Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter know, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Julie M. Longo 04/18/2011

Improving a Synthetic Riboswitch that Responds to an Herbicide

by

Julie Longo

Dr. Justin Gallivan
Adviser

Department of Chemistry

Dr. Justin Gallivan

Adviser

Dr. Lloyd Parker

Committee Member

Dr. Sarah McPhee

Committee Member

04/18/2011

Improving a Synthetic Riboswitch that Responds to an Herbicide

By

Julie Longo

Dr. Justin Gallivan

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Chemistry

2011

Abstract

Improving a Synthetic Riboswitch that Responds to an Herbicide
By Julie Longo

A riboswitch is a part of an mRNA molecule that controls gene expression through binding to a small molecule ligand. Riboswitches are found in the 5' untranslated region of mRNA and are made up of two components, the aptamer, which binds to a small molecule ligand, and the expression platform, which changes conformation in response to the ligand. Because riboswitches influence gene expression, they can be used to reprogram bacteria to perform novel functions. The Gallivan lab recently identified a riboswitch that recognizes atrazine, a harmful herbicide. However, this switch shows residual gene expression in the absence of the ligand and only a five fold activation ratio in the presence of the ligand. This project seeks to improve this riboswitch to give it a greater dynamic range, the difference in gene expression in the presence and absence of the ligand. Methods used to alter the switching ability of the original atrazine riboswitch include randomization of a section of the expression platform and deletion of fifteen nucleotides downstream of the aptamer. By designing a better switch, it will be possible to create bacteria that will be better able to metabolize residual atrazine in the environment.

Improving a Synthetic Riboswitch that Responds to an Herbicide

By

Julie Longo

Dr. Justin Gallivan

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Chemistry

2011

Acknowledgements

I would first like to thank my advisor, Dr. Justin Gallivan, for giving me the opportunity to work in his lab over the past two years. I have learned a lot from him and from the other members of the Gallivan lab. In particular, I'd like to thank Dr. Joy Sinha for being a wonderful mentor and for always laughing when things went wrong, a good reminder to put things in perspective. I'd also like to thank Dr. Adrianne Edwards and Dr. Dennis Mishler for all their help and encouragement while writing this honors thesis. In addition, I am grateful to the Emory SIRE Research Partners Program for funding my first year of work in the Gallivan lab and for providing an outlet to discuss research with my peers. Thanks, too, to the rest of my committee, Dr. Lloyd Parker and Dr. Sarah McPhee for their valuable time.

I owe immeasurable gratitude to the people in my life who have encouraged me in my pursuit of chemistry over the past four years. First, the professors who always supported me and were always there to talk about chemistry or anything else, Dr. Lloyd Parker and Ms. Brenda Harmon. My experience at Oxford would not have been the same without them there. In addition, I'd like to thank Dr. Geoff Coates at Cornell University for giving me the opportunity for my first research experience and sparking my interest continuing my education.

Last but not least, I want to thank my family and friends for all of their support and encouragement. Sam Whitehead, Leslie Munoz, and Adrienne Vinson have seen me through everything in my research career, celebrating my successes and giving me hope through my failures. In addition, my mom, dad, and sister have never failed to be my biggest cheering section, always believing in my abilities even when I didn't.

Table of Contents

List of Figures and Tables

Introduction

Controlling gene expression is an important challenge in molecular biology. In nature, even the simplest organisms must coordinate the expression of hundreds of genes to respond efficiently and quickly to specific signals and bring about their intended effects. The systems that control gene expression must interpret many chemical and physical signals while coordinating a number of tasks carried out by cells.[1] Much of this regulation occurs through protein-based control of gene expression, such as transcription factors, which bind to DNA to promote or repress transcription. However, there are a number of other factors that can influence gene expression post-transcriptionally, translationally, and at the level of mRNA processing. Recently, the importance of the regulatory properties of RNA in controlling gene expression has been recognized.[1]

Until recently, gene regulation was widely believed to be controlled primarily by proteins. The discovery of the *lac* repressor, a DNA-binding protein involved in the metabolism of lactose, led to a large amount of data describing the ability of protein factors to respond to various small molecules.[3] However, alternative speculation that RNA might also play a regulatory role existed. This was based on the "RNA world" hypothesis, which states that early life forms used RNA for both genetic storage and catalytic processes. If such a world did exist, it is possible that remnants of this type of regulatory control are still present in modern biological systems. The discovery of ribozymes, RNA molecules that perform catalysis, demonstrated that RNA could form complex shapes and catalyze critical and specific reactions.[4,5] The ability of RNA to form a multitude of different secondary structures suggests that these molecules could have the ability to control gene expression through conformational changes induced by

small molecule binding. In addition, a number of anomalies in protein based gene-control began to arise.

For example, conserved regions in the 5' untranslated region (UTR) of the *btuB* mRNA in the bacteria *Escherichia coli* and the *cob* mRNA in *Salmonella* Typhimurium were found to be essential to the regulation of the coenzyme B_{12}.[6,7,8] Protein factors that could account for this regulation could not be found in repeated attempts, and as a result, it seemed increasingly likely that there was direct interaction between the coenzyme and the mRNA itself.[9,7] Structural probing experiments showed that the RNA underwent a structural reorganization upon ligand binding, allowing the 5' UTR of the *btuB* and *cob* transcripts to regulate gene expression. The conserved region of the 5' UTR bound very selectively to B_{12}, even in the presence of similar analogs.[1] In a similar manner, complex folded domains were found in the 5' UTR of many genes involved in the regulation of a number of metabolites, including thiamin pyrophosphate (TPP)[10,11], flavin mononucleotide (FMN)[12,13], *S*-adenosylmethionine (SAM)[14], guanine[15,16], and lysine[17], among others (Figure 1).

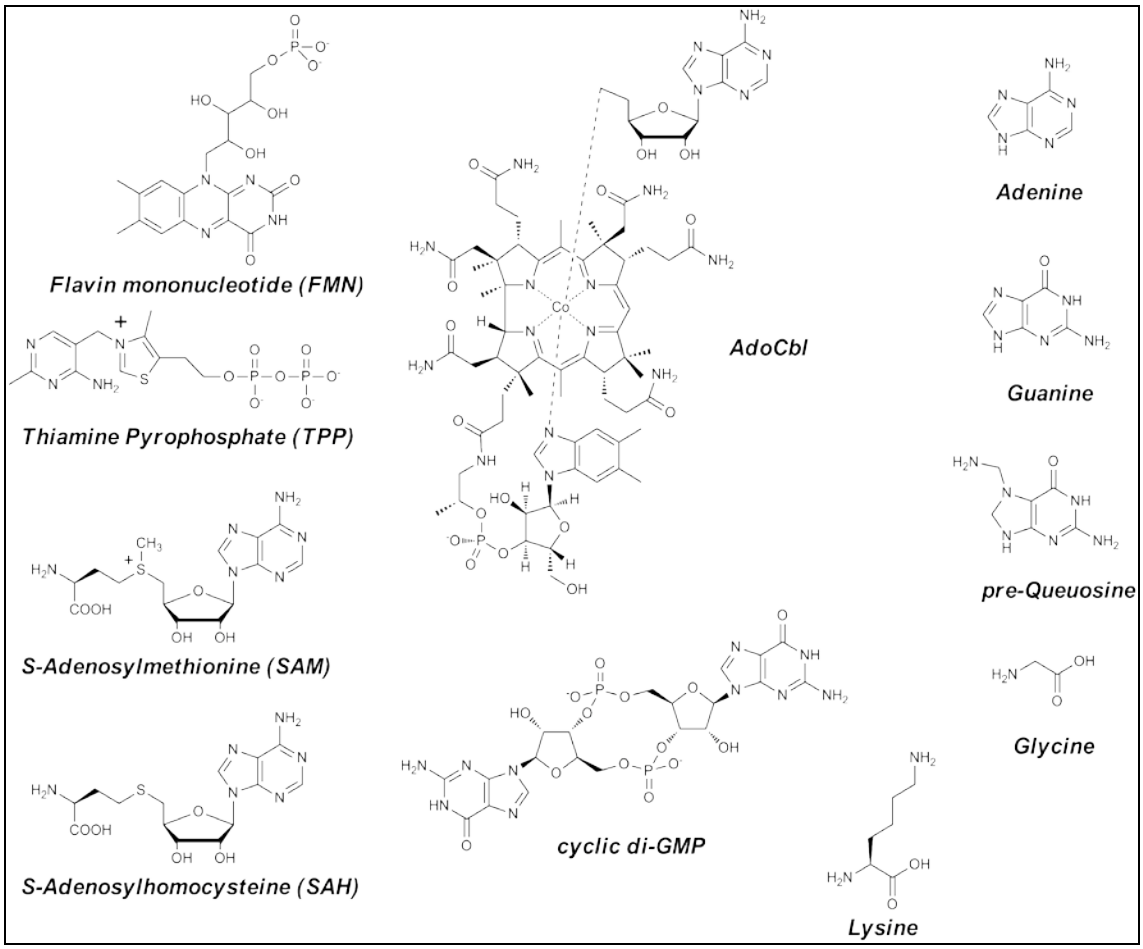

Figure 1. Small molecule ligands found in nature that interact directly with mRNA that can affect gene expression through binding.

