THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF CHEMICAL ENGINEERING

COMPUTATIONAL DESIGN AND EXPERIMENTAL VALIDATION OF RNA-BASED BIOSENSORS FOR THE DETECTION OF BIOMARKER PROTEINS

LIPIKA REDDY GADILA
SPRING 2018

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Chemical Engineering
with honors in Chemical Engineering

Reviewed and approved* by the following:

Howard Salis
Associate Professor of Chemical Engineering
Thesis Supervisor

Enrique Gomez
Associate Professor of Chemical Engineering
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

There is a need to detect medically relevant proteins such as biomarkers of human disease. Aptamers are nucleic acid-based recognition elements that have the ability to selectively bind to a large range of targets with high affinity, including proteins. Specifically, our objective is to engineer RNA-based riboswitches, which use aptamers to bind to a target ligand and thereby alter gene expression. Riboswitches can be potentially engineered as diagnostic sensors for protein targets. Compared to enzyme-linked immunosorbent assays (ELISAs) or surface plasmon resonance (SPR), riboswitch sensors have not been extensively tested for their ability to detect protein biomarkers of human disease. This thesis aims to validate a method of computationally designing riboswitches that can bind to and detect a specific protein of interest. Creating a riboswitch for a specific aptamer can be a slow process without computational modeling. As a result, the riboswitches in these experiments have been designed with the Riboswitch Calculator[1]. Riboswitches, were designed to detect three proteins of interest: MS2 coat protein, C-reactive protein (CRP), and Interleukin-32. Aptamers for these protein targets were identified and used to design the riboswitch sequences. The riboswitches and expression cassettes were tested in a coupled transcription-translation *in vitro* gene expression system known as TX-TL. Two designed riboswitches successfully detected the MS2 coat protein with statistically significant ($p < 0.05$) activation ratios (ARs) of 3.93 ± 2.18 and 4.90 ± 2.10. The riboswitches designed for CRP and IL-32 did not have statistically significant activation ratios in these experiments, due to insufficient folded protein levels within the TX-TL *in vitro* assay. The two successfully characterized riboswitches demonstrate that it is possible to computationally design a riboswitch to detect a specific protein in a cost-effective and efficient manner.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Howard Salis and Grace Vezeau for their guidance during this two year project. I have a new interest in synthetic biology because of them, and I know I have gained valuable skills. I have learned about a new promising field that seems to keep growing every year, and I feel humbled that I had the opportunity to work on an important project. Thank you to everyone in the Salis lab who has helped me and taught me what they know about this field.

Chapter 1

Introduction

Early detection of disease is a key aspect of medical progress. Biosensors, used in medicine and biotechnology, have been a useful tool for both detecting disease markers early and monitoring therapeutic progress[2]. These sensors have provided quick and quantitative diagnostic information to clinicians, allowing for better patient care during disease outbreaks[3].

A biosensor contains a recognition element and a signaling element to detect a specific molecule and display the recognition[2] with an optical, mechanical or electrical signal[4]. Current methods that can selectively detect these different biomarkers on a commercial scale are enzyme-linked immunosorbent assay (ELISA), surface plasmon resonance (SPR), and electroanalysis which all require some type of labeling. These recognition molecules can vary, but antibodies have been commonly used because of their high selectivity and affinity, characteristics that determine sensor performance[2]. However, antibody production and testing can take a very long time and is very expensive. Additionally, they are sensitive to temperature and are prone to irreversible denaturation[2].

As a result of these complications, other biological recognition elements are being considered for use in biomedical diagnosis. Aptamers are nucleic acid recognition elements that can selectively bind to a large range of targets with high affinity[4]. They are also more chemically stable than antibodies and have an indefinite shelf-life[5] and can be covalently modified to further improve their stability[6]. Some aptamer based biosensors, known as riboswitches, use aptamers to bind to target ligands, change shape, and alter gene expression[1]. Riboswitches can also easily be

synthesized *in vitro* and naturally work *in vivo* with high ligand specificity and affinity in response to a ligand[7].

There is a great interest in developing cheaper and more efficient aptamer-based biosensors for eventual use in the clinic. This thesis describes the computational design of riboswitches that can operate in an *in vitro* cell-free transcription-translation, suggesting an efficient way to create both functioning biosensors and cost-effective assaying methods.

Chapter 2

Current Use of Aptamer Based Biosensors

Recently, there has been extensive research of aptamers' role in biomedical diagnostics. Their small size, specificity, high stability, and high binding affinity have made them a strong candidate for biosensor development. Most aptamers are discovered through a method known as systematic evolution of ligands by exponential enrichment (SELEX) in which a random nucleic acid library is used for *in vitro* selection and amplification. A target ligand and the nucleic acid library bind, and the unbound DNA is filtered out, leaving the bound DNA behind for further amplification with PCR. SELEX provides researchers with numerous possible aptamer sequences for a specific ligand with high binding affinities (K_d)[2].

There are currently many different methods for designing aptamer- based biosensors. Aptamers change their shape once they bind to a target ligand so they can be tagged with fluorophores to identify if sensing is improved with an external quencher[2]. Two aptamers can even be combined in order to recognize a target and give a signal[2]. The designed aptamers can then be used as sensing elements in multiple different kinds of assays.

> **ELONA[8]:** This common method is used to measure protein levels in biological samples. In this test, a capture antibody will bind to and concentrate the protein biomarker onto a chip. Then a detection antibody, possibly tagged with a fluorescent dye, is used to detect the protein. Aptamers can replace the capture antibody in ELISA[2].

Surface Plasmon Resonance (SPR)[9]: SPR is a label free surface technique that detects molecules. A light is shone between a conductive metal and a dielectric which measures a refractive index (RI). After a ligand binds to the metal, the RI increase is measured. SPR can use DNA or RNA aptamers with nanoparticles to create a signal, but it is difficult for small molecules to detect a large signal with this method.

Electrochemical[10]: Electrochemical aptamer-based (E-AB) sensor are used to sense targets without the need to consider chemical reactivity. Once the sensor is included using voltammetry, the aptamer's conformational change is used to change the efficiency as a redox reporter at one end of the aptamer approaches an electrode conjugated to the other end of the electrode.

Colorimetric[11]: Detection of ligands using a colorimetric method involves a change in the color of the bound DNA-protein complex, generally with the use of gold nanoparticles (AuNP).

These methods of testing biosensors are only a few of many different tests. However, there has not been as much research on engineered riboswitch based sensors compared to these common methods. There are some natural riboswitches that can respond to cellular metabolites and cofactors in prokaryotes and eukaryotes, suggesting that riboswitches can detect ligands with specificity and sensitivity[12]. However, current methods of designing a riboswitch for a specific aptamer can be slow and inefficient. For example, a riboswitch was engineered to activate protein translation in response to 2,4-dinitrotoluene by creating a random plasmid library[13]. However, it is important to consider more efficient ways to design riboswitches for a specific aptamer.

Chapter 3

Computational Modeling to Design Riboswitches and Proteins

A large component of creating more functioning riboswitches involves making the design process more efficient. Mainly, researchers have used qualitative design, combinatorial library generation, and high throughput screening to design riboswitches, which has slowed down the process of creating more riboswitches. These methods for designing riboswitches could have been more efficient if there was a computational model for riboswitch design.

The Riboswitch Calculator is a model developed by Borujeni et al. to predict the sequence-structure function relationship for riboswitches that can activate gene expression within *in vitro* or *in vivo* assays[1]. Additionally, the model computationally designs different riboswitches and explains how aptamer structure, ligand affinity, switching free energy and macromolecular crowding affect riboswitch activation. Aptamer sequences are necessary for the Riboswitch Calculator to create the riboswitch sequence. The chosen aptamers (sequences in Appendix A), MS2 aptamer[14], C-reactive aptamer[15], and Interleukin-32 aptamer[16] sequences were obtained from previously published literature and were used to create the necessary riboswitch sequences. The Riboswitch Calculator model is described below[1].

The biophysical model considers the ribosome's binding free energy to the mRNA without and with a ligand. In the absence of a ligand, the mRNA folds into a structure with a minimum amount of free energy. It then binds to the ribosome with free energy in the off state ($\Delta G_{total,OFF}$). This total energy is calculated with the terms shown in Equation 1.

(Eq. 1) $$\Delta G_{total,OFF} = \Delta G_{mRNA:rRNA} + \Delta G_{start} + \Delta G_{spacing} + \Delta G_{standby} - \Delta G_{mRNA}$$

The $\Delta G_{mRNA-rRNA}$ free energy is the energy released when the Shine-Dalgarno (SD) sequence binds to the last nine nucleotides of the 16s rRNA. The energetic penalty due to a ribosome stretching or compressing is represented by the variable $\Delta G_{spacing}$. The ΔG_{start} variable is the energy released when a start codon is bound by a tRNA. There is also an energetic penalty by the single-stranded surface area of upstream standby sites represented by the variable $\Delta G_{standby}$. Finally, these free energies are subtracted by the energy needed to unfold inhibitory mRNA structures, or ΔG_{mRNA}. For the on state, the ligand is now bound to the riboswitch changing the variable $\Delta G_{mRNA:rRNA}$ into $\Delta G_{mRNA:ligand:rRNA}$. In this state the mRNA-ligand complex has stabilized itself with the free mRNA folded into its minimum free energy structure. From the free energies, the change in the riboswitch's translation initiation rate (TIR) can be found. The TIR is a rate limiting step in protein production so modeling this component can have a huge effect on the riboswitch's overall function. The relationship of the TIR (r) to the ΔG_{total} is shown in Equation 2[17].

(Eq. 2) $$r \propto \exp(-\beta \Delta G_{total})$$

The model also predicts the maximum increase in the riboswitch's translation initiation rate, or maximum activation ratio (AR_{max}), shown in Equation 3.

(Eq. 3) $$AR_{max} = \frac{r_{ON}}{r_{OFF}} = \exp(-\beta[\Delta G_{total,ON} - \Delta G_{total,OFF}])$$

The AR can be affected by ligand and mRNA concentrations. The AR_{conc} equation, displayed in Equation 4, is calculated using the volume fractions and molar volumes of mRNA and ligand.

