SYNTHETIC CAFFEINE RIBOSWITCHES AS GENETICAL AND ANALYTICAL TOOLS

by

Matt Sherman

Submitted in partial fulfillment of the
requirements for Departmental Honors in
the Department of Chemistry
Texas Christian University
Fort Worth, Texas

May, 2, 2014

SYNTHETIC CAFFEINE RIBOSWITCHES AS GENETICAL AND ANALYTICAL TOOLS

Project Approved:

Youngha Ryu, Ph.D.
Department of Chemistry
(Supervising Professor)

Jean-luc Montchamp, Ph.D.
Department of Chemistry

Giri Akkaraju, Ph.D.
Department of Biology

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I want to thank Dr. Youngha Ryu for helping me through two years of research and all the mistakes and questions I had along the way. As well Pradeep Budhathoki was a huge help to me in learning and trying to understand how to perform all the different procedures of the lab and source to turn to when I was utterly confused. I would like to thank the College of Science and Engineering Undergraduate Research Grant.

INTRODUCTION

Riboswitches are 5'-unstranslated regions of mRNA molecules that bind to specific ligands and regulate expression of downstream genes associated with the biosynthesis, transport, or degradation, of the specific molecule[1]. Natural riboswitches are found most commonly in bacteria but they are also known to exist in eukaryotes such as plants and fungi. Natural riboswitches have been identified for many primary cellular metabolites including thiamine pyrophosphate (TPP), flavin mononucleotide (FMN), vitamin B12, S-adenosylmethionine (SAM), glutamine, glycine, lysine, adenine, glucosamine-6-phophate (GlcN6P) and cyclic di-GMP. [2-13] Developed in the laboratory synthetic riboswitches offer a new tool to detect target ligands including small molecules, proteins, nucleic acids, and cells for analytical, biological, and therapeutic applications.[14]

A riboswitch contains two different domains, the aptamer domain that actually binds the specific ligand in question and the expression platform that regulates downstream gene expression.[15] The aptamer domain is different between different riboswitches as this is the site of binding. As the bases of mRNA interact with the ligand different ligands require a different organization of bases in its aptamer domain. Upon binding to its cognate ligand the aptamer domain undergoes a structural change. This change triggers the structural change in the expression platform. There are three different mechanisms by which the expression platform can regulate the downstream gene expression: (1) transcription termination/anti-termination; (2) ribosome binding/ sequestration; and (3) self-cleavage.[16] For example, the ribosome binding site (RBS) is blocked

in the expression platform, the translation of the downstream gene is stopped. Upon ligand binding the structural change in the aptamer domain can cause the RBS to be exposed and therefore the downstream gene is expressed.[16] Figure 1 shows the conformational change of the theophylline synthetic riboswitch that exposes the RBS upon the binding of theophylline to the aptamer domain.

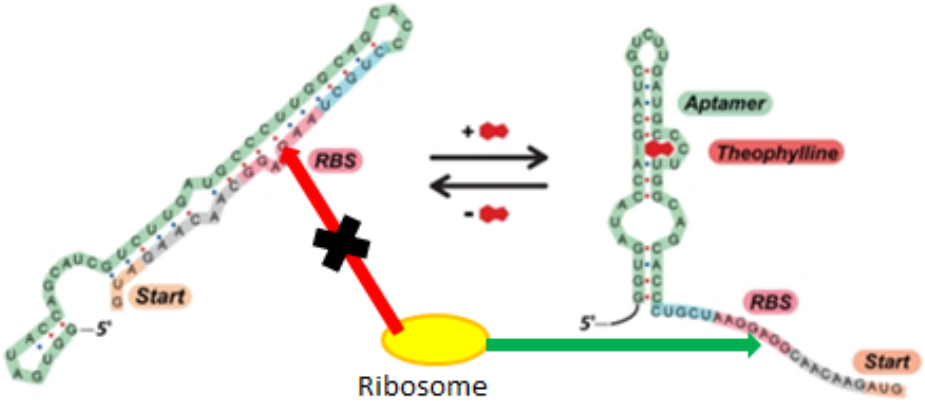


Figure 1: The structural change of the theophylline riboswitch upon biding to theophylline. Upon binding the RBS is exposed to activate the translation of the downstream gene.

