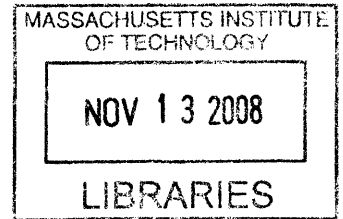# Design and Characterization of Artificial Transcriptional Terminators

by

Haiyao Huang

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

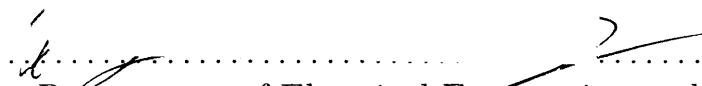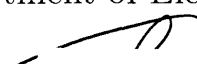Master of Engineering in Electrical Engineering and Computer Science
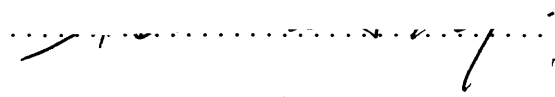
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2007

© Haiyao Huang, MMVII. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
August 21, 2007

Certified by...
Thomas F. Knight Jr.
Senior Research Scientist
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Design and Characterization of Artificial Transcriptional Terminators

by

Haiyao Huang

Abstract

Design and characterization of artificial transcriptional terminators. Ten new terminators were designed based on previous research of terminator structure and termiantion efficiency. The terminators were built by PCR extension, ligated into a BioBrick plasmid backbone, and transformed into TOP10 cells. Characterization devices were built to test the terminators. Input and output of the terminator were measured by expression of RFP and GFP. Charaterization devices were then placed into the E. coli strain CW2553/pJAT18, which hijacks the arabinose transport system to provide controlled input to the terminator. Of the ten terminators designed and tested, BBa_B1002, BBa_B1004, BBa_B1006 and BBa_B1010 proved to be strong terminators with termination efficiencies above 90%. These terminators may be obtained from the Registry of Standarized Parts at parts.mit.edu.

Thesis Supervisor: Thomas F. Knight Jr.
Title: Senior Research Scientist

.

Contents

List of Figures

9

List of Tables

Chapter 1

Introduction

Synthetic biology creates biological systems using engineering design principles. The goal of synthetic biology is to create and maintain a library of standardized and fully characterized biological parts for the construction of artificial biological systems. Currently, a library of biological parts used in synthetic biology can be found at http://parts.mit.edu. The goal of this project was to design and characterize artificial transcriptional terminators to further the advances of synthetic biology.

The registry possesses a collection of transcriptional terminators, but only have detailed information regarding the performance of a few. Of the forty plus terminators available before this project commenced, only five terminators in the collection were both available for use, and were classified as working. The most efficient and thus most commonly used terminator is B0015, which would appear a multitude of times in a large biological device. The presence of repeats of DNA segments has been known to cause unintended translocation of genetic elements, and may disrupt the carefully designed genetic machines. It would be most beneficial to design a group of terminators with high termination efficiency to lessen the likelihood of multiple repeats.

A collection of transcriptional terminators with variable termination efficiency may be used to control inputs to other genetic systems for purposes of characterizing other genetic parts. Currently, controlled inputs may be generated by hijacking transport systems already present, such as arabinose. Using transcriptional terminators to

control input instead of inherent metabolic systems allows separation of cell processes from the introduced genetic system under test.

Ten artificial transcriptional terminators were designed and characterized. The terminators were designed to achieve a range of termination efficiencies, and conformed to BioBrick standards for easy assembly with other genetic parts. The terminators were built by PCR extension, ligated into BioBrick backbones, and transformed into competent cells. Characterization devices testing the performance of the terminators utilized fluorescent proteins to measure input and output and altered the arabinose transport system to control inputs. The fluorescence produced by the characterization devices were then measured using flow cytometry to calculate the termination efficiency of the terminators.

Chapter 2

Background

2.1 Advances in Synthetic Biology

To simplify the task of engineering biological systems, the construction framework BioBricks and its associated abstraction hierarchy were developed. The abstraction hierarchy allows engineers working on one abstraction level to obscure everything in the abstraction levels below it. The BioBrick construction framework allows simple, easily repeatable assembly methods for creation of genetic machines [1].

2.1.1 The abstraction barrier and its use

The abstraction hierarchy defines abstraction levels and the interactions allowed between those levels [1]. The abstraction hierarchy and an ideal exchange of information between levels is shown in Figure 2-1. Complexity at each level is reduced because information not relevant to that level is obscured. For example, an engineer designing genetic devices would need to know how the devices are used in systems, and what parts are needed to construct the device, but would not need to know about DNA synthesis.
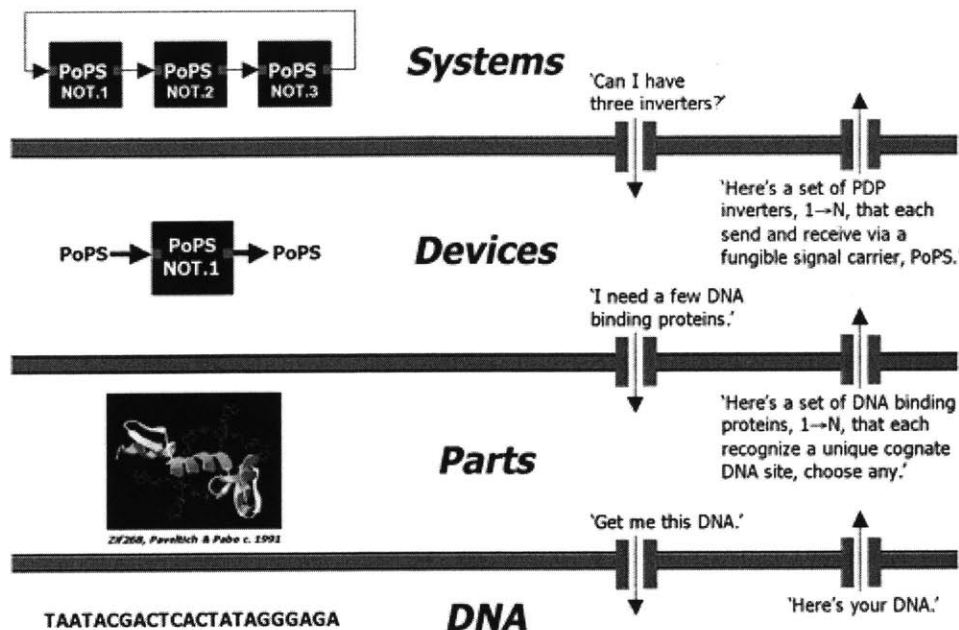
Figure 2-1: Abstraction barriers, shown in red, control the access of information across abstraction levels. A sample exchange is shown in green. This image was taken from http://parts.mit.edu.

2.1.2 The advance of BioBricks

The goal of BioBricks is to standardize the form of genetic components to allow idempotent reactions where the key structural elements of a component are unchanged by the reactions. The outputs of such reactions can be used as the starting point in subsequent reactions. BioBricks provide a standard method of assembling genetic components using specified prefixes and suffixes.

BioBrick prefixes and suffixes

Each BioBrick part contains the component of interest flanked upstream by EcoRI and XbaI restriction sites and downstream by SpeI and PstI restriction sites [2]. The component should not contain any of the restriction sites. When creating a new BioBrick part, PCR primers containing the BioBrick prefix and suffix are used to turn the component in question into a BioBrick. The resulting PCR product can then be cut with EcoRI and PstI, and ligated into the plasmid of choice. The primers
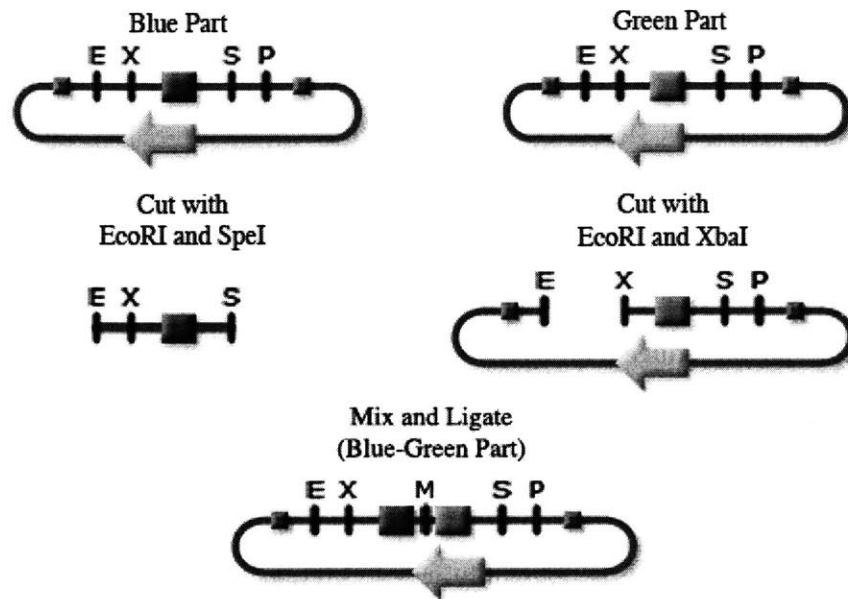
Figure 2-2: Two BioBrick parts, blue and green, are combined to form one part. This image was taken from http://parts.mit.edu.

contain extra bases beyond the restriction sites to allowed restriction enzymes to bind to the EcoRI and PstI sites.

BioBrick Prefix:

GAATTCGCGGCCGCTTCTAG for parts that start with *AGT*...

GAATTCGCGGCCGCTTCTAGAG for all other parts

BioBrick Suffix Primer:

TACTAGTAGCGGCCGCTGCAG for all parts

Assembly Methods

To combine two BioBrick parts A and B, part A is cut with EcoRI and SpeI while part B is cut with EcoRI and XbaI. The insert cut from part A and the insert from part B are ligated together to a new backbone. The SpeI and XbaI cut sites have comparable overhangs, and can be ligated together to create a mixed site that is not recognized by either restriction enzyme and cannot be cut. The resulting vector will still be flanked by the appropriate restriction sites, but will contain the parts A and

B. This construction process is shown in Figure 2-2.

There are two ways to create a construct containing N BioBrick parts. The first involves first joining two parts and then adding the subsequent parts in order. This process will take a total $N - 1$ constructions, and weeks of time that might not be able to be spared. The second method involves parallel assembly, shown in Figure x. By performing multiple pairwise constructions in parallel, the number of constructions can be reduced to $log_2(N)$ from N. If error occurs in one construction in parallel assembly, the failed part is ignored, and further construction continues with successful constructs. However, such errors in the standard assembly would require an additional stage to compensate for the failure. For these reasons, it is more convenient to use parallel assembly when making large constructs.

2.1.3   Evolution of the genetic inverter

The common signal used by BioBrick for gene expression is PoPS (polymerase per second) instead of a relying on a chemical signal. The PoPS level is set by the number of RNA polymerase molecules that move across a particular section of DNA [1]. Having a common signal means that any PoPS based device may be connected to any other PoPS device. An example of a PoPS based device is a genetic inverter, which takes a high input and returns a low output.

A PoPS based inverter fixes the main problem of a classic inverter: using proteins as signals. A classic genetic inverter receives as input the concentration of protein A, and through gene regulation, sends as output the concentration of the repressor B. The problem with a classic inverter is that any device upstream of it must output a concentration of protein A, and any device downstream from it must take as input repressor B. Therefore, inversion of signal using a classic inverter requires two proteins. A PoPS based inverter both takes PoPS as input and outputs PoPS and requires only one protein, the repressor, for an inverter. Unlike the classic inverter, the specific molecular interactions in a PoPS inverter are internal and can be hidden to reduce complexity [1]. A comparison of a classic inverter and a PoPS based inverter is shown in Figure 2-3.

24
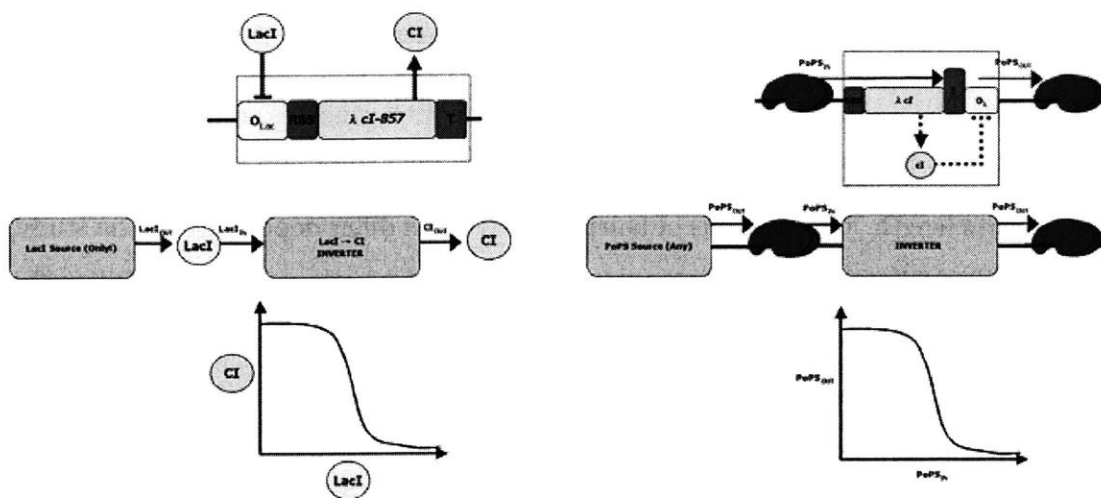
Figure 2-3: On left is a classic inverter, which uses LacI and cI as signals. A high concentration of LacI inhibits the expression of the downstream cI gene while a low concentration of LacI allows more transcription and expression of cI. On the right is a PoPS based inverter, which takes PoPS as both inputs and outputs. When the input is high, cI is produced and acts to the keep the output PoPS low. This image was taken from http://parts.mit.edu.

2.2 Rho-independent Transcriptional Terminators

Transcriptional termination in prokaryotes is a complex process that involves RNA polymerase and possible other protein factors. Terminators that can function in vitro with only the DNA substrate and RNA polymerase are called intrinsic or rho-independent terminators. This section will discuss the structure of rho-independent terminators, the effects of structure on termination efficiency, and early attempts to characterize BioBrick terminators. Unlike rho-dependent terminators, the characteristics of rho-independent terminators are well understood, making it possible to design new rho-independent terminators.

The majority of transcriptional terminators studied in the d'Aubenton Carafa [3] paper have a G+C rich stem of 7(+/-1)bp and a loop of 4(+/-1) nucleotides followed by a poly(T) tail. The most common loop size found was 4nt, with 55% of the terminators studied having a loop of that size. Of the tetra-nucleotide loops found, the most commonly repeated sequences were TTCG and GAAA, both of which are known to increase RNA hairpin stability.

25

2.2.1 Finding rho-independent terminators in prokaryote genomes

To identify rho-independent terminators in prokaryote genomes, one must find the sequences that have high likelihood for hairpin formation that also have a T-tail of an appropriate length. Calculations of hairpin predictions differ depending on the study, but usually involve minimizing free energy of the stem loop structure. To calculate the likelihood that a T-tail is present, several recent papers [4, 5] use a modified version of the algorithm presented in the d'Aubenton Carafa paper. The algorithm favors thymine residues closer to the stem loop, and penalizes presence of other nucleotides in a 15 nt sequence. The original algorithm is as follows.

The parameter n_T evaluates the importance of T residues going from 5' to 3'. To calculate n_T, a number x_n is calculated for the nucleotide at position n as follows:

$$x_n = \begin{cases} 0.9 * x_{n-1} & \text{if the } n\text{th nucleotide is a T} \\ 0.6 * x_{n-1} & \text{if the } n\text{th nucleotide is not a T} \end{cases}$$

The value of x_0, the first T residue is set to be 0.9.

To calculate n_T:

$$n_T = \sum x_n \text{ for all T residues in 15 residue segment}$$

In the original paper, the authors considered a tail score of 2.895 to be the minimum score for a real terminator.

The modified algorithm used by Ermolaeva in 2000 and Kingsford in 2007 calculates the tail score as follows to have a low score represent a T-rich tail.

$$n_T = -\sum_{n=1}^{15} x_n$$

where

$$x_n = \begin{cases} 0.9 * x_{n-1} & \text{if the } n\text{th nucleotide is a T} \\ 0.6 * x_{n-1} & \text{if the } n\text{th nucleotide is not a T} \end{cases}$$

for $n = 1 \ldots 15$ and $x_0 = 1$.