(Eq. 4) $$AR_{conc} = \frac{\upsilon_{mRNA,1}}{\upsilon_{mRNA,total}} + \frac{V_{mRNA}}{V_{ligand}}\left[\frac{\upsilon_{ligand,total} - \upsilon_{ligand,free}}{\upsilon_{mRNA,total}}\right] AR_{max}$$

The actual activation ratio, shown in Equation 5, is then calculated by taking the difference in free energy between the mRNA folded without the ligand and the alternate minimum free energy structure free energies ($\Delta\Delta G_{mRNA}$). The Boltzmann factor, β, and the AR at a specific ligand concentration, AR_{conc}, is also included in the calculation.

(Eq. 5) $$AR_{actual} = 1 + \frac{\exp(-\beta(\Delta\Delta G_{mRNA} + \Delta G_{ligand}))}{1+\exp(-\beta(\Delta\Delta G_{mRNA} + \Delta G_{ligand}))})(AR_{conc} - 1)$$

These thermodynamic characteristics should be considered when choosing a riboswitch with optimal activation ratios. The switching free energy ($\Delta\Delta G_{mRNA} + \Delta G_{ligand}$) must be negative in order for the bound ligand and mRNA to be thermodynamically favorable and stabilized. Additionally, choosing riboswitches with high predicted ARs may be a more efficient way of getting high actual ARs during experimentation.

The Riboswitch Calculator outputs many riboswitch sequences with varying predicted ARs, TIRs, and predicted free energies of interactions with a 30s ribosome. The model assumes that chemical equilibrium is reached faster than the mRNA's half-life, which means that the ligand must quickly bind to the mRNA. Additionally, the mRNA must quickly change shape after a ligand is bound. This model doesn't account for changes in mRNA stability due to the change in its conformation.

Using computational modeling can significantly shorten the time it takes to design riboswitches to achieve a specific activation ratio. Researchers can also use model-predicted TIRs in the on and off states in order to achieve a desired expression level of protein. The Riboswitch Calculator was used in this experiment to design riboswitches at varying TIRs.

Chapter 4

TX-TL *in vivo* System

Almost all genetic engineering is done *in vivo*. Developing pathway controls and genetically- encoded biosensors are examples of genetic engineering work that can be done *in vivo*. As mentioned earlier, current assaying methods are very expensive and time consuming. A large part of these problems involves the use of cells, which, although very useful and necessary for large scale production, can be unnecessary for small scale assays. Using *in vitro* methods can be extremely advantageous by increasing the flexibility to engineer[18] and providing more control over testing parameters, more quantitative results, and a larger parameter space that can be studied[19]. Additionally, cell free expression systems do not contain cell walls or membranes[20].

Recently, cell free expression systems have been able to mimic an *in vivo* environment and eliminate the problems mentioned earlier. One method, a coupled transcription-translation system called TX-TL, can produce equivalent amounts of protein as T7-based systems with a 98% cost reduction compared to similar systems[21]. This is an *Escherichia coli*-based cell-free expression system that is used for producing proteins and testing genetic circuits in a controlled environment.

This system contains three essential components: crude cell extract, amino acid solution, and energy solution[21]. The amino acid and energy solutions are essential to make the TX-TL buffer which provides a stable pH and substrate supply[20] while the crude cell extract provides the proteins required for the expression of protein from DNA. This expression system also contains Mg and K-glutamate for increased efficiency. A single batch can produce about 3000 reactions[21]

and was enough for the purposes of this experiment. TX-TL is ideal for prototyping and can be a cost effective and efficient way of discovering more medically relevant riboswitches.

Chapter 5

Results

MS2 Tests

The activation ratio of five MS2 riboswitches (M701, M718, M957, M958, M959) with varying TIRs (Appendix A) were found in the presence of an expression cassette producing the MS2 coat protein (2nM). The no aptamer control corrected fluorescence value in the presence of the protein was divided by the fluorescence value in the absence of the protein to provide the AR values of each riboswitch, as displayed in Figure 1. The low activation ratio (below 1) of the M957, M958, and M959 riboswitches suggests that the riboswitches were not able to turn on RFP production in the presence of the MS2 coat protein. As a result, these three riboswitches were not considered for varied expression cassette concentration testing. Two riboswitches, M701 and M718, had activation ratios of 1.58±0.46 and 1.59±0.68 respectively, suggesting they would be most successful in a test with varying MS2 expression cassette concentrations (TIR in Appendix B). The M701 and M718 riboswitches were tested with concentrations of MS2 expression cassette ranging from 1nM- 5nM, with 1 nM increments to determine if increasing the expression cassette concentration will lead to an increase in MS2 protein concentration within the TX-TL assay, resulting in a higher activation ratio.

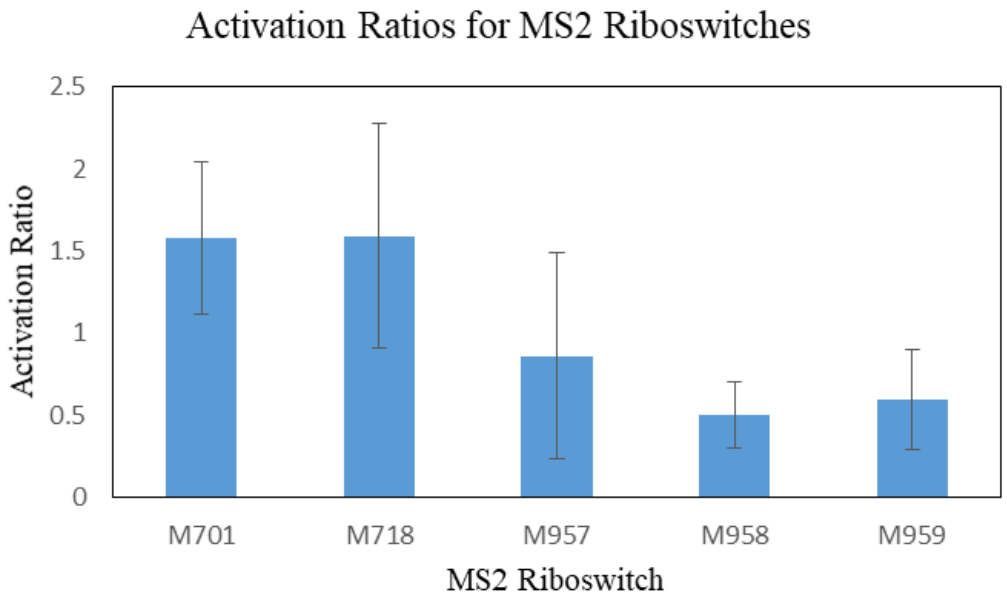

Figure 1: Activation ratios of five MS2 riboswitches (M701, M718, M957, M958, M959) were found in the presence of a 2nM expression cassette for MS2 coat protein.

The M701 and M718 riboswitch activation ratios increased as the expression cassette concentration increased, as shown in Figure 2 and Figure 3. The statistically significant ($p < 0.05$) activation ratios from 1nM-5nM expression cassette concentrations suggest that the riboswitches were successfully activating RFP production. The statistically significant maximum activation ratios of M701 and M718, 3.93 ± 2.18 and 4.90 ± 2.10 respectively, were obtained with a 5nM expression cassette concentration. The M718 riboswitch had higher activation ratio values when compared to the M701 riboswitch.

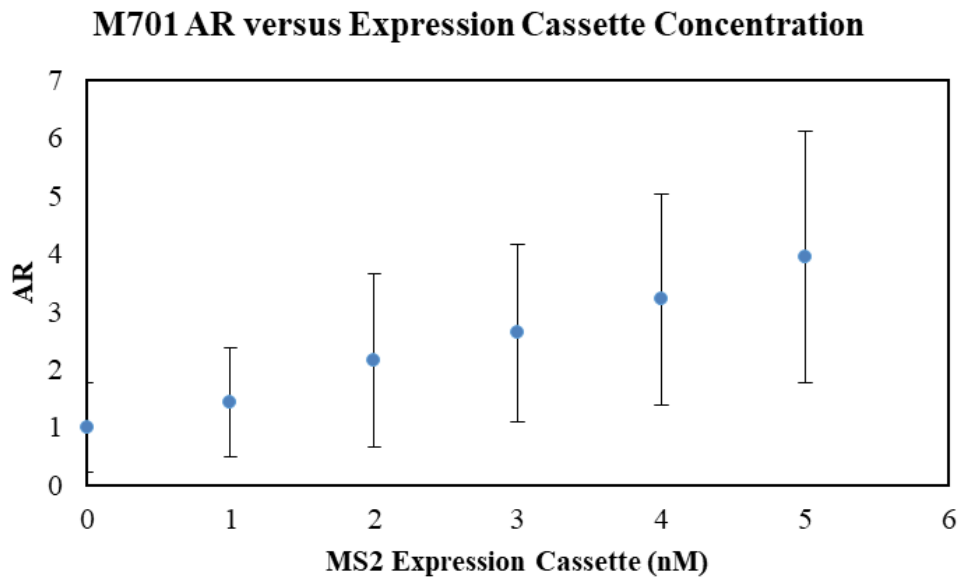

Figure 2: The activation ratios of the M701 riboswitch were found in the presence of varying concentrations (0nM, 1nM, 2nM, 3nM, 4nM, 5nM) of the expression cassette expressing MS2 coat protein. The highest activation ratio was obtained at a 5nM expression cassette concentration.

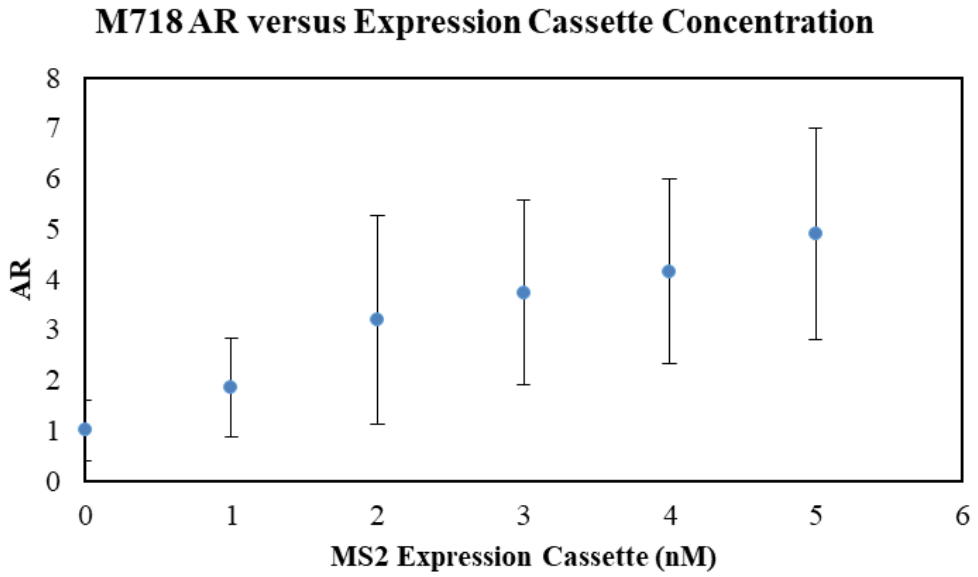

Figure 3: The activation ratios of the M718 riboswitch were found in the presence of the expression cassette at varying concentrations (0nM, 1nM, 2nM, 3nM, 4nM, 5nM). The highest activation ratio was obtained at a 5nM expression cassette concentration.