A riboswitch capable of binding to caffeine can be used to detect caffeine levels in the body fluids during the treatment of apnea of immaturity. Apnea of immaturity is a condition in some premature babies where they have not developed enough to breathe on their own. Caffeine is used as a stimulant to in essence help them remember to breathe. Monitoring the caffeine intake level is vital because of the danger of caffeine ingestion in newborns.[17,18] Synthetic caffeine riboswitches can also be used for caffeine toxicology and doping analysis[19] Caffeine levels can be used as a marker of raw sewage contamination

in fresh water supply. The only major source of caffeine in waste water is from human consumption and excretion. Caffeine along with fecal matter should be removed from waste water upon treatment. However in cases where there is a leak in pipes fecal matter can move into fresh water along with co-contaminant caffeine.[20]

RESULTS AND DISCUSSION

In order to develop a synthetic caffeine riboswitch a theophylline riboswitch was used as the starting point because of the molecular similarity between caffeine and theophylline. Figure 2 shows the molecular structures of the two molecules.

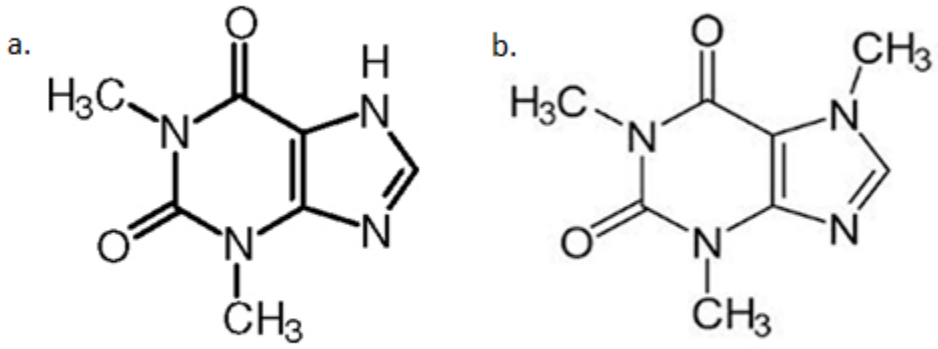


Figure 2: structure of theophylline(a) and caffeine (b)

The theophylline riboswitch is a synthetic riboswitch that was identified in 1994 and has 10,000 times more affinity for theophylline than caffeine as well as having been used as the starting point to develop many other synthetic riboswitches.[21]

According to previous binding studies eleven bases in the aptamer domain of the theophylline synthetic riboswitch interact strongly with theophylline. Figure 3 shows this interaction and the 11 bases that are randomized.

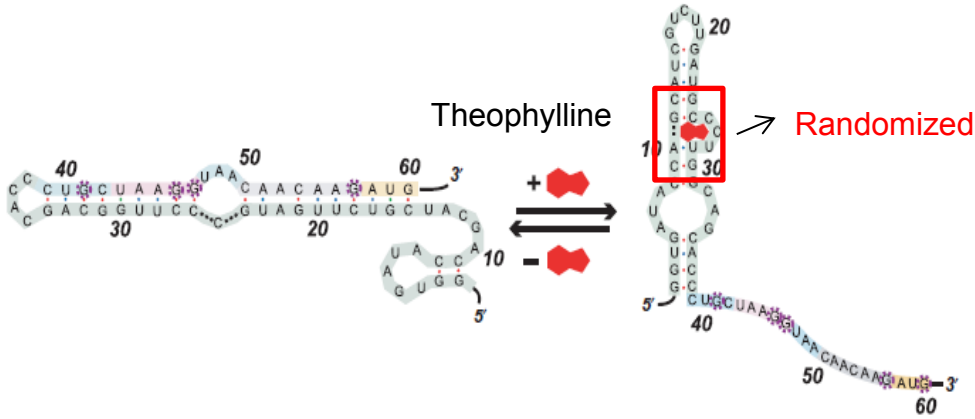


Figure 3: The eleven bases in the aptamer domain of the theophylline riboswitch were randomized to make the library.

These eleven bases were randomized to construct an aptamer library. The library is a collection of all possible base combinations in the selected eleven locations. Since DNA and RNA have 4 different bases in a given position the library contains 4^{11} or about 4.2 million possible variants.

The library was constructed by polymerase chain reaction. This modified PCR procedure uses two sets of primers to insert mutations into the sequence at specific points. [22]

The constructed library was inserted into pTrp-TheoRS-CatUpp that encodes for the *cat-upp* fusion gene under control of the trp promoter. The *cat* gene codes for chloramphenicol resistance and is the key feature of the positive selection. The *upp* part of the fusion protein functions in the negative selection.[23, 24] Figure 4 shows the riboswitch library and the trp promoter that

precedes it.

