Andrew Martens
04/07/08

The Use of Terminators and Anti-Terminators to Expand the Two-Spatula Burnt-Pancake Problem in *Escherichia coli*

Honors Thesis

Davidson College

First Reader: Dr. Campbell

Second Reader: Dr. Wessner

Abstract

The emerging field of synthetic biology provides a new experimental approach to biological research by allowing for the design and implementation of novel genetic systems. In 2006, a joint effort by researchers from Davidson College and Missouri Western State University developed a bacterial system that uses the Salmonella Hin/*hix* DNA recombination mechanism to solve *in vivo* the two-spatula, burnt-pancake mathematical problem. The problem asks to calculate the least number of flips required to sort a given stack of pancakes with two spatulas, from largest to smallest and with all golden sides facing up, assuming that each pancake is burnt on one side and golden on the other. The system worked as a proof-of-concept for using bacteria to solve mathematical problems *in vivo*, but was limited to pancake stacks with 3 pancakes. A new system that uses transcription anti-terminators could potentially extend the size of the problem, allowing for many more pancakes per stack. I developed a way to use the naturally occurring *E. coli trp* operon anti-terminator, downstream of the pLac promoter, to control GFP gene expression in tryptophan-rich and tryptophan-poor conditions. RNA tertiary structure computer simulations were used to predict if the anti-terminator sequence could be modified to accommodate Hin-mediated flipping. The success of using an anti-terminator could enable synthetic biologists to design hundreds of anti-terminators that would allow bacteria to solve burnt-pancake problems that were exponentially more difficult.

Introduction

Traditional biological research has focused on expanding our understanding of natural biological processes. With the advent of synthetic biology, it is becoming increasingly common to take advantage of our knowledge of molecular biology and biochemistry to engineer living systems to perform new functions. Evaluating an engineered system's functionality allows us to think of genetic and molecular events in terms of parameters that we can control. Comparing the behavior of the components in an engineered system with their behavior in natural biological processes can shed new insight into how these systems work in their natural environment (Keasling, 2008). Over the last two years, researchers at Davidson College and Missouri Western State University have focused on developing bacterial systems that can compute solutions to mathematical problems. In 2006, a DNA-based system was developed to solve the burnt-pancake problem in *E. coli* (Haynes *et al.*, 2007). The problem asks what is the minimum

required number of flips to convert an unsorted stack of pancakes into a sorted stack of pancakes, using two spatulas. Each pancake has a golden side and a burnt side, and the sorted pancake stack is arranged from smallest on top to largest on the bottom with all golden sides facing upwards (Figure 1).

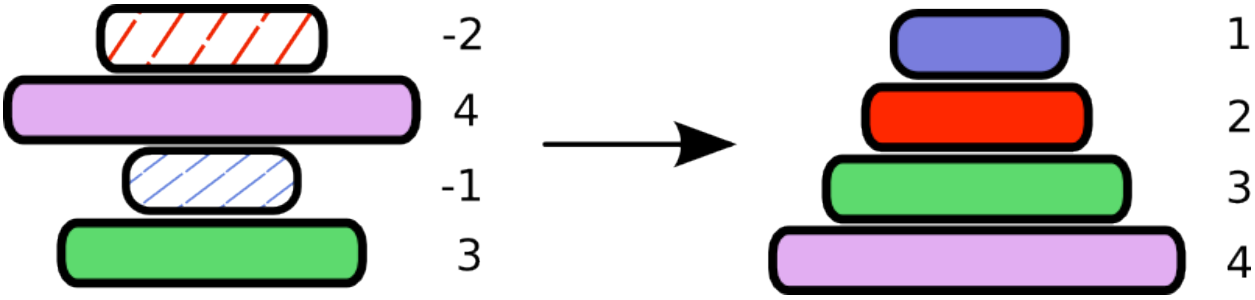


Figure 1: Visualization of the two-spatula burnt-pancake problem. Two stacks of pancakes, where each pancake is a different size (numbered 1 – 4) and burnt on one side. Negative signs and diagonal segments denote burnt-side-up pancakes. Using two spatulas, what is the fewest number of flips required to transform the stack on the left into the stack on the right?

Several properties of the pancake problem make it challenging to solve using traditional means. First, the problem grows exponentially in complexity with only linear increases in the size of the pancake stack. For example, adding 1 pancake to a stack of 4 pancakes increases the number of possible stacks by 3,456, compared to 336 new stacks when adding 1 pancake to a stack of 3 pancakes. Second, computing a solution benefits from the use of parallel computation, which is difficult for regular computers. Although this particular version of the pancake problem can be solved by computers efficiently in polynomial time, other versions, such as the one with only a single spatula, cannot be solved efficiently and are known as NP, for non-polynomial. These types of problems require computers to analyze every possible stack one at a time, and quickly become too difficult to solve. The development of a better bacterial computer for the two-spatula burnt-pancake problem may lead towards solving the more complex versions of the problem.

Bacterial computers have the potential to compute faster than the silicon computers we are used to, provided that they divide at exponential rates and double the total parallel processing ability with every cell division. Each bacterium would contain several copies of the mathematical problem in the form of an engineered plasmid. If each plasmid functions as a biological processor, then a single flask of bacteria would contain trillions of processors working independently, in parallel. A single plasmid may eventually develop a solution, which could be detected by screening for a specific phenotype. Such computation could take place during a single overnight incubation without any human intervention (Haynes *et al.*, 2007).

Our bacterial computation relies on the *Salmonella* Hin/*hix* system, which catalyzes the inversion and recombination of DNA sequences (Nanassy *et al.*, 1998; Haynes *et al*., 2007). DNA flanked by a pair of 26 bp *hixC* sites can be inverted by two dimers of Hin protein. If sequential regions are separated by *hixC* sites, Hin can catalyze the inversion of individual regions of DNA and invert more than one flanked segment as a single unit (Haynes *et al*., 2007). Thus it is possible to obtain every permutation of these regions using the Hin/*hix* system (Figure 2).

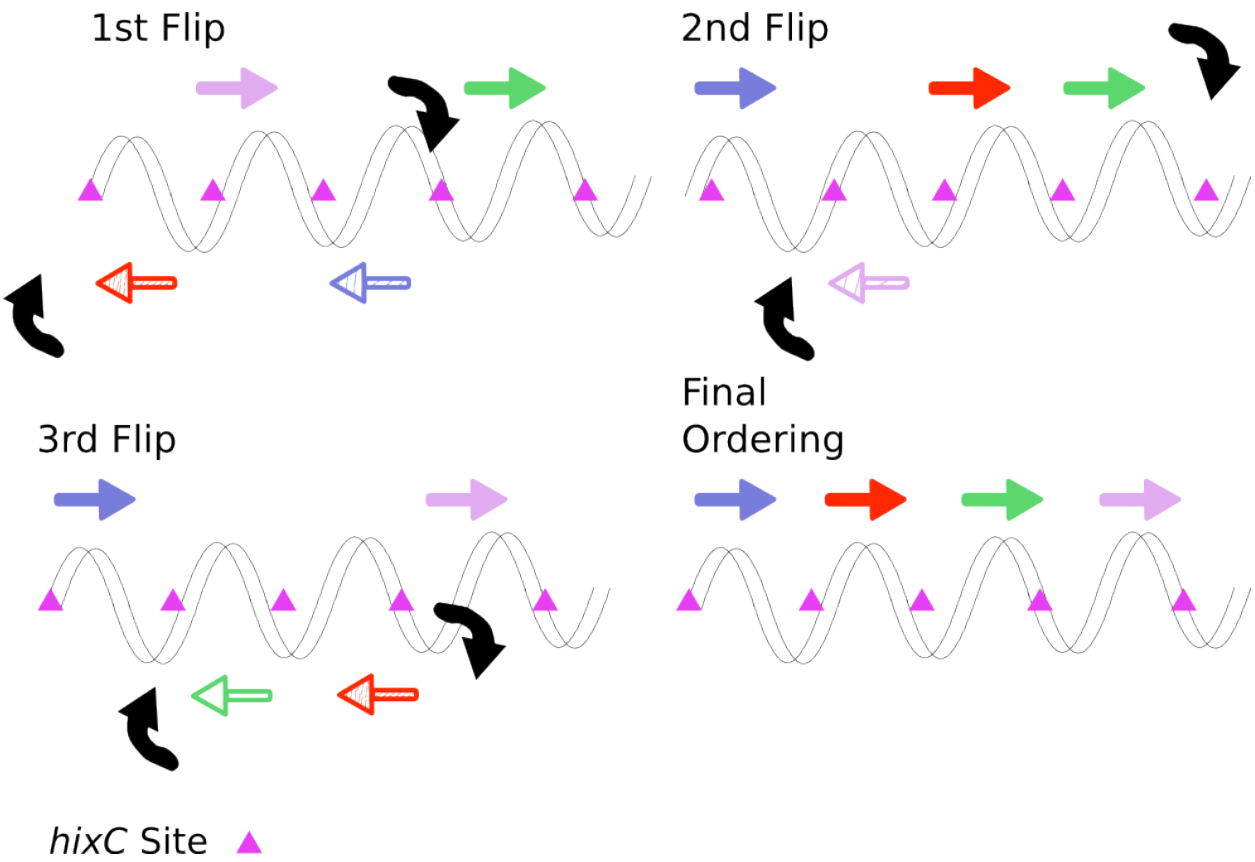


Figure 2: Visualization of the two-spatula burnt-pancake problem using DNA segments flanked by *hixC* segments to represent pancakes. The Hin/*hix* system can rearrange DNA segments into all possible permutations; each *hixC* site is represented by a purple triangle. The gene ordering on the top left, represented by arrows, may be rearranged into the ordering on the bottom right in a minimum of 3 flips.

The original system was designed such that each genetic pancake was represented by a DNA sequence with a specific function, and a solution was detected phenotypically. For example, one pancake could be the sequence for a promoter, another pancake could be a reporter gene, and a third pancake could be a transcriptional terminator. Only when the promoter comes first, followed by the reporter and transcriptional terminator, will the reporter gene phenotype be visible (Figure 3). DNA sequences not in this order would have to be rearranged using Hin-mediated recombination in order to achieve a sorted stack orientation. Although this idea was implemented successfully as a proof-of-concept for bacterial computation, the implementation

was limited in its ability to model the pancake problem. For example, it is not possible to have more than 3 DNA-based pancakes. Promoters typically will lead to the transcription of only a few genes by RNA polymerase. If each additional pancake is represented by an additional reporter gene, then the addition of extra pancakes to the problem will eventually extend beyond the reach of RNA polymerase. In addition, having two or more coding genes in a row permits more than one sorted stack to produce equivalent phenotypes and thus not a unique final outcome. If different pancake stacks give the same phenotype, then it is not possible to distinguish between them. This is problematic because it would be hard to screen for bacteria that have sorted their pancake stacks correctly (Figure 3).

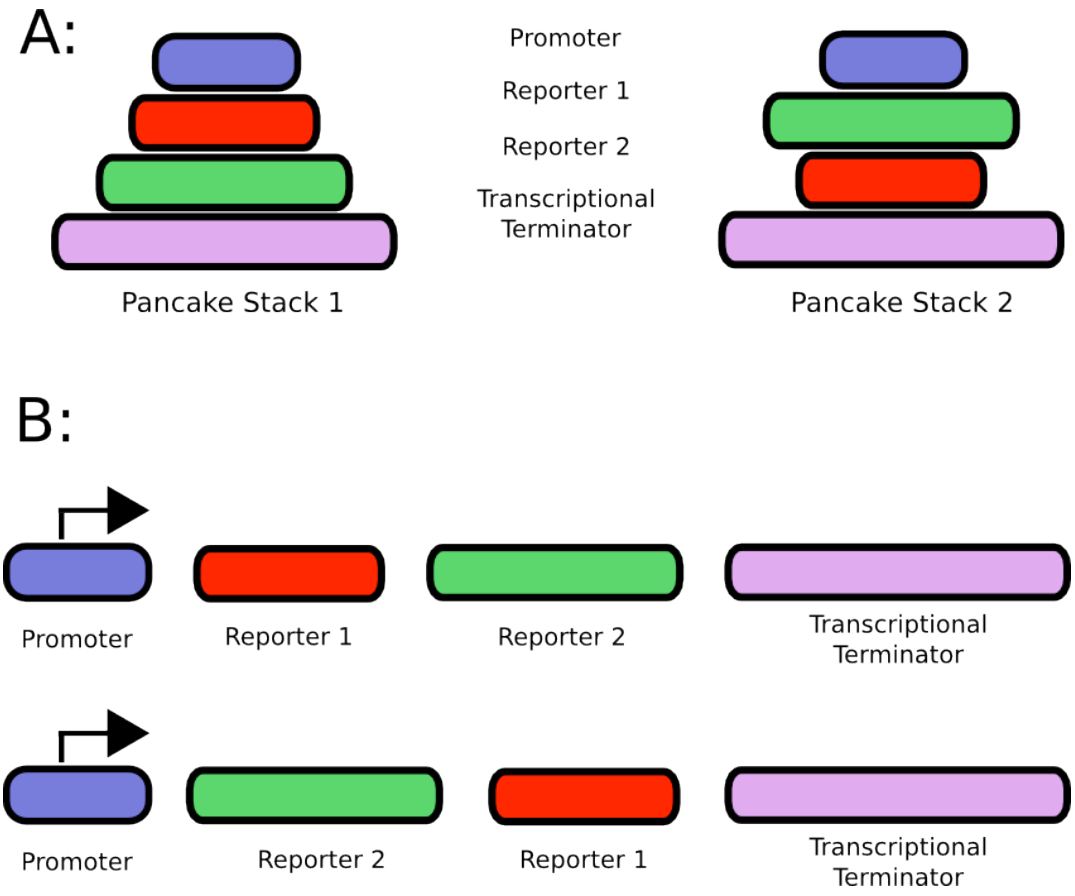


Figure 3: Design of the first genetic two-spatula burnt-pancake problem computer. Using the original pancake computer system, these two pancake stacks with different orderings produce the

same phenotype. **A:** Two stacks of pancakes. Stack 1 is sorted correctly, stack 2 is not. **B:** The two stacks as genetic information on a plasmid. Either DNA sequence will exhibit the phenotypes from reporters 1 and 2, yet the two stacks are different.

The Hin/*hix* system, used in the original *E. coli* burnt-pancake-problem computer, is relatively well understood. The primary goal of my research was to improve the computational system by using anti-terminators. Anti-terminators occur in nature to regulate transcription of a variety of genes, such as bacterial operons and phage genes (Nudler *et al*., 2002). Transcriptional termination is caused by the formation of an RNA stem-loop structure, the terminator, that forces RNA polymerase to detach from the mRNA (Figure 4 B). An anti-terminator is another RNA stem-loop structure that precedes and inhibits the formation of the terminator and its stem-loop structure, thereby preventing termination (Figure 4 A) (Nudler *et al.*, 2002). Although the anti-terminator is also a stem-loop structure, it does not terminate transcription. If an anti-terminator forms, a terminator cannot form, and transcription continues. This transcriptional attenuation mechanism allows for regulation of transcription by either selectively terminating transcription prematurely or by allowing it to proceed. The choice between transcriptional termination or anti-termination will change what is transcribed and modify gene regulation in response to environmental conditions.

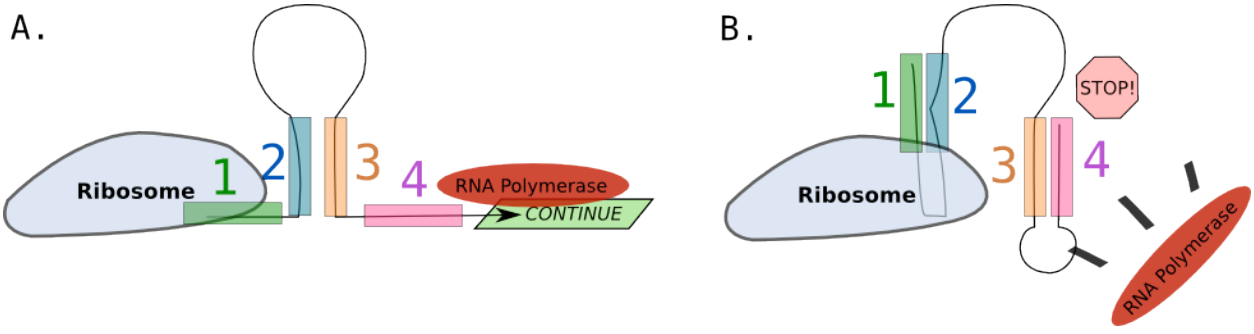


Figure 4: Mode of action of the *E. coli trp* operon anti-terminator. The position of the ribosome on the mRNA molecule, while both transcription and translation occur in tandem, affects the mRNA tertiary structure and decides between the formation of the anti-terminator or the

terminator. **A:** If the anti-terminator forms, shown by the pairing of regions 2 and 3, then the terminator cannot form, and transcription continues. **B:** If the anti-terminator does not form, then the terminator will form by pairing regions 3 and 4, and transcription will stop.

