

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Engineering Scalable Combinational Logic in Escherichia coli Using Zinc Finger Proteins

### Permalink

<https://escholarship.org/uc/item/05j372zt>

### Author

Holtz, William Joseph

### Publication Date

2011

Peer reviewed|Thesis/dissertation

Engineering Scalable Combinational Logic in *Escherichia coli*  
Using Zinc Finger Proteins

by

William Joseph Holtz

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering-Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jay D. Keasling, Co-chair  
Professor Michel M. Maharbiz, Co-chair  
Professor Murat Arcak  
Professor John Christopher Anderson

Spring 2011

Engineering Scalable Combinational Logic in *Escherichia coli* Using Zinc Finger Proteins

Copyright 2011

by

William Joseph Holtz

## Abstract

Engineering Scalable Combinational Logic in *Escherichia coli* Using Zinc Finger Proteins

by

William Joseph Holtz

Doctor of Philosophy in Engineering-Electrical Engineer and Computer Science

University of California, Berkeley

Professors Jay D. Keasling and Michel M. Maharbiz Co-chairs

Available to synthetic biologists are a wide range of genetic devices. Many of these devices are able to either sense or alter local conditions. The ability to sense a multitude of inputs combined with diverse outputs could enable engineered organisms that interact with their environment in new and complex ways. Currently the complexity of such systems has been limited by our ability to integrate several inputs into a desired output. Simple combinational logic functions, containing 1 to 3 logic gates, have been constructed in *Escherichia coli*, but more complex logic networks are needed to fully exploit the opportunities presented by these sensors and actuators. Use of genetic logic gates is constrained by the specific molecular interactions that are used to implement each gate. These interactions involve diffusible molecules that can move within the cytoplasm of the cell and therefore are not spatially separated from other gates. To make larger logic blocks, sets of gates that use unique molecular interactions with minimal crosstalk are required.

Zinc finger proteins (ZFPs) can be used to predictably create a large number of unique protein-DNA interactions. These proteins can then be used to build transcriptional activators or repressors in *E. coli*, but these methods are not well defined. Attempts at using ZFPs to make one-hybrid transcriptional activators have failed to give a fold activation of higher than 2. ZFP repressors based on steric hindrance of RNA polymerase performed better with fold repression values of up to 300. The positional dependence of the ZFP operator site within the promoter was investigated, and both position and dissociation constant were found to play important roles in determining the level of repression.

ZFP based repressors without cooperativity cannot be used to create logic gates. A new inverter topology using both ZFP based repressors and sRNA was designed. This topology uses a reference promoter to set the switching threshold of the gate. There are no cooperative interactions in this topology, but the maximum slope of the transfer function is similar to a Hill-equation with a coefficient of 10. The high slope and excellent transfer function of these gates make them robust to many types of parameter variation and noise.

A set of 27 validated ZFP repressors and 27 promoters with ZFP operator sites were created and tested for non-orthogonal interactions. A sub-set of 5 repressor-promoter pairs were found to have a high degree of orthogonality where the cognate pairs resulted in more than 73% attenuation of the promoter and non-cognate pairs gave less than 19%

attenuation. The ZFPs and promoter used in this task were far from optimal and these attenuation values could readily be improved.

The combination of these orthogonal repressor-promoter pairs and the new logic gate topology should enable more logic gates to be implemented in a single *E. coli* cell.

This dissertation is dedicated to my late father, who taught me the value of a strong work ethic. I know he would have been very proud to see this work come to completion.

## Table of Contents

Introduction.....	v
Acknowledgements.....	vii
Chapter 1 - Synthetic Biology.....	1
Biology as an engineering discipline .....	1
Functional composition.....	2
Orthogonality.....	3
Chapter 2 - C <sub>2</sub> H <sub>2</sub> -type Zinc Finger Proteins.....	4
Properties .....	4
Synthetic Zinc Finger Proteins.....	4
Rational design based on individual fingers.....	4
Pair wise finger design methods .....	5
DNA Assembly.....	6
PCR based methods.....	6
BglBricks.....	6
Golden Gate.....	8
Chapter 3 - Synthetic Zinc Finger Protein Transcription Factors.....	14
Existing methods.....	14
Two-hybrid .....	14
One-hybrid.....	15
Dimeric zinc finger proteins .....	15
Steric hindrance of RNA polymerase .....	16
Design of a synthetic promoter that can bind many zinc finger proteins.....	18
Creating the objective function.....	20
Analysis of the synthetic promoter P <sub>ZC35</sub> .....	23
Assay design for transcriptional attenuation experiment .....	25
Positional dependence in transcriptional attenuation experiment.....	29
Measurement of dissociation constants for zinc finger proteins .....	39
Purification of zinc finger proteins.....	41
Creating repressors for other promoters .....	43
Chapter 4 Topologies for building combinational logic gates.....	45
Desired properties of logic gates .....	45
Previous work.....	46
Methodology for designing new logic gate topologies .....	47

Positive feedback with competitive binder topology .....	50
Analysis of topology .....	50
Modeling and analysis .....	50
Implementation and testing.....	56
Threshold sensing topology using sRNA .....	65
Choice of reference levels .....	65
Using sRNAs .....	66
Overview of topology.....	67
Modeling and analysis .....	68
Implementation and testing.....	74
Chapter 5 Creating orthogonal sets of repressor-promoters pairs.....	77
Orthogonality defined .....	77
Why orthogonality is needed .....	77
Desired specification for sets of orthogonal promoter-repressor pairs .....	77
Creating orthogonal sets .....	80
Choosing zinc finger proteins.....	80
Choosing promoters.....	80
Orthogonality assay .....	82
Results and analysis.....	84
Metrics for orthogonal sets .....	85
Reduced sets .....	86
Chapter 6 - Multi-gate logic, putting it all together .....	88
Architecture .....	88
Methodology .....	89
Challenges .....	89
Construct stability.....	89
Signal fidelity .....	90
Chapter 7 – Bibliography .....	91
Chapter 8 – Appendix: sequence information .....	97
Empty vector Cole1/Kan .....	97
ZFP03 leak on SC101/Kan.....	98
ZFP5475 matched act/rep on SC101/Kan .....	100
ZFP02 matched atc/rep on SC101/Kan .....	104
ZFP5475 inverter on SC101/Kan .....	107



ZFP03 matched atc/rep on SC101/Kan .....	111
ZFP02 leak on SC101/Kan.....	114
Zif268 leak on SC101/Kan.....	116
Zif268 matched act/rep on SC101/Kan .....	119
ZFP03 inverter on SC101/Kan.....	122
ZFP5475 leak on SC101/Kan .....	126
ZFP03 activator p15A/Kan .....	128
ZFP5475 activator p15A/Kan .....	131
ZFP03 mismatched atc/rep on p15A/Kan .....	133
ZFP5475 mismatched act/rep on p15A/Kan.....	135
ZFP5475 leak on p15A/Kan.....	138
ZFP03 leak on p15A/Kan.....	140
ZFP5475 matched atc/rep on p15A/Kan .....	141
ZFP03 matched atc/rep on p15A/Kan .....	144
Reporter for TetR-sRNA inverter .....	147
Reporter for steric hindrance ( $P_{ZC35}$ ) .....	148
TetR and sRNA inverter (+sRNA) .....	150
RFP-ZFP fusion expression from $P_{lacUV5}$ .....	153
TetR inverter (-sRNA) .....	155
ZFP expression from $P_{lacUV5}$ for steric hindrance experiments.....	158
ZFP-65 expression from $P_{BAD}$ for steric hindrance and orthogonality testing.....	160
ZFP-3 expression from $P_{BAD}$ for steric hindrance and orthogonality testing .....	163
ZFP-39 expression from pET-29 for protein purification.....	165
Reporter for orthogonality testing: operator-3 and BioFAB promoter 83.....	168

## Introduction

Synthetic biology questions if the complexities of the cell, with all of its interacting and closely coupled elements, can be made manageable by the application of the principles of engineering? Over the past ten years the synthetic biology community has formally presented this question and tried to answer it. The output of the community has already demonstrated that design of biological systems can be made easier through the use of engineering fundamentals. Obviously, there is not a clear line where biology crosses into the realm of an engineering discipline. But now the question has evolved to focus on streamlining the design and implementation of these systems. Perhaps a better way to approach the question is to look at the rate of change in our ability to create more complex biological systems.

The field of synthetic biology is rife with analogies, and the majority of these analogies compare synthetic biology to the field of electrical engineering. There are many reasons for this choice including a desire to be associated with an industry that has been highly successful in building complex systems using abstraction, hierarchies, standards, modeling and automation. For the goals of attracting research funding and investment capital, it is useful to draw similarities to an industry that transformed the world, improved lives and created vast amounts of wealth. There are also similarities to the networks of interactions that are central to the work of electrical engineers and synthetic biologists; thus, analytical frameworks and tools developed by electrical engineers can be adopted to quickly solve problems in synthetic biology. But all analogies have their limits and there is a danger in trying to follow them too far.

As an electrical engineer entering into the field of synthetic biology, I have tried to be careful not to overstate what lessons can be applied across both of these areas. Determining where the crossover of ideas can be productive has been an exercise in comparing systems, methodologies, and assumptions from both fields and evaluating where they converge or diverge. Once it has been determined how well the concept transfers between the domains, an evaluation of the utility of the concept in the new domain is essential. One way to assess this utility is to consider the impact the implemented idea may have on the previously mentioned rate of change in our ability to create more complex biological systems. If the idea can withstand the transfer between domains and has utility in the new domain, then it does not over-extend the analogy between the fields and is a good candidate with which to proceed.

Digital logic has been one of the most powerful abstractions to come out of electrical engineering. By working with many instances of only a handful of highly characterized devices, extremely complex systems can be made with minimal understanding of the details of the inner workings of the devices. Some of the assumptions inherent to making useful digital logic are access to many instances of the device, globally defined states of high and low, and that the inputs and output of all devices must be compatible. Of these assumptions, the creation of many instances of a device provides the largest hurdle. The ongoing development of synthetic zinc finger proteins offered a possible route. If successful the utility of digital logic in synthetic biology could be huge. Every organism is filled with networks that combine multiple input signals into changes in gene expression, so there must be value in performing such functions. However, natural systems have been created

via evolution. Faster and better understood methods are required for us to engineer these types of functions. The abstractions of digital logic could provide these methods.

By working through the analysis outlined above, it appears that there is an opportunity to advance synthetic biology by applying concepts from digital logic in electrical engineering.

## Acknowledgements

I have had a lot of time in graduate school to accumulate people to whom I am indebted for their advice, assistance, support and friendship.

None of this would have been possible without the enthusiastic support of my primary advisor, Jay Keasling. I have been fortunate to work in a lab where students are allowed to run with their own ideas and don't have to worry about when funding will dry up. Your attitude has been inspirational to me and I doubt I could have pushed past my failed projects without your unwavering belief in the value of my creativity and ideas.

My co-advisor, Michel Maharbiz, has played the ever changing role in my graduate career: from mentor, to friend, and finally to co-advisor. Thank you for filling in for this latest role and buffering me from a department that I barely fit into anymore. Your humor has brightened many a difficult day of graduate school. Most of all I appreciate the fearlessness you bring to your research. Not enough researchers have your willingness to go after difficult projects that are unlikely to reach their final goal.

My initial advisor, Roger Howe, had to leave Berkeley before I finished my degree, but I greatly enjoyed our time together. Thank you for putting up with my adventures in biology.

Twice when I was between advisors, Andy Neureuther graciously helped me navigate the department bureaucracy. Thank you for being the most caring professor in the department.

The Howe group was full of great people who were always willing to help each other out. Together Frank Zendejas and Noel Arellano worked with me through prelims, qualifying exams, and the transition when Roger Howe left Berkeley. Sunil Bhave, Elliot Hui and Uthara Srinivasan taught me much through our discussions.

Turi Aytur and Emmanuel Quevy were always interesting and entertaining whether debating science, current events or hanging out after a long day in the lab.

I joined the Keasling Lab with minimal biology knowledge and no molecular or microbial biology experience. Many Keasling Lab members have taken their time to teach me including Tim Ham, Mario Oullette, Jack Newman, Sung Kuk Lee, David Laidlaw, Aindrila Mukhopadhyay, James Carothers, Jim Kirby, John Dueber, Eric Luning, Suzan Yilmaz, Brian Pflieger, Kai Wang, Josh Gilmore, Nathan Hillson and Wes Marnier. Many Keasling members contributed to making graduate school a more pleasant place by just being a fun and friendly group of people. These people include Karen Wong, John Haliburton, Greg Bokinsky, Clem Fortman, Adrienne McKee, Jeffery Dietrich, Sarah Rodriguez, Becky Rutherford, Eric Steen, Lane Weaver, John Thorne, Tina Mahatdejkul, Sam Meyers, Rachel Eachus, Rumi Asano, Doug Pitera, Sydnor Withers, Bernardo da Costa, Eric Paradise, Trina McMahon, Afshan Shaikh, Sandra Villa, Sundiep Tehara and Matt Mattozzi.

Julius Lucks of the Adam Arkin Lab directly contributed to this research by giving me his sRNA devices and advice on how to best use them.

Ruth Gjerde of the EECS staff always had a friendly and reassuring smile that matched her ability to help me through any administrative mess.

I never told my mother the cardinal rule of graduate school – never ask a graduate student how long until they are done. But she figured it out on her own. Thank you for your love and support.

Finally, to my fiancée, Kier, meeting you was the best thing that happened to me in graduate school. I greatly appreciated your understanding of both the time commitments and emotional burdens of doing research. Thank you for reading and commenting on my dissertation draft.

# Chapter 1 - Synthetic Biology

## *Biology as an engineering discipline*

The goal of synthetic biology is to make engineering biological systems easier. By taking general engineering principles and applying them to the design of biological systems, one can hope that there will be improvements in the ability to predict system performance before construction, the time and money needed to deliver working systems, and the maximum size and complexity of systems created. General engineering principles include the following concepts: abstraction, standardization and decoupling [1].

Abstraction in engineering deals with removing complexity. In systems there are often various hierarchies of components. A designer who builds components on one level of the hierarchy only needs to know the interface specifications for the sub-components on the next lower level of the hierarchy. They do not need to know all the details of how the sub-components operate internally, only the information to interface sub-components together. By hiding the details of all the lower levels of the system, the total amount of knowledge the component designer must master to be successful is reduced. Additionally, the amount of information about a specific project that a designer must be able to mentally juggle is also lowered. Abstraction can play a large role in automating design processes, as abstracted systems are often simple enough that complete sets of formal rules can be enumerated. This then allows for computer algorithms to find optimal solutions for a given set of constraints without the intervention of someone with an expert level of knowledge.

Standards leverage the work of a community by making it easier for each individual to integrate the advances others have made into their own work. Standards can be made for many types of things, information or processes. Standards have been made for interfaces between objects, formats for data exchange, measurement methods, and even the creation of new standards. In terms of organizational structure, standards reduce the effort required for a party to integrate the work of another party into their own. Without standards, integrating an element into N systems requires order N times more work than if the systems utilized a common standard. When these savings are multiplied by a large number of exchanges in a community, the effort saved can be transformative and enable tasks that otherwise would be cost prohibitive.

Decoupling is the ability to break a task into smaller tasks that can be worked on independently. Decoupling can occur on several different levels within an engineering project. On a high level, a division of labor, where people with specialized skill sets perform different functions such as design and quality control, is an example of decoupling. Decoupling can also be applied to the design of components. By making components that function independently of the context in which they are used, it becomes possible to split the design of components into individual problems that can be easily addressed in parallel.

The synthetic biology community is still in the early stages of fully integrating engineering principles into their workflows. Abstraction is evident in the hierarchy of parts, devices, chassis, and systems that has become common. However there are not yet clear boundaries between these levels and most practitioners do not currently work on only one level of this hierarchy.

Progress on the creation and adoption of standards is clearer. Numerous standards on DNA assembly methodology have been created and used including BioBricks[2], which has been widely used in the International Genetically Engineered Machine competition (iGEM), and BglBricks[3], which has been used in several labs at UC Berkeley. An informal standard exists for how parts that utilize transcription events should have their interface on the DNA between the promoter and the start of the transcribed region. This also led to a standardized unit of transcriptional activity, Polymerase Per Second (PoPS)[4]. Some work has also been done on standard measurement methods that can be used to characterize components including a promoter measurement technique [5]. There has also been the adoption of standards developed by related communities. For example, the Systems Biology Markup Language [6] (SBML) has become the primary format for moving models of synthetic biology systems between various design and simulation tools.

Decoupling of job functions is in its infancy within synthetic biology, but some instances can be found. In many projects there are multiple genetic networks that are involved. The work to create or modify each of these genetic networks can often take place in parallel. Creating components that are decoupled from their context has proven to be a difficult problem. It is probably not possible to create components that function independent of all contexts, and therefore the problem should probably be restated as components that function properly within some defined range of contexts. Some of the most promising work thus far on component decoupling has been on the introduction of orthogonal transcription and translation networks into *E. coli* [7].

### ***Functional composition***

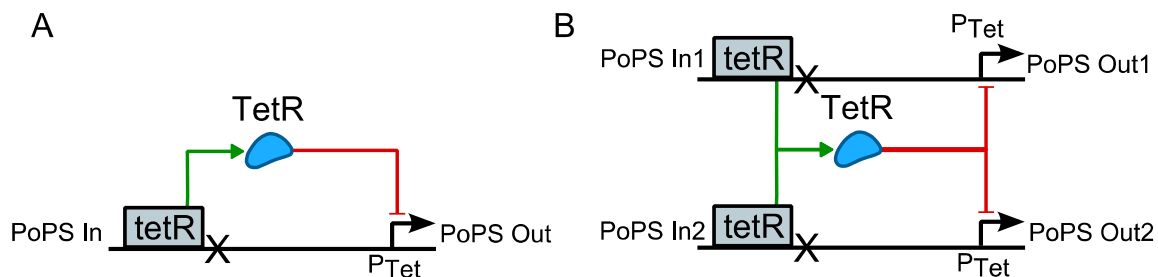
One of the biggest challenges facing synthetic biology is enabling functional composition. Functional composition is the ability to take multiple characterized elements that have been combined into a larger unit and be able to predict how the larger unit will behave.

An example of where functional composition has not yet been achieved is the interface between a 5' un-translated region (5'-UTR) and a gene coding sequence. Most of the control of translation rates is encoded within the 5'-UTR and therefore it could be expected that the 5'-UTR will contribute to the translation rate and the coding sequence will contribute the primary sequence of the protein; however, if the translation rate of an mRNA consisting of a 5'-UTR and a gene coding sequence is measured and then the coding sequence is swapped to contain a different gene, then not only will the resulting protein sequence change but the rate of translation will also change. This finding is mostly due to the mRNA secondary structure playing a significant role in determining the translation rate; therefore both the 5'-UTR and gene coding sequence can contribute to the mRNA secondary structure. The secondary structure of an mRNA influences translation rates primarily by modulating the ribosome's ability to access its binding site on the mRNA. When the ribosome binding site (RBS) of the mRNA is bound to other bases of the mRNA strand and is not single stranded, then the ribosome loses its ability to bind to the mRNA and initiate translation. Predicting the translation rate of an mRNA may be possible with the complete sequence of that mRNA, but it is unlikely that this prediction can be done with just a few characterized parameters from both the 5'-UTR and the gene coding sequence [8].

While functional composition may not be possible for the general case of any pair of 5'-UTR and gene coding sequence, there may be sets of 5'-UTRs and gene coding sequences that are amenable to functional composition. If such a set does exist, then it is likely that through creative design work the size of the set can be expanded. 5'-UTRs that contain strong hairpin structures may be less likely to form secondary structures that cross the 5'-UTR to gene coding sequence boundary, and therefore make it possible to design for functional composition between some subset of components. This loss of generalizability, where a specification only holds when the component is used within a range of conditions, is common throughout engineering. These tradeoffs can be made because the loss of flexibility is offset by gains in designer productivity.

### Orthogonality

In many domains of engineering, designing the first instance of a component is significantly harder than every additional instance. For example, designing a new ball bearing is a significant undertaking, but when a system requires four ball bearings each with the same specifications, then the first ball bearing design gets reused four times with minimal additional design work. In molecular biology, re-use is often more complicated because of diffusion and the limited constraints placed on the positions of molecules. All molecules in a cell have some ability to move relative to other molecules in that cell. There may be constraints placed on the relative movement of two molecules due to diffusion barriers such as cell or vesicle membranes, or the formation of a molecular complex may greatly restrict the movement of a molecule. Because genetic networks are primarily composed of molecular binding interactions and enzymatic reactions, the elimination of unwanted interactions and reactions must largely come from choosing or designing molecules that do not have compatible interfaces for binding. Spatial separation either with membranes[9] or scaffolds[10], [11] has also been demonstrated as methods of keeping reactions orthogonal, but it is unclear how to scale these methods to more complex systems.



**Figure 1-1 A) An inverter based on the tetR repressor only produces the protein TetR when the PoPS In level is high. When TetR is present it binds to the  $P_{Tet}$  promoter and turns it off. B) If two tetR based inverters are used within the same cell, the TetR protein produced by either input being high will bind to both output promoters and results in both outputs switching to a low PoPS level. Because the inverters are not orthogonal, the system behaves as a single NOR gate instead of two independent inverters.**



## Chapter 2 - C<sub>2</sub>H<sub>2</sub>-type Zinc Finger Proteins

### Properties

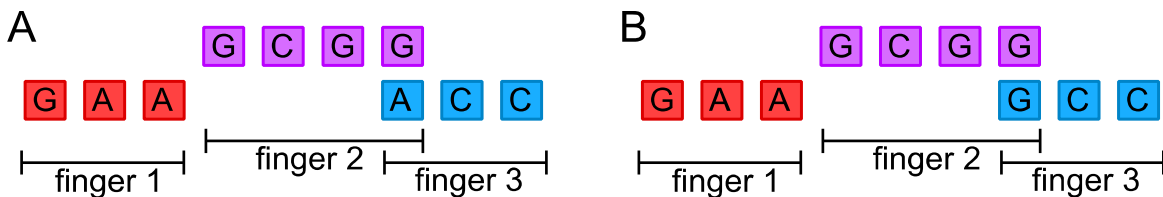
Zinc finger proteins (ZFPs) contain a fold coordinated by a zinc ion and commonly bind DNA, but some are able to bind other molecules such as RNA, proteins or small molecules. Zinc fingers are frequently found in transcription factors, especially in eukaryotes. There are several classes of zinc finger proteins and the most commonly used class in engineered systems are the C<sub>2</sub>H<sub>2</sub>-type. C<sub>2</sub>H<sub>2</sub>-type zinc fingers contain a Zn(II) ion coordinated by two cysteine and two histidine residues and a single C<sub>2</sub>H<sub>2</sub>-type zinc finger binds to 3-4 bases of double-stranded DNA (dsDNA). Natural C<sub>2</sub>H<sub>2</sub>-type zinc finger proteins generally contain 3 or more zinc fingers, allowing them to bind to 9 or more base pairs (bp) of DNA with dissociation constants commonly in the nanomolar range.

### Synthetic Zinc Finger Proteins

Several techniques have been used to create synthetic zinc finger proteins that can bind to a specified operator sequence. These methods can be broken into two broad classifications: screening or selection methods and rational methods. In the screening or selection methods a collection of coding sequences for individual fingers is created and then a large library of multi-finger coding sequences is generated with all or some of the fingers being randomly assigned from the collection. These coding sequences are then subjected to screening or selection by phage display[12-14] or an *in-vivo* binding assay[15], [16] to pick the best zinc finger protein in the library.

### Rational design based on individual fingers

In the rational methods individual zinc fingers have been characterized to determine the DNA sequence they bind[15], [17-20]. This characterization is usually done in the context of a 3-finger with the middle finger being the site investigated. Because some fingers interact with 4 base pairs of DNA, the identity of the neighboring fingers becomes important. When the middle finger interacts with 4 base pairs of DNA, the 4<sup>th</sup> base pair is also interacting with the neighboring finger. Fingers that bind to GNG sequences are known



**Figure 2-1 A) The DNA binding site for a zinc finger protein containing three fingers can clash if a GNG site is followed by an ANN or CNN site, because GNG sites interact with a fourth base pair. A protein created with these fingers would likely have poor binding characteristics. B) By changing the third finger to a GNN site the clash has been relieved and the protein is more likely to be a functional DNA binder.**

to interact with 4 base pairs, and the 4<sup>th</sup> base pair is required to be either G or T for a low dissociation constant and good specificity [21], as shown in Figure 2-1. Otherwise the final G of the middle finger and the initial A or C of the following finger will clash and result in a protein with a high dissociation constant and poor specificity.

In order to be able to rationally design zinc finger proteins that can bind to any arbitrary DNA sequence, fingers for each of the  $4^3=64$  triplets would need to be found and characterized. Thus far only 49 of the 64 possible triplets have high quality fingers [15], [17-20]. With 49 triplets, zinc finger proteins with 3 fingers can rationally be designed for  $(49/64)^3 \sim 45\%$  of all possible sites and 6 finger proteins can rationally be designed for  $(49/64)^6 \sim 20\%$  of all sites are possible.

Because rationally designed zinc finger proteins are easy to create, there was significant momentum behind their use, despite the limitations on the sequence space that can be targeted by rationally designed zinc finger proteins. Initial reports indicated that between 100% [22] and 60% [23] of rationally designed 3-finger proteins had the desired functionality. However these success rates were not universal. In 2008 a paper demonstrated that the initial reports were biased in the sequence space they targeted, and that only 24% of a less biased set of 3-fingers proteins worked correctly [24]. Extrapolating their data set to the completely unbiased case yields a 13% success rate on sites that can be targeted with the current set of good fingers. The rate of failure for a zinc finger protein was strongly correlated with the number of GNN fingers in the protein. All zinc finger proteins with no GNN fingers failed and 88%, 71%, and 41% the proteins containing 1, 2 and 3 GNN fingers respectively failed [24]. While this information reduced the expected utility of rationally designing zinc finger proteins, it also provided guidance on how to select binding sites with an increased likelihood of success. It is unclear if new sets of zinc fingers will be found that can overcome the current limitations that restrict the choice of operator site or require GNN fingers.

### **Pair wise finger design methods**

Because zinc finger proteins can interact with more than three base pairs of DNA, rational design based on individual fingers can have a high failure rate when neighboring fingers clash. Alternative design methods have been developed that use pairs of fingers in order to reduce these problematic interactions. Some of these methods are rational and use characterized pairs of fingers to build specific zinc finger proteins [12], [25]. Other methods use phage based selections to find the desired zinc finger protein from a large library of randomized candidates [26]. The phage based methods have been shown to produce high quality zinc finger proteins, but the methods are very time intensive. Rational pair based design is known to work better than rational methods based on single fingers, but public information about characterized pairs is sorely lacking. Only a few characterized pairs have been published. In order to match the diversity possible using single fingers, a total of  $49^2=2401$  pairs of fingers would need to be characterized. A proprietary dataset of characterized finger pairs exists and is owned by Sangamo. They will provide custom zinc finger proteins for a fee, but the costs are significant.

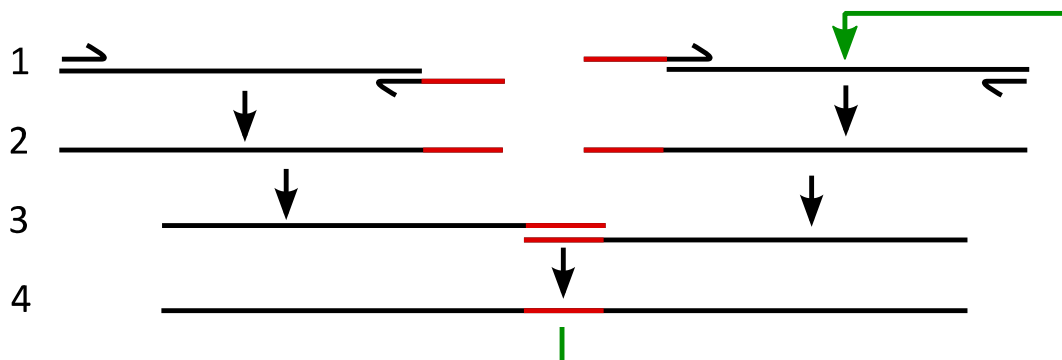
## DNA Assembly

### PCR based methods

Individual fingers can be assembled into multi-finger constructs using a PCR overlap extension method [21]. A schematic of the process is shown in Figure 2-2. Single fingers or groups of contiguous fingers are PCR amplified with primers containing 5' regions not homologous to the template DNA. These primer defined segments are designed to be homologous with regions added to PCR products for adjacent fingers. Purified PCR products for two neighboring fingers are then placed in a PCR mix and thermocycled in order to extend the two templates into a single product. This process can be iterated to join any number of fingers together.

### BglBricks

The majority of the zinc finger proteins used in this work were assembled using restriction enzymes within the context of the BglBrick standard. First, a set of plasmids each, containing a single zinc finger between BglBrick cloning sites, were constructed. These plasmids were made using the “Round-the-horn site-directed mutagenesis” method, which consists of a PCR using a plasmid template and phosphorylated primers with non-homologous 5' ends followed by a blunt end ligation [27]. This method results in the non-homologous region of the primers being inserted into the vector. The common sequence shared across all fingers along with the protein N and C terminal sequences are given in Table 2-1. The completed zinc finger proteins have ligation scars flanking each finger and the scar sequence (GGATCT) is not included in Table 2-1. The sequence of the variable region can be found in Table 2-5. The proteins constructed in this work all consisted of 6 zinc fingers. Because each zinc finger protein consisted of an N-terminal module, 6 finger modules, and a C-terminal module, a total of 7 junctions were ligated in order to create one zinc finger protein.



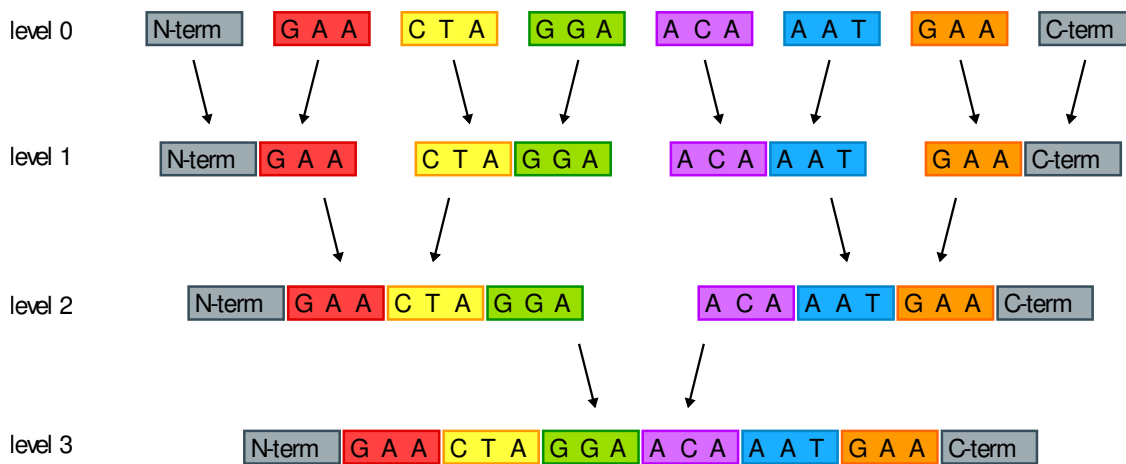
**Figure 2-2 PCR based assembly of zinc finger proteins. 1) Individual fingers or contiguous groups of fingers are PCR amplified with primers that contain a non-annealing region shown in red. 2) PCR products generated contain matching red regions. 3) Thermocycling allows single stranded DNA from the two products to anneal in the red region and extend to form a double stranded product. 4) The resulting double stranded DNA contains fingers from the two initial templates. This process can be iteratively performed as shown by the green arrow.**

**Table 2-1 Sequences of segments of zinc finger proteins for assembly with BlgBrick cloning**

Region	Coding Sequence	Amino Acid Sequence
N-terminus	CTGGAACCA	LEP
Finger, 5' of variable region	AAACCGTACAAATGTCCGGAATGT GGTAAATCCTTCTCC	KPYKCPECGKSFS
Finger, variable region	21 base pairs	7 amino acids
Finger, 3' of variable region	CATCAACGTACTCACA	HQRTHT
C-terminus	AAAACCTCT	KTS

By organizing the assembly as a full binary tree, the protein can be generated in 3 parallel assembly steps as shown in Figure 2-3. A best case time frame for completing such a 6 finger assembly would involve DNA digest, ligation and transformation on the first day, colony PCR and colony picking into liquid media on the second day, and DNA purification, digesting, ligation and transformation on the third day. If verification by sequencing is desired before proceeding to the next level of the assembly tree, then an additional day must be added for each level of the tree. Therefore if there are no failures and sequencing is not required, assembled and purified DNA could be in hand on the seventh day for a 6 finger assembly. If sequencing is required at each step, then the best case time to go from individual fingers a fully assembled 6 finger protein with verified sequence is 10 days.

The frequency of sequence verification should be related to the expected error rate, and for restriction enzyme based sub-cloning this rate can be relatively low. However the error rate is not only dependent on the method, but also on the sequence being

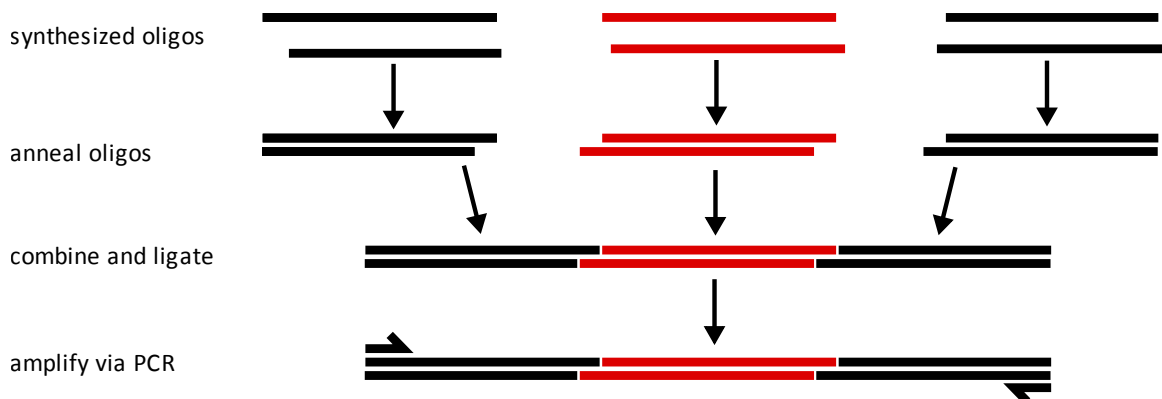


**Figure 2-3 Assembly of 6-zinc fingers into a single protein using restriction enzyme based sub-cloning. By maximizing the steps performed in parallel, the depth of the tree, and therefore the duration of the process, can be minimized. Each box represents a segment of DNA. The boxes containing 3 letters each correspond to one zinc finger and the letters indicate the sequence that finger binds to, not the sequence of the DNA being assembled.**

constructed. Constructs that are toxic, impose a high metabolic burden on the cell, or contain repeated sequence are likely to rapidly mutate away from the desired sequence [28], [29]. Initial construction of the zinc finger proteins was not done with a promoter present and therefore toxicity or metabolic burden were not an issue. Deletions due to recombination were a concern because a 63 base pair segment of DNA was repeated 5 times within each protein. To lower the rate of recombination, all work was done in *recA*-strains, primarily DH10B, DH5 $\alpha$  and BLR(DE3). However *recA* dependent recombination only occurs with repeats of 200 base pairs or more, and thus it is questionable if the *recA*-strains altered the stability of these zinc finger protein constructs [29]. Given the possibility of recombination, a hybrid sequencing strategy was employed. Colony PCR was always utilized and was able to catch most failures. Sequencing was also performed after every ligation step, but downstream cloning continued in parallel with the sequencing. With this process the majority of the sequences were good, and no time was spent waiting for sequencing results. In the rare case when a sequence came back incorrect, no more than a few hours had been wasted on further cloning with the bad DNA.

### Golden Gate

The Golden Gate cloning method allows for the rapid assembly of multiple pieces of DNA in a single pot [30], [31]. Type IIS restriction enzymes, which cut DNA a fixed distance away from their operator site, are essential to this method. Because the cut site does not overlap the operator site, it becomes possible to use one enzyme to generate many different DNA overhangs. Golden Gate cloning requires more design work and is more prone to failure than using BglBricks, but because it is not limited to joining only two pieces of DNA, it can considerably lower the time needed to assemble DNA. With aid of computer



**Figure 2-4 Individual zinc fingers for Golden Gate cloning have been constructed from oligonucleotides. Black oligonucleotides are shared across all zinc fingers that are designed for the same position within the final protein. Red oligonucleotides are shared across all zinc fingers that bind to the same DNA sequence. To use 49 binding triplets in 6 different finger positions requires making  $6 \times 49 = 294$  unique zinc fingers. Given the assembly strategy shown here, only 124 oligonucleoties were needed ( $2 \times 3 \times 6$  black +  $2 \times 49$  red) to construct and amplify all 294 zinc fingers.**



**Table 2-3 Effect of important parameters for Golden Gate cloning.**

Source of Variation	Effect	P-value
Digest duration	40	0.12
Digest temperature	-32	0.20
Ligation duration	23	0.35
Number of cycles	13	0.59

resulting plasmid was slightly smaller than the BsaI digested PCR product containing the vector backbone. The two ends of the digested PCR product did not have compatible overhangs, so circularization should not have occurred. Several of these resultant plasmids with no zinc fingers were sequenced. All of them were smaller than the digested PCR product for the vector backbone, and the missing base pairs were always from one or both of the ends of the digested PCR product. Because the digested PCR product was AT-rich at both ends, the digested PCR product may have frayed at the ends and created an opportunity for annealing and ligation within the frayed regions. The  $T_m$  of the 13 terminal *double stranded* bases on each end of the product were found to be 34°C and 33°C [33]. Because the Golden Gate cloning protocol includes a 37°C incubation step, it is likely that significant fraying of the digested DNA occurs (see protocol outline in Table 2-4).

An experiment was performed to determine if lowering the digestion temperature would result in higher efficiency cloning. It was hoped that lowering the digestion temperature would stop the ends of the digested DNA from fraying and give fewer incorrect ligations. A  $2^{4-1}$  factorial experiment [34] was designed to investigate the effects of digestion temperature, digestion duration, ligation duration, and number of cycles. Digestion was tested at 37°C and 27°C for 5 minutes or 2 minutes. Ligation duration was also tested at 5 minutes and 2 minutes. Number of cycles was tested at 12 and 24. The ligation temperature was uniformly reduced to 22°C, as this is the temperature recommended by Fermentas, the manufacturer of the T4 ligase used. In order to extract as much useful information from this experiment as possible, the objective function used was based on the total number of zinc fingers successfully inserted into the vector across all colonies. This metric could be problematic if some conditions were extremely efficient at inserting some but not all of the 6 fingers, while conditions that inserted all 6 fingers were less efficient. Such results were not observed; the number of colonies roughly correlated with the number of zinc finger successfully inserted.

**Table 2-4 Protocols for one pot Golden Gate cloning.**

	Published protocol	Modified Protocol
BsaI digestion	37°C for 5 minutes	27°C for 5 minutes
Ligation (T4)	16°C for 5 minutes	22°C for 5 minutes
	Repeat above 2 steps 25 times	Repeat above 2 steps 25 times
BsaI digestion	50°C for 5 minutes	50°C for 5 minutes
Enzyme inactivation	80°C for 5 minutes	80°C for 5 minutes

degree of cutting. Due to the low local melting temperature ( $T_m$ ) at the ends of the PCR products, fraying of the PCR product may occur, and this could lead to less cutting than would have occurred if the segment was contained in a plasmid.

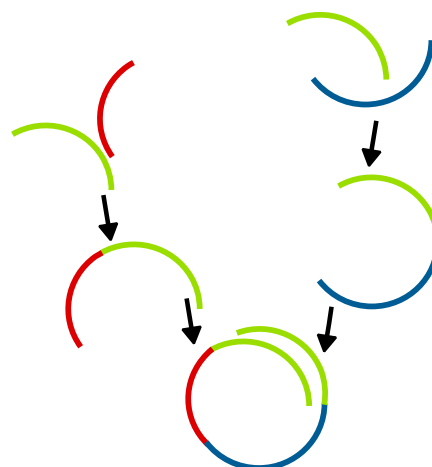
The first attempts at Golden Gate assembly of zinc finger proteins resulted in re-circularized vector backbone with none of the zinc fingers inserted. The size of the



This experiment indicated that digest duration is the most important parameter followed by digest temperature as shown in Table 2-3. This type of experiment only yields the local slope of the parameter space. Therefore it does not give an optimal set of parameters, but instead gives a direction to move towards. The absolute value of the effect indicates the strength of the dependence between the parameter and the objective function, and the sign of the effect indicates the direction the parameter should be moved to increase the objective function. The number of colonies that grew per experimental condition was low; therefore, it is not surprising that none of the p-values fell below the usual threshold of 0.05. Some of the conditions tested still gave more correct colonies than the original set of parameters, indicating these results are worth following up on even if the p-values were over the typical 0.05 threshold. The modified protocol that resulted from this experiment is shown in Table 2-4.

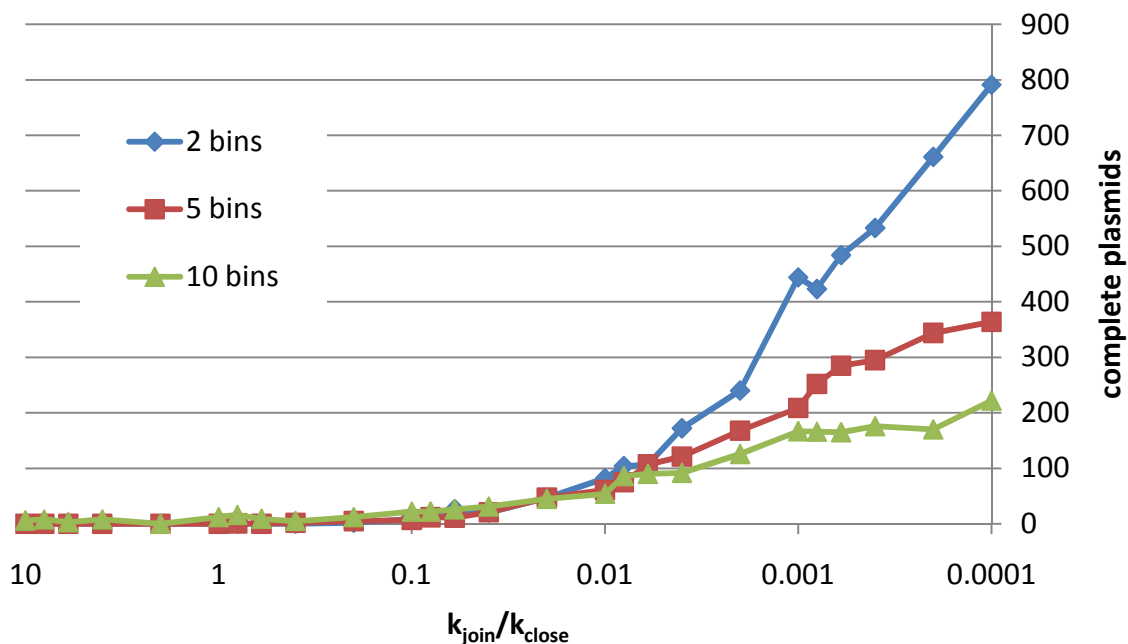
The simultaneous digestion and ligation of Golden Gate cloning along with the formation of irreversible junctions is supposed to drive all of the DNA to the final assembly product; however, it is possible for products to form that cannot contribute to the desired product. A simplistic example of this is shown in Figure 2-5.

Simulations were performed to determine the relation between number of pieces of DNA to assemble and the fraction of DNA segments that result in correct assemblies. Because of the many possible states in a multi-part assembly, analytical analysis methods were difficult to implement. To get past these limitations, a stochastic simulation method was used. Simulations were configured with assemblies of 2, 5 and 10 parts, and all parts present at equal molar concentrations. The primary parameters investigated in were the kinetic rates for two pieces of DNA joining and for a single piece of DNA to circularize and form the desired product. The rate of circularization is independent of the concentration of parts, but the rate of parts joining will increase with concentration; therefore, the ratio of these two parameters can be adjusted by setting the per-part concentration. The simulations were performed using the Stochastic Simulation Compiler v0.6 [35]. 1000 instances of each part were simulated in a single volume (no diffusion) until no further reactions were possible. The results are shown in Figure 2-6. All parts that did not end up in a complete plasmid were in assemblies larger than the desired plasmid.



**Figure 2-5 A three part Golden Gate assembly is shown where each part is in a different color. If parts 1 and 2 ligate together, parts 2 and 3 ligate together, and then these two products ligate to form a 4 part assembly. The resulting product is then larger than the desired product. This 4 part assembly could circularize into a plasmid that is an integer number of times larger than the desired product.**





**Figure 2-6 Stochastic simulations of a Golden Gate assembly showing how the relative rates of intra-molecular ( $k_{\text{join}}$ ) and inter-molecular ( $k_{\text{close}}$ ) reactions influence the number of correct plasmids created. Complete plasmids are out of a possible 1000. All reactions have perfect fidelity. The fraction of correct plasmids increases as the concentrations of parts decrease.**

Adjusting the concentration of parts resulted in the majority of parts going into the desired product for 2 part assemblies, but was not able to achieve similar results for the 5 or 10 part cases. A modified assembly method was investigated to increase the number of correct plasmids formed for the assemblies with more than 2 parts. To prevent the formation of assemblies larger than the desired plasmid, one of the restriction enzyme sites was changed to recognize a second type IIS restriction enzyme but create the same overhang as before. This second enzyme would not be added to the reaction mix until late in the reaction. Simulations showed that over 99% of the parts went into linear products of the desired size before the second restriction enzyme was added. Once the second restriction enzyme is added, both circularization and joining of segments can occur. By performing a dilution when the second restriction enzyme is added, the ratio of  $k_{\text{join}}/k_{\text{close}}$  can be independently set for the linear growth phase and the circularization phase.

**Table 2-5 Sequence of variable regions used in zinc fingers to determine binding specificity.**

<b>Binding Site</b>	<b>Amino Acid Sequence</b>	<b>Nucleotide Sequence used in BglBrick Assemblies</b>	<b>Nucleotide Sequence used in Golden Gate Assemblies</b>
AAA	QRANLRA	CAACGTGCAAATCTGCGTGCT	CAGCGCGCGAACCTGCGCGCG
AAC	DSGNLRV	GACTCTGGTAATCTGCGTGTT	GATAGCGGCAACCTGCGCGTG
AAG	RKDNLKN	CGCAAAGATAACCTGAAAAAC	CGCAAAGATAACCTGAAAAAC
AAT	TTGNLTV	ACTACTGGTAATCTGACCGTT	ACCACCGGCAACCTGACCGTG
ACA	SPADLTR	TCTCCAGCTGATCTGACTCGT	AGCCCGGCGGATCTGACCCGC
ACC	DKKDLTR	GACAAAAAGGACCTGACTCGT	GATAAAAAAGATCTGACCCGC
ACG	RTDTRLR	CGTACTGATACTCTGCGTGAT	CGCACCGATACCTGCGCGAT
ACT	THLDLIR	ACTCATCTGGATCTGATTTCGT	ACCCATCTGGATCTGATTTCG
AGA	QLAHLRA	CAACTGGCTCATCTGCGTGCA	CAGCTGGCGCATCTGCGCGCG
AGC	ERSHLRE	GAGCGTTCTCATCTGCGTGAA	GAACGCAGCCATCTGCGCGAA
AGG	RSDHLTN	CGTCCGATCATCTGACTAAT	CGCAGCGATCATCTGACCAAC
AGT	HRTTLTN	CATCGTACTACTCTGACTAAC	CATCGCACCCCTGACCAAC
ATA	QKSSLIA	CAGAAATCTTCCCTGATTGCC	CAGAAAAGCAGCCTGATTGCG
ATG	RRDELNV	CGTCGTGATGAACTGAATGTT	CGCCGCGATGAACTGAACGTG
ATT	HKNALQN	CATAAAAACGCACTGCAAAAC	CATAAAAACGCGCTGCAGAAC
CAA	QSGNLTE	CAATCTGGTAACCTGACTGAA	CAGAGCGGCAACCTGACCGAA
CAC	SKKALTE	TCTAAGAAAGCGCTGACTGAA	AGCAAAAAGCGCTGACCGAA
CAG	RADNLTE	CGTGACGATAACCTGACTGAA	CGCGCGGATAACCTGACCGAA
CAT	TSGNLTE	ACCTCTGGTAATCTGACTGAA	ACCAGCGGCAACCTGACCGAA
CCA	TSHSLTE	ACCTCTCATTCTCTGACTGAA	ACCAGCCATAGCCTGACCGAA
CCC	SKKHLAE	TCTAAGAAACACCTGGCCGAA	AGCAAAAACATCTGGCCGAA
CCG	RNDTLTE	CGCAATGATACTCTGACTGAA	CGCAACGATACCTGACCGAA
CCT	TKNSLTE	ACTAAAACTCCCTGACCGAA	ACCAAAAACAGCCTGACCGAA
CGA	QSGHLTE	CAATCTGGCCATCTGACTGAA	CAGAGCGGCCATCTGACCGAA
CGC	HTGHLE	CATACTGGTCATCTGCTGGAA	CATACCGGCCATCTGCTGGAA
CGG	RSDKLTE	CGTAGCGATAAACTGACTGAA	CGCAGCGATAAACTGACCGAA
CGT	SRRTCRA	TCTCGTCGTACTTGTCTGCA	AGCCGCCGCACCTGCCGCG
CTA	QNSTLTE	CAAAACTCTACCTGACTGAA	CAGAACAGCACCTGACCGAA
CTG	RNDALTE	CGTAATGATGCACTGACTGAA	CGCAACGATGCGCTGACCGAA
CTT	TTGALTE	ACTACTGGTGCTCTGACTGAA	ACCACCGCGCGCTGACCGAA
GAA	QSSNLVR	CAATCTTCCAATCTGGTTCGT	CAGAGCAGCAACCTGGTGC
GAC	DPGNLVR	GATCCAGGTAATCTGGTTCGT	GATCCGGGCAACCTGGTGC
GAG	RSDNLVR	CGTTCTGATAATCTGGTACGT	CGCAGCGATAACCTGGTGC
GAT	TSGNLVR	ACTTCTGGTAATCTGGTACGT	ACCAGCGGCAACCTGGTGC
GCA	QSGDLRR	CAATCTGGTGATCTGCGTCGT	CAGAGCGGCGATCTGCGCCG
GCC	DCRDLAR	GATTGTCGTGATCTGGCTCGT	GATTGCCGCGATCTGGCGCG
GCG	RSDDLVR	CGTTCTGATGATCTGGTTCGT	CGCAGCGATGATCTGGTGC
GCT	TSGELVR	ACTTCTGGTGAACCTGGTACGT	ACCAGCGGCAACTGGTGC
GGA	QRAHLER	CAGCGTGCTCATCTGGAACGT	CAGCGCGCGCATCTGGAACG
GGC	DPGHLVR	GATCCAGGCCATCTGGTTCGT	GATCCGGGCCATCTGGTGC
GGG	RSDKLVR	CGTTCTGATAAACTGGTTCGT	CGCAGCGATAAACTGGTGC
GGT	TSGHLVR	ACTTCTGGTCATCTGGTACGT	ACCAGCGGCCATCTGGTGC
GTA	QSSSLVR	CAATCTTCTTCTCTGGTTCGT	CAGAGCAGCAGCCTGGTGC
GTC	DPGALVR	GATCCTGGTGCACTGGTTCGT	GATCCGGGCGCGCTGGTGC
GTG	RSDELVR	CGCTCTGATGAACTGGTACGT	CGCAGCGATGAACTGGTGC
GTT	TSGSLVR	ACTTCTGGTTCTCTGGTTCGT	ACCAGCGGCGCCTGGTGC
TAG	REDNLHT	CGTGAAGATAACCTGCATACT	CGCGAAGATAACCTGCATACC
TGA	QAGHLAS	CAAGCTGGTCATCTGGCATCT	CAGGCGGGCCATCTGGCGAGC
TGG	RSDHLTT	CGTTCTGATCATCTGACTACT	CGCAGCGATCATCTGACCAAC

## Chapter 3 - Synthetic Zinc Finger Protein Transcription Factors

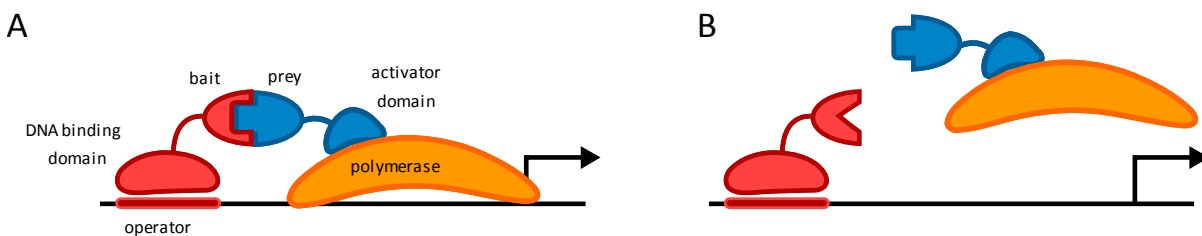
### Existing methods

#### Two-hybrid

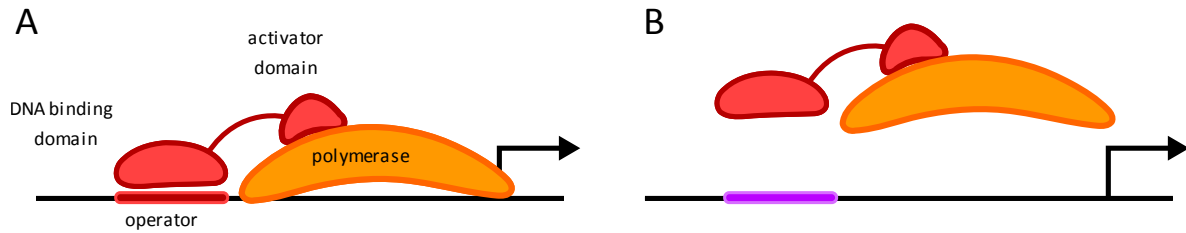
Two-hybrid systems are tools for probing protein-protein interactions *in-vivo*, and they consist of two hybrid proteins and a reporter plasmid. One of the proteins contains a DNA binding domain fused to a protein interaction domain, and the second protein consists of a second protein-protein interaction domain fused to a transcriptional activation domain. An operator binding site corresponding to the DNA binding domain is placed upstream of a promoter with low basal activity. When both of the hybrid proteins are present, the DNA binding domain will bind the operator site, and the activating domain will bind part of an RNA polymerase. If the protein-protein interaction domains of the two hybrid proteins successfully bind, then physical connection between the DNA operator site and the RNA polymerase results in RNA polymerase being recruited to promoter and the onset of transcription. If the protein-protein interaction domains do not bind each other, then RNA polymerase will not be recruited to the promoter, and transcription will not occur.

Two-hybrid systems were first created in *Saccharomyces cerevisiae* [36] and they continue to be used extensively in this organism. It is not surprising that yeast was the first organism to be used in a two-hybrid assay because *S. cerevisiae* is a model organism with many genetic tools, and it has a well studied set of modular domains for DNA binding, transcription regulation, and protein-protein interactions.

Two-hybrid systems are easier to implement in eukaryotic systems than prokaryotic ones because of the increased modularity in eukaryotic systems and because eukaryotic organisms more frequently use regulation based on transcriptional activation. Being able to perform two-hybrid assays in prokaryotic organisms such as *E. coli* would have advantages including faster experiments due to higher growth rate, higher transformation efficiency yielding larger libraries, and a different background of proteins than yeast [37]. Nine years after the first two-hybrid system was published for yeast, an *E.*



**Figure 3-1 A two-hybrid system. A) When a protein-protein interaction exists between the bait and prey domains, then the polymerase is localized to the promoter region downstream of the operator site and transcription occurs. B) When a protein-protein interaction does not exist between the bait and prey domains, then the polymerase is not recruited to the promoter region and transcription does not occur.**



**Figure 3-2 A one-hybrid system. A) If the DNA binding domain is able to bind to the operator site and the activator domain is able to bind an RNA polymerase, then the polymerase will be localized to the promoter region downstream of the operator site and transcription initiation occurs. B) If the DNA binding domain is not able to bind the operator site, then the polymerase is not localized and transcription does not occur.**

*coli* version was published [38]. The duration between these publication dates should serve as an indication of the challenges involved in performing two-hybrid assays in *E. coli*.

### One-hybrid

One-hybrid systems only contain a single hybrid protein that contains a DNA binding domain and an activation domain. These systems are typically used to investigate a protein-DNA interaction, but could also be used to assay for a protein-protein interaction between a protein and an RNA polymerase. One-hybrid assays have been developed in *E. coli* that use zinc finger proteins as the DNA binding domain [39], [40]. These one-hybrid proteins are synthetic transcriptional activators, and it may be possible to make sets of orthogonal transcriptional activators by modifying the zinc finger protein used for the DNA binding domain.

One-hybrid systems also have been developed for eukaryotic hosts, where a wider range of configurations are possible. Eukaryotic organisms have both transcriptional activation and repressions domains. Taking a DNA binding domain fused with an activation domain and swapping it for a repression domain will change the protein from a transcriptional activator to a transcriptional repressor. This swap cannot be done in prokaryotic systems because a modular repression domain has not yet been found.

### Dimeric zinc finger proteins

Many DNA binding proteins form homo-dimers or higher order complexes in order to bind DNA. These structures can result in cooperative binding, where the binding affinity of the first monomer is higher than the binding affinity of the second half of the dimer. Cooperativity yields a system with higher sensitivity to the presence of the proteins. Dimerization also increases the size of the operator site, and therefore decreases the likelihood of the operator site occur elsewhere in the genome by random chance. For example, if a protein has a 9 base pair operator, then this operator would be expected to appear once in a random sequence of  $4^9 \approx 2.6 \times 10^5$  base pairs, but if this protein can dimerize then the operator will likely be unique in a sequence of up to  $4^{18} \approx 6.8 \times 10^{10}$  base pairs. Synthetic zinc finger protein dimers have been created by fusing a leucine zipper

domain to a zinc finger based DNA binding domain [41], [42]. Leucine zippers are a well studied class of protein-protein interaction domains.

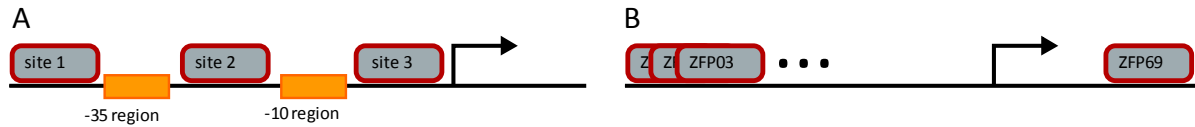
Many prokaryotic transcriptional repressors function as dimers that sterically prevent the RNA polymerase from interacting with a promoter. Because zinc finger proteins are excellent DNA binders and can be made to dimerize, they are good candidates to use in engineering synthetic transcriptional repressors.

Such a system has been constructed, but the level of transcriptional repression obtained was 2-3 fold [43]; and this is significantly lower than that of other commonly used systems where the fold attenuation ranges from 17-5050 fold [44], [45]. Expression of the dimeric zinc finger proteins used by Shetty caused a significant decrease in growth rate, and this was hypothesized to be the result of non-specific binding of the zinc finger domains [43]. When non-specific binding exists, there is a possibility of transcriptional repression occurring at off-target sites. These off-target sites may be expressing genes that are important for maximizing the cell's growth rate in the current environment, and repression of these genes will therefore lead to a reduction in growth rate. Shetty noted that the 14 base pair operator for her dimeric zinc finger protein does not occur in the MG1655 *E. coli* genome, nor does any 1 base pair mismatch. However, it is highly likely that the operator for a monomeric zinc finger does exist within the MG1655 genome, as a 7 base pair operator should appear about every  $4^7 = 16384$  base pairs. The MG1655 genome is more than 4.6 million base pairs, and therefore a 7 base pair sequence would be expected to occur more than 280 times. Alternatively, the toxicity may have been caused by the zinc finger protein being expressed to a concentration much higher than its dissociation constant. The protein needs to be expressed to a concentration above its dissociation constant in order to bind the majority of the operator sites, but as the concentration increases there is a tradeoff between tighter repression and minimizing non-specific binding. Even at concentrations well below the dissociation constant for non-specific binding there can be detrimental effects, as there are millions of non-specific sites and only a few key sites need to be occupied for the growth rate to be impacted.

Building repressors with dimeric zinc finger proteins has only had limited success so far, but they likely have potential that has not yet been completely explored. Further investigation into reducing the toxicity of dimeric zinc finger proteins by lowering expression levels may solve many of the current problems. But even if these problems are addressed, it is unclear how well dimeric zinc finger repressors will scale. Creating an orthogonal set of dimeric zinc finger repressors would require both sets of orthogonal DNA binding domains and orthogonal sets of protein-protein interaction domains. The human genome contains over 900 zinc finger proteins, and while these proteins may not all be orthogonal DNA binders, this provides a rough estimate for the productive limit of interacting zinc finger proteins. The human genome only contains 56 basic leucine zippers, and these zippers have been found to have a low degree of orthogonality [46]. Because of the low orthogonality of leucine zippers, it may be difficult to make more than 10 or 20 dimeric zinc finger proteins with strong orthogonality [47].

### ***Steric hindrance of RNA polymerase***

In prokaryotes there have been several demonstrations of attenuating transcription from a promoter by binding a protein near the promoter, specifically between -74 to +43



**Figure 3-3 A) Previous work on elucidating positional effects in transcriptional attenuation was based on moving an operator site between three main sites that were minimally constrained. Site 1 is directly 5' of the -35 region. Site 2 is between the -35 and -10 regions. Site 3 is 5' of the -10 site. The sites can be moved a few base pairs to get a locally high resolution, but they cannot move into the important -35 and -10 regions, and thus the resolution near these regions is limited to 12-20 base pairs depending on the operator length. B) By using synthetic zinc finger proteins, it may be possible to target a unique zinc finger to every position along the promoter including the -35 and -10 regions, yielding a 1 base pair resolution.**

relative to the transcriptional start site[44], [45], [48]. Despite the demonstrated ease of adding operator sites to existing promoters, there have been no reports of monomeric zinc finger proteins being used to sterically hinder RNA polymerase in prokaryotes. By using monomeric zinc finger proteins, the limitations on orthogonality due to the need for protein-protein interaction domains (leucine zippers) would be eliminated.

Previous work on the engineering of steric hindrance of RNA polymerase has been constrained by operator sites that cannot overlap with important promoter features. Because there was little or no flexibility in the sequence of the operator sites used, operators could only be introduced in non-critical sections of the promoter as shown in Figure 3-3-A. Replacing the -35 or -10 RNA polymerase binding sites with the sequence of an operator site would likely lead to a loss of activity for the promoter; thus, the insertion of operator sites has mainly been limited to 5' of the -35 region, between the -35 and -10 regions, or 3' of the -10 region [44], [45], [48]. Positions that overlap the -35 or -10 regions likely would yield high levels of transcriptional repression, but investigating these positions has not been possible with existing repressors. By using synthetic zinc finger proteins, it becomes possible to make a DNA binding protein that binds over the existing -35 or -10 sequence of a promoter. With this flexibility, new methods for investigating steric hindrance of RNA polymerase become feasible. Previous work on the effect of operator position on steric hindrance has used one repressor protein, and the operator for that protein was moved to several locations within a promoter [44], [45], [48]. With synthetic zinc finger proteins it becomes possible to keep the promoter sequence constant and design multiple zinc finger proteins that are targeted to unique operator sites at different positions along the promoter as shown in Figure 3-3-B. Ideally, every segment of DNA the size of an operator within the promoter would be targeted by a zinc finger protein, so that the position of the DNA binding protein could be investigated with single base pair resolution.

Due to the limitations on zinc finger protein operator sites that were discussed in Chapter 2, it is not possible to take an existing promoter and design zinc finger proteins that can bind at every position along that sequence. These limitations include a need for GNN binding fingers and that the operators only contain the well characterized pallet of 49

DNA triplets. Using bioinformatics, known promoter sequences could be searched for a promoter that contains a large number of high quality zinc finger protein binding sites. Given the data on GNN requirements [24] and that only 49 triplets have good fingers, the expected frequency of usable operator sites using 3 finger or 6 finger proteins respectively are:

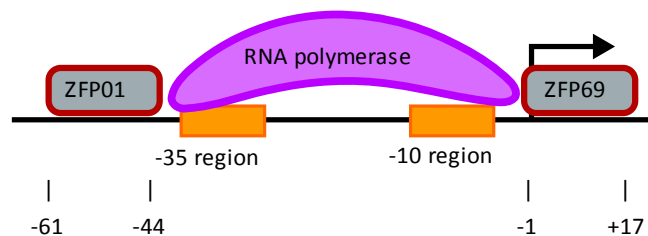
$$\left[ \left( \frac{64-16}{64} \right)^2 \frac{16}{64} 0.12 + \frac{64-16}{64} \left( \frac{16}{64} \right)^2 0.29 + \left( \frac{16}{64} \right)^3 0.59 \right]^{-1} \approx 37.7 \text{ base pairs/site}$$

$$\left[ \left( \frac{64-16}{64} \right)^5 \frac{16}{64} 0.12 + \left( \frac{64-16}{64} \right)^4 \left( \frac{16}{64} \right)^2 0.29 + \left( \frac{49}{64} \right)^3 \left( \frac{16}{64} \right)^3 0.59 \right]^{-1} \approx 153.6 \text{ base pairs/site}$$

For this calculation, 64 is the total number of triplets, 49 in the number of triplets for which good fingers exist, and 16 is the number of GNN fingers. The 6 finger sites most likely occur more often than shown here as the work on GNN requirements only considered 3 finger proteins [24]. Proteins with 4 or more GNN sites probably have a higher than 59% chance of being functional. A search of promoters should yield some sequences that are significantly more enriched for good zinc finger protein operator sites than random sequence. Even with a 10-fold enrichment, the resolution using 3 finger or 6 fingers proteins would be 3.77 base pairs/site and 15.36 base pairs/sites respectively. Achieving single base pair resolution using an existing promoter appears unlikely.

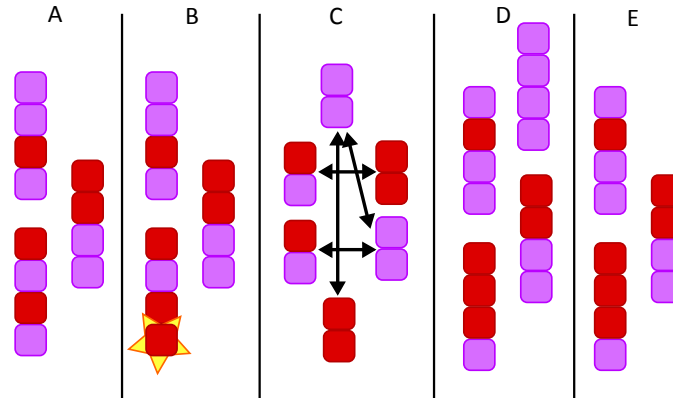
### Design of a synthetic promoter that can bind many zinc finger proteins

Since existing promoters were not expected to yield a high resolution of binding sites, a synthetic promoter was designed with single base pair resolution of high quality binding sites. The first step in the design of this promoter was deciding on the size of zinc finger proteins to use. With 4-finger proteins, about a quarter of all possible sequences would be found in the *E. coli* genome. Because the promoters have many constraints placed on their sequences, the operators needed to cover any promoter are probably over-represented in the genome. Therefore a factor of 4 in available sequence space did not seem adequate. By increasing the number of fingers to 5, the possible sequence space becomes 268 times the size of the *E. coli* genome. While 5 fingers would work well in the



**Figure 3-4 Defining the region to target with zinc finger proteins. The footprint of RNA polymerase extends from about -43 to -2 relative to the transcriptional start site. If zinc finger proteins with 18 base pair operator sites flank the RNA polymerase, then the region including these operator sites spans from -61 to +17. This region has been expanded by 4 base pairs on each end (-65 to +21) to accommodate any difference between the zinc finger protein operator site and footprint size and to give a margin of error.**





**Figure 3-5 Description of a genetic algorithm. Genetic algorithms are useful for searching through a large parameter space that may not be smooth. A) A starting population with diversity is randomly generated. B) Mutations introduce additional diversity by randomly modifying an attribute for some fraction of the population. A yellow star marks the site of a mutation in the figure. C) Sexual reproduction takes the attributes from two members of the population and combines them to form new members. C) The population after the sexual reproduction step. The figure does not show the parents persisting, but in most implementations they are still included in the population at this step. E) The fitness of each member is evaluated through the objective function. The objective function is a quantitative metric and is application specific. Low fitness members of the population are eliminated. In this example fitness is proportional to the number of red squares. Steps B-E are repeated until the maximum fitness remains constant over many iterations. Tuning key parameters for tuning the genetic algorithm are the mutation rate, the rate of sexual reproduction, and the population size.**

context of the *E. coli* genome, constructing a library of 5-finger proteins with restriction enzyme cloning requires the same number of steps as constructing a 6-finger protein library. 6-finger proteins will allow for more GNN fingers per protein, and the 18-finger proteins are large enough that their operators are unlikely to appear in the genome of any highly studied organism.

The next task was to determine how much flanking sequence around the core segment of -35 region through -10 region should be included. Footprinting measurements of *E. coli*'s RNA polymerase have indicated that up to 41 base pairs can be covered by the polymerase [49]. If RNA polymerase is centered on the -35 to -10 regions, then a 41 base pair footprint would extend from about the -43 to -2 positions. To cover all positions that could interfere with this footprint, the most upstream 18 base pair operator needs to extend from -44 to -61 and the most downstream 18 base pair operator needs to extend from -1 to +17 as shown in Figure 3-4. To allow for a margin of error, and the possibility of the zinc finger protein having a footprint that is larger than its operator site, the promoter region to be targeted by zinc finger proteins was defined to go from -65 to +21 relative to the approximate transcriptional start site.

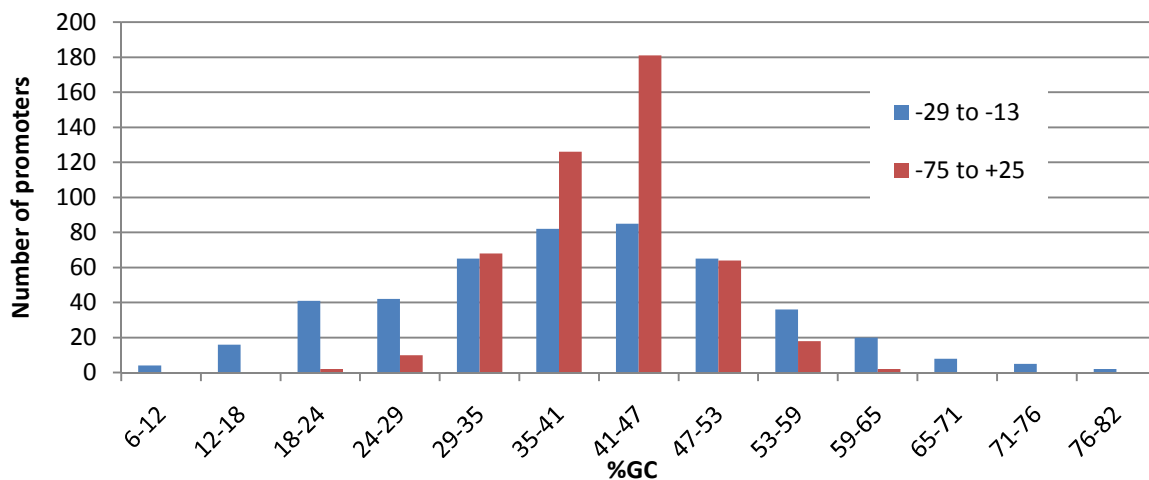


With the promoter region of interest defined to span from -65 to +21, the total possible sequence space is quite large at  $4^{86} \approx 6 \times 10^{51}$  possibilities. Because the sequence space is large, a brute force approach, where every possible sequence is evaluated, is not possible. A genetic algorithm was used to rapidly search through the possible sequence space. For an overview of genetic algorithms see Figure 3-5. Properly defining the objective function is crucial to obtaining a useful result. Because genetic algorithms are conceptually based on evolution, the common phrase from directed evolution applies here also, “you get what you select for.”

### Creating the objective function

The most important constraint encoded in the objective function is to keep the sequence close to that of a consensus promoter, because if the sequence fails to have constitutive promoter activity, then transcription attenuation assays cannot be performed. A published consensus *E. coli* promoter sequence [50] was used to evaluate the potential sequences (sequence shown in Figure 3-8). This consensus sequence was constructed from a subset of *E. coli* promoters and includes both  $\sigma^{70}$  promoters and promoters that use alternative sigma factors. Because the majority of *E. coli* promoters are  $\sigma^{70}$  promoters, this consensus sequence is biased towards the consensus  $\sigma^{70}$  promoter sequence. This set of promoter is also biased by the fact they are studied promoters and therefore may not be a representative population. To evaluate the similarity between the consensus and a potential promoter, the number of matching base pairs between the consensus and potential promoter were counted (Ns in the consensus did not count as matches as they do not specify any reduction in degrees of freedom).

Scoring the number and quality of the zinc finger protein operator sites was based on several metrics. For all of these metrics, every possible 18 base pair sequence of the promoter was evaluated under the metric, and the results for all segments were summed.



**Figure 3-6** The %GC content of 472 *E. coli* promoters is shown in blue for the 17 bases that normally fall between the -35 and -10 regions and in red for the whole promoter defined as from -75 to +25 relative to the transcriptional start site. For both regions the distribution is centered slightly below 50%.

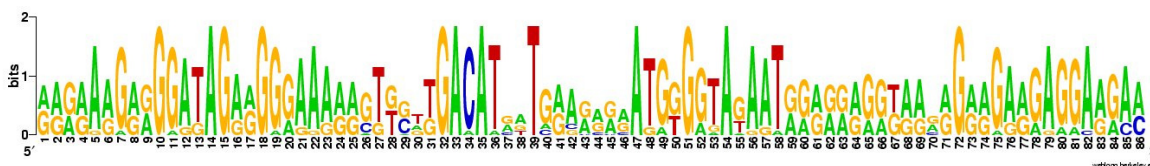
WTWNWNNNNAWWWNNTTCTTGACATNTTNNNNNNWTRTGRTATAATDNNNNNNRRTWW  
 -35 region -10 region

**Figure 3-8 A consensus *E. coli* promoter sequence<sup>47</sup>. This promoter sequence extends from -53 to +4 relative to the transcriptional start site. N can be any base pair. D can be A, C, or G. R can be A or G. W can be A or T.**

The number of triplets that correspond to fingers that can be rationally designed for was counted. The number of GNN triplets was also tallied. Both of these counts were a positive contribution to the objective function. Each operator site was checked to see if it occurred elsewhere in the promoter sequence or to see if it occurred in the MG1655 *E. coli* genome. Checks were also done against the genome to see if changing 1 or 2 base pairs in the operator site would yield a match within the genome. Hits for repetition of sites contribute negatively to the objective function.

