SYNTHETIC BIOMOLECULAR CONTROL SYSTEMS

TO REPROGRAM CELLULAR FUNCTIONS AND DYNAMIC RESPONSES

by

Ka-Hei Siu

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering

Winter 2019

SYNTHETIC BIOMOLECULAR CONTROL SYSTEMS

TO REPROGRAM CELLULAR FUNCTIONS AND DYNAMIC RESPONSES

by

Ka-Hei Siu

Approved: _____
Eric M. Furst, Ph.D.
Chair of the Department of Chemical Engineering

Approved: _____
Levi T. Thompson, Ph.D.
Dean of the College of Engineering

Approved: _____
Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Wilfred Chen, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Eleftherios T Papoutsakis, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Maciek R Antoniewicz, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Sharon Rozovsky, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

A great many memories come across my mind as I fill these pages to chronicle my journey of the past few years. Many more pages will be needed to adequately express my gratitude for whom these memories were made, more than I can possibly write. For the sake of clarity and brevity, I will give thanks to these people who have lent their helping hands on this long road using a somewhat chronological approach:

No amounts of words would suffice to acknowledge the unconditional support and patience provided to me by my family. Too often, I have failed to communicate much, if any, aspects and details of my whereabouts since leaving home a decade ago. Despite my failings, my family has always been accepting and supportive of my choices, even when they do not believe the choices are necessarily right for me. Special thanks must be given to my aunt and her family for taking my mom and me into their own homes and making us part of the family. I could not have left to pursue my studies without their continual care for the rest of the family. Last, but certainly not least, I need to recognize the outsized role that my mom has played in my journey thus far. There is nothing that I can say or do that will justify the sacrifices and toil that she has made for a child who has plainly not been easy to satisfy or raise. For all the troubles and heartaches this son has brought, I can only hope that my deeds, including the dedication of this work, will be enough to provide you some solace and pride in exchange for your unwavering devotion.

I have been fortunate enough to have been mentored by some truly inspiring figures throughout my studies. If my memories serve me correctly, the educators of

my younger years deserve a great deal of praise for their perseverance and beliefs in their missions to nurture a young mind, even in the face of language and cultural barriers. Beyond my teenage years, I am especially thankful for the mentors I met during my undergraduate and graduate studies. Thanks to Professor Matt DeLisa for giving me the opportunity to work in his lab and Dr. Jason Boock for being patient enough to teach the basics of scientific research to a clueless undergraduate.

Of course, none of this work would be possible without my graduate advisor, Dr. Wilfred Chen. Too many times have I walked into his office without a grain of useful data and too many times have I failed to heed his advice. Yet despite all the disappointments, his audacity to grant me the independence and trust in my work never faltered and he always provided the intellectual and emotional boost that I needed, even in the most grueling times. I am also grateful for the hours of discussions on science (and sports) that we have had and, hopefully, will continue to have beyond my time here.

Outside of my studies, I must thank the many friends and peers who have made the time during my graduate studies enjoyable. Likely, the times spent chatting and working and playing and eating with them will be my most enduring memories of this time. Two particular groups need to be explicitly acknowledged:

First, my fellow Chen lab members with whom I spent the majority of my waking hours for the past few years are to be commended for tolerating my occasional blunders and incessant preaching in and out of the lab. I thank Miso Park and Heejae Kim for helping me find my way around the lab. Special mentions are given to my office and bench mates over the years, Rebecca Chen, Emily Hartzell, Long Chen, Andrew Gaynor, and Daniel Yur for listening to my rambling anecdotes and

convoluted metaphors and still managing to laugh at them (sometimes). I am thankful for the meals and grievances that Emily Hartzell and I have shared in all the restaurants that I will never try otherwise. And there is scarcely any one who knows me more than Rebecca does, for which I am grateful.

The second group of friends to be recognized is my roommates at 123 Arielle Drive, past and present. Plainly speaking, they are my best friends, even if they like to call me their landlord instead. The many gatherings we have hosted and attended over the years remain some of the most memorable moments and best stories that I recall. I particularly commend the courage of Chris, Cameron, Andrew, and Chad for placing their trust in me when the house has not yet materialized. Watching sports, news, and movies has not been the same without Ed and Megan. Though the time they have spent at the house has not yet been long, Alex, Kamil, and Terry continue to take good care of me and keep me young. For these and countless other matters, my roommates, I promise that you will always be welcome into my home, wherever and whenever I may be, just as I did when you first moved into the house. I cannot wait to see where you are heading and am hopeful that I will continue to play a part on your journeys.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Living systems of all scales must constantly adapt to changing conditions by surveying their surroundings and executing appropriate responses. To adequately respond to the diverse environmental stimuli presented by nature, living organisms have evolved a vast repertoire of molecular circuits to direct their physiological behaviors. In contrast, synthetic biological systems are not yet equipped with robust, autonomous, and generalizable biomolecular regulatory elements to implement dynamic control circuits necessary for advanced biotechnology and human health applications. There is, thus, a need to develop new biomolecular tools to mimic and integrate various aspects of natural control systems into functional regulatory schemes.

The first step in any dynamic control scheme involves the recognition of a given stimuli by a sensor. While numerous sensing machineries have been found in nature that provide a rich toolkit of sensory components for many forms of stimuli, these naturally occurring components are not adapted for use in an artificial context. To bypass the often tedious process of engineering membrane receptors and intracellular signaling pathways, we envisioned a synthetic extracellular sensing circuit that can exploit pre-existing membrane receptors by input-induced reconstitution of native signaling peptide. The biochemical basis of such a sensing circuit was established by adapting intein-mediated reactions to reconstitute the well-known yeast mating pheromone peptide, α-factor, and exploiting the associated yeast mating pathway to direct cellular responses.

Biological control systems contain logic elements that receive and process signals from sensory elements to decide what responses need to be executed. A particular hurdle for integrating synthetic components into control circuits is the prevalence of unspecific interactions and crosstalk that inevitably arise from the crowded cellular milieu. In addition, the inability to utilize endogenous cellular information as inputs for activating or repressing responses prohibit the use of synthetic control elements from being used in autonomous regulatory schemes. These limitations drove the design and construction of a new class of riboregulators, termed toehold-gated guide RNA. The synthetic riboregulators can be programmed to respond to a very large variety of RNA sequences, including full-length mRNA, and control CRISPR/Cas9 activities for multiplexed gene regulation in *E. coli* with minimal cross-talk. The versatility of this platform was demonstrated by the use of endogenous RNA transcripts as triggers to activate Cas9 functions, allowing thgRNA to be used as an autonomous control elements in dynamic control schemes.

Molecular control circuits rely on precise protein-protein interactions to be made consistently to execute the appropriate responses. The actuations of biomolecular responses are tightly regulated by post-translational modifications such as phosphorylation, which are not readily applicable to synthetic output components like heterologous enzymes or transcription factors. One way to circumvent the use of post-translational modifications to control protein-protein interactions is to apply coiled-coil motifs as recombinant peptide tags on synthetic components to direct their interactions. Dynamic reconfiguration of protein complex can be achieved by taking advantage of coiled-coil motifs and their associated strand displacement reaction. Control over a heterologous enzyme complex and its productivity was demonstrated.

Chapter 1

INTRODUCTION

1.1 Dynamic Responses as Prerequisites for Natural or Engineered Biological Systems

Biology defines one central characteristic of life as the ability for a system to respond to stimuli (Reece et al. 1989). Natural living systems have evolved a broad range of complex molecular control circuits to survey external and internal states and execute appropriate responses. These molecular circuits contain functional subsystems that monitor environmental and cellular signals, perform signal processing operations, and actuate appropriate molecular responses (Lim 2010). Because environmental stimuli are endlessly diverse and dynamic, these naturally evolved control modules have also adapted to function in different contexts and time scales to enable individual organisms and populations to thrive against changing conditions and challenges.

Engineered biological systems present challenges similar to those encountered in the transient and mercurial natural environments and thus also require dynamic responses to efficiently fulfill their functions. Much as a chemical processing plant requires control devices to maintain high productivity, a cellular factory, whether it is applied in bioprocessing or human health, needs molecular control elements to sustain its operations (Figure 1.1). Reliable means of regulating biomolecular processes would be needed to fully translate the promise of synthetic biology of creating tailored-made organisms to tackle any given task (Cameron, Bashor, and Collins 2014).