C-Reactive Protein Tests

The activation ratio of seven CRP riboswitches (TIRs in Appendix A) were found at varying calcium levels (0mM, 2mM) and at a 2nM CRP expression cassette concentration (TIR in Appendix B), as shown in Figure 4. Calcium was added in order to promote folding of the protein, which has several binding sites for calcium ions[22]. The fluorescence values of three CRP riboswitches (C684, C693, C726) were measured on a TECAN m1000 microplate reader. Figure 4 displays that the activation ratios of the three riboswitches were below or equal to one, suggesting that the tests were not optimized. The highest activation ratio, (0.98±2.34), resulted from the C684 riboswitch with a calcium concentration of 2mM. Additionally, the added calcium did not consistently aid in increasing the AR.

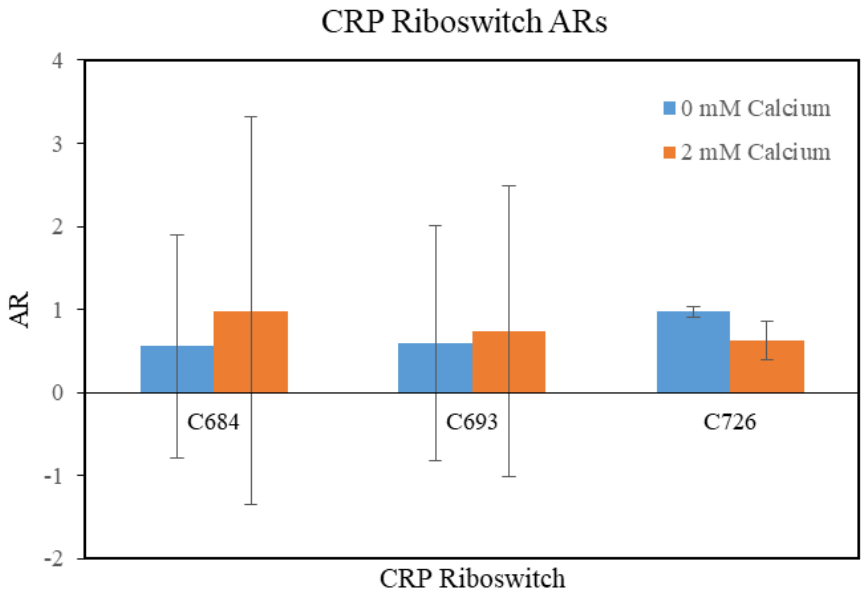

Figure 4: The activation ratios of three CRP riboswitches (C684, C693, C726) were found in the presence of a 2nM expression cassette expressing C-reactive protein. The highest activation ratio was found with the C684 riboswitch.

The fluorescence values of the next four CRP riboswitches (C624, C634, C642, and C648) were recorded with a SPARK microplate reader. The activation ratios, as shown in Figure 5, were slightly higher than those of the first three CRP riboswitches. The highest activation ratio, 1.34±0.07, was found with the C624 riboswitch at a calcium concentration of 2mM. Two of the seven riboswitches that had the highest activation ratios, C684 and C624, were utilized in an experiment with varying CRP protein concentrations.

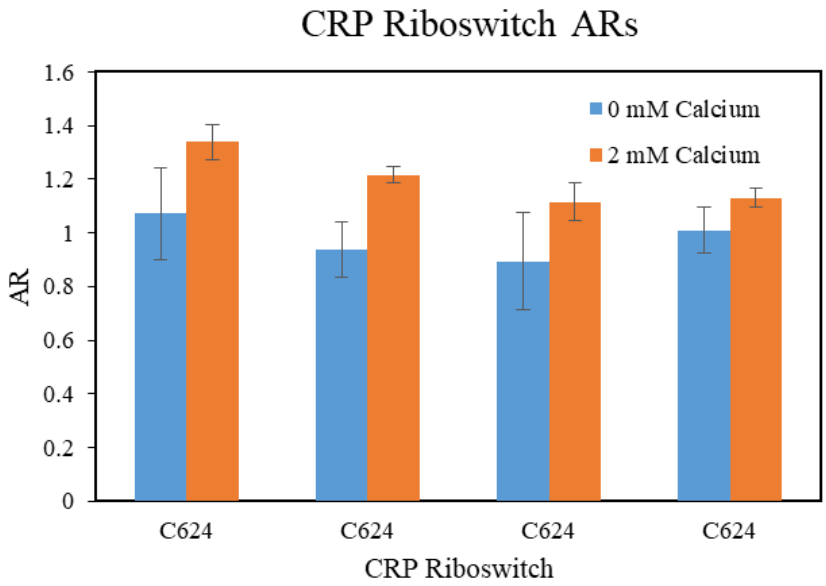

Figure 5: The activation ratios of an additional four CRP riboswitches (C624, C634, C642, C648) were found in the presence of a 2nM expression cassette expressing C-reactive protein. The highest activation ratio was obtained with C624.

The C624 and C684 riboswitches were both tested with varying concentrations of calcium (0mM, 2mM) and CRP expression cassette (0nM, 2nM, 3nM, 4nM, 5nM). The activation ratios of the C624 riboswitch at varying CRP expression cassette concentrations and 0mM calcium are shown in Figure 6. The highest activation ratio, 3.10± 8.26, resulted from a CRP expression cassette concentration of 3nM. The riboswitch was also tested with 2mM

calcium and the new ARs were found, as shown in Figure 7. The highest AR obtained at this calcium concentration was 1.46±2.98. The C624 riboswitch showed no clear trend of activation ratio with increasing protein concentration. Additionally, the 0mM calcium tests had better fluorescence values than the 2mM calcium tests. With the lack of a clear trend, there is no clear evidence that the C624 riboswitch turns on expression of RFP.

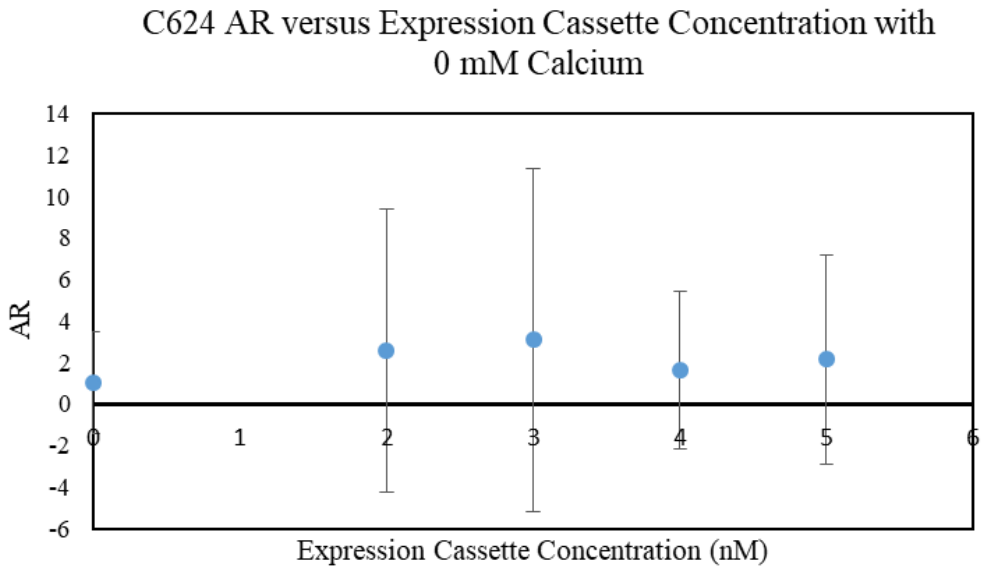

Figure 6: The activation ratios of the C624 riboswitch were found with varying concentrations (0nM, 2 nM, 3nM, 4nM, 5nM) of the expression cassette expressing C-reactive protein and with 0mM calcium. The highest activation ratio was obtained with a 2nM concentration.

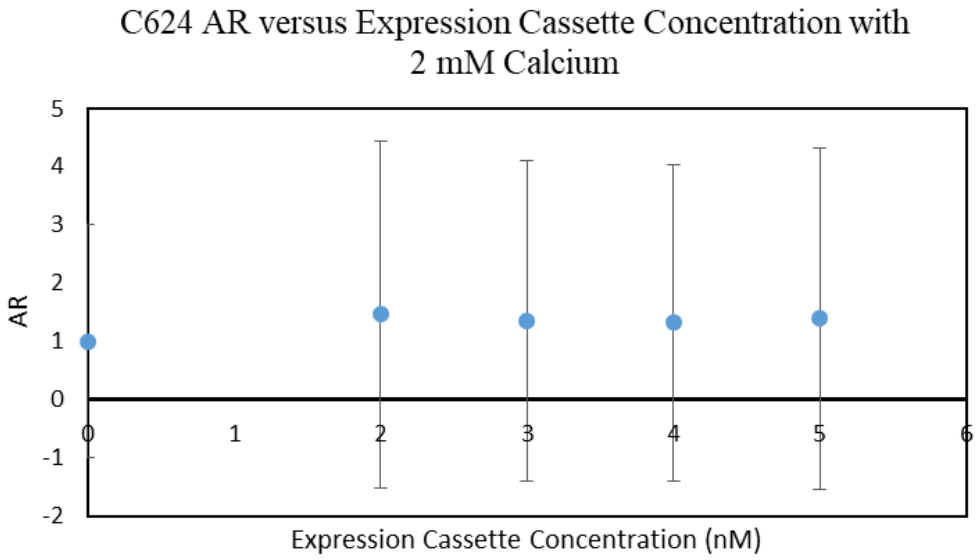

Figure 7: The activation ratios of C624 riboswitch were found in the presence of varying expression cassette concentrations (0nM, 2nM, 3nM, 4nM, 5nM) and with 2mM calcium. The highest activation ratio was obtained with a 2nM expression cassette.

