# Catching Rabies by the Toe:

An Investigation into the Toehold Switch as a Sensor for Rabies virus.

by

Nadja Hedrich

# A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE (HONS.) in the Departments of Biological and Physical Sciences (Biology)

This thesis has been accepted as conforming to the required standards by:

Donald Nelson (Ph.D.), Thesis Supervisor, Dept. Biological Sciences Bruno Cinel (Ph.D.), Co-Supervisor, Dept. Physical Sciences Jonathan Van Hamme (Ph.D.), Examining Committee member, Dept. Biological Sciences



Dated this 28th day of April, 2016, in Kamloops, British Columbia, Canada

## ABSTRACT

Rabies is a disease of the central nervous system caused by the Rabies virus. Current methods to test for Rabies are costly, time consuming, and not sensitive enough to identify the disease early enough for treatment. A novel mechanism of virus identification that has shown promise for Ebola and Zika virus detection is the toehold switch. Toehold switches are RNA-based constructs that allow the visual identification of genomic targets of interest. They are composed of a switch RNA and a trigger RNA. In the presence of the trigger, a change in secondary structure causes the reporter gene following the switch to be expressed. When the trigger is not present, the gene is repressed. This project designed a toehold switch for use as a Rabies virus sensor. First, a region of the Rabies genome was identified as a promising trigger sequence. From this sequence, a complementary switch was designed with GFP as a reporter gene. The designed switch was constructed and inserted into a pBluescript plasmid backbone. The next steps of this project would be to develop an assay to test the effectiveness of the switch to identify Rabies virus.

Thesis Supervisor: Dr. Don Nelson

# ACKNOWLEDGEMENTS

Thank you to my supervisors for their guidance and honesty, and making sure this project was very much my own, and to Thompson Rivers University for the labs and facilities. Thank you to the UREAP for funding this project, and to Jason, for proofreading all of these pages many times over. And finally, thank you to my family and Jon for all of their support and understanding.

# TABLE OF CONTENTS

ABSTRACT	1
ACKNOWLEDGEMENTS	2
LIST OF FIGURES	6
INTRODUCTION	8
MATERIALS AND METHODS	14
I.Identifying Targets in the Rabies Genome for Trigger Sequence	14
II. Design of Complementary Toehold Switches	15
III. Design of Final Toehold Construct	16
IV. Plasmid Production	17
A. Transformation	17
B. Restriction Digests and Fragment Isolation	18
C. Ligations and Visualization	18
D. Determination of Plasmid Identity	19
RESULTS AND DISCUSSION	20
I. Identifying Targets in the Rabies Genome for Trigger Sequence	20
II. Design of Corresponding Toehold Switches	26
III. Design of Final Toehold Construct	29
CONCLUSIONS AND FUTURE WORK	39
REFERENCES	42
APPENDIX	44
A: Sequences, locations and penalties assigned to the top 8 trigger sequences.	44
B. Sequences of the top 8 switches corresponding to 8 top triggers.	45
C: Results of transformation of DH5- $\alpha$ E. coli with three different plasmids (pBS, j and pToe, as well as a no DNA control.	pGFP, 46
D: Agarose gel quantification of plasmids isolated using the Qiagen maxiprep kit a protocol. Plasmids were linearized with Kpn1 and 1 $\mu$ L of DNA was loaded and 50 of lambda HindIII marker. The gel was run for 1.5 hrs at 65 V.	and )0 ng 46
E. Results of transformations done to test ligation reaction efficiency. Competent E were heat transformed and grown on Lb agar plates overnight.	E. coli 47
F: Annotated partial sequence of the constructed positive control plasmid.	48
G: Annotated partial sequence of the constructed experimental toehold plasmid.	50

#### LIST OF FIGURES

Figure 1. Schematic of the general Rabies virus lifecycle. Boxes indicate genes in the Rabies genome.

Figure 2. The mechanism of action of toehold switches. The toehold without the trigger does not express the gene as translation is blocked by the hairpin loop. When the trigger is added to the switch, the hairpin is disrupted, opening up the ribosome binding site and allowing translation to occur and the gene to be expressed.

Figure 3. Secondary structure at 37oC of the eight top trigger candidate sequences visualized using NUPACK software. Thermodynamic stability of the structure is shown below each trigger as free energy of secondary structure.

Figure 4. Trigger candidate dimer formation at 37oC visualized using NUPACK software. The free energy of the secondary structure is shown below each dimer.

Figure 5. Designed switch sequences corresponding to the top eight identified triggers. Switches were visualized using NUPACK software.

Figure 6. Secondary structure of the chosen trigger sequence, corresponding switch sequence and the trigger switch complex, visualized using NUPACK software.

Figure 7. A. General schematic of the ordered toehold switch construct as optimized by Green et al. 2014. B. Actual purchased nucleotide sequence containing the switch, and GFPuv as a reporter gene. Colours represent different portions of the sequence. Beige=stuffer nucleotides, blue=restriction site, orange= GGG site, grey=switch sequence, cyan=ribosome binding site, purple=start sequence, red=linker region, green=GFPuv, magenta=terminator loop. Figure 8. Plasmid maps of the three plasmids used in subcloning experiments, pBS, pGFPuv, and pToe. Important features and their location are shown on each map.

Figure 9. Cloning experiment to create the control plasmid. The plasmid was created by combining the pBS backbone and the GFP gene from pGFPuv using a restriction digest, fragment isolation, and ligation.

Figure 10. Cloning experiment to create the experimental toehold plasmid. The plasmid was created by combining the pBS backbone and the designed pToe insert using a restriction digest, fragment isolation, and ligation.

Figure 11. Restriction digests performed to support the identity of the constructed experimental toehold plasmid.

## INTRODUCTION

Rabies is a disease of the central nervous system caused by the Rabies virus. In humans, it is almost always transmitted by the bite of an infected animal. Rabies is most commonly transmitted by dogs, and is fatal unless treated immediately after infection. There are tens of thousands of human deaths worldwide annually due to Rabies, most in developing countries (Rupprecht). Once the virus enters the body, it infects the peripheral nervous system, then progressively travels through the nervous system cells to the central nervous system. It aggregates in the central nervous system at first, and then spreads to other organs late in the disease (Galveston, 1996). Symptoms are not seen until after the incubation period, which generally lasts from 30 to 60 days, but varies depending on the location and severity of the infection site (Rupprecht). Symptoms come in two forms, consisting of either paralysis or hyperactivity, but both forms result in comas and ultimately respiratory arrest Galveston, 1996).