26

2.2.2 Effects of structure on termination efficiency

The termination efficiency of a terminators depends on its particular stem loop structure and the length of the poly(T) tail. In general, disruption of stem loop structure lowers termination efficiency more than disruption of the poly(T) tail [6]. Destroying the G+C dyad symmetry of the stem by either creating mismatches or by replacing all G+C pairs with A+T pairs will greatly reduce the termination efficiency by up to 90% [6]. Decreasing the free energy of the stem does not guarantee an increase in termination efficiency [7, 8]. In some cases [7], replacing the loop of a terminator with the sequence TTCG will stabilize the RNA hairpin, and slightly increase termination efficiency.

The effects of disruption of the poly(T) tail is much more straight forward. The termination efficiency increases linearly with the number of thymine residues present up to around 7 residues. The addition of more residues does not further increase the termination efficiency [9].

2.2.3 Predicted termination efficiency given structure

The structure of a terminator can be used to predict its efficiency using a formula developed by d'Aubenton-Carafa. The formula was used to calculate a factor d, which described the likelihood of a given sequence being a terminator, but could also be used to predict termination efficiency.

$$d = n_T * 18.16 + Y * 96.59 - 116.87$$

where n_T is the tail score calculated as follows:

$$n_T = \sum x_n \text{ for all T residues in 15 residue segment}$$

$$x_n = \begin{cases} 0.9 & \text{if } n = 1 \\ 0.9 * x_{n-1} & \text{if the } n\text{th nucleotide is a T} \\ 0.6 * x_{n-1} & \text{if the } n\text{th nucleotide is not a T} \end{cases}$$

and

$$Y = \frac{-\Delta G}{L_H}$$

where $-\Delta G$ is the free energy of hairpin formation and L_H is the number of nucleotides in the entire stem loop

In general, the higher the d score, the higher the termination efficiency. Shown in Figure 2-4 is a figure taken from the d'Aubenton-Carafa paper that relates the d score of a variety of terminators to their measured termination efficiencies.
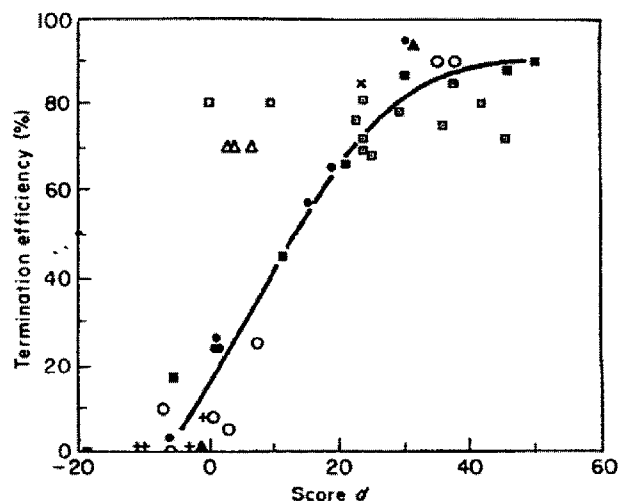
Figure 7. Diagram showing the correlation between the score *d* of some rho-independent terminators and their efficiency *in vitro*. We name the terminators by the name of the preceding gene or operon: (□) *rrnB T1* (Schmidt & Chamberlin, 1987), bacteriophage T7 *Te* (Chamberlin *et al.*, 1979); (▲) *ampL* attenuator and *ampL35A* mutant (Jaurin *et al.*, 1981); (○) *infC* (Sacerdot *et al.*, 1982; ref. (13) in Table 2A), *pheS* attenuator (Fayat *et al.*, 1983; ref. (65) in Table 2A), *his* attenuator (Frunzio *et al.*, 1981; ref. (68) in Table 2A), *trpt* and *trpC301* and *trpC302* mutants (Christie *et al.*, 1981), bacteriophage T3 *Te* (Neff & Chamberlin, 1980); (△) *tonB* (both directions) (Postle & Good, 1985; ref. (122) in Table 2A), *rplT* (Fayat *et al.*, 1983; ref. (65) in Table 2A); (●) *trp* attenuator, *trp a1419* and *trp a135* mutants (Christie *et al.*, 1981), *trpL77*, *trpL78*, *trpL80*, *trpL153* mutants (Stauffer *et al.*, 1978); (■) *thr* attenuator and *T2, T3, T4, T5, T6, T8* mutants in the poly(U) stretch (Lynn *et al.*, 1988); (▣) *thr* attenuator stem mutants *L135U, L138U, L139A, L140A, L151A, L151U, L153A, L153U, L153+G, L153−G, L156U* (Lynn *et al.*, 1988); (×) *rnpB* (Sakamoto *et al.*, 1983); (+) intracistronic signals in *cca* (this work).

Figure 2-4: This image, taken from the d'Aubenton-Carafa paper, illustrates the correlation between the calulated d score and the measured efficiency of a terminator. Termination efficiency increases linearly as the d score increases, and plateaus after the d score is higher that 40.

29

Chapter 3

Design and Construction of Artificial Transcriptional Terminators

This chapter describes the design and construction of ten artificial transcriptional terminations with a theoretical range of termination efficiencies from 10% to 90%. These ten terminators are in the registry of standard parts as BBa_B1001 to BBa_B1010.

3.1 Design

The ten terminators are split into two series based on their stem-loop sequence. Series A contains terminators BBa_B1001 through BBa_B1005, while series B contains BBa_B1006 through BBa_1010. Each series has five terminators with varying thymines in their poly(T) tails. As previous studies show it is easier to predict termination efficiency by truncating the poly(T) tail, I have chosen that method to vary the termination efficiency of these terminators. The stem loop and tail sequences of these terminators are shown in Table 3.1.

Table 3.1: This table shows the structure and sequence of the designed terminators.

Part Name	Stem-Loop Sequnce	Loop	Tail
BBa_B1001	CCCCGCTTCGGCGGGG	TTCG	TTTTTTTTT
BBa_B1002	CCCCGCTTCGGCGGGG	TTCG	TTTTTT
BBa_B1003	CCCCGCTTCGGCGGGG	TTCG	TTTTT
BBa_B1004	CCCCGCTTCGGCGGGG	TTCG	TTTT
BBa_B1005	CCCCGCTTCGGCGGGG	TTCG	TTT
BBa_B1006	CCCCGCCCCTGACAGGGCGGGG	CTGACA	TTTTTTTTT
BBa_B1007	CCCCGCCCCTGACAGGGCGGGG	CTGACA	TTTTTT
BBa_B1008	CCCCGCCCCTGACAGGGCGGGG	CTGACA	TTTTT
BBa_B1009	CCCCGCCCCTGACAGGGCGGGG	CTGACA	TTTT
BBa_B1010	CCCCGCCCCTGACAGGGCGGGG	CTGACA	TTT

3.1.1 Hairpin

All terminators in Series A have a stem of 6nt and a loop of 4nt. The stem seqence, GGGGCG, is the consensus sequence found for terminators with 4nt loops in the d'Aubenton-Carafa paper. The loop, TTCG, is one of two loop sequences known to stabilize the mRNA hairpin. Series B terminators are a variation of the thr terminator. All A-T pairs in the stem of the thr terminator were replaced by G-C pairs, and the first G-C pair was changed into a C-G pair. Removal of the A-T pairs should increase the stability of the stem-loop structure. Series B terminators have a stem of 8bp, and a loop of 6nt.

3.1.2 Length of tail

The lengths of the poly(T) tail in each series goes from a minimum of 3nt to a maximum of 9nt. In theory, additional T residues beyond 6nt would not further increase termination efficiency. Additionally, termination efficiency increases linearly as the poly(T) tail increases from 3nt to 6nt. Only one terminator in each series has a poly(T) tail greater than 6nt at 9nt. The other terminators have poly(T) tails ranging from 3nt to 6nt.

3.1.3 Predicted efficiency

The paper by d'Aubenton-Carafa suggests that the same algorithm used to predict the presence of a terminator can also be used to estimate termination efficiency of a given terminator. The termination efficiency can be predicted as a function of the total length of the terminator, the free energy of the stem loop structure, and number of thymine residues in the stretch of 15 nt after the stem loop:

$$d = n_T * 18.16 + Y * 96.59 - 116.87$$

where n_T is the tail score calculated as follows:

$$n_T = \sum x_n \text{ for all T residues in 15 residue segment}$$

$$x_n = \begin{cases} 0.9 & \text{if } n = 1 \\ 0.9 * x_{n-1} & \text{if the } n\text{th nucleotide is a T} \\ 0.6 * x_{n-1} & \text{if the } n\text{th nucleotide is not a T} \end{cases}$$

and

$$Y = \frac{-\Delta G}{L_H}$$

where $-\Delta G$ is the free energy of hairpin formation and L_H is the number of nucleotides in the entire stem loop structure.

Python scripts were written to calculate both d and n_T. The energy of hairpin formation for a given sequence was calculated using UNAFold [10]. A summary of the d scores and predicted efficiencies of terminators BBa_B1001 through BBa_B1010 is shown in Table 3.2.

Table 3.2: This table shows the calculated d scores and the predicted termination efficiency for BioBrick terminators B1001-B1010.

Part Name	$-\Delta G$	Length	Tail Score	d score	Predicted Efficiency
BBa_B1001	-12.6	16	5.68	62.33	95%
BBa_B1002	-12.6	16	4.35	38.25	90%
BBa_B1003	-12.6	16	3.78	27.78	80%
BBa_B1004	-12.6	16	3.16	16.05	55%
BBa_B1005	-12.6	16	2.48	4.22	25%
BBa_B1006	-16.2	22	5.68	57.39	95%
BBa_B1007	-16.2	22	4.35	33.32	80%
BBa_B1008	-16.2	22	3.78	22.84	70%
BBa_B1009	-16.2	22	3.16	11.56	40%
BBa_B1010	-16.2	22	2.48	0.72	10%

3.2 Construction

Five rounds of construction were needed to create these terminators. This count does not include rounds of construction that yielded no successful construct. The first round yielded BBa_B1004. BBa_1007, BBa_1005, and BBa_1001 were completed in rounds 2, 3, and 4 respectively. All other terminators were completed in round 5.

The terminators were made by overlapping primers and extending them by PCR. The PCR products were then purified, and cut with EcoRI and PstI. The BioBrick plasmid backbone was also made by PCR and cut with the same restriction enzymes. The insert and backbone were the ligated together and transformed into TOP10 cells.

3.2.1 Primer design

The process of making BioBrick parts from the designed sequences begins with creating PCR primers to turn the sequence from text on paper to a stretch of DNA . The forward and reverse primers overlap from 2nt before the loop to 2nt after the loop. This creates a 8bp overlap for series A terminators and a 10bp overlap for series B terminators. Ideally, the forward and reverse primers would only overlap at the loop, but a 4bp or 6bp overlap is not sufficient for binding. Extending the overlap 2nt

34

Table 3.3: This tables show the forward primers used to synthesize BioBrick terminators B1001-B1010.

Part Name	Forward Primer
BBa_B1001	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGAAAAAAAAACCCCGCTTCGGC
BBa_B1002	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCAAAAAACCCCGCTTCGGC
BBa_B1003	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCCAAAAACCCCGCTTCGGC
BBa_B1004	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGGCCGAAAACCCCGCTTCGGC
BBa_B1005	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCCGCAAACCCCGCTTCGGC
BBa_B1006	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGAAAAAAAAACCCCGCCCTGACAGG
BBa_B1007	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCAAAAAACCCCGCCCTGACAGG
BBa_B1008	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCCAAAAACCCCGCCCTGACAGG
BBa_B1009	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCCGAAAACCCCGCCCTGACAGG
BBa_B1010	GTTTCTTCGAATTCGCGGCCGCTTCTAGAGCGCCGCAAACCCCGCCCTGACAGG

Table 3.4: This tables shows the reverse primers used to synthesize BioBrick terminators B1001-B1010.

Part Name	Reverse Primer
BBa_B1001	GTTTCTTCCTGCAGCGGCCGCTACTAGTAAAAAAAAACCCCGCCGAAGC
BBa_B1002	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGAAAAAACCCCGCCGAAGC
BBa_B1003	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGGAAAAACCCCGCCGAAGC
BBa_B1004	GTTTCTTCCTGCAGCGGCCGCTACTAGTACGGCAAAACCCCGCCGAAGC
BBa_B1005	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGGCGAAACCCCGCCGAAGC
BBa_B1006	GTTTCTTCCTGCAGCGGCCGCTACTAGTAAAAAAAAACCCCGCCCTGTCAGGG
BBa_B1007	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGAAAAAACCCCGCCCTGTCAGGG
BBa_B1008	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGGAAAAACCCCGCCCTGTCAGGG
BBa_B1009	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGGCAAAACCCCGCCCTGTCAGGG
BBa_B1010	GTTTCTTCCTGCAGCGGCCGCTACTAGTAGCGGCGAAACCCCGCCCTGTCAGGG

each way beyond the loop should provide adequate strength for binding, but should not cause the primers to form a hairpin with themselves. Both forward and reverse primers have extra bases beyond the EcoRI and PstI cut sites to ensure the restriction enzyme can bind to the site effectively. The forward and reverse primers used to construct terminators BBa_B1001 through BBa_B1010 are shown in Tables 3.3 and 3.4.

3ul of each of the primers were added to 40ul of Fidelity PCR supermix. The primers would overlap and form the template, so no additional template DNA was required. The samples were placed in the thermocycler for 20 extension cycles. PCR products were then run on Metaphor 3.5% agarose gels to check for length. It took some experimentation to determine what percentage gel to use to separate the product from primers because the DNA fragments were only twice the size of the primers, and typical 1-1.5% agarose gels don't have the definition to separate 100bp products from 50bp primers.

3.2.2 DNA purification

The PCR product must be purified to remove all remaining primers and unused nucleotides. Purification was complicated by the relatively short length of the PCR product (90bp). Two methods of PCR used were gel purification and a modified version of the Qiagen PCR purification protocol.

Gel purification

The first attempts to purify the PCR were done with the Qiagen gel purification kit. The PCR products were run on Metaphor 3.5% gels, and the appropriate bands were cut and placed into 2ml tubes. The tubes were weighed, and QX1 buffer equal to 6 times the weight of the gel was added to each tube. Ten microliters of QIAEX II beads were added, and the tubes were incubated for 10 minutes at 50 degrees C. The tubes were vortexed every two minutes to keep the QIAEX II beads in suspension. After incubation, the tubes were centrifuged for 30 seconds, and the supernatant

removed. The remaining pellets were then washed with 500 ul of QXI buffer once, and 500 ul of PE twice. The pellets were then left to air dry for 15 minutes until they turned white. To elute the DNA from the QIAEX II beads, 50ul of Qiagen elution buffer was added to each tube, and the pellets were resuspended. After 5 minutes of incubation on the bench, the tubes were centrifuged, and the supernatant containing the DNA stored at -20 C.

For reasons yet unknown, not all uses of the kit yielded any DNA but this mistake was not noticed until the fourth round of construction. An alternative method was found for purifying PCR products in the fourth round of terminator construction.

Alternate protocol for PCR purification

This protocol used Qiagen solutions PB, PE, and EB as well as Qiagen miniprep columns. The sample and 5 times as much PB were applied to the miniprep column, and the column was spun for 1 minute at 3000G. The flowthrough was then applied to the column, and this procedure was repeated twice. After the third spin, the flowthrough was discarded, and 750 ul of PE was added to the column. The column was then spun for 1 minute at 3000G; the flow through was discarded, and the column spun again for 3 minutes at 17900G to remove all residual PE, which might hamper later reactions. After spinning, the column was transferred to a new tube, and 30 ul of EB was added. The column was incubated on the bench at room temperature for 5 minutes, and then spun for one minute at 6800G. Another 30 ul of EB was added, and the column was incubated for another 5 minutes. The column was spun for a final time for 5 minutes at 6800G to recover the PCR products.

3.2.3   Insertion into BioBrick backbone

After the PCR product is purified, it must be ligated into an approved BioBrick backbone to form a BioBrick part. BioBrick backbones carry one of kanamycin , chloramphenicol, or tetracycline resistance markers in addition to ampicillin resistance. Backbones can be created by miniprepping cultures carrying the backbone

plasmid or by building the backbone using a template and PCR primers. Both the backbone and the insert are cut with appropriate restriction enzymes, and ligated together.

Creation of backbone

I chose to use the backbone with the kanamycin resistance marker for the terminator BioBrick. I used PCR to build the backbone as this method generally yielded higher concentrations of backbone DNA than miniprepping cultures with the appropriate plasmid. The backbone was then purified using the protocol described earlier.