This growing collection of mRNA domains that regulate gene expression were given the name "riboswitches". Riboswitches are comprised of two key parts: the aptamer, which binds the ligand, and the expression platform, which controls gene expression (Figure 2). The aptamer acts as the chemical sensor, taking the place of the protein factor that would otherwise serve as the sensory element.[1] The aptamer must be able to selectively recognize a metabolite with suitable affinity. Unlike proteins, which have an alphabet of twenty amino acids from which to choose, allowing great structural diversity in the possibilities of structures, riboswitches, like all RNA molecules, only possess four different nucleotides (letters) with which to construct the metabolite

detector. However, because the aptamer can form a distinct secondary structure through base pairing, specificity of binding to metabolites can be achieved. In the presence of a

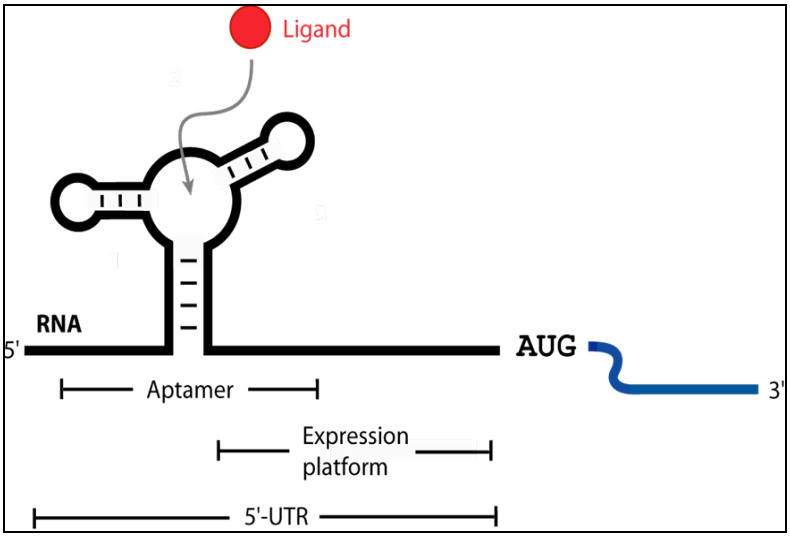

Figure 2. Schematic Representation of a Riboswitch

ligand, the expression platform, located directly downstream of the aptamer, undergoes a conformational change that affects gene expression. Gene expression is most commonly affected at the level of transcription elongation or translation initiation by influencing the binding of transcriptional or translational machinery.

The Gallivan lab has focused mostly on riboswitches that control gene expression at the translational level. Directly downstream of the expression platform is the ribosome binding site and the start codon (AUG), signaling the start of translation of the mRNA into protein. Genes located downstream from this start codon are thus controlled by the accessibility of the ribosome binding site, which, in turn, is controlled by the conformation of the expression platform. The binding of a ligand to the aptamer domain causes a conformational change that either opens up the ribosome binding site or obscures it, allowing for up-regulation or down-regulation of gene expression.

Previous Work

In addition to their prevalence in nature, riboswitches can also be artificially produced in the lab. For riboswitches to be viable as a way to manipulate gene expression, they must be able to respond to different ligands than those found in nature. Previous work by several groups has led to the development of synthetic riboswitches that respond to novel target molecules.[3] The Gallivan lab chose to investigate a completely new target, atrazine. Atrazine is a useful ligand for the development of riboswitches for three reasons. From an environmental standpoint, atrazine, which is toxic to humans, is one of the most widely used herbicides in the United States and is present in high concentrations in our soil and environment.[18,19] Secondly, atrazine is an attractive target for interaction with RNA because it contains both hydrogen donors and acceptors. Finally, the metabolic pathway of atrazine is well characterized, which means that the enzymes required for the catabolism of atrazine can be expressed and purified in *E.coli*, making it possible to engineer cells that can degrade atrazine into the less harmful hydroxyatrazine (Figure 3).[20]

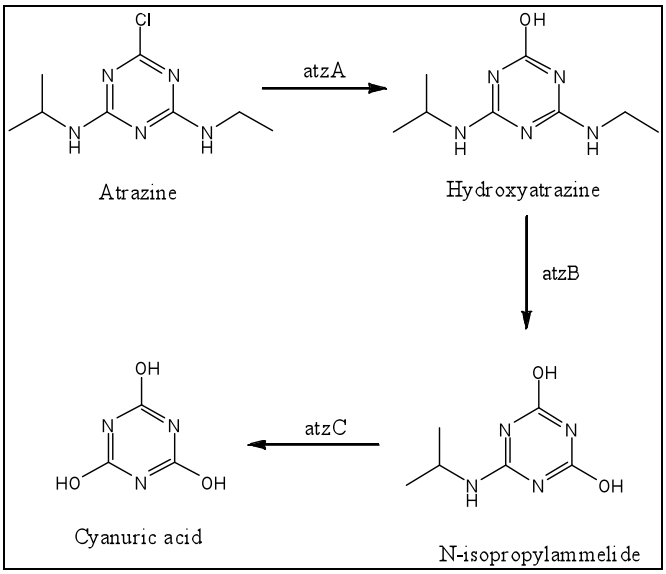

Figure 3. The catabolic pathway of atrazine

To select for RNA aptamers that specifically bind to atrazine, the Gallivan lab used Systematic Evolution of Ligands by Exponential Enrichment (SELEX), an *in vitro* technique (Figure 4).[2] In brief, an atrazine derivative was synthesized and coupled to a solid support. A library of DNA sequences comprising forty random nucleotides (N_{40}), flanked by a T7 promoter sequence at the 5' end and a constant sequence at the 3' end, was prepared from chemically synthesized oligonucleotides using PCR. The DNA library was transcribed to RNA using T7 RNA polymerase, and the RNA was gel purified and subjected to nine rounds of SELEX, using buffers to wash away the non-selective sequences. After the ninth round, a counter selection was performed by washing the column with the atrazine catabolite hydroxyatrazine to remove those molecules that were not selective. The remaining bound RNA was removed by washing

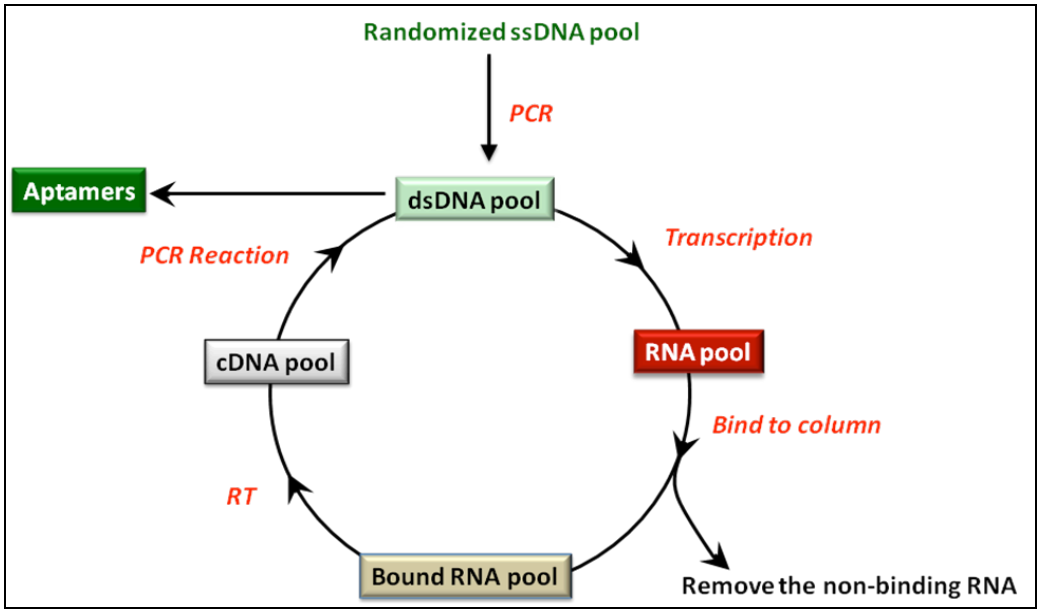

Figure 4. Schematic of the SELEX process. A randomized DNA pool is transcribed into RNA and applied to a column containing the metabolite of interest. Buffers are used to remove RNA that does not selectively bind to the column. The metabolite is used to elute the bound RNA, and reverse transcription (RT) converts the RNA back into DNA. The DNA is amplified using PCR and transcribed into RNA again, and the cycle is repeated several times for enrichment.

the column with atrazine. Figure 5 shows the amount of RNA bound to the column after each successive round of SELEX. The remaining RNA pool was subjected to two more rounds of SELEX, and after all twelve rounds of selection approximately 55% of the RNA in the pool bound to the column could be eluted with atrazine. At this point, sequencing of 33 individual clones showed that the pool remained diverse and that all of the sequences were unique. The SELEX experiments are designed to select the tightest-binding RNA, but for a riboswitch to be functional, it must also be able to undergo a