The virus itself is a member of the *Lyssavirus* genus and family *Rhabdoviridae*. It is rod-shaped and consists of a nucleocapsid (containing proteins and the RNA genome) and an outer envelope consisting of a lipid bilayer with glycoprotein spikes. The viral genome consists of single stranded, nonsegmented negative sense RNA; this means that it must be transcribed into positive mRNA before it can be translated into proteins. The genome is about 12kB in length, and codes for five proteins: L(RNA-polymerase), N(nucleoprotein), NS(RNA-polymerase associated) aka P (phosphase), M(matrix), and G(glycoprotein) (www.cdc.gov). Figure 1 shows the positions of these genes in the genome and the changes to the Rabies genome throughout its lifecycle. The G protein forms the glycoprotein spikes on the surface of the virus, which allow the virus to adsorb onto and penetrate other cells. After the virus has entered a host cell, it releases its nucleocapsid and the L protein transcribes the genomic RNA to make mRNAs. These mRNAs are then translated into proteins using the host cell's ribosomes. Once a certain amount of viral proteins are produced, the viral polymerase starts making full length positive copies of the RNA genome, from which more negative RNA strands are made. Then the proteins and RNA aggregate and form new viruses with the help of the N protein, which leave the cell and infect neighbouring cells. Once the virus reaches the brain, aggregation of the virus in brain tissue causes inflammation of the tissue, resulting in the psychological symptoms associated with Rabies virus infection. However, the exact physiological mechanism that causes this inflammation is not yet known (www.cdc.gov, www.who.int).



Figure 1: Schematic of the general Rabies virus lifecycle. Boxes indicate genes in the Rabies genome.

A vaccine does exist, and a dog vaccination program has been successful in drastically lowering the cases of rabies infection in developed countries; however, the virus is difficult to detect, as symptoms only appear late in the disease (Rupprecht). Traditional methods of screening for rabies include direct immunofluorescence tests on skin biopsies to look for rabies antigen. Another method of testing is culturing the virus from saliva samples in neuroblastoma cells or rats, but this method is expensive. Screening for rabies antibody in cerebrospinal fluid is also used. Finally, the most common testing method for Rabies in animals is to examine a post mortem brain tissue sample (<u>www.who.int</u>). However, all of these tests are limited, as they are not very sensitive, and may not detect the virus early on in the disease. As such, a method to screen for Rabies virus that can identify low virus concentrations, and does not require the death of the host organism is desirable. One method that seems like a promising solution to this problem due to its sensitivity and affordability is the toehold switch.

Toehold switches are a new form of translational gene regulators (Church et al., 2014). This means that they can regulate when proteins are being expressed by inhibiting or promoting translation of a certain gene, normally through changes in secondary structure. Toehold switches consist of two strands of RNA; a switch and a trigger. The switch RNA consists of a specific hairpin loop secondary structure that contains the ribosome binding site at the top of the loop and the start sequence, as well as a toehold region at the base of the hairpin loop. The trigger RNA sequence is complementary to this toehold sequence and part of the hairpin loop stem. When the switch is present in a cell without the trigger, the ribosome binding site is hidden and translation does not occur. However, when the trigger is present, it will bind to the toehold site and disrupt the hairpin, making the site accessible, thus allowing translation to occur (see Figure 2). The principle behind this mechanism is DNA strand displacement. When two strands are complementary to the same sequence, they will compete with each other to bind to that sequence. However, if one of the free strands is bound to a region adjacent to the contested portion (a toehold region), then it will have an advantage. The double stranded toehold portion

9

will increase the local concentration of the toehold mediated strand, allowing it to outcompete the other strand. Even when one strand is already bound, as in a toehold switch, the second strand will displace the first through branch migration. The opposite reaction is much slower because the lack of a toehold section does not allow for colocalization of the strand therefore the toehold strand will displace the incumbent strand (Zhang and Winfree, 2009).

These `toehold switches have a wide dynamic range and low off target effects (Green et al. 2014). Dynamic range refers to the difference between the smallest, and the largest visible signal, dependent on concentration of trigger. A wide dynamic range is favourable so that small amounts of viral genome will be identified as having turned the switch "on", and the amount of virus can be determined, giving an idea of disease progression. Off target effects refer to the ability of similar triggers to turn on switches other than the one they were designed to bind with. This is unfavourable because it can lead to the formation of false positive.



Trigger RNA

A Green et al., Cell, 2014

Figure 2. The mechanism of action of toehold switches. The toehold without the trigger does not express the gene as translation is blocked by the hairpin loop. When the trigger is added to the switch, the hairpin is disrupted, opening up the ribosome binding site and allowing translation to occur and the gene to be expressed (Green et al., 2014).

These switches have the potential to be modified to act as sensors to detect specific viral genomes (Pardee et al., 2014). To do this, the gene being repressed would be a reporter gene that codes for proteins whose presence is easily visualized, and the trigger used would be a piece of endogenous RNA, such as a part of a viral genome. When these modifications are made, this mechanism can be used to identify the presence of specific genetic sequences. Recently, toehold switches have been used to sense Ebola virus and Zika virus (Pardee et al., 2014, 2015). Virus sensing was done by using a portion of the viral genomes as the trigger and designing switches to report the presence of the viral genome using green fluorescent protein (GFP) or B-galactosidase as reporters. Toehold switches are sensitive assays that can also be applied to cell free systems which makes them quick and easy to use outside of a laboratory environment.

For my Honours project, I designed a toehold switch to detect Rabies RNA. This involved finding a portion of the Rabies genome to use as a trigger, and developing a switch that would express a reporter protein in the presence of Rabies virus genome, but would not express this reporter when Rabies virus is not present. To do this, I first identified potential targets in the Rabies viral genome for use as a unique trigger sequence, then designed a toehold switch and trigger sequence *in silico* from the chosen target. After the design was complete, I purchased the designed construct, and inserted it into a pBluescript plasmid backbone, through a series of transformations, restriction enzyme digests, and ligation reactions.