Digestion and ligation

Both the backbone and the insert were cut with EcoRI and PstI. Each digest was set up with 20 ul DNA, 1 ul of each of the enzymes, 5 ul of NEB buffer 2 at 10x concentration, 0.5 ul of BSA at 100x concentration and 22.5 ul of distilled water. These digests were then incubated at 37 degrees C for 2 hours for the enzymes to cut the DNA, and then heated to 80 degrees C to heat kill the enzymes. Cut DNA should be stored in TE buffer at -20 degrees C to reduce chances of cut ends degrading.

Once both inserts and backbones have been cut with the correct enzymes, they should be ligated to form a BioBrick part. A rough ratio of 3:1 insert to backbone should be present in the ligation reactions. Adding too much insert causes the inserts to ligate to each other, forming final products with repeats of three to five inserts in one backbone. It is also possible for the backbones to ligate to each other, but these constructs never form viable colonies, and thus is not a problem. The formula to calculate the amount of insert to add to a reaction is given by the following formula:

$$\text{ng insert} = 3 * frac{\text{ng vector} * \text{kb insert}}{\text{kb vector}}$$

The ligation mixtures contained 1ul 10x T4 DNA ligase buffer, 1 ul plasmid backbone, the amount of insert needed as calculated by the formula above, 0.3 ul T4 DNA ligase, and enough water to make a final volume of 10ul. The water is added first, and

the ligase added last. The final mixture should have equimolar amounts of plasmid backbone and insert DNA. The ligations are incubated at room temperature for 15 minutes, and then used immediately to transform competent cells.

3.2.4 Transformation into TOP10 competent cells

The finished ligation products are transformed into TOP10 competent cells, and the cultures with the correct BioBrick part are store for future use. The competent cells are thawed on ice, and 10 ul of cells are aliquoted for each transformation. The cells are then diluted with 40ul CMB800 to increase transformation efficiency. 1 ul of the appropriate ligation product is added to each aliquot, and the cells are kept on ice for thirty minutes. The cells are then heat shocked for 50 seconds at 42 degrees C, and put on ice again for two minutes. SOC media is added to the cultures, and the cells are incubated at 37 degrees C to recover antibiotic resistance before being plated out and grown overnight at 37 degrees C. Generally, one hour of incubation is enough to recover resistance to kanamycin or ampicillin, but two hours is preferable when using tetracycline or chloramphenicol as resistance markers.

At first, I added 1 ml of SOC to each culture, and plated out two plates per culture, one with 200ul and the other with 20ul. I found that these plates rarely yielded enough colonies, particularly the plates spread with 20ul. I switched to adding 250ul of SOC to each culture, and plating out the entire culture. This approach tended to yield at least one colony per transformation.

3.2.5 Plasmid recovery, verification, and storage

DNA was recovered from the transformants by minipreps, and was then verified for the presence of the correct BioBrick part. If the part was correct, then the cells would be prepared for storage.

Minipreps to recover plasmid DNA

The following day, one colony was picked from each plate of transformants and grown up for minipreps in LB broth. I used the Qiagen miniprep kit to perform these minipreps. First, 1.7 ml of culture were spun down to obtain a pellet of cells. The pellet was resuspended in 250 ul of P1 buffer, and vortexed for 30 seconds to lyse the cells. 250 ul of P2 buffer was then added, and the tube was gently inverted 6 to 10 times. The tube was not vortexed at this time as it would result in shearing of the genomic DNA. After the solution has turned an even shade of blue, 350 ul of N3 buffer was added to stop the reaction. Once the solution turned colorless and cloudy after inverting the tube 6-10 times, the tube was centrifuged at 17900G for 10 minutes.

The supernatant was then transferred to a miniprep spin column, and the column spun for 1 minute. The flowthrough was discarded, and 500ul of PB buffer was applied to the column. The column was spun again for 1 minute, and 750 ul of PE was added. The column was then spun twice to remove all PE, which could hamper future reactions. Afterwards, the column was transferred to a clean tube. 50 ul of EB was applied, and the column was incubated on the bench at room temperature for 5 minutes. The tube was then spun for 3 minutes at 6800G to recover the plasmid DNA. The DNA obtained from these minipreps is then used to verify the existence of the correct BioBrick part.

Verification of BioBrick part

The plasmid DNA recovered from the minipreps were first cut with one restriction enzyme and run on a 1.5% agarose gel to check for its presence. The concentration of the plasmid DNA was then measured with the NanoDrop spectrometer and recorded. PCR primers VF2 and VR were then used to measure the length of the insert using the plasmid DNA as the template. VF2 binds to the template before the BioBrick prefix, and VR binds to the template after the BioBrick suffix. The resulting PCR product was then run on an 1% agarose gel to measure its length. If the length was

correct, the plasmid DNA was then sent out for sequencing at the MIT Sequencing Center. Each sample contained 200ng plasmid DNA, 0.3 ul of either VF2 or VR, and enough water to make a final volume of 12 ul.

Storage of cultures

If the culture contains the correct BioBrick part, it will be prepared to be stored at -80 degrees C for future use. A solution of 85% cell culture and 15% glycerol is vortexed for 30 seconds to mix it well, and then incubated on the bench at room temperature for 30 minutes. The solution is then stored in the -80C freezer until it is needed. From past experience, incubation for periods longer than 30 minutes does not seem harm the cells. I use an 80% glycerol solution to make these frozen cultures since it is easier to accurately pipette than 100% glycerol.

3.2.6 Errors and Troubleshooting

A number of problems occurred during the construction phase including but not limited to: setting the thermocycler to the wrong temperatures for PCR extension when creating the parts, not being able to recover DNA with the gel extraction kit, and degraded cut ends on the backbone which caused the backbone to ligate shut. Perhaps the most perplexing instance was that of transformants which grew on antibiotic plates but lacked the plasmid and insert when miniprepped. In addition, when streaked out from prepared glycerols, the transformants again grew on antibiotic plates, but would not grow in liquid media. After trying in vain to recover plasmid DNA from these transformants for two weeks, these transformants were scrapped, and a new set of construction started.

A common problem faced with these constructions was the insertion of genomic DNA into the plasmid during the transformations. This was noticed when the Bio-Brick site was amplified with PCR and run on gels. For these terminators, the approximate size should be around 350 bp. When genomic DNA was inserted instead of the BioBrick part, the size of the insert ranged from 350bp to 750bp. If the length

was around 300, it meant that nothing had been inserted, and the backbone was ligated to itself.

Chapter 4

Design and Construction of Characterization Devices

The next goal after creating the terminators was to develop devices that would allow characterization of the new BioBrick terminators. The characterization devices for the artificial terminators use a GFP/RFP dual fluorescent system with the inputs to the system controlled by the promoter. The characterization devices were constructed using 3A assembly, described in 4.2.1, either performed by hand or by robot. The finished constructs were then transformed into E. coli strain CW2553.

4.1   Design of terminator characterization devices

The input to the characterization devices is controlled by the P_{araBAD} system. The terminator to be characterized is flanked by two fluorescent proteins, GFP and RFP,
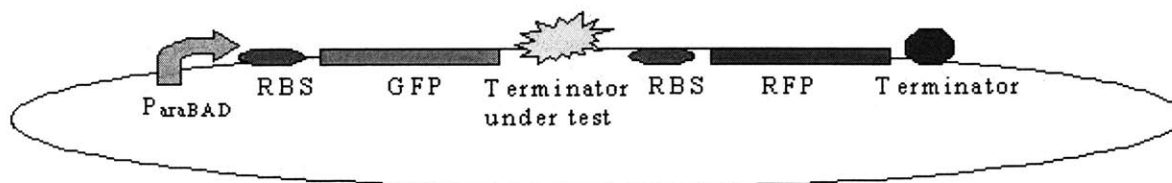
Figure 4-1: Characterization plasmid version 1: controlled by P_{araBAD}, with inputs measured by GFP expression and outputs measured by RFP expression.
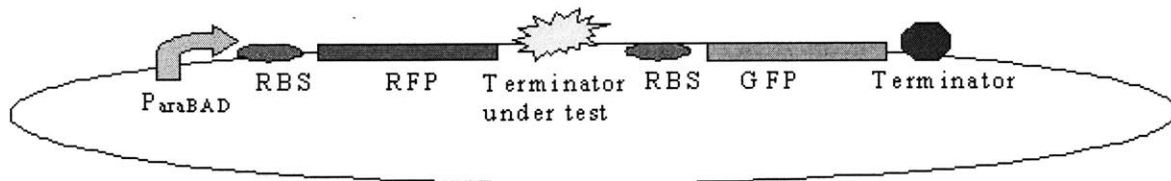
Figure 4-2: Characterization plasmid version 2: controlled by P_{araBAD}, with inputs measured by RFP expression and outputs measured by GFP expression.

Table 4.1: A list of the BioBrick parts needed to construct the terminator characterization plasmids and a short description of the function of those parts. Data for these parts were found on the Standard Registry at http://parts.mit.edu.

Part Name	Part Type	Description
BBa_I0500	promoter	inducible ParaBAD
BBa_E0034	RBS	strong RBS
BBa_E0040	reporter	generates GFP
BBa_E1010	reporter	generates RFP
BBa_B0015	terminator	terminator with high TE

which are used to measure the termination efficiency of the terminator. The characterization devices are made entirely from BioBrick parts found in the registry. Parts used to construct these devices are shown in Table 4.1. Figure 4-1 shows the version of the characterization device using GFP expression to measure input and RFP expression to measure output. Figure 4-2 shows a characterization device using the opposite measuring scheme.

Some of the parts used to construct the characterization devices are available in composite parts. Using the composite parts instead of the individual parts speed up
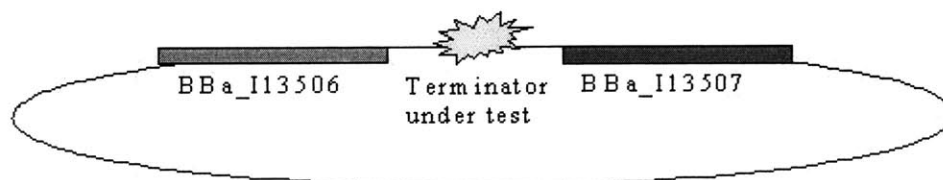
Figure 4-3: Characterzation plasmid version 1 constructed using composite BioBrick parts aviable in the Standard Registry.

44
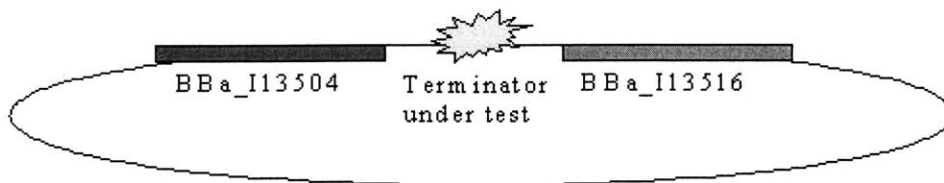
Figure 4-4: Characterzation plasmid version 2 constructed using composite BioBrick parts aviable in the Standard Registry.

Table 4.2: This table shows composite BioBrick parts availible from the registry (http://parts.mit.edu). These composite parts were used in the construction of the terminator characterization devices.

Part Name	Components
BBa_I13506	I0500, E0034, E0040
BBa_I13507	E0034, E1010, B0015
BBa_I13516	I0500, E0034, E1010
BBa_I13504	E0034, E0040, B0015

the construction process by reducing the number of assemblies needed. The composite parts used to construct the characterization devices are described in Table 4.2. The final constructs are shown in Figures 4-3 and 4-4.

4.1.1 PoPS input generator

The PoPS generator must be able to vary the input signal to produce a wide range of outputs for device characterization. One possible way to produce a wide range of PoPS inputs is to use an inducible promoter such as the arabinose promoter, P_{araBAD}. However, inducing the araBAD operon at subsaturation concentrations results a population of cells which exhibit linear behavior in response to chances in inducer concentration but individual cells with either be fully induced or not induced. Decoupling the arabinose transport gene araE from the P_{araBAD} promoter and putting it under the control of an arabinose independent promoter will remove the all-or-none effects and produce a population of cells that will exhibit linear behavior in P_{araBAD} expression as a function of arabinose concentration at an individual level with all cells in the population having a similar level of expression as shown in the paper by

45
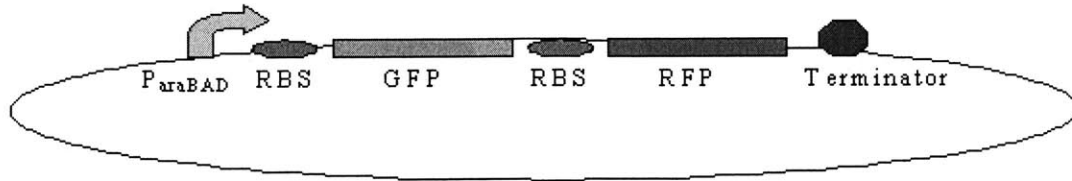
Figure 4-5: I13514: calibrates GFP input to RPF output

Khlebnikov et al [11].

4.1.2 Device under test

The new BioBrick terminators will be the devices under test for these measurement devices. A list of terminators tested is shown earlier in Table 3.1.

4.1.3 Dual Fluorescent system

The DUT is flanked by two fluorescent proteins, GFP and RFP. The fluorescent protein preceding the DUT measure inputs to the DUT while the fluorescent protein following the DUT measures the output.

The termination efficiency will by measured by the ratio of the first fluorescent protein produced to the second protein produced. If the terminator has a high termination efficiency, very little fo the second protein will be produced. If the terminator has low termination efficiency, there should be no difference in the levels of the first and second proteins. In the off chance that one of the terminators acts as a promoter, more the second protein will be produced than the first protein.

Two sets of the characterization devices were constructed. The first set has GFP flanking on the left of the DUT and RFP flanking on the right. The second set is reversed with RFP on the left, and GFP on the right. This allows calibration of the input and output measurements.

46

Table 4.3: This table shows the function and component parts of the control plasmids. All controls are availible from parts.mit.edu.

Part Name	Description	Componenets
BBa_I13514	Calibration of GFP input to RFP output	I0500, E0034, E0040, E0034, E1010, B0015
BBa_I13515	Calibration of RFP input to GFP output	I0500, E0034, E1010, E0034, E0040, B0015
BBa_I13521	Maximum RFP output	R0040, E0034, E0040, B0015
BBa_I13522	Maximum GFP output	R0040, E0034, E1010, B0015
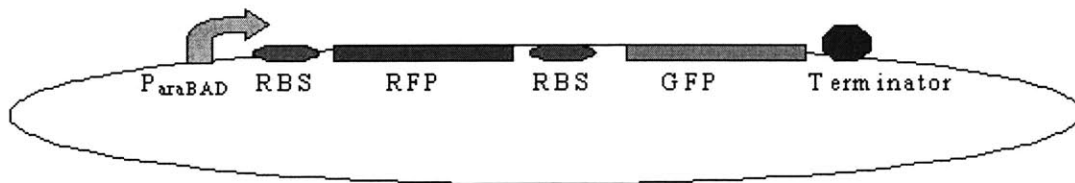
Figure 4-6: I13515: calibrates RFP input to GFP output.
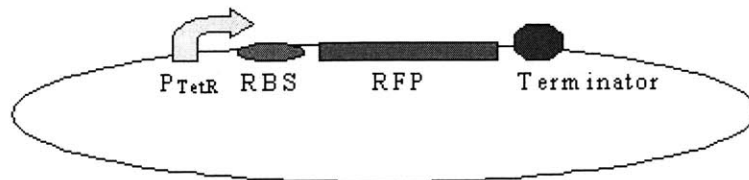
Figure 4-7: I13521: measures maximum RFP expression
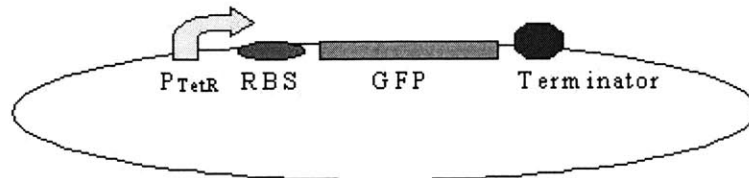
Figure 4-8: I13522: measures maximum GFP expression.

4.1.4 Controls

Controls were needed to calibrate the levels of GFP and RFP expression in the characterization constructs. A list of controls used is shown in Table 4.3, and all controls are available from the Standard Registry. The controls I13514 and I13515 have the same components as the characterization devices, but lack a DUT. These are used to calibrate the input and output between the two sets of characterization devices by showing normal GFP and RFP expression without interference from the DUT, and are shown in Figures 4-5 and 4-6.

