DEVELOPMENT OF NOVEL ELP-BASED TRANSCRIPTIONAL REGULATORS FOR IMPROVED BIOMANUFACTURING

by

Logan Rae Readnour

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science in Agricultural and Biological Engineering
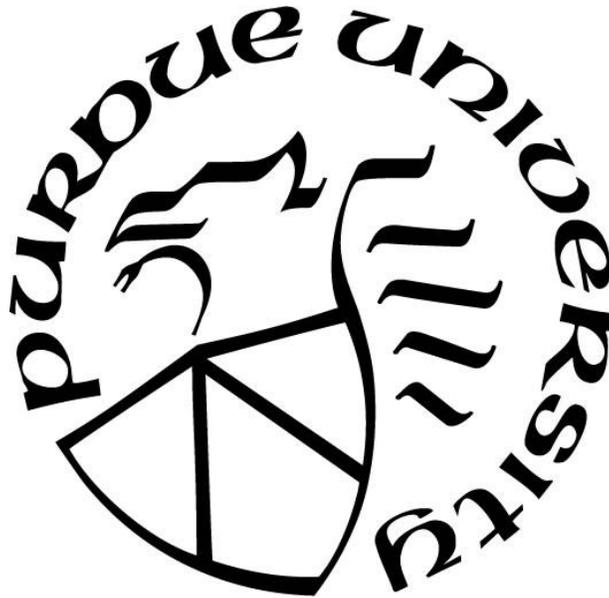
School of Agricultural and Biological Engineering

West Lafayette, Indiana

December 2018

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Kevin Solomon, Chair
>   Department of Agricultural and Biological Engineering

Dr. Kari Clase
>   Department of Agricultural and Biological Engineering

Dr. Julie Liu
>   Department of Chemical Engineering

Approved by:
>   Dr. Bernard Engel
>>   Head of the Graduate Program

This thesis is dedicated to my parents, Kimberly and Darin Readnour, for their constant love, support, and encouragement.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATc	Anhydrotetracycline
ATPS	Aqueous two-phase system
CIAP	Calf Intestinal Alkaline Phosphatase
d_2H_2O	Distilled deionized water
ds	Double stranded
DTT	Dithiothreitol
ELP	Elastin-like polypeptide
ELP-SF	Elastin-like polypeptide sigma factor fusion
FAEE	Fatty acid ethyl ester
FPP	Farnesyl pyrophosphate
GFP	Green fluorescent protein
GGPP	Geranylgeranyl diphosphate
IPTG	Isopropyl β-D-1-thiogalactopryanoside
ITC	Inverse transition cycling
LB	Luria-Bertani broth
MCFs	Microbial chemical factories
MW	Molecular weight
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PRe-RDL	Recursive directional ligation by plasmid reconstruction
RBS	Ribosome binding site
RFP	Red fluorescent protein
RNAP	RNA polymerase
RT	Room temperature
SD	Shine-Dalgarno
SF	Sigma factor

TB   Terrific Broth

T_t Transition temperature

WCL   Whole cell lysate

ABSTRACT

Author: Readnour, Logan, R. MSABE
Institution: Purdue University
Degree Received: December 2018
Title: Development of Novel ELP-Based Transcriptional Regulators for Improved
      Biomanufacturing
Major Professor: Dr. Kevin Solomon

Microbial chemical factories (MCFs) have become an attractive platform for producing valuable drugs, chemicals, and biofuels due to increasing environmental concerns, energy demands, and the difficulties associated with chemically synthesizing complex molecules. However, the potential of using microbes to produce many valuable products has not been fully realized due to low productivity and yields. Production may be enhanced through the dynamic control of metabolic pathways to alleviate metabolic imbalances and toxic product build-up, but very few tools are available to broadly implement this paradigm. I aim to expand this toolbox by developing tunable sensor-regulator devices that act as feedback controllers for toxic intermediate formation by responding to cues of cell health for improved production. Elastin-like polypeptide (ELP) will act as the sensing domain of the controller to indirectly sense toxic metabolite accumulation through changes in intracellular pH. ELPs make ideal sensors since they exhibit a sharp, inverse phase transition to indicators of cellular health such as pH and ionic strength, and external stimuli such as temperature. In this research, a library of ELPs that exhibit pH sensitivity and transition under various conditions was made and purified using a new organic solvent extraction method. It is hypothesized that fusion of ELP to orthogonal transcription factor will allow for the controlled expression of target genes in response to stimuli without disrupting native processes. As proof of concept, an ELP fusion to orthogonal sigma factor was designed to drive the expression of a fluorescent reporter protein. Initial designs successfully alter gene expression by 21% in response to temperature. To improve this response, an alternative feed-forward loop architecture was modelled, which predicted an improved response of 35% and increased ultrasensitivity. Refinement of this design and combinatorial construct libraries will generate various regulators with diverse outputs that may be integrated in bioproduction pathways for improved performance.

1. INTRODUCTION

1.1 Motivation

Microbes have the extraordinary biosynthetic potential to sustainably make the compounds that drive today's society, such as medicines, biofuels, and chemicals. However, microbial systems are not yet efficient enough to produce these compounds at economically feasible levels, especially for commodity chemicals and biofuels where margins are much lower compared to specialty chemicals or pharmaceuticals. Often, yields and titers can be limited due to metabolic imbalances and by the build-up of toxic pathway intermediates, which adversely affect cell health and productivity [1, 2]. These limitations can be addressed through the dynamic control of metabolic pathways in order to optimize sustainable biomanufacturing in microbes. While this has proven to be a promising strategy, there are very few tools available to broadly implement this paradigm. The main objective of this research is to expand this toolbox by developing tunable sensor-regulator devices that act as feedback controllers for toxic intermediate formation and can be universally applied to different pathways for improved cell health and production.

Since the emergence of the field, traditional metabolic engineering strategies have been employed to engineer microbial chemical factories (MCFs) for the production of valuable products. Early production was limited to what could be found in nature; by mutating microbes isolated from various environments, strains could be enhanced for the large scale production of amino acids, antibiotics, fatty acids, organic acids, and alcohols [3]. When genetic engineering arrived, and with it new tools for manipulating genes and pathways, it became possible to express heterologous products—those found in nature, but not naturally produced by the host. This opened the door to producing more complex products from MCFs that would otherwise be difficult to achieve through chemical synthesis. Some of these products, including the antimalarial drug artemisinin [4] and the chemicals 1,4-butanediol [5] and 1,3 propanediol [6], are now produced from MCFs at the commercial scale, yet numerous metabolically engineered compounds from scientific literature remain only at the μg/L or mg/L scale [7, 8]. Poor production titers, rates, and yields often prevent the affordable manufacturing of these products

from microbial platforms [1, 9, 10]. Even more problematic, is the commercialization of compounds for high-volume, low-value markets, like biofuels and commodity chemicals, where economic considerations are even more substantial. Despite this obstacle, interest in using MCFs to produce commodity chemicals and biofuels continues to grow due to its potential for reducing global CO_2 emissions, lessening society's dependence of fossil fuels, and improving energy security [7]. As we combat these challenges, we will need new tools and processes to better improve the efficiency of MCFs for heterologous pathway expression.

There are several challenges that metabolic engineers face when designing a system for optimal production. The expression of heterologous pathways can lead to metabolic imbalances where expression can be either too low, resulting in bottlenecks, or too high, placing a metabolic burden on the cell that results in an inefficient use of cellular resources. [1, 11, 12]. Overexpression of pathways can also lead to the buildup of toxic intermediates, enzymes, or products that inhibit growth and reduce productivity [2, 13]. Although conventional metabolic control methods to overexpress rate-limiting steps, knock out competing pathways, and drive flux towards product have been shown to improve titers and productivity, these methods are limited by their inability to modulate pathway flux in response to changing cellular conditions [14].

The arrival of dynamic control strategies for heterologous pathway expression offers a promising path forward as we search for new ways to improve yields and productivity. These strategies offer an advantage over conventional metabolic control methods, which generate static outputs that do not respond to transient cellular or environmental conditions. Such static forms of gene regulation make use of promoter libraries [15-17], mRNA stability [18], and ribosome binding sites (RBS) [19] to optimize expression for a particular condition or growth phase. The problem remains, however, that growth and bioreactor conditions tend to change, resulting in suboptimal product formation. Dynamic control, on the other hand, enables the organism to adapt its metabolic flux to perturbations in its environment or growth to sustain optimal production at all stages. This is accomplished through the implementation of dynamic sensor-regulator systems that can sense fluctuations in growth conditions and metabolite concentrations to regulate pathway expression as needed for high productivity [1, 20].

There are several cases where dynamic sensor-regulator systems have been implemented to regulate flux through a heterologous pathway for enhanced productivity. The earliest implementation of this strategy involved a regulatory circuit that used the intracellular metabolite acetyl phosphate to indirectly sense excess glucolytic flux in order to regulate lycopene biosynthesis. This control loop was able to alleviate the adverse effects of metabolic imbalance to significantly improve lycopene production [21]. Other studies have used direct sensors of metabolite formation. In one case, a naturally existing transcriptional regulator was rewired to dynamically control the supply and consumption of malonyl-CoA to significantly improved fatty acid biosynthesis [22]. In another, the production of the biodiesel fatty acid ethyl ester (FAEE) was enhanced when a naturally occurring transcription factor was used to regulate the expression of downstream pathway genes through sensing the buildup of fatty acids, a key pathway intermediate. By requiring a minimum fatty acid concentration to be met before product formation could begin, metabolic burden was balanced and a threefold increase in FAEE titer was observed [11].

While examples such as these show the potential of using dynamic control for improved production, these strategies are limited by the lack of available known sensors to detect toxic metabolites levels. Current approaches primarily rely on large screens to identify native sensors that are only specific to individual products. This was the case of Dahl et al. [23], where whole-genome transcript arrays were used to identify stress-responsive promoters to sense the accumulation of toxic intermediates. The use of native regulatory elements for targeted gene regulation, however, can also interact with native, nonproduction pathways in unforeseen and undesirable ways. Therefore, there is a growing need for dynamic control devices that can (1) be readily applied to an arbitrary pathway (universal), (2) not interfere with the native processes of the cell (orthogonal), and (3) be easily tuned to sense toxic intermediate formation (programmable). These properties encompass the hallmarks of an ideal controller of gene expression.

In light of this challenge, I propose novel elastin-like polypeptide (ELP) based transcriptional regulator devices that work as dynamic feedback controllers for toxic product accumulation. It has been suggested that methods of pathway toxicity are sometimes linked to the acid stress

response [23]. Therefore, it is predicted that changes in intracellular pH, caused by toxic product build up, will trigger the regulators to turn off product formation until the cell can recover and pH returns to normal. ELP will act as the sensing domain of the controller to indirectly sense toxic metabolite accumulation through changes in intracellular pH. This aspect makes the controllers universal, capable of being applied to any pathway where toxic metabolite formation causes a shift in intracellular pH. There is also potential to implement the controller in other hosts besides bacteria, such as yeasts, fungi, and plants. The fusion of ELP with orthogonal transcription factors will allow for the targeted control of genes for toxic intermediates while avoiding the disruption of native cellular processes. This system takes advantage of the reversible, inverse phase transition property of ELP, which allows it to aggregate or become soluble in response to certain stimuli, such as changes in pH, temperature, or ionic strength. The availability of transcription factor to control gene expression will then depend upon the aggregated or soluble state of the ELP. In this way, these devices recognize cues of cellular health to autoregulate the expression of key genes within a bioproduction pathway to enhance yield and productivity.

1.2   ELP-transcription factor fusions as tunable regulators of gene expression

Elastin-like polypeptides are artificial polymers that are derived from the hydrophobic domain of human tropoelastin, the precursor of elastin [24, 25]. ELPs consist of pentapeptide repeats in the form of **[VPGXG]$_n$,** where X represents a guest residue that can be any amino acid with the exception of proline, and n indicates the number of pentamer repeats. ELPs exhibit a reversible, inverse phase transition that can be triggered by a number of stimuli, such as changes in temperature, pH, or ionic strength [26-28]. Below a critical temperature, known as the transition temperature (T_t), ELP is structurally disordered and soluble in aqueous solvents. However, as temperature is increased, ELP aggregates to adopt a more ordered, compact globular structure. The basis for this mechanism has been attributed to hydrophobic collapse of the protein and the expulsion of water molecules from the polymer's nonpolar side-chains [29, 30]. The phase transition occurs over a narrow temperature range (2-3ºC) [26-28], making ELPs suitable as environmental sensors, while the ability of ELP to retain its activity upon fusion to other proteins [31], indicates the feasibility of ELP-transcription factor controllers.

A defining characteristic of controllers is programmability—the ability to produce a specific response to a defined set of conditions. By making changes in either the polymer composition or chain length, the phase transition of ELP can be fine-tuned to respond to specific stimuli. This property makes ELP ideally suited as a controller as rational design can be applied to produce ELPs with predictable transition temperatures [32, 33]. Certain trends have been observed; the T_t can be lowered by increasing overall polymer hydrophobicity through insertions of hydrophobic amino acids at the guest residue, or by increasing the number of pentamer repeats [34]. The phase transition can also be made pH responsive by introducing amino acids at the guest residue that are susceptible to ionization [28, 35, 36]. Therefore, it is possible to program pH sensitivity into ELPs via sequence design.

To make pH sensitive ELPs, I designed an "acidic" ELPA where X = V/I/E [1:3:1] and a "basic" ELPB where X = V/H/G/A [1:2:1:1] [35]. When pH is greater than pK_a, the acidic ELP is charged, resulting in a higher transition temperature. Decreasing the pH below pK_a, however, causes protonation of the glutamic acid residues, thereby making the acidic ELP neutral and significantly reducing the transition temperature. The basic ELP exhibits an opposing pH-dependent phase transition behavior where the charged histidine residues are neutralized once pH is raised above the pK_a. Therefore, a decrease in pH will cause the T_t of the acidic ELP to decrease but will increase the T_t of the basic ELP. In addition to the acidic and basic ELPs, I also designed a non-pH sensitive "neutral" ELP where X = V. The three ELP types will be fused to transcription factors to determine their ability to control gene expression in response to changes in pH and temperature.

1.2.1   ELP-transcription factor controller architectures

To determine whether fusions of ELP with transcription factors can be used to regulate gene expression, three different network architectures were designed. The first two architectures use simple activation and repression networks that consists only of two genes—a reporter gene and a gene encoding the ELP fusion protein. The third model uses a feed-forward loop (FFL) architecture, which has a three-gene pattern, to enhance the dynamic range and ultrasensitivity of the controller.

1.2.1.1 Simple activation architecture

The initial controller design features the regulation of a target gene by simple activation using orthogonal sigma factor (SF). As promoter recognition subunits of RNA polymerase (RNAP), sigma factors play a role in initiating gene transcription by recognizing and binding to the -10 and -35 regions in their target promoters. Using sigma factors that are orthogonal to the host's native regulation allows for the construction of genetic circuits that do not interfere with these processes, thereby avoiding undesirable interactions [37, 38]. By making fusions of ELP with sigma factor (ELP-SF), the phase behavior of ELP can be leveraged to control the availability of SF for gene activation. It is predicted that when ELP is soluble below its T_t, SF will be free to initiate gene expression. Once ELP aggregates in response to stimuli, however, it will prevent SF from interacting at the promoter, thereby preventing expression. The soluble and aggregated forms of ELP fusions are termed active and inactive respectively. In the simple activation architecture, ELP-SF is used to drive the expression of green fluorescent protein (GFP) (Figure 1). It is hypothesized that a change in stimuli to induce the inactive form of ELP-SF will cause a decrease in GFP expression, which can be measured as a decrease in fluorescence.

Targeted Gene Expression

Stimuli Induced Gene Inhibition

ΔT, ΔpH, ΔIonic Strength
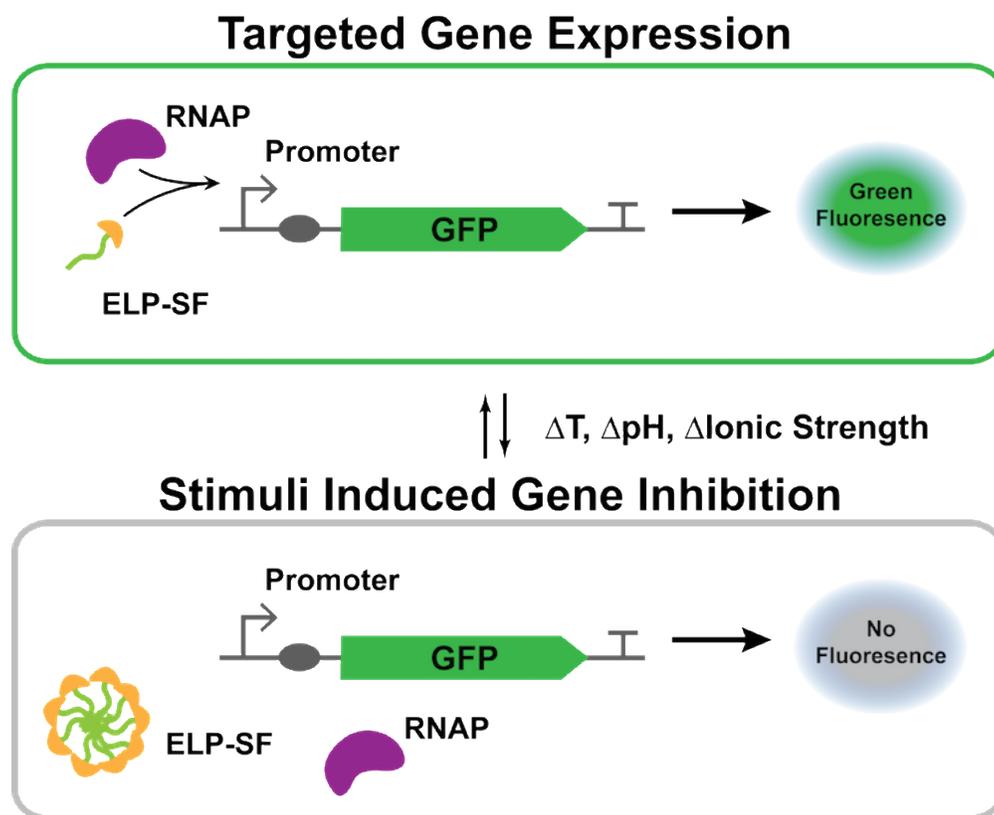
Figure 1. Simple activation architecture of the ELP-SF system. Environmental conditions induce the active or inactive state of ELP-SF, which in turn controls GFP expression.

1.2.1.2 Simple repression architecture

In addition to the simple activation architecture, a second system was explored in which Tet repressor (TetR) protein is used to control the expression of a target gene. This simple repression model uses TetR fused with ELP (TetR-ELP) to control the repression of red fluorescent protein (RFP), which is cloned under a constitutive Tet repressible promoter (Figure 2). The TetR-ELP fusion is constitutively expressed; it is expected that when TetR-ELP is active, below its T_t, TetR will be free to repress RFP production. In its inactive form, above its T_t, ELP will aggregate together to prevent the interaction of TetR with the promoter, allowing RFP expression to occur. It is also anticipated that the addition of anhydrotetracycline (ATc), which inhibits TetR by binding to it, will also lead to RFP expression.

Below Tt

Above Tt

Below Tt with aTc
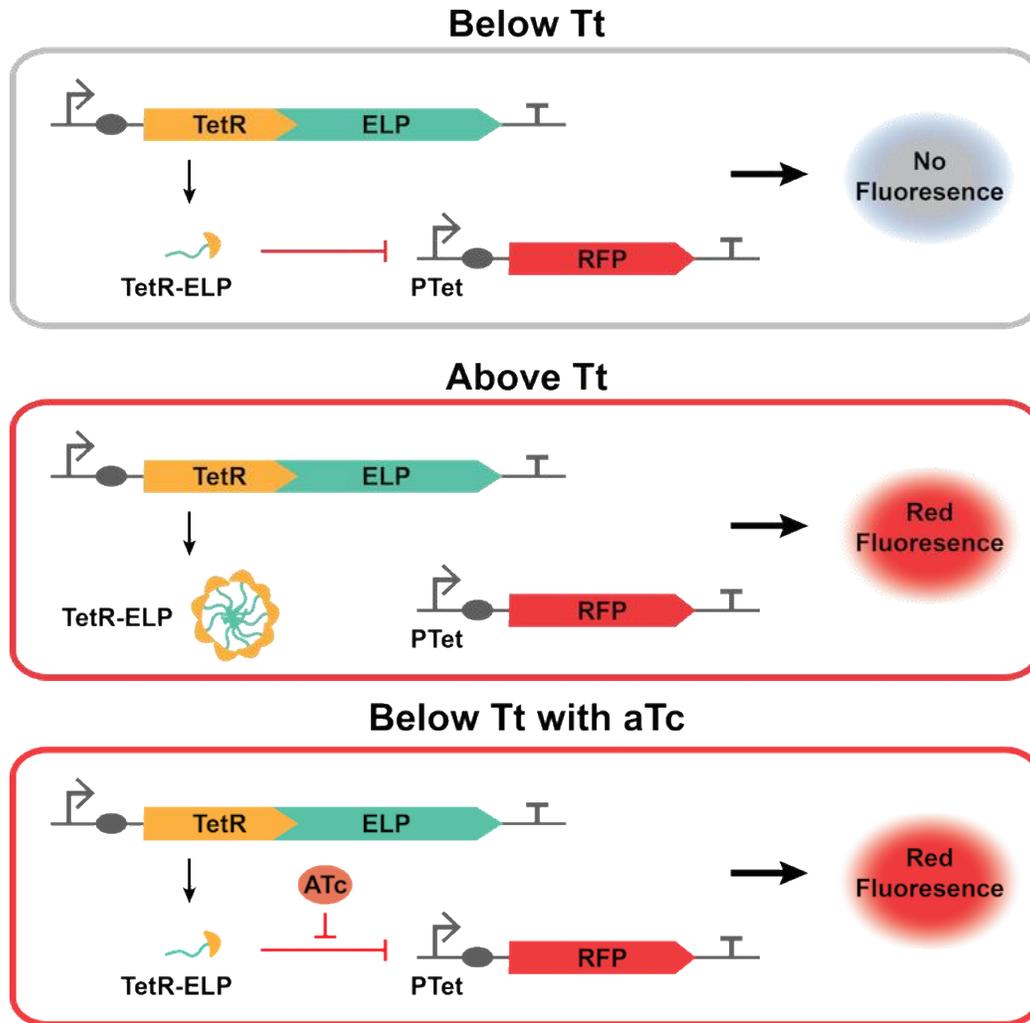
Figure 2. Simple repression architecture of TetR-ELP fusion. Active TetR-ELP repress RFP production (top), while inactive TetR-ELP prevents repression (middle). The addition of ATc to active TetR-ELP prevents repression of RFP (bottom).