11

## **MATERIALS AND METHODS**

I. Identifying Targets in the Rabies Genome for Trigger Sequence

To identify targets for the trigger sequence in the Rabies genome, the nucleotide sequence for the SAD B19 strain of Rabies from NCBI was used. The VIPR primer search engine was then used to search for targets with the following specifications (Pardee et al., 2016):

- 1. In L gene (5414 to 11797)
- 2. Length 30 nucleotides
- 3. 40-60% GC content
- 4. Melting temperature >41C
- 5. No runs of 4 or more of the same nucleotides
- 6. Sequence of SSW (or WSS, SWS) at the 13-15th nucleotides (where S=strong=G or C, and W=weak=A or T)
- 7. Minimal secondary structure
- 8. Minimal dimer formation
- 9. Higher AT at 3' end, higher GC at 5'end (6 nucleotides)

Criteria 1, 2, and 3 were entered into the primer engine search criteria, returning only primers containing those specifications. Fifty primers were returned at a time, and manually examined for criteria 4.

The returned sequences were then visualized using the NUPACK software system and examined for criteria 6 and 7 (http://www.nupack.org). NUPACK software is a free online platform that allows for the easy visualization of RNA secondary structure; a single sequence can be inputted to examine intramolecular structure, or several sequences can be inputted to examine binding between and within sequences. Several parameters can be set, including temperature, ion concentrations, and strand concentrations. In the software, the parameters were set to a temperature of 37°C, 1.0 M Na<sup>+</sup>, and 0.0 M Mg<sup>2+</sup> (Serra and Turner, 1995), and assumed strand concentrations of 100 nM. The sequences were also visually examined and ranked using criteria 8 and 9.

II. Design of Complementary Toehold Switches

Toehold switches complementary to the resulting trigger sequences were then designed, following Green et al. 2014. The switch contained a GGG sequence, complementary sequence to the trigger, a ribosome binding site, a linker sequence, and GFPuv as a reporter gene. The GFPuv gene has a known sequence (https://www.ncbi.nlm.nih.gov/nuccore/1490531), so no extra design was required. These designed switches were then examined using NUPACK software to look for unwanted secondary structures and unwanted stop codons. Unwanted secondary structure would include any binding of the toehold region to itself, or to the linker region, thereby blocking trigger binding. Stop codons would interfere with the expression of the GFP by prematurely terminating the ribosome. The binding of the trigger and sequences was examined in NUPACK using a temperature of 37°C, 1.0 M Na<sup>+</sup>, 0.0 M Mg<sup>2+</sup>, and assumed strand concentrations of 100 nM.

The designed switch and trigger pairs were then ranked according to the selection criteria mentioned above, with an emphasis on favourable secondary structure (no trigger secondary structure, minimal dimer formation, and no switch secondary structure in the toehold region). The most promising switch/trigger combination was then examined using NCBI BLAST (https://blast.ncbi.nlm.nih.gov) to determine uniqueness. The trigger sequence was searched against all sequences in the database excluding only Rabies virus. resulting sequences were then returned along with their corresponding similarities to the trigger sequence.

## III. Design of Final Toehold Construct

Once a suitable trigger/switch combination was determined using the previously mentioned steps, the linker sequence, GFPuv sequence, and a terminator loop sequence was added, according to Green et al. 2104 (See Figure 7). Restriction enzyme sites corresponding the EcorI and KpnI were added, EcoR1 on the 3' end, and KpnI at the 5' end, as well as six nucleotides at the terminal ends of the sequence to increase the restriction enzyme activity. Synthesis of the entire construct, now consisting of 878 nucleotides was ordered from Thermo Fisher Scientific Invitrogen (https://www.thermofisher.com) (full sequence in Appendix C).

14

#### IV. Plasmid Production

Two plasmids; a positive control plasmid and an experimental plasmid, were constructed once the ordered switch sequence arrived. The plasmids were constructed from previously isolated stock solutions of pBluescript, pGFPuv, and the ordered switch plasmid, which will be referenced to as pToe. The two plasmids were constructed using a series of transformation, fermentation, restriction reactions, fragment isolations, and ligation reactions according to standard molecular biological approaches. First, a control plasmid consisting of a pBluescript backbone with the GFP gene from the GFPuv plasmid was created. The GFP gene was inserted into the multiple cloning site using EcoRI and KpnI restriction sites. The experimental plasmid was constructed in the same way as the control, but the insert consisted of the ordered switch sequence and the GFP reporter gene.

#### A. Transformation

First, previously isolated pGFP and PBS as well as the purchased toehold plasmid (pToe) were transformed into DH5-α competent *E. coli* cells using standard heat transformation protocols (Fasman 1989). The transformed bacteria were plated on Lb +Amp plates and incubated overnight at 37 °C. The next day 5 mL of LB broth containing ampicillin was inoculated with one colony. The 5 mL broth was incubated overnight at 37 °C and then used to inoculate a 100 mL broth which was incubated overnight. A maxiprep using the Qiagen plasmid maxiprep kit and protocol was performed to isolate the plasmids. Following this, samples were

linearized using EcoR1 and visualized on a 1% agarose gel to determine concentration (Qiagen Kits).

#### B. Restriction Digests and Fragment Isolation

The isolated pBS, pToe, and pGFP were then both cut using a two-step restriction digestion. The first digestion was performed using either EcoRI or KpnI (Fasman 1989). The reactions were incubated at 37 °C for 1 hour. The linearized plasmid DNA was then isolated using isopropanol, spun down at 14,000 rpm for one minute, then washed with 70% ethanol, spun twice at 14000 rpm, and dried. A second restriction enzyme digestion was then performed using the either EcoRI or KpnI (the opposite as previously used).

The cut plasmids were then visualized on a 1% agarose gel, and the pBS backbone fragment, and GFP gene fragment were isolated using the QiaEx II fragment isolation kit and protocol (QIAEX® II Handbook). The isolated fragments were then visualized on a 1% agarose gel to determine efficiency of isolation and DNA concentration.