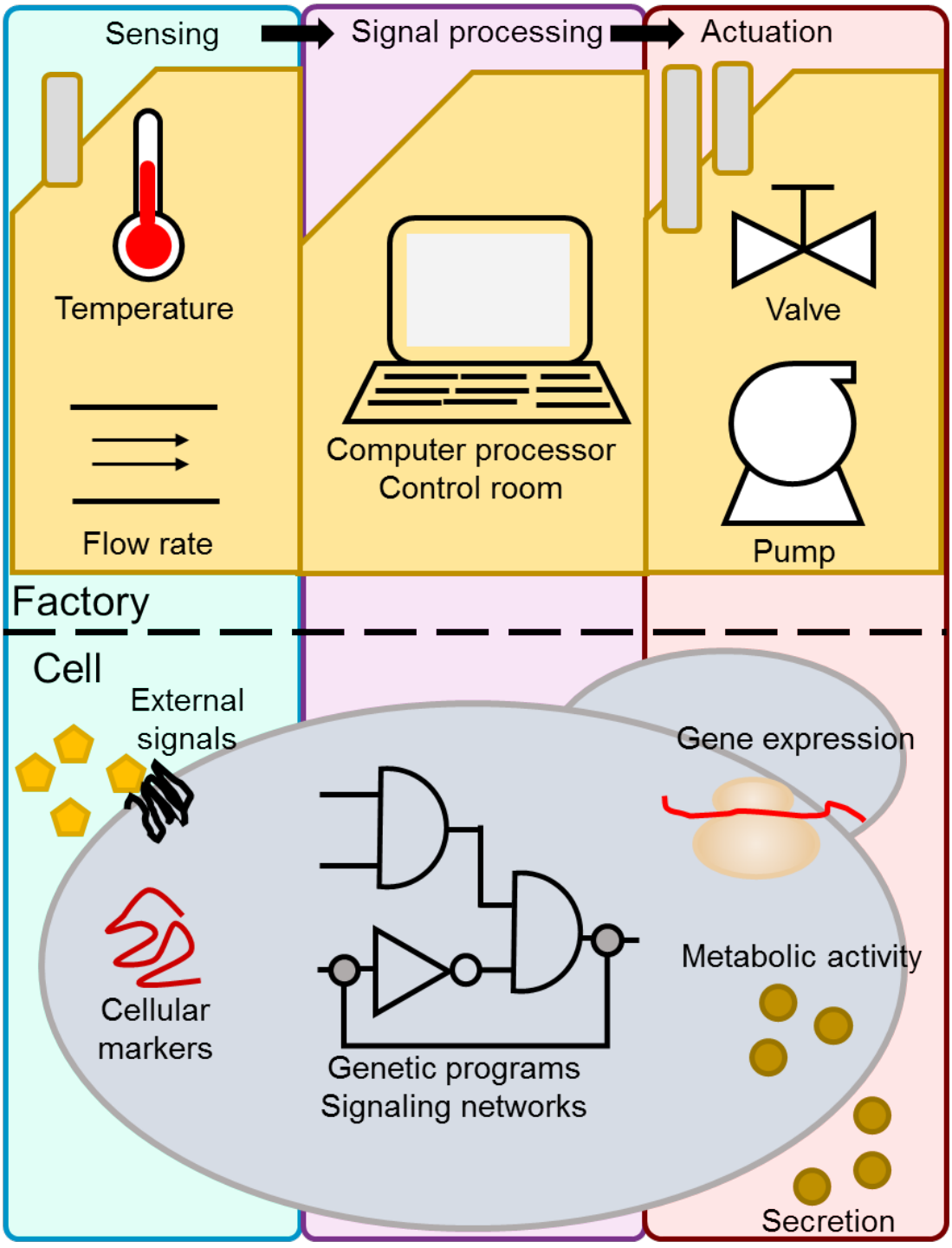


Figure 1.1 Analogous needs for control mechanisms in chemical processes operating in manufacturing plants and in cellular factories.

Tremendous progress has been made in making synthetic biological control elements inspired by electrical circuitry, ranging from the basic genetic toggle switch (Gardner, Cantor, and Collins 2000) and the repressilator (Elowitz and Leibier 2000) to the more complex memory elements (Friedland et al. 2009), oscillators (Elowitz and Leibier 2000; Atkinson et al. 2003; Stricker et al. 2008), and logic gates (Anderson, Voigt, and Arkin 2007; Maung and Smolke 2008). However, the application of these devices to engineer biological systems outside of the laboratory remains in its infancy (Khalil and Collins 2010). In contrast to naturally evolved biological control circuits that have been fine-tuned as interconnected networks over eons (Mitchell et al. 2009), current state-of-the-art synthetic biomolecular control systems often lack the capacities to sufficiently respond to dynamic changes and seamlessly interface with each other or native cellular machineries. These deficiencies limit the use of synthetic control modules as modern biotechnology and medicine require engineered biological systems to perform increasingly complex tasks with precision and accuracy.

To overcome these limitations, novel approaches to impose control need to be developed to allow access to information originating from both artificial and endogenous sources and perform the intended functions robustly across different platforms. The need for these novel approaches necessitates further engineering of biomolecules to mimic and integrate various aspects of natural control systems into functional control schemes. The following work represent our efforts to create new biomolecular control modules to serve as parts of a toolkit, categorized as sensors (Chapter 2), signal processors (Chapter 3), and actuators (Chapter 4), that can be used to customize engineered biological systems, as summarized in Figure 1.2.

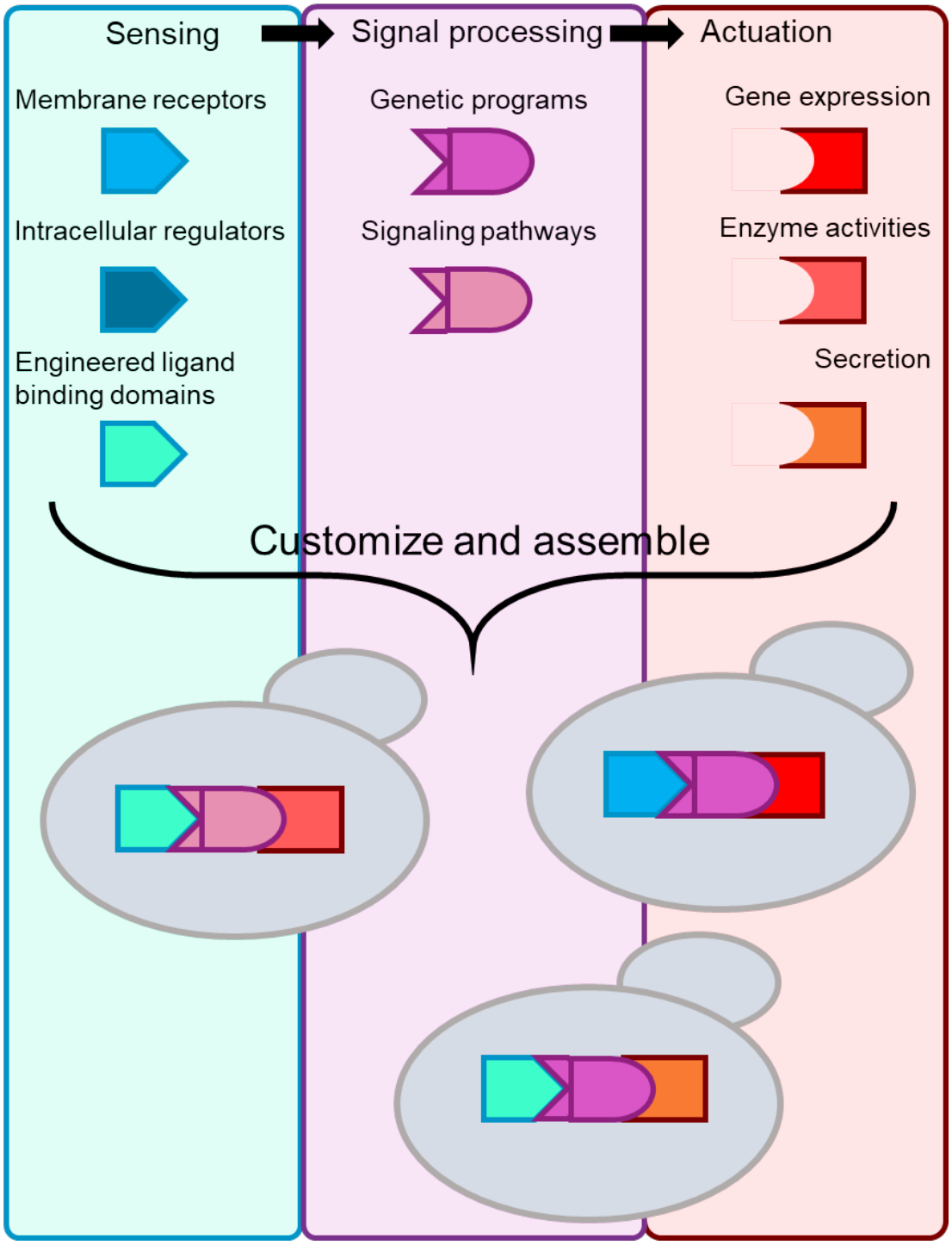


Figure 1.2 A toolkit of biomolecular control elements to tailor engineered organisms for specific functions.

1.2 Synthetic Sensing Circuits

Stimuli, both synthetically introduced and naturally occurring, need to be recognizable by sensory components to initiate any physiological responses (Doyle, Mangelsdorf, and Corey 2000; Koh 2002). In nature, living organisms are exposed to innumerable forms of stimuli, which has driven the evolution of countless sensing machineries. However, native sensors have not evolved to meet our needs in biotechnological and human health applications. As such, existing biosensors such as membrane receptors and intracellular factors must be further engineered to ensure robust performance outside of their native context (Conklin, Hsiao, and Claeysen 2008; Armbruster et al. 2007; Taylor et al. 2015).

Furthermore, synthetic applications present signals that are often difficult to detect with naturally evolved sensors because spatial barriers, such as cell membrane, may prohibit access of the relevant analytes, such as large polymeric molecules, or halt signal transduction to downstream processing circuits. One possible solution to circumvent these complications is to "export" the sensing functions served by membrane receptors and intracellular regulators to an entirely extracellular sensing circuit. This synthetic extracellular circuit must a) identify the presence of the target signal outside of the cell and b) transduce the recognition event into a form that is readily detected by the cell (Figure 1.3). Effectively, such a sensing circuit acts as a biomolecular "translator" that bridges the gap between a foreign input signal and native cellular responses, enabling the cell to react appropriately to the input, without the often tedious process of engineering membrane receptors or intracellular signaling pathways (Daringer et al. 2014). Because the central sensing functions occur outside of the cell, this strategy may avoid the potential pitfalls of undesirable crosstalk with

native signaling mechanisms that frequently plague intracellular sensory machineries (Zarrinpar, Park, and Lim 2003; Daringer et al. 2014).