Let's look at a specific example to illustrate more clearly how anti-terminators work (Figure 5). The tryptophan operon in *E. coli*, *trp*, is regulated using an anti-terminator and encodes for genes that synthesize the amino acid tryptophan (Figure 6). The first important step in the transcription of the operon takes place at the leader peptide sequence, trpL. In bacteria, translation can begin while transcription is still under way. In this particular case, it is critical that transcription and translation are synchronized, such that a specific distance between the RNA polymerase and ribosome is set. RNA polymerase begins transcribing the mRNA until it produces a stem-loop structure which causes the polymerase to pause, denoted by the yellow region. The pause by RNA polymerase gives the ribosome time to start translating the trpL mRNA sequence into a short, non-functional leader peptide and to catch up with RNA polymerase. Once the ribosome has caught up to the RNA polymerase, the secondary structure that caused the RNA polymerase to pause is altered, and RNA polymerase continues transcription. This pausing of RNA polymerase causes transcription and translation to become coupled and to happen in synchrony (Landick *et al.*, 1987).

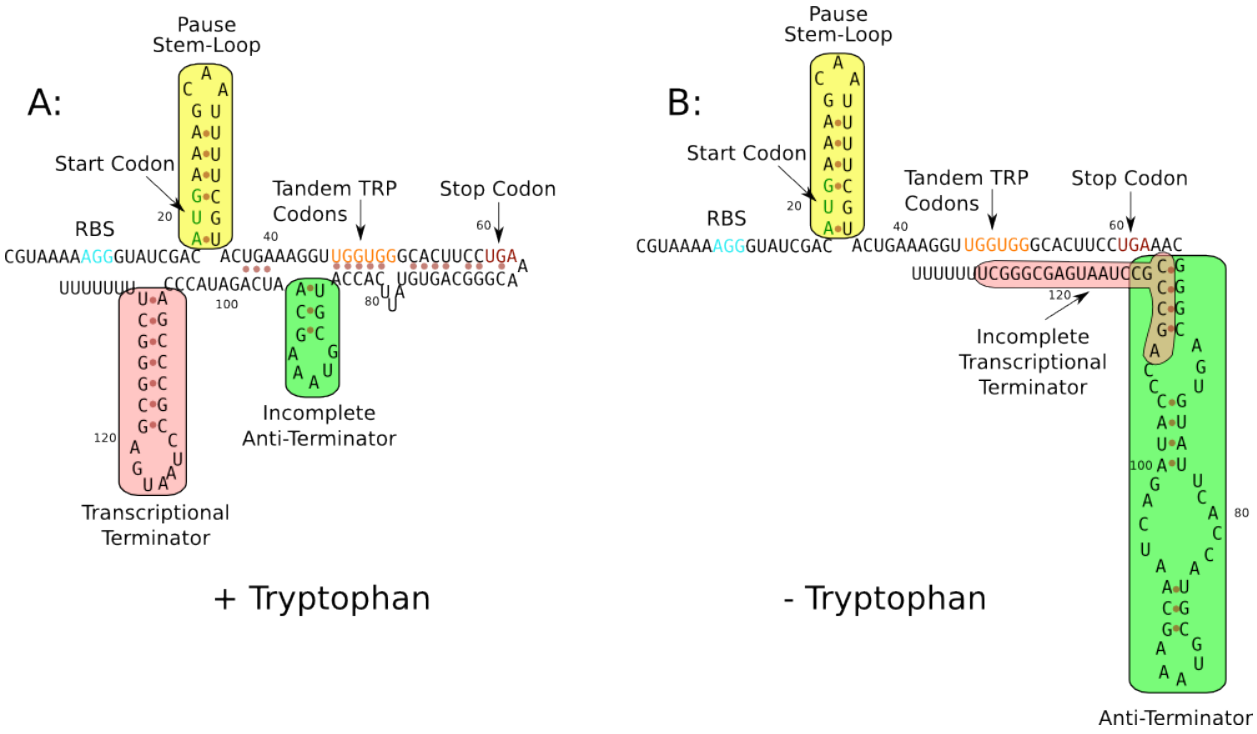


Figure 5: Two possible mRNA tertiary conformations, in either the presence (A) or absence (B) of tryptophan. Although transcription of the *trp* operon has not finished, translation of the nascent mRNA strand begins. Depending on the levels of tryptophan, the elongating mRNA can take on two different conformations. **A:** The conformation of the trp operon mRNA under normal conditions, when tryptophan is abundant. The transcription-pause loop, shown in yellow, pauses the RNA polymerase and allows the ribosome to catch up, synchronizing transcription and translation. The ribosome does not stall at the tandem tryptophan codons, and the anti-terminator cannot form. The transcriptional terminator does form, and transcription ends. **B:** Under tryptophan-starved conditions, the conformation of the mRNA is noticeably different. Again, transcription and translation are synchronized due to the presence of the pause stem-loop. However, the ribosome stalls at the tandem tryptophan codons because it is waiting for $tRNA^{Trp}$, which are now rare. Meanwhile, RNA polymerase moves further ahead. The anti-terminator then forms and prevents the downstream terminator from forming. Transcription does not end, and the entire *trp* operon is transcribed.

Before the leader peptide's stop codon, and near the beginning of the anti-terminator sequence, are two sequential codons that code for tryptophan, shown in orange. If tryptophan concentrations are low, the ribosome will pause at these two tryptophan codons until tRNAs carrying tryptophan arrive. Meanwhile, RNA polymerase moves further downstream, elongating the mRNA molecule. The elongated mRNA forms an anti-terminator stem-loop structure which

prevents the formation of a terminator stem-loop (Figure 5 B). The RNA polymerase continues mRNA elongation, the ribosome resumes once two $tRNA^{Trp}$ arrive and the entire *trp* operon is transcribed and translated (Figure 6) (Landick *et al.*, 1987).

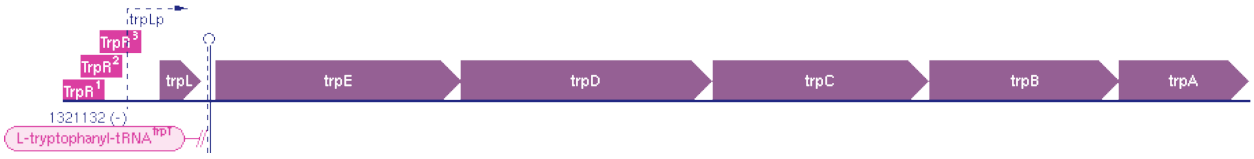


Figure 6: The *trp* operon of *E. coli*. trpL is the leader peptide, a nonfunctional 14 amino acid polypeptide whose translation regulates the expression of the rest of the operon. If tryptophan levels are low, anti-termination occurs and the remaining genes are transcribed so that *E. coli* can make its own tryptophan (BioCyc, 2008).

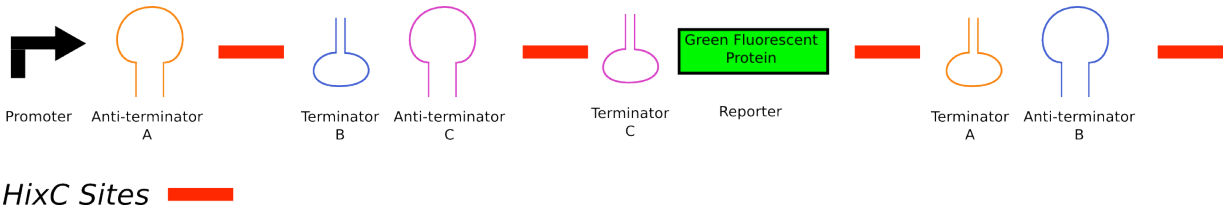
Conversely, when tryptophan is abundant, the ribosome does not pause at the pair of tryptophan codons because $tRNA^{Trp}$ are plentiful; translation continues unabated, and RNA polymerase does not distance itself from the ribosome (Figure 5 A). As the ribosome advances, it prevents the formation of the anti-terminator stem-loop structure through steric interactions. The lack of an anti-terminator structure leads to the formation of a new stem-loop structure, the transcriptional terminator, and causes the attenuation of the transcription. In this way, *E. coli* will not transcribe the *trp* operon unless it needs more tryptophan. The ribosome pauses temporarily only when there is a shortage of tryptophan, triggering the eventual synthesis of enzymes that will produce more tryptophan (Landick *et al.*, 1987).

Solving the Burnt-Pancake Problem

My goal was to design, build, and test a DNA and RNA-based system that uses the Hin/*hix* system, anti-terminators, and a single reporter gene to solve increasingly complex burnt pancake problems. My system would avoid the previous limitations of the first burnt-pancake problem computer and has the potential to be a more effective computer by providing the ability to solve problems involving larger pancake stacks and by simplifying solution detectability.

Terminators and anti-terminators are obligate pairs; a stem-loop structure from one anti-terminator only anti-terminates its specific terminator counterpart. Given this understanding, it might be possible to design hundreds of terminator and anti-terminator pairs. If I can successfully construct a proof of concept version, then each half could be separated from its partner using *hixC* sites (Figure 7). Recombination using Hin would be required to unite a disjoined anti-terminator/terminator pair. By using several different anti-terminator/terminator pairs, it might be possible to solve much more difficult burnt-pancake problems with more pancakes in a stack.

Since the goal of the pancake flipping is to achieve a sorted pancake stack, we know our desired outcome from the beginning. In a solution, each upstream anti-terminator must be paired with the appropriate downstream terminator, such that no terminator sequence will be able to form a secondary structure. There would be no termination, and the reporter gene would be transcribed and translated (Figure 8). Constructing a solved version that does not need any flipping will test whether all the anti-terminators are working properly, as it should already have



the correct ordering of anti-terminators and terminators.

Figure 7: An unsolved plasmid, with each anti-terminator/terminator pair designated by a particular color. In this example, the mRNA transcript will be attenuated, and there will be no GFP expression. To permit transcription, each anti-terminator must be immediately followed by its terminator partner.

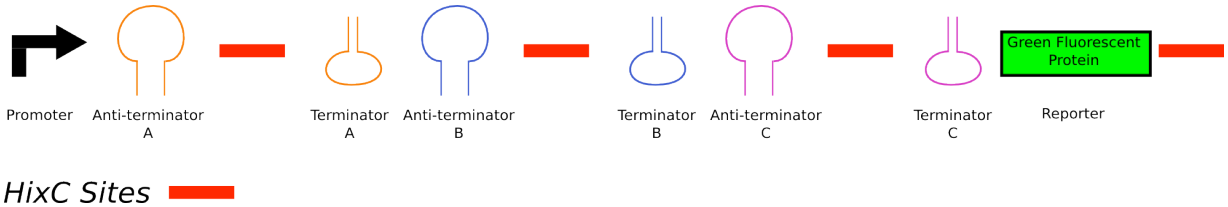


Figure 8: A solved plasmid. Each anti-terminator is immediately followed by its terminator partner. Transcription of this plasmid would result in the production of a full mRNA transcript and lead to GFP expression.

The burnt pancake problems are encoded on plasmids and transformed into *E. coli* along with the Hin expression system. If cells containing a solved plasmid exhibit the correct phenotype, then a plasmid containing an unsolved pancake problem can be constructed and tested. In my version of the unsolved plasmid, anti-terminator A is adjacent to terminator B. A and B are not of the same anti-terminator/terminator pair, and so anti-terminator A will not anti-terminate transcriptional terminator B (Figure 7). The terminator will successfully form a stem-loop structure, which will cause RNA polymerase to stop elongating the mRNA. The reporter gene will not be transcribed, and its phenotype will not be visible. Bacteria that possess the scrambled plasmid must therefore "flip" the DNA segments, using Hin, in order to rearrange these segments into the correct order. Cells with properly flipped plasmids will express the reporter's phenotype.

The first step to use the computer is to transform bacteria with two plasmids, an unsolved plasmid and a plasmid with the *Hin* gene, and to grow them in liquid culture. The culture can either be monitored continuously until the desired phenotype is visible, or fractions of the culture can be analyzed at set time intervals. In the case of the *trp* operon anti-terminator, tryptophan levels must be limited by growing cells in minimal media called M9. A flow cytometry machine can be used to measure individual cell fluorescence, and thus can distinguish a single fluorescent

cell from a large number of non-fluorescent cells, providing an accurate quantification of the proportion of fluorescing cells. If the amount of time required per flip can be estimated, then we should be able to calculate the number of flips that were required to reach the sorted pancake stack.

Materials and Methods

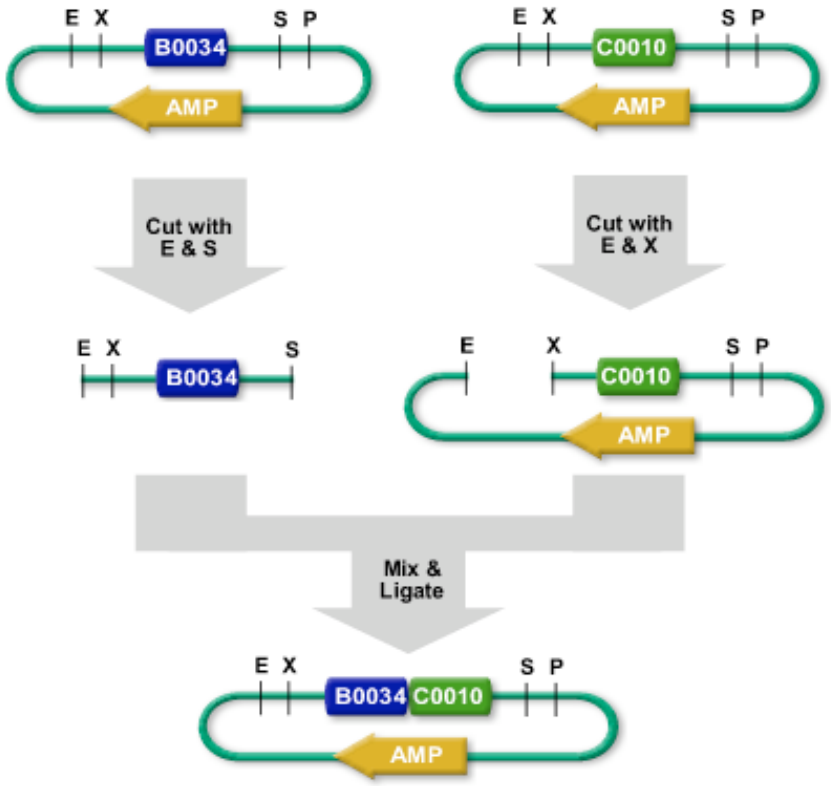
Bacterial cultures: JM109 and MC4100 *E. coli* cell cultures were grown on LB ampicillin agar plates or in liquid cultures at an antibiotic concentration of 100 µg/mL at 37 °C. To test growth in tryptophan-starvation conditions, cells were grown on M9G ampicillin agar plates.

Preparation of M9G media: Following the procedures outlined in Sambrook & Russell (2001), M9 media was prepared by adding 200 mL 5X M9 salts, 2 mL 1 M $MgSO_4$, 20 mL of 20% glucose, 0.1 mL 1 M $CaCl_2$, and sterilized, deionized water to a final volume of 980 mL. Filter-sterilized ampicillin antibiotic was added for a final concentration of 100 µg/mL. Filter-sterilized thiamine was added for a final concentration of 0.01%.

Preparation of M9 salts: M9 salts were prepared by dissolving 64 g Na_2HPO_4-$7H_2O$, 15 g KH_2PO_4, 2.5 g NaCl, and 5.0 g NH_4Cl to a final volume of 1000 mL with deionized water. 200-mL aliquots were sterilized by autoclaving.

Preparation of plasmid DNA: DNA minipreps were performed using the Promega Wizard Plus SV miniprep DNA purification system (catalog # A1460). (http://www.bio.davidson.edu/courses/Molbio/Protocols/miniprepPrmega.html)

DNA restriction-enzyme digestions: In preparation for the ligation of DNA segments, DNA was digested using two restriction enzymes simultaneously (Promega). Depending on the type of ligation, various combinations of EcoRI, XbaI, SpeI, and PstI were used, along with the proper buffer (http://parts.mit.edu/igem07/index.php/Davidson_Missouri_W/Double_Digest_Guide). The standard BioBrick assembly processed was used (http://parts.mit.edu/registry/index.php/Assembly:Standard_assembly).



The digestion of two separate plasmids with different pairs of restriction enzymes allows for the irreversible ligation of DNA. In the example shown, the front insert (blue) is digested with EcoRI and SpeI, and the front vector (green) is digested with EcoRI and XbaI. The two digested EcoRI restriction enzyme sites can ligate to each other, restoring a full EcoRI restriction site. The SpeI and XbaI sites can also ligate to each other, but they produce a 6 bp scar that cannot be re-digested. The resulting vector contains the blue part upstream of the green part. The process of flanking DNA sequences with these restriction enzyme sites is known as the construction of BioBricks, and the ligation of multiple BioBricks in sequence is known as BioBrick assembly.

Gel electrophoresis: The length of DNA segments was verified using gel electrophoresis. Determination of DNA strand length was done by comparing the distance traveled by the band with bands on the 1Kb ladder (catalog # 15615-016, Invitrogen).

DNA gel-purification: DNA was run on an agarose gel of an appropriate percentage. After excision of a gel slice, it was purified using the Qiagen QiAquick gel extraction kit (catalog # 28704) (http://www.bio.davidson.edu/courses/Molbio/Protocols/Qiagen_gelpure.html).

DNA ligation: T4 DNA ligase was used to ligate DNA segments previously digested with restriction enzymes (http://www.bio.davidson.edu/courses/Molbio/Protocols/ligation.html).