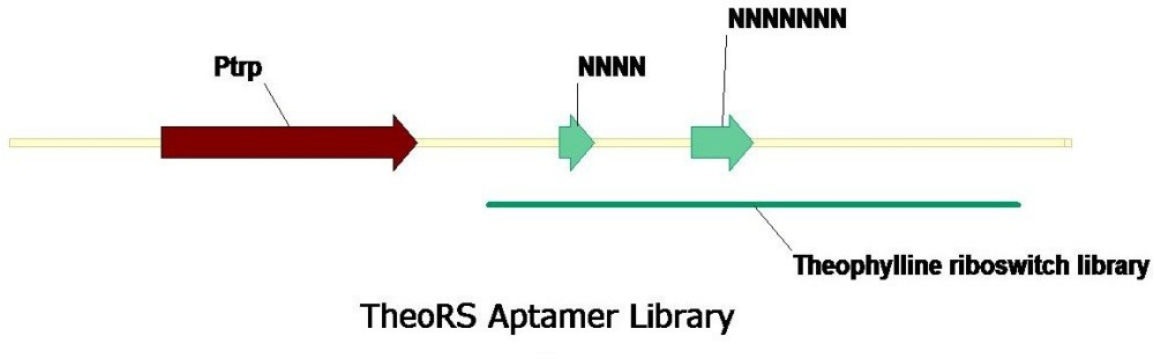


Figure 4: TheoRS Aptamer Library

In order to carry out the dual genetic screening procedure the library plasmid was transformed into the GH371 *E. coli* cells by electroporation. These cells have the ability to survive in high concentrations of 5-FU as they lack the *upp* gene.

Genetic screening, through positive and negative selection, determines if any of the possible 4.2 million variants can selectively bind caffeine. Screening ensures that only those cells harboring a functional caffeine riboswitch survives both the positive and negative selection.[23,24]

In positive selection only those cells containing a functional riboswitch survive because upon the binding of the ligand the downstream chloramphenicol resistance gene is expressed. However, during positive selection the switch can also be triggered by an endogenous cellular metabolite. The negative selection removes the cells containing the riboswitches that bind to any endogenous metabolites.

For positive selection GH371 *E.coli* cells transformed with the library were grown in the presence of caffeine (to trigger the riboswitch) and chloramphenicol as shown in Table 1. The cells were grown with 100µg/mL concentration of carbenicillin to maintain the plasmid. The rest of the agar media were standard components needed for cell growth.

Component	Final concentration
20XP	1 X
Glucose	0.2 %
$MgSO_4$	2 mM
Vitamin solution	1 X
Trace metal solution	1 X
Amino Acids	0.2 %
Carbenicillin	100 µg/mL
Caffeine	2.5 mM
Choloramphenicol	30 µg/mL

Table 1: Concentration of the components of the positive selection plate

Cells survive in the presence of chloramphenicol only if the riboswitch turns on the production of chloramphenicol acetlytransferase, giving the cell chloramphenicol resistance.

For negative selection the same transformed GH371 cells were grown in the presence of 5-fluorouracil. The components and final concentrations of the negative selection plate are shown in Table 2.

Component	Final concentration
20XP	1 X
Glucose	0.2 %
$MgSO_4$	2 mM
Vitamin solution	1 X
Trace metal solution	1 X
Amino Acids	0.2 %
Carbenicillin	100 µg/mL
5-FU	.5 µg/mL

Table 2: Concentration of components of the negative selection plate

During the negative selection the colonies containing the caffeine riboswitch can survive because no *cat-upp* gene is expressed in the absence of caffeine. In contrast in the colonies containing the riboswitches that bind to any endogenous ligands the UPRT protein is expressed from the *upp.* This protein converts 5-FU into 5-FUMP which is toxic to cells. The selection procedure is outlined in Figure 5.