1.2.1.3   Feed forward loop architecture

In an attempt to expand the dynamic range and improve the sensitivity of the ELP-SF controller, an alternative network architecture was explored. This alternate model best resembles a type 1 coherent feed-forward loop (FFL) network motif. A typical FFL has a three-gene pattern where two transcription factors, one of which regulates the other, are both required to jointly regulate the expression of a target gene. For a type 1 coherent FFL, the regulatory effect is positive, or activating, for each interaction, and can be represented as a logical AND gate. In this case, regulator X activates regulator Y, and both are needed to activate the expression of gene Z. This

type of genetic circuit exhibits sign-sensitive delay, in which a delay is experienced in one direction (after stimulation), but not in the other (when stimulation stops). The delay is a result of the AND input function for gene Z. First, a stimulus activates X production, which accumulates and binds to the Y and Z promoters, thereby activating Y production. Production of Z only begins once the concentration of Y surpasses the activation threshold necessary for the Z promoter, resulting in a delay following the initial signal. The higher the threshold, the longer the delay. No delay is experienced, however, once the stimulus stops, since the loss of either the X or Y inputs cause Z production to cease. This sign-sensitive delay behavior allows the FFL to protect the target gene from transient ON signals, only permitting persistent stimuli to lead to a response [39-41].

To expand upon the simple activation design, an FFL model was designed in which two inputs, active ELP-SF and orthogonal ribosome (o-ribosome), are both required to initiate GFP expression. Orthogonal ribosomes recognize altered Shine-Dalgarno (SD) sequences to translate genes that are not recognized by the host's native ribosomes, thereby partially decoupling gene translation from the cell's native translational machinery [42]. Several functional o-ribosome-mRNA pairs have already been engineered in *E. coli* [43-45]. In our FFL design, the promoters for o-ribosome and GFP are both regulated by the same orthogonal sigma factor, which is fused to ELP. Therefore, ELP-SF must be in its soluble, active form to allow the sigma factor to initiate expression. The RBS upstream of the GFP gene is an altered SD sequence that will only be recognized by its corresponding o-ribosome. Therefore, production of GFP will only begin once sufficient levels of o-ribosome have accumulated to carry out translation. Thus, the connectivity of the system is similar to a type 1 coherent FFL. Figure 3 (A) depicts the system in its simplest form, a logical AND gate, while (B) illustrates the details of the system.
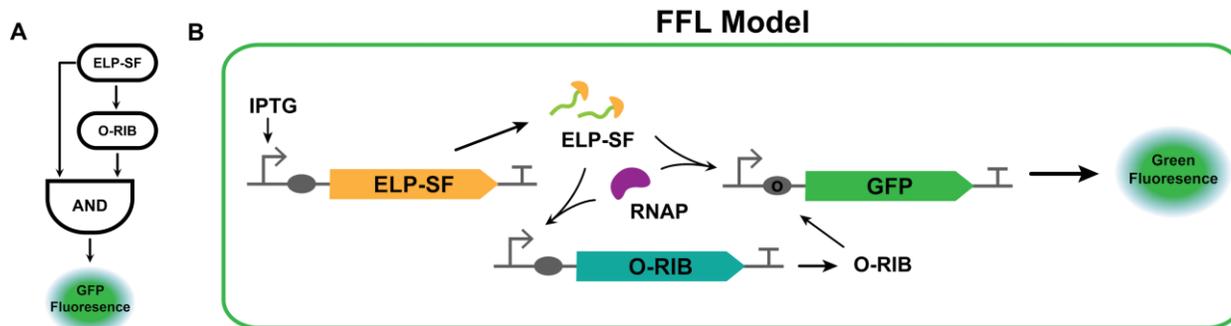
Figure 3. Feed forward loop architecture of the ELP-SF system. (A) Logical AND gate representation of (B) the full model where GFP expression requires both active ELP-SF and o-ribosome (O-RIB) as inputs.

A type 1 coherent FFL network motif was chosen as the basis of the FFL model design since it exhibits sign-sensitive delay. It is believed that this behavior would be beneficial as it would allow the control system to quickly respond to changes in stimuli caused by toxic intermediate formation. For example, a change in intracellular pH upon toxic intermediate formation will trigger the phase transition of a pH sensitive ELP, causing it to aggregate and quickly turn off expression of the target gene, which in this case is the toxic intermediate. This OFF step experiences no delay, which is critical since the cell will need to quickly recover from its unhealthy state in order to survive. The return to production, however, does experience delay since it first requires the obtainment of a minimum threshold for both active ELP-SF and o-ribosome. The need for persistent stimuli to turn on gene production will ensure that production only resumes once the cell has returned to a healthy state.

1.3   Project objectives

Three primary objectives for this project are as follows:

1. Design, build, and characterize a library of ELP constructs that exhibit pH sensitivity and phase transition under various conditions.
2. Determine the feasibility of ELP-transcription factor fusions as dynamic sensor-regulator devices for gene expression through model simulations and experimental results.
3. Explore alternative network architectures for enhanced controller performance and dynamic range.

A primary goal of this project was to develop various ELPs with different phase transition properties so that the feasibility of ELP-transcription factor controllers could be investigated. Therefore, the first objective of this research was to design, build, and characterize a library of ELP constructs that exhibit pH sensitivity and transition within a physiologically relevant range. Three ELP types at three different lengths were successfully cloned, establishing a library of nine constructs in total. Out of the nine, four ELPs were successfully purified and the transition behavior of three were characterized *in vitro*.

To determine the feasibility of ELP-transcription factor fusions as dynamic sensor-regulator devices for gene expression (objective 2), a simple activation model was first simulated using MATLAB, which produced predictions that were in line with preliminary studies. In addition, a simple repression model was designed and tested experimentally but failed to show dynamic control. Preliminary studies of the simple activation model showed a modest change in gene expression (21%) in response to temperature, which is likely too modest to be of use in a real production pathway. Therefore, an alternative network architecture was explored to improve the dynamic range of the controller (objective 3). To achieve this, a feed-forward loop model was designed, which is predicted to improve the response to 35% and increase sensitivity.

The long-term goal of this project is to develop dynamic control of a bioproduction pathway by implementing pH-sensitive ELP-transcription factor regulators to control the expression of toxic intermediates, enzymes, or products in order to improve cell productivity and yield. It is expected that certain toxic products will adversely affect cell health through shifting intracellular pH away from its normal operating range of 7.6-7.8. The controller will be able to detect this shift and respond by turning off production of the toxic product until pH returns to normal. More details for implementing the controllers into a bioproduction pathway can be found in the final chapter "Conclusions and Future Directions."

2. METHODS

2.1 Bacterial strains, general cloning, and culture media

All plasmid cloning was performed in the *E. coli* DH5α strain (New England Biolabs), while either the *E. coli* BL21(DE3) expression strain (Coli Genetic Stock Center) or *E. coli* MG1655 strain (Coli Genetic Stock Center) were used for protein expression. Standard cloning procedures were performed using DNA miniprep, gel purification, and PCR purification kits (Omega Bio-tek). All restriction enzymes used were FastDigest enzymes (Thermo Scientific) except for BseRI (New England Biolabs). All synthesized primers and oligonucleotides (Sigma-Aldrich) are listed in Table 4, appendix A. Cultures were grown in Luria-Bertani (LB) broth (Fisher BioReagents) for general cloning, while Terrific Broth (TB) (4 mL/L glycerol, 12 g/L tryptone, 24 g/L yeast extract, 2.31 g/L KH_2PO_4, and 12.54 g/L K_2PO_4) was used for ELP expression studies.

2.2 ELP library design and construction

In order to achieve ELPs capable of transitioning within a physiological relevant range, an ELP library was designed by altering the guest residue and number of pentamer repeats. The ELP library consists of nine constructs in total and contains three different ELP types at three different sizes (Table 1). The three ELP types include the following: a "neutral" ELPN where guest reside X = V, an "acidic" ELPA where X = V/I/E [1:3:1], and a "basic" ELPB where X = V/H/G/A [1:2:1:1]. The ratios chosen were based on another study by Mackay et al. [35] that achieved pH sensitive ELPs that transition within a physiological temperature range. ELPs were elongated through successive rounds of recursive directional ligation by plasmid reconstruction (PRe-RDL) [46] to achieve lengths of n = 40, 80, and 160. The expected molecular weight of each ELP protein was calculated using ExPASy ProtParam (Table 1).

Table 1. Characteristics of the designed ELP library including the guest residue ratios, number of pentamer repeats, amino acid sequence, and the expected molecular weight of each ELP.

ELP Type	Guest Residue (X)	Pentamer Repeats (GXGVP)	Amino Acid Sequence	Expected MW (kDa)
ELPN	**V**	40	M(G**V**GVPG**V**GVPG**V**GVPG**V**GVPG**V**GVP)$_8$GY	16.62
		80	M(G**V**GVPG**V**GVPG**V**GVPG**V**GVPG**V**GVP)$_{16}$GY	32.99
		160	M(G**V**GVPG**V**GVPG**V**GVPG**V**GVPG**V**GVP)$_{32}$GY	65.76
ELPA	**V/I/E** [1:3:1]	40	M(G**V**GVPG**I**GVPG**I**GVPG**E**GVPG**I**GVP)$_8$GY	17.19
		80	M(G**V**GVPG**I**GVPG**I**GVPG**E**GVPG**I**GVP)$_{16}$GY	34.15
		160	M(G**V**GVPG**I**GVPG**I**GVPG**E**GVPG**I**GVP)$_{32}$GY	68.06
ELPB	**V/H/G/A** [1:2:1:1]	40	M(G**V**GVPG**H**GVPG**G**GVPG**H**GVPG**A**GVP)$_8$GY	16.66
		80	M(G**V**GVPG**H**GVPG**G**GVPG**H**GVPG**A**GVP)$_{16}$GY	33.09
		160	M(G**V**GVPG**H**GVPG**G**GVPG**H**GVPG**A**GVP)$_{32}$GY	65.94

2.2.1   Modification of pET28a-RDG for PRe-RDL

A pET28a expression vector encoding the gene for protein aECM1 [47] was modified for compatibility with the ELP genes and the PRe-RDL method of elongation. The vector, relabeled as pET28a-RGD, was donated by Dr. Julie Liu at Purdue University and contains an RGD cell-binding domain along with an elastin-like domain (X = K/I, n = 60) interspersed with two small linker regions. Compatibility was accomplished by excising the aECM1 gene and inserting a linker sequence downstream of a T7 promoter that encodes the Shine-Dalgarno ribosomal binding site [5′-AGGAGG-3′] and sites for the type II restriction enzymes BseRI and AcuI. The linker sequence, which was also designed to have 3′ overhangs complimentary to XbaI and XhoI sites, was made by annealing together oligonucleotides 1 and 2 (Table 4, appendix A). Oligonucleotides were annealed by first heating 50 μL at 5 μM concentration of each strand in annealing buffer (100mM NaP, 150 mM NaCl, and 1 mM EDTA) at 95°C for 5 min and then allowing to cool gradually to room temperature (RT) over an hour. The vector pET28a-RGD was then digested with XbaI and XhoI, purified from a gel, and ligated with the annealed linker sequence for 15 min at RT using a Rapid DNA Ligation Kit (Thermo Scientific). The ligation product was transformed into DH5α, allowed to recover for 1 h at 37°C in LB, and then plated on LB supplemented with 50 mg/mL of kanamycin. The resulting colonies were screened by diagnostic digestion and positive clones were confirmed by DNA sequencing.

2.2.2   Insertion of ELP genes into modified pET28a

Following insertion of the linker sequence, genes encoding each ELP type were inserted into the modified pET28a vector. The genes were made by annealing together synthesized ELP oligomer pairs. To make the ELPA gene, the oligomers 5 and 6 (Table 4, appendix A) were annealed as previously described. This resulted in a double stranded (ds) DNA sequence with 2 bp, nonpalindromic, 3′ overhangs. Likewise, the ELPN gene was made by annealing oligomers 3 and 4 (Table 4, appendix A), and the ELPB gene was made by annealing oligomers 7 and 8 (Table 4, appendix A). The 5′ ends of each dsDNA pair were phosphorylated by incubating 20 µL of 20 pmol of dsDNA, 10 U T4 polynucleotide kinase, and 10 mM ATP for 20 min at 37°C and then purified.

The modified pET28a vector was digested with BseRI followed by dephosphorylation of the 5′ ends with 1 U of Calf Intestinal Alkaline Phosphatase (CIAP) (Invitrogen) during the last 30 min of digestion at 37°C. The linearized vector was then purified from a gel followed by ligation with the dsDNA ELP sequences, which resulted in products referred to as pET28-ELP. These products were transformed into DH5α, allowed to recover for 1 h at 37°C in LB, and then plated on LB supplemented with 50 mg/mL of kanamycin. The resulting colonies were screened by diagnostic digestion and positive clones were confirmed by DNA sequencing.

2.2.3   PRe-RDL

ELPs were elongated through successive rounds of PRe-RDL, which was carried out in the strain DH5α. A single round of PRe-RDL involves (1) digesting the parent vector with AcuI and BglI to isolate the ELP-containing DNA fragment, (2) digesting the parent vector with BseRI and BglI to isolate the ELP-containing DNA fragment, and (3) ligating the two purified halves together to reconstruct a functional plasmid, dimerizing the ELP sequences and doubling its length (Figure 4). Multiple rounds of PRe-RDL were performed to achieve lengths of 40, 80, and 160 pentamer repeats for each of the three ELP types. For reference, genes are labeled as ELP type followed by the number of pentamer repeats. For example, the acidic, neutral, and basic ELPs at n = 40 repeats are referred to as ELPA40, ELPN40, and ELPB40 respectively. All

constructs were screened by diagnostic digestion, and positive clones were confirmed by DNA sequencing.
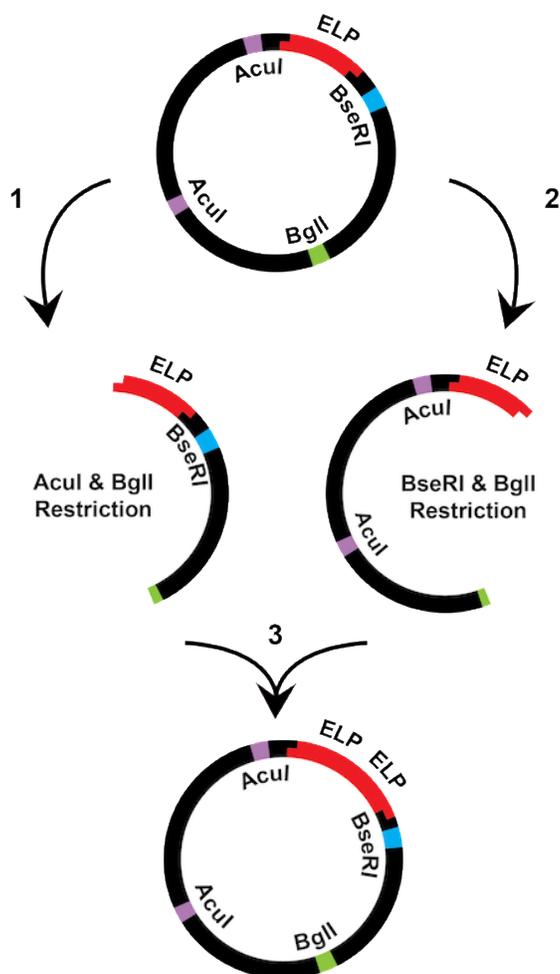
Figure 4. A single round of the PRe-RDL process. (1) The parent vector is digested with AcuI and BglI and the resulting ELP-containing DNA fragment is purified. (2) The parent vector is digested with BseRI and BglI and the resulting ELP-containing DNA fragment is purified. (3) The two purified halves are ligated together to reconstruct a functional plasmid, dimerizing the ELP sequences and doubling its length.

2.2.4   Transfer of the ELP library to pET32a

Each ELP gene (sizes n = 40, 80, and 160) in the pET28a vector was transferred to a pET32a expression vector gifted by Dr. Xi Ge from the University of California, Riverside. The vector, labeled as pET32a-GST-ELP180 consists of a GST tag fused to the C-terminus of an ELP (X = V, n = 180) all under a T7 promoter. To remove the ELP library genes, the pET28a-ELP vectors were digested with the restriction enzymes XhoI and NdeI followed by the addition of CIAP

during the last 30 min of digestion at 37°C and purification. The pET32a vector was digested with XhoI and NdeI to excise the GST-ELP180 gene, purified from a gel, and ligated with the purified ELP library genes, resulting in products referred to as pET32a-ELP. These products were transformed into DH5α and allowed to recover for 1 h at 37°C in LB before plating on LB supplemented with 100 mg/mL of ampicillin. The resulting colonies were screened by diagnostic digestion and positive clones were confirmed by DNA sequencing.

2.2.5   Making SF-ELP fusions

The orthogonal sigma factor (SF) ECF20_992 [37] was previously amplified from pVRa20_992 (Addgene) and fused to the C-terminus of ELP180 to make pET32a-SF-ELP180. A C-terminal fusion was chosen since other studies have shown superior expression of C-terminus fusions to ELP compared to N-terminal fusions [48]. Additionally, SF was fused to the C-terminus of ELPA40. To make the fusion, pET32-ELPA40 was first linearized with NdeI, followed by the addition of CIAP during the last 30 min of digestion at 37°C, and purified from a gel. Orthogonal sigma factor was then amplified from pVRa20_992 using primers 9 and 10 (Table 4, appendix A), digested with type IIS FauI, purified from a gel, and ligated with purified pET32-ELPA40. This resulted in a pET32-SF-ELPA40 vector that was then transformed into DH5α and allowed to recover for 1 h at 37°C in LB before plating on LB supplemented with 100 mg/mL of ampicillin. The resulting colonies were screened by diagnostic digestion and a positive clone was confirmed by DNA sequencing.

2.3   ELP expression and purification

2.3.1   Expression

Plasmid DNA containing the ELP constructs were transformed into BL21 (DE3) and plated on selective agar. Single colonies were used to inoculate 5 mL starter cultures of TB supplemented with appropriate antibiotic (100 mg/mL ampicillin for pET32a-ELP vectors or 50 mg/mL kanamycin for pET28a-ELP vectors). These cultures were then incubated at 37°C overnight at 250 RPM. The cultures were then resuspended in 1 mL of fresh TB and inoculated into 199 mL cultures of TB in a 1 L flask supplemented with antibiotic. Cultures were then grown for 24

hours at 250 rpm and pelleted by centrifugation at 8,000 RPM for 5 min. After decanting excess media, the cell pellets were flash frozen and stored at -80°C until purification.

2.3.2 Purification

To purify the ELP proteins, two types of methods were used: (1) sonication followed by inverse transition cycling (ITC) and (2) organic solvent extraction [49] followed by ITC. Additionally, a quick protein expression protocol was used to compare the whole cell lysates (WCL) of induced and noninduced samples of pET32a-ELP.