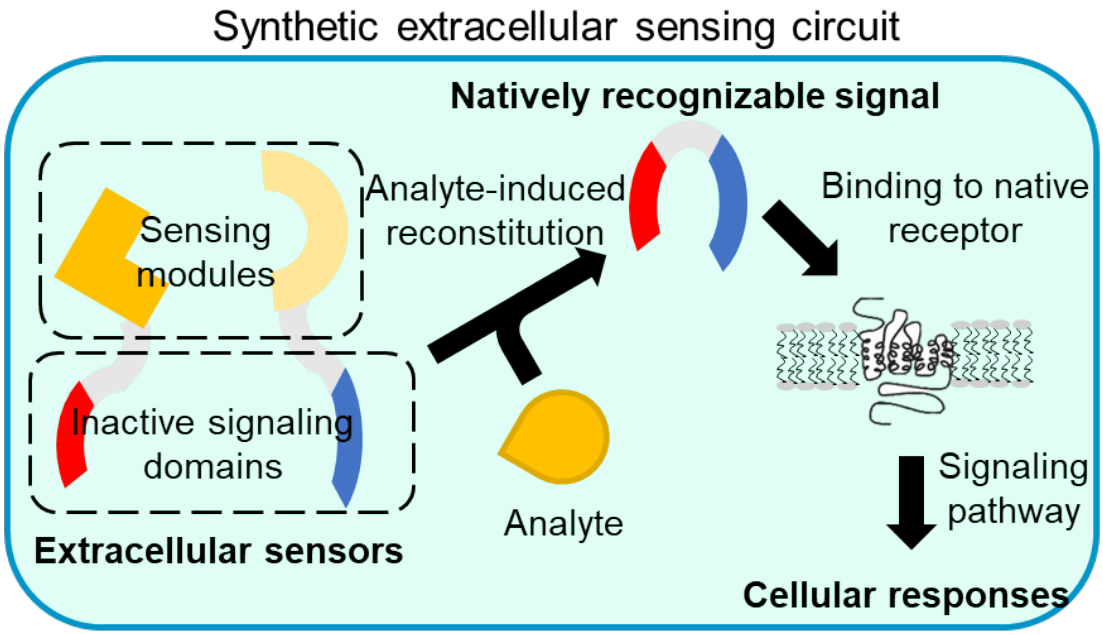


Figure 1.3 Synthetic extracellular sensing circuit to detect analyte of interest by reconstitution of natively recognizable signal.

Chapter 2 of this work details our attempts toward creating such an extracellular sensing circuit enabled by intein-mediated reactions. The primary challenge of constructing an easily adaptable extracellular sensor is resolved by the reconstitution of signaling peptides through the promiscuous post-translational intein-mediated reactions (K. H. Siu and Chen 2017). The reactions are largely independent of both the sensing function and the identity of the signal peptide, thus allowing this scheme to serve as the biochemical basis for the creation of synthetic sensing circuits in a variety of context.

1.2.1 Intein-mediated Post-translational Modifications

Protein splicing elements, termed inteins, are internal amino acid sequences embedded in various protein precursors and are excised by a post-translational, self-catalytic process (Perler et al. 1994). The removal of inteins from the primary translation product is concomitant with the formation of a new peptide bond joining the flanking sequences, called exteins, thus producing two separate proteins. Remarkably, this auto-catalytic process is able to proceed without any known cofactors, accessory proteins, or energy source. Many inteins have also been shown to retain their native activities when inserted into heterologous proteins (Lockless and Muir 2009). This robust and versatile ability to modify precursor proteins has been harnessed in a wide variety of applications, including purification (Wood et al. 1999), biosynthesis of cyclic proteins and peptides (Iwai, Lingel, and Pluckthun 2001; Tavassoli and Benkovic 2007), and gene delivery (J. Li et al. 2008).

In our study, we have not only used inteins to ligate split signaling peptides through protein *trans*-splicing, but also explored the use of a side reaction, N-terminal cleavage (Volkmann, Sun, and Liu 2009), to reconstitute functional signaling peptides. These biochemical reactions can be applied for the creation of logic-gated synthetic extracellular circuits, thus forming the mechanistic basis for activation of our synthetic extracellular sensors.

1.3 Biomolecular Integrated Control Elements for Signal Processing

Much like electronic systems, biological control systems contain logic elements that receive and process signals from sensory elements to decide what responses need to be executed (Khalil and Collins 2010; Lim 2010). However, unlike electronic circuit elements, which can be insulated spatially and electrically, biological

components constantly encounter and interact with one another within the complex cellular milieu. As a result, unspecific interactions and crosstalk between components inevitably arise and complicate the use of engineered control elements in biological systems. To address this fundamental limitation, new devices with high dynamic range, low system crosstalk, and flexible design are necessary to realize the multi-layered control circuits that underlie biotechnology and medicine.

Another hurdle for integrating synthetic control elements with native cellular machineries is the inability to utilize endogenous cellular information as inputs for activating or repressing synthetic responses. While many forms of conditional control systems have been developed, almost all rely on externally supplied inputs in the forms of small molecules (Spencer et al. 1993; Ho et al. 1996; Miyamoto et al. 2012) or light (Ballister et al. 2014). The requirement of external inputs prevents synthetic control elements from being implemented as autonomous regulatory mechanisms in engineered organisms that can robustly accomplish many tasks simultaneously.

In light of the limitations of currently available synthetic control schemes, we sought to develop a new class of control elements that can operate robustly in cellular environments. This led us to create a class of riboregulators based on gRNA scaffolds found in CRISPR/Cas9 systems termed toehold-gated gRNA (thgRNA) described in Chapter 3 (K.-H. Siu and Chen 2018) (Figure 1.4). These thgRNAs are controlled by toehold-mediated strand displacement. Activation of thgRNA is enabled by sequence-specific unblocking of the spacer region using a specific RNA trigger, which offers a high degree of orthogonality and very low system cross-talk with other un-related endogenous information. The flexibility to exploit endogenous RNAs for gene regulation further provides a simple interface between native signals and synthetic

transcriptional outputs. These valuable traits allow thgRNAs to be applied as flexible logic elements that are capable of integrating native cellular information in biological control systems.

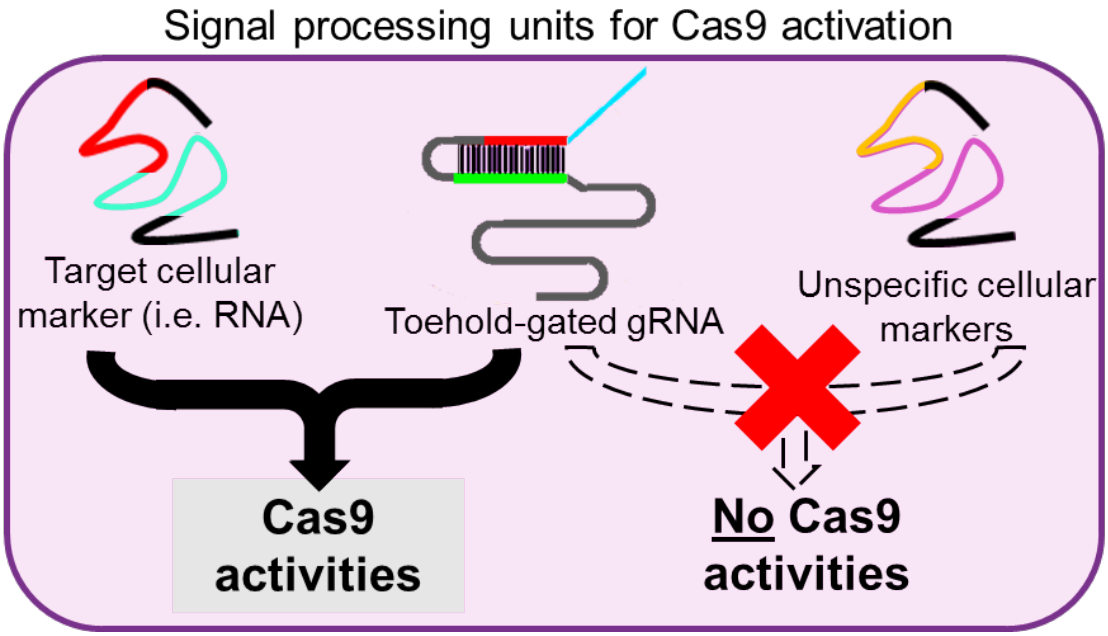


Figure 1.4 Application of thgRNAs as signal processing units for control of Cas9 functions.

1.3.1 CRISPR/Cas9 Functions and Applications

The Clustered Regularly Interspaced Short Palindromic Repeats, or CRISPR, system is a prokaryotic immune system that grants resistance to foreign genetic materials such as phage genomes and plasmids in a form of adaptive immunity (Horvath and Barrangou 2010; Sapranauskas et al. 2011; Jinek et al. 2012b). A simple subtype, CRISPR/Cas9, has gained immense popularity over the past few years because it only requires two molecules—the Cas9 protein and the sgRNA (Jinek et al.

2012b)— for it to be readily adapted as a host-independent gene-targeting platform. This simplicity has led to its widespread adoption as a versatile and efficient tool for genome engineering and gene regulation (Ran, Hsu, Wright, et al. 2013; Qi et al. 2013).

The ability of CRISPR/Cas9 to target DNA preceding a protospacer adjacent motif (PAM) using sequence complementarity of the spacer region of the sgRNA is well-documented (Jinek et al. 2012a; Nishimasu et al. 2014). However, the imperfect specificity of Cas9 activity (Hsu et al. 2013; Pattanayak et al. 2013) raises concerns over off-target effects that limit its usefulness in therapeutic and research applications. These concerns have led researchers to reduce off-target binding by either engineering the Cas9 protein (Ran, Hsu, Lin, et al. 2013; Tsai et al. 2014) or the sgRNA (Fu et al. 2014). As an alternative approach, some groups have focused on creating CRISPR/Cas9 systems that are dependent on post-translational control via an externally applied small molecule (Davis et al. 2015; W. Tang, Hu, and Liu 2017) or light (Nihongaki et al. 2015; Zhou et al. 2017).