Constructs I13521 and I13522 each have a fluorescent protein, RFP and GFP respectively, under the control of a constitutive promoter, TetR. These controls test the maximum levels of GFP and RFP expression and are shown in Figures 4-7 and 4-8.

4.2 Construction of Characterization Devices

The characterization devices were constructed using the triple antibiotic rolling assembly method performed either by hand or by robot. Finished constructs were sent out for sequencing to verify the existence of the correct construct. The complete list

Table 4.4: This table shows the parts used to constuct the terminator characterization devices.

Part Name	Left Part	DUT	Right Part
BBa_B3101	BBa_I13506	BBa_B1001	BBa_I13507
BBa_B3102	BBa_I13506	BBa_B1002	BBa_I13507
BBa_B3103	BBa_I13506	BBa_B1003	BBa_I13507
BBa_B3104	BBa_I13506	BBa_B1004	BBa_I13507
BBa_B3105	BBa_I13506	BBa_B1005	BBa_I13507
BBa_B3106	BBa_I13506	BBa_B1006	BBa_I13507
BBa_B3107	BBa_I13506	BBa_B1007	BBa_I13507
BBa_B3108	BBa_I13506	BBa_B1008	BBa_I13507
BBa_B3109	BBa_I13506	BBa_B1009	BBa_I13507
BBa_B3110	BBa_I13506	BBa_B1010	BBa_I13507
BBa_B3201	BBa_I13516	BBa_B1001	BBa_I13504
BBa_B3202	BBa_I13516	BBa_B1002	BBa_I13504
BBa_B3203	BBa_I13516	BBa_B1003	BBa_I13504
BBa_B3204	BBa_I13516	BBa_B1004	BBa_I13504
BBa_B3205	BBa_I13516	BBa_B1005	BBa_I13504
BBa_B3206	BBa_I13516	BBa_B1006	BBa_I13504
BBa_B3207	BBa_I13516	BBa_B1007	BBa_I13504
BBa_B3208	BBa_I13516	BBa_B1008	BBa_I13504
BBa_B3209	BBa_I13516	BBa_B1009	BBa_I13504
BBa_B3210	BBa_I13516	BBa_B1010	BBa_I13504

of all constructions is shown in Table 4.4.

The characterization devices were made in two steps. The first set of constructions and the resulting intermediate parts is shown in 4.5. Terminators were combined with either BBa_I13507 or BBa_I13504 to form the intermediate parts BBa_B11XX and BBa_B12XX. The intermediate parts were then combined with either BBa_I13506 or BBa_I13516, with the details of these constructions shown in Table 4.6.

The constructions shown in Table 4.5 were assembled using triple antibiotic assembly performed by hand. Constructions shown in Table 4.6 were assembled by the same methods using the robot. All constructions that failed in the first round were subsequently assembled by hand.

Table 4.5: This table shows the first construction step and the intermediate parts created in making the terminator characterzation devices.

Part Name	Left Part	Right Part	Description
BBa_B1101	BBa_B1001	BBa_I13507	B1001 + rfp
BBa_B1102	BBa_B1002	BBa_I13507	B1002 + rfp
BBa_B1103	BBa_B1003	BBa_I13507	B1003 + rfp
BBa_B1104	BBa_B1004	BBa_I13507	B1004 + rfp
BBa_B1105	BBa_B1005	BBa_I13507	B1005 + rfp
BBa_B1106	BBa_B1006	BBa_I13507	B1006 + rfp
BBa_B1107	BBa_B1007	BBa_I13507	B1007 + rfp
BBa_B1108	BBa_B1008	BBa_I13507	B1008 + rfp
BBa_B1109	BBa_B1009	BBa_I13507	B1009 + rfp
BBa_B1110	BBa_B1010	BBa_I13507	B1010 + rfp
BBa_B1201	BBa_B1001	BBa_I13504	B1001 + gfp
BBa_B1202	BBa_B1002	BBa_I13504	B1002 + gfp
BBa_B1203	BBa_B1003	BBa_I13504	B1003 + gfp
BBa_B1204	BBa_B1004	BBa_I13504	B1004 + gfp
BBa_B1205	BBa_B1005	BBa_I13504	B1005 + gfp
BBa_B1206	BBa_B1006	BBa_I13504	B1006 + gfp
BBa_B1207	BBa_B1007	BBa_I13504	B1007 + gfp
BBa_B1208	BBa_B1008	BBa_I13504	B1008 + gfp
BBa_B1209	BBa_B1009	BBa_I13504	B1009 + gfp
BBa_B1210	BBa_B1010	BBa_I13504	B1010 + gfp

Table 4.6: This table shows the second construction used to create the terminator characterization devices. The intermediate part from the first construction is used as the right part in this construction.

Part Name	Left Part	Right Part	Description
BBa_B3101	BBa_I13506	BBa_B1101	gfp + B1001 + rfp
BBa_B3102	BBa_I13506	BBa_B1102	gfp + B1002 + rfp
BBa_B3103	BBa_I13506	BBa_B1103	gfp + B1003 + rfp
BBa_B3104	BBa_I13506	BBa_B1104	gfp + B1004 + rfp
BBa_B3105	BBa_I13506	BBa_B1105	gfp + B1005 + rfp
BBa_B3106	BBa_I13506	BBa_B1106	gfp + B1006 + rfp
BBa_B3107	BBa_I13506	BBa_B1107	gfp + B1007 + rfp
BBa_B3108	BBa_I13506	BBa_B1108	gfp + B1008 + rfp
BBa_B3109	BBa_I13506	BBa_B1109	gfp + B1009 + rfp
BBa_B3110	BBa_I13506	BBa_B1110	gfp + B1010 + rfp
BBa_B3201	BBa_I13516	BBa_B1201	rfp + B1001 + gfp
BBa_B3202	BBa_I13516	BBa_B1202	rfp + B1002 + gfp
BBa_B3203	BBa_I13516	BBa_B1203	rfp + B1003 + gfp
BBa_B3204	BBa_I13516	BBa_B1204	rfp + B1004 + gfp
BBa_B3205	BBa_I13516	BBa_B1205	rfp + B1005 + gfp
BBa_B3206	BBa_I13516	BBa_B1206	rfp + B1006 + gfp
BBa_B3207	BBa_I13516	BBa_B1207	rfp + B1007 + gfp
BBa_B3208	BBa_I13516	BBa_B1208	rfp + B1008 + gfp
BBa_B3209	BBa_I13516	BBa_B1209	rfp + B1009 + gfp
BBa_B3210	BBa_I13516	BBa_B1210	rfp + B1010 + gfp
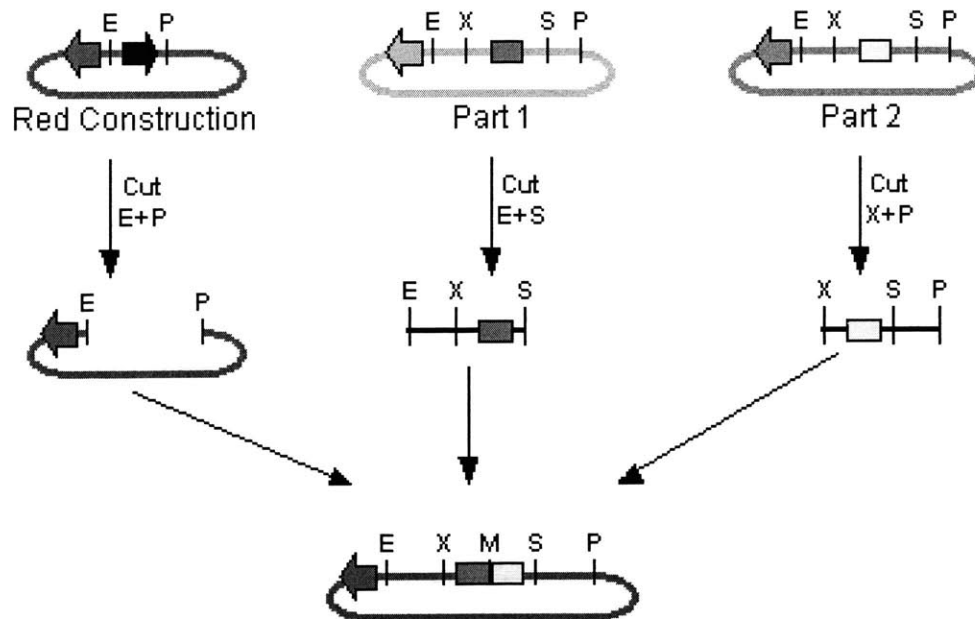
Figure 4-9: Triple antibiotic assembly is the method by which two BioBrick parts are combined into a new part with a differnt antibiotic marker which can then be used in futher constuctions. The image was taken from http://openwetware.mit.edu

4.2.1 Triple antibiotic assembly assembly

Triple antibiotic assembly is the method by which two BioBrick parts are combined to form a new BioBrick. The process in shown in Figure 4-9. Each BioBrick has a backbone which contains one of three different antibiotic markers: kanamycin, tetracycline, or chloramphenicol. Having a choice of three markers ensures that a new BioBrick part formed from two old parts will have a different resistance marker than its components and allows for selection of that new part.

The first step in triple antibiotic assembly is selecting the correct backbone for the assembly. The backbone is then cut with EcoRI and PstI. The two parts to be combined are also cut: the first part with EcoRI and SpeI, the second with XbaI and PstI. When combined, the SpeI and XbaI sites will form a mixed site which cannot be cut with either enzyme. The digested backbone and inserts are then combined and ligated with T4 ligase

In the ideal case, the backbone will combine with the two cut inserts to form the

new BioBrick part. However, conditions in the ligation mixture are far from ideal. Ideally, the only DNA present in the ligation mixture will be the backbone and the inserts. Realistically, additional cut fragments such as backbones from the cut inserts and sticky end fragments from the cut backbone are all present and will interfere with the formation of the BioBrick part. A construct with multiple copies of one or both inserts or constructs with no inserts may be possible. Constructs containing more than one backbone ligated together are not viable, and do not have to be taken into consideration. The resulting constructs are then transformed into Top10 cells.

4.2.2 Robotic assembly

Assemblies to make the constructs shown in 4.6 were performed on the epMotion5075 by Meagan Lizarazo. Assemblies that failed were then done by hand.

4.2.3 Verification

Colonies of constructs were picked onto an index plate, miniprepped, and sent out for sequencing. The constructs from the robotics assemblies arrived in plasmid form, were transformed into TOP10 cells, and then miniprepped as there was not enough initial plasmid DNA for sequencing. The miniprepped DNA was then sent off for sequencing.

The presence of full BioBrick prefix and suffix sequences in the sequencing results demonstrated successful construction of a part. Common failure modes included the inclusion of genomic DNA or the lack of an insert.

4.3 Transforming into an ideal strain

TOP10 cells are useful as competent cells for creating initial cell stocks but make a poor experimental strain as it has the wild type arabinose operon. For the P_{araBAD} promoter system to be fully functional, the characterization device must be in a strain which has the arabinose transport genes knocked out. One such strain is CW2553

developed by the Keating lab [11].

4.3.1   E. coli strain CW2553

The E. coli strain CW2553 (araE201, .araFGH::kan) has all arabinose transport genes either deleted or mutated. The araE gene must be under the control of an arabinose independent promoter to ensure homogenous induction of cells. In the study by Khlebnikov et al. [11], putting araE under the control of the P_{CP18} promoter resulted in cells being homogeneously induced by arabinose in the media as well as producing the highest concentrations of fluorescence. The pJAT18 plasmid contains araE under the control of P_{CP18}, and is included in the CW2553 strain obtained from the Endy lab. The pJAT18 plasmid uses gentamicin as a resistance marker, so all transforms will be grown on media containing gentamicin.

4.3.2   Making competent cells

A streaked plate containing CW2553 cells with pJAT18 was obtained from the Endy lab. A colony was picked off the plate and grown overnight in 5ml of LB media. In the morning, the culture was diluted into 500ml of LB, and grown until it reached an OD_{600} of 0.4. The culture was split into two 225ml falcon tubes and incubated on ice for 10 minutes.

The falcon tubes were centrifuged for 10 minutes at 3000rpm and 4 degrees C. The supernatant was then removed and discarded, and the cell pellets were resuspended in TSS buffer. TSS volume equal to 10% of the culture volume was used resuspend the cells TSS buffer was prepared by adding 5g PEG 8000, 1.5 ml 1M $MgCl_2$, and 2.5 ml of DMSO to LB media to a final volume of 50ml. The TSS buffer was then filter sterilized and chilled to 4 degrees C. The resulting cell solution was aliquoted into 1ml eppendorf tubes and stored at -80 degrees C.

4.3.3 Transformation of characterization devices into CW2553

Both the characterization devices and controls were transformed into CW2553/pJAT18. One ul of miniprepped DNA from a characterization device or control was added to 50ul of competent cells, and incubated on ice for 30 minutes. The cells were then heat shocked at 42 degrees C for 50 seconds, and put back on ice for two minutes. 250 ul of SOC media was added to each tube of cells, and the cells were incubated at 37 degrees C for two hours to recover antibiotic resistance. After incubation, the cells were spread on plates containing both gentamicin and ampicillin and grown overnight. A colony was picked off each plate and made into a glycerol for future use.

Chapter 5

Using Characterization Devices to Measure Termination Efficiency

The terminators were characterized by measuring the inputs and outputs of the characterization devices using a protocol developed by Jason Kelly of the Endy lab [12]. The characterization devices were grown in supplemented M9 media overnight, and then induced with arabinose. The following day, aliquots of the cultures were taken to the MIT Flow Lab, and the induced fluorescence was measured.

5.1   Materials

Setting up experimental cultures required preparation of a rich media such as M9, and a solution of 1% arabinose for induction. Cultures containing the characterization devices were streaked on plates and grown overnight to check for viability. One colony of each strain was used to start a new experimental culture. Five controls were used to gauge the accuracy and validity of the experimental cultures.

5.1.1 Media

M9 supplemented Media

Initially, 2L of 1x supplemented M9 media were made. The requirements from making a litre of supplemented M9 is listed below. This recipe was found on openwetware.

- 200 ml of M9 minimal salts at 5X concentration

- 34 ml of thiamine at 10mg/ml

- 10 ml of 40% glycerol

- 20 ml of 10% Casamino acids

- 2 ml of 1M $MgSO_4$

- 100 ul of 1M $CaCl_2$

- 733.9 ml of sterile deionized H_2O

The following instructions were followed to make 2L of supplemented M9 media To dissolve the M9 minimal salts, 22.6 grams of Bacto M9 minimal salts, in a concentration of 5X from Difco, was dissolved in 400 ml of H_2O. 780 mg of thiamine was dissolved in 78 ml of H_2O, and filter sterilized. An equal amount of H_2O was added to 20ml of 80% glycerol to dilute it to 40% glycerol. 20 g of Bacto Casamino acids from Difco was dissolved in 200 mL of H_2O to create a 10% solution of casamino acids. To make a 1M $MgSO_4$ solution, 24.65g of $MgSO_4 * 7H_2O$ was dissolved in 100ml of H_2O. Likewise, 14.7g of $CaCl_2 * 2H_2O$ was dissolved in 100 ml of H_2O to create a 1M solution. All except for the thiamine solution were autoclaved at 121 degrees C for 15 minutes to sterilize. The thiamine solution was filter sterilized. The ingredients were added to 733.9 ml of sterile deionized H_2O following the recipe listed above. The media was then split into 1 1L bottle and 2 500 ml bottles, and appropriate antibiotics added to each bottle. Both gentamicin and ampicillin were added to 1.5 L of media, and the remaining 500 ml had only gentamicin added to it.

Arabinose

An initial stock solution of 1% arabinose by weight was made by adding 1 g of L-arabinose to 100ml of H_2O and filter sterilized. The stock solution was then further diluted to create a 0.1% arabinose solution used to induce the characterization devices in the experimental cultures.

5.1.2 Characterization Devices

Cultures containing the characterization devices shown in 4.4 were streaked on plates to check for viability. A single colony from each plate was then used to start an experimental culture.

5.1.3 Controls

The experimental controls used to validate the experimental cultures were previously shown in 4.3. Controls I13514 and I13515 were used to calibrate the inputs and outputs; while controls I13521 and I13522 provided the maximum expression levels of RFP and GFP. The negative control for this experiment was a culture of CW2553/pJAT18 containing no BioBrick plasmids. Care had to be taken not to accidently grow the negative control in media unsuitable for it.