Transformation of competent cells: Zippy Z-competent JM109 (part #T3003, Zymo Research) cells were transformed with either miniprepped DNA or with DNA ligation mixtures. S.O.C.

medium (catalog # 15544-034, Invitrogen) was added to help spread on LB agar medium (http://www.bio.davidson.edu/courses/Molbio/Protocols/Zippy_Transformation.html).

Preparation and Transformation of Competent MC4100 *E. Coli* Cells using TSS Solution: Competent MC4100 *E. coli* cells were prepared according to the Biomedal guidelines (cat nos. RS-3215 and RS-3216). These were transformed with miniprepped DNA.

Colony PCR screening for ligations: colonies previously transformed with DNA ligations were screened for the presence of successful ligations. Multiple colonies were picked and added to individual solutions of 20 pmol forward primer VF2, 20 pmol reverse primer VR, 10 µl water, and 12 µl Promega GoTaq Green Master Mix (part # MZ12C). The elongation time was set in relation to the expected total length, such that one minute was allowed for every thousand base pairs. When estimating the size of an insert, 238 bp must be added to the length of the insert when the vector is pSB1A2 (http://parts.mit.edu/igem07/index.php/Davidson_Missouri_W/colony_PCR).

Sequence of primer VF2: `5' TGCCACCTGACGTCTAAGAA 3'`

Sequence of primer VR: `5' ATTACCGCCTTTGAGTGAGC 3'`

Oligonucleotide synthesis of double-stranded DNA: Double-stranded DNA constructs were assembled by annealing multiple, overlapping DNA strands. Lance Harden's computer software was used to generate the complementary oligonucleotide sequences, optimized for most similar melting points (Figure 15) (http://gcat.davidson.edu/IGEM06/oligo.html).

PCR amplification of chromosomal DNA: PCR primers were designed to amplify a desired segment of *E. coli* chromosomal DNA. The use of custom-written "gene-splitter" software (http://gcat.davidson.edu/iGEM07/genesplitter.html) facilitated the design such that desired restriction enzyme sites were incorporated into the primers. The PCR reaction was performed as in the colony PCR procedure on JM109 strain *E. coli* cells.

DNA sequence of forward primer:

`5' GCATGAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAATGAAA 3'`

DNA sequence or reverse primer:

`5' GCATCTGCAGCGGCCGCAACTAGTAAAAAAAAGCCCGCTCATTAGG 3'`

PCR amplification of *trp* chromosomal DNA with a *hixC* insertion: PCR primers were designed to amplify the desired *trp* operon regulatory sequence of *E. coli* chromosomal DNA and to introduce within the final product the *hixC* sequence. The process requires three separate PCR reactions. The first two amplify the left-hand side and right-hand sides of the sequence and incorporate the *hixC* sequence. The third PCR reaction merges the PCR products of the first two reactions, amplifying the final DNA construct (Figure 28). BioBrick ends were incorporated into the final construct design. The following parameters were used in the PCR machine:

Step 1: 94.0 º C, 10.0 minutes

Step 2: 94.0 º C, 15 seconds

Step 3: 46.0 º C, 15 seconds

Step 4: 66.0 º C, 15 seconds (elongation time)

Step 5: 30 times to step 2

Step 6: 21.0 º C indefinitely

Step 7: End

DNA sequence of forward primer for left-hand side reaction:

5' GCATGAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAATGAAA 3'

DNA sequence or reverse primer for left-hand side reaction:

5' CGGGCTGGGTATCTGATTGCTTTATCAAAAACCATGGTTTTTGATAATACGCATGGTGAATA
CACTGCC 3'

DNA sequence of forward primer for right-hand side reaction:

5' GGCAGTGTATTCACCATGCGTATTATCAAAAACCATGGTTTTTGATAAAGCAATCAGATACC
CAGCCCG 3'

DNA sequence or reverse primer for right-hand side reaction:

5' GCATCTGCAGCGGCCGCAACTAGTAAAAAAAAGCCCGCTCATTAGG 3'

UNAFold to predict RNA secondary structures: Prior to wet-lab experimentation, UNAFold version 3.4 software, with all default parameters, was used to predict RNA secondary structures (http://www.bioinfo.rpi.edu/applications/hybrid/download.php). This tool, along with information gathered from the literature, aided with the design of wild-type and experimental *trp* anti-terminator DNA sequences. For each sequence, two-dimensional structures, ranked by entropy, were generated. The following sequences were simulated:

The full wild-type trp sequence:
5′CGUAAAAAGGGUAUCGACAAUGAAAGCAAUUUUCGUACUGAAAGGUUGGUGGCGCACUUCCU
GAAACGGGCAGUGUAUUCACCAUGCGUAAAGCAAUCAGAUACCCAGCCCGCCUAAUGAGCGGGC
UUUUUUUU 3′

The full sequence with a *hixC* insertion:
5′CGUAAAAAGGGUAUCGACAAUGAAAGCAAUUUUCGUACUGAAAGGUUGGUGGCGCACUUCCU
GAAACGGGCAGUGUAUUCACCAUGCGUAUUAUCAAAAACCAUGGUUUUUGAUAAAGCAAUCAGA

UACCCAGCCCGCCUAAUGAGCGGGCUUUUUUUU 3′

The partial wild-type sequence, intended to show the full anti-terminator:
5′AACGGGCAGUGUAUUCACCAUGCGUAAAGCAAUCAGAUACCCAGCCCGCCUAA 3′

The partial sequence with a *hixC* insertion:
5′AACGGGCAGUGUAUUCACCAUGCGUAUUAUCAAAAACCAUGGUUUUUGAUAAAGCAAUCAGA
UACCCAGCCCGCCUAA 3′

Design of *trp* BioBrick sequences: Two sequences were designed to include the *trp* regulatory sequence, containing the anti-terminator and terminator sequences, as well as the appropriate BioBrick ends. The second sequence has the *hixC* insertion. Other important features are noted below.

5' GAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAATGAAAGCAATTTTCGTACTG
AAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTCACCATGCGTAAAGCAATCAGATAC
CCAGCCCGCCTAATGAGCGGGCTTTTTTTTTACTAGTTGCGGCCGCTGCAG 3'

5' GAATTCGCGGCCGCTTCTAGATTCTAGACGTAAAAAGGGTATCGACAATGAAAGCAATTTT
CGTACTGAAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTCACCATGCGTAAATAGTT
TTTGGTACCAAAAACTATTAGCAATCAGATACCCAGCCCGCCTAATGAGCGGGCTTTTTTTTTAC
TAGTTGCGGCCGCTGCAG 3'

The wild-type (top) and experimental (bottom) DNA sequences. Coloring denotes BioBrick ends (brown), predicted RBS (purple), start codon (green), tandem tryptophan codons (cyan), stop codon (red), insertion (blue), tandem Ts (orange).

Sequencing of DNA: 80 ng of miniprepped DNA samples were dried using a Savant DNA 110 Speed Vac at a high drying rate. VF2 was used as a sequencing primer at a concentration of 5 pmol. The sequencing reaction was performed by the CUGI sequencing center at Clemson University (https://www.genome.clemson.edu/cgi-bin/orders?page=serviceHome&service=sequencing).

Alignment of sequences for *trp* operon DNA: Sequences were aligned using the ClustalW multiple-alignment algorithm, using default parameters (http://www.ebi.ac.uk/Tools/clustalw2/index.html). The comparisons were made against the wild-type DNA sequence, spanning the entire BioBrick ends.

Measurement of DNA concentrations: DNA concentrations were measured using a Nanodrop ND-1000 machine. After blanking with the appropriate elution buffer or nuclease-free water, the absorbances at 260 nm for 2.0 µl of miniprepped or gel-purified DNA were measured, and a DNA concentration in ng/µl was derived.

Quantification of GFP Expression Using Flow Cytometry: Cultures were grown overnight at room temperature while shaking. These were first diluted such that the OD_{600} on the BIO-TEK ELx808 plate reader was approximately 0.2. They were then further diluted with PBS such that

the cellular flow rate was below 1000 cells per second on the Guava EasyCyte Plus flow cytometer. GFP expression was calibrated against MC4100 cells containing no plasmid. A total of 10,000 cells were counted for each sample.

Results

The RNA structures involved in the anti-termination mechanism of the *E. coli trp* operon were analyzed extensively. Ultimately, the goal of these analyses was to find a way to modify the *trp* operon regulatory sequence by inserting a *hixC* site, thereby providing the means for flipping, without impeding tryptophan-dependent transcriptional regulation. There are four regions responsible for forming the two conformations, the anti-termination conformation and the termination conformation. Three of these regions, outlined in Figure 9 in yellow, red, and green, span bases 66-77, 100-113, and 119-126. It was suggested that the yellow and green regions can each bind to the red region, but that these associations are mutually exclusive (Lee *et al.*, 1977). Further work showed that the region from base 44 to base 60, shown in blue, could base-pair with the yellow region, thus establishing the existence of four critical segments of RNA which cannot be modified (Figure 9) (Winkler *et al.*, 1982).

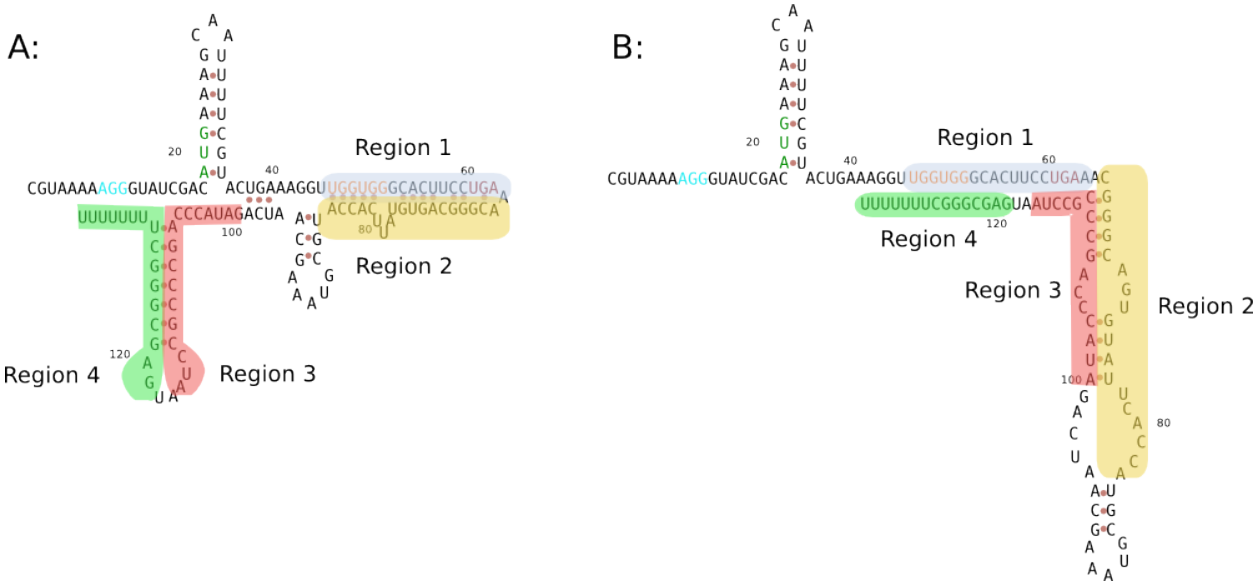


Figure 9: Diagram of the *trp* operon regulatory sequence's mRNA structure. The two possible complementary base-pairing scenarios, resulting in the termination of transcription (panel A), and the read-through due to anti-termination (panel B). Transcriptional termination occurs when region 1 binds region 2 and region 3 binds region 4. Anti-termination occurs when region 2 binds region 3.

The four regions shown in Figure 9 are necessary for initiating either termination or anti-termination. Another RNA base-pairing interaction in this RNA strand is also present, however. Das *et al.* (1983) postulated one such interaction in *S. typhimurium*, with a suggested regulatory role that inhibits the synthesis of leader peptide. By base-pairing to the beginning of the leader peptide sequence, the terminal end of the mRNA prevents excess translation of the leader peptide, itself a nonfunctional protein. The *E. coli* sequence may also have the same regulatory mechanism, as evidenced by a similar analysis of potential base-pairing between the two strands (Figure 10). Since specific base-pairing interactions are required for the proper termination, anti-termination and control of leader peptide synthesis, alterations in these regions were avoided. The *hixC* insertion, shown in Figure 14 and Figure 15, takes all these experimental results into consideration and does not obstruct any of these regions.

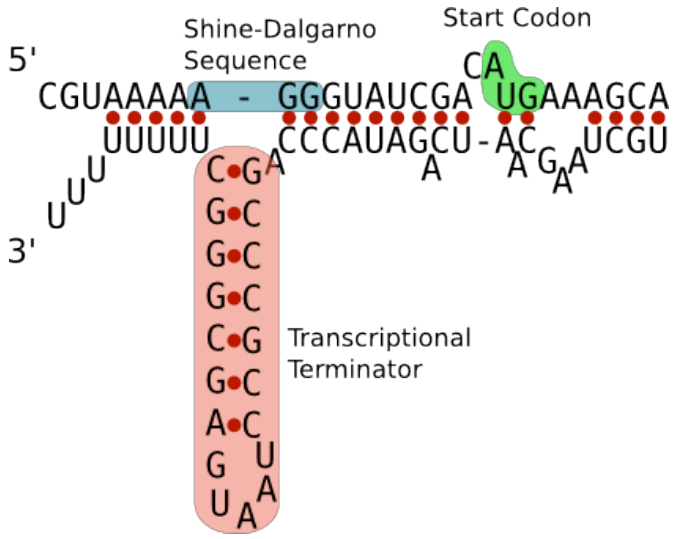


Figure 10: Predicted base-pairing between the beginning and end of the *trp* operon mRNA sequence. Based on the predictions by Das *et. al.*, a similar region was found in the *E. coli trp* operon sequence. The Shine-Dalgarno sequence, start codon, and transcriptional terminator are highlighted, and predicted base pairing is denoted by the red circles. It is quite likely that, similar to *S. typhimurium*, *E. coli* limits leader peptide synthesis in this manner.

To further help understand the secondary structures of the *trp* regulatory sequence, different sequences were tested in UNAFold. After inputting an RNA sequence, the software predicts the most likely secondary structures using a set of rules and parameters. The wild-type *E. coli trp* operon regulatory sequence was analyzed with the goal of simulating the stem-loop structures that have been described in the literature. If the software cannot reproduce what has been shown to be experimentally true, then it is not useful in understanding the secondary structures. However, accurate reproduction of these structures would help simulate the effects of any modifications to the sequence. As shown in Figure 11, the proper RNA secondary structures were predicted. The pause loop structure, responsible for synchronizing the RNA polymerase and the ribosome, spans between bases 20 and 37 (yellow). A second stem-loop structure, from base 46 to base 84, prevents the formation of the anti-terminator stem-loop. A small portion of the anti-terminator stem-loop lies between bases 85 and 95 (green). The full terminator has

formed and is visible between bases 108 and 126 (red).

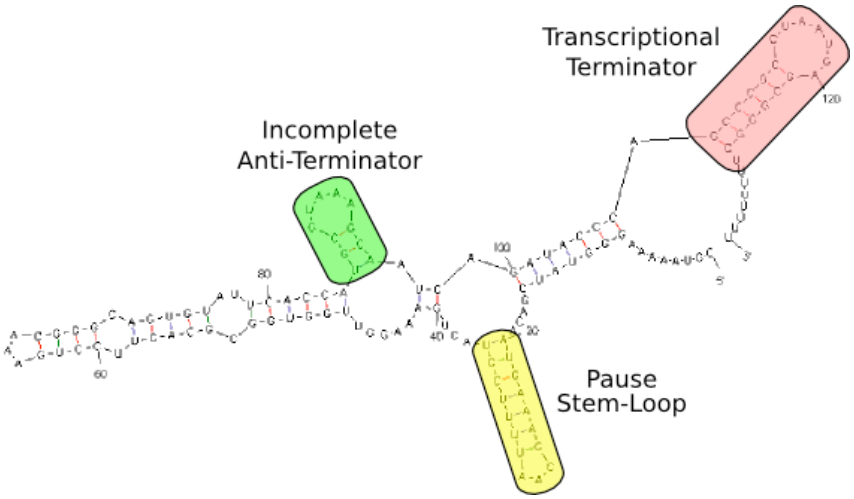


Figure 11: UNAFold prediction for secondary structure of wild-type *trp* operon leader peptide and anti-terminator mRNA sequence in the termination conformation. Coloring denotes the pause stem-loop (yellow), the partial anti-terminator stem-loop (green), and the terminator stem-loop (red).

After the various secondary structures were satisfactorily reproduced in UNAFold for wild-type sequences, an optimal *hixC* sequence insertion point was found by adding the appropriate bases and analyzing the effects on secondary structure. The ideal modification would have no effect on the formation of regions 1, 2, 3, or 4 (Figure 9), allowing the proper development of all stem-loops, depending on tryptophan availability. The *hixC* sequence would also allow for recombination with other DNA elements as catalyzed by Hin.

As shown in Figure 12, the additional bases, between positions 91 and 117, produce a prominent stem-loop structure, extending from the partial anti-terminator loop that results during the termination conformation. Although the effects of such an insertion are currently unknown, it is predicted that it will not affect important transcriptional mechanisms, as evidenced by the near-exact reproduction of all other structural elements throughout the sequence. However, it is feasible that the elongated stem-loop could slow down RNA polymerase, like the pause stem-

loop does. If this pause were too long, then under tryptophan starvation conditions the ribosome might have time to catch up and force the unwanted termination conformation.