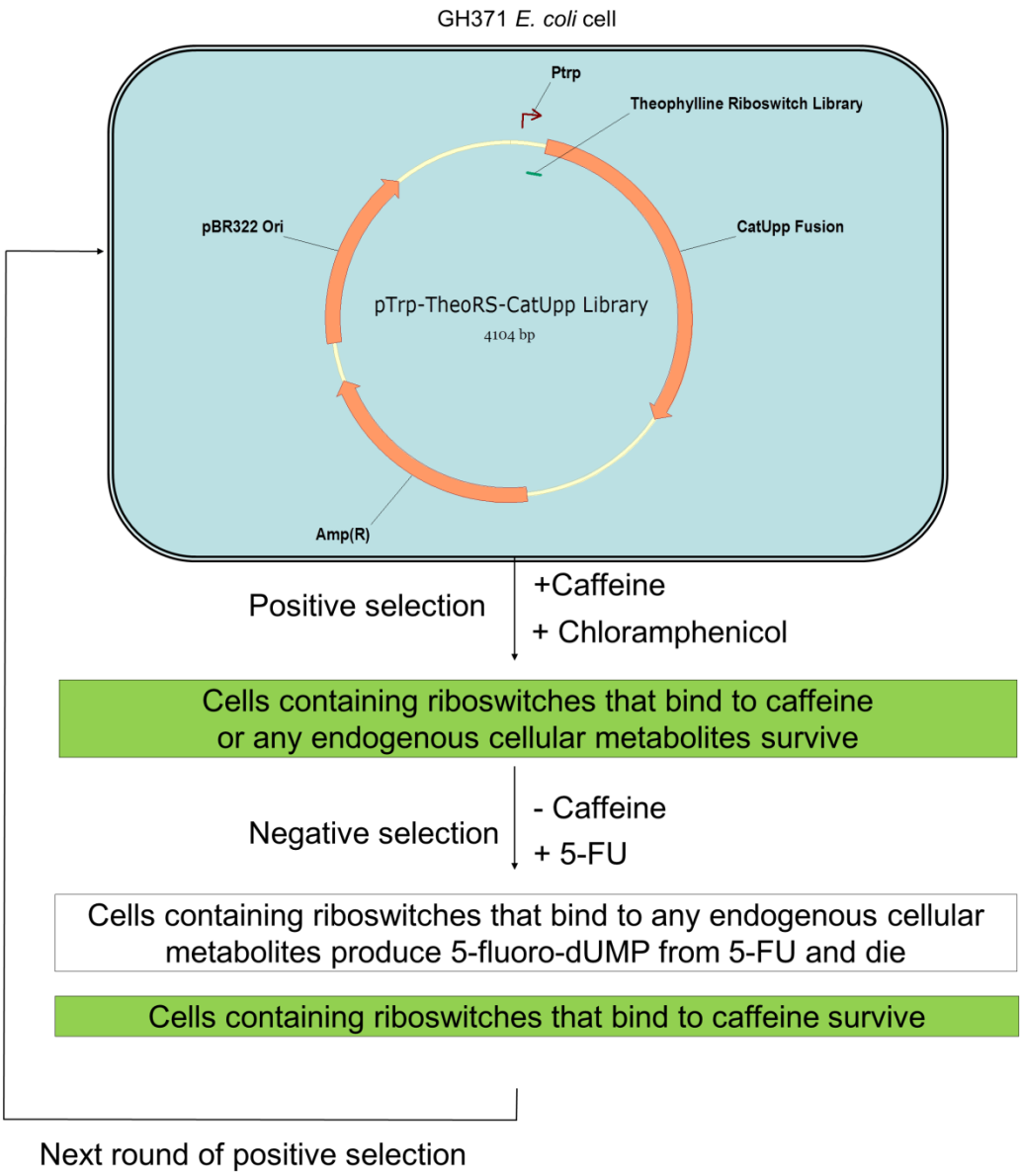


Figure 5: The dual genetic screening procedure

Following the dual genetic selection 192 colonies were randomly picked and grown in 96 well plates. They were then tested for the resistance to chloramphenicol(50 µg/mL), in the presence and absence of caffeine. These plates are shown in Figure 6.