While the balance between consensus sequence and coverage with zinc finger proteins was the most important part of the objective function, there were several other minor factors that had to be included to get a useful result. Because the resulting promoter was to be constructed from oligonucleotides, restrictions from oligonucleotide synthesis had to be included in the objective function. Runs of consecutive Gs longer than 5 base pairs were penalized, because these are difficult to synthesize. Sequences tended to be very G-rich because of the GNN constraint, but *E. coli* promoter sequences tend to be slightly AT rich [51]. The GC content of 472 *E. coli* promoters is shown in Figure 3-6. As the center of the GC content distribution is slightly below 50%, the objective function was set to target a GC content of 47% for the whole promoter.

A consequence of enforcing a specific AT content and rewarding GNN triplets were regions of AT at both ends of the promoter. A GNN triplet that was 17 base pairs or more away from the end of the promoter would occur in 6 different operators (each in a different finger position), but a GNN that was at either end of the promoter sequence would only contribute to one operator site. The objective function contains a term based on the number of GNN triplets summed across all operator sites. The genetic algorithm successfully found the optimum for these constraints and placed all the GNN triplets in the interior of the promoter and 17 base pairs of As and Ts at both ends of the promoter. While this solution is correct based on the objective function, the sequences are unlikely to function as promoters due to the high GC content between the -35 and -10 region where the DNA melting for open complex formation starts. Additionally, these sequences will not have uniform coverage with zinc finger protein operator sites, as they are more

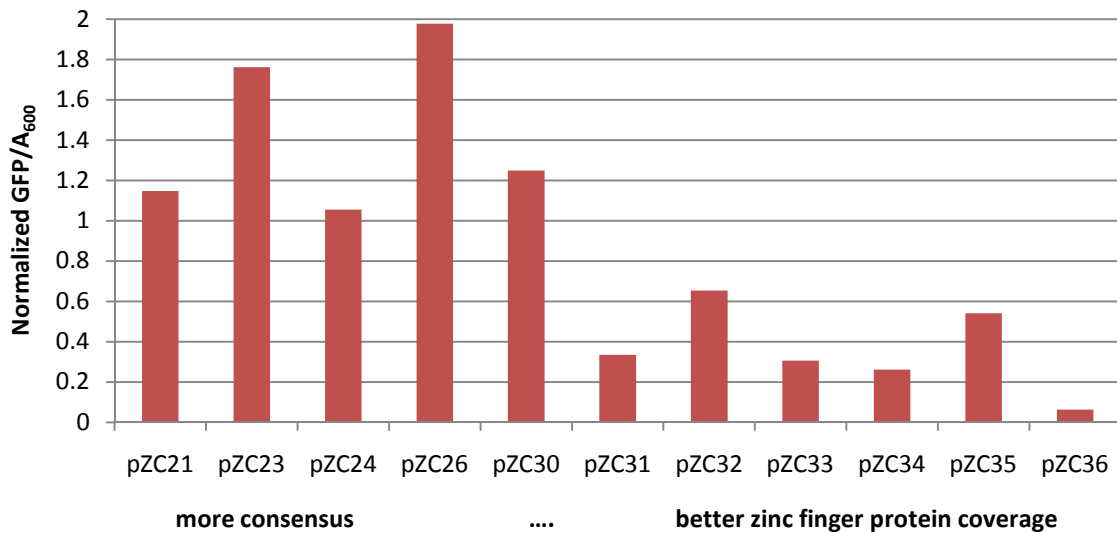


**Figure 3-7 A sequence logo<sup>49</sup> showing the diversity in the promoter library generated by the genetic algorithm.**

concentrated at the center of the sequence. Two approaches were taken to correct these issues. First, a new constraint was added to the objective function that targeted the percentage of GC content at 47% for the segment between the -35 region and -10 region. Second, a new penalty was placed in the objective function that is proportional to the highest percentage of GC content found in any 8 base pair segment of the sequence. Both of these corrections resulted in a more even distribution of the operator sites and a high likelihood of promoter activity.

The objective function used here primarily had to balance between keeping enough consensus promoter sequence that transcription still occurs in the absence of zinc finger proteins and including as many high quality zinc finger protein operator sites as possible. Promoters are generally rich in A and T bases because this lowers the energy needed to locally melt the DNA and allow for transcription, but a sequence optimized for the binding of multiple zinc finger proteins will have an elevated number of G bases. These two requirements for enrichment in both AT and G bases are in conflict and must be balanced. Predicting if a potential promoter sequence will have activity is difficult; thus, there was no good computational way to determine when the balance between consensus promoter sequence and zinc finger protein coverage was optimal.

A hybrid *in-silico/in-vivo* method was used to balance these two factors. The genetic algorithm was run several times with a range of weighting factors used in the objective function to tradeoff consensus promoter sequence versus zinc finger protein coverage. Each of these runs generated a promoter sequence, and this set of promoters was cloned 5' of green fluorescent protein (GFP, version mut3b) and assayed for constitutive promoter



**Figure 3-9 Evaluating constitutive activity of synthetic promoter library. Promoters are listed on the x-axis in rank order of their weighting toward either having sequence more similar to a consensus promoter sequence or having more high quality zinc finger protein operator sites with P<sub>ZC21</sub> being closest to consensus and P<sub>ZC36</sub> having the best set of operator sites. All values were normalized to the strength of P<sub>Ltet0-1</sub> in the absence of TetR.**

AAAAAAGGGGAAAGAGGGGAAAGAGGAGTGGTGGACATAGTGAAGGAATGGAGATAGTGGGAGAGAGGGGAAGGAGAGGAGGAAAAA  
 -35 region -10 region

**Figure 3-10** The sequence of the synthetic promoter  $P_{ZC35}$  is shown with the transcriptional start site (+1) marked by a purple dot. The base pairs shown in red are not in agreement with the *E. coli* consensus sequence given in Figure 3-8. 100% of the non consensus base pairs are Gs.

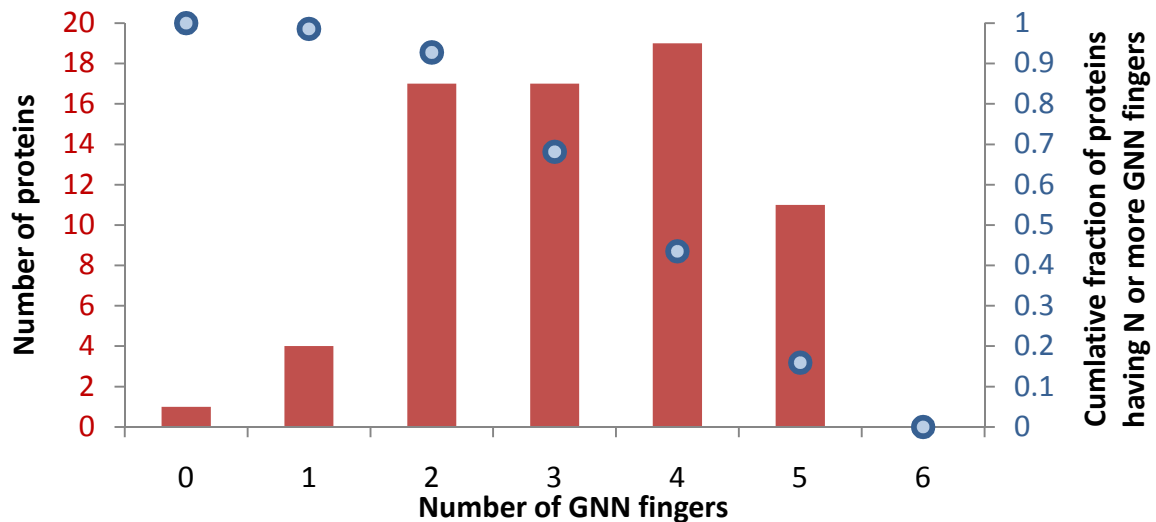
activity (no zinc finger proteins present). This library of candidate promoters has diversity at most positions outside of the -35 and -10 regions as shown in Figure 3-7. Adjusting the balance between the consensus promoter sequence and the coverage of zinc finger proteins caused alterations through out most of the promoter and did not involve just a few base pairs changing. This indicates that the two goals of having a completely consensus sequence and full coverage with zinc finger protein sites are conflicting goals.

### Analysis of the synthetic promoter $P_{ZC35}$

Promoter  $P_{ZC35}$  was chosen for further analysis. Of the promoters that gave enough signal for easy detection,  $P_{ZC35}$  had the highest rating for potential zinc finger protein coverage as shown in Figure 3-9.

The naming convention adopted for the operator sites is opX where X is the position of the most 5' base pair of the operator relative to the transcriptional start site. X is prefaced with a "+" for positions downstream of the transcriptional start site. This means the operator site furthest upstream is op-65 and the operator furthest downstream is op+4.

All DNA triplets contained within  $P_{ZC35}$  are represented in the set of good zinc finger

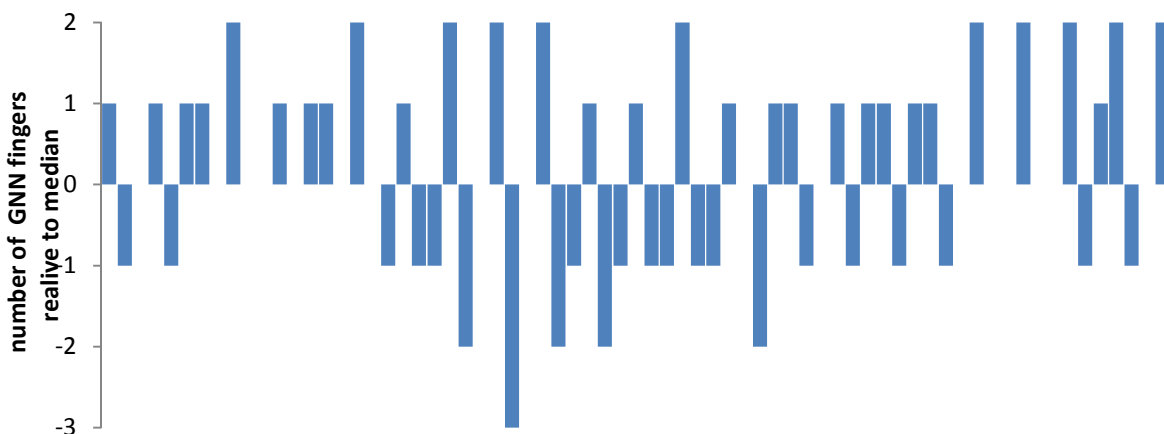


**Figure 3-11** The distribution of GNN binding zinc fingers. Function zinc finger proteins constructed with rational methods are more likely to be function of they contain multiple GNN binding fingers. Over two-thirds of the proteins used in this experiment had 3 or more GNN binding fingers.

for rational design. The number of GNN targeting fingers used within each of the 69 operator sites ranges from 0 to 5, with the majority of operators containing 3 or more GNN sites as shown in Figure 3-11. The probability that a rationally designed zinc finger protein will bind its operator site is dependent on the number of GNN binding fingers in the proteins and has been studied for 3-finger proteins. If having more than 3 GNN binding fingers does not increase the probability of getting a functional zinc finger protein, then about 33 of the 69 proteins should be functional. Alternatively, the trend of additional GNN binding fingers contributing to proteins being function could continue above 3 GNN binding fingers. Extrapolating this trend with a second order polynomial predicts that 45 of the 69 proteins should be functional.

As mentioned previously, the objective function used to create  $P_{ZC35}$  initially was biasing the ends of the promoter sequence to have few GNN binding fingers, but additional constraints were added to mitigate this bias. To see if the quality of the operator sites near the ends of the sequence are worse, the difference from the median number of GNN binding fingers was plotted for each position as shown in Figure 3-12. The operator sites near the ends of the promoter do not appear to be systematically lower in quality than the operator sites for other regions of the promoter.

Of the 69 operator sites none of their sequences are found in the MG1655 *E. coli* genome and if any one base pair of any operator site is changed, the modified site will still not be found in the MG1655 *E. coli* genome. If any two base pairs of any operator are changed, then for only two of the operators (op-59 and op-28) will the modified sites be found in the MG1655 *E. coli* genome. Op-59's closest matching site on the genome is in a non coding region about 30 base pairs downstream of *tsgA* and about 200 base pairs upstream of *nirB*. If a zinc finger protein were to bind at this site, the transcription of *tsgA*



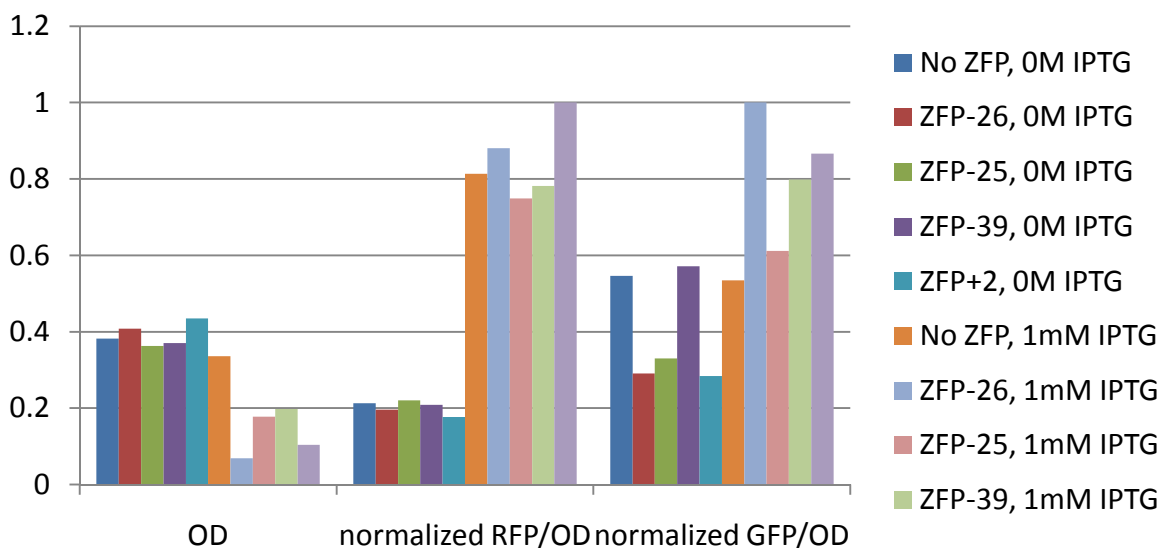
**Figure 3-12 Quality of operator sites as a function of position. The median number of GNN binding zinc fingers this set of proteins was 3. The difference from that median value is plotted here. The left most bar is for op-65 and the right most bar is for op+4. The ends of the sequence did not end up biased to low quality operator sites. The distribution of quality appears even except for between op-60 and op-48 where no operator sites fall below the median number of GNN binding fingers.**

should not be affected, but nirB transcription could be attenuated. Op-28's closest matching site on the genome is in the coding region of hcaR and about 225 base pairs upstream from hcaT. It is unclear if the binding of a zinc finger protein in the middle of gene can reduce transcription, so one or both of these genes could be affected by a protein binding at this site. Besides having limited hits in the genome, none of the 69 operator sites are with two base pair changes of any of the other 69 operators in P<sub>ZC35</sub>.

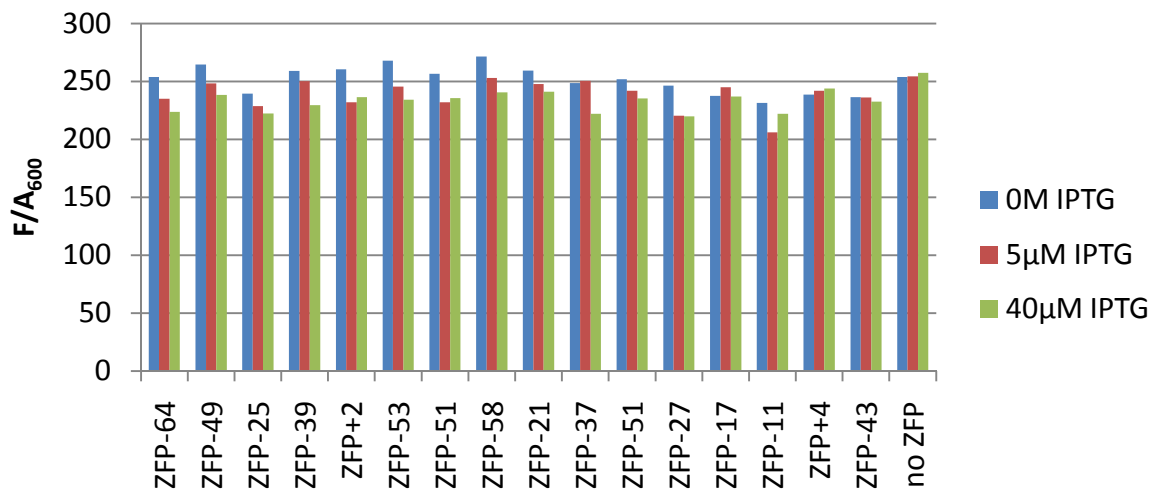
### Assay design for transcriptional attenuation experiment

With the P<sub>ZC35</sub> promoter selected, the remainder of the attenuation assay could be built. The 69 zinc finger proteins were assembled using BglBrick based restriction cloning and placed on the plasmid pBbA5c-RFP. This plasmid has a p15A origin of replication, chloramphenicol resistance, constitutive expression of LacI, and the P<sub>lacUV5</sub> promoter. A protein fusion was made between the zinc finger proteins and the red fluorescent protein (RFP), with the RFP on the C-terminal end of the protein. This fusion made it possible to easily verify that the induction system was working and that both transcription and translation of the zinc finger protein was occurring. A second plasmid contained the reporter operon of P<sub>ZC35</sub> driving GFPmut3b along with a Cole1 origin of replication and kanamycin resistance. The reporter plasmid along with one zinc finger protein-RFP expression plasmid was transformed into DH10B.

To perform attenuation assays, 5mL overnight cultures were grown in LB medium, and then diluted by 20:1 into fresh M9 medium containing 1% glucose to a final volume of 155  $\mu$ L in a 96 well plate. Inducer was added at the time of dilution into M9 medium. After growing to stationary phase at 37C, the absorbance at 600nm (A<sub>600</sub>), RFP, and GFP were



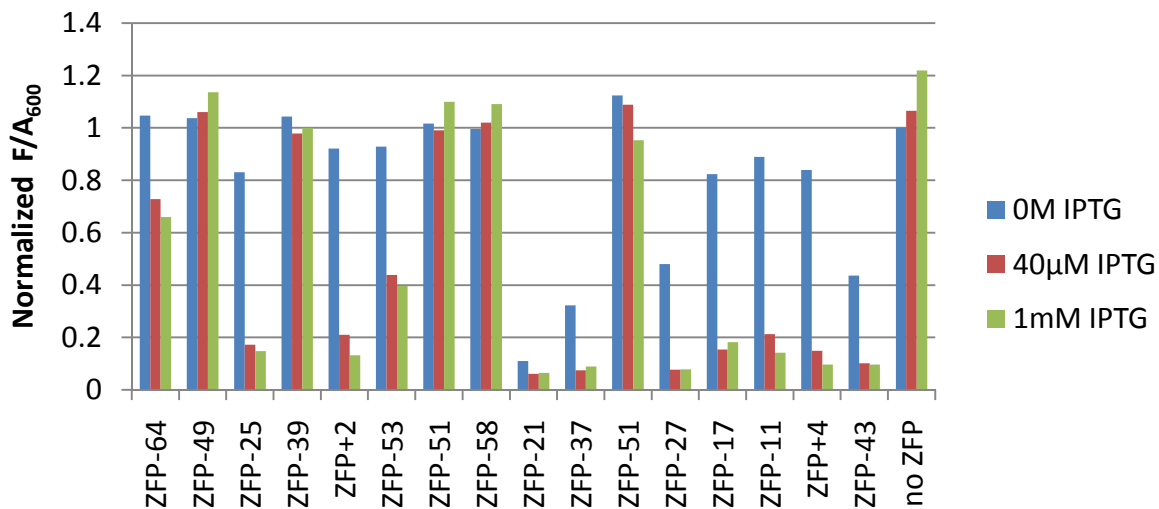
**Figure 3-13 Initial experiment with ZFP-RFP fusion repressing GFP expressing operon. RFP/OD increased with induction indicating the ZFP-RFP fusions were expressed, but ZFP-RFP induction also lowered the OD and failed to repress the GFP expressing operon.**



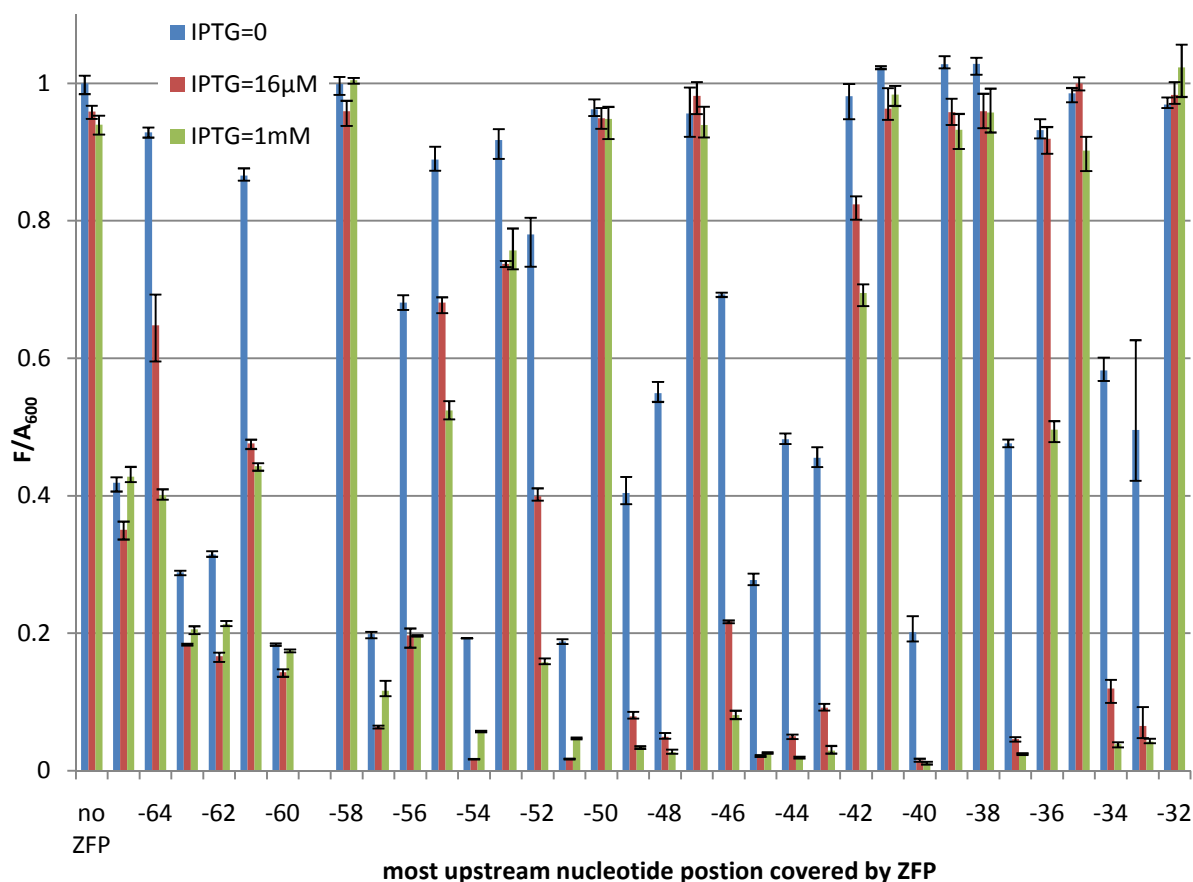
**Figure 3-14 Early experiment with ZFP no longer fused to RFP. No growth inhibition was detected with these lower induction levels.**

measured on a microplate reader.

Initial experiments with 1mM IPTG inductions showed that the zinc finger protein-RFP was expressed, because induction increased the RFP/ $A_{600}$  over 3-fold relative to the un-induced controls (Figure 3-13). However the expression of GFP on the reporter plasmid was not attenuated by the presence of the zinc finger protein-RFP fusion, and the cultures expressing the zinc finger protein-RFP fusion did not reach the same density as the un-induced cultures. The zinc finger protein-RFP fusions could have failed to fold correctly because they were in the context of a fusion, or they might have folded correctly but the



**Figure 3-15 Demonstration of steric hindrance assay. Using mCherry as the reporter. All values have been normalized to the no ZFP 0M IPTG value.**

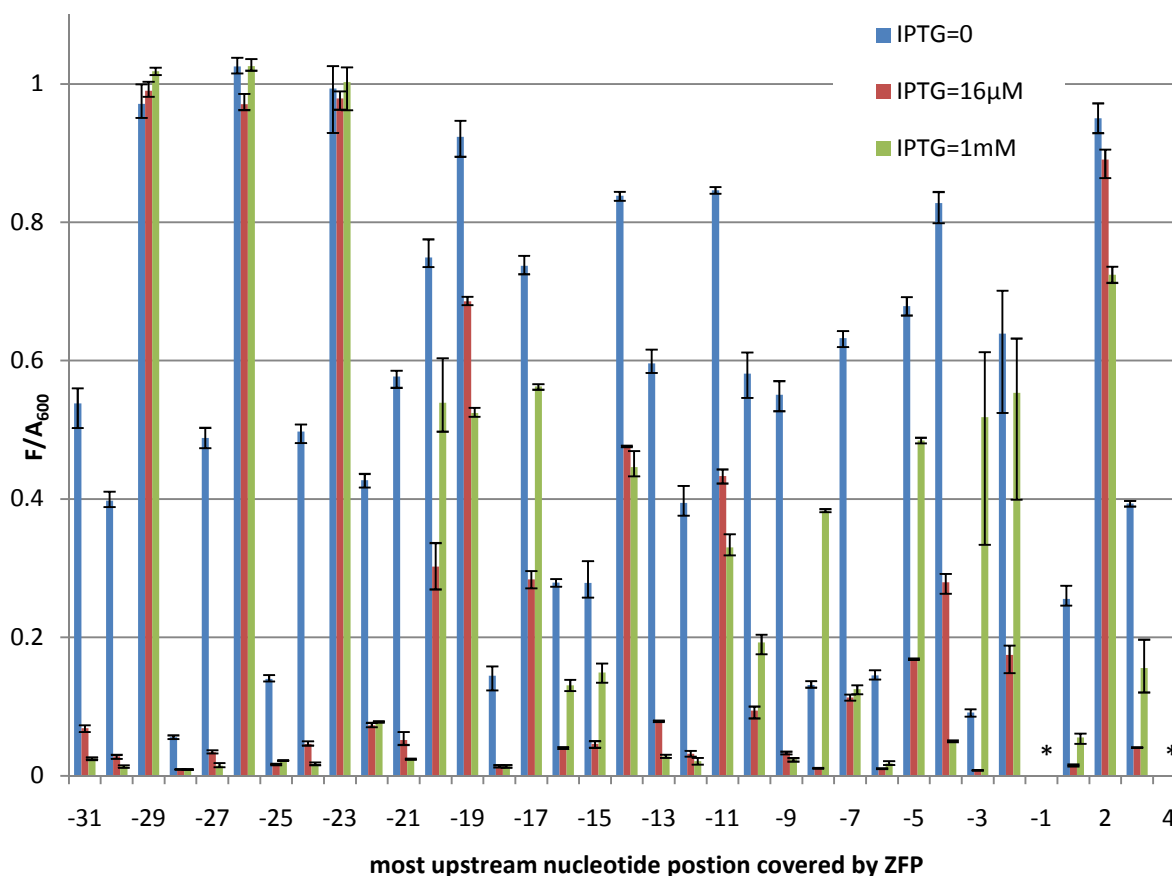


**Figure 3-16** Transcriptional attenuation from the binding of zinc finger proteins to the promoter  $P_{ZC35}$  at various positions. The  $RFP/A_{600}$  values have been normalized to the un-induced (IPTG=0) no zinc finger protein control. Error bars are the minimum and maximum of 3 biological replicates. The positions listed on the x-axis are relative to the approximate transcriptional start site. No data was measured for position -59. For this data set the zinc finger proteins were expressed by a  $P_{lacUV5}$  promoter on plasmids containing the p15A origin of replication. Data for positions -32 through 4 is shown in the next figure.

zinc finger protein domain was prevented from binding to DNA because of the attached RFP. Since the expression of the RFP had already shown that the zinc finger proteins were successfully being transcribed and translated, the RFP was no longer needed. Thus RFP was removed from the zinc finger protein to eliminate the possibility that the fusion protein was causing the zinc finger protein domain to fail to bind the DNA. The slow growth of the zinc finger protein-RFP expressing cells may indicate the fusions are not folding properly or that the proteins are toxic.

The follow up experiment, without RFP, was performed with a range of sub-saturating inductions. Because the cultures grew slowly on M9 medium, the M9 medium was dropped in favor of EZ-Rich medium (Teknova). This medium has a low auto-fluorescence like M9, but also has amino acids and trace metals. All steric hindrance assays

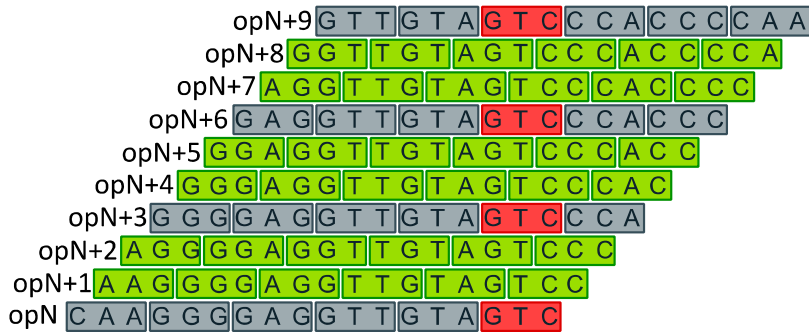




**Figure 3-17** Transcriptional attenuation from the binding of zinc finger proteins to the promoter  $P_{ZC35}$  at various positions. The  $RFP/A_{600}$  values have been normalized to the un-induced (IPTG=0) no zinc finger protein control. Error bars are the minimum and maximum of 3 biological replicates. The positions listed on the x-axis are relative to the approximate transcriptional start site. No data was measured for positions -1 and 4. For this data set the zinc finger proteins were expressed by a  $P_{lacUV5}$  promoter on plasmids containing the p15A origin of replication. Data for the negative control (no ZFP) and positions -65 through -32 is shown in the previous figure.

done from this experiment on were done with EZ-Rich medium. The reduced induction levels did remove the growth inhibition, but the attenuation of  $GFP/A_{600}$  was only 10% or less (Figure 3-14). Unfortunately there was not a control without a GFP plasmid to reference as a minimum signal level.

Because *E. coli* fluoresces in the same range as GFP, GFP is not a good choice for measuring low signal levels. Therefore the GFP on the reporter plasmid was replaced with mCherry, a RFP variant that has been codon optimized for *E. coli*[52] and has less overlap with the fluorescence of *E. coli*. An experiment done with both saturating (1mM) and non-saturating (40μM) IPTG concentrations gave promising results. Saturating inductions of 5 of the 16 zinc finger proteins resulted in attenuations of over 90% (Figure 3-15).



**Figure 3-18 Zinc finger proteins that fail to attenuate transcription are often separated by multiples of 3. The set of 6 fingers that make up a single zinc finger protein are shown on each line and are labeled with the DNA triplet each finger targets. Adjacent rows contain operator sites that differ in position by one base pair, and the labels on the left give their relative positions. Proteins that only contain fingers with functional sets of fingers are shown in green, and proteins that contain a non-functional set of fingers are shown in grey with the non-binding finger in red. In every third row the non-binding finger occurs.**

Additionally, the growth rate in exponential phase was similar for all cultures including the negative control, thus demonstrating no toxicity exists from the expression of zinc finger proteins at these levels.

### Positional dependence in transcriptional attenuation experiment

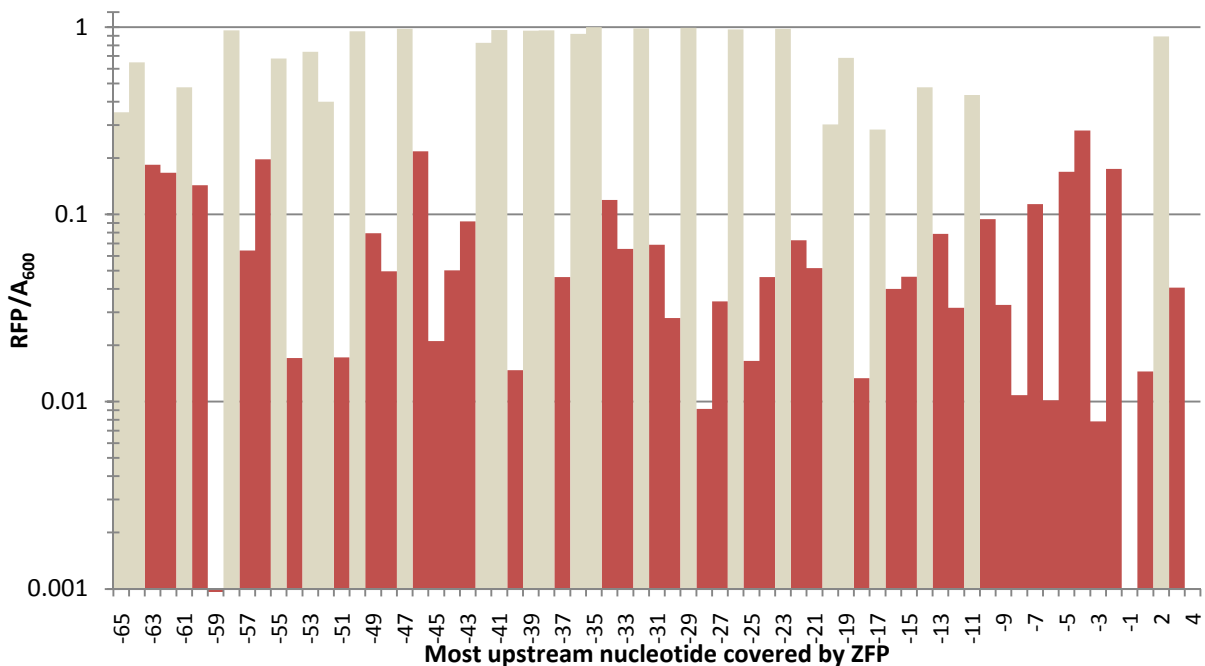
At this point the transcription attenuation assay appeared to be working well enough to run an experiment with the full set of 69 zinc finger proteins done in biological triplicate. Three colonies were picked from each zinc finger protein transformant into 800µL of LB media in a deep 96 well plate. The plate was sealed with Aeroseal breathable seals (no. B-100, Excel Scientific) and incubated at 37°C and 300 RMP overnight. Each well of the LB plates were diluted 20:1 into new deep well plates containing 800µL of EZ-Rich media and either 0, 40µM, or 1mM of IPTG. After 24 hours 150 µL of culture was transferred to black walled, clear bottom 96 well plates for reading.

Interpretation of the data from this experiment involved some subtleties and resulted in additional questions being posed. The transcriptional attenuation data is split across Figure 3-16 and Figure 3-17. This data shows that over 100 fold attenuation of transcription is possible from zinc finger protein binding (position -3 for 16µM IPTG, position -28 for 16µM or 1mM IPTG). At many positions the leaky expression of the zinc finger protein from the P<sub>lacUV5</sub> promoter was enough to cause the reporter to be attenuated by 50% or more, indicating that these zinc finger proteins likely have small dissociation constants. Several of the zinc finger proteins failed to give substantial attenuation at any expression level. These proteins either were targeted to insensitive locations on the promoter or these zinc finger proteins failed to bind their intended targets with a strong enough interaction.

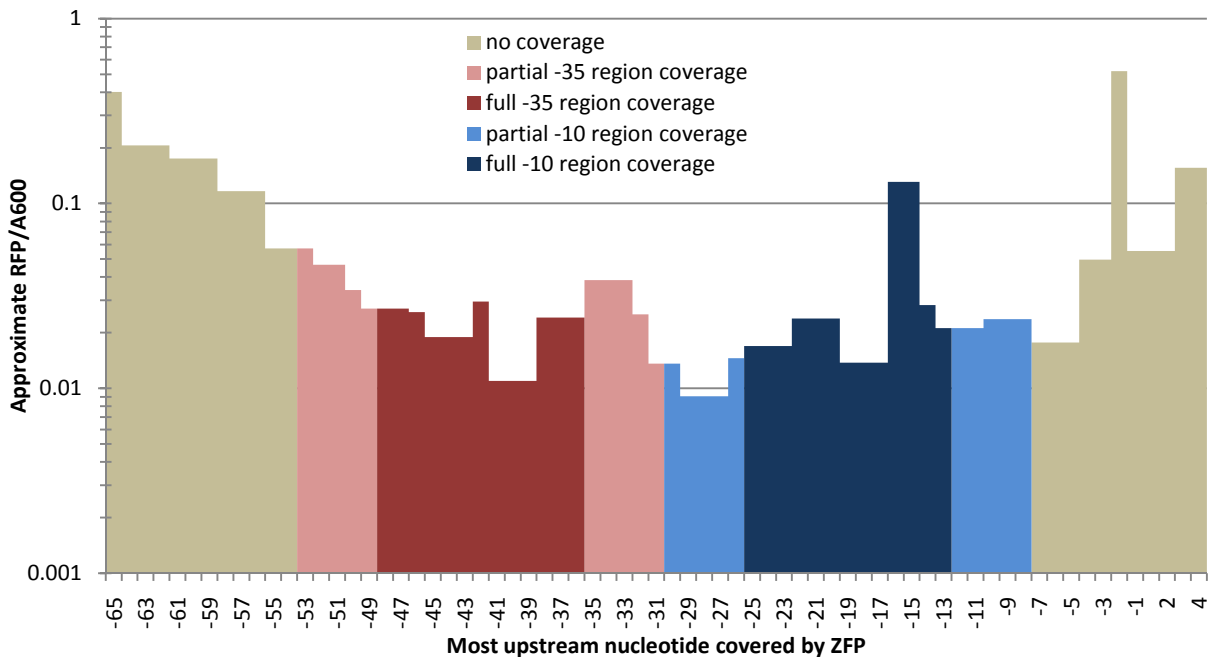
Many of the zinc finger proteins that failed to attenuate transcription have operators separated by a multiple of 3 base pairs. This is not surprising given how rationally designed

zinc finger proteins are constructed. If a zinc finger is used in a context that renders it a poor binder, then there will be between 3 and 5 other proteins that also use the offending zinc finger protein in the problematic context (see Figure 3-18). The exact number of proteins affected will depend on the amount of context required for the finger to fail. If both flanking fingers are required for the detrimental context, then only 3 proteins will contain the stretch of three fingers necessary, but if only one neighboring finger contributes to the context, then there will be 5 proteins that have the problematic pair of fingers.

Several of the zinc finger proteins failed to attenuate transcription by more than 20% when fully induced and are separated by multiples of 3. Specifically the set op-41, op-38, op-35, op-32, op-29, op-26, and op-23 all gave minimal attenuation but one or more of their neighboring sites had over 90% attenuation. The strong repression at neighboring sites makes it unlikely that these minimally attenuating zinc finger proteins failed to cause more attenuation because proteins tightly bound at their positions do not sterically hinder RNA polymerase. Instead, the more likely scenario is that this set of proteins has relatively high dissociation constants for their intended operator sites and this weak binding was not able to hinder RNA polymerase. With the data presented here, it is not possible to prove this hypothesis, but it should be possible to make draw stronger conclusions once



**Figure 3-19** Transcriptional attenuation as a function of operator sites position relative to the approximate transcriptional start site with weak induction (16µM IPTG). Success rates for rationally designed zinc finger proteins indicate that 23 of the 66 zinc finger proteins assayed here should fail to bind their target operator site. Therefore the data for the 23 worst performing zinc finger proteins is shown in grey. The contour from the remaining red data points shows the attenuation as a function of position for the zinc finger proteins that are likely functional. Values were not measured for positions -59, -1, and 4.

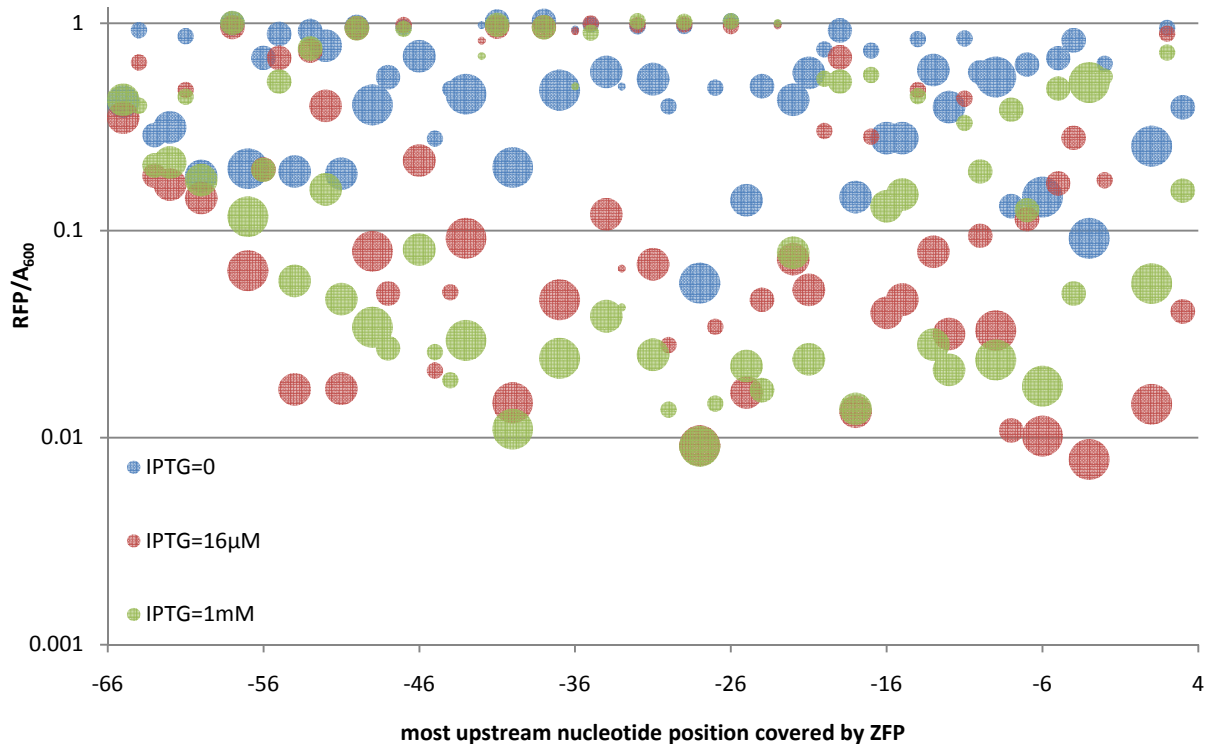


**Figure 3-20 A lower resolution approximation of the attenuation as a function of position for saturating induction (IPTG=1mM). The value plotted at each position is the minimum of the value measured for that position and the two neighboring positions. The colors indicate if a zinc finger protein bound at that position would cover all 6 base pairs (full coverage) of either the -10 or -35 region of the promoter or 1-5 base pairs (partial coverage) of the promoter.**

disocation constants have been measured.

If these poor attenuators are an artifact of non-functional zinc finger proteins, then they should not be considered when determining the amount of attenuation possible at a position due to the binding of a functional zinc finger protein. Earlier in this chapter it was estimated that 24 of the 69 zinc finger proteins that target  $P_{ZC35}$  would fail to bind their intended operator sites. Only 66 of the zinc finger proteins were used in this experiment, and thus the number of failures should scale to 23. The worst performing set of 23 zinc finger proteins can be found by using a threshold of less than 43% attenuation. Evaluating the results of the experiment minus the data from these 23 low attenuators will likely give a more informative view of how attenuation varies with operator position. A plot with these 23 low attenuators differentially colored is shown in Figure 3-19.

The local attenuation maxima (positions -51, -40, -28, -18) in this plot have some periodicity with an average spacing of 11 base pairs. The B-DNA geometry completes a turn in 10.5 base pairs and the Z-DNA geometry has 12 base pairs to a turn. B-DNA is the form most commonly found in cells, but the Z-DNA form occurs in DNA with alternating purine-pyrimidine sequences[53]. A short segment of alternative purine-pyrimidine sequence occurs in the  $P_{ZC35}$  promoter with the sequence ACATA from positions -33 to -29. Of the 3 periodic segments listed above, only the 12 base pair segment (positions -40 to -28) contains a possible Z-DNA region. The remaining two segments have an average spacing of



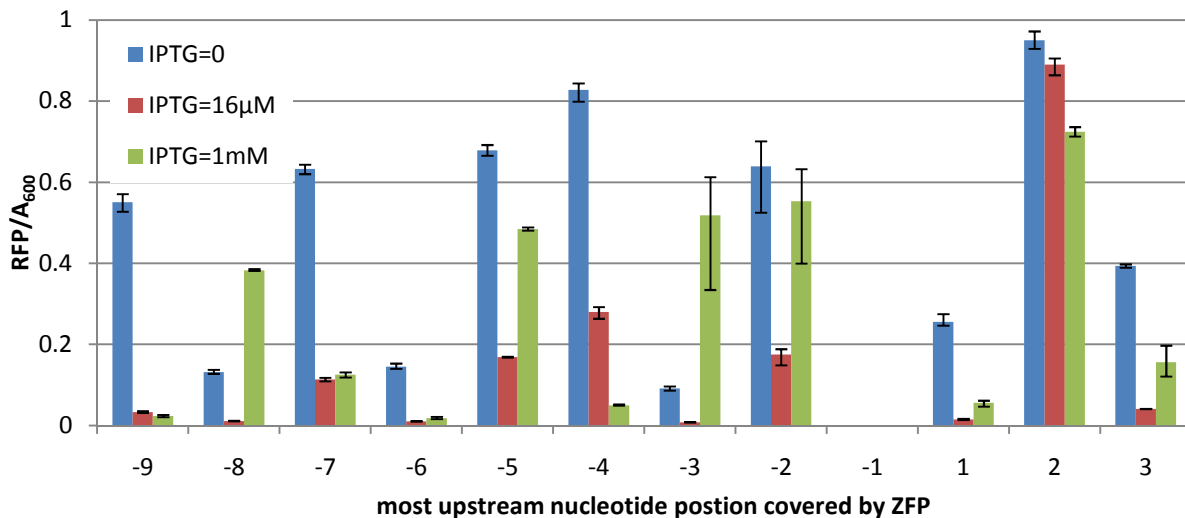
**Figure 3-21** Transcriptional attenuation as a function of both position and number of GNN binding fingers in a zinc finger protein. The width of the bubble is proportional to the number of GNN binding fingers contained in each zinc finger protein. Many, but not all, of the weak attenuators have a low number of GNN fingers.

10.5 base pairs. A similar pattern, of approximately 10 base pairs extending from -54 to -6, has been observed in DNA footprinting analysis of RNA polymerase[54].

When the position of these local attenuation maxima are expressed as the most upstream base pair covered by the zinc finger protein, they occur at positions -34, -23, -11 and -1. Two of these maxima are near the center of the important -35 and -10 regions where contact between the  $\sigma$  factor and DNA occur. While it is not surprising the local maxima are spaced out by the length of a DNA twist, it is unexpected that the local maxima would have partial coverage of the -35 or -10 region. Maximum attenuation was expected to happen when the zinc finger protein was centered on the -35 or -10 region. This result may have to do with the kinetics of RNA polymerase scanning of the DNA and a potential directional bias, but this is well beyond the scope of this study.

An alternative way of thinking about which of the data points are likely to be non-functional zinc finger proteins is to consider how many GNN fingers each protein contains. In Figure 3-21 the average attenuation data for each position is shown in a bubble plot where the size of the width of the bubble is proportional to the number of GNN fingers.

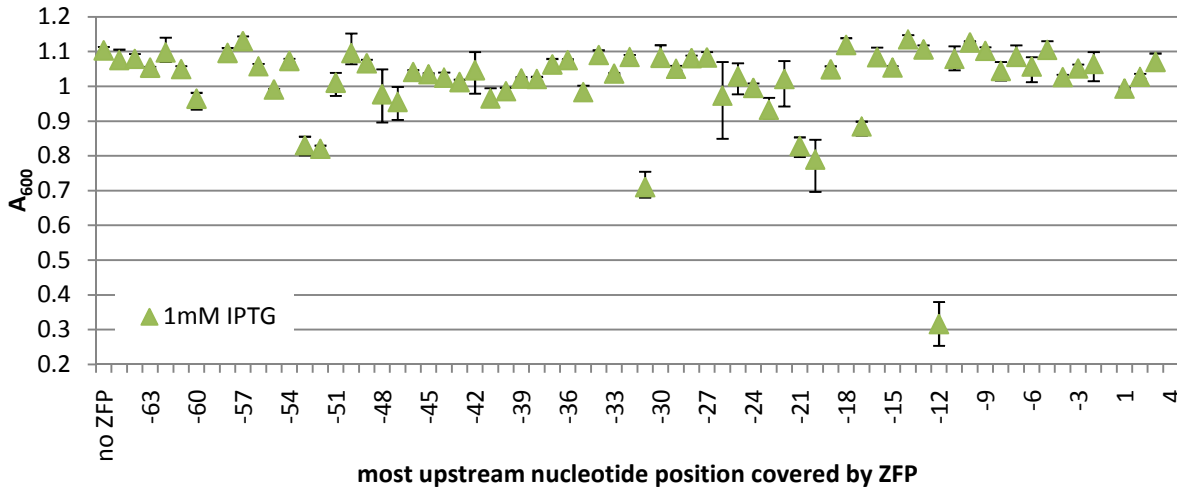
Because of the non-functional zinc finger proteins, this experiment cannot give single base pair resolution of the attenuation as a function of position; however, an



**Figure 3-22 A portion of the transcriptional attenuation data showing the attenuation is not monotonic with induction when the zinc finger proteins are expressed from a  $P_{lacUV5}/p15A$  plasmid.**

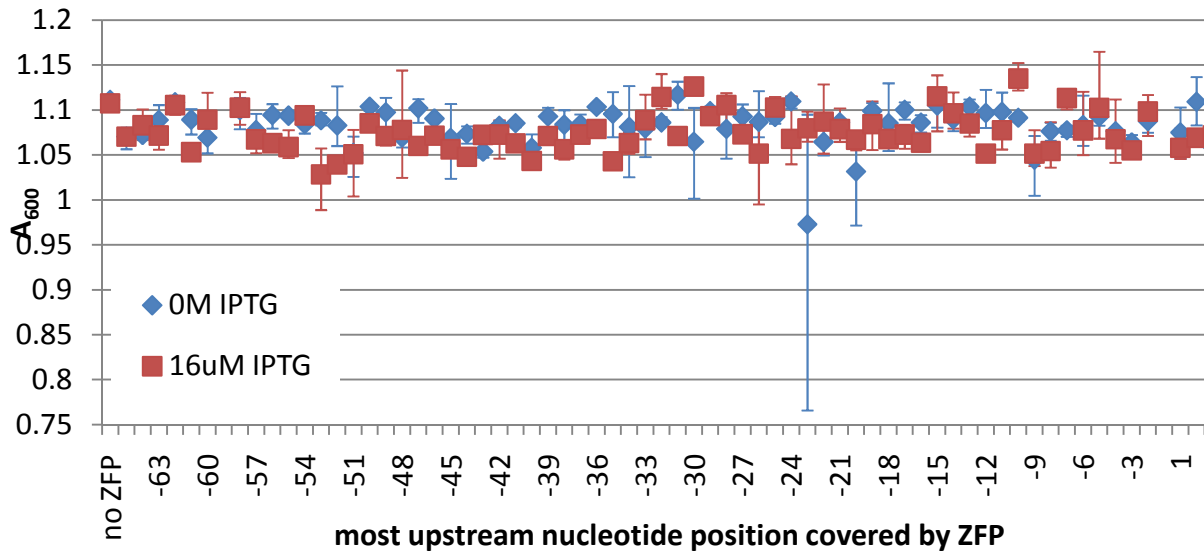
approximation of a lower resolution is possible. Neighboring positions will likely have similar attenuation levels when bound with zinc finger proteins having equal dissociation constants. There are possible discontinuities in the attenuation function. For example, at the transition from the zinc finger protein binding just outside of the -35 region to having 1 base pair of the -35 region covered by the zinc finger protein. Even at these transition points, the protein footprint will probably extend beyond the operator site and the energy required to displace this protein overhang will likely drop off with distance from the operator site. Therefore, even at this very binary transition from covering the operator to partially covering the operator, there can be a gradient response spread over several base pair positions. To approximate a lower resolution attenuation function, the value for a position was set to the minimum of the values measured for that position or either of the neighboring positions. The reduced resolution data is shown in Figure 3-20. This approximation may overstate the attenuation possible at some positions, but if the actual attenuation possible at each position does not have large discontinuities, then this approximation, while not inherently conservative, should be at least be close to the actual value.

Some of the zinc finger proteins displayed a non-monotonic behavior where attenuation of the reporter did not always increase with increasing induction of the zinc finger protein. This is most prevalent from positions -20 to 3 and a portion of region is shown in Figure 3-22. Some of these non-monotonic zinc finger proteins appear toxic at high expression levels, but the two lowest  $A_{600}$  values are from zinc finger proteins with monotonic responses. The source of this non-monotonic behavior for the zinc finger proteins that do not display toxicity is unclear. Because only final  $A_{600}$  measurements were done and not growth rate measurements, it is possible that the non-monotonic zinc finger proteins do inhibit the growth rate but not the final cell density.

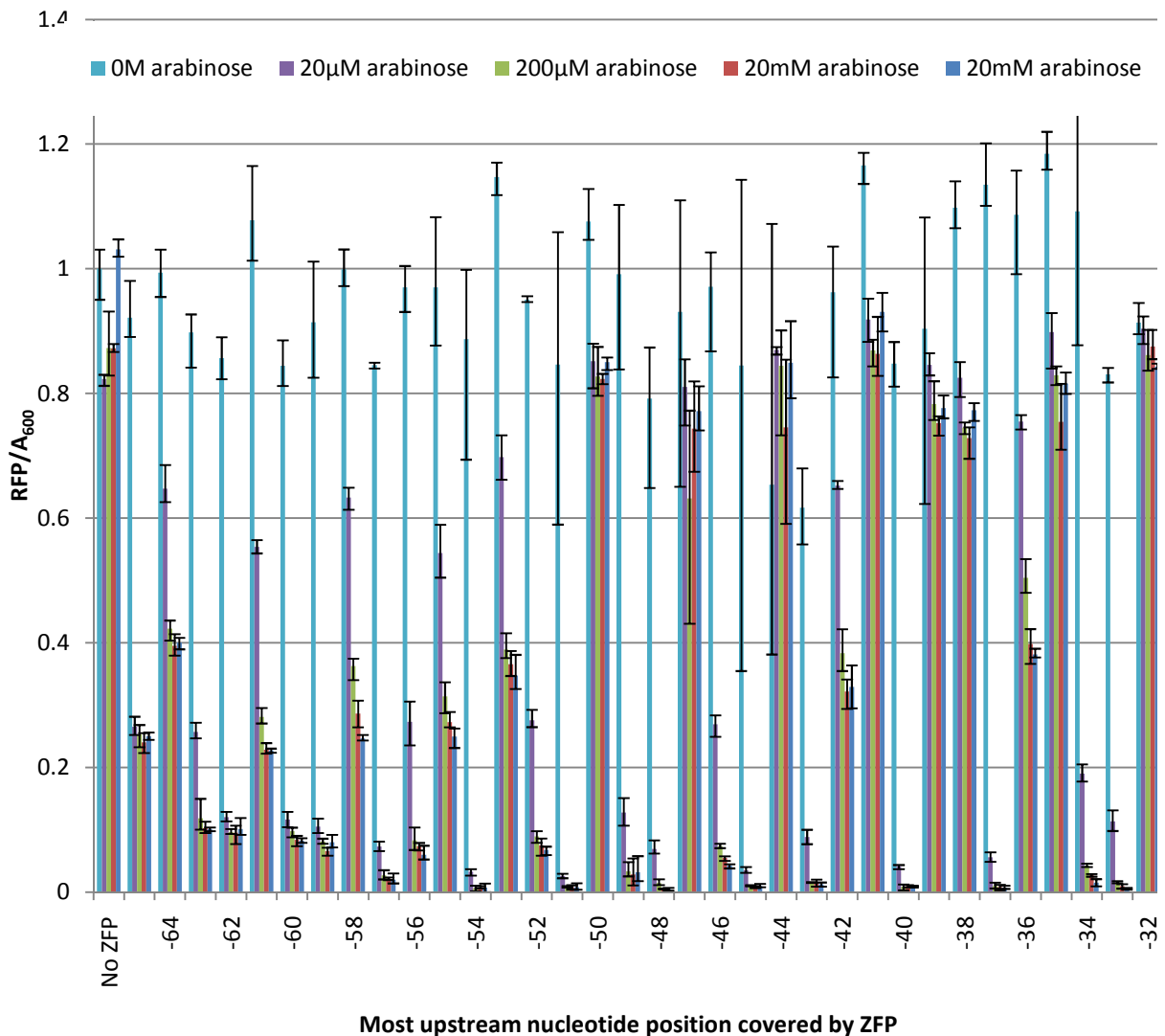


**Figure 3-23 Final absorbance values with saturating (1mM IPTG) induction in a transcriptional attenuation experiment. Several of the zinc finger proteins appear to be toxic at high induction levels.**

The combination of the toxicity issue combined with the high leakage of the P<sub>lacUV5</sub> promoter motivated a move to different zinc finger protein expression plasmid. By moving from a P<sub>lacUV5</sub> promoter on a p15A plasmid to a P<sub>BAD</sub> promoter on an SC101\*\* plasmid, it



**Figure 3-24 Final absorbance values for low inductions in transcriptional attenuation experiment. At these induction levels, there does not appear to be strong toxicity, although there may be a reduction in growth rate due to weak toxicity that would not be captured by this data. The low error bar for position -23 is only due to one data point and is likely not reproducible. The error bars are the maximum and minimum value for 3 biological replicates.**

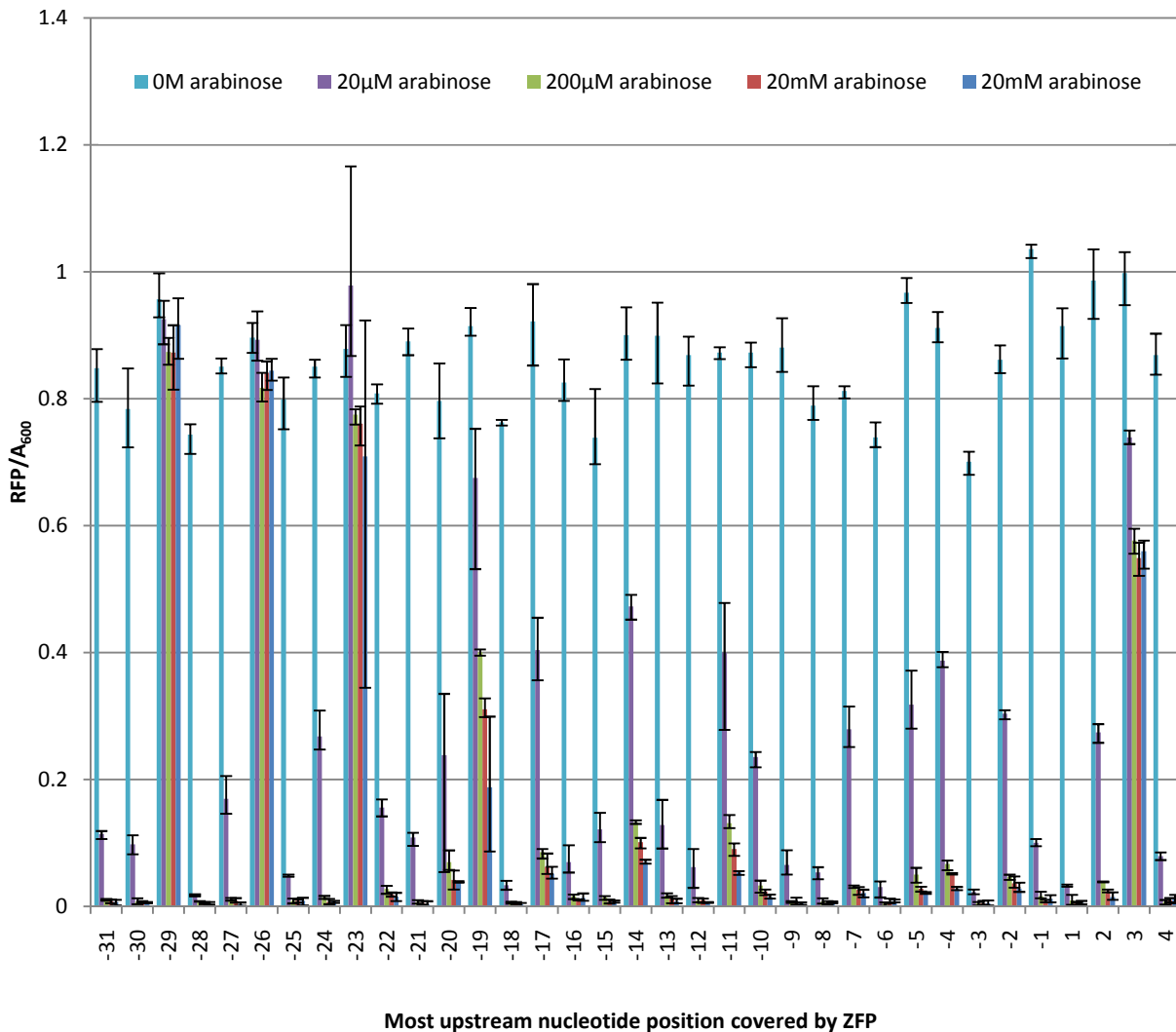


**Figure 3-25 Transcriptional attenuation measurements with the zinc finger protein expressed from a P<sub>BAD</sub>/SC101\*\* plasmid. The other half of this data set can be found in the next figure. Error bars are min/max of 3 biological replicates.**

was hoped that no toxicity would be apparent, the reporter would be sensitive to a larger portion of the induction range, and that the non-induced state would be more similar to the negative control where no zinc finger protein is present. The host strain for P<sub>BAD</sub> based experiments was DP10[55]. This strain has been modified to improve the linearity P<sub>BAD</sub>'s induction.

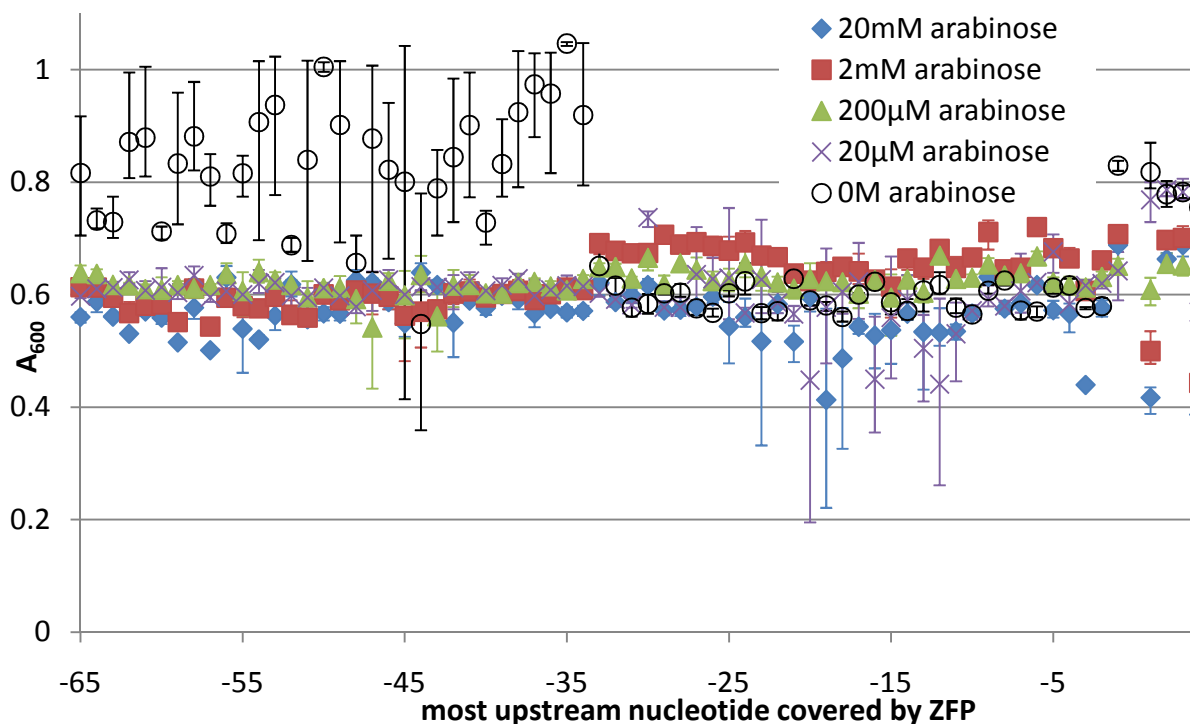
The final A<sub>600</sub> values for the P<sub>BAD</sub>/SC101\*\* plasmid are lower than from the P<sub>lacUV5</sub>/p15A plasmid, but this is due to the plasmid backbone and not the expression of the zinc finger proteins (Figure 3-27). Even though the copy number of the plasmid was reduced, the concentration of chloramphenicol was not reduced from 34 µg/mL, and this resulted in slower growth and lower final cell densities.





**Figure 3-26 Transcriptional attenuation measurements with the zinc finger protein expressed from a P<sub>BAD</sub>/SC101\*\* plasmid. The other half of this data set can be found in the previous figure. Error bars are min/max of 3 biological replicates.**

The final cell density data contains a significant amount of noise that makes observing toxicity difficult. One 96 well plate of samples shows higher cell densities and larger variability between biological replicates. This plate likely was on the top of a stack of plates loaded onto the plate reader, and it was able to evaporate for about 30 minutes while the other plates in the stack were read. Future experiments using stacks of plates should be done with an empty plate placed on the top of the stack to prevent evaporation. Additional controls that are replicated across the plates would also help to clearly identify plate-to-plate variation and allow for systematic plate variations to be normalized out.



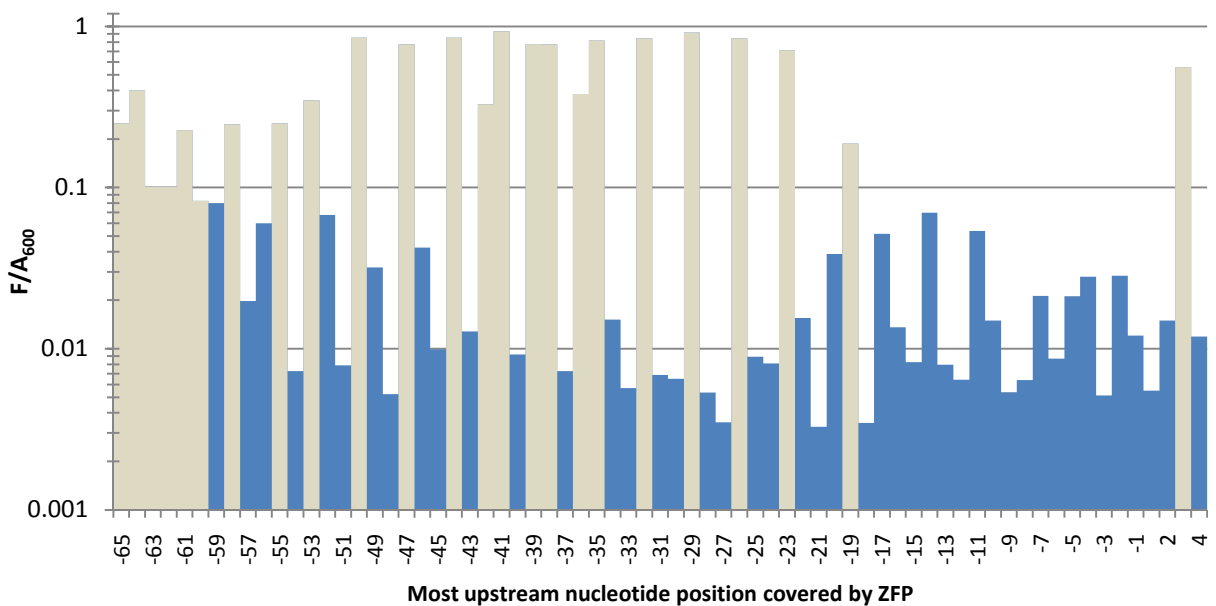
**Figure 3-27 Final absorbance measurements with zinc finger proteins expressed from a  $P_{BAD}/SC101^{**}$  expression plasmid. Despite the technical noise in this assay, it appears only three zinc finger proteins, at positions -3, 1, and 4, show significant toxicity. Error bars are min/max of 3 biological replicates.**

A saturating induction of the  $P_{BAD}/SC101^{**}$  expression plasmids shows reduced toxicity relative to the previously used  $P_{lacUV5}/p15A$  expression plasmids. Significant toxicity was only observed for positions -3, 1, and 4, in contrast to the previous experiment where saturating induction of the  $P_{lacUV5}/p15A$  plasmid resulted in reduction of final  $A_{600}$  levels by over 10% for 8 different zinc finger proteins. Surprisingly, these two sets of zinc finger proteins that have been toxic in some contexts are non-overlapping. Such a result would be possible if a zinc finger protein was binding the origin of replication of a plasmid and causing a reduction in the copy number of that plasmid and the antibiotic resistance gene it contains. While this explanation is possible, the probability that 8 of the 69 zinc finger proteins all have binding sites within the  $\sim 4000$  base pairs of the  $P_{lacUV5}$  expression plasmid is extremely low.

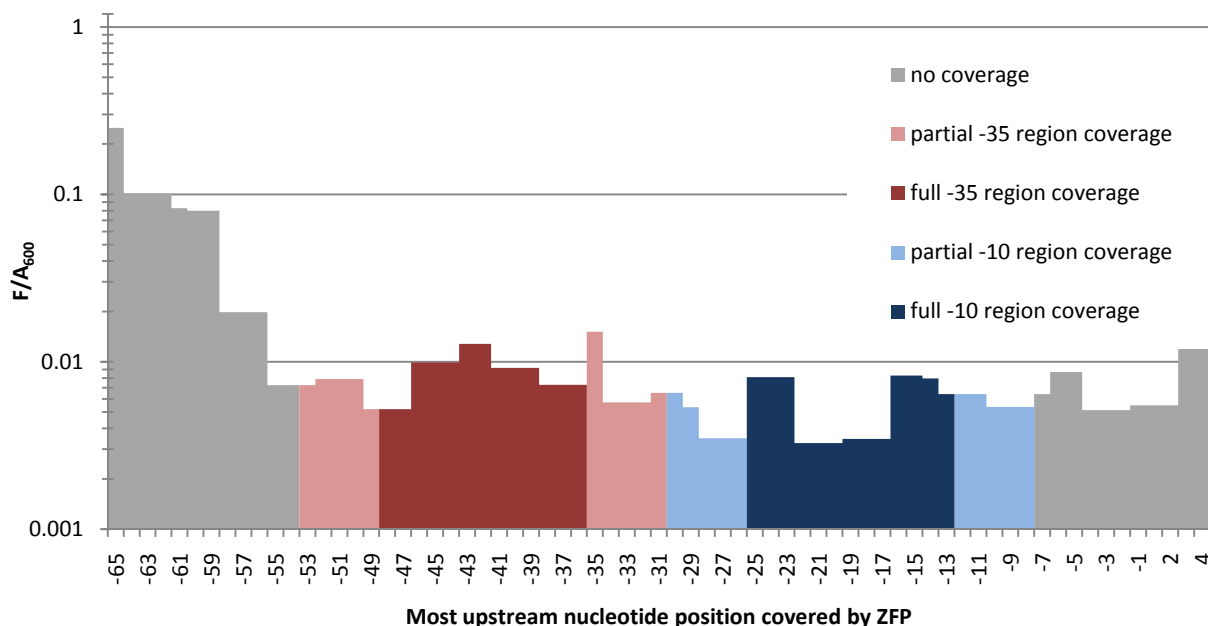
The transcriptional attenuation data generated by expressing the zinc finger proteins from  $P_{BAD}/SC101^{**}$  plasmids is similar to the  $P_{lacUV5}/p15A$  data set. The attenuation data is shown in Figure 3-25 and Figure 3-26. The analysis done on the  $P_{lacUV5}/p15A$  data set was repeated with this new data set. The maximum attenuation was over 300 fold and was achieved at position -21. Some of the zinc finger proteins failed to give substantial attenuation, and because these low attenuation proteins have positions that differ by multiples of 3, they are likely due to a zinc finger not functioning in this context. Of the set of 69 zinc finger proteins tested, 24 of them are expected to fail.

In Figure 3-28 the data representing the 24 lowest attenuating zinc finger proteins are displayed in grey in order to highlight the remaining data points that have a higher probability of being from functional zinc finger proteins. Alternatively, by plotting the minimum of the value measured for a position and its two neighboring positions, a plot with a lower resolution approximation of the attenuation possible can be generated as shown in Figure 3-29. From this plot it is clear that if a functional zinc finger protein is used, then more than 80% attenuation is probably possible for positions -53 to 2 and attenuation of over 99% may be possible for the majority of these positions.

By considering the probability that each zinc finger protein is functional, it becomes easier to notice relationships within the data set. In Figure 3-30 it is striking that from -65 to -53 all of the zinc finger proteins result in over 59% attenuation. But from -50 to -23, 10 of the zinc finger proteins gave less than 30% attenuation. There are several possible explanations for this difference, but unfortunately they can not be elucidated with the current data. One explanation is that the consensus promoter sequence is more constrained in the region covered by the -50 to -23 positions, and these sequence constraints resulted in a higher probability of non-functional zinc finger proteins. Another possibility is that poor binders may be able to cause over 59% attenuation in the region from -65 to -53 while similarly poor binders can only cause up to 30% attenuation in the region from -50 to -23. This would make mechanistic sense as -53 is the border between where there is no coverage of the -35 and -10 regions. Positions upstream of the -53 site may only be disrupting the RNA polymerase holoenzyme sliding along the DNA while



**Figure 3-28** Transcriptional attenuation with zinc finger proteins expressed from the P<sub>BAD</sub>/SC101\*\* plasmid and a saturating induction (20mM arabinose). The 24 zinc finger proteins giving the least attenuation are displayed in grey because they have a high probability of being non-functional zinc finger proteins. For this data set, the cutoff attenuation value separating the 24 poor attenuators from the 45 good attenuators was 0.0825.