The highest activation ratio for C684 in Figure 8, 2.90±7.70, resulted from a 2nM CRP expression cassette concentration with 0mM calcium. There is no clear trend of activation ratio with increasing expression cassette concentration. The highest activation ratio found with the presence of 2mM calcium was 1.12±2.83, as shown in Figure 9. The experiments with no calcium resulted in higher expression of RFP when compared to the addition of 2mM calcium suggesting that the calcium repressed the protein's RFP expression. Again, with no clear trend of activation ratios with increasing expression cassette concentrations, it is not clear that C684 riboswitch is functional.

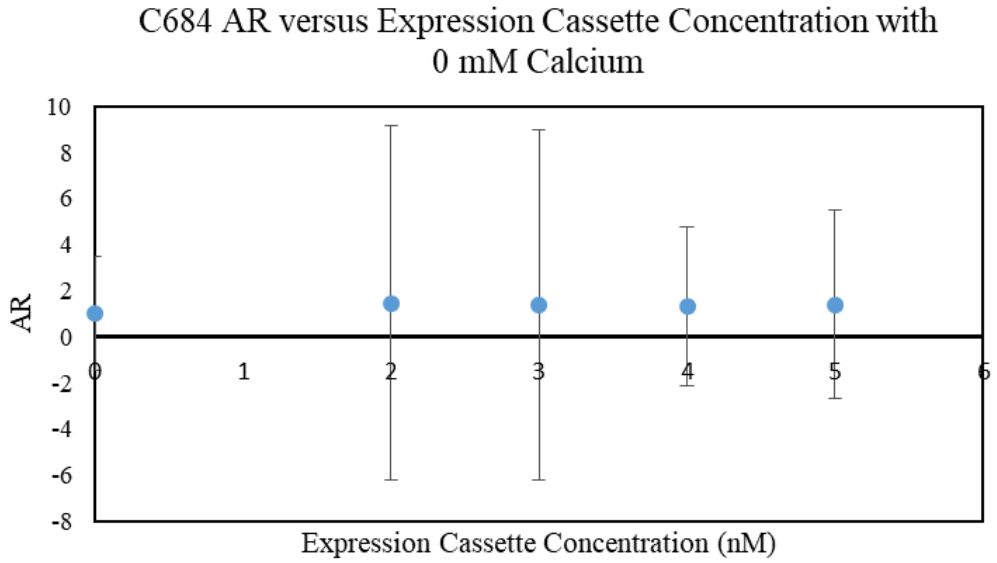

Figure 8: The activation ratios of C684 riboswitch were found in the presence of varying expression cassette concentrations (0nM, 2nM, 3nM, 4nM, 5nM) and with 0mM calcium. The highest activation ratio was obtained at 2nM expression cassette.

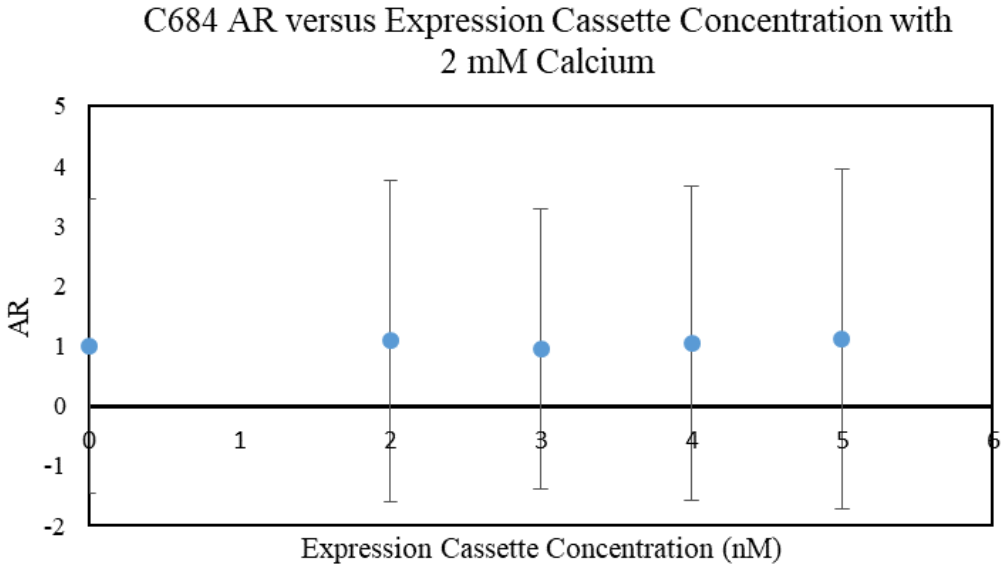

Figure 9: The activation ratios of C684 riboswitch were found in the presence of varying concentrations of the expression cassette expressing C-reactive protein with added 2mM calcium. The highest activation ratio was obtained with a 5nM expression cassette concentration.

Interleukin-32 Tests

Five IL-32 riboswitch plasmids (I949, I972, I1000, I1003, I1054) at varying TIRs (Appendix A) were tested in the presence of 2nM IL-32 protein-expressing plasmid (TIR in Appendix B), as shown in Figure 10. The highest ARs of, 1.61±0.50 and 1.51±0.4, were obtained by the I1003 and I1000 riboswitches, respectively. As a result, these riboswitches were further experimented in a test varying IL-32 expression cassette concentrations.

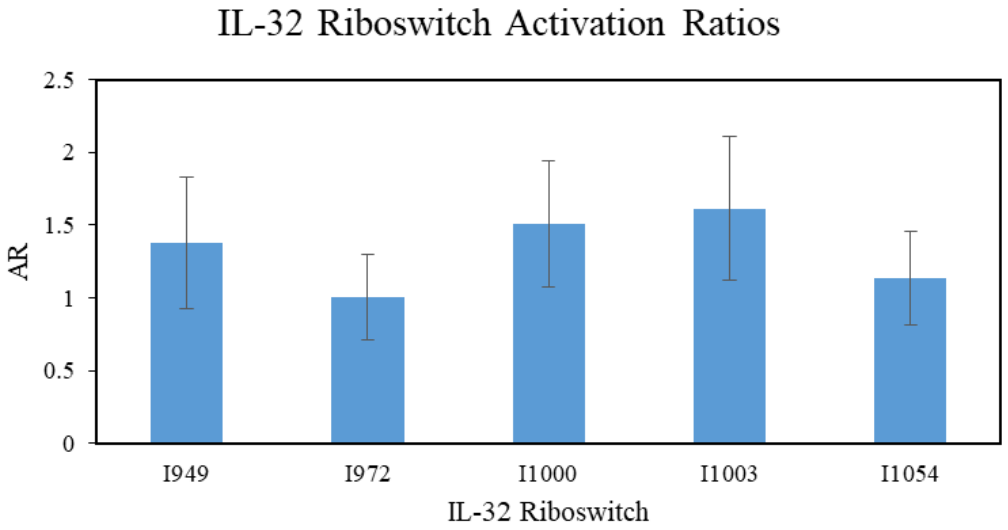

Figure 10: The activation ratios of five IL-32 riboswitches (I949, I972, I1000, I1003, I1054) were found in the presence of a 2nM expression cassette expressing the IL-32 protein. The highest activation ratios were obtained with the I1000 and I1003 riboswitches.

The I1000 and I1003 riboswitches were tested at IL-32 expression cassette concentrations varying from 1nM-5nM, with 1 nM increments. The ARs were then found in order to evaluate how protein concentration affects the riboswitches. The maximum ARs for I1000 and I1003 were 1.97±3.95 and 1.27±2.55, respectively. The results, displayed in Figure 11 and Figure 12, show no clear trend with AR and increasing expression cassette concentration. There is no clear

trend with the designed IL-32 riboswitches for activation ratios and increasing expression cassette concentrations, suggesting that it is not clear if the riboswitches are functional.

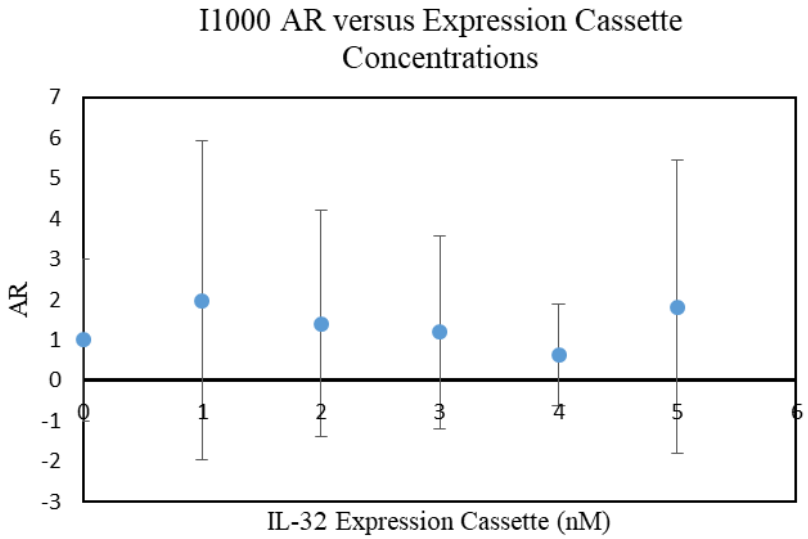

Figure 11: The activation ratio of I1000 riboswitch were found in the presence of varying expression cassette concentrations (0nM, 1nM, 2nM, 3nM, 4nM, 5nM). The highest activation ratio was obtained with a 1nM expression cassette.