To enable autonomous control over transcriptional activities inside the cell, a method to control Cas9 activities without the need of external input is desirable. Binding of Cas9-sgRNA complex to its DNA target depends heavily on the availability of the spacer, particularly the so-called “seed” region, of the sgRNA to hybridize with the targeted strand (Gong et al. 2018). Sequestering the spacer region with a RNA hairpin structure should therefore prevent binding and inhibit Cas9 activities. The reconfiguration of the sequestered structure to unblock the spacer region would provide us the ability to control Cas9 activities. This was accomplished

by exploiting a non-enzymatic reaction known as toehold-mediated strand displacement described below.

1.3.2 Toehold-mediated Strand Displacement

While traditional nucleic acids-based nanotechnology has focused on the design and application of static, self-assembled structures, recent advances in computational tools and our understanding of nucleic acids' biophysical properties have enabled the design and construction of assemblies with dynamic behaviors. In particular, the thermodynamic properties of nucleic acids allow a complementary single-stranded oligonucleotide to invade and displace a pre-hybridized strand (D. Y. Zhang and Seelig 2011). The displacement reaction occurs when one of the pre-hybridized strands contains an extended region to which only the invading strand can hybridize; this extended complementary region is called a toehold as it serves as the initiating site of the displacement reaction. Once the invading strand binds onto the toehold region of a pre-hybridized strand, it begins to unwind the sequestered strand from the pre-hybridized complex by means of "branch migration", eventually resulting in full displacement (Figure 1.5) (D. Y. Zhang and Winfree 2009). The strand displacement reaction rate can be quantitatively controlled simply by varying the length of a toehold region (Zhang and Seelig 2011) or creating mismatched complementary base-pairs in the toehold region (Broadwater and Kim 2016). The ability to control the rate and efficiency of displacement allows the creation of highly functional nucleic acid circuits, even in an intracellular setting (Groves et al. 2015).

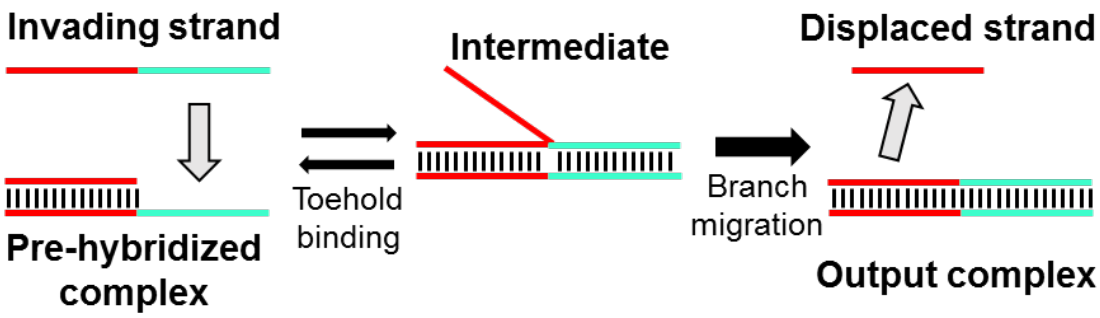


Figure 1.5 Toehold-mediated strand displacement.

We were interested in exploiting this remarkable level of control based solely on sequence complementarity between two strands. In particular, we sought to exploit the abilities of toehold-mediated strand displacement to change RNA configurations inside cells and potentially distinguish different RNA strands down to the single-nucleotide level. In conjunction with the appealing properties of the CRISPR/Cas9 system discussed previously, these traits of the displacement reaction enabled us to design intracellular riboregulators in the form of our thgRNA capable of using RNA transcripts as triggers.

1.4 Actuation through Control of Intracellular Protein Interactions

Interactions between proteins play a dominant role in all aspects of dynamic responses in biological systems. Precise connections between components are essential to actuate the appropriate cellular machineries and execute any particular responses (Bashor et al. 2008; Dueber et al. 2003; R. Chen et al. 2014). Consequently, protein interactions must be tightly regulated in time and space within the confines of the cell while remaining responsive to changes in cellular conditions. These regulatory mechanisms typically take the form of conditional activation or inhibition of binding

motifs through post-translational modifications such as phosphorylation and ubiquitination, resulting in changes of the binding motifs' characteristics.

Despite the recognition that dynamic control can only be achieved with stimuli-responsive elements, synthetic modules designed to bridge the gap between two given cellular pathways have remained mostly static. The primary challenge in creating dynamic interaction domains is the lack of auxiliary regulatory mechanisms that can reconfigure the synthetic modules' binding characteristics. Without the ability to directly change protein conformations through post-translational modifications (e.g. phosphorylation, ubiquitination, etc.), artificial interactions can only be controlled by either addition of external inputs such as chemicals, light-induced dimerization, or turnover of static interaction motifs. None of these mechanisms are capable of autonomous, reversible adjustments of protein interactions that define dynamic control systems. To address these deficiencies, we have employed a class of protein structural motifs known as coiled-coils as synthetic actuators and exploited their capacity to partake in strand displacement to impose dynamic controls over protein interactions in Chapter 4 (Figure 1.6).

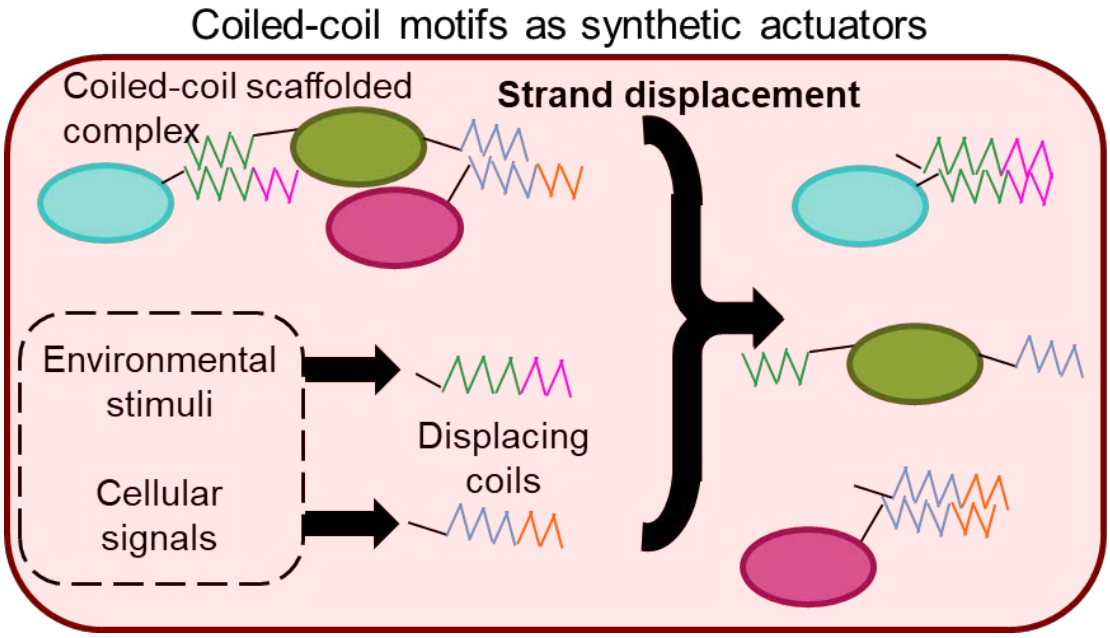


Figure 1.6 Coiled-coil strand displacement as mechanism for actuating dynamic response

1.4.1 Coiled-coil Motifs: Structure, Design, and Strand Displacement

Coiled-coil structural motifs were first independently proposed by Linus Pauling and co-workers (Pauling, Corey, and Branson 1951) as well as Crick (Crick 1952) as theoretical recurring domains present in α-keratins. These structures are made of heptad amino acid repeats, usually labeled *abcdefg*, where *a* and *d* are the hydrophobic positions often occupied by isoleucine, leucine, or valine. These hydrophobic residues forms a "stripe" that coils around the α-helical secondary structure upon folding of the repeating heptad pattern, forming an amphipathic helix. For such amphipathic helices to arrange themselves in aqueous solution, the most thermodynamically favorable conformation is to wrap the hydrophobic "stripes" against each other sandwiched between the hydrophilic amino acids. (Pauling, Corey, and Branson 1951) Thus, the oligomerization of coiled-coils is thermodynamically

driven by the minimization of hydrophobic surfaces in contact with bulk solution (Crick 1952).

A substantial body of work in coiled-coil motifs has followed the first *de novo* coiled-coil design reported by Hodges and colleagues (Hodges et al. 1981; Kohn and Hodges 1998; O'Shea, Lumb, and Kim 1993; Burkhard, Ivaninskii, and Lustig 2002; Reinke, Grant, and Keating 2010; Gradišar and Jerala 2011). These studies have yielded progressively more clear principles and guidelines that are now being incorporated into *in silico* algorithms for designing new coiled-coil motifs (Harbury et al. 1998; Negron and Keating 2014). Aided by these advanced computational tools, it is now possible to design nanostructures of considerable sophistication that self-assemble from different coiled-coil motifs *a priori* (MacPhee and Woolfson 2004; W. M. Park et al. 2017). This deep level of understanding facilitates the use of these peptide motifs to dictate interactions with unusually high reliability and predictability, a unique feature that we exploit in our design for a synthetic actuator to control protein proximity.

We also sought to apply dynamic coiled-coil systems that enable the controlled and reversible dissociation of complexes in an intracellular environment. Recently, Gröger et al. demonstrated reversible coiled-coil association using pairs of heterodimeric coiled-coil based on the Hodges EK peptides by modulating the lengths of the individual coils (Gröger, Gavins, and Seitz 2017). The varying lengths of the coiled-coils resulted in a range of dissociation constants of μM to sub-nM, allowing longer coils to dynamically displace shorter coils akin to strand displacement for nucleic acids discussed in Chapter 1.3.2 above (D. Y. Zhang and Winfree 2009). Kinetic studies on the displacement reaction have further elucidated the characteristics

of this dynamic phenomenon (Groth et al. 2018). Accordingly, we set out to expand the use of coiled-coil motifs for dynamic control schemes in Chapter 4.