2.3.2.1 Purification by sonication and inverse transition cycling (ITC)

Initial efforts to purify ELP from both the pET28a and pET32a vectors involved cell lysis by sonication followed by several rounds of ITC. A typical round of ITC involves (1) adding small quantities of ammonium sulfate to induce the phase transition of ELP, (2) centrifuging at max speed at 40°C for 10 min to collect the protein pellet, (3) resuspending the pellet in 200 µL PBS, and (4) centrifuging at max speed at 4°C for 10 min to collect the ELP containing supernatant (Figure 5). Ideally, each round of ITC improves the purity of the sample.
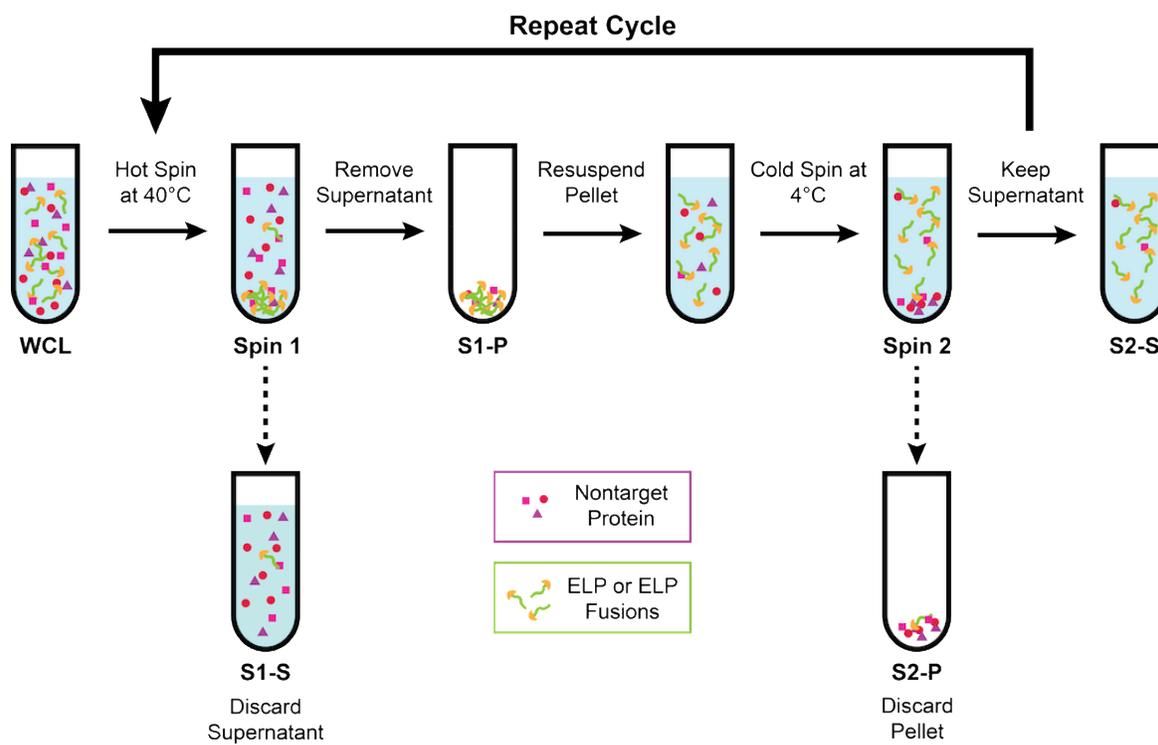
Figure 5. One cycle of inverse transition cycling (ITC).

Cultures were first centrifuged for 5 min at 8,000 RPM to collect cell pellets, which were resuspended in 3-4 mL of phosphate buffered saline (PBS) (Fisher BioReagents) with added Pierce protease inhibitor tablets (Thermo Scientific) and DNase (NEB). Cells were lysed using multiple cycles of 15 s of sonication followed by 30 s on ice. Once the cellular debris was removed by centrifugation, ITC was performed. Protein concentration was quantified using a Pierce BCA Protein Assay Kit (Thermo Scientific). The ELP proteins were visualized using SDS-PAGE by diluting 10 μg of protein in 16 μL of sterile water, 20 μL of Novex Tris-Glycine SDS Sample Buffer (2X) (Thermo Scientific), and 2 μL of 1 M dithiothreitol (DTT) (Fisher Bioreagents) and heating for 5 min at 95°C. The samples were ran on a 12% Tris-HCl polyacrylamide gel at 120 V and stained overnight using SYPRO Ruby protein gel stain (Invitrogen). The gels were visualized using an Azure Biosystems c400 imaging system.

2.3.2.2   Purification by organic solvent extraction followed by ITC

In collaboration with Craig Sweet, ELP protein expressed from the pET32a vector was purified using an organic solvent extraction method adapted by VerHeul *et al*. [49] followed by 2-3 rounds of ITC. Several different organic solvent combinations were used to screen the ELP library for optimal purity (data not shown). All organic solvents were mixed in a 1:1 (v/v) ratio. From this data, it was concluded that the best combinations to use were butanol + ethanol for ELPN40, and isopropyl alcohol + butanol for ELPA constructs. Using this method, the ELPN80, ELPN160, and ELPB proteins were unable to be purified.

First, 3 mL of organic solvent mixtures were added per 1 g of cell pellet, vortexed, and centrifuged at 8,000 RPM for 5 min to collect the supernatant. A 1:1 (v/v) ratio of d_2H_2O was then added and mixed with the supernatant followed by the addition and mixing of ethyl acetate in a 1:1 (v/v) ratio. The solutions were left alone for 10 min or until two layers became visible. The lower aqueous phase was then separated from the upper organic phase and allowed to evaporate either passively on an orbital shaker at 37°C, or actively under a constant stream of air at room temperature. Once evaporated, 200 μL of d_2H_2O was added to resuspend the dried protein. To further purify the ELP protein, 2-3 rounds of ITC were performed. Samples were then visualized via SDS-PAGE as described previously (Figure 6). Although the SYPRO Ruby protein gel stain worked ideally for the ELPA proteins, it proved to be a poor choice for

visualizing ELPN40. It was discovered that colloidal Coomassie stain was a better option for ELPN40, however, the bands tend to disappear when stained longer than a couple hours.

2.3.2.3   Quick protein expression protocol

A quick expression experiment was conducted to compare the whole cell lysates of induced versus noninduced samples of the pET32a-ELP library (Figures 18-20, appendix A). Single colonies were used to inoculate 5 mL TB cultures supplemented with 100 mg/mL ampicillin and grown overnight at 37°C at 250 rpm. 100 μL of overnight culture was used to inoculate two fresh 5 mL TB cultures/sample with antibiotic; one was induced with IPTG to a final concentration of 10 mM and the other was not. Cultures were grown for ~4 h before 1 mL of cells at OD600 = 0.8 was collected and centrifuged to remove the supernatant. Pellets were then resuspended in 100 μL of Novex Tris-Glycine SDS Sample Buffer and lysed by heating to 95°C for 10 min. Once samples were cooled to RT, they were centrifuged at max speed for 5 min to remove the cellular debris. Finally, 10 μL of each sample WCL was visualized by SDS-PAGE and stained using SYPRO Ruby protein gel stain.

2.4   Characterizing ELP phase transition

2.4.1   Characterizing ELP T_t *in vitro*

For each purified ELP, the transition temperature was determined *in vitro* by finding the maximal change in absorbance over a range of temperature. First, ELP was purified using methods previously described, dialyzed overnight to remove residual salt from ITC, and lyophilized. Absorbance readings at 600 nm were then taken over a temperature range at a rate of 0.5 °C/min using a Thermo Scientific Evolution 260 Bio UV-Vis spectrophotometer equipped with a Peltier thermostatting accessory. For ELPN40, lyophilized protein was resuspended in d_2H_2O to final concentrations of 30 μM, 45 μM, 60 μM, and 200 μM. Absorbance was plotted against temperature to produce the characteristic S-shape curves (Figure 22, appendix A). The change in absorbance was then plotted against temperature to determine the T_t values (Figure 23, appendix A). Finally, T_t was plotted against concentration and a linear trend was applied (Figure

7). This data was used to interpolate values of concentration for physiologically relevant transition temperatures (Table 3).

To characterize the pH sensitivity of the acidic ELPs, T_t was observed as function of pH. The lyophilized ELPA80 and ELPA160 proteins were resuspended at 100 μM in PBS at pH 3, 4.5, 6, 7.5, and 9 and absorbance readings were taken over a range of temperature. As a control, absorbance readings of 100 μM of ELPN40 in PBS at pH 3, 4.5, 7.5, and 9 were also taken. The change in absorbance was first plotted against temperature in order to determine the T_t for each sample (Figures 24-26, appendix A). The calculated T_t was then plotted against pH (Figure 8). To find the best fit, nonlinear least squares was performed, and data was fit to the following function:

(1) $T_t = \frac{T_t max - T_t min}{1 + e^{-k(pH - pH_t)}} + T_t min$

The above function was based on a logistic function that produces a common "S-shape" curve. The parameters are as follow: $T_t$max represents the maximum observed T_t, $T_t$min represents the minimal observed T_t, k defines steepness of the curve, pH_t represents the pH at the midpoint of the curve, and pH indicates the pH value in the range from 2-10.

2.4.2   Characterizing ELP T_t *in vivo*

2.4.2.1   Making GFP-ELP fusion protein

To establish the phase transition behavior of ELP *in vivo*, we aimed to make GFP-ELP protein fusions so that the transition could be observed under a microscope. A GFP-ELP fusion protein was made by fusing mNeonGreen GFP to the C-terminus of ELP180. The mNeonGreen GFP gene was first PCR amplified from pNCS-mNeonGreen (Allele Biotech) [50] using primers 16 and 17 (Table 4, appendix A). The PCR product was then digested with NdeI and KpnI and ligated into the corresponding sites on the pET32a-GST-ELP180 vector, removing the GST tag. The resulting pET32a-mNeonGreen-ELP180 vector was transformed into DH5α and allowed to recover for 1 h at 37°C in LB before plating on LB supplemented with 100 mg/mL of ampicillin. The resulting colonies were screened by diagnostic digestion and a positive clone was confirmed by DNA sequencing. The confirmed pET32a-mNeonGreen-ELP180 vector was then transformed

into the BL21(DE3) expression strain and selected on LB supplemented with 100 mg/mL of ampicillin

2.4.2.2   Sample preparation for imaging

To express the GFP-ELP fusion protein, 5 mL cultures of TB supplemented with 100 mg/mL ampicillin were inoculated from a freezer stock of BL21(DE3) cells carrying the pET32-mNeonGreen-ELP180 vector and grown overnight at 37°C at 250 rpm. Fresh 5 mL TB cultures were inoculated from the starter culture to an OD600 of 0.6 and induced with IPTG to final concentrations of 0 mM, 0.01mM, and 0.05 mM. After two hours of growth, cells were fixed to microscope slides and imaged within an hour of fixing.

The method for fixing cells to the microscope slides was adapted from Ge *et al.* [51]. To prevent movement, the cells were immobilized using a 10% polyacrylamide gel. The gel was prepared by centrifuging 500 μL of cell culture and resuspending in 377 μL of d_2H_2O, 119 μL of 40% acrylamide/bis solution (Bio-Rad), 3 μL of 10% ammonium persulfate (Invitrogen), and 0.3 μL of TEMED (Invitrogen). After mixing, 5 μL of solution were transferred onto microscope slides, covered with a cover glass, and left to polymerize. To prevent drying, nail polish was used to seal the edges of the cover glass. All slides were imaged within 1 h of sampling. Samples were imaged on a Nikon A1Rsi confocal microscope using a 20X objective lens. To image the GFP-ELP fusion protein, the laser source Argon ion (488nm) was used along with a transmitted light overlay. A temperature-controlled box was used to adjust the temperature of the slides between 24-33°C.

2.5   Characterizing simple activation and simple repression systems

2.5.1   Simple activation system construction and expression

The pET32a-SF-ELP180 plasmid was transformed into the *E. coli* strain MG1655 along with pN565 (Addgene) [37], which encodes a gene for T7RNAP, and pVRb20_992 (Addgene) [37], which encodes superfolder GFP under the P20_992 promoter that is activated by SF. As a control, the vectors pVRa20_992, pN565, and pVRb20_992 were transformed together into

MG1655. Single colonies were used to inoculate 4 mL starter cultures of LB supplemented with 100 mg/mL ampicillin, 50 mg/ml spectinomycin, and 50 mg/mL kanamycin and incubated at 37°C overnight at 250 RPM. Both cultures were then diluted to 32 mL of LB with antibiotic, 4 mL of which were transferred to two new tubes/sample. Each sample was induced with IPTG and grown at either 30°C or 40°C for 1 h. Samples were collected by centrifuging 1 mL of culture/tube and resuspending in 1 mL of PBS. 100 μL of this solution was then transferred to a 96-well plate. A plate reader was then used to collect readings for absorbance at OD600 and fluorescence using the parameters $\lambda_{ex} = 485$ nm and $\lambda_{em} = 528$ nm. Fluorescence readings were first normalized by OD600, averaged, and then normalized by the average fluorescence of SF-ELP180 at 40°C. To determine the significance of the change in expression, a two-tailed t-test assuming unequal variances using a significance level of 0.05 was used. The percent change in gene expression between 30°C and 40°C was calculated using equation 2.

$$(2) \; Percent \; change \; in \; expression = \frac{GFP \; expression \; at \; 40°C - GFP \; expression \; at \; 30°C}{GFP \; expression \; at \; 30°C} * 100\%$$

2.5.2 Simple repression system construction and expression

2.5.2.1 Cloning of the simple repression system constructs

To make the simple repression system, three vectors were cloned. The vector pET28a-TetR-RDG was made by fusing TetR with the aECM1 protein under an *E. coli* constitutive promoter. A control vector, pET28a-TetR, was made by excising aECM1and inserting TetR under the same promoter. The vector pVRb-Ptet-mCherry was made by cloning the reporter gene, mCherry RFP, under the control of a Tet repressible promoter.

First, the TetR sequence was amplified from pCas9cr4 [52] using primers 11 and 12 (Table 4, appendix A), followed by a second round of PCR using primers 12 and 13 (Table 4, appendix A) to insert the promoter sequence. To make the control plasmid, the PCR product and the pET28-RDG vector were both digested with BglII and XhoI, purified from a gel, and ligated together. The pET28a-TetR-RDG vector was made in a similar way but was instead digested with BglII and NheI. Ligation products were transformed into DH5α and allowed to recover for 1 h at 37°C

in LB before plating on LB supplemented with 50 mg/mL of kanamycin. To clone the pVRb-Ptet-mCherry vector, primers 14 and 15 (Table 4, appendix A) were used to amplify mCherry from pRSETb-mCherry [53]. This PCR product and modified pVRb vector were digested with BamHI and BglII, purified from a gel, ligated, and transformed into DH5α. The transformant was allowed to recover for 1 h at 37°C in LB before plating on LB supplemented with 25 mg/mL of chloramphenicol. Colonies were screened by diagnostic digestion and positive clones were confirmed by DNA sequencing.

2.5.2.2   Expression of the simple repression system

The simple repression system constructs were transformed into the strains BL21(DE3) and MG1655 and plated on selective agar. Single colonies were used to inoculate 5 mL starter cultures of LB supplemented with appropriate antibiotic, which were grown at 37°C at 250 rpm. Starter cultures were used to inoculate fresh 5 mL LB cultures at an OD600 of 0.3. For some conditions, ATc was added to a final concentration of 0.05 mg/mL. Each sample was grown in triplicate at 30°C and 40°C, shaking at 250 rpm, and the OD600 and fluorescence intensity were measured after 2 and 3 hours of growth using a Biotek plate reader. For each sample, fluorescence intensity was first normalized by OD600 and then averaged. For each time point, the averaged samples were normalized by the average fluorescence of TetR + ATc at 40°C to produce graphs that showed relative fluorescence.

2.6   Simple activation and FFL model simulations

Models for the simple activation and feed-forward loop systems were simulated in MATLAB. For both models, equations 3-5 were used to describe the phase transition behavior of the ELP-SF fusion protein in order to determine the portion that is in the active form. The active form (ELPsfA) is defined as the fraction of total ELP-SF fusion protein (ELPsfT) that is soluble and capable of gene activation, while the inactive form (ELPsfI) is defined as the aggregated fraction. In equation 3, the total ELP-SF protein available at steady state is described by a Hill function where gene activation depends upon IPTG concentration. Total ELP-SF also depends upon the max transcription rate (β_1), mRNA degradation rate (k_{dm}), translation rate (K), and dilution of the protein due to cell growth (μ). Due to the stability of ELP, protein degradation was assumed

negligible. Equation 4 describes the amount of inactive ELP-SF present using a modified logistic equation, which produces an "S-shape" curve that resembles the curve of absorbance versus temperature. This equation was fit to transition data collected for ELPN40 *in vitro* in order to estimate the steepness coefficient, k_1. It was assumed that ELP-SF is 100% inactive when fully aggregated and 100% active when fully soluble. Due to conservation, the active portion can be determined by subtracting the inactive portion from the total amount present (equation 5).

(3) $ELPsfT_{SS} = \frac{K}{\mu} * \frac{\beta_1}{k_{dm}} * \frac{IPTG^{n_1}}{A^{n_1} + IPTG^{n_1}}$

(4) $ELPsfI = ELPsfT * \frac{1}{1 + e^{-k_1(T - T_t)}}$

(5) $ELPsfA = ELPsfT - ELPsfI$

The simple activation architecture was modelled using the following equations assuming steady state:

(6) $mRNA_{SS} = \frac{\beta_2}{k_{dm}} * \frac{ELPsfA^{n_2}}{K_{m1}^{n_2} + ELPsfA^{n_2}}$

(7) $GFP_{SS} = \frac{K}{\mu} * mRNA$

The feed-forward loop architecture was modelled using the following equations assuming steady state:

(8) $ORIB_{SS} = \frac{\beta_3}{k_{dm}} * \frac{ELPsfA^{n_2}}{K_{m1}^{n_2} + ELPsfA^{n_2}}$

(9) $mRNA_{SS} = \frac{\beta_2}{k_{dm}} * \frac{ELPsfA^{n_2}}{K_{m1}^{n_2} + ELPsfA^{n_2}}$

(10) $GFP_{SS} = \frac{K}{\mu} * \frac{ORIB^{n_3}}{K_{m2}^{n_3} + ORIB^{n_3}} * mRNA$

Equations 6 and 7 describe mRNA and GFP production for the simple activation model. The production of mRNA (equation 6) is a function of max transcription rate from the P20_992 promoter, mRNA degradation, and gene activation by the substrate ELPsfA using Hill kinetics. The production of GFP (equation 7) is then a function of mRNA concentration, translation rate, and dilution due to growth. The degradation of GFP was assumed negligible due to the stability of the protein. The equations for the FFL model are similar to those for the simple activation model but include the added o-ribosome (ORIB) element. Equation 8 describes ORIB

production, which is a function of max transcription rate from the P20_992 promoter, rRNA degradation rate, and gene activation by the substrate ELPsfA using Hill kinetics. The production of mRNA (Equation 9) is the same as before. The production of GFP (equation 10) is a function of mRNA concentration, ORIB concentration and assembly following Hill kinetics, and translation rate of mRNA by ORIB.

Equation parameters can be found in Table 2. For all equations describing RNA production, the basal transcription rate was assumed to be 0 and therefore omitted from the above steady state equations. Typically, *E. coli* can transcribe a gene at 80 bp/s [54]. Maximal transcription rates (β) for ELP-SF, GFP, and ORIB were calculated according to the length of each gene. The hill coefficient (n) governs the steepness of the hill function. A value of n = 1 indicates independent substrate binding, while n > 1 indicates positive cooperativity. The exact value of n_2 is not known, but activation by the ELPsfA substrate was assumed to be slightly cooperative. It was assumed that ORIB assembly with mRNA is an independent binding event and n_3 is therefore equal to 1. The activation coefficient (K_m) is defined as the concentration of inducer needed to significantly activate expression; its value depends upon the chemical affinity of the inducer for the promoter site, in addition to other factors. The K_{m1} value for activation by ELPsfA is not known but was estimated to be within a reasonable range and was adjusted so that the simple activation model aligned with observed data at 40°C. Dilution rate due to cell growth was estimated assuming a doubling time of 30 mins and was considered negligible for mRNA production due to differences in timescales. The percent change for predicted gene expression between 30°C and 40°C was calculated using equation 1. The MATLAB code used to model the two systems can be found in Appendix B.

Table 2. Model parameters.

Parameter	Value	Description
A	1.3 μM	IPTG activation coefficient [55]
n_1	2	Hill coefficient describing activation by IPTG [56]
n_2	1.2	Hill coefficient describing activation by ELPsfA (estimated)
n_3	1	Hill coefficient describing ORIB assembly with mRNA (estimated)
k_{dm}	0.83 1/h	mRNA degradation rate [57]
K	300 uM protein/uM mRNA/h	Translation rate of mRNA [58]
μ	1.39 1/h	Growth rate assuming doubling time of 30 min
T_t	39°C	Transition temperature of phase change

Table 2. continued

k_1	2.4	Coefficient for steepness (determined experimentally)
β_1	0.15 μM/h	Max transcription rate of ELPsf from T7 promoter (estimated assuming 80 bp/s [54])
β_2	0.68 μM/h	Max transcription rate of GFP from P20_992 promoter (estimated assuming 80 bp/s [54])
β_3	0.089 μM/h	Max transcription rate of ORIB from P20_992 promoter (estimated assuming 80 bp/s [54])
T	30-45°C	Temperature range
K_{m1}	1.1 μM	ELPsfA activation coefficient (estimated)
K_{m2}	1 μM	ORIB activation coefficient [59]

3.   RESULTS & DISCUSSION

3.1   ELP library characterization

To meet the first objective of this study, an ELP library was successfully designed and cloned that exhibits pH sensitivity and transitions under various conditions. PRe-RDL was used to achieve lengths of 40, 80, and 160 pentamer repeats for each of the three designed ELP types: a "neutral" ELPN where guest reside X = V, an "acidic" ELPA where X = V/I/E [1:3:1], and a "basic" ELPB where X = V/H/G/A [1:2:1:1]. ELP expression from an initial pET28a vector proved to be poor but improved once the library was transferred to a pET32a vector. I worked in collaboration with Dr. Dave Thompson's lab at Purdue University to achieve high purity yields of ELPA40, ELPA80, ELPA160, and ELPN40 using a recently developed organic solvent extraction method [49]. These purified ELP products were then used for *in vitro* phase transition characterization studies to determine the effect of concentration and pH on transition temperature. Additionally, *in vivo* studies were conducted to gain a better understanding of ELP phase transition behavior within a cellular environment.