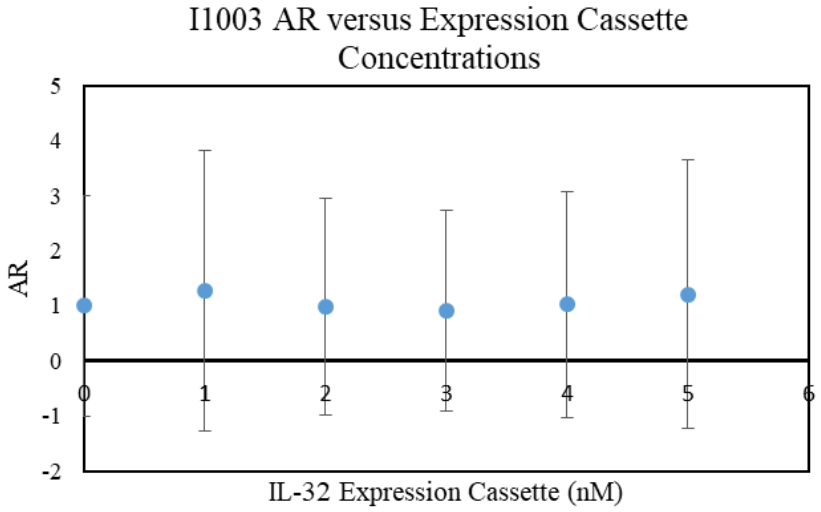

Figure 12: The activation ratios of I1003 riboswitch were found in the presence of varying concentrations (0nM, 1nM, 2nM, 3nM, 4nM, 5nM) of the expression cassette expressing the IL-32 protein. The highest activation ratio was obtained with a 1nM expression cassette.

Chapter 6

Discussion

The chosen proteins tested, with the exception of MS2 coat protein, had medical relevance. MS2 coat protein is known to repress translation of the phase replicase gene[23]. The MS2 coat protein aptamer was used as a control to test if a computationally designed riboswitch for an aptamer with known functionality could control the expression of a desired protein in an *in vitro* transcription-translation system. C-reactive protein is a clinical biomarker of inflammation and tissue damage[22], making it extremely relevant for current diagnostics. CRP can be detected with methods such as ELISA and SPR[22], but there has been little research on expressing CRP inside TX-TL and detecting their presence with riboswitches within *in vitro* tests. The Interleukin-32 cytokine has also been related to many inflammatory disorders such as rheumatoid arthritis. It is known to act as a defense against pathogens like influenza and HIV, making it an important protein to study[24]. Currently, there is little research about riboswitch sensors that can detect this protein.

Some of the riboswitches tested were able to express their respective proteins better than others. The difference in their expression levels can be explained by using the model of riboswitch regulation to predict the energetics of different riboswitch and protein interactions. As mentioned in Chapter 3, the stability of the riboswitch bound to the ligand can determine if it is thermodynamically favorable. In order for a stable ligand:mRNA complex to form, the switching free energy should be negative. The more negative the free energy, the more stable the complex. Additionally, the riboswitch will easily be able to switch from an off to an on state. The

transcription, translation, and degradation rates also affect the total protein production, and therefore the ARs. Decreased efficiency of these three components can reduce the number of ribosomes that can complete translation. The expression level of the protein target inside the TX-TL assay is an important value[1]. The CRP and IL-32 expression cassettes utilizes ribosome binding sites with low translation initiation rates of 477 and 179.1 respectively, causing their protein expression levels inside the TX-TL assay to be rather low. Comparably, the MS2 expression cassette used a ribosome binding site with a high translation initiation rate, and therefore was expressed at a much higher level. As a result, more MS2 coat protein was synthesized in the TX-TL reaction, allowing for the MS2 riboswitches to easily detect the presence of proteins. However, the lower protein expression levels from the CRP and IL-32 expression cassettes likely caused the riboswitches to have low ARs. The low protein production may also explain why adding calcium didn't increase the AR values of the CRP riboswitches tested.

The Riboswitch Calculator model predictions were compared to the experimental values. As mentioned in Chapter 3, the translation initiation rate is related to ΔG_{tot} as shown in Equation 2. The natural log of the measured mRFP1 protein expression level in the OFF state is plotted against the ΔG_{tot} where β is the slope and ΔG_{tot} is represented by x shown in Figure 13.

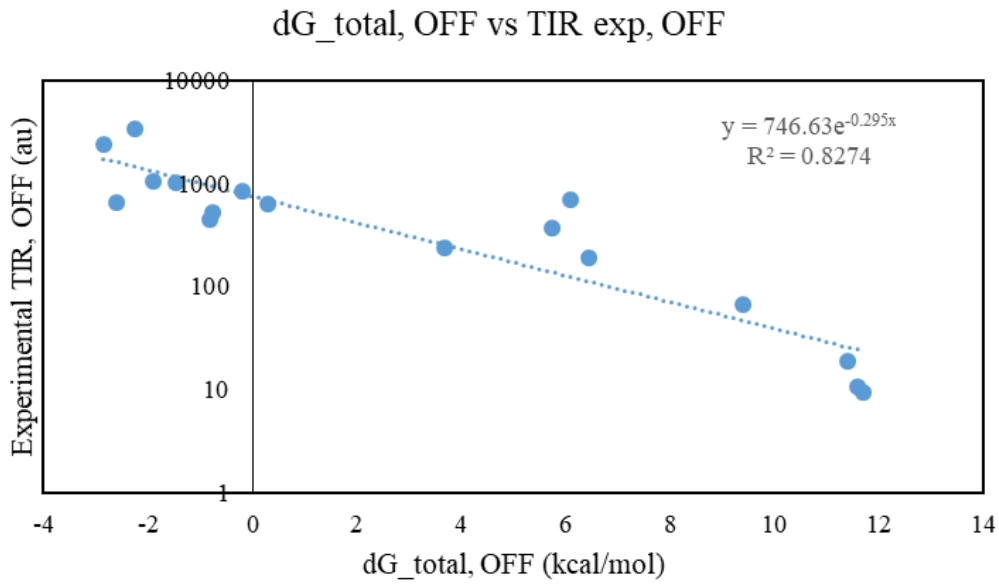

Figure 13: The log scale of the experimental value was plotted against the model predicted dG_total, OFF values for all the riboswitches tested. A slope of -0.295 was obtained with a R2 value of 0.83.

After plotting the experimental TIRs of the five MS2 riboswitches, seven CRP riboswitches, and five IL-32 riboswitches, a slope of -0.295 was obtained with a R^2 value of 0.8274. The slope represents the apparent thermodynamic beta. The value of β can be calculated with the Boltzmann constant (0.001987 kcal/(mol*K)) and the reaction temperature from Equation 6.

(Eq. 6) $$\beta = \frac{1}{k_B T}$$

At a reaction temperature of 302.15 K, the calculated Beta value is 1.67 which is higher than the model predicted slope. However, this calculation is assuming an ideal solution which is not the case. The model predicted value considers a compressibility factor that can be used to measure the deviation from ideality as shown in Equation 7.

(Eq. 7) $$\beta = \frac{1}{Zk_BT}$$

The compressibility factor is calculated to be 5.65. The difference could be due to changing mRNA levels over time during the TX-TL tests. As a result, the actual translation rate will be different than the one predicted by the model.

Additionally, the compressibility factor considers the volume fraction (φ) of a host. At a volume fraction of 0.4 (efficiency lowers as the degree of crowding is increased above this value[25]), the resulting β will be closer to the slope value. However, this calculation is assuming there is a large amount of macromolecular crowding, which is not a valid assumption for the TX-TL system since it is a dilute system.

The TX-TL system was useful for creating a controlled *Escherichia coli* based environment to test the riboswitches. TX-TL systems, as mentioned in Chapter 4, have many benefits. Additionally, this *in vitro* system doesn't run into problems dealing with cellular environments as *in vivo* systems do which allows for faster optimization during each experiment. Most proteins cannot pass through cellular membranes, and therefore could not be detected by a riboswitch expressed inside the cell. The CRP experiment, for example, was manipulated with the addition of calcium. In an *in vivo* system, it would be difficult to calculate how much calcium goes into the cell, which would add more time in experimenting. However, in the TX-TL system, calcium was a necessary component that could easily be added to the mix and directly bind with the protein. However, there are a few challenges with using a cell free expression system. For example there may be some batch to batch variability for the crude cell extract, amino acid solution and energy solution. Additionally, protein yields may vary based on different strains of bacteria[21]. Although TX-TL has these limitations, the effectiveness of this system to model an *in*

vivo environment allows experimenters to easily change experimental variables and quickly obtain results.

Expressing heterologous proteins in an *Escherichia coli* based system presents some challenges. The expression of enzymes in *Escherichia coli* is affected by the promoter sequence, the ribosome binding site's translation initiation rate, and plasmid copy number[26]. Heterologous protein expression in Escherichia coli based system also changes based on protein folding[26], so properly cloning the expression cassettes is important. The CRP protein was difficult to clone possibly because there has been research showing CRP is insoluble when expressed in *Escherichia coli*[27]. Cloning a CRP expression cassette with a high TIR value also created many mutations during cloning. As a result of this, a RBS library was necessary in order to efficiently clone the expression cassettes at varying TIRs. This method was also used for IL-32. The expression cassettes, which had no mutations, had lower TIR values. However, as mentioned earlier, a lower TIR created another problem of lower protein production, resulting in lower ARs. If the TIR isn't optimal, not enough proteins will be made, and the engineered riboswitches cannot be used to sense the proteins.

Chapter 7

Conclusion

The MS2 coat protein, C-reactive protein, and Interleukin-32 protein were tested to determine if computationally designed riboswitches can detect specific ligands of interest. There were two MS2 riboswitches, M701 and M718, that were able to turn on gene expression of the MS2 coat protein with statistically significant ($p < 0.05$) activation ratios of 3.93 ± 2.18 and 4.90 ± 2.10, respectively. The two CRP riboswitches further tested in dose-response experiments, C624 and C684, had maximum activation ratios of 3.10 ± 8.26 and 2.90 ± 7.70, respectively. The two IL-32 riboswitches tested in dose-response experiments, I1000 and I1003, had activation ratios of 1.97 ± 3.95 and 1.27 ± 2.55, respectively. Both the CRP riboswitches and IL-32 were unable to significantly activate gene expression due to low protein TIRs (477 for CRP and 179 for IL-32). The low TIRs cause a lower protein concentration in the TX-TL system so the riboswitches may have had few proteins to detect, causing minimal or no activation of gene expression.