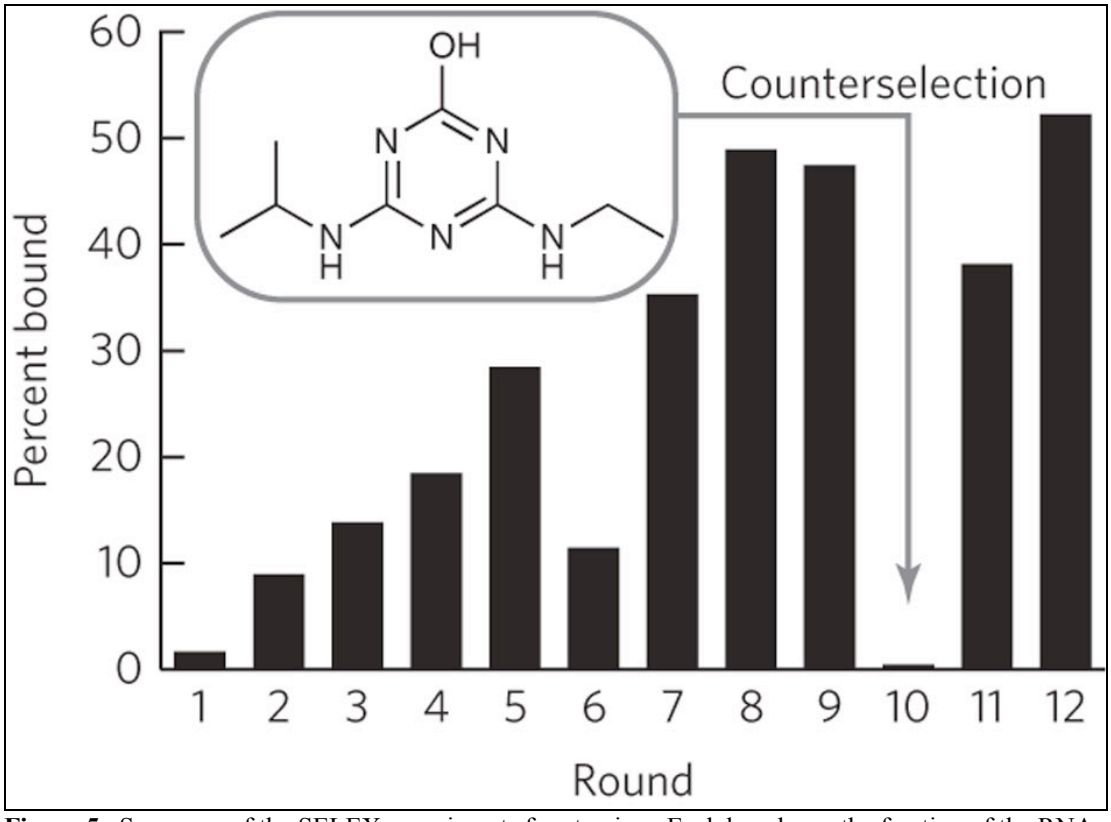

Figure 5. Summary of the SELEX experiments for atrazine. Each bar shows the fraction of the RNA pool bound to the atrazine-derivatized column after washing, with the counter selection occurring after round nine.[2]

favorable conformational change on a biologically relevant time scale. For this reason, instead of choosing a single aptamer that bound very tightly to atrazine, a library of aptamers showing affinities towards atrazine were chosen. This library was then cloned in front of a random N_{10} RNA sequence in the 5'-UTR, upstream of the *cheZ* gene, which controls cell motility, and putative riboswitches were screened for using bacterial motility screening methods.[2]

This work utilizing SELEX led to the discovery of an atrazine riboswitch that switches translation "on" in the presence of the ligand.[2] In this type of regulation, the accessibility of the ribosome binding site is increased by the binding of the aptamer to the ligand and the resulting conformational change in the expression platform. For an "on" switch, this means that when the ligand is present, ribosomes bind more efficiently to the mRNA and translation increases (Figure 6).

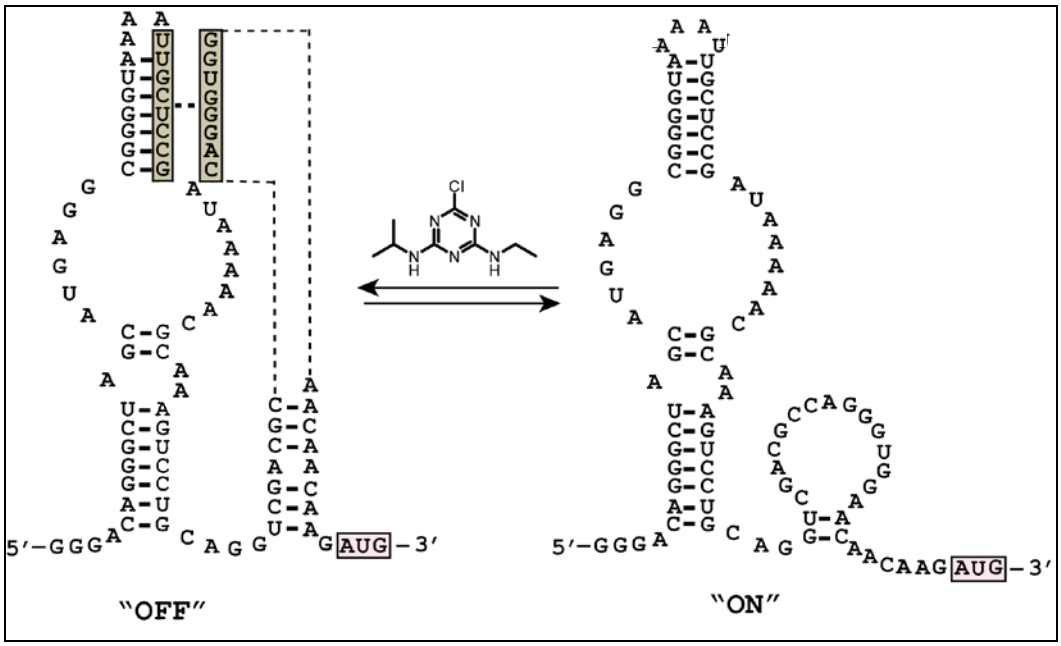

Figure 6. Secondary structure of the atrazine "on" switch at the translational level, showing the increased accessibility of the ribosome binding site upstream of the start codon (AUG) in the presence of the ligand, which allows for increased translation to occur. [2]

The drawbacks to this original switch, however, are that it does not effectively inhibit gene expression in the absence of the ligand, and that there is only a five-fold activation ratio in the presence of atrazine. While this riboswitch regulates gene expression as a result of the presence or absence of atrazine, an increased activation ratio would be beneficial for more demanding screening and application purposes. The main purpose of my research project is to improve the existing atrazine riboswitch to provide a greater dynamic range, the difference in gene expression in the presence and absence of the ligand, and to make it more suitable for eventual real world applications. If the binding of atrazine can cause a significant change in the expression of reporter genes, it can also be made to control the expression of *atzA*, the gene that encodes atrazine chlorohydrolase, which degrades atrazine to the less harmful hydroxyatrazine.[2] By designing a switch with a larger dynamic range, it is possible to create bacteria that would more efficiently metabolize residual atrazine in the environment through the riboswitch control of *atzA*. Using the riboswitch to control this gene, rather than having the gene expressed all the time, allows the cell to produce atrazine chlorohydrolase only when it is needed in the environment to degrade atrazine.

Improving the Atrazine Riboswitch

There are a number of ways to improve a synthetic riboswitch. If little is known about the secondary structure of the mRNA aptamer, one method involves randomizing the number of bases in the expression platform (Figure 7). The Gallivan lab has previously shown that the randomization of a number of nucleotides in a theophylline riboswitch expression platform coupled with a high throughput screening selection improved the activation ratio of the switch by an order of magnitude.[21,22,23] The goal behind this randomization is to optimize the conformational changes that occur when the aptamer binds to a ligand. Instead of trying to design an ideal expression platform one sequence at a time, a library of possible combinations allows large scale screening for efficient switches through the use of bacterial motility screenings, β-galactosidase activity assays, or other reporter assays.

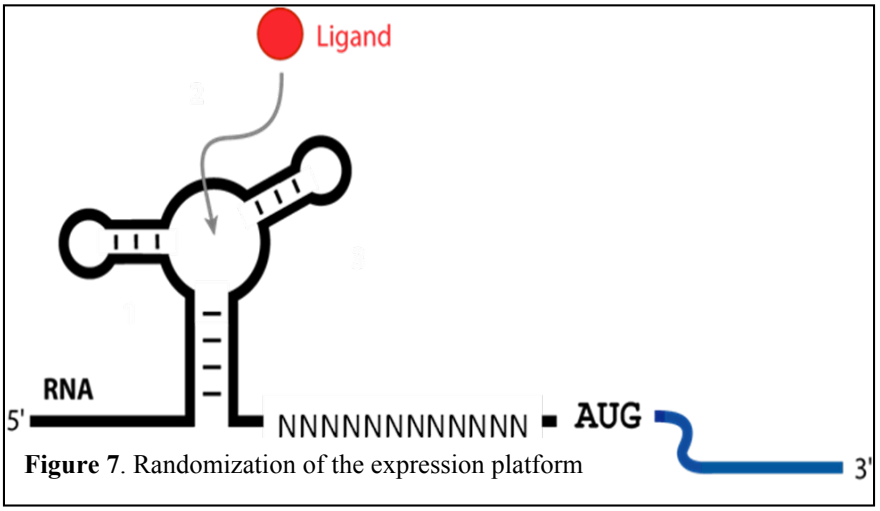

Figure 7. Randomization of the expression platform

The initial screening step for selection of an expression platform takes advantage of chemotaxis, a process in which bacteria direct their movements according to specific chemicals present in their environment. Several proteins are involved in bacterial chemotaxis, and of these proteins, *cheZ* is unique because knocking out this gene prevents bacterial chemotaxis all together.[24] Using this method, the library obtained from randomization of the expression platform can be cloned in front of *cheZ* and transformed

into *E. coli* cells.[22] The cells are then plated in the absence of atrazine, and cells that are immobile are picked and grown on a plate containing atrazine (Figure 8B). In the positive selection step, the cells that move the furthest in the presence of atrazine are picked as these likely contain a functional riboswitch (Figure 8A). These putative riboswitches can then be isolated and cloned upstream of the β-galactosidase (*lacZ)* gene.