3.1.1   ELP library expression and purification

3.1.1.1   Troubleshooting ELP expression from the pET28a vector

The ELP library was originally expressed from a pET28a vector, but expression and purification proved to be challenging. It was unclear whether the problem was an issue of protein expression, ELP toxicity to the cell, or inclusion body formation. First, plasmid integrity was checked by expressing them in MG1655 and BL21(DE3), purifying them, and having them sequenced. Results did not show any significant mutations that would have affected expression. Initially, purification of ELP was performed only by ITC. Difficulties in achieving adequate purity of ELP samples made drawing conclusions based on SDS-PAGE results problematic; often, the results were inconclusive. To circumvent this issue, mCherry was cloned in place of ELP in the pET28a vector and expression was compared to a similar vector, pETM6 which also expresses mCherry from a T7 promoter and has a similar plasmid copy number. These similarities indicate that differences in expression are due to construct design. Significantly higher fluorescence was

observed from pETM6 (Figure 16, appendix A), suggesting that the issue was not a result of protein toxicity or inclusion body formation, but due to poor protein expression.

It was hypothesized that the folding of the mRNA secondary structure was making the RBS inaccessible to the ribosome, thus interfering with translation and leading to poor expression. To investigate this possibility, the mfold web server [60] was used to predict the mRNA secondary structure at the RBS site upstream of the ELPB gene in the pET28a vector and was compared to the RBS sites for ELP genes in two other pET vectors that have previously been shown to have good expression (Figure 17, appendix A). These predictions show that the entire RBS site upstream of ELPB is completely buried within the stem of a hairpin loop, while the first adenine in the RBS sites of the other two ELP genes is freer and not hydrogen bonded to another base. It was therefore concluded that this difference in conformation is most likely the reason for poor expression from the pET28a vector. To improve ELP expression, the ELP library was transferred to a pET32a vector that had previously showed good expression of GST-ELP180; the GST-ELP180 gene was removed and the designed ELP genes were cloned downstream of its T7 promoter and RBS site. This change in vector improved ELP expression as the phase transition during ITC became more prominent.

3.1.1.2   ELP Purification

Multiple expression and purification techniques were tested to confirm ELP expression and to obtain enough pure isolated protein to be used for *in vitro* phase transition characterization studies. A quick expression method without purification was used to compare induced and uninduced sample whole cell lysates, but ELP presence was mostly indiscernible from other cellular proteins (Figures 18-20, appendix A). To isolate the ELPs, the ITC method of purification was initially used, but faced challenges in obtaining high purity samples. Lingering contaminants from ITC often made it difficult to discern which bands were ELP and which were non-target protein on SDS-PAGE gels. A second issue when using ITC is that each round diminishes the final protein yield.

To improve the efficiency and purity of ELP isolation, a new, alternative ELP purification method was adapted from VerHeul et al. [49] that uses organic solvents to simultaneously lyse

cells and extract the ELP. This organic solvent extraction method speeds up the recovery time and allows for the isolation of higher yields of ELP with greater than 95% purity. I worked in collaboration with Dr. Dave Thompson's lab to screen the ELP library against various organic solvents and solvent blends to find the optimal purification conditions. It was discovered that purer isolates of ELPN40 and ELPA could be obtained using blends of butanol + ethanol and isopropyl alcohol + butanol respectively. This extraction method followed by 2-3 rounds of ITC was used to successfully purify four of the nine designed proteins, ELPA40, ELPA80, ELPA160, and ELPN40.

The purified ELPAs along with their expected weights can be seen in Figure 6, while ELPN40 data is not shown. Organic solvent screens of ELPB, ELPN80, and ELPN160, however, did not yield purified product. It is hypothesized that the larger, more hydrophobic, ELPN80 and ELPN160 had a lower solubility than ELPN40 when resuspended in d_2H_2O and is the reason why purification was unsuccessful. For ELPB, it is unclear whether modifications need to be made to the extraction method to purify these proteins or if protein expression is poor. In an attempt to show expression, cell disruption by sonication followed by three rounds of ITC was used to try to isolate ELPB160 (Figure 21, appendix A), which was chosen since it is expected to have the lowest T_t out of the three ELPB sizes. Unfortunately, the results were inconclusive as ELPB160 could not be distinguished from non-target protein. It is hypothesized that the T_t of ELPB160 may still exceed the 37-40°C max that can be achieved during ITC, and thus may be getting lost in the first round. In the future, this experiment will be repeated with a buffer at a higher pH to further depress the T_t below this upper limit.
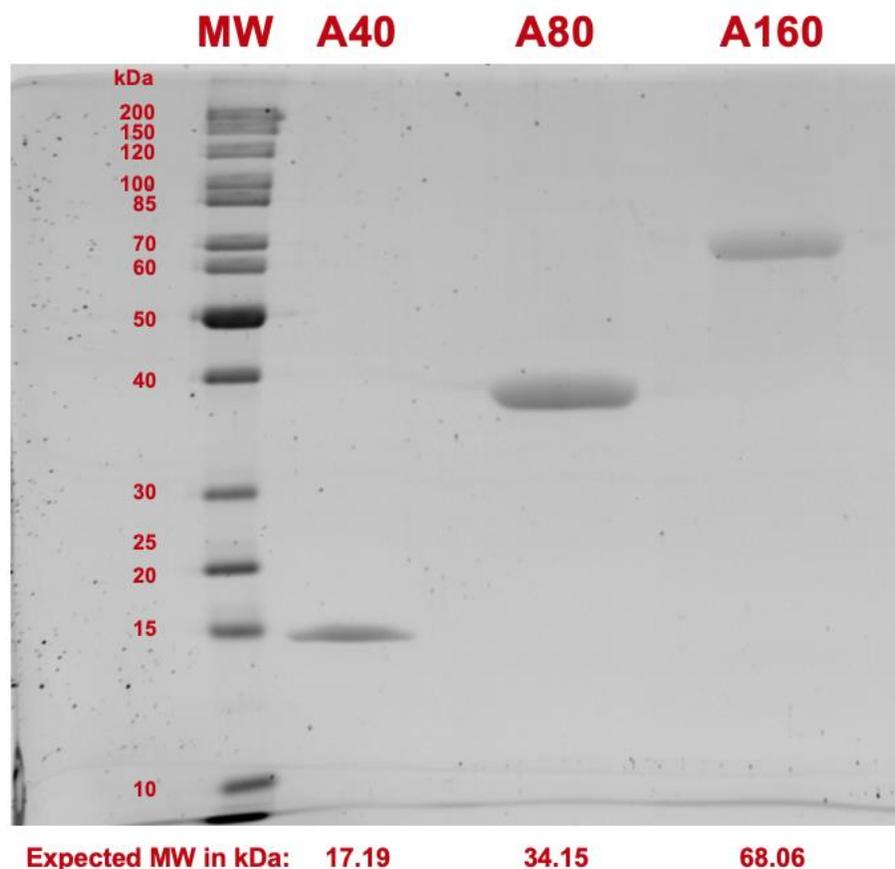
Figure 6. SDS-PAGE demonstrating expression and purity of the ELPA protein at n = 40, 80, & 160. The left lane is a molecular weight (MW) standard labeled in kDa and the expected molecular weights of each ELPA protein is listed below their respective lane.

3.1.2 Characterizing the phase transition of ELP *in vitro* and *in vivo*

3.1.2.1 Characterizing ELP T_t *in vitro*

Phase transitions of purified ELP were determined *in vitro* by tracking the change in absorbance over a range of temperature using a spectrophotometer. The observed T_t values for ELPN40 were plotted against concentration (Figure 7), the values of which are listed in Table 3. The relationship between T_t and concentration appears to follow a linear trend, where T_t decreases as concentration increases. This relationship was used to interpolate concentration values for physiologically relevant transition temperatures (Table 3). Since larger ELPs have lower transition temperatures, it can be assumed that the concentrations of ELPN80 and ELPN160 necessary to achieve the same T_t will be lower than those predicted for ELPN40.

A previous experiment yielded 3.3 mg/400 mL culture of purified ELPN40, which corresponds to a cellular concentration of ~496 μM. This value is within the same order of magnitude as the estimated concentrations (149-188 μM) needed for the T_t of ELPN40 to be within a physiologically relevant range (39-35°C), suggesting that phase transitions in this range are feasible *in vivo*. In the case that cellular concentration is too high, resulting in transitions below this temperature range, shorter ELPs can be made to raise the T_t at the same concentration.
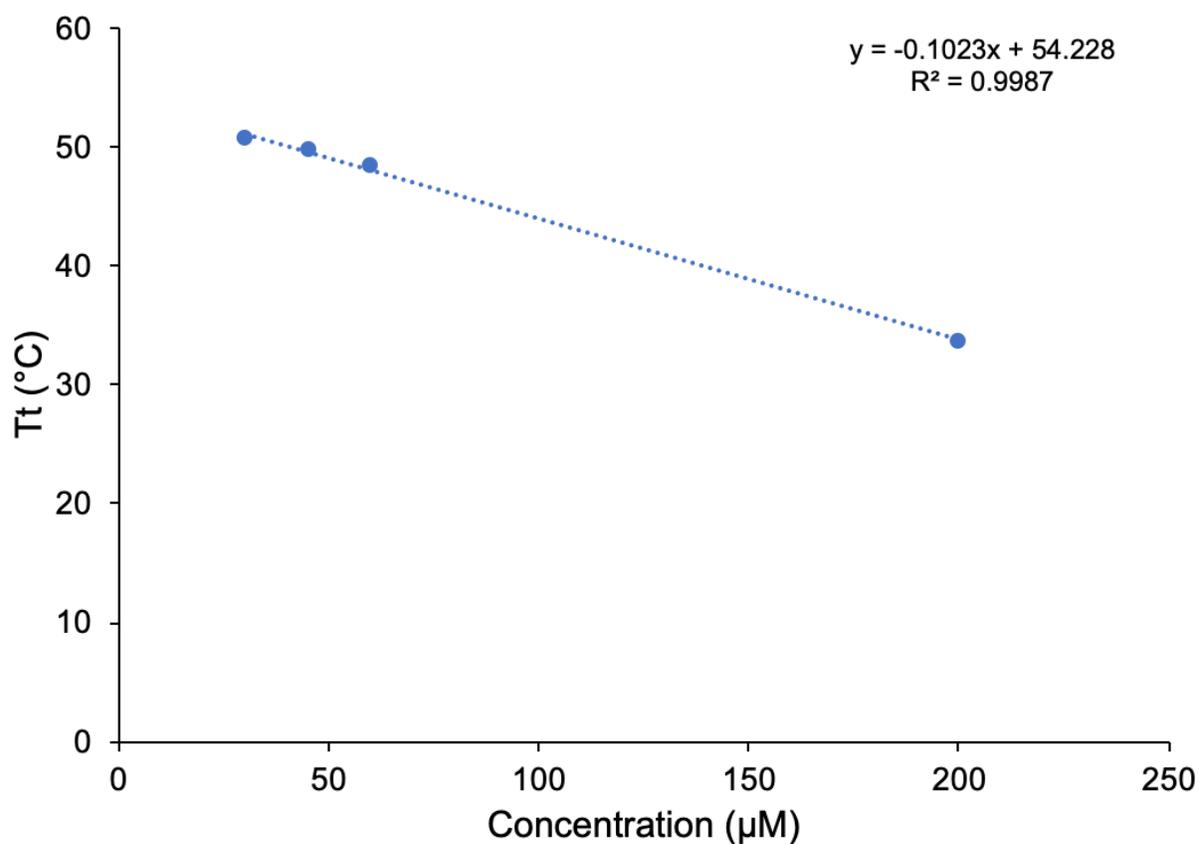
Figure 7. The observed phase transition temperatures of ELPN40 at 30 μM, 45 μM, 60 μM, and 200 μM.

Table 3. Estimated concentrations for physiologically relevant transition temperatures of ELPN40.

Tt (°C)	Estimated Concentration (μM)
39	148.84
37	168.39
35	187.94

To demonstrate the pH sensitivity of ELPA, observed T_t was plotted against pH (Figure 8). We observed a nonlinear trend for T_t with respect to pH that follows a very similar pattern to what has previously been observed by Mackay et al. [35]. This trend indicates that lowering the pH of the solution depresses the T_t for our acidic ELPs. We also observe that the transition temperatures for ELPA160 were overall lower than those of ELPA80 indicating that the larger the ELP is, the lower its T_t will be. It can therefore be assumed that ELPA40 will have higher transition temperatures than ELPA80. As predicted, the ELPN40 control was not sensitive to changes in pH.
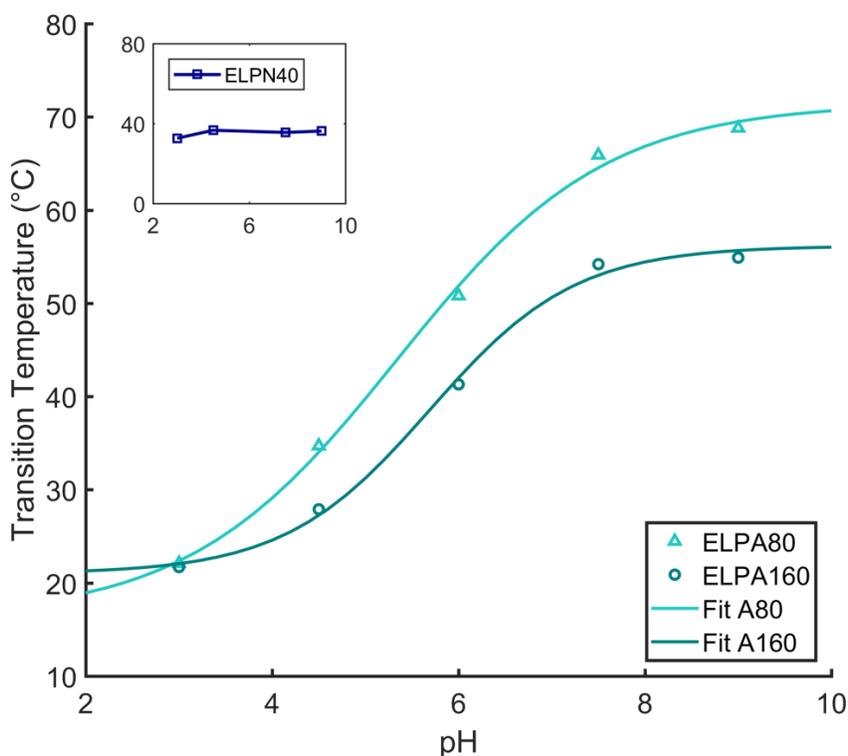
Figure 8. The transition temperature (T_t) dependence on pH for ELPA80, ELPA160, and ELPN40 at 100 μM.

According to the observed trend, there appears to be an upper limit for T_t once the pH has exceeded 7.5. Above this pH, the changes in T_t are minimal (0.7-2.9°C). The change is more dramatic, however, between pH 7.5 and 6 (12.9-15.1°C), which indicates that small drops in pH from *E. coli's* normal operating range (7.6-7.8) [61] will be able to trigger the phase transition response of the controller. It is predicted that toxic intermediate build-up will cause a drop in intracellular pH below this range, resulting in ELP aggregation. This response will then turn off expression of the toxic intermediate until the pH can return to healthy levels. For ELPA160 at 100 μM, the pH drop from 7.5 to 6 caused the T_t to change from 54.2°C to 41.3°C, which is higher than the typical 37°C growth temperature of *E. coli*. For the controller to operate as desired, the T_t would need to be above 37°C at a healthy pH (7.6-7.8) and below 37°C when pH drops below 7.6. The cellular concentration of ELP is expected to be ~496 μM, which may lower the T_t enough to fall within the desired physiological range. However, it is also probable that the *in vitro* phase transition behavior of pure ELP does not directly translate to its behavior within the cellular environment. Therefore, *in vivo* studies need to be performed to better characterize the phase transition within the cell.

Future work will involve characterizing the T_t of ELPA40 with respect to pH to see how it compares to ELPA80 and ELPA160. The rest of the ELP library will also be characterized once the issue of purification has been resolved. Since the fusion of ELP to transcription factor will alter the T_t, any ELP fusion proteins made will also need to be characterized *in vitro*.

3.1.2.2   Characterizing ELP T_t *in vivo*

Although characterizing the T_t *in vitro* offers a convenient way to establish the transition properties of pure ELP, it cannot be assumed to accurately reflect ELP behavior in a cellular environment. *In vitro* data may provide a good starting basis, but further *in vivo* studies are needed to better predict this behavior. To characterize the phase transition of ELP *in vivo*, the transition of a GFP-ELP fusion was monitored using a confocal microscope with a temperature control device. It was hypothesized that soluble GFP-ELP fusions will appear uniformly distributed throughout the cell, while the aggregation of GFP-ELP will cause the protein to polarize to the ends of the cell. The temperature at which this change occurs will be considered the transition temperature. This prediction was based on another study that observed this

behavior in which ELP induced the formation of an aqueous two-phase system (ATPS) in *E. coli* cells [51].

Although the temperature control device could easily heat the samples, it was limited in its ability to cool the samples below room temp. Thus, a lower limit of 24°C was reached, making it impossible to detect a transition below this temperature. The fusion protein GFP-ELP180 was expressed in *E. coli* using an IPTG induction concentration of 0.05 mM and then observed at 25°C and 33°C with the microscope (A, B from Figure 9). At both temperatures, the polarization of GFP-ELP180 to the ends of the cell can be seen, suggesting that the T_t was lower than what could be observed. Since the transition temperature is concentration dependent, we tried to raise the T_t above room temperature by decreasing the amount of GFP-ELP180 protein produced. Using lower induction levels, however, produced such low quantities of protein that it was difficult to image the cells, resulting in poor quality images. The sample that was not induced with IPTG did not produce enough protein to be detected by the microscope. Although the quality of the images for the sample induced at 0.01 mM IPTG is poor, polarization of GFP-ELP can still be seen at both 24°C and 28°C (C, D from Figure 9).

From this study, it was concluded that the decrease in protein concentration was still not enough to raise the T_t above the lower limit of the microscope, making the phase transition of the GFP-ELP180 fusion unobservable. Expressing the protein at lower concentrations also results in poor quality images. Going forward, it will be necessary to find a way to cool the system below 24°C in order to observe the T_t of larger ELPs that have low transition temperatures. It may be possible, however, to observe the phase transitions of shorter ELPs with higher transition temperatures using the current set up. In the future, fusions of GFP with ELPN40, ELPB40, and ELPA40, which have the highest transition temperatures of the ELP library, will be made to see if their transitions can be detected *in vivo*.
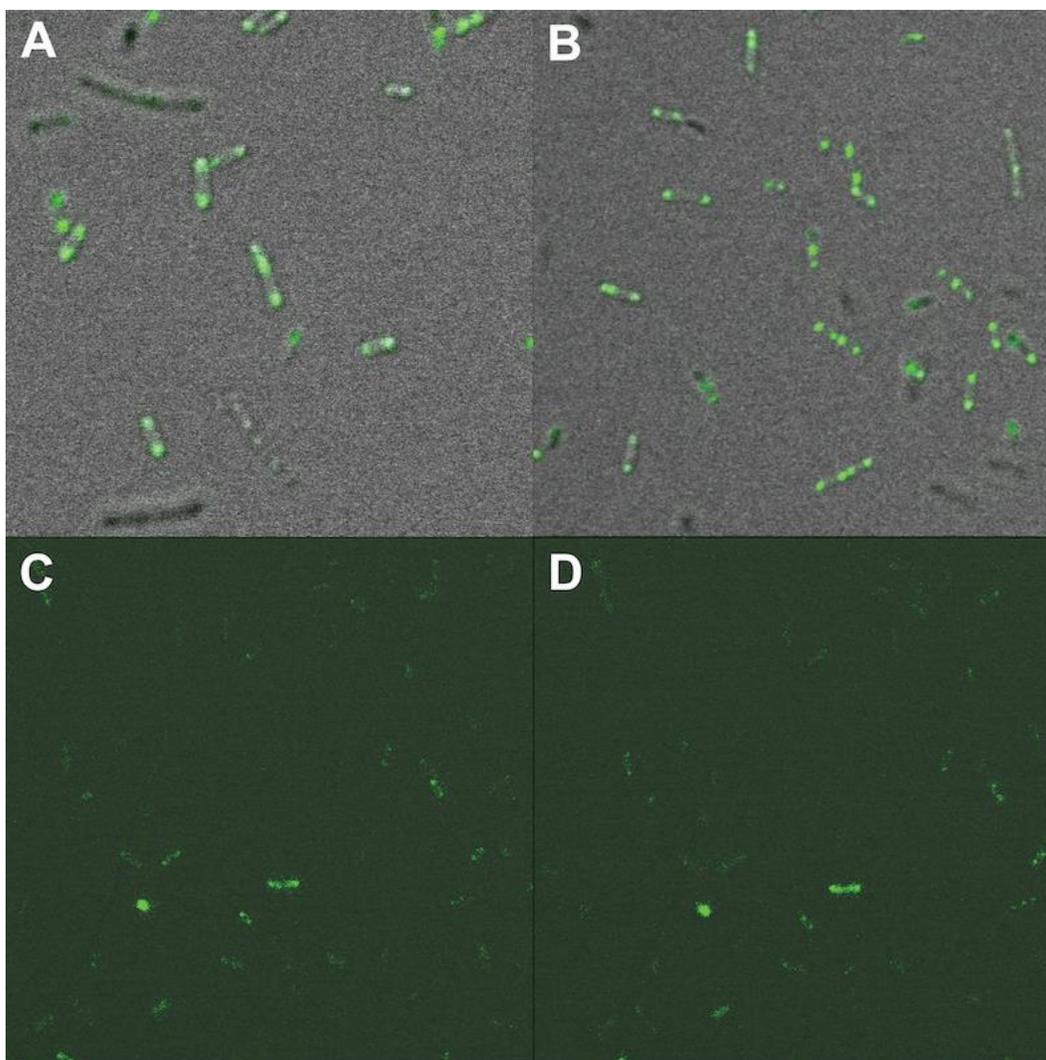
Figure 9. Confocal microscope images of GFP-ELP180 expressed in *E. coli.* (A, B) samples were induced with IPTG to a concentration of 0.05 mM and imaged at (A) 25°C and (B) 33°C. (C, D) samples were induced to a concentration of 0.01 mM and imaged at (C) 24°C and (D) 28°C.