Chapter 2

SYNTHETIC EXTRACELLULAR SENSING CIRCUIT BY INTEIN-MEDIATED RECONSTITUTION OF YEAST MATING FACTOR

This Chapter is adapted with permission from Siu, Ka Hei, and Wilfred Chen. 2017. "Control of the Yeast Mating Pathway by Reconstitution of Functional α-Factor Using Split Intein-Catalyzed Reactions." *ACS Synthetic Biology*. doi:10.1021/acssynbio.7b00078. (Appendix B.1)

2.1 Summary

One of the most common mechanisms for surveying environmental stimuli and adapting to changing conditions employed by natural living systems is the use of signaling pathways that are able to sense an external input of interest, transmit the signal inside the cell, and actuate appropriate cellular responses. Mimicking these natural mechanisms, synthetic biosensors have been devised to detect signals in and out of the cell. However, synthetic control strategies using signaling peptides to regulate and coordinate cellular behaviors in multicellular organisms and synthetic consortia remain largely underdeveloped because of the complexities necessitated by heterologous peptide expression. Using recombinant proteins that exploit split intein-mediated reactions, we present here a new strategy for reconstituting functional signaling peptides capable of eliciting desired cellular responses in *S. cerevisiae*. These designs can potentially be tailored to any signaling peptides to be reconstituted, as the split inteins are promiscuous and both the peptides and the reactions are

amendable to changes by directed evolution and other protein engineering tools, thereby offering a general strategy to implement synthetic control strategies in a large variety of applications.

2.2 Introduction

Living systems of all scales, from unicellular microbes to multi-species communities, must constantly adapt to changing conditions by continually surveying their surroundings for relevant stimuli. One of the most common mechanisms for this continual process is the use of signaling pathways that are able to sense an external input of interest, transmit the signal inside the cell, and actuate appropriate cellular responses such as transcriptional upregulation (Lim 2010). Often, this transduction mechanism involves a signaling molecule and its cognate receptor, as frequently seen in the myriad of G-protein coupled receptors (GPCRs) present in eukaryotes (Venkatakrishnan et al. 2013). While a limited number of signaling molecules and their associated pathways (i.e. quorum sensing) have been successfully engineered into different species (M.-T. Chen and Weiss 2005; Collins, Leadbetter, and Arnold 2006), more complex signals such as peptide hormones have not yet been widely incorporated into synthetic control schemes. As opposed to the small molecules used by bacterial quorum sensing systems, peptide signaling molecules are often difficult to secrete, and engineering of their cognate membrane bound GPCRs has proven to be even more challenging because of their very poor expression in recombinant hosts (Sarramegna et al. 2006; Lundstrom et al. 2006). As a result, the use of GPCRs and their associated signaling pathways for orthogonal control of cellular behaviors remain elusive.

Despite these difficulties, there are tremendous scientific and clinical interests to adopt these peptide-mediated signaling pathways as synthetic extracellular sensing circuits because of their importance in cell-to-cell communication and disease progression (Rozengurt, Sinnett-Smith, and Kisfalvi 2010; Lappano and Maggiolini 2011). We present here a new generalizable framework for designing synthetic extracellular sensing circuits that could be adapted for any signaling peptide of interest. Our design involves the use of split inteins to facilitate the functional reconstitution of inactive signaling peptides through either protein *trans*-splicing (PTS) or N-terminal cleavage (NTC) reactions (Figure 2.1). In PTS, two polypeptides, called the N- and the C-inteins, are able to associate with each other and trigger a series of biochemical reactions that result in the self-excision of the intein sequences from the protein complex and, concomitantly, the formation of a new peptide bond between their flanking sequences, called the exteins (Figure 2.2, left) (Muir 2003). By mutating a key residue at the catalytic site of the intein complex (N146A), the PTS reaction pathway can be redirected to NTC in which the N-extein is cleaved from the complex instead of being ligated to the C-extein. (Figure 2.2, right) (Amitai et al. 2009; Shah et al. 2012). These reactions are applied to reconstitute the yeast mating pheromone signal peptide, α-factor, from recombinant proteins. The reactions thus serve as the basic biochemical mechanisms to activate yeast mating GPCR signaling pathway and control cellular responses.

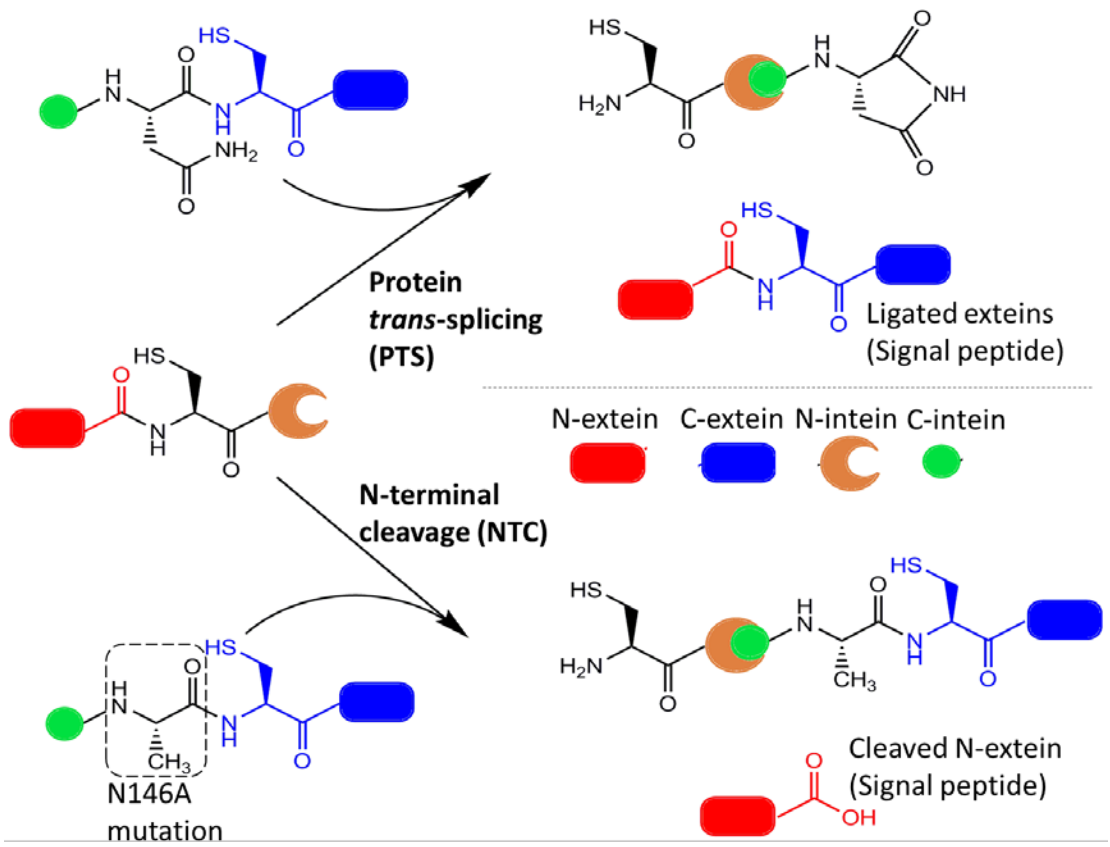


Figure 2.1 Protein *trans*-splicing (PTS) and N-terminal cleavage (NTC) reactions catalyzed by split inteins.

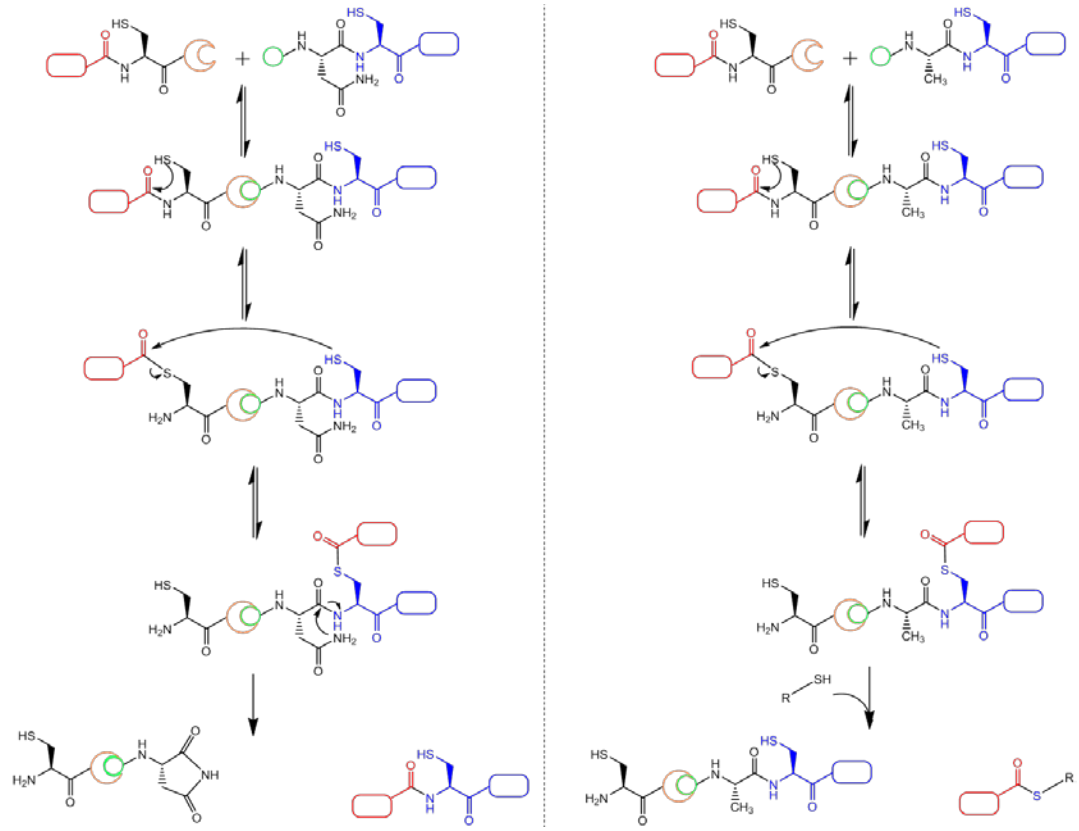


Figure 2.2 Reaction mechanisms for protein *trans*-splicing (PTS, left) and N-terminal cleavage (NTC, right). Note the end products for PTS result in the formation of a new peptide bond between the N- and C-exteins while the N-extein of NTC reactions are cleaved from the reaction intermediate by a nucleophile that forms a new thioester bond. In our study, the nucleophile is a reducing agent (DTT), which leads to subsequent hydrolysis, leaving a carboxyl group at the C-terminus for the newly released N-extein.