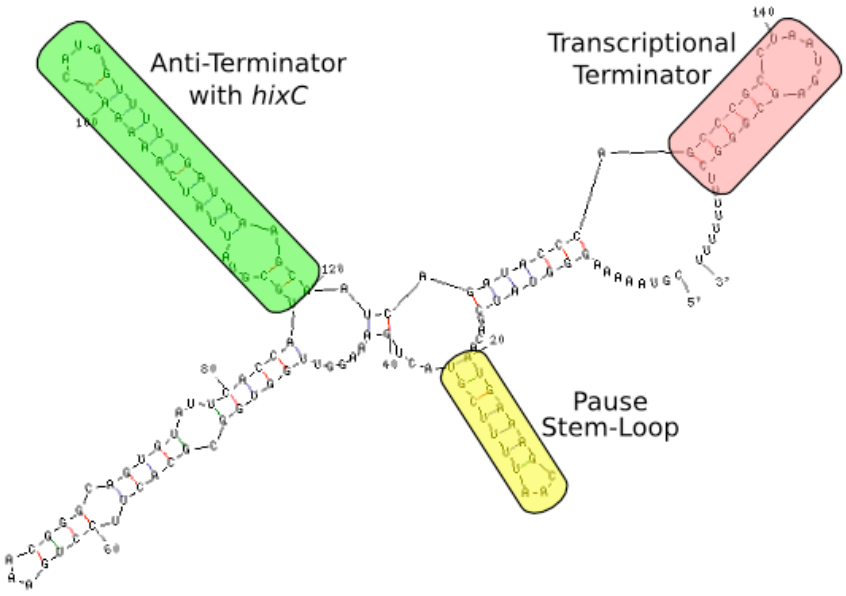


Figure 12: UNAFold prediction for the secondary structure of the *trp* operon leader peptide and anti-terminator mRNA sequence, with the *hixC* insertion, in the termination conformation. Coloring denotes the pause stem-loop (yellow), the partial anti-terminator stem-loop with a *hixC* insertion (green), and the terminator stem-loop (red).

Although the above simulations demonstrated that a *hixC* site insertion into the anti-terminator stem-loop has potential for success, as it probably will not interfere with secondary structures during termination, it is also important to show that this same insertion will not interfere with the secondary structures present during anti-termination. The removal of bases before the beginning and after the end of the anti-terminator bases attempts to simulate the anti-terminator structure (Figure 13A). The insertion of the *hixC* sequence extends the stem-loop structure but does not disrupt its integrity (Figure 13B). This insertion point is appropriate as it is not expected to disrupt either the termination nor the anti-termination secondary structures. A schematic of the entire sequence was drawn, outlining the desirable termination and anti-

termination conformations containing a *hixC* insertion (Figure 14).

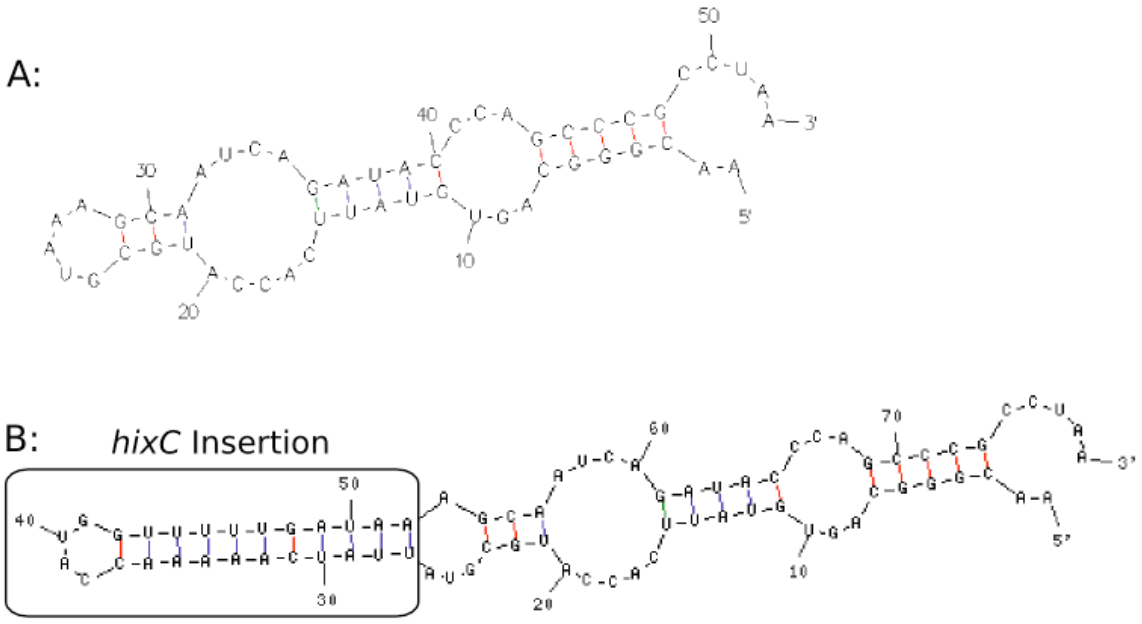


Figure 13: A: UNAFold prediction for secondary structure of the wild-type *trp* operon anti-terminator mRNA stem-loop structure. This structure was generated by removing bases before and after the structure to simulate the presence of a ribosome. **B:** UNAFold prediction for the secondary structure of *trp* operon mRNA anti-terminator stem-loop structure with a *hixC* insertion. The integrity of the anti-terminator stem loop is maintained, although the stem-loop is now longer.

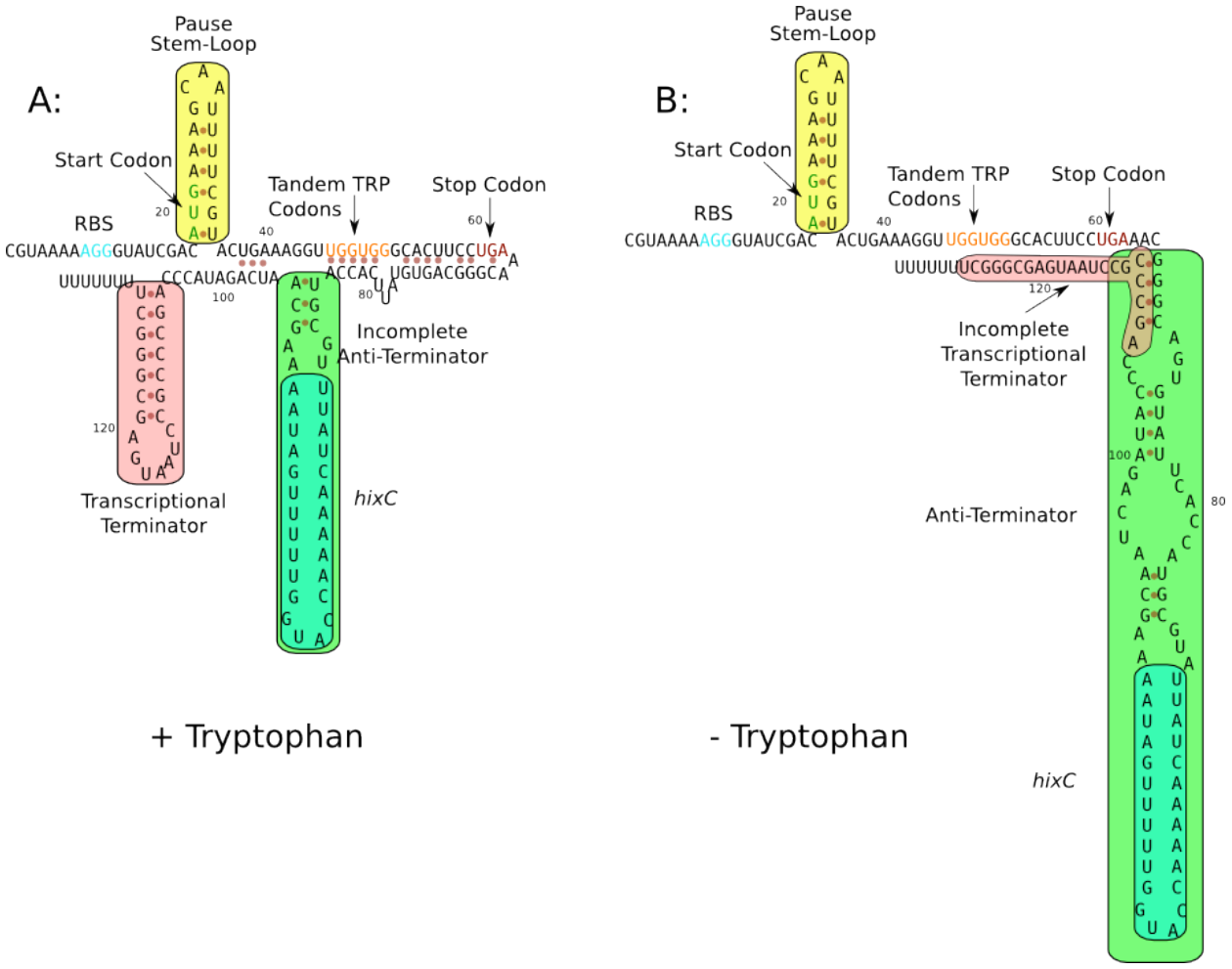


Figure 14: Hypothetical mRNA conformations of the regulatory sequence with the *hixC* insertion. **A:** The mRNA conformation when tryptophan is available. The *hixC* site, shown in blue, protrudes from the incomplete anti-terminator, shown in green. The integrity of the pause loop, in yellow, or the terminator, in red, are not affected. **B:** The mRNA conformation during tryptophan-starvation conditions. The *hixC* site protrudes from a complete anti-terminator stem-loop. The integrity of the pause loop, in yellow is not affected; however, the terminator cannot form, as desired.

Once the DNA sequences and secondary structures were analyzed, and a potential *hixC* insertion location was found, it was necessary to determine how much of the *E. coli* DNA sequence to use in the synthetic construct. Specifically, the entire leader peptide sequence is required, and all the bases including the poly-Us at the end are also necessary (Das *et al.*, 1983). The chosen wild-type sequence from *E. coli* includes a predicted ribosomal binding site

containing the Shine-Dalgarno bases A-G-G, followed by the leader peptide sequence, and ends after eight tandem Us. The modified construct has base 91 replaced with the *hixC* sequence. Both constructs are flanked on either side by the appropriate restriction enzyme sites to create a BioBrick part, making the DNA sequences compatible with all the other sequences in the Registry of Standard Biological Parts.

The DNA sequences were to be initially constructed by the annealing of several overlapping oligonucleotides (Figure 15). Colony PCR and restriction enzyme digests with EcoRI and PstI were used to determine the lengths of the DNA products resulting from the annealing reaction. Initially, it seemed as though the bands in lanes 3, 4, 5, and 7 had bands approximately the correct length of 372 bp (Figure 16). These annealing products were ligated into vector pSB1A2 which previously contained a DNA insert 173 bp in length.

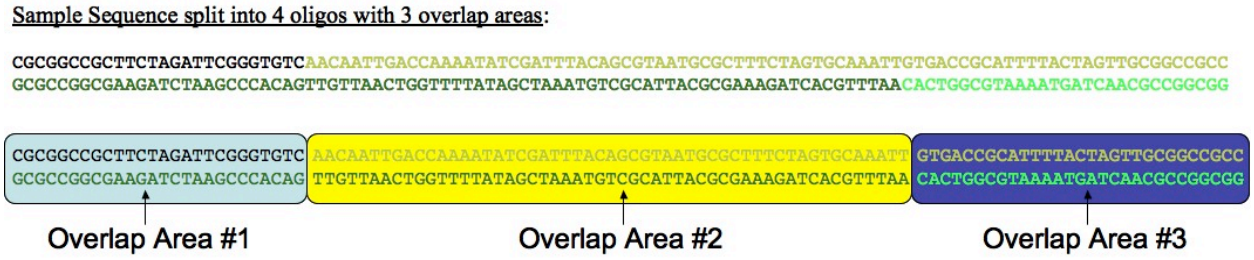


Figure 15: An example of how double-stranded DNA synthesis through oligonucleotide annealing works (image courtesy of Lance Harden). Several short DNA oligonucleotides are synthesized such that their sequences contain overlapping regions that can base-pair. Annealing these oligonucleotides at an elevated temperature forms a final DNA molecule containing the oligonucleotides in the proper order.

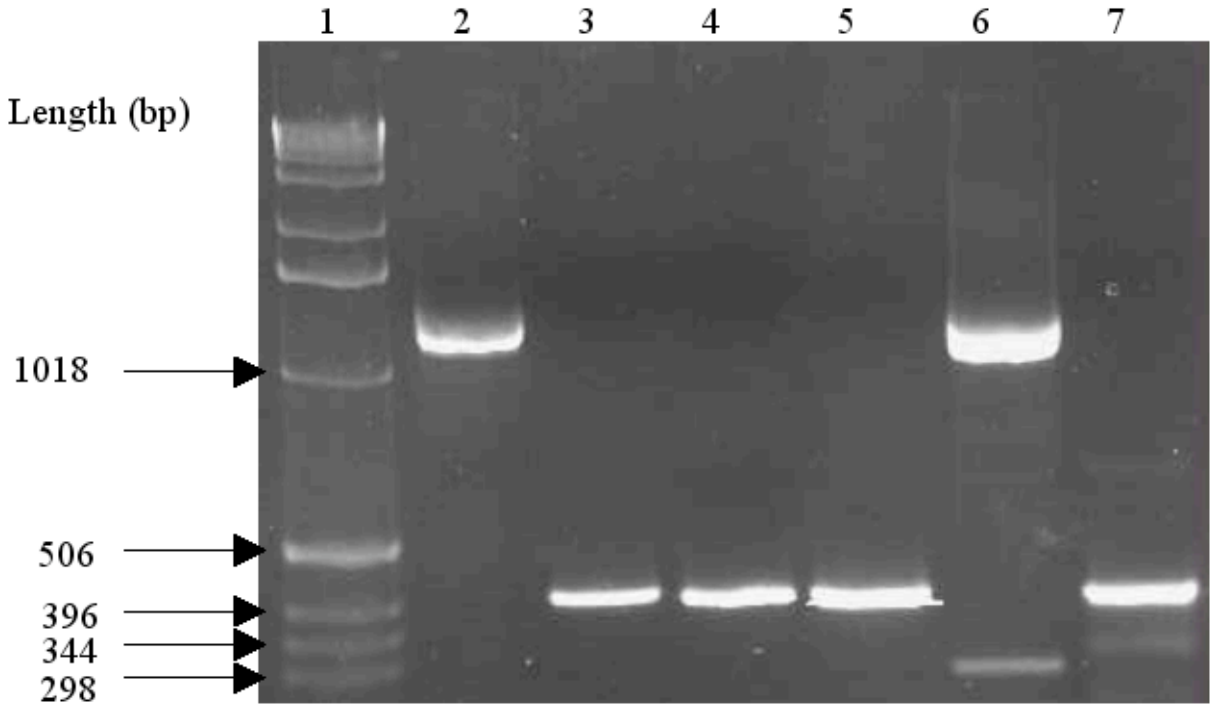


Figure 16: Colony PCR results for the ligation of *trp* DNA assembled by annealing oligonucleotides. Lanes 2 through 7 contain the PCR product from a colony transformed with a ligation of the annealed product and vector pSB1A2. Bands in lanes 2 and 6 were too large, but the bands in lanes 3, 4, 5 and 7 were promisingly close to 372 bp. However, the prior insert in the vector was similar in size to the new insert, and it was too difficult to distinguish between them. There is also an unexplained, shorter band in lane 6. The gel is 1.6% agarose.

Template plasmid from PCR reactions in lanes 3, 4, and 5 of Figure 18 was digested with EcoRI and PstI to verify the length of the insertion. Although the colony PCR procedure is useful for screening many colonies after a ligation, it is more error-prone and less reliable than double-digestions with EcoRI and PstI (Registry, 2007). As a control, the prior insert in the vector was also digested for comparison. The bands in lanes 2, 3 and 4 in Figure 19 are all the same length, close to the 200 bp mark. For unknown reasons, the control in lane 5 shows no band for the insert even though there is a band for the vector at around 2Kb. The length was too similar to the original insert, which was known to be 178 bp, and thus it was too difficult to tell if my desired product had been successfully ligated into the vector (Figure 17).