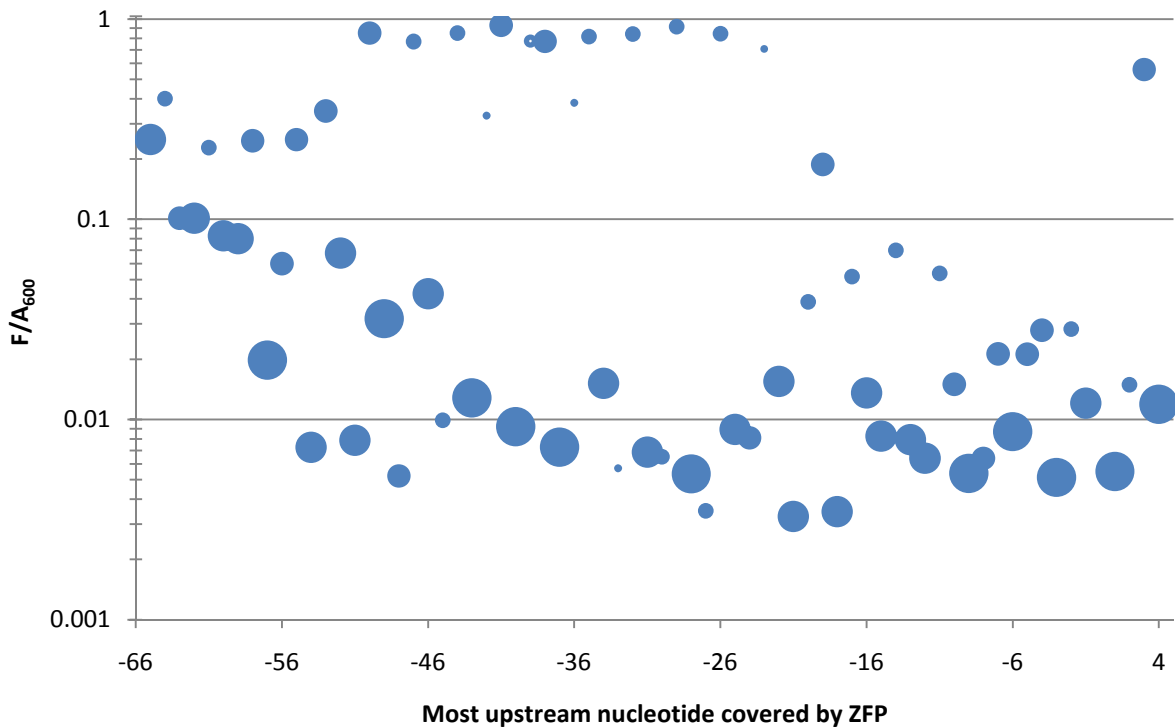
**Figure 3-29 A lower resolution approximation of the attenuation as a function of position for saturating induction (20mM arabinose). The value plotted at each position is the minimum of the value measured for that position and the two neighboring positions. The colors indicate if a zinc finger protein bound at that position would cover all 6 base pairs (full coverage) of either the -10 or -35 region of the promoter or 1-5 base pairs (partial coverage) of the promoter.**

weakly bound. This contrasts with the region from -50 to -23 where the zinc finger protein must compete with the stronger and specific binding of the  $\sigma^{70}$  portion of the RNA polymerase holoenzyme. If the weakly bound zinc finger protein can disrupt the non-specific interactions between the RNA polymerase holoenzyme but not the specific interaction with a promoter, then such a discontinuity would be expected.

The region from -20 to 2 also has a different behavior than either the -65 to -53 region or the -50 to -23 region. Within the -20 to 2 region all zinc finger proteins created over 80% of attenuation, and all but one zinc finger protein caused over 93% attenuation. These zinc finger proteins all have coverage within the -11 to 2 segment where DNA strand separation begins and the transcription initiation bubble occurs[56]. Because all 6 of the zinc finger proteins in the -20 to 2 region that contain exactly 2 GNN fingers were able to cause more than 93% attenuation, it is likely that even poor DNA binders are able to compete against the RNA polymerase attempting to form the transcription initiation bubble.

### Measurement of dissociation constants for zinc finger proteins

The attenuation data in the previous section indicates there are many positions along a promoter where binding a zinc finger protein has the potential to attenuate transcription. Drawing stronger conclusions about the relative amount of repression possible at different positions will require more data. Measuring dissociation constants for



**Figure 3-30** Transcriptional attenuation as a function of both position and number of GNN binding fingers in a zinc finger protein. The  $P_{BAD}/SC101^{**}$  zinc finger protein plasmids were induced with 20mM arabinose. The width of the bubble is proportional to the number of GNN binding fingers contained in each zinc finger protein. Open circles represent proteins without any GNN fingers. Many, but not all, of the weak attenuators have a low number of GNN fingers.

the zinc finger proteins with their target operator sites should yield additional insight. Measuring the dissociation constants for all 69 zinc finger proteins would be ideal, and this will require a fairly rapid method.

There are many methods for measuring dissociation constants. The gold standard is to use radioactively labeled DNA in an electrophoretic mobility shift assay (EMSA) where a DNA-protein complex runs slower on a gel than non-complexed DNA. These assays are relatively time consuming and the electrophoresis buffer and gel environment may not be representative of *in-vivo* conditions.

Fluorescence anisotropy assays use polarized light to measure the tumbling rate of fluorescently labeled DNA. When the protein binds to the DNA the mass increases and the tumbling rate slows. Fluorescence anisotropy assays can be performed in any buffer and can be rapidly performed on a plate reader. Published reports on using this assay with zinc finger proteins differ on how to setup the experiment with one source claiming the need for internal labeling of the operator site[57] and another source end labeling the DNA but fusing maltose binding protein to the zinc finger protein[58]. When internally labeling the operator site, it is unclear if the difference in the tumbling rate of the fluorophore is due to reduced tumbling of the DNA-protein complex or if the linker between the fluorophore and

DNA has reduced mobility when the zinc finger protein is bound. The cost to create a library of internally labeled DNA strands is high because a single primer with a fluorophore cannot be re-used to create many operators.

Equilibrium dialysis has been reported as a general method for measuring dissociation constants, but it has not yet been done with zinc finger proteins. The zinc finger protein is placed in a dialysis membrane which it cannot penetrate but can be crossed by DNA operators. Then a measurement of the equilibrium DNA concentration outside of the membrane can be used to calculate the dissociation constant. This method can be performed in any buffer and does not require the use of labeled DNA.

A variation on equilibrium dialysis is equilibrium filtration where an equilibrated mixture of protein and DNA is placed in a spin filter that retains the protein, and ~10-20% of the solution is filtered [59]. The DNA concentration in both the eluted and retained volumes is measured and used to calculate the dissociation constant. Equilibrium filtration is better suited for small samples and potentially faster than equilibrium dialysis. The filtration method does assume that the inverse of the off rate will be much larger than the duration of the assay.

Some work has been done on a computational method for determining the dissociation constant of zinc finger proteins [58]. However this study was limited to 3-finger proteins, and it is unclear if the method will work equally well on 6-finger proteins, where the geometric constraints on the protein backbone are different [22].

Using an *in-vivo* method it is possible to distinguish high and low affinity zinc finger proteins [39]. By using a one-hybrid system to activate the transcription of an antibiotic, it was possible to determine if the dissociation constant of a zinc finger protein was above or below 1 $\mu$ M. Another study using a one-hybrid system found a strong correlation between transcription rates and zinc finger protein strength [60]; however because this work was performed with Zif268 and binding site variations, it is unclear if similar results would be found for rationally designed zinc finger proteins and their target operator sites.

Because the *in vivo* and computational methods are not yet well developed and tested enough to determine the dissociation constants of the zinc finger proteins in this study, *in vitro* methods were pursued. All of the *in vitro* methods require purifying the zinc finger protein first.

### **Purification of zinc finger proteins**

Zinc finger proteins were cloned into a pET-29 expression plasmid. This vector contains a pBR322 origin, confers kanamycin resistance and the cloning site has an N-terminal S-tag along with a C-terminal 6xHis-tag. The 6xHis-tag enables fast purification of the protein using Qiagen's Ni-NTA spin columns. The S-tag can be used to quickly quantify the amount of protein in whole cell lysate using an enzymatic assay with a colorimetric output. Additionally, if the 6xHis-tag purification fails, a slower purification method can be done using a column with affinity for the S-tag.

Developing a purification protocol for these zinc finger proteins was difficult because they are toxic at high expression levels. The *E. coli* strain BLR(DE3) was used for protein over-expression because the zinc finger proteins contain repeated sequence and this strain is recA-; therefore this strain is less likely to delete repeated sequences through homologous recombination. The keys to obtaining substantial amounts of protein were to

minimize pre-induction expression and to induce near the relatively high cell density of  $A_{600}=1.6$ . Including the plasmid pLysS helped reduce leakage by producing T7 lysozyme, which inhibits T7 polymerase. This lysozyme is also helpful to have for the cell lysis step of the purification where a sonication step would be extremely time consuming. At all stages, both on solid media and for liquid culture, the cells were grown on LB supplemented with 1% glucose. The presence of glucose increases the inhibition of the Lac promoter driving expression of the T7 polymerase.

Care was taken to avoid having the cells enter stationary phase any more than necessary. A fresh transformation was made before each purification and was only grown long enough to obtain easy to pick colonies. Colonies were picked into 5 mL of LB with 1% glucose and grown for 8 hours at 37°C. Cultures were then placed at 4°C overnight. This was done to keep them from entering stationary phase. The following day a 50:1 dilution was done into 5mL of fresh LB with 1% glucose and the cultures were grown at 37°C until they reached  $A_{600}=1.6$ . Then they were induced with 1mM of IPTG. Non-saturating inductions done at lower cell densities were found to produce less of the desired protein. The cells were returned to 37°C shaker for 3 hours post-induction.

Cells were harvested and lysed using Novagen's Bugbuster according to the manufacturer's suggested protocol, including the optional benzonase and lysozyme additions. Protease inhibitors were found to be unnecessary, but if they are used they must be carefully selected to avoid chelating agents such as EDTA that may remove the zinc ions from the proteins.

The cell lysates were mixed in a 1:1 ratio with the Qiagen NPI-20 buffer and then purified using the manufacturer's suggested protocol. Using cell cultures of larger than 5mL was found to reduce the purity of the protein. Purified protein was stored at -20°C in a solution of 1mM DTT, 0.02% (w/v) sodium azide and 25% (w/v) glycerol. Once again, chelating agents are to be avoided.

Only 2 of the 69 zinc finger proteins have been purified thus far. Using a Bradford assay with BSA as the standard, the protein concentrations were measured. The cognate protein for op-39 has relatively low toxicity, and 6nmol was obtained from a 5mL culture. This is 50% of the maximum binding capacity of the spin column. The cognate protein for op+1 has relatively high toxicity, and 2 nmol was purified from a 5 mL culture. The calculated mass of the zinc finger protein for op-39 is ~24.6kDa, which includes 6 zinc atoms. SDS-PAGE using a MOPS buffer showed a strong band at 30kDa. These zinc finger proteins are proline rich, and proline rich proteins are known to migrate slowly on SDS-PAGE gels due to the reduced flexibility of the protein backbone near proline residues [61]. An S-tag assay using Novagen's S-tag rapid assay kit confirmed this 30kDa band was indeed the desired protein.

Measurements of the protein-DNA dissociation constants have not yet been completed but are needed to finish this work. Both equilibrium filtration and fluorescence anisotropy measurements will be attempted. To increase the mass of the zinc finger protein, it may be necessary to create a protein fusion with maltose binding protein. This increased mass is advantageous with both methods because both methods rely on the mass and size difference between the DNA molecule and the protein.

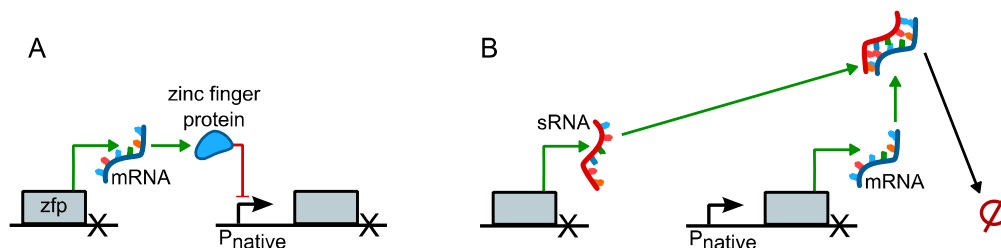
## Creating repressors for other promoters

By designing zinc finger proteins that target native promoters, it should be possible to attenuate expression of native genes. If the zinc finger protein is introduced on a plasmid, then it should be possible to modulate gene expression from the chromosome without making any modifications to the chromosome. This would allow for the inducible and temporary attenuation of required genes that are lethal to knockout. Those promoters targeted by zinc finger protein repressors would still remain under the control of the native regulation when the zinc finger protein repressor was not present.

Similar results may be possible by designing sRNAs to target native genes, but currently there is not an efficient method for designing sRNAs to target a specific gene of interest. Additionally the metabolic burden of a zinc finger protein repressor approach differs from the burden imposed by an sRNA based system. For the zinc finger protein repressor system, the targeted gene does not get transcribed but there is the cost of translating the repressor protein. With an sRNA system the burden of the target transcript would still exist, but there would be no translational burden. There are advanced sRNA systems where the sRNA binds the target mRNA before it is completely transcribed and introduces an early terminator, thus reducing the transcriptional load [62]. However, engineering such a system to work with an arbitrary native mRNA would be an impressive feat. Engineered sRNAs repressing the translation (and not transcription) of native mRNA is the more likely approach.

Whether an sRNA or protein repressor based system causes more overall metabolic burden will depend on the operon being targeted as shown in Figure 3-31. Weakly transcribed mRNA can be attenuated with less load by an sRNA, compared to highly transcribed genes where eliminating those transcripts could outweigh the cost of translating a repressor protein.

In metabolic engineering there are often native pathways that take away precursors or intermediate metabolites from an engineered pathway. These native pathways cannot be knocked out because they are required for growth. However, in metabolic engineering it is common practice to maximize the productivity of a culture by growing to a desired cell density without product formation and then inducing the engineered pathway. During the post-induction phase, the goal is push as much carbon through the product formation



**Figure 3-31 Differences in metabolic load for methods of repressing native genes. A) Use of a zinc finger protein repressor does not expend energy on the transcription of the target gene, but energy is used to translate the repressor. B) Regulation with sRNA does not use any energy for translation, but energy is used to translate the mRNA for the target gene. The system with the lowest metabolic burden depends on the rate of transcription from the native promoter.**



pathway as possible and to divert as little carbon as possible to growth. By using zinc finger protein repressors the induction of the product formation pathway can also induce zinc finger proteins to turn off pathways that are related to cell growth. This may result in both higher productivities and increased yields.

## Chapter 4 Topologies for building combinational logic gates

Combinational logic is defined by three attributes. The inputs and output are digital, meaning they can be represented by a series of two-state values. These two states are commonly referred to as “high” and “low” or 1 and 0. The second attribute is the relationship between the inputs and outputs consists of Boolean functions. These functions have binary (two-state) inputs and outputs and perform operations composed of AND, OR, and NOT operators. The third attribute is that the output of combinational logic is solely a function of the current input. There is no dependence on previous states of the system.

The primitive element of combinational logic is the logic gate. Logic gates take one or more binary inputs and produce a single binary output based on Boolean functions. The two simplest logic gates have only a single input and are called the buffer and the inverter. For a buffer gate the output state is the same as the input state. Because buffer gates only propagate an existing state, they are not needed in idealized systems; however, in real implementations they are used to limit the propagation of non-ideal effects. Inverters output the opposite state as the input. Inverters are also called a NOT gate. The basic logic gate symbols and their truth tables (input/output definitions) are shown in Figure 4-1. Both the NOR gate and the NAND gate are universal gates, meaning any combinational logic function can be implemented by only using one type of these gates. In Figure 4-1, NOR implementations of the basic logic gates are depicted in the right most column.

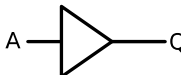
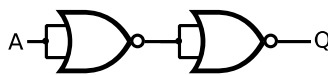
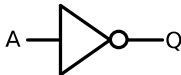
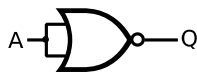

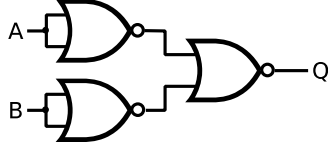

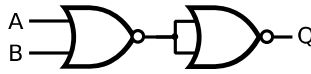

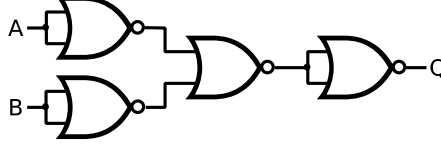

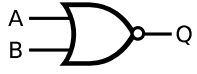

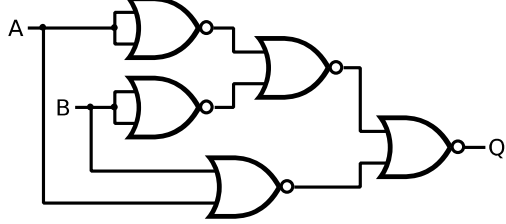
### *Desired properties of logic gates*

A recent paper by Clancy and Voigt lists some important characteristics that a system of combinational logic should have [63]:

- Scalability – It should be possible to build many orthogonal gates, enabling the implementation of large logic functions.
- Extensibility – Making useful logic functions requires connecting the output of one gate to the input of another gate; therefore, the input and output signals must have compatible specifications.
- Modularity – In order for combinational logic to be useful in real applications, the inputs and output of the logic block must be easy to interface with other common devices such as sensors and gene expression devices.
- Robustness – Combinational logic blocks should produce the desired outputs over a range of contexts and environmental conditions and in the presence of typical amounts of noise. Additionally, the impact on the host organism should be minimized.
- Speed – Every application will have different constraints on speeds, but generally faster is better. As logic blocks get larger and the longest path from an input to an output grows, the speed for an individual gate to transition will become increasingly important.

### Previous work

Implementing combinational logic in engineered organisms has been pursued by the synthetic biology community for over a decade. Some of the earliest work in this area was done by Ron Weiss during his Ph.D. including both simulation [64] and physical

Name	Symbol	Definition	NOR Implementation
BUFFER		$\begin{array}{c c} A & Q \\ \hline 0 & 0 \\ 1 & 1 \end{array}$	
NOT		$\begin{array}{c c} A & Q \\ \hline 0 & 1 \\ 1 & 0 \end{array}$	
AND		$\begin{array}{cc c} A & B & Q \\ \hline 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{array}$	
OR		$\begin{array}{cc c} A & B & Q \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{array}$	
NAND		$\begin{array}{cc c} A & B & Q \\ \hline 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{array}$	
NOR		$\begin{array}{cc c} A & B & Q \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{array}$	
XOR		$\begin{array}{cc c} A & B & Q \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{array}$	

**Figure 4-1** The basic set of two input logic gates. A and B are binary input signals, and Q is the binary output signal. The “Definition” column shows the truth table for each gate. This table defines the input/output behavior of each gate. The “NOR Implementation” column demonstrates how any of these basic gate functionalities can be implemented with a set of NOR gates. With enough NOR (or NAND) gates it is possible to implement any combinational logic function.

implementation [65]. The first synthetic inverters were design, built and characterized by Weiss using commonly studied repressor-promoter pairs such as LacI/P<sub>lac</sub>, TetR/P<sub>Ltet0-1</sub>, and cI/P<sub>R</sub>. By using repressor proteins with cooperativity (Hill coefficients > 1) and small dissociation constants, Weiss' devices had transfer curves that were good enough for use in many applications. Since then several gates have been designed that use various small molecules as their inputs [66], [67]. Because of their dependence on specific small molecules, use of these gates is limited to applications where the small molecule used as inputs are relevant to the task at hand. Other work has strived to be more re-usable and has embraced the standard of PoPS based input and outputs [68], [69]. Having a standardized system of interconnect between gates, sensors, and output devices is essential to making devices re-usable in other applications. Libraries of mutant LacI repressors along with promoter variants have been engineered to create small sets of orthogonal gates [69]; however their device characterization left much to be desired as no transfer curves were reported. Tamsir reported that multi-strain circuits could allow for gate re-use by physically separating the multiple instances of the gate with cell membranes [9]. Repressor proteins have not been the only type of molecules used to create combinational logic. Several groups have created RNA based logic gates [70], [71]. While RNA has many appealing properties to use in logic gate design, the performance obtained so far has trailed behind protein based gates. Some groups are looking beyond the creation of gates and are designing software tools that can take a collection of components and design a gene network that will perform a given logic function [63], [72].

The previous work on creating combinational logic has been successful at demonstrating and advancing four of the five characteristics desired for combinational logic systems. Extensibility was demonstrated by Weiss' ability to connect gates together [65] Modularity was shown by Anderson when he was able to replace components connected to the input and output of his AND gate [68]. A degree of scalability was engineered by Zhan when he created his mutant LacI-promoter library [69]. Little has been published on testing the robustness of logic gates at the bench, but Tamsir does present some data on parameter variation and noise in the supplementary material [9]. The speed of logic gates has not yet received much attention. It is expected that RNA based gates have the ability to switch faster than gates based the combined transcription and translation of repressors proteins, but this has not yet been experimentally shown. Creating faster gates will likely remain a low priority until the other desired characteristics have all been successfully combined within a single system of logic gates.

### ***Methodology for designing new logic gate topologies***

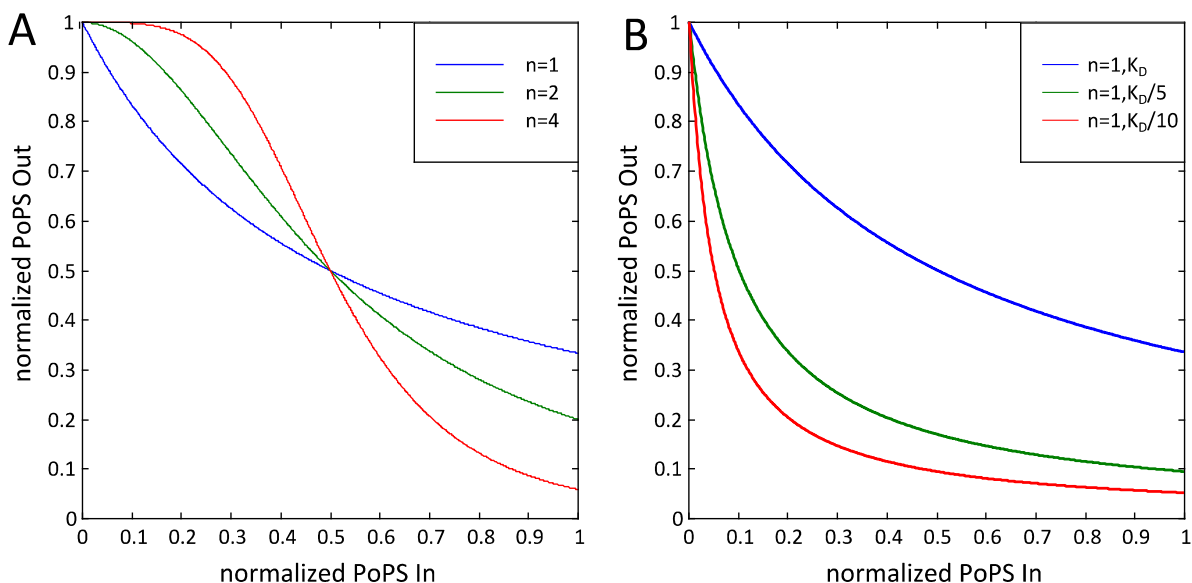
For transcription-based logic systems the problems of extensibility and modularity have largely been solved through the use of PoPS based inputs and outputs. The next challenge to overcome was building logic systems that can scale to 5 gates and well beyond that while maintaining robust behavior.

Due to the advances in synthetic zinc finger protein engineering, it appeared that utilizing zinc finger proteins would allow for highly scalable logic; however, it was not obvious how to create logic gates from zinc finger proteins. For transcription-based logic, there will need to be a transcription factor, and there are multiple ways to create transcription factors from a DNA binding domain:

- A DNA binding domain can prevent RNA polymerase from binding or transcribing through steric hindrance
- Attaching a repression domain and binding upstream of a promoter to inhibit transcription
- Attaching an activation domain and binding upstream of a promoter to activate transcription

Creating logic gates from transcriptional repressors that operate through steric hindrance of the RNA polymerase requires cooperativity as shown in Figure 4-2. The transfer function obtained from the non-cooperative binding of a monomeric protein to a promoter is not a reverse sigmoid and therefore not suited for creating logic gates. Creating a logic gate from steric hindrance requires a Hill-coefficient greater than 1 and for reasonable performance should be at least 2. Because simple synthetic zinc finger proteins do not dimerize, they cannot be used as transcriptional repressors to create logic gates. Synthetic zinc finger proteins have been made to dimerize by adding leucine zipper domains [73], but due to the small number of orthogonal leucine zipper domains this topology has scaling limitations.

In eukaryotes, synthetic transcriptional repressors have been created by fusing a DNA binding domain to a domain that represses transcription. These repression domains



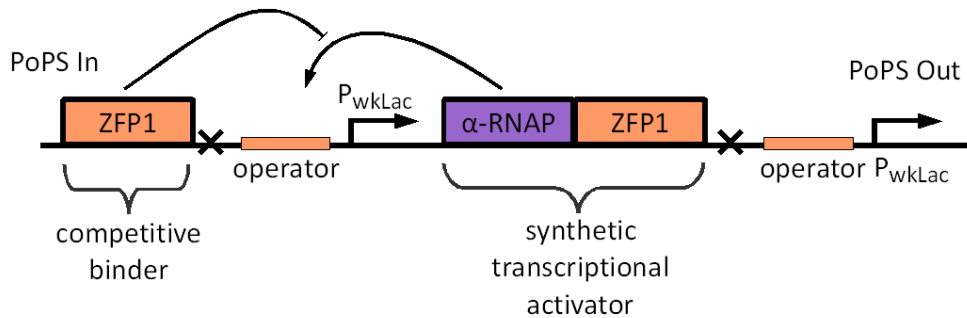
**Figure 4-2 Idealized transfer curves for inverters constructed with transcriptional repressors. These are simply plots of the Hill equation A) Transfer curves are shown for three different amounts of cooperativity where n is the Hill-coefficient. Based on the number of monomers complexed, a TetR based repressor should have a Hill-coefficient of 2 and 4 for a LacI based repressor. Only once the Hill-coefficient is greater than 1 does the curve become a reverse sigmoid. B) Transfer curves where the Hill-coefficient is 1 but the dissociation constant is varied by a factor of 10. With the largest dissociation constant the off state is far from 0, but with the smaller dissociation constants the on state is too small.**

from eukaryotes are unlikely to function in prokaryotic hosts, and similar modular repression domains have not yet been identified and studied in prokaryotes. Basic science work is needed for this approach to become feasible.

As discussed in earlier chapters, building transcriptional activators from zinc finger proteins in *E. coli* is a demonstrated technology. Using transcriptional activators, it is very easy to build OR and AND gates, but these gates do not allow for signal inversion. How to make a universal gate (NAND or NOR) from transcriptional activators is not obvious.

To determine how to make universal logic gates from zinc finger proteins the following process was followed:

1. Brainstorm possible topologies involving zinc finger proteins
2. Construct model using JDesigner
3. Run ODE based simulations using Systems Biology Workbench
4. Evaluate topologies heuristically
5. Eliminate topologies with poor evaluations
6. Solve ODE model for analytical solution at steady state
7. Evaluate sensitivity of topology to parameter variation
8. Eliminate topologies that are hyper-sensitive
9. Build and test topologies



**Figure 4-3 Positive feedback with competitive binder inverter topology.** A synthetic transcriptional activator consisting of the  $\alpha$ -subunit of RNA polymerase fused to a zinc finger based DNA binding domain is used to create a positive feedback loop by activating its own transcription. The same DNA binding domain without the  $\alpha$ -subunit competes to bind to the same operator sites as the transcriptional activator. Because both proteins have the same DNA binding domain and thus the same dissociation constant, the output is proportional to the fraction of DNA binding domains with the  $\alpha$ -subunit over the total number of DNA binding domains. Two copies of the  $P_{wkLac}$  promoter are used to maintain a pure PoPS output. If a single  $P_{wkLac}$  promoter was used to generate a multi-cistronic containing the transcriptional activator and the downstream gene connected to the output of the device, then it is possible for both RNA polymerase and ribosomes to cross the output interface. This possibility is eliminated by using a second promoter to drive the output. Additionally, the output of a PoPS device should ideally occur at the transcriptional start site, so that the PoPS generator is distinct from the mRNA coding.

The goal of this process was to rapidly test and eliminate topologies, thus allowing even farfetched or unusual ideas to be considered with a minimal investment of time. Using JDesigner from the Systems Biology Workbench [74], models of molecular interactions can quickly be designed using a graphical user interface to draw the molecules and their reactions. JDesigner can then create an ODE model from the drawn network. This was found to be faster and considerably less error prone than drawing a topology on paper, manually converting this to a system of ODEs and then putting these equations into simulation software.

At the time this work was performed, JDesigner and the Systems Biology Workbench were found to be the best tools for this task. Since then TinkerCell [75] has become available and has surpassed JDesigner and the Systems Biology Workbench for these tasks.

### ***Positive feedback with competitive binder topology***

Several rounds of iteration through the methodology listed above resulted in a candidate topology that seemed worthy of trying to implement. One of the keys to finding a candidate topology was realizing that it may be of use to have both transcriptional activators and pure DNA binding domains in the same device. This topology consists of a positive feedback loop and an inhibitor of the positive feedback loop. The positive feedback loop sets the inverter output value, and the inverter input is proportional to the inhibitor. Auto induction of a one-hybrid transcriptional activator creates the positive feedback loop, and a DNA binding domain competing for the same operator site as the one-hybrid activator creates the inhibition. The topology is shown in Figure 4-3.

### ***Analysis of topology***

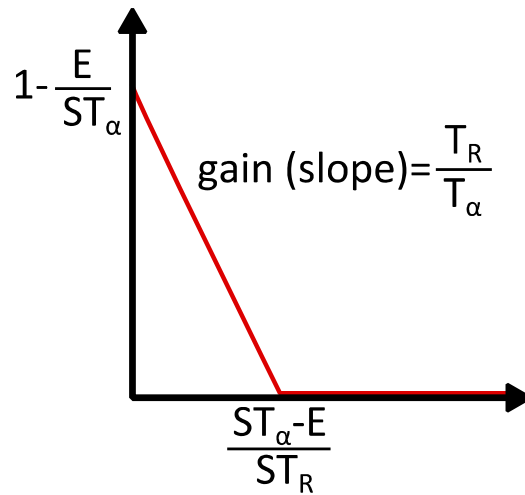
#### **Modeling and analysis**

An ODE model was constructed for the purpose of solving for a steady state transfer function. Two ODEs were used to represent the total concentrations of activator and repressor. A third equality is used to define the output value in terms of the activator and repressor concentrations. It will be assumed that the concentrations of DNA-activator and DNA-repressor complexes are functions of the total concentrations of activator and repressor. This assumption may not be good for simulating dynamics of the system, but should be accurate for steady state results.

$$\frac{d\alpha}{dt} = P T_{\alpha} P_T \frac{\alpha}{K_D + \alpha + z} - D\alpha \quad \mathbf{1}$$

$$\frac{dz}{dt} = Input P_T T_R - Dz \quad \mathbf{2}$$

$$Output = P \frac{\alpha}{K_D + \alpha + z} \quad \mathbf{3}$$



**Figure 4-4 The transfer curve for the positive feedback with competitive binder inverter. The gain is determined by the ratio of strengths of the ribosome binding sites for the competitive binder and the transcriptional activator. This transfer function is not a reverse sigmoid, but because the gate has a gain of greater than one and the low state reaches zero, this gate can be used to amplify digital signals.**

$\alpha$  is the concentration of transcriptional activator –  $\alpha$ -subunit of RNA polymerase fused to a zinc-finger protein.  $z$  is the concentration of transcriptional repressor – a zinc-finger protein that is competitively binds to the same site as the activator.  $P$  is the maximum rate of production of mRNA from one copy of  $P_{wkLac}$ . Units are  $t^{-1}$ .  $P_T$  is the concentration of the  $P_{wkLac}$ .  $K_D$  is the disassociation constant for the zinc-finger protein, both the activator and repressor, to the site upstream of the  $P_{wkLac}$  promoter. Input is the number of polymerases entering the device per unit of time.  $D$  is the degradation rate of both the activator and repressor proteins, with units of  $t^{-1}$ .  $T_\alpha$  and  $T_R$  represent the respective strengths of the activator and repressor RBSes. These variables are unit-less. Output is the number of polymerase exiting the device per unit of time. The variables of the system are non-dimensionalized as follows:

$$A = \frac{\alpha}{K_D}; R = \frac{z}{K_D}; S = \frac{P_T}{K_D}; \tau = tP; I = \frac{Input}{P}; E = \frac{D}{P}; O = \frac{Output}{P} \quad 4$$

Substituting Equation 4 into Equations 1, 2, and 3 yields non-dimensionalized equations:



$$\frac{dA}{d\tau} = \frac{ST_{\alpha}A}{1+A+R} - EA \quad 5$$

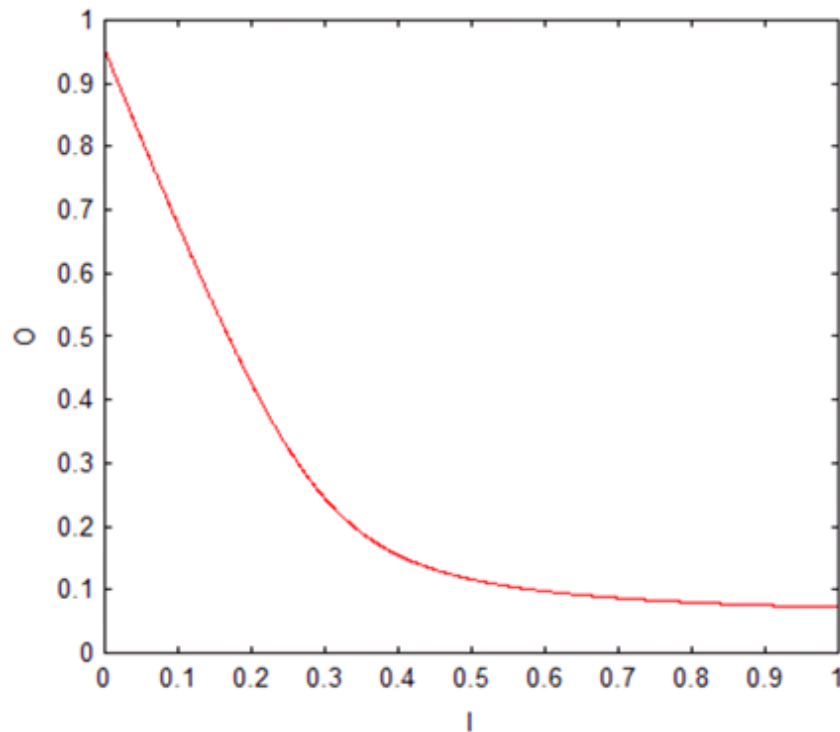
$$\frac{dR}{d\tau} = IST_R - ER \quad 6$$

$$0 = \frac{A}{1+A+R} \quad 7$$

The steady state transfer function can then be found by setting the ODEs equal to zero, solving for A and R, and substituting those expressions into Equation 7.

$$0 = \frac{\frac{ST_{\alpha} - E - IST_R}{E}}{1 + \frac{ST_{\alpha} - E - IST_R}{E} + \frac{IST_R}{E}} = 1 - \frac{E}{ST_{\alpha}} - \frac{T_R}{T_{\alpha}} I \quad 8$$

For regions where Equation 8 is negative, the transfer function should be set equal to 0. The transfer function can also be calculated while taking into account the non-zero leakage of the  $P_{wklac}$  promoter, but the simplified equation derived above provide more insight into how the parameters shape the transfer function. To derive the transfer function with leakage included,  $P_L$  is defined as the rate of mRNA production by the  $P_{wklac}$  promoter while



**Figure 4-5 Transfer function of the inverter with leakage included. The parameter values used for this plot are:  $L=0.05$ ;  $S=20$ ;  $E=1$ ;  $T_{\alpha}=1$ ;  $T_R=3$ .**

in a non-activated state. Equation 1 is replaced with:

$$\frac{d\alpha}{dt} = PT_{\alpha}P_T \frac{\alpha}{K_D + \alpha + z} - D\alpha + \left[1 - \frac{\alpha}{K_D + \alpha + z}\right] P_L T_{\alpha} P_T \quad 9$$

And Equation 1 is replaced by:

$$Output = P \frac{\alpha}{K_D + \alpha + z} + \left[1 - \frac{\alpha}{K_D + \alpha + z}\right] P_L \quad 10$$

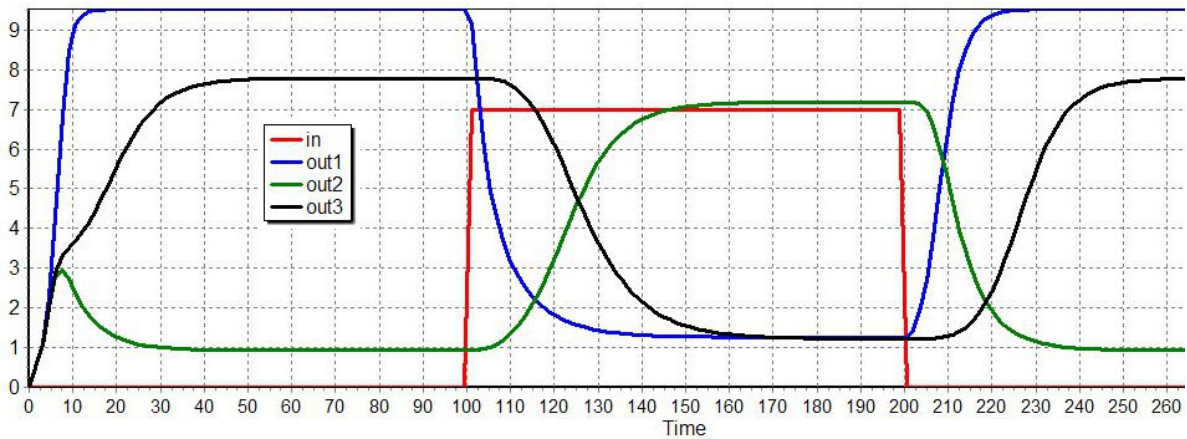
The new term  $P_L$  is non-dimensionalized as:

$$L = \frac{P_L}{P} \quad 11$$

Proceeding as above, a closed form analytical solution for the transfer function can be found; however, it is a large expression and doesn't offer much intuition.

To demonstrate that this unusual transfer function can pass digital signals, an ODE model of three inverters in series was simulated with a pulse function as the input. The results of this simulation are shown in Figure 4-6.

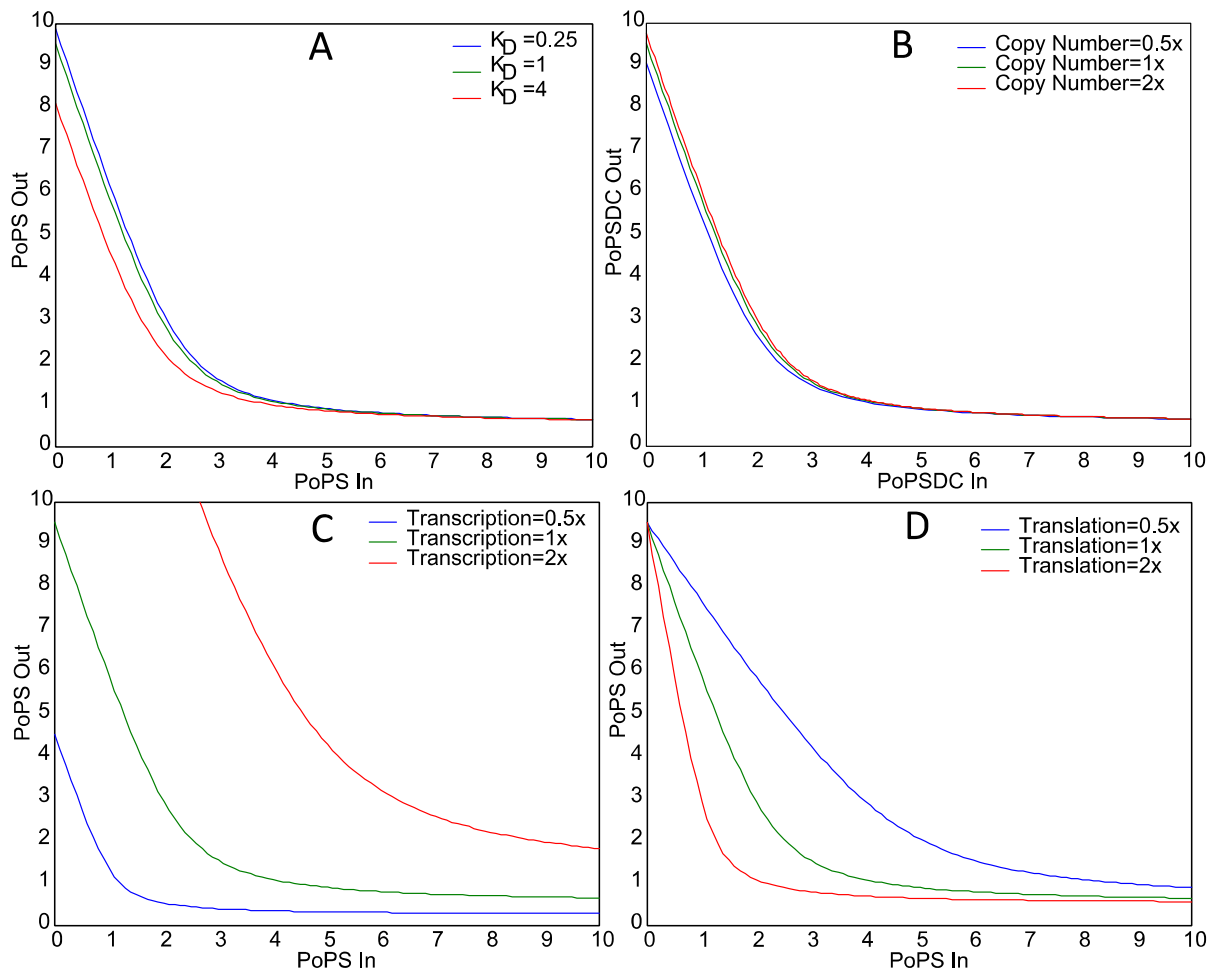
Using the model with leakage included, the sensitivity of the steady state transfer curve to various noise sources and parameter variations was investigated as shown in Figure 4-7. The performance under these sets of conditions was mixed with some parameters having almost no effect on the transfer function and other parameters causing large changes. The inverter is very insensitive to the dissociation constant being made smaller, but increasing the dissociation constant can slightly alter the performance. If a set



**Figure 4-6 Simulation of three “positive feedback with competitive binder” inverters in series. An ODE based simulation was performed with  $L=0.05$  and  $T_R=T_{\alpha}$ . The input to the first gate was pulsed and while none of the inverter outputs are able to reach 0, there are two clearly defined regions for high and low signals.**

of logic gates was constructed with this topology, it would be critical to have the dissociation value of all the zinc finger proteins be below some threshold value in order to ensure the gates have similar transfer functions. A change in the number of DNA copies of a gate has almost no effect on the transfer function. This is an important test because during every cell replication cycle, the DNA concentration in a cell must vary by a factor of 2.

For the copy number analysis, the transfer function is shown with the input and output units as Polymerase Per Second per DNA Copy (PoPSDC). This unit should always be used for transcriptional logic instead of just PoPS, but it is slightly more confusing and only becomes important when considering copy number effects. By using PoPSDC, the average

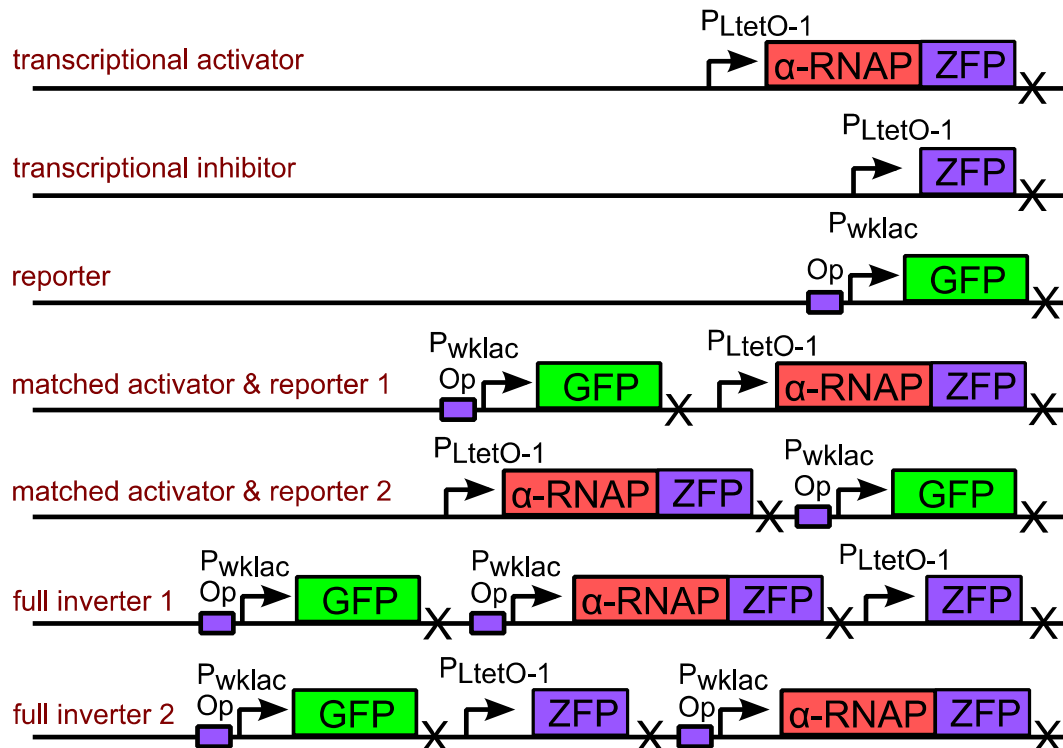


**Figure 4-7 Transfer curves for the “positive feedback with competitive binder” inverters with various perturbations. A) The  $K_D$  of both the activator and competitive binder were simultaneously modified. B) The number of copies of DNA containing the circuit was varied. The units on this plot are Polymerase Per Second Per DNA Copy (PoPSDC) as measuring on a per copy basis becomes important when investigating copy number effects. C) Varying the rate of transcription has devastating effects on the transfer function. D) Translation rate changes also cause significant changes in the transfer function.**

amount of polymerase crossing a junction on a piece of DNA is considered rather than the total amount of polymerase crossing all copies of a junction.

Changes to either the transcriptional or translational rates are quite detrimental to the transfer function. Altering the transcriptional rate shifts the whole transfer function and results in changes to the leakage level, the maximum output, and the position of the knee in the curve. Because changing the transcriptional rate scales both the input and output axes similarly, gates connected in series should be able to pass signals without a loss of fidelity. However the other devices connected to the logic block might react to a change in transcriptional rates differently than the gates, and thus signals may not properly propagate across these heterogeneous interfaces. Changes in translation rate only scale the transfer function along the input dimension. For small changes in translation rate, gates in series will still correctly propagate values, but their ability to tolerate noise will be reduced. If the reduction in translation rate is large enough that the gain of the gate drops below 1, then the gates will fail. For large increases in translation rate, the gain of the gate will increase while decreasing the input value needed to switch the output of the gate, and at high enough translation rates the width of the high state will vanish into the noise and the gates will fail.

Decreases in both transcriptional and translational rates are detrimental, and both can be caused by high metabolic burdens. For gates with these characteristics it will be



**Figure 4-8 Inverter construct and associated controls. Each of these constructs is on a plasmid with a SC101 origin of replication and kanamycin resistance. A brief description of each construct is on the left in red text.**

**Table 4-1 Zinc finger proteins and their operator sites**

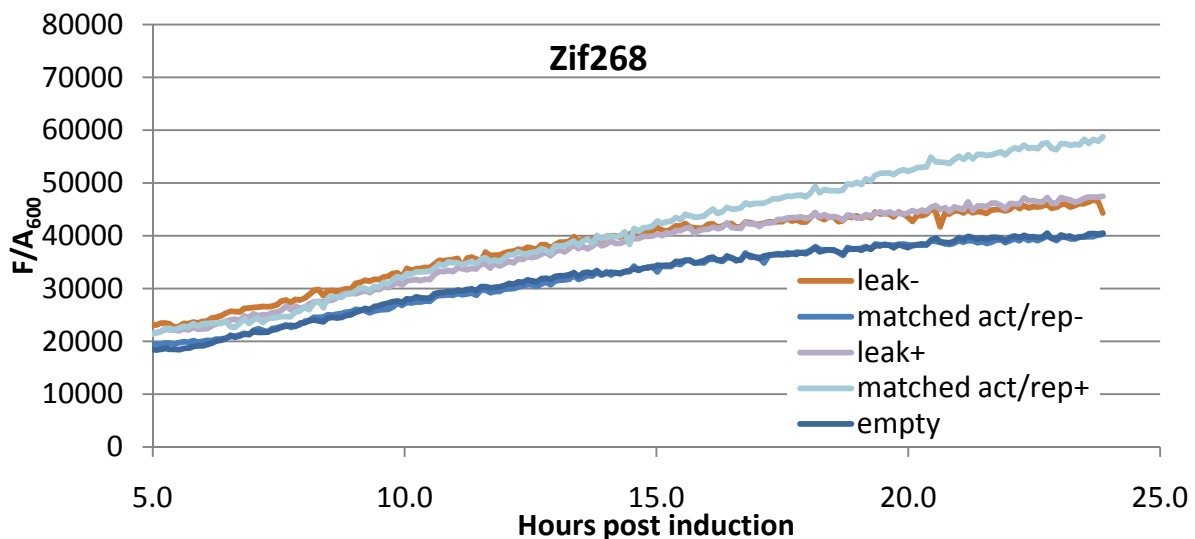
Zinc finger protein name	Operator site used
ZFP02	ATA ATG GAC CTA GGA GCT
ZFP03	GGT AGA CGT CTA GTA ACT
Zif268	GCG TGG GCG
ZFP5475	ACC CGG GTT CCC CTC GGG

vital to design and test the gates at conditions that will have similar growth rates, and therefore similar metabolic loads, as the conditions these gates will be ultimately used in. Unfortunately theory, tools and metrics for quantifying metabolic burden are lacking, and growth rate is the only good quantitative metric for comparing metabolic loads.

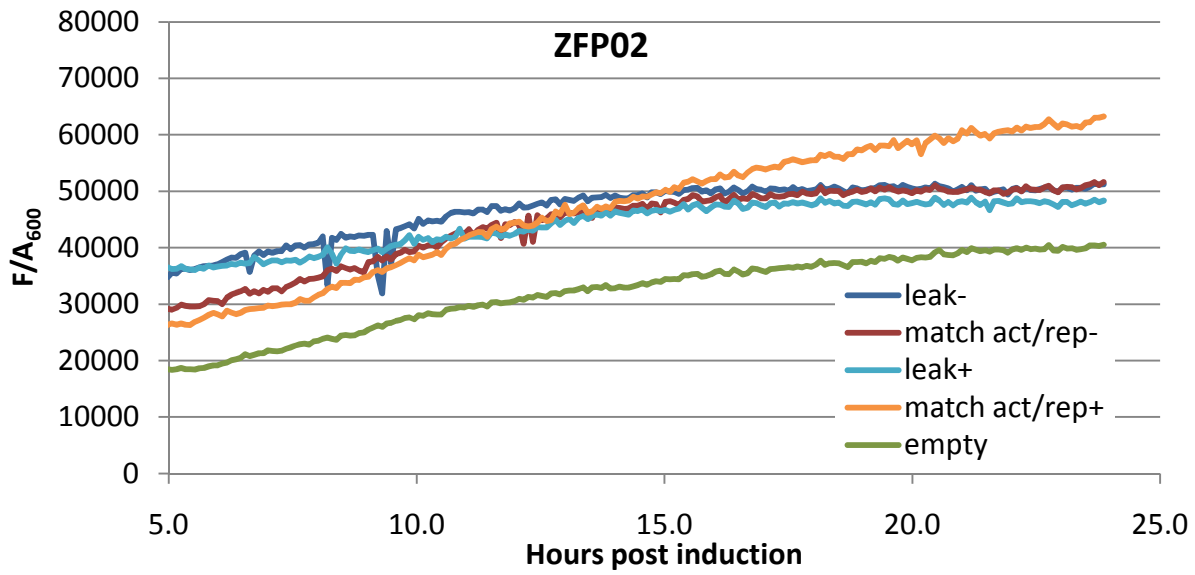
### Implementation and testing

The transcriptional activator architecture used was based on the work of Meng [40]. The  $\alpha$ -subunit of RNA polymerase was attached to the N-terminus of a 6-zinc finger protein. Meng used a linker of AAAPRVRTGS, but in order to facilitate BglBrick based cloning this amino acid sequence was changed to GSAPRVRTGS. The promoter to be activated by this fusion was a weakened version of  $P_{lac}$  called  $P_{wklac}$ . The operator site was placed upstream of the -68 position relative to the transcriptional start site. This is the same promoter used by Meng except the bases from positions -68 to -62 were altered to allow for BglBrick cloning of the operator.

A series of controls and inverters were constructed and tested. The constructs are shown in Figure 4-8. These construct were designed to verify the transcriptional activator was functional and to determine the amounts of transcriptional leakage and background

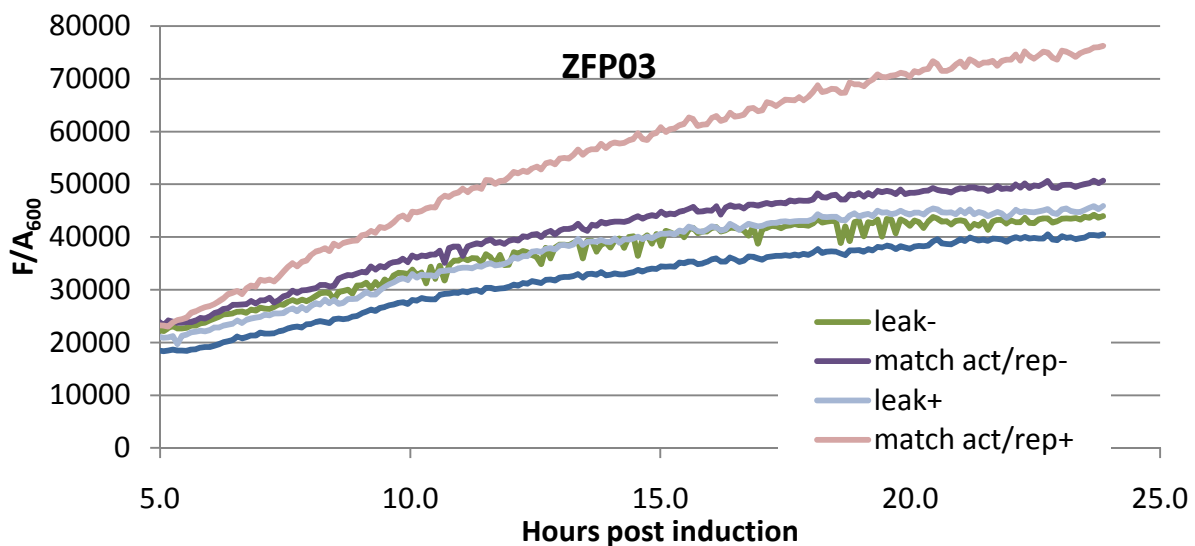


**Figure 4-9 Results from Zif268 controls. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched activator and reporter are of type 1.**

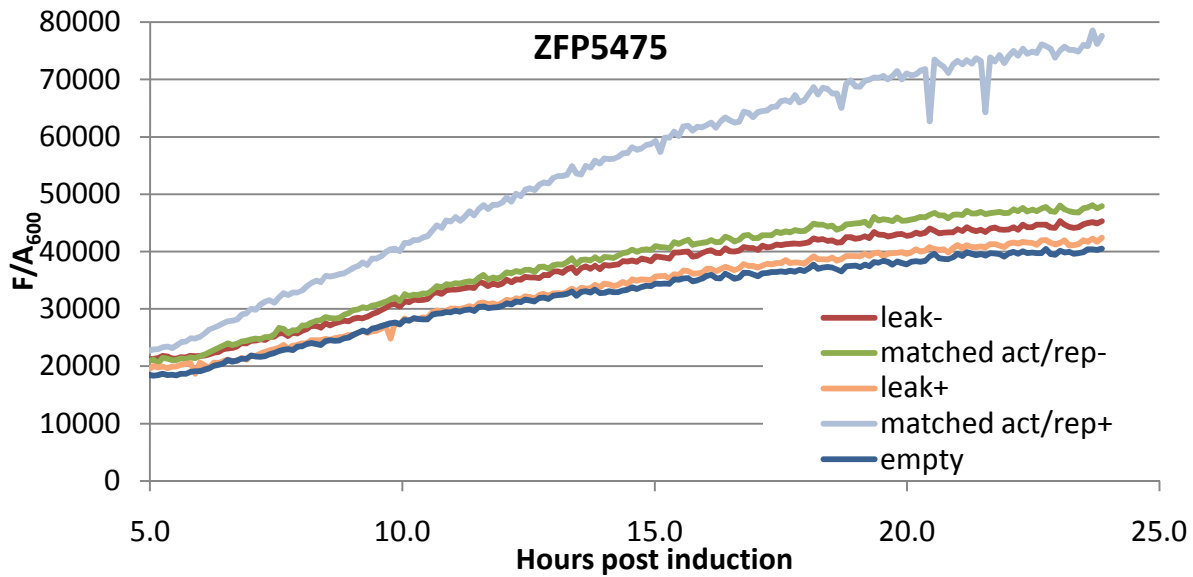


**Figure 4-10 Results from ZFP02 controls. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched activator and reporter are of type 2.**

fluorescence that should be expected. The two controls that do not contain GFP were used to determine if the expression of the zinc finger protein or transcriptional activator had toxic effects on the cell. The reporter plasmid was used to find the baseline level of fluorescence due to leakage from the P<sub>wklac</sub> promoter. The matched activator and reporter confirm the transcriptional activator is functional.



**Figure 4-11 Results from ZFP03 controls. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched activator and reporter are of type 1.**

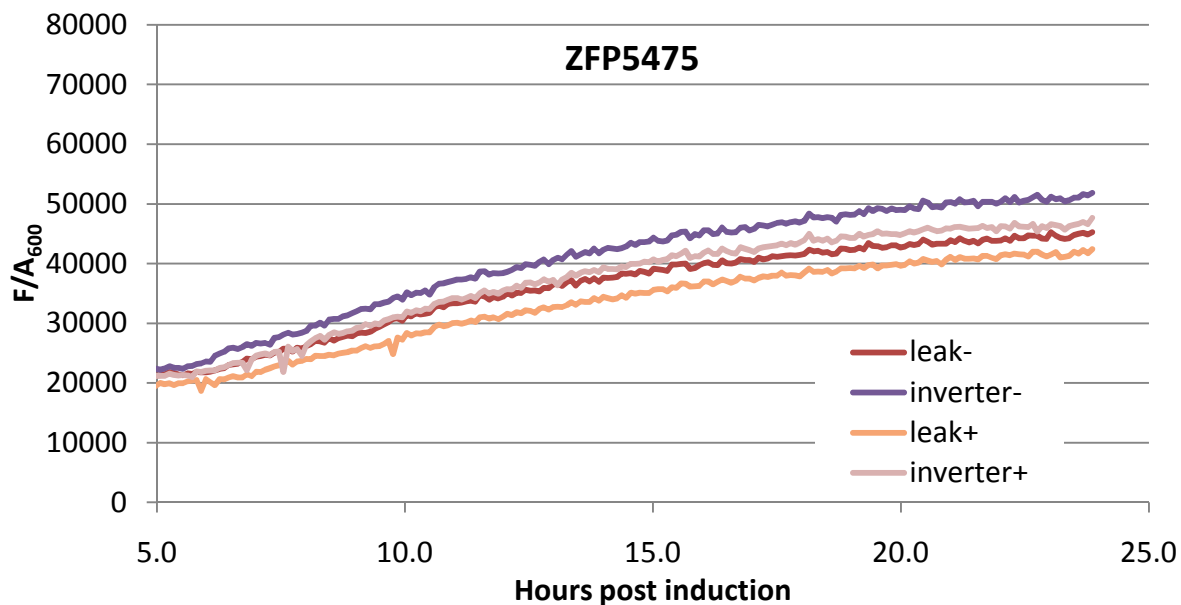


**Figure 4-12 Results from ZFP5475 controls. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched activator and reporter are of type 1.**

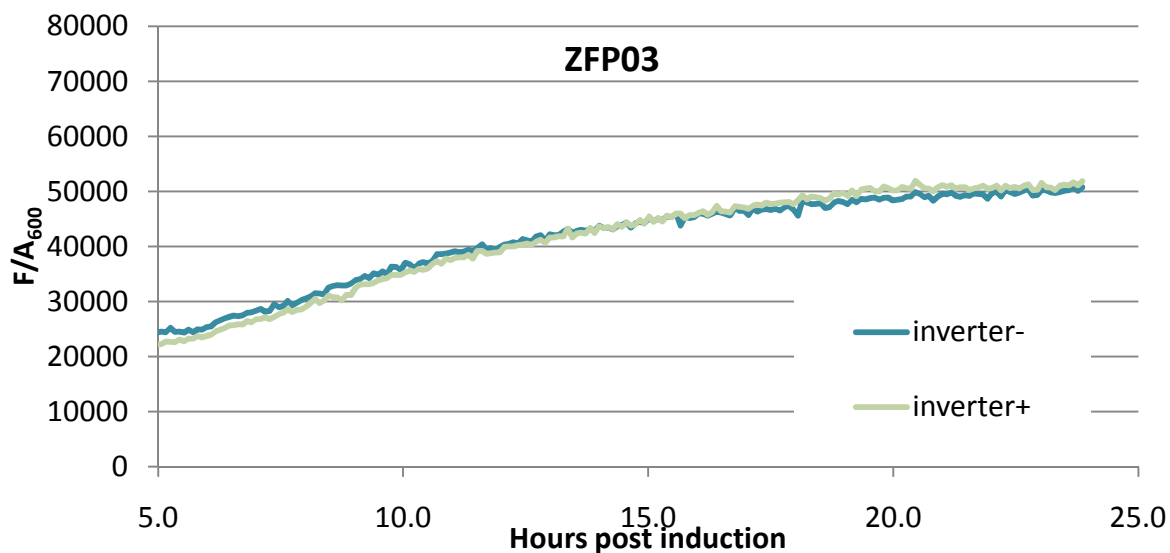
These constructs were built using 4 different zinc finger proteins. ZFP02 and ZFP03 are 6-finger proteins designed as part of this work. ZFP5475 is a previously published 6-finger protein [76], and Zif268 is a heavily studied mammalian transcription factor with 3 fingers. The apparent dissociation constants for ZFP5475 and Zif268 are reported to be 70pM [76] and 19pM[60] respectively. The operator sites for these zinc finger proteins are listed in Table 4-1.

The first round of experiments presented here were done with a supplemented M9 media [40] in 24 well plates (Labnet, P9835) with a gas permeable seal (Thermo Scientific, AB-0718). The host DH5 $\alpha$ -Z1 was used, and it expresses both LacI and TetR constitutively from the chromosome [44]. Induction was done by adding anhydrous tetracycline (aTc) at a saturating level (50 ng/mL).

Initial experiments indicated that some of the constructs were functional, but none of them performed as well as desired. The matched activator and reporter constructs containing ZFP03 and ZFP5475 both showed activation of the reporter, but in both cases the induced  $F/A_{600}$  was only about 50% higher than the  $F/A_{600}$  for the uninduced. For the constructs containing ZFP02 and Zif268, there may be slight activation but only when the cells are well into stationary phase (15 hours post-induction). The full inverter construct containing ZFP03 displayed no difference between the induced and uninduced cultures. The ZFP5475 inverter shows a slight inversion of signal, but it is unclear if this is significant or not as the induced and uninduced leakage constructs show about the same degree of difference. The  $F/A_{600}$  for the leakage constructs should be independent of the inducer levels, and therefore their difference can be used as a proxy for the noise in this assay.



**Figure 4-13 Results from ZFP5475 controls and inverter. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The inverters are of type 2.**



**Figure 4-14 Results from ZFP03 inverter. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the construct see Figure 4-8. The inverters are of type 1.**

The fold activation levels from the transcriptional activators did not seem to be large enough to use in making inverters. The measured activation was only about 1.5-fold, but 3-fold or more activation was desired. Additionally, the activation did not reach its highest levels until well into saturation. Ideally, gates will function in both exponential

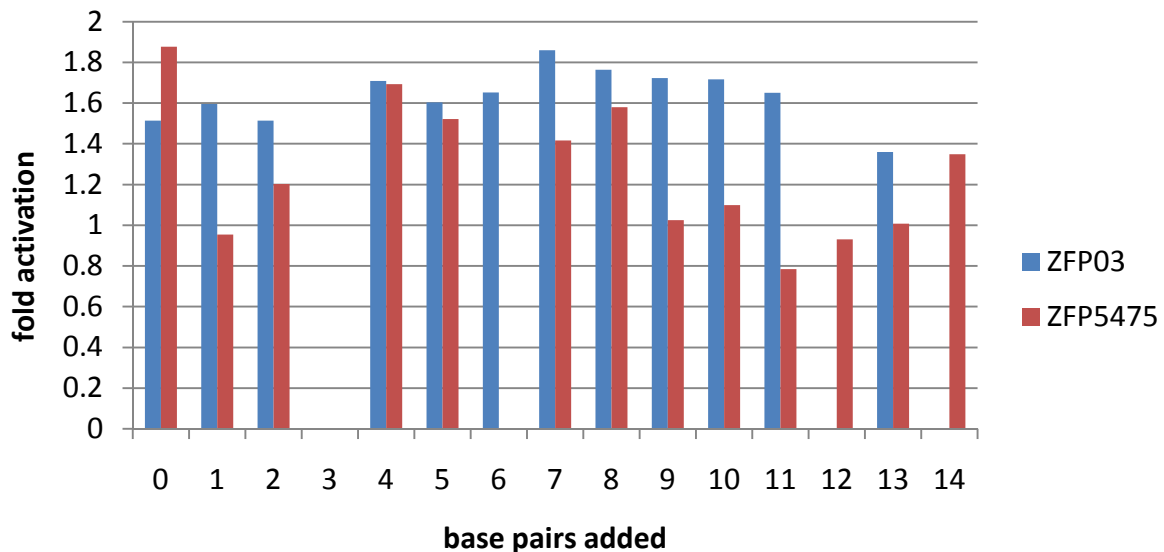


**Table 4-2 Alternate linkers for one-hybrid activators.**

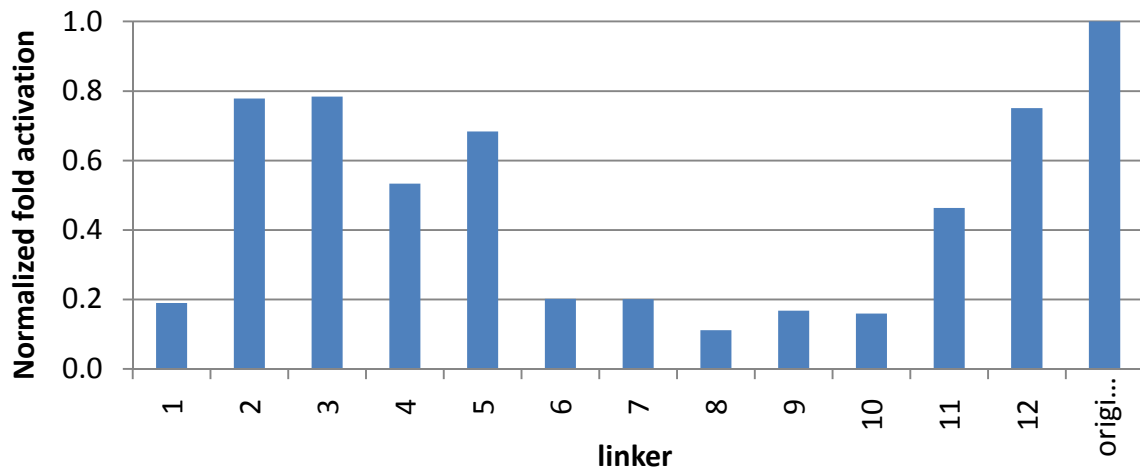
Number	Linker Amino Acid Sequence	Length	%G	%S	%A
1	AAADYKDDDDKFRTGSKTPPHRS	23	4	9	13
2	AAGGGGSGGGSGGGGSTAAA	21	57	14	24
3	PAAGAAGAGAAP	12	25	0	58
4	PAAGAAGAGAAGAAP	15	27	0	60
5	PAAGAAGAGAAGAAGAAP	18	28	0	1
6	PAAGAAGAAGAGAAGAAGAAP	21	28	0	62
7	PAAGAAGAAGAGAAGAAGAAGAAP	24	29	0	63
8	PGSGSGSGSGSP	12	50	33	0
9	PGSGSGSGSGSGSP	15	47	40	0
10	PGSGSGSGSGSGSGGGP	18	50	39	0
11	PGGSGSGSGSGSGSGSGGGP	21	52	38	0
12	PGGSGSGSGSGSGSGSGSGGGP	24	54	38	0

phase and stationary phase. Given a choice between functionality in only one growth phase, exponential would be preferred for ease of measurement and modeling.

Most one and two-hybrid systems are used as a selection in screening libraries.



**Figure 4-15 Modulating the distance between the promoter and operator. Fluoresce per A<sub>600</sub> was measured after 17.5 hours, and the induced value divided by the uninduced value is given here.**

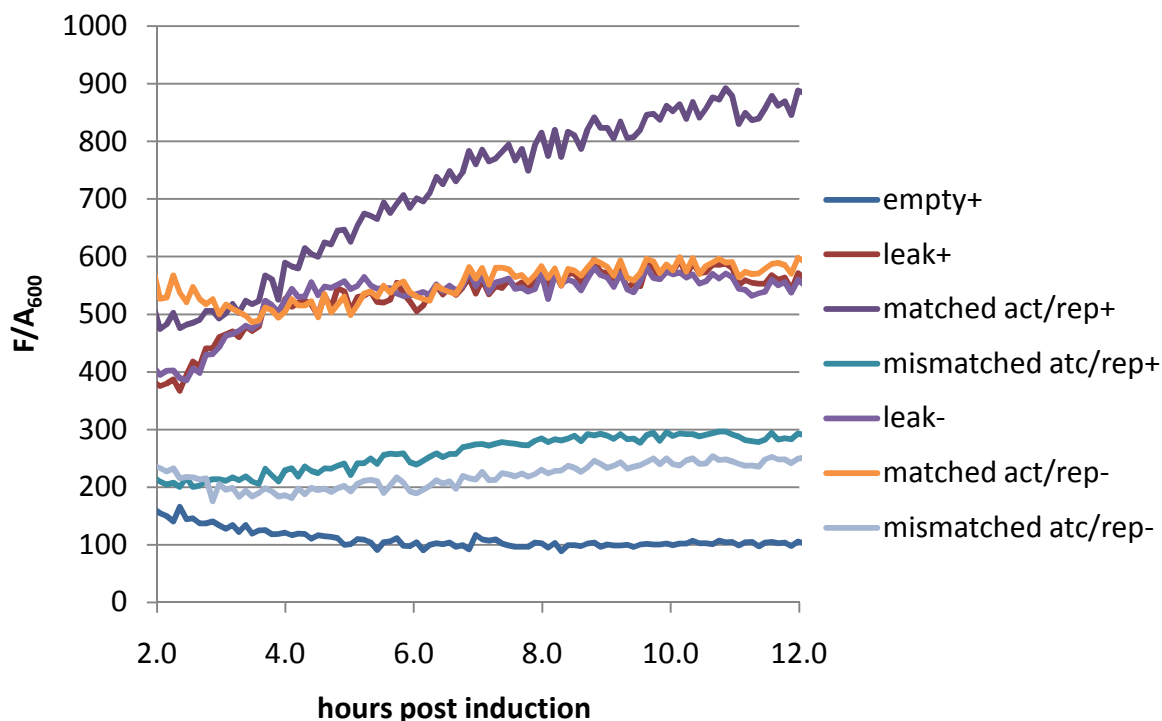


**Figure 4-16 Fold activation as a function of linker. Twelve different linkers were used between the  $\alpha$ -subunit of RNA polymerase and the DNA binding domain ZFP5475. All values are normalized to the fold activation of the original linker.**

These selections are plate based, and only cells that that can activate transcription with the hybrid protein are able to grow. The results of these selections are reported as a number of colonies and not as fold activation. One paper has used a one-hybrid system in *E. coli* to drive the expression of GFP, and they claim fold activations of 8 to 22 fold [39]. But from the data presented it appears that 8 fold is their limit.

To attempt to achieve closer to 8-fold activation, the distance between the promoter and operator site was modulated. Because DNA completes a rotation about every 10.5 base pairs, by adding 0 to 14 base pairs of DNA, the operator site should complete more than 1 rotation around the DNA. One of those 15 positions should present the operator site on the same side of the DNA as the RNA polymerase. Because the previous experiment gave the best results with ZFP03 and ZFP5475, only these two DNA binding domains were used in this experiment. As shown in Figure 4-15, ZFP5475 did not have increased fold activation at other positions. ZFP03 did slightly increase its activation at several positions, with a maximum of 1.86-fold activation with 7 base pairs added, but this was still a lower fold activation than needed.

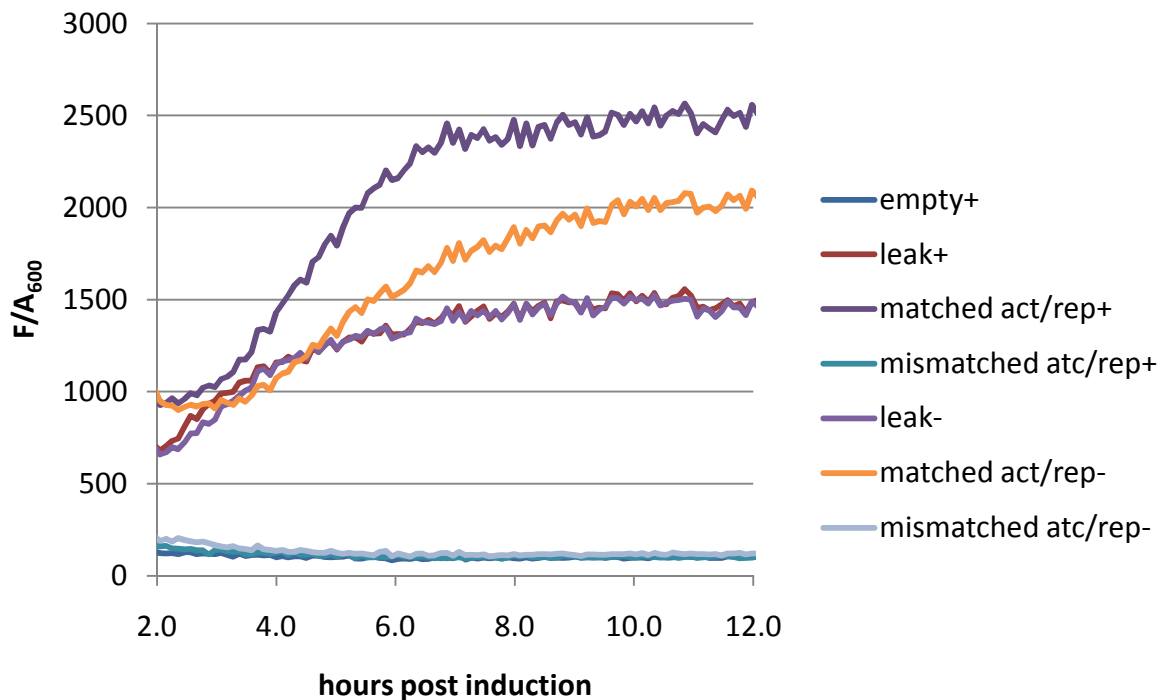
Another possible source of poor alignment between the one-hybrid activator and the RNA polymerase is the linker within the activator which connects the  $\alpha$ -subunit to the zinc finger protein. Several new linkers were designed and tested. The linkers varied in not only length but also composition and can be found in Table 4-2. The first two linkers are from the one-hybrid literature[39], [40]. The remainder of the new linkers were designed to have different lengths and G/S or G/A ratios that promote stable folding of the linker region[77]. The fold activation for each linker was determined at 20 hours post induction. All of the linkers failed to give a higher fold activation than the original linker.