2.3 Materials and Methods

2.3.1 Peptide Design and Preparation.

A small library of synthetic analogues of α–factor were designed and listed in Table 2.1.

Table 2.1 Library of synthetic α-factor analogues screened for growth arrest activities

	Sequence[a]	Number of residues	Clearing formation[b]	EC50 (nM) [c]
1	WHWLRLG^CGQPMY	13	Y	100
1-B[d]	MWHWLRLG^CGQPMY	14	Y	200
2	WHWLRL^CAGQPMY	13	Y	150
2-B[d]	MWHWLRL^CAGQPMY	14	Y	425
3	WHWLRLG^CFQPMY	13	N	N.D.
4	WHWLRLGA^CGQPMY	14	Y	N.D.
5	WHWLRLG^CAGQPMY	14	Y	N.D.
6	WHWL^CLSPGQPMY	13	Y	450
7	WHWL^CLQPGQPMY	13	Y	400

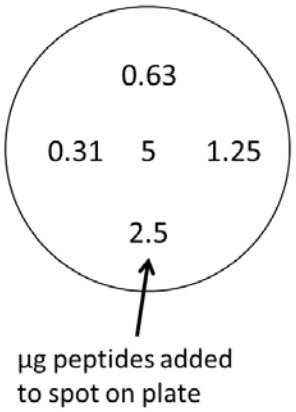
[a] Residues that differ from WT are underlined. ^ mark the intended split site of the synthetic analogue when fused onto split inteins. [b] Clearing formation observed from agar plate-based growth arrest assays. [c] EC50 defined as the approximate concentration required to achieve 50% growth inhibition of BY4741 bar1Δ cells growing in YPD (Figure 2.6). [d] "B" denotes an N'-Methionine added variant of the original peptide sequence to account for expression of the final recombinant proteins in *E. coli*.

All peptides were purchased from GenScript, Piscataway, NJ. The chemically synthesized peptides were purified, lyophilized, and analyzed by MS and HPLC to ensure ≥70% purity by the company. Prior to use, lyophilized peptides were reconstituted in 100% DMSO to stock concentrations of 10 mg/mL. Sterile water was used to dilute stock peptide solutions where needed.

2.3.2 Screening of Synthetic α–factor Analogues.

Individual yeast BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 bar1::Kan^R) colonies were picked from YPD (10g/L yeast extract, 20g/L peptone, 20g/L dextrose) agar plates and grown overnight in 3mL YPD. Overnight cultures were used to inoculate 3mL of fresh YPD at an initial OD_{600} of ~0.05 and grow to exponential phase (OD_{600} of ~0.3-0.5). The growing yeast cultures were thoroughly mixed into sterile, melted 0.75% agar at 50°C at an OD_{600} of ~0.005 and quickly poured over

YPD agar plates. Once the soft overlay containing yeast solidified, 5µL of α–factor analogue solutions of desired concentrations were added (Figure 2.3) and air-dried. Dried plated were incubated at 30°C for 36hrs prior to visual inspection and imaging. Images were taken using the Gel Doc Ez system (Bio-Rad Laboratories, Inc., Hercules, CA, USA).



	Sequence[a]	Number of residues	MW (g/mol)	5µg = __nmol	[peptide] (mM, 5µg in 5µL)
1	WHWLRLG^CGQPMY	13	1646.94	3.04	0.607
1-B	MWHWLRLG^CGQPMY	14	1778.14	2.81	0.562
2	WHWLRL^CAGQPMY	13	1660.97	3.01	0.602
2-B	MWHWLRL^CAGQPMY	14	1792.17	2.79	0.558
3	WHWLRLG^CFQPMY	13	1737.07	2.88	0.576
4	WHWLRLGA^CGQPMY	14	1718.02	2.91	0.582
5	WHWLRLG^CAGQPMY	14	1718.02	2.91	0.582
6	WHWL^CLSPGQPMY	13	1617.9	3.09	0.618
7	WHWL^CLQPGQPMY	13	1658.95	3.01	0.603

Figure 2.3 Layout of plate-based growth arrest assay and relevant information. Each peptide variant is spotted onto YPD-agar plates and air-dried prior to addition of a soft agar overlay containing growing yeast cells. The plates are then incubated at 30°C for 36hrs prior to imaging and analysis.

All synthetic analogues that resulted in clearings on plates were further subjected to a liquid growth assay to quantify their biological activities. Similar to the plate-based assay, individual colonies were picked and grown overnight. Overnight cultures were used to inoculate 3mL of fresh YPD at an initial OD_{600} of ~0.05 and grow to early exponential phase (OD_{600} of ~0.2-0.25). Varying concentrations of α-

factor analogues were added to these cultures, which were then incubated at 30°C for 16 hours with shaking before OD_{600} were measured (Figure 2.4). The final OD_{600} were normalized to the OD_{600} of control samples with no peptides added. By plotting the normalized OD_{600} against the concentrations of peptides added on a semi-log plot, the half maximal effective concentration (EC_{50}) of each analogue could be identified.

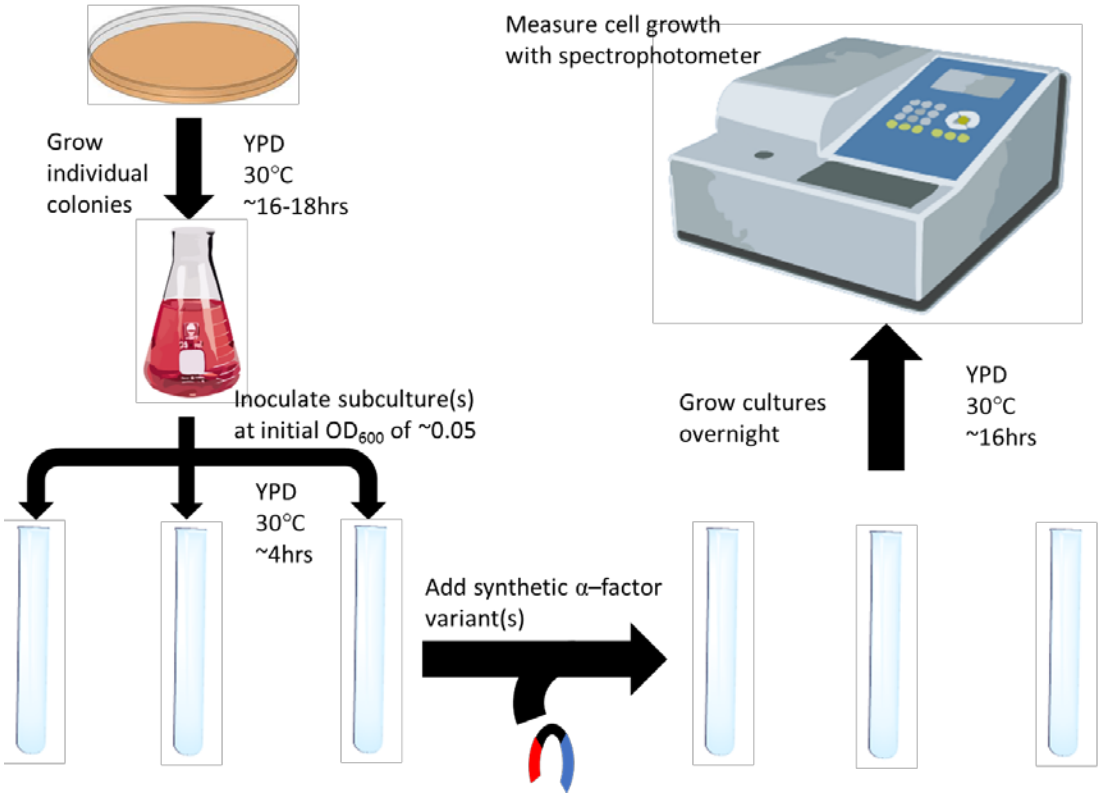


Figure 2.4 Workflow of growth arrest assay in liquid media.

2.3.3 Construction of Expression Plasmids.

The plasmids encoding all constructs were created by standard subcloning techniques and are listed in Table A.1.1. Briefly, genes for the *Npu* split inteins were

amplified from pI plasmids from Ramirez et al. (Ramirez et al. 2013) using primers with additional flanking sequences encoding split α–factors, His-tag, and restriction sites as needed (Table 2.2).

