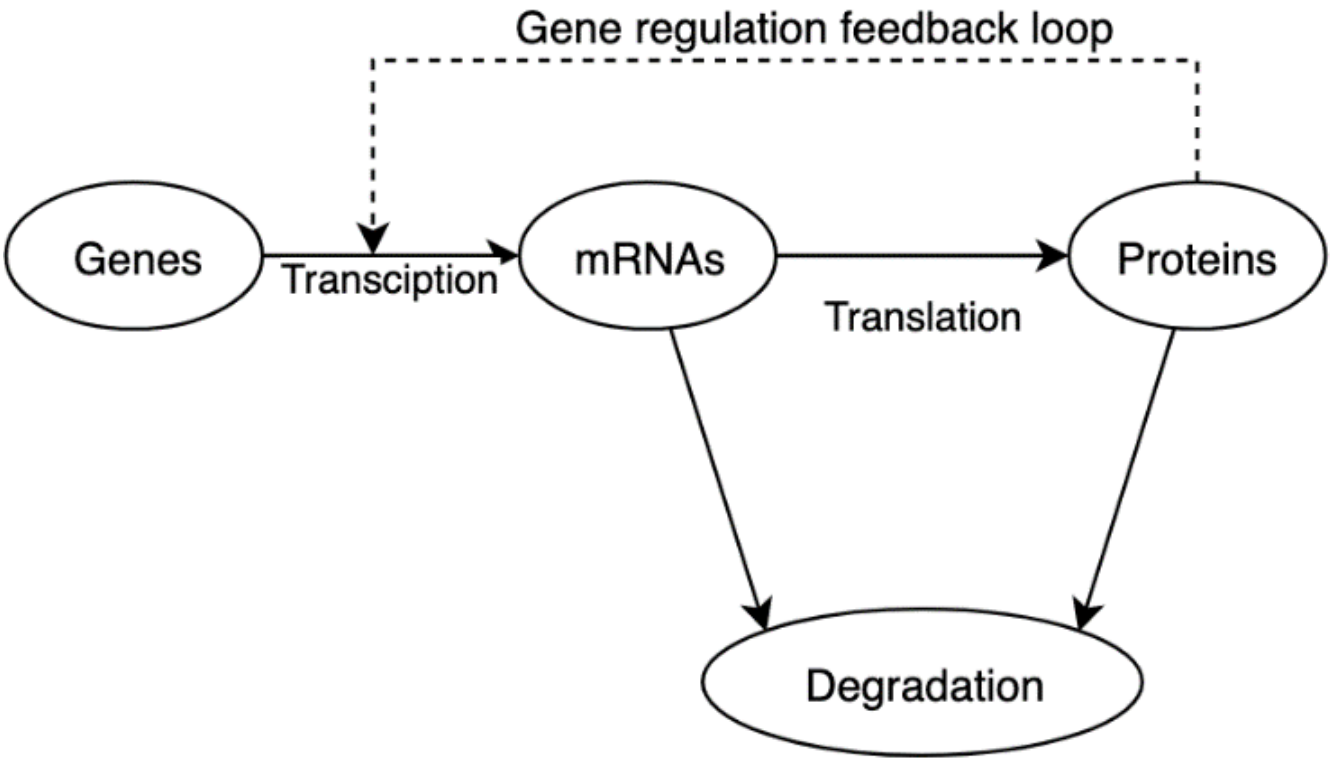

Figure 1.4: Simplified regulation of gene expression.

A process schematic showing transcription of a gene which produces an mRNA that is translated to protein. The mRNA and protein experience natural decay at different rates. The proteins might act as positive (induce) or negative (repress) feedback on the same or other genes in the cell.

1.6.2 Mathematical modeling of genetic regulatory network

A significant effort has been placed on modeling to simulate the processes in biological cells[45], [47], [53]–[58]. The most successful are the models that are developed through thermodynamic or kinetic differential equations [47], [57]. Thermodynamic equations are typically used to predict the interactions between molecules in the cells and has been frequently used for predicting metabolic pathways[59]–[61]. Kinetic equations are often used for predicting gene expression and for examining different behaviors in genetic circuits[62]. Using the model shown in **Figure 1.4**, the interactions between the different components in the system is written as a system of differential equations (**Table 1.1**).

Table 1.1 – Set of equations for modeling gene expression

Abstraction	Equation
Equation (1.2) Constitutive RNA transcription	$\frac{d}{dt}x_{RNA}(t) = \alpha_0 + \alpha - \gamma \cdot x_{RNA}(t)$
Equation (1.3) Inducer regulated RNA transcription	$\frac{d}{dt}x_{RNA}(t) = \alpha_0 + \alpha \cdot \frac{u_g(t)^n}{K_{hill}^n + u_g(t)^n} - \gamma \cdot x_{RNA}(t)$
Equation (1.4) Repressor regulated RNA transcription	$\frac{d}{dt}x_{RNA}(t) = \alpha_0 + \alpha \cdot \frac{K_{hill}^n}{K_{hill}^n + u_r(t)^n} - \gamma \cdot x_{RNA}(t)$
Equation (1.5) Protein translation	$\frac{d}{dt}x_P(t) = \sigma \cdot x_{RNA}(t) - \beta \cdot x_P(t)$
Equation (1.6) Hill-function	$hill(x, K_{hill}, n) = \frac{x^n}{K_{hill}^n + x^n}$
Equation (1.7) Cell growth	$\frac{d}{dt}cells(t) = \alpha_{cells} \cdot cells(t) \cdot (1 - cells(t)) - \beta_{cells} \cdot cells(t)$

Where:

$x_{RNA}(t)$ - Amount of RNA in the system
$x_P(t)$ - Amount of protein in the system
α_0 - Leaky transcription rate
α - RNA transcription rate
β_{cells} - Cells decay rate
γ - RNA decay rate
σ - Protein translation rate
β - Protein decay rate
α_{cells} – Cells growth rate
u_g and u_r – Concentration of inducer (green light) and repressor (red light)
cells(t) – normalized cell density [0,1]; '1' represents maximum cell density

Note: All amounts are in arbitrary units (a.u); translation, decay and death rates are in s^{-1}; transcription rate is in $a.u \cdot s^{-1}$; cell growth/decay rates are in $s^{-1} \cdot a.u^{-1}$; the amount of mRNA, proteins, and cells were chosen to be arbitrary (a.u) since the real amounts were not predicted.

To simplify the explanation, it is assumed that the output and state of the system are identical. The state of the system is a continuous time function that represents the amount (or concentration) of every element in the system at any given point in time. The model consists of multiple first order differential equations such that each one explicitly defines the rate of change of a specific element in the system. The practicality of mathematically solving the system of equations (numerically or analytically) is to find the amount of each element in the system at any point in time. For example, gene expression is described by a set of equations (**Table 1.1)**. Equations 1.2 – 1.4 describes the transcription process and equation 1.5 describes the translation process. Equation 1.5 depends on the state (or solution) of equation 1.2 and 1.3 (or 1.4) -i.e. $x_{RNA}(t)$ - the amount of mRNA and $x_P(t)$ - the amount of protein. Degradation of the gene plays a critical role in the regulatory network and is usually modeled as a first-order decay process. The decay rate coefficients are a statistical value that represent the probability of an element to break down and to be excreted from the system. At the same time, the protein translation process is modeled under the assumption that there is an unlimited amount of ribosomes in the system which means higher amount of mRNA will directly result in a higher rate of proteins synthesis. The transcription equation can also be expanded to model an inducible promoter that is affected by the presence of an inducer or repressor element in the system (equation 1.3 or 1.4). The dependency on the inducer or repressor is based on the Hill function (equation 1.6). The output of the Hill-function is a value between 0 to 1 that represents the probability of the inducer/repressor to bind with the promoter. The multiplication of the Hill-function by the maximum transcription rate describes the actual transcription rate as a function of inducer/repressor amount (or concentration). The Hill-function has two key parameters (1) the dissociation constant (K_{hill}) that defines the threshold-amount of the inducer/repressor in which the output of the function equals to 0.5 (50 % activated/repressed) and (2) is the cooperativity coefficient (n) that defines the sensitivity of the binding process, to the inducer/repressor changes around the threshold amount. **Figure 1.5** graphically demonstrates the effect of those parameters. Finally, we also describe cell growth by the difference between the growth and death rates , (equation 1.7), similar to other studies[58], [63]. In this equation, α_{cells} and β_{cells} represents the specific growth and death rates of the cells respectively. The parameter cells(t) is the normalized cell density such that a value of '1' represents the maximum level of cell growth.

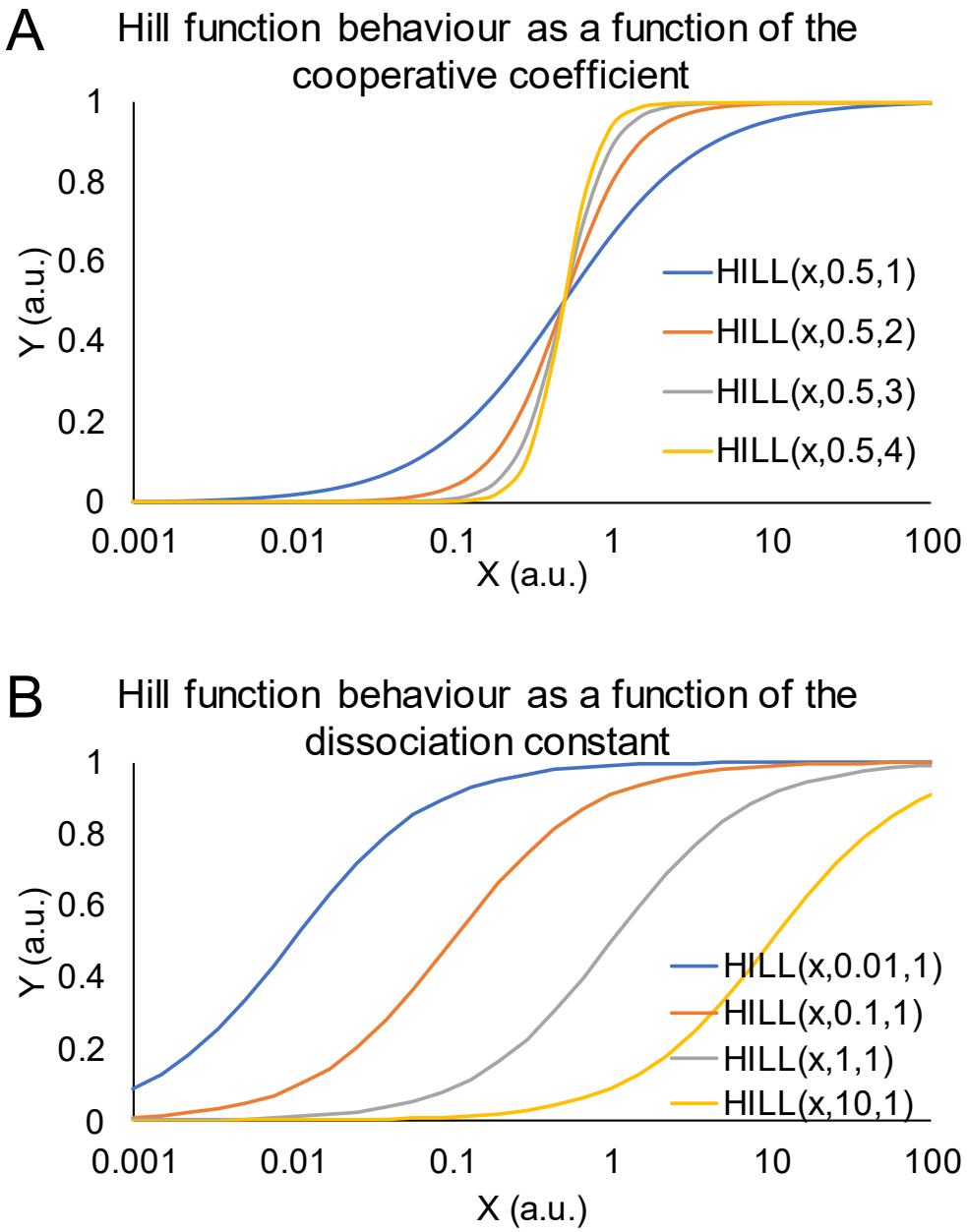

Figure 1.5: The effects of K_{hill} and n over the Hill function transfer function.

(A) Shows the Hill function curves with dissociation constant of (K_{hill} = 0.5) and different values of cooperative coefficient (n). Increasing the cooperative coefficients increases the sensitivity of the function to around the dissociation value. **(B)** Shows the Hill function curves with cooperative coefficient of (n = 1) and different values of dissociative constant (K_{hill}). The dissociative constant defines the threshold amount of inducer/repressor to cause the function output to be 0.5.

1.7 Objectives

The main goal of this study is to develop a closed-loop optogenetics system that is capable of controlling gene expression in bacterial cells. To achieve this goal, this work is divided into two specific aims shown below.

1. Design an optogenetic setup for well-plates and microfluidic devices. The setup will be capable of automating light delivery to the cells, and capable of obtaining analytical measurements (e.g., absorbance and fluorescence measurement) for open- and closed-loop control.
2. As a proof-of-principle, show closed-loop control of GFP expression in bacterial cells through modeling and experimentation.

2 Materials and Methods

In this chapter, I describe the materials and methodologies used in the optogenetics experiments. I will also describe methodologies (and software) used in my simulation models.

2.1 Reagents and plasmids

Unless specified otherwise, general-use chemicals were purchased from Sigma Aldrich (St. Louis, MO). Competent *E. coli* DH5α strains were stored at -80 °C before transformation procedures. Three plasmids were used for this study: pSR43.6r and pSR58.6 (Addgene plasmid # 63197 and #63176) and pFAB4876 was donated from the Joint BioEnergy Institute. LB (lysogeny broth) bacterial growth media was prepared with 1% w/v tryptone, 0.5% w/v yeast extract and 0.5% w/v NaCl and sterilized by autoclave. Kanamycin (50 mg/mL ddH_2O), spectinomycin (25 mg/mL ddH_2O) and chloramphenicol (50 mg/mL 95% EtOH) antibiotic stock solutions were prepared and stored at -20 °C. For experiments antibiotics were added into the bacterial culture at indicated dilutions.

2.2 Plasmid construction, purification, and extraction

Plasmid pSR58.6j was generated by modifying pSR58.6 to include a kanamycin resistance gene amplified from pFAB4876. Construction of the plasmid was achieved through PCR amplification and Golden Gate assembly using protocols described elsewhere[64], [65]. All primers and DNA fragments for PCR amplification were designed using Benchling software with T_m values of approximately 55 °C. Primers contained a 5' BsaI cut site followed by a 4-base region forming compatible sticky ends upon BsaI digestion. For each DNA part, four 50 μL PCR reactions were setup using Phusion DNA polymerase (Thermofisher cat no. F530S) and executed using the recommended thermal-cycling schedules provided by the manufacturer. The fragment of the original pSR58.6 that contains the *sfGFP*, promoter and ribosomal binding site (RBS) was amplified using primers OG_1 and OG_4. A kanamycin resistance gene was amplified from pFAB4876 using primers OG_2 and OG_3. After PCR, 1 μL of DpnI (Invitrogen cat no. IVGN0106) was added to each 50 μL reaction. After a 2-hour incubation at 37 °C, the DNA fragments were purified using 0.8% agarose gel DNA electrophoresis, excised and extracted using Biobasic prep kits (Biobasic cat no. BS353).

Table 2.1: Primer List

No.	Name	Full 5'
OG_1	OG_pSR58.6_R	cacaccaGGTCTCC**catc**tagtatttctcctcttttaaaaatgcgatc
OG_2	OG_Kan_F	cacaccaGGTCTCC**gatg**attgaacaagatggattgcacg
OG_3	OG_Kan_R	cacaccaGGTCTCC**tcag**aagaactcgtcaagaaggc
OG_4	OG_sfGFP_F	cacaccaGGTCTCC**atgc**gtaaaggcgaagagc
OG_5	OG_pSR58.6_F	cacaccaGGTCTCC**ctga**tgatgataataatctagaccaggcatca
OG_6	OG_pSR58.6_colony_F	GCGTTCACCGACAAACAACAGATA
OG_7	OG_pSR58.6_colony_R	CAGGTCTGGGTTAATTCATGTTGC

Note: The 5' lower case bases are a cap to improve BsaI enzyme anchoring. Capital letters represent the BsaI recognition site. The **bolded** bases are the four-base overhangs. The 3' lower case bases are the annealing region which primes the template DNA during PCR.

For Golden Gate assembly, 40 fmol of both fragments were combined in a 10 μL Golden Gate reaction containing 0.5 μL T4 DNA ligase and 1 μL T4 ligase buffer (NEB cat. no. M0202S), 1 μL BsaI (NEB cat no. R0535S) and 7.5 μL of water. Golden Gate thermal-cycler temperature were scheduled as follows: 25 cycles (37 °C 3 min, 16 °C 3 min) then kept at 4 °C or −20 °C until transformation. The 10 μL reaction was then transformed into 50 μL of competent cells by heat shock. Following transformation, colonies were picked and resuspended in 20 μL for colony PCR to verify the presence of the KanR gene insert. For colony PCR, the Phusion PCR reaction mentioned above was scaled down to 20 μL with 1 μL of resuspended colony added as template and using primers OG_6 and OG_7.

Transformed bacteria were grown overnight in 5 mL of LB with corresponding antibiotics (diluted to 1:1000), shaken at 200 rpm and 37 °C. Extraction of DNA was conducted using Biobasic prep kits (Biobasic cat no. BS353), according to manufacturer's instructions. DNA concentration and quality were assessed by running a 0.8% agarose gel and measuring concentration on a plate reader (TECAN Infinite M200).

2.3 Bacterial cryopreservation, preparation, culture, and transformation

Single bacterial colonies grown on petri dishes with appropriate antibiotics for selection were picked into 30 mL of LB supplemented with 30 µL of spectinomycin or 30 µL chloramphenicol and incubated overnight at 37 °C with 200 rpm shaking. The bacteria culture was mixed at a 1:1 ratio with sterile 50 % glycerol, aliquoted, and stored at -80 °C.

DH5α stored at -80 °C was streaked-out on a plate containing 8 % agar LB and grown overnight. After 24 h, one colony was picked from the plate and inoculated in 5 mL of LB and incubated overnight at 37 °C, 200 rpm. Then, 5 mL of the culture was added into 500 mL of fresh LB contained in a 1 L baffled flask and incubated at 37 °C with 200 rpm shaking until the OD_{600} measured 0.45. The bacteria were transferred to 50 mL centrifuge tubes, and recovered by centrifugation at 2000 g for 10 min at 4 °C. After centrifugation, the supernatant was discarded, and the pellet was re-suspended in 50 mL of ice-cold solution containing 75 mM of $CaCl_2$ and 15 % glycerol and incubated on ice for 5 min. The centrifugation and resuspension steps were then repeated. Lastly, the cell mixture was transferred to 50 µl aliquots in 1.5 mL centrifuge tubes, snap frozen in liquid nitrogen, and stored at -80 °C.

Chemically competent DH5α were co-transformed with pSR43.6r and either pSR58.6 or pSR58j and were used in the optogenetic experiments. Prior to each experiment, DH5α cells were incubated overnight at 200 rpm and 37 °C, and then diluted at 1:10 with LB supplemented with 1:1000 spectinomycin and 1:1000 chloramphenicol antibiotics at a final volume of 1 mL. The typical range of OD_{600} of the diluted bacteria was made to 0.3 to 0.4. Higher or lower initial OD_{600} were made by different dilutions.

Before transformation experiments, 50 µL aliquots of competent cells were thawed on ice for 10 min. 50-100 ng of each plasmid (pSR43.6, pSR58.6 or p58.6j) were added to the competent cells and mixed gently by flicking the tubes and incubated on ice for 30 min. After incubation, the cell mixture was heat shocked at 42 °C for 30 s and then immediately placed back in the ice-bath for 3 min. After incubation, 950 µL of super optimal broth with catabolite repression (SOC) was added to the cell mixture and incubated at 37 °C for one hour, allowing cells to recover. Following recovery, cells were plated on LB-agar selection plates containing both spectinomycin and chloramphenicol. Plates were stored overnight at 37 °C. After 24 h, single colonies were picked for experiments described below.

2.4 Microfluidic device and fabrication

A 3D-printed mold containing a U-channel (height = 1.5 mm; width 1.0 mm) was printed on PLA material using an Ultimaker 2+ Extended (Ultimaker, Netherlands). After printing, the mold was filed and coated with acetone to achieve a smooth surface. We used soft-lithography procedures to create the microfluidic device, similar to published procedures[66]. Briefly, PDMS was prepared by mixing 10:1 with curing agent and was casted onto the mold. The mold with the PDMS was placed in a vacuum chamber for an hour to eliminate air bubbles in the PDMS. Final curing was done at 45 °C overnight (~12 h). The PDMS was gently peeled off the mold and it was plasma treated (Corona-SB, Elveflow, France) for 20 s. The PDMS was then bonded to a plasma-treated glass slide and baked at 160 °C for 20 min to improve adhesion between the two pieces. The device was coated with Aquapel^{TM} (Aquapel) for 1 min and rinsed with water by pipetting to prepare for optogenetic experiments.

2.5 Light delivery system, operation, and calibration

The light delivery system consisted of 2 x 2 matrix of RGB LEDs (APTF3216QBDZGSURKC, Kingbright) soldered onto the LED board. The four red (650 nm) and green (520 nm) LEDs are independently controlled, while the four blue (480 nm) LEDs have one common activation control. The illumination of red and green LED was controlled by an Arduino Uno that is serially connected to a control board containing solid-state switches (AQW216, Panasonic), a 9 kΩ resistor, and the LED. The four blue LEDs were driven by an H-bridge (MC33886, NXP) with a 1.5 Ω series resistor and by using a PWM signal that was generated by the Arduino. A 3D-printed optical-filter selector containing an excitation filter (470 nm, 67027, Edmund Optics) with multiple sized apertures (D = 1 mm) was placed 1 mm above the LED board. A well-plate or a microfluidic device was placed above the optical-filter selector (with the LEDs aligned below to the respective wells) and bacteria cultures were aliquoted into the individual wells. To obtain fluorescence and absorbance measurements, a long-pass filter (515nm, 66085, Edmund Optics) and μEye camera (USB2.0 DCC1645C, Thorlabs) were placed 40 cm above the well-plate or microfluidic device. The light-delivery system was also integrated with a heating element to control temperature (TE technology INC, Texas, USA) and a servo (MG90S, Power Pro) to select the optical filters and apertures. Both components were connected to a control board supplying 5 V and 12 V. To control the CcaS-CcaR system, the green-to-red LEDs were activated/deactivated under continuous illumination conditions (with different ratio of green to red activation time over a period of 10 min). To control the blue LEDs, a pulse-

width modulation signal (PWM) at a frequency of 1 kHz with a duty cycle of 16 % to 35 % was applied to the samples before fluorescence measurement.

All 3D printed parts: the plate holder, filter-selector, heat-pump inlet and outlet funnels, were designed using Creo Elements/Direct (PTC, Boston, Massachusetts, USA). All 3D printing .stl files were converted to GCODE using Cura 3D conversion (nozzle - 0.4 mm, material – PLA, and profile - high quality). All 3D parts were printed using the Ultimaker 2+ Extended machine and PLA material (Shop3D, Mississauga, ON, Canada). A detailed schematic of the optogenetics setup was shown in **Figure 3.1-3.3**.