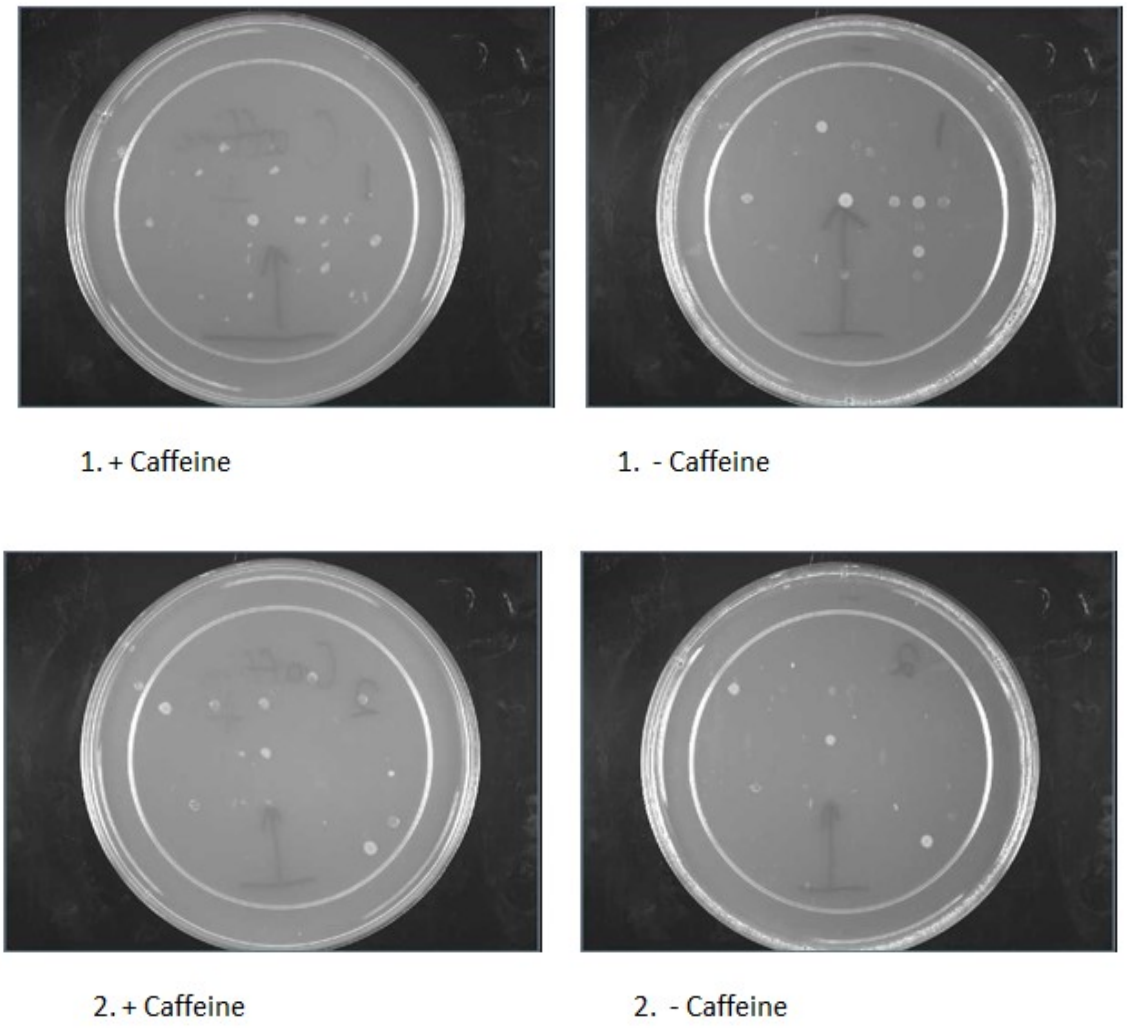


Figure 6: The replica plates of 192 in the presence and absence of caffeine.

The possible hits were the colonies that grew on the agar plate containing caffeine (positive plate) but did not grow on the agar plate without caffeine (negative plate). The 192 colonies from the well plates and the outcome of replica plating are shown in Figure 7. Those that had either minimal growth on the positive plate or growth on both plates are marked purple. These colonies were able to live on both plates possibly because they responded to endogenous

ligands. The two colonies that appeared just on the negative plate are marked red. The remainder of the colonies failed to show growth on either plate. The large number of colonies that failed to grow was probably due to the increased chloramphenicol concentration. The four colonies on the positive plate that showed no growth on the negative plate are marked in light blue.

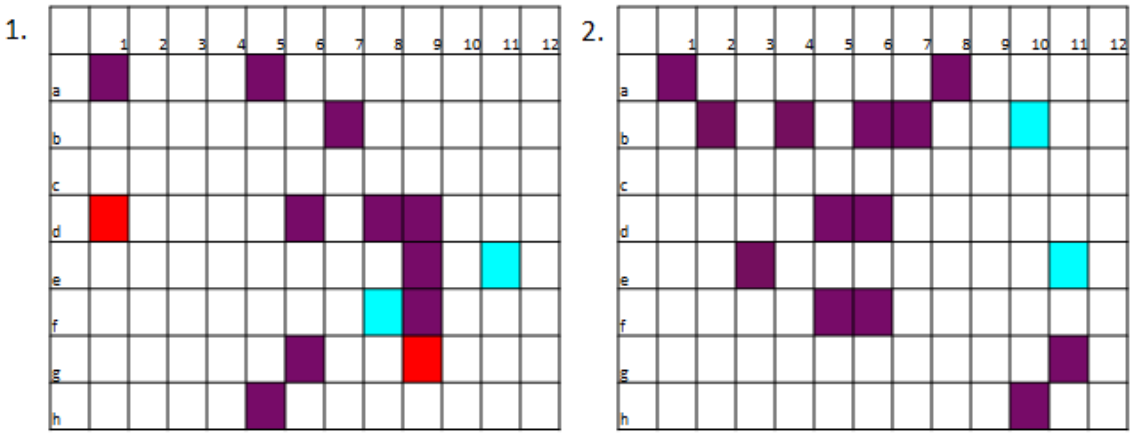


Figure 7: Diagram of two 96 well plates and indication of growth after plating. Light blue wells were identified as possible hits. White wells failed to grown on either plate. Purple wells grew on both plates. Red wells grown on only the negative.