## C. Ligations and Visualization

Next the isolated fragments were combined in a ligation reaction (Fasman 1989). The reactions were incubated overnight at room temperature. The ligated plasmids were then heat transformed into competent DH5- $\alpha$  *E. coli* cells. The plates were incubated at 37°C overnight, and then examined the following day for the presence of white colonies. Several controls were performed testing for linearization efficiency, and transformation efficiency.

#### D. Determination of Plasmid Identity

In order to determine whether the intended plasmids were successfully produced, white colonies (that do not have a working lacZ gene) indicating successful insertion into the pBS multiple cloning site from the ligated plasmid plates were restreaked, and then used to inoculate 5mL Lb + Amp broth. After being incubated at 37°C overnight, a miniprep was performed on 1 mL of the 5mL culture, and 1 mL was used to inoculate 100 mL of Lb +Amp broth which was incubated at 37°C overnight. The miniprep was performed using Qiagen Plasmid prep kit buffers and standard protocol (Fasman 1989).

A plasmid maxiprep was then performed using the 100 mL culture, following the protocol and using the contents of the Qiagen Plasmid maxiprep kit (Qiagen kit). The isolated plasmids from the maxiprep were then visualized on a 1% agarose gel to determine a rough estimate of yield.

A series of restriction digests were then performed on the isolated plasmids following the previously discussed protocol. The fragments were then visualized on a 1% agarose gel with ethidium bromide under UV light.

### **RESULTS AND DISCUSSION**

The purpose of this project was to design and create a novel toehold switch to be used as a sensor for the Rabies virus.

I. Identifying Targets in the Rabies Genome for Trigger Sequence

The first step in this project was to find targets in the Rabies genome that would be promising for use as trigger sequences. Several criteria that were known to optimize the trigger binding to and "turning on" the switch sequence were used as parameters to identify promising targets (Green et al,. 2013).

The first of these parameters was that it was found in the L gene. The L gene was chosen because it is highly conserved between Rabies species, suggesting that the trigger would we common to multiple strains of Rabies (Tordo et al. 2017). The length of 30 nucleotides was found to produce the best strand displacement, as shorter toeholds showed less successful strand displacement, and longer toeholds did not significantly increase displacement. The GC content is important because GC bonds are stronger than AT bonds. Stronger binding to the toehold increases strand displacement (Xu et al., 2013). Similarly the melting temperature optimizes strand binding and displacement (Green et al, 2014).

After inputting these initial criteria into the ViPR primer search engine, approximately 300 primers were examined. Nucleotide runs of four or more were eliminated because they are unstable and increase transcription and translation errors (Ackermann et al., 2006). Next the sequence of two strong and one weak base at the 13th-15th nucleotides were examined. This sequence has strong enough binding for the stopp loop structure to be stable on its own, but weak

enough to allow for optimal strand displacement when the trigger binds the toehold. Using those criteria to eliminate primers, the top eight primers were examined using NUPACK software. NUPACK allows for easy visualization of RNA secondary structure, so it was used to examine secondary structure of the eight trigger sequences.



Figure 3. Secondary structure at 37oC of the eight top trigger candidate sequences visualized using NUPACK software. Thermodynamic stability of the structure is shown below each trigger as free energy of secondary structure.

The results from the NUPACK software inputting only the trigger sequences are seen in Figure 3 above. Secondary structure varied from no nucleotide pairs in trigger 4, to eight nucleotide pairs in trigger 1. From this initial visualization, trigger 4 seemed most promising, as it contained no secondary structure that would interfere with the binding of the trigger to the switch. Triggers 3 and 5 also showed some promise, as they had minimal secondary structure (only two paired nucleotides), and these paired nucleotides were not close to the start of the sequence, so the trigger would still potentially bind with the switch RNA. The other switches showed increased secondary structure; and in some cases, as in trigger 7, had large stem loop structures that would potentially inhibit binding of the trigger to the toehold switch.

These eight trigger sequences were then examined again using NUPACK software but this time for dimer formation, the binding of one trigger sequence to another identical trigger molecule. This is important to consider because an ideal trigger will bind only to its corresponding switch. If the trigger binds to another copy of itself, forming a dimer, there will be less binding of the trigger to the switch, thereby making the sensor less sensitive. In the context of the entire Rabies genome, this is less of a concern, however, dimer formation could lead to problems when testing the switches with shorter sequences (oligonucleotides), giving potential false negatives.



Figure 4. Trigger candidate dimer formation at 37oC visualized using NUPACK software. The free energy of the secondary structure is shown below each dimer.

When the trigger sequences were examined for dimerization, there was a range of bonding seen, from no dimerization, to 22 nucleotide pairs (See Figure 4). As dimerization tendency increases, the availability of the trigger sequence to bind to the switch sequence decreases, as some of the trigger molecules are bonded to each other. Therefore, a trigger sequence that shows no dimer formation would be ideal.

As the trigger is a portion of the Rabies genome, this is not a large concern in practical applications, as dimerization of the entire Rabies virus genome is unlikely, however, it may be significant in a testing assay using oligonucleotides, where dimerization is more probable. Triggers 2 and 4 showed no dimer formation, suggesting that the trigger would preferentially bind to the switch sequence than another trigger. Trigger 5 and 6 both had fourteen bonds in their dimers, with several bubbles where no bonds were present, and no bonds near the ends, suggesting that binding with the switch sequence may still be preferred over dimerization. Trigger 8 had only six bonds in its dimer form, but the stem loops seen in the secondary structure of the trigger sequence were also present. The other three triggers had more bonds in the dimers, up to 22 bonds in trigger 1. These would not be as promising as trigger sequences because there may be a significant decrease in trigger/switch binding due to high levels of dimerization. From the examination of trigger sequence secondary structure and dimerization tendency, trigger 4 seemed the most promising, as it contained no secondary structure, and did not form dimers. Trigger 4 also had no dimer formation, but contained significant secondary structure, a stem-loop of seven nucleotide pairs that would potentially inhibit its binding to the switch sequence. Triggers 3 and 5 had minimal secondary structure (only two bonds), but both showed some dimer formation, with 18 and 14 bonds in the dimer form, respectively. These are still promising

23

trigger sequences, but the possibility of dimerization may decrease switch interactions in a testing assay.