To calibrate our system for OD_{600} measurements, four different OD dilutions (0.35, 0.4, 0.6, and 0.8) were prepared and 100 µL of each dilution was added to each well (that are directly in-line with the LED). To measure the OD_{600} using our optogenetics system, the red LED was driven at low intensity power of 25 µW for 650 ms. An image was captured showing the four detected samples and was used to calculate their respective pixel intensities. Using Open-CV2 (open source computer vision program, The OpenCV Library, Bradski, G., https://pypi.org/project/opencv-python/), the image was converted into an HSV (hue, saturation, value) color-space. The V channel (light intensity of the color value) of all pixels that reside inside the region of interest (20 x 20 pixels) were averaged. The H and S were set to H = 0 ± 20, S = 250 ± 60. Each averaged value was reference corrected such that the calculated value represents the amount of light that came out of the well and captured by the camera. The data from our system was correlated to absorbance measurements measured using a well-plate reader (TECAN Infinite M200). The reader settings used for absorbance measurements was shown in **Table 2.2.**

Table 2.2: TECAN configuration for OD measurements

Parameter	Value	Unit
Multiple Reads per Well (Circle (filled))	3 x 3	
Multiple Reads per Well (Border)	500	µm
Wavelength	600	nm
Bandwidth	9	nm
Number of Flashes	25	
Settle Time	0	ms

Similarly, for fluorescence calibration, a set of six fluorescein dilutions (1:12, 1:16, 1:20, 1:24, 1:28 and 1:32) were prepared by diluting 1 g of fluorescein (46955-100G-F, Sigma Aldrich) in 100 mL in 1M NaOH. This stock solution was further diluted in distilled water to the final target dilution. 100 µL of each

dilution was placed in each of the four wells and their fluorescence intensities were taken in a well-plate reader (TECAN Infinite M200) and compared to the values obtained from the optogenetics system at 20 % excitation power. All wells were calibrated to a single reference well (arbitrarily chosen) by setting a gain correction value for each of the three wells.

To increase the dynamic range of our fluorescence measurement, we derived an excitation power compensation function (equation 2.3). The amplification function, $\mathrm{Amp}(\mathrm{Ex_{pwr}})$, was modeled using equation 2.1. Experimental fluorescence measurements ($\mathrm{F_{raw}}$) from 1:16 fluorescein dilutions were measured using an excitation power ranging from 4 - 50 % in 2 % increments (n = 4). Next, the normalized amplification function (equation 2.2) was derived by dividing each measurement by the reference measurement taken at 20 % excitation power ($\mathrm{F_{raw}}(20\%)$). The excitation compensation function was fit to the inverse of the normalized amplification function as described in equation 2.3. During the experiment, the compensated fluorescence value ($\mathrm{F_{comp}}$) was calculated for each measurement in real-time. Each uncompensated fluorescence measurement was multiplied by the calculated compensation factor by using the respective excitation power as described in equation 2.4. This equation enabled an approximate calculation of the total amount of protein for a given compensated fluorescence measurement (equation 2.5).

Table 2.3 – Equations for Power Compensation

Process	Equation
Equation (2.1) Fluorescence measurement model	$\mathrm{F}_{raw}(Ex_{\mathrm{pwr}}) = \mathrm{G_{state}} \cdot \mathrm{Amp}(Ex_{\mathrm{pwr}}) = \mathrm{G_{state}} \cdot \mathrm{K} \cdot \mathrm{Amp}_{norm}(Ex_{\mathrm{pwr}})$
Equation (2.2) Normalized amplification	$\mathrm{Amp}_{norm}(Ex_{\mathrm{pwr}}) = \frac{\mathrm{F}_{raw}(Ex_{\mathrm{pwr}})}{\mathrm{F}_{raw}(20\%)}$
Equation (2.3) Excitation compensation	$\mathrm{ExComp}(\mathrm{Ex_{pwr}}) = [\mathrm{Amp_{norm}}(\mathrm{Ex_{pwr}})]^{-1} \cong 470.023 \cdot \mathrm{Ex_{pwr}}^{-2.04}$
Equation (2.4) Compensated fluorescence measurement	$\mathrm{F}_{comp} = \mathrm{F}_{raw}(Ex_{\mathrm{pwr}}) \cdot \mathrm{ExComp}(\mathrm{Ex_{pwr}})$
Equation (2.5) Estimated amount of GFP	$\mathrm{G}_{state} = \mathrm{K}^{-1} \cdot \mathrm{F_{raw}}(Ex_{\mathrm{pwr}}) \cdot \mathrm{Amp_{norm}}(\mathrm{Ex_{pwr}})^{-1}$

Where:
$\mathrm{Ex_{pwr}}$ is the excitation power used during the fluorescence measurement (in %)
$\mathrm{F_{raw}}$ is the fluorescence intensity value as captured by the light detector (camera)
$\mathrm{F_{comp}}$ is the compensated fluorescence measurement value
$\mathrm{G_{state}}$ is the real amount of GFP in the system – not directly measured

K is the factor that links between the compensated fluorescence value to amount of GFP (K = 1 in our experiments).
Amp is the function that relates between G_{state} and the excitation power to fluorescence intensity
Amp_{norm} is the normalized Amp function.
ExComp is the excitation power compensation function

For all fluorescence measurements, an image was captured showing the four detected samples and was used to calculate their respective pixel intensities. Quantifying the fluorescence levels was via a captured image (similar to absorbance measurements) except the H and S parameters were set to: H = 60 ± 20, S = 254 ± 60 for the green channel. The averaged value represented the amount of green light that came out of the well and captured by the camera. The reader settings used for fluorescence measurements is shown in **Table 2.4.**

Table 2.4: TECAN configuration for fluorescence measurements

Parameter	Value	Unit
Multiple Reads per Well (Circle (filled))	2 x 2	
Multiple Reads per Well (Border)	500	µm
Excitation Wavelength	488	nm
Emission Wavelength	530	nm
Excitation Bandwidth	9	nm
Emission Bandwidth	20	nm
Gain	80	Manual
Number of Flashes	25	
Integration Time	20	µs
Lag Time	0	µs
Settle Time	0	ms

2.6 Description of the CcaS-CcaR-V2 model

The overall system model was developed following the Model Based Design methodology[67], [68]. First, the mathematical model represented the biological system (i.e. the plant) was developed (see below) which was approximated to be a linear-time-invariant (LTI). Following the plant design, the entire system (PID controller, feedback and plant) was modeled together in the continuous s-domain. The model was transformed to the digital z-domain using two methods: the CcaSR model was transformed using the Zero-Order-Hold (ZOH) method and the controller, differentiator and the low-pass-filter were transformed using the Euler's backward method. A root-locus method was used to tune the controller for stability (**Figure A.1 and A.2**). MATLAB Simulink Control System toolbox was used to convert the continuous time controller to discrete time controller and to perform the root-locus analysis.

The general model of the CcaSR system followed the methodology for genetic regulation modeling using the equations 1.2 - 1.7. It consisted of six first order, non-linear and time varying differential equations. Inputs of the model were the red and green lights intensity – designated as $u_r(t)$ and $u_g(t)$. Outputs of the model were the total amount of GFP (GFP(t)) and living cells in the system (cells(t)). The general model simulated only the phosphorylation and dephosphorylation of the CcaR protein by light[37], [69] since it was assumed that these processes occurred simultaneously and were governed by the same input (i.e. the induced light). The dependency on cell count was approximated by setting the rate of the GFP RNA transcription to be proportional to the number of cells in the system – i.e. the assumption that RNA transcription was done in parallel by all cells in the system. All modeling was done by an in-house Python script and the solver used a fixed-step backward Euler method with a time step of 5 min.

2.6.1 The non-linear, time-variant, single-in model

We adapted the general model and modified it to create a non-linear, time-variant, single-in model (see **Table 2.5** for equations). This model was used as an intermediate step toward the conversion of LTI model. The input (or inducer) was defined to be the ratio between the green and red irradiation time which was normalized to be between [0,1] (where 0 is maximum red and 1 maximum green) and the outputs of the model were represented by (1) the total amount of GFP and (2) number of living cells in the system. The inducer was directly controlling the Hill function which was directly affecting the GFP RNA expression rate. The equations for CcaR RNA, CcaR and phosporylated CcaR were removed because of redundancy.

Table 2.5 – Equations for the non-linear, time-variant, single-in model

Process	Equation
Cell growth	$\frac{\mathrm{d}}{\mathrm{dt}}\mathrm{cells(t)} = \alpha_{\mathrm{cells}} \cdot \mathrm{cells(t)} \cdot (1 - \mathrm{cells(t)}) - \beta_{\mathrm{cells}} \cdot \mathrm{cells(t)}$
Regulated transcription of GPF mRNA	$\frac{\mathrm{d}}{\mathrm{dt}}\mathrm{RNA_{GFP}(t)} = \mathrm{cells(t)} \cdot [\alpha_{0\mathrm{GFP}} + \alpha_{\mathrm{GFP}} \cdot \frac{u_{\mathrm{i}}(\mathrm{t})^{\mathrm{n}}}{k_{\mathrm{hill}}{}^{\mathrm{n}} + u_{\mathrm{i}}(\mathrm{t})^{\mathrm{n}}} - \gamma_{\mathrm{GFP}} \cdot \mathrm{RNA_{GFP}(t)}]$
Translation of GFP	$\frac{\mathrm{d}}{\mathrm{dt}}\mathrm{GFP(t)} = \sigma_{\mathrm{GFP}} \cdot \mathrm{RNA_{GFP}(t)} - \beta_{\mathrm{GFP}} \cdot \mathrm{GFP(t)}$

Where:
α_{0GFP} = 0.002 **-** GFP RNA leaky transcription rate
α_{GFP} = 0.08 **-** GFP RNA transcription rate
Υ_{GFP} = 0.00032 **-** GFP RNA decay rate
σ_{GFP} = 0.000036 **-** GFP Translation rate
β_{GFP} = 0.0000002 **-** GFP decay rate
K_{hill} = 0.38 **–** Hill dissociation factor
n = 2 – Hill cooperative factor
σ_{Cells} = 0.0009 **-** Cells growth rate
β_{Cells} = 0.000015 **-** Cells death rate
$u_i(t)$ – Inducer (input)
$RNA_{GFP}(t)$ - Amount of RNA GFP
GFP(t) - Amount of GFP (output 1)
cells(t) – Normalized cell density between [0,1] (output 2)

Note: All amounts are in arbitrary units (a.u); translation, decay and death rates are in s^{-1}; transcription rate is in $a.u \cdot s^{-1}$; cell growth/decay rates are in $s^{-1} \cdot a.u^{-1}$; the amount of mRNA, proteins, and cells were chosen to be arbitrary (a.u) since the real amounts were not predicted.

2.6.2 The linear and time-invariant, single in single out model

In order to use the root-locus method, the model was further approximated to a linear and time-invariant form. The Hill function was linearized for three inducer equilibrium points: 0.1, 0.35 and 0.6. The slope of the Hill function was calculated at each point. The slope multiplied by the input signal replaced the Hill function. To simplify the model, the leaky factor of the GFP RNA was set to zero. The transformation of the continuous time model to discrete time was done using MATLAB Simulink Control System toolbox (Mathworks). The state-space representation of the model was (such that $dx/dt = Ax(t) + Bu_i(t)$ and $y(t) = GFP(t)$):

Equation (2.4) Systems state matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{RNA_{GFP}(t)} \\ \mathbf{GFP(t)} \end{bmatrix}$$

Equation (2.5) State transition matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{-\gamma\_GFP} & \mathbf{0} \\ \mathbf{\sigma\_GFP} & \mathbf{-\beta\_GFP} \end{bmatrix}$$

Equation (2.6) Input matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{Hill_{slop} \cdot \alpha\_GFP} \\ \mathbf{0} \end{bmatrix}$$

Equation (2.7) Output matrix

$$\mathbf{C} = [\mathbf{0} \quad \mathbf{1}]$$

Equation (2.8) Feedforward matrix

$$\mathbf{D} = \mathbf{0}$$

Where:

$\boldsymbol{\alpha_{GFP}}$ = 0.08 - GFP RNA transcription rate
$\boldsymbol{\Upsilon_{GFP}}$ = 0.00032 - GFP RNA decay rate
$\boldsymbol{\sigma_{GFP}}$ = 0.000036 - GFP Translation rate
$\boldsymbol{\beta_{GFP}}$ = 0.0000002 - GFP decay rate
$\mathbf{Hill_{slope}}$ = 1.211, 1.419, 0.681 – Linearized values around inducer values of 0.1, 0.36 0.6 respectively

$\mathbf{u_i(t)}$ – Induced (input)
$\mathbf{RNA_{GFP}(t)}$ - Amount of GFP RNA
GFP(t) - Amount of GFP (output)

Note: All amounts are in arbitrary units (a.u); translation, decay and death rates are in s^{-1}; transcription rate is in $a.u \cdot s^{-1}$; cell growth/decay rates are in $s^{-1} \cdot a.u^{-1}$; the amount of mRNA, proteins, and cells were chosen to be arbitrary (a.u) since the real amounts were not predicted.

2.7 Open-loop experiments

Prior to the experiment, the pSR58.6 and pSR43.6r that were co-transformed in DH5α were diluted 1:10 (to OD_{600} = ~0.3) using LB supplemented with 1:1000 (1 μg/mL) spectinomycin and 1:1000 (1 μg/mL)

chloramphenicol. The optogenetic setup was connected to a PC by a USB connection and to a 12 V and 5 V power supply to drive the LEDs and heating element. 100 μL of the cells were aliquoted into the well-plate and incubated at 37 °C. 20 μL of silicon oil was loaded into each well to reduce evaporation of the samples. During the experiment, each well was exposed to different ratios of green-to-red light illumination conditions. Three open-loop experiments were performed using three different illumination schemes: (1) constant, (2) pulse, and (3) random illumination. For the constant illumination scheme, the wells were exposed different ratios (0, 30, 45 %) of green-to-red light for 6 h. For the pulse illumination scheme, cultures were exposed to a 40 % green-to-red light ratio for 5 h. For the random illumination scheme, cultures were exposed to a range of ratios (0 - 90 %) of green-to-red light intensity for 16 h. For all illumination schemes, cultures were initially exposed to red light for 1 h. Fluorescence and OD_{600} values were measured periodically every 30 min and obtained using image-based techniques (as described above). All fluorescence measurements were initiated by excited the samples using the blue LED set at 20 % excitation power. The fluorescence and OD_{600} values were periodically obtained by image-based techniques (see calibration section) by the main python program and recorded into a log file.

For open-loop microfluidic experiments, the media used in the device were identical to the those used for the well-plate (see above). The device was mounted on the plate holder inside the optogenetic setup and incubated at 37 °C for 21 h. At 0 h, the device was filled with co-transformed CcaSR bacterial cells (at an OD_{600} of 0.35) using a handheld pipet. Each microfluidic well was exposed to one of the ratios of green-to-red light – 10 %, 20 %, 40 % and 60 %. Fluorescence was detected by exciting the samples using the blue LED set at 85 % excitation power in each microfluidic well. The fluorescence and OD_{600} values were recorded every 30 min and logged similar to the well-plate experiments.

2.8 Closed-loop experiments

Protocols to prepare samples, to load samples into the well-plate or microfluidic device, and to setup the optogenetics system followed the procedures of the open-loop experiment. An in-house Python script was written to implement the PID control algorithm similar to our previous work [70]. The PID controller contains three gain coefficients: K_P, K_I, and K_D for the proportional, integral, and derivative terms in the PID controller equation (see **Equation 1.1**). For our closed-loop experiments, the measurements and controller's output were calculated every 30 min. The gains for each experiment was set to the following: K_P = 6 or 7 and K_I = 0.5 or 0.2 which were chosen by manual tuning. K_D was set to 0 since the derivative term represents anticipatory control which was not required for our experiments. **Table 2.6** describes all the parameters used for each closed-loop experiment. Mathematically, the closed-loop control followed a five-step algorithm:

Step 1: Compensation for excitation power
$$F_{comp}[n] = F_{raw}[n] \cdot ExComp(EX[n])$$

Step 2: Low pass filtering the measurement
$$F_{filtered}[n] = \frac{F_{comp}[n] + F_{comp}[n-1]}{2}$$

Step 3: Discrete differentiation
$$FB[n] = \frac{F_{filtered}[n] - F_{filtered}[n-1]}{T_s}$$

Step 4: The error signal
$$e[n] = SP[n] - FB[n]$$

Step 5: The PID controller
$$Y[n] = K_p \cdot e[n] + \sum_{0}^{n} K_i \cdot e[n] \cdot T_s + K_d \cdot \frac{e[n] - e[n-1]}{T_s}$$

Where:

F_{raw}[n] is the fluorescence intensity as was taken by the camera.
F_{comp}[n] is the compensated fluorescence measurement (for excitation power).
$F_{filtered}$[n] is the filtered fluorescence measurement (averaged of the last two samples).
ExComp(x) is the compensation function (see **Equation 2.1**)
EX[n] is the excitation power in which the fluorescence[n] measurement was taken at (in %).
FB[n] is the fluorescence rate of change (a.u./hour).
e[n] is the error between the setpoint and the feedback signal.
T_s is the control loop sampling period time (0.5 hour).
SP[n] is the expression rate target value (a.u./hour).
K_p, K_i, K_d are the proportional, integrative and differential gain coefficients.
Y[n] is the controller's output.

Table 2.6– Parameters and conditions used for closed-loop experiments

Experiment	Gain coefficients	Target condition	Experimental conditions	Figure
Low expression rate	$K_P = 6.0$, $K_I = 0.5$, $K_D = 0$	8 a.u/h	Initial $OD_{600} = 0.35$ Excitation power = 20 %	3.14A
High expression rate	$K_P = 6.0$, $K_I = 0.5$, $K_D = 0$	22 a.u/h	Initial $OD_{600} = 0.35$ Excitation power = 14 %	3.14D
Reducing overshoot by PID retune	$K_P = 10.0$, $K_I = 0.04$, $K_D = 0$	12 a.u/h	Initial $OD_{600} = 0.35$ Excitation power = 18 %	3.15A
Reducing overshoot by gradual transition	$K_P = 7.0$, $K_I = 0.2$, $K_D = 0$	20 a.u/h	Initial $OD_{600} = 0.35$ Excitation power = 18 %	3.15B
Multiple setpoints	$K_P = 6.0$, $K_I = 0.2$, $K_D = 0$	14,28 a.u/h	Initial $OD_{600} = 0.3$ Excitation power = 14 %	3.16A, B
Different starting OD values	$K_P = 7.0$, $K_I = 0.2$, $K_D = 0$	7 a.u/h	Initial $OD_{600} = 0.35, 0.55$ Excitation power = 20 %	3.17
Microfluidic	$K_P = 7.0$, $K_I = 0.2$, $K_D = 0$	6 a.u/h	Initial $OD_{600} = 0.35$ Excitation power = 85 %	3.18

3 Results

This section describes novel methods developed for the project as well as the results from conducted experiments.

The setup of the optogenetics system consists of the following components: (1) design of optical setup - holder, filters, light sources, LED driving circuit, (2) characterizing the setup - excitation power compensation, crosstalk between wells, crosstalk with excitation light, calibration of OD, calibration of fluorescence, (3) software design, (4) ppen-loop - model and experiment, (5) controller design, (6) controller implementation and simulation, and (7) closed-loop - model and experiment.

3.1 Optogenetics System: Design and Characterization

The optogenetics system is shown in **Figure 3.1A**. It consists of four main modules: (1) a camera with a long-pass filter to measure the pixel intensities of the imaged bacterial cells, (2) a well-plate (or microfluidic) platform to culture and to incubate the bacterial cells, (3) an optical filter module placed along the illumination path that is used for fluorescence and absorbance measurements, and (4) an automated light delivery system to control the illumination profiles. As illustrated in **Figure 3.1B**, the filter module contains a band pass filter (λ_{ex} = 470 nm ±17 nm), and through-holes of different sizes to direct the light to the cell samples and to prevent light crosstalk between the different control-lights that are targeted to the different incubation areas. The OD diffuser is used to spread the light on the samples to obtain an accurate reading of the OD. In line with the light path, an adaptor is 3D printed to hold the LED board and the well-plate (or device) (**Figure 3.1C**).

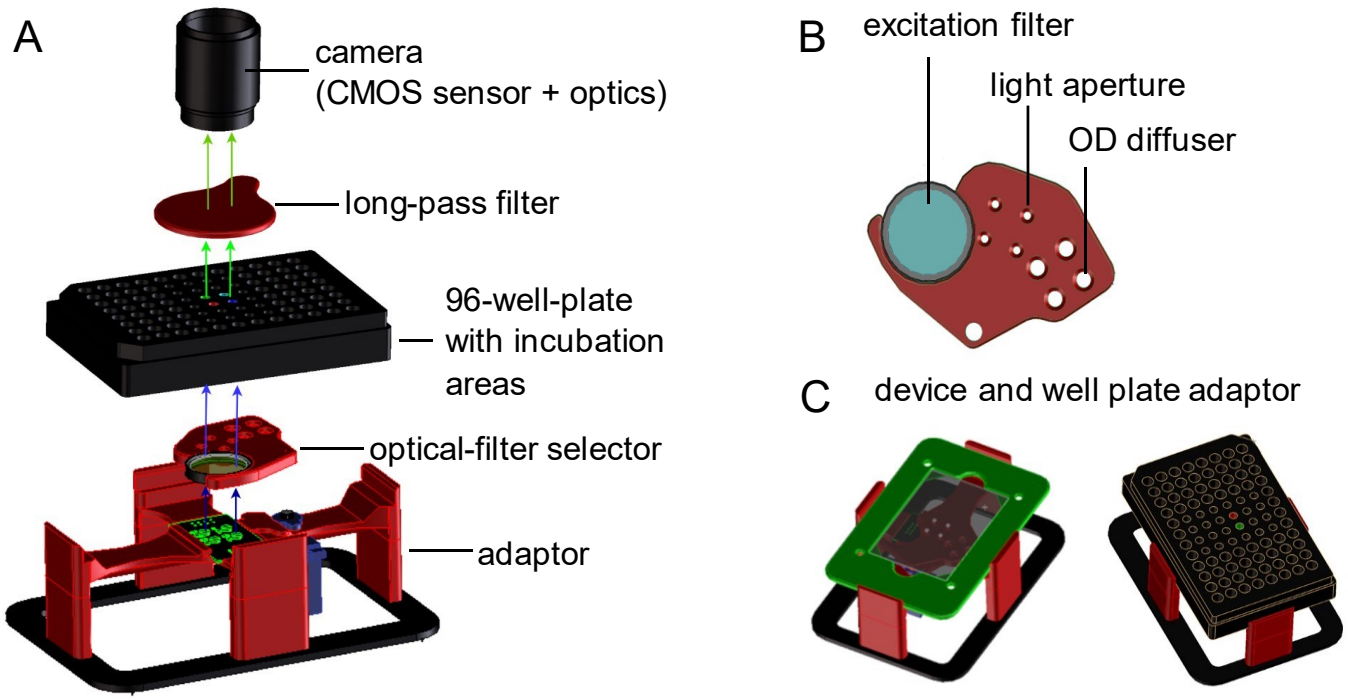

Figure 3.1: Optogenetic optical design

(A) Real-time Optogenetic (RT-OGENE) setup. The setup consisted of a camera, long-pass filter, a well-plate or microfluidic device, optical filter selector, and an adaptor for the platform. Blue arrows indicated the path of a blue (470 nm) light beam originating from the LED. Green arrows indicated the fluorescent (520 nm) signal emitted from the cells and detected by the camera. For OD_{600} measurements, the filter selector was rotated to the OD diffuser position. Blue arrows indicated the path of a red (650 nm) light beam originating from the LED. Green arrows indicated the scattered red light signal from the cells and detected by the camera. For CcaSR control, the filter selector was rotated to the light apertures. Blue arrows indicated the path of a red or green (650 nm/520 nm) light originating from the LED. No images were captured in this mode. **(B)** The 3D printed filter selector contained a filter slot for the 480 nm excitation filter, four control light apertures, and four OD diffusers. **(C)** A microfluidic device (left) or well plate (right) was mounted on the plate-holder. The plate holder contained the filter selector mechanism and the LED board (not shown).

The four modules are enclosed inside an incubation chamber that includes a heating element (used to maintain a constant temperature of 37 °C) and a temperature sensor to measure the temperature in real-time (**Figure 3.2**). The automated light delivery system consists of a LED driver and control board (**Figure 3.3**) which is powered by a 12 V power supply and controlled by a Python based application running on a host computer.