The four potential hits are then re-plated on to the + and – caffeine plates to check for chloramphenicol resistance (Figure 8). The plates are split into quadrants for easy identification of differences in growth. The clones 1 and 2 were most likely hits because these variants only showed chloramphenicol resistance in the presence of caffeine and limited to no resistance in its absence. The clone-1 showed decent growth in the presence of caffeine while the clone-2

showed slightly less growth. The clones 3 and 4 showed a growth pattern opposite of the desired outcome.

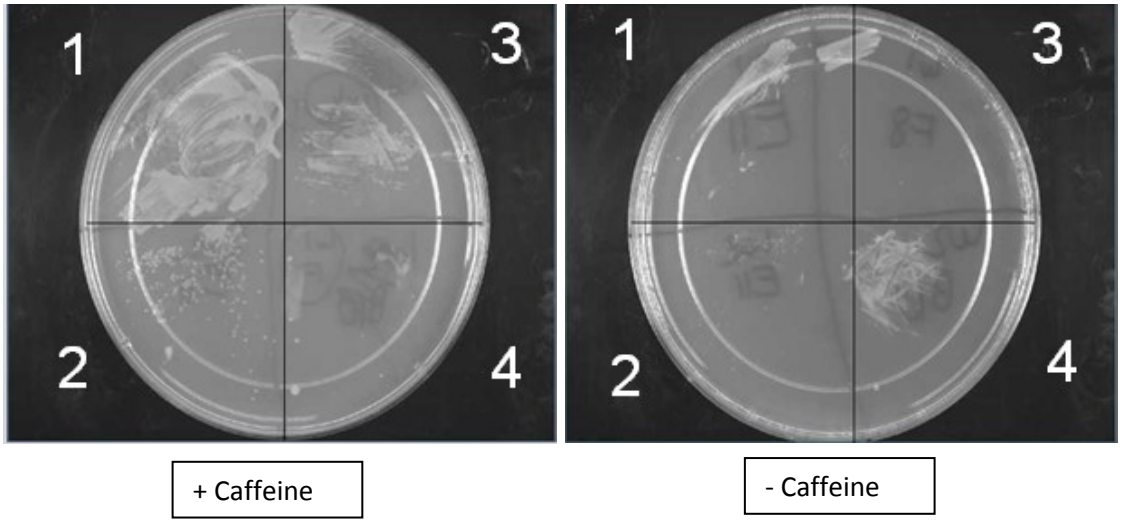
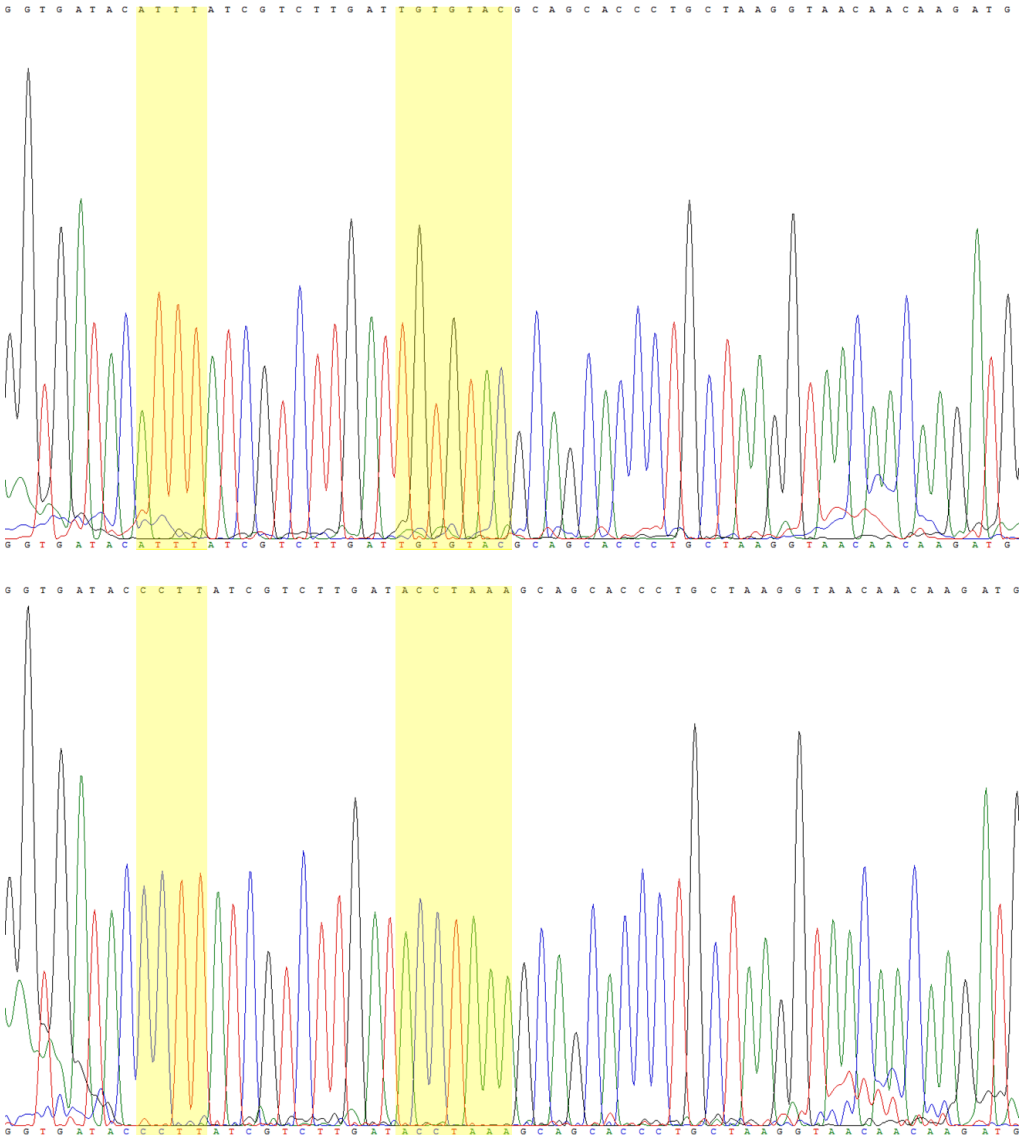