5.2 Protocols

Each characterization device and control were first streaked out on plates. One colony from each plate was grown in 5ml M9 with the appropriate antibiotic was grown for 24 hours at 37 degrees C. The OD600 of each culture was then measured and recorded. Each culture was diluted to an OD600 of approximately 0.07, which contains around 10^4 CFU.

The diluted cultures were then grown from 2 hours at 37 degrees C, and fluorescent protein expression was induced with arabinose. Studies have shown that the best range for arabinose induction is between 0.0001% and 0.01% (Khlebinokov). All

samples were induced with 5ul of 0.1% arabinose in 5ml of culture, creating a final arabinose concentration of 0.001%.

The induced samples were then grown overnight for 12-14 hours to maximize fluorescent protein expression. The following morning, 1ml aliquots of each cultures were placed in falcon 3026 polypropylene tubes on ice to stop further growth. The aliquots were taken the MIT Flow lab, and GFP and RFP expression were measured.

Chapter 6

Results

6.1 Controls

Five controls were measured in this experiment. CW2553/pJAT18 was used as the negative control to determine the ranges of background fluorescence. I13521 and I13522 constitutively expressed RFP and GFP respectively. I13514 and I13515 measured the GFP and RFP expression of the two versions of the characterization plasmids, but lacked the internal terminator under test. The mean GFP and RFP fluorescence and the standard deviations of these controls are shown in Table 6.1.

6.1.1 Constitutive expression of RFP and GFP

I13521 and I13522 provide a baseline measurement of reasonable ranges of RFP and GFP fluorescence. Figure 6-1 shows the measured fluorescence of these controls. The sample I13521 has a mixture of fluorescent and nonfluorescent cells. For purposes of calculating the mean and standard deviation of the cell population in I13521, only cells expressing sufficient fluorescence, defined as being above 4std of the negative control, were included.

As expected, I13521 has negligible GFP expression and I13522 has negligible RFP expression. The mean RFP of I13521 was 83.92, compared to the negative control of 2.17. The mean GFP of I13522 was 15.12, compared to a negative control of 2.86. It

Table 6.1: This table shows the average GFP and RFP expression of the negative control, I13514, I13515, I13521 and I13522. I13521 and I13522 constitutively express RFP and GFP respectively. I13514 and I13515 are used to calibrate input and output measurements of the characterization devices. In cases of I13514 and I13521, which have two distinct populations of cells, the cells which do not express sufficient fluorescence are discounted from the mean.

Sample	Description	Mean GFP	Std GFP	Mean RFP	Std RFP
CW2553	no fluorescence	2.86	1.32	2.18	0.81
I13521	constitutive RFP only	3.81	1.96	83.92	77.25
I13522	constitutive GFP only	15.12	13.07	2.17	0.60
I13514	inducible GFP/RFP	20.73	23.05	17.2	14.65
I13515	inducible RFP/GFP	210.30	102.48	2.22	0.62

Figure 6-1: This figure shows the measured GFP and RFP of controls I13521 and I13522 as compared to the negative control CW2553/pJAT18. Controls I13521 and I13522 respectively express RFP and GFP constitutively. As expected, I13521 has negligible GFP expression and I13522 has negligible RFP expression. The sample I13521 contained a population of cells that produced neither GFP nor RFP, and those cells were ignored when calculating the mean RFP expression. The mean RFP of I13521 was 83.92, compared to the negative control of 2.17. The mean GFP of I13522 was 15.12, compared to a negative control of 2.86. It is not know why constitutive GFP expression was much lower than constitutive RFP expression.
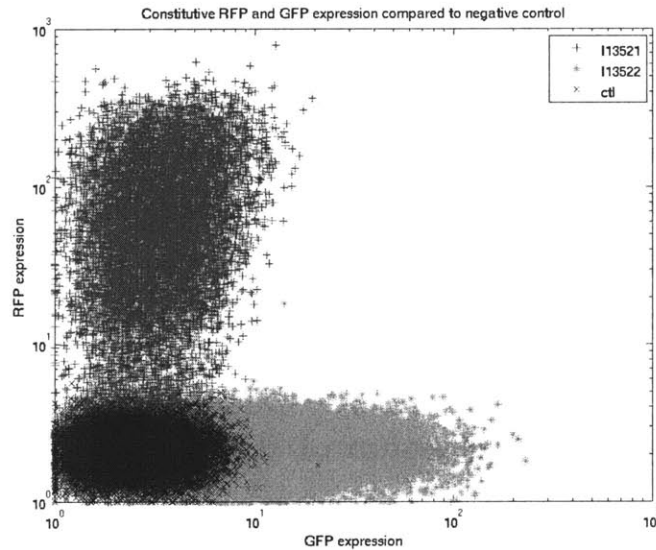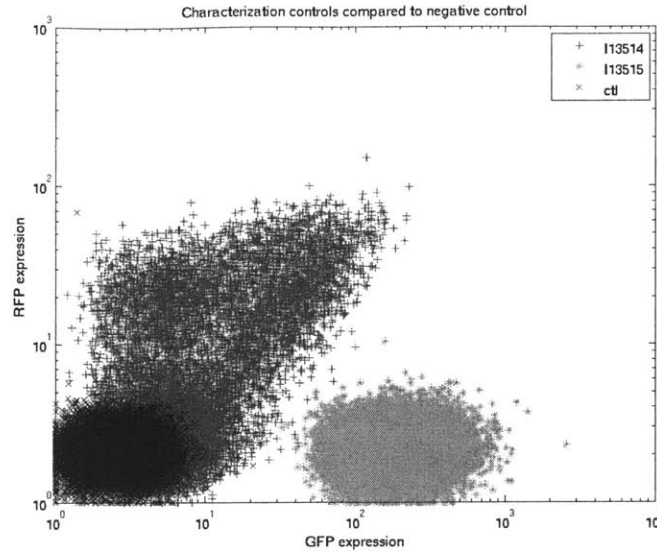
Figure 6-2: This figure shows the measured GFP and RFP of controls I13514 and I13515 as compared to the negative control CW2553/pJAT18. Ideally, I13514 and I13515 should have the same levels of GFP and RFP. The majority of cells with the plasmid I13514 produced no significant amounts of GFP or RFP. Of the cells producing significant fluorescence, the mean GFP expression was 20.73, and the mean RFP expression was 17.2. As the majority of cells produced neither GFP nor RFP, I13514 cannot be used to accurately calibrate the ratio of input to output of a terminator under test in version 1 of the characterization plasmid. Due the possible presence of an RNAse cut site in the RFP coding region, the control I13515 produced negligible RFP. The mean GFP expression for I13515 was 210.3.

is not know why constitutive GFP expression was much lower than constitutive RFP expression.

6.1.2 Expression of RFP and GFP from empty characterization plasmids

I13514 and I13515 are the empty versions of the characterization plasmids, lacking the terminator under test. I13514 has GFP followed by RFP under the control of the arabinose promoter P_{araBAD}. I13515 is similar, having RFP followed by GFP under the control of the same promoter. Figure 6-2 shows the measured fluorescence of these controls.

63

I13514

The cell population of I13514 contains a mixture of nonfluorescent cells, cells only expressing RFP, and cells expressing both RFP and GFP. The majority of the cells express no significant levels of fluorescence. Of the cells producing significant fluorescence, the mean GFP expression was 20.73, and the mean RFP expression was 17.2. As such, the measurements taken from I13514 cannot be accurately used to calibrate the input (in terms of GFP) to the output (in terms of RFP).

I13515

The cell population of I13515 does not express significant levels of RFP, but expresses high levels of GFP, with a mean GFP expression of 210.30. This may be due to the possible existence of an internal RNAse site in the RFP coding region, which causes fast degradation of RFP mRNA. In effect, the characterization plasmid would only have GFP to measure output of the terminator, instead of RFP to measure input and GFP to measure output.

6.2 Terminators

Ten terminators were characterized using two versions of the characterization plasmid. Version 1, shown previously in Figure 4-1, contained GFP, followed by the terminator under test and RFP under the control of the arabinose promoter P_{araBAD}. Version 2 switched the locations of the RFP and GFP coding regions, but was otherwise the same, and was shown previously in Figure 4-2. Under ideal circumstances, if a strong terminator was placed into characterization plasmid 1, the only output should be GFP. Likewise, if a strong terminator was present in characterization plasmid 2, only RFP should be visible as the output.

Figure 6-3: This figure shows the GFP vs RFP expression of each cell in the sample of B3101 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3101 is 5.71 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3101, 2.17, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3101 produces both low levels of GFP and RFP, contrary to the initial belief that it would produce levels of GFP close to that of I13514, but low RFP. The terminator tested, B1001, is unlikely to be a strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.

Figure 6-4: This figure shows the GFP vs RFP expression of each cell in the sample of B3102 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3101 is 22.42 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3102, 2.23, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3102 produces a similar amount of GFP compared to the control I13514, but negligible RFP, and is likely to be a strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.
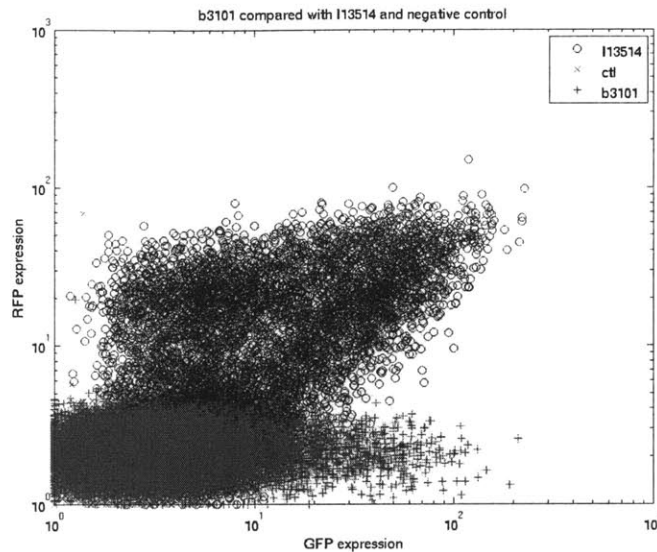
Figure 6-5: This figure shows the GFP vs RFP expression of each cell in the sample of B3103 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3103 is 11.18 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3103, 2.17, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3103 produces only 54% of GFP of the control I13514, but negligible RFP, and is unlikely to be strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.
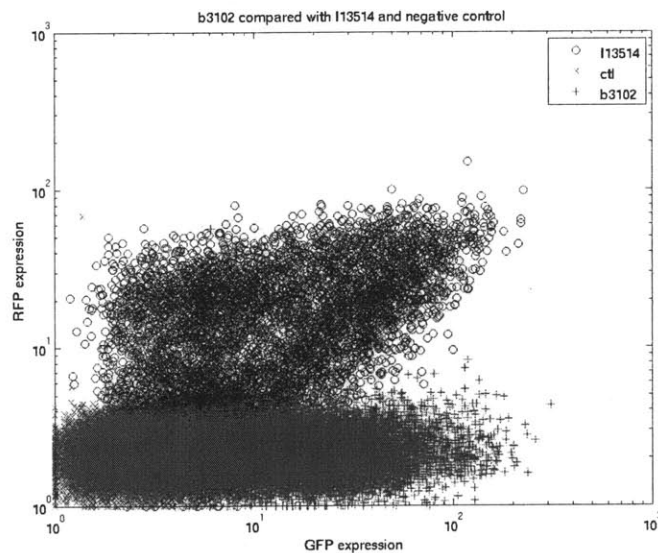
Figure 6-6: This figure shows the GFP vs RFP expression of each cell in the sample of B3104 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3101 is 22.47 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3104, 2.16, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3104 produces a similar amount of GFP compared to the control I13514, but negligible RFP, and is likely to be a strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.
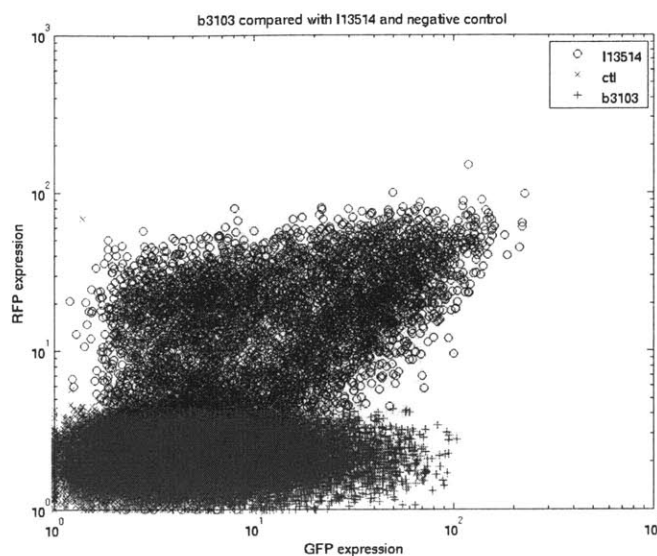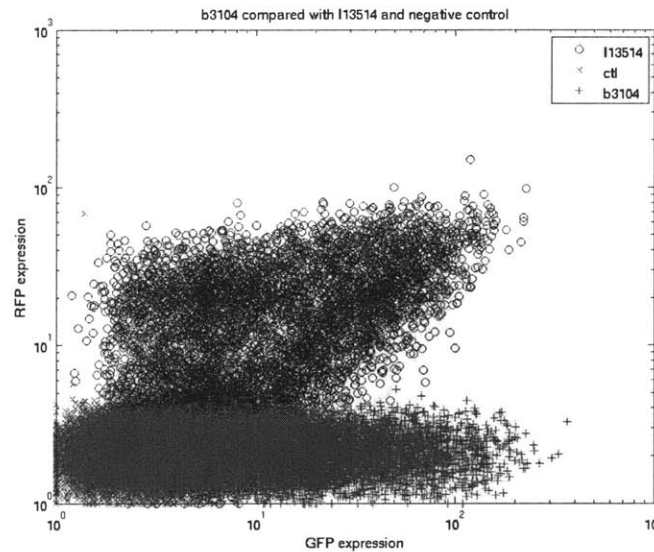
Figure 6-7: This figure shows the GFP vs RFP expression of each cell in the sample of B3105 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3105 is 3.69 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3101, 2.19, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3105 produces both low levels of GFP and RFP, contrary to the inital belief that it would produce levels of GFP close to that of I13514, but low RFP. The terminator tested, B1005, is unlikely to be a strong terminator as it affects input as well as output.