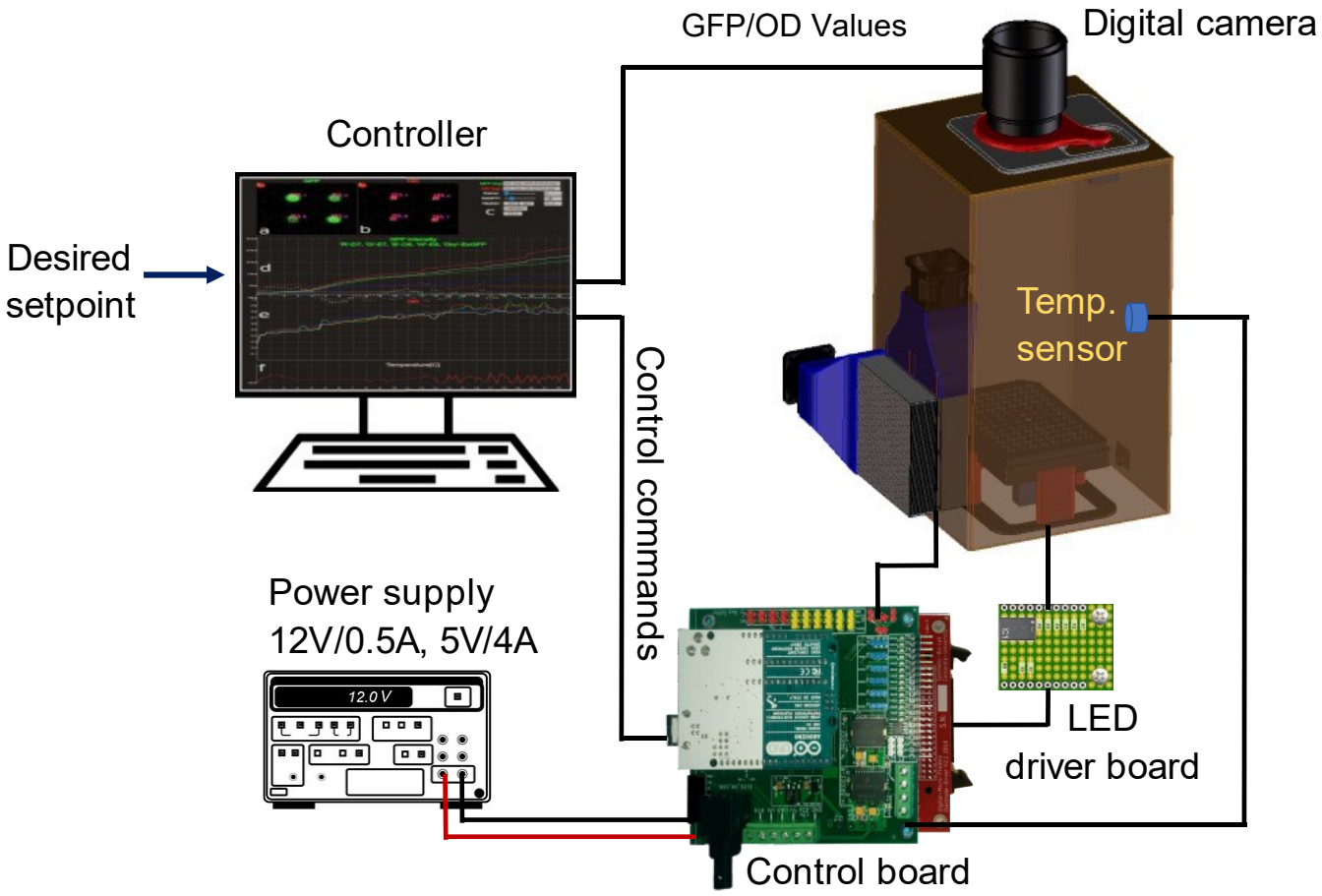

Figure 3.2: Automation system connectivity

A digital camera is located on top of the incubation chamber with a clear view over the incubation area. A plate holder module holding a well-plate (or microfluidic device) with bacterial cells is enclosed inside the incubator chamber. The plate-holder module includes a light delivery system consists of LED board that is driven by the LED driver and control board which are powered by a 12 V and 5 V power supplies (located outside of the chamber). The incubation chamber is insulated from ambient light and includes a heating element and temperature sensor used to maintain a constant temperature. The camera and control board are connected to the main controller (PC) that runs a software to control, monitors and log the entire experiment.

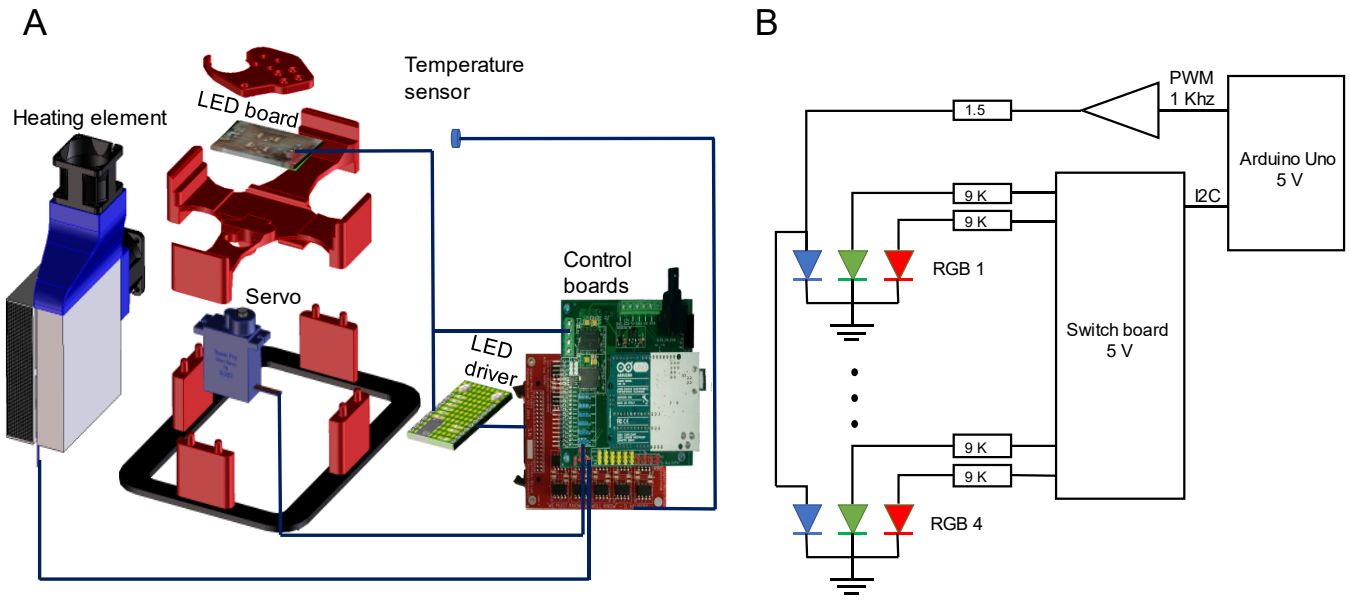

Figure 3.3: Light delivery connectivity

(A) Illustration of the electrical components. The ArduShield is connected to the switch-board, servo motor, heating element and temperature sensor. The switch board is connected to the LED driver that is connected to the LED board. **(B)** The light delivery electrical diagram. The four red and green LEDs are connected to the switch-board through a 9 kΩ serial resistor. The four blue excitation LEDs are connected in parallel to 5V, 1kHz PWM signal through a 1.5 Ω serial resistor. The excitation power is controlled by the duty cycle of the PWM signal.

A special feature with this setup is that the optogenetics setup can be remotely monitored and controlled using a webpage (**Figure 3.4**). As shown, the webpage shows the real-time images of the cells that are being illuminated by two wavelengths (480 nm and 650 nm) to measure the GFP and growth rate of the cells (**Figure 3.4A-C**). Further, the program is also capable of logging the experiment telemetry and plot the fluorescence and optical density in real-time **(Figure 3.4D and E**) while maintaining the temperature at the desired target (**Figure 3.4F**). We call this system "RT-OGENE" (Real-Time Optical GENEtics).

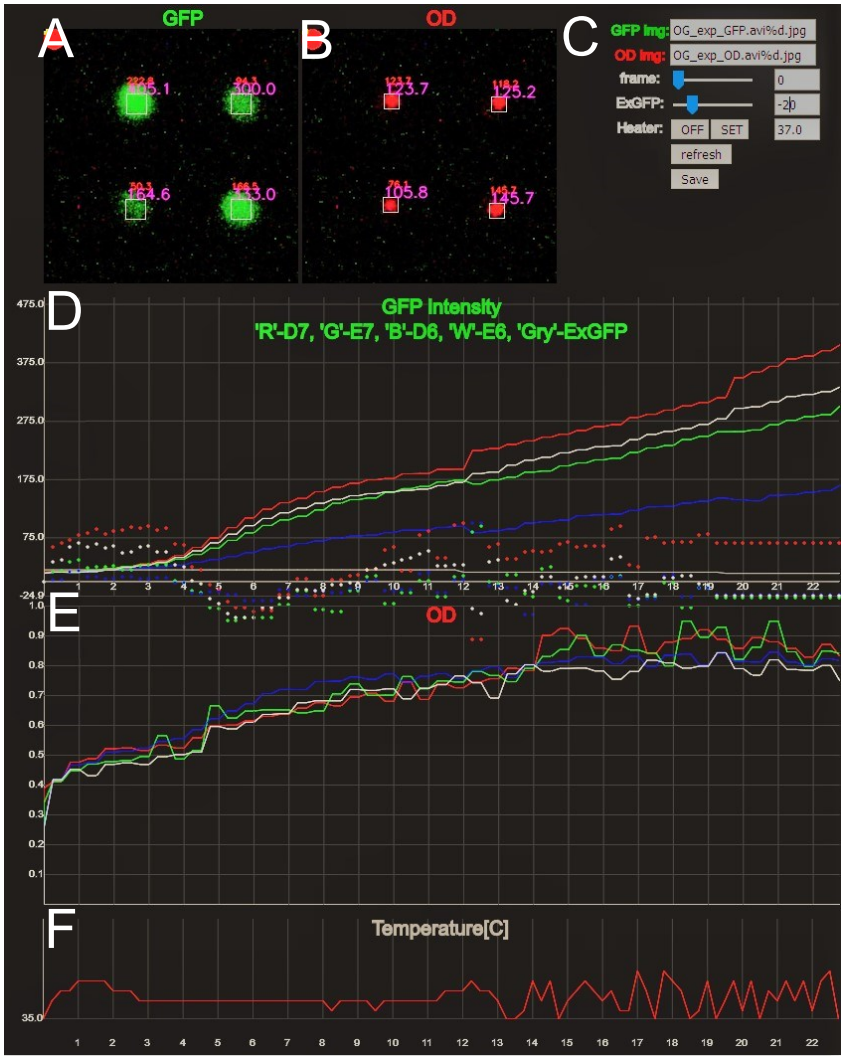

Figure 3.4: Real-time monitoring of the optogenetics experiment

A snapshot of the JavaScript based webpage to remotely monitor and control the optogenetics experiment. It consists of a real-time view of the **(A)** fluorescence and **(B)** OD_{600} pixelated values from the bacterial culture, **(C)** a parameter control toolbar consisting of fluorescence excitation power set and previous images navigation, **(D)** the fluorescence intensity and **(E)** the OD_{600} measurements over time of each sample, and **(F)** incubator temperature.

3.2 Software design

The software was build using Python2.7 programming language and three open-source libraries (Python 2.7 (32-bit), PySerial and OpenCV2). The main program, named Multi_Detection_RT, performs all the necessary activities to run the experiment (such as image capturing, optical filter selection, control-light activation and incubation temperature control). The program executes the main gene expression controller and maintains a log file that contains all the experiment's telemetry data. The software communicates with the GSOF_ArduShield board (by the author) that includes an Arduino Uno R3 SBC. This setup is a generic solution that is also used in ShihLab for DMF-based experiments. For this optogenetic-system it was used to control the system's electrical components (i.e. LEDs, heater and the servo motor of the filter selector). The communication with the GSOF_ArduShield was done using the GSOF_ArduBridge library (by the author). The USB camera was accessed by using the μEye camera library (by IDS Imaging Development Systems). The graphical webpage to control the experiment and present its telemetry data was programed in Java-script and used the P5 graphical library.

3.3 Light delivery system

In the RT-OGENE system, we used a three-colour sources system to activate different components of our system and to perform fluorescence and OD measurements. We used a green and red wavelength (520 and 650 nm) to activate and to de-activate the CcaS-CcaR system respectively and a blue (480 nm) light to excite the GFP (the gene that we want to control via the RT-OGENE system). Since the excitation (blue) and emission spectrum (green) of the GFP overlaps with the green light activation of the CcaS-CcaR system (**Figure 3.5A and B**), we hypothesized that the activation and emission of the CcaS-CcaR system under blue light excitation (instead of the green light) will also affect protein expression.

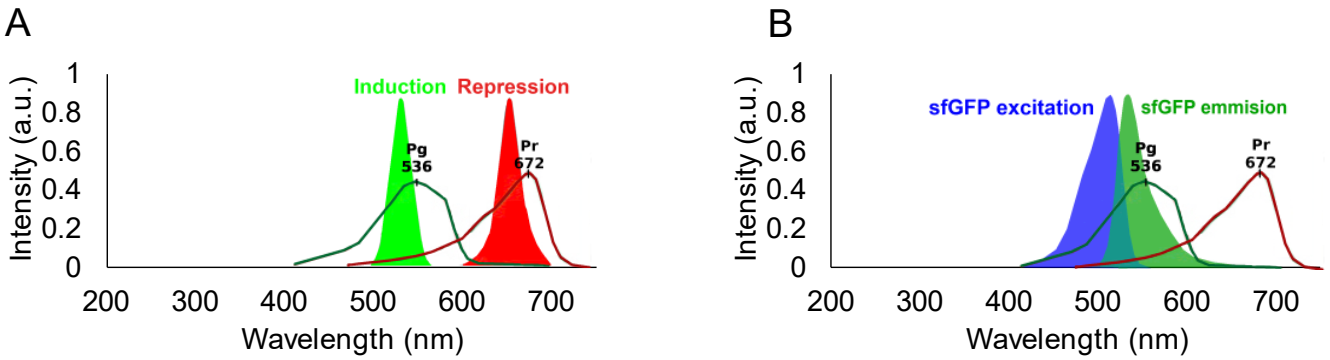

Figure 3.5: Light delivery spectrum

(A) Showing the spectrum domain to control the optogenetic regulation. The CcaS-CcaR activation and deactivation wavelengths are indicated by green (Pg) and red (Pr) lines, respectively. The green and red LED wavelengths are indicated by solid peaks. **(B)** Showing the spectrum domain for GFP measurement and the potential crosstalk with the optogenetic activation spectrum. The CcaS-CcaR activation and deactivation wavelengths are indicated by green and red lines, respectively. The blue (sfGFP excitation spectrum), and green (sfGFP emitted spectrum) are indicated by solid peaks.

As an initial test, we periodically switched on the blue and red light every five hours, while fluorescence measurements are obtained every hour. As shown in **Figure 3.6A**, GFP shows a clear upward trend while the CcaS-CcaR is being regulated by the blue light. We also verified this phenomenon by calculating the rate of protein expression. As shown in **Figure 3.6B**, there is an upward trend when the blue light is activated (generally a 2- to 4-fold increase in expression compared to when the red light is activated). We also normalized the fluorescence measurement by dividing the values by the OD_{600} measurement (**Figure 3.6C**) to determine if the increase in fluorescence is due to growth – and as expected, growth minimally affected the fluorescence values at the measured time points (0-5, 5-10, 10-15, 15-20 and 20-25 h) ($P < 0.05$; **Figure 3.6D**). Hence, to minimize the effect of the blue light on the CcaS-CcaR system, we only used the blue light (for 650 ms) to measure the GFP expression and continued to use green and red lights to activate and de-activate the CcaS-CcaR system.

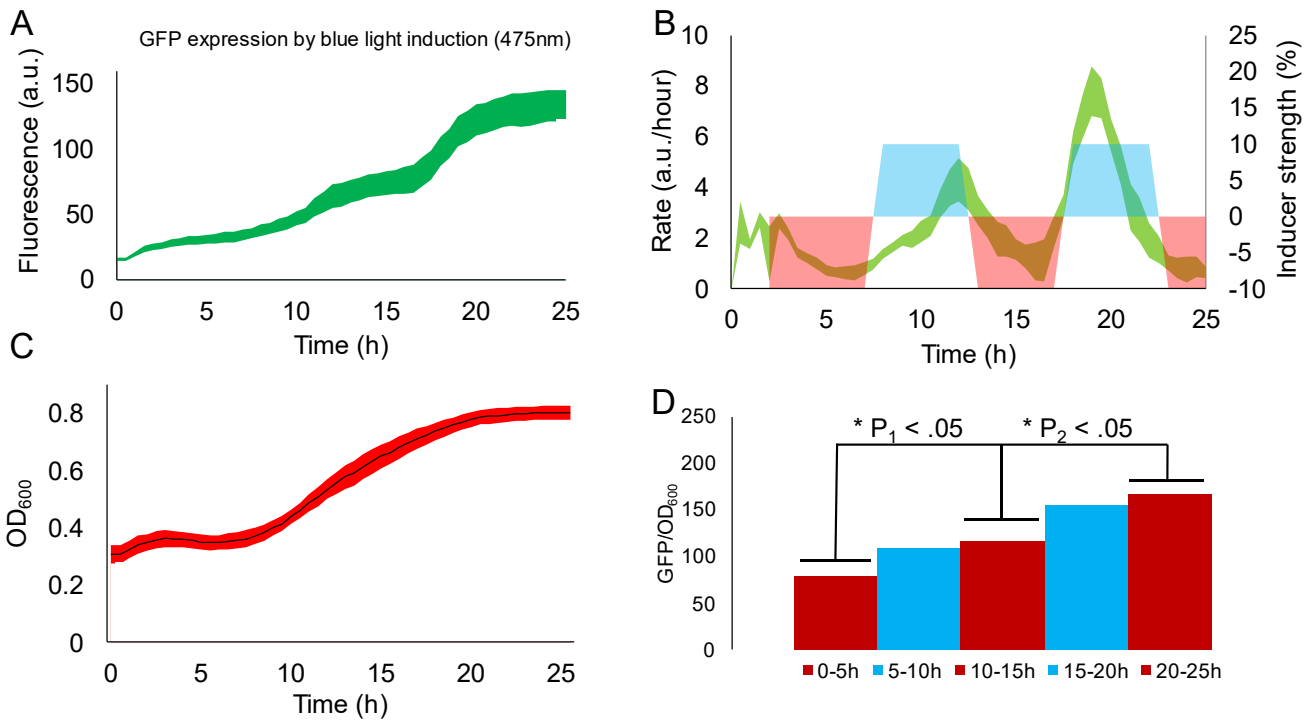

Figure 3.6: CcaS-CcaR sensitivity to blue light characterization

(A) Fluorescence accumulation of GFP over 24 h due to periodic activation (blue) and deactivation (red) of the CcaSR system every 5 h. **(B)** The illumination profile: induction (shaded blue) and repression (shaded red) cycles and rate of change in GFP expression rate (shaded green). The shaded green area represents the average and ± 1 S.D with n = 4. **(C)** Graph showing changes in OD during the experiment. (**D**) The normalized amount of accumulated GFP over OD represented as a bar plot. A fluorescence and OD_{600} value were obtained at the end of every red or blue light irradiation cycle. A Student's t-test ($p = 0.02 < 0.05$) was performed on the normalized fluorescence values over the red light cycles to show the effect of blue light.

A critical feature of the RT-OGENE system is that it is capable of simultaneous control of multiple samples which will require simultaneous activation of multiple lights. In our system, a LED is directed to one sample well which is in close proximity (~9 mm) to an adjacent sample well. Given the close distance, we evaluated the light leakage that is in the adjacent sample well when the sample well is light activated. As shown in **Figure 3.7C**, we measured the light intensity in the adjacent sample well for the green and red lights. When green and red light is activated on the target sample well, we observed a cross talk of 0.05 and 0.48 % in the adjacent well respectively by dividing the light intensity in the adjacent well by the light intensity in the target well. To reduce the cross talk in the adjacent well, we designed a beam narrower (see **Figure 3.7A and B**) that directs the light into the sample wells. With the beam narrower, we observe a significant reduction in the leakage in the adjacent well. Both green and red lights showed a reduction by a factor of at least five in the adjacent wells. To test leakage effect with the beam narrower on living cells, strains transformed with the CcaS-CcaR system were placed in individual wells (and a negative control that was in the dark) (**Figure 3.7D**). As expected, the measured fluorescence intensities at six hours are higher (at least 7x) than the initial fluorescence reading for the samples that are directly illuminated with the green light. We also measured the fluorescence in the adjacent wells and observed a lower fold value (~ 4x) compared to the direct illuminated samples which is similar to the negative control (~3x increase). We hypothesize that the leaky transcription of the system[41] together with cells growth during the six hours, increased the fluorescence in the samples from control and adjacent wells. We propose that other optical means like direct integration of fibers or collimating optics onto our platform can also further reduce the light leakage in adjacent wells.

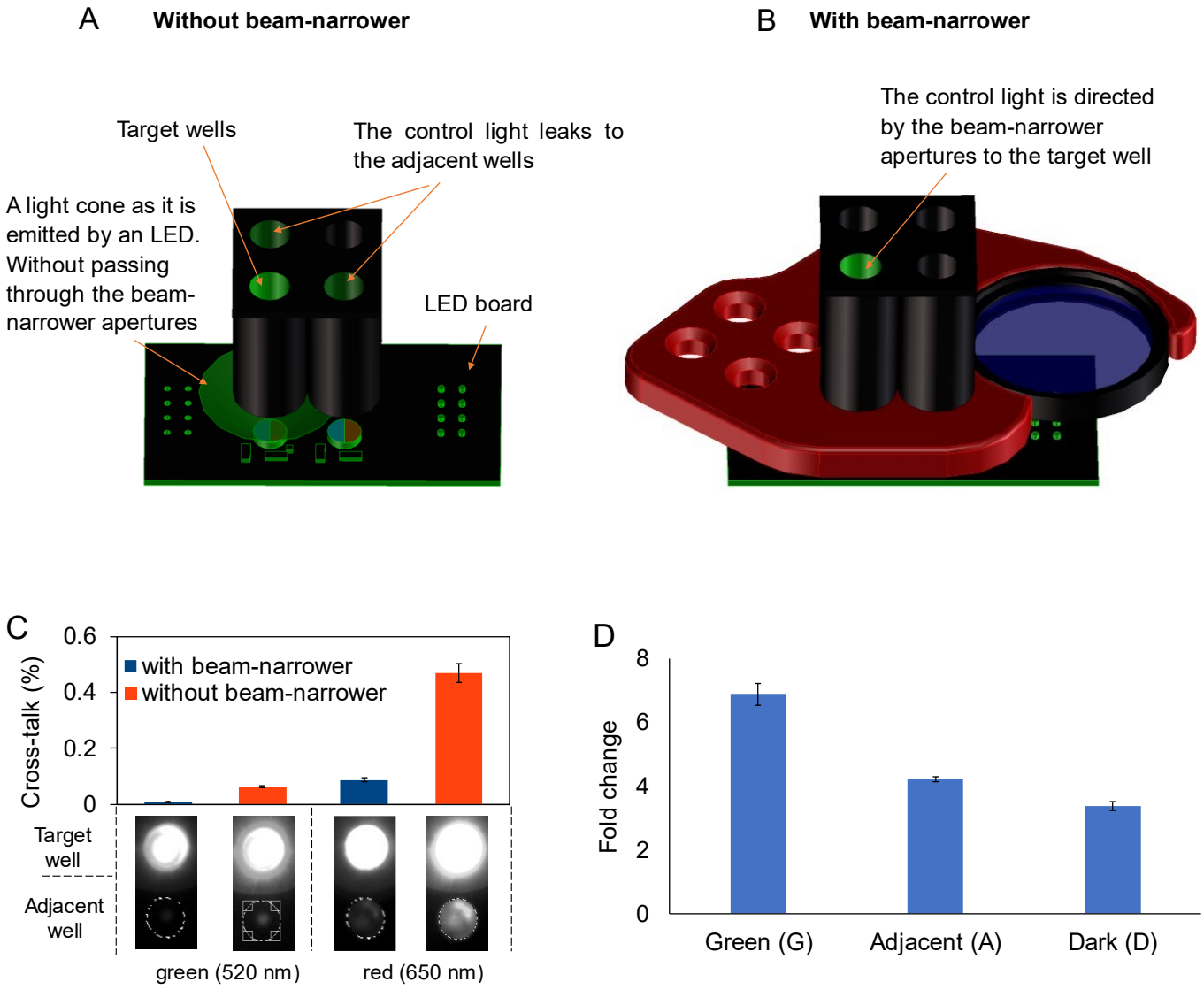

Figure 3.7: Beam narrower setup

(A) The light-beam (green) coming out of the LED at a wide angle and spreads to the adjacent wells. **(B)** The small diameter holes (apertures) in the filter-selector, direct the light-beam to the target well and reduces its spread to the adjacent wells. **(C)** Evaluation of the effect on using a beam narrower. A target well was illuminated by green or red light and with or without a beam narrower. The % cross-talk value was calculated by dividing the light intensity from the adjacent well by the light intensity from the target wells multiplied by 100 %. **(D)** Evaluating the cross-talk effect on bacterial cells. The quantification of the change in fluorescence values after being irradiated with green light for six hours in the target, adjacent, and a control 'dark' well. A 'dark' well was located 55 mm away from the target well. All error bars represent ± 1 S.D.