Figure 8: Re-plating of the four possible hits. The clones 1 and 2 likely contain the caffeine riboswitch.

The clones 1 and 2 were sequenced to determine the bases in the riboswitch binding pocket. The sequencing traces with the eleven bases in the aptamer domain are shown in Figure 9. Tthe sequences of the randomized region for the caffeine riboswitch hits along with the theophylline riboswitch are shown in Table 3. Both of the possible hits have a different sequence than the original theophylline riboswitch. Both sequences however can function as caffeine riboswitches.



Graph 9: The sequencing traces of the caffeine riboswitch hits.

Riboswitch	Positions 9-12	Positions 24-30
TheoRS	CAGC	GCCCTTG
Caffeine RS Hit 1	ATTT	TGTGTAC
Caffeine RS Hit 2	CCTT	ACCTAAA

Table 3: The DNA sequence of the caffeine riboswitch hits

CONCLUSION AND FUTURE WORK

Two caffeine specific synthetic riboswitches were identified by the modification of a theophylline riboswitch. Both riboswitches could be used as genetic and analytical tools. The two variants will be tested for colorimetric and fluorescence assays using β-galactosidase (Lac Z) and the green fluorescent protein as the reporter genes. The β-galactosidase reporter gene is used in the blue white screening procedure with x-gal where the proportion of blue colonies to white identifies gene expression. The same procedure is then true for the use of GFP as the reporter gene. In this case the production of fluorescent cells notes the ability of caffeine binding to change gene expression.

REFERENCES

1. Nudler, E., Mironov, A. S. (2004) The Riboswitch control of bacterial metabolism, *Trends. Biochem. Sci. 29,* 11-17
2. Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R. R. (2007) Control of alternative RNA splicing and gene expression by eukaryotic riboswitches, *nature 447,* 497-500
3. Sudarsan, N., Barrick, J.E., and Breaker, R. R. (2003) Metabolite binding RNA domains are present in the gene of eukaryotes, *RNA 9,* 644-647
4. Winkler, W.C., Cohen-Chalamis, S., and Breaker, R. R. (2002) An mRNA structure that control gene expression by binding FMN, *Proc. Natl. Acad. Sci. U.S.A 99*, 15908-15913
5. Winkler, W., Nahvi, A., and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression, *Nature 419,* 952-956
6. Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K. L. and Breaker , R.R. (2002) Genetic control by a metabolite binding mRNA, *Chem. Biol. 9,* 1043
7. Winkler, W., Nahvi, A., Sudarsan, N., Barrick, J. E., and Breaker, R. R. (2003) An mRNA structure that controls gene expression by binding S-adenosylmethionine, *Nat. Sruct. Biol. 10,* 701-707
8. Wang, J.X., Lee, E. R. Morales, D. R., Lim, J., and Breaker, R.R. (2008) Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling, *Mol. Cell 29,* 691-702

9. Sudarsan, N., Lee, E. R., Weinberg, X., Moy, R.H., Kim, J.N., Link, K. H., Breaker , R.R (2008) Riboswitches in eubacteria sense the second messenger cycling di GMP, *Science 321,* 411-413.