Figure 6-8: This figure shows the GFP vs RFP expression of each cell in the sample of B3106 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3106 is 24.86 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3106, 2.16, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3106 produces a similar amount of GFP compared to the control I13514, but negligible RFP, and is likely to be a strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.
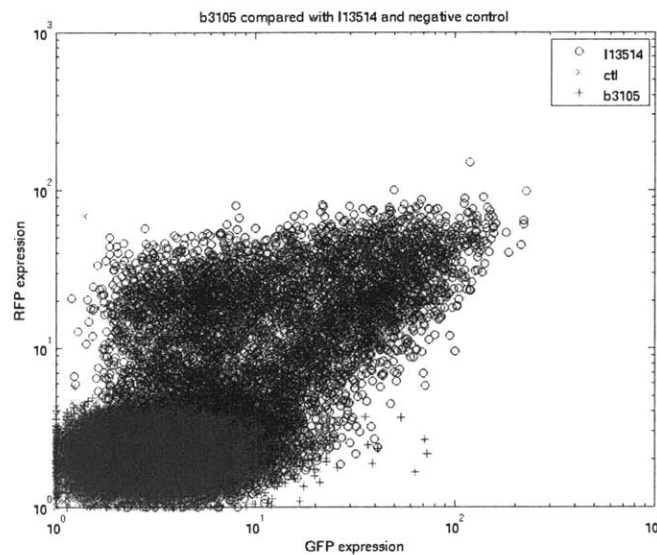
Figure 6-9: This figure shows the GFP vs RFP expression of each cell in the sample of B3107 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3107 is 14.34 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3107, 2.18, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3107 produces only 69% of GFP of the control I13514, but negligible RFP, and is unlikely to be strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.
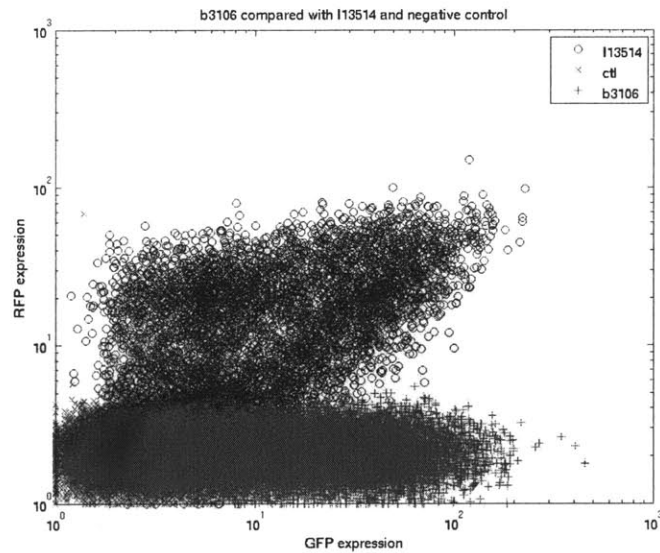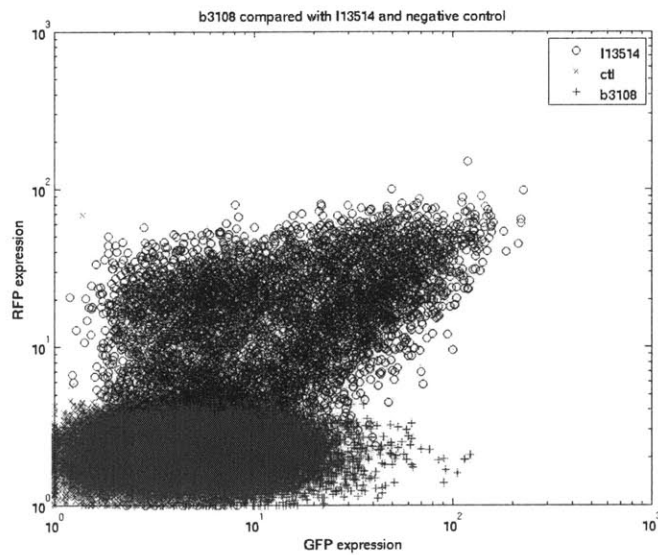
Figure 6-10: This figure shows the GFP vs RFP expression of each cell in the sample of B3108 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3108 is 8.59 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3108, 2.18, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3108 produces only 41% of GFP of the control I13514, but negligible RFP, and is unlikely to be strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.
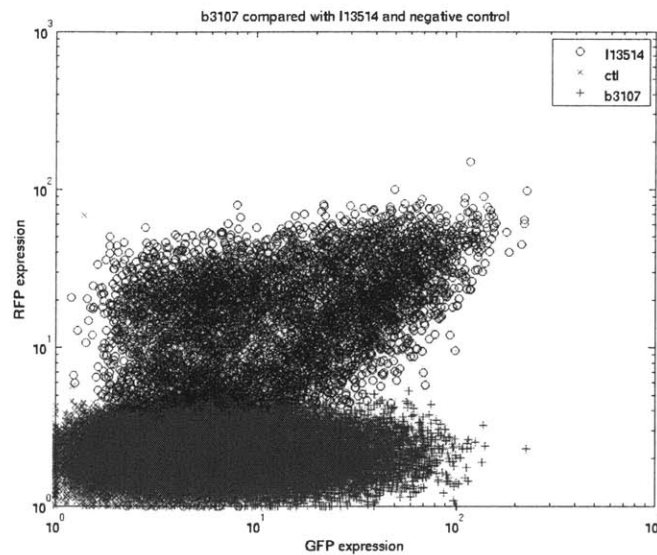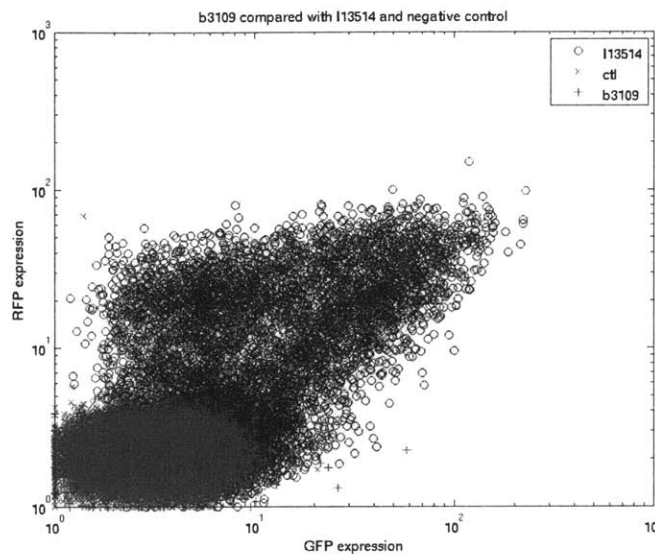
Figure 6-11: This figure shows the GFP vs RFP expression of each cell in the sample of B3109 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3101 is 3.57 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3109, 2.18, is the same as the negative control, and significantly less than the mean RFP of I13514, 17.2. B3109 produces both low levels of GFP and RFP, contrary to the initial belief that it would produce levels of GFP close to that of I13514, but low RFP. The terminator tested, B1009, is unlikely to be a strong terminator as it affects input as well as output.
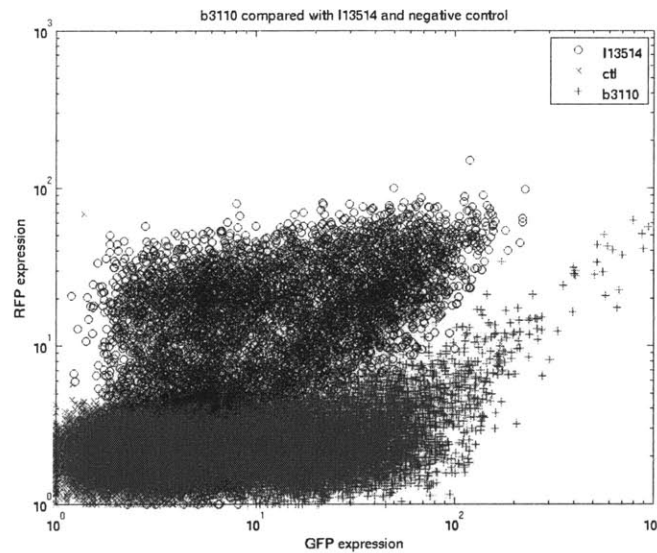
Figure 6-12: This figure shows the GFP vs RFP expression of each cell in the sample of B3110 compared to the negative control CW2553 and the empty characterization plasmid I13514. The mean GFP expression for B3110 is 25.48 compared to 2.86 for the negative control and 20.73 for I13514. The mean RFP for B3110, 2.60, is the close the negative control, and significantly less than the mean RFP of I13514, 17.2. B3110 produces a similar amount of GFP compared to the control I13514, but negligible RFP, and is likely to be a strong terminator. The actual termination efficiency cannot be determined as the control I13514 has a majority of cells that produce no significant fluorescence, and cannot be used to accurately calibrate input measurements to output measurements.

Table 6.2: This table shows the mean GFP and RFP expression of characterization plasmids B3101 through B3110. The mean GFP and RFP expression of the negative control CW2553/pJAT18 and I13514 are shown for comparison.

Sample	Mean GFP	std GFP	Mean RFP	std RFP
CW2553/pJAT18	2.86	1.32	2.18	0.81
I13514	20.73	23.05	17.2	14.65
B3101	5.71	9.65	2.17	0.62
B3102	22.42	19.39	2.23	0.78
B3103	11.18	8.92	2.17	0.61
B3104	22.47	26.24	2.16	0.59
B3105	3.69	2.52	2.19	0.0.62
B3106	24.86	23.88	2.16	0.60
B3107	14.34	11.67	2.18	0.60
B3108	8.59	5.71	2.17	0.60
B3109	3.57	1.73	2.18	0.62
B3110	25.48	33.25	2.60	1.92

6.2.1 Results of characterization plasmid 1

Table 6.2 shows the mean GFP and RFP expression of characterization plasmids B3101 through B3110 as compared to both the negative control, CW2553/pJAT18 and the empty characterization plasmid I13514. The mean RFP expression of all the characterization plasmids was negligible when compared to the negative control. Mean GFP expression ranged from negligible compared to the negative control to close to the maximum indicated by I13514. The range in mean GFP of the different characterization plasmids was unexpected as the presence of the terminator under test should only affect the coding region downstream from it. The exact fluorescence measurements of each characterization plasmid are shown in Figures 6-3 through 6-12.

Of the ten terminators tested with this version of the characterization device, four behaved in such a way that indicated high termination efficiency. Devices containing terminators B1002, B1004, B1006, and B1010 all expressed minimal levels of RFP and high levels of GFP. Devices containing terminators B1001, B1005, and B1009 expressed minimal levels of both GFP and RFP, and as such the termination efficiency of those terminators cannot be accurately judged with these results.

Table 6.3: This table shows the mean GFP and mean RFP expression of characterization plasmids B3201 through B3210. The mean GFP and RFP expression of the negative control CW2553/pJAT18 and I13515 are shown also for comparison.

Sample	Mean GFP	std GFP	Mean RFP	std RFP
CW2553/pJAT18	2.86	1.32	2.18	0.81
I13515	210.30	102.48	2.22	0.62
B3201	37.95	40.89	2.82	1.46
B3202	3.14	1.61	2.17	0.60
B3203	33.87	41.70	2.34	0.81
B3204	13.28	14.71	2.59	1.17
B3205	29.70	32.85	3.16	2.02
B3206	4.20	4.98	2.18	0.67
B3207	35.85	29.08	3.02	1.55
B3208	9.90	8.04	2.19	0.62
B3209	12.50	9.56	2.20	0.62
B3210	9.01	6.17	2.18	0.60

6.2.2 Results of characterization plasmid 2

Table 6.3 shows the mean GFP and RFP expression of characterization plasmids B3201 through B3210. All these characterization plasmids had the same flaw as the control plasmid I13515. The possible presence of an RNAse site in the RFP coding region made it such that there was limited RFP expression in all the characterization plasmids, and input to the terminators could not be accurately measured. Figures 6-13 through 6-22 show the exact fluorescence of these characterization plasmids.

For these characterization devices, it is necessary to ignore the RFP measurements, as these devices can only accurately measure the GFP output. A strong terminator characterized by one of these devices will show low levels of GFP output, while a weak terminator will show high levels. The average GFP measured on the empty plasmid I13515 was 210.30. Of the terminators tested with these characterization plasmids, B1002 and B1006 cause the greatest decrease of mean GFP expression to 3.14 and 4.2 respectively.

Figure 6-13: This figure shows the GFP vs RFP expression of each cell in the sample of B3201 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1001 reduces the mean GFP output from 210.30 as measured by the control I13515 to 37.95, decreasing the output by 81%