3.4 Fluorescence and OD_{600} measurement characterization

As a final characterization of the RT-OGENE system, we wanted to test the uniformity and the dynamic range of our optical density and fluorescence measurements. In these experiments, different cell densities were measured by obtaining an image of the samples and measuring their pixel intensities (**Figure 3.8A**). For OD measurements, the light intensity values were compared to OD_{600} values from a well-plate reader. As shown in **Figure 3.8B**, the measured intensities for one well showed excellent linearity with the OD_{600} measured by a standard well-plate reader ($R^2 = 0.9897$). In addition, we conducted the measurements for other wells and similarly showed the same trends (**Figure 3.8C-E**). Furthermore, given that our system can measure up to a maximum of pixel intensity of 255, we measured a limit of detection of ~ 0.29 $O.D_{600}$., which is suitable for the open- and closed-loop experiments.

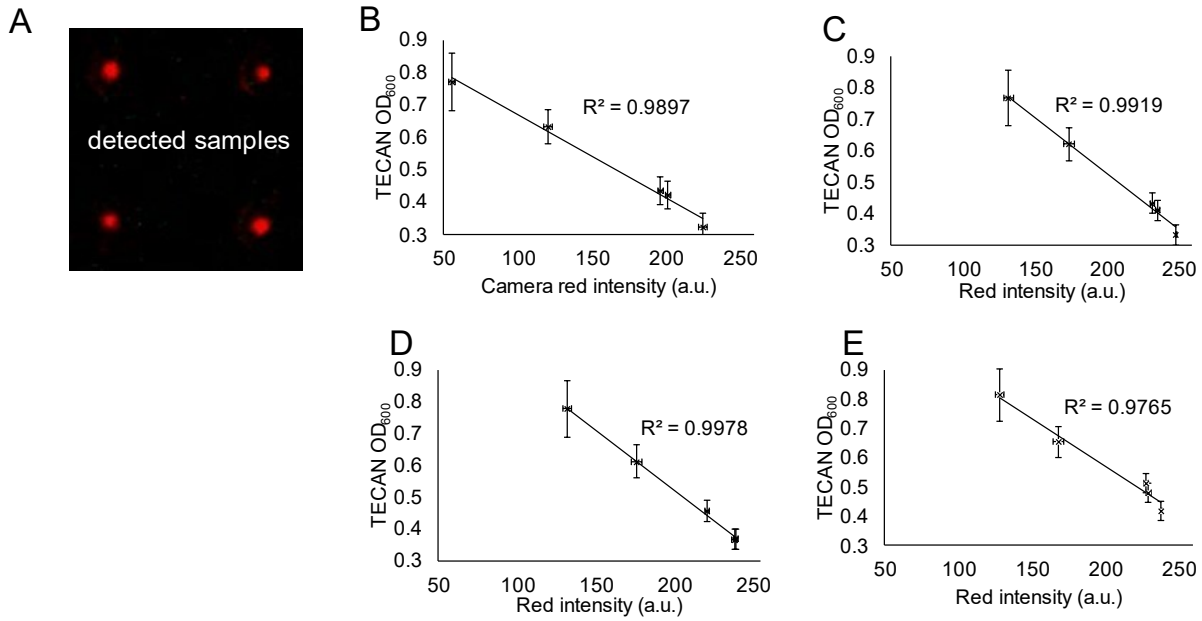

Figure 3.8: OD_{600} calibration of wells

(A) Images of four bacterial colonies illuminated by a red light using the RT-OGENE system. High intensity of red color indicated regions with lower cell-density. (**B-E**) Standard curves showing the relations between the measured red-light intensity using the RT-OGENE system and the OD_{600} from a well-plate reader for each well. The black lines represent the conversion function between the measured red pixels intensity to OD_{600}. All error bars represent ± 1 S.D for n = 4.

Next, we calibrated the fluorescence measurements using 6 fluorescein concentrations (1:6, 1:8, 1:10, 1:12, 1:14, 1:16) and measured their fluorescence at 20% excitation power. We compared the measured pixel intensity values with fluorescence values measured in a standardized well-plate reader. From initial experiments, higher concentration samples (e.g., 1:6 and 1:8) showed higher variability between the different wells (**Figure 3.9A**). One reason for the variability is due to slight differences in excitation power between the individual LED under each well. These differences translated into variations in the emitted excitation light and directly affected the fluorescence measurement. To alleviate the high variability between the wells, we corrected the deviation by multiplying a scaling factor which corrected the slopes to an arbitrarily chosen reference well (see Methods for equations and sample calculations). As shown in **Figure 3.9D**, the scaling factor reduces the variation between well-to-well measurements.

During calibration we noticed that the TECAN sensitivity and dynamic range are higher than our system (**Figure 3.9A**). The chosen approach to increase the sensitivity of the fluorescence measurement is to increase the excitation power, especially when the sample is expressing low fluorescence values (low concentration of GFP). This motivated us to determine the dynamic range of the samples as a function of excitation power (**as described in Equations 2.1-2.5**). We tested five different fluorescein concentrations (1:1, 3:4, 1:2, 1:4, 1:8 and 1:16) and measured their pixel intensities (i.e. fluorescence) as a function of excitation power. As shown in **Figure 3.9E**, fluorescence measurements of five different fluorescein dilution, were taken before applying the excitation power compensation. High fluorescein concentration shows steeper conversion curve between excitation intensity to fluorescence intensity and as a result the fluorescence measurement exhibits saturation at lower excitation power. On the contrary, low fluorescein concentrations show close to flat conversion curve and saturation is exhibited at higher excitation power. **Figure 3.9F** shows fluorescence measurements of the same five different fluorescein dilution after applying the excitation power compensation (**Equation 2.3**). The measurement of each dilution shows close to constant readings over a working zone (green) of excitation power values. The green region marks the range of excitation power suitable for each fluorescence intensity and GFP concentration. The area above the working zone marks that the camera's sensor is saturated. Below 6 % excitation power, the fluorescence intensity is too low and larger errors in the compensation algorithm were measured. The importance of the excitation compensation algorithm is that it enables us to compare readings taken at different excitation powers and dynamically lower the excitation power as the GFP concentration increases during the experiment. The compensated fluorescence value is calculated for each measurement in real-time.

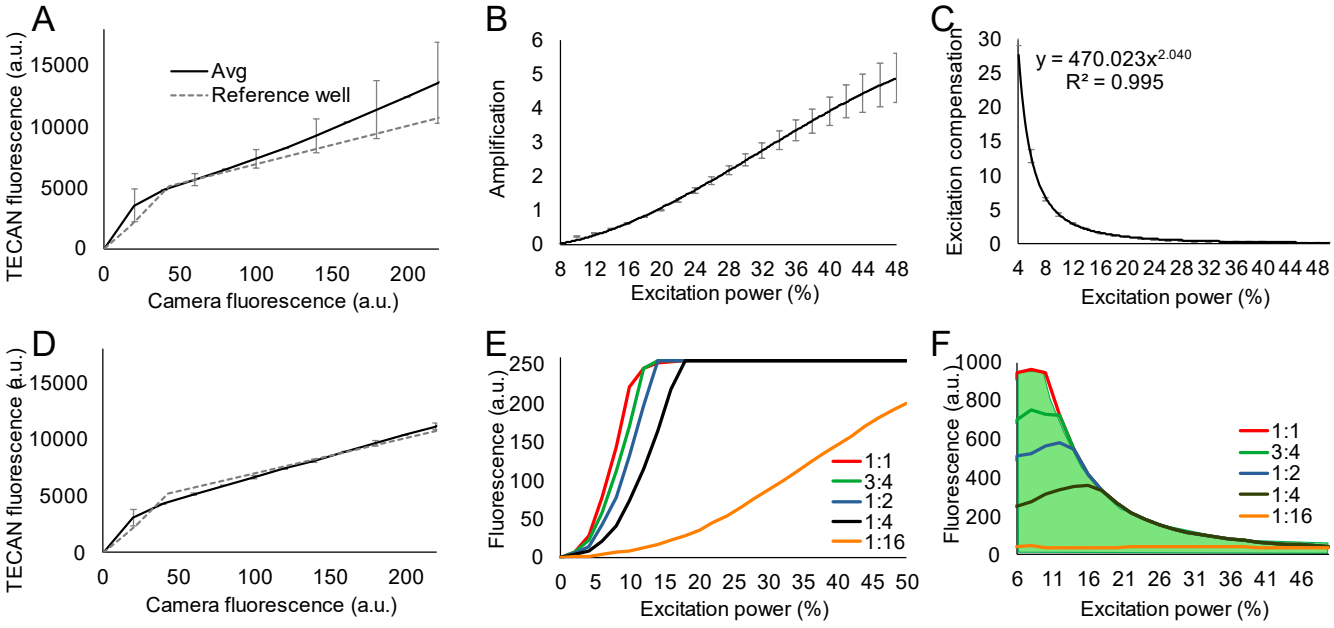

Figure 3.9: Fluorescence calibration results

Characterization of fluorescence measurements using the RT-OGENE system. Standard curves compared the relation between the measured fluorescence intensity using the RT-OGENE system and from a well-plate reader **(A)** before and **(D)** after independently calibrating each well to a reference well. The dashed line represents an arbitrarily chosen reference well. **(B)** The curve that represents the relation between the amplification of the fluorescence measurement as a function of the excitation power. **(C)** Plot showing the excitation compensation function. It was derived by calculating the inverse of the amplification function and approximated by the indicated function that was found using a power-fit function. **(E, F)** Evaluating the excitation compensation **(E)** before and **(F)** after applying excitation power compensation. Five fluorescein dilutions were analyzed under different excitation power. For compensation, the highlighted green region shows the range of excitation power suitable for each fluorescence intensity. The camera was saturated when the fluorescence values were outside the working zone (i.e. white area). All error bars represent ± 1 S.D for n = 4.

3.5 Modeling the biological system

As a first step in designing and building a control system is to determine the robustness of the current system. We have constructed an *in-silico* approach to obtain an estimation of the GFP production using a system of six first order differential equations. The system of equations is designed to predict the behaviour of the CcaS-CcaR system (see **Table 3.1** for constants). Equations 3.1-3.3 represents the constitutive expression of CcaR RNA, the protein and phosphorylation expression of CcaR system using the red and green light as two independent input variables. Equation 1.4 represents cell growth and equations 3.5-3.6 represents the RNA and protein expression of GFP. Using these set of equations, we have modeled the output of GFP production under different illumination profiles: constant, pulsed, and random. The cells growth was modeled in means of growth and death rate as was presented in other studies[58], [63].

Table 3.1 – Parameters and equations used for closed-loop experiments

Process	Equation
Equation (3.1) Constitutive CcaR mRNA transcription	$\frac{d}{dt}\mathrm{RNA_{CcaR}}(t) = \alpha_{0\mathrm{CcaR}} + \alpha_{\mathrm{CcaR}} - \gamma_{\mathrm{CcaR}} \cdot \mathrm{RNA_{CcaR}}(t)$
Equation (3.2) Transcription of de phosphorylated CcaR	$\frac{d}{dt}\mathrm{CcaR}(t) = \sigma_{\mathrm{CcaR}} \cdot \mathrm{RNA_{CcaR}}(t) - \beta_{\mathrm{CcaR}} \cdot \mathrm{CcaR}(t) + \mathrm{U_{red}}(t) \cdot \mathrm{CcaR_p}(t) - \mathrm{U_{grn}}(t) \cdot \mathrm{CcaR}(t)$
Equation (3.3) Phosphorylation of CcaR	$\frac{d}{dt}\mathrm{CcaR_p}(t) = -\beta_{\mathrm{CcaR}} \cdot \mathrm{CcaR_p}(t) + \mathrm{u_{grn}}(t) \cdot \mathrm{CcaR}(t) - \mathrm{u_{red}}(t) \cdot \mathrm{CcaR_p}(t)$
Equation (3.4) Cell growth	$\frac{d}{dt}\mathrm{cells}(t) = \alpha_{\mathrm{cells}} \cdot \mathrm{cells}(t) \cdot (1 - \mathrm{cells}(t)) - \beta_{\mathrm{cells}} \cdot \mathrm{cells}(t)$
Equation (3.5) Regulated transcription of GPF mRNA	$\frac{d}{dt}\mathrm{RNA_{GFP}}(t) = \mathrm{Cells}(t) \cdot (\alpha_{0\mathrm{GFP}} + \alpha_{\mathrm{GFP}} \cdot \frac{\mathrm{CcaR_p}(t)^n}{(\mathrm{CcaR_{max}} \cdot \mathrm{K_{hill}})^n + \mathrm{CcaR_p}(t)^n} - \gamma_{\mathrm{GFP}} \cdot \mathrm{RNA_{GFP}}(t))$
Equation (3.6) Translation of GFP	$\frac{d}{dt}\mathrm{GFP}(t) = \sigma_{\mathrm{GFP}} \cdot \mathrm{RNA_{GFP}}(t) - \beta_{\mathrm{GFP}} \cdot \mathrm{GFP}(t)$
Equation (3.7) Steady state value of CcaR	$\mathrm{CcaR_{max}} = \frac{\alpha_{\mathrm{cells}}}{\beta_{\mathrm{cells}}} \cdot \frac{\alpha_{0\mathrm{CcaR}} + \alpha_{\mathrm{CcaR}}}{\gamma_{\mathrm{CcaR}}}$

Where:
α_{0CcaR} = 0.002 - *CCaR RNA* leaky transcription rate
α_{CcaR} = 0.01508 - *CCaR RNA* transcription rate
Υ_{CcaR} = 0.0007 - *CCaR RNA* decay rate
σ_{CcaR} = 0.01508 - CCaR Translation rate
β_{CcaR} = 0.01508 - CCaR decay rate
α_{0GFP} = 0.002 - *GFP RNA* leaky transcription rate
α_{GFP} = 0.08 - *GFP RNA* transcription rate
Υ_{GFP} = 0.00032 - *GFP RNA* decay rate
σ_{GFP} = 0.000036 - GFP Translation rate
β_{GFP} = 0.0000002 - GFP decay rate
K_{hill} = 0.38 - Hills dissociation factor
n = 2 - Hills cooperative factor
σ_{Cells} = 0.0009 - Cells growth rate
β_{Cells} = 0.000015 - Cells death rate

$U_{red}(t)$ - Induced red light (input 1)
$U_{grn}(t)$ - Induced green light (input 2)
$RNA_{CcaR}(t)$ - Amount of *CcaR* RNA
$RNA_{GFP}(t)$ - Amount of *GFP* RNA
CcaR(t) - Amount of dephosphorylated CcaR
$CcaR_p$ (t) - Amount of phosphorylated CcaR
GFP(t) - Amount of GFP (output 1)
Cells(t) - Amount of cells (output 2)

Note: All amounts are in arbitrary units (a.u); translation, decay and death rates are in s^{-1}; transcription rate is in $a.u \cdot s^{-1}$; cell growth/decay rates are in $s^{-1} \cdot a.u^{-1}$; the amount of mRNA, proteins, and cells were chosen to be arbitrary (a.u) since the real amounts were not predicted.

Open-loop experiments

For the constant illumination profile, green light intensities of 0 %, 30 % and 45 % for 7 hours and are used to model the effect on GFP production. As shown in **Figure 3.10A**, when the green light is increased in intensity, we observe an increase in GFP production over the span of 7 hours. We also observe a slight increase in fluorescence when 0 % green light (or 100 % red light) is used. We have verified our model by experiment – i.e. inoculating a low starting density of bacteria (~0.2) and exposing the samples to the same intensities of green and red light. Strikingly, we observe the same trends as the model (i.e. increasing green light intensities will increase GFP production). The rate of expression for 0, 30, 45 % green light intensity are found to be 1.16, 4.3, and 9.5 a.u./h respectively (N = 4). We also measured the OD during this period (**Figure 3.10B**) to determine if the low rate of expression at 0 % green light is due to growth or due to leakiness in the promoter[71]. By comparing their normalized GFP/OD values at 0, 1, 3, and 5 h, we observe insignificant normalized values (16.5, 18.5 and 20 a.u.). This suggests that the increase is not due to the growth of the cells.

Having established a model and experiment to test open-loop response, we also tested two other types of illumination profiles: pulsed and random. For the pulsed experiments, cells where kept under red light for the first two hours from plating. Then, kept under green and red-light cycle at 40 % duty cycle

(with period time of 10 min) for four hours. Lastly, the cells where under red-light for two additional hours. We measured the fluorescence values and normalized it to OD values and observed that the experiments showed a similar trend compared to our model (**Figure 3.10C and D**). When the inducer is on, we observe an increase in fluorescence. When the inducer is off, we observe a 'levelling' plateau in the fluorescence, indicating the GFP expression is arrested. Furthermore, comparing the pulsed to constant illumination, we notice that the fluorescence at the end of 7 hours is similar in both cases. This indicates that a pulsed response is just as effective as a constant illumination at the same inducer percentage, which means that a constant illumination is not necessary to achieve a specific level of protein expression. Although we observe similar trends between the model and experiment, there are still some variations between them. For example, when red light is turned on (and green light is turned off), we notice that there was a longer delay (~ 1 h) for the fluorescence to 'level' off compared to the model. We estimate that the reason for this phenomenon is due to the translation processes that are still ongoing even after the transcriptional activation is arrested or that the dephosphorylation is slower than what we used in our model.

Finally, we also tested the robustness of our model by implementing a random illumination profile, which is testing the system against several different step responses. As shown in **Figure 3.10E**, the behavior of the random illumination is similar to the constant and pulsed profiles – i.e. when the illumination intensity increases, we observe increases in the fluorescence intensity or protein expression. In fact, when we change the ratio between green light to red light time ratio from 40 to 60 %, we immediately observe a two-fold increase in the fluorescence intensity (while the OD is constant – **Figure 3.10F**). When the green illumination is turned off (between 10-12 h), the fluorescence intensity remains constant. Interestingly, the experiment (and model) is able to track changes in the illumination profile which demonstrates the robustness of our system (and model) for the different types of illumination profiles

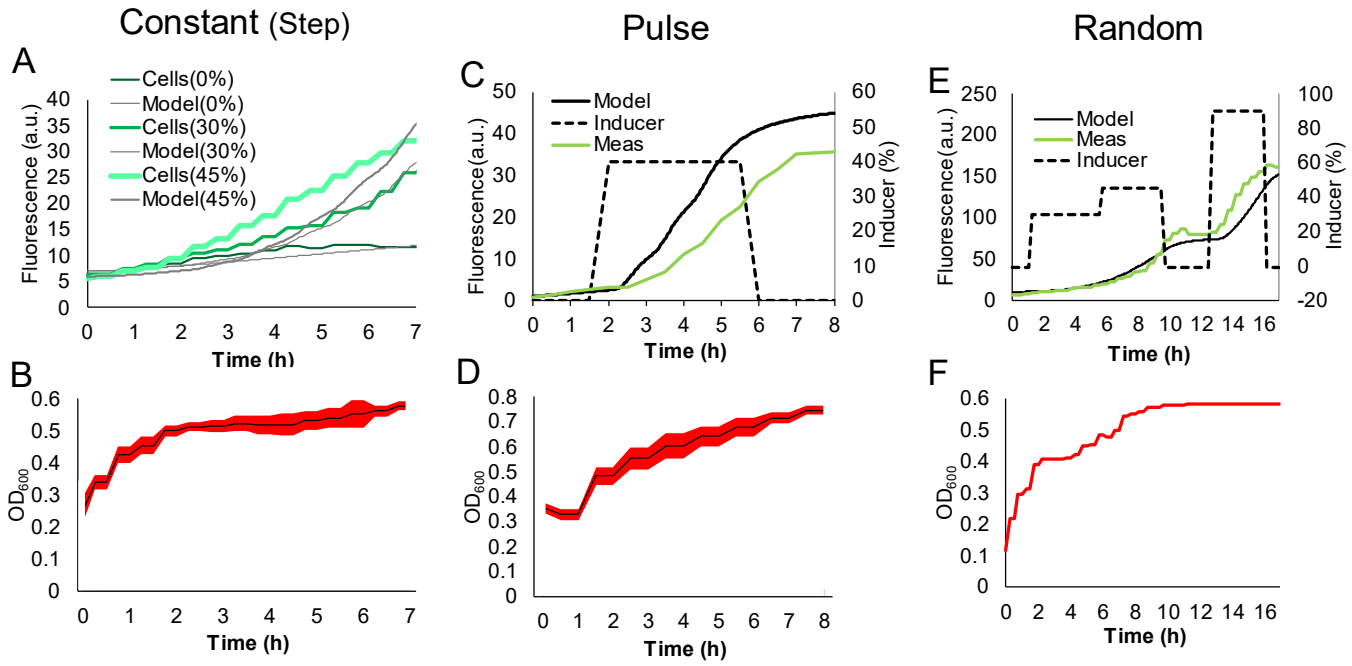

Figure 3.10: Open loop step response

The fluorescence and absorbance profiles of transformed CcaSR bacteria grown under different green- and red-light illumination profiles. **(A)** Constant: experimental and simulated profiles of fluorescence when bacteria were grown under three conditions: 0 % green light, 30 % green light (i.e. cycles 3 min green and 7 min red), and 45 % green light (i.e. cycles of 4.5 min green and 5.5 min red) for the full duration of the experiment of 7 h. The model was manually tuned to fit the experiment results **(B)** OD_{600} profile for constant inducer experiment (n = 3). **(C)** Pulsed: experimental (green solid) and simulated (black solid) profiles of fluorescence when bacteria were grown under green and red-light at 40 % duty cycle with period time of 10 min for 4 h. The activation pulse is indicated by a dotted black line. **(D)** OD_{600} profile for the pulsed experiment (n = 4). **(E)** Random: experimental (green solid) and simulated (black solid) profiles of fluorescence when bacteria were grown under green and red-light at 40 % - 60 % duty cycle with period time of 10 min for the full duration of the experiment of 16 h. The activation pulse is indicated by a dotted black line. **(F)** OD_{600} profile for the pulsed experiment (n = 4). For all absorbance curves, the solid line represents the average of the absorbance measurements and the red shaded region indicate ± one standard deviation for the data.

3.6 Controller design

After modifying the model to meet with the open-loop step response experiments, it was used to develop the controller. The *in-silico* approach is a convenient method to avoid lengthy trial-and-error optimization experiments that requires frequent validation and tuning a closed loop controller. Hence, we created a model that is linear and time invariant by linearizing the Hill function for 36 % induced strength and setting the time variant variables to their steady state values (e.g., cells(t) = 1) (**Figure 3.11**). The linear and time

invariant model determines the type of controller needed to achieve system stability at zero steady state error.