**Figure 4-17 ZFP03 based controls on a p15A plasmid with mCherry as the reporter. In the legend, “-” indicates no induction and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched/mismatched activator and reporter constructs are of type 1.**

Neither moving the operator site nor changing the linker resulted in a large change in the fold activation, but these two variables are not independent. Finding the activation maximum probably requires a full search of these two variables and not just independent excursions along each axis. Because neither variable allowed for much improvement, the 2D search was not expected to find the large improvement that was needed. A poor fold activation can result from the activated state being too low or from the non-activated state being too high. Thus far the experiments have focused on pushing the activated state higher. A significant amount of auto-fluorescence from the *E. coli* made it difficult to assess the lower concentrations of GFP. Subtraction of auto-fluorescence can be tricky as it is not linearly related to  $A_{600}$ . Because non-accurate subtraction of background fluorescence can greatly inflate fold activation numbers, no subtraction of background fluorescence has been performed.  $F/A_{600}$  values have been calculated by dividing the raw F value by  $A_{\text{sample}} - A_{\text{media}}$ . This is quite conservative, but greatly reduces the subjectivity and the possibility of overstating a fold activation.

To reduce the auto-fluorescence and increase the reporter strength GFP was replaced with mCherry (a red fluorescent protein) and the constructs were moved to p15A and ColE1 based plasmids with higher copy numbers. The excitation and emission filters used to measure GFP are similar to the auto-fluorescent range for *E. coli*. mCherry has longer wavelengths for both excitation and emission, and these longer wavelengths have

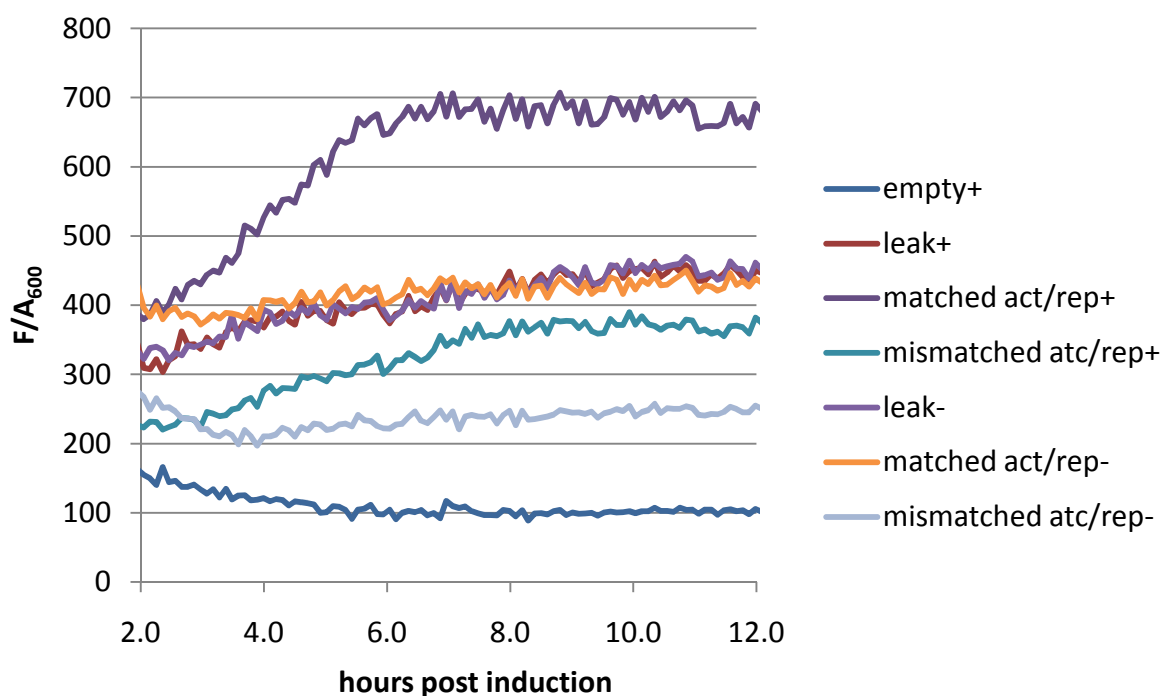


**Figure 4-18 ZFP03 based controls on a Cole1 plasmid with mCherry as the reporter. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched/mismatched activator and reporter constructs are of type 1.**

less overlap with the *E. coli* auto-fluorescence spectrum. Additionally, the constructs were moved to higher copy plasmids as this should increase the ratio between the low fluorescence signal from the reporter and the cell’s auto-fluorescence.

Moving from GFP to mCherry greatly reduced the background fluorescence. In Figure 4-17 the data for ZFP03 constructs on a p15A based plasmid. The empty vector control, representing the auto-fluorescence of the cell, falls well below the leakage controls. This demonstrates the low fold induction was not an artifact of the high auto-fluorescence. The fold activation is still around 1.5. The uninduced matched activator and reporter construct nicely shows the same level of expression as the both the induced and uninduced leakage controls. The mismatched activator and reporter did not behave as expected. First, if the mismatched activator and reporter did not interact, then for both the induced and uninduced cases the mismatched control should behave like the leakage control; but instead, the mismatch controls express mCherry at about half the strength as the leakage control, and there is a slight activation with induction.

The behavior of the ZFP5475 controls on p15A plasmids was similar to the behaviour of the ZFP03 controls on p15A plasmids (Figure 4-19). The mismatched activator and reporter still gave less expression than the leakage controls, but the difference between the uninduced and induced mismatched control was larger. The fold change for the mismatched control was about 1.5. This induction of the mismatched

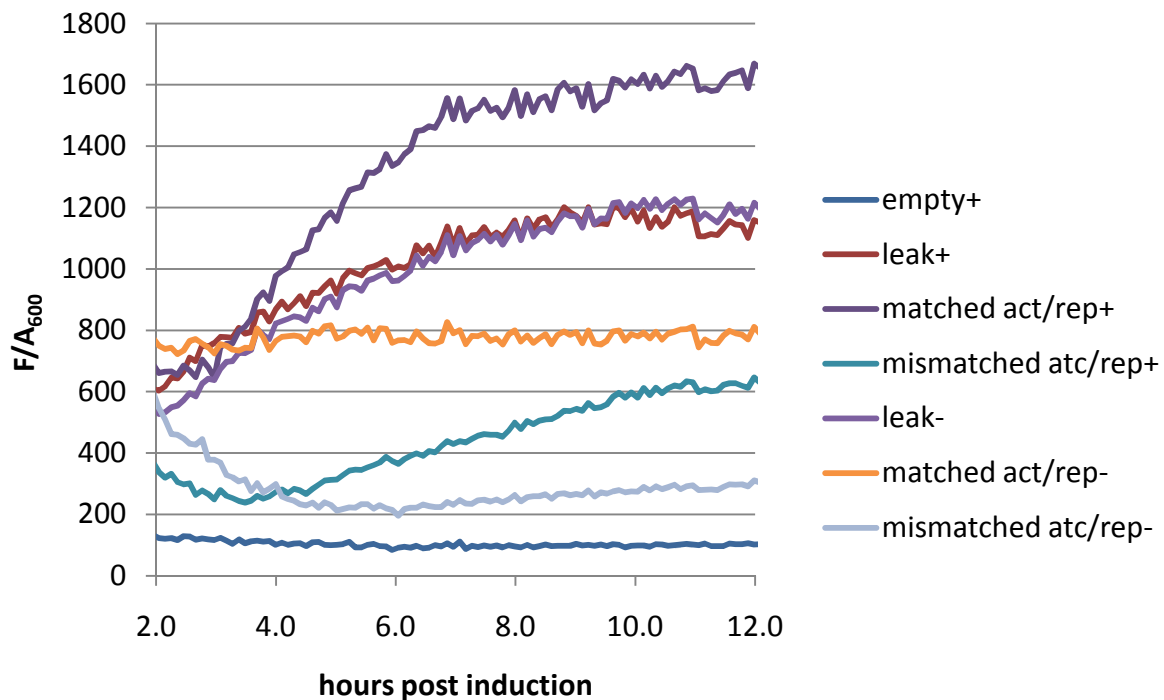


**Figure 4-19 ZFP5475 based controls on a p15A plasmid with mCherry as the reporter. In the legend, “-” indicates no induction and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched/mismatched activator and reporter constructs are of type 1.**

activator and reporter could be due to some non-specific interaction with the  $\alpha$ -subunit. To test this hypothesis a version of the mismatched activator and reporter construct was made without the zinc fingers. This should prevent the extra  $\alpha$ -subunit from specifically activating the reporter. This “free”  $\alpha$ -subunit control displayed a higher activation than the mismatched activator and reporter constructs.

Moving the constructs to a Cole1 plasmid proved rather detrimental to their performance. The ZFP03 constructs are shown in Figure 4-18. The fold activation for the matched activator and reporter decreased because the output from the uninduced culture became significantly higher than the output from the leakage construct. This simply could be that the leakage from the PLtetO-1 promoter, driving the one-hybrid protein, is now high enough to partially activate the reporter. Interestingly, the mCherry expression from both the uninduced and induced mismatched activator and reporter drop to the same level as the empty plasmid.

One-hybrid transcriptional activators continued to be problematic with low fold activation and unexpected behavior from non-cognate activators. Even if this set of controls were tweaked such that they behaved as desired, it is questionable if the system would be able to robustly scale to multiple gates. Transcriptional activator based gates were abandoned when a more promising topology was devised.



**Figure 4-20 ZFP5475 based controls on a ColeE1 plasmid with mCherry as the reporter. In the legend, “-” indicates no induction, and “+” indicates a saturating induction with aTc. For a description of the constructs see Figure 4-8. The matched/mismatched activator and reporter constructs are of type 1.**

### ***Threshold sensing topology using sRNA***

#### **Choice of reference levels**

In electronic based combinational logic there are two power supply voltages, and signals are measured by their voltage relative to these supply voltage. The supply voltages span the range of voltage signals passed by the gates, and the voltage that represents the boundary between a high and low signal is typically close the mid-point between the supply voltages. In a PoPS based logic framework the reference low signal level is obvious, as there is a minimum PoPS value of 0. Defining a reference high level is not as straight forward. The maximum PoPS value of a promoter could be used. As the environmental conditions and metabolic load experienced by the cell varies, this maximum PoPS value could change. If the boundary between high and low states is half this maximum PoPS value, then gates have to be able to dynamically determine this half-way point. This is not an easy task.

Instead of using a promoter as a reference for approximately twice the switching value, the promoter could directly reference the switching value. Such an approach will reduce the complexity of highly robust gates and allows the switching value to be more precisely determined. By definition, a PoPS value should directly correspond to the

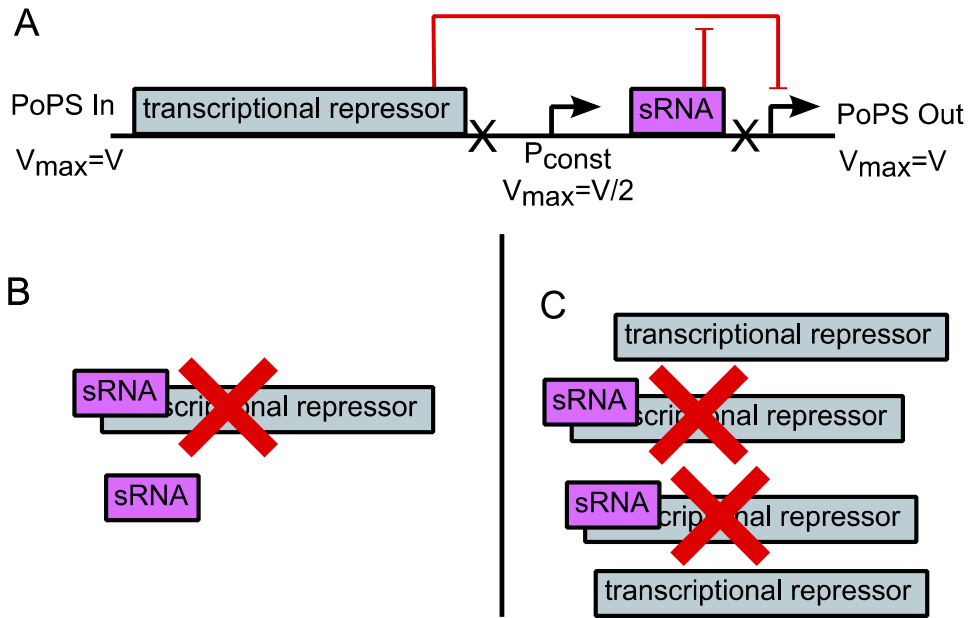
number of mRNA produced by a device. This mRNA then gets translated to protein by a stochastic process. Thus sensing protein levels can provide a relative estimate of the PoPS, but sensing mRNA creation directly gives the true PoPS value. Because of this difference, it would be better to use mRNA rather than protein levels to sense the strength of a promoter.

### Using sRNAs

Recent work on small non-coding RNA (sRNA) indicates they might be well suited for the task of sensing differences in PoPS values. If a sRNA and mRNA contain enough complementary sequence, about 15 base pairs, then they can bind together and form a region of double stranded RNA [78]. If this bond is strong enough or if utilizes all base pairs on the mRNA, leaving no single stranded bases, then the mRNA cannot be translated because the ribosome cannot access the single coding strand. Double stranded RNAs are targeted for degradation and usually have a shorter half-life than their single stranded counterparts [79]. This pairing of mRNA with sRNA and then combined degradation results in a 1:1 removal of mRNA and sRNA. Because mRNA and sRNA binding can happen faster than translation initiation, this process should result in either the mRNA or sRNA existing in a single stranded form. The molecule present in the lower concentration will become completely complexed with the molecule present at the higher concentration; therefore, the presence or absence of single stranded mRNA can be used to indicate the relative strengths of the promoters producing the mRNA and sRNA.

The number of well studied prokaryotic sRNAs is small, and few synthetic sRNAs have been published. Many sRNAs work by binding within the 5'-UTR and altering the accessibility of the RBS. The functionality of these sRNAs is highly dependent on specific RNA secondary structures forming. Because RNA structure prediction has not been perfected, designing this type of sRNA would require some trial and error. A report indicates that sRNA do not have to bind within the 5'-UTR in order to cause reduced translation of the mRNA. Instead the sRNA can bind within the coding region of the mRNA [78]. Designing sRNA to bind within the coding region of an mRNA involves fewer constraints related to secondary structure.

Additionally, sRNAs that bind within the coding region are likely to be a scalable technology. Silent mutations within the coding regions enables the creation of orthogonal sRNA and coding sequence pairs even when the set of coding sequences all translate to the same amino acid sequence. The binding position of the sRNA must cover the 5<sup>th</sup> codon in the mRNA. There is a degree of freedom here too as the 5<sup>th</sup> codon can be placed at various positions within the sRNA. For a sRNA that binds a 15 base pair segment of the mRNA, there will be 13 positions that completely cover the 5<sup>th</sup> codon. Some tricks could also be played to increase the scalability of sRNA orthogonality. For example, additional codons could be added to the N-terminus of the protein so that there is greater possible sequence diversity in this critical region. Care would need to be taken to insure this addition does not disrupt the folding of the remainder of the protein. Knowledge of where the N-terminus of the unmodified protein folds and the inclusion of a glycine-serine linker between the original protein and the new N-terminal could help to maintain activity of the protein. While the scalability of sRNAs has not yet been demonstrated, it appears that 10s of



**Figure 4-21 Overview of the threshold sensing topology using sRNA. A) Schematic of an inverter based on the threshold sensing topology using sRNA. A constitutive sRNA weakly attenuates translation of repressor mRNA. B) For low inputs there are more sRNAs than mRNAs and therefore no mRNA can be translated. C) For high inputs there are more mRNAs than sRNAs. Non-complexed mRNAs are translated.**

orthogonal sRNAs should be possible. This is backed up by the fact that 62 native sRNAs have been found in *E. coli* and experimentally verified [80].

### Overview of topology

The threshold sensing topology is based on a traditional transcriptional repressor based inverter with the addition of a constitutively expressed sRNA that attenuates expression of the transcriptional repressor. The topology is shown in Figure 4-21. The constitutive promoter will create a pool of sRNA. When the pool of mRNA is smaller than the sRNA pool, all of the mRNA will bind with sRNA and there will be no free mRNA available to be translated. When the pool of mRNA is larger than the sRNA pool, all of the sRNA will bind with mRNA and the excess mRNA will be translated. Only once the PoPS input is large enough to push the mRNA levels over the sRNA threshold does the output switch.

The constraints on the transcriptional repressor in this topology are fewer than the constraints when building a traditional repressor based inverter without sRNA. As stated previously, a Hill coefficient of greater than 1 is required to build a traditional repressor based inverter, and a Hill coefficient of 2 or more is needed for good performance. With the threshold sensing topology using sRNA, the sigmoidal shape of the transfer function mostly comes from the interaction between the mRNA and sRNA. Thus transcriptional repressors



with Hill coefficients of only 1 can be used to create well performing gates with this new topology. Therefore single zinc finger proteins, that do not form dimers, can be used to create orthogonal sets of these gates. It was hoped that creating repressor based gates would be easier than activator based ones because there is more of a precedent of wide spread use of engineered repressible promoters (such as  $P_{Ltet0-1}$ ,  $P_{Llac0-1}$ ).

### Modeling and analysis

An ODE model was created to investigate the behavior of this topology.

$$\frac{dm}{dt} = In - \gamma_m m - k_{on} ms \quad 12$$

$$\frac{ds}{dt} = V_R n - \gamma_s s - k_{on} ms \quad 13$$

$$\frac{dr}{dt} = mp - \gamma_r r \quad 14$$

$$0 = \frac{V}{1 + \frac{r}{K_D}} \quad 15$$

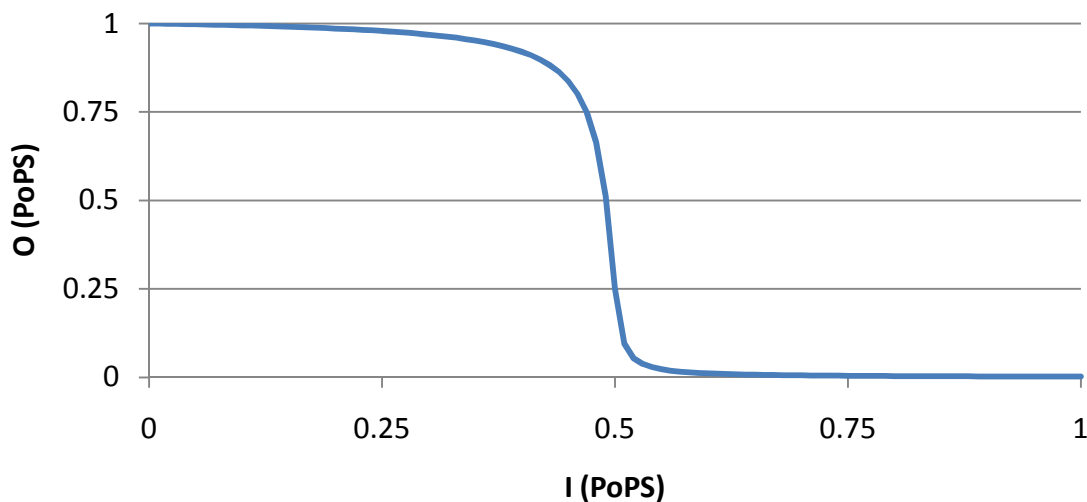
$m$ ,  $s$ , and  $r$  are respectively the concentrations of the transcriptional repressor mRNA, the concentration of the sRNA, and the concentration of the transcriptional repressor protein.  $p$  is the maximum rate of production of repressor protein from one copy of mRNA. Units are  $t^{-1}$ .  $n$  is the concentration DNA containing the device.  $K_D$  is the disassociation constant for the transcriptional repressor, to the promoter.  $I$  is the number of polymerases entering the device per unit of time.  $\gamma_m$ ,  $\gamma_s$  and  $\gamma_r$  are the degradation rate for mRNA, sRNA and repressor protein respectively with units of  $t^{-1}$ .  $k_{on}$  is the apparent on rate for mRNA-sRNA binding.  $V_R$  and  $V$  are the maximum rate of RNA production for the sRNA producing and output promoters respectively.  $O$  is the number of polymerase exiting the device per unit of time.

Generally  $V$  should be twice  $V_R$  in order to equalize the input ranges corresponding to high and low outputs. Such a configuration allows the inverter to maximally tolerate noise while in either state.

The complexes of mRNA and sRNA are assumed to be tightly bound and are rapidly degraded. This simplification means a differential equation is not needed for the complexes.

The equation for  $O$  (Equation 16) is based on a monomeric repressor protein. Modifying this equation for a dimeric repressor or one that forms higher order complexes is easy, but the analysis with a monomer better illustrates the advantages of this topology.

To find the steady state transfer function the ODEs are set equal to zero and then the variables  $m$ ,  $r$  and  $s$  are eliminated. This results in:



**Figure 4-22 Example transfer function of threshold sensing topology using sRNA. Using parameters  $\gamma_s=0.0023 \text{ s}^{-1}$ ,  $\gamma_m=0.012 \text{ s}^{-1}$ ,  $\gamma_r=0.00039 \text{ s}^{-1}$ ,  $K_D=20 \text{ nM}$ ,  $k_{on}=69 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $n=15 \text{ nM}$ ,  $p=0.005 \text{ s}^{-1}$ .**

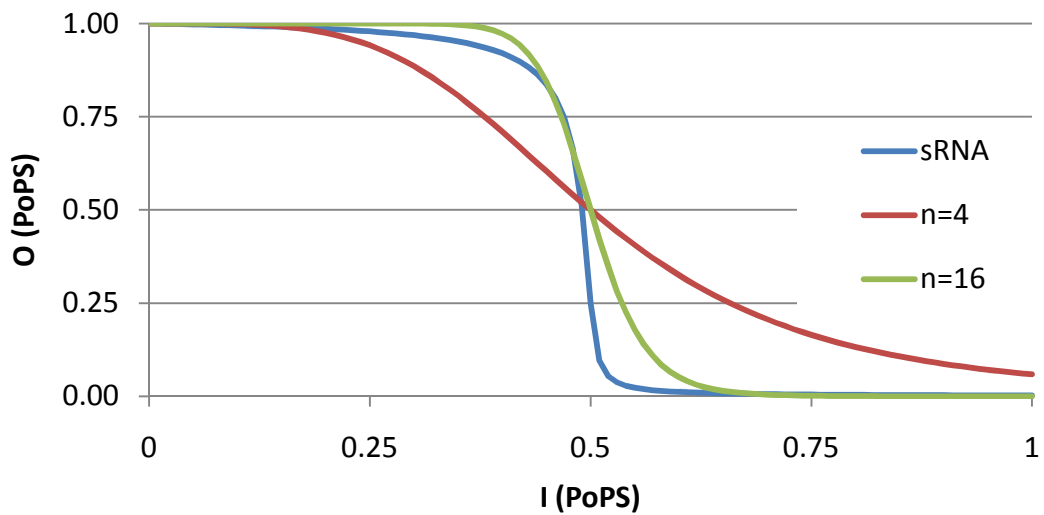
$$O = \frac{2\gamma_m\gamma_r k_{on}V K_D}{p\sqrt{[k_{on}n(I - V_R) + \gamma_m \gamma_s]^2 + 4\gamma_m\gamma_s k_{on}n V_R} + k_{on}n(I - V_R) + \gamma_m(2\gamma_r k_{on}K_D - p\gamma_s)} \quad 16$$

A plot of this function is shown in Figure 4-22. The  $k_{on}$  value is a published value for a specific sRNA-mRNA interaction occurring in the presence of the helper protein Hfq[81]. Hfq is a native *E. coli* protein that has been shown to increase sRNA-RNA annealing rates for several different sRNAs.

This gate topology is quite robust to parameter variations, environmental conditions and noise. This robustness is mostly due to the excellent shape of the transfer function. The transfer function has three characteristics that contribute to its robustness. First, the switching point can be set to the middle of the input range. This maximizes the range of possible reference values that will not lead to an incorrect output. Secondly, the maximum slope of the transfer function is large, and this prevents noise on a signal from propagating to downstream gates. Finally, for most of the input range, the output is close to the minimum and maximum values, and this also helps to suppress noise propagation.

The transfer function does not have the exact form of a Hill equation, but the Hill equation still provides a useful comparison. Traditional gates made with TetR or LacI are expected to have Hill-coefficients of about 2 and 4 because TetR functions as a dimer and LacI functions as a tetramer. With the right parameters, the threshold sensing topology using sRNA can far exceed the maximum slope of a Hill equation with a coefficient of 4 as shown in Figure 4-24.

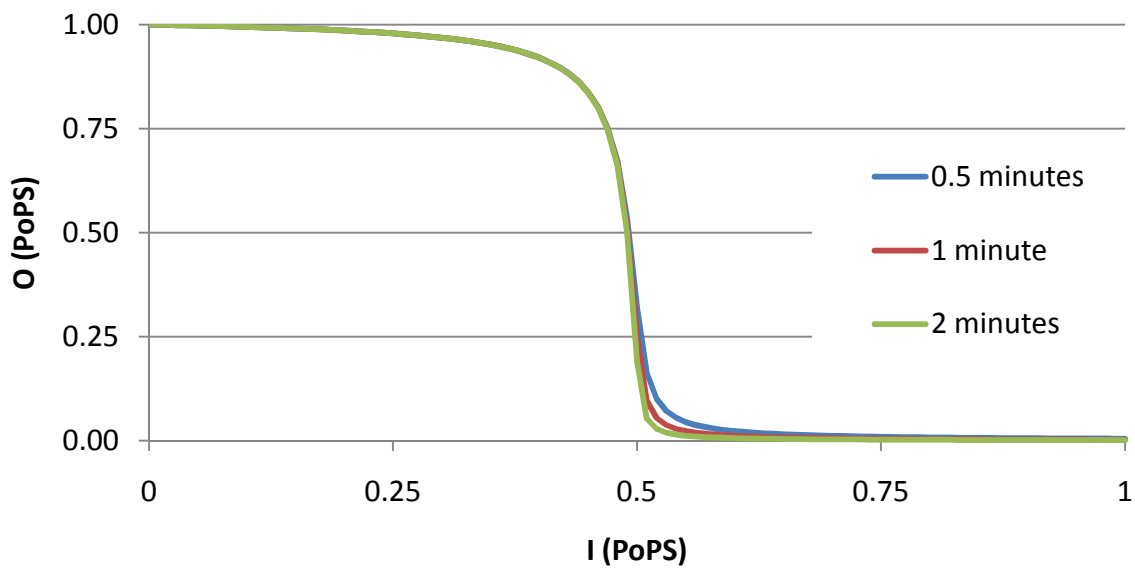
To demonstrate the robustness of the transfer function to parameter variations, the transfer function was plotted with each parameter at the nominal value shown in Figure 4-22 and twice and half that value. The sensitivity to mRNA degradation rate variation is



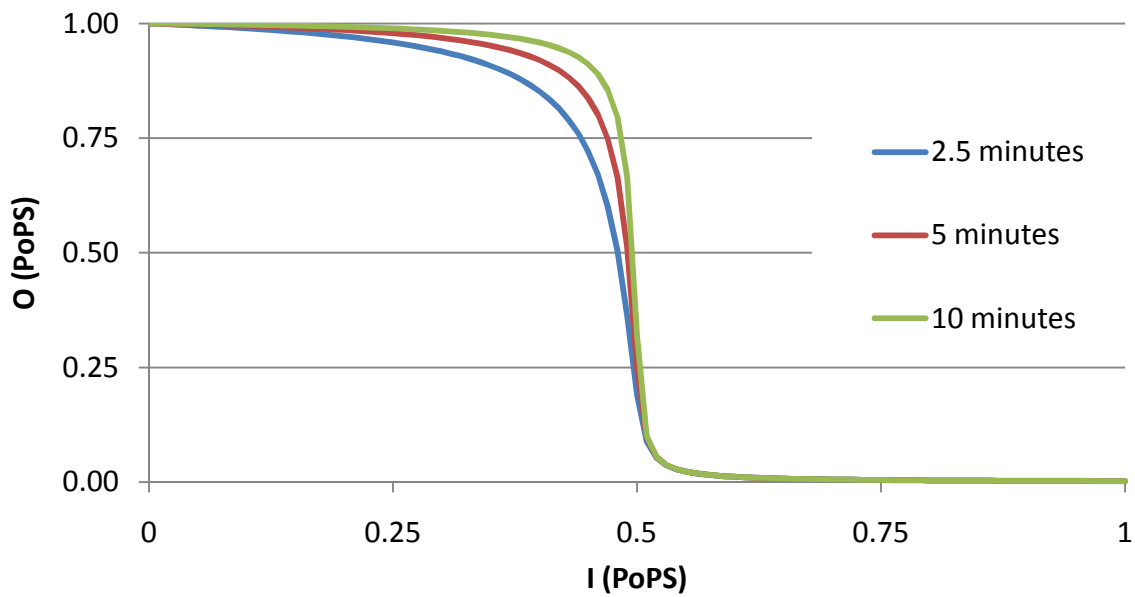
**Figure 4-24 Comparison of Hill-equations with threshold sensing topology using sRNA. Lines labeled with a n parameter are Hill-equations where the n value is the Hill-coefficient used. The line labeled with sRNA represents the threshold sensing topology using sRNA. The parameters used for the sRNA line are the same as in Figure 4-22..**

shown in Figure 4-23. The mRNA half-life must be kept small in order to have a strong off output state.

The sensitivity to sRNA degradation rate variation is shown in Figure 4-25. The sRNA half-life must be large relative to the mRNA half-life for the inverter to have a strong



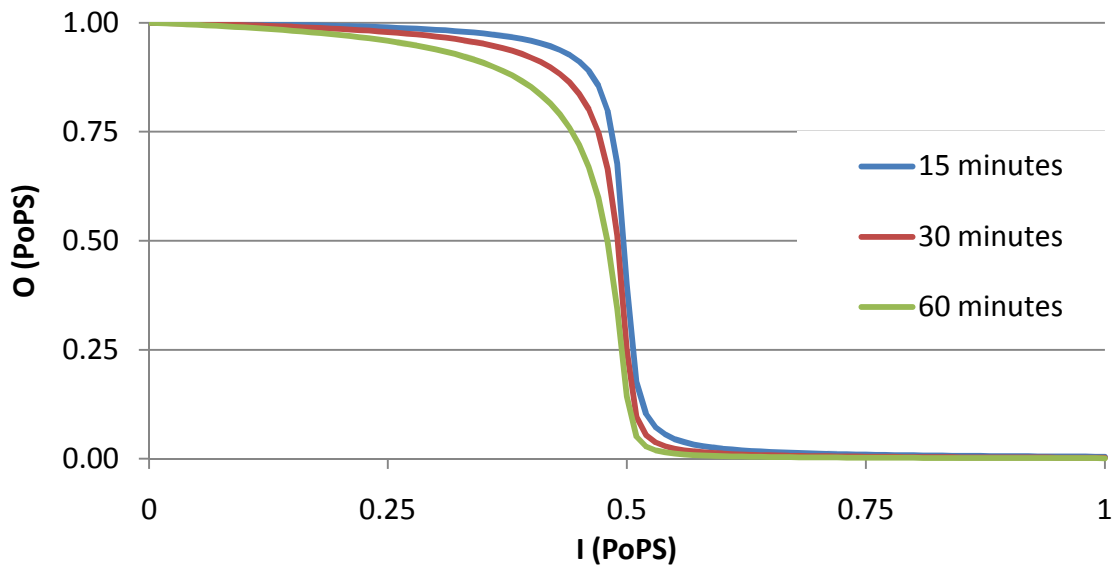
**Figure 4-23 Sensitivity of transfer function to mRNA half-life variation.**



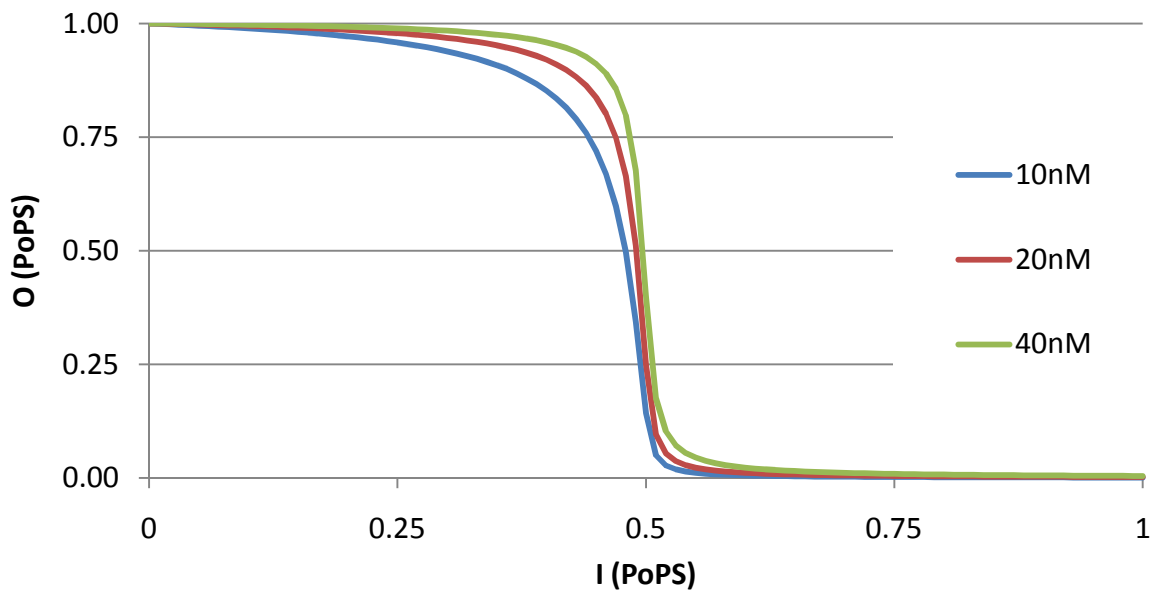
**Figure 4-25 Sensitivity of transfer function to sRNA half-life variation.**

on output state.

The sensitivity to protein degradation rate variation is shown in Figure 4-26. The protein degradation rate affects the steady state transfer function by modulating the size of the steady state pool of repressor protein. A small amount of repressor can get translated when the input is low, and this repressor will stick around longer with low degradation rates.

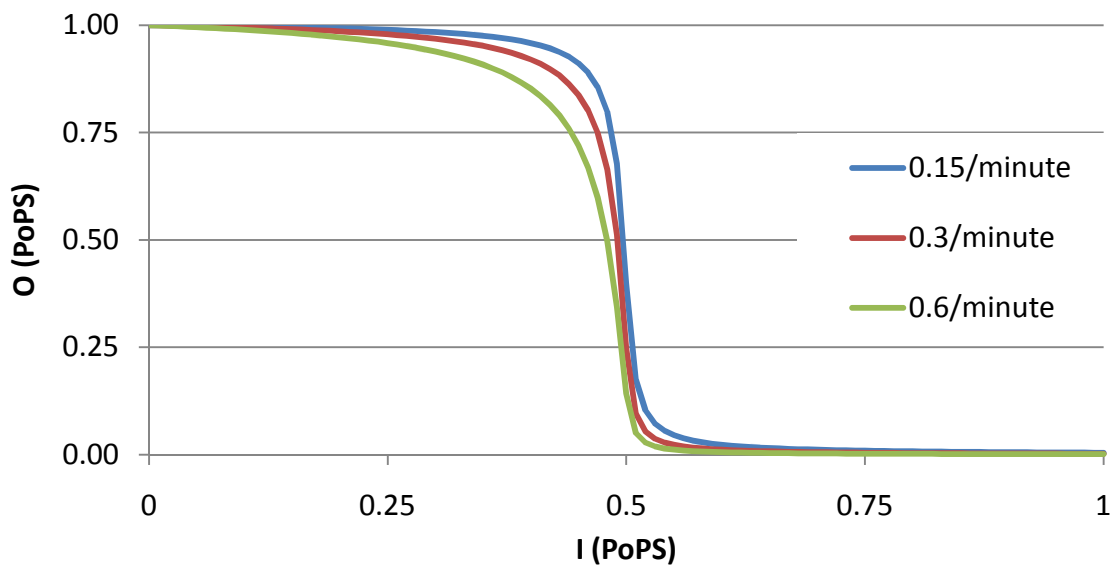


**Figure 4-26 Sensitivity of transfer function to protein half-life variation.**

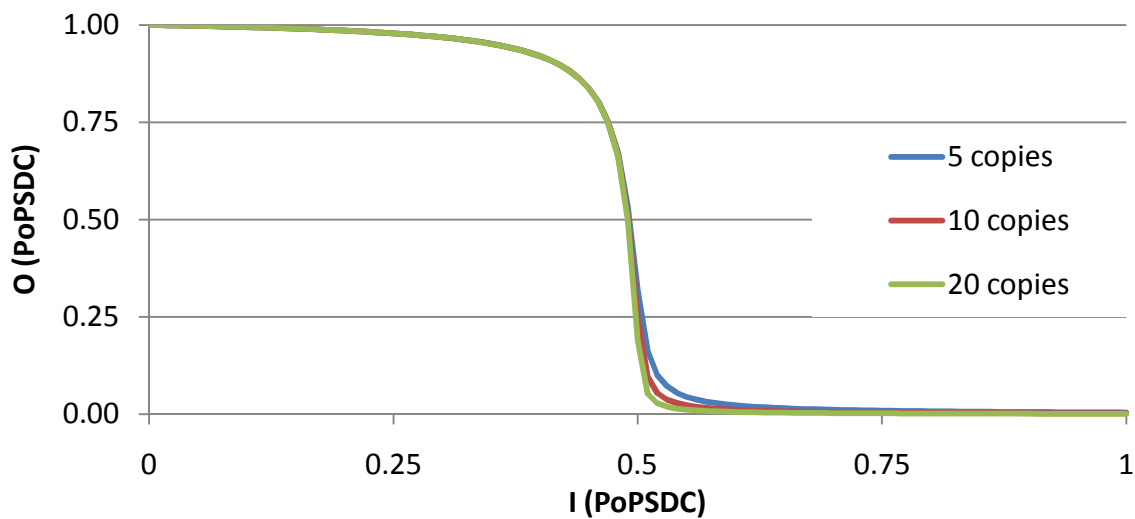


**Figure 4-28 Sensitivity of transfer function to transcriptional repressor  $K_D$  variation.**

The sensitivity to  $K_D$  (protein-DNA dissociation constant) variation is shown in Figure 4-28. Larger  $K_D$  values result in a strong on output state, but slightly weaken the strength of the off output state. A large  $K_D$  value can mitigate a small amount of repressor being translated while the input is low because a larger pool of repressor is required to cause down regulation of the output promoter. A large  $K_D$  weakens the off output state for



**Figure 4-27 Sensitivity of transfer function to translational initiation rate variation. The values are rates of translation initiation for a single mRNA.**



**Figure 4-29 Sensitivity of transfer function to DNA copy number variation. The number of copies listed are per cell values.**

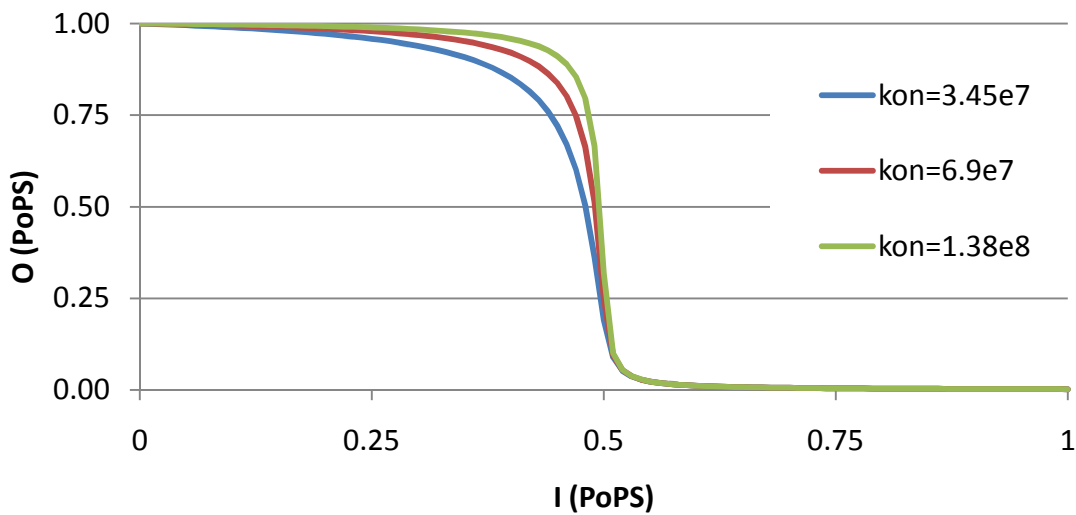
the same reason, but the impact is less because the ribosomes are no longer competing with the sRNA to bind with mRNA.

The sensitivity to translation initiation rate variation is shown in Figure 4-27. Because these transfer functions are based on steady state conditions, an increase in the protein production rate is equivalent to a decrease in the protein degradation. Hence the arguments given above about the protein degradation rates also apply here, but with the impacts on the transfer function reversed.

The sensitivity to DNA copy number variation is shown in Figure 4-29. As discussed in the in the previous chapter, the axes of this plot are in PoPSDC (Polymerase Per Second per DNA Copy) to remove any ambiguity. Copy number has almost no impact on the transfer function. This is because the number of operator sites to repressor is equal to the copy number, but the number of promoters producing mRNA is also equal to the copy number. Because the number of repressors produced per copy number is constant and the number of operator sites per copy number is also constant, thus the PoPS output per copy number will also be constant. Because a continuous ODE model was used here, stochastic effects due to small number of DNA copies are not modeled. In the limit as the copy number approaches 1 DNA copy per cell, these stochastic effects can have a large impact on performance.

The sensitivity to  $k_{on}$  variation is shown in Figure 4-30. The parameter  $k_{on}$  is one of the most important parameters of this system, but also is poorly understood. The sharpness of the transfer function depends on the competition between sRNA and the ribosomes to bind with the mRNA. A higher  $k_{on}$  value leaves less opportunity for the ribosomes to interact with the mRNA. Values of  $k_{on}$  have been measured for some sRNA [62] and many are near  $10^6 \text{ M}^{-1}\text{s}^{-1}$ , but the presence of Hfq can increase some  $k_{on}$  rates by 100 fold [81].

The transfer curves presented thus far are an important part of the topology characterization, but they are only representative of steady state conditions. The dynamics of the gate are also a crucial element of the performance. The primary dynamics parameter



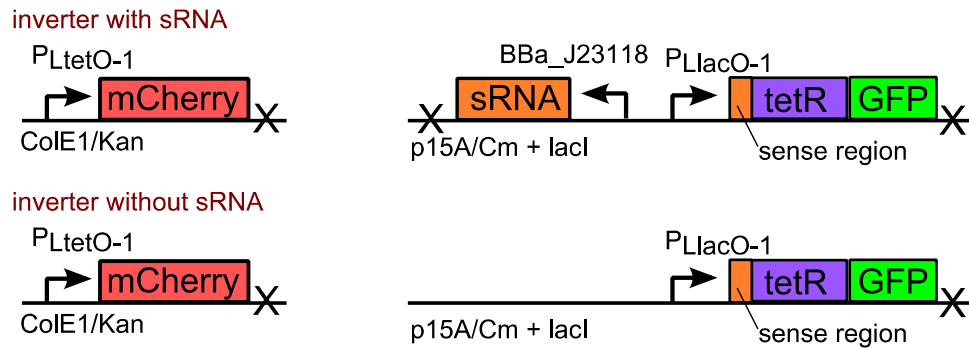
**Figure 4-30 Sensitivity of transfer function to  $k_{on}$  (mRNA-sRNA annealing rate) variation. Units for the legend are  $M^{-1}s^{-1}$ .**

to consider is the delay between the input changing and the output updating to reflect this change. For this topology the dynamics of switching the output in a downward direction is quite different from switching in an upward direction. When the output is falling, enough mRNA must be produced to overcome the pool of sRNA and then the mRNA must be translated to create repressor protein in excess of the  $K_D$ . When the output is rising, any free mRNA must be degraded or bound to sRNA in order to stop the translation of new repressor proteins. Then the pool of existing repressor proteins must be degraded. For most reasonable cases the duration for protein degradation to occur will dominate the switching time.

When a promoter is fully repressed, the concentration of repressor protein concentration will be 100 or more times the  $K_D$ ; therefore the time to lower the repression concentration to below the  $K_D$  can be many times the half-life of the protein. Increasing the protein degradation rate can be expensive to the cell because the protein product rate has to be increased by the same factor in order to maintain the same steady state levels of protein. Thus there is a direct trade off between the speed these gates can propagate signals and the metabolic burden they impose on the cell. This issue can be slightly mitigated when cooperative repressors are used because they can more tightly attenuate a promoter for the same multiple of concentration above the  $K_D$ . In an advance gate design feedback could be implemented to limit the maximum concentration of repressor protein that can accumulate and thereby reduce the duration needed to degrade the repressor protein when switching the output of the gate.

### Implementation and testing

This gate topology was first implemented using TetR as the transcriptional repressor. The topology was conceived with zinc finger based repressors in mind, but as of that time a zinc finger based repressor had not been demonstrated to work in *E. coli*.



**Figure 4-31 sRNA based Inverter construct and sans-sRNA control. BBa\_J23118 is a synthetic constitutive promoter made by Chris Anderson. The sRNA binds to the sense region of the tetR-gfp mRNA.**

An sRNA device and its cognate mRNA was obtained from Julius Lucks and Adam Arkin (not yet published). This device is based on the pT181 sense-antisense pair [62]. The previous description of the topology describes the sRNA as stopping translation of the mRNA, but for pT181 the presence of the sRNA causes an early terminator to form within the mRNA. This early terminator results in transcription stopping before the coding sequence can be transcribed. While this mechanism is different than the one given earlier, the previous ODE model does correctly encompass both mechanisms. The model only requires that mRNA and sRNA bind and then are never translated.

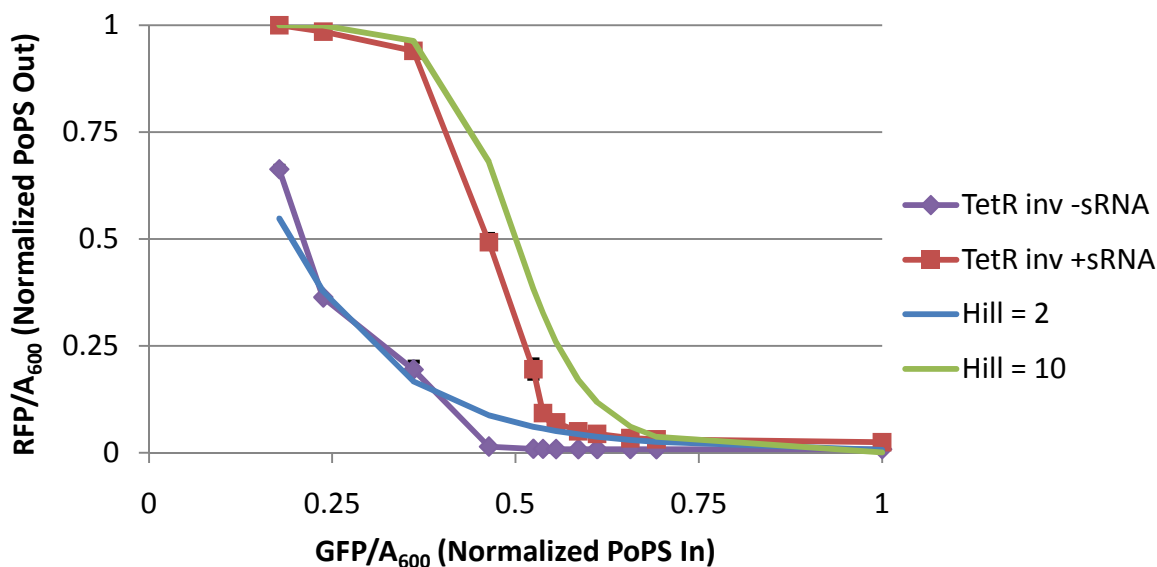
The approach used to find parameters that gave inverter behavior was as follows:

1. Construct TetR based inverter with an inducible  $P_{LlacO-1}$  on the input and mCherry on the output. GFP was also fused to TetR. So that input levels could be monitored.
2. Create diversity in the RBS for TetR-GFP and then screen for an inverter that switches with upon induction of the  $P_{LlacO-1}$  promoter with IPTG.
3. Express the sRNA from a library of constitutive promoters and screen for inverters that retain the ability to switch upon induction of the  $P_{LlacO-1}$  promoter with IPTG.
4. Measure full transfer curves for the inverters that passed the previous screen and pick the one with the best transfer curve.

In the context of the models, it is easy to produce transfer functions of PoPS in to PoPS out. Directly measuring PoPS in a cell is quite difficult, and so fluorescent reporters are used as proxies for PoPS. If the fluorescence of a single reporter molecule is known along with its degradation rate and average number of proteins produced per mRNA, then it is possible to calculate the a PoPS value from the fluorescence values. A more direct route would be to use RT-qPCR to measure the mRNA concentration and then use the mRNA degradation to determine the rate of mRNA production. For this proof of concept work, only relative PoPS values were needed. The extra precision that may be achieved with RT-qPCR was not deemed worth the additional effort required.

Constructs were created for a traditional TetR based inverter and a TetR-sRNA based inverter. The constructs are shown in Figure 4-31. These constructs were exactly the same except for the addition of the sRNA producing operon. The TetR-GFP mRNA is the





**Figure 4-32 Measured transfer function for inverter made with tetR and sRNA. “TetR inv -sRNA” does not contain the sRNA operon, and the “TetR inv +sRNA” construct does contain the sRNA operon. The Hill=2 and Hill=10 lines are Hill equations where the Hill coefficient are 2 and 10 respectively. Due to the leakiness of the P<sub>LlacO-1</sub> promoter, inputs below 0.17 could not be generated with these constructs.**

same for both constructs and includes a sense region in the 5'-UTR where the sRNA can bind. The host strain for these experiments was DH10B.

Cultures for this experiment were grown overnight in 5mL LB medium and then diluted 20:1 into fresh EZ-Rich (Teknova) medium containing 1% glucose, 50 mg/L kanamycin and 33 mg/L chloramphenicol for a total volume of 800  $\mu$ L in a deep 96-well plate. Plates were sealed with Aeresal breathable seals (no. B-100, Excel Scientific). Induction with IPTG was performed at the same time as the dilution. After 24 hours of growth at 37°C, 150  $\mu$ L of culture was transferred to a black wall, clear bottom 96-well plate (Costar) for reading on a SpectraMax M2 plate reader (Molecular Devices).

The measured transfer curve shows the presence of the sRNA shifts the switching point to a higher input as expected (Figure 4-32). The maximum slope of the sRNA-TetR inverter is similar to the slope of a Hill-equation with a Hill-coefficient of 10. The low corner of this curve is sharper than the equation for the Hill-coefficient of 10, but the high corner of the curve performs slightly worse than the equation for the Hill-coefficient of 10. The TetR inverter without sRNA control should have a Hill-coefficient of 2 or less because TetR functions as a dimer. However, the TetR inverter without sRNA had a higher slope than expected. This may be due to nonlinearities in the GFP response that result in distortions of the input axis.

The next step in this work is to demonstrate that a zinc finger protein based transcriptional repressor can be used in place of TetR. Being able to use a zinc finger protein will allow this topology to scale ten or more gates. For at least the near term the limiting factor for scaling will be the number of orthogonal sRNAs.

## Chapter 5 Creating orthogonal sets of repressor-promoters pairs

### ***Orthogonality defined***

Orthogonal sub-systems have behaviors that are uncorrelated to each other. In other words, the state of one sub-system does not impact the state of another orthogonal sub-system. In synthetic biology it is difficult to declare any two sub-systems within a single cell to be completely orthogonal. All sub-systems in a cell draw energy and common co-factors from the same pool, and depleting any of these shared resources will impact all other sub-systems in that cell. Therefore in synthetic biology, orthogonality must be stated in terms of the conditions of the test and the degree of correlation observed. When considering orthogonality for groups of more than two sub-systems it is often useful to define a threshold of orthogonality, where a set of sub-systems is considered orthogonal if all possible pairwise correlation coefficients for that set are below some defined threshold value.

### ***Why orthogonality is needed***

The level of orthogonality required between sub-systems is highly dependent on the functions performed by the sub-systems and the application context they are used in. Some sub-system, such as metabolic pathways, may be able to perform their intended function even when used in tandem with non-orthogonal sub-systems. Generally, the performance of the metabolic pathway will degrade, but not stop functioning completely in the presence of competing pathways. However, the class of sub-systems that contains cell signaling and information processing sub-systems are not as robust to non-orthogonal interactions. These sub-systems generally have a bimodal behavior where the performance is minimally impacted by increasing amounts of non-orthogonality up to a limit. After this limit has been exceeded the performance drops off rapidly and often catastrophically.

Orthogonality not only enables functional ensembles of sub-systems, but also plays an important role in reducing complexity. When sub-systems have material amounts of non-orthogonality an interaction term must be included between every pair of sub-systems. As the number of sub-systems increases, the number of interaction terms goes up with the square of the number of sub-systems, therefore, if non-orthogonal interaction terms can be kept small enough that they can be safely ignored, then a large amount of complexity can be dropped from the system. Removing this complexity gives an engineer a greater ability to think about the behavior of a larger system than would be possible if the interaction terms had to be considered. The reduced complexity also simplifies the design process, resulting in faster design cycles.

### ***Desired specification for sets of orthogonal promoter-repressor pairs***

The primary goal of creating orthogonal promoter-repressor pairs is to use them to implement logic gates. In this context orthogonality mainly implies that the transcriptional repressor from one promoter-repressor pair will not attenuate transcription from the promoter of a second promoter-repressor pair. The degree of orthogonality necessary for logic gates is not fixed. Generally a logic gate can tolerate a specific amount of non-orthogonal interactions. The total amount of non-orthogonal interactions a gate is

subjected to is the sum of the individual non-orthogonal interactions the gate has with every other gate in the system. This implies that the strength of the individual non-orthogonal interactions must decrease as the number of gates in the system increases.

The degree of orthogonality required is also a function of the fidelity required from the logic block being implemented. If the logic block specifications allow for the output to be incorrect for some percentage of inputs, then the orthogonality requirements can be reduced. Consider the case of a system involving 10 promoter-repressor pairs. The worst case scenario in terms of orthogonality is that 9 of the 10 repressors are being expressed. The promoter whose cognate repressor is not being expressed will be subjected to the non-orthogonal interactions with the 9 other repressors. There are a total of  $2^{10}=1024$  states that can be encoded by the presence or absence of 10 repressor proteins. Of these 1024 states, 10 of them, or less than 1%, are of this worst case class where 9 of 10 repressors are expressed. If the system is designed where it produces an incorrect output if 9 of the 10 repressors are expressed at once, but gives the correct output for cases where 8 of the 10 repressors are expressed at once, then the amount of orthogonality required for each gate can be decreased.

This tradeoff between orthogonality and fidelity can be taken further, but the number of states that give incorrect outputs grows faster than the orthogonality requirements decrease. For example, if the system is allowed to give incorrect outputs for cases when 9 or 8 of the 10 repressors are expressed, then the number of states where 9 or 8 of the repressors are expressed is given by:

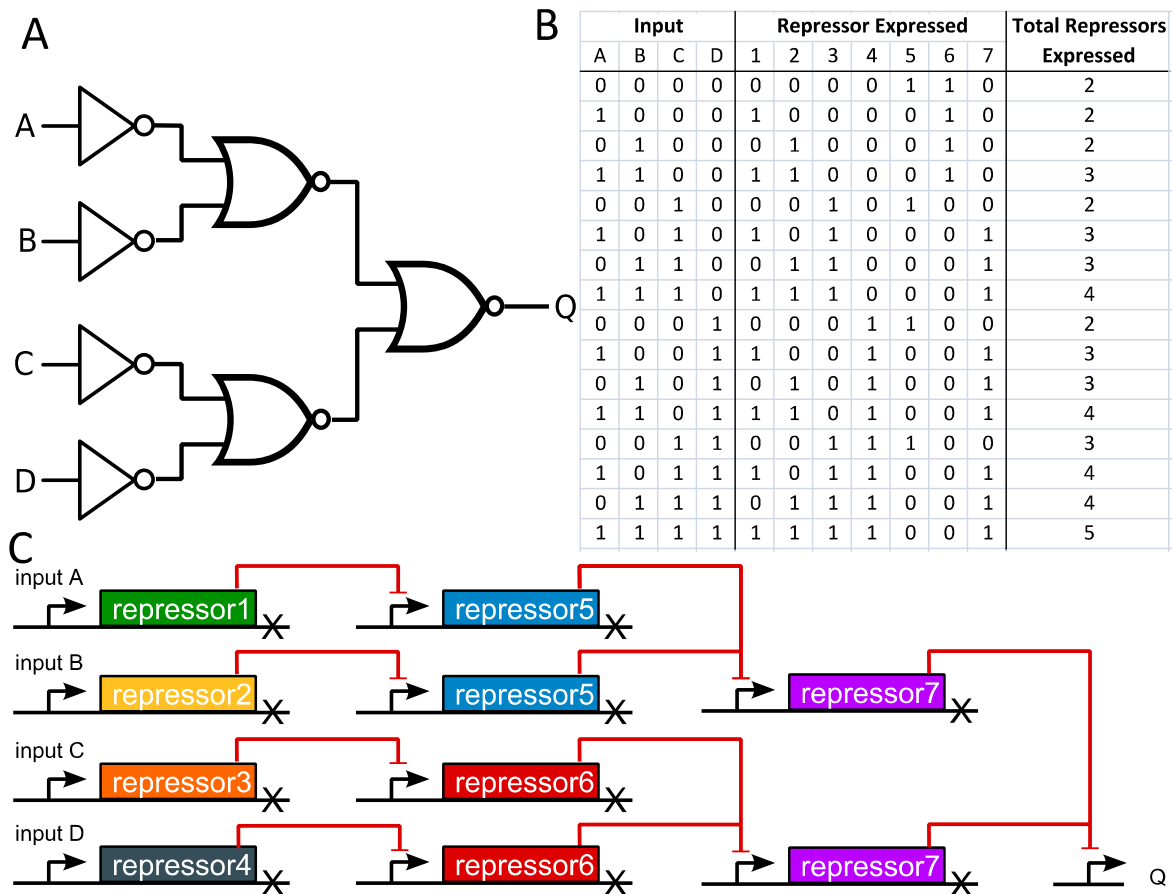
$$\binom{10}{9} + \binom{10}{8} = 10 + 45 = 55 \quad 17$$

This is over 5% of the possible 1024 states.

Additionally, in most logic networks the combination of states taken on by the individual gates is a subset of the possible states the individual gates can take on when not connected. These unutilized states can allow for reduced orthogonality between a set of gates without a reduction of fidelity. For example, for the 4 inverter, 3 NOR gate circuit shown schematically in Figure 5-1A and as a genetic network in Figure 5-1C, no more than 5 of the 7 repressors proteins will ever be expressed simultaneously as shown in the table in Figure 5-1B. The worst possible condition for a generalized 7 repressor circuit is 6 of the 7 repressors being simultaneously expressed, but for this specific circuit that case cannot happen. With this knowledge, the allowable strength of interaction between non-cognate pairs can be increased.

In summary, the degree of orthogonality required for logic circuits will depend on the fidelity required by the application and the specific logic network being implemented. Generally the strength of non-cognate interactions must decrease as the size of the logic network increases. When designing sets of orthogonal promoter-repressor pairs for general use, the strength of the non-cognate interactions relative to the cognate interactions should be as small as possible.

For the specific goal of constructing a 5 or more gate logic block as a technology demonstration, the ceiling for the allowable orthogonality can be estimated. If the sum of non-cognate attenuation values for a promoter exceeds 50% of the promoter's maximal activity, then the output of the gate may switch incorrectly. The choice of a 50% limit



**Figure 5-1** Not all combinations of repressor states can be reached by a logic circuit. **A)** An example logic circuit with 4 inverters and 3 NOR gates. **B)** The truth table for the circuit in **A** along with the state of each of the repressors proteins for every possible set of inputs. **C)** A genetic network that implements the same circuit as **A**.

assumes the gate's switching threshold is centered in the gate's input range and that none of the cognate repressor is present. Neither of these assumptions is likely to be true, but other parts of this analysis are overly conservative and offset these liberal assumptions. For a 5 repressor system, there are 4 non-cognate repressors. For each promoter the sum of the attenuation percentages from the non-cognate repressors should sum to less than 50%. Assuming all non-cognate interaction strengths are equal, then a non-cognate repressor should cause 12.5% of attenuate or less.

The above calculation does not place any constraints on the strength of cognate repression. There is a simple constraint of the cognate repressor causing over 50% repression in order for the output to switch. Beyond orthogonality, the output range of each gate ideally would be equal. This implies the maximum activity of each promoter in the set should be similar.

The strength of the promoters also should be strong enough to produce pools of proteins 10 times larger than the  $K_D$  of the repressors ( $\sim 10\text{nM}$ ) while cells are growing exponentially (30 minutes/cell division  $\rightarrow$  effective degradation rate of  $3.9 \times 10^{-4}\text{s}^{-1}$ ).

$$\frac{10K_D\gamma}{\text{cell volume}} = \frac{10(10^{-8}\text{M})3.9 \times 10^{-4}\text{s}^{-1}}{1.5 \times 10^{15}\text{L}} = 0.026 \text{ proteins/second/cell} \quad \mathbf{18}$$

Depending on the number of proteins generated per mRNA, the required PoPS per cell should also be near this number. This represents a weak promoter. Too strong of a promoter is also detrimental as the metabolic burden on the cell will increase with promoter strength and proteins are more likely to be toxic at high expression levels. Placing an upper limit on the promoter strength is hard because there are limited analytical tools for investigating metabolic burden.

### ***Creating orthogonal sets***

Testing for orthogonality between  $N$  pairs of items generally requires  $N^2$  evaluations. Because the number of evaluations grows quickly, there is a tradeoff between keeping the number of evaluations in the orthogonality experiment small and searching a large enough library of pairs to find the needed orthogonality. Assaying promoter-repressor pairs for orthogonality was done on 96-well plates, therefore it made sense to pick a library size based on how many pairs can be tested by an  $A \times B$  number of 96-well plates. To reduce pipetting errors, the plates were tiled and the utilized wells were all in an  $N \times N$  square. Therefore  $9 \times 9$  test would use 2 96-well plates (because they are limited to 8 wells in one direction) even though one plate contains enough wells. Based on plate geometry,  $8 \times 8$ ,  $16 \times 16$ ,  $24 \times 24$ ,  $32 \times 32$  and  $40 \times 40$  sized squares are efficient choices. Setting up a  $40 \times 40$  sized experiment without using robotics would be fatiguing and push the limits of what one person can complete on each day of the experiment. A  $32 \times 32$  square (12 plates) experiment appeared as though it could be carried out by a single person and hopefully would be large enough to get 5 pairs with good orthogonality.

### **Choosing zinc finger proteins**

32 zinc finger proteins from the steric hindrance experiment were selected based on their ability to strongly attenuate transcription in that experiment. This provided experimental validation that these zinc finger proteins are able to bind to the operator site they were designed for and that they are not highly toxic when expressed at low levels. Because pre-existing zinc finger proteins were used, the set of operator sites was also fixed. This limits the placement of the operator sites to areas mostly outside of the -35 and -10 regions.

### **Choosing promoters**

A set of orthogonal promoter-repressor pairs could be constructed from a set of promoters that only differ by the operator site for the zinc finger protein. However, if the end goal is to use the orthogonal promoters within a single cell, then minimizing sequence homology between the promoters is important for reducing homologous recombination and maintaining the desired DNA sequence. To achieve some sequence diversity and

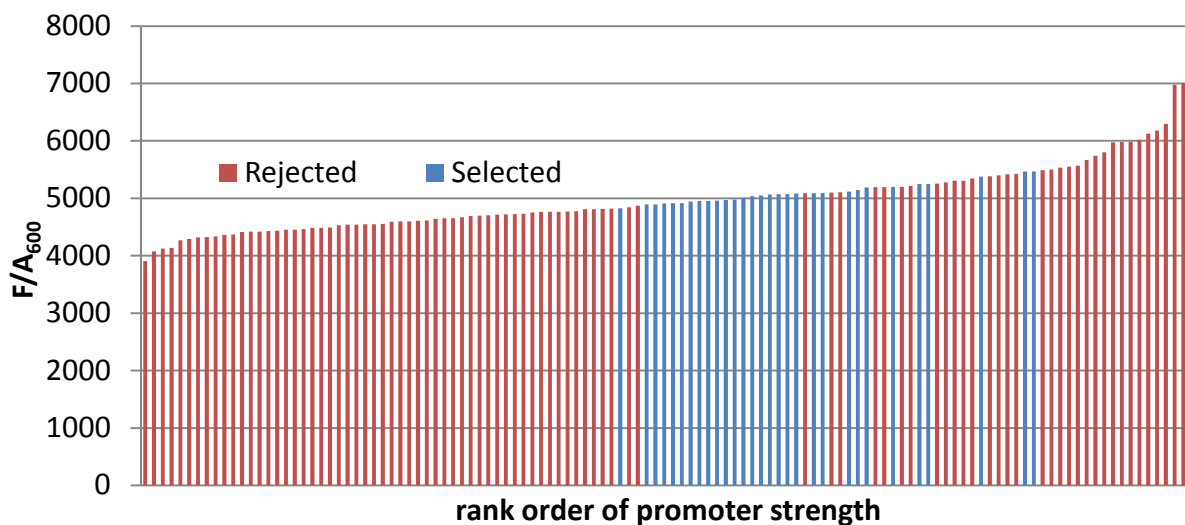
**Table 5-1 BioFAB promoters**

RPL #	Promoter sequence
66	TTGACAATTAATCATCGGCTCATATGGTCTGTGGA
69	TTGACAATTAATCATCGGCTCGTAGGGTATGTGGA
81	TTGACAATTAATCTCCGGCTCTACGGTATGTGGA
82	TTGACAATTAATCATCGGCTCCTATGGTGTGTGGA
83	TTGACAATTAATCATCGGCTCATAACCTTTGTGGA
104	TTGACAATTAATCATCGGCTCTTAGGTTCTGTGGA
106	TTGACAATTAATCATCGGCTCGTAGATTTTGTGGA
125	TTGACAATTAATCATCCGGCTCATAGCGTGTGTGGA
126	TTGACAATTAATCATCGGCTCGTAGGCTGTGTGGA

obtain similar maximum promoter activities, a set of promoters were picked. These promoters came from the BioFAB's Randomized Promoter Library (RPL) version 1. A set of 9 promoters with a similar but low constitutive activities were selected. These promoters are listed in Table 5-1. The 32 operators corresponding to the pool of 32 zinc finger proteins were cloned into the 9 promoters directly upstream of the -35 region. These constructs were built on the BioFAB plasmid pFAB217. The complete combinatorial set of  $9 \times 32 = 288$  promoter-operators was attempted. A randomly picked colony from each cloning reaction was sequenced and 120 correct insertions were found, representing 30 of the 32 operator sites. This was deemed a large enough pool to move forward with.

The decision to place the operator just upstream of the -35 region was made before the results of the steric hindrance experiments had concluded. In the steric hindrance experiments this corresponds to the operator at position -54. The ZFP-54 was able to cause strong attenuation at this position, but in general the region around -54 had only moderate repression. Better approaches are certainly possible if a new set of zinc finger proteins were constructed, as then they could overlap the -35 region and extend downstream from there. Alternatively, it may be possible to use the current set of zinc finger proteins and place the operators between the -35 and -10 regions. This may involve extending this region from 17 to 18 base pairs. From the steric hindrance data, it appears that binding directly downstream of the -10 region could give good results. However, placing an 18bp operator downstream of the -10 region means the operator will extend into the sequence coding for the mRNA. This is not desirable from an engineering perspective because it restricts the flexibility to place any mRNA coding sequence downstream of the promoter. This also violates the spirit of a PoPS interface as nothing but the polymerase should cross the interface between the two connected components.

Constitutive activity of the promoter library was measured and the results are shown in Figure 5-2. From this data 30 promoters were chosen. This set contains all 30 of the zinc finger protein operator sites that were successfully cloned and has a low level of variation in promoter strength.



**Figure 5-2 Strength of 120 promoters were measured. 30 promoters with similar strength were selected such that none of these 30 promoters contain the same zinc finger protein operator site.**

### **Orthogonality assay**

The selected promoters were moved to the same mCherry-based reporter plasmid that was used for the steric hindrance of RNA polymerase experiments. 3 of the 30 promoters failed to clone into this reporter plasmid and were dropped from the study. Once again the DP10 strain was used [55]. Even with a saturating induction, the DP10 strain is useful because it removes the up regulation of arabinose import upon induction. This positive feedback loop can result in a heterogeneous induction when a subset of the population sequesters all the arabinose.

Each evaluation of a promoter-repressor pairing required transforming a zinc finger expression plasmid and a reporter plasmid into *E. coli*. For a  $27 \times 27$  size orthogonality experiment, this required 729 transformations. To save time and effort these transformations were performed without plating on solid media. The sources of the plasmids going into the transformations were from mini-preps of plasmids that had been previously sequence verified, thus there was no need to obtain single colonies.

Initially a double transformation of the reporter with a ColE1 origin and the zinc finger protein expression plasmid with a SC101\*\* was attempted with chemically competent DP10 cells. After growing the transformation in LB and antibiotics overnight, the  $A_{600}$  of the cultures was measured to verify the cultures were growing and the transformations were successful. This method was tried several times with various tweaks to the transformation protocol, but the transformation success rate never exceeded 65%. This was likely due to the SC101\*\* plasmid. When plasmids based on this vector were transformed and plated on solid media, colonies would not appear until 18 to 24 hours after the transformation. Other plasmids often show colonies in as little as 12 hours. This

difference may be due to the antibiotic cassette on the SC101\*\* plasmid not having been optimized for such low copy numbers.

The successful method was to perform single plasmid transformations to place the SC101\*\* plasmid in DP10 to make 27 strains, and then make these 27 strains competent and transform in the ColeE1 reporter strain. The transformation protocol used to do the ColeE1 transformations without solid media was performed as follows:

1. Prepare and freeze chemically competent cells[82]
2. Place sterile deep 96-well plates on ice
3. Thaw competent cells on ice
4. Dispense 100µL into each well of a deep 96-well plate
5. Pipette 100ng of plasmid into each well
6. Leave 96-well plates on ice for 10 minutes

**Table 5-2 27×27 orthogonality matrix, left hand half.**

promoter	op	zinc finger protein												
		-57	-54	-51	-49	-48	-45	-44	-43	-40	-37	-33	-31	-30
83	-57	.12	.24	.12	.92	.37	.73	.89	.88	.95	.89	.81	.9	.38
82	-54	.44	.31	.35	1.	.33	.71	.85	.99	.89	.78	.82	.83	.68
125	-51	.37	.44	.3	.76	.39	.52	.85	.73	.82	.78	.82	.8	.67
125	-49	.56	.57	.49	.63	.43	.57	.85	.73	.87	.82	.79	.81	.74
82	-48	.45	.59	.38	.81	.43	.41	.7	.56	.75	.77	.84	.85	.73
83	-45	.42	.26	.52	.75	.59	.52	.8	.59	.73	.73	.71	.72	.75
126	-44	.67	.75	.59	.77	.6	.62	.69	.57	.62	.74	.67	.67	.8
69	-43	.66	.77	.71	.72	.67	.69	.69	.3	.64	.7	.75	.73	.75
125	-40	.67	.74	.71	.81	.7	.88	.75	.52	.11	.45	.71	.76	.78
69	-37	.7	.77	.73	.75	.65	.7	.81	.71	.47	.2	.7	.61	.94
83	-33	.77	.83	.72	.81	.73	.75	.78	.8	.82	.77	.6	.5	.79
83	-31	.84	.86	.85	.9	.8	.81	.91	.86	.88	.88	.52	.42	.51
83	-30	.78	.82	.75	.84	.74	.75	.8	.83	.82	.78	.6	.47	.14
83	-28	.89	.88	.79	.9	.74	.87	.86	.93	1.	.81	.89	.67	.9
83	-27	.8	.77	.63	.73	.58	.79	.69	.7	.8	.8	.66	.63	.38
104	-25	.84	.86	.81	.73	.74	.87	.78	.89	.96	.88	.64	.72	.33
82	-24	.48	.89	.77	.88	.85	.86	.87	.82	.84	.97	.83	1.	.49
106	-22	.61	.95	.79	.96	.77	1.	.87	.85	.88	.9	.84	.89	.59
69	-18	.96	.91	.71	.92	.8	.87	.97	.93	.92	.92	.97	.83	.87
83	-16	.96	.89	.7	.82	.72	.87	.87	.96	.88	.88	.9	.85	.92
83	-13	1.	.98	.75	.97	.76	.91	.86	.93	.94	.91	.9	.94	.92
125	-12	.96	.94	.76	1.	.76	.93	.94	.95	.95	.96	.87	1.	.96
104	-10	.99	1.	.45	.89	.57	.92	.62	.9	.89	.95	.67	.74	.56
83	-9	.89	.92	.81	.91	.78	.83	.92	.95	.91	.87	.87	.84	.95
66	-6	.56	.69	.55	.46	.54	.58	.57	.52	.64	.49	.65	.51	.59
83	-3	.78	.82	.74	.99	.67	.76	.95	.97	.95	.97	.94	.91	.88
126	+4	.69	.77	.65	.89	.61	.69	.92	.79	.9	.94	.9	.89	1.



7. Seal 96-well plate with Aeroseal breathable seals (no. B-100, Excel Scientific)
8. Incubate at 37°C with shaking for 1 hour
9. Add 900µL of LB media and antibiotics to each well
10. Re-seal plates and incubate at 37°C with shaking overnight

From this point on the assay is the same as the steric hindrance assay. Only saturating inductions were performed with 20mM arabinose, and F/A<sub>600</sub> levels were measured 24 hours after dilution and induction into EZ-Rich media (Teknova).