Figure 6-14: This figure shows the GFP vs RFP expression of each cell in the sample of B3202 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1002 reduces the mean GFP output from 210.30 as measured by the control I13515 to 3.14, decreasing the output by 99%
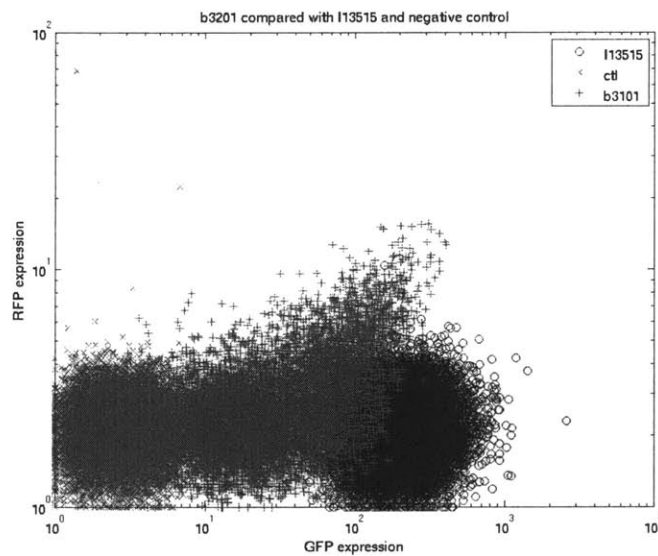
Figure 6-15: This figure shows the GFP vs RFP expression of each cell in the sample of B3203 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1003 reduces the mean GFP output from 210.30 as measured by the control I13515 to 33.87, decreasing the output by 83%
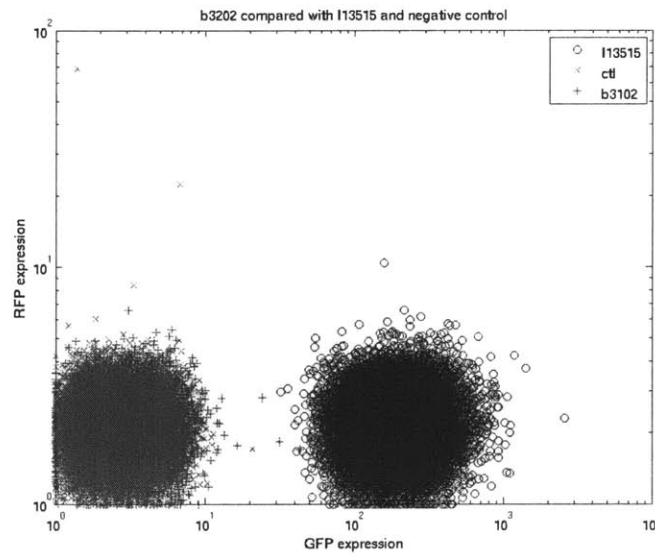
Figure 6-16: This figure shows the GFP vs RFP expression of each cell in the sample of B3204 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1004 reduces the mean GFP output from 210.30 as measured by the control I13515 to 13.28, decreasing the output by 94%
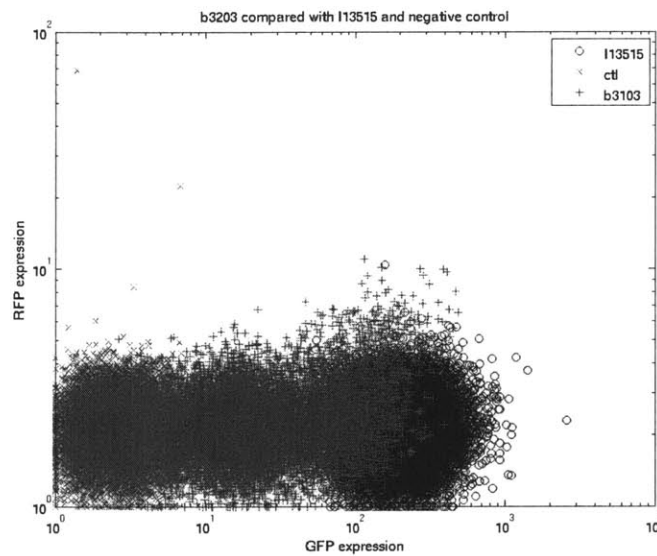
Figure 6-17: This figure shows the GFP vs RFP expression of each cell in the sample of B3205 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1005 reduces the mean GFP output from 210.30 as measured by the control I13515 to 29.70, decreasing the output by 86%
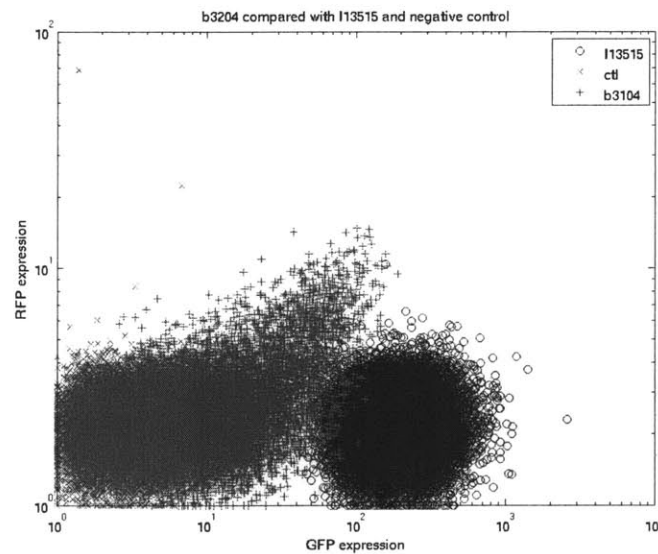
Figure 6-18: This figure shows the GFP vs RFP expression of each cell in the sample of B3206 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1006 reduces the mean GFP output from 210.30 as measured by the control I13515 to 4.20, decreasing the output by 98%
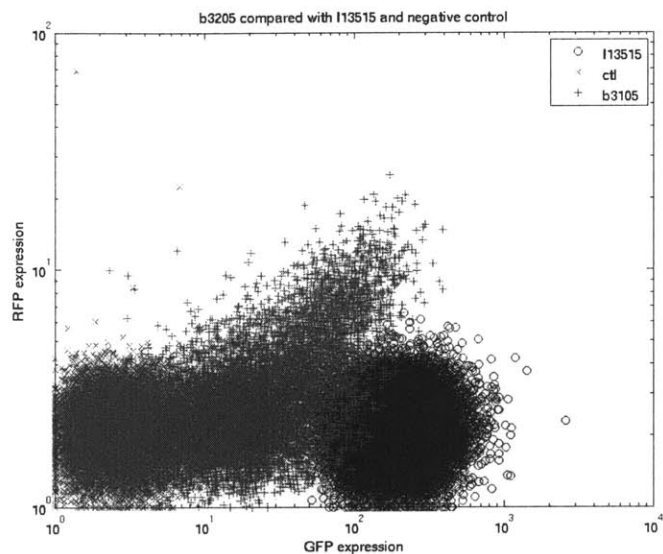
Figure 6-19: This figure shows the GFP vs RFP expression of each cell in the sample of B3207 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1007 reduces the mean GFP output from 210.30 as measured by the control I13515 to 35.85, decreasing the output by 83%
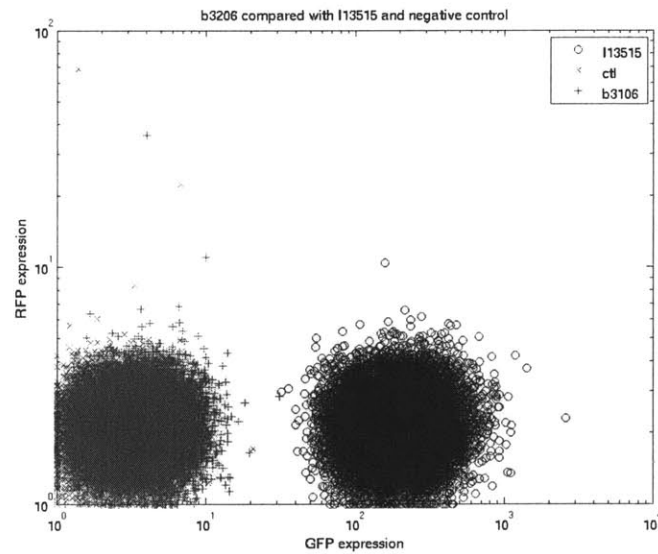
Figure 6-20: This figure shows the GFP vs RFP expression of each cell in the sample of B3208 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1008 reduces the mean GFP output from 210.30 as measured by the control I13515 to 9.90, decreasing the output by 95%
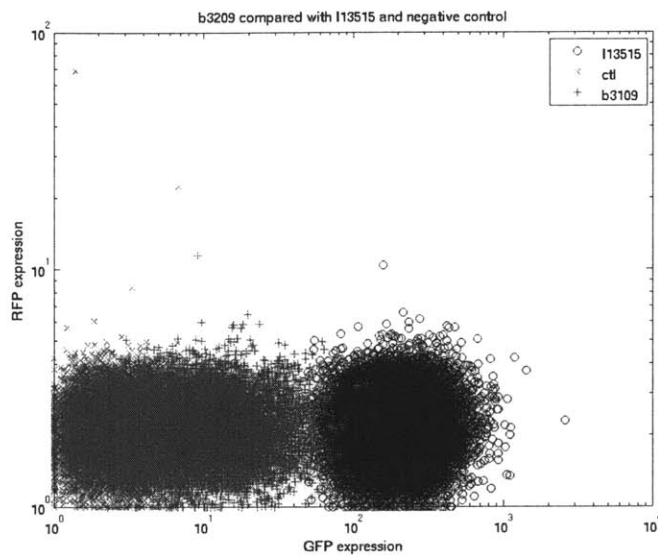
Figure 6-21: This figure shows the GFP vs RFP expression of each cell in the sample of B3209 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1009 reduces the mean GFP output from 210.30 as measured by the control I13515 to 12.50, decreasing the output by 94%
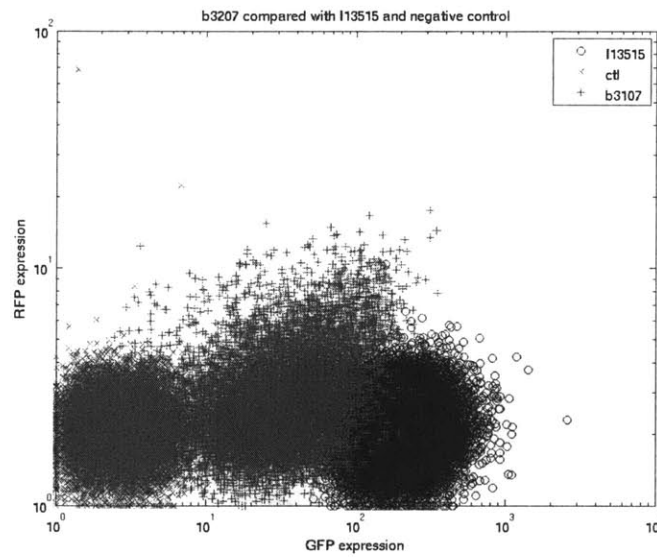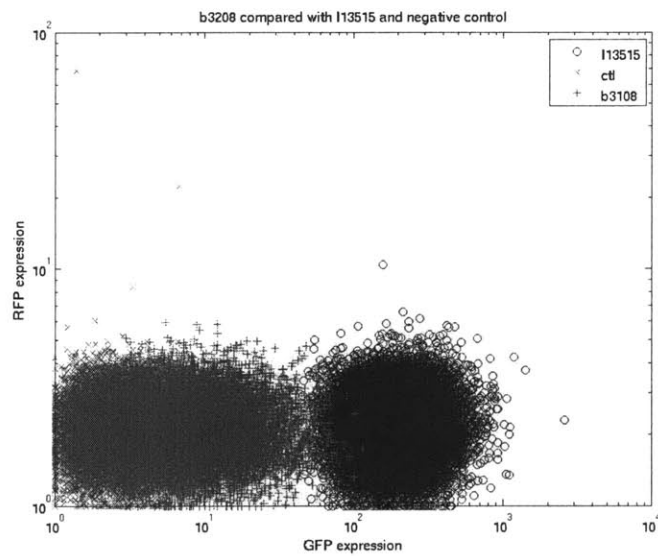
Figure 6-22: This figure shows the GFP vs RFP expression of each cell in the sample of B3210 compared to the negative control CW2553 and the empty characterization plasmid I13515. Due to the presence of RNAse cut site in the RFP coding region upstream of the terminator, only the output of the terminator can be measured. The presence of terminator B1010 reduces the mean GFP output from 210.30 as measured by the control I13515 to 9.01, decreasing the output by 96%
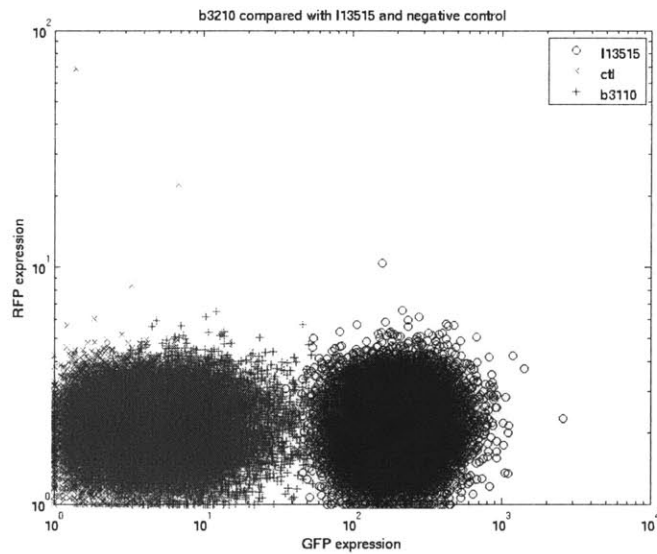
Table 6.4: This table shows the termination efficiencies of the new BioBrick terminators B1001 through B1010. The strongest terminators are B1002 and B1006. The weakest terminator is B1001.

Terminator	TE
B1001	0.81
B1002	0.99
B1003	0.83
B1004	0.93
B1005	0.86
B1006	0.98
B1007	0.83
B1008	0.95
B1009	0.94
B1010	0.95

6.3 Termination Efficiency

Only the results from the second set of characterization devices were used to calculate termination efficiency. The control for the first set, I13514, did not have enough cells with significant GFP or RFP expression to accurately measure input and output of the terminator under test. Calculations of termination efficiency can be performed with only the output of the terminators, measured by the second set of characterization devices. Termination efficiency would be measured by the ratio of the mean GFP of a characterization device to the mean GFP of control I13515. The mean TE was calculated by average the TE of each cell in the sample population. Termination efficiency was calculated by the following formula.

Table 6.4 shows the average termination efficiency of the artificial BioBrick terminators B1001 through B1010 while Figures 6-23 through 6-32 show the histograms of the TE of the terminators as measured by the second set of characterization devices. Terminators B1002 and B1006 are the strongest terminators with mean % TE of .99 and .98 respectively. Other strong terminator with a % TE above .9 are B1004,

Figure 6-23: The figure shows a histogram of the termination efficiency of the terminator B1001, calculated using the data from B3201. The average termination efficiency is 0.81.

Figure 6-24: The figure shows a histogram of the termination efficiency of the terminator B1002, calculated using the data from B3202. The average termination efficiency is 0.99.

Figure 6-25: The figure shows a histogram of the termination efficiency of the terminator B1003, calculated using the data from B3203. The average termination efficiency is 0.83.

Figure 6-26: The figure shows a histogram of the termination efficiency of the terminator B1004, calculated using the data from B3204. The average termination efficiency is 0.94.
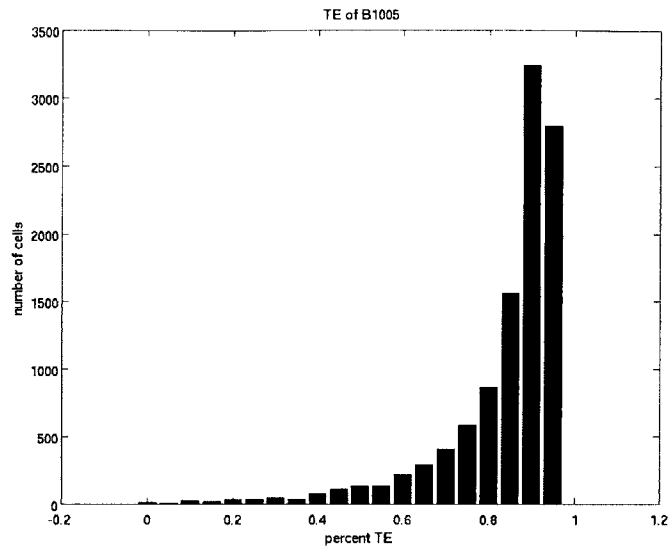
Figure 6-27: The figure shows a histogram of the termination efficiency of the terminator B1005, calculated using the data from B3205. The average termination efficiency is 0.86.
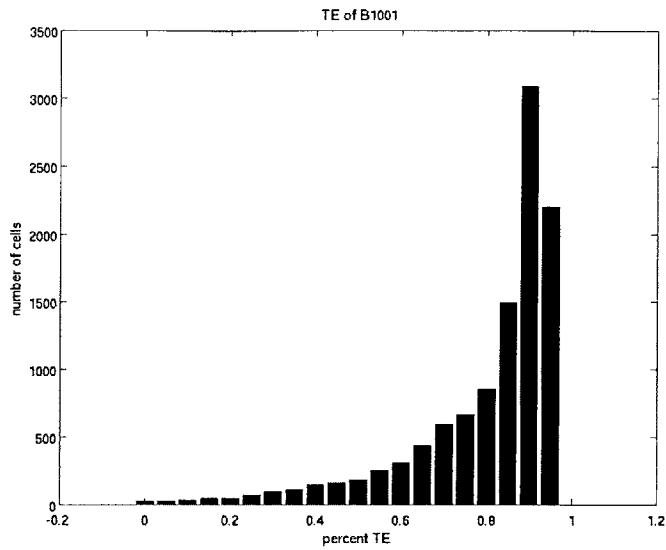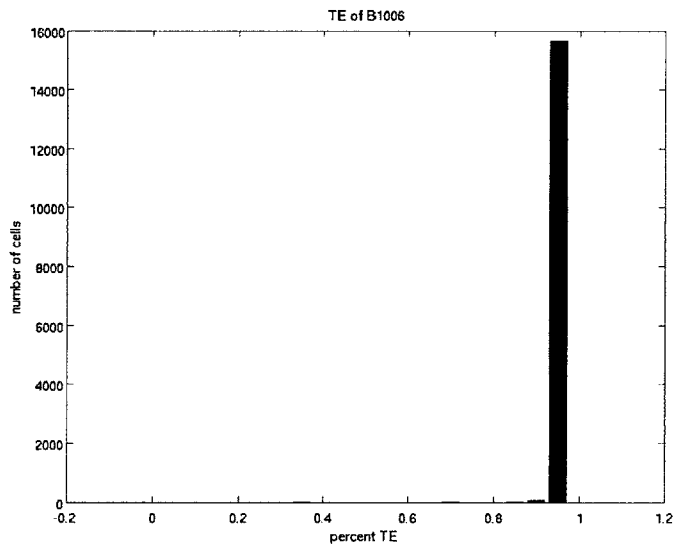

Figure 6-28: The figure shows a histogram of the termination efficiency of the terminator B1006, calculated using the data from B3206. The average termination efficiency is 0.98.
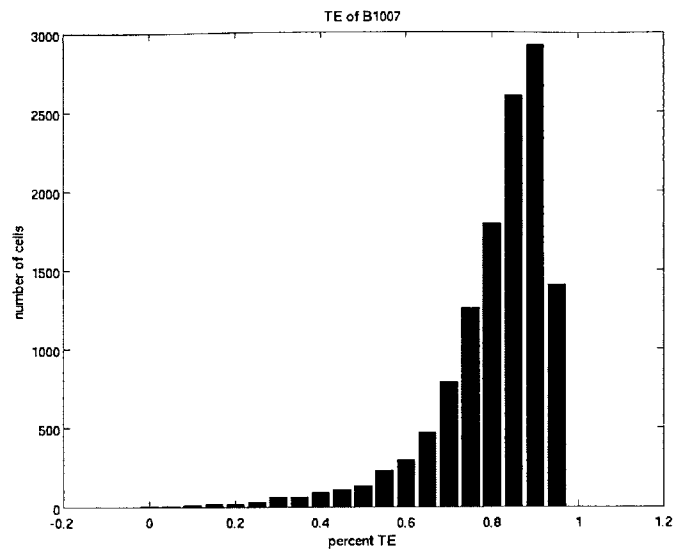
Figure 6-29: The figure shows a histogram of the termination efficiency of the terminator B1007, calculated using the data from B3207. The average termination efficiency is 0.83.
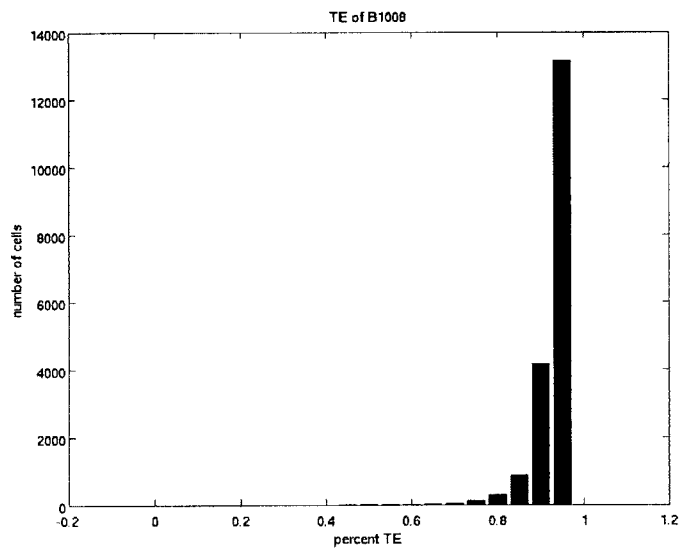
Figure 6-30: The figure shows a histogram of the termination efficiency of the terminator B10, calculated using the data from B3208. The average termination efficiency is 0.95.