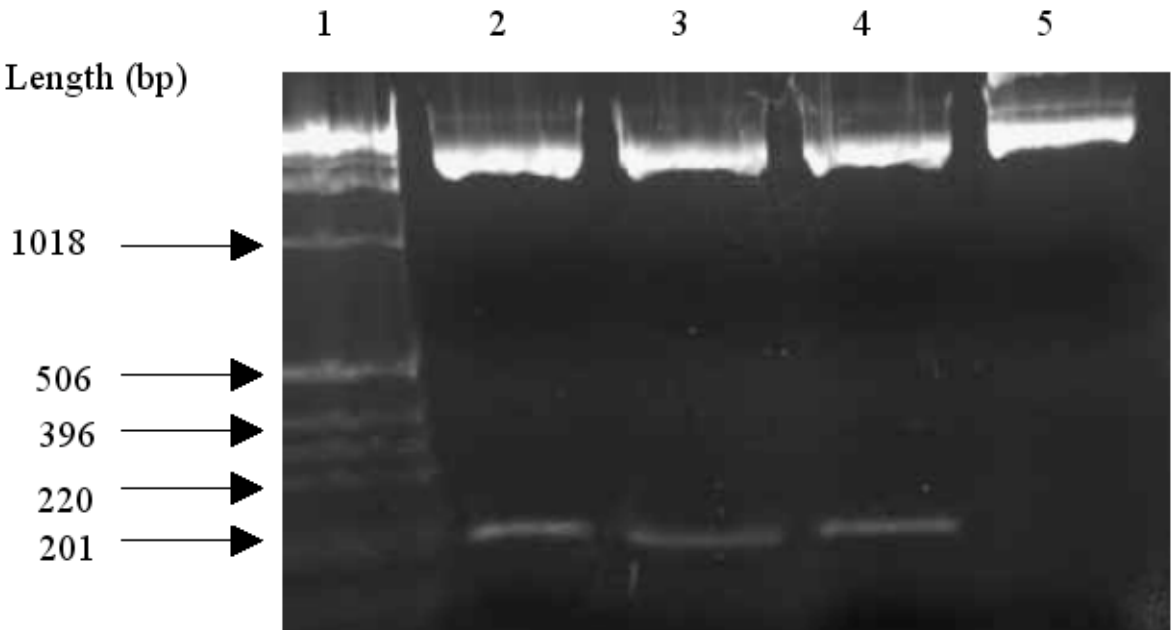


Figure 17: EcoRI and PstI digestions of vectors thought to contain *trp* operon DNA inserts. Lanes 2, 3, and 4 contain digested DNA of the annealed part and pSB1A2. Lane 5 contains digested DNA of the vector before the annealed product was inserted, as was expected to have an insert 178 bp long. The gel is 1.9% agarose.

Due to the difficulty of distinguishing between the previous 178 bp insert and the 134 bp annealing product, the process was attempted with a different vector, one that previously contained the 932 bp red fluorescent protein coding sequence (RFP), making it easier to distinguish on a gel and also permitting screening for ligation success by color. If the vector did not digest completely, then it would contain its original insert, and not the new, desired insert. Any cell transformed with such plasmids would express RFP and be red in color; these cells can be screened against by visual inspection.

Analysis of the transformation results showed that a vast majority of the transformants were red, indicating that most of the cells that grew had vectors which were not fully digested and still contained the original RFP insert. When verified by a PCR screen, none of the ligation

products from white colonies were the correct length, which was expected to be close to 400 bp. Instead, they either had no inserts, shown by the bands with a length near 238 bp in lanes 2, 3, 6 and 7, or the inserts were close to 1Kb in length, in lanes 4 and 5. Lanes 4 and 5 also show faint bands with lengths of approximately 298 bp (Figure 18). The samples from lanes 4 and 5 were verified using digestion with EcoRI and PstI and confirmed to be too large; the faint bands were no longer visible.

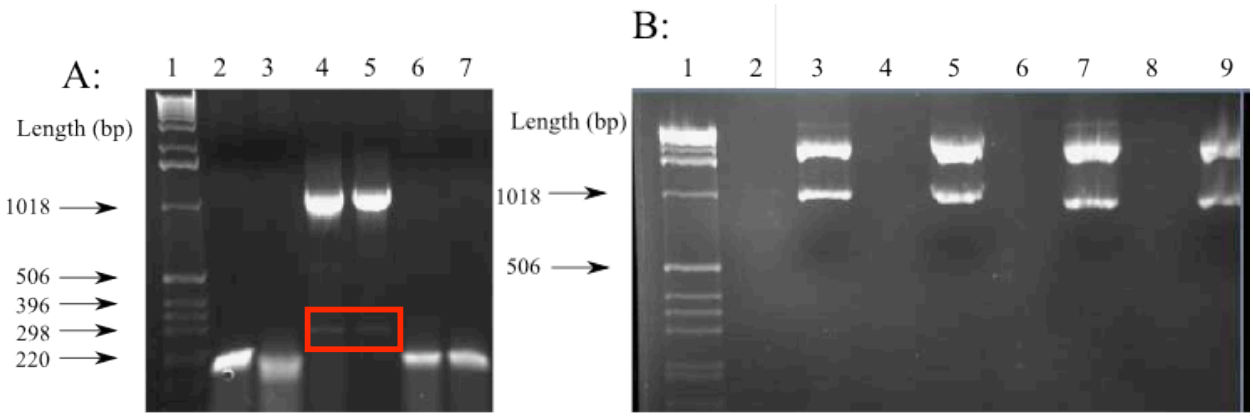


Figure 18: A: Colony PCR results for the ligation of the *trp* annealing DNA segment with vector pSB1A2. The PCR products in lanes 2, 3, 6 and 7 show no insert, and the products in lanes 4 and 5 are much larger than the expected length. Note the faint bands in lanes 4 and 5 at around 298 bp, boxed in red. The gel is 1.1% agarose. **B:** EcoRI and PstI digestion of the DNA from lanes 4 and 5 on the first gel. The inserts are close to 1000 bp, and the faint bands previously visible are no longer there. The gel is 1.7% agarose.

Since the first attempt to assemble the DNA construct failed, synthesis of the wild-type double-stranded DNA was performed by using PCR primers designed to amplify chromosomal DNA from *E. coli*. As seen in Figure 19 A and B, this procedure was successful, with lengths close to the expected 164 bp and 370 bp. Digestion with EcoRI and PstI confirmed the lengths in part B to be the correct length (Figure 20).

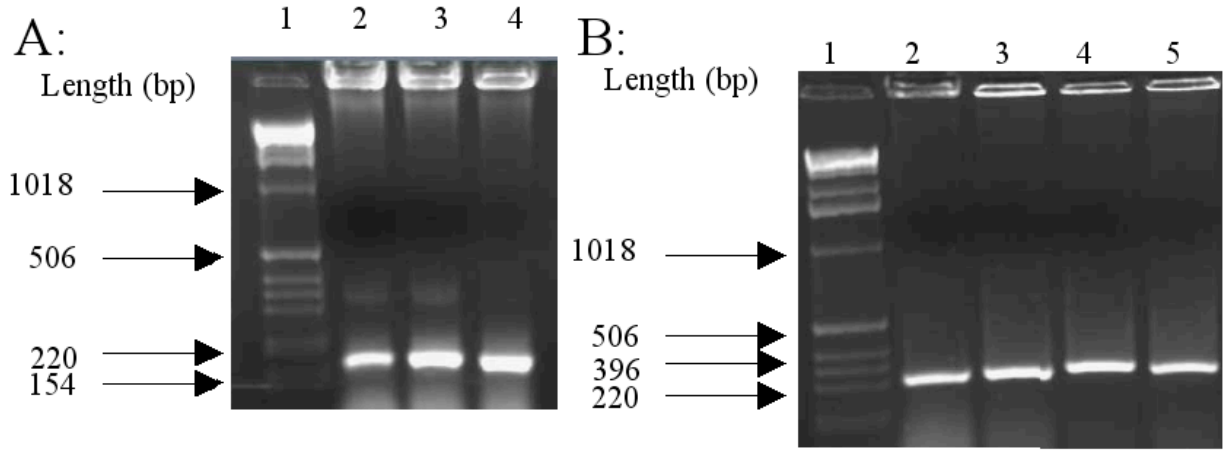


Figure 19: A: Colony PCR for 3 separate reactions to clone the desired tryptophan operon sequence from *E. coli*. The band in each lane is the correct length, lying between the bands at 220 bp and 154 bp. The gel is 1.8% agarose. **B:** The PCR products were gel-purified and subsequently ligated into vector pSB1A2. Colony PCR on the resultant transformations show inserts with lengths between 298 bp and 344 bp, although the expected was closer to 370 bp. The gel is 1.6% agarose.

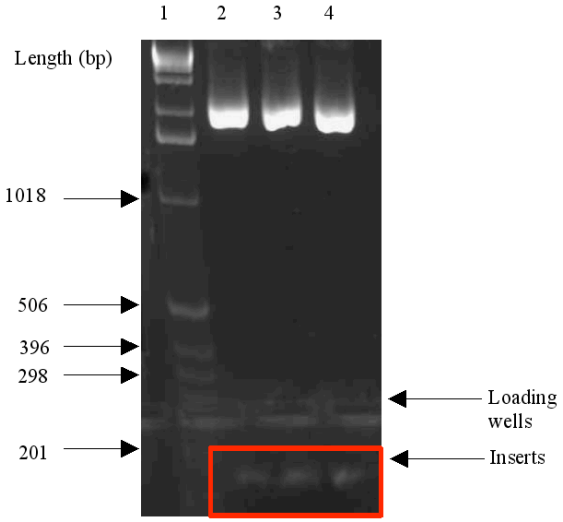


Figure 20: Digestion with restriction enzymes EcoRI and PstI show that multiple, distinct colonies have the correct insert in their plasmids, with the lengths between 201 bp and 154 bp, near the correct length of 164 bp. The gel is 1.8% agarose. The wells are shown in the picture because the gel was allowed to run for an extra amount of time, giving a better size resolution.

Once the lengths of the PCR products were confirmed to be approximately correct using gel electrophoresis, purified DNA of each of 3 different products was sequenced to confirm that

the sequences were correct and that there were no point mutations introduced. All three sequences were perfect matches to the desired sequence (Figure 21).

```
a      -TAGGCGTATCACGAGGCAGAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAA 59
ref    ------------------GAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAA 41
b      ATAGGCGTATCACGAGGCAGAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAA 60
c      ATAGGCGTATCACGAGGCAGAATTCGCGGCCGCTTCTAGACGTAAAAAGGGTATCGACAA 60
                        ****************************************

a      TGAAAGCAATTTTCGTACTGAAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTC 119
ref    TGAAAGCAATTTTCGTACTGAAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTC 101
b      TGAAAGCAATTTTCGTACTGAAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTC 120
c      TGAAAGCAATTTTCGTACTGAAAGGTTGGTGGCGCACTTCCTGAAACGGGCAGTGTATTC 120
       ************************************************************

a      ACCATGCGTAAAGCAATCAGATACCCAGCCCGCCTAATGAGCGGGCTTTTTTTTACTAGT 179
ref    ACCATGCGTAAAGCAATCAGATACCCAGCCCGCCTAATGAGCGGGCTTTTTTTTACTAGT 161
b      ACCATGCGTAAAGCAATCAGATACCCAGCCCGCCTAATGAGCGGGCTTTTTTTTACTAGT 180
c      ACCATGCGTAAAGCAATCAGATACCCAGCCCGCCTAATGAGCGGGCTTTTTTTTACTAGT 180
       ************************************************************

a      TGCGGCCGCTGCAGATGCAGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 239
ref    TGCGGCCGCTGCAG--------------------------------------------- 175
b      TGCGGCCGCTGCAG------GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 234
c      TGCGGCCGCTGCAG------GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 234
       **************
```

Figure 21: ClustalW alignment of the three sequences generated using PCR and the desired reference sequence. A, B, and C are the sequencing results, and ref is the sequence that was expected. The asterisks show the bases where all four sequences match. The desired sequence begins with the EcoRI restriction enzyme site GAATTC and ends with the PstI restriction enzyme site CTGCAG.

Once the wild-type DNA had been isolated, it was necessary to begin testing its function as a regulator of gene expression. Use of the anti-terminator sequence requires the addition of an upstream promoter to initiate transcription and the presence of a reporter gene downstream. An "on-switch" device was designed to test the functionality of the anti-terminator. Under normal conditions, when tryptophan is available, the *trp* operon regulatory sequence will attenuate transcription prematurely. The reporter gene will not be transcribed, and there will be no visible phenotype; this is considered the "off" position of the switch. During tryptophan-starvation conditions, however, the anti-termination mRNA conformation will develop, and transcription of the entire reporter coding sequence will follow, resulting in a visible phenotype, or the "on"

position of the switch (Figure 22).

The two chosen promoters were pLac and pBad. pLac has been shown to initiate transcription of GFP enough to get ample fluorescence, and is also inducible by IPTG if more transcription is desired (Haynes *et al.*, 2007). pBad is inducible by arabinose and also had been tested extensively in the past, although it is generally weaker than pLac (OpenWetware, 2007). Downstream of the anti-terminator sequence is a ribosomal binding site with strong ribosomal binding affinity, the GFP coding sequence with an LVA tag for rapid protein degradation, and a transcriptional terminator. The LVA-induced degradation reduces the time required to return to an "off" state from an "on" state, which provides faster detection of changes in the system's environment and allows GFP to degrade on the order of hours rather than days (Wendland and Bumann, 2002).

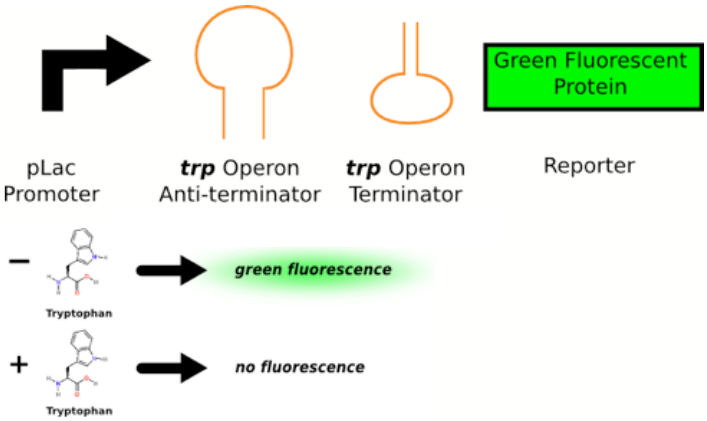


Figure 22: The device used to test the wild-type anti-terminator. Under tryptophan starvation conditions, such as growth on M9 media, the anti-terminator loop should form, and the transcription will result in the production of GFP mRNA and green fluorescence. Otherwise, in the presence of tryptophan, no GFP will be produced.

An analogous device was designed to test the sequence modified with the *hixC* site insertion. If the *hixC* insertion does not disrupt the secondary structures critical to anti-termination and termination, then the anti-terminator/terminator pair can be used in a pancake

computer. This device cannot be constructed until I have developed a means to generate the modified *trp* regulatory sequence with the *hixC* insertion (Figure 23).

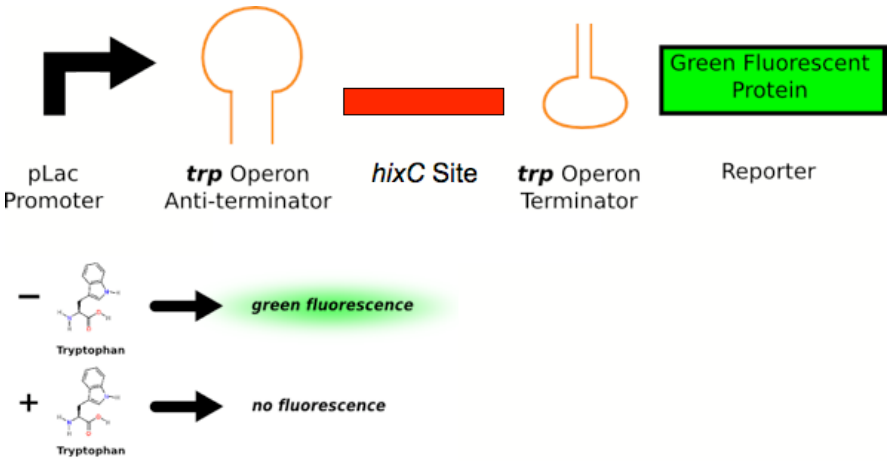


Figure 23: The device used to test the anti-terminator with a *hixC* insertion. The device is supposed to act identically as the wild-type device, but the additional *hixC* sequence will enable Hin-mediated DNA inversion, giving the option of separating the anti-terminator from the terminator.

System Construction and Implementation

The entire device with the pLac promoter has been assembled, as shown by the following series of gel images. First, a ribosomal binding site was added upstream of GFP.lva-TT (Figure 24), and the tryptophan operon DNA sequence has been ligated upstream of RBS-GFP.lva-TT. Next, the promoter was ligated upstream of Trp-RBS-GFP.lva-TT, allowing for testing. In total, 4 new parts were successfully built and added to the Registry of Standard Biological Parts (Table 1).