### Results and analysis

The data is shown in Table 5-2 and Table 5-3. Each row of these tables corresponds to a single promoter and each column belongs to a zinc finger protein. The promoter

**Table 5-3 27×27 orthogonality matrix, left hand half.**

zinc finger protein													promoter	op	
-28	-27	-25	-24	-22	-18	-16	-13	-12	-10	-9	-6	-3			4
.62	.58	.88	.71	.77	.89	.85	.84	.85	.85	1.	.71	.5	.58	83	-57
.86	.73	.93	.63	.84	.91	.89	.86	.81	.77	.92	.77	.75	.76	82	-54
.87	.66	.78	.6	.78	.73	.74	.76	.76	.76	.92	1.	.65	.73	125	-51
.92	.74	.83	.68	.72	.8	.82	.79	.78	.78	.86	1.	.65	.87	125	-49
.67	.74	.72	.54	.83	.82	.5	.8	.66	.8	.97	1.	.62	.93	82	-48
.74	.79	.81	.81	.42	.81	.7	.77	.71	.76	.92	1.	.7	.88	83	-45
.69	.67	.66	.7	.66	.72	.74	.68	.71	.78	.85	.98	.76	1.	126	-44
.77	.78	.82	.75	.71	.66	.74	.79	.66	.79	.9	.92	.67	1.	69	-43
.83	.92	.82	.89	.76	.8	.84	.75	.82	.85	1.	.94	.76	.87	125	-40
.8	.93	.91	.86	.85	.81	.83	.83	.78	.79	.9	1.	.85	.89	69	-37
.64	.75	.76	.78	.75	.75	.75	.83	.81	.74	.89	1.	.79	.84	83	-33
.68	.6	.88	.57	.86	.86	.84	.85	.81	.82	1.	.94	.83	.87	83	-31
.15	.36	.73	.24	.77	.76	.83	.83	.8	.85	1.	.99	.76	.64	83	-30
.27	.7	.75	.62	.87	.81	.93	.88	.84	.92	.99	.98	.73	.73	83	-28
.29	.42	.35	.46	.64	.75	.8	.75	.71	.74	.83	1.	.78	.78	83	-27
.14	.53	.4	.37	.74	.92	.81	.82	.85	.82	.93	1.	.9	.91	104	-25
.73	.57	.71	.38	.65	.73	.73	.67	.73	.78	.99	.99	.84	.85	82	-24
.86	.64	.8	.52	.63	.7	.82	.86	.82	.85	.97	.98	.91	.91	106	-22
.89	.88	.85	.82	.88	.2	.8	.86	.86	.84	.95	1.	.88	.94	69	-18
.84	.87	.76	.78	.82	.78	.64	.83	.72	.8	.8	1.	.82	.97	83	-16
.91	.83	.83	.89	.78	.8	.81	.62	.22	.81	.4	.86	.84	.88	83	-13
.8	1.	.89	.93	.87	.84	.9	.89	.25	.81	.42	.86	.87	.93	125	-12
.83	.62	.69	.74	.72	.67	.6	1.	.48	.64	.39	.59	.54	.57	104	-10
.83	.94	.84	.84	.79	.79	.76	.81	.4	.61	.27	.62	.81	1.	83	-9
.69	.67	.66	.66	.66	.59	1.	.55	.49	.53	.46	.27	.2	.57	66	-6
.95	1.	.94	.85	.87	.84	.87	.84	.78	.79	.66	.37	.08	.27	83	-3
.94	.76	.88	.83	.91	.85	.83	.91	.63	.86	.87	.94	.33	.29	126	4

column lists the base promoter that the operator was added to. These number reference Table 5-1. The operator column lists the operator that is upstream of the -35 region of the promoter. These values are from the  $P_{ZC35}$  promoter in Chapter 3 and indicate the most upstream base pair of the operator site relative to the transcriptional start site. The numeric column headings also reference the  $P_{ZC35}$  promoter. These numbers tell what operator size along  $P_{Z35}$  the zinc finger protein was designed to bind.

The numeric data points at the intersection of a row and column are normalized  $F/A_{600}$  values, and they are normalized to the maximum value found in that row. This assumes that for each promoter tested, at least one of the 27 zinc finger proteins tested will not interact with that promoter. This is likely true, but a no zinc finger protein control should be included for each promoter if this experiment is repeated. Each cell of the table is shaded proportionally to the value it contains. The shading for a value of 1 is completely white and for a value of 0 the shading is dark blue/grey. Locations on the table that represent a cognate pair have a black boarder. These locations all fall on a diagonal line. Ideally this diagonal line would only contain values of 0 (dark shading) and the off diagonal locations should all contain values of 1 (white).

In the steric hindrance experiment where the zinc finger protein was expressed using  $P_{BAD}$ , one third of the zinc finger proteins were able to reduce the mCherry expression to below 1% of the nominal value. In the orthogonality experiment the strongest cognate pair repression only reduced expression to the 8% level. The cognate pair repression would probably increase if the operator sites were moved downstream.

At several regions along the diagonal there are groups of non-cognate high repression. For example the group -57, -54 and -51 all have more than 50% repression from the non-cognate pairings. These three proteins also have a four finger segment in common. It appears these common segments of fingers can weakly bind to a portion of their cognate operator and cause attenuation. If new zinc finger proteins were designed for the purpose of making sets of orthogonal promoter-repressor pairs, then sharing segments of several zinc fingers between proteins would not be advised.

### **Metrics for orthogonal sets**

This set of orthogonality data begs the question, if a set of N orthogonal promoter-repressor pairs are needed, then how should a subset of pairs be selected from the full set? A metric is needed for evaluating the quality of a set of orthogonal pairs. Such a metric can be quite simple. For example the expression difference between the lowest non-cognate pair and the highest cognate pair could be used. This is a very conservative metric that only considers the worst case. In fact, this metric cannot differentiate between great set with one bad interaction and a set where all the interactions are bad. A less conservative metric would be to take the average expression level for non-cognate pairs and from this subtract the average expression level for cognate pairs. Such a metric encapsulates information about every interaction within the set. Variations on this metric can also be done where the difference is between one standard deviation below the average expression level for non-cognate pairs and one standard deviation above the average expression level for cognate pairs. This metric is more conservative than simply subtracting the means, but still includes information from all of the data points.

For more fine tuned control of the metric, the three methods listed above can be combined. These three methods are difference between minimum non-cognate and maximum cognate, difference between means, and difference between a standard deviation below the mean to a standard deviation above the mean. By using a weighted sum the three differences can be rolled into a single metric. The relative weighting of these three components in the summation can be modified to suit the application.

At the beginning of this chapter, the orthogonality requirements for logic gates hinted at some useful metrics of orthogonality. With logic gates, the sum of non-cognate interactions for an individual promoter is important. This can be expanded to a whole set of orthogonality data by taking the maximum of these per-promoter sums to arrive at a single orthogonality metric for the data set. Alternatively, some of these interactions may not be relevant to a specific logic network, as depicted in Figure 5-1. For these situations the orthogonality data related to a promoter can be evaluated by considering each possible set of inputs and then summing the non-cognate interaction strengths for only the repressors that can be simultaneously present in that case. Thus each promoter will have a sum for every possible input. The maximum of these sums is then the orthogonality term for that promoter. The orthogonality metric for the whole logic network would be the maximum of the per-promoter orthogonality terms. To take this optimization to its conclusion, the orthogonality metric for the whole logic network is a function of where each promoter is used in the logic network, and therefore this metric can be calculated for all permutations of promoter assignments to find the optimal configuration. This optimal configuration will be the one with the lowest orthogonality metric.

### Reduced sets

Using metrics like the ones described above, useful subset of promoter-repressor pairs can be pulled out of the orthogonality experiment data. A set of 7 orthogonal promoter-repressors pairs is shown in Table 5-4. This data comes from Table 5-2 and Table 5-3 and the labels and values have the same meaning here. The data has been re-normalized to the maximum in each row. This set has a few strong non-cognate interactions that will limit its usefulness. However, for applications where some combinations of repressors will never co-exist, this set of repressors may be usable.

**Table 5-4 Orthogonal set of 7 promoter-repressor pairs**

		zinc finger protein						
promoter	op	-57	-40	-37	-30	-18	-12	-3
83	-57	0.12	1.00	0.94	0.40	0.93	0.90	0.53
125	-40	0.81	0.14	0.55	0.95	0.97	1.00	0.93
69	-37	0.75	0.50	0.21	1.00	0.87	0.83	0.91
83	-30	0.95	1.00	0.94	0.17	0.93	0.97	0.93
69	-18	1.00	0.96	0.96	0.91	0.20	0.90	0.92
125	-12	1.00	0.99	1.00	0.99	0.87	0.26	0.90
83	-3	0.81	0.98	1.00	0.91	0.87	0.81	0.08

**Table 5-5 Highly orthogonal set of 5 promoter-repressor pairs**

promoter	op	zinc finger protein				
		-40	-30	-18	-12	-3
125	-40	0.14	0.95	0.97	1.00	0.93
83	-30	1.00	0.17	0.93	0.97	0.93
69	-18	1.00	0.95	0.21	0.94	0.96
125	-12	1.00	1.00	0.88	0.26	0.91
83	-3	1.00	0.92	0.88	0.82	0.08

For applications with more stringent requirements, a set of 5 promoter-repressor pairs has considerably better orthogonality and is shown in Table 5-5. This data comes from Table 5-2 and Table 5-3 and the labels and values have the same meaning here. The data has been re-normalized to the maximum in each row.

## Chapter 6 - Multi-gate logic, putting it all together

This project came very close to attempting to assemble a multi-gate logic network in a single cell. Key components towards this task were developed including:

- A method for creating synthetic transcriptional repressors in *E. coli*
- A set of orthogonal promoters and transcriptional repressors
- A robust topology for making logic gates with high Hill-coefficients

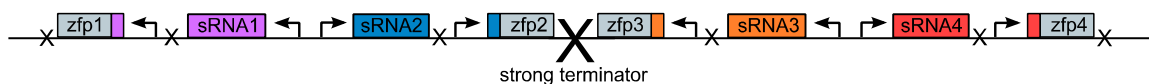
Additionally an important lesson was learned about the relative difficulty level of engineering transcriptional activators compared to engineering transcriptional repressors in prokaryotes.

Even though a multi-gate logic network was not built, a considerable amount of time was spent thinking about the challenges that will exist in implementing and testing combinational logic networks.

### Architecture

A multi-gate logic network will contain many operons and these operons must function independently of context on the DNA. Neighboring operons will need to be well terminated to prevent transcriptional read through. There are currently few strong and well proven transcriptional terminators. Basic design rules about terminators are also lacking. For example, when two operons are placed one after another on the same strand of DNA, how much space should be left between the terminator of the first operon and the promoter of the second operon?

There are some simple strategies that can be used to minimize the number of good terminators used. Divergent operons are the best way to maintain independence between operons; however with any more than two operons, terminators are required to maintain operon independence. For the threshold sensing topology using sRNA, the constitutive sRNA operon can be placed upstream of the gene it regulates and on the same strand as shown multiple times in Figure 6-1. A terminator should be used to separate the sRNA operon from the repressor it regulates. If the terminator is not 100% efficient then some additional mRNA will be produced, but this extra mRNA is proportional to the sRNA produced and therefore should be eliminated through mRNA-sRNA binding. Because the extra mRNA gets eliminated, the poor terminator does not result in a change in the repressor concentration. Instead the sRNA concentration is lowered from binding to the extra mRNA. The sRNA operons are driven by constitutive promoters, thus the lowering of sRNA concentration due to the termination can be compensated for with a stronger constitutive promoter.



**Figure 6-1 An eight operon construct with 1 strong terminator and 6 weak terminators. Operons expressing sRNA are placed upstream of the gene they regulate. The terminators between the sRNAs and zinc finger proteins can be poor and will have minimal impact on the circuit. The only terminator that needs to be strong in this construct is the middle terminator depicted with a large X.**

4 of these 8 operon pieces shown in Figure 6-1 could be constructed, with each one reusing the same 6 weak terminators and 1 strong terminator. These four segments of DNA could then be placed in 3 compatible plasmids and the 4<sup>th</sup> segment could easily be placed on the chromosome. This method would yield 32 synthetic operons, which should be enough to construct 10 2-input NOR gates.

## ***Methodology***

Individual gates will need to be built and characterized. Once proven these gates can then be assembled into multi-gate structures. Care should be taken to design the single gate structures so that they can easily be used as parts in a multi-gate assembly. The multi-gate assembly process will be performed with either CPEC [83] or Golden Gate cloning [30], [31]

Testing of logic blocks with many inputs can be challenging because there are only a handful of inducible promoters that can be used to rapidly titrate input levels. Constitutive promoters can also be used as fixed inputs, but they should be used to mimic the output of imperfect gates and not as ideal high or low signals. Several inputs can be held by constitutive promoters while a single inducible input is wiggled and the output observed.

## ***Challenges***

### **Construct stability**

The individual gates that make up a multi-gate logic system should have similar characteristics so that they are fairly interchangeable. The easiest way to make gates have similar behavior is to make them out of similar components and thus similar DNA. But if many similar sequences need to be placed in the cell to build a logic network, then the probability of homologous recombination and loss of function increases. Codon optimization can give new sequence space and relieve some of this pressure, but a better understanding of homologous recombination is needed too. Some sources claim repeats need to be separated by 10 kbp to avoid homologous recombination [29] while other sources say only 1 to 2 kbp are needed between repeats to prevent homologous recombination [84]. Solid rate information on homologous recombination as a function of repeat separation distance is needed if stable multi-gate constructs are to be stable.

The metabolic burden of a logic network could be quite large if expression levels are not carefully tuned to minimize the load placed on the host organism. The rate mutations take over a population is exponentially related to the metabolic burden removed by the mutation [28], therefore the metabolic burden must be kept in check. Often engineering low metabolic burden systems can be annoying because the signal to noise ratio of assays drops as the amount of gene expression decreases. Using high expression levels may appear to make experiments easier, but in the long run systems with lower mutation rates are easier to understand and work with. Currently there are few metrics for measuring metabolic burden besides growth rate and no good methods for predicting metabolic burden from a DNA sequence. Due to this lack of tools, engineers are mostly reacting to effects of metabolic burden instead of anticipating and avoiding them.

## **Signal fidelity**

Much has already been said on orthogonality and crosstalk between promoter-repressor pairs. As logic blocks increase in size there will be continuing pressure to reduce the orthogonality of gates.

Most electrical logic circuits exist as part of a state machine. These circuits are synchronous, meaning that a global clock signal indicates when the output of the circuit has completed transitioning and may be reliably sampled. Biological logic circuits have not yet been designed with such a clock and this exposes the system to temporary outputs that are not representative of the input state. Some techniques known as hazard-free logic have been developed for electrical circuits to deal with this issue, but they will not be cheap to implement in biological systems.

## Chapter 7 – Bibliography

- [1] D. Endy, "Foundations for engineering biology," *Nature*, vol. 438, no. 7067, pp. 449-453, Nov. 2005.
- [2] T. Knight, "Draft Standard for Biobrick Biological Parts." [Online]. Available: <http://dspace.mit.edu/handle/1721.1/45138>. [Accessed: 27-Mar-2011].
- [3] J. C. Anderson et al., "BglBricks: A flexible standard for biological part assembly," *Journal of Biological Engineering*, vol. 4, no. 1, p. 1, 2010.
- [4] B. Canton, A. Labno, and D. Endy, "Refinement and standardization of synthetic biological parts and devices," *Nat Biotech*, vol. 26, no. 7, pp. 787-793, Jul. 2008.
- [5] J. Kelly et al., "Measuring the Activity of BioBrick Promoters Using an In Vivo Reference Standard," *Journal of Biological Engineering*, vol. 3, no. 1, p. 4, Mar. 2009.
- [6] M. Hucka et al., "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models," *Bioinformatics (Oxford, England)*, vol. 19, no. 4, pp. 524-531, Mar. 2003.
- [7] W. An and J. W. Chin, "Synthesis of orthogonal transcription-translation networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 21, pp. 8477-8482, May. 2009.
- [8] W. J. Holtz and J. D. Keasling, "Engineering Static and Dynamic Control of Synthetic Pathways," *Cell*, vol. 140, no. 1, pp. 19-23, Jan. 2010.
- [9] A. Tamsir, J. J. Tabor, and C. A. Voigt, "Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'," *Nature*, vol. 469, no. 7329, pp. 212-215, Jan. 2011.
- [10] T. S. Moon, J. E. Dueber, E. Shiue, and K. L. J. Prather, "Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*," *Metabolic Engineering*, vol. 12, no. 3, pp. 298-305, May. 2010.
- [11] J. E. Dueber et al., "Synthetic protein scaffolds provide modular control over metabolic flux," *Nature Biotechnology*, Aug. 2009.
- [12] M. Isalan, A. Klug, and Y. Choo, "A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter," *Nature Biotechnology*, vol. 19, no. 7, pp. 656-660, Jul. 2001.
- [13] M. Isalan and Y. Choo, "Rapid, high-throughput engineering of sequence-specific zinc finger DNA-binding proteins," *Methods in Enzymology*, vol. 340, pp. 593-609, 2001.
- [14] C. V. Lund, P. Blancafort, M. Popkov, and C. F. Barbas 3rd, "Promoter-targeted phage display selections with preassembled synthetic zinc finger libraries for endogenous gene regulation," *Journal of Molecular Biology*, vol. 340, no. 3, pp. 599-613, Jul. 2004.
- [15] P. Blancafort, L. Magnenat, and C. F. Barbas 3rd, "Scanning the human genome with combinatorial transcription factor libraries," *Nature Biotechnology*, vol. 21, no. 3, pp. 269-274, Mar. 2003.
- [16] L. Magnenat, P. Blancafort, and C. F. Barbas 3rd, "In vivo selection of combinatorial libraries and designed affinity maturation of polydactyl zinc finger transcription factors for ICAM-1 provides new insights into gene regulation," *Journal of Molecular Biology*, vol. 341, no. 3, pp. 635-649, Aug. 2004.
- [17] D. J. Segal, B. Dreier, R. R. Beerli, and C. F. Barbas 3rd, "Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the



- 5'-GNN-3' DNA target sequences," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 6, pp. 2758-2763, Mar. 1999.
- [18] B. Dreier, D. J. Segal, and C. F. Barbas 3rd, "Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains," *Journal of Molecular Biology*, vol. 303, no. 4, pp. 489-502, Nov. 2000.
- [19] B. Dreier, R. R. Beerli, D. J. Segal, J. D. Flippin, and C. F. Barbas 3rd, "Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors," *The Journal of Biological Chemistry*, vol. 276, no. 31, pp. 29466-29478, Aug. 2001.
- [20] B. Dreier et al., "Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors," *The Journal of Biological Chemistry*, vol. 280, no. 42, pp. 35588-35597, Oct. 2005.
- [21] R. R. Beerli, D. J. Segal, B. Dreier, and C. F. Barbas 3rd, "Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14628-14633, Dec. 1998.
- [22] D. J. Segal et al., "Evaluation of a Modular Strategy for the Construction of Novel Polydactyl Zinc Finger DNA-Binding Proteins†," *Biochemistry*, vol. 42, no. 7, pp. 2137-2148, Feb. 2003.
- [23] K.-H. Bae et al., "Human zinc fingers as building blocks in the construction of artificial transcription factors," *Nat Biotech*, vol. 21, no. 3, pp. 275-280, Mar. 2003.
- [24] C. L. Ramirez et al., "Unexpected failure rates for modular assembly of engineered zinc fingers," *Nat Meth*, vol. 5, no. 5, pp. 374-375, May. 2008.
- [25] M. Moore, Y. Choo, and A. Klug, "Design of polyzinc finger peptides with structured linkers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 4, pp. 1432-1436, Feb. 2001.
- [26] M. Moore, Aaron Klug, and Y. Choo, "Improved DNA Binding Specificity from Polyzinc Finger Peptides by Using Strings of Two-Finger Units," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 4, pp. 1437-1441, Feb. 2001.
- [27] "'Round-the-horn site-directed mutagenesis - OpenWetWare." [Online]. Available: [http://openwetware.org/index.php?title=%27Round-the-horn\\_site-directed\\_mutagenesis&oldid=458415](http://openwetware.org/index.php?title=%27Round-the-horn_site-directed_mutagenesis&oldid=458415). [Accessed: 31-Mar-2011].
- [28] S. C. Sleight, B. A. Bartley, J. A. Lieviant, and H. M. Sauro, "Designing and engineering evolutionary robust genetic circuits," *Journal of Biological Engineering*, vol. 4, p. 12, 2010.
- [29] S. T. Lovett, "Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences," *Molecular Microbiology*, vol. 52, no. 5, pp. 1243-1253, Jun. 2004.
- [30] C. Engler, R. Kandzia, and S. Marillonnet, "A One Pot, One Step, Precision Cloning Method with High Throughput Capability," *PLoS ONE*, vol. 3, no. 11, p. e3647, Nov. 2008.

- [31] C. Engler, R. Gruetzner, R. Kandzia, and S. Marillonnet, "Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes," *PLoS ONE*, vol. 4, no. 5, p. e5553, May. 2009.
- [32] "Main Page - j5." [Online]. Available: [http://jbei-exwebapp.lbl.gov/j5/index.php/Main\\_Page](http://jbei-exwebapp.lbl.gov/j5/index.php/Main_Page). [Accessed: 31-Mar-2011].
- [33] K. J. Breslauer, R. Frank, H. Blöcker, and L. A. Marky, "Predicting DNA duplex stability from the base sequence.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 11, pp. 3746-3750, Jun. 1986.
- [34] D. C. Montgomery, *Introduction to Statistical Quality Control*, 6th ed. Wiley, 2008.
- [35] M. Lis, M. N. Artyomov, S. Devadas, and A. K. Chakraborty, "Efficient stochastic simulation of reaction-diffusion processes via direct compilation," vol. 25, no. 17, pp. 2289-2291, Sep. 2009.
- [36] S. Fields and O. Song, "A novel genetic system to detect protein-protein interactions," *Nature*, vol. 340, no. 6230, pp. 245-246, Jul. 1989.
- [37] J. C. Hu, M. G. Kornacker, and A. Hochschild, "Escherichia coli one- and two-hybrid systems for the analysis and identification of protein-protein interactions," *Methods (San Diego, Calif.)*, vol. 20, no. 1, pp. 80-94, Jan. 2000.
- [38] S. L. Dove and A. Hochschild, "Conversion of the omega subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target," *Genes & Development*, vol. 12, no. 5, pp. 745-754, Mar. 1998.
- [39] S. Durai, A. Bosley, A. B. Abulencia, S. Chandrasegaran, and M. Ostermeier, "A bacterial one-hybrid selection system for interrogating zinc finger-DNA interactions," *Combinatorial Chemistry & High Throughput Screening*, vol. 9, no. 4, pp. 301-311, May. 2006.
- [40] X. Meng, M. H. Brodsky, and S. A. Wolfe, "A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors," *Nature Biotechnology*, vol. 23, no. 8, pp. 988-994, Aug. 2005.
- [41] S. A. Wolfe, E. I. Ramm, and C. O. Pabo, "Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers," *Structure (London, England: 1993)*, vol. 8, no. 7, pp. 739-750, Jul. 2000.
- [42] J. L. Pomerantz, S. A. Wolfe, and C. O. Pabo, "Structure-Based Design of a Dimeric Zinc Finger Protein†," *Biochemistry*, vol. 37, no. 4, pp. 965-970, Jan. 1998.
- [43] R. Shetty, "Applying engineering principles to the design and construction of transcriptional devices," Massachusetts Institute of Technology, 2008.
- [44] R. Lutz and H. Bujard, "Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements," *Nucleic Acids Research*, vol. 25, no. 6, pp. 1203-1210, Mar. 1997.
- [45] M. Lanzer and H. Bujard, "Promoters largely determine the efficiency of repressor action," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, no. 23, pp. 8973-8977, Dec. 1988.
- [46] C. D. Deppmann, R. S. Alvania, and E. J. Taparowsky, "Cross-Species Annotation of Basic Leucine Zipper Factor Interactions: Insight into the Evolution of Closed Interaction Networks," *Molecular Biology and Evolution*, vol. 23, no. 8, pp. 1480 -1492, Aug. 2006.

- [47] J. R. S. Newman and A. E. Keating, "Comprehensive Identification of Human bZIP Interactions with Coiled-Coil Arrays," *Science*, vol. 300, no. 5628, pp. 2097-2101, Jun. 2003.
- [48] M. Besse, B. von Wilcken-Bergmann, and B. Müller-Hill, "Synthetic lac operator mediates repression through lac repressor when introduced upstream and downstream from lac promoter.," *The EMBO Journal*, vol. 5, no. 6, pp. 1377-1381, Jun. 1986.
- [49] B. Krummel and M. J. Chamberlin, "Structural analysis of ternary complexes of Escherichia coli RNA polymerase. Deoxyribonuclease I footprinting of defined complexes," *Journal of Molecular Biology*, vol. 225, no. 2, pp. 239-250, May. 1992.
- [50] M. De Mey, J. Maertens, G. Lequeux, W. Soetaert, and E. Vandamme, "Construction and model-based analysis of a promoter library for E. coli: an indispensable tool for metabolic engineering," *BMC Biotechnology*, vol. 7, no. 1, p. 34, 2007.
- [51] R. Hershberg, G. Bejerano, A. Santos-Zavaleta, and H. Margalit, "PromEC: An updated database of Escherichia coli mRNA promoters with experimentally identified transcriptional start sites," *Nucleic Acids Research*, vol. 29, no. 1, p. 277, Jan. 2001.
- [52] R. S. Cox 3rd, M. J. Dunlop, and M. B. Elowitz, "A synthetic three-color scaffold for monitoring genetic regulation and noise," *Journal of Biological Engineering*, vol. 4, p. 10, 2010.
- [53] A. Rich, A. Nordheim, and A. H. J. Wang, "The Chemistry and Biology of Left-Handed Z-DNA," *Annual Review of Biochemistry*, vol. 53, no. 1, pp. 791-846, Jun. 1984.
- [54] K. S. Murakami, S. Masuda, E. A. Campbell, O. Muzzin, and S. A. Darst, "Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex," *Science (New York, N.Y.)*, vol. 296, no. 5571, pp. 1285-1290, May. 2002.
- [55] D. J. Pitera, C. J. Paddon, J. D. Newman, and J. D. Keasling, "Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli," *Metabolic Engineering*, vol. 9, no. 2, pp. 193-207, Mar. 2007.
- [56] R. M. Saecker, M. T. Record Jr., and P. L. deHaseth, "Mechanism of Bacterial Transcription Initiation: Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis," *Journal of Molecular Biology*, vol. In Press, Uncorrected Proof.
- [57] D. Jantz and J. M. Berg, "Probing the DNA-Binding Affinity and Specificity of Designed Zinc Finger Proteins," *Biophysical Journal*, vol. 98, no. 5, pp. 852-860, Mar. 2010.
- [58] J. D. Sander, P. Zaback, J. K. Joung, D. F. Voytas, and D. Dobbs, "An affinity-based scoring scheme for predicting DNA-binding activities of modularly assembled zinc-finger proteins," *Nucleic Acids Research*, vol. 37, no. 2, pp. 506-515, Feb. 2009.
- [59] J. M. Carothers, S. C. Oestreich, and J. W. Szostak, "Aptamers Selected for Higher-Affinity Binding Are Not More Specific for the Target Ligand," *Journal of the American Chemical Society*, vol. 128, no. 24, pp. 7929-7937, Jun. 2006.
- [60] J. S. Kang, "Correlation between functional and binding activities of designer zinc-finger proteins," *Biochemical Journal*, vol. 403, no. 1, pp. 177-182, Apr. 2007.
- [61] J. F. Timoney, J. Walker, M. Zhou, and J. Ding, "Cloning and sequence analysis of a protective M-like protein gene from Streptococcus equi subsp. zooepidemicus.," *Infection and Immunity*, vol. 63, no. 4, pp. 1440-1445, Apr. 1995.

- [62] S. Brantl and E. G. H. Wagner, "Antisense RNA-mediated transcriptional attenuation: an *in vitro* study of plasmid pT181," *Molecular Microbiology*, vol. 35, no. 6, pp. 1469-1482, 2000.
- [63] K. Clancy and C. A. Voigt, "Programming cells: towards an automated [ ]Genetic Compiler'," *Current Opinion in Biotechnology*, vol. 21, no. 4, pp. 572-581, Aug. 2010.
- [64] R. Weiss, G. E. Homsy, T. F. Knight, and Jr, "Toward *in vivo* Digital Circuits," 1999.
- [65] R. Weiss, "Cellular computation and communications using engineered genetic regulatory networks," Massachusetts Institute of Technology, 2001.
- [66] D. J. Sayut, Y. Niu, and L. Sun, "Construction and enhancement of a minimal genetic and logic gate," *Applied and Environmental Microbiology*, vol. 75, no. 3, pp. 637-642, Feb. 2009.
- [67] Z. Li, M. A. Rosenbaum, A. Venkataraman, T. K. Tam, E. Katz, and L. T. Angenent, "Bacteria-based AND logic gate: a decision-making and self-powered biosensor," *Chemical Communications*, vol. 47, no. 11, p. 3060, 2011.
- [68] J. C. Anderson, C. A. Voigt, and A. P. Arkin, "Environmental signal integration by a modular AND gate," *Molecular Systems Biology*, vol. 3, p. 133, 2007.
- [69] J. Zhan et al., "Develop reusable and combinable designs for transcriptional logic gates," *Mol Syst Biol*, vol. 6, Jul. 2010.
- [70] M. N. Win and C. D. Smolke, "Higher-order cellular information processing with synthetic RNA devices," *Science (New York, N.Y.)*, vol. 322, no. 5900, pp. 456-460, Oct. 2008.
- [71] N. Sudarsan et al., "Tandem riboswitch architectures exhibit complex gene control functions," *Science (New York, N.Y.)*, vol. 314, no. 5797, pp. 300-304, Oct. 2006.
- [72] M. A. Marchisio and J. Stelling, "Automatic Design of Digital Synthetic Gene Circuits," *PLoS Comput Biol*, vol. 7, no. 2, p. e1001083, Feb. 2011.
- [73] S. A. Wolfe, R. A. Grant, and C. O. Pabo, "Structure of a designed dimeric zinc finger protein bound to DNA," *Biochemistry*, vol. 42, no. 46, pp. 13401-13409, Nov. 2003.
- [74] H. M. Sauro et al., "Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration," *OmicS: A Journal of Integrative Biology*, vol. 7, no. 4, pp. 355-372, 2003.
- [75] D. Chandran, F. T. Bergmann, and H. M. Sauro, "Computer-aided design of biological circuits using TinkerCell," *Bioengineered Bugs*, vol. 1, no. 4, pp. 274-281, 2010.
- [76] S. Tan et al., "Zinc-finger protein-targeted gene regulation: Genomewide single-gene specificity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 21, pp. 11997-12002, Oct. 2003.
- [77] C. R. Robinson and R. T. Sauer, "Optimizing the stability of single-chain proteins by linker length and composition mutagenesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 11, pp. 5929-5934, May. 1998.
- [78] M. Bouvier, C. M. Sharma, F. Mika, K. H. Nierhaus, and J. Vogel, "Small RNA Binding to 5' mRNA Coding Region Inhibits Translational Initiation," *Molecular Cell*, vol. 32, no. 6, pp. 827-837, Dec. 2008.
- [79] S. Gottesman, "Micros for microbes: non-coding regulatory RNAs in bacteria," *Trends in Genetics*, vol. 21, no. 7, pp. 399-404, Jul. 2005.
- [80] J. Vogel et al., "RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria," *Nucleic Acids Research*, vol. 31, no. 22, pp. 6435 -6443, Nov. 2003.

- [81] J. F. Hopkins, S. Panja, and S. A. Woodson, "Rapid binding and release of Hfq from ternary complexes during RNA annealing," *Nucleic Acids Research*, Mar. 2011.
- [82] C. T. Chung and R. H. Miller, "Preparation and storage of competent *Escherichia coli* cells," *Methods in Enzymology*, vol. 218, pp. 621-627, 1993.
- [83] J. Quan and J. Tian, "Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways," *PLoS ONE*, vol. 4, no. 7, p. e6441, Jul. 2009.
- [84] M. Bzymek and S. T. Lovett, "Instability of repetitive DNA sequences: The role of replication in multiple mechanisms," *Proceedings of the National Academy of Sciences*, vol. 98, no. 15, pp. 8319 -8325, Jul. 2001.

## Chapter 8 – Appendix: sequence information

The following sequence information is in GenBank format. While GenBank is currently the most widely supported annotated sequence format, the file format has not been well standardized. The files listed here were created with the tool GENTle. This tool is freely available at <http://gentle.magnusmanske.de/>.

In Chapter 3, several constructs are tested on a p15A/Kan backbone and a ColE1/Kan backbone. Only the p15A/Kan versions are listed below. The ColE1 version can be obtained by transferring the sequence between the BglII and BamHI sites from the p15A plasmid to between the BglII and BamHI sites on the “Empty vector ColE1/Kan” plasmid listed below

### **Empty vector ColE1/Kan**

LOCUS 17-36 2051 bp DNA circular

FEATURES Location/Qualifiers  
gene complement(1106..1900)  
/gene="KanR"  
rep\_origin complement(287..969)  
/gene="ColE1"  
terminator 25..153  
/gene="double term"  
terminator 975..1080  
/gene="Terminator"

BASE COUNT 481 a 590 c 545 g 435 t

#### ORIGIN

```
1 ggatcctaac tcgagtaagg atctccaggc atcaaataaa acgaaaggct cagtcgaaag
61 actgggcctt tcgttttatc tgttgttgtg cggatgaacgc tctctactag agtcacactg
121 gctcaccttc gggtagggcct ttctgcgttt atacctaggg cgttcggctg cggcgagcgg
181 tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat aacgcaggaa
241 agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg
301 cgttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga
361 ggtggcga aa cccgacagga ctataaagat accaggcgtt tcccctgga agtcacctc
421 tgcgctctcc tgttccgacc ctgccgctta ccgatacct gtccgccttt ctccttcgg
481 gaagcgtggc gtttctcat agctcacgct gtaggtatct cagttcgggtg taggtcgctc
541 gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg
601 gtaactatcg tcttgagtcc aaccggtaa gacacgactt atgccactg gcagcagcca
661 ctggtaacag gattagcaga gcgaggtatg taggcgggtg tacagagttc ttgaagtgg
721 ggctaacta cggctacact agaaggacag tatttggtat ctgctgctctg ctgaagccag
781 ttaccttcgg aaaaagagtt gtagctctt gatccggcaa acaaacacc gctgtagcgg
841 gtgggttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc
901 cttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt
961 tggatcatgac tagtgcttgg atttcacca ataaaaaacg cccggcggca accgagcgtt
1021 ctgaacaaat ccagatggag ttctgaggtc attactggat ctatcaacag gagtccaagc
1081 gagctctcga acccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga
1141 tgcgctgcga atcgggagcg gcgatacctg aaagcagcag gaagcgggtca gccattcgc
```

1201 cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtctgatag cgggccgcca  
 1261 caccagccg gccacagtcg atgaatccag aaaagcggcc atttccacc atgatattcg  
 1321 gcaagcaggc atcgccatgg gtcacgacga gatcctcgcc gtcgggcatg cgcgccttga  
 1381 gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcatcctgat  
 1441 cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgttc gcttgggtgt  
 1501 cgaatgggca ggtagccgga tcaagcgtat gcagccgccc cattgcatca gccatgatgg  
 1561 atactttctc ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgccc  
 1621 atagcagcca gtccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc  
 1681 ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg  
 1741 acaggtcggc cttgacaaaa agaaccgggc gccctgcgc tgacagccgg aacacggcgg  
 1801 catcagagca gccgattgtc tgttggtccc agtcatagcc gaatagcctc tccaccaag  
 1861 cggccggaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat cctcatcctg  
 1921 tctcttgatc agatcatgat ccctgcgcc atcagatcct tggcggcaag aaagccatcc  
 1981 agtttacttt gcagggcttc ccaaccttac cagagggcgc cccagctggc aattccgaat  
 2041 tcatgagatc t

//

**ZFP03 leak on SC101/Kan**

LOCUS 12-10 5022 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(2596..4748)  
     /gene="Ori SC101"  
 CDS 1628..2422  
     /gene="KanR"  
     /codon\_start="0"  
 terminator complement(4813..4986)  
     /gene="Terminator"  
     and Thermal Tm unaffected"  
 terminator complement(2443..2558)  
     /gene="Terminator"  
 terminator 4767..4812  
     /gene="Terminator"  
 CDS 166..885  
     /gene="GFPmut3b"  
     /codon\_start="0"  
 terminator 1..46  
     /gene="term BBa\_B0011"  
 prot\_bind 145..159  
     /gene="RBS"  
 prot\_bind 53..70  
     /gene="ZFP03 binding site"  
 promoter 77..138  
     /gene="pwklac"  
 BASE COUNT 1516 a 1075 c 1168 g 1263 t

## ORIGIN

1 agagaatata aaaagccaga ttattaatcc ggcttttta ttatttgat ctggtagacg  
61 tctagtaact ggatctccgt attctttaca ctttatgctt cgggctcgt tgttggtc  
121 accgagcggga taacaattgg atctattaaa gaggagaaag gatctatgcg taaaggagaa  
181 gaactttca ctggagtgt cccaattctt gttgaattag atggtgatgt taatgggcac  
241 aaattttctg tcagtggaga ggtgaaggt gatgcaacat acggaaaact tacccttaa  
301 tttatttca ctactggaaa actacgtgt ccatggccaa cactgtcac tacttcggt  
361 tatggtgtc aatgctttgc gagataccca gatcatatga aacagcatga cttttcaag  
421 agtgccatgc ccgaaggta tgtacaggaa agaactatat tttcaaaga tgacgggaac  
481 tacaagacac gtgctgaagt caagttgaa ggtgataccc ttgtaatag aatcgagta  
541 aaaggtattg atttaaaga agatggaaac attctggac acaattgga atacaactat  
601 aactcacaca atgtatacat catggcagac aaacaaaaga atggaatcaa agttaactc  
661 aaaattagac acaacattga agatggaagc gttcaactag cagaccatta tcaacaaat  
721 actccaattg gcgatggccc tgcctttta ccagacaacc attacctgc cacacaatc  
781 gccctttcga aagatccca cgaaaagaga gaccacatgg tccttctga gtttgaacc  
841 gctgctggga ttacacatgg catggatgaa ctatacaaat aataaggatc tagagaat  
901 aaaaagccag attattaatc cggctttttt attatttga tctaactcg agtgcaggc  
961 gtctcgatc gagaaggaca cggttaatac taggctgct ggctggaat cgccagcagg  
1021 ctttttatt tgggggagag ggaagtcatg aaaaaactaa ctttgaat tcgatcca  
1081 ccacatcagc tctgaagcaa cgtaaaaaa cccgccccg cgggttttt tataccgta  
1141 gtatccccac ttatctaca tagctgtcct taattaagg tgaataaata aaaacagccg  
1201 ttgccagaaa gaggcacggc tgttttatt tttagtgag accgggagca gtaaacgca  
1261 gaaaggccca cccgaagggt agccagtgtg actctagtag agagcgttca ccgacaaca  
1321 acagataaaa cgaaaggccc agtctttcga ctgagccttt cgttttatt gatgctggt  
1381 tattattatt tgtacagctc atccatgcca ccggtcaaca ggagtccaag cgaccggtg  
1441 ttgatgtct agctagctag aacaggacta gctaattggt tcttagacgt cggattgcc  
1501 agctggggcg ccctctgta aggttgggaa gccctgcaa gtaaactgga tggcttctt  
1561 gccccaagg atctgatggc gcaggggatc aaaatctgat caagagacag gatgaggatc  
1621 gtttcgatg attgaacaag atggattgca cgcaggttct cggccgctt gggaggagag  
1681 gctattcggc tatgactggg cacaacagac aatcggctgc tctgatccg ccgtgtccg  
1741 gctgtcagcg caggggcgcc cggttcttt tgtcaagacc gacctgtcc gtgcctgaa  
1801 tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacggcg tccttgcc  
1861 agctgtgctc gacgttgc ctgaagcggg aagggactgg ctgctattgg gcgaagtgc  
1921 ggggcaggat ctctgtcat ctacctgc tctgcccag aaagtatca tcatggctga  
1981 tgcaatgccc cggtgcata cgctgatcc ggctacctgc ccattcgacc accaagcgaa  
2041 acatcgcac gagcagcac gtactcggat ggaagccggt cttgtcgc atggatgatc  
2101 ggacgaagag catcaggggc tcgcccagc cgaactgtc gccaggctca aggcgcgat  
2161 gcccagcgc gaggatctc tcgtaccca tggcgatcc tcttgccga atatcatggt  
2221 ggaaaatggc cgctttctg gattcatcga ctgtggccg ctgggtgtg cggaccgta  
2281 tcaggacata gcgttgcta cccgtgat tgctgaagag cttggcggc aatgggctga  
2341 ccgcttctc gtctttac gtatcggc tcccgatcc cagcgcac cttctatc  
2401 cttcttgac gattctct gagcgggact ctgggttc agagctcgt tggactcctg  
2461 ttgatagatc cagtaatgac ctcaactc catctggatt tttcagaac gctcgggtg  
2521 cgccggcgt ttttattg tgagaatca agcactagg acagtaagac gggtaagcct  
2581 gttgatgata ccgctgcct actgggtgca ttagccagtc tgaatgacct gtcacggat  
2641 aatccgaagt ggtcagact gaaaatcaga gggcaggaac tgctgaacag caaaaagtc



2701 gatagcacca catagcagac ccgcataaa acgccctgag aagcccgtga cgggcttttc  
2761 ttgtattatg ggtagtttcc ttgcatgaat ccataaaagg cgcctgtagt gccatttacc  
2821 cccattcact gccagagccg tgagcgcagc gaactgaatg tcacgaaaaa gacagcgact  
2881 caggtgcctg atggtcggag acaaaaggaa tattcagcga ttgcccagag cttgcgaggg  
2941 tgctacttaa gccttaggg tttaaggtc tgttttagtag aggagcaaac agcgtttgcg  
3001 acatcctttt gtaatactgc ggaactgact aaagtagtga gttatacaca gggctgggat  
3061 ctattctttt tatctttttt tattctttct ttattctata aattataacc acttgaatat  
3121 aaacaaaaaa aacacacaaa ggtctagcgg aatttacaga gggcttagca gaatttaca  
3181 gttttccagc aaaggtctag cagaatttac agatacccac aactcaaagg aaaaggacta  
3241 gtaattatca ttgactagcc catctcaatt ggtatagtga ttaaaatcac cttagaccaat  
3301 tgagatgtat gtctgaatta gttgttttca aagcaaata actagcgatt agtcgctatg  
3361 acttaacgga gcatgaaacc aagctaattt tatgctgtgt ggcactactc aaccccacga  
3421 ttgaaaacc tacaaggaaa gaacggacgg tatcgttcac ttataacaa tacgctcaga  
3481 tgatgaacat cagtagggaa aatgcttatg gtgtattagc taaagcaacc agagagctga  
3541 tgacgagaac tgtggaaatc aggaatcctt tggtaaagg ctttgagatt tccagtgga  
3601 caaactatgc caagttctca agcgaaaaat tagaattagt ttttagtgaa gagatattgc  
3661 cttatctttt ccagttaaaa aaattcataa aatataatct ggaacatggt aagcttttg  
3721 aaaacaaata ctctatgagg atttatgagt ggtattaaa agaactaaca caaaagaaaa  
3781 ctacaaggc aaatatagag attagccttg atgaattta gttcatgtta atgcttgaaa  
3841 ataactacca tgagttaaaa aggcttaacc aatgggtttt gaaaccaata agtaaagatt  
3901 taaactta cagcaatatg aaattggtgg ttgataagcg aggccgccc actgatacgt  
3961 tgattttcca agtgaacta gatagacaaa tggatctcgt aaccgaactt gagaacaacc  
4021 agataaaaat gaatggtgac aaaatacaca caaccattac atcagattcc tacctacgta  
4081 acggactaag aaaaacta cacgatgctt taactgcaa aattcagctc accagtttg  
4141 aggcaaaatt tttgagtac atgcaaagta agcatgatct caatggttcg ttctcatggc  
4201 tcacgaaaa acaacgaacc aactagaga acatactggc taaatacga aggatctgag  
4261 gttcttatgg ctcttgatc tatcagtgaa gcatcaagac taacaaaca aagtagaaca  
4321 actgttcacc gtagatatac aaagggaaaa ctgtccatat gcacagatga aaacggtgta  
4381 aaaaagatag atacatcaga gcttttacga gtttttggtg catttaaagc tgttccat  
4441 gaacagatcg acaatgtaac agatgaacag catgtaacac ctaatagaac aggtgaaacc  
4501 agtaaaaca agcaactaga acatgaaatt gaacacctga gacaactgt tacagctcaa  
4561 cagtacaca tagacagcct gaaacaggcg atgctgctta tcgaatcaa gctgccgaca  
4621 acacgggagc cagtgcgcc tcccgtgggg aaaaaatcat ggcaattctg gaagaaatag  
4681 cgctttcagc cggcaaact gaagccggat ctgctgattct gatacaaac tagcaacacc  
4741 agaacagcgc ggccgaaaa ggaaaagatc cggcaaaaa accaccgtt gtagcgggtg  
4801 ttttttgtg ttgatcgaca atcttcgtaa gcgtcatcaa taagcgtaaa aaaaccgggc  
4861 aatgcccggt ttttaata gaaatttac ctgtcgtagc cgccaccatc cggcaaagaa  
4921 gcatacaagg cttttggctt atagctacgt agcgcattgc gtcgcagcac aatcccggca  
4981 ccgatcaagt cttcgcgatg attattagaa ttcattgatg ct

//

**ZFP5475 matched act/rep on SC101/Kan**

LOCUS 12-16 6497 bp DNA circular

FEATURES Location/Qualifiers

```

rep_origin complement(4327..6479)
      /gene="Ori SC101"
CDS      3359..4153
      /gene="KanR"
      /codon_start="0"
terminator complement(4174..4289)
      /gene="Terminator"
CDS      426..1145
      /gene="GFPmut3b"
      /codon_start="0"
terminator 257..302
      /gene="term BBa_B0011"
prot_bind 405..419
      /gene="RBS"
promoter 1204..1257
      /gene="pLtetO-1"
CDS      1285..2028
      /gene="RNAP alpha subunit"
      /codon_start="0"
CDS      2035..2052
      /gene="linker"
      /codon_start="0"
prot_bind 309..330
      /gene="ZFP5475 binding site"
promoter 337..398
      /gene="pwklac"
CDS      2059..2616
      /gene="ZFP5475"
      /codon_start="0"
BASE COUNT 1893 a 1469 c 1560 g 1575 t

```

ORIGIN

```

1 gatccggcaa acaaaccacc gttgtagcg gtggttttt tgttggatc gacaatctt
61 gtaagcgta tcaataagcg taaaaaac gggcaatgcc cggttttta atgagaaat
121 ttacctgtcg tagccgccac catccggcaa agaagcatac aaggctttg gcttatagc
181 acgtagcgca ttgcgtcgca gcacaatccc ggcaccgatc aagtcttcgc gatgattat
241 agaattcaaa agatctagag aatataaaaa gccagattat taatccggct ttttattat
301 ttgatctac acccgggttc cctcgggtg ggatctccgt attcttaca cttatgctt
361 cggctcgtg tgttgtgtcg accgagcggg taacaattgg atctattaa gaggagaaag
421 gatctatgcg taaaggagaa gaactttca ctggagtgt cccaattctt gttgaattg
481 atggtgatgt taatgggcac aaattttctg tcagtggaga gggatgaagg gatgcaacat
541 acgaaaaact taccttaaa tttattgca ctactggaaa actacctgtt ccatggccaa
601 cactgtcac tacttccggt tatggtgttc aatgctttgc gagatacca gatcatatga
661 aacagcatga cttttcaag agtgccatgc ccgaaggta tgtacaggaa agaactatat
721 tttcaaaga tgacgggaac tacaagacac gtgctgaagt caagttttaa ggtgataccc
781 ttgtaatag aatcgagta aaaggtattg attttaaaga agatggaaac attctggac
841 acaaattgga atacaactat aactcacaca atgtatacat catggcagac aaacaaaaga

```

901 atggaatcaa agttaacttc aaaattagac acaacattga agatggaagc gttcaactag  
961 cagaccatta tcaacaaaat actccaattg gcatggccc tgcctttta ccagacaacc  
1021 attacctgtc cacacaatct gccctttcga aagatcccaa cgaaaagaga gaccacatgg  
1081 tccttcttga gtttgaacc gctgctggga ttacacatgg catggatgaa ctatacaaat  
1141 aataaggatc tagagaatat aaaaagccag attattaatc cggctttttt attatttgga  
1201 tcttcctat cagtataga gattgacatc cctatcagtg atagagatac tgagcacgga  
1261 tctattaaag aggagaaagg atctatgcag ggttctgtga cagagtttct aaaaccgcg  
1321 ctggttgata tcgagcaagt gagttcgacg cacgccaagg tgacccttga gcctttagag  
1381 cgtggctttg gccatactct gggtaacgca ctgcccgtta ttctgctctc atcgatgccg  
1441 ggttgcgagg tgaccgaggt tgagattgat ggtgtactac atgagtacag caccaaagaa  
1501 ggcgttcagg aagatatcct ggaatcctg ctcaacctga aagggtggc ggtgagagt  
1561 caggcгааag atgaagtat tcttacctg aataaatctg gcattggccc tgtgactgca  
1621 gccatatca cccacgacgg tgatgtcga atcgtcaagc cgcagcacgt gatctgccac  
1681 ctgaccgatg agaacgcgtc tattagcatg cgtatcaaag ttcagcggcg tcgtggttat  
1741 gtgccggctt ctaccgaaat tcttcggaa gaagatgagc gcccaatcgg ccgtctgctg  
1801 gtcgacgat gctacagccc tgtggagcgt attgcctaca atgttgaagc agcgcgtgta  
1861 gaacagcgt cgcacctgga caagctggtc atcgaaatgg aaaccaacgg cacaatgat  
1921 cctgaagagg cgattcgtc tgcggcaacc attctggctg aacaactgga agctttcgtt  
1981 gacttacgtg atgtactca gcctgaagt aaagaagaga aaccagaggg atctgccccg  
2041 cgagtcgga cggatctgg ggtaccgcc atggcggaga ggccttcca gtgtcgaatc  
2101 tgatgcgta acttcagtgc tagtgaccac ctgagccggc acatccgac cacacaggg  
2161 gagaagcctt ttgcctgtga catttgggg aggaaattg ccgacaaccg ggaaccgaca  
2221 aagcacaac agatacacac gggcggacag cggccgtacg catgcctgt cgagtctctg  
2281 gatcggcgt tttctgacag gaagacact atcgagcata tccgatcca caccggtcag  
2341 aagccttcc agtgtcgaat ctgcatgcgt aacttcagta ccagcagcgg cctgagccgc  
2401 cacatccgca cccacacagg atctcagaag ccctccagt gtcgaatctg catgctaac  
2461 ttcagtctga gtgaccacct gagcgaacac atccgaccc acacaggcga gaagccttt  
2521 gcctgtgaca tttgtgggag gaaattgcc accagcagcg accgcacaaa gcatacag  
2581 atacactgc gccaaaaaga tgcggcccgg taataaggat ctagagaata taaaaagca  
2641 gattattaat cggcttttt tattatttgg atcctaactc gactgcaggt cgtctcggat  
2701 cgagaaggac acggttaata ctaggcctgc tggctggtaa tcgccagcag gccttttat  
2761 ttgggggaga gggaagcat gaaaaacta accttgaaa ttcgatctcc accacatcag  
2821 ctctgaagca acgtaaaaaa acccgccccg gcgggttttt ttataccctg agtatccca  
2881 ctatctaca atagctgtcc ttaattaagg ttgaataaat aaaacagcc gttgccagaa  
2941 agaggcacgg ctgtttttat ttctagtga gaccgggagc agttaaagc agaaaggccc  
3001 accgaaagt gagccagtgt gactctagta gagagcgttc accgacaaac aacagataaa  
3061 acgaaaggcc cagtcttcg actgagcctt tcgtttatt tgatgctgg ttattattat  
3121 ttgtacagct catccatgcc accggtcaac aggagtcaa gcgaccggtg gttgatgtc  
3181 tagctagcta gaacaggact agtaatggt ttcttagacg tcggaattgc cagctggggc  
3241 gccctctggt aaggtggga agcctgcaa agtaaactgg atggcttct tgcgccaag  
3301 gatctgatgg cgcaggggat caagatctga tcaagagaca ggatgaggat cgtttcgtat  
3361 gattgaacaa gatggattgc acgaggttc tccggcgtc tgggtggaga ggctattcgg  
3421 ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgtcc ggctgtcagc  
3481 gcagggggc cgggtcttt ttgtcaagac cgacctgtcc ggtgcctga atgaactgca  
3541 ggacgaggca gcgcggtat cgtggctggc cacgacgggc gttccttgcg cagctgtgct  
3601 cgacgttgc actgaagcgg gaaggactg gctgctattg ggcgaagtgc cggggcagga

3661 tctcctgtca tctcaccttg ctctgccga gaaagtatcc atcatggctg atgcaatgcg  
3721 gcggtgcat acgcttgatc cggctacctg cccattcgac caccaagcga aacatcgcat  
3781 cgagcgagca cgtactcgga tggagccgg tcttgtcga caggatgatc tggacgaaga  
3841 gcatcagggg ctcgccag ccgaactgtt gcgcaggctc aaggcgcgca tggccgacgg  
3901 cgaggatctc gtcgtgacc atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg  
3961 ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct atcaggacat  
4021 agcgttggt acccgtgata ttgctgaaga gcttggcggc gaatgggctg accgcttct  
4081 cgtgctttac ggtatcgccg ctcccattc gcagcgcac gccttctatc gccttctga  
4141 cgagttcttc tgagcgggac tctggggttc gagagctcgc ttggactcct gttgatagat  
4201 ccagtaatga cctcagaact ccatctggat ttgttcagaa cgctcgggtg ccgcccggcg  
4261 tttttattg gtgagaatcc aagcactagg gacagtaaga cgggtaagcc tgttgatgat  
4321 accgctgcct tactgggtgc attagccagt ctgaatgacc tgtcacggga taatccgaag  
4381 tggtcagact ggaaaatcag agggcaggaa ctgctgaaca gcaaaaagtc agatagcacc  
4441 acatagcaga cccgcataa aacgccctga gaagcccgtg acgggctttt cttgtattat  
4501 gggtagttc cttgatgaa tccataaaag gcgcctgtg tgccatttac cccattcac  
4561 tgccagagcc gtgagcgcag cgaactgaat gtcacgaaaa agacagcgac tcaggtgcct  
4621 gatggtcgga gaaaaagga atattcagcg atttgccga gcttgcgagg gtgctactta  
4681 agccttagg gtttaaggt ctgtttgta gaggagcaaa cagcgttgc gacatcctt  
4741 tgaataactg cggaactgac taaagtagt agttatacac agggctggga tctattctt  
4801 ttatctttt ttattcttc tttattctat aaattatac cactgaata taaacaaaa  
4861 aaacacacia aggtctagcg gaatttacag aggtctagc agaatttaca agtttccag  
4921 caaaggtcta gcagaatta cagataccca caactcaaag gaaaaggact agtaattatc  
4981 attgactagc ccatctcaat tggatagtg attaaaatca ctagaccaa ttgagatgta  
5041 tgtctgaatt agttgtttc aaagcaatg aactagcgat tagtcgctat gacttaacgg  
5101 agcatgaaac caagctaatt ttatgctgtg tggcactact caaccacg attgaaaacc  
5161 ctacaaggaa agaacggacg gtatcgttca cttataacca atacgctcag atgatgaaca  
5221 tcagtaggga aatgcttat ggtgtattag ctaaagcaac cagagagctg atgacgagaa  
5281 ctgtggaaat caggaatcct ttggttaaag gctttgagat tttccagtgg acaactatg  
5341 ccaagtctc aagcgaaaaa ttagaattag ttttagtga agagatattg cttatctt  
5401 tccagttaa aaaattcata aaatataatc tggaaatgt taagtcttt gaaaacaaat  
5461 actctatgag gatttatgag tggttattaa aagaactaac aaaaagaaa actcacaagg  
5521 caaatataga gattagcctt gatgaattta agttcatgtt aatgcttgaa aataactacc  
5581 atgagttaa aaggctaac caatgggtt tgaaccaat aagtaaagat taaactct  
5641 acagcaatat gaaattggtg gttgataagc gaggccgcc gactgatacg ttgatttcc  
5701 aagttgaact agatagacia atggatctcg taaccgaact tgagaacaac cagataaaaa  
5761 tgaatggtga caaaatacca acaaccatta catcagattc ctacctagt aacggactaa  
5821 gaaaaaact acacgatgct ttaactgcaa aaattcagct caccagttt gaggcaaat  
5881 ttttagtga catgcaaagt aagcatgatc tcaatggttc gttctcatgg ctcacgcaa  
5941 aacaacgaac cactagag aacatactgg ctaaatacgg aaggatctga ggttcttatg  
6001 gctcttgat ctatcagta agcatcaaga ctaacaaca aaagtagaac aactgttcac  
6061 cgttagatat caaaggaaa actgtccata tgcacagatg aaaacggtg aaaaaagata  
6121 gatacatcag agcttttac agtttttggc gatttaaag ctgttaccac tgaacagatc  
6181 gacaatgtaa cagatgaaca gcatgtaaca ctaatagaa caggtgaaac cagtaaaaa  
6241 aagcaactag aacatgaaat tgaacacctg agacaactg ttacagctca acagtcacac  
6301 atagacagcc tgaacagggc gatgctgctt atcgaatcaa agctgccgac aacacgggag  
6361 ccagtgacgc ctcccgtggg gaaaaaatca tggcaattct ggaagaaata gcctttcag

6421 ccggcaaacc tgaagccgga tctgcgattc tgataacaaa ctagcaacac cagaacagcg  
6481 cggccgcaaa aggaaaa

//

**ZFP02 matched atc/rep on SC101/Kan**

LOCUS 12-27 6472 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(4354..34)  
/gene="Ori SC101"  
CDS 3386..4180  
/gene="KanR"  
/codon\_start="0"  
terminator complement(99..272)  
/gene="Terminator"  
terminator complement(4201..4316)  
/gene="Terminator"  
terminator 53..98  
/gene="Terminator"  
gene 1924..2643  
/gene="GFPmut3b"  
terminator 309..354  
/gene="term BBa\_B0011"  
prot\_bind 421..435  
/gene="RBS"  
promoter 361..414  
/gene="pLtetO-1"  
CDS 442..1185  
/gene="RNAP alpha subunit"  
/codon\_start="0"  
CDS 1192..1209  
/gene="linker"  
/codon\_start="0"  
promoter 1835..1896  
/gene="pwklac"  
CDS 1216..1746  
/gene="ZFP02"  
/codon\_start="0"  
prot\_bind 1811..1828  
/gene="ZFP02 binding site"  
BASE COUNT 1909 a 1438 c 1543 g 1582 t  
ORIGIN  
1 gattctgata acaaactagc aacaccagaa cagcgcggcc gcaaaaggaa aagatccggc  
61 aaacaaacca ccgttgtag cggtagttt tttgttggga tcgacaatct tcgtaagcgt  
121 catcaataag cgtaaaaaa ccgggcaatg cccggtttt taatgagaaa tttacctgt

181 cgtagccgcc accatccggc aaagaagcat acaaggcttt tggcttatag ctacgtagcg  
241 cattgcgtcg cagcacaatc ccggcaccga tcaagtcttc gcgatgatta ttagaattca  
301 tgagatctag agaataaaa aagccagatt attaatccgg cttttttatt atttggatct  
361 tccctatcag tgatagagat tgacatccct atcagtgata gagatactga gcacggatct  
421 attaaagagg agaaaggatc tatgcagggt tctgtgacag agtttctaaa accgcgcctg  
481 gttgatatcg agcaagtgag ttcgacgcac gccaaggtga cccttgagcc ttagagcgt  
541 ggctttggcc atactctggg taacgcactg cgccgtattc tgctctcadc gatgccgggt  
601 tgcgcggtga ccgaggtga gattgatggt gtactacatg agtacagcac caaagaaggc  
661 gttcaggaag atatcctgga aatcctgctc aacctgaaa ggctggcgggt gagagttcag  
721 ggcaaagatg aagttattct taccttgaat aaatctggca ttggccctgt gactgcagcc  
781 gatatcacc acgacggtga tgcgaaaatc gtcaagccgc agcacgtgat ctgccacctg  
841 accgatgaga acgctctat tagcatgctg atcaaagttc agcgcggtcg tggttatgtg  
901 ccggcttcta cccgaattca ttcggaagaa gatgagcgc caatcggccc tctgctggtc  
961 gacgcatgct acagccctgt ggagcgtatt gcctacaatg ttgaagcagc gcgtgtagaa  
1021 cagcgtaccg acctggacaa gctggtcadc gaaatggaaa ccaacggcac aatcgatct  
1081 gaagaggcga ttcgtcgtgc ggcaaccatt ctggctgaac aactggaagc ttcgttgac  
1141 ttacgtgatg tacgtcagcc tgaagtgaag gaagagaaac cagagggatc tgccccgcga  
1201 gtccggaccg gatctatgct ggaaccaggt gaaaagcctt ataatgtcc ggaatgtggc  
1261 aaaagcttct ccaccagcgg tgagctgggt cgtcaccaac gtactcacac tggtgagaag  
1321 ccgtacaaat gcctgaatg cggtaaagac ttagccagc gtgcgcacct ggaacgcat  
1381 cagcgcacc ataccggcga gaaaccgtat aaatgcccgg aatgtggtaa gtcttttcc  
1441 cagaacagca ccctgactga acaccagcgc actcacaccg gcgaaaaacc atacaagtgc  
1501 ccggagtgcg gcaaatcctt ctctgacca ggtaatctgg ttcgtcatca acgtactcat  
1561 acgggtgaaa agccgtataa atgtccgga tgcggtaaat cttcagccg ccgtgatgaa  
1621 ctgaacgtac atcaacgtac ccactgggt gagaaaccgt acaagtgcc agaatgcggc  
1681 aaatctttt ctcagaagtc cagcctgac gcgcaccagc gtactcacac cggtaaagaa  
1741 acttctaat aaggatctag agaataaaa aagccagatt attaatccgg cttttttatt  
1801 atttggatct ataatggacc taggagctgg atctccgtat tctttact ttatgcttc  
1861 ggctcgtatg ttgtgtcgc cgagcggata acaattggat ctattaaaga ggagaaagga  
1921 tctatgcgta aaggagaaga acttttact ggagttgtcc caattctgt tgaattgat  
1981 ggtgatgta atgggcacaa atttctgtc agtggagagg gtgaaggtga tgcaacatac  
2041 ggaaaactta ccctaaatt tattgact actggaaaac tacctgtcc atggccaaca  
2101 cttgtcacta cttcggta tgggttcaa tgctttgcga gataccaga tcatatgaaa  
2161 cagcatgact tttcaagag tgccatgccc gaaggttatg tacaggaaag aactatatt  
2221 ttcaagatg acgggaacta caagacacgt gctgaagtca agtttgaagg tgatacctt  
2281 gttaatagaa tcgagttaaa aggtattgat tttaaagaag atggaaacat tcttgacac  
2341 aaattggaat acaactataa ctacacaat gtatacatca tggcagacaa acaaaagaat  
2401 ggaatcaaag ttaactcaa aattagacac aacattgaag atggaagcgt tcaactagca  
2461 gaccattatc aacaaaatac tcaattggc gatggccctg tcctttacc agacaacct  
2521 tacctgtcca cacaatctgc ctttcgaaa gatccaacg aaaagagaga ccacatggc  
2581 cttcttgagt ttgtaaccg tgctgggatt acacatggca tggatgaact atacaaata  
2641 taaggatcta gagaataaa aaagccagat tattaatccg gctttttat tattggatc  
2701 ctaactcgag tgcaggtcgt ctggtatcga gaaggacacg gtaatacta ggctgctgg  
2761 ctggtaatcg ccagcaggcc tttttattg ggggagaggg aagtcatgaa aaaactaacc  
2821 ttgaaattc gatctccacc acatcagctc tgaagcaacg taaaaaac cgccccggc  
2881 ggtttttta taccgtagt atcccactt atctacaata gctgtccta ataaaggtg

2941 aataaataaa aacagccgtt gccagaaaga ggcacggctg ttttatttt ctagtgagac  
3001 cgggagcagt taaacgcaga aaggcccacc cgaaggtgag ccagtgtgac tctagtagag  
3061 agcgttcacc gacaaacaac agataaacg aaaggcccag tctttcgact gagcctttcg  
3121 ttttatttga tgcctgggta ttattatttg tacagctcat ccatgccacc ggtcaacagg  
3181 agtccaagcg accggtgggt gcatgtctag ctagctagaa caggactagc taatggttc  
3241 ttagacgtcg gaattgccag ctggggcgcc ctctggtaag gttgggaagc cctgcaaagt  
3301 aaactggatg gctttcttgc cgccaaggat ctgatggcgc aggggatcaa gatctgatca  
3361 agagacagga tgaggatcgt ttcgcatgat tgaacaagat ggattgcacg caggttctcc  
3421 ggccgcttgg gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc  
3481 tgatccgcc gtgtccggc tgctcagcga ggggcgccc gttcttttg tcaagaccga  
3541 cctgtccggt gccctgaatg aactgcagga cgaggcagcg cggctatcgt ggctggccac  
3601 gacggggcgt ccttgcgag ctgtgctcga cgttgcact gaagcgggaa gggactggct  
3661 gctattgggc gaagtgccgg ggcaggatct cctgtcatct caccttctc ctgccgagaa  
3721 agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgcc  
3781 attcgaccac caagcgaac atcgcatcga gcgagcacgt actcggatgg aagccggtc  
3841 tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc  
3901 caggctcaag gcgcatcgc ccgacggcga ggatctcgtc gtgacctatg gcgatgcctg  
3961 cttgccgaat atcatgggtg aaaatggccg cttttctgga ttcacgact gtggccggct  
4021 ggggtgtggcg gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct  
4081 tggcggcgaa tgggctgacc gcttctcgt gctttacggg atcggcgtc ccgattcgca  
4141 gcgcatcgcc ttctatgcc ttcttgacga gttcttctga gcgggactct ggggttcgag  
4201 agctcgttg gactcctgtt gatagatcca gtaatgacct cagaactcca tctggattg  
4261 ttcagaacgc tcggttgcg ccgggcgttt tttattggtg agaatccaag cactagggac  
4321 agtaagacgg gtaagcctgt tgatgatacc gctgccttac tgggtgcatt agccagtctg  
4381 aatgacctgt cacgggataa tccgaagtgg tcagactgga aatcagagg gcaggaactg  
4441 ctgaacagca aaaagtcaga tagcaccaca tagcagacc gccataaac gcctgagaa  
4501 gcccgtgacg ggcttttctt gtattatggg tagtttctt gcatgaatcc ataaaaggcg  
4561 cctgtagtgc cattacccc cattactgc cagagccgtg agcgcagcga actgaatgc  
4621 acgaaaaaga cagcgactca ggtgctgat ggtcggagac aaaaggaata ttcagcgatt  
4681 tgcccagct tgcgagggtg ctacttaagc ctttagggtt ttaaggtctg tttgtagag  
4741 gagcaaacag cgtttgcgac atcctttgt aatactcgg aactgactaa agtagtgagt  
4801 tatacacagg gctgggatct attctttta tctttttta ttcttctt atctataaa  
4861 ttataaccac ttgaatataa acaaaaaaaaa cacacaaagg tctagcggaa tttacagagg  
4921 gtctagcaga atttacaagt tttccagcaa aggtctagca gaattacag ataccacaa  
4981 ctcaaaggaa aaggactagt aattatcatt gactagccca tctcaattgg tatagtgatt  
5041 aaaatcacct agaccaattg agatgtatgt ctgaattagt tgtttcaaaa gcaaatgaac  
5101 tagcgattag tcgctatgac ttaacggagc atgaaaccaa gctaattta tgctgtgtgg  
5161 cactactcaa cccacgatt gaaaacccta caaggaaaga acggacggta tcgttactt  
5221 ataaccaata cgctcagatg atgaacatca gtagggaaaa tgcttatggt gtattagcta  
5281 aagcaaccag agagctgatg acgagaactg tggaaatcag gaatccttg gtaaagggt  
5341 ttgagatttt ccagtggaca aactatgcc agttctcaag cgaaaaatta gaattagtt  
5401 ttagtgaaga gatattgcct tatctttcc agttaaaaa attcataaaa tataatctgg  
5461 aacatgttaa gtctttgaa acaataact ctatgaggat ttatgagtgg ttattaaag  
5521 aactaacaca aaagaaaact cacaaggcaa atatagagat tagccttgat gaatttaagt  
5581 tcatgttaat gcttgaanaa aactaccatg agtttaaaag gcttaacca tgggtttga  
5641 aaccaataag taaagattta aacttaca gcaatatgaa attggtggtt gataagcgag

5701 gccgcccgcac tgatacgttg atttccaag ttgaactaga tagacaaatg gatctcgtaa  
 5761 ccgaacttga gaacaaccag ataaaaatga atggtgacaa aataccaaca accattacat  
 5821 cagattccta cctacgtaac ggactaagaa aaacactaca cgatgcttta actgcaaaaa  
 5881 ttcagctcac cagttttgag gcaaaatfff tgagtgcacat gcaaagtaag catgatctca  
 5941 atggttcggt ctcattggctc acgcaaaaac aacgaaccac actagagaac atactggcta  
 6001 aatacgggaag gatctgaggt tcttatggct cttgtatcta tcagtgaagc atcaagacta  
 6061 acaaacaaaa gtagaacaac tgttcaccgt tagatatcaa agggaaaact gtccatgatg  
 6121 acagatgaaa acggtgtaaa aaagatagat acatcagagc ttttacgagt ttttgggtgca  
 6181 tttaaagctg ttcacatga acagatcgac aatgtaacag atgaacagca tgtaaacacct  
 6241 aatagaacag gtgaaccag taaaacaaag caactagaac atgaaattga acacctgaga  
 6301 caactgttta cagctcaaca gtcacacata gacagcctga aacaggcgat gctgcttacc  
 6361 gaatcaaagc tgccgacaac acgggagcca gtgacgcctc ccgtggggaa aaaatcatgg  
 6421 caattctgga agaaatagcg ctttcagccg gcaaacctga agccgatct gc

//

**ZFP5475 inverter on SC101/Kan**

LOCUS 12-28 7233 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(5063..7215)  
     /gene="Ori SC101"  
 CDS 4095..4889  
     /gene="KanR"  
     /codon\_start="0"  
 terminator complement(4910..5025)  
     /gene="Terminator"  
 CDS 426..1145  
     /gene="GFPmut3b"  
     /codon\_start="0"  
 terminator 257..302  
     /gene="term BBa\_B0011"  
 prot\_bind 405..419  
     /gene="RBS"  
 promoter 1204..1257  
     /gene="pLtetO-1"  
 CDS 2021..2764  
     /gene="RNAP alpha subunit"  
     /codon\_start="0"  
 CDS 2771..2788  
     /gene="linker"  
     /codon\_start="0"  
 prot\_bind 309..330  
     /gene="ZFP5475 binding site"  
 promoter 337..398  
     /gene="pwklac"



```

CDS      1288..1845
         /gene="ZFP5475"
         /codon_start="0"
terminator 1152..1197
         /gene="term BBa_B0011"
promoter   1852..1897
         /gene="term BBa_B0011"
terminator 3359..3404
         /gene="term BBa_B0011"
CDS      2795..3352
         /gene="ZFP5475"
         /codon_start="0"
promoter   1932..1993
         /gene="pwklac"
prot_bind  1904..1925
         /gene="ZFP5475 binding site"
prot_bind  1264..1278
         /gene="RBS"
prot_bind  2000..2014
         /gene="RBS"
BASE COUNT 2081 a 1679 c 1741 g 1732 t
ORIGIN
  1 gatccggcaa acaaaccacc gttgtagcg gtggttttt tgttggatc gacaatcttc
 61 gtaagcgta tcaataagcg taaaaaaacc gggcaatgcc cggttttta atgagaaatt
121 ttacctgtcg tagccgccac catccggcaa agaagcatac aaggctttg gcttatagct
181 acgtagcgc tgcgtcgca gcacaatccc ggcaccgatc aagtcttcgc gatgattatt
241 agaattcaaa agatctagag aatataaaaa gccagattat taatccggct ttttattat
301 ttggatctac acccgggttc ccctcgggtg ggatctccgt attctttaca cttatgctt
361 cgggctcgta tgttgtgctg accgagcgga taacaattgg atctattaaa gaggagaaag
421 gatctatgcg taaaggagaa gaactttca ctggagtgtg cccaattctt gttgaattag
481 atggtgatgt taatgggcac aaatctctg tcagtggaga gggatgaaggt gatgcaacat
541 acggaaaact taccttaaa tttattgca ctactggaaa actacctgtt ccatggccaa
601 cacttgtcac tactttcggg tatggtgttc aatgctttgc gagataccca gatcatatga
661 aacagcatga cttttcaag agtgccatgc ccgaaggta tgtacaggaa agaactatat
721 tttcaaaga tgacgggaac tacaagacac gtgctgaagt caagtttgaa ggtgatacc
781 ttgtaatag aatcgagta aaaggtattg attttaaaaga agatggaaac attcttggac
841 acaaattgga atacaactat aactcacaca atgtatacat catggcagac aaacaaaaga
901 atggaatcaa agttaacttc aaaattagac acaacattga agatggaagc gttcaactag
961 cagaccatta tcaacaaaat actccaattg gcgatggccc tgcctttta ccagacaacc
1021 attacctgtc cacacaatct gccctttcga aagatcccaa cgaaaagaga gaccatagc
1081 tccttcttga gtttgaacc gctgctggga ttacacatgg catggatgaa ctatacaaat
1141 aataaggatc tagagaatat aaaaagccag attattaatc cgctttttt attatttgg
1201 tcttcctat cagtgataga gattgacatc cctatcagt atagagatac tgagcacgga
1261 tctattaaag aggagaaagg atctatgggg gtaccgccca tggcggagag gccttccag
1321 tgtcgaatct gcatgcgtaa cttcagtcgt agtgaccacc tgagccggca catccgcacc
1381 cacacaggcg agaagccttt tgctgtgac atttgtggga ggaaatttgc cgacaaccg