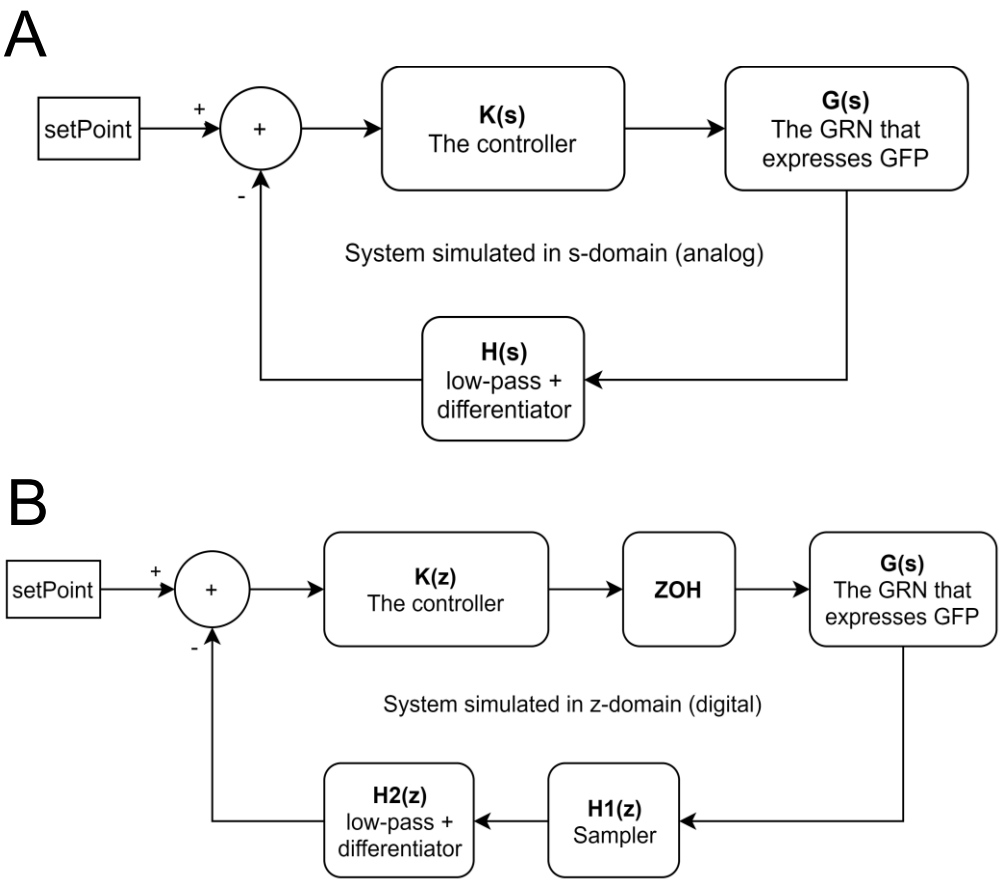

Figure 3.11: Modeling the system in the discrete time domain

(A) The system block diagram in the continuous s-domain. The plane has a continuous nature represented as G(s). The controller K(s) is implemented in software and will be modeled as a discrete function K(z). **(B)** The system block diagram in the discrete z-domain. A zero-order-hold and a sampler were used to mathematically 'match' between the continuous nature of the plant, G(s), to the discrete nature of the real-time controller K(z). By using the ZOH and a sampler, G(s) can be transformed into G(z). This way, the entire system can be simulated accurately in the z-domain.

The controller had to control the rate of GFP production rather than the total amount of GFP produced. For that, we choose to add a digital differentiator to the fluorescence intensity reading (as captured by the camera) in order to calculate the fluorescence rate of change. The calculated fluorescence rate value was used as the controller's feedback. The differentiator lowered the plant type from type-1 to type-0 (no poles as s=0 or z=1). This suggests that a type-1 controller will achieve zero steady error with a constant set-point (that in this case directly represents the expression rate). Therefore, a PID controller was chosen (The PID

was later reduced to PI by nulling the D coefficient). It is worth mentioning that, according to system control-theory, the same type of controller (PI or PID) can also be used in the absence of the differentiator (which would leave the plant as type-1). Adding the controller will result in a type-2 system that can track a ramp signal. Under these conditions, the input to the controller would be a ramp signal with a slop equal to the wanted expression rate. The advantage of adding the differentiator is that it removes the need to generate the ramp signal, thus simplifies the development and testing of the system. Next, the controller together with the differentiator and plant were modeled in the continuous s-domain (**Figure 3.11A**) and were further transformed to the digital z-domain by using two methods (**Figure 3.11B**): The GRN model was transformed using the Zero-Order-Hold (ZOH) method. The controller, differentiator and the low-pass-filter were transformed using the Euler's backward method. The sampling time for the discrete conversion was chosen to be 30 minutes. This time frame was set for two main reasons: The first was to minimize the cross activation of the promoter by the GFP-signal itself. The second was to meet the minimum necessary sampling demand for the system's dynamics. The time it took for the GFP signal to reach 90 % of its maximal values was approximately 7 hours, therefore a sampling period of 0.5-1 hours should be maximum possible value. Moreover, since the duration of GFP measurement was 1 second, the ratio between the control signal period to GFP sampling period was 1:1800 seconds, which equals to 0.055 % of the total sampling period. Lastly, the values for K_P and K_I were found (K_P=6.0, K_I=0.5) and the system was tested for stability at different working-points using the root-locus method. The criteria for choosing these PID values were the ability to maintain stability with a gain margin of four (the maximal difference observed in system response) and without saturating the controller output at all working points. To validate the controller's performances, we simulated the system responses to different set points. The set points were defined as the target production rate of GFP. All the closed-loop simulations were done using the LTI model, ignoring cell growth and nonlinear components. The controller's output was plotted (that is, how much green/red light to use) (**Figure 3.12A**), the expected GFP accumulation (**Figure 3.12B**), and the GFP production rate (**Figure 3.12C**). According to these results, the controller should be able to stabilize and maintain the defined rate of gene expression.

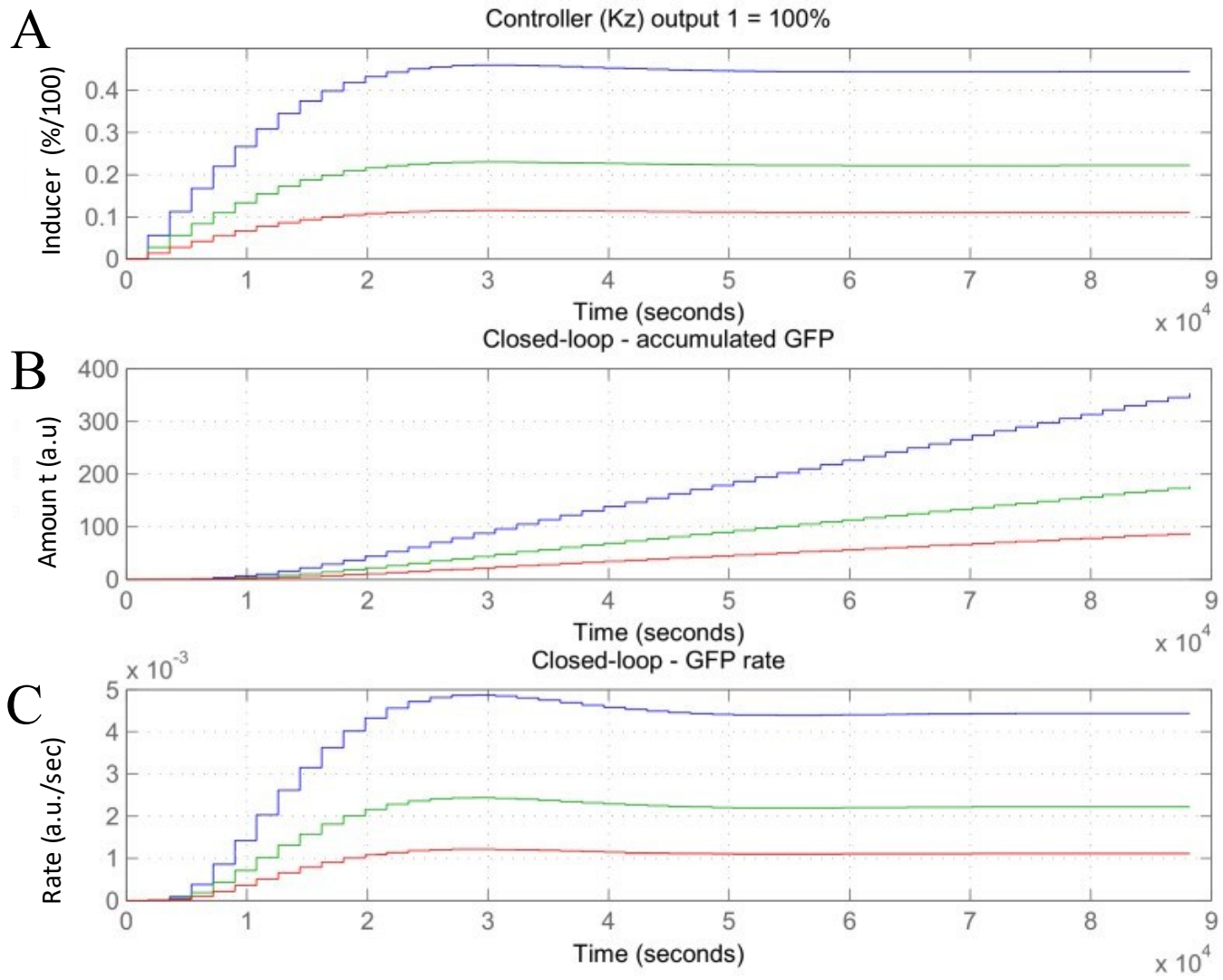

Figure 3.12: Simulation results using the LTI model

Three target expression rates were simulated: 4 a.u/h (red), 8 a.u/h (green), and 16 a.u/h (blue). Three plots were generated: (**A)** The controller's output. A value of 0.5 corresponds to 50 % inducer strength. **(B)** The simulated GFP accumulation and **(C)** the simulated GFP expression rates. The controller's coefficients were set to 6, 0.5, and 0 for K_p, K_i, and K_d respectively.

3.7 Controller implementation and real-time simulator of the entire system

After the appropriate controller was found **(Figure 3.13A)**, it was implemented. The low sampling period of the controller (30 minutes) removed any need for hard-real-time demands. To simplify the design, the controller was implemented in software using Python 2.7 as part of the main program. All numerical calculations were done using the default double precision floating point standard in order to maximize precision and dynamic range. The software development was done in a simulated environment of the entire system. The simulated environment allowed to test and verify each component of the software, in a deterministic, convenient and quick way. As for the simulated biological process, the originally developed non-LTI model was used and operated in an accelerated mode. The simulator is constructed from three independent programs. The first, is the GRN (biological plant) simulator that simulates the biological behavior of the system by solving the model. The simulator also includes a stochastic behavior by adding random noise to all the coefficients of the model. The inducer and repressor lights are the inputs and the GFP protein amount is the output of the simulation. The inducer / repressor (inputs) values are received remotely via UDP messages. For each simulation step a synthetic image of a colony that expresses GFP is generated. The intensity of the green color of the colony represents the simulation output value. To save development time, the simulation was configured to run in accelerated mode, meaning that every real-time second is equivalent to one hour of biological process. The second program is the controller that captures the synthetic image that was generated by the GRN simulator and applies the image-processing algorithm to quantify the amount of green (GFP) in each area of interest on the captured image. This value is the controller's feedback. The controller calculates the next state equation and output (inducer) value. Then, the output value is sent, as a UDP packet, to the specific IP address and PORT number of to the GRN simulator. In addition, the controller managed a user interface (console) that allows the user to change the set-point and control-loop parameters (Kp, Ki, Kd) in real-time. The third, is a real-time plot and data logging program (optional) - Receives the telemetry (as a UDP packet) from the controller and displays a real-time-plot while also logging all the results to a specified file.

The three application were configured to communicate with each other and this way they formed a fully simulated environment of a virtual biological experiment (**Figure 3.13B**).

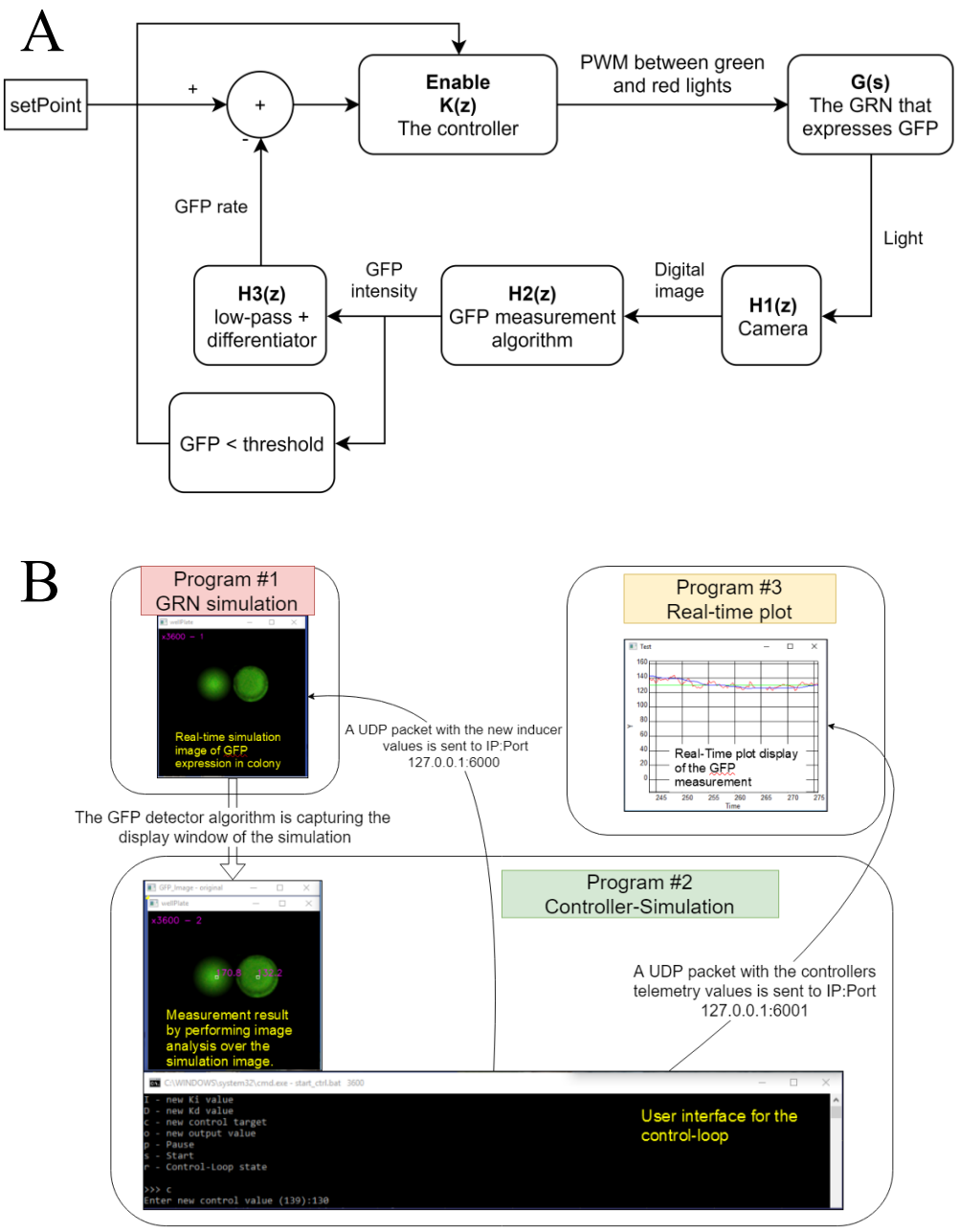

Figure 3.13: Controller design and simulation

(A) The block diagram of the final controller with the plant. as was implemented in software. The plant G(s) was used only for simulation purposes during the software development. **(B)** The GRN simulation architecture. The GRN simulator is constantly calculating the next step of the gene expression by solving the mathematical model. The gene expression result controls the intensity of the colony's green colour in the synthetic image. The controller captures the synthetic image (that emulates image capture by the digital camera) and runs the image processing algorithm on it. The result is used as the feedback signal of the control-loop. The controller calculates the new (next step) output value (amount of inducer / repressor) and sends it back to the biological simulator by using an IP: PORT 127.0.0.1:6000. The controller also sends telemetry data to the Real-time plot application by using the IP: PORT 127.0.0.1:6001

3.8 Closed-loop experiments

In the previous set of experiments, we have demonstrated the use of optogenetics as an effective modality to control protein expression. In many organisms (like *E.coli*), there is a fine balance in fine-tuning the timing and expression levels of enzymes that are involved in producing valuable products[72]–[74]. With this motivation, we describe the integration of our RT-OGENE system with a standard PID controller to control the rate of protein expression. Our first goal is to implement a setpoint (or target) for the rate of GFP production and determine if the PID controller can regulate towards this target fluorescence rate. We set the expression level of GFP to 8 a.u/h for 14 h and monitored the rate of fluorescence and the illumination profiles. As shown in **Figure 3.14A**, the rate immediately starts to increase to the setpoint by activating the inducer (i.e. green light). At 3 h, it overshoots the target by ~90 %, however the system activates the repressor (i.e. red light) to enable the rate to be reduced back to the target (**Figure 3.14B**). After 6 h, the rate starts to swing (± 30 %) around the setpoint. We also tested the PID controller tracking performance by increasing the setpoint to 12 a.u/h (at ~ 14 h) and immediately we observe a rate increase towards 12 a.u/h. During the 24 h, we observe an accumulation in the fluorescence (**Figure 3.14B**) and that the growth rates (~0.07 h^{-1}) are similar to the previous open-loop control experiments (**Figure 3.14C**). To test the limits of our system, we set the target rate to a higher expression rate of 22 a.u/h. As expected, the GFP accumulated rapidly towards the set point, and reaches the target after two hours (**Figure 3.14D**). At this time, the rate continues to increase (while the repressor is activated) creating a high overshoot (~190 %). (**Figure 3.14E**). For the case of this biological plant, the high overshoot and fast rise time is an expected result for high gain settings. There are solutions for reducing the overshoot – by tuning the PID parameters (**Figure 3.15A**), changing the setpoint gradually (**Figure 3.15B**) or changing to a different controller[75]. At any rate, despite the overshoot, we do obtain smooth outputs after 10 hours along with stable growth of the cells (**Figure 3.14F**).

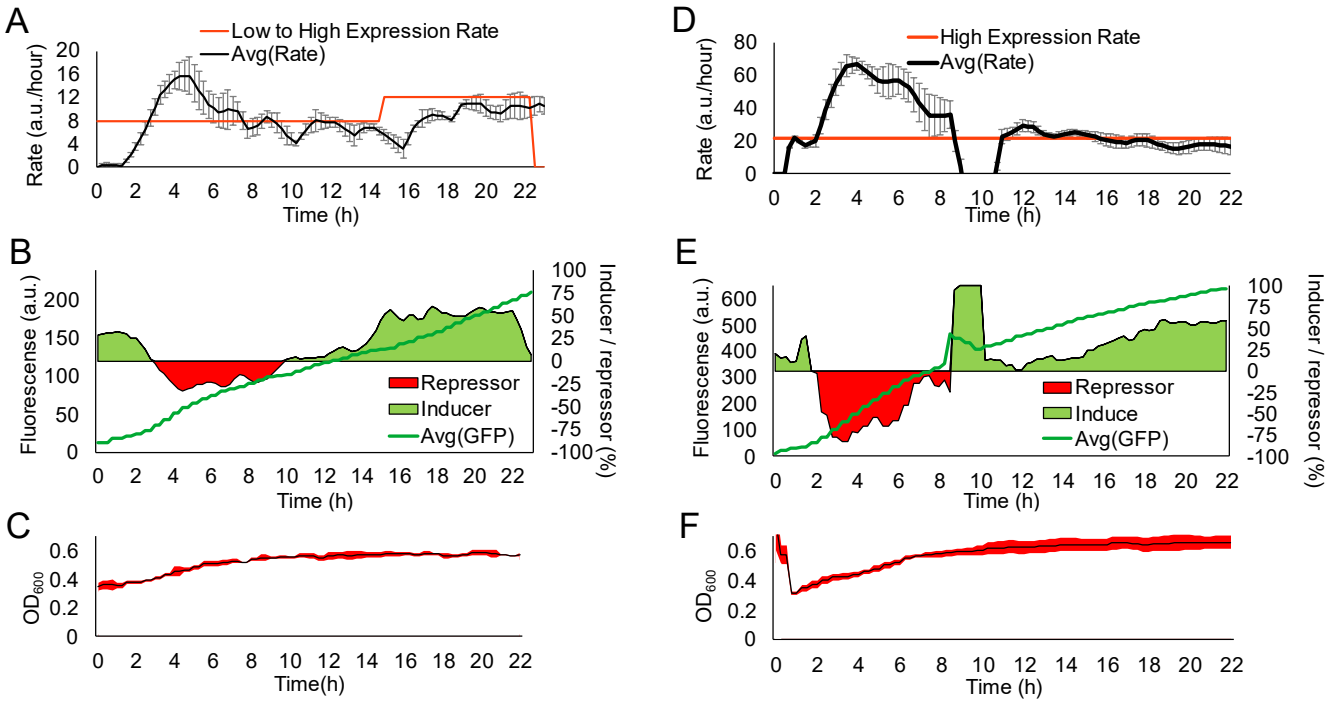

Figure 3.14: Closed loop step response

Comparison of fluorescence and absorbance profiles of transformed CcaSR bacteria grown under closed-loop control for a (**A-C**) low target setpoint and a (**D-F**) high target setpoint. **(A)** The expression rate setpoint of GFP was set to 8 a.u./h for 14 h and then increased to 12 a.u./h (red line; n = 4). The real-time tracking of GFP expression rate is indicated by a black line. **(B)** The overall accumulation of GFP in the system (green curve) and the illumination profile (shaded green and red areas) for the low rate setpoint. **(C)** An OD_{600} profile for the bacteria cells controlled at low rate setpoint **(D)** The expression rate setpoint of GFP was set to 22 a.u/h (red line; n = 4). The real-time tracking of GFP expression rate is indicated by a black line. **(E)** The overall accumulation of GFP in the system (green curve) and the illumination profile (shaded green and red areas) for the high rate setpoint. (**F**) An OD_{600} profile for the bacteria cells controlled at high rate setpoint. For all absorbance curves, the solid line represents the average of the absorbance measurements and the red shaded region indicate ± one standard deviation for the data. Error bars also represent ± 1 S.D with n = 4.

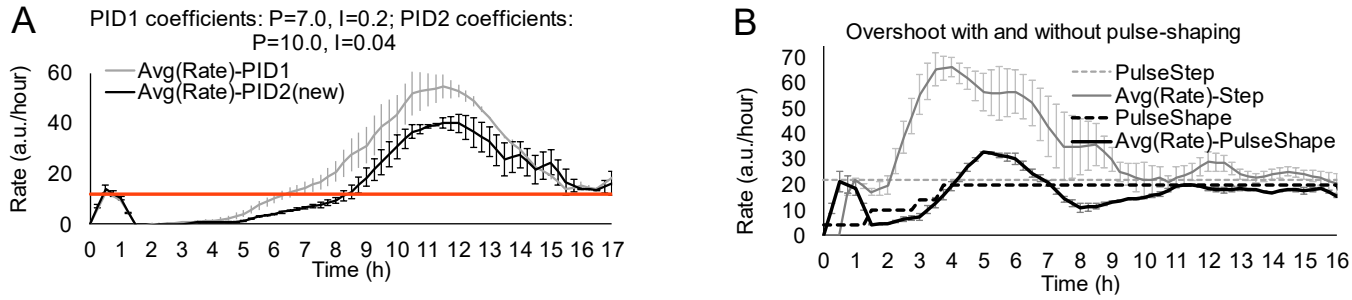

Figure 3.15: Reducing overshoot

(A) Closed-loop regulation using two sets of PID coefficients for a set point of 10 a.u./hour. Values of GFP expression rate when using PID controller (P=10.0, I=0.04, D=0) are shown in black line. Values in the case of PID controller (P=10.0, I=0.04, D=0) are shown in gray line (n=2). **(B)** Closed-loop regulation using gradual steps approach. Graph showing the changes in GFP expression rate for the gradual (black) vs step (gray) step transition for set points of 21 a.u./hour. In the case of the gradual transition the overshoot and settling time were significantly reduced. Error bars also represent ± 1 S.D with n = 2.