Although there were some challenges presented with the selected proteins, there is data supporting that further experimentation on these medically relevant proteins should be done. For future experiments, more research should be done in order to clone a protein with a high translation initiation rate. This will allow for more protein production during the TX-TL experiments, increasing the chances that the riboswitch can successfully turn on gene expression.

There are many other aptamers that bind to biomarkers of human disease that can be used to create specific riboswitches that will detect them. There has been some research for RNA

aptamers for human influenza B virus[28], Sialyl Lewis X[29], and αvβ3 integrin[30]. Riboswitches that can successfully detect proteins are a great alternative to the current expensive methods for creating protein assays. Eventually they could become scalable and used in medical lab settings.

Chapter 8

Materials and Methods

Computational Sequence Design

The Riboswitch Calculator[1] was used to design pre-aptamer and post-aptamer sequences with varying predicted TIRs in OFF and ON states, activation ratios, and thermodynamic properties. This model also provided predictions for the riboswitches' activation ratios. The Operon Calculator[31] was used to optimize the MS2 and CRP protein coding sequences for expression in *Escherichia coli*. Following optimization of each protein, the RBS Library Calculator[32] was used to create a library of RBS's, varying the protein's expression level. The RBS Calculator was then used to confirm the TIR for each specific RBS.

Cloning Riboswitches and Protein Coding Sequences

The inserts, the protein operon and riboswitches, were cloned into pFTV1-mRFP plasmid vectors. The designed oligonucleotides were amplified with PCR. The PCR product was then digested with restriction enzymes; XbaI and SacI were used for the riboswitches while XbaI and NotI were used for the protein coding sequences. Ligation of the insert and backbone resulted in the final desired expression cassette and riboswitch plasmids. These plasmids were finally transformed into chemically competent *Escherichia coli,* strain DH10B, which is typically used for large insert DNA library clones[33].

TX-TL Prep and Set Up

The TX-TL crude cell extract was prepped with E.coli BL21 Rosetta 2 cells and a 2xYT+P+40 ug/mL solution while S30A buffer (6 mL/g cell pellet) was used as a wash. The cells were centrifuged with continuous flow centrifuge and resuspended with S30A buffer (6mL/g cell pellet). The cells were then centrifuged again, and the supernatant was decanted. The wash was repeated. The cells was resuspended with S30A (1 mL/g cell pellet) and run through a microfluidizer twice. The resulting lysate was centrifuged for one hour ($4°C$). The supernatant was incubated at 37 $°C$, open to air for 80 min, and was then centrifuged for one hour (4 $°C$). The final pellet was filtered through a 0.1 µm filter. The filtrate was then diafiltrated for 22 hours with S30B buffer and the final protein concentration was measured (20-30 mg/mL). The amino acid solution was created using a sigma amino acid kit. Stocks of the twenty amino acids were made (168mM except for leucine at 140mM) and combined together to create a final concentration of 5mM for leucine and 6mM for the remaining amino acids. These solutions were titrated to obtain a pH of 7 and stored in 500uL aliquots. The energy solution was created using the relevant protocol created by Sun and Hayes[21]. For each test, 2% PEG-8000 (0.25µL), 4mM Magnesium glutamate (0.2µL), 80 mM Potassium glutamate (0.15µL), 1mM DTT (0.01µL), and water (µL needed to fill rest of 5µL volume) were essential for the functionality of this *in vitro* test. In order to improve the folding of the CRP protein, 1mM and 2mM calcium was also used in the tests. Along with the 33% extract, 25% amino acid solution, and 7% energy solution, these components were all included with the riboswitch and respective protein in a 5 µL reaction.

Measurement of Results

In order to measure the activation of the MS2 and CRP riboswitches, a TECAN microplate reader measured the fluorescence levels of each experiment run. The TX-TL experiments were run for twelve hours at a temperature of 29 $°C$. The TECAN microplate reader measured the fluorescence from the bottom of each plate well and provided the fluorescence and OD values as a function of time.

Appendix A

Additional Riboswitch Data

The sequences of the riboswitches determine how they function. The aptamer sequences, shown in Table 1, for MS2, CRP, and IL-32 were input into the Riboswitch Calculator to create the riboswitch sequences with varying outputs such as predicted actual AR, switching free energy, TIR in the OFF and ON states, and the free energy values in the OFF and ON state.

Table 1: Aptamer sequences of MS2, CRP, and IL-32.

Aptamer Name	Aptamer Sequence	ΔG_{ligand} (kcal/mol)
MS2	ACATGAGGATCACCCATGT	-13.18
CRP	GCCTGTAAGGTGGTCGGTGTGGCGAGTGTGTTAGG AGAGATTGC	-8.41
IL-32	GGGTTCACTGCAGACTTGACGAAGCTTCCGGAGAGA AGGGTCAAAGTTGTGCGGGAGTGTGTTGTGGAATGG ATCCACATCTACGAATTC	-9.82

The aptamer minimum free energy structures were obtained with Vienna RNAfold. All three aptamers have a negative minimum free energy which is more stable.

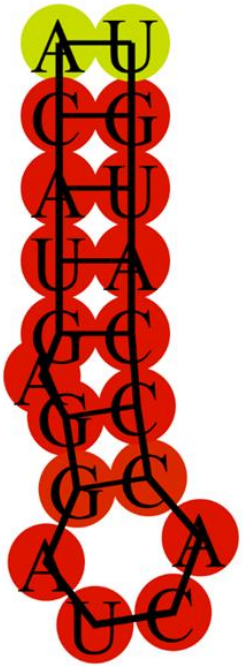

Figure 14: The minimum free energy structure of the MS2 aptamer was obtained from RNAfold.

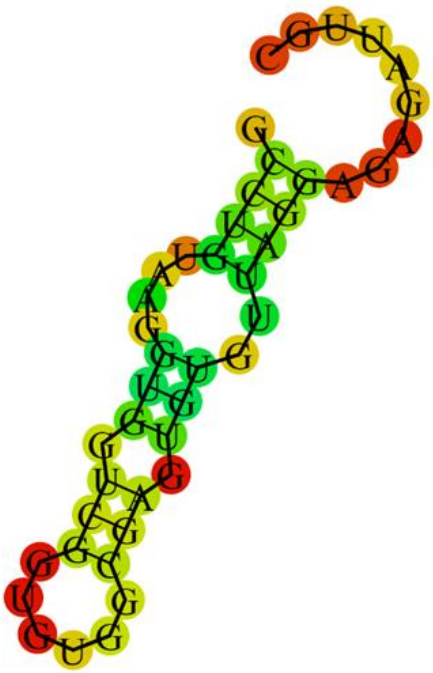

Figure 15: The minimum free energy structure of the MS2 aptamer was obtained from RNAfold.

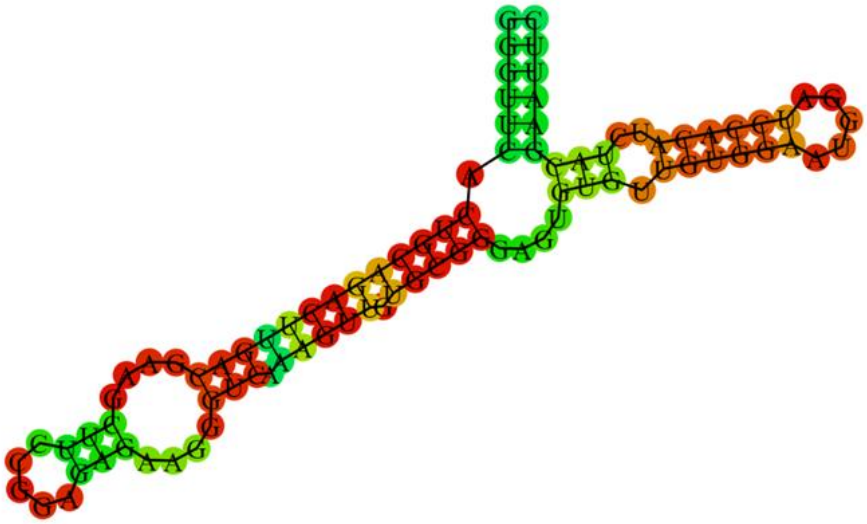

Figure 16: The minimum free energy structure of the IL-32 aptamer was obtained from RNAfold.

The Riboswitch Calculator creates a pre-aptamer and post-aptamer sequence to create the entire riboswitch sequence displayed in Tables 2, 3, and 4.

Table 2: MS2 riboswitch pre-aptamer and post-aptamer sequences were obtained with the Riboswitch Calculator.

MS2 Riboswitch	Pre-Aptamer Sequence	Post-Aptamer Sequence
M701	TCTAGACTTAAAATGAGAC	CTGGAGACAGGCAAGGAGGGTGATA
M718	TCTAGACAAAGAATGAAAC	CTGGAGACAGACAAGGAGGGTGAGT
M957	TCTAGACAAAAAATGAAAC	CTGGAGACAGACAAGGAGGGTGAGT
M958	TCTAGAAATGAGAAGAGAC	CTGAGAAAAGATAAGGAGGGTGAGG
M959	TCTAGACAAGTAAAGAGAC	CTGAAGTAAGATAAGGAGGGTGAGG

Table 3: CRP riboswitch pre-aptamer and post-aptamer sequences were obtained with the Riboswitch Calculator.

CRP Riboswitch	Pre-Aptamer Sequence	Post-Aptamer Sequence
C624	TCTAGAAATAAGCTAATAC	CCATTGGCTAATAAGGAGGGTGAAA
C634	TCTAGACAAAAGCAGTTAC	TCACGGGTCAATAAGGAGGGTGGGT
C642	TCTAGACATATACAGCGAC	TCACAGGCCAATAAGGAGGGTGAGG
C648	TCTAGAAGTATTGCAGCAC	GACAAGGCTAATAAGGAGGGTGAGA
C726	TCTAGAAAAGCTGAAGTAC	CGCAAGGCCATGAAGGAGGCGTAGG
C684	TCTAGATATGCTGAAGTAC	GTCAAGGCCATTAAGGAGGAGTACG
C693	TCTAGAGAAACTGAAATTC	GACAAGGCCATTAAGGAGGTGAAAG

Table 4: IL-32 riboswitch pre-aptamer and post-aptamer sequences were obtained with the Riboswitch Calculator.