TE of B1009
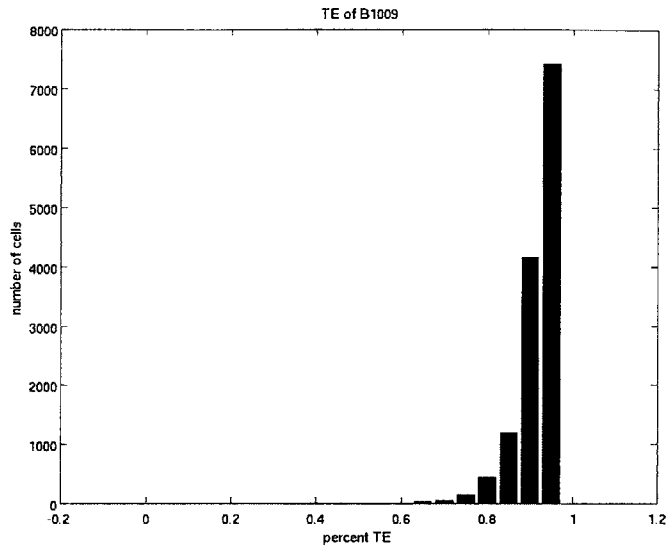
Figure 6-31: The figure shows a histogram of the termination efficiency of the terminator B1009, calculated using the data from B3209. The average termination efficiency is 0.94.
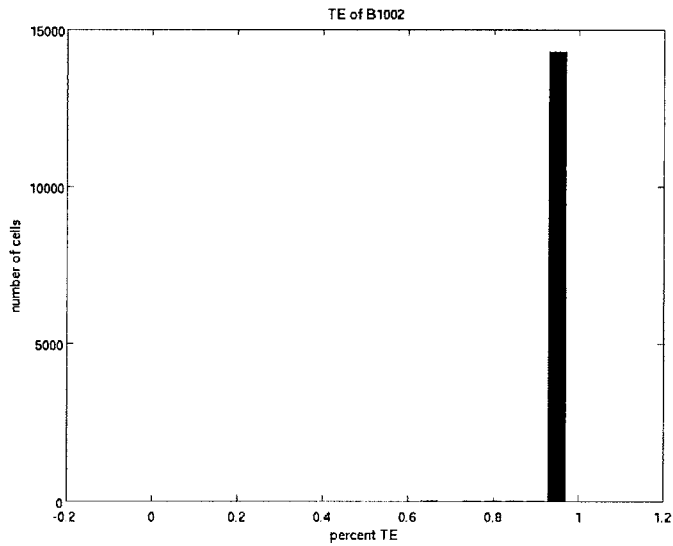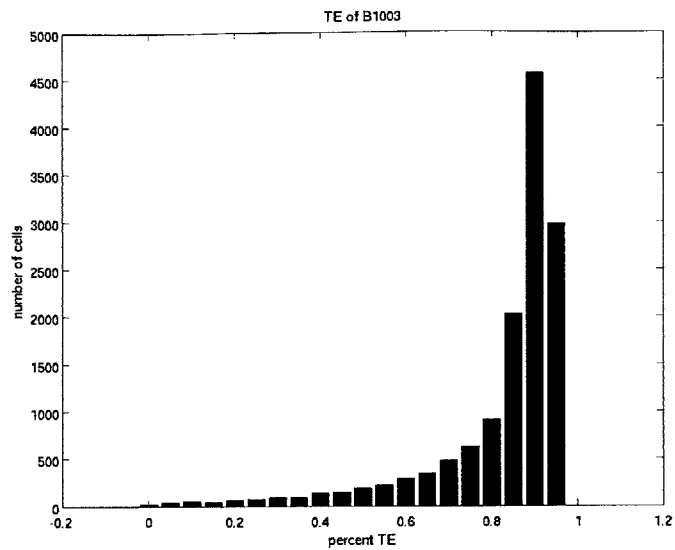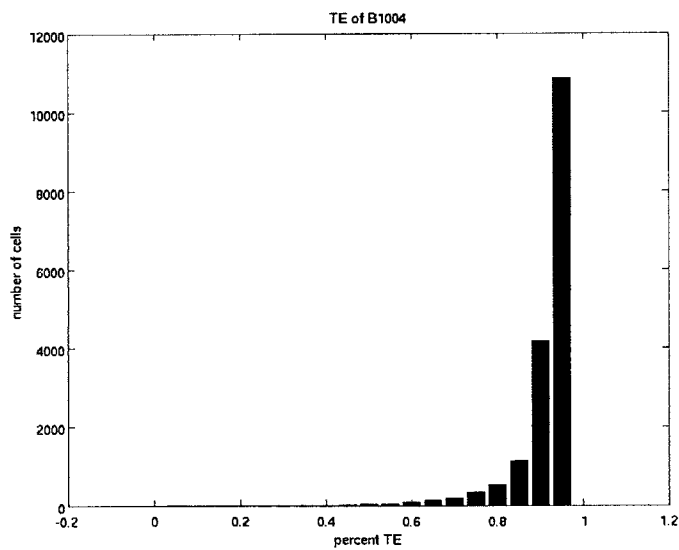

TE of B1010
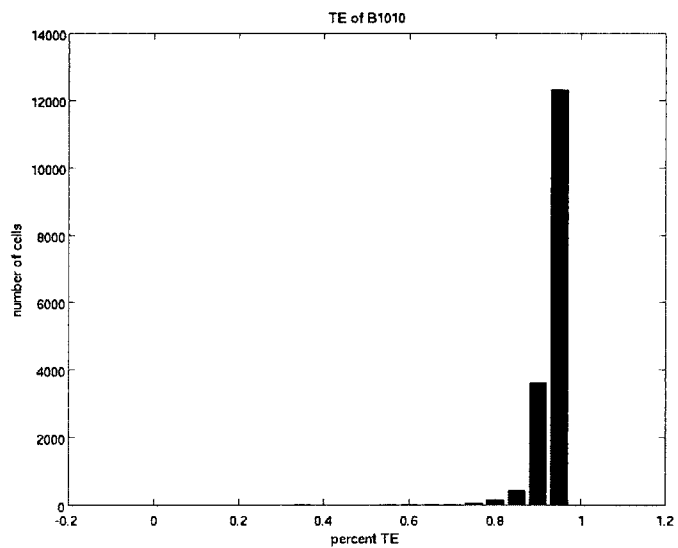
Figure 6-32: The figure shows a histogram of the termination efficiency of the terminator B1010, calculated using the data from B3210. The average termination efficiency is 0.96.

B1008, B1009, and B1010. The remaining four terminators, B1001, B1003, B1005 and B1007 are all weaker, with % TE under .86. As these termination efficiencies were only calculated with the data from the second set of characterization plasmids, no final conclusions can be made until the behavior of the terminators is verified with the first set of characterization plasmids.

Chapter 7

Discussion

Differences in data from the two different characterization devices need to be resolved. Several terminators, when tested with characterization plasmid 1, reduced the expression of the upstream GFP as well as the downstream RFP but testing with characterization plasmid 2 shows that they have high termination efficiency. In addition, the actual termination efficiencies were very different from the predicted values.

7.1 Effects of mRNA stability

The presence of an RNAse site in the RFP coding region would destabilize the mRNA for both proteins and result in minimal RFP expression in all samples. A strong hairpin of a terminator would help stabilize the mRNA after it has been cut, and slow the rate of degradation. When using the version of the characterization plasmid with GFP upstream of the terminator and RFP downstream, the terminator under test would have the job of stabilizing the remaining mRNA. A strong terminator would be able to slow the degradation of the remaining mRNA, so the resulting system would have high levels of GFP expression. A weak terminator would be unable to protect the remaining mRNA, causing the GFP coding region to be degraded as well. The resulting systems would then produce neither GFP nor RFP.

Of the terminators tested, B1002, B1004, B1006, and B1010 proved to be strong

Table 7.1: This table shows the termination efficiencies of the new BioBrick terminators B1001 through B1010 as well as the mRNA stabilization ability. mRNA stabilization is based on much GFP was produced a terminator was tested with version 1 of the characterization plasmid as compared to control I13514. Strong terminators should also be able to stabilize mRNA better than weak terminators. B1008 and B1009 have high % TE, but are unable to stabilize mRNA. As the data from the two different characterization plasmids conflict for these two terminators, no conclusions can be made about them. The best terminators are B1002, B1004, B1006 and B1010.

Terminator	TE	GFP produced
B1001	0.81	0.28
B1002	0.99	1.08
B1003	0.83	0.54
B1004	0.93	1.08
B1005	0.86	0.18
B1006	0.98	1.20
B1007	0.83	0.69
B1008	0.95	0.41
B1009	0.94	0.18
B1010	0.95	1.23

enough to prevent degradation of the GFP coding region. Terminators B1001, B1005, and B1009 proved to be poor at mRNA stabilization as shown by the fact that they have the lowest levels of GFP expression compared to the control I13514. The remaining terminators B1003, B1007 and B1008 provided a moderate amount of protection and allowed approximately 50% GFP expression as compared to the control.

Table 7.1 shows a summary of the BioBrick terminators and their termination efficiencies, as well as their ability to stabilize mRNA. With the exception of B1008 and B1009, the stong terminators were able to prevent degradation of the GFP mRNA in the first characterization plasmid. Conclusions of the termination efficiency of B1008 and B1009 cannot be made as their behavior in the two characterization plasmids contradict each other. The four best terminators are B1002, B1004, B1006, and B1010.

Table 7.2: This table shows sequences, predicted % TE, measured % TE, and the error in the prediction. The strongest terminators are B1002, B1006, B1010, and B1004. The formula used to predict termination efficiency was most accurate when the terminator had approximately 6 thymine residues in the tail. The most accurately predicted terminators were B1002, B1003, B1006 and B1007..

Terminator	Predicted TE	Measured TE	error
BBa_B1001	0.95	0.81	0.17
BBa_B1002	0.90	0.99	0.09
BBa_B1003	0.80	0.83	0.04
BBa_B1004	0.55	0.93	0.40
BBa_B1005	0.25	0.86	0.70
BBa_B1006	0.95	0.98	0.03
BBa_B1007	0.80	0.83	0.04
BBa_B1008	0.70	N/A	N/A
BBa_B1009	0.40	N/A	N/A
BBa_B1010	0.10	0.95	0.89

7.2   Accuracy of predicted termination efficiencies

The most accurate predictions of termination efficiency occurred when the terminator in question had a poly(T) tail of approximately 6nt. The formula was least accurate when predicting termination efficiencies of terminators with tails less than 5nt long. predicted termination efficiencies of the new BioBrick terminators are shown in Table 7.2.

The most surprising result was that B1010 proved to be one of the most effective terminators while B1001 had the lowest termination efficiency. Since B1010 only had a 3nt T tail and a long stem loop structure, its predicted termination efficiency was only around 0.1 but its actual termination efficiency was measured to be 0.95. The terminator B1004 was another terminator that proved to be much more successful than predicted, with a predicted % TE of .55 but an actual % TE of 0.93. B1005, while not as effective as B1010 or B1004, had a measured % TE of .86, but was predicted to have a % TE of .25. B1001, contrary to initial expectations, was a poor terminator despite a poly(T) tail of 9nt and a high t score. It is not known at this time why these terminators behaved in this manner.

Of the remaining terminators, B1002, B1003, B1006 and B1007 all behaved as expected, with the predicted termination efficiency coming within 10% of the measured termination efficiency. Terminators B1008 and B1009 could not be accurately characterized, and as such, no comparisons could be made between the measured termination efficiency and the predicted. This leads to the conclusion that the formula used to predict termination efficiency is most accurate when applied to terminators with moderate poly(T) tails of approximately 6nt in length.

7.3   Future works

Several things can be done to improve the terminators in the future. The first thing to do would be to verify and retest the control plasmid I13514 to determine why only a minority of the cells carrying that plasmid produce sufficient GFP and RFP. This may be due to mRNA instability, but further tests should be performed, and, if necessary, a new control designed. The control plasmid I13515 should also be updated to prevent mRNA stability by removing mRNA3 cut sites from the RFP coding region or use hairpins to stabilize the remaining mRNA if the cut site cannot be moved. The constitutive controls of I13521 and I13522 should also be reviewed to see if the fluorescence of these controls could be increased.

In future experiments, calibration beads will be run on the flow cytometer, so data from different days of measurements can be compared. Three sets of measurements were taken during when the terminators were characterized, but only one set could be analyzed. While the results were overall consistent across multiple days of measurements, minor differences in the setup of the machine made combining those measurements ill advised. Running calibration beads at the start of each session of flow cytometry will provide a baseline to compare the performance of the flow cytometer across different days.

Further studies into designing new terminators will include using the device characterization plasmid designed by Endy Lab to characterize any new terminators. Using the Endy Lab plasmid will result in reducing the number of constructions nec-

essary to prepare a BioBrick part for characterization. New terminators will have more varied stem loops in an attempt to vary termination efficiency, and will have thymine tails of approximately 6nt to maximize the effectiveness of the formula used to predict termination efficiency.

Chapter 8

Conclusion

Ten new terminators were designed based on previous research of terminator structure and termination efficiency. The terminators were built by PCR extension, ligated into a BioBrick plasmid backbone, and transformed into TOP10 cells. Characterization devices were built to test the terminators. Input and output of the terminator were measured by expression of RFP and GFP. Characterization devices were then placed into the E. coli strain CW2553/pJAT18, which hijacks the arabinose transport system to provide controlled input to the terminator.

Of the ten terminators designed and tested, B1002, B1004, B1006 and B1010 proved to be strong terminators with termination efficiencies above 90%. Terminators B1001, B1003, B1005 and B1007 were weaker, all with termination between 80% and 86%. Due to conflicting data, no conclusions could be made about terminators B1008 or B1009.

The algorithm used to predict termination efficiency based on terminator sequence was most accurate for terminators with poly(T) tails of around 6nt. The error between the predicted and actual termination efficiencies were under 10% for terminators B1002, B1003, B1006, and B1007. Predictions for termination efficiencies for terminators with the shortest tails, B1005 and B1010 were least accurate, with errors above 80%.

All terminations characterization in this project were added to the Registry of Standardized Parts at http://parts.mit.edu. Future studies will include the design of

better controls and measurement methods as well as attempting to expand the range of termination efficiencies for designed terminators

Bibliography

[1] Endy, D. iGEM 2005 Supplement. Available at
http://web.mit.edu/endy/www/igem/iGEM.supplement.pdf. Accessed May 1,
2005.

[2] Knight, T. Idempotent Vector Design for Standard Assembly of Biobricks.
DSpace at MIT Web site. 2001. Available at:
http://dspace.mit.edu/bitstream/1721.1/21168/1/biobricks.pdf Accessed May
8, 2005.

[3] d'Aubenton Carafa Y, Brody E, and Thermes C. *Prediction of rho-independent
Escherichia coli transcription terminators. A statistical analysis of their RNA
stem-loop structures.* J Mol Biol 1990 Dec 20; 216(4) 835-58.

[4] Ermolaeva MD, Khalak HG, White O, Smith HO, and Salzberg SL. *Prediction
of transcription terminators in bacterial genomes.* J Mol Biol 2000 Aug 4; 301(1)
27-33.

[5] Kingsford CL, Ayanbule K, Salzberg SL. *Rapid, accurate, computational discov-
ery of Rho-independent transcription terminators illuminates their relationship
to DNA uptake.* Genome Biol. 2007;8(2):R22.

[6] Abe H and Aiba H. *Differential contributions of two elements of rho-independent
terminator to transcription termination and mRNA stabilization.* Biochimie
1996; 78(11-12) 1035-42.

[7] Wilson KS and von Hippel PH. *Transcription termination at intrinsic termina-tors: the role of the RNA hairpin.* Proc Natl Acad Sci U S A 1995 Sep 12; 92(19) 8793-7.

[8] Cheng SW, Lynch EC, Leason KR, Court DL, Shapiro BA, and Friedman DI. *Functional importance of sequence in the stem-loop of a transcription terminator.* Science 1991 Nov 22; 254(5035) 1205-7.

[9] Lynn SP, Kasper LM, and Gardner JF. *Contributions of RNA secondary struc-ture and length of the thymidine tract to transcription termination at the thr operon attenuator.* J Biol Chem 1988 Jan 5; 263(1) 472-9.

[10] Markham N R, Zuker M. *DINAMelt web server for nucleic acid melting predic-tion.* Nucleic Acids Res. 2005 Jul 1;33(Web Server issue):W577-81.

[11] Khlebnikov A, Skaug T, Keasling JD. *Modulation of gene expression from the arabinose-inducible araBAD promoter.* J Ind Microbiol Biotechnol. 2002 Jul;29(1):34-7.

[12] Kelly, J. Design and Evolution of Engineered Biological Systems. DSpace at MIT Web site. 2005. Available at
http://dspace.mit.edu/bitstream/1721.1/21169/1/JasonKelly.ThesisProposal.pdf
Accessed May 5, 2005.