```

1441 gaccgcacaa agcataccaa gatacacacg ggccgacagc ggccgtacgc atgccctgtc  
1501 gagtctctgcg atcgccgctt ttctgacagg aagacactta tcgagcatat ccgcatccac  
1561 accggtcaga agcccttcca gtgtcgaatc tgcatgcgta acttcagtac cagcagcggc  
1621 ctgagccgcc acatccgcac ccacacagga tctcagaagc cttccagtg tcgaatctgc  
1681 atgcgtaact tcagtcgtag tgaccacctg agcgaacaca tccgcacca cacaggcgag  
1741 aagccttttg cctgtgacat ttgtgggagg aaatttgcca ccagcagcga ccgcacaaag  
1801 cataccaaga tacacctgcg ccaaaaagat gcggcccggg aataaggatc tagagaatc  
1861 aaaaagccag attattaatc cggctttttt attatttga tctacaccg ggtcccctc  
1921 ggggtgggac tccgtattct ttacacttta tgcttccggc tcgtatgttg tctcgaccga  
1981 gcgataaca attggatcta ttaaagagga gaaaggatct atgcagggtt ctgtgacaga  
2041 gtttctaaaa ccgcgctgg ttgatatcga gcaagtgagt tcgacgcacg ccaaggtgac  
2101 cttgagcct ttagagcgtg gctttggcca tactctgggt aacgcactgc gccgtattct  
2161 gctctcatcg atgccgggtt gcgcggtgac cgaggttgag attgatggtg tactacatga  
2221 gtacagcacc aaagaaggcg ttcaggaaga tatcttgaa atcttctca acctgaaagg  
2281 gctggcggtg agagttcagg gcaaagatga agttattctt acctgaata aatctggcat  
2341 tggcctctg actgcagccg atatcacca cgacgggat gtcgaaatcg tcaagccgca  
2401 gcacgtgac tgccacctga ccgatgagaa cgcgtctatt agcatgcgta tcaaagttca  
2461 gcgcggtcgt gggtatgtc cggcttctac ccgaattcat tcggaagaag atgagcggc  
2521 aatcggcctg ctgctggctg acgcatgcta cagcctgtg gagcgtattg cctacaatgt  
2581 tgaagcagc cgtgtagaac agcgtaccga cctggacaag ctggatcatg aatggaaac  
2641 caacggcaca atcgatcctg aagaggcgat tcgtcgtgcg gcaaccattc tggctgaaca  
2701 actggaagct ttcgtgact tacgtgatgt acgtcagcct gaagtgaag aagagaaacc  
2761 agagggatct gcccgcgag tccggaccgg atctggggta cccgcatgg cggagaggcc  
2821 cttccagtg cgaatctgca tgcgtaactt cagtcgtagt gaccacctga gccggcacat  
2881 ccgacccac acaggcgaga agcctttgct ctgtgacatt tgtgggagga aatttgcca  
2941 caaccgggac cgcaaaaagc ataccaagat acacacgggc ggacagcggc cgtacgatg  
3001 ccctgtcag tctcgcgac gccgctttc tgacaggaag acacttatcg agcatatccg  
3061 catcacacc ggtcagaagc cttccagtg tcgaatctgc atgcgtaact tcagtaccag  
3121 cagcggcctg agccgcaca tccgcacca cacaggatct cagaagcct tccagtgtcg  
3181 aatctgatg cgtaactta gtcgtagtga ccacctgac gaacacatcc gcaaccacac  
3241 aggcgagaag cttttgctg gtgacattg tgggaggaaa tttgccaca gcagcgaccg  
3301 caaaagcat accaagatac acctgcgcca aaaagatgcg gcccggaat aaggatctag  
3361 agaataaaa aagccagatt attaatccgg ctttttatt atttgatcc taactcgagt  
3421 gcaggtcgtc tcggatcag aaggacagc ttaatactag gcctgctggc tggtaatcgc  
3481 cagcaggcct tttatttgg gggagagga agtcatgaaa aactaacct ttgaaattcg  
3541 atctccacca catcagctct gaagcaactg aaaaaaaccc gcccgggcg gttttttat  
3601 acccgtagta tccccacta tctacaatag ctgtcctta ttaaggttga ataaataaaa  
3661 acagccgtg ccagaaagag gcacggctgt tttattttc tagtgagacc gggagcagtt  
3721 aaacgcagaa agggccacc gaaggtgac cagtgtgact ctagtagaga gcgttcaccg  
3781 acaacaaca gataaacga aaggccagc cttcactg agccttctg ttatttgat  
3841 gcctggttat tattattgt acagctcat catgccacc gtcaacagga gtccaagcga  
3901 ccggtggtg catgtctag tagctagaac aggactagc aatggttct tagacgtcgg  
3961 aattgccagc tggggcgccc tctgtaagg ttgggaagcc ctgcaaagta aactggatg  
4021 ctttctgccc gcaaggatc tgatggcgca ggggatcaag atctgatcaa gagacaggat  
4081 gaggatcgt tcgatgatt gaacaagat gattgcacgc aggttctcc gccgcttggg  
4141 tggagaggct attcggctat gactgggcac aacagacaat cggctgctct gatgccgcc

4201 tgttccggct gtcagcgag gggcgcccg ttcttttgt caagaccgac ctgtccggtg  
4261 ccctgaatga actgcaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc  
4321 cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg  
4381 aagtgccggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca  
4441 tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgccc ttcgaccacc  
4501 aagcgaaca tcgcatcgag cgagcacgta ctcgatgga agccggtctt gtcgatcagg  
4561 atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg  
4621 cgcgcatgcc cgacggcgag gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata  
4681 tcatggtgga aatggcgcg ttttctggat tcatcgactg tggccggctg ggtgtggcgg  
4741 accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat  
4801 gggctgaccg cttcctcgtg ctttacggta tcgccgctcc cgattcgag cgcatcgcct  
4861 tctatcgcct tcttgacgag ttcttctgag cgggactctg gggttcgaga gctcgttgg  
4921 actcctgttg atagatccag taatgacctc agaactccat ctggatttgt tcagaacgct  
4981 cggttgccgc cgggcgtttt ttattggtga gaatccaagc actagggaca gtaagacggg  
5041 taagcctggt gatgataccg ctgccttact ggggtgcatta gccagtctga atgacctgct  
5101 acgggataat cgaagtggg cagactggaa aatcagaggg caggaaactgc tgaacagcaa  
5161 aaagtcagat agcaccacat agcagaccg ccataaaacg cctgagaag cccgtgacgg  
5221 gcttttcttg tattatgggt agtttcttg catgaatcca taaaaggcgc ctgtagtgc  
5281 atttaccctt attcactgcc agagccgtga gcgcagcgaa ctgaatgtca cgaaaaagac  
5341 agcgactcag gtgcctgatg gtcggagaca aaaggaatat tcagcgattt gccgagctt  
5401 gcgagggtgc tacttaagcc ttagggttt taaggctctg tttgtagagg agcaaacagc  
5461 gtttgcgaca tccttttga atactgcgga actgactaaa gtagtgagtt atacacaggg  
5521 ctgggatcta ttcttttat cttttttat tcttcttta ttctataaat tataaccact  
5581 tgaataataa caaaaaaac acacaaagg ctagcggaa ttacagaggg ttagcagaa  
5641 ttacaagt ttccagcaa ggtctagcag aatttacaga taccacaac tcaaaggaaa  
5701 aggactagta attatcattg actagccat ctcaattggt atagtgatta aatcaccta  
5761 gaccaattga gatgtatgc tgaattagtt gtttcaaag caaatgaact agcgattagt  
5821 cgctatgact taacggagca tgaaccaag ctaattttat gctgtgtggc actactcaac  
5881 cccagattg aaaaccctac aaggaaagaa cggacggat cgttactta taaccaatac  
5941 gctcagatga tgaacatcag tagggaaaat gcttatggtg tattagctaa agcaaccaga  
6001 gagctgatga cgagaactgt ggaaatcagg aatccttgg ttaaaggctt tgagatttc  
6061 cagtggaca actatgcaa gttctcaag caaaaattag aattagttt tagtgaagag  
6121 atattgcctt atctttcca gttaaaaaa ttataaaat ataactgga acatgttaag  
6181 tttttgaaa acaatactc tatgaggatt tatgagtgg tattaaaaga actaacaaa  
6241 aagaaaactc acaaggcaa tatagagatt agccttgatg aatttaagtt catgttaatg  
6301 ctgaaaata actaccatga gttaaaagg ctaaccaat gggttttgaa accaataagt  
6361 aaagatttaa acacttacg caatatgaaa ttggtggtg ataagcgagg ccgcccgact  
6421 gatcgttga tttccaagt tgaactagat agacaaatgg atctcgtaac cgaacttgag  
6481 aacaaccaga taaaaatgaa tggtgacaaa ataccaaaa ccattacatc agattcctac  
6541 ctacgtaacg gactaagaaa aactacac gatgctttaa ctgcaaaaat tcagctcacc  
6601 agttttgagg caaaatttt gagtgacatg caaagtaagc atgatctca tggttcgttc  
6661 tcatggctca cgaaaaaca acgaaccaca ctagagaaca tactggctaa atacggaagg  
6721 atctgaggtt ctatggctc ttgtatctat cagtgaagca tcaagactaa caaacaag  
6781 tagaacaact gttcaccgtt agatatcaaa gggaaaactg tccatagca cagatgaaaa  
6841 cgggtgaaaa aagatagata catcagagct ttacgagtt tttggtgcat taaagctgt  
6901 tccatgaa cagatcgaca atgtaacaga tgaacagcat gtaacaccta atagaacagg

6961 tgaaccagt aaaacaagc aactagaaca tgaattgaa cacctgagac aacttgttac  
7021 agctcaacag tcacacatag acagcctgaa acaggcgatg ctgcttatcg aatcaaagct  
7081 gccgacaaca cgggagccag tgacgcctcc cgtggggaaa aaatcatggc aattctggaa  
7141 gaaatagcgc tttcagccgg caaacctgaa gccggatctg cgattctgat aacaaactag  
7201 caacaccaga acagcgcggc cgcaaaagga aaa

//

**ZFP03 matched atc/rep on SC101/Kan**

LOCUS 12-29 6469 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(4351..34)  
/gene="Ori SC101"  
CDS 3383..4177  
/gene="KanR"  
/codon\_start="0"  
terminator complement(99..272)  
/gene="Terminator"  
terminator complement(4198..4313)  
/gene="Terminator"  
terminator 53..98  
/gene="Terminator"  
CDS 474..1193  
/gene="GFPmut3b"  
/codon\_start="0"  
terminator 309..354  
/gene="term BBa\_B0011"  
prot\_bind 453..467  
/gene="RBS"  
promoter 1252..1305  
/gene="pLtetO-1"  
CDS 1333..2076  
/gene="RNAP alpha"  
/codon\_start="0"  
CDS 2083..2100  
/gene="linker"  
/codon\_start="0"  
prot\_bind 361..378  
/gene="ZFP03 binding site"  
CDS 2107..2640  
/gene="ZFP03"  
/codon\_start="0"  
promoter 385..446  
/gene="pwklac"  
BASE COUNT 1901 a 1438 c 1507 g 1623 t

## ORIGIN

1 gattctgata acaaactagc aacaccagaa cagcgcggcc gcaaaaggaa aagatccggc  
61 aaacaaacca ccgttggtag cgggtggttt tttgtttgga tcgacaatct tcgtaagcgt  
121 catcaataag cgtaaaaaaa ccgggcaatg cccggttttt taatgagaaa ttttacctgt  
181 cgtagccgcc accatccggc aaagaagcat acaaggcttt tggcttatag ctacgtagcg  
241 cattgcgtcg cagcacaatc ccggcaccga tcaagtcttc gcgatgatta ttagaattca  
301 tgagatctag agaataataa aagccagatt attaatccgg cttttttatt atttggatct  
361 ggtagacgtc tagtaactgg atctccgtat tctttacact ttatgcttcc ggctcgtatg  
421 ttgtgtcgac cgagcggata acaattggat ctattaaaga ggagaaagga tctatgcgta  
481 aaggagaaga acttttact ggagttgtcc caattctgt tgaattagat ggtgatgta  
541 atgggcacaa attttctgtc agtggagagg gtgaaggtga tgcaacatac ggaaaactta  
601 ccctaaatt tatttgact actggaaaac tacctgttcc atggccaaca cttgtcacta  
661 cttcgggta tgggtttcaa tgctttgca gataccaga tcatatgaaa cagcatgact  
721 tttcaagag tgccatgccc gaaggttatg tacaggaaag aactatattt tcaaaagatg  
781 acgggaacta caagacacgt gctgaagtca agtttgaagg tgataccctt gtaatagaa  
841 tcgagttaa aggtattgat ttaaagaag atggaaacat tcttgacac aaattggaat  
901 acaactataa ctacacaaat gtatacatca tggcagacaa acaaaagaat ggaatcaaag  
961 ttaacttcaa aattagacac aacattgaag atggaagcgt tcaactagca gaccattatc  
1021 aacaaaatac tccaattggc gatggccctg tcctttacc agacaacct tacctgtcca  
1081 cacaatctgc ctttcgaaa gatcccaacg aaaagagaga ccacatggtc cttcttgagt  
1141 ttgtaaccgc tgctgggatt acacatggca tggatgaact atacaaataa taaggatcta  
1201 gagaataata aaagccagat tattaatccg gctttttat tattggatc ttcctatca  
1261 gtgatagaga ttgacatccc tatcagtgat agagatactg agcacggatc tattaagag  
1321 gagaaaggat ctatgcaggg ttctgtgaca gagtttctaa aaccgcgct ggtgatatc  
1381 gagcaagtga gttcgacgca cgccaaggtg acccttgagc cttagagcg tggcttggc  
1441 catactctgg gtaacgcact gcgccgtatt ctgctctcat cgatgccggg ttgcgcggtg  
1501 accgaggtg agattgatgg gtactacat gactacagca ccaaagaagg cgttcaggaa  
1561 gatcctctgg aatcctgct caacctgaaa gggctggcgg tgagagtca gggcaaagat  
1621 gaagttattc ttacctgaa taaatctggc attggcctg tgactgcagc cgatcacc  
1681 cagcaggtg atgtcgaaat cgtcaagccg cagcacgtga tctgccacct gaccgatgag  
1741 aacgcgtcta ttagcatgcg tatcaaagtt cagcgcggtc gtggttatgt gccggttct  
1801 accgaattc attcggaga agatgagcgc ccaatcggcc gtctgctggt cgacgcatgc  
1861 tacagccctg tggagcgtat gcctacaat gttgaagcag cgcgtgtaga acagcgtacc  
1921 gacctggaca agctggtcat cgaatggaa accaacggca caatcgatcc tgaagaggcg  
1981 attcgtctg cggaaccat tctggctgaa caactggaag cttcgttga cttactgat  
2041 gtacgtcagc ctgaagtga agaagagaaa ccagaggat ctgccccgcg agtccggacc  
2101 ggatctctgg aaccaggatc taaaccgtac aaatgtccgg aatgtgtaa atccttctcc  
2161 actcatctgg atctgattcg tcatcaactg actcacactg gatctaaacc gtacaaatg  
2221 ccggaatgtg gtaaactctt ctccaatct tcttctctgg ttcgtcatca acgtactcac  
2281 actggatcta aaccgtaca atgtccgga tgtggtaaat cttctccca aaacttacc  
2341 ctgactgaac atcaacgtac tcactgga tctaaaccgt acaaatgtcc ggaatgtggt  
2401 aaatccttct cctctcgtcg tacttgcgt gcacatcaac gtactcacac tggatctaaa  
2461 ccgtacaaat gtccggaatg tggtaaactc ttctccaac tggctcatct gcgtgcacat  
2521 caacgtactc aactggatc taaaccgtac aaatgtccgg aatgtgtaa atccttctcc  
2581 acttctggtc atctggtac tcatcaactg actcacactg gatctaaaac ctcttaataa  
2641 ggatctagag aatataaaaa gccagattat taatccgct tttttattat ttggatccta

2701 actcgagtgc aggtcgtctc ggatcgagaa ggacacgggt aatactaggc ctgctggctg  
2761 gtaatcgcca gcaggccttt ttattgggg gagaggggaag tcatgaaaa actaaccttt  
2821 gaaattcgat ctccaccaca tcagctctga agcaacgtaa aaaaacccgc cccggcgggt  
2881 tttttatac ccgtagtata cccacttata tacaatagct gtccttaatt aaggttgaat  
2941 aaataaaaac agccgtgcc agaaagaggc acggctgttt ttatttcta gtgagaccgg  
3001 gagcagttaa acgcagaaag gccacccga aggtgagcca gtgtgactct agtagagagc  
3061 gttcaccgac aaacaacaga taaaacgaaa ggccagctct ttcgactgag ccttcggtt  
3121 tatttgatgc ctggttatta ttattgtac agtcatca tgccaccggt caacaggagt  
3181 ccaagcgacc ggtggttga tcttagcta gtagaacag gactagctaa tggtttcta  
3241 gacgtcggaa ttgccagctg gggcgcctc tggtaagggt gggaaagcct gcaaagtaa  
3301 ctggatggct ttctgccc caaggatctg atggcgagg ggatcaagat ctgatcaaga  
3361 gacaggatga ggatcgttc gcatgattga acaagatgga ttgcacgag gttctccggc  
3421 cgctgggtg gagaggctat tcggctatga ctggcaca cagacaatcg gctgcttga  
3481 tgccgctg tccgctgt cagcgagg gcgcccgtt cttttgtca agaccgact  
3541 gtccggtgcc ctgaatgaa tcagggaca ggacgcgg ctatcgtggc tggccagac  
3601 gggcgttct tgcgagctg tgctcagct tgcactgaa gcgggaagg actggtgct  
3661 attggcgaa gtccggggc aggatctct gcatctcac ctgctctg ccgagaaagt  
3721 atccatcatg gctgatgaa tggcggtg gcatcgtt gatccgcta cctgccatt  
3781 cgaccacaa gcgaaatc gcatcgagc agcacgtact cggatggaag ccggtctgt  
3841 cgatcaggat gatctggac aagagcatca ggggctcgc ccagccgaac tttcggccg  
3901 gctcaaggcg cgatgccc acggcgagga tctcgtctg acctatggcg atgctgctt  
3961 gccaatatc atggtgaaa atggcgtt tctggattc atcactgtg gccggtggg  
4021 tgtggcgac cgctacagg acatagcgtt ggctaccgt gatattgctg aagagcttg  
4081 cggcaatgg gctgaccgt tctcgtct ttacggtat gccgtccc attcgcagc  
4141 catgccttc tatgcctc ttgacgagt cttctgagc ggactctggg gttcagagc  
4201 tcgcttgac tctgttgat agatccagta atgacctc aactccatct ggattgttc  
4261 agaacgctg gttgccgccc ggcgttttt attggtgaga atccaagcac tagggacagt  
4321 aagacgggta agcctgtga tgatacgtt gccttactgg gtgcattagc cagtctgaat  
4381 gacctgtac gggataatc gaagtgtca gactggaaaa tcagagggca ggaactgctg  
4441 aacagcaaaa agtcagatag caccatag cagaccgcc ataaaacgcc ctgagaagcc  
4501 cgtgacggc tttcttga ttatggtag tttcttgca tgaatcata aaaggcct  
4561 gtatgcat ttacccat tactgccag agcgtgagc gcagcgaact gaatgtcac  
4621 aaaaagacag cgactcaggt gcctgatgg cggagacaaa agaatatc agcgattgg  
4681 ccgagcttg gagggtgcta ctaagcct tagggttta aggtctgtt ttagaggag  
4741 caaacagct ttgcgacatc ctttgtaat actgcggaac tgactaaagt agtgattat  
4801 acacagggt ggtatctatt cttttatct tttttatt tttttatt ctataaata  
4861 taaccattg aatataaca aaaaaaac acaaaggtct agcggattt acagagggtc  
4921 tagcagaatt tacaagttt ccagcaagg ttagcagaa ttacagata ccacaactc  
4981 aaaggaaaag gactagtaat tatcattgac tagccatct caattggtat agtgattaaa  
5041 atcacatga ccaattgaga tgatgtctg aattagttt ttcaaagca aatgaactg  
5101 cgattagtc ctatgacta acggagcatg aaaccaagct aatttatgc tgtgtggc  
5161 tactcaacc cacgattgaa aacctaca ggaaagaac gacggtatc ttcactata  
5221 accaatacgc tcagatgat aacatcagta gggaaaatgc ttatggtgta ttagctaaag  
5281 caaccagaga gctgatgac agaactgtg aatcaggaa tctttggtt aaaggcttg  
5341 agattttca gggacaaac tatccaagt tctcaagca aaaattagaa ttagtttta  
5401 gtgaagagat attgcctat ctttcagt taaaaaatt cataaatat aatctggaac

5461 atgtaagtc tttgaaaac aaatactcta tgaggattta tgagtggta ttaaaagaac  
 5521 taacacaaa gaaaactcac aaggcaaata tagagattag cttgatgaa ttaagtca  
 5581 tgtaatgct tgaaaataac taccatgagt ttaaaggct taaccaatgg gtttgaac  
 5641 caataagtaa agatttaaac acttacagca atatgaaatt ggtggttgat aagcgaggcc  
 5701 gcccgactga tacgttgatt ttcaagttg aactagatag acaaatggat ctgtaaccg  
 5761 aactgagaa caaccagata aaaatgaatg gtgacaaaat accaacaacc attacatcag  
 5821 attcctacct acgtaacgga ctaagaaaaa cactacacga tgctttaact gcaaaaattc  
 5881 agctcaccag tttgaggca aaatthtga gtgacatgca aagtaagcat gatctcaatg  
 5941 gttcgttctc atggctcacg caaaaacaac gaaccacact agagaacata ctggctaaat  
 6001 acggaaggat ctgaggttct tatggctctt gtatctatca gtgaagcatc aagactaaca  
 6061 aacaaaagta gaacaactgt tcaccgttag atatcaaagg gaaaactgtc catatgcaca  
 6121 gatgaaaacg gtgtaaaaaa gatagataca tcagagcttt tacgagtttt tgggtcattt  
 6181 aaagctgttc accatgaaca gatcgacaat gtaacagatg aacagcatgt aacaccta  
 6241 agaacagtg aaaccagtaa acaaagcaa ctagaacatg aaattgaaca cctgagacaa  
 6301 cttgttacag ctcaacagtc acacatagac agcctgaaac aggcgatgct gcttatcgaa  
 6361 tcaaagctgc cgacaacacg ggagccagtg acgcctcccg tggggaaaaa atcatggcaa  
 6421 ttctggaaga aatagcgctt tcagccggca aacctgaagc cggatctgc

//

### **ZFP02 leak on SC101/Kan**

LOCUS 12-30 4970 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(2544..4696)  
     /gene="Ori SC101"  
 CDS 1576..2370  
     /gene="KanR"  
     /codon\_start="0"  
 terminator complement(4761..4934)  
     /gene="Terminator"  
 terminator complement(2391..2506)  
     /gene="Terminator"  
 terminator 4715..4760  
     /gene="Terminator"  
 CDS 114..833  
     /gene="GFPmut3b"  
     /codon\_start="1"  
 terminator 840..885  
     /gene="term BBa\_B0011"  
 prot\_bind 93..107  
     /gene="RBS"  
 promoter 25..86  
     /gene="pwklac"  
 prot\_bind 1..18  
     /gene="ZFP02 operator site"

BASE COUNT 1498 a 1069 c 1161 g 1242 t

ORIGIN

1 ataatggacc taggagctgg atctccgtat tctttacact ttatgcttcc ggctcgtatg  
61 ttgtgtcgac cgagcggata acaattggat ctattaaaga ggagaaagga tctatgcgta  
121 aaggagaaga acttttact ggagttgtcc caattcttgt tgaattagat ggtgatgta  
181 atgggcacaa attttctgtc agtggagagg gtgaaggatga tgcaacatac ggaaaactta  
241 cccttaaatt tatttgcaact actggaaaac tacctgttcc atggccaaca cttgtcacta  
301 ctttcggtta tgggtttcaa tgctttgcga gataccaga tcatatgaaa cagcatgact  
361 ttttcaagag tgccatgccc gaaggttatg tacaggaaag aactatattt tcaaaagatg  
421 acgggaacta caagacacgt gctgaagtca agtttgaagg tgataccctt gtaatagaa  
481 tcgagttaaa aggtattgat ttaaagaag atggaaacat tcttggacac aaattggaat  
541 acaactataa ctacacaat gtatacatca tggcagacaa acaaaagaat ggaatcaaag  
601 ttaacttcaa aattagacac aacattgaag atggaagcgt tcaactagca gaccattatc  
661 acaaaaatac tccaattggc gatggccctg tccttttacc agacaacat tacctgtcca  
721 cacaatctgc cctttcgaaa gatcccaacg aaaagagaga ccacatggtc cttcttgagt  
781 ttgtaaccgc tgctgggatt acacatggca tggatgaact atacaataa taaggatcta  
841 gagaatataa aaagccagat tattaatccg gctttttat tatttggatc ctaactcgag  
901 tgcaggtcgt ctcggatcga gaaggacacg gtaataacta ggctgctgg ctggtaatcg  
961 ccagcaggcc ttttatttg ggggagaggg aagtcatgaa aaaactaacc tttgaaattc  
1021 gatctccacc acatcagctc tgaagcaacg taaaaaacc cgccccggcg ggtttttta  
1081 taccctagt atccccactt atctacaata gctgtcctta attaagggtg aataaataaa  
1141 aacagccgtt gccagaaaga ggcacggctg ttttattt ctagtgagac cgggagcagt  
1201 taaacgcaga aaggcccacc cgaaggtag cagtgtagac tctagtagag agcgttcacc  
1261 gacaaacaac agataaaacg aaaggccag tctttcgact gagcctttcg ttttattga  
1321 tgcttggtta ttattattg tacagctcat ccatgccacc ggtcaacagg agtccaagcg  
1381 accggtggtt gcatgtctag ctagctagaa caggactagc taatggttc ttagacgtc  
1441 gaattgccag ctggggcgcc ctctggtaag gttgggaagc cctgcaaagt aaactggatg  
1501 gctttctgc cgccaaggat ctgatggcgc aggggatcaa gatctgatca agagacagga  
1561 tgaggatcgt ttcgatgat tgaacaagat ggattgcacg caggttctcc ggccgcttg  
1621 gtggagaggc tattcggcta tgactggca caacagacaa tcggctgctc tgatgccgc  
1681 gtgtccggc tgcagcgca ggggcgccc gttcttttg tcaagaccga cctgtccggt  
1741 gccctgaatg aactgcagga cgaggcagc cggctatcgt ggctggccac gacgggcgtt  
1801 ccttgccag ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattggc  
1861 gaagtccgg ggcaggatct cctgtcatct caccttgctc ctgccagaa agtatccatc  
1921 atggctgatg caatcggcg gctgcatac cttgatccg ctacctgcc attcgaccac  
1981 caagcgaac atcgcacga gcgagcacgt actcggatgg aagccggtct tctgatcag  
2041 gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag  
2101 gcgcatcgc ccgacggcga ggatctcgtc tgacctatg gcgatgctg cttgccgaat  
2161 atcatggtg aaaatggccg ctttctgga tcatcgact gtggccggt ggtgtggcg  
2221 gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa  
2281 tgggctgacc gcttctcgt gctttacggt atcggcctc ccgattcgca gcgcatgcc  
2341 ttctatgcc ttcttgacga gttcttctga gcgggactct ggggttcgag agctcgctt  
2401 gactcctgtt gatagatcca gtaatgacct cagaactcca tctggattg ttcagaacgc  
2461 tcggttgccg ccgggcgtt tttattggtg agaatccaag cactagggac agtaagacgg  
2521 gtaagcctgt tgatgatacc gctgccttac tgggtgcatt agccagtctg aatgacctg  
2581 cacgggataa tccaagtgg tcagactgga aatcagagg gcaggaactg ctgaacagca



2641 aaaagtcaga tagcaccaca tagcagaccc gccataaac gccctgagaa gccctgacg  
2701 ggcttttctt gtattatggg tagtttctt gcatgaatcc ataaaaggcg cctgtagtgc  
2761 catttaccct cattactgc cagagccgtg agcgcagcga actgaatgtc acgaaaaaga  
2821 cagcgactca ggtgcctgat ggtcggagac aaaaggaata ttcagcgatt tgcccagact  
2881 tgcgaggggtg ctacttaagc ctttaggggtt ttaaggtctg tttgtagag gagcaaacg  
2941 cgtttgcgac atcctttgt aatactgcgg aactgactaa agtagtgagt tatacacagg  
3001 gctgggatct attctttta tctttttta ttctttctt attctataaa ttataaccac  
3061 ttgaatataa acaaaaaaaaa cacacaaagg tctagcggaa ttacagagg gtctagcaga  
3121 atttacaagt ttccagcaa aggtctagca gaatttacag ataccacaa ctcaaaggaa  
3181 aaggactagt aattatcatt gactagccca tctcaattgg tatagtattg aaaatcacct  
3241 agaccaattg agatgtatgt ctgaattagt tgttttcaa gcaaatgaac tagcgattg  
3301 tcgctatgac ttaacggagc atgaaaccaa gctaatttta tgctgtgtgg cactactcaa  
3361 cccacgatt gaaaacccta caaggaaaga acggacggta tcgttactt ataaccaata  
3421 cgctcagatg atgaacatca gtagggaaaa tgcttatggg gtattagcta aagcaaccag  
3481 agagctgatg acgagaactg tggaaatcag gaatcctttg gttaaaggct ttgagattt  
3541 ccagtggaca aactatgcca agtctcaag cgaaaaatta gaattagttt ttagtgaaga  
3601 gatattgcct tatctttcc agttaaaaa attcataaaa tataatctgg aacatgttaa  
3661 gtcttttgaa acaaaatact ctatgaggat ttatgagtgg ttattaaaag aactaacaca  
3721 aaagaaaact cacaaggcaa atatagagat tagccttgat gaatttaagt tcatgttaat  
3781 gcttgaanaa aactaccatg agtttaaaag gcttaaccaa tgggttttga aaccaataag  
3841 taaagattta aactctaca gcaatagaa attgggtggg gataagcgag gccgcccac  
3901 tgatacgttg attttcaag ttgaactaga tagacaaatg gatctcgtaa ccgaacttga  
3961 gaacaaccag ataaaaatga atggtgacaa aataccaaca accattacat cagattccta  
4021 cctacgtaac ggactaagaa aacactaca cgatgcttta actgcaaaaa ttcagctcac  
4081 cagttttgag gcaaaatttt tgagtacat gcaaagtaag catgatctca atggttcggt  
4141 ctcatggctc acgcaaaaac aacgaaccac actagagaac atactggcta aatacgggaag  
4201 gatctgaggt tcttatggct cttgtatcta tcaagtgaag atcaagacta acaaacaaaa  
4261 gtagaacaac tgttcaccgt tagatatcaa agggaaaact gtccatagc acagatgaaa  
4321 acggtgtaaa aaagatagat acatcagagc ttttacgagt ttttgggca ttaaagctg  
4381 ttcacatga acagatcagc aatgtaacag atgaacagca tgtaaacct aatagaacag  
4441 gtgaaaccag taaaacaaag caactagaac atgaaattga acacctgaga caactgtta  
4501 cagctcaaca gtcacacata gacagcctga aacaggcgat gctgcttatc gaatcaaagc  
4561 tgccgacaac acgggagcca gtgacgcctc ccgtggggaa aaaatcatgg caattctgga  
4621 agaaatagcg ctttcagccg gcaaactga agccggatct gcgattctga taacaaacta  
4681 gcaacaccag aacagcgcgg ccgcaaaagg aaaagatccg gcaaacaaac caccgttggg  
4741 agcgggtggt ttttggttg gatcgacaat cttcgtgtaag gtcataata agcgtaaaaa  
4801 aaccgggcaa tgcccgggtt ttaatagaga aattttacct gtcgtagccg ccaccatccg  
4861 gcaagaagc atacaaggct tttggcttat agctacgtag cgcattgcgt cgcagcaaa  
4921 tcccggcacc gatcaagtct tcgcatgat tattagaatt catgagatct

//

**Zif268 leak on SC101/Kan**

LOCUS 12-31 4963 bp DNA circular

FEATURES Location/Qualifiers

```

rep_origin complement(2537..4689)
        /gene="Ori SC101"
CDS      1569..2363
        /gene="KanR"
        /codon_start="0"
terminator complement(4754..4927)
        /gene="Terminator"
terminator complement(2384..2499)
        /gene="Terminator"
terminator 4708..4753
        /gene="Terminator"
CDS      107..826
        /gene="GFPmut3b"
        /codon_start="0"
terminator 833..878
        /gene="term BBa_B0011"
prot_bind 86..100
        /gene="RBS"
promoter 18..79
        /gene="pwklac"
prot_bind 1..11
        /gene="Zif268 binding site"
BASE COUNT 1492 a 1068 c 1163 g 1240 t
ORIGIN
1  tgcgtgggcg gggatctccg tattctttac actttatgct tccggctcgt atgttgtgc
61  gaccgagcgg ataacaattg gatctattaa agaggagaaa ggatctatgc gtaaaggaga
121 agaactttc actggagttg tccaattct tgttgaatta gatggtgatg ttaatgggca
181 caaattttc gtcagtggag aggggtgaagg tgatgcaaca tacggaaaac ttacccttaa
241 atttattgc actactggaa aactacctgt tccatggcca aacttgta ctactttcgg
301 ttatggtgtt caatgctttg cgagataccc agatcatatg aacagcatg acttttcaa
361 gagtgccatg cccgaagggt atgtacagga aagaactata ttttcaaag atgacgggaa
421 ctacaagaca cgtgctgaag tcaagttga agtgatacc cttgtaata gaatcgagtt
481 aaaaggatt gattttaaag aagatggaaa cattcttga cacaaattgg aatacaacta
541 taactcacac aatgtataca tcatggcaga caaacaagaag aatggaatca aagttaactt
601 caaaattaga cacaacattg aagatggaag cgttcaacta gcagaccatt atcaacaaaa
661 tactccaatt ggcgatggcc ctgtcctttt accagacaac cttacctgt ccacacaatc
721 tgccctttcg aaagatcca acgaaaagag agaccacatg gtccttctg agtttgtaac
781 cgctgctggg attacacatg gcatggatga actatacaaa taataaggat ctagagaata
841 taaaaagcca gattattaat cgggctttt tattatttgg atcctaactc gtagtcaggt
901 cgtctcgat cgagaaggac acggttaata ctaggcctgc tggctggtaa tcgccagcag
961 gcctttttat tgggggaga gggaagtcac gaaaaaacta acctttgaaa ttcgatctcc
1021 accacatcag ctctgaagca acgtaaaaaa acccgccccg gcgggtttt ttataccggt
1081 agtatcccca cttatctaca atagctgtcc ttaattaagg ttgaataaat aaaaacagcc
1141 gttgccagaa agaggcacgg ctgtttttat tttctagtga gaccgggagc agttaaacgc
1201 agaaaggccc acccgaaggt gagccagtgt gactctagta gagagcgttc accgacaaac
1261 aacagataaa acgaaaggcc cagtcttctg actgagcctt tcgttttatt tgatgcctgg

```

1321 ttattattat ttgtacagct catccatgcc accggtcaac aggagtccaa gcgaccggtg  
1381 gttgcatgtc tagctagcta gaacaggact agctaattggt ttcttagacg tcggaattgc  
1441 cagctggggc gccctctggt aaggtggga agccctgcaa agtaaaactgg atggctttct  
1501 tgccccaag gatctgatgg cgcaggggat caagatctga tcaagagaca ggatgaggat  
1561 cgtttcgcat gattgaacaa gatggattgc acgcaggttc tccggccgct tgggtggaga  
1621 ggctattcgg ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgtcc  
1681 ggctgtcagc gcaggggagc cgggttcttt ttgtcaagac cgacctgtcc ggtgcctga  
1741 atgaactgca ggacgaggca gcgcggctat cgtggctggc cagcagggc gttccttgcg  
1801 cagctgtgct cgactgttc actgaagcgg gaagggactg gctgctattg ggcgaagtgc  
1861 cggggcagga tctcctgtca tctcaccttg ctctgccga gaaagtatcc atcatggctg  
1921 atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac caccaagcga  
1981 aacatcgcat cgagcgagca cgtactcggg tggaaagcgg tctgtcgat caggatgatc  
2041 tggacgaaga gcatcagggg ctgcgccag ccgaactgtt cgccaggctc aaggcgcga  
2101 tgcccagcg cgaggatctc gtcgtgacct atggcgatgc ctgcttgccg aatatcatgg  
2161 tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct  
2221 atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc gaatgggctg  
2281 accgcttctc cgtgcttac ggtatcggc ctcccattc gcagcgcac gccttctatc  
2341 gccttctga cgagtcttc tgagcgggac tctggggttc gagagctcgc ttgactctc  
2401 gttgatagat ccagtaatga cctcagaact ccatctggat ttgttcagaa cgctcgggtg  
2461 ccgcccggcg tttttattg gtgagaatcc aagcactagg gacagtaaga cgggtaagcc  
2521 tgttgatgat accgctgct tactgggtgc attagccagt ctgaatgacc tgtcacggga  
2581 taatccgaag tggtcagact ggaaaatcag agggcaggaa ctgctgaaca gcaaaaagtc  
2641 agatagcacc acatagcaga cccgccataa aacgcctga gaagcccgtg acgggctttt  
2701 cttgtattat gggtagtttc cttgcatgaa tccataaaag gcgcctgtag tgccattac  
2761 cccattcac tgccagagcc gtgagcgcag cgaactgaat gtcacgaaaa agacagcgac  
2821 tcaggtgctc gatggtcggg gacaaaagga atattcagcg atttgcccga gcttgcgagg  
2881 gtgctactta agccttagg gtttaaggt ctgtttgta gaggagcaaa cagcgtttgc  
2941 gacatcttt tgaataactg cggaactgac taaagtagtg agttatacac agggctggga  
3001 tctattcttt ttatctttt ttattcttc tttattctat aaattataac cactgaata  
3061 taaacaaaaa aaacacacaa aggtctagcg gaatttacag agggctagc agaattaca  
3121 agttttccag caaaggctca gcagaatta cagatacca caactcaaag gaaaaggact  
3181 agtaattatc attgactagc ccatctcaat tggatagtg attaaaatca cctagaccaa  
3241 ttgatagta tgtctgaatt agttgtttc aaagcaaatg aactagcgat tagtcgctat  
3301 gacttaacgg agcatgaaac caagctaatt ttatgctgtg tggcactact caaccccag  
3361 attgaaaacc ctacaaggaa agaacggagc gtatcgttca cttataacca atacgctcag  
3421 atgatgaaca tcagtaggga aatgcttat ggtgtattag ctaaagcaac cagagagctg  
3481 atgacgagaa ctgtggaaat caggaatcct ttggttaaag gctttgagat tttcagtg  
3541 acaaatatg ccaagttctc aagcgaaaaa ttagaattag ttttagtga agagatattg  
3601 ccttatcttt tccagtaaa aaaattcata aatataatc tggaaatgt taagtcttt  
3661 gaaaacaaat actctatgag gatttatgag tggttattaa aagaactaac acaaaagaaa  
3721 actcacaagg caaatataga gattagcctt gatgaattta agttcatgtt aatgcttga  
3781 aataactacc atgagtttaa aaggctaac caatgggttt tgaaccaat aagtaaagat  
3841 taaacactt acagcaatat gaaattggtg gttgataagc gaggccgcc gactgatacg  
3901 ttgattttcc aagttgaact agatagacaa atggatctcg taaccgaact tgagaacaac  
3961 cagataaaaa tgaatggtga caaaatacca acaaccatta catcagattc ctacctacgt  
4021 aacggactaa gaaaacact acacgatgct ttaactgcaa aaattcagct caccagttt

4081 gaggcaaaat ttttgagtga catgcaaagt aagcatgata tcaatggtc gttctcatgg  
 4141 ctacgcaaa aacaacgaac cacactagag aacatactgg ctaatacgg aaggatctga  
 4201 ggttcttat gctcttgat ctatcagtga agcatcaaga ctaacaaaca aaagtagaac  
 4261 aactgttac cgtagatat caaagggaaa actgtccata tgcacagatg aaaacgggtg  
 4321 aaaaaagata gatacatcag agcttttacg agtttttggg gcatttaaag ctgttcacca  
 4381 tgaacagatc gacaatgtaa cagatgaaca gcatgtaaca cctaatagaa caggtgaaac  
 4441 cagtaaaaca aagcaactag aacatgaaat tgaacacctg agacaactg ttacagctca  
 4501 acagtacac atagacagcc tgaacaggc gatgctgctt atcgaatcaa agctgccgac  
 4561 aacacgggag ccagtgacgc ctcccgtggg gaaaaaatca tggcaattct ggaagaaata  
 4621 gcgctttcag cggcgaacc tgaagccgga tctgcgattc tgataacaaa ctagcaacac  
 4681 cagaacagcg cggccgcaa aggaaaagat cggcgaaca aaccaccgtt ggtagcgggtg  
 4741 gttttttgt ttggatcagc aatcttcgta agcgtcatca ataagcgtaa aaaaaccggg  
 4801 caatgcccg tttttaatg agaaatttta cctgtcgtag ccgccaccat cggcgaaga  
 4861 agcatacaag gcttttgct tatagctacg tagcgcattg cgtcgcagca caatcccggc  
 4921 accgatcaag tcttcgcat gattattaga attcatgaga tct

//

**Zif268 matched act/rep on SC101/Kan**

LOCUS 12-33 6216 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(3790..5942)  
     /gene="Ori SC101"  
 CDS 2822..3616  
     /gene="KanR"  
     /codon\_start="0"  
 terminator complement(6007..6180)  
     /gene="Terminator"  
 terminator complement(3637..3752)  
     /gene="Terminator"  
 terminator 5961..6006  
     /gene="Terminator"  
 CDS 159..878  
     /gene="GFPmut3b"  
     /codon\_start="0"  
 terminator 1..46  
     /gene="term BBa\_B0011"  
 prot\_bind 138..152  
     /gene="RBS"  
 promoter 937..990  
     /gene="pLtetO-1"  
 CDS 1018..1761  
     /gene="RNAP alpha subunit"  
     /codon\_start="0"  
 CDS 1768..1785

```

        /gene="linker"
        /codon_start="0"
promoter    70..131
        /gene="pwklac"
prot_bind   53..63
        /gene="Zif268 binding site"
CDS         1792..2076
        /gene="Zif268"
        /codon_start="0"
BASE COUNT 1818 a 1376 c 1486 g 1536 t
ORIGIN
    1 agagaatata aaaagccaga ttattaatcc ggcttttita ttatttggat cttgcgtggg
    61 cggggatctc cgtattcttt acactttatg cttccggctc gtatgttggtg tcgaccgagc
   121 ggataacaat tggatctatt aaagaggaga aaggatctat gcgtaaagga gaagaacttt
   181 tcaactggagt tgtcccaatt cttgttgaat tagatgggtga tgtaaatggg cacaaatttt
   241 ctgtcagtgg agagggtgaa ggtgatgcaa catacggaaa acttaccctt aaatattttt
   301 gcactactgg aaaactacct gttccatggc caacacttgt cactactttc ggttatgggtg
   361 ttcaatgctt tgcgagatac ccagatcata tgaacagca tgactttttc aagagtgccca
   421 tgcccgaagg ttatgtacag gaaagaacta tatttttcaa agatgacggg aactacaaga
   481 cacgtgctga agtcaagttt gaaggtgata cccttgtaa tagaatcgag ttaaaaggta
   541 ttgattttaa agaagatgga aacattcttg gacacaaatt ggaatacaac tataactcac
   601 acaatgtata catcatggca gacaacaaa agaatggaat caaagttaac tcaaaaatta
   661 gacacaacat tgaagatgga agcgttcaac tagcagacca ttatcaaca aatactcaa
   721 ttggcgatgg ccctgtcctt ttaccagaca accattacct gtccacaca tctgccttt
   781 cgaaagatcc caacgaaaag agagaccaca tggccttct tgagtttgta accgctgctg
   841 ggattacaca tggcatggat gaactataca aataataagg atctagagaa tataaaaagc
   901 cagattatta atccggcttt ttattattt ggatcttccc tatcagtgat agagattgac
   961 atccctatca gtgatagaga tactgagcac ggatctatta aagaggagaa aggatctatg
  1021 cagggttctg tgacagagt tctaaaaccg cgctgggtg atatcgagca agtgagttcg
  1081 acgcacgccca aggtgaccct tgagccttta gagcgtggct ttggccatac tctgggtaac
  1141 gcactgcgcc gtattctgct ctcatcgatg ccgggttgcg cggtgaccga ggttgagatt
  1201 gatggtgtac tacatgagta cagcaccaaa gaaggcgttc aggaagatat cctggaaatc
  1261 ctgctcaacc tgaagggtt ggctgtgaga gttcagggca aagatgaagt tattcttacc
  1321 ttgaataaat ctggcattgg ccctgtgact gcagccgata tcaccacga cggatgatgc
  1381 gaaatcgta agccgcagca cgtgatctgc cacctgaccg atgagaacgc gtctattagc
  1441 atgcgtatca aagttcagcg cggctgtggt tatgtgccgg ctctaccgg aattcattcg
  1501 gaagaagatg agcgccaat cggccgtctg ctggtcgacg catgctacag ccctgtggag
  1561 cgtattgcct acaatgttga agcagcgcgt gtagaacagc gtaccgacct ggacaagctg
  1621 gtcatcgaaa tggaaaccaa cggcacaatc gatcctgaag aggcgattcg tcgtcgggca
  1681 accattctgg ctgaacaact ggaagcttcc gttgacttac gtgatgtacg tcagcctgaa
  1741 gtgaaagaag agaaaccaga gggatctgcc ccgcgagtcc ggaccggatc taagacaccc
  1801 cccatgaac gcccatatgc ttgccctgtc gactcctgcg atcggcgtt ttctcgtctg
  1861 gatgagctta cccgcatat ccgcatccat accggtcaga agcccttcca gtgtcgaatc
  1921 tgcatgcgta acttcagtgc tagtgaccac cttaccacc acatccgcac ccacacaggc
  1981 gagaagcctt ttgcctgtga catttgggg aggaagttg ccaggagtga tgaacgcaag
  2041 aggcatacca aatccattt aagacagaag gacgtctagg gatctagaga atataaaaaa

```

2101 ccagattatt aatccggctt ttttattatt tggatcctaa ctcgagtga ggtcgtctcg  
2161 gatcgagaag gacacggta atactaggcc tgctggctgg taatcgccag caggcctttt  
2221 ttttggggg agaggaagt catgaaaaa ctaaccttg aaattcgatc tccaccacat  
2281 cagctctgaa gcaacgtaaa aaaacccgcc cggcgggtt ttttataacc cgtagtatcc  
2341 ccacttatct acaatagctg tccttaatta aggtgaata aataaaaaca gccgttgcca  
2401 gaaagaggca cggctgttt tttttctag tgagaccggg agcagttaa cgcagaaagg  
2461 cccaccgaa ggtgagccag tgtgactcta gtagagagcg ttcaccgaca aacaacagat  
2521 aaaacgaaag gccagctt tcgactgagc ctttcgtttt attgatgcc tggttattat  
2581 tttttgaca gtcacatcat gccaccggtc aacaggagtc caagcgaccg gtggttgc  
2641 gtctagctag ctagaacagg actagctaat ggtttcttag acgtcggaat tgccagctgg  
2701 ggcgcctct ggtaaggtg ggaagcctg caaagtaaac tggatggctt tcttgccgcc  
2761 aaggatctga tggcgcagg gatcaagatc tgatcaagag acaggatgag gatcgtttc  
2821 catgattgaa caagatgat tgacgcagg ttctccggcc gcttgggtgg agaggctatt  
2881 cggctatgac tgggcacaac agacaatcgg ctgctctgat gccccgtgt tccggctgc  
2941 agcgcagggg cgcccgttc ttttgtcaa gaccgacctg tccggtgcc tgaatgaact  
3001 gcaggacgag gcagcgcggc tctcgtggct ggccacgacg ggcgttcctt gcgcagctgt  
3061 gctcgcgctt gtcactgaag cgggaaggga ctggctgcta tgggcgaag tgccggggca  
3121 ggatctctg tcatctacc ttgctctgc cgagaaagta tccatcatgg ctgatgcaat  
3181 gcggcggctg catacgttg atccggctac ctgcccattc gaccaccaag cgaacatcg  
3241 catcgagcga gcacgtactc ggtggaagc cggcttctg gatcaggatg atctggacga  
3301 agagcatcag gggctcgcgc cagccgaact gttcggcagg ctcaaggcgc gcatgcccga  
3361 cggcgaggat ctgctcgtga cccatggcga tgcctgcttg ccgaatatca tgggtgaaaa  
3421 tggccgctt tctggattca tcgactgtgg cggctgggt gtggcggacc gctatcagga  
3481 catagcgtg gctaccctg atattgctga agagcttggc ggcaatggg ctgaccgctt  
3541 cctcgtgctt tacggtatcg ccgctcccga ttcgacgcgc atgccttct atgccttct  
3601 tgacgagttc ttctgagcgg gactctgggg ttcgagagct cgcttgact cctgttgata  
3661 gatccagtaa tgacctaga actccatctg gatttgtca gaacgctcgg ttgccgccg  
3721 gcgttttta tgggtgagaa tccaagcact agggacagta agacgggtaa gctgttgat  
3781 gataccctg ccttactggg tgcattagcc agtctgaatg acctgtcacg ggataatccg  
3841 aagtgtcag actggaaaat cagagggcag gaactgctga acagcaaaaa gtcagatagc  
3901 accacatagc agaccgcca taaaacgcc tgagaagccc gtgacgggct tttctgtat  
3961 tatggtagt ttcttgcac gaatcataa aaggcgcctg tagtgccatt tacccttatt  
4021 cactgccaga gccgtgagcg cagcgaactg aatgtcacga aaaagacagc gactcaggtg  
4081 cctgatggtc ggagacaaaa ggaatattca gcgattgcc cgagcttgcg aggggtgctac  
4141 ttaagcctt agggtttaaa ggtctgttt gtagaggagc aaacagcgtt tgcgacatcc  
4201 ttttgaata ctgcggaact gactaaagta gtgagttata cacagggtg ggatctattc  
4261 ttttatctt ttttattct ttcttattc tataaattat aaccactga atataaaca  
4321 aaaaaacaca caaaggtcta gcggaatta cagaggtct agcagaattt acaagtttc  
4381 cagcaaaggc ctacgagaat ttacagatac ccacaactca aaggaaaagg actagtaatt  
4441 atcattgact agccatctc aattgtata gtgattaaaa tcacctagac caattgagat  
4501 gtatgtctga attagttgt ttcaaagcaa atgaactagc gattagtcgc tatgacttaa  
4561 cggagcatga aaccaagcta atttatgct gtgtggcact actcaacccc acgattgaaa  
4621 accctacaag gaaagaacgg acggtatcgt tcaattataa ccaatacgtc cagatgatga  
4681 acatcagtag ggaaaatgct tatgggtgat tagctaaagc aaccagagag ctgatgacga  
4741 gaactgtgga aatcaggaat ccttggtta aaggcttga gatttccag tggacaaact  
4801 atgccaagtt ctcaagcga aaattagaat tagtttttag tgaagagata ttgccttatc

4861 tttccagtt aaaaaaattc ataaaatata atctggaaca tgtaagtct tttgaaaaca  
 4921 aatactctat gaggattat gagtggttat taaaagaact aacacaaaag aaaactcaca  
 4981 aggcaaatat agagattagc cttgatgaat ttaagttcat gttaatgctt gaaaataact  
 5041 accatgagtt taaaaggctt aaccaatggg ttttgaacc aataagtaaa gatttaaaca  
 5101 cttacagcaa tatgaaattg gtggttgata agcgaggccg cccgactgat acgttgattt  
 5161 tccaagttga actagataga caaatggatc tcgtaaccga acttgagaac aaccagataa  
 5221 aatgaatgg tgacaaaata ccaacaacca ttacatcaga ttctaccta cgtaacggac  
 5281 taagaaaaac actacacgat gctttaactg caaaaattca gtcaccagt tttgaggcaa  
 5341 aatttttgag tgacatgcaa agtaagcatg atctcaatgg ttcgttctca tggctcacgc  
 5401 aaaaacaacg aaccacacta gagaacatac tggctaaata cggaaggatc tgaggttctt  
 5461 atggctcttg tatctatcag tgaagcatca agactaaca acaaaagtag aacaactgtt  
 5521 caccgttaga tatcaaaggg aaaactgtcc atatgcacag atgaaaacgg tgtaaaaaag  
 5581 atagatacat cagagctttt acgagttttt ggtgcattta aagctgttca ccatgaacag  
 5641 atcgacaatg taacagatga acagcatgta acacctaata gaacaggtga aaccagtaaa  
 5701 acaaagcaac tagaacatga aattgaacac ctgagacaac ttgttacagc tcaacagtca  
 5761 cacatagaca gctgaaaca ggcatgctg cttatcgaat caaagctgcc gacaacacgg  
 5821 gagccagtga cgctcccgt ggggaaaaaa tcattggcaat tctggaagaa atagcgtttt  
 5881 cagccggcaa acctgaagcc ggatctgcga ttctgataac aaactagcaa caccagaaca  
 5941 gcgcggccgc aaaaggaaaa gatccggcaa acaaccacc gttggtagcg gtggtttttt  
 6001 tgtttggatc gacaatcttc gtaagcgtca tcaataagcg taaaaaaacc gggcaatgcc  
 6061 cggttttttg atgagaaatt ttactgtcg tagccgccac catccggcaa agaagcatac  
 6121 aaggcttttg gcttatagct acgtagcgc ttcgctcgc gcacaatccc ggcaccgatc  
 6181 aagtcttcgc gatgattatt agaattcatg agatct

//

### **ZFP03 inverter on SC101/Kan**

LOCUS 12-34 7125 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(4699..6851)  
     /gene="Ori SC101"  
 CDS 3731..4525  
     /gene="KanR"  
     /codon\_start="0"  
 CDS 1005..1748  
     /gene="RNAP alpha subunit"  
     /codon\_start="0"  
 CDS 114..833  
     /gene="GFPmut3b"  
     /codon\_start="0"  
 CDS 1779..2312  
     /gene="ZFP03"  
     /codon\_start="0"  
 CDS 2455..2988  
     /gene="ZFP03"

```

        /codon_start="0"
terminator complement(6916..7089)
        /gene="Terminator"
terminator complement(4546..4661)
        /gene="Terminator"
gene      25..86
        /gene="pwklac"
gene      916..977
        /gene="pwklac"
promoter  2371..2424
        /gene="pLtetO-1"
terminator 6870..6915
        /gene="Terminator"
terminator 2319..2364
        /gene="term BBa_B0011"
terminator 840..885
        /gene="term BBa_B0011"
terminator 2995..3040
        /gene="term BBa_B0011"
prot_bind 892..909
        /gene="ZFP03 binding site"
gene      1755..1772
        /gene="linker"
gene      91..107
        /gene="RBS"
gene      984..998
        /gene="RBS"
gene      2431..2445
        /gene="RBS"
BASE COUNT 2081 a 1611 c 1626 g 1807 t
ORIGIN
1  ggtagacgtc tagtaactgg atctccgtat tctttacact ttatgcttcc ggctcgtatg
61  ttgtgtcgac cgagcggata acaattggat ctattaaaga ggagaaagga tctatgcgta
121 aaggagaaga acttttctact ggagttgtcc caattcttgt tgaattagat ggtgatgta
181 atgggcacaa attttctgtc agtggagagg gtgaaggtga tgcaacatac ggaaaactta
241 cccttaaatt tatttgact actggaaaac tacctgttcc atggccaaca cttgtcacta
301 ctttcggtta tgggtttcaa tgctttgcga gataccaga tcatatgaaa cagcatgact
361 ttttcaagag tgccatgccc gaaggttatg tacaggaaag aactatattt tcaaagatg
421 acgggaacta caagacacgt gctgaagtca agtttgaagg tgataccctt gtaatagaa
481 tcgagttaa aggtattgat ttaaagaag atggaacat tcttggacac aaattggaat
541 acaactataa ctacacaaat gtatacatca tggcagacaa acaaaagaat ggaatcaaag
601 ttaacttcaa aattagacac aacattgaag atggaagcgt tcaactagca gaccattatc
661 acaaaaatac tccaattggc gatggccctg tcctttacc agacaacat tacctgtcca
721 cacaatctgc ccttgcgaaa gatcccacg aaaagagaga ccacatggtc cttcttgagt
781 ttgtaaccgc tgctgggatt acacatggca tggatgaact atacaataa taaggatcta
841 gagaatataa aaagccagat tattaatccg gctttttat tatttggatc tggtagacgt

```



901 ctagtaactg gatctccgta ttctttacac tttatgcttc cggctcgtat gttgtgctga  
961 ccgagcggat aacaattgga tctattaaag aggagaaagg atctatgcag ggttctgtga  
1021 cagagtttct aaaaccgcgc ctggttgata tcgagcaagt gagttcgacg cacgccaagg  
1081 tgaccttga gccttagag cgtggctttg gccatactct gggtaacgca ctgcgccgta  
1141 ttctgctctc atcgatgccg ggttgcgcgg tgaccgaggt tgagattgat ggtgtactac  
1201 atgagtacag caccaaagaa ggcgttcagg aagatctct ggaaatctcg ctcaacctga  
1261 aagggctggc ggtgagagtt cagggcaaag atgaagttat tcttaccttg aataaatctg  
1321 gcattggccc tgtgactgca gccgatatca ccacgacgg tgatgtcgaa atcgtcaagc  
1381 cgcagcacgt gatctgccac ctgaccgatg agaacgcgtc tattagcatg cgtatcaaag  
1441 ttcagcgcgg tctgtggtat gtgccggctt ctaccgaaat tcattcggaa gaagatgagc  
1501 gcccaatcgg ccgtctgctg gtcgacgcat gctacagccc tgtggagcgt attgcctaca  
1561 atgttgaagc agcgcgtgta gaacagcgta ccgacctgga caagctggtc atcgaaatgg  
1621 aaaccaacgg cacaatcgat cctgaagagg cgattcgtcg tgcggcaacc attctggctg  
1681 aacaactgga agcttctgtt gacttacgtg atgtacgtca gcctgaagtg aaagaagaga  
1741 aaccagaggg atctgccccg cgagtccgga ccggatctct ggaaccagga tctaaacctg  
1801 acaaatgtcc ggaatgtggt aaatccttct cactcatct ggatctgatt cgtcatcaac  
1861 gtactcacac tggatctaaa ccgtacaaat gtccggaatg tggtaaattc ttctccaat  
1921 cttcttctct ggttcgtcat caactactc acactggatc taaacctac aaatgtccgg  
1981 aatgtggtaa atccttctcc caaaactcta ccctgactga acatcaactg actcacactg  
2041 gatctaaacc gtacaaatgt ccggaatgtg gtaaactctt ctctctcgt cgtactgtc  
2101 gtgcacatca acgtactcac actggatcta aaccgtacaa atgtccggaa tgtggtaa  
2161 cttctccca actggctcat ctgcgtcac atcaactgac tcacactgga tctaaacctg  
2221 acaaatgtcc ggaatgtggt aaatccttct cacttctgg tcacttggtg cgtcatcaac  
2281 gtactcacac tggatctaaa acctcttaat aaggatctag agaataataa aagccagatt  
2341 attaatccgg ctttttatt atttggatct tcctatcag tgatagagat tgacatcct  
2401 atcagtgata gagatactga gcacggatct attaaagagg agaaaggatc tatgtcggaa  
2461 ccaggatcta aaccgtacaa atgtccggaa tgtggtaaat cttctccac tcactggat  
2521 ctgattcgtc atcaactgac tcacactgga tctaaacctg acaaatgtcc ggaatgtggt  
2581 aaatccttct ccaatcttc ttctctggtt cgtcatcaac gtactcacac tggatctaaa  
2641 ccgtacaaat gtccggaatg tggtaaattc ttctccaaa actctaccct gactgaacat  
2701 caactgactc acactggatc taaacctgac aaatgtccgg aatgtggtaa atccttctcc  
2761 tctcgtcgta cttgtcgtgc acatcaactg actcacactg gatctaaacc gtacaaatgt  
2821 ccggaatgtg gtaaactctt ctcccaactg gctcatctgc gtgcacatca acgtactcac  
2881 actggatcta aaccgtacaa atgtccggaa tgtggtaaat cttctccac ttctggatc  
2941 ctggtacgtc atcaactgac tcacactgga tctaaaacct ctaataagg atctagagaa  
3001 tataaaaagc cagattatta atccggcttt ttattattt ggatcctaac tcgagtgcag  
3061 gtcgtctcgg atcgagaagg acacggtaa tactaggcct gctggctggt aatcgccagc  
3121 aggcctttt atttggggga gagggaagtc atgaaaaaac taaccttga aattgatct  
3181 ccaccacatc agctctgaag caactgaaaa aaaccgccc cggcggggtt tttataccc  
3241 gtatgatacc cacttatcta caatagctgt ccttaattaa ggtgaataa ataaaaacg  
3301 ccgttgccag aaagaggcac ggctgtttt attttctagt gagaccggga gcagttaaac  
3361 gcagaaaggc ccaccgaag gtgagccagt gtgactctag tagagagcgt tcaccgacaa  
3421 acaacagata aaacgaaagg cccagtctt cactgagcc tttcgttta ttgatgcct  
3481 ggttattatt atttgtacag ctcatccatg ccaccgtca acaggagtcc aagcgaccgg  
3541 tggttgcatg tctagctagc tagaacagga ctagctaatg gtttctaga cgtcggaa  
3601 gccagctggg gcgccctctg gtaaggttgg gaagccctgc aaagtaaact ggatggctt

3661 cttgcccca aggatctgat ggcgagggg atcaagatct gatcaagaga caggatgagg  
3721 atcgtttcgc atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga  
3781 gaggctattc ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt  
3841 ccggctgtca ggcgaggggc gcccggttct tttgtcaag accgacctgt ccggtgcctt  
3901 gaatgaactg caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg  
3961 cgcagctgtg ctcgacgttg tcaactgaag gggaaggac tggctgctat tgggcgaagt  
4021 gccggggcag gatctctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc  
4081 tgatgcaatg cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc  
4141 gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttctcg atcaggatga  
4201 tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg  
4261 catgccgac ggcgaggatc tcgtcgtgac ccatggcgat gctgcttgc cgaatatcat  
4321 ggtggaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg  
4381 ctatcaggac atagcgttgg ctaccctga tattgctgaa gagcttggcg gcgaatgggc  
4441 tgaccgctc ctcgtgctt acggtatcgc cgctcccgat tcgacgcga tcgccttcta  
4501 tcgcttctt gacgattct tctgagcggg actctggggg tcgagagctc gcttggactc  
4561 ctgtgatag atccagtaat gacctcagaa ctccatctgg atttgttcag aacgctcgg  
4621 tgccgccggg cgtttttat tgggtgagaat ccaagcacta gggacagtaa gacgggtaag  
4681 cctgttgatg ataccgctgc ctactgggt gcattagcca gtctgaatga cctgtcacgg  
4741 gataatccga agtggtcaga ctggaaaatc agagggcagg aactgctgaa cagcaaaaag  
4801 tcagatagca ccacatagca gaccgccat aaaacgccct gagaagcccg tgacgggctt  
4861 ttctgtatt atgggtagt tcttgcagc aatccataaa aggcgctgt agtgccattt  
4921 accccattc actgccagag ccgtgagcgc agcgaactga atgtcacgaa aaagacagcg  
4981 actcaggtgc ctgatggctg gagacaaaag gaatattcag cgatttccc gagcttgcga  
5041 ggggtgctact taagccttta gggtttaag gtctgtttg tagaggagca aacagcgttt  
5101 gcgacatcct tttgtaatac tgcggaactg actaaagtag tgagttatac acagggctgg  
5161 gatctattc tttatcttt tttattctt tctttattc ataaattata accacttgaa  
5221 tataaacaac aaaaacacac aaaggtctag cggaattac agaggggtcta gcagaattta  
5281 caagtttcc agcaaaggtc tagcagaatt tacagatacc cacaactcaa aggaaaagga  
5341 ctagtaatta tcattgacta gccatctca atgggtatag tgattaaaat cacctagacc  
5401 aattgagatg tatgtctgaa ttagttgtt tcaaagcaaa tgaactagcg attagtcgct  
5461 atgacttaac ggagcatgaa accaagctaa tttatgctg tgtggcacta ctcaaccca  
5521 cgattgaaaa cctacaagg aaagaacgga cggatcgtt cactataac caatagctc  
5581 agatgatgaa catcagtagg gaaaatgctt atgggtgatt agctaaagca accagagagc  
5641 tgatgacgag aactgtggaa atcaggaatc ctttggtaa aggctttag atttccagt  
5701 ggacaaacta tgccaagttc tcaagcgaac aattagaatt agtttttagt gaagagatat  
5761 tgccttatct tttccagta aaaaaattca taaaatataa tctggaacat gttaagtctt  
5821 ttgaaaacaa atactctatg aggatttatg agtggttatt aaaagaacta acacaaaaga  
5881 aaactacaaa ggcaaatata gagattagcc ttgatgaatt taagttcatg ttaatgctt  
5941 aaaataacta ccatgagttt aaaaggctta accaatgggt tttgaaacca ataagtaaag  
6001 attaaacac ttacagcaat atgaaattgg tggttgataa gcgaggccgc ccgactgata  
6061 cgttgatttt ccaagttgaa ctagatagac aatggatct cgtaaccgaa cttgagaaca  
6121 accagataaa aatgaatggg gacaaaatac caacaacat tacatcagat tctacctac  
6181 gtaacggact aagaaaaaca ctacacgatg cttaactgc aaaaattcag ctaccagtt  
6241 ttgaggcaaa atttttgagt gacatgcaaa gtaagcatga tctcaatggg tcgttctcat  
6301 ggctcacgca aaaacaacga accacactag agaacatact ggctaaatac ggaaggatct  
6361 gaggttctta tggctctgt atctatcagt gaagcatcaa gactaacaac caaaagtaga

6421 acaactgttc accgttagat atcaaagggg aaactgtcca tatgcacaga tgaaaacggt  
 6481 gtaaaaaaga tagatacatc agagctttta cgagtttttg gtgcatttaa agctgttcac  
 6541 catgaacaga tcgacaatgt aacagatgaa cagcatgtaa cacctaatag aacaggtgaa  
 6601 accagtaaaa caaagcaact agaacatgaa attgaacacc tgagacaact tgttacagct  
 6661 caacagtcac acatagacag cctgaaacag gcgatgctgc ttatcgaatc aaagctgccg  
 6721 acaacacggg agccagtgac gcctcccgtg gggaaaaaat catggcaatt ctggaagaaa  
 6781 tagcgctttc agccggcaaa cctgaagccg gatctgcat tctgataaca aactagcaac  
 6841 accagaacag cgcggccgca aaaggaaaag atccggcaaa caaaccaccg ttgtagcgg  
 6901 tggttttttt gtttgatcg acaatcttcg taagcgtcat caataagcgt aaaaaaacgg  
 6961 ggcaatgccc ggttttttaa tgagaaattt tacctgctgt agccgccacc atccggcaaa  
 7021 gaagcataca aggcttttgg cttatagcta cgtagcgcg tgcgtcgcag cacaatcccg  
 7081 gcaccgatca agtcttcgcg atgattatta gaattcatga gatct

//

### **ZFP5475 leak on SC101/Kan**

LOCUS 12-9 5026 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(2600..4752)  
     /gene="Ori SC101"  
 CDS 1632..2426  
     /gene="KanR"  
     /codon\_start="0"  
 CDS 170..889  
     /gene="GFPmut3b"  
     /codon\_start="0"  
 terminator complement(4817..4990)  
     /gene="Terminator"  
 terminator complement(2447..2562)  
     /gene="Terminator"  
 promoter 81..142  
     /gene="pwklac"  
 terminator 4771..4816  
     /gene="Terminator"  
 terminator 1..46  
     /gene="term BBa\_B0011"  
 terminator 896..941  
     /gene="term BBa\_B0011"  
 prot\_bind 53..74  
     /gene="ZFP5475 operator"  
 prot\_bind 149..163  
     /gene="RBS"  
 BASE COUNT 1512 a 1081 c 1171 g 1262 t  
 ORIGIN  
     1 agagaatata aaaagccaga ttattaatcc ggctttttta ttatttggat ctacaccggg

61 gttcccctcg ggtgggatct ccgtattctt tacactttat gcttccggct cgtatgttg  
121 gtcgaccgag cggataacaa ttggatctat taaagaggag aaaggatcta tgcgtaaagg  
181 agaagaactt ttcactggag ttgtcccaat tcttgttgaa ttagatggtg atgttaatgg  
241 gcacaaattt tctgtcagtg gagagggtga aggtgatgca acatacggaa aacttacct  
301 taaatttatt tgcactactg gaaaactacc tgttccatgg ccaacacttg tcactacttt  
361 cggttatggg gttcaatgct ttgcgagata cccagatcat atgaaacagc atgacttttt  
421 caagagtgcc atgccgaag gttatgtaca ggaaagaact atattttca aagatgacgg  
481 gaactacaag acacgtgctg aagtcaagtt tgaaggatg acccttgta atagaatcga  
541 gttaaaagg attgatttta aagaagatgg aacattctt ggacacaaat tggaaataca  
601 ctataactca cacaatgtat acatcatggc agacaaacaa aagaatggaa tcaaaagtaa  
661 cttcaaaatt agacacaaca ttgaagatgg aagcgttcaa ctgacagacc attatcaaca  
721 aaatactcca attgcgatg gccctgtcct tttaccagac aaccattacc tgtccacaca  
781 atctgcctt tcgaaagatc ccaacgaaaa gagagaccac atggctctt ttgagttgt  
841 aaccgctgct gggattacac atggcatgga tgaactatac aaataataag gatctagaga  
901 atataaaaag ccagattatt aatccggctt ttttattatt tggatcctaa ctcgagtga  
961 ggtcgtctcg gatcgagaag gacacggta atactaggcc tgctggctgg taatgccag  
1021 caggcctttt tatttggggg agagggaagt catgaaaaaa ctaaccttg aaattcgatc  
1081 tccaccacat cagctctgaa gcaacgtaaa aaaaccgcc ccggcgggtt ttttatacc  
1141 cgtagtatcc ccacttatct acaatagctg tccttaatta aggtgaata aataaaaaca  
1201 gccgttgcca gaaagaggca cggctgtttt tttttctag tgagaccggg agcagttaa  
1261 cgcagaaagg cccaccgaa ggtgagccag tgtgactcta gtagagagcg tccaccgaca  
1321 aacaacagat aaaacgaaag gccagctt tcgactgagc cttcgtttt atttgatgcc  
1381 tggttattat tattgtaca gctcatccat gccaccggtc aacaggagtc caagcgaccg  
1441 gtggttgc atgtagctag ctagaacagg actagctaat ggtttcttag acgtcgaat  
1501 tgccagctgg ggcgcctct gtaagggtg ggaagccctg caaagtaaac tggatggctt  
1561 tcttgccgc aaggatctga tggcgaggg gatcaagatc tgatcaagag acaggatgag  
1621 gatcgttcg catgattgaa caagatggat tgcacgcagg tctccggcc gcttgggtg  
1681 agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat gccgcctgt  
1741 tccgctgtc agcgcagggg cggcgggtc tttttgcaa gaccgacctg tccggtgcc  
1801 tgaatgaact gcaggacgag gcagcgggc tatcgtggct ggccacgagc ggcgttctt  
1861 gcgcagctgt gctcagctt gtcactgaag cgggaaggga ctggctgcta ttggcggaag  
1921 tgccggggca ggatctctg tcactcacc ttgctctgc cgagaaagta tccatcatgg  
1981 ctgatgcaat gcggcggtg catacgttg atccggctac ctgcccattc gaccaccaag  
2041 cgaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgct gatcaggatg  
2101 atctggacga agagcatcag gggctcgcg cagccgaact gttcgccagg ctcaaggcgc  
2161 gcatgcccga cggcgaggat ctgctctgga cccatggcga tgcctgcttg ccgaatatca  
2221 tgggtgaaaa tggccgctt tctggattca tcgactgtgg cggctgggt tggcgacc  
2281 gctatcagga catagcgtt gctaccgtg atattgctga agagcttggc ggcaatggg  
2341 ctgaccgtt cctcgtgct tacggtatcg ccgtcccga ttcgacgcg atcgccttct  
2401 atcgccttct tgacgagttc ttctgagcgg gactctgggg ttcgagagct cgcttgact  
2461 cctgttgata gatccagtaa tgacctcaga actccatctg gatttgtca gaacgctcg  
2521 ttccgcccgg gcgttttta ttggtgagaa tcaagcact agggacagta agacgggtaa  
2581 gcctgtgat gataccgctg cttactggg tgcattagcc agtctgaatg acctgtcacg  
2641 ggataatccg aagtgtcag actggaaaat cagagggcag gaactgctga acagcaaaaa  
2701 gtcagatagc accacatagc agaccgcca taaaacgccc tgagaagccc gtgacggct  
2761 tttctgtat tatggtagt ttcttgc ataatccataa aaggcgctg tagtgccatt

2821 taccgccatt cactgccaga gccgtgagcg cagcgaactg aatgtcacga aaaagacagc  
2881 gactcaggtg cctgatggtc ggagacaaaa ggaatattca gcgattgcc cgagcttgcc  
2941 aggggtgctac ttaagccttt agggttttaa ggtctgtttt gtagaggagc aaacagcgtt  
3001 tgcgacatcc ttttgtaata ctgcggaact gactaaagta gtgagttata cacagggctg  
3061 ggatctattc ttttatctt tttttattct ttctttattc tataaattat aaccacttga  
3121 atataaacia aaaaaacaca caaaggctca gcggaattta cagagggctc agcagaattt  
3181 acaagtttcc cagcaaaggt ctgacagaat ttacagatac ccacaactca aaggaaaagg  
3241 actagtaatt atcattgact agcccatctc aattggata gtgattaaaa tcacctagac  
3301 caattgagat gtatgtctga attagtgtt ttcaaagcaa atgaactagc gattagtccg  
3361 tatgacttaa cggagcatga aaccaagcta atttatgct gtgtggcact actcaacccc  
3421 acgattgaaa accctacaag gaaagaacgg acggtatcgt tcactataa ccaatacgtc  
3481 cagatgatga acatcagtag ggaaaatgct tatgggtgat tagctaaagc aaccagagag  
3541 ctgatgacga gaactgtgga aatcaggaat cttttggta aaggcttga gattttccag  
3601 tggacaaact atgccaagtt ctcaagcgaa aaattagaat tagtttttag tgaagagata  
3661 ttgccttacc tttccagtt aaaaaattc ataaaatata atctggaaca tgtaagtct  
3721 ttgaaaaca aatactctat gaggattat gagtggttat taaaagaact aacacaaaag  
3781 aaaactcaca aggcaatat agagattagc ctgatgaat ttaagttcat gttaatgctt  
3841 gaaaataact accatgagtt taaaaggctt aaccaatggg tttgaaacc aataagtaa  
3901 gatttaaca cttacagcaa tatgaaattg gtggttgata agcgaggccg cccgactgat  
3961 acgttgattt tccaagttga actagataga caaatggatc tcgtaaccga acttgagaac  
4021 aaccagataa aatgaatgg tgacaaaata ccaacaacca ttacatcaga ttctaccta  
4081 cgtaacggac taagaaaaac actacacgat gctttaactg caaaaattca gctcaccagt  
4141 tttgaggcaa aatttttgag tgacatgcaa agtaagcatg atctcaatgg ttcgttctca  
4201 tggctcacgc aaaaacaacg aaccacacta gagaacatac tggctaaata cggaaggatc  
4261 tgaggttctt atggctcttg tatctatcag tgaagcatca agactaaca acaaaagtag  
4321 aacaactgtt caccgttaga tatcaaaggg aaaactgtcc atatgcacag atgaaaacgg  
4381 tgtaaaaaag atagatacat cagagctttt acgagttttt ggtgcattta aagctgttca  
4441 ccatgaacag atcgacaatg taacagatga acagcatgta acacctaata gaacaggtga  
4501 aaccagtaaa acaagcaac tagaacatga aattgaacac ctgagacaac ttgttacagc  
4561 tcaacagtca cacatagaca gcctgaaaca ggcgatgctg cttatcgaat caaagctgcc  
4621 gacaacacgg gagccagtga cgctcccgt ggggaaaaaa tcatggcaat tctggaagaa  
4681 atagcgcttt cagccggcaa acctgaagcc ggatctgcga ttctgataac aactagcaa  
4741 caccagaaca gcgcgccgc aaaaggaaaa gatccggcaa acaaaccacc gttgtagcg  
4801 gtggtttttt tgttggatc gacaatcttc gtaagcgtca tcaataagcg taaaaaac  
4861 gggcaatgcc cggttttta atgagaaatt ttacctgtcg tagccgccac catccggcaa  
4921 agaagcatac aaggcttttg gcttatagct acgtagcgca ttgcgtcgca gcacaatccc  
4981 ggcaccgatc aagtcttcgc gatgattatt agaattcatg agatct