Next, the switch sequences corresponding to these triggers were designed according to layouts determined by previous research (Green et al., 2014), and the switch sequences were examined NUPACK software.

# II. Design of Corresponding Toehold Switches

Toehold switches corresponding to the top eight trigger sequences were designed according to previous specifications (Green et al. 2014). These switches were then examined using NUPACK software for ideal stem-loop structure, and any in frame stop sequences.



4

7



Free energy of secondary structure: -23.80 kcal/mol



2

5

8



Free energy of secondary structure: -23.60 kcal/mol

Free energy of secondary structure: -21.40 kcal/mol

6

3



Free energy of secondary structure: -24.60 kcal/mol



Free energy of secondary structure: -20.70 kcal/mol



Free energy of secondary structure: -25.30 kcal/mol

A C G U



Free energy of secondary structure: -26.70 kcal/mol

Free energy of secondary structure: -24.20 kcal/mol



When the switch sequences corresponding to the top eight trigger sequences were visually examined, several different RNA secondary structures were seen (see Figure 5). Switches 4 and 7 showed a secondary structure consisting of the ideal stem-loop structure with 15 nucleotides bonded, a three nucleotide bubble at the start site, and no secondary structure in the toehold section before the loop (see Figure 7). The small stem-loop in the linker sequence would not inhibit trigger binding, or subsequent ribosome action. Switches 1 and 8 had small stem loops in the toehold regions, but the small amount of bonding would likely still allow for successful trigger binding. Some of the other switch sequences had less favourable secondary structures however, with switch 2 showing the toehold region binding to the linker region, which would likely decrease the ability of the trigger to bind. Switch 6 has a large stem-loop structure in the toehold section, which again would likely decrease trigger bonding, making it an unfavourable candidate. Finally, switches 3 and 5 had shifted stem loops and some unwanted secondary structure in their toehold regions, also making them less than ideal candidates for a successful trigger/switch combination.

Using this information, along with trigger secondary structure and dimer formation, trigger 4 and its corresponding switch was chosen as the best candidate for use as a Rabies virus sensor. The free energies of the secondary structure, indicating the stability of secondary enthalpies of the dimers and trigger sequences were also compared (see Figures 3, 4, and 5).

The chosen trigger sequence was then examined in NCBI BLAST, and no other sequences outside of the Rabies genome were found to match the sequence 100% when queried against the entire NCBI database excluding other Rabies viruses. The closest matches had <50% similarity, and none had the toehold region the same as the trigger sequence. Within different Rabies virus

26

strains, the sequence was conserved, indicating that many different strains of Rabies virus could be detected with this trigger.

III. Design of Final Toehold Construct

Using the sequence of trigger 4 and its corresponding switch, which were found to be unique using NCBI blast, contained no unwanted secondary structure, and formed no dimers, a final toehold construct was designed following a previously optimized design (Green et al., 2014).



Free energy of secondary structure: -24.60 kcal/mol

Figure 6. Secondary structure of the chosen trigger sequence, corresponding switch sequence and the trigger switch complex, visualized using NUPACK software.

Figure 6 (above) shows the secondary structure of trigger 4, its corresponding switch, and the trigger/switch complex more closely. As was mentioned before, the trigger shows no secondary structure and no dimer formation, making it an ideal candidate because there was nothing present

that would inhibit the trigger bonding to the switch. The switch sequence also does not contain any unwanted secondary structure in the toehold region, again making it ideal for binding to the trigger; additionally, the stem loop structure follows the optimized design discussed by Green et al. 2014. Finally, the trigger switch combination shows the trigger fully binding to the switch, and linearizing the stem loop structure. There is some secondary structure seen in the area behind the trigger binding, but it does not seem to be likely to inhibit ribosome binding as it consists of very few nucleotides, and will not be very stable. There are also no stop sequences present in the switch sequence, making this a very promising candidate for continuing the project. Using this switch sequence as a starting point, a construct was designed to sense Rabies virus. Behind the linker sequence, the sequence for GFP was added, and a terminator loop was included behind the GFP. GFP is an easily visualized reporter gene. It does not require anything more than a UV light to be visualized, and is a protein, so no other substrates are necessary, unlike some other reporters such as LacZ. Also, because GFP is a protein the amount of green fluorescence is proportional to the amount of GFP produced, and to the amount of trigger present (Feilmeier). The terminator loop will end translation by the RNA polymerase after the GFP is expressed. At the 3' end of the sequence, an EcoRI restriction site was added, and on the 5' end a KpnI restriction site was added, along with 6 more stuffer nucleotides at each end (see Figure 7). The restriction sites allow for the sequence to be easily moved between plasmids using standard molecular cloning techniques. Two different restriction sites were used so that the sequence can only be inserted in one orientation, eliminating the possibility of a reversed insert. The stuffer nucleotides are necessary for the restriction enzyme to effectively cut the DNA at the restriction site.



GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTG GTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACATTGAAGATGGATCC CCAGAACCTGGCGGCAGCGCAAAAGATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAAT GATTACACATGGCATGGATGAGCTCTACAAATAATAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTT CGACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTG CGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAAACTCGAGTACAACTATAACTCACACAA CTATATCTTTCAAAGATGACGGGAACTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTA AGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCC TIGGAATTCATCCAG TTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGC4 AAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTCTCTTATGGTGTTCAATGC

of the sequence. Beige=stuffer nucleotides, blue=restriction site, orange= GGG site, grey=switch sequence purchased nucleotide sequence containing the switch, and GFPuv as a reporter gene. Colours represent different portions cyan=ribosome binding site, purple=start sequence, red=linker region, green=GFPuv, magenta=terminator loop Figure 7. A. General schematic of the ordered toehold switch construct as optimized by Green et al. 2014. B. Actual The switch construct, as seen in Figure 7 was designed and then ordered from Thermo Fisher Scientific. There, they combined pre-prepared synthetic oligonucleotides into the desired nucleotide sequence and inserted it into a plasmid containing ampicillin resistance and a ColE1 promoter. The sequence was examined to ensure that it was the same as the ordered sequence, and the construct was delivered as an insert in a plasmid termed pToe (see Figure 8).