Having established a closed-loop controller with the RT-OGENE system to simultaneously regulate expression for all samples with the same setpoint, we applied the RT-OGENE system to regulate expression at two different setpoints to investigate our system's performance. We have prepared four bacterial cultures at a starting OD of 0.3 a.u. and two wells were set to a target rate of 14 a.u/h and two different wells were set to a rate of 28 a.u/h. As far as we are aware, this is the first report of an optogenetic method that is designed to control two different setpoints simultaneously. As shown in **Figure 3.16 A and B**, the system steadily increased towards their setpoints with an overshoot of 30 % and 80 % respectively. As expected, the higher setpoint showed higher overshoot and required a longer duration to stabilize near the target. The stabilization is more prominent from the light input - as more oscillations (between inducer and repressor) are shown when the set point is reached (**Figure 3.16 C and D**). The GFP accumulation also shows increases in both cases with the higher target exhibiting higher fluorescence values. We extended our experiments to track and to control four different setpoints (0, 4, 6, and 9 a.u/h) and similar results in terms of overshoot and accumulation have been observed (**Figure 3.16E and 3.16F**). Unfortunately, we do observe rate changes when the target is set to 0 a.u/h. This is presumably due to limitations of our optical system (i.e. cross-contamination from the lights aimed to the other wells , the original CcaS/CcaR system (i.e. a transcriptional leaky system)[41]. Thus, we tested the behavior of the cells by inoculating the starting culture at different ODs to determine if the starting density will affect the tracking behavior. Two starting cultures of OD 0.35 and 0.55 with the same setpoint of 8 a.u/h were evaluated using our RT-OGENE with PID system. When samples contained more bacteria at the start (**Figure 3.17A**), it shows a steeper increase in GFP expression rate during the first five hours and it also reached its set point faster (9 h vs. 10 h). Furthermore, we also observed a smaller overshoot (120 %) in contrast to samples with lower starting OD (250 %; **Figure 3.17B**). The main reason for this is that the GFP production rate is a function of cell density. Hence, the starting OD inevitably changes how the controller will regulate different expression rates and will change the dynamic behavior of the cells which will lead to different responses in the tracking profiles with small steady state error[76].

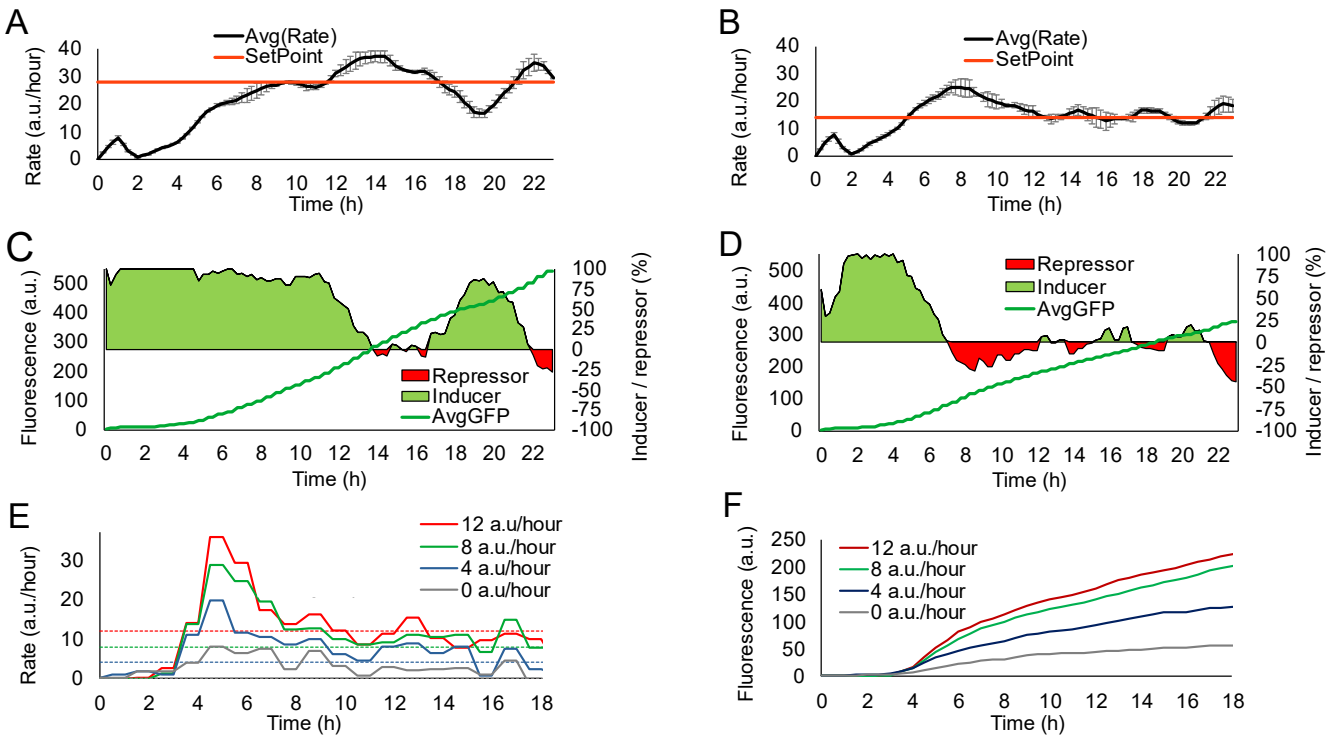

Figure 3.16: Simultaneous control of samples with different setpoints

Evaluation of controlling multiple setpoints in parallel using the RT-OGENE system. Controlling two setpoints: the expression rate of GFP was set to **(A)** 28 a.u./h in one pair of wells and **(B)** 14 a.u./h in another pair of wells (red lines). The averaged expression rate is indicated by a black line. (**C, D**) The accumulation of GFP in the system (solid green line) and the illumination profile (shaded regions) for the closed loop control for **(A)** and **(B)** respectively. Controlling four setpoints: **(E)** the expression rate of GFP (solid lines) is the closed-loop response to the setpoint (dotted lines) for each sample (matched by colour). **(F)** The accumulation of GFP in each sample. Error bars also represent ± 1 S.D with n = 4.

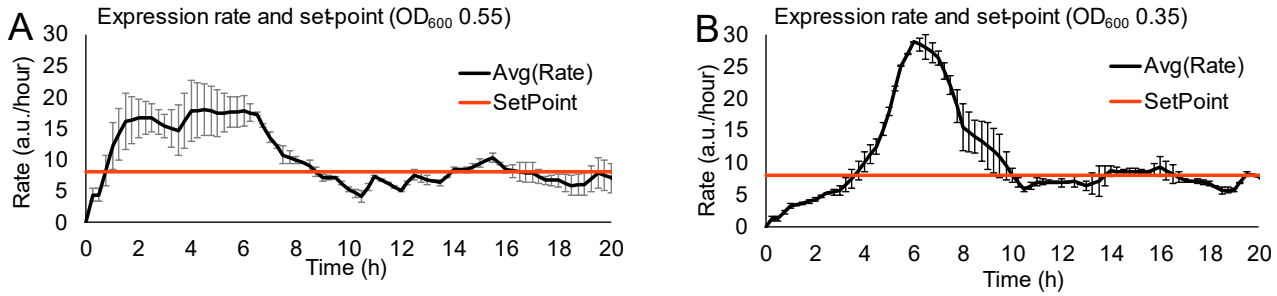

Figure 3.17: Regulating protein expression rate when starting at different OD600 values

Graphs showing the closed-loop regulation of GFP expression rate in wells containing bacteria culture at different initial OD_{600} conditions. The expression rate of GFP was set to 8 a.u./hour in both experiments. The averaged expression rate is indicated by a black line. **(A)** bacteria starting OD_{600} of 0.55. **(B)** bacteria starting OD_{600} of 0.35. Error bars also represent ± 1 S.D with n = 2.

3.9 Microfluidic platform

Microfluidics offer several advantages in terms of low volumes, smaller samples, automation, scalability, and portability. Due to these advantages, there have been numerous applications that have been integrated with microfluidics: cell culturing and analysis[77], [78], single cell-dynamics[79], [80], point-of-care diagnostics[81], [82], and genomic-based analysis[54], [83]. Here, we want to add to the repertoire of microfluidic applications which motivated us to design a microfluidic device that can be used for optogenetic analysis. In addition, we want to show the versatility of our system – i.e. the RT-OGENE is well-suited for well-plates and adaptable to other platforms like microfluidics. We fabricated a device (**Figure 3.18A**) that consists of one continuous channel, shaped in a U-configuration with four incubation chambers that are 1.5 mm in height (translates to ~ 2.25 μL of culture media). Two challenges arise when we use microfluidics for optogenetic analysis: (1) the significant reduction in volume can reduce the detection sensitivity of our analyte (fluorescence), and (2) the reduction in height can also reduce the accuracy and sensitivity of our OD measurements since we are reducing the path length by 3 Below we describe two experiments: open- and closed-loop experiments to investigate these challenges.

In open-loop experiments, we applied our RT-OGENE system to monitor the rate of gene expression of GFP and cell growth in the microfluidic chambers. Specifically, we developed an open-loop microfluidic protocol that will monitor the cell's fluorescence, by using 85 % excitation power, under different illumination profiles ranging from 10 to 60 % inducer strength and measuring the OD over time. As far as we are aware, this represents the first report of a microfluidic method for optogenetic analysis under open-loop (and closed-loop) conditions. As expected, fluorescence measurements at 60 % inducer strength produced the highest production rate after 20 h and at 10 % the fluorescence values showed the lowest production rate (**Figure 3.18B**). In addition, we have measured the OD for each well and have observed high variations in the average OD (**Figure 3.18C**). We speculate that the variations in the OD measurements may be caused by a number of factors. The most likely is that the microfluidic channel is connected and there is no separation between the wells. This design will allow bacteria to diffuse to other wells in the channel, which can vastly affect the OD readings in each well. Despite these drawbacks, a microfluidic device can be used to optogenetically regulate gene expression in microscale volumes.

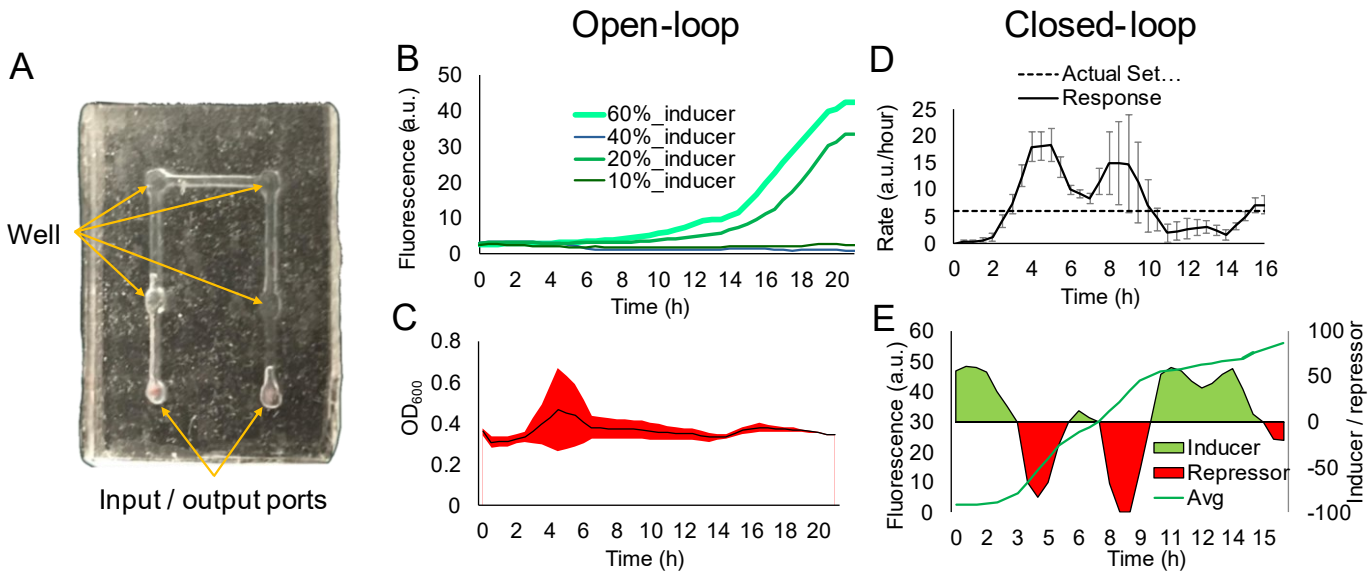

Figure 3.18: Microfluidic platform

Testing the RT-OGENE system with a microfluidic platform. **(A)** An image of the microfluidic device used for open- and closed-loop experiments. Open-loop response. **(B)** GFP accumulation in response to different inducer strengths (10, 20, 40, and 60 %) in the four microfluidic wells. 10 % represents cycles of 1min green and 9 min red **(C)** An averaged OD_{600} profile for the bacteria cells in response to 10 %, 20 % and 60 % inducer. Closed-loop response. **(D)** GFP expression rate (solid line) in four microfluidics at a target setpoint of 6 a.u./h (dotted line). **(E)** The overall accumulation of GFP in the system (green curve) and the illumination profile (shaded green and red areas) for the closed-loop control experiment. Error bars also represent ± 1 S.D with n = 4

In closed-loop experiments, we integrated the microfluidics with the PID controller to control gene expression in the microfluidic wells. We initially exposed the samples to green light and set a target rate of 6 a.u/h. Although the volumes of the samples were 40 times smaller, they all reached the target set point after ~8 h, which is similar to the trends seen in the well-plate experiments (**Figure 3.18D**). The similarities in the overshooting and steady state-error between the microfluidics and well-plates indicates that further improvement to the controller and the camera's sensitivity should be considered. Although improvements to the device itself are necessary, we propose alternatives to reduce both leakage and bacterial diffusion by using a mother machine microfluidic device that is capable of examining optogenetic circuits in single-cell format[71], [84] or the use of digital microfluidics given their inherent ability for individual addressability and control[70], [85]. Further, the experiment presented here was designed to show a comparison between well-plate and microfluidic formats for open and closed loop optogenetic control, using the RT-OGENE system. It is likely that by improving the optics in the RT-OGENE system or designing a different

controller, the variability in the measurements could be decreased while still allowing many samples to be analyzed.

4 Discussion

In this thesis I present a fully automated system that can monitor and regulate protein expression, based on light controlled GRN and a fluorescent output. This work reflects the implementation of a multidisciplinary approach in the construction of an optogenetic-synthetic system.

Though attempts to develop such systems were done before[42], [44], [86], [87], as far as I know, this set-up is the only one that is camera based, fits microfluidics system applications as well as well-plates, controls the rate of GFP expression in a non-chemostat environment and can simultaneously monitor several areas of interest. This system can be used as a versatile experimental platform for micro-manipulation of cell cultures, using standard culture plates.

4.1 Aspects in the design of the electro-mechanical system

4.1.1 Advantages of camera use in cell monitoring

In this work I chose to use a high-resolution digital camera to measure light intensity (i.e. fluorescence and OD) over other common approaches such as discrete photodiode sensors. The use of camera has several advantages; It allows greater flexibility in defining the location and size of the areas-of-interest. In addition, it enables the user to monitor several regions in parallel. These give the system higher flexibility regarding the shape or type of the mounted plate and the number and location of the target wells/areas of interest. Though quantification is done for light intensity in defined areas, using a camera provides a wide view of the plate, which can offer more information regarding the ongoing experiment (for example leaks in a microfluidic device or droplet movement in a digital microfluidic device).

In this system the camera was set to sample each pixel with 256 values (8 bits). This was done in order to allow compatibility with different types of cameras. To increase camera sensitivity and/or dynamic range the camera can also operate in a special acquisition mode of 1024 values (10 bits) per pixel. This is expected to increase the camera's sensitivity and dynamic range by a factor of four, thus increase its sensitivity to fluorescence measurements and potentially will allow to reduce the exposure of cells to excitation light. In addition, since the system was designed with monochromatic filters, the camera can be replaced with a high sensitivity black and white module to improve the system's sensitivity even further.

Lastly, using color camera allows the detection of multiple wavelengths in parallel and could be used with GRNs in which the output consists of more than one type of fluorescent signal (e.g. GFP and RFP).

4.1.2 Two- versus three- color approach in GRN activation

As part of the characterization of the biological system, I demonstrated the activation of gene expression in the GRN by blue light which is the same wavelengths that was used during the excitation of the fluorescent protein. As far as I know, this is the first time that the effect of blue light on the CcaS-CcaR GRN was demonstrated (**Figures 3.5 and 3.7**). Activation by blue light, in combination with the red light, which was used for transcription inactivation and OD measurements, can allow the use of a two-color instead of a three-color approach in GRN activation and sampling. This activation of the GRN and sampling of GFP signal by the same light source, could simplify the system electronics, wires and components needed. In addition, it can also further localize the controlled area since two LEDs can be located closed together (relative to three LEDs). Yet, I've decided not to use this two-color approach for several reasons. Transcription activation in the CcaS-CcaR GRN by blue light is less efficient than by green light, which results in the need for higher exposure intensity. In addition, the GFP emitted light is also found in the activation spectrum range of the GRN and could affect expression (though the effect of GFP emitted light on GRN activation was not directly tested in this work). The activation of the GRN by its own output could influence the system's response which is based on the starting conditions rather than the intensity of the activation light. That is, at low GFP concentrations the GRN activation is expected to be directly affected by the intensity of the control (blue) light. On the other hand, at high GFP concentrations low excitation (blue) light may induce a larger activation of the GRN than anticipated by the GFP emitted (green) light. Moreover, constant illumination with blue light could cause GFP bleaching, reducing the fluorescent signal or even damage the cells. All these directly affect the responsiveness of the system to the input. Therefore, I've chosen to use the three-color approach (green for GRN activation, red for GRN deactivation and blue for GFP excitation) and limited the sampling time to 0.055 % relative to the time the cells are under control-light.

4.1.3 Crosstalk between adjacent wells

One of the main features of the system is its ability to regulate gene expression in several regions of interest in an independent manner and in parallel. One of the challenges was designing and positioning the light source and filters to focus the light on the target wells. For this purpose, I've used a pinhole strategy, in which an opaque board with a small hole was place between the light source and the well plate, allowing a narrow beam to reach the cells. Using this strategy had significantly reduced the crosstalk of the activation light between adjacent wells, though not completely solved this problem. Yet, using the pinhole is an easy to use and modify, and therefore was used in this setting.

An alternative approach for the pinhole could be the use of optical lenses that can precisely focus the light beam onto the cells. In comparison to former approach, the use of lenses is more expensive and requires adjustments to allow precise focusing onto the focal plane. Further experiments are needed to determine the efficiency of lenses in comparison to pinhole in the reduction of crosstalk between wells.

4.2 Factors affecting set-point/expression rate stabilization

Most closed-loop experiments presented here reach and stabilize at their set point after approximately 8 hours, many times following a large overshoot and damped oscillatory motion. Since transcription and translation processes occur on the scale of minutes, why does it take so long to reach target expression rate by the controller? Interestingly, during the overshooting period at the beginning of most experiments the activation light was turned off and the cells were exposed to red inactivation light.

There are several possible explanations for this delayed response:

(1) Light contamination could lead to un-controlled expression patterns. Yet, this option is less likely to cause interferences in my system since the system was built in a dark box and light contamination was tested and controlled for.

(2) Another possible explanation could be auto-activation by GFP emitted light following excitation. At the beginning of each experiment GFP is produced as a result of the short activation pulse, as well as the promoter leakiness. Since GFP is a relatively stable protein, its accumulation is expected to produce an increasingly higher activation signal upon excitation (i.e. short exposure to blue light), independently from the controller regulated illumination by green activation light. In this case overshooting in GFP production rate is expected to dependent directly on the starting OD of the bacterial culture (number of cells), the levels of promoter leakiness and the length of the first activation pulse. In addition, if there is an auto-activation, replacing the GFP with other fluorophore should solve the problem. All these could be verified in future experiments.

(3) Lastly, improving the controller's performances could also contribute improvement in the system's kinetics. I performed a few preliminary experiments showing that changing the PID coefficients and taking a pulse shaping filter approach, in which the set point is gradually changed until reached to the desired value, could greatly reduce overshooting and shorten the lag time until reaching the set point.

4.3 Using a non-chemostat environment

A chemostat is an instrument designed to keep certain growth conditions in a culture by continuously replacing portions of the growth culture medium containing metabolites and cells with fresh medium. This is extremely important in systems designed to produce large amounts of proteins or other biological products, as it can keep the cells in an exponential growth phase, while maintaining the culture volume over long periods of time. On the other hand, in the microscale, using non-chemostat conditions are more commonly used for many laboratory applications, including cell-based (eukaryotic or bacterial) and cell-free systems, which examine dynamic changes over time or in response to stimuli. Including, the investigation of interactions between proteins (for example, ELISA essays), activation of promoters (for example, luciferase-reporter assays or activation of photo-regulated promoters as demonstrated here), dynamic changes in culture (used for example to examine aspects in cell cycle, or the response of cells to drugs), research focused on cellular (cytoskeletal based) movement, etc. Therefore, the system presented here can be modified to provide a simple and inexpensive solution for image-based, continues analysis of cultures using standard labware for many of the commonly used lab applications.

4.4 Future work and improvements

The designed system described here can be further improved and modified to give solution for future applications and improved performances.

4.4.1 The CcaS-CcaR optogenetic TCS

Recently, the CcaSR TCS was further optimized and published[88]. The new CcaSR-V3 has a folding factor of 600 (the ratio between transcription in the active versus inactive states). This is mainly due to a sharp reduction in the promoter leakiness. The CcaSR-V3 is highly recommended for future research.

4.4.2 GFP measurement dynamic-range versus resolution

Good sensitivity was achieved by setting the camera for long exposure-time (650ms) and GFP excitation power to maximum (but still around x10 less compared to the TECAN).

In order to increase the measurements dynamic range three approaches can be considered:

- Dynamic exposure adaptation - As the GFP expression intensity gets higher, lower exposure time can be used and linear compensation can be added to the sampled result.
- Dynamic GFP excitation power - As the GFP expression intensity gets higher, lower excitation intensity can be used and proper compensation can be added to the sampled result.
- Setting the camera to capture with 10-bit resolution mode – This will increase the dynamic range by four. Other options were discussed earlier and include changing to black and white CMOS camera of the same model, which is expected to increase the sensitivity by a factor of ~3. (Model: USB2.0 DCC1545M by THORLABS)

4.4.3 GRN light-control

There are several optimizations that can be done to improve the light path, reduce crosstalk between wells and focus the light onto the wells. This includes painting the filter selector with black 2.0v (to reduce light reflections) and consider embedding miniature lens on the optical selector.

4.4.4 Cell growth control

Genetic modifications to the GRN can be done to diverse and expend its functions. One such option, for example, is to control cell growth via a 'killer gene' or controlling through an antibiotic resistance gene. Introducing such a genetic module could be used in the context of my designed system to monitor and control growth rate of the cell culture.

4.5 Conclusions

In this work I present an electro-mechanical setup and associated software and user-interface, which can monitor and control protein expression and rate, based on fluorescent signal, using standard labware.

5 References

[1] S. A. Benner and A. M. Sismour, "Synthetic biology," *Nat. Rev. Genet.*, vol. 6, no. 7, pp. 533–543, 2005.

[2] A. S. Khalil and J. J. Collins, "Synthetic biology: applications come of age.," *Nat. Rev. Genet.*, vol. 11, no. 5, pp. 367–79, 2010.

[3] C. J. Bashor and J. J. Collins, "Understanding biological regulation through synthetic biology," *Annu. Rev. Biophys.*, vol. 47, no. 1, pp. 399–423, 2018.