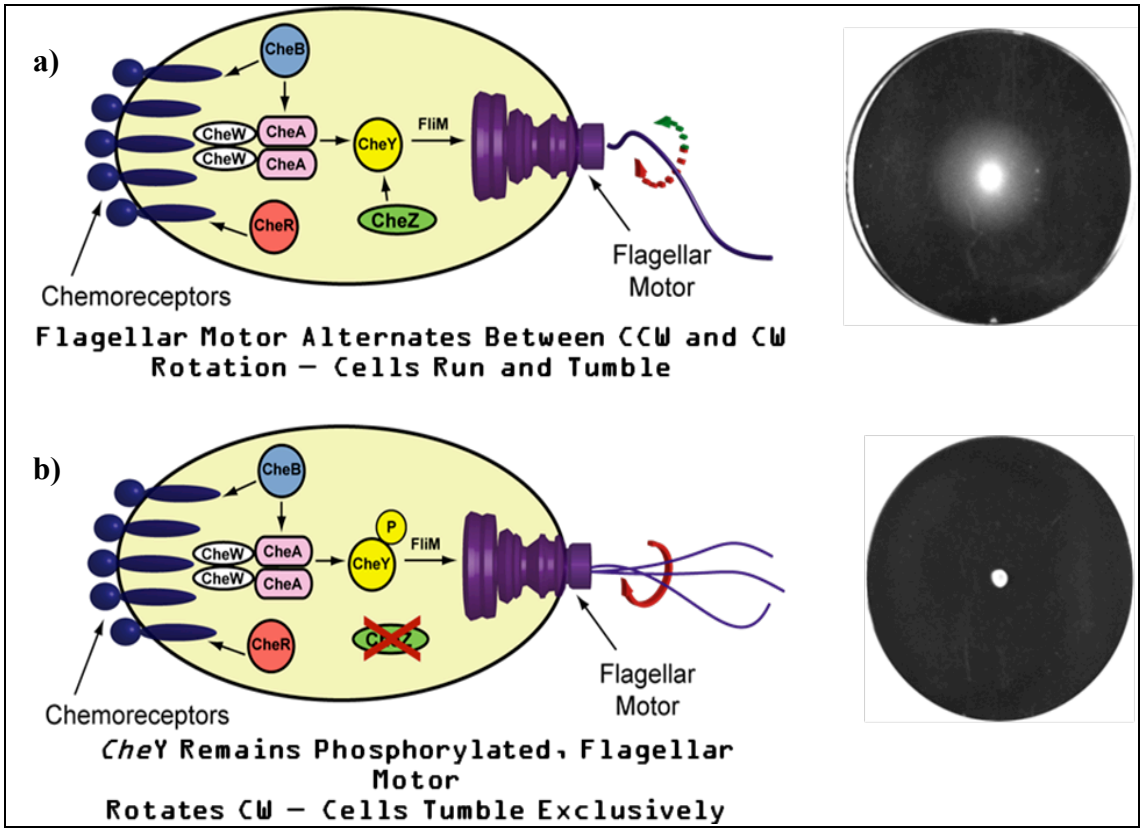

Figure 8. The functionality of *cheZ* determines whether or not the bacterial cells move or simply tumble in place. Riboswitches can be used to control the expression of this gene, and therefore the motility of the cells. **a)** When *cheZ* is expressed, the flagella can rotate counterclockwise, allowing movement of the cells from the center of the plate. **b)** When *cheZ* is knocked out, however, the flagella can only rotate clockwise and cells can only tumble in place.

The next step is to quantify this riboswitch activity using a β-galactosidase activity assay. Here, cells are plated in the presence of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), the substrate for the enzyme β-galactosidase, which causes the

colonies to develop a blue color if β-galactosidase is present. In the absence of atrazine, cells harboring the improved switches collected from the chemotaxis assay are expected to show very little *lacZ* activity (little to no β-galactosidase is produced) and thus these colonies are white. By analogy to the bacterial motility screening, colonies can be picked and grown in the presence or absence of atrazine to determine a change in gene expression. In this method, the cells grown in both the presence and absence of the ligand are subjected to a *lacZ* assay, in which the cells are lysed, and the addition of ortho-nitrophenyl-β-galactoside (ONPG) causes the cells to turn yellow if the *lacZ* gene is expressed (Figure 9).[25]

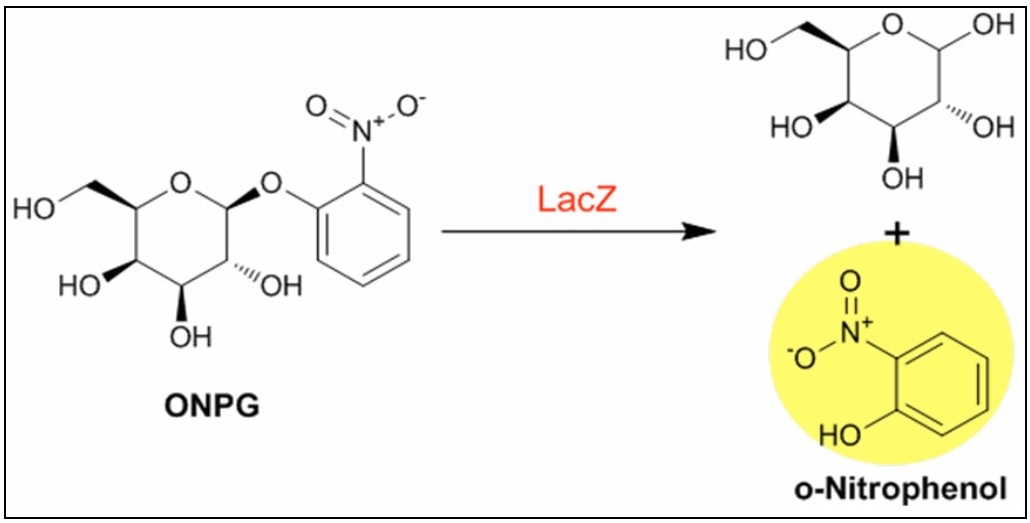

Figure 9. The addition of ONPG in β-galactosidase assays. The presence of LacZ (β-galactosidase) causes ONPG to degrade into o-Nitrophenol, a yellow compound that absorbs light at 420 nm.

Riboswitches are quantified in a *lacZ* assay through the calculation of Miller units (MU), which take into account the optical density of the cells, as shown in the equation below. The ratio of Miller units in the presence versus the absence of the ligand shows effectively how much the ligand causes gene expression to change.

$$MU = OD_{420} / (OD_{600} \times \text{hydrolysis time} \times [\text{volume of cell lysate/total volume}])$$

Another method of improving a synthetic riboswitch by increasing its dynamic range may involve using characteristics of known riboswitches to design specific mutations to the original riboswitch. In this project in particular, the spacing between the aptamer and the ribosome binding site (the length of the expression platform) was explored as shown below in the sequence of the original atrazine riboswitch (Figure 10).

5'Primer Binding Site N40 Region Deleted Nucleotides N10 Linker 3' Primer Binding Site IS10 Region

.

GGGAGAGGGCTAGCATGAGGCGGGGTAAAATTGCTCCGATAAAAACGCAAAGTCCTGCAGGTCGA CGCCAGGGTGGAACAACAAG
ATGTGCGAACTCGATATTTTACACGACTCTCTTTACCAATTCTGCCCCGAATTACACTTAAAACG
ACTCAACAGCTTAACGTTGGCTTGCCACGCATTACTTGACTGTAAAACTCTCACTCTTACCGAAC
TTGGCCGTAACCTGCCAACCAAAGCGAGAACAAAACATAACATCAAACGAATCGCCAAGCTTGGG
GGGATCGATCCC

Figure 10. Original atrazine riboswitch sequence. The nucleotides between the aptamer (the N_{40} region) and the N_{10} linker —highlighted in purple— were deleted one at a time to attempt to optimize the riboswitch for higher activation ratios in the presence of atrazine.

In natural systems, control of gene expression is most efficient when certain parameters are met, and this length of the expression platform is one such parameter. In the original atrazine riboswitch there are fifteen nucleotides between the end of the aptamer and the beginning of the ribosome binding site. Individual constructs can be made in which one base pair at a time is deleted from the expression platform, and these constructs can then be subjected to *lacZ* assays, as with the randomized library.

Experimental

General Considerations

Synthetic oligonucleotides were purchased from Integrated DNA technologies (Coralville, IA). Culture media was obtained from EMD Biosciences, and ampicillin was purchased from Fisher. *E. coli* TOP10F' cells were used for initial transformations, and all cultures were grown at 37 °C in LB broth supplemented with ampicillin (50 μg/mL). Cultures were grown with orbital shaking at 250 rpm until OD_{600} = 0.5-0.7. Transformed cells were permitted to recover for 90 minutes in SOB media under the same conditions before being plated on LB agar with 50 μg/mL ampicillin. These plated cells were incubated at 37 °C overnight, and colonies were picked and grown in LB at 37 °C with orbital shaking at 250 rpm overnight. Plasmids were purified using the QIAprep Spin Miniprep Kit (Qiagen, Inc., Valencia, CA). Enzymes were received from New England Biolabs and were used in the manufacturer's supplied buffers under conditions specified by the manufacturers. Products from enzymatic manipulations were purified using kits from Qiagen. Gels were imaged using a GelDoc imager (Bio-Rad). New constructs were verified by DNA sequencing at MWG Operon (Huntsville, AL). *E. coli* JW1870 cells with the *cheZ* gene knocked out, used for migration experiments, were obtained from Keio collection.[26]

Library Construction

The initial manipulation of the original five-fold atrazine riboswitch began with the construction of a twelve-nucleotide randomized (N_{12}) expression platform library

following the atrazine aptamer obtained after twelve rounds of SELEX.[2] Because the initial riboswitch was created using an N_{10} randomized expression platform, two additional nucleotides were randomized in this experiment in order to increase the number of possible combinations found in the library and thereby increase the potential for positive conformational changes. The constructed library was obtained through three polymerase chain reaction (PCR) experiments. Figure 11 illustrates a general strategy for constructing a riboswitch library using oligonucleotide-based cassette mutagenesis. The assembled riboswitch library is digested with *Kpn*I and *Hind*III, gel purified, and cloned into a vector containing the appropriate reporter gene which has been digested with the same enzymes. To prevent recircularization of the vector, the digested plasmid is dephosphorylated with CIP (Calf Intestinal Phosphatase) prior to the ligation.