# **IV.** Plasmid Production



Figure 8. Plasmid maps of the three plasmids used in subcloning experiments, pBS, pGFPuv, and pToe. Important features and their location are shown on each map.

# Figure 8 above shows the main elements of each of the three plasmids used in the subcloning

experiment. Each plasmid has a CoIE1 origin of replication, and ampicillin resistance. The pBS plasmid also contains a LacZ gene spanning a multiple cloning site (MCs), which contains EcorI and KpnI restriction enzyme sites. Normally, the LacZ gene will cause the bacterial colonies to appear blue when plated with IPTG and X-gal, however, when a fragment is inserted in the MCS, the LacZ gene will not be functional, and any colonies plates on IPTG and X-gal will not be blue. The pBS also contains a T7 RNA polymerase promoter region, allowing for selected promotion of the region following. This is why pBS was chosen as a backbone for the constructed plasmids.

pGFP has a ColE1 site, Amp resistance, and the GFP gene flanked by KpnI and EcoRI restriction sites, and pToe is very similar, but it contains the designed toehold construct between the restriction sites.





Figure 9. Cloning experiment to create the control plasmid. The plasmid was created by combining the pBS backbone and the GFP gene from pGFPuv using a restriction digest, fragment isolation, and ligation.

# Figure 9 shows the cloning experiment to create the positive control plasmid. Both pBS and

pGFPuv were cut with EcoRI and KpnI. Following this, the backbone of pBS, and the GFp insert of pGFPuv were isolated and ligated. The positive control contains all of the elements of pBS such as Amp resistance, ColE1 origin, and the T7 promoter. However, the LacZ gene has been disrupted, and GFP has been inserted. This plasmid works as a positive control because when induced by the T7 promoter, GFP should be expressed, just as the activated trigger and switch complex should express GFP.





Figure 10. Cloning experiment to create the experimental toehold plasmid. The plasmid was created by combining the pBS backbone and the designed pToe insert using a restriction digest, fragment isolation, and ligation.

Figure 10 shows the cloning experiment to create the experimental plasmid. This time, pBS and

pToe were cut with EcoRI and KpnI. Following this, the backbone of pBS, and the toehold construct insert of pToe were isolated and ligated. Again. the experimental plasmid contains all of the elements of pBS such as Amp resistance, ColE1 origin, and the T7 promoter and a disrupted LacZ gene. However, this time the toehold construct has been inserted into the MCS.

Lambda EcoRI HindIII BgIII Kpnl Kpnl Clal Clal Kpnl Clal EcoRI EcoRI



Figure 11. Restriction digests performed to support the identity of the constructed experimental toehold plasmid.

To determine whether the constructed experimental toehold plasmid was successfully created, a series of restriction digests was performed: KpnI, EcoRI, ClaI, HindIII, EcoRI and KpnI, EcoRI and ClaI, KpnI and ClaI, and BglII. The results of these restriction enzyme reactions are seen in Figure 11 above. The HindIII and BglII lanes show only closed circular plasmid, as expected because there should be no BglII or HindIII sites present in the plasmid. The EcoRI, KpnI, and ClaI all linearized the plasmid, as expected for enzymes with only a single restriction site present in the plasmid. The single band seen in the EcoRI/ClaI digest is also as we would expected, with a single large fragment seen, very slightly smaller than the linearized plasmid. The second fragment would only consist of approximately 10 nucleotides, so we do not expect to see that

fragment. The KpnI/ClaI and KpnI/EcoRI lanes look very similar, with two bands, one corresponding to the toehold and GFP insert, and the other, larger band corresponding to the pBS backbone. However, we expect the KpnI/ClaI was expected to have a slightly larger GFP insert, as was seen. Therefore, these restriction enzyme reaction provide support that the toehold plasmid was successfully created.

## **CONCLUSIONS AND FUTURE WORK**

Due to the shortcomings of present forms of Rabies virus identification, the development of a novel detection method is of considerable interest. Toehold switches are a promising novel mechanism for virus identification that have recently been used in the Ebola and Zika virus epidemics with success. Not only are they sensitive and affordable, they can be portable, which makes them especially convenient for third world countries. This is important for Rabies virus, because it is third world countries, lacking in proper vaccination programs, that would most benefit from this sensor. This project made strides to developing a new way of sensing Rabies virus using toehold switch technology.

Following the guidelines set out by Green et al. 2014, a sequence was selected from the Rabies genome for use as a trigger. The chosen sequence was theoretically ideal, lacking in secondary structure, and unable to dimerize, ensuring that the trigger and switch binding would not be hindered by unwanted secondary structure. The sequence was also examined using NCBI BLAST, and found to be unique to Rabies virus, though conserved between different Rabies virus strains. Due to these qualities the sequence was chosen as a trigger. From this, a corresponding switch was designed *in silico*, and it was found to lack unwanted secondary structure in its toehold region and any in-frame stop codons, making it a promising candidate for a switch sequence. GFPuv was chosen to be a reporter gene because it is easily visualized, and has previously been used as a reporter for toehold switches with success. Restriction enzyme sites were added to the terminal ends of the trigger sequence in order to make subsequent ligation into a pBS backbone easier, and the whole construct was purchased from Invitrogen.

A series of transformation, restriction, and ligation experiments successfully inserted the purchased toehold construct into a pBS backbone. The pBS was chosen as a backbone because of the presence of a T7 RNA polymerase promoter region, allowing for selective expression of the toehold sequence.

There were some issues with DNA amount due to low yield returned from the plasmid preps, making some visualization of the restriction and ligation reactions unsuccessful. However, there was sufficient DNA for the reactions to occur. Redoing the plasmid prep early in the experiment, or doing two preps and concentrating the DNA may have avoided these problems.

In order to test whether the toehold switch successfully senses the Rabies sequence, a series of testing assays will then be performed to examine the ability of the trigger to induce expression of GFP by binding with the switch. Initially, an assay using oligonucleotide sequences of the 30 nucleotide trigger sequence will be performed. In order to test out toehold function with an oligonucleotide, DNA-RNA bonding in a cell free system will be used. Cell free systems will be created through lysis of E. coli cells through sonication, and then centrifugation to remove cell membranes and organelles, leaving only cytoplasm, and cell machinery related to

39

transcription and translation. In this way, the trigger oligonucleotide can be introduced directly to the cell free extract, also containing the switch RNA and ribosomes, and the extract can then be tested for the presence of GFP with UV light. This test would be performed using the experimental toehold insert, and the control insert, once it is successfully produced. Following this test, another test will be performed examining RNA-RNA interactions and switch function.