Table 2.2 Oligonucleotides used for construction of plasmids

#	Sequence	Purpose
1	tatacatatgTGGCATTGGTTGAGATTGGGTTGTTTAAGCTATGAAACGGA	Forward primer for α^{N1}-*Npu*N
2	atatatgagctcATTCGGCAAATTATCAACCCGCATC	Reverse primer for *Npu*N
3	agcagccatggggctcccagTACAAATTAATCCTTAATGGTAAAACATTG	Forward primer for GB1 (*Npu*N fusions)
4	tggtggtggtggtgctcgagTTCAGTAACTGTAAAGGTCTTAGTCGC	Reverse primer for GB1 (*Npu*N fusions)
5	gatatacatatgTGGCATTGGCTGCGCCTGTGTTTAAGCTATGAAACGGAAATATTG	Forward primer for α^{N2}-*Npu*N
6	gatatacatatgAAAATCGAAGAAGGTAAACTGGTAATC	Forward primer for MBP
7	atgccacatatgCGAGCTCGAATTAGTCTGCG	Reverse primer for MBP
8	atacatatgcaccaccaccaccaccaccagTACAAATTAATCCTTAATGGTAAAACATTG	Forward primer for HHHHHH-GB1 (*Npu*C fusions)
9	ttacgtgtggctattttgatggagccTTCAGTAACTGTAAAGGTCTTAGTCGC	Reverse primer for GB1 (*Npu*C fusions)
10	ctttacagttactgaaggctccaTCAAAATAGCCACACGTAAATATTTAGGC	Forward primer for *Npu*C
11	acgtgctcgagtcaATACATTGGTTGACCACAATTAGAAGCTATGAAGCCATT	Reverse primer for *Npu*C-α^{C1}
12	acgtgctcgagtcaATACATCGGCTGGCCCGCGCAATTAGAAGCTATGAAGCCATTTTTG	Reverse primer for *Npu*C-α^{C2}
13	gcctcgagacaagaacaAGCAGAAGCTATGAAGCCATTTTTGAG	Reverse primer for *Npu*C(N146A)

Underlined portions represents restriction sites. Upper case letters indicate binding region to gene of interest.

Amplified products were purified by gel electrophoresis. Purified products and pET-24a(+) vector (MilliporeSigma, St. Louis, MO, USA) were digested using

restriction enzymes (New England BioLabs®, Inc., Ipswich, MA, USA) as needed. Recombinant constructs were made by ligating digested vectors and inserts using T4 ligase (New England BioLabs®, Inc., Ipswich, MA, USA), followed by heat-shock transformation into NEB5®α cells (New England BioLabs®, Inc., Ipswich, MA, USA). Transformed cells were screened by restriction digests and the sequences of each expression construct were further confirmed by sequencing.

2.3.4 Protein Expression and Purification.

E. coli BL21-Gold (DE3) cells (Agilent Technologies, Cedar Creek, TX) were transformed with individual expression plasmids. Transformed cells were grown in 3mL TB (Terrific Broth, 12g/L tryptone, 24g/L yeast extract, 4mL/L glycerol, 17mM KH_2PO_4, 72mM K_2HPO_4) at 37°C overnight. Overnight cultures were used to inoculate 25mL of fresh TB to an initial OD_{600} of ~0.05 and grow to mid-exponential phase (OD_{600}~0.75) at 37°C. Protein expression was induced by addition of 200μM isopropyl-β-thiogalactopyranoside (IPTG) at 20°C for ~16hrs. Induced cells were then harvested by centrifugation at 3000*g* for 10 min, resuspended in lysis buffer (50 mM Tris-Cl, 150 mM NaCl. pH 7.5) to an OD_{600} of ~15, and lysed by sonication. Insoluble materials were removed by centrifugation at 10000*g* for 30 min. Desired fusion proteins were purified from soluble supernatants by affinity chromatography using His-Bind resin (MilliporeSigma, St. Louis, MO, USA) columns. Purified proteins were dialyzed against Reaction Buffer (50mM Tris-HCl, 150mM NaCl, 1mM EDTA, 10% v/v glycerol, 5mM TCEP, pH 7.4). Dialyzed protein solutions were analyzed by SDS-PAGE and Bradford assays to estimate their concentrations and purities. These protein solutions were used directly in *in vitro* reactions as described below.

2.3.5 *In vitro* Reconstitution of Functional α–factor from Split Intein Fusions.

All *in vitro* reactions were carried out in Reaction Buffer with TCEP added immediately prior to the start of each reaction to a final concentration of 5 mM. Fusion proteins were mixed into the reaction mixtures at concentrations as indicated. The assembled reaction mixtures were incubated at room temperature on a rotator. Samples were collected at times as noted.

For reactions involving constructs **N1B**/**C1**, **N2B**/**C2**, and **N3B**/**C3**, where further analyses by SDS-PAGE were needed, the reactions were quenched by addition of Gel Loading Buffer (5x, 300 mM Tris-HCl, 10% w/v SDS, 25% v/v glycerol, 0.5% w/v bromophenol blue, 200 mM β-mercaptoethanol, pH 6.8) to collected samples and heated to 95°C for 5 minutes. After quenching, the reaction products were loaded onto 10% or 12% polyacrylamide gels followed by electrophoresis. Standard Coomassie staining or Western Blotting techniques were used to visualize the results. Densitometry analysis was performed using ImageJ.

2.3.6 Yeast Growth Arrest Assays using Products of Intein-catalyzed Reactions.

Yeast BY4741 bar1Δ cells (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 bar1::Kan^R) were grown from single colonies in 3mL YPD at 30°C overnight. Overnight cultures were used to inoculate 25mL of fresh YPD to an initial OD_{600} of ~0.05 and grow to early exponential phase (OD_{600}~0.2-0.25) at 30°C. 900μL of this culture was induced with 100μL of reaction products, resulting in 10-fold dilutions of final concentration of α–factor where products were present. The resultant mixtures were incubated at 30°C for 16 hours to induce growth arrest. OD_{600} of these cultures were then measured and normalized to the average of cultures with only reaction buffer added.

2.3.7 Transcriptional Induction of GFP in Yeast.

Yeast W303-derived cells with GFP under the control of pFUS1 mating responsive promoter (MATa far1Δ mfa2::pFUS1-GFP bar1::Nat^R his3 trp1 leu2 ura3, courtesy of Prof. Wendell Lim of UCSF) were grown and induced using the same procedures as in the growth arrest assays described above. Instead of 16 hours of induction, OD_{600} and fluorescence (Ex: 475nm, Em: 515nm) of the induced cultures were measured using a microplate reader after 3 hours of induction. The specific fluorescence were determined by dividing the fluorescence reading by OD_{600} measurements. Samples were also examined using flow cytometry. Collected samples were further pelleted, lysed by disruption with glass beads, and analyzed with SDS PAGE and blotted against anti-GFP. For testing NTC reactions in living cultures, 1μM of purified proteins in Reaction Buffer were added directly into the culture medium with growing yeast at OD_{600} ~ 0.5. Specific GFP fluorescence were determined using the same protocol as noted above 3 hours after the addition of the proteins.

2.4 Results and Discussions

2.4.1 Active Synthetic α-factor Analogues that Accommodates Intein Reactions can be Created

To demonstrate the feasibility of our strategy, the well-characterized α-factor (Figure 2.5a), which activates the yeast mating pheromone response pathway (Figure 2.6) upon binding onto its native receptor, Ste2p (Naider and Becker 1986; Naider and Becker 2004), was used as a model signaling peptide. This pathway is ideally suited to establish and validate our approach as it has been used as a model for all eukaryotic GPCR signaling pathway and engineered to perform a variety of synthetic functions with tunable behaviors (S.-H. Park, Zarrinpar, and Lim 2003; Youk and Lim 2014).

The yeast α-factor is a 13-amino acid peptide composed of three distinct domains: residues that initiate signaling in the N terminus, residues that mediate binding to Ste2p in the C-terminus, and a flexible loop region in the middle to orient the signaling and binding domains (Figure 2.5a) (Naider and Becker 2004). Since mutations to either the N- or C- terminal region essentially abolish GPCR activation (S.-H. Park, Zarrinpar, and Lim 2003; Youk and Lim 2014), these results indicate that both GPCR binding and activation are necessary for correct pathway signaling. We hypothesized that split α-factor fragments composed of either the N- or C- terminal region that are incapable of activating the mating pathway can be created for intein-mediated functional reconstitution.