IL-32 Riboswitch	Pre-Aptamer Sequence	Post-Aptamer Sequence
I949	TCTAGAATTTAAAAAGA AAGATCTGT	CTTAGATGATATAAGGAGGGTG AAA
I972	TCTAGATAGAAAGATAA ACGGTAAGA	CGTGCAAGAGATAAGGAGGGTG AAA
I1000	TCTAGAATGAATAAATA AGAGTTACA	GCTACATTGGATAAGGAGGGTG AGT
I1003	TCTAGAATTAATTAATAT GAGTGACG	GCTACTGTGGATAAGGAGGGTG AGG
I1054	TCTAGAATGCAATTGATA AAGTGTCA	GCTACACTAAATAAGGAGGGTG AAA

All the cloned riboswitch sequences and their relevant predicted ARs, $\Delta\Delta G_{standby}$, and TIR (OFF and ON) are displayed in Tables 5, 6, and 7.

Table 5: The MS2 riboswitch model predicted ARactual, $\Delta\Delta G_{standby}$, and TIR (OFF and ON) were obtained from the Riboswitch Calculator

MS2 Riboswitch	Predicted AR_{actual}	$\Delta\Delta G_{standby}$	TIR, OFF (au)	TIR, ON (au)
M701	570.65	-4.62	137.38	248017.8
M718	164.77	-2.75	477.20	248017.8
M957	2462.77	-7.27	12.93	100827.4
M958	2818.64	-7.27	13.52	120714.5
M959	2576.11	-7.27	14.80	120714.5

Table 6: The CRP riboswitch model predicted ARactual, $\Delta\Delta G_{standby}$, and TIR (OFF and ON) were obtained from the Riboswitch Calculator.

CRP Riboswitch	Predicted AR_{actual}	$\Delta\Delta G_{standby}$	TIR, OFF (au)	TIR, ON (au)
C624	10.64	-2.33	5865.73	248017.8
C634	9.17	-2.81	6889.32	248017.8
C642	7.91	-0.51	8104.65	248017.8
C648	7.17	-3.99	9036.59	248017.8
C726	225.14	-7.11	36.59	35173.93
C684	167.36	-6.9	190.10	135699.1
C693	157.18	-6.2	161.66	108355.5

Table 7: The IL-32 riboswitch model predicted AR_actual, $\Delta\Delta G_{standby}$, and TIR (OFF and ON) were obtained from the Riboswitch Calculator.

IL-32 Riboswitch	Predicted AR_{actual}	$\Delta\Delta G_{standby}$	TIR, OFF (au)	TIR, ON (au)
I949	24.22	-3.12	2177.78	248017.8
I972	19.59	-3.23	2714.37	248017.8
I1000	15.25	-2.71	3529.36	248017.8
I1003	14.93	-1.2	3607.73	248017.8
I1054	11.30	0	4855.48	248017.8

Appendix B

Additional RBS Library Data

The translation initiation rate (TIR) of the target protein can influence the activation of the riboswitch during coexpression in TX-TL. Therefore, it is necessary to obtain protein clones (of MS2, CRP, IL-32) at varying TIRs. The RBS Library Calculator can shorten the cloning process by creating a degenerate library of RBS sequences. After inputting the original protein RBS into the RBS Library Calculator, the calculator will design a new RBS containing degenerate nucleotides (represented by Y, K, S, W, M, N, D, and R) highlighted in Table 8 below.

Table 8: The RBS Library sequence for the MS2, CRP, and IL-32 expression cassettes contained degenerate nucleotides and was obtained with the RBS Library Calculator.

Protein	RBS Library Sequence	# of sequences	Max TIR (au)
MS2	CAGCACAATCTTAGAGSCCYKTASGGWGSTATCT	64	98791.5
CRP	GWTAGAATCTCKTCAAACDAKGTAAGSAGAGTAMA	96	78116.5
IL-32	CMCTAACTTTACNTTAGGAAACSAARGAKGTAAAG	64	198159.8

These degenerate nucleotides were replaced with either A, C, T, or G. During transformation, each colony had a different RBS sequence because the degenerate nucleotide could be replaced with any one of the four DNA nucleotides. As a result of this, four

successfully cloned colonies had different TIRs. Each RBS Library plasmid was tested with the respective riboswitches, allowing the highest expression levels of the protein to be obtained.

A TECAN experiment was performed to evaluate the effects of each RBS for each riboswitch. For MS2 tests, 4 RBS constructs were tested (M1, M3-1, M3-7, M5-1) with the presence of a RFP reporter. Figure 17 displays the expression levels of each MS2 protein and shows that the M1 construct had the best expression level, suggesting this RBS should be tested with each MS2 riboswitch.

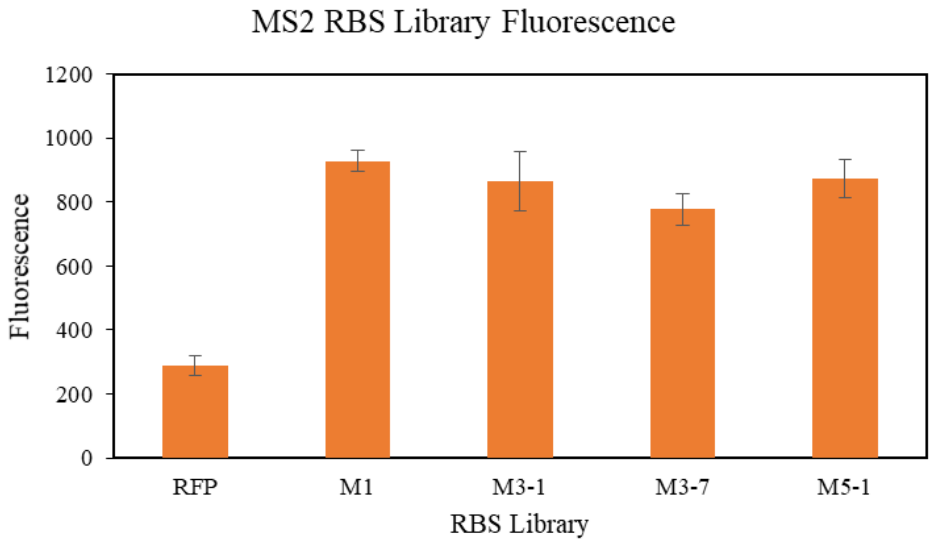

Figure 17: The fluorescence values of four MS2 RBS library expression cassettes (M1, M3-1, M3-7, M5-1) at 2nM concentrations were found in the presence of a fluorescent report protein, RFP.

Four different CRP RBS constructs (C3-3, C3-3 5nM, C3-7, C3-8, and C5-1) were also tested at 2nM concentrations unless otherwise specified in the presence of RFP. The fluorescence values are shown in Figure 18. The largest fluorescence value was obtained by the C5-1 RBS sequence. However, the results in this test were not significantly different than the other RBS constructs. The RBS Library, C3-3 (2nM), was used for the dose response tests.

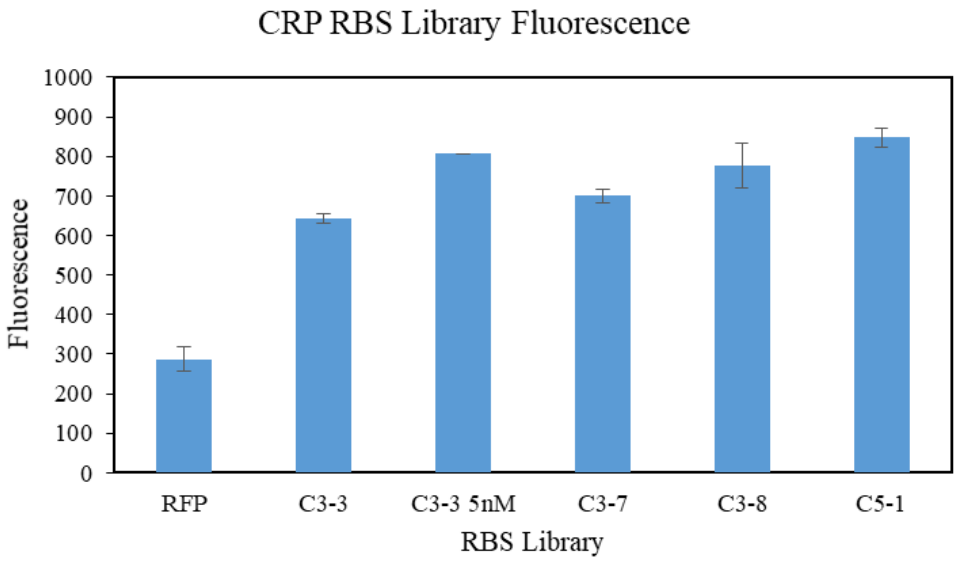

Figure 18: The fluorescence values of four CRP RBS library expression cassettes at 2nM concentrations (C3-3, C3-7, C3-8, C5-1) and one expression cassette at 5nM (C3-3) were found in the presence of RFP

Three RBS Library sequences (ILC-1, ILC-12, ILC-13) for the IL-32 expression cassette were tested at 2nM concentrations in the presence of RFP, and are shown in Figure 19. Out of the three RBS constructs, ILC-12 had the best fluorescence but the results were not significantly different from the other RBS sequences. Additionally, the no protein control was not plotted, because it was assumed that the protein and RFP tests would have a larger fluorescence value as

seen in the cases of the MS2 and CRP RBS libraries. The ILC-13 RBS Library was used for dose response testing due to its low standard deviation.