## REFERENCES

- Ackermann, Martin, and Lin Chao. "DNA Sequences Shaped By Selection For Stability". *PLoS Genetics* 2.2 (2006): e22. Web. 27 Apr. 2017.
- "BLAST: Basic Local Alignment Search Tool". *Blast.ncbi.nlm.nih.gov.* N.p., 2017. Web. 27 Apr. 2017.
- Center for Disease Control. "The Rabies Virus." Accessed at <u>http://www.cdc.gov/rabies/transmission/virus.html</u> Sept. 22, 2016
- Church G. M., Elowitz M. B., Smolke, C. D., Voigt C. A., and R. Weiss. VIEWPOINT "Realizing the potential of synthetic biology." *Nature Reviews. Molecular Cell Biology*. 15 (2014): 289-294. Web.
- "Cloning Vector Pbad-Gfpuv, Complete Sequence Nucleotide NCBI". *Ncbi.nlm.nih.gov*. N.p., 2017. Web. 27 Apr. 2017.
- Fasman, Gerald D. *Practical Handbook Of Biochemistry And Molecular Biology*. 1st ed. Boca Raton, Fla.: CRC Press, 1989. Print.
- Feilmeier, Bradley J. et al. "Green Fluorescent Protein Functions as a Reporter for Protein Localization in *Escherichia Coli*." *Journal of Bacteriology* 182.14 (2000): 4068–4076. Print.
- "Geneart Gene Synthesis And Services | Thermo Fisher Scientific". *Thermofisher.com*. N.p., 2017. Web. 27 Apr. 2017.
- Green A.G. et al. "Toehold Switches: De-Novo-Designed Regulators of Gene Expression." *Cell* 159 (2014): 1–15. Web
- "QIAGEN Plasmid Mini, Midi, And Maxi Kits (EN)". *Qiagen.com*. N.p., 2017. Web. 25 Apr. 2017.

Pardee K. et al. "Paper-Based Synthetic Gene Networks." Cell 159.4 (2014): 940-54. Web.

- Pardee, K. et al. "Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components."*Cell* 165.5 (2016): 1255-266. Web.
- Rupprecht C. E. "Rhabdoviruses: Rabies Virus." In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 61. Available from: <u>https://www.ncbi.nlm.nih.gov/books/NBK8618/</u>
- Tordo, Noël et al. "Completion Of The Rabies Virus Genome Sequence Determination: Highly Conserved Domains Among The L (Polymerase) Proteins Of Unsegmented Negative-Strand RNA Viruses". *Virology* 165.2 (1988): 565-576. Web. 27 Apr. 2017.
- Xu, Pingping, Fujian Huang, and Haojun Liang. "Real-Time Study Of A DNA Strand Displacement Reaction Using Dual Polarization Interferometry". *Biosensors and Bioelectronics* 41 (2013): 505-510. Web.
- World Health Organization. "Rabies". Accessed at <u>http://www.who.int/mediacentre/factsheets/fs099/en/</u> Sept 22, 2016
- Zhang, David Yu, and Erik Winfree. "Control Of DNA Strand Displacement Kinetics Using Toehold Exchange". *Journal of the American Chemical Society* 131.47 (2009): 17303-17314. Web. 27 Apr. 2017.

# APPENDIX

Appendix A: Sequences, locations and penalties assigned to the top 8 trigger sequences.

#	Trigger sequence	Start in the genome	Concentrations of strands for dimerization
1	CTGGAGTCTAGTCAGCTTATTGATGATAG A	7519	0.081mM single 0.0096 mM double
2	GGATCTATTATTCAGACAGATCAGACCTC A	7389	n/a
3	GAATAGACCTTATCGGATGACTCTAACAG A	5584	0.075mM single 0.013mM double
4	CTCAATACTTACCAGTCTCATCTTCTACTC	10037	n/a
5	ATATATTCTCCTGAGGCTAGATAACCATCC	9778	0.075mM single 0.012mM double
6	TCTCATATATTCTCCTGAGGCTAGATAACC	9774	0.097mM single 0.0015mM double
7	CTCACTGGATCAGGTTGATTTACAAGATA G	11632	0.096mM single 0.0016mM double
8	GGTTGAGAGAAACCTATCTAAGAGTATGA G	10072	0.093mM single 0.0035mM double

Appendix B. Sequences of the top 8 switches corresponding to 8 top triggers.

Number	Switch Sequence
1	GGGTCTATCATCAATAAGCTGACTAGACTCCAGATACAGAAACAGAGG AGATATCTGAUGTCTAGTCAGAACCUGGCGGCAGCGCAAAG
2	GGGTGAGGTCTGATCTGTCTGAATAATAGATCCATACAGAAACAGAGG AGATATGGAAUGATTATTCAGAACCUGGCGGCAGCGCAAAG
3	GGGTCTGTTAGAGTCATCCGATAAGGTCTATTCATACAGAAACAGAGG AGATATGAAAUGACCTTATCGAACCTGGCGGCAGCGCAAAG
4	GGGGAGTAGAAGATGAGACTGGTAAGTAATGAGATACAGAAACAGAG GAGATATCTCAUGACTTACCAGAACCUGGCGGCAGCGCAAAG
5	GGGGGATGGTTATCTAGCCTCAGGAGAATATATATACAGAAACAGAG GAGATATATAAUGTCTCCTGAGAACCUGGCGGCAGCGCAAAG
6	GGGGGTTATCTAGCCTCAGGAGAATATATGAGAATACAGAAACAGAG GAGATATTCTAUGATATTCTCCAACCUGGCGGCAGCGCAAAG
7	GGGCTATCTTGTAAATCAACCTGATCCAGTGAGATACAGAAACAGAGG AGATATCTCAUGGGATCAGGTAACCUGGCGGCAGCGCAAAG
8	GGGCTCATACTCTTAGATAGGTTTCTCTCAACCATACAGAAACAGAGG AGATATGGTAUGGAGAAACCTAACCUGGCGGCAGCGCAAAG

Sample	Volume plated	Number of Colonies	Ampicillin
Control	100uL	carpet	Ν
Control	100uL	none	Y
pBS	100uL	carpet	Ν
pBS	100uL	carpet	Y
pBS	10uL	>400	Y
pGFPuv	100uL	~100	Y
pGFPuv	10uL	9	Y
рТое	100uL	~100	Y
рТое	10uL	~30	Y

Appendix C: Results of transformation of DH5- $\alpha$  E. coli with three different plasmids (pBS, pGFP, and pToe, as well as a no DNA control.