3.2 Determining the feasibility of ELP-transcription factor controllers

To meet the second objective of this study—determining the feasibility of ELP-transcription factor fusions as dynamic sensor-regulator devices for gene expression—simple activation and repression architectures were designed and tested *in vivo* to control the expression of a fluorescent reporter gene. These studies produced conflicting results; the simple activation system showed evidence for control of fluorescent gene expression, while the simple repression system did not. The simple activation system study produced a significant reduction in gene

expression upon ELP aggregation, suggesting that SF-ELP fusions can control expression in a stimuli dependent manner. The observed 21% reduction in expression, however, is likely too modest to be used in a real production pathway, indicating that alternative network architectures need to be explored to improve the dynamic range of the controller. To meet this challenge, a feed-forward loop architecture was modelled, which predicted an improved 35% reduction in expression and an increased ultrasensitive response.

3.2.1   Simple activation architecture demonstrates control of gene expression

In a preliminary study, a SF-ELP180 fusion protein and a SF control were used to drive the expression of GFP. *E. coli* MG1655 cultures harboring GFP and either the SF-ELP180 or control plasmids were grown at a low temperature (30°C) and a high temperature (40°C) to induce the phase transition of ELP180, which was presumed to occur in this range. It was hypothesized that fluorescence would decrease at 40°C as a result of the aggregation of SF-ELP180, while the SF control would be unaffected. Our results support this hypothesis, as a 21% reduction in fluorescence was observed at the elevated temperature for SF-ELP180 while an 8% increase in fluorescence was observed for the control (Figure 10). A two-tailed t-test using a significance level of 0.05 revealed that the change in fluorescence between 30°C and 40°C for SF-ELP180 was significant ($p = 0.006$), while the difference for the control was not ($p = 0.061$). While these results indicate that SF-ELP fusions can be used to regulate gene expression in a predictable, stimuli dependent manner, the dynamic range of the controller is modest at best. To address this issue, a feed-forward loop system was designed and is predicted to have a 35% reduction in expression, thus improving the dynamic range (Figure 13).
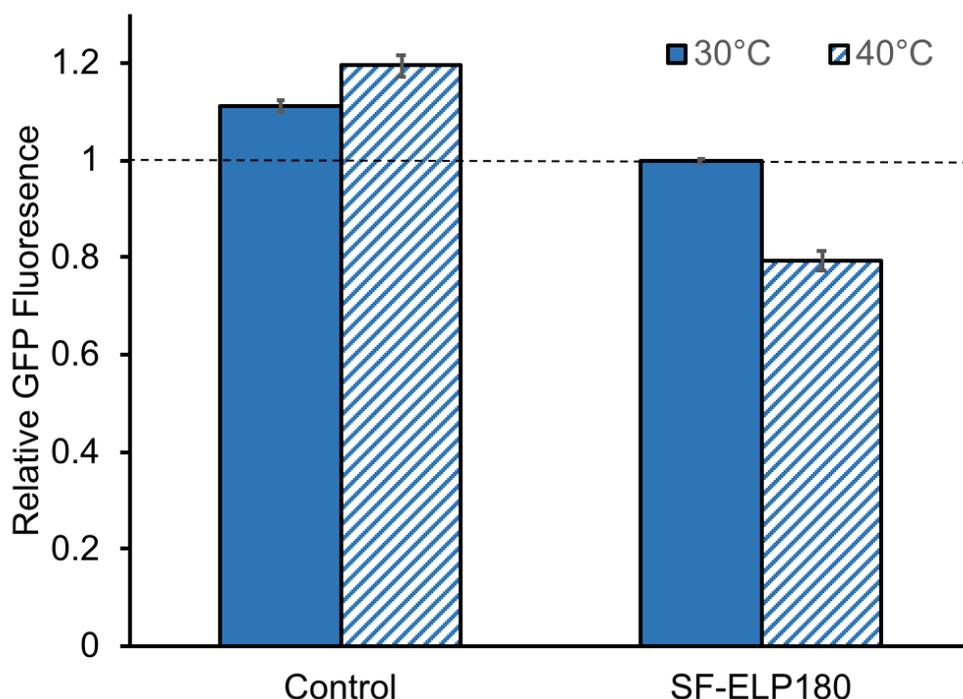
Figure 10. Relative GFP fluorescence in response to temperature for SF-ELP180 compared to a sigma factor control. The SF-ELP180 (X = V, n = 180) construct was expressed from pET32-SF-ELP180 and the SF control was expressed from pVRa20_992.

3.2.2 Simple repression architecture fails to control gene expression

In addition to the simple activation system, a simple repression system consisting of a TetR-ELP fusion and TetR control was tested for its ability to control the expression of a reporter mCherry RFP gene, which was cloned under a Tet repressible promoter. *E. coli* cultures transformed with RFP and either the TetR-ELP or TetR control plasmids were grown at a low temperature (30°C) and a high temperature (40°C) to induce the phase transition of ELP; the system was expressed in both MG1655 and BL21(DE3) strains. At the lower temperature, the constitutively expressed TetR-ELP is predicted to be in its soluble, active form, which will repress mCherry fluorescence. At the elevated temperature, it was hypothesized that TetR-ELP will aggregate, preventing TetR from repressing mCherry production and causing an increase in fluorescence. The control, meanwhile, will not be affected by temperature and continue to repress mCherry production. It was also anticipated that the addition of ATc, which inhibits TetR by binding to it, would result in high mCherry expression for both TetR-ELP and the TetR control.

Unfortunately, the expected trends were not observed, indicating that the simple repression system does not control gene expression as anticipated. In the BL21(DE3) strain (Figure 11), the addition of ATc to TetR causes a drastic increase in fluorescence, which was expected since ATc prevents TetR from repressing expression. This dramatic increase is not observed, however, for TetR-ELP at 40°C, which also should have increased expression due to ELP aggregation. Therefore, the observed expression levels for TetR-ELP were attributed as background/noise. It can also be noted that the expression for TetR-ELP was higher at 30°C than 40°C at both time points, which is the opposite of what was expected. It was also expected that the addition of ATc to TetR-ELP would cause an increase in expression comparable to TetR + ATc. Although an increase was observed, it was significantly lower than that of TetR + ATc.
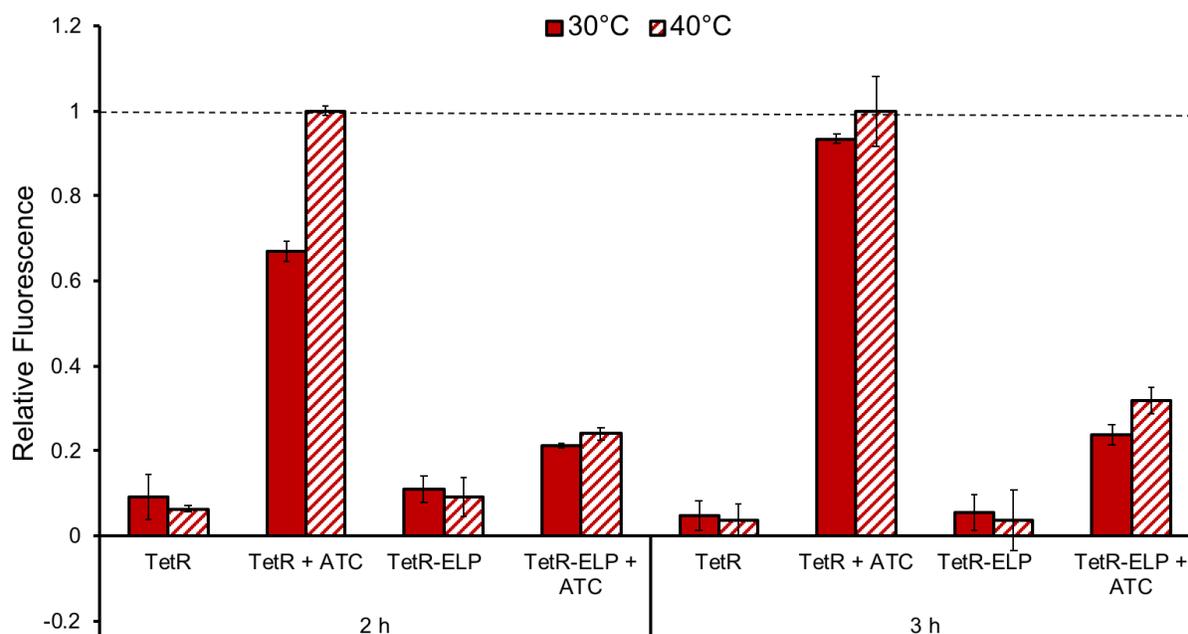
Figure 11. Expression of simple repression system in response to temperature in *E. coli* BL21(DE3). The TetR-ELP (X = K/I, n = 60) construct was expressed from the pET28-TetR-ELP vector.

Similar trends were observed when the system was expressed in MG1655 (Figure 12). Again, expression was higher at 30°C compared to 40°C for both TetR and TetR-ELP. Expression also dramatically increased once ATc was added, but TetR-ELP at 40°C remained low. A noticeable

difference from the BL21(DE3) strain occurred, in which expression of TetR-ELP + ATc at 3 h at 40°C drastically increased compared to the other conditions.
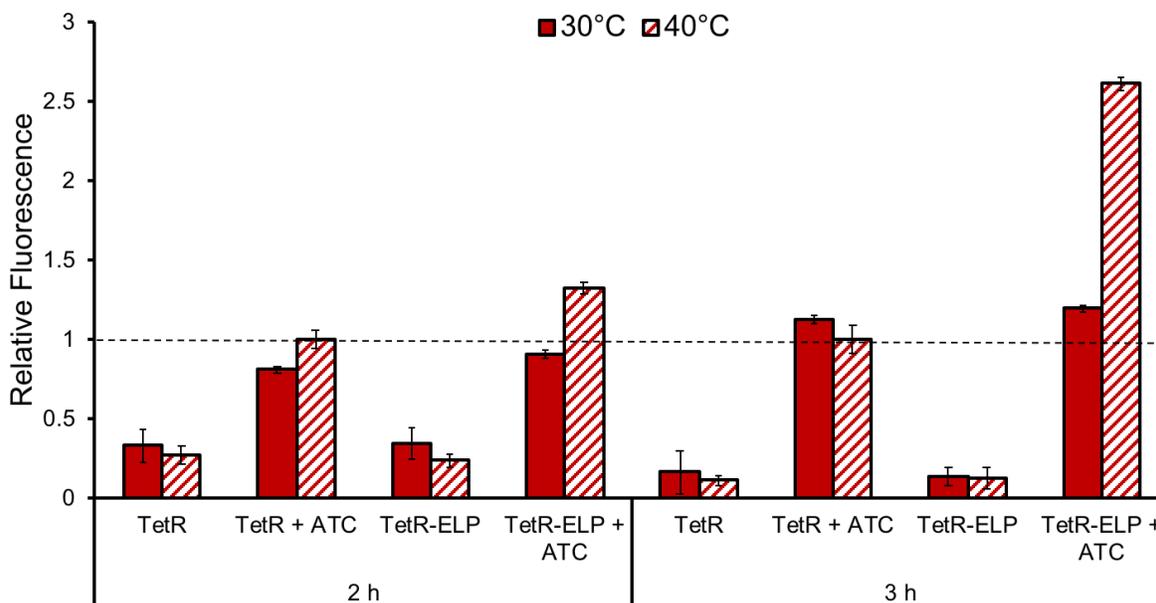
Figure 12. Expression of simple repression system in response to temperature in E. coli MG1655. The TetR-ELP (X = K/I, n = 60) construct was expressed from the pET28-TetR-ELP vector.

From the given results, it is unclear what caused the observed trends and why the simple repression system did not work as expected. It is hypothesized that failure to control gene expression may be due to TetR having a stronger binding affinity for DNA than ELP has for aggregation. As a result, ELP aggregation would not prevent TetR from binding to the promoter and thus could explain why ELP aggregation does not appear to increase mCherry expression at the elevated temperature. A second theory is that the conformational shape of the TetR-ELP fusion protein is not ideal to allow the assembly of aggregated ELP nanostructures. Alternatively, the problem could simply be that the T_t *in vivo* was higher than 40°C.

3.3 Modeling of alternative networks for enhanced controller performance

To explore alternative networks for improved dynamic range and ultrasensitivity of the ELP-transcription factor controller, a feed-forward loop architecture was modelled and compared to a

model for the simple activation architecture (Figure 13). Additionally, experimental results of the simple activation system were compared to the model in order to better predict the behavior of the system. The simple activation model response was aligned to show the same 21% reduction in GFP expression at 40°C as observed experimental data. The FFL model predicts a 35% reduction in expression at the same temperature, thus showing an improved dynamic range compared to the simple activation architecture. Additionally, the FFL model predicts an improved ultrasensitive response, meaning that the transition from high to low expression occurs over a smaller temperature range.
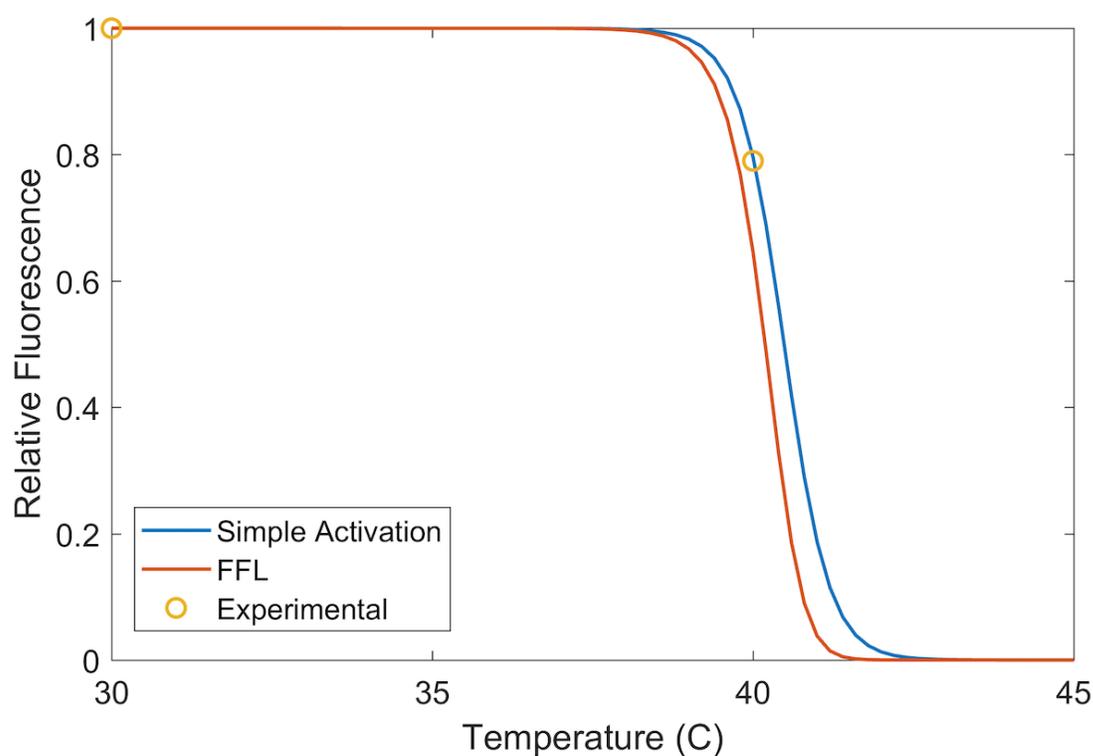
Figure 13. Estimated responses for simple activation and FFL architecture models compared to observed simple activation experimental data.

Both models predict that GFP expression will eventually drop to zero given a high enough temperature is reached past the T_t. This effect was not observed experimentally likely due to two reasons. First, it is possible that the T_t of ELP-SF *in vivo* may have been close to either 30°C or 40°C and thus the full reduction in GFP expression was not observed. Second, the model assumes an ideal condition where ELP-SF is 100% inactive when fully aggregated. This is likely not the case in a real system as some level of ELP-SF interaction with the P20_992 promoter

probably still occurs even when aggregated. Although the FFL predicts an improved dynamic range compared to the simple activation model, it is uncertain if a 35% reduction will be significant enough to allow for dynamic control in a production pathway. It may be necessary to explore additional networks options to increase the range even further. In the meantime, the physical FFL constructs will be built to determine the actual performance of the system *in vivo*.

4. CONCLUSIONS AND FUTURE DIRECTIONS

4.1 Summary of current progress and future work

This research aimed to expand the toolbox of stimuli-responsive control devices of gene expression through the development of novel, tunable ELP-based transcriptional regulators. This work presents the conceptual foundation for the use of these devices for the dynamic control of toxic intermediates within a bioproduction pathway. These devices encompass the hallmarks of an ideal controller in the following ways: 1) rational sequence design of ELP allows for behavior programmability, 2) the use of orthogonal transcription factor devices avoids unwanted interactions with the host's native processes, 3) the use of pH sensitive ELPs makes the controller responsive to changes in intracellular pH, and 4) the detection of a general indicator of cell stress allows for wider applicability. The potential of these devices to control gene expression in a programable, stimuli-dependent manner was demonstrated through the *in vivo* testing of simple network architectures and model simulations.

The first objective of this study was to design, build, and characterize a library of ELP constructs that exhibit pH sensitivity and phase transition under various conditions. To accomplish this, three ELP types were designed, ELPA, ELPB, and ELPN, at three different lengths, n = 40, 80, and 160, resulting in a total of nine constructs. Of the nine constructs, four were able to be successfully purified, including ELPA40, ELPA80, ELPA160, and ELPN40. Purification was achieved using a new organic solvent extraction method, which improves ELP yield and purity compared to traditional ITC. SDS-PAGE was used to confirm protein size, while phase transition behavior was confirmed via spectrophotometry by tracking changes in turbidity in response to temperature. For the remaining five constructs, more investigation is needed to determine whether the problem is a result of poor expression, protein toxicity, inclusion body formation, or inappropriate purification method and staining.

The phase transition behavior of ELP was characterized by *in vitro* and *in vivo* studies. For *in vitro* studies, purified ELP was resuspended to known concentrations and the T_t was determined

by tracking the change in absorbance over a range of temperature. A linear trend was observed for T_t in response to ELP concentration, indicating that T_t decreases as concentration increases. Additionally, a nonlinear, sigmoidal trend was observed for T_t in response to pH for the three ELPA constructs in which T_t decreases with decreasing pH. This behavior demonstrates that ELP functionality is retained after purification by organic solvent extraction. It remains to be seen, however, how those transition temperatures differ *in vivo* as compared to *in vitro*.

A new method to establish the T_t of ELP *in vivo* was explored in this study. Based on ELP's ability to form an ATPS upon aggregation in the cell, it was hypothesized that a change in temperature could drive ATPS formation. The temperature at which this change occurs could then be taken as the transition temperature. To explore this possibility, the behavior of a GFP-ELP180 fusion was observed under a microscope at a lower limit of 24°C and as high as 33°C. The fusion protein appeared to polarize to the ends of the cell at all temperatures, suggesting that the T_t was lower than what could be observed. Although ATPS formation was observed, more testing is needed to prove this hypothesis. Future work will involve purifying the remaining five ELP constructs and characterizing their transitions *in vitro*. Fusions of GFP to the ELP library will also be made for *in vivo* transition characterization studies, starting with the shorter ELPs that are predicted to have the highest transition temperatures.