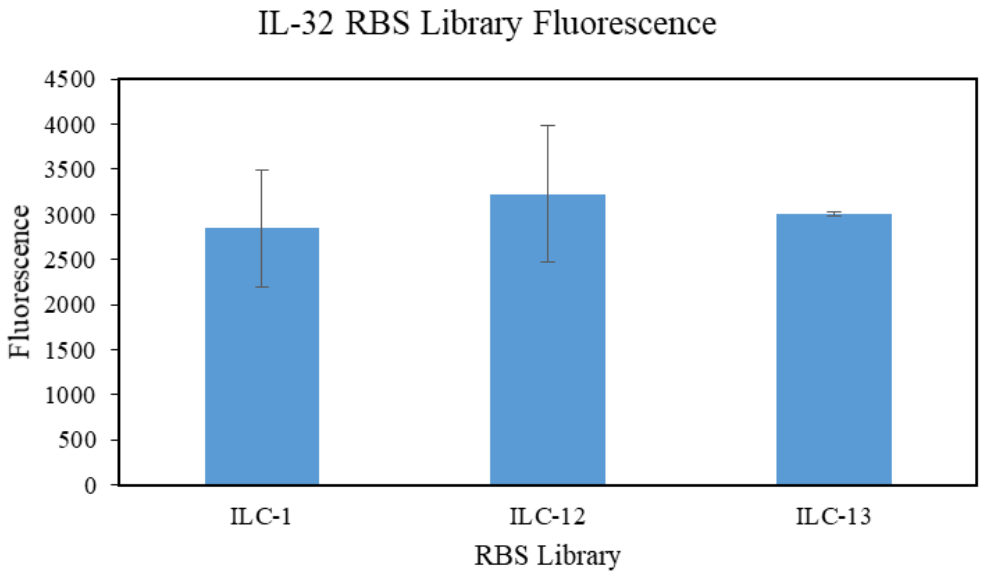

Figure 19: The fluorescence values of three IL-32 RBS library expression cassettes (ILC-1, ILC-12, ILC-13) at 2nM concentrations were found in the presence of RFP.

The chosen RBS sequences and TIRs for each protein are shown in Table 9.

Table 9: The selected protein RBS's are displayed with their sequences and TIR values.

Protein RBS	Selected RBS Sequence	TIR (au)
M1	ACGAACGAAATTATAAGGAGGTAAA	117150.0
C3-3	GTTAGAATCTCGTCAAACGATGTAAGCAGAGTAAA	477.0
ILC-13	CCTAACTTTACTTTAGGAAACGAAAGAGGTAAAG	179.1

BIBLIOGRAPHY

1. Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W. & Salis, H. M. Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic Acids Res.* **44,** 1–13 (2016).
2. Zhou, W., Jimmy Huang, P.-J., Ding, J. & Liu, J. Aptamer-based biosensors for biomedical diagnostics. *Analyst* **139,** 2627 (2014).
3. Corrie, S. R., Coffey, J. W., Islam, J., Markey, K. A. & Kendall, M. A. F. Blood, sweat, and tears: developing clinically relevant protein biosensors for integrated body fluid analysis. *Analyst* **140,** 4350–4364 (2015).
4. Luo, X. & Davis, J. J. Electrical biosensors and the label free detection of protein disease biomarkers. *Chem. Soc. Rev.* **42,** 5944 (2013).
5. Yu, Y. *et al.* Molecular selection, modification and development of therapeutic oligonucleotide aptamers. *Int. J. Mol. Sci.* **17,** 1–19 (2016).
6. Osborne, S. Aptamers as therapeutic and diagnostic reagents: problems and prospects. *Curr. Opin. Chem. Biol.* **1,** 5–9 (1997).
7. Hallberg, Z. F., Su, Y., Kitto, R. Z. & Hammond, M. C. Engineering and In Vivo Applications of Riboswitches. *Annu. Rev. Biochem* **86,** 515–39 (2017).
8. Stoltenburg, R., Krafčiková, P., Víglaský, V. & Strehlitz, B. G-quadruplex aptamer targeting Protein A and its capability to detect Staphylococcus aureus demonstrated by ELONA. *Sci. Rep.* **6,** (2016).
9. Cappi, G. *et al.* Label-Free detection of tobramycin in serum by transmission-localized surface plasmon resonance. *Anal. Chem.* **87,** 5278–5285 (2015).
10. Arroyo-Currás, N. *et al.* Real-time measurement of small molecules directly in awake, ambulatory animals. *Proc. Natl. Acad. Sci.* **114,** 645–650 (2017).
11. Han, X. *et al.* Gold nanoparticle based photometric determination of tobramycin by using new specific DNA aptamers. *Microchim. Acta* **185,** 4 (2017).
12. Michener, J. K., Thodey, K., Liang, J. C. & Smolke, C. D. Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.* **14,** 212–222 (2012).
13. Davidson, M. E., Harbaugh, S. V., Chushak, Y. G., Stone, M. O. & Kelley-Loughnane, N. Development of a 2,4-dinitrotoluene-responsive synthetic riboswitch in E. coli Cells. *ACS Chem. Biol.* **8,** 234–241 (2013).
14. Johansson, H. E., Liljas, L. & Uhlenbeck, O. C. RNA recognition by the MS2 phage coat protein. *Seminars in Virology* **8,** 176–185 (1997).
15. Wang, M. S., Black, J. C., Knowles, M. K. & Reed, S. M. C-reactive protein (CRP) aptamer binds to monomeric but not pentameric form of CRP. *Anal. Bioanal. Chem.* **401,** 1309–1318 (2011).
16. Kim, S. *et al.* Generation of antagonistic RNA aptamers specific to proinflammatory cytokine interleukin-32. *Bull. Korean Chem. Soc.* **31,** 3561–3566 (2010).
17. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27,** 946–950 (2009).

18. Forster, A. C. & Church, G. M. Synthetic biology projects in vitro. *Genome Research* **17,** 1–6 (2007).
19. Noireaux, V., Bar-Ziv, R. & Libchaber, A. Principles of cell-free genetic circuit assembly. *Proc. Natl. Acad. Sci.* **100,** 12672–12677 (2003).
20. Jewett, M. C. & Swartz, J. R. Mimicking the Escherichia coli Cytoplasmic Environment Activates Long-Lived and Efficient Cell-Free Protein Synthesis. *Biotechnol. Bioeng.* **86,** 19–26 (2004).
21. Sun, Z. Z. *et al.* Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *J. Vis. Exp.* (2013). doi:10.3791/50762
22. Bini, A., Centi, S., Tombelli, S., Minunni, M. & Mascini, M. Development of an optical RNA-based aptasensor for C-reactive protein. *Anal. Bioanal. Chem.* **390,** 1077–1086 (2008).
23. Shtatland, T. Interactions of Escherichia coli RNA with bacteriophage MS2 coat protein: genomic SELEX. *Nucleic Acids Res.* **28,** 93e–93 (2000).
24. Keswani, A. *et al.* Differential expression of interleukin-32 in chronic rhinosinusitis with and without nasal polyps. *Allergy* **67,** 25–32 (2012).
25. Tadmor, A. D. & Tlusty, T. A coarse-grained biophysical model of E. coli and its application to perturbation of the rRNA operon copy number. *PLoS Comput. Biol.* **4,** (2008).
26. Fakruddin, M., Mohammad Mazumdar, R., Bin Mannan, K. S., Chowdhury, A. & Hossain, M. N. Critical Factors Affecting the Success of Cloning, Expression, and Mass Production of Enzymes by Recombinant E. coli. *ISRN Biotechnol.* **2013,** 590587 (2013).
27. Dortay, H., Schmöckel, S. M., Fettke, J. & Mueller-Roeber, B. Expression of human c-reactive protein in different systems and its purification from Leishmania tarentolae. *Protein Expr. Purif.* **78,** 55–60 (2011).
28. Gopinath, S. C. B., Sakamaki, Y., Kawasaki, K. & Kumar, P. K. R. An efficient RNA aptamer against human influenza B virus hemagglutinin. *J. Biochem.* **139,** 837–846 (2006).
29. Jeong, S., Eom, T., Kim, S., Lee, S. & Yu, J. In vitro selection of the RNA aptamer against the Sialyl Lewis X and its inhibition of the cell adhesion. *Biochem. Biophys. Res. Commun.* **281,** 237–43 (2001).
30. Mi, J. *et al.* Targeted inhibition of αvβ3 integrin with an RNA aptamer impairs endothelial cell growth and survival. *Biochem. Biophys. Res. Commun.* **338,** 956–963 (2005).
31. Halper, S. M., Cetnar, D. P. & Salis, H. M. An automated pipeline for engineering many-enzyme pathways: Computational sequence design, pathway expression-flux mapping, and scalable pathway optimization. in *Methods in Molecular Biology* **1671,** 39–61 (2018).
32. Farasat, I. *et al.* Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Mol. Syst. Biol.* **10,** 731–731 (2014).
33. Durfee, T. *et al.* The complete genome sequence of Escherichia coli DH10B: Insights into the biology of a laboratory workhorse. *J. Bacteriol.* **190,** 2597–2606 (2008).

ACADEMIC VITA

Academic Vita of Lipika Reddy Gadila
lipi567@gmail.com

Education

B.S., Chemical Engineering, 2018, The Pennsylvania State University, University Park, PA
Major: Chemical Engineering
Honors: Dean's List (2014, 2015, Spring 2016, 2017)

Thesis Title: Computation Design and Experimental Validation of RNA-Based Biosensors for the Detection of Biomarker Proteins

Thesis Supervisor: Dr. Howard Salis

Work Experience

- Merck Vaccines Manufacturing Intern — Summer 2017
  West Point, PA
  Supervisor: Richard Spotts
  - Created a job aide to standardize work for operators performing quarterly cleaning, projected to lower departmental overtime spend for quarterly cleans by 10%
  - Designed a training tool to reduce human error deviations from a process by providing an understanding on the background and problems in previous processes
  - Authored an atypical investigation that resulted from a human related mistake, saving a batch from discard
  - Ensured all site-wide products remained compliant by manually monitoring temperature of CTUs following companywide cyber-attack
  - Worked with a cross functional team to support execution of a process improvement by ensuring materials and equipment were available, allowing the process to be executed by the target timeline

- Research Assistant — Fall 2015-Current
  State College, PA
  Principal Investigator: Dr. Howard Salis
  - Led project to design, construct, and characterize biosensors to detect biomarkers of human disease
  - Optimized protocol to increase efficiency with the Riboswitch Calculator and Operon Calculator

- Analyzed data to improve experiments by measuring fluorescence intensity of C-reactive and MS2 proteins

Leadership/ Activities

- Engineering Orientation Network Mentor (August 2016- Fall 2017)
  State College, PA
  - Mentored six incoming freshman engineers and helped guide their transition to college
  - Facilitated group collaboration of students during design competition
- Rules and Regulations Committee Member (THON) (Fall 2015-Spring 2017)
  State College, PA
  - Maintained expectations of authorities and attendants within security administrating role for an event that has raised $147 million for children with pediatric cancer

Grants

- SHC Academic Excellence Scholarship (Fall 2014- Current)
- The Penn State University Four Year Provost Fund (Fall 2014- Current)