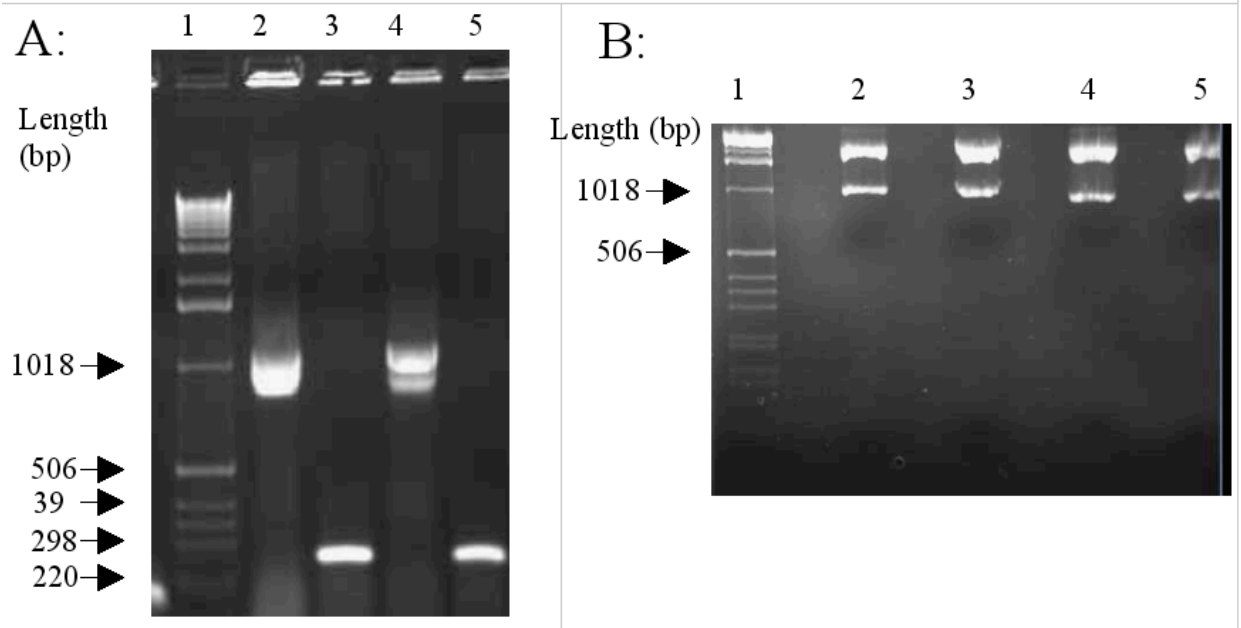


Figure 24: A: PCR colony screening showing the successful ligation of the ribosomal binding site to GFP.lva-TT in lanes 2 and 4; the expected length was 1130 bp. The gel is 1.1% agarose. B: these results were confirmed by digestions with EcoRI and PstI, showing the bands in lanes 4 and 5 to be below 1000 bp, close to the expected 908 bp. The gel is 1.7% agarose.

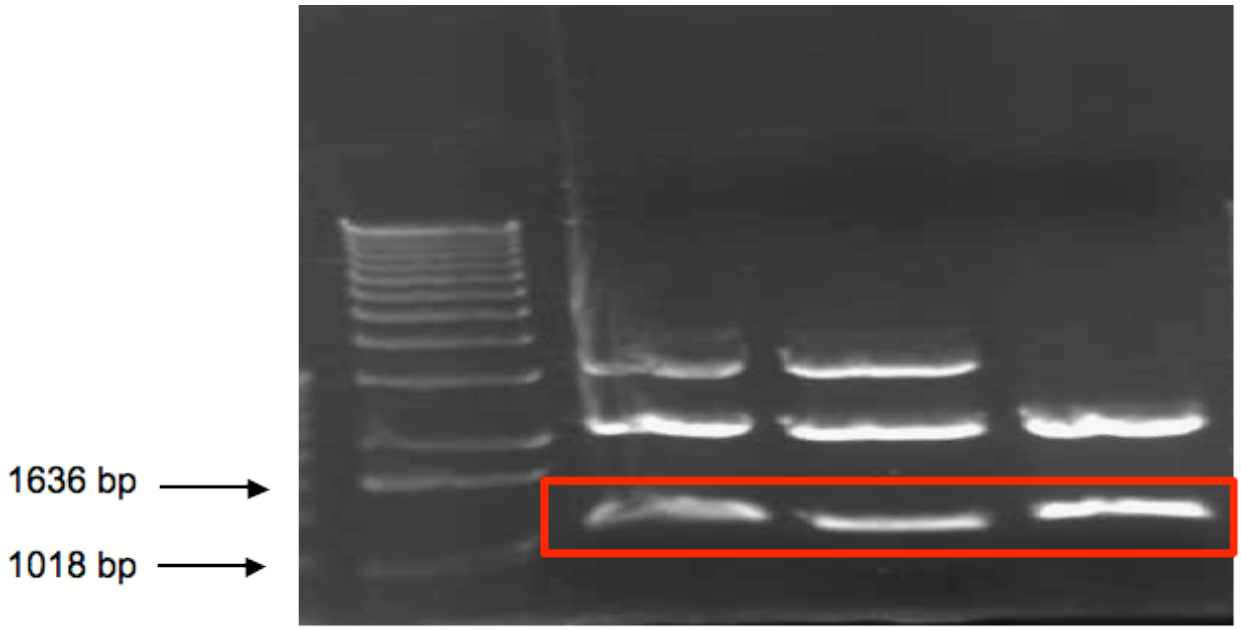


Figure 25: EcoRI and PstI digestion of miniprepped DNA for the ligation of the *trp* regulatory sequence in front of RBS-GFP.lva-TT. The expected length was 1056 bp, and all 3 lanes have a band in the correct area. The gel is 0.7% agarose.

Testing of the *trp* regulatory sequence requires the finalized assembly of the GFP reporter system. The last part required was the addition of a promoter sequence. Two different promoters were chosen to be added upstream, pLac and pBad. The promoters were digested as vectors, with expected base pair lengths of 2279, in lanes 2 through 5, and 2209, in lanes 6 through 10 on the first row. The back insert was Trp-RBS-GFP.lva-TT, with an expected length of 1056 bp, in lanes 2 through 4 on the second row. These digestions were confirmed successful on a 0.6% agarose gel (Figure 26).

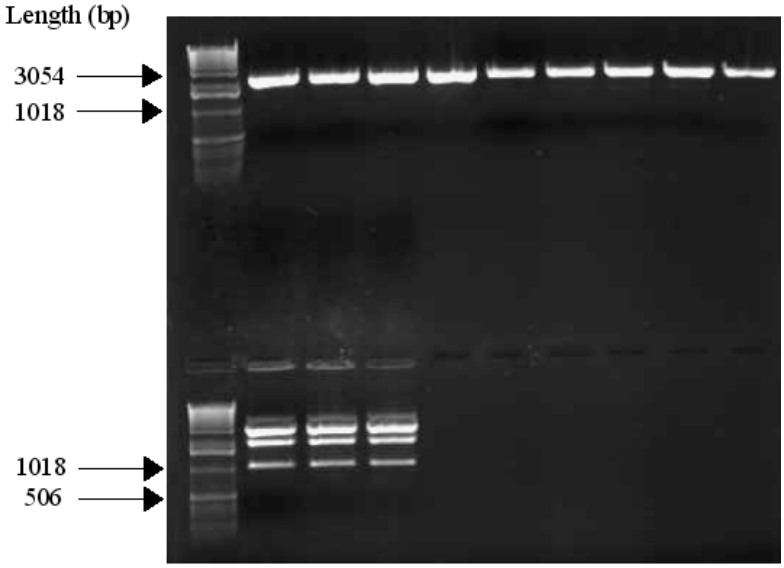


Figure 26: Digestions of pLac as a back vector, lanes 2-5, pBad as a back vector, lanes 6-10, and Trp-RBS-GFP.lva-TT as a back insert on the second row. The digested vectors were supposed to have lengths greater than 2kb, and the insert was expected to be 1056 bp long. The bands are the appropriate lengths, and they were gel-purified. The gel is 0.6% agarose.

After gel-purifying the excised gel fragments, two separate ligations were performed. In the first ligation, the insert was placed downstream of the pLac promoter. In the second ligation, the insert was placed downstream of the pBad promoter. The ratios of the negative control colonies, with no insert, compared to the experimental colonies, with a ligation product, were

similar, indicating that the efficiency of the ligation and transformation was poor. However, several colonies showed green fluorescence on the plate with the pLac promoter vector. These colonies were picked and miniprepped, and the DNA was digested with restriction enzymes EcoRI and PstI for length verification. The desired length of a successful ligation was 1264 bp, and all five bands were the correct length (Figure 27). The band in lane 4 was the brightest, and so this culture was frozen down and kept.

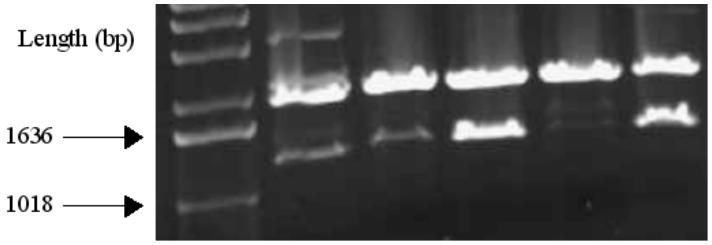


Figure 27: Digestion of the five miniprepped DNA samples from the plate containing ligation products for the ligation of pLac with Trp-RBS-GFP.lva-TT. The desired insert length was 1264 bp, which was found in each lane. The gel is 0.5% agarose.

Although pLac-Trp-RBS-GFP.lva-TT was successfully built, the part with the pBad promoter was not. Rather than performing many colony PCR screens with the hope of finding a ligation that worked, an attempt was made to plate out 24 individual colonies on minimal media with the hope that at least one would fluoresce green. Some green fluorescence may be expected because the *trp* regulatory sequence would turn on GFP expression in minimal media. This screen would provide a quicker way of screening many colonies at once. The screen did not find any colonies with green fluorescence, and so it was decided that none was a successful ligation. However, in retrospect it was realized that the pBad promoter is strongly induced by arabinose, which was not added to the growth media. Thus a lack of green fluorescence may have been due to poor promotion, and not the failure of the ligation reaction.

In addition to building the reporter plasmid with the wild-type *trp* regulatory sequence

and testing the ability of this sequence to attenuate transcription, it was also necessary to see how this sequence, modified with a *hixC* insertion, would behave. Ideally, the modified sequence would regulate gene expression identically as the unmodified sequence, but this must be verified experimentally. The first method of assembling this part through oligonucleotide annealing failed, so another method was required. It was determined that it may be possible to complete the assembly using PCR and specially designed primers (Figure 28; Davies *et al.*, 2007). Amplifying the sequences to the left and to the right of the *hixC* insertion alone would not give the hixC insertion. However, if one of the PCR primers in each of these reactions was concatenated with the entire *hixC* sequence and a part of the opposite side's sequence, then a third reaction combining the PCR products of the first two reactions would, in theory, yield the entire sequence with the *hixC* sequence inside.

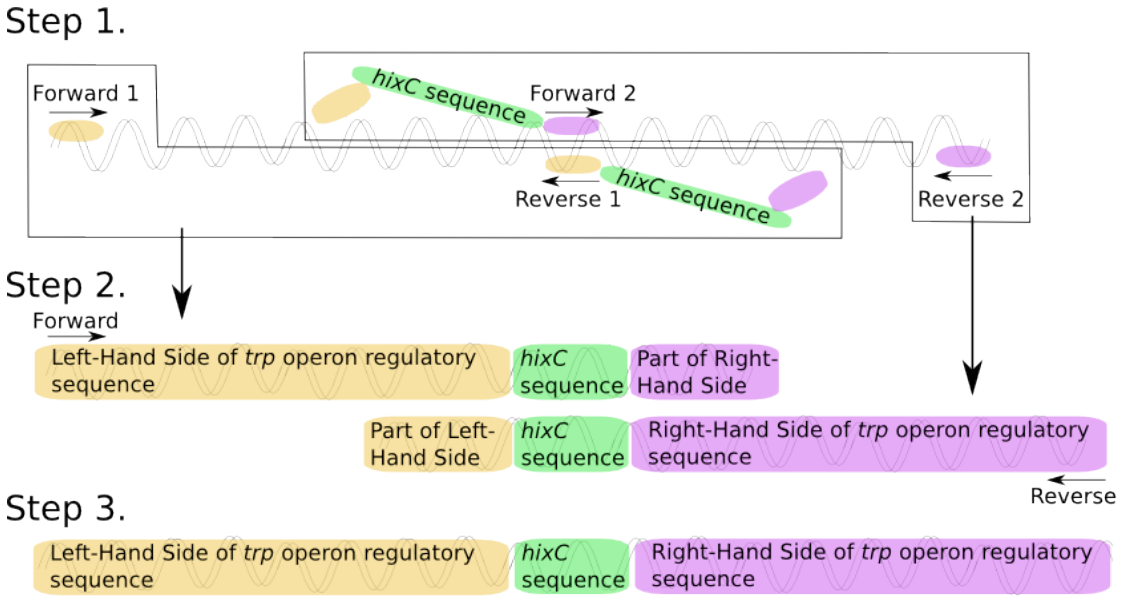


Figure 28: Procedure for the synthesis of the *trp* regulatory sequence with the *hixC* sequence insertion. In the first step, two separate PCR reactions are performed, amplifying a different part of the *trp* sequence. The left-hand side is shown in orange, and the right-hand side is shown in purple. In addition, the *hixC* sequence is appended to one of the extremities, followed by a short segment of the other side of the *trp* sequence. This is done on the reverse primer of the left-hand

side reaction, and the forward primer of the right-hand side reaction. When the resulting PCR products are isolated, mixing them together in another PCR reaction should result in the amplification of the entire *trp* sequence with the *hixC* insertion.

This method to produce the modified *trp* regulatory sequence was attempted multiple times. After the first attempt, it was realized that several mistakes were made. First, the gel was hard to read because the ladder faded too soon, but no more PCR product was left, and so the gel could not be run again (Figure 29). Second, the parameters on the PCR reaction were set to too few cycles.

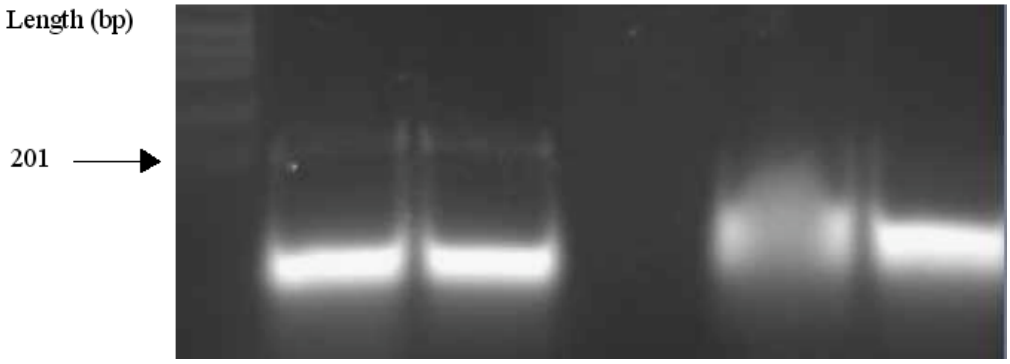


Figure 29: The PCR products of the first attempt to synthesize the *trp* operon sequence with the *hixC* insertion. The ladder is too faint to determine the lengths of the DNA products, so this gel cannot be interpreted with enough certainty. The gel is 2.0% agarose.

These issues were taken into account, and the reaction was repeated. However, this time other mistakes were made. When loading the gel, loading dye instead of 1 kb ladder was loaded into the first lane, which made it nearly impossible to decipher base-pair lengths. Furthermore, rather than producing distinct bands, a smear of many bands was in each lane (Figure 30). This indicated that it would be necessary to increase the temperature so as to require more specific binding.

Figure 30: This agarose gel shows both the digestion of the pBad promoter as a back vector in psb1A2 and the PCR products for the first reaction in the attempt to amplify the *trp* regulatory sequence with the *hixC* sequence insertion. Instead of 1 kb ladder, loading dye was mistakenly added in the first well. Since lengths of the PCR products could not be determined, and since there were smears, and not definite bands, the PCR reaction was redone with altered parameters. The left-hand side PCR is in the lane to the left of the right-hand side reaction, showing that both produced similar smears. The gel is 2.0% agarose.

The PCR protocol was modified by increasing the temperature for step 4 from 62º C to 66º C. The reaction was performed in duplicate for the left-hand side reaction, lanes 2-3, and in duplicate for the right-hand side reaction, lanes 4-5. The PCR products were run on a 2.0% agarose gel (Figure 31). In each of these lanes, there are multiple bands instead of the single desired band. Several bands were excised based on the predicted lengths of the desired PCR products, thought to be 116 bp and 69 bp, respectively. However, it was later realized that, when taking into account the BioBrick ends, these lengths should really be 162 bp and 115 bp. Lane 6 contains the PCR product resulting from the reaction using the forward primer for the left-hand side reaction and the reverse primer of the right-hand side reaction. These were the two primers used for the original chromosomal amplification of the *trp* operon sequence, and serve as a positive control to show that the primers and PCR reaction were done properly. Lane 7 contains the PCR product for the use of the forward primer of the right-hand side reaction and the reverse primer of the left-hand side reaction. No bands were expected in lane 7, but if any were to

appear, it may indicate that there are issues with unspecific binding. There are no bands in this lane, showing that there was no unspecific binding.

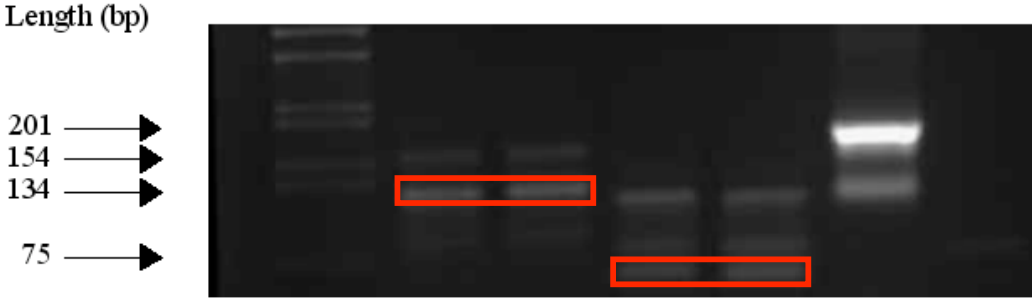


Figure 31: The first step for the third attempt at the PCR amplification of the *trp* operon regulatory sequence with an added *hixC* sequence. Lane 1 contains the 1 kb ladder. Lanes 2 and 3 contain the PCR products for the left-hand side reaction. Lanes 4 and 5 contain the PCR products for the right-hand side reaction. Lane 6 serves as a positive control using the same forward and reverse primers used to amplify the wild-type trp operon regulatory sequence. Lane 7 serves as a negative control, using the forward primer for the right-hand reaction and the reverse primer of the left-hand reaction, which is not expected to give any PCR products. The bands highlighted in red were excised. The gel is 2.0% agarose.