The second objective of this study was to determine the feasibility of ELP-transcription factor fusions as dynamic sensor-regulator devices for gene expression. To do this, simple activation and repression architectures were designed that use fusions of ELP with SF or TetR to drive or repress the expression of a fluorescent reporter gene. These studies produced conflicting results; the simple activation architecture showed evidence for control, while the simple repression architecture did not. A preliminary study for the simple activation architecture was performed, which showed a significant 21% reduction in GFP expression in response to a temperature increase from 30°C to 40°C, indicating that ELP aggregation controls expression as expected. The simple repression architecture, on the other hand, did not behave as we anticipated, requiring further investigation. It is hypothesized that binding affinity may play a role and Gibbs free energy calculations will be used to determine whether the binding of affinity of TetR to DNA is more favorable than ELP aggregation.

In addition to the preliminary study, the simple activation architecture was modelled in MATLAB and was adjusted to reflect the same 21% reduction in expression as the experimental data. Since expression was only measured at 30°C and 40°C, the preliminary study will be repeated in the range between 30°C-40°C to see the effect temperature has on expression and if it is similar to what was predicted by the simple activation model. The simple activation architecture has yet to be tested using any of the ELPs from the designed library; therefore, SF will be fused to the ELP library and expression will be tested above and below the predicted transition temperatures. One fusion, SF-ELPA40, has already been cloned, but has not been tested yet. It is noted that the observed 21% reduction in expression is likely too modest to be used in a real production pathway and therefore an alternative feed-forward loop design was explored for enhanced dynamic range.

The third objective of this study was to explore alternative network architectures for enhanced controller performance and dynamic range. This was accomplished by the development of a feed-forward loop model, which was predicted to reduce gene expression by 35% at 40°C, thus showing an improved dynamic range and ultrasensitive response compared to the simple activation model. Future work will include cloning the physical FFL constructs and testing gene expression to validate the model. To simplify testing of the genetic circuits, control of gene expression will first be characterized *in vitro* using cell-free protein synthesis. This method of characterization will not only be faster and simpler but will also eliminate the possibility of protein inclusion body formation or the need to optimize growth conditions. If *in vitro* control of gene expression can first be proven, then *in vivo* testing can be explored.

4.2   ELP-transcription factor devices for dynamic pathway control

The long-term goal of this project is to implement the pH sensitive ELP-transcription factor sensor-regulators into a production pathway to show enhanced production via dynamic control of toxic intermediates. Eventually, the devices will be introduced into a heterologous mevalonate isoprenoid pathway to improve yields of taxadiene, a precursor for Taxol. Taxol, or paclitaxel, is a billion-dollar anticancer drug originally isolated from the Pacific yew tree (*Taxus brevifolia*) [62]. Its complex structure and size make chemical synthesis infeasible; less than 30 mg of the

compound have been produced by total synthesis [63]. As a result, current Taxol production has been dominated by biological methods, however, these methods still face challenges that make them difficult to scale. Production from plant cell-cultures has been explored but is hindered by low yields and high production costs. The current commercial supply of Taxol relies on semi-synthetic production from yew tree derived baccatin, making it dependent on the supply of yew tree materials. This reliance makes yields unpredictable as production is impacted by climate, region, tree age, sex, and other variables [64-66]. Thus, microbial-based production of Taxol and/or its precursors may be an attractive alternative to current production platforms as microbial cultures offer more predictability and could easily be scaled to meet worldwide demand. So far, the commercial scale production of Taxol or its precursors have yet to be achieved microbially despite the fact that researchers have been able to optimize expression of a heterologous mevalonate pathway in *E. coli* to make g/L titers of other isoprenoid-derived drugs [67, 68].
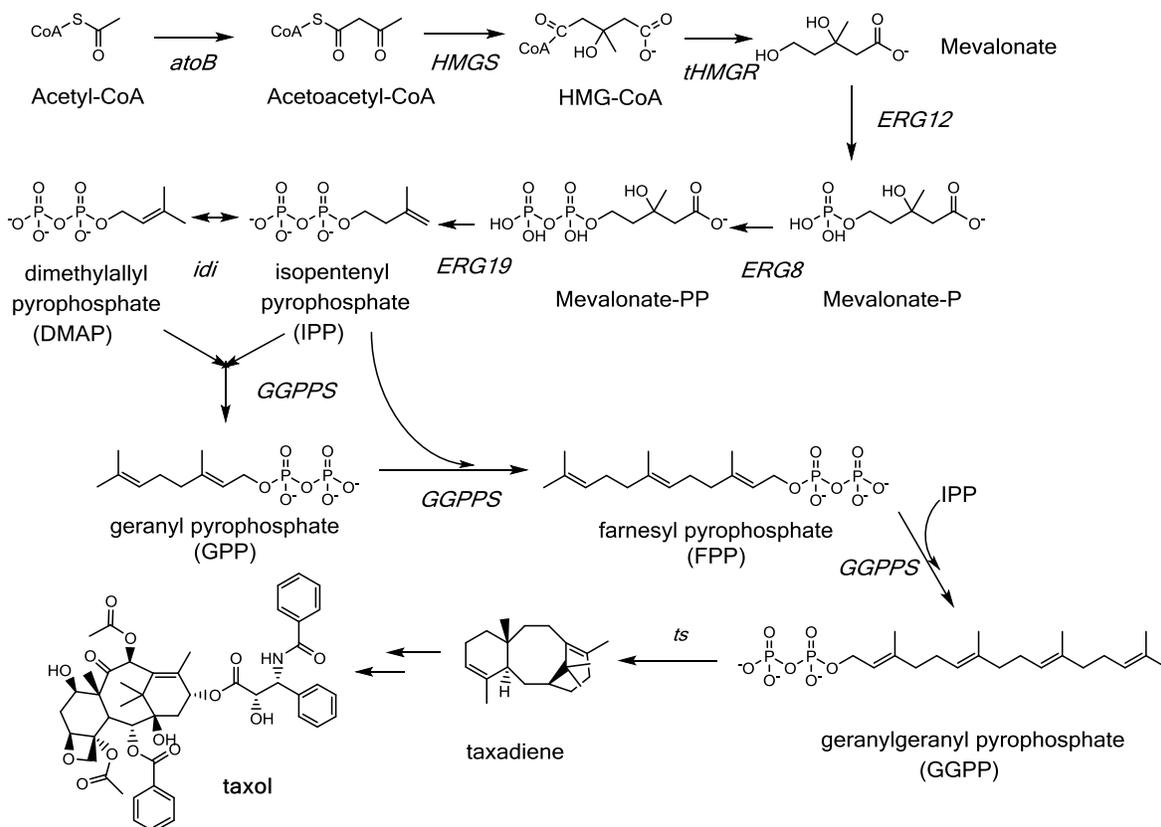

Figure 14. The proposed schema to taxadiene via the mevalonate isoprenoid pathway. Image credit: Dr. Kevin Solomon.

The first committed step to Taxol biosynthesis is the cyclization of geranylgeranyl diphosphate (GGPP) to taxadiene via a taxadiene synthase [69] (Figure 14); thus, any high yielding Taxol process must first produce high yields of the precursor taxadiene. For this reason, we chose taxadiene as the final product of our heterologous mevalonate isoprenoid pathway in *E. coli*, which is currently being constructed and optimized by another Solomon Lab member. So far, a combinatorial library of the first three genes through mevalonate in the pathway has been constructed in *E. coli* MG1655(DE3) and validated for mevalonate production (data not shown). Future work will entail cloning the remaining genes to taxadiene production.

We aim to improve taxadiene production through the dynamic control of farnesyl pyrophosphate (FPP), a known toxic pathway intermediate linked to the acid-stress response. Researchers have used GadE, an acid inducible transcription factor [70], as a sensor-regulator for FPP accumulation to control the expression of the mevalonate operon by a P_{GadE} promoter to improve amorphadiene titers twofold over static gene expression. This strain also showed improved cell growth and carbon utilization efficiency [67]. We speculate that the mechanism of toxicity for FPP is via dysregulation of intracellular pH. Therefore, it is predicted that FPP build-up will cause a drift in intracellular pH, which will then trigger the phase transition of the pH sensitive ELPs. Expression of FPP will be controlled by an ELP-transcription factor fusion protein, which once aggregated, will halt FPP production until pH returns to normal.

To determine how ELP transition behavior correlates with changes in intracellular pH, fusions of GFP to the pH sensitive ELPs will be made and expressed in *E. coli* DZ3, a proton transport mutant defective in pH homeostasis whose intracellular pH is correlated with extracellular pH [61]. These cells will be exposed to extracellular medium of varying pH, by titration of dilute NaOH or HCl, and then observed under a microscope to monitor aggregation. To measure the change in intracellular pH, a pHRED fluorescent protein whose excitation and emission spectra varies with pH [71], will be integrated into the genome of DZ3 and expressed at low constitutive levels. This will allow for the correlation of intracellular pH with ELP aggregation. Additionally, pHRED will be integrated into the genome of MG1655(DE3) and expressed along with the taxadiene pathway. The library, which produces variable amounts of mevalonate and thus toxic FPP, will be assayed for taxadiene production via GC-MS. Production can then be correlated

with intracellular pH as measured via pHRED. From these experiments, the appropriate SF-ELP controllers can be chosen for implementation into the taxadiene pathway. Finally, this will allow for the main hypothesis to be tested—whether SF-ELP regulatory devices can be used to dynamically control production in response to toxic intermediate accumulation for improved titers and productivity. A timeline outlining these experiments can be found in Figure 15.

Estimated Timelne for Project	2018 Q1 Q2 Q3 Q4	2019 Q1 Q2 Q3 Q4	2020 Q1 Q2 Q3 Q4
Characterize ELP library and extend dynamic range			
1) Troubleshoot expression/purification for remaining ELPs			
2) Make SF fusions to ELP library			
3) Clone FFL constructs			
4) Characterize fusion proteins *in vitro*			
5) Test dynamic range of systems			
Assess controller ability to respond to intracellular pH			
1) Make GFP fusions to pH sensitive ELP library			
2) Integrate pHRED in *E. coli* DZ3 mutant			
3) *In vivo* characterization of GFP-ELP with pHRED			
Establish cues for dynamic control in taxadiene pathway			
1) Finish pathway to taxadiene			
2) Integrate pHRED in MG1655(DE3)			
3) Assay taxadiene production and intracellular pH			
4) Assess SF-ELP ability to control toxicity and improve titer			

Figure 15. Estimated Timeline for Project.

APPENDIX A

Table 4. List of primers and oligonucleotides used in this study. The forward and reverse strands are represented by the beginning letter "f" or "r" respectively.

Number	ID	Sequence (5′ → 3′)
1	fREIIS	ctagaaataattttaaggaggagtacatatgggctactgataatgatcttcagc
2	rREIIS	tcgagctgaagatcattatcagtagcccatatgtactcctccttaaaattattt
3	fELPN	cgtgggtgttccgggcgttggtgtcccaggtgttggcgtaccgggcgttggtgttcctggtgttggcgtgccggg
4	rELPN	cggcacgccaacaccaggaacaccaacgcccggtacgccaacacctgggactccaacgcccggaacacccacgcc
5	fELPA	cgtgggcgttccgggtatcggtgttccgggtatcggtgttccgggtgaaggtgttccgggtatcggtgtgccggg
6	rELPA	cggcacaccgatacccggaacaccttcacccggaacaccgatacccggaacaccgatacccggaacgcccacgcc
7	fELPB	cgtgggtgttccgggccacggtgtcccaggtggcggcgtaccgggccacggtgttcctggtgctggcgtgccggg
8	rELPB	cggcacgccagcaccaggaacaccgtggcccggtacgccgccacctgggactccgtggcccggaacacccacgcc
9	fSF	agatcccgcatcttatgagcggcctggaagttctgtt
10	rSF	atcatatgttgcgggccaccgccacccggtttgcgacgaccgctca
11	fcp1TetR	gatgaatgatataataggaaagtactgttttgtttaactttaagaaggagatatacatatgatgtctagattagataaaagtaaagtgat
12	rXNTetR	tgatctcgaggctagcaccgccaccagacccactttcacatttaagtt
13	fcp2TetR	atcaagatctatcaccgagtttattcttgacacctgatgcgatgaatgatataataggaaagtactgttt
14	fCherry	gtcaagatctaaagaggagaaaggatctatggtgagcaagggcga
15	rCherry	tgacggatccttacttgtacagctcgtccatgc

Table 4. continued

Number	ID	Sequence (5′ → 3′)
16	fGFPELP	Actgcatatggcagtgagcaagggcgaggaggata
17	rGFPELP	tataggtaccccaccgcccttgtacagctcgtccat
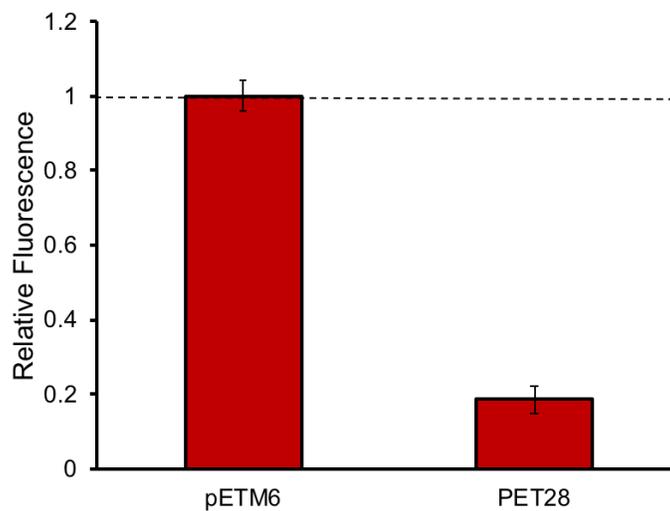
Figure 16. Expression of mCherry from pET28 was 81% lower compared to pETM6. Fluorescence intensity was first normalized by OD600, averaged, and then normalized by pETM6 average fluorescence.