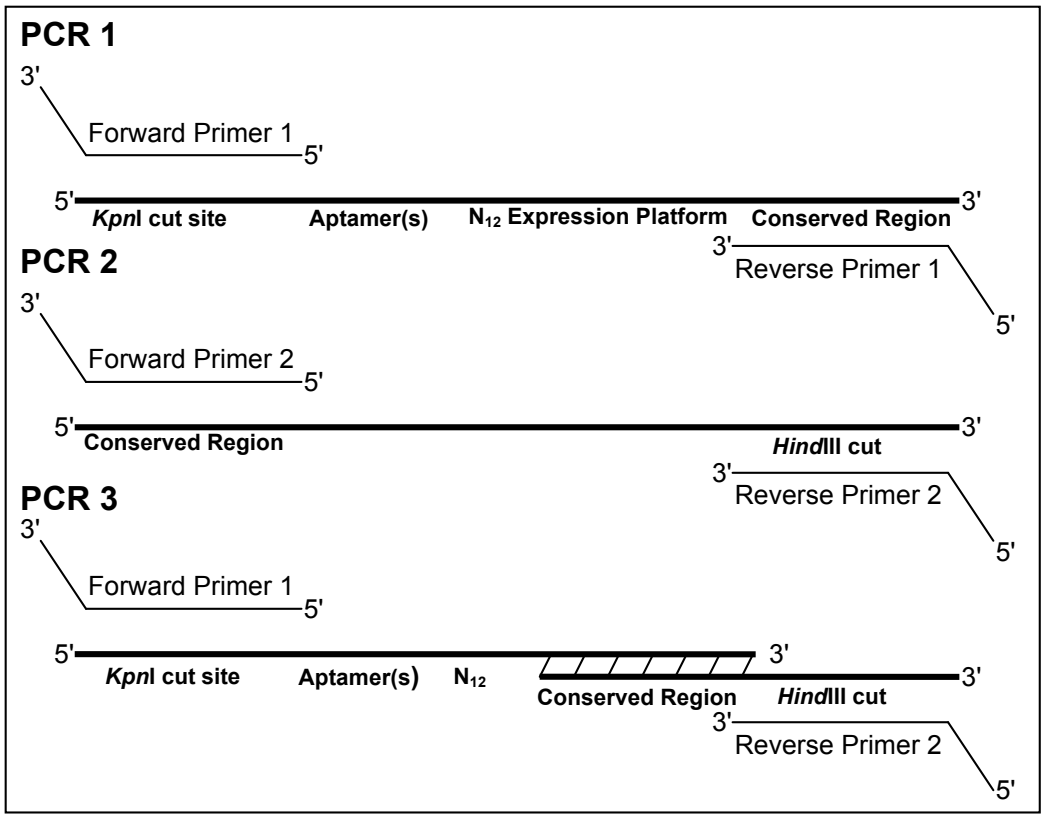

Figure 11. Scheme of three-part PCR reactions. PCR 1 and PCR 2 amplified the two parts of the riboswitch, and the third PCR assembled both previous PCR products to give the final construct containing the randomized region of the expression platform.

In addition to an N_{12} randomization using only the final aptamer, an N_{10} randomized construct was cloned into a pool of possible aptamers. The pool of aptamers came from the ninth round of the original SELEX experiments, before the counter selection with hydroxyatrazine. Returning to the earlier round of selection allowed an increase in the diversity of aptamer possibilities from the N_{12} construct. As with the N_{12} construct, the N_{10} randomization was done in a series of three PCR reactions. The first PCR amplified a *Kpn*I restriction site, the pool of aptamers, the randomized expression platform and a conserved region, the second amplified the conserved region again plus a *Hind*III restriction site, and the third provided the assembly of the two pieces (Figure 12).

a) Forward Primer JS-20 : 5'-CCCCGGTACCGGGACAGGGCTAGC-3'
Reverse Primer SKD-56: 5′-CGACGGGATCGATCCCCCC-3′

b) Reverse Primer JS-112: 5′-GCCATCTTGTTGNNNNNNNNNNNNACTTTGCG-3′
Forward Primer JS-28: 5′-CAACAAGATG-3′

c) Reverse Primer JS-129
5′-CGAGTTCGCACATGCGTCGACCTGCAG-3′
Reverse Primer with Randomized region JS-130
5′-CGAGTTCGCACATNNNNNNNNNNGCGACCTGCAG-3′
Fwd primer for IS-10: JS-131
5′-ATGTGCGAACTCGATATTTTACACG-3′

Figure 12. Primers used in the construction of the N_{12} and N_{10} randomized constructs. **a)** Forward and reverse primers containing the restriction sites which allow the constructs to be inserted into plasmids digested at those same sites. JS-20 contains a *Kpn*I cut site, and SKD-56 contains a *Hind*III cut site. **b)** Reverse primer for the N_{12} amplification (JS-112) and forward primer for the second PCR reaction in the N_{12} construct (JS-28). **c)** N_{10} randomization primers. Reverse primer JS-129 contains the original 3' primer binding site and no randomization, reverse primer JS-130 contains the N_{10} randomization, and forward primer JS-131 amplifies the second region for the N_{10} randomization, containing an IS-10 promoter.

Prior to transformation into plasmids containing the desired reporter gene, the ligation reactions were precipitated with 1-butanol, and then washed with cold 70 % ethanol.

Finally, the pellet was dissolved in 10 μL water, and 1 μL of the ligation reaction was used to transform 40 μL of TOP10F' cells via electroporation.

Screening via bacterial migration experiments

To perform the bacterial migration screening, the N_{10} construct was first transformed into TOP10F' *E. coli* cells and plated on LB agar. These cells were then scraped off into LB media and incubated with shaking at 250 rpm at 37 °C for one hour. After the incubation, the plasmids were purified using 3 mL of the cells and then transformed again, this time into JW1870 cells. These cells were plated again on LB agar and incubated overnight at 37 °C. For the first negative selection, the cells were scraped from the plate using Tryptone broth and incubated with shaking at 250 rpm and 37 °C for 30 minutes. One hundred μL of this culture was diluted to 5 mL, and incubated under the same conditions until the cells reached the mid-log phase (OD_{600} = 0.5). The cells were then diluted to OD_{600} = 0.2, 3 μL were applied to the center of a plate containing selective media (Tryptone broth with 0.24 % agar containing no atrazine), and the plate was incubated at 30 °C for 12-16 hours. At this time, the plate was imaged using white light with the GelDoc station to determine the motility of the cells. Because this selection did not contain atrazine, cells that were not motile were desirable for further screening for improved "on" switches. The very center of the ring of cells was therefore picked, inoculated in 5 mL Tryptone broth, and incubated with shaking as before until the cells reached the mid-log phase (OD_{600}=0.5). The second negative selection was done in exactly the same way as the first, starting from the dilution of the OD_{600} value from 0.5 to

0.2. Again, the very center of the ring of cells was picked for the next round, positive selection. This time, in addition to plating cells in the absence of atrazine, cells were

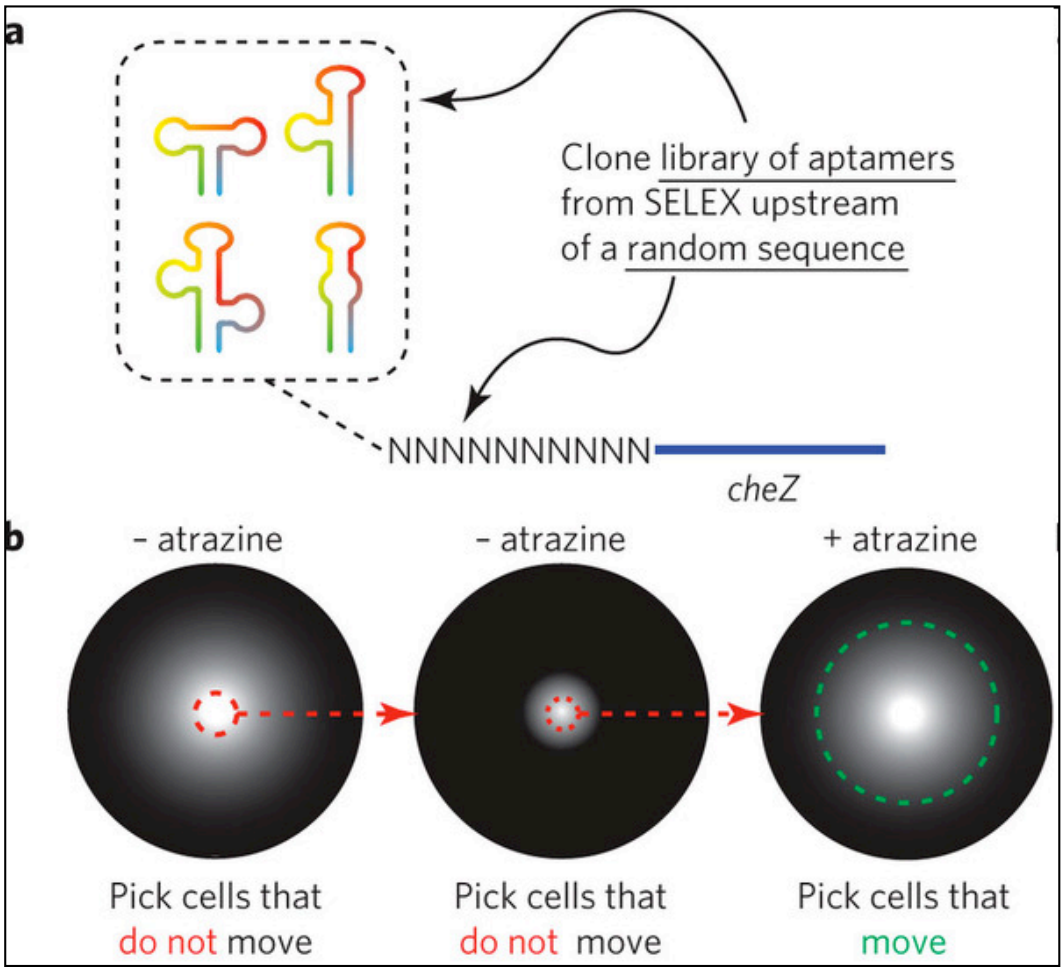

Figure 13. Selection of riboswitches. a) To select riboswitches, a library of atrazine-binding aptamers were cloned in front of an N_{10} randomized expression platform in the 5'-UTR of the *cheZ* gene that controls cell motility. b) Schematic of the selection scheme. Two rounds of negative selection, in which the cells that did not move were picked were followed by a round of positive selection, in which the cells that moved were picked.[2]

plated on selective media agar containing 500 μM atrazine and incubated again for 12-16 hours. This gave a direct comparison for the general movement of cells in the presence and absence of atrazine. The cells in the outermost ring on the atrazine plates were then collected and inoculated in LB media and the plasmids isolated (Figure 13). As a further chemotaxis screening mechanism, the N_{10} library was subjected to a motility screening

using hydroxyatrazine to measure the selectivity of the aptamer pool for atrazine over hydroxyatrazine. This screening was performed in the same way as the positive selection for atrazine, using hydroxyatrazine in place of atrazine. The pool of aptamers showed a greater affinity for atrazine than hydroxyatrazine, and as a result, this pool was further screened using β-galactosidase assays.