After performing the first step of the PCR reaction to produce the *trp* regulatory sequence with the *hixC* insertion, several gel fragments were excised and gel-purified. These were then mixed to have primer concentrations of 0.5 pmol/µl. The PCR protocol was modified to have an annealing temperature of 78º C, as computed by the Proligo melting temperature website. The reaction was performed in triplicate, and the resulting PCR products were run on a 2.0% agarose gel (Figure 32). The final PCR products of a successful reaction were predicted to be 208 bp long. No PCR products were detected in the agarose gel, so it was concluded that the reaction did not successfully amplify the DNA as desired.

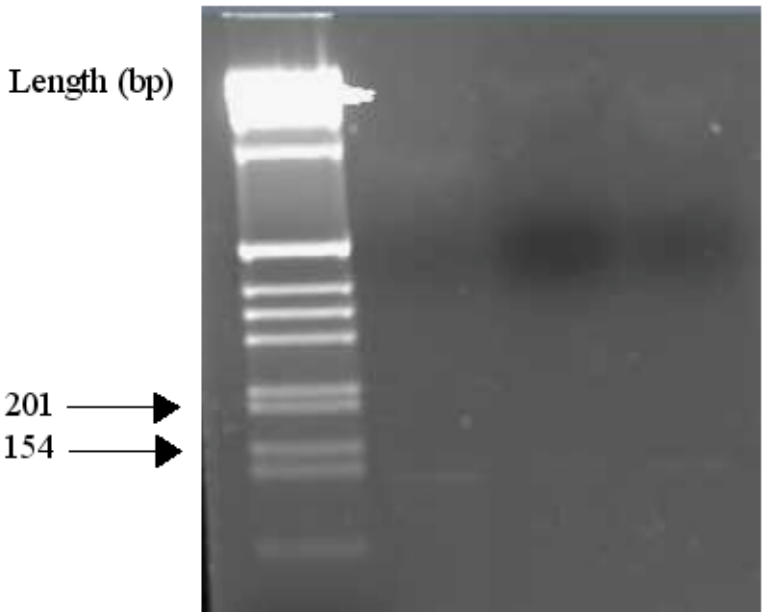


Figure 32: PCR products for the second step in the attempt to synthesize the *trp* regulatory sequence with the *hixC* insertion. Although faint bands can be seen under the 154 bp marker, these are shorter than the expected 208 bp and are probably showing some of the original primer. Since the proper bands did not show, the reaction was not considered successful.

Since the second step in the PCR assembly of the modified *trp* regulatory sequence failed, it is possible that the excised bands from the first step were the wrong ones. The first step was repeated, and the larger bands were excised and purified (Figure 33). Lanes 2-4 contain PCR product for the left-hand side reaction, and lanes 5-6 contain the PCR products for the right-hand side reaction. Lane 7 contains the reaction using the forward primer of the left-hand side reaction and the reverse primer of the right-hand side reaction, serving as a positive control, and producing the wild-type *trp* sequence. The band in lane 4 was not excised because it was too faint. The bands are closer to the desired 162 bp and 115 bp than those excised in the previous attempt.

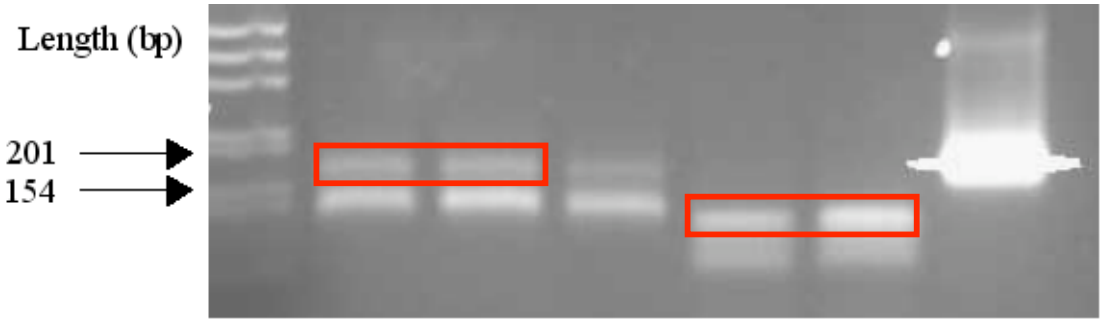


Figure 33: A repeat of the first step of the PCR amplification of the *trp* regulatory sequence with the *hixC* sequence insertion. Lanes 2-4 contain the PCR products of the left-hand side reaction, and lanes 5-6 contain the PCR products for the right-hand side reaction. Lane 7 serves as a positive control using the forward primer of the left-hand side and the reverse primer of the right-hand side. The bands closest to the 162 bp and 115 bp marks were excised (boxed in red). The gel is 2.0% agarose.

After the PCR products from the first step were isolated and purified, the second PCR reaction step was performed. The reaction was done in triplicate, with an elongation time extended to 20 seconds, up from 15 seconds, and the annealing temperature lowered to 70.0º C from 78.0ºC. The longer elongation time gives more time for the entire DNA strand to be elongated, and the lower melting temperature allows for base-pair complementation to occur more easily. These modifications were done with the hope that they would enhance the chance of amplifying more PCR product. Unfortunately, there were no bands in either of the three lanes (Figure 34). Thus the second step failed again.

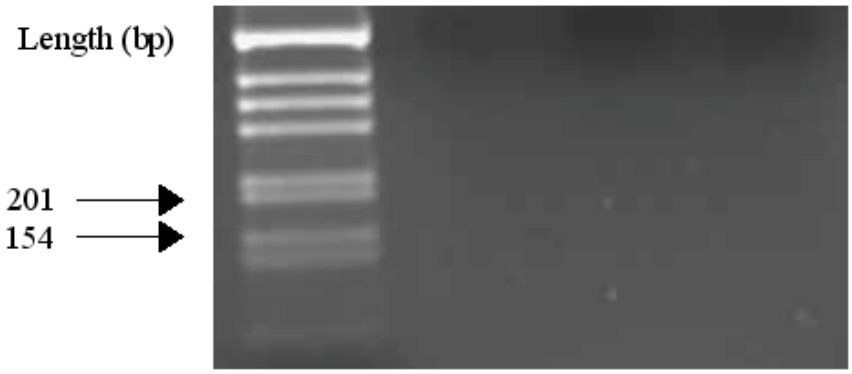


Figure 34: The PCR products for the new attempt at the second step of the assembly of the *trp* regulatory sequence with the *hixC* sequence insertion. The desired final product would have had a length of approximately 208 bp, but there were no bands in any of the lanes. The gel is 2.0%

agarose.

While trying to construct the *trp* sequence with the *hixC* insertion, other attempts were made to put the *trp* regulatory sequence and GFP in front of the pBad promoter so as to gain more flexibility for testing and evaluating GFP expression. For this new attempt, the ligation and transformation were screened by colony PCR to check for plasmids with the correct length, which was expected to be 1432 bp. None of the colonies screened had the correct plasmid, so the ligation was deemed a failure (Figure 35).

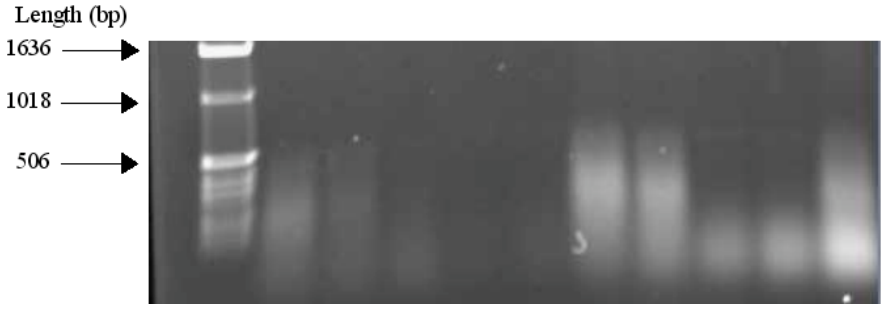


Figure 35: The colony PCR results for the ligation of pBad with Trp-RBS-GFP.lva-TT show that none of the ligations screened were successful. The expected length of a successful ligation in a PCR screen is 1432 bp. The gel is 0.6% agarose.

Due to the bad ratios of the number of colonies on the negative control plate compared to the number of colonies on the ligation plate, shrimp alkaline phosphatase was used to reduce the occurrence of unwanted plasmid self-ligation. SAP removes the phosphate from the 5' end of the plasmid, so the only DNA that can ligate to the vector is the desired insert. The ligation and transformation reactions were done with the dephosphorylated vector. This time, the ratios between the control and experimental plates were ideal, as there was no growth on the negative control plate but lots of growth on the experimental plate. Colony PCR screens were used to check for successful ligations, but none was found (Figure 36).

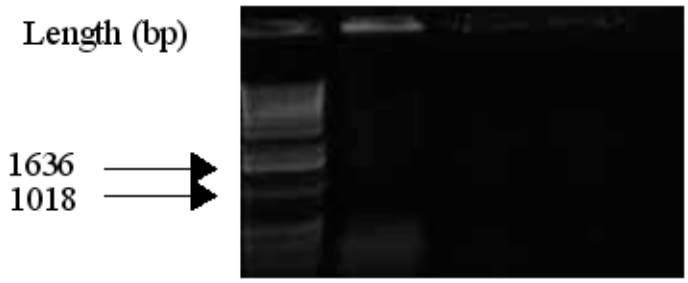


Figure 36: Colony PCR screen for the ligation of the PBAD promoter with Trp-RBS-GFP.lva-TT. Although SAP was used, and the ratios of colony counts on the experimental and negative control plates were ideal, no successful ligations were found. The length of a successful ligation product was 1432 bp. The gel is 0.7% agarose.

Only two colonies were screened the first time, under the assumption that such good ratios would guarantee a successful ligation in nearly every picked colony. It was possible, however, that more needed to be picked. Six more colonies were picked and screened using the same colony PCR protocol (Figure 37). Again, none of the colonies had the desired product with the length of 1432 bp. There were no further attempts made at assembling this part, and it remains unclear why the ratios were good yet no ligation products were found.

Figure 37: Colony PCR screen for six colonies expected to have the PBAD and Trp-RBS-GFP.lva-TT ligation product. A successful ligation would appear as a band with a length of 1432 bp, which does not occur in any of the lanes. The gel is 0.7% agarose.

Once the entire part with the *trp* regulatory sequence, a promoter sequence, and GFP sequence had been completed, initial testing on phenotypes was undertaken on LB media and on minimal M9 media. These initial tests did not show any fluorescence in cells containing the *trp* regulatory sequence plasmid, either on M9 or LB media.

In case the LVA tag on GFP was reducing expression too much on minimal media, the design and construction of a GFP reporter plasmid, with the *trp* regulatory sequence but without the LVA tag, was begun. This assembly required two separate ligations: putting the GFP sequence downstream of the *trp* regulatory region, and then putting this new sequence downstream of a promoter. The RBS-GFP-TT part was digested with restriction enzymes XbaI and PstI as a back insert (Figure 38). The ligation and transformation were verified by colony PCR screening and were successful, as indicated by the bands in lanes 3-5 (Figure 39). Although this part requires the addition of an upstream promoter for it to provide the desired function, it was later decided that the part would not be necessary, and so synthesis was stopped.

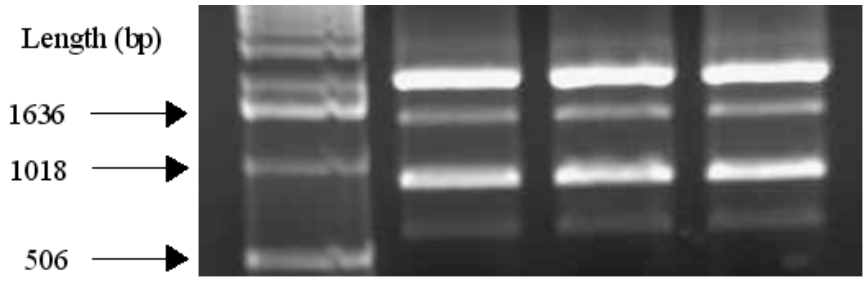


Figure 38: Digestion of RBS-GFP-TT with XbaI and PstI as a back insert. The length of the sequence is 875 bp, so the bands near the 1 kb region were excised and gel-purified. There were other bands present in each lane, but their identities are unknown. The gel is 0.7% agarose.

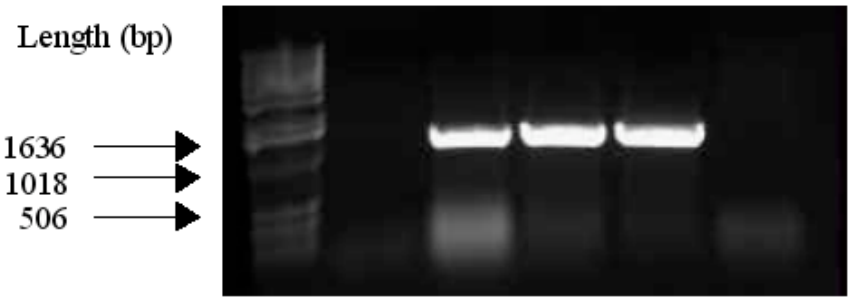


Figure 39: Colony PCR screen results for the ligation of the *trp* regulatory sequence with RBS-GFP-TT. The three bands are in the correct vicinity for the desired product length of 1255 bp. This version of GFP does not have an LVA tag, and could potentially give better fluorescence under minimal growth conditions. The gel is 0.7% agarose.

Once the colony PCR screen revealed some successful ligations, these cultures were

grown overnight and miniprepped. The DNA was digested with restriction enzymes EcoRI and SpeI for the final length verification step and run on a gel (Figure 40). Two of the samples do contain a band with the correct length, in lanes 2 and 3. However, they also contain other bands with different lengths whose identities are unknown. Lane 4 is also hard to interpret, as the entire lane is smeared, and there are no distinct bands. One possible explanation for the extra bands is that other DNA products were introduced into two of the samples. Purification of the bands with the correct length and re-ligation into a vector would yield only the desired DNA products. Once this is done, then the intermediate part will be fully assembled; an upstream promoter would still be needed to complete a functional part.

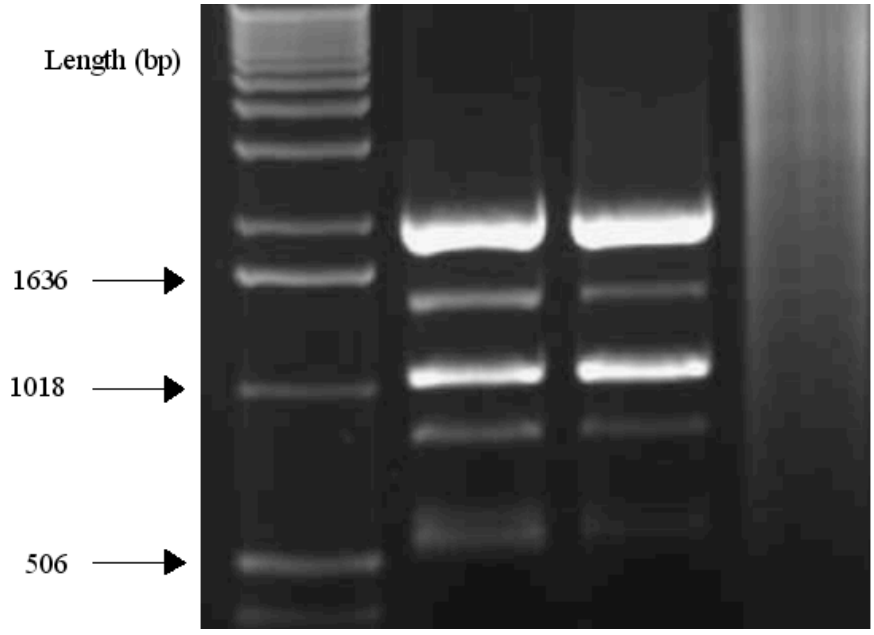


Figure 40: EcoRI and PstI digestions of the samples previously confirmed to contain the ligation of the *trp* regulatory sequence with RBS-GFP-TT. The expected length of the final product was 1017 bp. Bands with this length are in lanes 2 and 3, although there are also other unwanted bands with other lengths in these lanes. Lane 4 has no distinct bands, but instead has a large smear. The gel is 0.7% agarose.

Table 1: Name, description, and build status of new parts added to the Registry.