//

### **ZFP03 activator p15A/Kan**

LOCUS 21-61 3569 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(1896..2607)  
/gene="p15A"

```

terminator 1761..1889
            /gene="double terminator"
terminator 2698..2803
            /gene="Terminator"
terminator 244..289
            /gene="term BBa_B0011"
prot_bind 356..370
            /gene="RBS"
promoter 296..349
            /gene="pLtetO-1"
CDS      1127..1144
            /gene="linker"
            /codon_start="0"
CDS      complement(2829..3569)
            /gene="KanR"
            /codon_start="0"
terminator 1761..1889
            /gene="double term"
CDS      377..1120
            /gene="RNAP alpha subunit"
            /codon_start="0"
CDS      1151..1684
            /gene="ZFP03"
            /codon_start="0"
terminator 1691..1736
            /gene="term BBa_B0011"
BASE COUNT 900 a 963 c 871 g 835 t
ORIGIN

```

```

1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa
61 cgatcctcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg
121 caagaaagcc atccagtta cttgacaggg cttccaacc ttaccagagg gcgccccagc
181 tggcaattcc gaattcatga gatctcacat ggcatggatg aactatacaa ataataagga
241 tctagagaat ataaaaagcc agattattaa tccggctttt ttattattg gatcttccct
301 atcagtgata gagattgaca tcctatcag tgatagagat actgagcacg gatctattaa
361 agaggagaaa ggatctatgc agggttctgt gacagagttt ctaaaaccgc gcctggttga
421 tatcgagcaa gtgagttcga cgcacgcaa ggtgaccctt gagcctttag agcgtggctt
481 tggccatact ctgggtaacg cactgcgccg tattctgctc tcatcgatgc cgggttgcgc
541 ggtgaccgag gttgagattg atggtgtact acatgagtac agcaccaaag aaggcgttca
601 ggaagatata ctgaaatcc tgctcaacct gaaagggctg gcggtgagag ttcagggcaa
661 agatgaagtt attcttacct tgaataaatc tggcattggc cctgtgactg cagccgatat
721 caccacgac ggtgatgtcg aaatcgtcaa gccgcagcac gtgatctgcc acctgaccga
781 tgagaacgcg tctattagca tgcgatcaa agttcagcgc ggtcgtggtt atgtgccggc
841 ttctaccgca attcattcgg aagaagatga gcgcccaatc ggccgtctgc tggtcgacgc
901 atgctacagc cctgtggagc gtattgcta caatgttgaa gcagcgcgtg tagaacagcg
961 taccgacctg gacaagctgg tcatcgaaat ggaaccaac ggcaaatcg atcctgaaga
1021 ggcgattcgt cgtgcggcaa ccattctggc tgaacaactg gaagcttctg ttgacttacg

```

1081 tgatgtacgt cagcctgaag tgaagaaga gaaaccagag ggatctgccc cgcgagtccg  
1141 gaccggatct ctggaaccag gatctaaacc gtacaaatgt ccggaatgtg gtaaactcct  
1201 ctccactcat ctggatctga ttcgtcatca acgtactcac actggatcta aaccgtaca  
1261 atgtccggaa tgtggtaaat ccttctcca atcttctct ctggttcgtc atcaactgac  
1321 tcactgga tctaaacctg acaaatgtcc ggaatgtggt aaatccttct cccaaaactc  
1381 taccctgact gaacatcac gtaactcac tggatctaaa ccgtacaaat gtccggaatg  
1441 tggtaaatcc ttctctctc gtcgtacttg tcgtgcacat caactactc aactggatc  
1501 taaaccgtac aatgtccgg aatgtggtaa atccttctcc caactggctc atctgcgtgc  
1561 acatcaactg actcactg gatctaaacc gtacaaatgt ccggaatgtg gtaaactcct  
1621 ctccacttct ggcatctgg tacgtcatca acgtactcac actggatcta aaactctta  
1681 ataaggatct agagaatata aaaagccaga ttattaatcc ggcttttta ttattggat  
1741 ctaactcga gtaaggatct ccaggcatca aataaacga aaggctcagt cgaaagactg  
1801 ggccttctg tttatctgtt gttgtcggg gaacgctctc tactagagtc aactggctc  
1861 acctcgggt ggccttct gcgtttata ctagggatat atccgctc ctcgctcact  
1921 gactcgctac gctcggctg tcgactcgg cgagcggaaa tggcttacga acggggcgga  
1981 gatttctgg aagatgccag gaagatactt aacagggag tgagagggcc gcggcaaagc  
2041 cgttttcca taggctccgc cccctgaca agcatcacga aatctgacgc tcaaatcagt  
2101 ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctggc ggctcctcg  
2161 tgcgctctc tgttctgcc tttcggttta ccggtgcat tccgctgta tggccgctg  
2221 tgtctcatt cacgctgac actcagttcc gggtaggcag ttcgctcaa gctggactg  
2281 atgcacgaac ccccgttca gtccgaccgc tgcgcttat ccggttaacta tcgtcttag  
2341 tcaaccgg aaagacatgc aaaagcacca ctggcagcag cactggtaa ttgattaga  
2401 ggagttagtc ttgaagcat gcgccggtta aggctaaact gaaaggacaa gtttggta  
2461 ctgcgctct ccaagccagt tacctcgtt caaagagttg gtagctcaga gaacctcga  
2521 aaaaccgcc tgcaaggcgg ttttctgtt tcagagcaa gagattacgc gcagacaaa  
2581 acgatctcaa gaagatcct ttattaatca gataaaat ttctagattt cagtcaatt  
2641 tatcttca aatgtagc ctgaagtcag cccatacga tataagttgt tactagtgt  
2701 tggatttca ccaataaaaa acgcccggcg gcaaccgagc gttctgaaca aatccagatg  
2761 gagttctgag gtcattactg gatctatcaa caggagtcca agcgagctc cgaacccag  
2821 agtcccgtc agaagaactc gtcaagaagg cgatagaagg cgatgcgctg cgaatcggga  
2881 gcggcgatac cgtaaagcac gaggaagcgg tcagccatt cgccgccaag ctctcagca  
2941 atatcaggg tagccaacgc tatgtctga tagcggctc ccacaccag ccggccacag  
3001 tcgatgaatc cagaaaagcg gccatttcc accatgatat tcggcaagca ggcacgcca  
3061 tgggtcacga cgatcctc gccgtcgggc atgcgctcct tgagcctggc gaacagttc  
3121 gctggcgca gccctgatg ctctctgccc agatcctc gatcgacaag accggcttc  
3181 atccgagtac gtgctgctc gatcgatgt ttcgcttggg ggtcgaatgg gcaggtagcc  
3241 ggatcaagcg tatgcagcc cgcatgca tcagccatga tggatactt ctcggcagga  
3301 gcaaggtgag atgacaggag atctgccc ggacttcgc ccaatagcag ccagtcctt  
3361 cccgctcag tgacaacgtc gagcacagct gcgcaaggaa cgcccgtcgt ggccagccac  
3421 gatagccgc ctgctcgtc ctgagttca ttcaggcac cggacaggtc ggtctgaca  
3481 aaaagaaccg ggcgccctg cgctgacagc cggaacacgg cggcatcaga gcagccgatt  
3541 gtctgttg cccagtcata gccgaatg

//

## ZFP5475 activator p15A/Kan

LOCUS 21-62 3593 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(1920..2631)  
/gene="p15A"  
terminator 2722..2827  
/gene="Terminator"  
terminator 244..289  
/gene="term BBa\_B0011"  
prot\_bind 356..370  
/gene="RBS"  
promoter 296..349  
/gene="pLtetO-1"  
CDS 1127..1144  
/gene="linker"  
/codon\_start="0"  
CDS complement(2853..3593)  
/gene="KanR"  
/codon\_start="0"  
terminator 1785..1913  
/gene="double term"  
CDS 377..1120  
/gene="RNAP alpha subunit"  
/codon\_start="0"  
CDS 1151..1708  
/gene="ZFP5475"  
/codon\_start="0"  
terminator 1715..1760  
/gene="term BBa\_B0011"

BASE COUNT 893 a 988 c 923 g 789 t

### ORIGIN

```
1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa
61 cgatcctcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg
121 caagaaagcc atccagtta ctttgcaggg ctccaacc ttaccagagg gcgcccaggc
181 tggcaattcc gaattcatga gatctcatat ggcatggatg aactatacaa ataataagga
241 tctagagaat ataaaaagcc agattattaa tccggctttt ttattatttg gatcttcct
301 atcagtgata gagattgaca tcctatcag tgatagagat actgagcacg gatctattaa
361 agaggagaaa ggatctatgc agggttctgt gacagagttt ctaaaccgc gcctgggtga
421 tatcgagcaa gtgagttcga cgcagccaa ggtgaccctt gagccttag agcgtggctt
481 tggccatact ctgggtaacg cactgcgccg tattctgctc tcatcgatgc cgggttgccg
541 ggtgaccgag gttgagattg atggtgtact acatgagtac agcaccaaag aaggcgttca
601 ggaagatatc ctggaaatcc tgctcaacct gaaagggctg gcggtgagag ttcagggcaa
661 agatgaagtt attcttacct tgaataaatc tggcattggc cctgtgactg cagccgatat
721 caccacgac ggtgatgtcg aaatcgtcaa gccgcagcac gtgatctgcc acctgaccga
```



781 tgagaacgcg tctattagca tgcgtatcaa agttcagcgc ggtcgtggtt atgtgccggc  
841 ttctaccgca attcattcgg aagaagatga gcgccaatc ggccgtctgc tggcgcagc  
901 atgctacagc cctgtggagc gtattgccta caatgttgaa gcagcgcgtg tagaacagcg  
961 taccgacctg gacaagctgg tcatcgaat ggaaccaac ggcaaatcg atcctgaaga  
1021 ggcgattcgt cgtgcggcaa ccattctggc tgaacaactg gaagcttctg ttgacttacg  
1081 tgatgtacgt cagcctgaag tgaagaaga gaaaccagag ggatctgcc cgcgagtccg  
1141 gaccggatct ggggtaccg ccatggcgga gaggccttc cagtgtcga tctgcatgcg  
1201 taacttcagt cgtagtacc acctgagccg gcacatccgc acccacagc gcgagaagcc  
1261 ttttgcctgt gacatttggt ggaggaaatt tgcgacaac cgggaccgca caaagcatac  
1321 caagatacac acgggcggac agcggccgta cgcattccct gtcgagtcct gcgatgccg  
1381 ctttctgac aggaagacac ttatcgagca tatccgcatc cacaccggtc agaagcctt  
1441 ccagtgtcga atctgcatgc gtaacttcag taccagcagc ggctgagcc gccacatccg  
1501 caccacaca ggatctcaga agccttcca gtgtcgaatc tgcattcgtg acttcagtcg  
1561 tagtgaccac ctgagcgaac acatccgac ccaacagcc gagaagcctt tgcctgtga  
1621 catttggtgg aggaatttg ccaccagcag cgaccgcaca aagcatacca agatacacct  
1681 gcgcaaaaa gatcggccc ggtaataagg atctagagaa tataaaaagc cagattatta  
1741 atccggcttt ttattattt ggatcctaac tcgagtaagg atctccaggc atcaataaa  
1801 acgaaaggct cagtgcgaag actgggcctt tcgtttatc tgtgtttgt cgggtaacgc  
1861 tctctactag agtcacactg gtcaccttc ggggtggcct tctgcgtt atacctaggg  
1921 atatattccg ctctctgct cactgactcg ctacgctcgg tcgttcgact gcggcgagcg  
1981 gaaatggctt acgaacgggg cggagatttc ctggaagatg ccaggaagat acttaacagg  
2041 gaagtgcgag ggccgcgga aagccgttt tccataggct ccgccccct gacaagcatc  
2101 acgaaatctg acgctcaat cagtgggtgc gaaaccgac aggactataa agataccagg  
2161 cgtttcccc tggcggtcc ctctgctgct ctctgttcc tgccttccg tttaccggtg  
2221 tcattccgct gttatggccg cgtttgtctc attccacgcc tgacactcag ttccgggtag  
2281 gcagttcgtc ccaagctgga ctgtatgcac gaacccccg ttcagtcga ccgctgcgcc  
2341 ttatccggtg actatcgtct tgagtccaac ccggaagac atgcaaaagc accactggca  
2401 gcagccactg gtaattgatt tagaggagt agtcttgaag tcatgcgcc gtttaaggcta  
2461 aactgaaagg acaagtttg gtgactgcgc tctccaagc cagttacctc ggttcaaaaga  
2521 gttgtagct cagagaacct tcgaaaaacc gccctgcaag gcggttttt cgtttcaga  
2581 gcaagagatt acgagcagc caaacgac tcaagaagat catctatta atcagataaa  
2641 atatttctag atttcagtc aatttatctc ttcaaatgta gcacctgaag tcagccccat  
2701 acgataaag ttgttactag tgcttggtt ctaccaata aaaaacgccc ggcggcaacc  
2761 gagcgtctg acaaatcca gatggagttc tgaggtcatt actggatccta tcaacaggag  
2821 tccaagcgag ctctgaacc ccagagtcct gctcagaaga actcgtcaag aaggcgatg  
2881 aaggcgatgc gctgcgaatc gggagcggcg ataccgtaaa gcacgaggaa gcggtcagcc  
2941 cattcggcg caagctctc agcaatatca cgggtagcca acgctatgct ctgatagcgg  
3001 tccgacacac ccagccggc acagtcgatg aatccagaaa agcggccatt ttccaccatg  
3061 atattcggca agcaggcatc gccatgggtc acgacgagat cctcggcgtc gggcatgcgc  
3121 gccttgagcc tggcgaacag ttcggctggc gcgagccct gatgctctc gtccagatca  
3181 tctgatcga caagaccggc ttcatccga gtacgtgctc gctcagtcg atgttctgct  
3241 tgggtgctga atgggcaggt agccggatca agcgtatgca gccgccgat tgcattcagcc  
3301 atgatggata cttctcggc aggagcaagg tgagatgaca ggagatcctg ccccgccact  
3361 tcgccaata gcagccagtc cttcccgtc tcagtgaca cgtcagcagc agctgcgca  
3421 ggaacgccc tcgtggccag ccacgatagc cgcgctgctc cgtcctgag ttattcagg  
3481 gcaccgaca ggtcgtctt gacaaaaaga accggcgcc cctgcgctga cagccggaac

3541 acggcggcat cagagcagcc gattgtctgt tgtgccagc catagccgaa tag  
//

**ZFP03 mismatched atc/rep on p15A/Kan**

LOCUS 22-20 4413 bp DNA circular

FEATURES Location/Qualifiers  
CDS 1221..1964  
/gene="RNAP alpha subunit"  
/codon\_start="0"  
rep\_origin complement(2740..3451)  
/gene="p15A"  
CDS 371..1081  
/gene="mCherry"  
/codon\_start="0"  
CDS 1995..2528  
/gene="ZFP03"  
/codon\_start="0"  
terminator 2605..2733  
/gene="double term"  
terminator 3542..3647  
/gene="Terminator"  
promoter 282..343  
/gene="pwklac"  
promoter 1140..1193  
/gene="pLtetO-1"  
terminator 206..251  
/gene="term BBa\_B0011"  
CDS 1971..1988  
/gene="linker"  
/codon\_start="0"  
prot\_bind 258..275  
/gene="mismatched binding site (V2)"  
prot\_bind 350..364  
/gene="RBS"  
CDS complement(3673..4413)  
/gene="KanR"  
/codon\_start="0"  
terminator 2605..2733  
/gene="double term"  
terminator 1088..1133  
/gene="term BBa\_B0011"  
terminator 2535..2580  
/gene="term BBa\_B0011"  
BASE COUNT 1135 a 1146 c 1095 g 1037 t

## ORIGIN

1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa  
61 cgatectcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg  
121 caagaaagcc atccagtta ctttgcaggg cttccaacc ttaccagagg gcgccccagc  
181 tggcaattcc gaattcaaaa gatctagaga atataaaaag ccagattatt aatccggctt  
241 ttttattatt tggatctccg ccgcccccg cccgcgatc tccgtattct ttacacttta  
301 tgcttccggc tcgtatgttg tgtcgaccga gcggataaca attggatcta ttaaagagga  
361 gaaaggatct atgcgtaaag gagaagaaga taacatggct atcattaaag agttcatgcg  
421 cttcaaagtt cacatggagg gttctgttaa cggtcacgag ttcgagatcg aaggcgaagg  
481 cgagggcctg ccgtatgaag gacccagac cgccaaactg aaagtacta aaggcggccc  
541 gctgcctttt gcgtgggaca tcctgagccc gcaatttatg tacggttcta aagcgtatg  
601 taaacacca gcggatatcc cggactatct gaagctgtct tttccggaag gtttcaagtg  
661 ggaacgcgta atgaatttg aagatgggtg tgtcgtgacc gtcactcagg actcctcct  
721 gcaggatggc gagttcatct ataaagtaa actgcgtggg actaatttc catctgatgg  
781 cccggtgatg cagaaaaaga cgatgggttg ggaggcgtct agcgaacgca tgotccgga  
841 agatggtgcg ctgaaaggcg aattaaaca ggcctgaaa ctgaaagatg gcggccatta  
901 tgacgctgaa gtgaaaacca cgtacaaagc caagaaacct gtgcagctgc ctggcgcgta  
961 caatgtgaat attaaactgg acatcacctc tcataatgaa gattatacga tcgtagagca  
1021 atatgagcgc gcggagggtc gtcattctac cggtggcatg gatgaactat acaataata  
1081 aggatctaga gaatataaaa agccagatta ttaatccggc tttttatta tttggatctt  
1141 ccctatcagt gatagagatt gacatcccta tcagtgatag agatactgag cacggatcta  
1201 ttaaagagga gaaaggatct atgcagggtt ctgtgacaga gtttctaaaa ccgpcctgg  
1261 ttgatatcga gcaagtgagt tcgacgcacg ccaaggtgac ccttgagcct ttgagcgtg  
1321 gctttggcca tactctgggt aacgcactgc gccgtattct gctctcatcg atgccgggtt  
1381 gcgpcgtgac cgaggtgag attgatggtg tactacatga gtacagcacc aaagaaggcg  
1441 ttcaggaaga tctctggaa atcctgctca acctgaaagg gctggcgggt agagttcagg  
1501 gcaaagatga agttattctt acctgaata aatctggcat tggcctgtg actgcagccg  
1561 atatcacca cgacggtgat gtcgaaatcg tcaagccgca gcacgtgatc tgccactga  
1621 ccgatgagaa cgcgtctatt agcatgcgta tcaaagtca gcgpcgtcgt ggttatgtgc  
1681 cggcttctac ccgaattcat tcggaagaag atgagcgcgc aatcggcctg ctgctggtcg  
1741 acgcatgcta cagcctgtg gagcgtattg cctacaatgt tgaagcagcg cgtgtagaac  
1801 agcgtaccga cctggacaag ctggtcatcg aatggaaac caacggcaca atcgtcctg  
1861 aagaggcgt tcgtcgtgcg gcaaccattc tggctgaaca actggaagct ttcgttact  
1921 tacgtgatgt acgtcagcct gaagtgaag aagagaaacc agagggatct gcccgcgag  
1981 tccggaccgg atctctggaa ccaggatcta aaccgtaca atgtccggaa tgtgtaaat  
2041 cttctccac tcatctggat ctgattcgtc atcaacgtac tcactgga tctaaacctg  
2101 acaaatgtcc ggaatgtggt aatccttct ccaatcttc ttctctggtt cgtcatcaac  
2161 gtactcacac tggatctaaa ccgtacaaat gtcggaatg tggtaaatcc ttctccaaa  
2221 actctacct gactgaacat caacgtactc aactggatc taaacctgac aatgtccgg  
2281 aatgtggtaa atccttctcc tctcgtcgtg cttgtcgtg acatcaactg actcacactg  
2341 gatctaaacc gtacaaatgt ccggaatgtg gtaaatcctt ctcccaactg gctcatctgc  
2401 gtgcacatca acgtactcac actggatcta aaccgtaca atgtccggaa tgtgtaaat  
2461 cttctccac ttctggtcat ctggtacgtc atcaacgtac tcactgga tctaaaacct  
2521 ctaataagg atctagagaa tataaaaagc cagattatta atccggcttt tttattatt  
2581 ggatcctaac tcgagtaagg atctccaggc atcaataaaa acgaaaggct cagtcgaaag  
2641 actgggcctt tcgtttatc tgtgtttgt cggtaacgc tctctactag agtcacactg

2701 gctcaccttc ggggtggcct tctgctgtt atacctaggg atatattccg cttcctcgct  
 2761 cactgactcg ctacgctcgg tcgttcgact gcggcgagcg gaaatggctt acgaacggggg  
 2821 cggagatttc ctggaagatg ccaggaagat acttaacagg gaagtgagag ggccgcggca  
 2881 aagccgtttt tccataggct ccgccccct gacaagcatc acgaaatctg acgctcaaat  
 2941 cagtgtggc gaaacccgac aggactataa agataccagg cgtttcccc tggcggctcc  
 3001 ctctgctgct ctctgttcc tgccttcgg ttaccggtg tcattccgct gttatggccg  
 3061 cgttgtctc attccacgcc tgacactcag ttccgggtag gcagttcgct ccaagctgga  
 3121 ctgtatgcac gaacccccg ttcagtccga ccgctgcgcc ttatccgta actatcgtct  
 3181 tgagtccaac ccgaaagac atgcaaaagc accactggca gcagccactg gtaattgatt  
 3241 tagaggagt agtcttgaag tcattgcgccc gtttaaggta aactgaaagg acaagtttg  
 3301 gtgactgcgc tctccaagc cagttacctc ggttcaaaga gttgtagct cagagaacct  
 3361 tcgaaaaacc gcctgcaag gcggttttt cgtttcaga gcaagagatt acgcgcagac  
 3421 caaacgata tcaagaagat catctatta atcagataaa atatttctag atttcagtgc  
 3481 aatttatctc tcaaatgta gcacctgaag tcagcccat acgatataag ttgttactag  
 3541 tgcttgatt ctaccaata aaaaacgcc ggcggcaacc gagcgttctg acaaatcca  
 3601 gatggagttc tgaggtcatt actggatcta tcaacaggag tccaagcgag ctctgaacc  
 3661 ccagagtccc gtcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc  
 3721 gggagcggcg atacctaaa gcacaggaa gcggtcagcc cattcggcg caagctcttc  
 3781 agcaatatca cgggtagcca acgctatgct ctgatagcgg tccgccacac ccagccggcc  
 3841 acagtcgatg aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc  
 3901 gccatgggta acgacgagat cctcggctg ggcgatgcgc gccttgagcc tggcgaacag  
 3961 ttcggctggc gcgagcccct gatgctctc gtccagatca tctgatcga caagaccggc  
 4021 ttccatccga gtacgtgctc gctcgtgctg atgttctgct tgggtgctga atgggcaggt  
 4081 agccggatca agcgtatgca gccgccgat tgcattcagc atgatggata ctttctcggc  
 4141 aggagcaagg tgagatgaca ggagatcctg ccccggcact tcgccaata gcagccagtc  
 4201 cttccccgct tcagtgaaa cgtcagcac agctgcgcaa ggaacgcccg tcgtggccag  
 4261 ccacgatagc cgcgctgctc cgtcctgcag ttattcagg gcaccggaca ggtcggctt  
 4321 gacaaaaaga accgggcgcc cctcgtgta cagccggaac acggcggcat cagagcagcc  
 4381 gattgtctgt tgtgccagt catagccgaa tag

//

**ZFP5475 mismatched act/rep on p15A/Kan**

LOCUS 22-22 4437 bp DNA circular

FEATURES Location/Qualifiers  
 CDS 1221..1964  
     /gene="RNAP alpha subunit"  
     /codon\_start="0"  
 rep\_origin complement(2764..3475)  
     /gene="p15A"  
 gene 371..1081  
     /gene="mCherry"  
 terminator 2629..2757  
     /gene="double term"  
 terminator 3566..3671

```

        /gene="Terminator"
promoter  1140..1193
        /gene="pLtetO-1"
terminator  206..251
        /gene="term BBa_B0011"
CDS      1971..1988
        /gene="linker"
        /codon_start="0"
prot_bind  258..275
        /gene="mismatched binding site (V2)"
prot_bind  350..364
        /gene="RBS"
promoter  282..343
        /gene="pwklac"
CDS      complement(3697..4437)
        /gene="KanR"
        /codon_start="0"
CDS      1995..2552
        /gene="ZFP5475"
        /codon_start="0"
terminator  1089..1134
        /gene="term BBa_B0011"
terminator  2560..2605
        /gene="term BBa_B0011"
BASE COUNT  1128 a 1171 c 1147 g 991 t
ORIGIN
1  cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa
61  cgatcctcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg
121 caagaaagcc atccagtta cttgcaggg cttccaacc ttaccagagg ggcgcccgcc
181 tggcaattcc gaattcaaaa gatctagaga atataaaaag ccagattatt aatccggctt
241 ttttattatt tggatctccg ccgcccccg cccgcggatc tccgtattct ttacacttta
301 tgcttccggc tcgtatgttg tgcgaccga gcggataaca attggatcta ttaaagagga
361 gaaaggatct atgcgtaaag gagaagaaga taacatggct atcattaaag agttcatgcg
421 cttcaaagtt cacatggagg gttctgttaa cggtcacgag ttcgagatcg aaggcgaagg
481 cgagggccgt ccgtatgaag gcaccagac gcctaaactg aaagtgacta aaggcggccc
541 gctgcctttt gcgtgggaca tcctgagccc gcaatttatg tacggttcta aagcgtatgt
601 taaacacca gcggatatcc cggactatct gaagctgtct tttccggaag gtttcaagtg
661 ggaacgcgta atgaatttg aagatgggtg tgcgtgacc gtcactcagg actcctcct
721 gcaggatggc gagttcatct ataaagtaa actgcgtggt actaatttc catctgatgg
781 cccggtgatg cagaaaaaga cgatgggttg ggaggcgtct agcgaacgca tgtatccgga
841 agatggtgcg ctgaaaggcg aattaaaca ggcctgaaa ctgaaagatg gcggccatta
901 tgacgctgaa gtgaaaacca cgtacaaagc caagaaacct gtcagctgc ctggcgcgta
961 caatgtgaat attaaactgg acatcacctc tcataatgaa gattatacga tcgtagagca
1021 atatgagcgc gcggagggtc gtcattctac cgggtggcatg gatgaactat acaaataata
1081 aggatctaga gaatataaaa agccagatta ttaatccggc tttttatta tttggatctt
1141 ccctatcagt gatagagatt gacatcccta tcagtgatag agatactgag cacggatcta

```

1201 ttaaagagga gaaaggatct atgcagggtt ctgtgacaga gtttctaaaa ccgcgcttg  
1261 ttgatatcga gcaagtgagt tcgacgcacg ccaaggtgac ccttgagcct ttagagcgtg  
1321 gctttggcca tactctgggt aacgcactgc gccgtattct gctctcatcg atgccgggtt  
1381 gcgcggtgac cgaggtgag attgatgggt tactacatga gtacagcacc aaagaaggcg  
1441 ttcaggaaga tatcctggaa atcctgctca acctgaaagg gctggcggtg agagttcagg  
1501 gcaaagatga agttattctt acctgaata aatctggcat tggcctgtg actgcagccg  
1561 atatcaccca cgacgtgat gtcgaaatcg tcaagccgca gcacgtgatc tgccacctga  
1621 ccgatgagaa cgcgtctatt agcatgctga tcaaagttca gcgcggtcgt ggttatgtgc  
1681 cggcttctac ccgaattcat tcggaagaag atgagcggcc aatcggcctg ctgctggctg  
1741 acgcatgcta cagccctgtg gagcgtattg cctacaatgt tgaagcagcg cgtgtagaac  
1801 agcgtaccga cctggacaag ctggatcatg aatggaaac caacggcaca atcgatcctg  
1861 aagaggcgat tcgtcgtgcg gcaaccattc tggctgaaca actggaagct ttcgtgact  
1921 tacgtgatgt acgtcagcct gaagtgaag aagagaaacc agagggatct gccccgag  
1981 tccggaccgg atctgggta cccgccatgg cggagaggcc cttccagtgt cgaatctgca  
2041 tgcgtaactt cagtcgtagt gaccacctga gccggcacat ccgcaccac acaggcgaga  
2101 agccttttgc ctgtgacatt tgtgggagga aatttgccga caaccgggac cgcacaaagc  
2161 ataccaagat acacacgggc ggacagcggc cgtacgcatg cctgtcag tctcgcgatc  
2221 gccgctttc tgacaggaag acattatcg agcatatccg catccacacc ggctagaagc  
2281 cttccagtg tcgaatctgc atgcgtaact tcagtaccag cagcggcctg agccgccaca  
2341 tccgaccca cacaggatct cagaagccct tccagtgtcg aatctgcatg cgtaactca  
2401 gtcgtagtga ccactgagc gaacacatcc gcaccacac aggcgagaag cttttgcct  
2461 gtgacattg tgggaggaaa tttgccacca gcagcgaccg cacaaagcat accaagatac  
2521 acctgcgcca aaaagatgcg gcccggtaat aaggatctag agaataaaa aagccagatt  
2581 attaatccgg ctttttatt atttgatcc taactcgagt aaggatctcc aggcatcaaa  
2641 taaaacgaaa ggctcagtcg aaagactggg ctttctgtt tatctgttgt ttgtcggta  
2701 acgctctcta ctgagtcac actggctcac ctccgggtgg gcctttctgc gttatacct  
2761 agggatatac tccgcttct cgctactga ctgctacgc tcggtcgttc gactcggcg  
2821 agcggaaatg gttacgaac gggcgaggaga tttctggaa gatgccagga agatacttaa  
2881 cagggaagt agagggccgc ggcaaagccg ttttccata ggctccgcc cctgacaag  
2941 catcacgaaa tctgacgctc aatcagtg tggcgaacc cgacaggact ataaagatac  
3001 caggcgttc ccctggcgg ctccctcgtg cgctctctg ttctgcctt tcggttacc  
3061 ggtgtcattc cgctgtatg gccggttg tctattcca cgctgacac tcagttccgg  
3121 gtaggcagtt cgctccaagc tggactgtat gcacgaacc cccgttcagt ccgaccgctg  
3181 cgcctatcc ggtaactatc gtcttgagtc caaccggaa agacatgcaa aagcaccact  
3241 ggcagcagcc actggttaatt gatttagagg agttagtctt gaagtcatgc gccggttaag  
3301 gctaaactga aaggacaagt tttggtgact gcgctcctcc aagccagtta cctcgttca  
3361 aagagttgg agctcagaga acctcgaaa aaccgcctg caaggcgtt tttcgttt  
3421 cagagcaaga gattacgcg agacaaaac gatctcaaga agatcatctt attaacaga  
3481 taaaatatt ctgatttca gtgcaatta tctttcaaa ttagcacct gaagtacgc  
3541 ccatacgata taagttgta ctagtcttg gattctcacc aataaaaaac gcccgcggc  
3601 aaccgagcgt tctgaacaaa tccagatgga gttctgaggt cactactgga tctatcaaca  
3661 ggagtccaag cgagctctcg aacccagag tccgctcag aagaactcgt caagaaggcg  
3721 atagaaggcg atgcgtgcg aatcgggagc ggcgataccg taaagcacga ggaagcggc  
3781 agccattcgc ccgcaagct cttagcaat atcacgggta gccaacgcta tgtctgata  
3841 gcggtccgcc acaccagcc ggccacagtc gatgaatca gaaaagcggc catttccac  
3901 catgatattc ggcaagcagg catcgcatg ggtcacgac agatctcgc cgtcgggcat

3961 gcgcgcttg agcctggcga acagttcggc tggcgcgagc cctgatgct cttcgtccag  
4021 atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgtcga tgcgatgttt  
4081 cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc gcattgcatc  
4141 agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat cctgccccgg  
4201 cacttcgccc aatagcagcc agtccttcc cgcttcagtg acaacgtcga gcacagctgc  
4261 gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct gcagttcatt  
4321 cagggcaccg gacaggtcgg tcttgacaaa aagaaccggg cgcccctgcg ctgacagccg  
4381 gaacacggcg gcatcagagc agccgattgt ctgttgcc cagtcatagc cgaatag

//

**ZFP5475 leak on p15A/Kan**

LOCUS 22-2 2970 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(1297..2008)

/gene="p15A"

CDS 375..1085

/gene="mCherry"

/codon\_start="0"

terminator 1162..1290

/gene="double term"

terminator 2099..2204

/gene="Terminator"

promoter 286..347

/gene="pwklac"

terminator 206..251

/gene="term BBa\_B0011"

prot\_bind 258..279

/gene="ZFP5475 binding site"

prot\_bind 354..368

/gene="RBS"

promoter 286..347

/gene="pwklac"

CDS complement(2230..2970)

/gene="KanR"

/codon\_start="0"

terminator 1162..1290

/gene="double term"

terminator 1092..1137

/gene="term BBa\_B0011"

BASE COUNT 749 a 778 c 761 g 682 t

ORIGIN

1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa  
61 cgatcctcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg  
121 caagaaagcc atccagtta ctttgcaggg cttcccaacc ttaccagagg gcgccccagc

181 tggcaattcc gaattcatga gatctagaga atataaaaag ccagattatt aatccggctt  
241 ttttattatt tggatctaca cccgggttcc cctcgggtgg gatctccgta ttctttacac  
301 tttatgcttc cggctcgtat gttgtgtcga ccgagcggat aacaattgga tctattaaag  
361 aggagaaagg atctatgctg aaaggagaag aagataacat ggctatcatt aaagagttca  
421 tgcgcttcaa agttcacatg gagggttctg ttaacggta cgagttcgag atcgaaggcg  
481 aaggcgaggg ccgtccgtat gaaggcacc agaccgcaa actgaaagt actaaaggcg  
541 gcccgtgcc ttttgcgtgg gacatcctga gcccgaatt tatgtacggg tctaaagcgt  
601 atgttaaaca ccagcggat atcccggact atctgaagct gtcttttccg gaaggttca  
661 agtgggaacg cgtaatgaat tttgaagatg gtgggtcgt gaccgtcact caggactct  
721 ccctgcagga tggcgagttc atctataaag ttaaactgcg tggactaat tttccatctg  
781 atggcccggg gatgcagaaa aagacgatgg gttgggaggg gtctagcga cgcagtatc  
841 cggaagatgg tgcgctgaaa ggcgaaatta aacagcgcct gaaactgaaa gatggcggcg  
901 attatgacgc tgaagtgaaa accacgtaca aagccaagaa acctgtgag ctgctggcg  
961 cgtacaatgt gaatattaaa ctggacatca ccttcataa tgaagattat acgatcgtag  
1021 agcaatatga gcgcgaggag ggtcgtcatt ctaccgggtg catggatgaa ctatacaaat  
1081 aataaggatc tagagaatat aaaaagccag attattaatc cggtttttt attatttggg  
1141 tctaaactc agtaaggatc tccaggcatc aaataaaacg aaaggctcag tcgaaagact  
1201 gggcctttc tttatctgt tgtttgctgg tgaacgtct ctactagagt cacactggct  
1261 cacctcggg tgggccttc tgcgttata ctagggata tattccgctt cctcgtcac  
1321 tgactcgta cgctcggctg ttcactgcg gcgagcggaa atggcttac aacggggcgg  
1381 agatttctg gaagatgcca ggaagatact taacagggaa gtgagagggc cgcggcaaaag  
1441 ccgttttcc ataggctccg ccccctgac aagcatcacg aatctgacg ctcaaatcag  
1501 tggggcgaa accgacagg actataaaga taccaggctg tccccctgg cggctcctc  
1561 gtgcgctctc ctgttctgc ctttcggtt accggtgca ttcgctgtt atggccgct  
1621 ttgtctcatt ccacgcctga cactcagttc cgggtaggca gttcgtcca agctggactg  
1681 tatgcagaa cccccgttc agtccgaccg ctgctcctta tccgtaact atcgtctga  
1741 gtccaaccg gaaagacatg caaaagcacc actggcagca gccactggta attgattag  
1801 aggagttagt cttgaagta tgcgccggtt aaggctaaac tgaaggaca agttttggg  
1861 actgcgctc tccaagccag ttacctcgt tcaaagagtt gtagctcag agaactctg  
1921 aaaaaccgcc ctgcaaggcg gtttttctg tttcagagca agagattacg cgcagacaa  
1981 aacgatctca agaagatcat ctattaatc agataaaata ttctagatt tcagtcaat  
2041 ttatcttcc aatgtagca cctgaagtca gcccatac atataagtg ttactagtgc  
2101 ttgatttcc accaataaaa aacgcccggc ggcaaccgag cgttctgaac aatccagat  
2161 ggagtctga ggtcattact ggatctatca acaggagtcc aagcgagctc tcgaaccca  
2221 gactcccgt cagaagaact cgtcaagaag gcgatagaag gcgatgctc gcgaatcggg  
2281 agcggcgata ccgtaaagca cgaggaagcg gtcagcccat tcgcccaa gctcttcagc  
2341 aatatcacgg gtagccaacg ctatgtctg atagcggcc gccacacca gccggccaca  
2401 gtcgatgaat ccagaaaagc ggccatttc caccatgata ttcggcaagc aggcacgcc  
2461 atgggtcacg acgagatcct cgccgtcggg catgctcggc ttgagcctgg cgaacagttc  
2521 ggctggcgcg agcccctgat gctcttcgtc cagatcatcc tgatcgaaa gaccggcttc  
2581 catccgagta cgtgctcgt cgatgcgat tttcgttgg tggtcgaatg ggcaggtagc  
2641 cggatcaagc gtatgcagcc gccgattgc atcagccatg atggatactt tctcggcagg  
2701 agcaaggtga gatgacagga gatcctgcc cggcacttcg cccaatagca gccagtcct  
2761 tcccgttca gtgacaacgt cgagcacagc tgcgcaagga acgcccgtc tggccagcca  
2821 cgatagccgc gctgctcgt cctgcagttc attcaggga ccggacaggt cgttctgac  
2881 aaaaagaacc gggcgcctt gcgtgacag ccggaacacg gcggcatcag agcagccgat



2941 tgtctgttgt gccagtcac agccgaatag  
//

**ZFP03 leak on p15A/Kan**

LOCUS 22-3 2966 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(1293..2004)  
/gene="p15A"  
CDS 371..1081  
/gene="mCherry"  
/codon\_start="0"  
terminator 2095..2200  
/gene="Terminator"  
terminator 206..251  
/gene="term BBa\_B0011"  
prot\_bind 258..275  
/gene="ZFP03 binding site"  
prot\_bind 350..364  
/gene="RBS"  
promoter 282..343  
/gene="pwklac"  
CDS complement(2226..2966)  
/gene="KanR"  
/codon\_start="0"  
terminator 1158..1286  
/gene="double term"  
terminator 1088..1133  
/gene="term BBa\_B0011"  
BASE COUNT 752 a 772 c 759 g 683 t  
ORIGIN

1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa  
61 cgatecctcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg  
121 caagaaagcc atccagtta ctttgcaggg cttccaacc ttaccagagg gcgccccagc  
181 tggcaattcc gaattcatga gatctagaga atataaaaag ccagattatt aatccggcct  
241 ttttattatt tggatctggt agacgtctag taactggatc tccgtattct ttacacttta  
301 tgcttccggc tcgtatgttg tgtcgaccga gcggataaca attggatcta ttaaagagga  
361 gaaaggatct atgcgtaaag gagaagaaga taacatggct atcattaaag agttcatgcg  
421 cttcaaagtt cacatggagg gttctgttaa cggtcacgag ttcgagatcg aaggcgaagg  
481 cgagggccgt ccgtatgaag gcaccagac cgccaaactg aaagtgacta aaggcggccc  
541 gctgcctttt gcgtgggaca tctgagccc gcaatttatg tacggttcta aagcgtatgt  
601 taaacacca gcggatatcc cggactatct gaagctgtct tttccggaag gtttcaagt  
661 ggaacgcgta atgaatttg aagatgggtg tgtcgtgacc gtcactcagg actcctcct  
721 gcaggatggc gagttcatct ataaagtaa actgcgtggt actaatttc catctgatgg  
781 cccggtgatg cagaaaaaga cgatgggttg ggaggcgtct agcgaacgca tgtatccgga

841 agatggtgcg ctgaaaggcg aaattaaaca ggcctgaaa ctgaaagatg gcgccatta  
901 tgacgctgaa gtgaaaacca cgtacaaagc caagaaacct gtgcagctgc ctggcgcgta  
961 caatgtgaat attaaactgg acatcacctc tcataatgaa gattatacga tcgtagagca  
1021 atatgagcgc gcggagggtc gtcattctac cgggtggcatg gatgaactat acaataata  
1081 aggatctaga gaatataaaa agccagatta ttaatccggc tttttatta tttggatcct  
1141 aactcgagta aggatctcca ggcatcaaat aaaacgaaag gctcagtcga aagactgggc  
1201 ctttcgtttt atctgttgtt tgcggtgaa cgctctctac tagagtcaca ctggctcacc  
1261 ttcgggtggg ctttctgcg ttataccta gggatatatt ccgcttctc gctcactgac  
1321 tcgctacgct cggtcgttcg actgcggcga gcggaaatgg cttacgaacg gggcggagat  
1381 ttctggaag atgccaggaa gatacttaac aggggaagtga gagggccgcg gcaaagccgt  
1441 tttccatag gctccgccc cctgacaagc atcacgaaat ctgacgctca aatcagtgtt  
1501 ggcgaaacc gacaggacta taaagatacc aggcgtttcc ccttggcggc tcctcgtgc  
1561 gctctctgt tctgccttt cggttaccg gtgtcattcc gctgttatgg ccgctttgt  
1621 ctattccac gctgacact cagttccggg taggcagttc gctccaagct ggactgtatg  
1681 cacgaaccc ccgttcagtc cgaccgctgc gccttatccg gtaactatcg tcttgatgc  
1741 aaccgggaaa gacatgcaaa agcaccactg gcagcagcca ctggtaattg attagaggaa  
1801 gtagtcttg aagtcagtc cgggtaagg ctaaactgaa aggacaagt ttggtgactg  
1861 cgctctcca agccagttac ctcggtcaa agagttgta gctcagagaa ccttcgaaaa  
1921 accgcctgc aaggcggtt tttcgtttc agagcaagag attacgcgca gacaaaacg  
1981 atctcaagaa gatcatctta ttaacagat aaaatattc tagatttcag tgcaattat  
2041 ctctcaaat gtagcacctg aagtcagccc catacगतat aagttgttac tagtgcttg  
2101 attctacca ataaaaacg cccggcggca accgagcgtt ctgaacaaat ccagatggag  
2161 ttctgaggtc attactggat ctatcaacag gactccaagc gagctctcga acccagagt  
2221 cccgctcaga agaactcgtc aagaaggcga tagaaggcga tgcgctcga atcgggagc  
2281 gcgatacctg aaagcacgag gaagcggcga gccattcgc cgcaagctc ttcagcaata  
2341 tcacggtag ccaacgctat gtcctgatag cggccgcca caccagccg gccacagtcg  
2401 atgaatccag aaaagcggcc atttccacc atgatattc gcaagcaggc atcgccatgg  
2461 gtcacgacga gatcctcgc gtcgggcatg cgcgcctga gcttggcga cagttcggct  
2521 ggcgcgagcc cctgatgctc ttcgtcaga tcctctgat cgacaagacc ggcttccatc  
2581 cgagtacgtg ctcgctcgt cgcgatgttc gcttgggtgt cgaatgggca ggtagccgga  
2641 tcaagcgtat gcagccgag cattgcatca gccatgatgg atacttctc ggcaggagca  
2701 aggtgagatg acaggagatc ctgccccggc acttcgcca atagcagcca gtccttccc  
2761 gcttcagtga caacgtcag cacagctgcg caaggaacgc ccgtcgtggc cagccacgat  
2821 agccgcgctg cctcgtctg cagttcattc agggcaccgg acaggtcgtt cttgacaaaa  
2881 agaaccgggc gccctcgcg tgacagccgg aacacggcgg catcagagca gccgattgc  
2941 tgtgtgccc agtcatagcc gaatag

//

**ZFP5475 matched atc/rep on p15A/Kan**

LOCUS 22-4 4441 bp DNA circular

FEATURES Location/Qualifiers  
CDS 1225..1968  
/gene="RNAP alpha subunit"  
/codon\_start="0"

```

rep_origin complement(2768..3479)
      /gene="p15a"
CDS      375..1085
      /gene="mCherry"
      /codon_start="0"
terminator 2633..2761
      /gene="double term"
terminator 3570..3675
      /gene="Terminator"
promoter 286..347
      /gene="pwklac"
promoter 1144..1197
      /gene="pLtetO-1"
terminator 206..251
      /gene="term BBa_B0011"
prot_bind 258..279
      /gene="ZFP5475 binding site"
CDS      1975..1992
      /gene="linker"
      /codon_start="0"
prot_bind 354..368
      /gene="RBS"
CDS      complement(3701..4441)
      /gene="KanR"
      /codon_start="0"
CDS      1999..2556
      /gene="ZFP5475"
      /codon_start="0"
terminator 1093..1138
      /gene="term BBa_B0011"
terminator 2564..2609
      /gene="term BBa_B0011"
BASE COUNT 1128 a 1166 c 1151 g 996 t
ORIGIN
  1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa
 61 cgatcctcat cctgtctctt gatcagatca tgatcccctg cgcatcaga tccttggcgg
121 caagaaagcc atccagtta cttgacaggg cttccaacc ttaccagagg gcgccccagc
181 tggcaattcc gaattcatga gatctagaga atataaaaag ccagattatt aatccggctt
241 tttattatt tggatctaca cccgggttcc cctcgggtgg gatctccgta ttctttacac
301 tttatgcttc cggctcgtat gttgtgctga ccgagcggat aacaattgga tctattaaag
361 aggagaaagg atctatgctg aaaggagaag aagataacat ggctatcatt aaagagtcca
421 tgcgcttcaa agttcacatg gagggttctg ttaacggctc cgagttcgag atcgaaggcg
481 aaggcgaggg ccgtccgtat gaaggcacc agaccgcaa actgaaagtg actaaaggcg
541 gcccgtgcc ttttgcgtgg gacatctga gcccgcaatt tatgtacggt tctaaagcgt
601 atgttaaaca ccagcggat atcccggact atctgaagct gtctttccg gaaggttca
661 agtgggaacg cgtaatgaat tttgaagatg gtggtgtcgt gaccgtcact caggactcct

```

721 ccctgcagga tggcgagttc atctataaag ttaaactgcg tggactaat tttccatctg  
781 atggcccggg gatgcagaaa aagacgatgg gttgggaggc gtctagcгаа cgcатgtatc  
841 cggaaatgg tgcgctgaaa ggcгааatta aacagcgcct gaaactgaaa gatggcggcc  
901 attatgacgc tgaagtгaaa accacgtaca aagccaagaa acctgtgсag ctgcctggcg  
961 cgtacaatgt gaatattaa ctggacatca cctctcataa tгаagattat acgatcgtag  
1021 agcaatatga gcgcgсggag ggtcgtcatt ctaccggтgg catggatgaa ctatacaat  
1081 aataaggatc tagagaatат aaaaagccag attattaatc cggctttttt attatttгga  
1141 tcttcctat cagtгataga gattgacatc cctatcagtг atagagatac tgagcacгga  
1201 tctattaaag aggagaaagg atctatgcag gttctgtga cagagtttct aaaaccgcgc  
1261 ctggttgata tcgagcaagt gagttcгacg cacгccaagg tgacccttga gcctttagag  
1321 cgtggctttg gccatactct gggtaacгca ctgcгccгta ttctгctctc atcатgсcg  
1381 ggttgcгcgg tgaccgaggt tgagattgat ggtgtactac atgagtacag caccaaгaa  
1441 ggcgttcagg aagatатcct ggaaatcctg ctcaacctga aaggгctggc ggtgagagtt  
1501 caggгcaaag atгаagttat tcttacctg aataaatctg gcattggccc tgtgactгca  
1561 gccgatatca cccacгacгg tgatgtcгаа atcgtcaagc gcгagcacgt gatctgсcaс  
1621 ctgaccgatg agaacгcгtc tattagcatg cgtatcaaag ttcagcгcгg tcгtgгttat  
1681 gtgсcгgctt ctaccgaaт tcatcгgгаа gaagatgagc gcccaatcгg ccгctгctg  
1741 gtcгacгcat gctacagccc tgtggagcgt attgсctaca atgttгаagc агcгcгtgta  
1801 gaacagcгta ccgacctgga caagctгgtc atcгааatgg aaaccaacгg cacaatcгat  
1861 cctгааagagg cгattcгtcг tgcгgcaacc attctгgctg aacaactgga агctttcгtt  
1921 gacttacгtg atgtacгta gcctгааtg aaгаagaga aaccagaggg atctgccccг  
1981 cгagtcгgga ccgгatctгg ggtaccгcc atggcгgгa ggcсctcca gtгtcгаatc  
2041 tgatгcгta acttcagtcг tagtgaccac ctgagccгgс acatccгcaс ccaсacagгc  
2101 gagaagcctt ttgсctгtga catttгtggg агgaaattg ccгacaaccг ggaccгcaс  
2161 aagcatacсa агatacacac gggcгgacag cгgсcгtacг catгcсctgt cгagtcctгc  
2221 gatгcгcгct tttctgacag gaagacatt atcгagcата tccгatcca caccgгtcг  
2281 aagсcttcc агtгtcгаat ctгcatгcgt aacttcagta ccagcagcгg cctgagccгc  
2341 cacatccгca cccacacagг atctcгааag cccttcagт gtcгаatctg catгcгtaac  
2401 ttcagtcгta gtgaccact gagcгааacac atccгcaccс acacagгcга gaagсcttt  
2461 gcctgtgaca tttgtgggag gaaatttгcc accagcagcг accгcaсaaa gcataccaag  
2521 atacactгc gccaaaaга tgcгgсccгg таатаaggat ctagagaata таaaaгсca  
2581 gattattaat cggcttttt tattatttгg atcctaact гagтаaggat ctccagгcat  
2641 caaataaaac gaaaggтca gtcгааagac tggгcctttc gttttatctg ttgtttгtcг  
2701 gtгаacгtc tctactagag tcacactгgс tcaccttcгg gtggгccttt ctгcгttat  
2761 acctagгgat atattcгct tctctгctca ctgactcгct acгctcгgtc gttcгactгc  
2821 ggcгagcгga aatggcttac gaacгggгcг gagatttctt gгаagatгcс agгаagatac  
2881 ttaacagгga агtgagagгg ccгcгgсaaa gccgttttc cataggctcc gccccctga  
2941 caagcatcac gaaatctгac gctcaaatca gtggtggcга aaccгacag gactataaг  
3001 ataccagгcг tttcccctg gcгgctcсct cgtгcгctct cctgttctg ccttcгgtt  
3061 taccggtгtc atccгctgt tatggсcгcг tttgtctcat tccacгcctg aactcagtt  
3121 ccgggtaggc агttcгctcc aagctгgact gtatгcacга accccccгtt cagtccгacc  
3181 gctгcгcctt atccгgтаac tatгtcttg агtccaaccс gгаaagacat gcaaaгcac  
3241 cactggcagc агcactгgt aattgatta gaggagttag tctгаagtc atгcгcгgt  
3301 таaggctaaa ctгааaggac aagtttгgt gactгcгctc ctccaagсca gttactcгg  
3361 ttcaaagagт tggtagctca gaaacctc gaaaaccгc cctгcaaggc ggtttttcг  
3421 tttcagagc aagagattac gcгcagacca aaacгatctc aгаagatca tcttattaat

3481 cagataaaat atttctagat ttcagtcaa tttatctctt caaatgtagc acctgaagtc  
 3541 agccccatac gatataagtt gttactagtg cttggattct caccaataaa aaacgcccgg  
 3601 cggcaaccga gcgttctgaa caaatccaga tggagtctg aggtcattac tggatctatc  
 3661 aacaggagtc caagcgagct ctcgaacccc agagtcccgc tcagaagaac tcgtcaagaa  
 3721 ggcgatagaa ggcgatgcmc tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc  
 3781 ggtcagccca ttcgcccca agctcttcag caatatcacg ggtagccaac gctatgtcct  
 3841 gatagcggtc cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatfff  
 3901 ccaccatgat attcggcaag caggcatcmc catgggtcac gacgagatcc tcgccgtcgg  
 3961 gcatgcgcmc cttgagcctg gcgaacagtt cggctggcmc gagcccctga tgctctcgt  
 4021 ccagatcatc ctgatcgaca agaccggctt ccatccgagt acgtgctcmc tcgatcgat  
 4081 gtttcgctg gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc cgccgattg  
 4141 catcagccat gatggatact ttctcggcag gagcaagggt agatgacagg agatcctgcc  
 4201 ccggcacttc gcccaatagc agccagtccc ttcccgctc agtgacaacg tcgagcacag  
 4261 ctgcgcaagg aacgcccgtc gtggccagcc acgatagccc cgctgcctcg tcctgcagtt  
 4321 cattcagggc accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca  
 4381 gccggaacac ggcggcatca gagcagccga ttgtctgttg tgcccagtca tagccgaata  
 4441 g

//

**ZFP03 matched atc/rep on p15A/Kan**

LOCUS 22-5 4413 bp DNA circular

FEATURES Location/Qualifiers  
 CDS 1221..1964  
     /gene="RNAP alpha subunit"  
     /codon\_start="0"  
 rep\_origin complement(2740..3451)  
     /gene="p15A"  
 CDS 371..1081  
     /gene="mCherry"  
     /codon\_start="0"  
 CDS 1995..2528  
     /gene="ZFP03"  
     /codon\_start="0"  
 terminator 2605..2733  
     /gene="double term"  
 terminator 3542..3647  
     /gene="Terminator"  
 promoter 282..343  
     /gene="pwklac"  
 promoter 1140..1193  
     /gene="pLtetO-1"  
 terminator 206..251  
     /gene="term BBa\_B0011"  
 CDS 1971..1988

```

        /gene="linker"
        /codon_start="0"
prot_bind   258..275
        /gene="ZFP03 binding site"
prot_bind   350..364
        /gene="RBS"
CDS         complement(3673..4413)
        /gene="KanR"
        /codon_start="0"
terminator  2605..2733
        /gene="double term"
terminator  1088..1133
        /gene="term BBa_B0011"
terminator  2535..2580
        /gene="term BBa_B0011"
BASE COUNT 1138 a 1135 c 1097 g 1043 t
ORIGIN
1 cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa tcatgcgaaa
61 cgatcctcat cctgtctctt gatcagatca tgatcccctg cgccatcaga tccttggcgg
121 caagaaagcc atccagtta cttgacaggg ctccaacc ttaccagagg gcgccccagc
181 tggcaattcc gaattcatga gatctagaga atataaaaag ccagattatt aatccggctt
241 ttttattatt tggatctggt agacgtctag taactggatc tccgtattct ttacacttta
301 tgcttccggc tcgtatgtt tgcgaccga gcggataaca attggatcta ttaaagagga
361 gaaaggatct atgcgtaaag gagaagaaga taacatggct atcattaaag agttcatgcg
421 cttcaaagtt cacatggagg gttctgttaa cggtcacgag ttcgagatcg aaggcgaagg
481 cgagggccgt ccgtatgaag gcaccagac cgccaaactg aaagtgacta aaggcggccc
541 gctgcctttt gcgtgggaca tcctgagccc gcaatttatg tacggttcta aagcgtatgt
601 taaacacca gcggatatcc cggactatct gaagctgtct tttccggaag gtttcaagtg
661 ggaacgcgta atgaatttg aagatgggtg tgcctgacc gtcactcagg actcctcct
721 gcaggatggc gagttcatct ataaagttaa actgcgtggt actaattttc catctgatgg
781 cccggtgatg cagaaaaaga cgatgggttg ggaggcgtct agcgaacgca tgtatccgga
841 agatggtgcg ctgaaaggcg aaattaaaca gcgcctgaaa ctgaaagatg gcggccatta
901 tgacgctgaa gtgaaaacca cgtacaaagc caagaaacct gtgcagctgc ctggcgcgta
961 caatgtgaat attaaactgg acatcacctc tcataatgaa gattatacga tcgtagagca
1021 atatgagcgc gcggagggtc gtcattctac cgggtggcatg gatgaactat acaaataata
1081 aggatctaga gaatataaaa agccagatta ttaatccggc tttttatta tttggatctt
1141 ccctatcagt gatagagatt gacatcccta tcagtgatag agatactgag cacggatcta
1201 ttaaagagga gaaaggatct atgcaggggt ctgtgacaga gtttctaaaa ccgcgcctgg
1261 ttgatatcga gcaagtgagt tcgacgcacg ccaaggtgac ccttgagcct ttgagcgtg
1321 gctttggcca tactctgggt aacgcactgc gccgtattct gctctcatcg atgccgggtt
1381 gcgcggtgac cgaggtgag attgatggtg tactacatga gtacagcacc aaagaaggcg
1441 ttcaggaaga taccctggaa atcctgtca acctgaaagg gctggcggtg agagttcagg
1501 gcaaagatga agttattctt acctgaata aatctggcat tggccctgtg actgcagccc
1561 atatcaccca cgacggtgat gtcgaaatcg tcaagccgca gcacgtgatc tgccacctga
1621 ccgatgagaa cgcgtctatt agcatgcgta tcaaagtta cgcgcggtcgt ggttatgtgc
1681 cggcttctac ccgaattcat tcggaagaag atgagcggcc aatcggccgt ctgctggtgc

```

1741 acgcatgcta cagccctgtg gagcgtattg cctacaatgt tgaagcagcg cgtgtagaac  
1801 agcgtaccga cctggacaag ctggatcatcg aaatggaaac caacggcaca atcgatcctg  
1861 aagaggcgat tcgtcgtgcg gcaaccattc tggctgaaca actggaagct ttcgttgact  
1921 tacgtgatgt acgtcagcct gaagtgaag aagagaaacc agagggatct gccccgag  
1981 tccggaccgg atctctggaa ccaggatcta aaccgtacaa atgtccggaa tgtggtaaat  
2041 cttctccac tcatctggat ctgattcgtc atcaacgtac tcacactgga tctaaaccgt  
2101 acaaatgtcc ggaatgtggt aaatccttct cccaatctt tctctgggt cgatcaaac  
2161 gtactcacac tggatctaaa ccgtacaaat gtccggaatg tggtaaatcc ttctccaaa  
2221 actctaccct gactgaacat caacgtactc aactggatc taaaccgtac aaatgtccgg  
2281 aatgtggtaa atccttctcc tctcgtcgtc ctgtcgtgc acatcaactg actcacactg  
2341 gatctaaacc gtacaaatgt ccggaatgtg gtaaactctt ctccaactg gctcatctgc  
2401 gtgcacatca acgtactcac actggatcta aaccgtacaa atgtccggaa tgtggtaaat  
2461 cttctccac ttctggatc ctggtacgtc atcaacgtac tcacactgga tctaaaacct  
2521 ctaataagg atctagagaa tataaaaagc cagattatta atccggcttt ttattattt  
2581 ggatcctaac tcgagtaagg atctccaggc atcaataaaa acgaaaggct cagtcgaaag  
2641 actgggcctt tcgtttatc tgtgtttgt cggatgaacgc tctctactag agtcacactg  
2701 gtcaccttc ggggtggcct ttctgcgtt atacttaggg atatattccg cttctcgtc  
2761 cactgactcg ctacgtcgg tcgttcgact gggcgagcg gaaatggctt acgaaacggg  
2821 cggagatttc ctggaagatg ccaggaagat acttaacagg gaagtgaag ggccgaggca  
2881 aagccgttt tcataggct ccgccccct gacaagcatc acgaaatctg acgtcaaat  
2941 cagtggggc gaaacccgac aggactataa agataccagg cgtttcccc tggcggctcc  
3001 ctctgctcgt ctctgttcc tgccttcgg ttaccgggtg tcattccgct gttatggccg  
3061 cgtttgtctc attccagcc tgacactcag ttccgggtag gcagttcgtc ccaagctgga  
3121 ctgtatgcac gaacccccg ttactgcca ccgctgcgcc ttatccggtg actatcgtc  
3181 tgagtccaac ccggaagac atgcaaaagc accactggca gcagccactg gtaattgatt  
3241 tagaggagt agtctgaag tcatgcgcc gtaaggcta aactgaaagg acaagtttg  
3301 gtgactgctc tctccaagc cagttactc gttcaaaaga gttgtagct cagagaacct  
3361 tcgaaaacc gcctgcaag gcggttttt cgtttcaga gcaagagatt acgcgcagac  
3421 caaacgac tcaagaagat catcttatta atcagataaa atatttctag atttcagtgc  
3481 aatttatctc tcaaatgta gcacctgaag tcagccccat acgatataag ttgttactag  
3541 tgcttgatt ctaccaata aaaaacgcc ggcggcaacc gagcgttctg acaaatcca  
3601 gatggagttc tgaggtcatt actggatcta tcaacaggag tcaagcagc ctctgaacc  
3661 ccagagtcct gctcagaaga actcgtcaag aaggcagatg aaggcagatc gctcgaatc  
3721 gggagcggcg ataccgtaa gcacaggaa gcggtcagcc cattcggcg caagctctc  
3781 agcaatatca cggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggc  
3841 acagtcgatg aatccagaaa agcggccatt ttccacctg atattcggca agcaggcatc  
3901 gccatgggtc acgacgagat cctcggcgtc gggcatgctc gccttgagcc tggcgaacag  
3961 ttcggctggc gcgagccct gatgctctc gtccagatca tctgatcga caagaccggc  
4021 ttccatcca gtactgctc gctcgtgct atgttctgct tgggtgctga atgggcaggt  
4081 agccgatca agcgtatgca gccgccgat tgcacagcc atgatggata ctttctcggc  
4141 aggagcaagg tgagatgaca ggagatctg ccccgact tcgccaata gcagccagtc  
4201 cttccccgct tcagtgaca cgctgagcac agctgcgcaa ggaacgccc tcgtggccag  
4261 ccacgatagc cgcgtgctc cgtcctgag ttcattcagg gcaccggaca ggtcgtctt  
4321 gacaaaaaga accggcgcc cctcgtgctg cagccggaac accggcgcat cagagcagcc  
4381 gattgtctg tgtgccagt catagccgaa tag

//

## Reporter for TetR-sRNA inverter

LOCUS 24-29 2859 bp DNA circular

FEATURES Location/Qualifiers  
gene complement(2053..2847)  
/gene="Kan"  
CDS 234..941  
/gene="mCherry"  
/codon\_start="1"  
rep\_origin complement(1234..1916)  
/gene="ColE1 origin"  
terminator 972..1100  
/gene="double term"  
terminator 1922..2027  
/gene="Terminator"  
promoter 140..193  
/gene="pLtetO-1"  
misc\_feature 194..194  
/gene="transcriptional start site (+1)"

BASE COUNT 715 a 759 c 766 g 619 t

### ORIGIN

```
1 tcacctctg tcttgatcag atcatgatcc cctgcgcat cagatccttg gcggaagaa
61 agccatccag ttactttgc agggcttccc aacctacca gagggcgccc cagctggcaa
121 ttccgaattc atgagatctt ccctatcagt gatagagatt gacatcccta tcaagtatag
181 agatactgag cacggatcta ggaaaaagct catataacta gagtaagagg tcaatggttt
241 ccaagggcga ggaggataac atggctatca ttaaagagtt catgcgcttc aaagttcaca
301 tggaggggtc tgtaacggt cacgagttcg agatcgaagg cgaaggcgag ggccgtccgt
361 atgaaggcac ccagaccgcc aaactgaaag tgactaaagg cggcccgtcg cctttgctg
421 gggacatcct gagcccga tttatgtacg gttctaaagc gtatgttaa caccagcgg
481 atatcccgga ctatctgaag ctgtctttc cggaagggtt caagtgggaa cgcgtaatga
541 atttgaaga tgggtggtgc gtgaccgtca ctcaggactc ctcctgcag gatggcgagt
601 tcactataa agttaaactg cgtggtacta atttccatc tgatggccc gtgatgcaga
661 aaaagacgat gggttgggag gcgtctagcg aacgcatgta tccggaagat ggtgcgctga
721 aaggcgaat taaacagcgc ctgaaactga aagatggcgg ccattatgac gctgaagtga
781 aaaccacgta caaagccaag aaacctgtgc agctgcctgg cgcgtacaat gtgaatata
841 aactggacat cacctctcat aatgaagatt atacgatcgt agagcaatat gagcgcgagg
901 aggtctgca ttctaccggt ggcatggatg agctgtaca ataataagga tcctaactcg
961 agtaaggatc tccaggcatc aaataaaacg aaaggctcag tcgaaagact gggcctttcg
1021 tttatctgt tgtttgctcg tgaacgctct ctactagagt cacttggtc cacctcggg
1081 tggcctttc tgcgttata ctagggcgt tcggctgcgg cgagcggat cagctcactc
1141 aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc
1201 aaaaggccag caaaaggcca ggaaccgtaa aaaggccgag ttgctggcgt tttccatag
1261 gctccgccc cctgacgagc atcaaaaaa tcgacgctca agtcagaggt ggcgaaaccc
1321 gacaggacta taaagatacc aggcgtttcc cctggaagc tcctcgtgc gctctcctg
1381 tccgacctg ccgcttaccg gatacctgct gcctttctc ccttcgggaa gcgtggcgt
```



1441 ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg  
 1501 ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc ttatccgga actatcgtct  
 1561 tgagtccaac ccggaagac acgacttatc gccactggca gcagccactg gtaacaggat  
 1621 tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtgggtggc ctaactacgg  
 1681 ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta cttcggaaa  
 1741 aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt  
 1801 ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatctttc  
 1861 tacggggctct gacgctcagt ggaacgaaaa ctcacgtaa gggattttgg tcatgactag  
 1921 tgcttgatt ctcaccaata aaaaacgcc ggccggcaacc gagegttctg aacaaatcca  
 1981 gatggagttc tgaggtcatt actggatcta tcaacaggag tccaagcgag ctctcgaacc  
 2041 ccagagtccc gtcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc  
 2101 gggagcggcg atacctaaa gcacgaggaa gcggtcagcc cattcggcg caagctcttc  
 2161 agcaatatca cgggtagcca acgctatgct ctgatagcgg tccgccacac ccagccggcc  
 2221 acagtcgatg aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc  
 2281 gccatgggtc acgacgagat cctcggctc gggcatgcgc gccttgagcc tggcgaacag  
 2341 ttcggctggc gcgagcccct gatgctcttc gtccagatca tctgatcga caagaccggc  
 2401 ttccatccga gtacgtgctc gctcgtatgc atgtttcgtc tgggtggtcga atgggcaggt  
 2461 agccggatca agcgtatgca gccgccgat tgcattcagc atgatggata ctttctcggc  
 2521 aggagcaagg tgagatgaca ggagatcctg ccccggcact tcgccaata gcagccagtc  
 2581 cttcccctc tcagtgaaa cgtcagcac agctgcgcaa ggaacgcccg tcgtggccag  
 2641 ccacgatagc cgcgctgcct cgtcctgcag ttattcagg gcaccggaca ggtcggctt  
 2701 gacaaaaaga accgggccc cctcgcgtga cagccggaac acggcggcat cagagcagcc  
 2761 gattgtctgt tgtgccagt catagccgaa tagcctctcc acccaagcgg ccggagaacc  
 2821 tgcgtgcaat ccatcttgt caatcatgcg aaacgatcc

//

**Reporter for steric hindrance ( $P_{ZC35}$ )**

LOCUS pWH24-30 2927 bp DNA circular

FEATURES Location/Qualifiers  
 gene complement(2063..2857)  
     /gene="KanR"  
 rep\_origin complement(1244..1926)  
     /gene="ColE1"  
 terminator 982..1110  
     /gene="terminator"  
 terminator 1932..2037  
     /gene="Terminator"  
 CDS 195..899  
     /gene="mCherry"  
     /codon\_start="1"  
 gene 174..188  
     /gene="RBS"  
 promoter 82..167  
     /gene="PZC35"

```

misc_feature 147..147
      /gene="approximate transcriptional start site (+1)"
terminator 912..957
      /gene="terminator BBa_B0011"
BASE COUNT 752 a 749 c 798 g 628 t
ORIGIN
1 aaagccatcc agtttacttt gcagggcttc ccaaccttac cagagggcgc cccagctggc
61 aattccgaat tcatgagatc taaaaaaggg gaagagggga agaggagtgg tggacatagt
121 ggaaggaatg ggagatagtg ggagagaggg aaggagagga ggaaaaagga tctattaaag
181 aggagaaagg atctatgctg aaaggagaag aagataacat ggctatcatt aaagagtca
241 tgcgcttcaa agttcacatg gagggttctg ttaacggta cgagttcgag atcgaaggcg
301 aaggcgaggg ccgtccgtat gaaggcacc agaccgcaa actgaaagt actaaaggcg
361 gcccgtgcc ttttgcgtgg gacatcctga gcccgaatt tatgtacggt tctaaagcgt
421 atgttaaaca ccagcggat atcccggact atctgaagct gtctttccg gaaggttca
481 agtgggaacg cgtaatgaat tttgaagatg gtgggtcgt gaccgtcact caggactct
541 ccctgcagga tggcgagttc atctataaag ttaaactgcg tggactaat tttcatctg
601 atggcccggg gatgcagaaa aagacgatgg gttgggaggc gtctagcga cgcagtatc
661 cggaagatgg tgcgctgaaa ggcgaaatta aacagcgcct gaaactgaaa gatggcggcg
721 attatgacgc tgaagtgaaa accacgtaca aagccaagaa acctgtgcag ctgctggcg
781 cgtacaatgt gaatattaaa ctggacatca ccttcataa tgaagattat acgatcgtag
841 agcaatatga gcgcgaggag ggtcgtcatt ctaccggagg catggatgaa ctatacaat
901 aataaggatc tagagaatat aaaaagccag attattaatc cggctttttt attatttgg
961 tcctaactcg agtaaggatc tccaggcatc aaataaaacg aaaggctcag tcgaaagact
1021 gggcctttcg ttttatctgt tgtttgctgg tgaacgctct ctactagagt cacactggct
1081 caccttcggg tgggcctttc tgcgtttata ctagggcgt tcggctgcgg cgagcggat
1141 cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga
1201 acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt
1261 tttccatag gctccgccc cctgacgagc atcaaaaaa tcgacgctca agtcagaggt
1321 ggcgaaacc gacaggacta taaagatacc aggcgtttcc cctggaagc tcctcgtgc
1381 gctctctgt tccgacctg ccgcttaccg gatacctgtc cgccttttc ccttcgggaa
1441 gcgtggcgt ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgt
1501 ccaagctggg ctgtgtcac gaacccccg ttcagcccga ccgctgcgc ttatccggta
1561 actatcgtct tgagtccaac ccgtaagac acgacttacc gccactggca gcagccactg
1621 gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtcttg aagtggggc
1681 ctaactacgg ctactactaga aggacagtat ttggtatctg cgctctgctg aagccagta
1741 ccttcgaaa aagagttggt agctcttgat ccggcaaaaca aaccaccgct ggtagcggg
1801 gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt
1861 tgatctttc tacgggtct gacgctcagt ggaacgaaaa ctcacgtaa gggatttgg
1921 tcatgactag tgcttgatt ctaccaata aaaaacgcc ggcggcaacc gagcgttctg
1981 aacaaatcca gatggagttc tgaggtcatt actggatcta tcaacaggag tccaagcgag
2041 ctctgaacc ccagagtccc gctcagaaga actcgtcaag aaggcgatag aaggcgatgc
2101 gctgcgaatc gggagcggcg atacgtaaa gcacgaggaa gcggtcagcc cattcggcg
2161 caagctctc agcaatatca cgggtagcca acgctatgct ctgatagcg tccgccac
2221 ccagccggcc acagtcgatg aatccagaaa agcggccatt tccaccatg atattcgga
2281 agcaggcatc gccatgggtc acgacgagat cctcgccgtc gggcatgcgc gccttgagc
2341 tggcgaacag ttcggctggc gcgagccct gatgctcttc gtccagatca tctgatcga

```

```

2401 caagaccggc ttccatccga gtacgtgctc gctcgatgcg atgtttcgct tgggtggtcga
2461 atgggcaggt agccggatca agcgtatgca gccgccgcat tgcatacagcc atgatggata
2521 ctttctcggc aggagcaagg tgagatgaca ggagatcctg ccccgactcgcgccaata
2581 gcagccagtc cttcccgtc tcaagtacaa cgtcagacac agctgcgcaa ggaacgcccg
2641 tcgtggccag ccacgatagc cgcgctgctt cgtcctgcag ttattcagg gcaccggaca
2701 ggtcggctt gacaaaaaga accgggcgcc cctgcgctga cagccggaac acggcggcat
2761 cagagcagcc gattgtctgt tgtgccagt catagccgaa tagcctctcc accaagcgg
2821 ccggagaacc tgcgtgcaat ccatcttgtt caatcatgcg aaacgatcct catcctgtct
2881 cttgatcaga tcatgatccc ctgcgccatc agatccttgg cggcaag
//

```

***TetR and sRNA inverter (+sRNA)***

LOCUS 24-81 5517 bp DNA circular

```

FEATURES             Location/Qualifiers
    gene             complement(4418..5500)
                    /gene="lacI"
    CDS              1784..2500
                    /gene="GFPmut3b"
                    /codon_start="0"
    rep_origin       complement(2661..3372)
                    /gene="p15a"
    gene             complement(3584..4243)
                    /gene="KanR"
    CDS              1157..1777
                    /gene="tetR"
                    /codon_start="0"
    terminator       complement(6..299)
                    /gene="terminator"
    terminator       2526..2654
                    /gene="double term"
    terminator       3463..3568
                    /gene="terminator"
    gene             1041..1136
                    /gene="repC miniCistron"
    misc_feature     854..944
                    /gene="UT sense RNA"
    misc_feature     complement(321..411)
                    /gene="sRNA"
    promoter         800..853
                    /gene="pLlacO-1"
    prot_bind        1143..1150
                    /gene="RBS"
    promoter         complement(412..446)
                    /gene="BBa_J23102"