Screening via ß-galactosidase assays

The purified plasmids collected after motility screening were digested again with *Kpn*I and *Hind*III, ligated into pSAL172, a plasmid containing the *lacZ* gene, and transformed back into TOP10F' cells, after which they were incubated with shaking overnight and the resulting plasmids were purified. The library of possible riboswitches was then subjected to β-galactosidase assays to screen for improved atrazine switches. For this second type of screening, cells were plated on LB agar containing X-gal (25 mg/mL DMF) and grown overnight at 37 °C to perform an initial blue/white screen. For an "on" switch, low gene expression is desired in the absence of the ligand, atrazine, meaning that the white colonies should be picked for further screening. However, colonies that show no color at all, and therefore no reporter gene expression, are most often re-circularized vectors that do not contain the inserted construct. For this reason, ninety-five light blue (indicating a small amount of gene expression) colonies were picked by hand from each plate and cultured overnight in 200 μL of LB media in a 96-well plate.

For the β-galactosidase assay, 2.5 μL of each overnight culture was added to 200 μL of LB media (containing either 750 μM atrazine or no atrazine) and incubated at

37 °C until it reached mid-log phase. At this point, the OD_{600} values were taken for all of the cultures, and the 96-well plates were placed on ice for ten minutes to prevent cell growth. Cells were lysed by adding 20 µL of Pop Culture™ to each well and pipeting up and down three times. The plates were then incubated for 10 minutes at 30 °C, 15 µL of the lysed cells were added to 133 µL of Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, and 50 mM β-mercaptoethanol, pH 7), and again the plates were incubated for 10 minutes at 30 °C. At this point, 29 µL of ONPG (4 mg/mL in 0.1 M phosphate buffer, pH 7) was added to the cells, and the plates were allowed to incubate for approximately 20 minutes, or until a yellow color appeared in some or all of the wells. Finally, the reaction was stopped by the addition of 75 µL of 1M Na_2CO_3, and the OD_{420} readings were taken. The relative expression of the *lacZ* gene was then calculated in the presence and absence of atrazine by determining the Miller Units of each well for both conditions using the equation above. The activation ratios of the potential switches were determined by calculating the ratio of the MU of each well in the presence and absence of atrazine.

For any of the wells that indicated the presence of a riboswitch with an activation ratio higher than 2, β-galactosidase assays were repeated on a larger scale and in triplicate. Two µL from these wells were added to 5 mL of LB media and incubated with shaking overnight. From these overnight cultures, 2 µL were added to 1.5 mL of LB media and to 1.5 mL LB containing 750 µM atrazine. Cells were incubated with shaking until they reached the mid log phase (OD_{600} = 0.5). At this point, 200 µL of the cells were added to 800 µL Z-buffer, 20 µL chloroform, and 10 µL 0.1% sodium dodecyl sulfate (SDS) and vortexed in order to lyse the cells. After incubation at 30 °C for ten

minutes, 200 µL of ONPG (4 mg/mL) were added to the mixture and each culture was again incubated at 30 °C until yellow color formed in the tubes, at which point 500 µL of Na_2CO_3 were added in order to stop the reaction. OD_{600} and OD_{420} values were obtained and the Miller Units calculated and activation ratio determined.

Controlling the spacing between the aptamer and the start codon

In addition to the randomized constructs, a series of constructs were created in which the manipulation of the expression platform was deliberate rather than random. Specifically, these constructs explored the range of spacing possible between the aptamer and the ribosome binding site. Each construct in the series was made to contain one less nucleotide in this region than the previous construct (Figure 14).

a)
JS 149: 5′-CATCTTGTTGTTCCACCCTCCGTCGACCTGCAGG-3′
JS 150: 5′-CATCTTGTTGTTCCACCCTCGTCGACCTGCAGG-3′
JS 151: 5′-CATCTTGTTGTTCCACCCTCTCGACCTGCAGG-3′
JS 152: 5′-CATCTTGTTGTTCCACCCTCCGACCTGCAGG-3′
JS 153: 5′-CATCTTGTTGTTCCACCCTCGACCTGCAGG-3′
JS 154: 5′-CATCTTGTTGTTCCACCCTCACCTGCAGG-3′
JS 155: 5′-CATCTTGTTGTTCCACCCTCCCTGCAGGACTTTGC-3′
JS 156: 5′-CATCTTGTTGTTCCACCCTCCTGCAGGACTTTGC-3′
JS 157: 5′-CATCTTGTTGTTCCACCCTCGCAGGACTTTGC-3′
JS 158: 5′-CATCTTGTTGTTCCACCCTCCAGGACTTTGC-3′
JS 159: 5′-CATCTTGTTGTTCCACCCTCAGGACTTTGC-3′
JS 160: 5′-CATCTTGTTGTTCCACCCTCGGACTTTGC-3′
JS 161: 5′-CATCTTGTTGTTCCACCCTCGACTTTGC-3′
JS 162: 5′-CATCTTGTTGTTCCACCCTCACTTTGC-3′
JS 163: 5′-CATCTTGTTGTTCCACCCTCCTTTGC-3′

b)
JS 147: 5′-CAACAAGATGTGCGAACTCG-3′

Figure 14. Primers specific to the constructs to decrease the length of the expression platform. **a)** Reverse primers for the first PCR step to adjust the length of the expression platform. The underlined region indicates the N_{10} randomized region from the original atrazine riboswitch. Beginning from the nucleotide directly following this region, each subsequent primer has one fewer nucleotide than the last. **b)** Forward primer for the second PCR step in adjusting the length of the expression platform. This second section was conserved for all of the constructs.

Like the randomized constructs, these were created using PCR, however, each specific construct was made separately. The first PCR for all the constructs, as in the N_{12} randomized construct, amplified the aptamer of the original atrazine riboswitch along with a short conserved region. The second PCR for each construct was different, employing the same forward primer (JS147) for each, but unique reverse primers that systematically deleted one nucleotide at a time during the PCR reaction (Figure 14). The third PCR reaction, like the randomized constructs, assembled the two pieces with *Kpn*I and *Hind*III restriction sites at the ends. Like the other constructs, these were then digested, ligated into pSAL172, and transformed into TOP10F' *E.coli* cells. Because only one sequence was present for each of these constructs, bacterial motility screening was not necessary. Instead, β-galactosidase activity assays were performed as described above on the larger scale for these constructs to screen for improved atrazine riboswitches.

Results and Discussion

In vivo screening for switches

The first step towards improving the existing atrazine riboswitch was to create a library with an N_{12} randomized expression platform cloned downstream of the aptamer discovered through the SELEX process. This N_{12} library allowed for greater flexibility in the formation of the expression platform and therefore a greater possibility that some randomized combination could give a more favorable conformational change upon the binding of the ligand to the aptamer. This library was inserted into a plasmid containing the *lacZ* gene such that the gene was under the direct control of the library of possible riboswitches. The production of protein from the *lacZ* gene was measured using a β-galactosidase assay. By monitoring the expression of the gene upon addition of ONPG, it is possible to monitor the switching abilities of the randomized library as well. The screening process for this assay allows for 95 possibilities to be screened at a time on a small scale using a 96-well plate with one negative control.

The first plate of 95 colonies indicated the presence of two potential "on" switches, one of which appeared to be a two-fold switch, and the other a four-fold switch (Table 1). Further investigation of these possibilities on a larger scale, however, did not show any switching, both showing only about 0.1 times more gene expression

Activation Ratios of Potential Switches

	Activation Ratio	
	Construct 1	Construct 2
Trial 1	4.2	2.6
Trial 2	3.8	2.8

Table 1. Activation ratios of the two potential switches found in the initial β-galactosidase assays for the N_{12} construct

in the presence of atrazine than in the absence. Two more plates of 95 colonies were screened using the same β-galactosidase assay procedures, but no potential riboswitches were discovered. Because the plate assays take place on such a small scale, it is difficult to be sure that the cells are being completely lysed, as the wells cannot be vortexed without contamination from each other. As a result of the potentially uneven lysing of the cells, it is not uncommon for this type of assay to overestimate the effects of the riboswitch on the production of the *lacZ* protein. For this reason it is important to repeat these assays on a larger scale and with multiple replicates if any switching is suspected.

Although it is not possible to screen every single combination of nucleotides possible in a randomization as large as N_{12}, multiple screenings should give some indication of whether or not the method is effective in improving the existing riboswitch. As nearly 300 colonies showed no improvement over the existing switch, a different method was attempted next. Rather than creating a library containing a single aptamer and a randomized pool of expression platforms, one was created containing both a pool of possible aptamers and a pool of possible expression platforms. Because of the greater diversity of the N_{10} library, a bacterial motility assay was performed as an initial step to narrow the focus of the screening. JW1870 contains a *cheZ* mutation and is unable to swim. In this case, our constructed inserts use the riboswitch to control this gene. This means that if an "on" switch is present, cells should not move in the absence and move in the presence of atrazine. Returning to the 9^{th} round of the SELEX experiments (before the counter selection with hydroxyatrazine) reintroduced some diversity into the aptamer pool, allowing for a wider range of possibilities from which to screen. In addition, a control was constructed which contained the aptamer with no expression platform at all.