10. Mandal, M., and Breaker, R. R. (2004) Adenine riboswitches and gene activation by disruption of a transcription terminator, *Nat. Struct. Mol. Biol. 11,* 29-35

11. Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., and Breaker, R.R. (2003) Riboswitches control fundamental biochemical pathways in *Bacillus Subtilis* and other bacteria, *Cells 113,* 577-586

12. Mandal, M., Lee, M., Barrick, J. E., Weinberg, Z., Emilsson, G. M., Ruzzo, W.L., and Breaker, R. R. (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression, *Science 306,* 275-279

13. Sudarsan, N., Wickiser, J.K., Nakamura, S., Ebert, M.S., and Breaker, R.R. (2003) An mRNA structure in bacteria that controls gene expression by binding lysine, *Genes Dev, 17,* 2688-2697

14. Topp, S. and Gallivan, J (2010). Emerging Applications of Riboswitches in Chemical Biology, ACS Chem Biol., January 15; 5(1): 139-148.

15. Nudler, E. (2006) Flipping Riboswitches, Cell 126, 19-22

16. Barrick, J.E., and Breaker, R.R. (2007) The distribution, mechanisms and structures of metabolite-binding riboswitches, *Genome Biol. 8,* R239.

17. Schmidt, B., Roberts, R. S., Davis, P., Doyle, L. W., Barrington, K. J., Ohlsson, A., Solimano, A., Tin, W., (2006) Caffeine Therapy for Apnea of Prematurity, *N. Engl. J. Med. 354,* 2112-2121

18. Martin, R.J., et al. (2002) Pathophysiologic Mechanisms Underlying Apnea of Prematurity, *NeoReviews 3(4) e*59-e65

19. NCAA. (2013) NCAA Drug-Testing Program 2013-14, 4-7.

20. Buerge, L. L., Poiger, T., Muller, M.D., Buser, H.R., (2003) Caffeine, an anthropogenic marker for wastewater contamination of surface waters, *Environ. Sci. Technol. 37(4),* 691-700

21. Jenison, R., Gill, S., Pardi, A., and Polinksy, B., (1994) High-resolution molecular discrimination by RNA, *Science 263,*1425-1429.

22. Higuchi, R., Krummel, N., and Saiki, R. K., (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acid Res, 16(15),* 7351-7367.

23. Rackham, O.; Chin, J. W. (2005) A network of orthogonal ribosome mRNA pairs, *Nat. Chem. Biol. 1,* 159-166

24. Melancon III, C. E.; Shcultz, P. G. (2009) One Plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases, *Bioor. Med. Chem. Lett., 19* 3845-3847

ABSTRACT

The goal of this project was to identify a synthetic riboswitch for caffeine. The synthetic caffeine riboswitch will be useful to regulate gene expression in response to caffeine and detect caffeine in complex biological samples such as urine and blood. A riboswitch is a sequence of non-coding RNA that can specifically binds a ligand and thereby controls expression of genes associated with the biosynthesis, transport, or degradation of the specific ligand. In order to develop the caffeine synthetic riboswitch, we modified the theophylline-binding domain of a synthetic theophylline riboswitch. Caffeine is structurally similar to theophylline differing only by one N-methyl group. A total of 11 bases that showed significant interactions with theophylline in the theophylline riboswitch were randomized in the library design. The library was constructed by overlapping extension polymerase chain reaction and inserted into pTrp-TheoRS-CatUpp, a plasmid that encodes the upp-cat fusion gene under control of pTrp promoter. GH371 E. coli cells were transformed with the library and subject to a few rounds of the positive selection on chloramphenicol in the presence of caffeine and the negative election on 5-fluorouracil in the absence of caffeine. From the genetic screening, 192 individual single colonies were selected and tested for chloramphenicol resistance in the presence and absence of caffeine. Two clones clearly showed chloramphenicol resistance only in the presence of caffeine. These two caffeine riboswitch variants were sequence and will be further tested for calorimetric and fluorescence assays using beta-galatosidase (LacZ) and green fluorescence protein respectively as reporter genes.