ELPV24 **GST-ELP180** **ELPB**
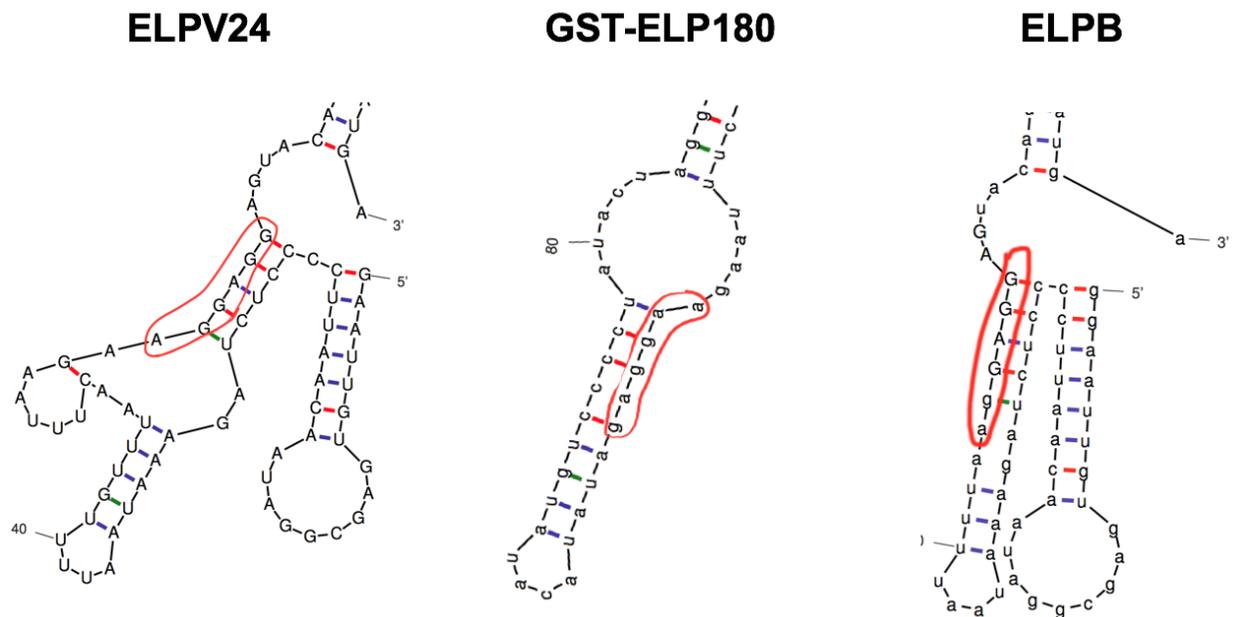
Figure 17. The mfold predictions of mRNA secondary structure at the RBS site for pET25b-ELPV24, pET32a-GST-ELP180, and pET32a-ELPB. The RBS sites [5′-AGGAGG-3′] and [5′-AAGGAG-3′] are circled in red.
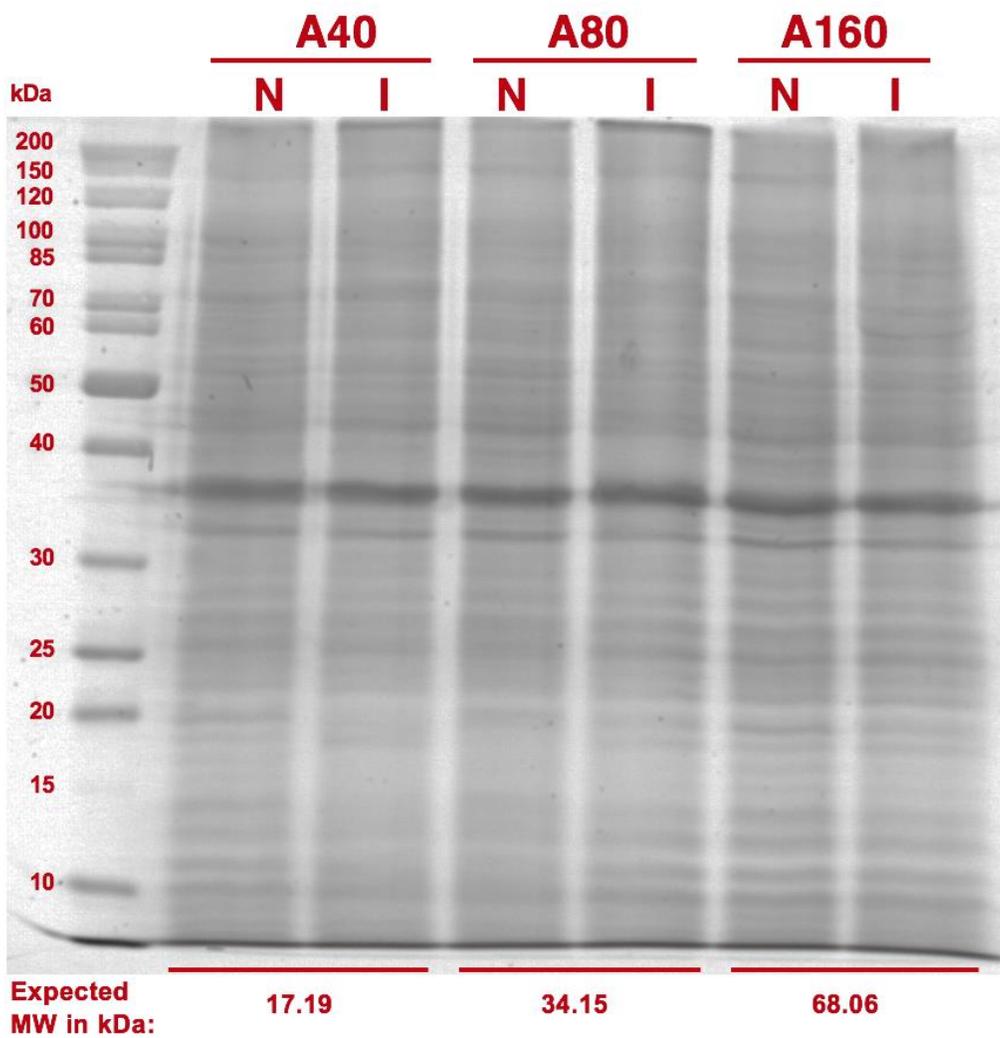
Figure 18. Expression check of induced (I) versus noninduced (N) samples of ELPA40, ELPA80, and ELPA160 using SDS-PAGE.
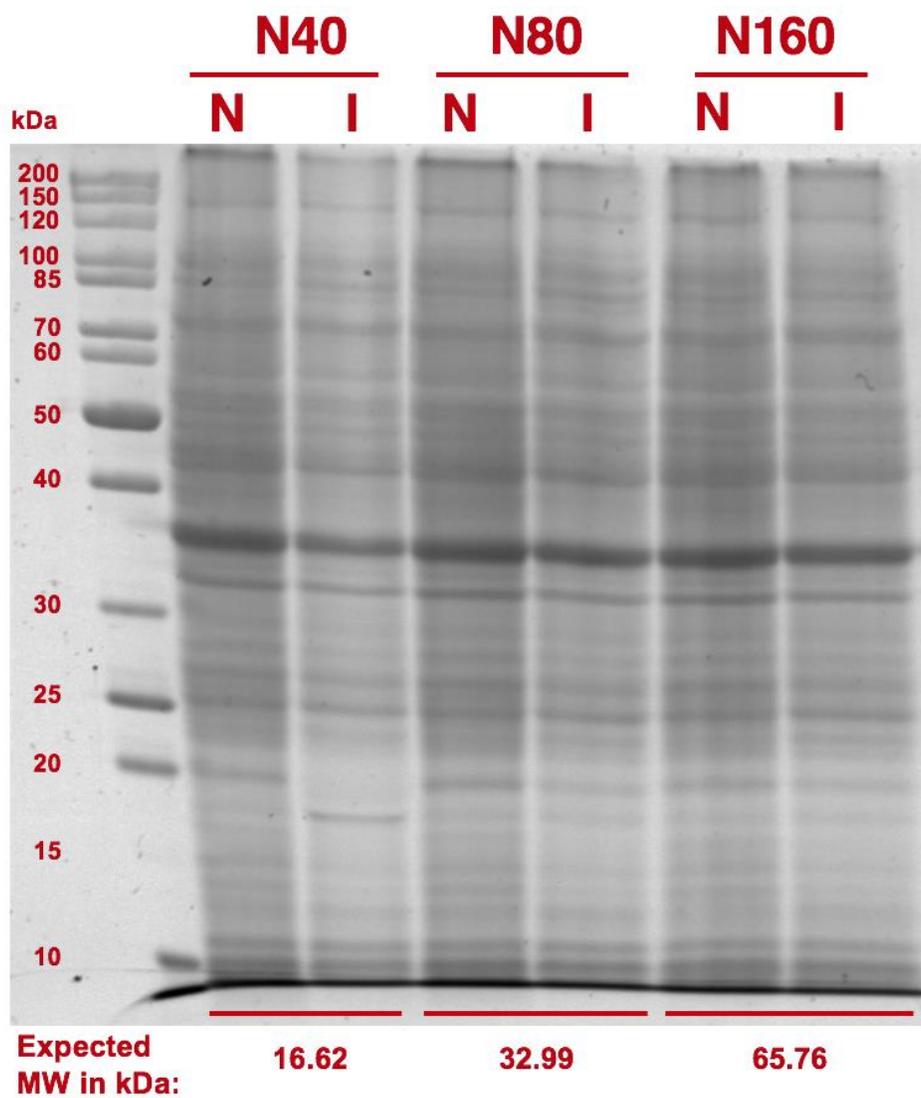
Figure 19. Expression check of induced (I) versus noninduced (N) samples of ELPN40, ELPN80, and ELPN160 using SDS-PAGE.
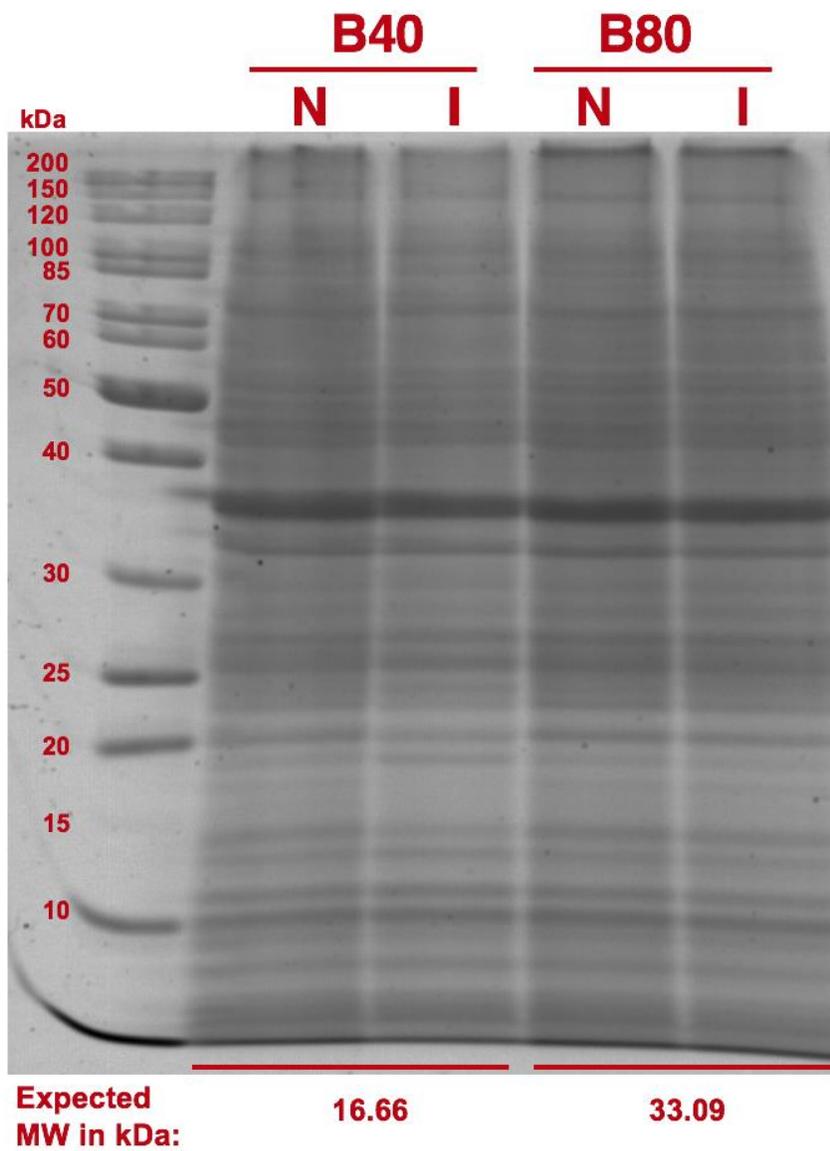
Figure 20. Expression check of induced (I) versus noninduced (N) samples of ELPB40 and ELPB80 using SDS-PAGE.
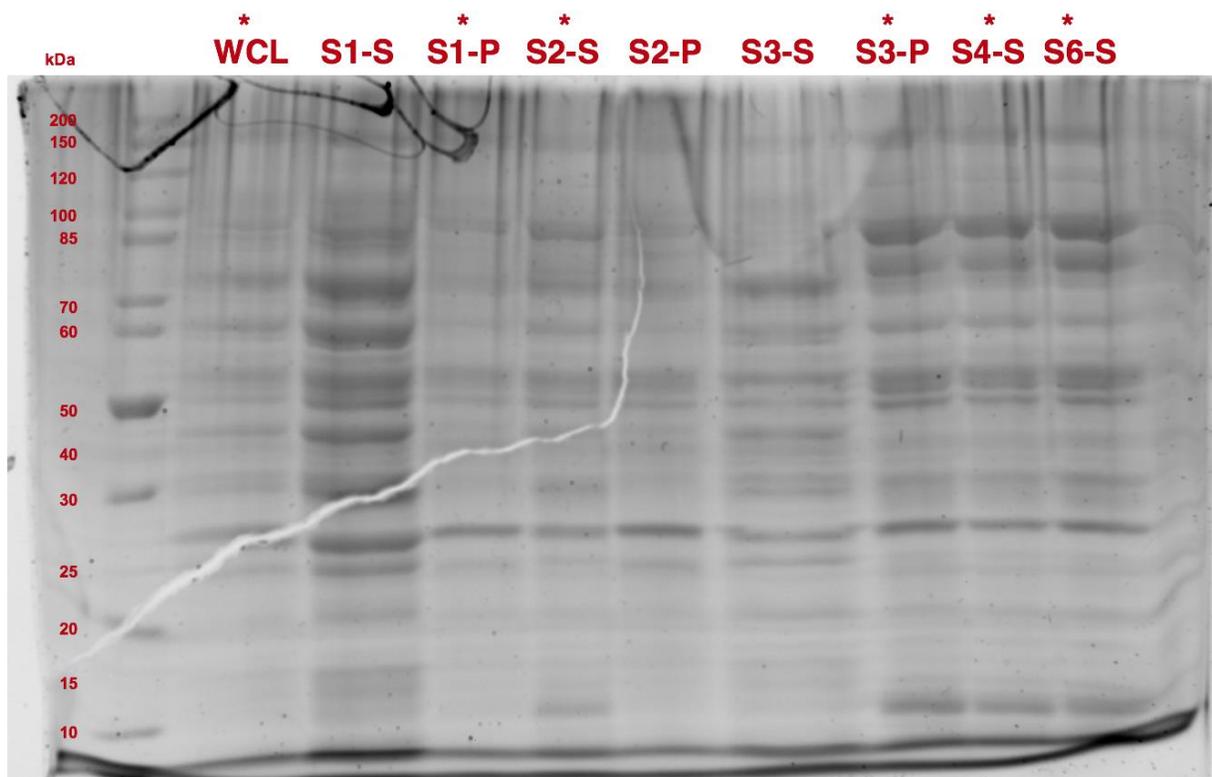
Figure 21. Expression check of ELPB160 using SDS-PAGE after multiple rounds of ITC. Odd numbers indicate hot spins, while even numbers indicate cold spins. The "-S" refers to supernatant, while "-P" refers to pellet. The * symbol indicates samples that should contain ELP protein. The predicted MW of ELPB160 is 65.94 kDa.
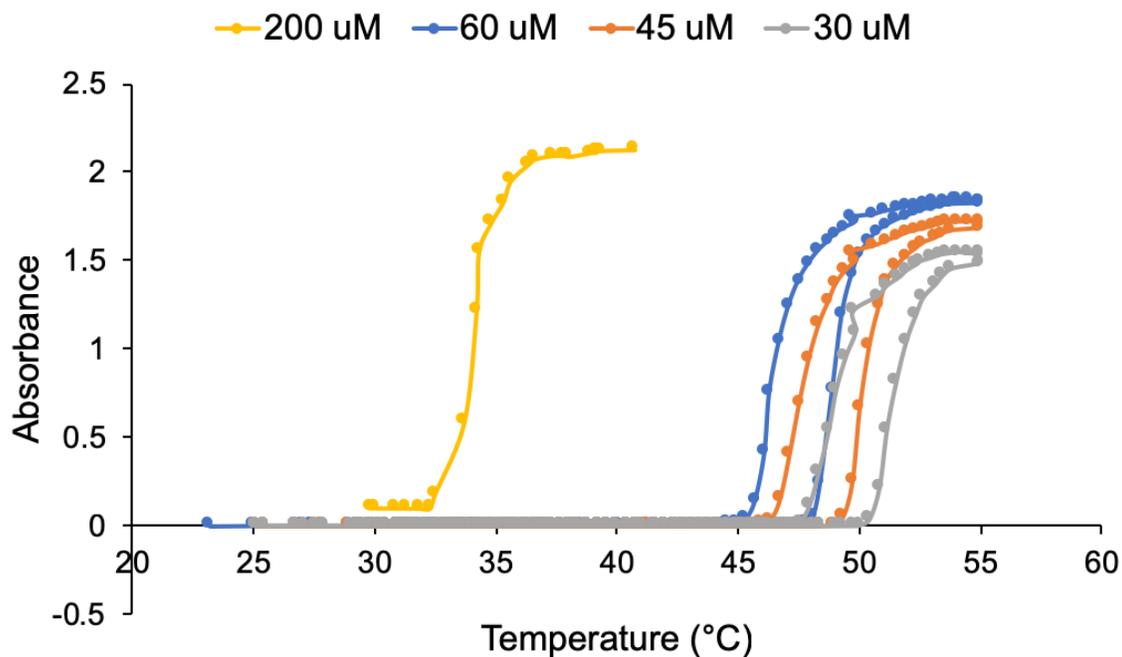
Figure 22. Absorbance versus temperature for ELPN40 at 30 µM, 45 µM, 60 µM, and 200 µM.
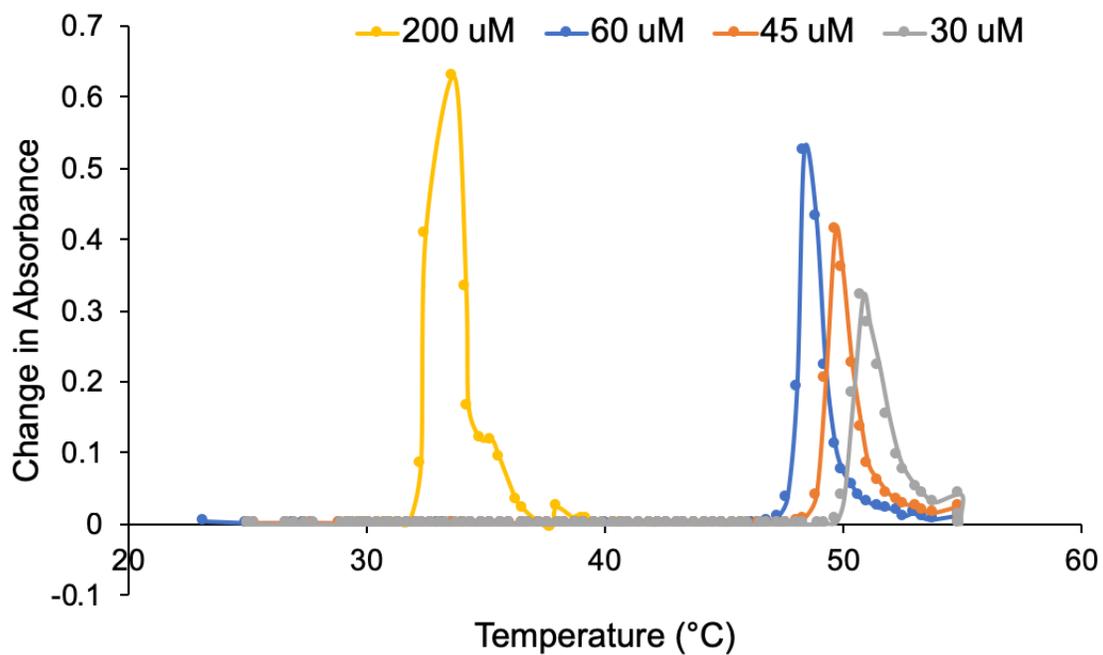
Figure 23. Change in absorbance versus temperature for ELPN40 at 30 µM, 45 µM, 60 µM, and 200 µM.
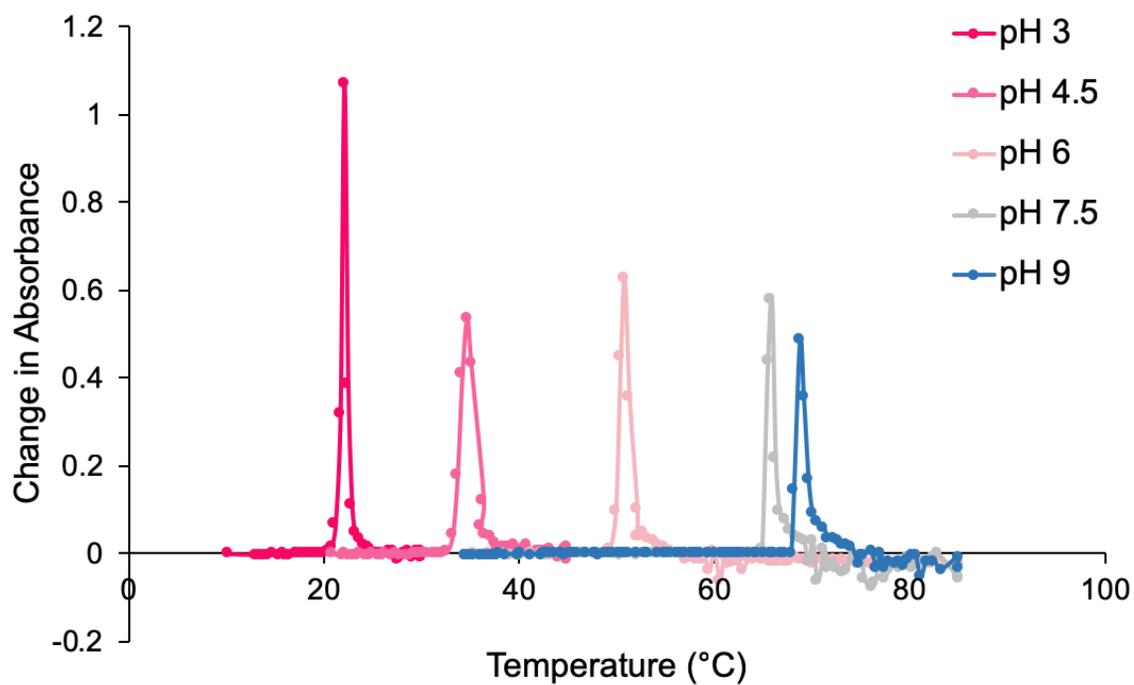
Figure 24. Change in absorbance versus temperature for 100 µM of ELPA80 at pH 3, 4.5, 6, 7.5, and 9.
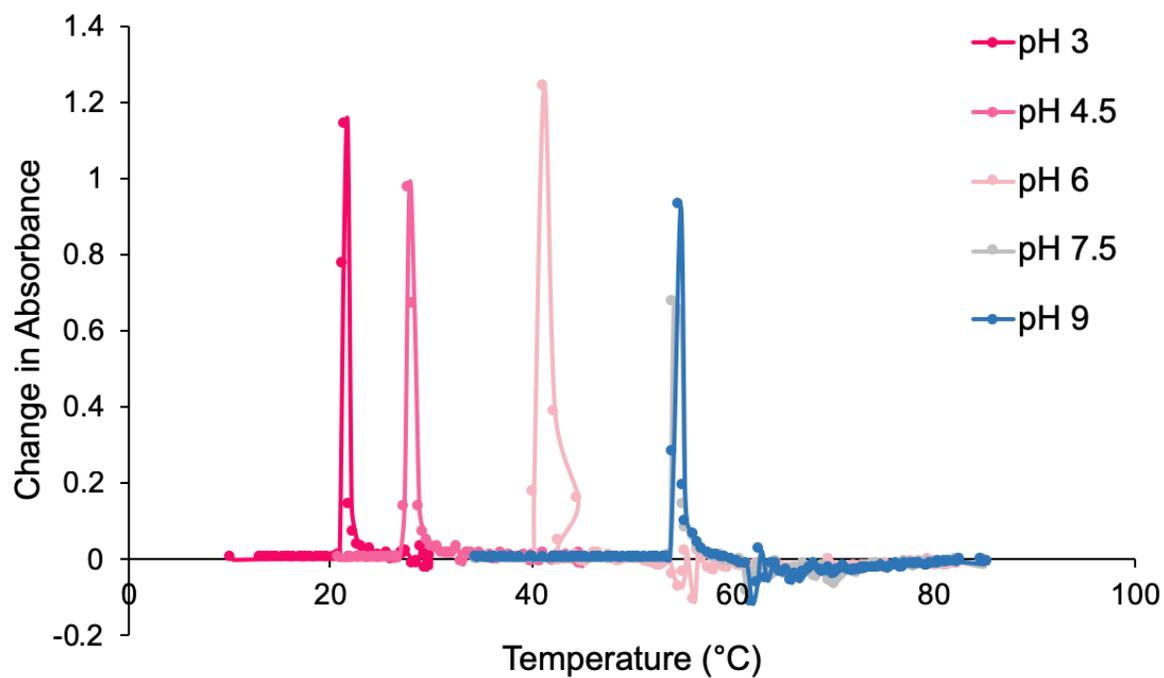
Figure 25. Change in absorbance versus temperature for 100 µM of ELPA160 at pH 3, 4.5, 6, 7.5, and 9.
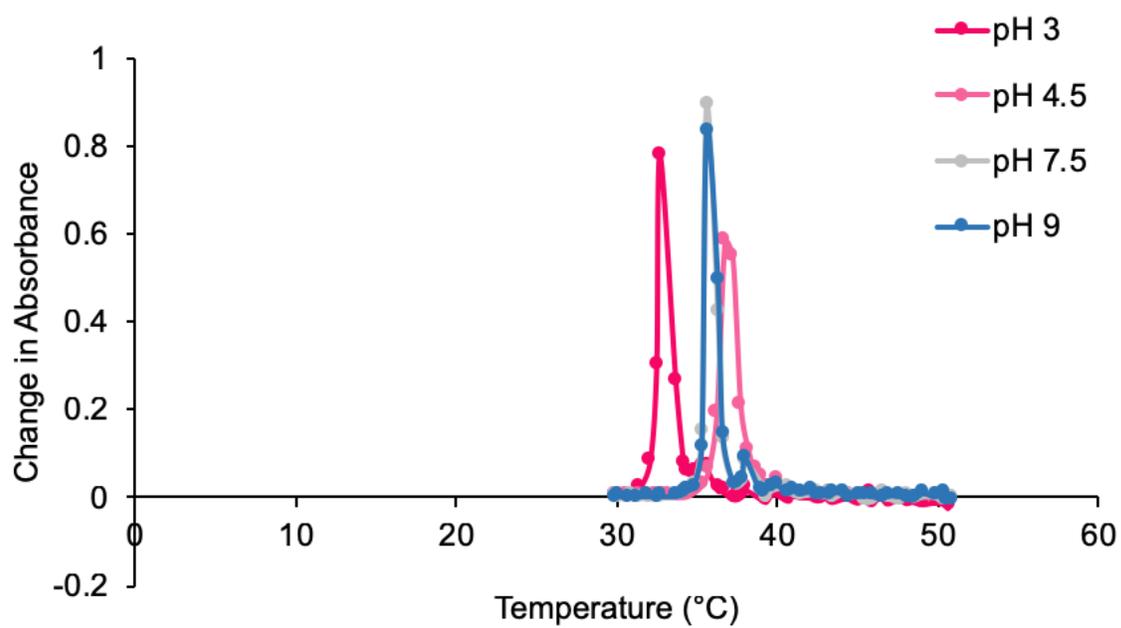
Figure 26. Change in absorbance versus temperature for 100 µM of ELPN40 at pH 3, 4.5, 7.5, and 9.