This control was set up to confirm that the expression platform was necessary in making a difference in the expression of the genes.

Both the control experiment and the N_{10} randomization were subjected to a bacterial motility screening as an initial way of determining what was present in the plasmid pool and selecting those constructs which were more likely to respond to atrazine for the expression of the associated genes. In both cases, the cells were plated in the absence of atrazine for two negative selections, and the non-motile cells were picked. After two rounds of selecting cells that did not move in the absence of atrazine, cells were plated again the presence and absence of atrazine. In both cases, cells moved more in the presence of the ligand than in the absence (Figure 15).

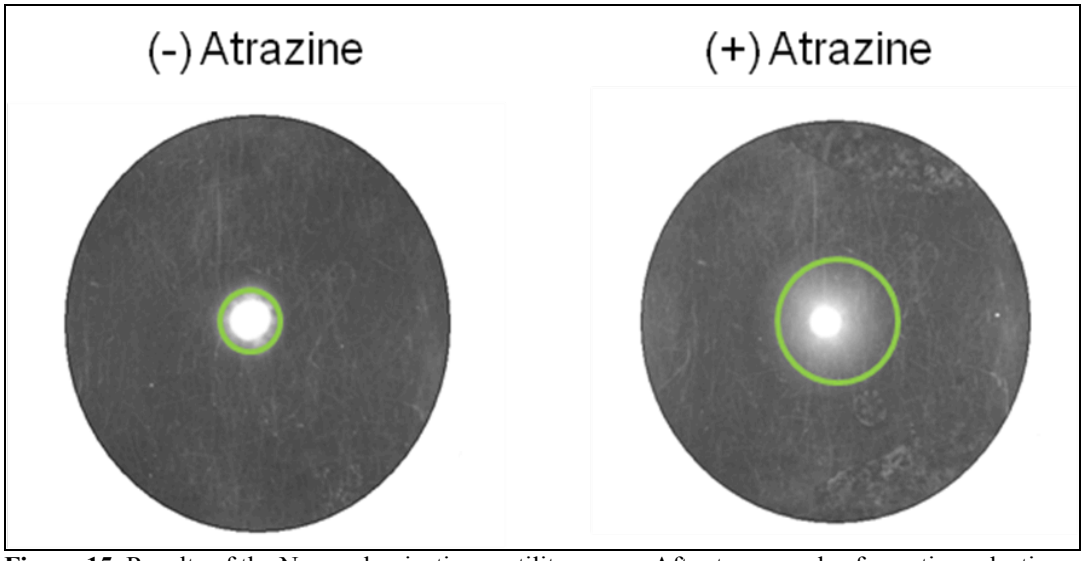

Figure 15. Results of the N_{10} randomization motility assay. After two rounds of negative selection, cells showed greater motility in the presence of atrazine than in the absence.

After the positive selection, the cells from the outer ring were picked for further screening via β-galactosidase assays. In the same way as the N_{12} construct, these new libraries were subjected to multiple rounds of screenings. The control experiment, as

expected, did not show any evidence of significant switching, further confirming the importance of the expression platform in controlling gene expression.

Although the chemotaxis assay for the N_{10} randomization indicated the presence of candidate riboswitches, the first β-galactosidase assays did not indicate any improved switches (data not shown). One possible explanation comes from a possible weakness in the method of chemotaxis screening. The gene which is controlled by the riboswitch, *cheZ*, controls the rotation of the flagella of the bacteria. When *cheZ* is not expressed in the cells, the flagella rotates counterclockwise, causing the cells to simply tumble in place. When the gene is expressed, the flagella rotate clockwise, resulting in movement of the cells. This is why controlling the expression of *cheZ* allows movement to be regulated in the chemotaxis screenings. However, if *cheZ* is over-expressed, the cells never tumble and get stuck in the agar. This means that within the library plated in these motility screenings, there could be constructs which respond very favorably to atrazine, but which were not selected because the movement was retarded by over-expression of the gene in the presence of atrazine. In addition, because the manipulation of the existing riboswitch was completely random, there is not a way to look for possibilities that might be more favorable without including all possibilities, many of which probably do not provide any favorable changes.

Controlling the spacing between the aptamer and the start codon

Another possibility for improving the riboswitch is to take characteristics already known about riboswitches that work well in nature and to manipulate nucleotides in a purposeful manner to mimic those qualities. In this case, the distance between the

aptamer and the ribosome binding site was manipulated, altering the length of the expression platform. This region is important because for the switch to have a meaningful effect on gene expression, it must be an optimal distance from the site where translation of the genes will begin.

Fifteen separate constructs were created in which the aptamer from the original SELEX experiments was cloned in front of an expression platform that became one nucleotide shorter with each successive construct. In this case, rather than ending up with an entire library of possibilities in one experiment, fifteen separate individual constructs were obtained and tested. The average activation ratio of each construct was between 1.3 and 2, lower than the original riboswitch derived from the SELEX experiments (Figure 16).

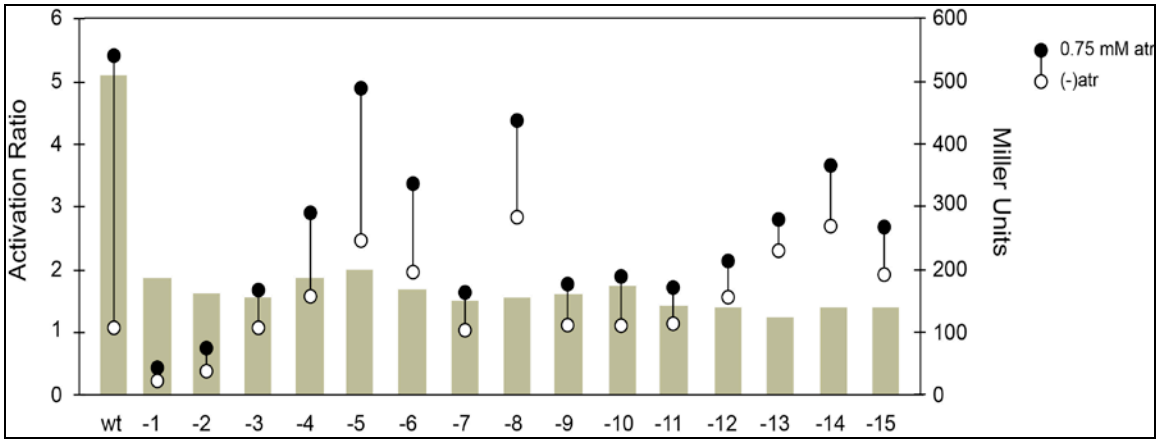

Figure 16. Lynch-Gallivan plot for deletion experiments. As nucleotides were deleted, the activation ratio of the riboswitch dropped initially, then remained unchanged for the rest of the deletions. Each activation ratio is an average of three β-galactosidase assays performed simultaneously.

A possible explanation for this decrease in activation ratio goes back to the secondary structure of the mRNA. As the distance between the aptamer and the ribosome binding site decreases, the ribosome binding site gets increasingly close to the aptamer, increasing the possibility that the bulk of the aptamer's secondary structure could physically

interfere with the binding of the ribosome to the mRNA. This would result in a decrease in the overall gene expression as well as a decrease in the change in gene expression seen upon addition of the ligand (Figure 17).

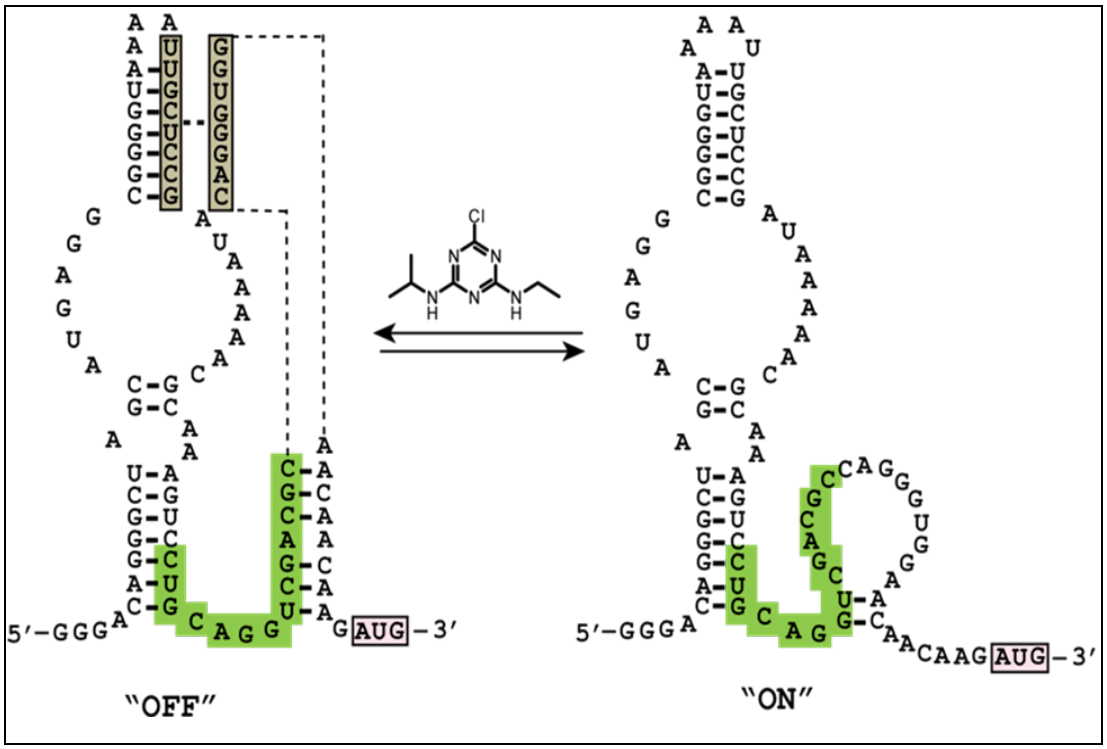

Figure 17. Secondary structure of the aptamer in the absence ("off") and presence ("on") of atrazine. The fifteen nucleotides highlighted in green show the nucleotides deleted one at a time between the aptamer and the start codon (AUG). Deleting nucleotides, besides bringing the start codon and ribosome binding site (highlighted in grey in the "off" state) closer to the aptamer, could also disrupt the base pairing of the ribosome binding site in the "off" state, which could account for the lower activation ratios observed upon deletion.