Part Number	Description	Status

BBa_S03903	RBS-GFP.lva-TT	Built, length verified
BBa_I715052	Wild-type *trp* operon sequence	Built, sequence-verified
BBa_I715053	*Trp* operon sequence with a *hixC* insertion	Assembly failed
BBa_S03912	wild-type *trp*-RBS-GFP.lva-TT	Built, length verified
BBa_S03913	*Trp*/*hixC*-RBS-GFP.lva-TT	Planning
BBa_I715054	pLac-wild-type *trp*-RBS-GFP.lva-TT	Built, length verified
BBa_I715055	pLac-*Trp*/*hixC*-RBS-GFP.lva-TT	Planning
BBa_I715059	pBad-wild-type *trp*-RBS-GFP.lva-TT	Assembly failed
BBa_I715060	pBad-*Trp*/*hixC*-RBS-GFP.lva-TT	Planning
BBa_S03921	wild-type *trp*-RBS-GFP-TT	In progress

The turning on and off of the *trp* operon regulatory sequence requires the ability to control for the presence or absence of tryptophan. The current method to do this is to deprive cells of tryptophan by growing them on minimal M9 media, turning on the device, and to provide cells with tryptophan by growing them on LB media or by adding the amino acid directly, thereby turning off the device. The effectiveness of this type of regulation was tested in multiple ways. First, different combinations of growth media, cell strains, and GFP reporter plasmids were grown on agar plates and analyzed (Table 2). There were no conclusive patterns in GFP expression, although the MC4100 *E. coli* strain consistently proved to grow better on minimal media and expressed GFP more readily than the JM109 strain. Most importantly, however, was that GFP expression profiles were not consistent within these parameters. For example, MC4100 cells grown on M9 with the GFP-LVA plasmid showed increased GFP expression as colony density decreased. This inconsistency in gene expression is problematic because it becomes more difficult to know when GFP expression should be expected. If the positive controls do not behave as expected, then it is not possible to interpret how the experimental plasmids with the *trp* regulatory sequence should behave. In other cases, the transfer of individual colonies that

expressed GFP to new media plates resulted in a loss of GFP expression. Again, this is problematic because a clonal colony should maintain its phenotype.

Table 2: Growth phenotypes for either the MC4100 or JM109 *E. coli* strains in either nutrient-rich LB media or minimal M9 media. Cells had either GFP, GFP.lva, or Trp-GFP.lva plasmids, all with the pLac promoter.

Strain	Plasmid	Phenotype on LB	Phenotype on M9
JM109	GFP.lva	Green fluorescence	No growth
MC4100	GFP.lva	Green fluorescence	Lawn not green, individual colonies fluoresce green
JM109	GFP	Ambiguous	No green fluorescence
MC4100	GFP	Green fluorescence	Green fluorescence
JM109	Trp-GFP.lva	No green fluorescence	No green fluorescence

Due to the differences in growth ability on M9 between the two *E. coli* strains, the plasmid containing the *trp* regulatory sequence was transformed into competent MC4100 cells. It is possible that the poor growth of JM109 cells in M9 hides the activity of the *trp* regulatory sequence. GFP expression might be turned on, but remain limited and invisible because of nutrient deficiencies, rather than regulatory mechanisms. Changing cell strains might address these questions as the MC4100 cells tended to grow better in M9 media.

In order to gain a more quantitative grasp of GFP expression, a flow cytometer was used to count individual cells that expressed GFP. Thus a more accurate idea of GFP expression can be developed, and even extremely low levels of expression can be detected (Table 3). These data show a decrease in GFP expression for JM109 cells carrying either the unregulated GFP or GFP.lva plasmids when they are grown in M9, compared to when they area grown in LB. This may be due to slower growth and selective pressure on minimizing GFP expression to save

resources. Conversely, MC4100 cells with these same two plasmids show some increase in GFP expression in M9 compared to in LB. This result is contrary to expectations and contrasts with JM109 behavior. The addition of the *trp* regulatory sequence has a marked impact on GFP expression in JM109 cells. Whereas only 0.10% of JM109 cells in LB expressed GFP, 16.43% of cells in M9 expressed GFP. This remarkable difference was not replicated in MC4100 cells, where there is essentially no difference in GFP expression between growth conditions.

Table 3: Number of green fluorescent cells per 10,000 counted by the flow cytometer. The most striking result is the GFP expression of the Trp-GFP.lva plasmid in LB and in M9 media, a more than one hundred-fold difference. Also noteworthy is the lack of difference for this same plasmid for the MC4100 strain, despite the otherwise strong GFP expression in MC4100.

Strain	Plasmid	Percent Green Fluorescent Cells, LB	Percent Green Fluorescent Cells, M9
JM109	GFP.lva	35.31	0.66
MC4100	GFP.lva	20.45	23.91
JM109	GFP	84.16	0.66
MC4100	GFP	10.55	94.27
JM109	Trp-GFP.lva	**0.10**	**16.43**
MC4100	Trp-GFP.lva	1.93	1.73

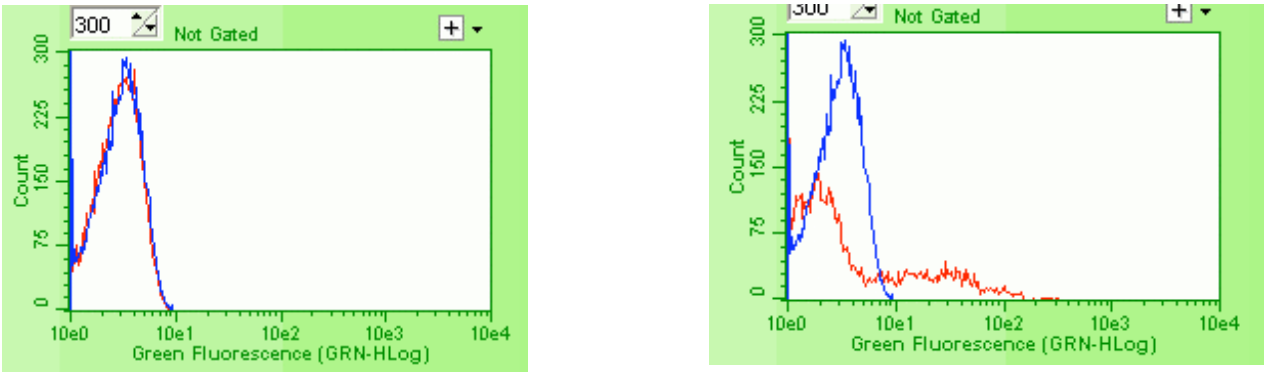


Figure 41: Flow cytometry data for JM109 cells grown in the presence (left) and absence (right) of tryptophan. In both conditions, cells without GFP plasmid do not exhibit green fluorescence (blue). Cells with GFP regulated by the *trp* operon regulatory sequence do not express GFP when there is tryptophan present, but do express GFP when there is no tryptophan (red).

Discussion

The use of bacteria as a means to solve challenging mathematical problems *in vivo* holds much promise. Due to the scaling limitations of the first bacterial system designed to solve the two-spatula burnt-pancake problem, a new system that solves larger stacks is desirable. Such a system has been designed using anti-terminators and could expand the size of pancake stacks beyond the range currently manageable by computers. This design is currently being tested using the *E. coli trp* operon anti-terminator. A *hixC* insertion point that is predicted to not

interfere with RNA secondary structures was found (Figures 12, 13, 14). This insertion has been substantiated both by software predictions and references to past experimental results.

It was not possible to assemble DNA sequences of the wild-type or modified *E. coli* tryptophan operon leader peptide and anti-termination region using oligonucleotide assembly. Due to the highly repetitive nature of the DNA sequence, and the fact that it is self-complementary, it is probable that several oligonucleotides could bind to more than one other oligonucleotide. I developed alternative PCR methods for the amplification of both the wild-tpe and modified sequences; the wild-type sequence amplification worked, but the *hixC*-insertion sequence amplification did not. It is not clear which step of the process caused the failure, but more time spent optimizing the PCR parameters might help. If this method was incapable of providing the desired *hixC* insertion, then it may be necessary to find another way of doing it. One possibility, though costly, could be to order the entire sequence synthesized. Another way of assembling this DNA sequence could be to use a new method, called circular assembly amplification, which has been shown capable of assembling highly repetitive DNA sequences (Bang *et al.*, 2008).

Despite these issues, we have learned several important lessons about several DNA sequence assembly methods. In the future, it may be necessary to enhance the oligonucleotide assembly software to check for possible alternative, undesirable binding between oligonucleotides. Software modifications may permit more intelligent oligonucleotide design to avoid assembly errors or blockage, making the process applicable to more sequences. Although PCR was an excellent way to amplify the wild-type *trp* regulatory sequence from the *E. coli* chromosome, I have not been able to modify this technique to insert the *hixC* sequence. If

enough time were spent trying to get the PCR reactions to work, this method may become a new useful tool in synthetic biology for the development of novel DNA sequences.

PCR amplification of the *trp* operon DNA from the chromosome was successful. A new BioBrick reporter part, RBS-GFP.lva-TT was produced, and the ligation of the tryptophan sequence and the pLac promoter to this reporter was completed. The RBS-GFP.lva-TT DNA sequence was cloned into both the JM109 and MC4100 *E. coli* strains. Although analysis of GFP expression profiles on agar plates proved to be mostly inconclusive (Table 2), flow cytometry has shown some promising results (Table 3, Figure 41). JM109 cells with the regular GFP.lva reporter plasmid showed GFP expression in both LB but not in M9. However, the plasmid with the added *trp* regulatory sequence greatly reduced GFP expression when grown in LB compared to growth in M9, which lacks tryptophan. These data indicate that the *trp* regulatory sequence worked as designed in regulating gene expression in response to tryptophan concentrations for cell type JM109.

In MC4100 cells, GFP was expressed in both LB and M9. The effects of the LVA degradation tag become apparent, as there is more green fluorescence exhibited by cells with the regular GFP than by cells with rapidly degrading GFP. Unlike in JM109 cells, GFP expression with the *trp* regulatory sequence was not substantially different in LB and M9 media. Perhaps MC4100 cells are more capable of synthesizing tryptophan in minimal media growth conditions, and thus do not trigger the formation of the anti-terminator conformation, thereby preventing transcription. Alternatively, the faster-growing MC4100 cells may not express as much GFP compared to the JM109 cells due to the kinetics of gene expression as dependent on growth rate (Leveau *et al.*, 2001). Cells that grow faster have more dilute protein concentrations, which in

the case of GFP expression would mean less fluorescence.

It is curious that under 1% of cells with the unregulated GFP plasmids in JM109, grown in M9, expressed GFP, but that nearly 20% of JM109 cells with the regulated GFP plasmid expressed GFP (Table 3) in M9 media. One would expect that cells with plasmids containing the *trp* regulatory sequence would have lower GFP expression than those with plasmids without the regulatory sequence, as the regulation is only known to turn off transcription, not enhance it. If this observation is replicated after more experimentation, it may be indicate that the *trp* regulatory sequence not only inhibits gene expression when tryptophan is abundant, but also enhances expression when tryptophan is lacking.

It is desirable to show more definitively that tryptophan concentration, and not other variables in the LB media, is responsible for repression. At the time of this writing, I am currently comparing GFP expression of cells grown in M9 media with different concentrations of added tryptophan; these will be analyzed using flow cytometry. Unfortunately, the first several measuring attempts have not shown any green fluorescence in either the positive controls or the experimental samples. I will continue to work on the growth conditions.

We have come closer to implementing the new bacterial pancake-problem system. The ability to regulate gene expression with the unmodified *trp* sequence means that one premise of the bacterial computer, that the anti-terminator can be used in a synthetically designed biological system, has been validated. It remains to be seen whether or not the addition of the *hixC* sequence will maintain regulation while allowing for Hin-mediated DNA inversion, and whether it is possible to design mutually incompatible anti-terminator/terminator pairs to extend the size of the pancake stack. However, the first step of this three-step process was completed and I have

hypothesized a new function for a biological control mechanism. My hypothesis demonstrates one of the advantages of synthetic biology, the power to discover new biological properties of natural systems. Synthetic biology offers two ways to learn, by building new biological devices and by uncovering new functions due to the engineering approach of the discipline.

Acknowledgments

I thank Dr. Campbell for his guidance over the course of the year. I also thank Dr. Wessner for his feedback as a second reader, and for the attentiveness and support of the attendees at the weekly lab-meetings, specifically: Drs. Haynes and Heyer, Will Deloache, Erin Zwack, and Mike Waters. I could not have performed flow cytometry assays without the help of Dr. Sarafova, for which I am grateful. I also recognize the importance of the punishment jar, for both providing entertainment and allowing us to coin new puns.

References

Bang, Duhee and George M. Church. 2008. "Gene Synthesis by Circular Assembly Amplification." *Nature Methods* **5**(1), 37-39.

Bertrand, Kevin, Laurence Korn, Frank Lee, Terry Platt, Catherine L. Squires, Craig Squires, and Charles Yanofsky. 1975. "New Features of the Regulation of the Tryptophan Operon: A new type of regulatory site has been studied." *Science* **189**, 22-26.

BioCyc. "Escherichia coli trpLEDCBA." <http://biocyc.org/ECOLI/NEW-IMAGE?type=OPERON&object=TU00067> Accessed 04/06/2008.

Das, Anathbandhu, Joseph Urbanowski, Herbert Weissbach, John Nestor, and Charles Yanofsky. 1983. "*In vitro* Synthesis of the Tryptophan Operon Leader Peptides of *Escherichia coli*, *Serratia marcescens*, and *Salmonella typhimurium*." *Proceedings of the National*

Academy of Sciences of the United States of America **80**(10), 2879-2883.

Davies, Wayne L., Livia S. Carvalho, and David M. Hunt. 2007. "SPLICE: A technique for generating *in vitro* spliced coding sequences from genomic DNA." *BioTechniques* **43**, 785-789.

Harden, Lance. "Oligo Cuts Optimization Program." <http://gcat.davidson.edu/IGEM06/oligo.html> Accessed 9/12/07.

Haynes, Karmella A., Marian L. Broderick, Adam D. Brown, Trevor L. Butner, Lance Harden, Lane Heard, Eric Jessen, Kelley Malloy, Brad Ogden, Sabriya Rosemond, Samantha Simpson, Erin Zwack, A. Malcolm Campbell, Todd Eckdahl, Laurie J. Heyer, Jeffrey L. Poet. 2007. "Computing with Living Hardware." *IET Synthetic Biology*.

Keasling, Jay D. 2008. "Synthetic Biology for Synthetic Chemistry." *ACS Chemical Biology* **3**(1), 64-76.

Landick, Robert, Jannette Carey, and Charles Yanofsky. 1985. "Translation Activates the Paused Transcription Complex and Restores Transcription of the *trp* Operon Leader Region." *Proceedings of the National Academy of Sciences of the United States of America* **82**(14), 4663-4667.

Landick, Robert, Jannette Carey, and Charles Yanofsky. 1987. "Detection of Transcription-Pausing *in vivo* in the *trp* Operon Leader Region." *Proceedings of the National Academy of Sciences of the United States of America* **84**(6), 1507-1511.

Lee, Frank, and Charles Yanofsky. 1977. "Transcription Termination at the *trp* Operon Attenuators of *Escherichia coli* and *Salmonella typhimurium*: RNA Secondary Structure and Regulation of Termination." *Proceedings of the National Academy of Sciences of the*

United States of America **74**(10), 4365-4369.

Leveau, Johan H. J., and Steven E. Lindow. 2001. "Predictive and Interpretive Simulation of Green Fluorescent Protein Expression in Reporter Bacteria." *Journal of Bacteriology* **183**(23), 6752-6762.

Nanassy, Oliver Z. and Kelly T. Hughes. 1998. "*In Vivo* Identification of Intermediate Stages of the DNA Inversion Reaction Catalyzed by the *Salmonella* Hin Recombinase." *Genetics* **149**, 1649-1663.

Nudler, E. and M. E. Gottesman. 2002. "Transcription termination and anti-termination in *E. coli*." *Genes to Cells* **7**, 755-768.

OpenWetware. "Titratable control of pBad and lac promoters in individual *E. coli* cells." <http://openwetware.org/wiki/Titratable_control_of_pBad_and_lac_promoters_in_individual_E._coli_cells> Accessed 12/20/2007.

Registry of Standard Biological Parts. "Problems with PCR using VR/VF2." <http://parts.mit.edu/registry/index.php/Problems_with_PCR_using_VR/VF2> Accessed 12/20/2007.

Roesser, James R. and Charles Yanofsky. 1991. "The effects of leader peptide sequence and length on attenuation control of the *trp* operon of *E. coli*." *Nucleic Acids Research* **19**, 795-800.

Russell, P. J. 2006. "Regulation of Gene Expression in Bacteria and Bacteriophages." In *iGenetics: A Molecular Approach* (2nd ed., pp. 528-532). San Francisco, CA: Pearson Benjamin Cummings.

Sambrook, Joseph and David W. Russell. 2001. "Molecular Cloning: A Laboratory Manual."

3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Stroynowski, Iwona, Mitzi Kuroda, and Charles Yanofsky. 1983. "Transcription Termination *in vitro* at the Tryptophan Operon Attenuator is Controlled by Secondary Structures in the Leader Transcript." *Proceedings of the National Academy of Sciences of the United States of America* **80**(8), 2206-2210.

Wendland, Meike and Dirk Bumann. 2002. "Optimization of GFP levels for analyzing *Salmonella* gene expression during an infection." *Federation of European Biochemical Societies Letters*. 521:105-108.

Winkler, Malcolm E., Kary Mullis, Jennifer Barnett, Iwona Stroynowski, and Charles Yanofsky. 1982. "Transcription Termination at the Tryptophan Operon Attenuator is Decreased *in vitro* by an Oligomer Complementary to a Segment of the Leader Transcript." *Proceedings of the National Academy of Sciences of the United States of America* **79**(7), 2181-2185.