[4] M. Xie and M. Fussenegger, "Designing cell function: assembly of synthetic gene circuits for cell biology applications," *Nat. Rev. Mol. Cell Biol.*, vol. 19, no. 8, pp. 507–525, 2018.

[5] J. C Venter *et al.*, "The sequence of the human genome.," *Science (80-.).*, vol. 291, no. 5507, pp. 1304–1351, 2001.

[6] R. Breitling, P. Armengaud, A. Amtmann, and P. Herzyk, "Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments," *FEBS Lett.*, vol. 573, no. 1–3, pp. 83–92, 2004.

[7] J. R. Karr *et al.*, "Supplementary material," vol. 13, no. 8, p. iv, 2012.

[8] D. K. Ro *et al.*, "Production of the antimalarial drug precursor artemisinic acid in engineered yeast," *Nature*, vol. 440, no. 7086, pp. 940–943, 2006.

[9] S. K. Lee, H. Chou, T. S. Ham, T. S. Lee, and J. D. Keasling, "Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels," *Curr. Opin. Biotechnol.*, vol. 19, no. 6, pp. 556–563, 2008.

[10] O. N. Ruiz, D. Alvarez, G. Gonzalez-Ruiz, and C. Torres, "Characterization of mercury bioremediation by transgenic bacteria expressing metallothionein and polyphosphate kinase," *BMC Biotechnol.*, vol. 11, no. 1, p. 82, 2011.

[11] J. C. Anderson, E. J. Clarke, A. P. Arkin, and C. A. Voigt, "Environmentally controlled invasion of cancer cells by engineered bacteria," *J. Mol. Biol.*, vol. 355, no. 4, pp. 619–627, 2006.

[12] J. Pitha, P. M. Pitha, and P. O. P. Ts'o, "Poly (1-vinyluracil): the preparation and interactions with adenosine derivatives," *Biochim. Biophys. Acta (BBA)-Nucleic Acids Protein Synth.*, vol. 204, no. 1, pp. 39–48, 1970.

[13] K. C. Schneider and S. A. Benner, "Oligonucleotides containing flexible nucleoside analogs," *J. Am. Chem. Soc.*, vol. 112, no. 1, pp. 453–455, 1990.

[14] P. E. Nielsen, "DNA analogues with nonphosphodiester backbones," *Annu. Rev. Biophys. Biomol.*

Struct., vol. 24, no. 1, pp. 167–183, 1995.

[15] H. O. Smith, C. A. Hutchison, C. Pfannkoch, and J. C. Venter, "Generating a synthetic genome by whole genome assembly: φX174 bacteriophage from synthetic oligonucleotides," *Proc. Natl. Acad. Sci.*, vol. 100, no. 26, pp. 15440–15445, 2003.

[16] D. G. Gibson *et al.*, "Creation of a bacterial cell controlled by a chemically synthesized genome," *Science (80-.).*, vol. 329, no. 5987, pp. 52–56, 2010.

[17] C. A. Hutchison *et al.*, "Design and synthesis of a minimal bacterial genome," *Science (80-.).*, vol. 351, no. 6280, p. aad6253, 2016.

[18] S. R. Scott, M. O. Din, P. Bittihn, L. Xiong, L. S. Tsimring, and J. Hasty, "A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis," *Nat. Microbiol.*, vol. 2, no. June, pp. 1–9, 2017.

[19] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli," *Nature*, vol. 403, no. 6767, p. 339, 2000.

[20] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.

[21] M. Ptashne, A. D. Johnson, and C. O. Pabo, "A Genetic Switch in a Bacterial Virus," *Sci. Am.*, vol. 247, no. 5, pp. 128–141, 1982.

[22] J. C. Dunlap, "Molecular bases for circadian clocks," *Cell*, vol. 96, no. 2, pp. 271–290, 1999.

[23] L. Potvin-Trottier, N. D. Lord, G. Vinnicombe, and J. Paulsson, "Synchronous long-term oscillations in a synthetic gene circuit," *Nature*, vol. 538, no. 7626, pp. 514–517, 2016.

[24] M. Tigges, T. T. Marquez-Lago, J. Stelling, and M. Fussenegger, "A tunable synthetic mammalian oscillator," *Nature*, vol. 457, no. 7227, pp. 309–312, 2009.

[25] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty, "A fast, robust and tunable synthetic gene oscillator," *Nature*, vol. 456, no. 7221, pp. 516–519, 2008.

[26] O. Mondragón-Palomino, T. Danino, J. Selimkhanov, L. Tsimring, and J. Hasty, "Entrainment of a population of synthetic genetic oscillators," *Science (80-.).*, vol. 333, no. 6047, pp. 1315–1319, 2011.

[27] I. Hoteit, N. Kharma, and L. Varin, "Computational simulation of a gene regulatory network implementing an extendable synchronous single-input delay flip-flop," *BioSystems*, vol. 109, no. 1, pp. 57–71, 2012.

[28] P. Hillenbrand, G. Fritz, and U. Gerland, "Biological Signal Processing with a Genetic Toggle Switch," *PLoS One*, vol. 8, no. 7, 2013.

[29] B. H. Weinberg *et al.*, "Large-scale design of robust genetic circuits with multiple inputs and outputs

for mammalian cells," *Nat. Biotechnol.*, vol. 35, no. 5, pp. 453–462, 2017.

[30] G. Chhetri, P. Kalita, and T. Tripathi, "An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli," *MethodsX*, vol. 2, pp. 385–391, 2015.

[31] B. C. Stanton, A. A. K. Nielsen, A. Tamsir, K. Clancy, T. Peterson, and C. A. Voigt, "Genomic mining of prokaryotic repressors for orthogonal logic gates.," *Nat. Chem. Biol.*, vol. 10, no. 2, pp. 99–105, 2014.

[32] A. K. Urbanek, W. Rymowicz, and A. M. Mirończuk, "Degradation of plastics and plastic-degrading bacteria in cold marine habitats," *Appl. Microbiol. Biotechnol.*, vol. 102, no. 18, pp. 7669–7678, Sep. 2018.

[33] P. P. Peralta-yahya and J. D. Keasling, "Advanced biofuel production in microbes," pp. 147–162, 2010.

[34] A. Tyagi, A. Kumar, S. V. Aparna, R. H. Mallappa, S. Grover, and V. K. Batish, *Synthetic Biology: Applications in the Food Sector*, vol. 56, no. 11. 2016.

[35] J.-Y. Trosset and P. Carbonell, "Synthetic biology for pharmaceutical drug discovery," *Drug Des. Devel. Ther.*, vol. 9, p. 6285, 2015.

[36] F. Jacob and J. Monod, "Genetic regulatory mechanisms in the synthesis of proteins," *J. Mol. Biol.*, vol. 3, no. 3, pp. 318–356, 1961.

[37] Y. Hirose, R. Narikawa, M. Katayama, and M. Ikeuchi, "Cyanobacteriochrome CcaS regulates phycoerythrin accumulation in Nostoc punctiforme, a group II chromatic adapter," *Proc. Natl. Acad. Sci.*, vol. 107, no. 19, pp. 8854–8859, Apr. 2010.

[38] J. J. Tabor, A. Levskaya, and C. A. Voigt, "Light-inducible promoter - E coli - 2011," vol. 405, no. 2, pp. 315–324, 2012.

[39] O. N. T. and T. J. J., "A Miniaturized Escherichia coli Green Light Sensor with High Dynamic Range," *ChemBioChem*, vol. 0, no. 0, 2018.

[40] E. J. Olson, L. A. Hartsough, B. P. Landry, R. Shroff, and J. J. Tabor, "Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals," *Nat. Methods*, vol. 11, no. 4, pp. 449–455, 2014.

[41] S. R. Schmidl, R. U. Sheth, A. Wu, and J. J. Tabor, "Refactoring and optimization of light-switchable Escherichia coli two-component systems," *ACS Synth. Biol.*, vol. 3, no. 11, pp. 820–831, Nov. 2014.

[42] G. Fiore, G. Perrino, M. Di Bernardo, and D. Di Bernardo, "In Vivo Real-Time Control of Gene Expression: A Comparative Analysis of Feedback Control Strategies in Yeast," *ACS Synth. Biol.*, vol. 5, no. 2, pp. 154–162, Feb. 2016.

[43] J. Melendez, M. Patel, B. L. Oakes, P. Xu, P. Morton, and M. N. McClean, "Real-time optogenetic control of intracellular protein concentration in microbial cell cultures," *Integr. Biol. (United Kingdom)*, vol. 6, no. 3, pp. 366–372, 2014.

[44] C. Fracassi, L. Postiglione, G. Fiore, and D. Di Bernardo, "Automatic Control of Gene Expression in Mammalian Cells," *ACS Synth. Biol.*, vol. 5, no. 4, pp. 296–302, 2016.

[45] A. Ay and D. N. Arnosti, "Mathematical modeling of gene expression: A guide for the perplexed biologist," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 46, no. 2. pp. 137–151, Apr-2011.

[46] J. Ang, E. Harris, B. J. Hussey, R. Kil, and D. R. McMillen, "Tuning response curves for synthetic biology," *ACS Synth. Biol.*, vol. 2, no. 10, pp. 547–567, 2013.

[47] T. Chen, H. L. He, and G. M. Church, "Modeling gene expression with differential equations.," *Pac. Symp. Biocomput.*, pp. 29–40, 1999.

[48] J. R. Karr *et al.*, "A whole-cell computational model predicts phenotype from genotype," *Cell*, vol. 150, no. 2, pp. 389–401, 2012.

[49] E. Fischer-Friedrich, G. Meacci, J. Lutkenhaus, H. Chaté, and K. Kruse, "Intra-and intercellular fluctuations in Min-protein dynamics decrease with cell length," *Proc. Natl. Acad. Sci.*, vol. 107, no. 14, pp. 6134–6139, 2010.

[50] E. Dekel and U. Alon, "Optimality and evolutionary tuning of the expression level of a protein," *Nature*, vol. 436, no. 7050, pp. 588–592, Jul. 2005.

[51] J. Bonnet, P. Subsoontorn, and D. Endy, "Rewritable digital data storage in live cells via engineered control of recombination directionality," *Proc. Natl. Acad. Sci.*, vol. 109, no. 23, pp. 8884–8889, 2012.

[52] A. E. Friedland, T. K. Lu, X. Wang, D. Shi, G. Church, and J. J. Collins, "Synthetic gene networks that count," *Science (80-.).*, vol. 324, no. 5931, pp. 1199–1202, 2009.

[53] H.-M. Kaltenbach, S. Dimopoulos, and J. Stelling, "Systems analysis of cellular networks under uncertainty," *FEBS Lett.*, vol. 583, no. 24, pp. 3923–3930, 2009.

[54] R.-S. S. Wang, A. Saadatpour, and R. R. Albert, "Boolean modeling in systems biology: an overview of methodology and applications," *Phys. Biol.*, vol. 9, no. 5, p. 55001, 2012.

[55] A. Szabo, C. M. Perou, M. Karaca, L. Perreard, J. F. Quackenbush, and P. S. Bernard, "Statistical modeling for selecting housekeeper genes," *Genome Biol.*, vol. 5, no. 8, p. R59, 2004.

[56] A. Krogh, M. Brown, I. S. Mian, K. Sjölander, and D. Haussler, "Hidden Markov models in computational biology: Applications to protein modeling," *J. Mol. Biol.*, vol. 235, no. 5, pp. 1501–

1531, 1994.

[57] F. X. Wu, "Gene Regulatory Network modelling: a state-space approach," *Int. J. Data Min. Bioinform.*, vol. 2, no. 1, p. 1, 2008.

[58] M. J. Dunlop, J. D. Keasling, and A. Mukhopadhyay, "A model for improving microbial biofuel production using a synthetic feedback loop," *Syst. Synth. Biol.*, vol. 4, no. 2, pp. 95–104, 2010.

[59] D. Dugar and G. Stephanopoulos, "Relative potential of biosynthetic pathways for biofuels and bio-based products," *Nat. Biotechnol.*, vol. 29, no. 12, p. 1074, 2011.

[60] B. R. Brooks *et al.*, "CHARMM: the biomolecular simulation program," *J. Comput. Chem.*, vol. 30, no. 10, pp. 1545–1614, 2009.

[61] S. M. Halper, D. P. Cetnar, and H. M. Salis, "An automated pipeline for engineering many-enzyme pathways: Computational sequence design, pathway expression-flux mapping, and scalable pathway optimization," in *Synthetic Metabolic Pathways*, Springer, 2018, pp. 39–61.

[62] M. N. Win and C. D. Smolke, "A modular and extensible RNA-based gene-regulatory platform for engineering cellular function," *Proc. Natl. Acad. Sci.*, vol. 104, no. 36, pp. 14283–14288, 2007.

[63] L. You, R. S. Cox, R. Weiss, and F. H. Arnold, "Programmed population control by cell–cell communication and regulated killing," *Nature*, vol. 428, no. 6985, pp. 868–871, 2004.

[64] P. Vladimir *et al.*, "Optimization of Golden Gate assembly through application of ligation sequence-dependent fidelity and bias profiling," *BioRxiv*, p. 322297, 2018.

[65] A. I. Andreou and N. Nakayama, "Mobius Assembly: A versatile Golden-Gate framework towards universal DNA assembly," *PLoS One*, vol. 13, no. 1, 2018.

[66] N. Bhattacharjee, A. Urrios, S. Kang, and A. Folch, "The upcoming 3D-printing revolution in microfluidics," *Lab Chip*, vol. 16, no. 10, pp. 1720–1742, 2016.

[67] J. Reedy and S. Lunzman, "Model based design accelerates the development of mechanical locomotive controls," SAE Technical Paper, 2010.

[68] J. C. Jensen, D. H. Chang, and E. A. Lee, "A model-based design methodology for cyber-physical systems," in *2011 7th International Wireless Communications and Mobile Computing Conference*, 2011, pp. 1666–1671.

[69] Y. Hirose, T. Shimada, R. Narikawa, M. Katayama, and M. Ikeuchi, "Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein," *Proc. Natl. Acad. Sci.*, vol. 105, no. 28, pp. 9528–9533, 2008.

[70] E. Moazami, J. M. Perry, G. Soffer, M. C. Husser, and S. C. C. C. Shih, "Integration of World-to-Chip Interfaces with Digital Microfluidics for Bacterial Transformation and Enzymatic Assays,"

Anal. Chem., vol. 91, no. 8, pp. 5159–5168, 2019.

[71] E. J. Olson and J. J. Tabor, "Optogenetic characterization methods overcome key challenges in synthetic and systems biology," *Nat. Chem. Biol.*, vol. 10, no. 7, p. 502, 2014.

[72] S. A. Nicolaou, S. M. Gaida, and E. T. Papoutsakis, "A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation," *Metab. Eng.*, vol. 12, no. 4, pp. 307–331, 2010.

[73] J. Winkler and K. C. Kao, "Tools for developing tolerance to toxic chemicals in microbial systems and perspectives on moving the field forward and into the industrial setting," *Curr. Opin. Chem. Eng.*, vol. 6, pp. 9–17, 2014.

[74] P. K. Ajikumar *et al.*, "Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli," *Science (80-.).*, vol. 330, no. 6000, pp. 70–74, 2010.

[75] G. Fiore, G. Perrino, M. Di Bernardo, and D. Di Bernardo, "In Vivo Real-Time Control of Gene Expression: A Comparative Analysis of Feedback Control Strategies in Yeast," *ACS Synth. Biol.*, vol. 5, no. 2, pp. 154–162, 2016.

[76] S. Kawasaki *et al.*, "Gene expression profiles during the initial phase of salt stress in rice," *Plant Cell*, vol. 13, no. 4, pp. 889–905, 2001.

[77] V. Lecault *et al.*, "High-throughput analysis of single hematopoietic stem cell proliferation in microfluidic cell culture arrays," *Nat. Methods*, vol. 8, no. 7, pp. 581–586, 2011.

[78] W. Li and J.-M. Lin, "Single-Cell Culture and Analysis on Microfluidics," in *Microfluidics for Single-Cell Analysis*, Springer, 2019, pp. 53–84.

[79] E. Eriksson *et al.*, "Optical manipulation and microfluidics for studies of single cell dynamics," *J. Opt. A Pure Appl. Opt.*, vol. 9, no. 8, p. S113, 2007.

[80] W. Zhang, D. S. Choi, Y. H. Nguyen, J. Chang, and L. Qin, "Studying cancer stem cell dynamics on PDMS surfaces for microfluidics device design," *Sci. Rep.*, vol. 3, no. 1, pp. 1–8, 2013.

[81] W. Su, X. Gao, L. Jiang, and J. Qin, "Microfluidic platform towards point-of-care diagnostics in infectious diseases," *J. Chromatogr. A*, vol. 1377, pp. 13–26, 2015.

[82] F. B. Myers and L. P. Lee, "Innovations in optical microfluidic technologies for point-of-care diagnostics," *Lab Chip*, vol. 8, no. 12, pp. 2015–2031, 2008.

[83] R. A. Kellogg, R. Gómez-Sjöberg, A. A. Leyrat, and S. Tay, "High-throughput microfluidic single-cell analysis pipeline for studies of signaling dynamics," *Nat. Protoc.*, vol. 9, no. 7, p. 1713, 2014.

[84] J.-B. Lugagne and M. J. Dunlop, "Cell-machine interfaces for characterizing gene regulatory network dynamics," *Curr. Opin. Syst. Biol.*, 2019.

[85] M. C. Husser, P. Q. N. Vo, H. Sinha, F. Ahmadi, and S. C. C. Shih, “An automated induction microfluidics system for synthetic biology,” *ACS Synth. Biol.*, vol. 7, no. 3, pp. 933–944, 2018.

[86] G. Fiore, G. Perrino, M. Di Bernardo, and D. Di Bernardo, “In-vivo real-time control of gene expression: a comparative analysis of feedback control strategies in yeast.”

[87] H. Wang and Y. T. Yang, “Mini Photobioreactors for in Vivo Real-Time Characterization and Evolutionary Tuning of Bacterial Optogenetic Circuit,” *ACS Synth. Biol.*, vol. 6, no. 9, pp. 1793–1796, Sep. 2017.

[88] N. T. Ong and J. J. Tabor, “A Miniaturized Escherichia coli Green Light Sensor with High Dynamic Range,” *ChemBioChem*, vol. 19, no. 12, pp. 1255–1258, 2018.

Appendix A – Controller's stability using root-locus method

The digital transfer function of the LTI model:

All results are derived by using MATLAB with system control toolbox.

The plant TF:

Gz =

$$\frac{4.317\,z + 3.564}{z^2 - 1.562\,z + 0.5619}$$

zeros at: -0.8257

poles at: 1.0, 0.5621

The differentiator TF:

Hz =

$$\frac{z - 1}{1800\,z}$$

zeros at: 1.0

poles at: 0.0

Plant and feedback TF:

The transfer function of the GRN in series with the differentiator (The differentiator cancels the plant integrator as intended):

GzHz =

$$\frac{z + 0.8257}{1800\,z - 1012}$$

series(Kz,GzHz)

$$\frac{0.2\,z^2 + 35.97\,z + 29.56}{1800\,z^2 - 2812\,z + 1012}$$

Root-locus stability around input=0.1:

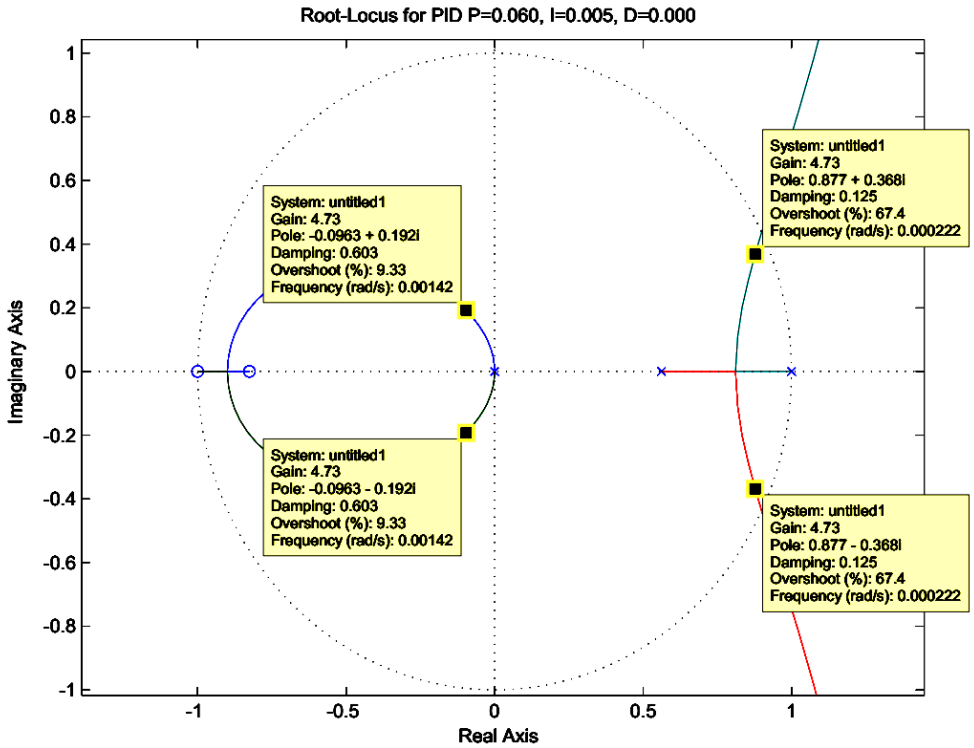

Figure A.1: The root-locus plot for the linearized system at input point 0.1.

Root-locus stability around input=0.35:

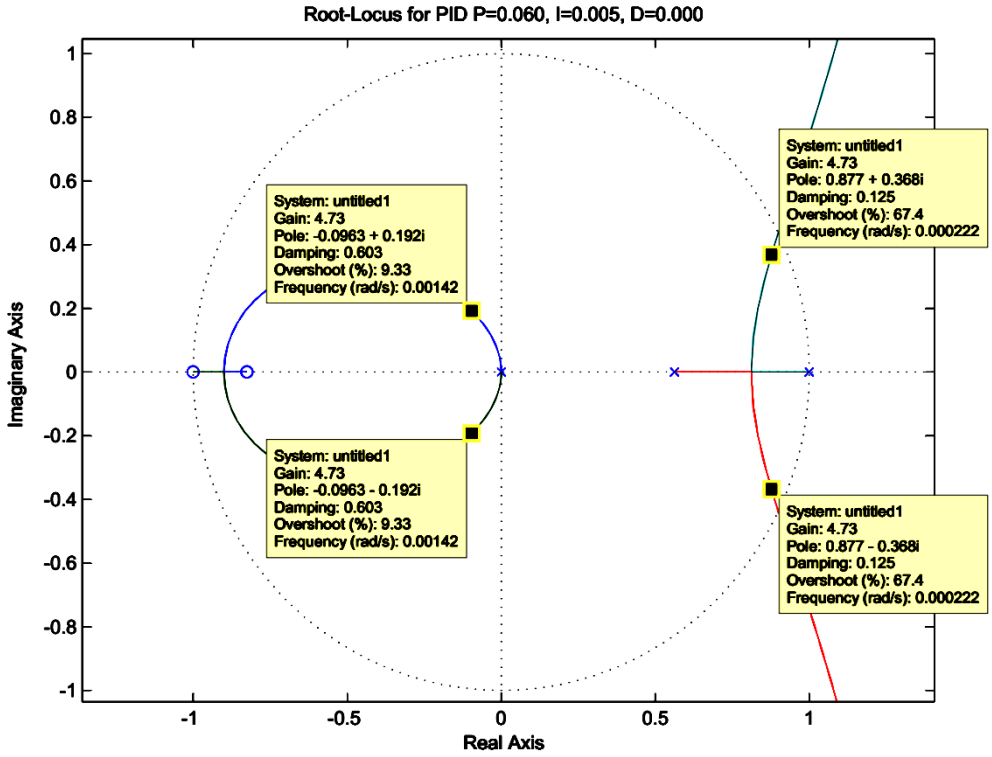

Figure A.2: The root-locus plot for the linearized system at input point 0.35.