```

BASE COUNT 1508 a 1327 c 1290 g 1392 t

ORIGIN

1 actctgcagt tcctactct cgcattgggga gacccacac taccatcggc gctacggcgt  
61 ttacttctg agttcggcat ggggtcaggt gggaccaccg cgctactgcc gccaggcaaa  
121 ttctgttta tcagaccgct tctgcgttct gatttaatct gtatcaggct gaaaatcttc  
181 tctcatccgc caaacagcc aagctggaga ccgtttaaac tcaatgatga tgatgatgat  
241 ggtcgacggc gctattcaga tcctctctg agatgagttt ttgttcgggc ccaagcttct  
301 cactactcga gtaaggatcc aaataaaaag gagtcgctct gtcctcgc aaagttgcag  
361 aacgacatca tcaaaagaaa aaaacactga gttgtttta taatcttgta tgctagcaca  
421 gtacctagga ctgagctagc tgtcaactct agtacttccg ggcgctatca tgccataccg  
481 cgaaggttt tgcgccattc gatggtgtcc gggatctcga cgctctcct tatgcgactc  
541 ctgcattagg aagcagccca gtagtaggtt gaggccgtt agcaccgccg ccgcaaggaa  
601 tgggtgatgc aaggagatgg cgcccaacag tccccggcc acggggcctg ccacataacc  
661 cacgccgaaa caagcgtca tgagcccga gggcgagcc cgatcttccc catcggatgat  
721 gtcggcgata taggcgccag caaccgcacc tgtggcgccg gtgatgccgg ccacgatcgc  
781 tccggcgtag aggatcgaga attgtgagcg gataacaatt gacatttgta gcggataaca  
841 agatactgag cacaataaaa aaggagtcgc tctgtccctc gccaaagtg cagaacgaca  
901 tcattcaag aaaaaaac tgagttgtt ttataatct gtatattag atattaaacg  
961 atattaaat atacataaag atatatatt gggtgagcga ttccttaac gaaattgaga  
1021 ttaaggagtc gctcttttt atgtataaaa acaatcatgc aatcattca aatcattgg  
1081 aaaatcacga ttagacaat ttttcaaaa ccggctactc taatagccgg ttgtaaggat  
1141 ctaggaccaa ggatctatgt ccagattaga taaaagtaaa gtgattaaca gcgcattaga  
1201 gctgcttaat gaggtcggaa tcgaaggtt aacaaccgt aaactcggcc agaagctagg  
1261 tgtagagcag cctacattgt attggcatgt aaaaaataag cgggctttgc tcgacgcctt  
1321 agccattgag atgtagata ggaccatac tctttttgc ctttagaag gggaaagctg  
1381 gcaagatttt ttacgtaata acgtaaaaag ttttagatgt gctttactaa gtcacgcga  
1441 tggagcaaaa gtacatttag gtacacggcc tacagaaaaa cagtatgaaa ctctgaaaa  
1501 tcaattagcc ttttatgcc aacaaggtt ttactagag aatgcattat atgcactcag  
1561 cgctgtgggg cattttact taggttgcgt attggaagat caagagcatc aagtcgctaa  
1621 agaagaaagg gaaacaccta ctactgatag tatgccgcca ttattacgac aagctatcga  
1681 attattgat caccaaggtg cagagccagc ctcttattc ggccttgaat tgatcatatg  
1741 cggattagaa aaacaactta aatgtgaaag tgggtccgga tctatgagta aaggagaaga  
1801 acttttact ggagttgtcc caattctgt tgaattagat ggtgatgta atgggcacaa  
1861 attttctgtc agtggagagg gtgaaggtga tgcaacatac ggaaaactta ccctaaatt  
1921 tatttgact actggaaaac tacctgttcc atggcaaca cttgtcacta cttgactta  
1981 tgggtttcaa tgctttcaa gataccaga tcatatgaaa cggcatgact tttcaagag  
2041 tgccatgcc gaaggtatg tacaggaaag aactatatt tcaaaagatg acgggaacta  
2101 taagacacgt gctgaagtca agtttgaagg tgatacactt gtaaatagaa tcgagttaaa  
2161 aggtattgat ttaaagaag atggaaacat tcttgacac aagttggaat acaactataa  
2221 ctacacaaat gtatacatca tggcagacaa acaaaagaat ggaatcaag ttaactcaa  
2281 aattagacac aacattgaag atggaagcgt tcaactagca gaccattatc acaaaaatac  
2341 tccaattggc gatggcctg tcctttacc agacaacat tacctgtcca cacaatctgc  
2401 ctttcgaaa gatccaacg aaaagagaga ccacatggc cttcttgagt ttgtaacagc  
2461 tgccgggatt acacatggca tggatgaact atacaaatag ggatccaaac tcgagtaagg  
2521 atctccaggc atcaataaaa acgaaaggct cagtcgaaag actgggcctt tcgtttatc  
2581 tgtgtttgt cggtgaacgc tctctactag agtcacactg gtcacctc gggtgggcct

2641 ttctgcgttt atacctaggg atatattccg cttcctcgct cactgactcg ctacgctcgg  
2701 tcgttcgact gcggcgagcg gaaatggctt acgaacgggg cggagatttc ctggaagatg  
2761 ccaggaagat acttaacagg gaagtgagag ggccgcggca aagccgtttt tccataggct  
2821 ccgccccct gacaagcatc acgaaatctg acgctcaaat cagtggtaggc gaaacccgac  
2881 aggactataa agataaccagg cgtttcccc tggcggctcc ctcgtgcgct ctctgttcc  
2941 tgcctttcgg tttaccggtg tcattccgct gttatggccg cgtttgtctc attccacgcc  
3001 tgacactcag ttccgggtag gcagttcgtc ccaagctgga ctgtatgac gaacccccg  
3061 ttcagtccga ccgctgcgcc ttatccggtg actatcgtct tgagtccaac ccggaagac  
3121 atgcaaaagc accactggca gcagccactg gtaattgatt tagaggagtt agtcttgaag  
3181 tcatgcgccc gttaaggcta aactgaaagg acaagtttg gtgactgcgc tctccaagc  
3241 cagttacctc ggtcaaaga gttgtagct cagagaacct tcgaaaacc gcctgcaag  
3301 gcggtttttt cgttttcaga gcaagagatt acgcgcagac caaacgac tcaagaagat  
3361 catcttata atcagataa atatttctag atttcagtgc aatttatctc ttcaaatgta  
3421 gcacctgaag tcagccccat acgatataag ttgttactag tgcttgatt ctaccaata  
3481 aaaaacgccc ggcggcaacc gagcgttctg aacaatcca gatggagttc tgaggtcatt  
3541 actggatcta tcaacaggag tccaagcgag ctcgatatca aattacgccc gccttgcca  
3601 ctatcgcag tactgttga attcattaag cattctgccg acatggaagc catcacaac  
3661 ggcattgatga acctgaatcg ccagcggcat cagcaccttg tcgccttgcg tataatattt  
3721 gccatggtg aaaacggggg cgaagaagt gtccatattg gccacgtta aatcaaaact  
3781 ggtgaaactc acccagggat tggctgagac gaaaaacata ttctcaata acccttagg  
3841 gaaataggcc aggtttcac cgtaacacgc cacatctgc gaatatatgt gtagaaactg  
3901 ccgaaatcg tcgtggtatt cactccagag cgatgaaaac gtttcagttt gctcatggaa  
3961 aacggtgtaa caagggtgaa cactatccca taccaccg tcaccgtctt tcatggcat  
4021 acgaaatcc ggatgagcat tcatcaggcg ggcaagaatg tgaataaagg ccgataaaa  
4081 ctgtgctta ttttctta cggctttaa aaaggccgta atatccagct gaacggtctg  
4141 gttataggta cattgagcaa ctgactgaaa tgcctcaaaa tgttcttac gatgccattg  
4201 ggatataatca acggtggtat atccagtgat tttttctcc atttagctt ccttagctc  
4261 tgaatctc gataactcaa aaaatacgc cggtagtat cttatttcat tatggtgaa  
4321 gttggaacct cttacgtgcc gatcaacgtc tcattttcgc cagatcga cgctcgtgcc  
4381 taatgagtga gctaactac attaattgcg ttgcgctcac tgcccgttt ccagtcggga  
4441 aacctgtcgt gccagctgca ttaatgaat ggccaacgcg cggggagagg cggttgctg  
4501 attggcgccc aggtggtt tctttcac cagtgagac ggcaacagct gattgcctt  
4561 caccgctgg ccctgagaga gttgagcaa gcggtccacg ctggttgcc ccagcaggcg  
4621 aaaatctgt ttgatggtg ttaacggcgg gatataacat gagctgtctt cggtatcgtc  
4681 gtatccact accgagatgt ccgaccaac gcgcagccc gactcgtaa tggcgcgct  
4741 tgcgccagc gccatctgat cgttggaac cagcatcga gtgggaacga tgcctcatt  
4801 cagcattgc atggttgtt gaaaaccgga catggcactc cagtcgctt ccgctcgc  
4861 taccgctga attgattgc gactgagata tttatgccag ccagccagac gcagacgcg  
4921 cgagacagaa ctaatgggc ccgtaacag gcgatttgc tggtagacca atgcgaccg  
4981 atgctccag ccagtcgcg taccgtctc atgggagaaa ataactgt tgatgggtg  
5041 ctggtcagag acatcaagaa ataacgccg aacattagt caggcagctt ccacagcaat  
5101 ggcattctg tcattcagcg gatagtaat gatcagcca ctgacgcgtt gcgcgagaag  
5161 attgtgacc gccgcttac aggtctgac gccgctcgt tctaccatg acaccacc  
5221 gctggcacc agttgatcgg cgcgagattt aatcggcg acaattgcg acggcgcgtg  
5281 caggccaga ctggaggtg caacccaat cagcaacgac tgttgccc ccagttgtg  
5341 tgccacgccc ttgggaatg aattcagctc cgccatgcc gcttccatt tttcccgt

5401 tttcgagaa acgtggctgg cctggttcac cacgcgggaa acggtctgat aagagacacc  
5461 ggcatactct gcgacatcgt ataacgttac tggtttcaca ttcaccacc tgaattg  
//

**RFP-ZFP fusion expression from  $P_{lacUV5}$**

LOCUS 26-51 4700 bp DNA circular  
DEFINITION 4160 bp DNA circular pBbA5c-RFP  
pBbA5c-RFP.1  
.

holtz

FEATURES Location/Qualifiers  
CDS complement(3323..4405)  
/gene="LacI"  
/codon\_start="0"  
rep\_origin complement(1566..2277)  
/gene="p15a"  
CDS 189..866  
/gene="RFP"  
/codon\_start="0"  
misc\_feature 1431..1559  
terminator 2368..2473  
/gene="Terminator"  
promoter 74..149  
/gene="PlacUV5"  
prot\_bind 169..188  
/gene="RBS"  
terminator 1431..1559  
/gene="double term"  
gene complement(2489..3148)  
/gene="CmR"  
CDS 873..1400  
/gene="ZFP-65"  
/codon\_start="1"  
CDS 189..1400  
/gene="RFP-ZFP fusion"  
/codon\_start="0"  
BASE COUNT 1188 a 1244 c 1133 g 1135 t  
ORIGIN

1 gatataggcg ccagcaaccg cacctgtggc gccggtgatg ccggccacga tgcgtccggc  
61 gtagaggatc gagatcggtt aggcaccca ggcttacac tttatgcttc cggctcgtat  
121 aatgtgtgga attgtgagcg gataacaatt tcagaattca aaagatcttt taagaaggag  
181 atatacatat ggcgagtagc gaagacgta tcaaagagtt catgcgttc aaagttcgtg

241 tggaaaggttc cgtaaacggt cacgagttcg aaatcgaagg tgaaggtgaa ggtcgtccgt  
301 acgaaggtac ccagaccgct aaactgaaag ttaccaaagg tgggccgctg ccgttcgctt  
361 gggacatcct gtccccgag ttccagtacg gttccaaagc ttacgttaaa cacccggctg  
421 acatcccga ctacctgaaa ctgtccttcc cggaaggttt caaatgggaa cgtgttatga  
481 acttcgaaga cggtggtggt gttaccgta cccaggactc ctccctgcaa gacggtgagt  
541 tcattctaaa agttaaactg cgtggtacca acttcccgtc cgacgggtccg gttatgcaga  
601 aaaaaacat gggttgggaa gcttccaccg aacgtatgta cccggaagac ggtgctctga  
661 aaggtgaaat caaatgcgt ctgaaactga aagacgggtg tccactacgac gctgaagtta  
721 aaaccaccta catggctaaa aaaccgggtc agctgccggg tgcttataaa accgacatca  
781 aactggacat cacctcccac aacgaagact acaccatcgt tgaacagtac gaacgtgctg  
841 aaggtcgtca ctccaccggt gcttaaggat ctctggaacc aggatctaaa ccgtacaaat  
901 gtccggaatg tggtaaactc ttctcccgtt ctgataaact ggctcgtcat caacgtactc  
961 aactggatc taaaccgtac aaatgtccgg aatgtggtaa atccttctcc cgcttctgata  
1021 atctggtacg tcatcaactg actcactg gatctaaacc gtacaaatgt ccggaatgtg  
1081 gtaaactcct ctcccaatct tccaatctgg ttcgtcatca acgtactcac actggatcta  
1141 aaccgtacaa atgtccgga tgtggtaaat cttctcccg ttctgataaa ctggttcgtc  
1201 atcaactgac tcactgga tctaaaccgt acaaatgtcc ggaatgtggt aaatccttct  
1261 cccaactgac aaatctgctg gctcatcaac gtactcacac tggatctaaa ccgtacaaat  
1321 gtccggaatg tggtaaactc ttctccaac gtgcaaatct gcgtgctcat caacgtactc  
1381 aactggatc taaaactct taataaggat cctgactcga gtaaggatct ccaggcatca  
1441 aataaacga aaggctcagt cgaaagactg ggccttctg tttatctggt gtttgcggt  
1501 gaacgtctc tactagagtc aactggctc acctcgggt gggccttct gcgttatac  
1561 ctaggatata attccgttc ctgctcact gactcgtac gctcggctg tgcactgagg  
1621 cgagcggaaa tggcttacga acggggcgga gatttctg aagatgccag gaagatactt  
1681 aacagggag tgagagggc gcggcaaagc cgttttcca taggctccgc cccctgaca  
1741 agcatcacga aatctgacg tcaaatcagt ggtggcgaaa cccgacagga ctataaagat  
1801 accaggcgtt tcccctggc ggctccctc tgcgcttcc tttctctgcc tttcggttta  
1861 ccggtgcat tccgtgta tggccgctt tctctctc cacgcctgac actcagttcc  
1921 gggtaggcag ttcgtccaa gctggactg atgcacgaac ccccggtca gtcgaccgc  
1981 tgcgccttat ccgtaacta tctcttgag tccaaccgg aaagacatgc aaaagcacca  
2041 ctggcagcag cactggtaa tgattaga ggagtagtc tgaagtcac gcgccggtta  
2101 aggtctaaact gaaaggaca gtttgggtga ctgcgtcct ccaagccagt tacctcgggt  
2161 caaagagttg gtagctcaga gaacctcga aaaaccgccc tgcaaggcgg tttttcgtt  
2221 tcagagcaa gagattacg gcagaccaa acgatctca gaagatcac ttattaatca  
2281 gataaaatat tttagattt cagtcaatt tatcttca aatgtagcac ctgaagtcag  
2341 cccatacga tataagttg tactagtgt tggatttca ccaataaaaa acgcccggcg  
2401 gcaaccgagc gttctgaaca aatccagatg gagttctgag gtcattactg gatctataa  
2461 caggagtcca agcagctc atatacaatt acgcccccc ctgacctca tgcagctact  
2521 gttgtaattc attagcatt ctgccgacat ggaagccatc acaaacggca tgatgaacct  
2581 gaatgccag cggcatcagc acctgtcgc cttgcgtata atattgccc atggtgaaaa  
2641 cggggggcga gaagttgtcc atattggcca cgtttaaact aaaactggtg aaactcacc  
2701 agggattggc tgagacgaaa acaatattct caataaacc ttagggaaa taggccaggt  
2761 tttaccgta acagccaca tcttgcaat atatgttag aaactgccg aaatcgtct  
2821 ggtattact ccagagcgt gaaaacgtt cagttgctc atggaaaac gtgtaacaag  
2881 ggtgaact atccatc accagctc cgtcttcat tgccatacga aattccgat  
2941 gagcattcat caggcgggca agaattgaa taaaggccg ataaaactg tgcttattt

3001 tctttacggt ctttaaaaag gccgtaatat ccagctgaac ggtctggta taggtacatt  
 3061 gagcaactga ctgaaatgcc tcaaatggt ctttacgatg ccattgggat atatcaacgg  
 3121 tggatatcc agtgatTTTT tctccattt tagcttcctt agctcctgaa aatctcgata  
 3181 actcaaaaaa tacgcccggg agtgatctta tttcattatg gtgaaagtg gaacctctta  
 3241 cgtgccgatc aacgttcat ttcgccaga tatcgacgtc ggtgcctaag gagtgagcta  
 3301 acttacatta attgcgttgc gtcactgcc cgctttccag tcgggaaacc tgtcgtgcca  
 3361 gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg ggcgccaggg  
 3421 tggTTTTtct ttcaccagt gagacgggca acagctgatt gccctcacc gcctggcctt  
 3481 gagagagttg cagcaagcgg tccacgtggt ttgccccag caggcgaaaa tctgtttga  
 3541 tggTggtaa cggcgggata taacatgagc tgtcttcggt atcgtcgtat cccactaccg  
 3601 agatgtccgc accaacgcgc agcccggact cgtaaatggc gcgcattgag cccagcgcca  
 3661 tctgatcgtt ggcaaccagc atcgcagtgg gaacgatgcc ctcattcagc atttgatgg  
 3721 tttgttgaag accggacatg gcaactcagt cgcttcccg ttccgctatc ggctgaattt  
 3781 gattgcgagt gagatatta tgccagccag ccagacgcag acgcgccgag acagaactta  
 3841 atgggcccgc taacagcgcg atttgctggt gaccaatgc gaccagatgc tccacgcca  
 3901 gtcgcgtacc gtcttcatgg gagaaaataa tactgttgat gggTgtctgg tcagagacat  
 3961 caagaaataa cgccggaaca ttagtgagg cagcttcac agcaatggca tctgtgtcat  
 4021 ccagcggata gttaatgac agcccactga cgcgttgcgc gagaagattg tgcaccgccc  
 4081 ctttacaggc ttcgacgcc cttcgttcta ccatcgacac caccacgtg gcaccagtt  
 4141 gatcggcgcg agatttaac gccgcgaca tttgcgacgg cgcgtgcagg gccagactgg  
 4201 aggtggcaac gccaatcagc aacgactgtt tgcccgccag ttgtgtgcc acgcggttgg  
 4261 gaatgaatt cagctccgc atcgcgcgtt cacttttc ccgcgtttc gcagaaacgt  
 4321 ggctggcctg gttcaccag cgggaaacgg tctgataaga gacaccggca tactctgcga  
 4381 catcgataa cgttactggt ttacattca ccaccctgaa ttgactctct tccgggcgct  
 4441 atcatgcat accgcgaaag gttttgcgc attcgtatgg gtccgggatc tcgacgctt  
 4501 cccttatgag actcctgcat taggaagcag cccagtagta ggttgaggcc gttgagcacc  
 4561 gccgccgcaa ggaatgggtg atgcaaggag atggcgcca acagtcccc ggccacgggg  
 4621 cctgccacca taccacgcc gaacaagcg ccatgagcc cgaagtggcg agcccgatct  
 4681 tcccacgg tgatgcggc

//

**TetR inverter (-sRNA)**

LOCUS 26-58 5391 bp DNA circular

FEATURES Location/Qualifiers  
 gene complement(3767..4849)  
 /gene="lacI"  
 CDS 1133..1849  
 /gene="GFPmut3b"  
 /codon\_start="0"  
 rep\_origin complement(2010..2721)  
 /gene="p15a"  
 gene complement(2933..3592)  
 /gene="KanR"  
 CDS 506..1126



```

        /gene="tetR"
        /codon_start="0"
terminator   complement(4872..5165)
        /gene="terminator"
terminator   1875..2003
        /gene="double term"
terminator   2812..2917
        /gene="terminator"
gene         390..485
        /gene="repC miniCistron"
misc_feature 203..293
        /gene="UT sense RNA"
promoter    149..202
        /gene="pLlacO-1"
prot_bind   492..499
        /gene="RBS"
BASE COUNT  1464 a 1301 c 1266 g 1360 t
ORIGIN
1  caccataccc acgccgaaac aagcgctcat gagcccgaag tggcgagccc gatcttccc
61  atcggatgatg tcggcgatat aggcgccagc aaccgcacct gtggcgccgg tgatgccggc
121  cacgatgcgt ccggcgtaga ggatcgagaa ttgtgagcgg ataacaattg acattgtgag
181  cggataacaa gatactgagc acaataaaaa aggagtcgct ctgtccctcg ccaaagttgc
241  agaacgacat cattcaaaga aaaaaaact gagttgttt tataatcttg tatatttaga
301  tattaaacga tatttaaata tacataaaga tatatattg ggtgagcgt tccttaaacg
361  aaattgagat taaggagtcg ctcttttta tgtataaaaa caatcatgca aatcattcaa
421  atcatttga aaatcacgat ttagacaatt tttctaaaac cggctactct aatagccggt
481  tgtaaggatc taggaccaag gatctatgtc cagattagat aaaagtaaag tgattaacag
541  cgcattagag ctgcttaatg aggtcggaat cgaaggttta acaaccgta aactcgccca
601  gaagctaggt gtagagcagc ctacattgta ttggcatgta aaaaataagc gggctttgct
661  cgacgcctta gccattgaga tgtagatag gcaccatact cacttttgcc cttagaagg
721  ggaaagctgg caagatttt tacgtaataa cgctaaaagt ttagatgtg cttactaag
781  tcatcgcat ggagcaaaag tacatttagg tacacggcct acagaaaaac agtatgaaac
841  tctcggaaat caattagcct tttatgcc acaaggtttt ttagatgaga atgcattata
901  tgcactcagc gctgtggggc atttacttt aggttcgta ttggaagatc aagagcatca
961  agtcgctaaa gaagaaagg aaacacctac tactgatagt atgccccat tattacgaca
1021  agctatcgaa ttattgatc accaaggtgc agagccagcc ttctattcg gccttgaatt
1081  gatcatatgc ggattagaaa aacaactta atgtgaaagt ggtccggat ctatgagtaa
1141  aggagaagaa ctttactg gagttgtccc aattctgtt gaattagatg gtgatgtaa
1201  tgggcacaaa tttctgtca gtggagaggg tgaaggatg gcaacatacg gaaaacttac
1261  cttaaattt atttgacta ctggaaaact acctgttcca tggccaacac ttgtcactac
1321  tttgacttat ggtgttcaat gcttttaag ataccagat catatgaaac ggcattgact
1381  ttcaagagt gcatgcccg aaggttatgt acaggaaaga actatatttt tcaaagatga
1441  cgggaactat aagacacgtg ctgaagtaa gttgaagg gatacacttg ttaatagaat
1501  cgagttaaaa ggtattgatt taaagaaga tggaaacatt cttggacaca agttggaata
1561  caactataac tcacacaatg tatacatcat ggagacaaa caaagaatg gaatcaaagt
1621  taacttcaaa attagacaca acattgaaga tggaaagcgtt caactagcag accattatca

```

1681 acaaaataact ccaattggcg atggccctgt cttttacca gacaaccatt acctgtccac  
1741 acaatctgcc ctttcgaaag atcccaacga aaagagagac cacatgggcc ttcttgagtt  
1801 tgtaacagct gccgggatta cacatggcat ggatgaacta tacaatagg gatccaaact  
1861 cgagtaagga tctccaggca tcaataaaaa cgaaggctc agtcgaaaga ctgggccttt  
1921 cgttttatct gttgtttgc ggtgaacgt ctctactaga gtcactggt ctcaccttcg  
1981 ggtgggcctt tctgcgtta tacctaggga tatattccgc ttctcgtc actgactgc  
2041 tacgctcgtt cgctcgactg cggcgagcgg aaatggctta cgaacggggc ggagattcc  
2101 tggaagatgc caggaagata ctaacaggg aagtgagagg gccgcggcaa agccgtttt  
2161 ccataggctc cgccccctg acaagcatca cgaaatctga cgctcaaac agtggtgagg  
2221 aaaccgaca ggactataaa gataccaggc gttccccct gccggctccc tcgtcgtc  
2281 tctgttctt gccttcggt ttaccggtt cattccgctg ttatggccgc gttgtctca  
2341 ttccacgct gacactcagt tccgggtagg cagttcgtc caagctggac tgatgcacg  
2401 aacccccgt tcagtccgac cgctgcgct tatccggtta ctatcgtt gagtccaac  
2461 cggaaagaca tgcaaaagca cactggcag cagccactgg taattgatt agaggagta  
2521 gtcttgaagt catgcgccg ttaaggctaa actgaaagga caagtttgg tgactcgtc  
2581 cctcaagcc agttacctc gttcaagag ttggtagctc agagaacct cgaaaaaccg  
2641 ccctgaagg cggtttttc gtttcagag caagagatta cgcgcagacc aaaacgatc  
2701 caagaagatc atcttataa tcagataaaa ttttctaga tttcagtca atttatctt  
2761 tcaaatgtag cacctgaagt cagccccata cgatataagt tttactagt gcttgattc  
2821 tcaccaataa aaaacgccc gggcaaccg agcgttctga acaatccag atggagtct  
2881 gaggtcatta ctggatctat caacaggagt ccaagcagc tcgatatcaa attaccccc  
2941 gccctgccac tcatcgact actgtttaa ttcattaagc attctgccga catggaagc  
3001 atcaaaacg gcatgatgaa cctgaatgc cagcggcatc agcacctgt cgccttgcg  
3061 ataattttg cccatggtg aaacgggggc gaagaagtt tccatattg ccacgttaa  
3121 atcaaaactg gtgaaactca ccagggtt ggctgagacg aaaaacatat tctcaataa  
3181 cccttaggg aataggcca ggtttcacc gtaacacgcc acatcttgc aatatatgt  
3241 tagaaactgc cggaaatcgt cgtggtatt actccagagc gatgaaaacg tttcagttg  
3301 ctcatgaaa acggtgtaac aagggtgta actatccat atcaccagc caccgtctt  
3361 cattgcata cgaaattcc gatgagcatt catcaggcgg gcaagaatg gaataaaggc  
3421 cggataaaac ttgtgcttat ttttcttac ggtctttaa aaggccgtaa tatccagctg  
3481 aacggtctgg ttataggta attgagcaac tgaactgaaat gcctcaaat gttctttacg  
3541 atgccattg gatatacaa cgggtgtata tccagtatt ttttctca ttttagctc  
3601 cttagctcct gaaaatctc ataactcaaa aaatagccc ggtagtgatc ttatttcatt  
3661 atggtgaaag ttggaacctc ttactgccc atcaactct cattttccc agatctgac  
3721 gtcggtgct aatgagtgag ctaactaca ttaattgct tgcgctcact gcccgttc  
3781 cagtcgggaa acctgctg ccagctgatc taatgaatc gccaacgcg ggggagaggc  
3841 ggtttgcgta tgggcgcca ggtggttt tctttcacc agtgagacgg gcaacagctg  
3901 attgccctc accgctggc cctgagagag tgcagcaag cggccacgc tggttgccc  
3961 cagcaggcga aaatctggt tgatggtggt taacggcggg atataacatg agctgtctc  
4021 ggtatcgtc tatccacta ccgagatgc gcaccaacg cgcagcccgg actcgtaat  
4081 ggcgcgcat gcgccagc ccatctgatc gttggcaacc agcatcgcag tgggaacgat  
4141 gccctcattc agcatttga tggttgtg aaaaccggac atggcactc agtcgcctc  
4201 ccgtccgct atcggctgaa ttgattgag agtgagatat ttatgccagc cagccagacg  
4261 cagacgcgc gagacagaac ttaatgggccc cgtaacagc gcgatttgc ggtgaccaa  
4321 tgcgaccaga tgctccagc ccagtcgct accgtctca tgggagaaaa taactgtt  
4381 gatgggtgct tggctagaga catcaagaaa taacgccga acattagtgc aggcagctc

4441 cacagcaatg gcatcctggt catccagcgg atagttaatg atcagcccac tgacgcgttg  
 4501 cgcgagaaga ttgtgcaccg ccgctttaca ggcttcgacg ccgcttcggt ctaccatcga  
 4561 caccaccacg ctggcaccca gttgatcggc gcgagattta atcgccgcga caatttgca  
 4621 cggcgcgtgc agggccagac tggaggtggc aacgccaatc agcaacgact gtttggcccg  
 4681 cagttgtgt gccacgcggt tgggaatgta attcagctcc gccatcgccg ctccacttt  
 4741 ttcccgcgtt ttcgagaaa cgtggctggc ctggttcacc acgcgggaaa cggctgata  
 4801 agagacaccg gcatactctg cgacatcgta taacgttact ggttcacat tcaccacct  
 4861 gaattgactc tgcagtccc tactctcgca tggggagacc ccactactacc atcggcgcta  
 4921 cggcgtttca ctctgagtt cggcatgggg tcaggtggga ccaccgcgct actgccgcca  
 4981 ggcaaattct gtttatcag accgcttctg cgttctgatt taatctgat caggctgaaa  
 5041 atcttctctc atccgcaaaa acagccaagc tggagaccgt ttaaactcaa tgatgatgat  
 5101 gatgatggtc gacggcgcta ttcagatcct ctctgagat gagttttgt tggggccaa  
 5161 gcttctcatt actcgagtaa ggatcctct agtacttccg ggcgctatca tgccataccg  
 5221 cgaaagggtt tgcgccattc gatggtgtcc gggatctcga cgctctcct tatgcgactc  
 5281 ctgcattagg aagcagccca gtagtaggtt gaggccgttg agcaccgccc cgcaaggaa  
 5341 tggatcatgc aaggagatgg cgccaacag tccccggcc acggggcctg c

//

**ZFP expression from  $P_{lacUV5}$  for steric hindrance experiments**

LOCUS 27-14 4019 bp DNA circular

FEATURES Location/Qualifiers  
 CDS complement(2642..3724)  
     /gene="LacI"  
     /codon\_start="0"  
 rep\_origin complement(885..1596)  
     /gene="p15a"  
 misc\_feature 750..878  
 terminator 1687..1792  
     /gene="Terminator"  
 promoter 74..149  
     /gene="PlacUV5"  
 prot\_bind 169..188  
     /gene="RBS"  
 terminator 750..878  
     /gene="double term"  
 gene complement(1808..2467)  
     /gene="CmR"  
 CDS 189..719  
     /gene="ZFP-65"  
     /codon\_start="1"

BASE COUNT 998 a 1071 c 963 g 987 t

ORIGIN

1 gatataggcg ccagcaaccg cacctgtggc gccggtgatg ccggccacga tgcgtccggc  
 61 gtagaggatc gagatcgttt aggcaccca ggctttacac tttatgcttc cggctcgtat

121 aatgtgtgga attgtgagcg gataacaatt tcagaattca aaagatcttt taagaaggag  
181 atatacatat gctggaacca ggatctaaac cgtacaaatg tccggaatgt ggtaaactct  
241 tctcccgttc tgataaactg gttcgtcatc aacgtactca cactggatct aaacctgaca  
301 aatgtccgga atgtggtaaa tccttctccc gttctgataa tctggtagct catcaacgta  
361 ctcacactgg atctaaaccg tacaatgtc cggaatgtgg taaatccttc tcccaatctt  
421 ccaatctggt tcgtcatcaa cgtactcaca ctggatctaa accgtacaaa tgtccggaat  
481 gtggtaaadc cttctcccgt tctgataaac tggttcgtca tcaacgtact cacactggat  
541 ctaaaccgta caaatgtccg gaatgtggta aatccttctc ccaacgtgca aatctgctg  
601 ctcatcaacg tactcacact ggatctaaac cgtacaaatg tccggaatgt ggtaaactct  
661 tctccaacg tgcaaatctg cgtgctcatc aacgtactca cactggatct aaaacctctt  
721 aataaggatc ctgactcgag taaggatctc caggcatcaa ataaaacgaa aggctcagtc  
781 gaaagactgg gccttctgtt ttatctgttg tttgtcggtg aacgtctct actagagtca  
841 cactggctca ccttcgggtg ggccttctg cgttataacc taggatata ttccgcttc  
901 tcgctactg actcgtacg ctcggtcgtt cgactgcggc gagcggaaat ggcttacgaa  
961 cggggcggag atttctgga agatgccagg aagatactta acagggagt gagagggccc  
1021 cggcaaagcc gttttccat aggctccgc ccctgacaa gcatcacgaa atctgacgct  
1081 caaatcagtg gtggcgaac cgcacaggac tataaagata ccaggcgttt cccctggcg  
1141 gctccctcgt gcgtctctt gttctgcct ttcggtttac cgggtcatt ccgctgttat  
1201 ggccgcgttt gtctcattc acgctgaca ctcagttccg ggtaggcagt tcgctccaag  
1261 ctggactgta tgcacgaacc cccggtcag tccgaccgct gcgccttacc cggtaactat  
1321 cgtcttgagt ccaaccgga aagacatgca aaagcaccac tggcagcagc cactggtaat  
1381 tgatttagag gagttagtct tgaagcatg cgccgggtta ggctaaactg aaaggacaag  
1441 ttttggtagc tgcgtctc caagccagtt acctcggctc aaagagttgg tagctcagag  
1501 aaccttcgaa aaaccgccct gcaaggcggg ttttctgtt tcagagcaag agattacgag  
1561 cagacaaaa cgatctcaag aagatcatct tattaatcag ataaaatatt tctagatttc  
1621 agtgcaattt atctctcaa atgtagcacc tgaagtcagc cccatagat ataagttgtt  
1681 actagtgtt ggattctc caataaaaaa cgccggcgg caaccgagc tctgaacaa  
1741 atccagatgg agttctgagg tcattactgg atctatcaac aggagtcaa gcgagctcga  
1801 tatcaaatta cccccccc tgccactcat cgcagtactg ttgtaattca ttaagcattc  
1861 tgccgacatg gaagccatca caaacggcat gatgaacctg aatcgccagc ggcatcagca  
1921 cctgtcgc ttgcgtataa tattgccc tggtgaaaac gggggcgaag aagttgtcca  
1981 tattggccac gtttaaatca aaactgggta aactcacca gggattggct gagacgaaaa  
2041 acatattctc aataaacct ttagggaaat aggccagggt tcaccgtaa cacgccat  
2101 cttgcaata tatgtgtaga aactgccgga aatcgtcgtg gtattcactc cagagcagatg  
2161 aaaacgttct agtttctca tggaaaacgg tgaacaagg gtgaacacta tccatatca  
2221 ccagctcacc gtcttctt gccatagca atccggatg agcattcatc agcgggcaa  
2281 gaatgtgaat aaaggccgga taaaactgt gcttatttt cttacggctc ttaaaaagg  
2341 ccgtaaatc cagctgaacg gtctggttat aggtacattg agcaactgac tgaatgctc  
2401 caaatgttc ttacgatgc cattgggata tatcaacggg ggtatatcca gtgattttt  
2461 tctccattt agcttctta gctctgaaa atctgataa ctcaaaaaat acgcccgga  
2521 gtgatcttatt cattatgg tgaagttgg aacctcttac gtccgatca acgtctcatt  
2581 ttcgccagat atcgacgtc gtgcctaag agtgagctaa cttacattaa ttgcgttcg  
2641 ctactgccc gcttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcgcca  
2701 acgcgccggg agaggcgggt tgcgtattgg gcgccagggt ggttttctt tcaccagtg  
2761 agacgggcaa cagctgattg ccttcaccg cctggccctg agagagttgc agcaagcgg  
2821 ccacgtggt ttgcccagc aggcgaaat cctgtttgat ggtggttaac ggcgggat

2881 aacatgagct gtcttcgga tcgtcgtatc ccactaccga gatgtccgca ccaacgcgca  
 2941 gcccgactc ggtaatggcg cgcattgcgc ccagcgcctat ctgatcgttg gcaaccagca  
 3001 tcgcagtggg aacgatgcc tcattcagca ttgcatggt ttgttgaaa cggacatgg  
 3061 cactccagtc gccttcccg tccgctatcg gctgaatttg attgcgagtg agatattat  
 3121 gccagccagc cagacgcaga cgcgccgaga cagaacttaa tgggcccgt aacagcgca  
 3181 ttgctggtg accaatgcg accagatgct ccagcccag tcgcgtaccg tcttcatggg  
 3241 agaaaataat actggtgatg ggtgtctggt cagagacatc aagaaataac gccggaacat  
 3301 tagtgcaggc agcttcaca gcaatggcat cctggtcatc cagcggatag ttaatgatca  
 3361 gccactgac gcgttgccg agaagattgt gcaccgccg tttacaggct tcgacgccg  
 3421 ttcgtttac catcgacacc accacgctgg caccagttg atcggcgca gattaatcg  
 3481 ccgcgacaat ttgcgacggc gcgtgcaggg ccagactgga ggtggcaacg ccaatcagca  
 3541 acgactgtt gcccgccagt tgttgcca cgcggtggg aatgtaattc agtccgcca  
 3601 tcgcccttc cacttttcc cgcgtttcg cagaaacgtg gctggcctgg tcaccacgc  
 3661 gggaaacggt ctgataagag acaccggcat actctgcgac atcgtataac gttactggtt  
 3721 tcacattcac caccctgaat tgactctct cggggcgcta tcatgccata ccgcgaaagg  
 3781 ttttgcgcca ttcgatggtg tccgggatct cgacgctct cttatgca ctctgcatt  
 3841 aggaagcagc ccagtagtag gttgaggccg ttgagcaccg ccgccgcaag gaatggtgca  
 3901 tgcaaggaga tggcgccca cagtccccg gccacggggc ctgccacat acccagccg  
 3961 aaacaagcgc tcatgagccc gaagtggcga gcccgatctt cccatcggg gatgtcggc  
 //

**ZFP-65 expression from  $P_{BAD}$  for steric hindrance and orthogonality testing**

LOCUS 29-15 5072 bp DNA circular

FEATURES Location/Qualifiers  
 rep\_origin complement(976..3204)  
     /gene="SC101\*\*"  
 CDS complement(4118..4996)  
     /gene="araC"  
     /codon\_start="0"  
 gene complement(3326..3985)  
     /gene="CmR"  
 terminator 846..974  
     /gene="double terminator"  
 terminator 3205..3310  
     /gene="Terminator"  
 prot\_bind 163..201  
     /gene="Protein binding"  
 promoter complement(75..103)  
     /gene="Promoter"  
 promoter 200..227  
     /gene="Promoter"  
 prot\_bind 111..132  
     /gene="Protein binding"  
 prot\_bind 5025..5042

```

        /gene="Protein binding"
misc_feature 154..167
prot_bind   265..284
        /gene="RBS"
CDS         285..815
        /gene="ZFP-65"
        /codon_start="1"
BASE COUNT 1330 a 1192 c 1043 g 1507 t
ORIGIN
    1 tactggctct tctcgtaac caaacggta accccgcta ttaaagcat tctgtaaca
    61 agcgggacca aagccatgac aaaaacgctg aacaaaagtg tctataatca cggcagaaaa
    121 gtccacattg attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata
    181 agattagcgg attctacctg acgcttttta tcgcaactct ctactgtttc tccatacccg
    241 ttttttggg aattcaaaag atctttaag aaggagatat acatatgctg gaaccaggat
    301 ctaaaccgta caaatgtccg gaatgtgta aatccttctc ccgttctgat aaactggttc
    361 gtcacaaacg tactcacact ggatctaaac cgtacaaatg tccggaatgt ggtaaatcct
    421 tctccggttc tgataatctg gtacgtcact aacgtactca cactggatct aaaccgtaca
    481 aatgtccgga atgtggtaaa tccttctccc aatcttcaa tctggttcgt catcaacgta
    541 ctcaactggg atctaaaccg tacaatgtc cggaatgtgg taaatccttc tcccgttctg
    601 ataaactggg tcgtcatcaa cgtactcaca ctggatctaa accgtacaaa tgtccggaat
    661 gtggtaaatc ctttcccaa cgtgcaaatc tgcgtgctca tcaacgtact cactggat
    721 ctaaaccgta caaatgtccg gaatgtgta aatccttctc ccaacgtgca aatctgcgtg
    781 ctcatcaacg tactcacact ggatctaaaa cctcttaata aggatccaaa ctcgagtaag
    841 gatctccagg catcaaataa aacgaaaggc tcagtcgaaa gactgggcct ttcgtttat
    901 ctggtgttg tcggtgaacg ctcttacta gactcacact ggctcacctt cgggtgggcc
    961 tttctgctt tatacctagg gtacgggtt tgctgcccgc aaacgggctg ttctgggtg
    1021 gctagttgt tatcagaatc gcagatccgg cttagccgg tttgccgct gaaagcgcta
    1081 tttctccag aattgccatg atttttccc cacgggaggc gtactggct cccgtgtgt
    1141 cggcagcttt gattcgataa gcagcatcgc ctgttccagg ctgtctatgt gtgactgtg
    1201 agctgtaaca agttgtctca ggtgttcaat tcatgttct agttgcttg tttactggt
    1261 ttcacctgtt ctattagggt ttacatgctg tcatctgtt acattgtcga tctgttcag
    1321 gtgaacagct ttgaatgcac caaaaactcg taaaagctct gatgtatcta tctttttac
    1381 accgttttca tctgtgcata tggacagttt tccctttgat atgtaacggt gaacagttgt
    1441 tctactttt tttgtagtc ttgatgctc actgatagat acaagagcca taagaacct
    1501 agatccttcc gtatttagcc agtatgttct ctagtgtgtg tcgttgttt tgcgtgagcc
    1561 atgagaacga accattgaga tcatacttac tttgcatgct actcaaaaat tttgctcaa
    1621 aactggtgag ctgaatttt gcagttaaag catcgtgtag tgttttctt agtccgttat
    1681 gtaggtagga atctgatgta atggtgttg gtattttgtc accattcatt tttatctgg
    1741 tgttctcaag ttcggttacg agatccattt gtctatctag ttcaacttg aaaatcaacg
    1801 tatcagtcgg gcggcctcgc ttatcaacca ccaatttcat attgctgtaa gtgtttaa
    1861 ctttacttat tggtttcaaa acccattggt taagcctttt aaactcatgg tagttattt
    1921 caagcattaa catgaacta aattcatcaa ggctaacttc tatattgccc ttgtgagtt
    1981 tcttttgtg tagttcttt aataaccact cataaatcct catagagat ttgtttcaa
    2041 aagacttaac atgttccaga ttatatttta tgaattttt taactggaaa agataaggca
    2101 atatcttct actaaaaact aattctaatt ttcgcttga gaactggca tagttgtcc
    2161 actggaaaat ctcaaagcct ttaaccaaag gattctgat ttccacagtt ctgctcatca

```

2221 gctctctggg tgctttagct aatacacat aagcattttc cctactgatg ttcacatct  
2281 gagcgtattg gttataagtg aacgataccg tccgttcttt ccttgtaggg tttcaatcg  
2341 tggggttgag tagtgccaca cagcataaaa ttagcttggg ttcagctcc gttaatgcat  
2401 agcgactaat cgctagtca tttgcttga aaacaactaa ttcagacata catctcaatt  
2461 ggtctagggt attttaatca ctataccaat tgagatgggc tagtcaatga taattactag  
2521 tcctttccc ggggatctg ggtatctgta aattctgcta gacctttgct ggaaaactg  
2581 taaattctgc tagaccctct gtaaattccg ctgaccttt gtgtgtttt tttgtttata  
2641 ttcaagtggg tataatttat agaataaaga aagaataaaa aaagataaaa agaatagatc  
2701 ccagccctgt gtataactca ctactttagt cagttccgca gtattacaaa aggatgtcgc  
2761 aaacgctgtt tgctcctca caaacagac cttaaaacc taaaggctta agtagcacc  
2821 tcgcaagctc gggcaaatcg ctgaatattc cttttgtctc cgacctcag gcacctgag  
2881 cgctgtcttt ttcgtgacat tcagttcgtc gcgctcacgg ctctggcagt gaatgggggt  
2941 aaatggcact acaggcgcct tttatggatt catgcaagga aactacctat aatacaagaa  
3001 aagccccgca cgggcttctc agggcgtttt atggcgggtc tgctatgtgg tgctatctga  
3061 cttttgctg ttcagcagtt cctgcctct gattttccag tctgacctc tggattatc  
3121 ccgtgacagg tcattcagac tggctaagc acccagtaag gcagcggat catcaacagg  
3181 cttaccctgc ttactgtccc tagtgcttgg attctacca ataaaaaacg cccggcggca  
3241 accgagcgtt ctgaacaaat ccagatggag ttctgaggtc attactggat ctatcaacag  
3301 gagtccaagc gagctcgata tcaaattacg ccccgccctg cactcatcg cagtactgtt  
3361 gtaattcatt aagcattctg ccgacatgga agccatcaca aacggcatga tgaacctgaa  
3421 tcgccagcgg catcagcacc ttgtgcctt cgtataata tttgccatg gtgaaaacgg  
3481 gggcgaagaa gttgtccata ttggccactg taaatcaaaa actggtgaaa ctcaccagg  
3541 gattggctga gacgaaaaac atattctcaa taaaccttt agggaaatag gccaggttt  
3601 caccgtaaca cgccacatct tgcgaatata tgtgtagaaa ctgccgaaa tctcgtggt  
3661 attcactcca gagcgtgaa aacgtttcag ttgctcatg gaaaacggtg taacaagggt  
3721 gaacactatc ccatatcacc agctcaccgt ctttcattgc catacgaaat tccggatgag  
3781 cattcatcag gcgggcaaga atgtgaataa aggccggata aaacttgtgc ttattttct  
3841 ttacggtctt taaaaaggcc gtaatatcca gctgaacggt ctggttatag gtacattgag  
3901 caactgactg aaatgcctca aaatgttctt tacgatgcca ttgggatata tcaacggtgg  
3961 tatatccagt gattttttc tccattttag cttccttagc tctgaaaaat ctcgataact  
4021 caaaaaatac gcccgtagt gatcttatt cattatgggtg aaagttgga cctctacgt  
4081 gccgatcaac gtctcattt cgcagatat cgacgtctta tgacaactg acggctacat  
4141 cattcattt ttctcaciaa ccggcacgga actcgtcgg gctggccccg gtgcatttt  
4201 taaatacccg cgagaaatag agttgatcgt caaaaccaac attgcgaccg acggtggcga  
4261 taggcacccg ggtggtgctc aaaagcagct tgcctggct gatacgttg tctcgcgcc  
4321 agcttaagac gctaaccct aactgctggc ggaaaagatg tgacagacgc gacggcgaca  
4381 agcaaacatg ctgtgcgacg ctggcgatat caaaattgct gtctgccagg tgatcgtga  
4441 tgactgaca agcctcgcgt acccgattat ccatcgggtg atggagcgc tcgttaatcg  
4501 ctccatgcg ccgcagtaac aattgtcaa gcagattat cggcagcagc tccgaatagc  
4561 gcccttccc ttcccggcg ttaatgatt gcccaacag gtcgctgaaa tgcggctggt  
4621 gcgcttcatc cgggcgaaag aaccccgat tggcaaatat tgacggccag ttaagccatt  
4681 catgccagta ggcgcgcgga cgaaagtaaa cccactggtg ataccattcg cgagcctccg  
4741 gatgacgacc gtagtgatga atctctctg gcgggaacag caaaatatca cccggtcggc  
4801 aaacaaattc tctcctcga tttttacca cccctgacc gcgaatggtg agattgagaa  
4861 tataacctt cattcccagc ggtcggctga taaaaaatc gagataaccg ttggcctcaa  
4921 tcggcgttaa acccgcacc agatgggcat taaacagta tcccggcagc aggggatcat

```
4981 ttgcgcttc agccatactt ttcatactcc cgccattcag agaagaaacc aattgtccat
5041 attgcatcag acattgccgt cactgctct tt
//
```

**ZFP-3 expression from  $P_{BAD}$  for steric hindrance and orthogonality testing**

LOCUS 29-77 5072 bp DNA circular

FEATURES Location/Qualifiers  
rep\_origin complement(976..3204)  
/gene="SC101\*\*"  
CDS complement(4118..4996)  
/gene="araC"  
/codon\_start="0"  
gene complement(3326..3985)  
/gene="CmR"  
terminator 846..974  
/gene="double terminator"  
terminator 3205..3310  
/gene="Terminator"  
prot\_bind 163..201  
/gene="Protein binding"  
promoter complement(75..103)  
/gene="Promoter"  
promoter 200..227  
/gene="Promoter"  
prot\_bind 111..132  
/gene="Protein binding"  
prot\_bind 265..284  
/gene="RBS"  
prot\_bind 5025..5042  
/gene="Protein binding"  
CDS 285..815  
/gene="ZFP-3"  
/codon\_start="0"

BASE COUNT 1329 a 1192 c 1042 g 1509 t

ORIGIN

```
1 tactggctct tctcgtaac caaacggta acccgccta ttaaagcat tctgtaaca
61 agcgggacca aagccatgac aaaaacgctg acaaaaagtg tctataatca cggcagaaaa
121 gtccacattg attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata
181 agattagcgg attctacctg acgcttttta tcgcaactct ctactgtttc tccatacccg
241 ttttttggg aattcaaaaag atctttaaag aaggagatat acatagctg gaaccaggat
301 ctaaaccgta caaatgtccg gaatgtggta aatccttctc ccgttctgat aatctggtac
361 gtcacaaacg tactcacact ggatctaaac cgtacaaatg tccggaatgt ggtaaactct
421 tctccgcttc tgataatctg gtacgtcatc aacgtactca cactggatct aaaccgtaca
481 aatgtccgga atgtggtaaa tccttctccc agcgtgctca tctggaacgt catcaacgta
```



541 ctcacactgg atctaaaccg tacaaatgtc cggaatgtgg taaatccttc tcccaatctt  
601 ccaatctggg tcgtcatcaa cgtactcaca ctggatctaa accgtacaaa tgtccggaat  
661 gtggtaaadc cttctcccgt tccgatcacc tgactaatca tcaacgtact cacactggat  
721 ctaaaccgta caaatgtccg gaatgtggta aatccttctc ccgttctgat aatctggtac  
781 gtcacacag tactcacact ggatctaaaa cctcttaata aggatccaaa ctcgagtaag  
841 gatctccagg catcaaataa aacgaaaggc tcagtcgaaa gactgggcct ttcgtttat  
901 ctgttgtttg tcggtgaacg ctcttacta gagtcacact ggctcacctt cgggtgggccc  
961 tttctgcgtt tatacctagg gtacggggtt tgctgcccgc aaacgggctg ttctggtgtt  
1021 gctagtttgt tatcagaatc gcagatccgg cttcagccgg ttgcccggct gaaagcgcta  
1081 tttctccag aattgccatg atttttccc cacgggaggc gtcaactggct cccgtgtgt  
1141 cggcagcttt gattcagata gcagcatcgc ctgttcagg ctgtctatgt gtgactgtg  
1201 agctgtaaca agttgtctca ggtgttcaat ttcattgttct agttgctttg tttactggt  
1261 ttcacctgtt ctattagtg ttacatgctg ttcattctgtt acattgtcga tctgttcag  
1321 gtgaacagct ttgaatgac caaaaactcg taaaagctct gatgtatcta tctttttac  
1381 accgttttca tctgtcata tggacagttt tccctttgat atgtaacggg gaacagttgt  
1441 tctactttt ttgttagtc ttgatgctc actgatagat acaagagcca taagaacct  
1501 agatccttc gtatttagc agtatgttct ctagtgtgg tcgttgttt tgcgtgagcc  
1561 atgagaacga accattgaga tcatacttac ttgcatgct actcaaaaat ttgctca  
1621 aactgggtgag ctgaatttt gcagttaaag catcgtgtag tgttttctt agtccgttat  
1681 gtaggtagga atctgatga atggtgttg gtattttgc accattcatt tttatctgg  
1741 tgttctcaag ttcggttac agatccatt gtctatctag ttaacttg aaatcaacg  
1801 tatcagtcgg gcgccctgc ttatcaacca ccaattcat attgctgtaa gtgtttaa  
1861 cttacttat tggtttcaa acccattgt taagccttt aaactcatgg tagttattt  
1921 caagcattaa catgaacta aattcatca ggctaatctc tatattgccc ttgtgagtt  
1981 tctttgtgt tagttcttt aataaccact cataaatcct catagagat ttgtttcaa  
2041 aagactaac atgttcaga ttatattta tgaattttt taactggaaa agataaggca  
2101 atatctctc actaaaaact aattctaatt ttcgcttga gaactggca tagttgtcc  
2161 actggaaaat ctcaaagcct ttaacaaag gattctgat tccacagtt ctgctatca  
2221 gctctctgt tgcttagct aataccat aagcatttc cctactgat ttcatcatc  
2281 gagcgtattg gttataagt aacgataccg tccgttctt cctgtaggg tttcaatcg  
2341 tggggtgag tagtccaca cagcataaaa ttagcttgg ttcatgctcc gtaagtcat  
2401 agcactaat cgtagttca tttgcttga aaacaacta ttcagacata catctcaat  
2461 ggtctagtg attttaatc ctataccaat tgagatgggc tagtcaatga taactactg  
2521 tcttttccc ggggatctg ggatctgta aattctgta gaccttgc gaaaactg  
2581 taaattctg tagacctc gtaaattccg ctgacctt gtgtgtttt tttgttata  
2641 tcaagtgt tataattat agaataaaga aagaataaaa aaagataaaa agaatagatc  
2701 ccagccctgt gtataactc ctacttagt cagttccga gtattacaaa aggatgctc  
2761 aaacgctgt tgctcctca caaacagac cttaaaacc taaaggctta agtagcacc  
2821 tcgcaagtc gggcaaatc ctgaatatt ctttctctc cgacctcag gcacctgag  
2881 cgctgtctt ttcgtgacat tcagttcgt gcgctcagg ctctggcag gaatggggg  
2941 aatggcact acaggcctc tttatggatt catgcaagga aactacct aatacaagaa  
3001 aagcccgta cgggcttctc agggcgttt atggcgggtc tgctatgtg tgctatctga  
3061 cttttgctg ttcagcagtt cctgccctc gatttccag tctgacct tcggattac  
3121 ccgtgacagg tcattcagac tggctaatc accagtaag gcagcggat catcaacag  
3181 ctaccctg ttaactgccc tagtgcttg attctacca ataaaaaac cccggcggca  
3241 accgagcgt ctgaacaaat ccagatggag ttctgaggtc attactggat ctatcaacg

3301 gagtccaagc gagctcgata tcaaattacg ccccgccctg cactcatcg cagtactgtt  
 3361 gtaattcatt aagcattctg cgcacatgga agccatcaca aacggcatga tgaacctgaa  
 3421 tcgccagcgg catcagcacc ttgtcgcctt gcgataata ttgcccatg gtgaaaacgg  
 3481 gggcgaagaa gttgtccata ttggccacgt taaatcaaaa actggtgaaa ctacccagg  
 3541 gattggctga gacgaaaaac atattctcaa taaacctttt agggaaatag gccaggtttt  
 3601 caccgtaaca cgccacatct tgcgaatata tgtgtagaaa ctgccggaaa tcgtcgtggt  
 3661 attcactcca gagcgaatga aacgtttcag ttgctcatg gaaaacgggt taacaagggt  
 3721 gaacactatc ccatatcacc agctcaccgt ctttcattgc catacgaat tccggatgag  
 3781 cattcatcag gccggcaaga atgtgaataa aggccggata aaacttgtgc ttatTTTTT  
 3841 ttacggtctt taaaaaggcc gtaatatcca gctgaacgggt ctggttatag gtacattgag  
 3901 caactgactg aatgcctca aatgttctt tacgatgcca ttgggatata tcaacgggtg  
 3961 tataccagt gattTTTTT tccatttag cttccttagc tctgaaaaat ctcgataact  
 4021 caaaaaatac gcccggtagt gatcttattt cattatggtg aaagttggaa cctcttacgt  
 4081 gccgatcaac gtctcatTTT cgccagatat cgacgtctta tgacaacttg acggctacat  
 4141 cattcacttt ttctcaciaa ccggcacgga actcgtcctg gctggccccg gtgcattttt  
 4201 taaatacccg cgagaaatag agttgatcgt caaaaccaac attgacgacc acggtggcga  
 4261 taggcatccg ggtggtgctc aaaagcagct tcgcctggct gatacgttgg tctcgcgcc  
 4321 agcttaagac gctaaccct aactgctggc ggaaaagatg tgacagacgc gacggcgaca  
 4381 agcaaacatg ctgtgcgacg ctggcgatat caaaattgct gtctgccagg tgatcgtga  
 4441 tgtactgaca agcctcgcgt acccgattat ccatcggtgg atggagcgac tcgttaatcg  
 4501 ctccatgcg ccgagtaac aattgtcaa gcagatttat cgccagcagc tccgaatagc  
 4561 gcccttccc ttgcccgcg ttaatgattt gccaaacag gtcgctgaaa tgcggctggt  
 4621 gcgcttcac cgggcgaaag aaccccgat tggcaaatat tgacggccag ttaagccatt  
 4681 catgccagta ggcgcgcgga cgaaagtaaa cccactggtg ataccattcg cgagcctccg  
 4741 gatgacgacc gtagtgatga atctctctg gcgggaacag caaataatca cccggtcggc  
 4801 aaacaaattc tcgtccctga ttttccaca cccctgacc gcgaatggtg agattgagaa  
 4861 tataaccttt cattcccagc ggtcggtcga taaaaaatc gagataaccg ttggcctcaa  
 4921 tcggcgtaa acccgccacc agatgggcat taaacgagta tcccgcgac aggggatcat  
 4981 ttgctcttc agccatactt ttcatactcc cgccattcag agaagaaacc aattgcat  
 5041 attgcatcag acattgccgt cactgcgtct tt

//

**ZFP-39 expression from pET-29 for protein purification**

LOCUS 30-15 5871 bp DNA circular

FEATURES Location/Qualifiers  
 CDS complement(4483..5565)  
 /gene="lacI"  
 CDS complement(1528..2343)  
 /gene="KanR"  
 /codon\_start="0"  
 rep\_origin 2465..3053  
 /gene="pMB1"  
 /direction="left"  
 misc\_feature 176..220

/gene="N-terminal S-tag coding sequence"  
 misc\_feature 104..128  
 /gene="lac operator"  
 misc\_feature 85..104  
 /gene="T7 promoter primer "  
 misc\_feature 236..253  
 /gene="thrombin cleavage site tag"  
 gene complement(812..829)  
 /gene="6xHis tag"  
 promoter 85..101  
 /gene="T7 Promoter"  
 misc\_feature 39..54  
 /gene="pET upstream primer"  
 CDS 269..796  
 /gene="ZFP-39"  
 /codon\_start="0"  
 CDS 173..829  
 /gene="full protein"  
 /codon\_start="1"  
 BASE COUNT 1420 a 1583 c 1463 g 1405 t  
 ORIGIN

1 cagcaaccgc acctgtggcg ccggtgatgc cggccacgat gcgtccggcg tagaggatcg  
 61 agatcgatct cgatcccgcg aaattaatac gactcactat aggggaattg tgagcggata  
 121 acaattccc tctagaata atttgttta actttaagaa ggagatatac atatgaaaga  
 181 aaccgtgct gctaaattcg aacgccagca catggacagc ccggatctgg gtaccctggt  
 241 gccacgcggt tccgaattca aaagatctct ggaaccagga tctaaaccgt acaaatgtcc  
 301 ggaatgtggt aaatccttct cccgcaaaga taacctgaaa aaccatcaac gtactcacac  
 361 tggatctaaa ccgtacaaat gtccggaatg tggtaaacc ttctccggtt ctgatcatct  
 421 gactactcat caacgtactc aactggatc taaaccgtac aaatgtccgg aatgtggtaa  
 481 atccttctcc cgtgaagata acctgcatac tcatcaacgt actcacactg gatctaaacc  
 541 gtacaaatgt ccggaatgtg gtaaactcct ctctctcca gctgatctga ctcgtcatca  
 601 acgtactcac actggatcta aaccgtacaa atgtccggaa tgtggtaaact ctttctccc  
 661 ttctgatcat ctgactactc atcaacgtac tcaactgga tctaaaccgt acaaatgtcc  
 721 ggaatgtggt aaatccttct cccgttctga tcatctgact actcatcaac gtactcacac  
 781 tggatctaaa acctctggat ccggctcgag ccaccaccac caccaccact gagatccggc  
 841 tgctaacaaa gcccgaaagg aagctgagtt ggctgctgcc accgctgagc aataactagc  
 901 ataaccctt ggggcctcta aacgggtcct gaggggtttt ttgctgaaag gaggaactat  
 961 atccggattg gcgaatggga cgcgcctgt agcggcgcac taagcgcggc ggggtggtg  
 1021 gttacgcgca gcgtgaccgc tacacttgc agcgccttag cggcctcc tttcgcttc  
 1081 ttccttctt ttctcgccac gttcggcgc tttcccgc aagctctaaa tcgggggctc  
 1141 ctttaggggt tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt  
 1201 gatggtcac gtagtgggccc atcgcctga tagacggttt ttcgccttt gacgttggag  
 1261 ccacgttct ttaatagtg actctgttc caaactggaa caaactcaa ccctatctcg  
 1321 gtctattctt ttgattata agggatttg ccgatttcgg cctattggtt aaaaaatgag  
 1381 ctgattaac aaaaattaa cgcgaatgtt acaaaaatat taacgtttac aatttcaggt  
 1441 ggcactttc ggggaaatgt gcgcggaacc cctattgtt tatttttcta aatacattca

1501 aatatgtatc cgctcatgaa ttaattctta gaaaaactca tcgagcatca aatgaaactg  
1561 caatttattc ataccagat tatcaatacc atatttttga aaaagccgtt tctgtaatga  
1621 aggagaaaac tcaccgaggc agttccatag gatggcaaga tcttggtatc ggtctgcgat  
1681 tccgactcgt ccaacatcaa tacaacctat taattcccc tcgtcaaaaa taaggttatc  
1741 aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa gtttatgcat  
1801 ttctttccag acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcac  
1861 aaccaaaccg ttattcattc gtgattgctg ctgagcgaga cgaaatacgc gatcgtggt  
1921 aaaaggacaa ttacaacag gaatcgaatg caaccggcgc aggaactg ccagcgcac  
1981 aacaatattt tcacctgaat caggatattc ttctaatacc tggaatgctg ttttccggg  
2041 gatcgcagtg gtgagtaacc atgcatcacc aggagtacgg ataaaatgct tgatggctgg  
2101 aagaggcata aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc  
2161 aacgctacct ttgcatggt tcagaaacaa ctctggcgca tcgggcttcc catacaatg  
2221 atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc catataatc  
2281 agcatccatg ttggaattta atcgcggcct agagcaagac gtttccggtt gaatatggct  
2341 cataacaccc cttgtattac tgtttatga agcagacagt ttattgttc atgacaaaa  
2401 tcccttaacg tgagtcttcg ttccactgag cgtcagaccc cgtagaaaag atcaaggat  
2461 cttcttgaga tctttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccg  
2521 taccagcggg ggtttgttg ccggatcaag agctaccaac tcttttccg aaggtaactg  
2581 gcttcagcag agcgcagata ccaatactg tcttctagt gtagccgtag ttaggccacc  
2641 actcaagaa ctctgtagca ccgctacat acctcgtct gctaatctg ttaccagtgg  
2701 ctgctgccag tggcgataag tctgtctta ccgggttga ctcaagacga tagttaccgg  
2761 ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcga  
2821 cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgc acgctccc  
2881 aaggagaaaa ggcggacagg tatccgtaa gcggcagggt cggaacagga gagcgcacga  
2941 gggagctcc aggggaaac gcctggtatc tttatagtcc tctcgggtt cccacctct  
3001 gacttgagcg tctgtttg tgatgctct cagggggcg gagcctatgg aaaaacgcca  
3061 gcaacgggc cttttaccg ttctggcct tttgctggc tttgctcac atgttcttc  
3121 ctgcttatc cctgattct gtgataacc gtattaccgc ctttgagtga gctgataccg  
3181 ctgccgcag ccgaacgacc gagcgcagc agtcagttag cgaggaagc gaagagcgc  
3241 tgatcggta ttttctct acgcatctgt gcggtattc acaccgata tatggtgcac  
3301 tctcagtaca atctgctct atgccgata gtaagccag tatacactcc gctatcgta  
3361 cgtgactggg tcatggctg gcccgacac ccgcaacac ccgctgacg ccctgacgg  
3421 gctgtctgc tcccggcatc cgcttacaga caagctgta ccgtctccg gagctgcat  
3481 tgcagaggt ttcaccgct ataccgaaa cgcgcgagc agctcggta aagctcatca  
3541 gcgtgctct gaagcattc acagatgct cctgttcat ccgctccag ctggtgagt  
3601 ttctccagaa gcgttaatgt ctggttctg ataaagcggg ccatgttaag ggcggtttt  
3661 tctgtttg tcaactgatc ctccgtgaa ggggatttc tttcatggg ggtaatgata  
3721 ccgatgaaac gagagaggat gctcagata cgggttactg atgatgaaca tgcccggta  
3781 ctggaacgtt gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc  
3841 actcagggtc aatccagc cttcgtaat acagatgtag gtgtccaca gggtagccag  
3901 cagcatctg cgatgcagat ccggaacata atggtgcagg gcgctgactt ccgctttcc  
3961 agacttacg aaacacgaa accgaagacc atcatgttg ttgctcaggt cgcagacgtt  
4021 ttgacgagc agtcgctca cgttcgctc cgtatcggg attcattctg ctaaccagta  
4081 aggcaacccc gccagcctag ccgggtctc aacgacagga gcacgatcat gcgacccgt  
4141 ggggccgcca tgccgggat aatggcctgc ttctcggca aacgtttgtt ggcgggacca  
4201 gtgacgaagg cttgagcag ggcgtgcaag attccgaata ccgcaagcga caggccgatc

4261 atcgtcgcgc tccagcga aa gcggtcctcg ccgaaaatga cccagagcgc tgccggcacc  
 4321 tgcctacga gttgcatgat aaagaagaca gtcataagtg cggcgacgat agtcatgccc  
 4381 cgcgccacc ggaaggagct gactgggttg aaggctctca agggcatcgg tcgagatccc  
 4441 ggtgccta at gagtgagcta acttacatta attgctgtgc gctcactgcc cgctttccag  
 4501 tcgggaaacc tgcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg  
 4561 ttgctattg ggcgccaggg tggttttct tttaccagt gagacgggca acagctgatt  
 4621 gcccttacc gcctggcct gagagagttg cagcaagcgg tccacgctgg ttgccccag  
 4681 caggcga aaa tctgtttga tgggtggtta cggcgggata taacatgagc tgccttcggt  
 4741 atcgtcgtat cccactaccg agatgtccgc accaacgcgc agcccggact cggtaatggc  
 4801 gcgcatgcg cccagcgcca tctgatcgtt ggcaaccagc atcgcagtgg gaacgatgcc  
 4861 ctattcagc attgcatgg tttgtgaaa accggacatg gcaactcagt cgccttccc  
 4921 ttccgctatc ggctgaattt gattgagcgt gagatattta tgccagccag ccagacgag  
 4981 acgcgccgag acagaactta atgggcccgc taacagcgcg attgtctggt gaccaatgc  
 5041 gaccagatgc tccacgcca gtcgctacc gtctcatgg gagaaaataa tactgttgat  
 5101 ggggtgtctg tcagagacat caagaaataa cgccggaaca ttagtgagg cagcttccac  
 5161 agcaatggca tctgtgtcat ccagcggata gtaatgatc agcccactga cgcgttgcg  
 5221 gagaagattg tgcaccgccc cttacagggc ttcgacgccc ctctgttcta ccatcgacac  
 5281 caccacgctg gcaccagtt gatcggcgcg agatttaac gccgcgaca tttgagcgg  
 5341 cgcgtgagg gccagactgg aggtggcaac gcaatcagc aacgactgtt tgcccggcag  
 5401 ttgttgccc acgcggttgg gaatgtaatt cagctccgcc atcggcgtt cacttttc  
 5461 ccgcttttc gcagaaactg ggctggcctg gttcaccagc cgggaaacgg tctgataaga  
 5521 gacaccggca tactctgca catctataa cgttactggt ttcacattca ccacctgaa  
 5581 ttgactctc tccgggctc atcatgcat accgcgaaag gttttgcgcc attcagtggt  
 5641 gtccgggatc tcgacgctc cccttatgcg actcctgcat taggaagcag cccagtagta  
 5701 ggttgaggcc gttgagcacc gccgccgca ggaatggtgc atgcaaggag atggcgcca  
 5761 acagtcctcc ggccacgggg cctgccacca taccacgcc gaaacaagcg ctcatgagcc  
 5821 cgaagtggcg agcccgatct tccccatcgg tgatgtcggc gatataggcg c

//

**Reporter for orthogonality testing: operator-3 and BioFAB promoter 83**

LOCUS 32-32 3039 bp DNA circular

FEATURES Location/Qualifiers  
 gene complement(411..1205)  
 /gene="KanR"  
 rep\_origin complement(2631..274)  
 /gene="ColE1"  
 terminator 2369..2497  
 /gene="terminator"  
 terminator 280..385  
 /gene="Terminator"  
 CDS 1582..2286  
 /gene="mCherry"  
 /codon\_start="1"  
 terminator 2299..2344

```

        /gene="terminator BBa_B0011"
prot_bind 1501..1518
        /gene="ZFP-3 binding site"
promoter 1519..1553
        /gene="BioFAB RPL-83"
misc_feature 1554..1581
        /gene="Bujard 5'-UTR"
misc_feature 1357..1500
        /gene="insulator"
BASE COUNT 773 a 794 c 799 g 673 t
ORIGIN
1 ggtgctacag agttcttgaa gtggtggcct aactacggct aactagaag gacagtattt
61 ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc
121 ggcaaaaaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca gattacgcbc
181 agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggctctga cgctcagtgg
241 aacgaaaact cacgttaagg gattttggtc atgactagtg cttggattct caccaataaa
301 aaacgcccgg cggcaaccga gcttctgaa caaatccaga tggagtctg aggtcattac
361 tggatctatc aacaggagtc caagcgagct ctcgaacccc agagtcccgc tcagaagaac
421 tcgtcaagaa ggcatagaa ggcatgcbc tgcgaatcgg gagcggcgat accgtaaagc
481 acgaggaagc ggtcagccca ttcgcccca agctcttcag caatatcacg ggtagccaac
541 gctatgtcct gatagcggtc cgccacacc agccggccac agtcgatgaa tccagaaaag
601 cggccatttt ccacatgat attcggcaag caggcatcgc catgggtcac gacgagatcc
661 tcgccgtcgg gcatgcbcgc cttgagcctg gcgaacagtt cggctggcgc gagcccctga
721 tgctcttctg ccagatcctc ctgatcgaca agaccggctt ccatccgagt acgtgctcgc
781 tcgatgcgat gtttcgcttg gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc
841 cgccgcattg catcagccat gatggatact ttctcggcag gagcaagggtg agatgacagg
901 agatcctgcc ccggcacttc gcccaatagc agccagtccc ttcccgttc agtgacaacg
961 tcgagcacag ctgcaagagg aacgcccgc gtggccagcc acgatagccg cgctgcctcg
1021 tctgcagtt cattcagggc accggacagg tcggtcttga caaaaagaac cgggcgcccc
1081 tgcgctgaca gccggaacac ggccgcatca gagcagccga ttgtctgttg tgcccagtca
1141 tagccgaata gcctctccac ccaagcggcc ggagaacctg cgtgcaatcc atctgttca
1201 atcatgcgaa acgatctca tctgtctct tgatcagatc atgatcccct gcgcatcag
1261 atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac cttaccagag
1321 ggcgccccag ctggcaattc cgaattcatg agatcttaag taagtaagag tatacgtata
1381 tcggctaata acgtattaag gcgcttcggc gcctttttt atgggggtat tttcatcca
1441 atccacacgt ccaacgcaca gcaaacacca cgtcgaccct atcagctgcg tgctttctat
1501 gagagggaag gagaggagtt gacaattaat catcggctca taaccttgt ggaacaattc
1561 attaaagagg agaaaggtag catgcgtaaa ggagaagaag ataacatggc tatcattaa
1621 gagttcatgc gttcaaaagt tcacatggag gtttctgtta acggtcacga gttcagatc
1681 gaaggcgaag gcgagggccc tccgtatgaa ggcaccaga ccgcaaaact gaaagtgact
1741 aaaggcggcc cgctgccttt tgctggtggc atcctgagcc cgcaattat gtacggttct
1801 aaagcgtatg ttaaaccacc agcggatata ccggactatc tgaagctgtc tttccggaa
1861 ggtttcaagt gggaacgcgt aatgaattt gaagatggtg gtgtcgtgac cgtcactcag
1921 gactcctccc tgcaggatgg cgagttcatc tataaagtta aactgcgtgg tactaatttt
1981 ccatctgatg gcccggtgat gcagaaaaag acgatgggtt gggaggcgtc tagcgaacgc
2041 atgatccgg aagatggtgc gctgaaaggc gaaattaaac agcgcctgaa actgaaagat

```

2101 ggcggccatt atgacgctga agtgaaaacc acgtacaaag ccaagaaacc tgtgcagctg  
2161 cctggcgcgt acaatgtgaa tattaactg gacatcacct ctcataatga agattatacg  
2221 atcgtagagc aatatgagcg cgcggagggt cgtcattcta ccggtggcat ggatgaacta  
2281 tacaataat aaggatctag agaataataa aagccagatt attaatccgg ctttttatt  
2341 attggatcc taactcgagt aaggatctcc aggcatacaa taaaacgaaa ggctcagtcg  
2401 aaagactggg ctttcgttt tatctgttgt ttgctgggta acgctctcta ctagagtcac  
2461 actggctcac cttcgggtgg gcctttctgc gttatacct agggcgttcg gctgcggcga  
2521 gcggtatcag ctactcaaa ggcggaataa cggttatcca cagaatcagg ggataacgca  
2581 ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg  
2641 ctggcgttt tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt  
2701 cagaggtggc gaaaccgac aggactataa agataccagg cgtttcccc tggagctcc  
2761 ctctgctcct ctctgttcc gaccctgcc cttaccggat acctgtccgc ctttctcct  
2821 tcgggaagcg tggcgtttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc  
2881 gttcgtcca agctgggctg tgtgcacgaa cccccgttc agccgaccg ctgctccta  
2941 tccgtaact atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca  
3001 gccactgta acaggattag cagagcgagg tatgtaggc

//