In addition, there is a possibility that the deletion of these nucleotides, while allowing new secondary structures that could have been more favorable for switching, instead caused a disruption in the original secondary structure that was unfavorable. For example, in the "off" state, the structure shows a great amount of base pairing of the ribosome binding site and the aptamer (highlighted in grey), and deleting base pairs upstream of this region could possibly disturb the paired region. This, in turn, might allow greater accessibility to the ribosome binding site in the absence of atrazine,

resulting in more gene expression and a lower activation ratio. Alternatively, the number of nucleotides between the ribosome binding site and start codon could be altered, instead of the number between the aptamer and the ribosome binding site. In nature, the ideal spacing between the ribosome binding site and the start codon is 8-10 nucleotides. However, there is a possibility that altering this length could allow a secondary structure that would provide a better riboswitch.

Conclusion and Future Directions

Despite early indications that proteins were the sole method of gene regulation, advances in the knowledge of RNA have shown that there are a multitude of ways in which RNA can affect gene expression, possibly hinting back to the idea that at one time, RNA performed all the functions of cell metabolism, rather than functioning as just genetic storage (the "RNA world" hypothesis). One such way is through riboswitch binding to specific metabolites. Previous work led to the creation of a riboswitch that responds to the herbicide, atrazine.[2] However, because this riboswitch showed only an approximately 5-fold increase in gene expression upon addition of atrazine, the original construct was manipulated to try to improve the switch, so that it could have the useful application of removing excess herbicide from the environment. Manipulations to the switch via randomization of the expression platform not only showed no increase in the switching abilities of the original switch, but in fact showed a decrease in activation ratio. Similarly, purposeful manipulation of the distance between the aptamer and the ribosome binding site caused a sharp decrease in the performance of the riboswitch, rather than an improvement. However, these manipulations are not the only possibilities that could improve the riboswitch.

Specific engineering of the riboswitch can also take advantage of the secondary structure of the aptamer and expression platform. In addition to the ribosome binding site and start codon (AUG) that are in frame with the *lacZ* gene, the original atrazine switch contains a second start codon which is out of frame, along with a possible ribosome binding site upstream from this codon (Figure 18). By deleting the first start codon, along with the two preceding nucleotides, the second start codon can be put in frame with

the *lacZ* gene. In the original riboswitch, this second start codon and ribosome binding site are base-paired with other nucleotides when atrazine is present. In the absence of atrazine, these structures are less tightly bound into a secondary structure. This presents the opportunity to screen for an "off" switch rather than an "on" switch, which could be useful for potential screening applications.

5′PBS-N40 Region-N10 linker-3′PBS-IS10 Region

GGGAGAGGGCTAGCATGAGGCGGGGTAAAATTGCTCCGATAAAAACGCAAAGTCCTGCAGGT
CGACGCCAGGGTGGAACAACAAG
ATGTGCGAACTCGATATTTTACACGACTCTCTTTACCAATTCTGCCCCGAATTACACTTAAA
ACGACTCAACAGCTTAACGTTGGCTTGCCACGCATTACTTGACTGTAAAACTCTCACTCTTA
CCGAACTTGGCCGTAACCTGCCAACCAAAGCGAGAACAAAACATAACATCAAACGAATCGCC
AAGCTTGGGGGGATCGATCCC

Figure 18. Creation of an "off" switch from the original atrazine riboswitch. The original atrazine switch utilizes the start codon shown in light blue, which is in frame with the IS10 promoter region that contains the *lacZ* gene. However, there is another possible start codon, shown in purple, which could allow for the creation of an "off" switch. Deleting the original start codon, plus the two nucleotides preceding it (in yellow) puts the new start codon in frame with the IS10 region. This start codon and the AG rich region upstream of it that would serve as the ribosome binding site are more tightly bound in the presence of atrazine than in the absence, allowing the possibility of an "off" switch.

In this scenario, the *cheZ* gene would be expressed in the absence of atrazine, allowing the *E.coli* to move around in its environment until atrazine is present. At this point, the switch would change conformations, turning off expression of *cheZ*, and causing the cells to stop moving. While the bacteria are stopped, they can be made to express the atrazine chlorohydrolase (*atzA*) gene through riboswitch control, which catabolizes atrazine to hydroxyatrazine. When the atrazine is depleted, *cheZ* would again be expressed, and the cells would move on until atrazine was again present in the environment.

Riboswitches were only recently discovered as a method of gene-expression control, and the many functions of RNA in the cell are still being uncovered.

Biochemical systems are extremely complicated and are influenced by a multitude of factors, some of which are yet to be understood. Although initial trials did not find a switch with a higher activation ratio than the current switch, further manipulation of this mRNA could not only improve this riboswitch, but improve the understanding of what factors affect the regulatory abilities of riboswitches in general. Improvement of the atrazine riboswitch through these or other methods of manipulation could have a potential long term application of removing an extremely toxic pollutant from soil and thereby cleaning our environment.

References

1. Mandal, M. & Breaker, R. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* **5**, 451-463 (2004).
2. Sinha, J., Reyes, S. & Gallivan, J. Reprogramming bacteria to seek and destroy an herbicide. *Nat. Chem. Biol.* **6**, 464-470 (2010).
3. Walter Gilbert & Muller-Hill, B. Isolation of Lac Repressor. *Proc. Natl. Acad. Sci. U S A* **56**, 1891 (1966).
4. Kruger, K. et al. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31**, 147-157 (1982).
5. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849-857 (1983).
6. Lundrigan, M.D., Köster, W. & Kadner, R.J. Transcribed sequences of the Escherichia coli btuB gene control its expression and regulation by vitamin B12. *Proc. Natl. Acad. Sci. U S A* **88**, 1479 -1483 (1991).
7. Ravnum, S. & Andersson, D.I. Vitamin B_{12} repression of the *btuB* gene in *Salmonella typhimurium* is mediated via a translational control which requires leader and coding sequences. *Mol. Microbiol.* **23**, 35-42 (1997).
8. Richter-Dahlfors, A.A., Ravnum, S. & Andersson, D.I. Vitamin B_{12} repression of the cob operon in Salmonella typhimurium: translational control of the cbiA gene. *Mol. Microbiol.* **13**, 541-553 (1994).
9. Nou, X. & Kadner, R.J. Adenosylcobalamin inhibits ribosome binding to btuB RNA. *Proc. Natl. Acad. Sci. U S A* **97**, 7190 -7195 (2000).
10. Miranda-Rios, J., Navarro, M. & Soberon, M. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl. Acad. Sci. U S A* **98**, 9736-9741 (2001).
11. Webb, E., Claas, K. & Downs, D. thiBPQ encodes an ABC transporter required for transport of thiamine and thiamine pyrophosphate in Salmonella typhimurium. *J. Biol. Chem.* **273**, 8946-8950 (1998).
12. Gelfand, M., Mironov, A., Jomantas, J., Kozlov, Y. & Perumov, D. A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* **15**, 439-442 (1999).
13. Mironov, V., Perumov, D., Kraev, A., Stepanov, A. & Skryabin, K. Unusual structure in the regulation region of the Bacillus-subtilis riboflavin biosynthesis operon. *Mol. Bio.* **24**, 221-225 (1990).
14. Grundy, F. & Henkin, T. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol. Microbiol.* **30**, 737-749 (1998).
15. Ebbole, D. & Zalkin, H. Cloning and characterization of a 12-gene cluster from Bacillus-subtilis encoding 9 enzymes for denovo purine nucleotide synthesis. *J. Biol. Chem.* **262**, 8274-8287 (1987).
16. Christiansen, L., Schou, S., Nygaard, P. & Saxild, H. Xanthine metabolism in Bacillus subtilis: Characterization of the xpt-pbuX operon and evidence for purine- and nitrogen-controlled expression of genes involved in xanthine salvage and catabolism. *J. Bacteriol.* **179**, 2540-2550 (1997).

17. Vold., B., Szulmajster, J. & Carbone, A. regulation of dihydrodipicolinate synthase and aspartate kinase in Bacillus-subtilis. *J. Bacteriol.* **121**, 970-974 (1975).
18. ATSDR - Toxic Substances - Atrazine. at <http://www.atsdr.cdc.gov/substances/toxsubstance.asp?toxid=59>
19. Duhigg, C. Debating How Much Weed Killer Is Safe in Your Water Glass in *New York Times* A1 (New York, August 23, 2009).
20. de Souza, M.L., Wackett, L.P. & Sadowsky, M.J. The atzABC Genes Encoding Atrazine Catabolism Are Located on a Self-Transmissible Plasmid in Pseudomonas sp. Strain ADP. *Appl. Environ. Microbiol.* **64**, 2323-2326 (1998).
21. Desai, S. & Gallivan, J. Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.* **126**, 13247-13254 (2004).
22. Topp, S. & Gallivan, J. Guiding bacteria with small molecules and RNA. *J. Am. Chem. Soc.* **129**, 6807-6811 (2007).
23. Gallivan, J. Toward reprogramming bacteria with small molecules and RNA. *Curr. Opin. Chem. Bio.* **11**, 612-619 (2007).
24. Topp, S. & Gallivan, J.P. Random Walks to Synthetic Riboswitches—A High-Throughput Selection Based on Cell Motility. *ChemBioChem.* **9**, 210-213 (2008).
25. Miller, J. H. (1972) *Experiments in molecular genetics,* p. xvi, 466 p. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, (2006).