Appendix B – Microfluidic

Historical background of miniaturization:

Shifting away from traditional glass tubes and dishes started in the 1950s when Dr. G. Takatsy described the first microliter well-plates. The new well plates were designed to have predefined volumetric capacities (in the range of microliters to milliliters) and the wells were arranged in a standardized, rectangular matrix. Takatsy laid the foundations for early bench-top miniaturization and his contemporary methods are the base for today's high-throughput biological applications.

Microfluidics:

Microfluidics is a relatively recent field. It is defined as the use of micro- to pico-volume fluid handling systems. It is used to further minimize and automize many biological protocols. Microfluidics devices are often described as 'labs-on-a-chip'. The three leading techniques to handle fluids on the micro scale are:

1. Channel microfluidics.
2. Droplet in channel microfluidics.
3. Digital (droplet) microfluidics.

The difference between the three are in throughput, flexibility and manufacturing price.

Where the most flexible ones are the digital microfluidics but the ones that have the most throughput are channel based. Other than that, all microfluidics platforms benefit from small volumes and high surface area to volume ratio.

The conservation of reagent is important when these are difficult or costly to obtain, as it is for many chemically synthesized potential pharmaceuticals, or in limited availability (such as blood samples from newborn infants). The high ratio of surface area to volume accelerates heat transfer and mass transfer rates, resulting in the greater efficiency of exothermic and endothermic reactions. The increase of surface-area-to-volume ratios has long been a method of engineering the enhancement of biosensors and reaction cells for various applications.

Channel microfluidics is the most widespread methods. In this method the fluid is pumped or socked into a small diameter channels (~1 μm), in most areas of the chip the Reynol's numbers are small which means that the dominated forces are the viscosity forces. This results in laminar flow in which multiple fluid streams can flow in parallel within the same channels without mixing together. When mixing is needed one of the techniques is to designate a special channel with higher Reynolds number.

Equation 1: Reynold's Number

$$\Re = \frac{\text{Net Inertial Forces}}{\text{Net Viscous Forces}} = \frac{\rho v L}{\upsilon}$$

where:

ρ is the fluid density (kg/m3)

ν is the mean velocity (m/s)

L is the characteristic length of the system

υ is the kinematic viscosity (m2/s)

In those areas mixing will happen in relatively shorter amount of time.

Another popular microfluidics technique is droplet microfluidics, like channel microfluidics but with water droplets in oil channels. The droplets are often generated in a junction were the water flow is crossed by oil flow. As a result, water droplets are pinched off from the water stream and are separated by the oil medium. This technique allows ultra-rapid generation of droplets for high-throughput applications (10-100 kHz). Each droplet is analogous to a test tube, especially in systems that can perform operations downstream of droplet formation such as the sorting or merging of droplets to perform a reaction.

DMF background and introduction:

Digital-Micro-Fluidic (DMF) is also known by the term Electro-Wetting-On-Dielectric (EWOD).

The usage of electro-wetting-based actuation of droplets for microfluidics application was first reported in 2000 by Pollack et al.166

DMF is the manipulation of fluids as discrete droplets on an array of electrodes. The advantages of DMF is that the user has full control over the droplet movement, time of movement and in some cases the volume of the droplet (as shown by the chipArray and other).

The movement of the droplet is governed by electrical forces. The electrical forces are a direct result of electrical field that is produced by an electrode or combination of electrodes (**Figure B.1**).

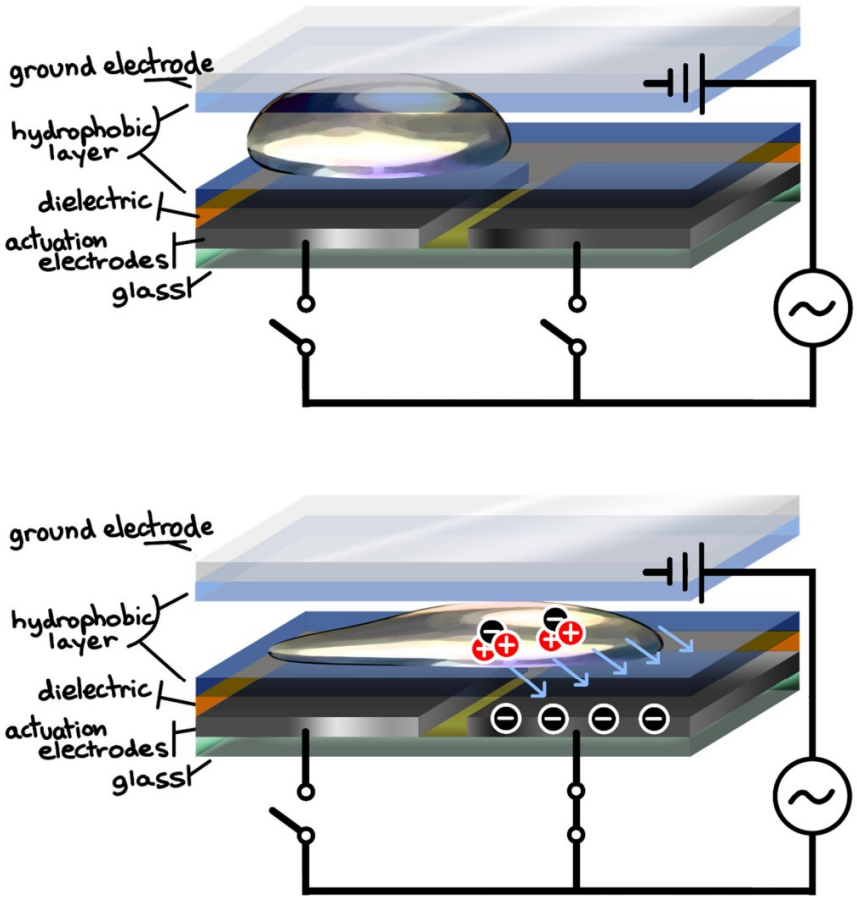

Figure 5B.1: Schematic description of Digital-Micro-Fluidic device

A droplet is being moved by electrical forces, which are the direct result of electrical field that is produced by an electrode or combination of electrodes. Taken from Laura Leclerc thesis.

The study of fluid dynamics on EWOD has resulted in the latest electro-mechanical model equation which describes the effect of the applied voltages on the force acting on the droplet. (**Equation B.2**)

Equation B.2:

$$E(x,L,V) := \frac{L}{2}\cdot\left[x\cdot\sum_{i=0}^{N}\frac{\varepsilon_0\cdot\varepsilon_{r\_liq_i}\cdot\left(V_{i,0}\right)^2}{d_i} + (L-x)\cdot\sum_{i=0}^{N}\frac{\varepsilon_0\cdot\varepsilon_{r\_fil_i}\cdot\left(V_{i,1}\right)^2}{d_i}\right]$$

where:

x - is the droplet displacement (m)

L - is the (m)

V - is the applied voltage (V)

d - is a matrix that contain the thickness values of the coated layers (m)

ε_{r_fil}- is a matrix that contain the dielectric values of all the materials that are surrounding the droplet (usually the chip layers and air or oil) (dimensionless)

ε_{r_liq}- is a matrix that contain the dielectric (relative permittivity) values of all the materials of the droplet (usually the chip layers and water) (dimensionless)

ε_0- is the dielectric constant in vacume 8.854*10^-12(V/m)

Basic DMF chip structure:

The typical set of movements needed to achieve a fully functional DMF chip is as follow (**Figure A.2**):

1. Lateral movement – the ability to move an individual droplet from one point to the other.
2. Split – the ability to split droplet into two droplets.
3. Marge – the ability to combine two droplets into one.
4. Dispense – the ability to generate a single droplet out of a bigger liquid reservoir.
5. Mix – the ability to mix the content of a droplet. Usually by lateral or circular movement.

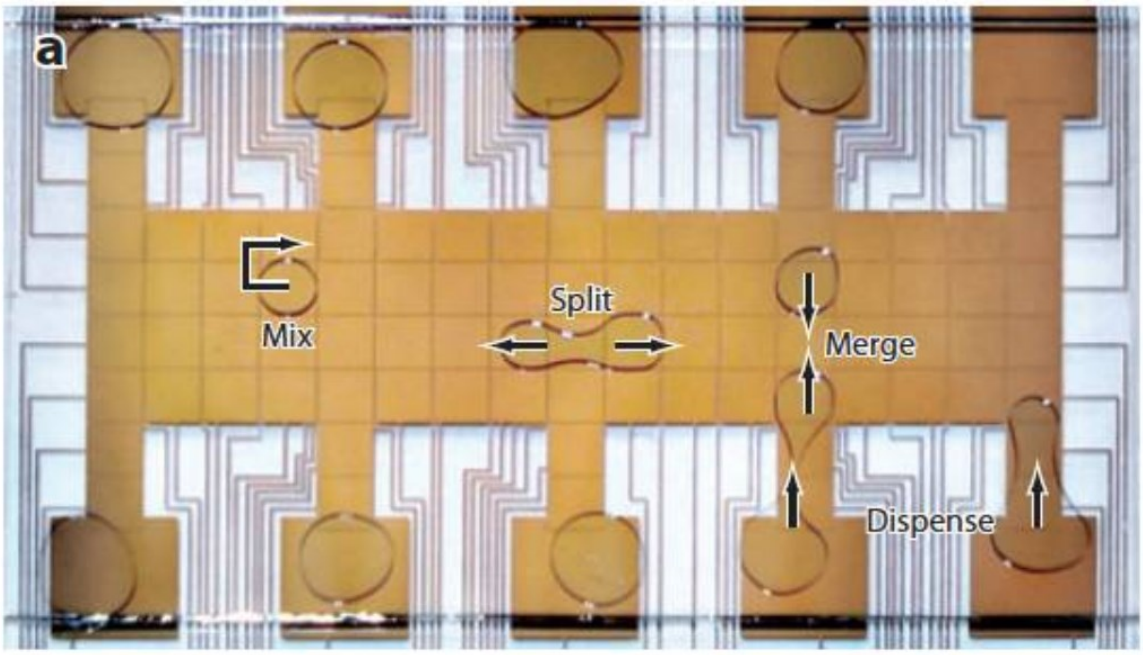

Figure B.2: Droplet manipulations using MDF chip

Image of an MDF microfluidic device showing basic droplet manipulations. Taken from Choi et al. 46

Current DMF challenges and limitations:

Evaporation – Because typical DMF chip structure is not hermetically sealed water and reagents evaporation occurs. In room temperature, a single droplet of 0.1uL will evaporate after less than 1 hour. Moreover, during the evaporation period, the concentration of the reagents in the droplet is increasing. Since most common incubation periods requires reagents at 37°C, evaporation occurs even faster, and a single droplet can evaporate after 10 minutes. Not addressing this issue makes a long term fully automated protocol impossible to conduct.

Manufacturing costs:

Some DMF chips are designed with a very high definition of fixtures and very small gaps between the electrodes. This results in high manufacturing costs and delivery times, since in most cases scientists need to design, manufacture and test it by themselves. This procedure is relatively expensive and time consuming. In many cases, a well-trained person can produce a single chip in ~6 hours. The manufacturing equipment needed includes:

1. Clean room.
2. Chemicals.
3. Glass chips.
4. Mask (usually prepared by an external company).
5. Ultraviolet exposure machine.
6. Spin-Coater or Pyrlen-Coater.
7. Chemical disposing facility or services.

Reliability and repeatability:

Most chips will experience movement faulting after a few hours of usage. Different designs and manufacturing methods can increase the chip reliability, nonetheless, the chip performance is degrading over time and usage. Not addressing this issue by real-time droplet tracking techniques will significantly increase the amount of experimental failures, especially on long term protocols that involve many droplets movements.

Availability:

As was mentioned, there is no simple way to order or manufacture a batch of identical chips in a reasonable amount of time or costs.

Automation and programming:

A typical DMF chip has ~100 electrode. Those electrodes must be activated and deactivate in a precipice order and timing in order to produce the desired droplet movement. This means that the automation system needs to be programmed with different electrode sequences for different chip designs. Moreover, every biological experiment has its unique set of droplet operations, timing and external activities (e.g. temperature change, magnet activation, LED activation).

This requires programming skills from the scientist running the experiment.

The development of a new and advanced DMF setup:

To address most of the DMF limitations and challenges that were presented, a new set of electronic hardware and software were developed. Key features of the new system are as follow:

1. Reliable.
2. Modular.
3. Generic and open for changes.
4. Simple to understand, maintain, debug and use.
5. Expend to support external peripherals and equipment.

The current system can support the following fully automated biological experiment as shown by the users:

1. Basic droplet movements such as (Merge, mix, split and move).
2. Basic movements of different sizes of droplets (as presented in the chipArray).
3. Hybrid DMF and micro fluidic device.
4. Heat shock transformation.
5. Automatic evaporation monitoring and compensation (James Perry from Shih lab is currently working on this feature).
6. Purifying a protein by using magnetic actuation.

Basic hardware setup:

The solution is based on the ArduShield board and ArduBridge software that serves as protocol bridge between the host PC and the switching control-boards (**Figure B.3)**. The communication between the PC and the Arduino is done over standard USB-RS232 serial device. The ArduBridge firmware is decoding the requests and sends them over the I2C bus to the MAX3700 which in response activates the specific high-voltage switch.

The entire biological-protocol and the electrode sequencing are done solely on the PC side.

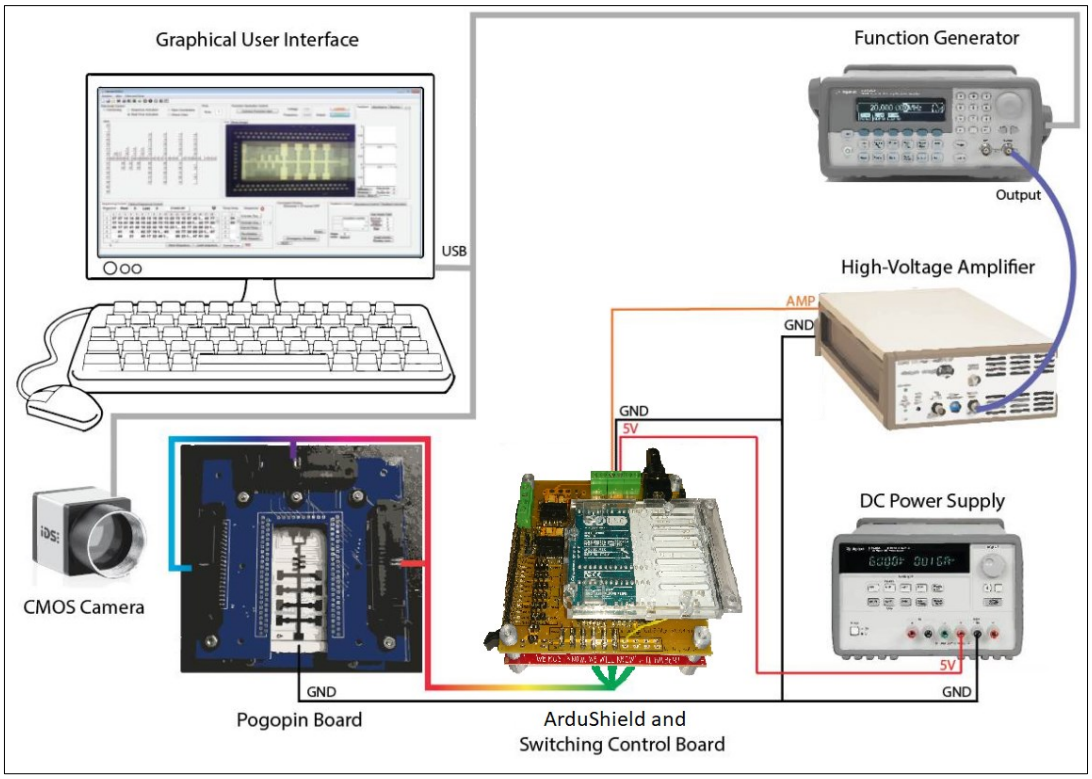

Figure B.3: A schematic description of a typical digital microfluidic setup for moving droplets

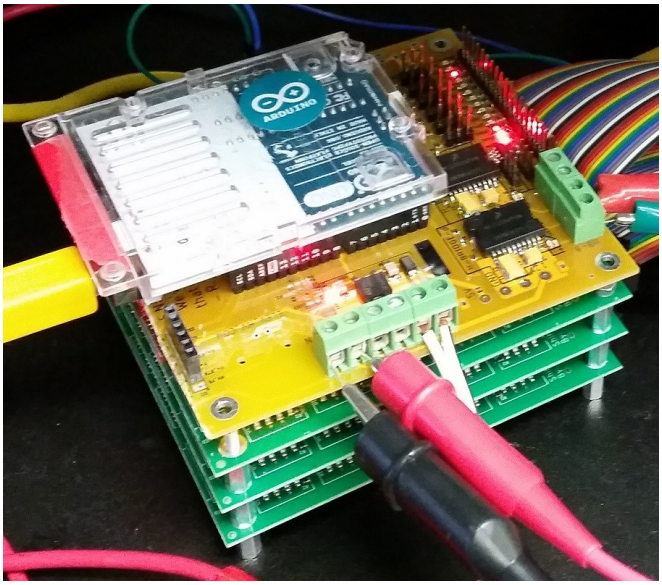

Figure B.4: Stack up arrangement of ArduShield and three electrode-driver boards

ArduShield board:

I designed The ArduShield board to meet the following demands:

1. Simplify the connections between the Arduino to any external component.
2. Simplify the connections to the switch-Board-Stack by connecting on top and using the same connection type and pin-out.
3. Includes two 16v/1A FET transistors to activate external devices such as fan, LED, heater or TEC.
4. Includes two 16v/2A H-Bridge drivers that are connected to the Arduino PWM channels. This can be used to control the intensity of a TEC, DC-motor or any other load.
5. Capable of measuring droplet impedance on every electrode.

SwitchBoard:

I improved the existing lab's SwitchBoard for the lab (**Figure B.5**). It is based on an old design (by a former student). Its main functionality is to activate the high voltage electrodes on the DMF device. Its specifications are as follow:

1. Each board can address 40 electrodes.
2. A stack-up of up to 4 boards can be build (160 electrodes)
3. Communication with the stack is done via I2C channel (MAX7300)
4. The high voltage switches are rated to 600V/40mA (420Vrms)
5. Input voltage 5v to 12v with reverse polarity protection
6. High-voltage switch bandwidth of DC to 1Mhz

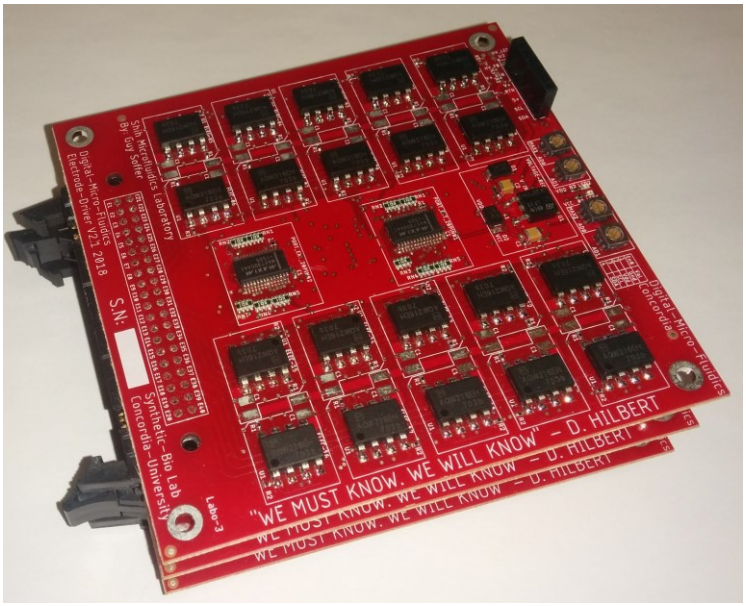

Figure B.5: A stack of three high-voltage electrode switch boards

Software:

The solution is made from two codes that each run on different computers (**Figure B.6**).

1. The GSOF_ArduBridge and additional libraries, written in Python and running on the PC side. It is fully cross platform over Windows, Linux and Mac operating systems. It can also run on Raspberry-Pie machines.
2. The GSOF_ArduCtrl_FW, written in C and running on the Arduino-Uno MCU (Atmel 8bit).

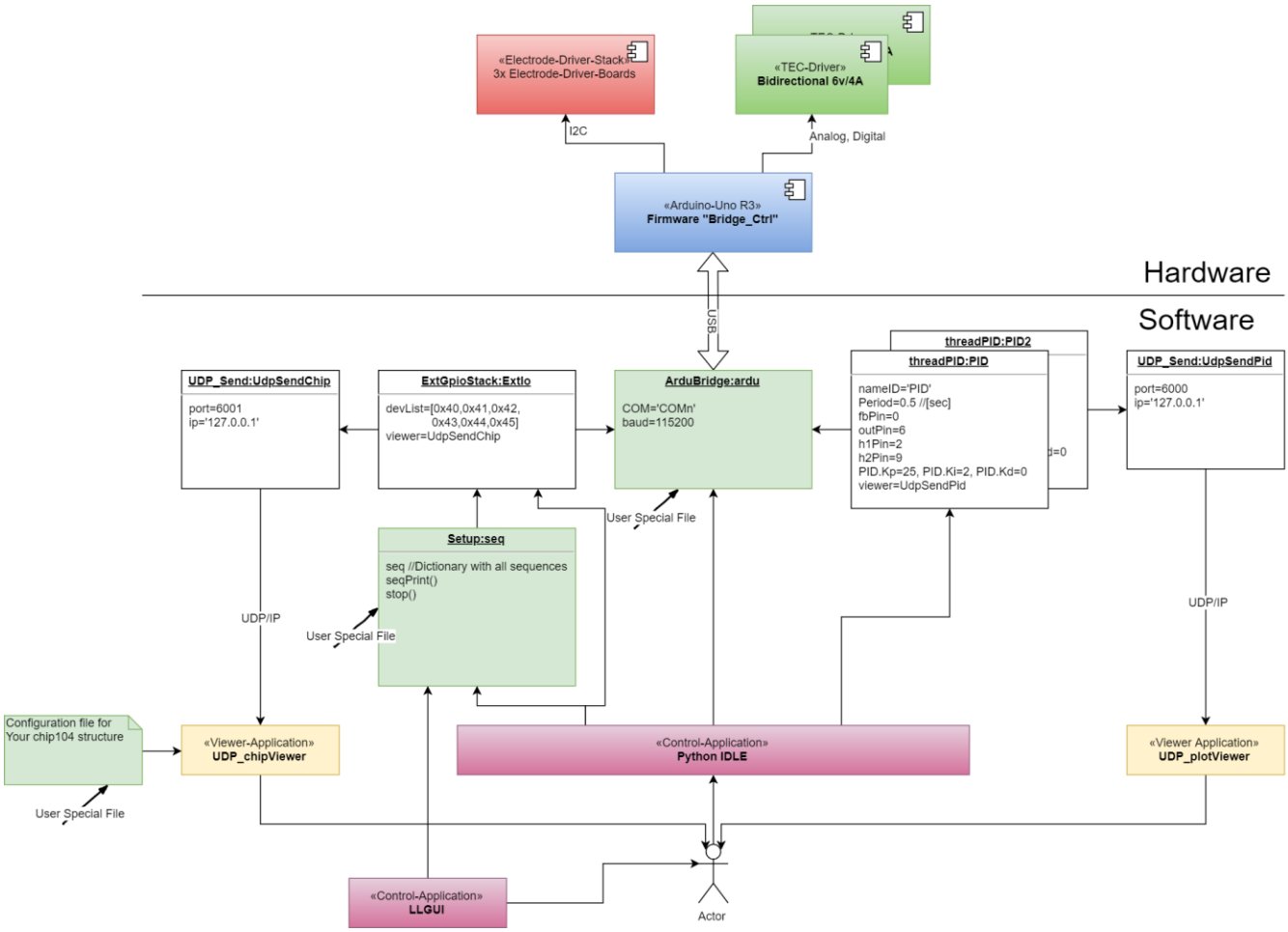

Figure B.6: GSOF_ArduBridge application architecture

written in Python2.7 using the modules pySerial and GSOF_ArduBridge.