APPENDIX B

The following is the MATLAB code for both the simple activation and feed-forward loop

models:

```
% Script file: ELP_SF_Models_LR
%
% Purpose:
%   This program compares different architectures for gene expression
%   control by elastin-like polypeptide (ELP) fused to sigma factor (SF) in
%   response to temperature.
%
% Record of revisions:
%   Date        Programmer          Description of change
%   ====        ==========          ===========================
% 10/15/18    Logan Readnour      Original code
% 10/17/18    Logan Readnour      Defined parameters and moved estimation to
%                                 seperate code.
% 10/18/18    Logan Readnour      Made relative fluorescence graph and added
%                                 experimental data
% 10/23/18    Logan Readnour      Updated parameters
%
% Model Assumptions:
% - Assumes Quasi steady state.
% - protein degradation assumed negligible for stable proteins.
% - dilution due to growth assumed negligible for mRNA production due to
% differences in timescales.
% -ELPsf assumed 100% inactive when fully aggregated and 100% active when
% fully soluble.
%-------------------------------------------------------------------------
%-------------------------------------------------------------------------

clear all;

%Parameters for system:

T = 30:.2:45; % temperature range in C
Tt = 39; % Transition temperature of ELPsf
IPTG = 100; % Concentration of IPTG in uM
A = 1.3; % IPTG activation coefficient in uM (Setty et al. 2003)
n1 = 2; % Hill coefficient describing activation by IPTG (iGem website)
n2 = 1.2; % Hill coefficient describing activation by ELPsfA
n3 = 1; % Hill coefficient describing ORIB assembly/16S interaction with mRNA
k1 = 2.4; % Coefficient for steepness (predicted from data fit)
K = 300; % Translation rate of mRNA in 1/h/mRNA (Pai et al. 2009)
mu = 1.39; % Dilution due to growth in 1/h (estimated for doubling time of 30
min)
kd_m = 0.83; % mRNA degradation rate in 1/h (Kim et al. 2011)
beta1 = 0.15; % Max transcription rate from T7 promoter in uM/h (estimated
            % assuming 80 bp/s, see p. 111 notebook 3)
beta2 = 0.68; % Max transcription rate from P20_992 promoter in uM/h
(estimated
```

```matlab
                    % assuming 80 bp/s, see p. 111 notebook 3)
beta3 = 0.089; % Max transcription rate from P20_992 promoter in uM/h
(estimated
                    % assuming 80 bp/s, see p. 112 notebook 3)
Km1 = 1.1; % Activation coefficient in uM for ELPsfA
Km2 = 1; % Activation coefficient for ORIB assembly in uM


%------------------------------------------------------------------------

% The following equations solve for the fraction of active/soluble ELPsf:
% Total ELPsf present at steady state
ELPsfT = (K/mu)*(beta1/kd_m)*(IPTG^n1/(A^n1 + IPTG^n1));
% Uses logistic function to describe "S-shape" curve of ELP transition.
ELPsfI = ELPsfT .* (1./(1 + exp(-k1*(T-Tt))));
% Active ELPsfA can be found due to conservation
ELPsfA = ELPsfT - ELPsfI;

%Plots the active ELPsfA fraction
figure (2)
plot(T,ELPsfA,'Linewidth',1.5);
xlabel('Temperature (C)');
ylabel('Active/Soluble ELPsf (uM)');

%------------------------------------------------------------------------
% SIMPLE ACTIVATION EQUATIONS AT SS

% mRNA production at steady state.
SA_mRNA = (beta2/kd_m) .* (ELPsfA.^n2./(Km1^n2 + ELPsfA.^n2));
% GFP production at steady state.
SA_GFP = (K/mu)* SA_mRNA;

%Simple Activation plot GFP and mRNA versus temperature
figure(3)
subplot(1,2,1);
plot(T,SA_GFP,'Linewidth',1.5);
title('SA GFP');
xlabel('Temperature (C)');
ylabel('GFP (uM)');

subplot(1,2,2);
plot(T,SA_mRNA,'Linewidth',1.5);
title('SA mRNA');
xlabel('Temperature (C)');
ylabel('mRNA (uM)');

%------------------------------------------------------------------------
% FEED FORWARD LOOP (FFL) EQUATIONS AT SS

% o-ribosome production at steady state.
FFL_ORIB = (beta3/kd_m) .* (ELPsfA.^n2./(Km1^n2 + ELPsfA.^n2));
% mRNA production at steady state.
FFL_mRNA = (beta2/kd_m) .* (ELPsfA.^n2./(Km1^n2 + ELPsfA.^n2));
% GFP production at steady state.
FFL_GFP = (K/mu) .* ((FFL_ORIB.^n3)./(Km2.^n3 + FFL_ORIB.^n3)).*FFL_mRNA;
```

```matlab
% FFL Plots of GFP, mRNA, and ORIB versus Temperature
figure(4)
subplot(2,2,1);
plot(T,FFL_GFP,'Linewidth',1.5);
title('FFL GFP');
xlabel('Temperature (C)');
ylabel('GFP (uM)');

subplot(2,2,2);
plot(T,FFL_mRNA,'Linewidth',1.5);
title('FFL mRNA');
xlabel('Temperature (C)');
ylabel('mRNA (uM)');

subplot(2,2,3);
plot(T,FFL_ORIB,'Linewidth',1.5);
title('FFL ORIB');
xlabel('Temperature (C)');
ylabel('ORIB (uM)');

%--------------------------------------------------------------------
% The following plots relative GFP and compares the different
% architectures.

% Normalizes GFP concentration:
% Simple activation:
norm_SA_GFP = SA_GFP./SA_GFP(1,1);
% FFL:
norm_FFL_GFP = FFL_GFP./FFL_GFP(1,1);

% Experimental data (Relative fluorescence controlled by ELP180sf was
compared at
% 30C and 40C):
T_exp = [30, 40];
GFP_exp = [1, 0.79];

% Plots relative GFP:
figure(5);
plot(T,norm_SA_GFP,T,norm_FFL_GFP,'Linewidth',1.5);
hold on
%Plots observed data
plot(T_exp,GFP_exp,'o','markersize',8,'linewidth',1.5);
legend('Location','southwest');
legend('Simple Activation','FFL','Experimental');
xlabel('Temperature (C)');
ylabel('Relative Fluorescence');
set(gca,'fontsize',14);
```

REFERENCES

1.   Wu G, Yan Q, Jones JA, et al (2016) Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. Trends in Biotechnology 34:652–664. doi: 10.1016/j.tibtech.2016.02.010

2.   Keasling JD (2010) Manufacturing Molecules Through Metabolic Engineering. Science 330:1355–1358. doi: 10.1126/science.1193990

3.   Keasling JD (2012) Synthetic biology and the development of tools for metabolic engineering. Metabolic Engineering 14:189–195. doi: 10.1016/j.ymben.2012.01.004

4.   Paddon CJ, Westfall PJ, Pitera DJ, et al (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 496:528–532. doi: 10.1038/nature12051

5.   Yim H, Haselbeck R, Niu W, et al (2011) Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat Chem Biol 7:445–444. doi: 10.1038/nchembio.580

6.   Nakamura CE, Whited GM (2003) Metabolic engineering for the microbial production of 1,3-propanediol. Current Opinion in Biotechnology 14:454–459. doi: 10.1016/j.copbio.2003.08.005

7.   Chubukov V, Mukhopadhyay A, Petzold CJ, et al (2016) Synthetic and systems biology for microbial production of commodity chemicals. npj Syst Biol Appl 2:253–11. doi: 10.1038/npjsba.2016.9

8.   Van Dien S (2013) From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. Current Opinion in Biotechnology 24:1061–1068. doi: 10.1016/j.copbio.2013.03.002

9.   Woolston BM, Edgar S, Stephanopoulos G (2013) Metabolic Engineering: Past and Future. Annu Rev Chem Biomol Eng 4:259–288. doi: 10.1146/annurev-chembioeng-061312-103312

10.   Lee SY, Kim HU (2015) Systems strategies for developing industrial microbial strains. Nature Biotechnology 33:1061–1072. doi: 10.1038/nbt.3365

11.   Zhang F, Carothers JM, Keasling JD (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nature Biotechnology 30:354–359. doi: 10.1038/nbt.2149

12.   Glick BR (1995) Metabolic load and heterologous gene expression. Biotechnology Advances 13:247–261. doi: 10.1016/0734-9750(95)00004-A

13.   Leonard E, Ajikumar PK, Xiao W-H, et al (2010) Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. PNAS 107:13654–13659. doi: 10.1073/pnas.1006138107/-/DCSupplemental

14.   Holtz WJ, Keasling JD (2010) Engineering Static and Dynamic Control of Synthetic Pathways. Cell 140:19–23. doi: 10.1016/j.cell.2009.12.029

15.   De Mey M, Maertens J, Lequeux GJ, et al (2007) Construction and model-based analysis of a promoter library for E. coli: an indispensable tool for metabolic engineering. BMC Biotechnol 7:34–14. doi: 10.1186/1472-6750-7-34

16.   Cox RS, Surette MG, Elowitz MB (2007) Programming gene expression with combinatorial promoters. Mol Syst Biol 3:597–11. doi: 10.1038/msb4100187

17.   Hammer K, Mijakovic I, Jensen PR (2006) Synthetic promoter libraries – tuning of gene expression. Trends in Biotechnology 24:53–55. doi: 10.1016/j.tibtech.2005.12.003

18.   Smolke CD, Martin VJJ, Keasling JD (2001) Controlling the Metabolic Flux through the Carotenoid Pathway Using Directed mRNA Processing and Stabilization. Metabolic Engineering 3:313–321. doi: 10.1006/mben.2001.0194

19.   Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nature Biotechnology 27:946–950. doi: 10.1038/nbt.1568

20.   Jones JA, Toparlak ÖD, Koffas MA (2015) Metabolic pathway balancing and its role in the production of biofuels and chemicals. Current Opinion in Biotechnology 33:52–59. doi: 10.1016/j.copbio.2014.11.013

21.   Farmer WR, Liao JC (2000) Improving lycopene production in Escherichia coli by engineering metabolic control. Nature Biotechnology 18:533–537. doi: 10.1038/75398

22.   Xu P, Li L, Zhang F, et al (2014) Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. Proceedings of the National Academy of Sciences 111:11299–11304. doi: 10.1073/pnas.1406401111

23.   Dahl RH, Zhang F, Alonso-Gutierrez J, et al (2013) Engineering dynamic pathway regulation using stress-response promoters. Nature Biotechnology 31:1039–1046. doi: 10.1038/nbt.2689

24.   Kristensen JH, Karsdal MA (2016) Chapter 30 - Elastin. Biochemistry of Collagens, Laminins and Elastin 197–201. doi: 10.1016/B978-0-12-809847-9.00030-1

25.   Chilkoti A, Christensen T, MacKay A (2006) Stimulus responsive elastin biopolymers: applications in medicine and biotechnology. Current Opinion in Chemical Biology 10:652–657. doi: 10.1016/j.cbpa.2006.10.010

26.   Despanie J, Dhandhukia JP, Hamm-Alvarez SF, MacKay JA (2016) Elastin-like polypeptides: Therapeutic applications for an emerging class of nanomedicines. Journal of Controlled Release 240:93–108. doi: 10.1016/j.jconrel.2015.11.010

27.   Floss DM, Schallau K, Rose-John S, et al (2010) Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. Trends in Biotechnology 28:37–45. doi: 10.1016/j.tibtech.2009.10.004

28.   Kowalczyk T, Hnatuszko-Konka K, Gerszberg A, Kononowicz AK (2014) Elastin-like polypeptides as a promising family of genetically-engineered protein based polymers. World J Microbiol Biotechnol 30:2141–2152. doi: 10.1007/s11274-014-1649-5

29.   Li B, Alonso DOV, Bennion BJ, Daggett V (2001) Hydrophobic Hydration Is an Important Source of Elasticity in Elastin-Based Biopolymers. J Am Chem Soc 123:11991–11998. doi: 10.1021/ja010363e

30.   Li B, Alonso DOV, Daggett V (2001) The molecular basis for the inverse temperature transition of elastin. Journal of Molecular Biology 305:581–592. doi: 10.1006/jmbi.2000.4306

31.   Meyer DE, Chilkoti A (1999) Purification of recombinant proteins by fusion with thermally-responsive polypeptides. Nature Biotechnology 17:1112–1115. doi: doi.org/10.1038/15100

32.   MacKay JA, Callahan DJ, FitzGerald KN, Chilkoti A (2010) Quantitative Model of the Phase Behavior of Recombinant pH-Responsive Elastin-Like Polypeptides. Biomacromolecules 11:2873–2879. doi: 10.1021/bm100571j

33.   Christensen T, Hassouneh W, Trabbic-Carlson K, Chilkoti A (2013) Predicting Transition Temperatures of Elastin-Like Polypeptide Fusion Proteins. Biomacromolecules 14:1514–1519. doi: 10.1021/bm400167h

34.   Urry DW (1997) Physical Chemistry of Biological Free Energy Transduction As Demonstrated by Elastic Protein-Based Polymers. J Phys Chem B 101:11007–11028. doi: 10.1021/jp972167t

35.   MacKay JA, Callahan DJ, FitzGerald KN, Chilkoti A (2010) Quantitative Model of the Phase Behavior of Recombinant pH-Responsive Elastin-Like Polypeptides. Biomacromolecules 11:2873–2879. doi: 10.1021/bm100571j

36.   Girotti A, Reguera J, Arias FJ, et al (2004) Influence of the Molecular Weight on the Inverse Temperature Transition of a Model Genetically Engineered Elastin-like pH-Responsive Polymer. Macromolecules 37:3396–3400. doi: 10.1021/ma035603k

37.   Rhodius VA, Segall-Shapiro TH, Sharon BD, et al (2013) Design of orthogonal genetic switches based on a crosstalk map of σs, anti-σs, and promoters. Mol Syst Biol 9:1–13. doi: 10.1038/msb.2013.58

38. Bervoets I, Van Brempt M, Van Nerom K, et al (2018) A sigma factor toolbox for orthogonal gene expression in Escherichia coli. Nucleic Acids Research 335:308–12. doi: 10.1093/nar/gky010

39. Mangan S, Alon U (2011) Structure and function of the feed-forward loop network motif. Proceedings of the National Academy of Sciences 100:11980–11985. doi: 10.1073/pnas.2133841100

40. Mangan S, Zaslaver A, Alon U (2003) The Coherent Feedforward Loop Serves as a Sign-sensitive Delay Element in Transcription Networks. Journal of Molecular Biology 334:197–204. doi: 10.1016/j.jmb.2003.09.049

41. Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8:450–461. doi: 10.1038/nrg2102

42. Rao CV (2012) Expanding the synthetic biology toolbox: engineering orthogonal regulators of gene expression. Current Opinion in Biotechnology 23:689–694. doi: 10.1016/j.copbio.2011.12.015

43. Chubiz LM, Rao CV (2008) Computational design of orthogonal ribosomes. Nucleic Acids Research 36:4038–4046. doi: 10.1093/nar/gkn354

44. Rackham O, Chin JW (2005) A network of orthogonal ribosome·mRNA pairs. Nat Chem Biol 1:159–166. doi: 10.1038/nchembio719

45. Rackham O, Chin JW (2005) Cellular Logic with Orthogonal Ribosomes. J Am Chem Soc 127:17584–17585. doi: 10.1021/ja055338d

46. McDaniel JR, MacKay JA, Quiroz FG, Chilkoti A (2010) Recursive Directional Ligation by Plasmid Reconstruction Allows Rapid and Seamless Cloning of Oligomeric Genes. Biomacromolecules 11:944–952. doi: 10.1021/bm901387t

47. Liu JC, Heilshorn SC, Tirrell DA (2004) Comparative Cell Response to Artificial Extracellular Matrix Proteins Containing the RGD and CS5 Cell-Binding Domains. Biomacromolecules 5:497–504. doi: 10.1021/bm034340z

48. Christensen T, Amiram M, Dagher S, et al (2009) Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. Protein Science 18:1377–1387. doi: 10.1002/pro.157

49. VerHeul R, Sweet C, Thompson DH (2018) Rapid and simple purification of elastin-like polypeptides directly from whole cells and cell lysates by organic solvent extraction. Biomaterials Science 6:863–876. doi: 10.1039/c8bm00124c

50. Shaner NC, Lambert GG, Chammas A, et al (2013) A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. Nat Methods 10:407–409. doi: 10.1038/nmeth.2413

51.   Ge X, Conley AJ, Brandle JE, et al (2009) In Vivo Formation of Protein Based Aqueous Microcompartments. J Am Chem Soc 131:9094–9099. doi: 10.1021/ja902890r

52.   Reisch CR, Prather KLJ (2015) The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli. Scientific Reports 1–12. doi: 10.1038/srep15096

53.   Rajendran M, Claywell B, Haynes EP, et al (2018) Imaging pH Dynamics Simultaneously in Two Cellular Compartments Using a Ratiometric pH-Sensitive Mutant of mCherry. ACS Omega 3:9476–9486. doi: 10.1021/acsomega.8b00655

54.   Alon U (2006) An Introduction to Systems Biology. Chapman & Hall/CRC

55.   Setty Y, Mayo AE, Surette MG, Alon U (2003) Detailed map of a cis-regulatory input function. Proceedings of the National Academy of Sciences 100:7702–7707. doi: 10.1073/pnas.1230759100

56.   Iadevaia S, Mantzaris NV (2006) Genetic network driven control of PHBV copolymer composition. Journal of Biotechnology 122:99–121. doi: 10.1016/j.jbiotec.2005.08.030

57.   Haseong Kim, Gelenbe E (2011) Stochastic Gene Expression Modeling with Hill Function for Switch-Like Gene Responses. IEEE/ACM Trans Comput Biol and Bioinf 9:973–979. doi: 10.1109/TCBB.2011.153

58.   Pai A, You L (2009) Optimal tuning of bacterial sensing potential. Mol Syst Biol 5:1–11. doi: 10.1038/msb.2009.43

59.   Snoep JL, Westerhoff HV, Rohwer JM, Hofmeyr JHS (2006) Is there an optimal ribosome concentration for maximal protein production? IEE Proc Syst Biol 153:398–3. doi: 10.1049/ip-syb:20060023

60.   Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31:3406–3415. doi: 10.1093/nar/gkg595

61.   Zilberstein D, Agmon V, Schuldiner S, Padan E (1984) Escherichia coli Intracellular pH, Membrane Potential, and Cell Growth. J Bacteriol 158:246–252.

62.   Jacoby M (2005) Top Pharmaceuticals: Taxol. Chemical Engineering News

63.   Mendoza A, Ishihara Y, Baran PS (2011) Scalable enantioselective total synthesis of taxanes. Nature Chemistry 4:21–25. doi: 10.1038/nchem.1196

64.   Hao X, Pan J, Zhu X (2013) Taxol Producing Fungi. In: Ramawat KG, Mérillon J-M (eds) Natural Products: Phytochemistry, Botany and Metabolism of Alkaloids, Phenolics and Terpenes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 2797–2812

65.   Nadeem M (2002) Taxol content in the bark of Himalayan Yew in relation to tree age and sex. Phytochemistry 60:627–631. doi: 10.1016/S0031-9422(02)00115-2

66. Hook I, Poupat C, Ahond A, et al (1999) Seasonal variation of neutral and basic taxoid contents in shoots of European Yew (Taxus baccata). Phytochemistry 52:1041–1045. doi: 10.1016/S0031-9422(99)00264-2

67. Dahl RH, Zhang F, Alonso-Gutierrez J, et al (2013) Engineering dynamic pathway regulation using stress-response promoters. Nature Biotechnology 31:1039–1046. doi: 10.1038/nbt.2689

68. Martin VJJ, Pitera DJ, Withers ST, et al (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nature Biotechnology 21:796–802. doi: 10.1038/nbt833

69. Koepp AE, Hezari M, Zajicek J, et al (1995) Cyclization of Geranylgeranyl Diphosphate to Taxa-4(5),11(12)-diene Is the Committed Step of Taxol Bioysnthesis in Pacific Yew. The Journal of Biological Chemistry 270:8686–8690.

70. Hommais F (2004) GadE (YhiE): a novel activator involved in the response to acid environment in Escherichia coli. Microbiology 150:61–72. doi: 10.1099/mic.0.26659-0

71. Tantama M, Hung YP, Yellen G (2011) Imaging Intracellular pH in Live Cells with a Genetically Encoded Red Fluorescent Protein Sensor. J Am Chem Soc 133:10034–10037. doi: 10.1021/ja202902d