Appendix D: Agarose gel quantification of plasmids isolated using the Qiagen maxiprep kit and protocol. Plasmids were linearized with Kpn1 and 1  $\mu$ L of DNA was loaded and 500 ng of lambda HindIII marker. The gel was run for 1.5 hrs at 65 V.

Lambda pToe pBS pGFP

Sample	Amount of DNA (ng)	# Colonies	Ampicillin
CCC pBS +lig	2.8	Carpet	Ν
CCC pBS +lig	2.8	~300	Y
CCC pBS -lig	2.8	0	Y
CCC pGFP +lig	2.8	Carpet	Ν
CCC pGFP +lig	2.8	>400	Y
CCC pGFP - lig	2.8	0	Y
pBS e-k +lig	4.7	Carpet	Ν
pBS e-k +lig	4.7	20	Y
pBS e-k -lig	4.7	0	Y
pBS k-e +lig	4.7	50	Y
pBS k-e -lig	4.7	0	Y
pGFP e-k +lig	4.7	9	Y
pGFP e-k -lig	4.7	0	Y
pGFP k-e +lig	4.7	~100	Y
pGFP k-e -lig	4.7	0	Y
Construct	28.0	Carpet	Ν
Construct	2.8	40	Y
Construct	28.0	~300 ~50 white	Y
Construct	28.0	~300	Y
Construct	28.0	~300	Y

Appendix E. Results of transformations done to test ligation reaction efficiency. Competent E. coli were heat transformed and grown on Lb agar plates overnight.

Appendix F: Annotated partial sequence of the constructed positive control plasmid.

AGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATA 630 TCCCCCTACACGACGTTCCGCTAATTCAACCCATTGCGGTCCCAAAAGGGTCAGTGCTGCAACATTTTGCTGCCGGTCACTTAACATTAT Pro Pro His Ala Ala Leu Arg Asn Leu Gin Thr Val Giv Pro Asn Giu Trp Asp Arg Arg Gin Leu Val Val Ala Leu Ser Asn Tyr Tyr M13 fed T7 promoter Acc651 Agel KpnI CGACTCACTATAGGGCGAATTGGGTACCGGTAGAAAAAATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAAT \*\*\*\*\* 720 **GCTGAGTGATATCCCGCTTAACC**CATGGCCATCTTTTTTACTCATTTCCTCTTGAAAAGTGACCTCAACAGGGTTAAGAACAACTTA Ser Glu Ser Tyr Pro Ser Asn Pro Val lacZo. T7 promoter > MCS \*\*\*\*\* 810 StyI NcoI NdeI MscI BspE1\* GCACTACT00AAAACTACCT0TTCCAT00CCAACACTT0TCACTACTTCTCTTAT00T0TTCAAT0CTTTTCCC0TTATCC00ATCATA \*\*\*\*\* 900 COTGATGACCTTTTTGATGGACAAGGTACCGGTTGTGAACAGTGATGATGAAGGAATACCACAAGTTACGAAAAGGGCAATAGGCCTAGTAT So 50 55 60 65 70 75 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His BarGI TGAAACGGCATGACTTTTTCAAGAGTGCCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGAACTACAAGA 990 ACTTTGCCGTACTGAAAAAGTTCTCACGGTACGGGCTTCCAATACATGTCCTTGCGTGATATAGAAAGTTTCTACTGCCCTTGATGTTCT 80 85 90 95 100 105 105 Met Lys Arg His Asp Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Gly Asn Tyr Lys MiuI COCOTOCT0AAGTCAAGTTT0AAG0T0ATACCCTT0TTAATCGTATC0AGTTAAAAG0TATT0ATTTTAAAGAA0AT00AAAACATTCTC0 \*\*\*\* 1080 0C0CACGACTTCAGTTCAAACTTCCACTATGGGAACAATTAGCATAGCTCAATTTTCCATAACTAAAATTTCCTTCTACCTTTGTAAGAGC Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu





Appendix F: Annotated partial sequence of the constructed positive control plasmid.

	EcoO1091	EcoRI Psti TspHI Smal Spel
ATTACACATGGCATGGCATGAGCTCTACAAATAATA	CATAACCCCTT0000CCTCTAAAC000TCTT0A0000TTTTT	IGGAATTCCTOCAGCCCGGGGGATCCACTA
TAATGTGTACCGTACCTACTCGAGATGTTTATTAT	COTATTOGGGAACCCCGGAGATTTGCCCAGAACTCCCCCAAAAAA	ACCTTAAGGACGTCGGGGCCCCCTAGGTGAT
230		
GEPUV	17 terminator	MCS
• V H C P H A R C I I	ку G R P A E L R T K L P K K n frame with lacZo)	s N R C G P P D V L
Not		1997 No. 1997 No. 1997 No. 1
XbaI Eagl Alel SacII BstXI		
CTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCT	ITTOTTCCCTTTA0T0A0G0TTAATTCC0A0CTT00C0TAATCA	ATGGTCATAGCTGTTTCCTGTGTGAAATTG
AGATCTCGCCGGCGGTGGCGCCACCTCGAGGTCGA	AAACAAGGGAAATCACTCCCAATTAAGGCTCGAACCGCATTAGT	TACCAGTATCGACAAAGGACACACTTTAAC
25	1 1 15 1 1 10 1 1 1 5 1 1	
E LAAAVATSSWS	KNGKTLTLEBSPTIP Ngữa	Interaction of the second seco
MCS	< T3 promoter	M13 rev