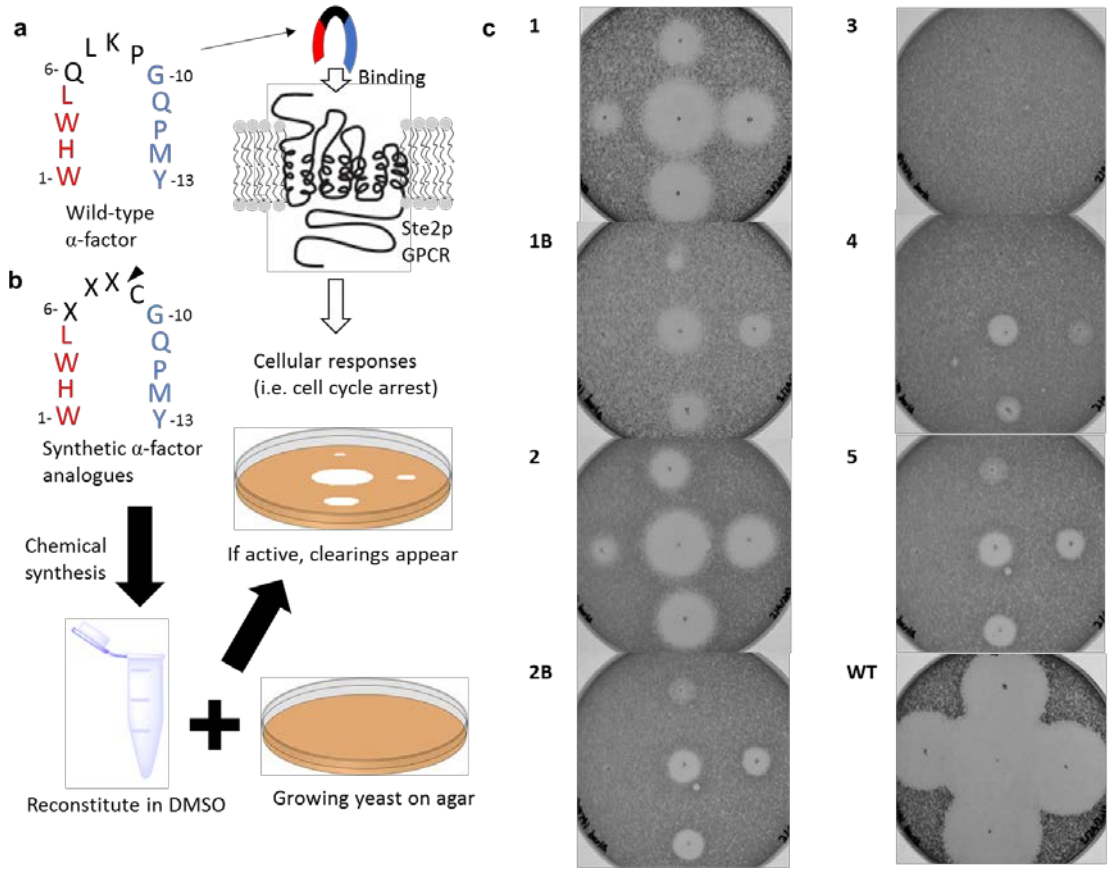


Figure 2.5 Adapting yeast mating pheromone, α-factor, for synthetic control. a) Wild-type α-factor sequence binds to its native receptor and triggers downstream cascade that results in many cellular responses, including cell cycle arrest. b) A small library of α-factor analogues (see Table 2.1 for complete list) with substitutions (marked by X) to insert the Cys residue required for PTS was designed and chemically synthesized, with the prospective split site between residues 8 and 9 (marked by ▲). The synthesized analogues were added to growing yeast on agar in increasing amounts to evaluate their activity. c) Analogues 1 and 2 displayed the highest activities relative to the wild-type (WT). For variants 1B and 2B, an N-terminal Met residue was added to peptides 1 and 2 to mimic the final spliced product after fusion proteins were expressed from *E. coli*.

For effective intein ligation, an absolute requirement is a Cys residue at the +1 position after the C-extein (Manfredi et al. 1996). Since a large variety of residues

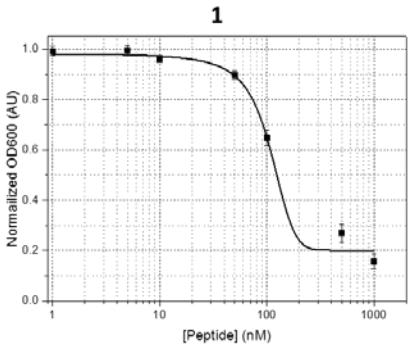
have been substituted into the loop domain of α-factor without affecting biological activity (Manfredi et al. 1996), we hypothesized that a synthetic α-factor analogue that includes the required residues for PTS while also retaining biological activities could be created. To test this hypothesis, we designed a small library of synthetic α-factor analogues (Table 2.1) and screened them using a simple plate-based growth arrest assay (Figure 2.3 and Figure 2.5b) based on the induction of cell cycle arrest in the G1 phase upon activating the mating pathway (Ishii et al. 2006). By exploiting the easily observable growth arrest phenotype, we successfully identified two synthetic analogues (1 and 2) that retain significant levels of biological activities by correlating the size of the clearing zone around a lawn of growing yeast cells (Figure 2.5c) (Manfredi et al. 1996).

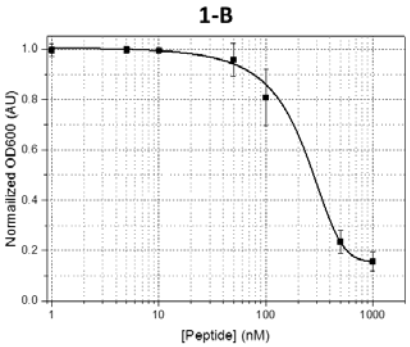
Based on the results of the semi-quantitative plate assay, we further characterized and quantified the biological activity of each synthetic analogue using a liquid-based growth assay (Table 2.1, Figure 2.4 and Figure 2.6). While we observed that almost all synthetic analogues displayed a varying degree of activity, analogues 1 and 2 were again significantly more active than others. Additionally, the subset of analogues that mimicked the length of the wild type peptides (1, 2, and WT) was at least 2 times more active than their extended counterparts (1B, 2B, and WT-B) (Table 2.1). It is likely that the longer peptides distorted and/or displaced their termini from effective interactions with the Ste2p receptor.

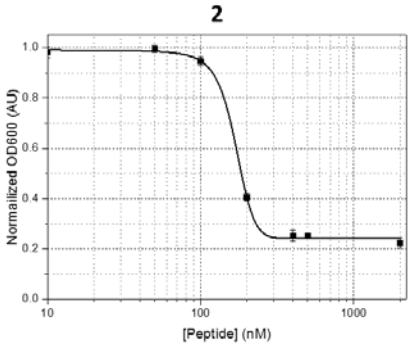
2.4.2 Reconstitution of Split α-factor Fragments by Protein *Trans*-splicing

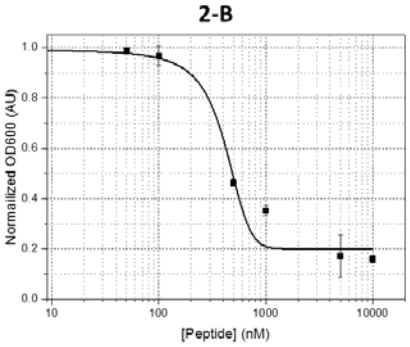
We chose 1 and 2 to be genetically split and fused onto the naturally split *Npu* DnaE inteins (Zettler, Schütz, and Mootz 2009) as constructs N1/C1 and N2/C2 (an N-terminus Met is added resulting in 1B and 2B as the final reconstituted peptide

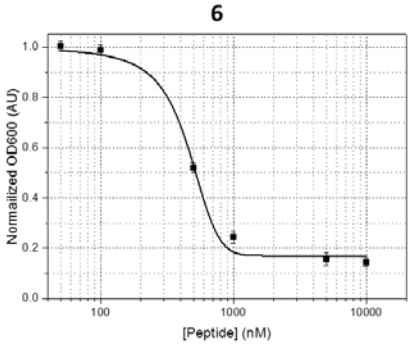
products) (Figure 2.7). To ensure these split fragments have no biological activity, all four fusion proteins were expressed in *E. coli* and purified using the flanking His6 tag. None of the split fragments was shown to elicit growth arrest as assessed by the liquid-based growth assay (Figure 2.9), an observation consistent with the requirement of both binding and activation for proper signaling.

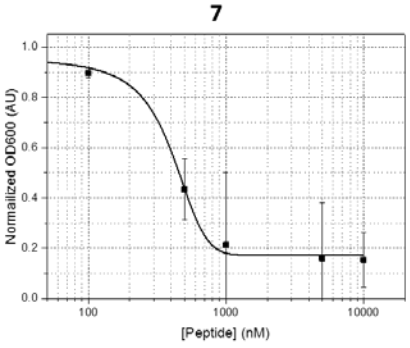

1
Normalized OD600 (AU)
[Peptide] (nM)

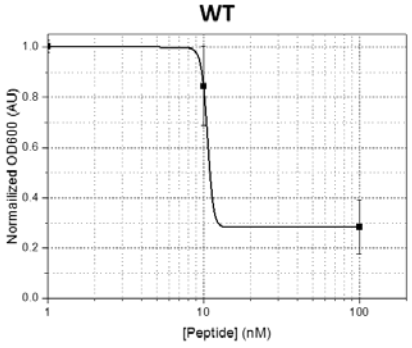

1-B
Normalized OD600 (AU)
[Peptide] (nM)


2
Normalized OD600 (AU)
[Peptide] (nM)


2-B
Normalized OD600 (AU)
[Peptide] (nM)


6
Normalized OD600 (AU)
[Peptide] (nM)


7
Normalized OD600 (AU)
[Peptide] (nM)


WT
Normalized OD600 (AU)
[Peptide] (nM)

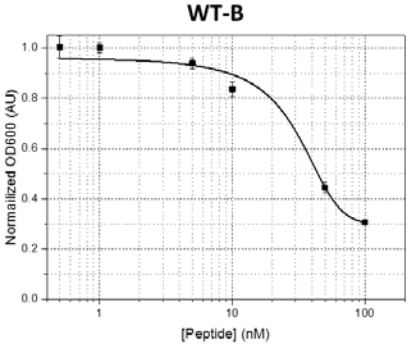
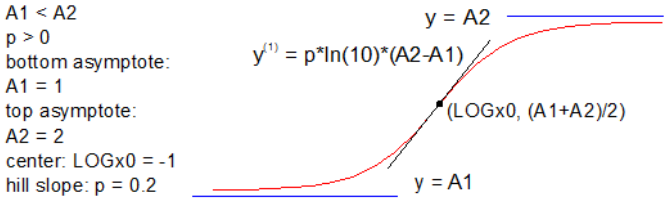

WT-B
Normalized OD600 (AU)
[Peptide] (nM)

Figure 2.6 Data and dosage curves for synthetic peptide variants tested. Dosage curves were fitted to the equation:

$$y = A1 + \frac{A2 - A1}{1 + 10^{(LOGx0 - x)p}}$$



using Origin software with a Chi-squared tolerance of 1e-9. Because of the use of normalized growth data, A2 (top asymptote) is fixed at 1. EC50 values are defined as the center of the curve (LOGx0).

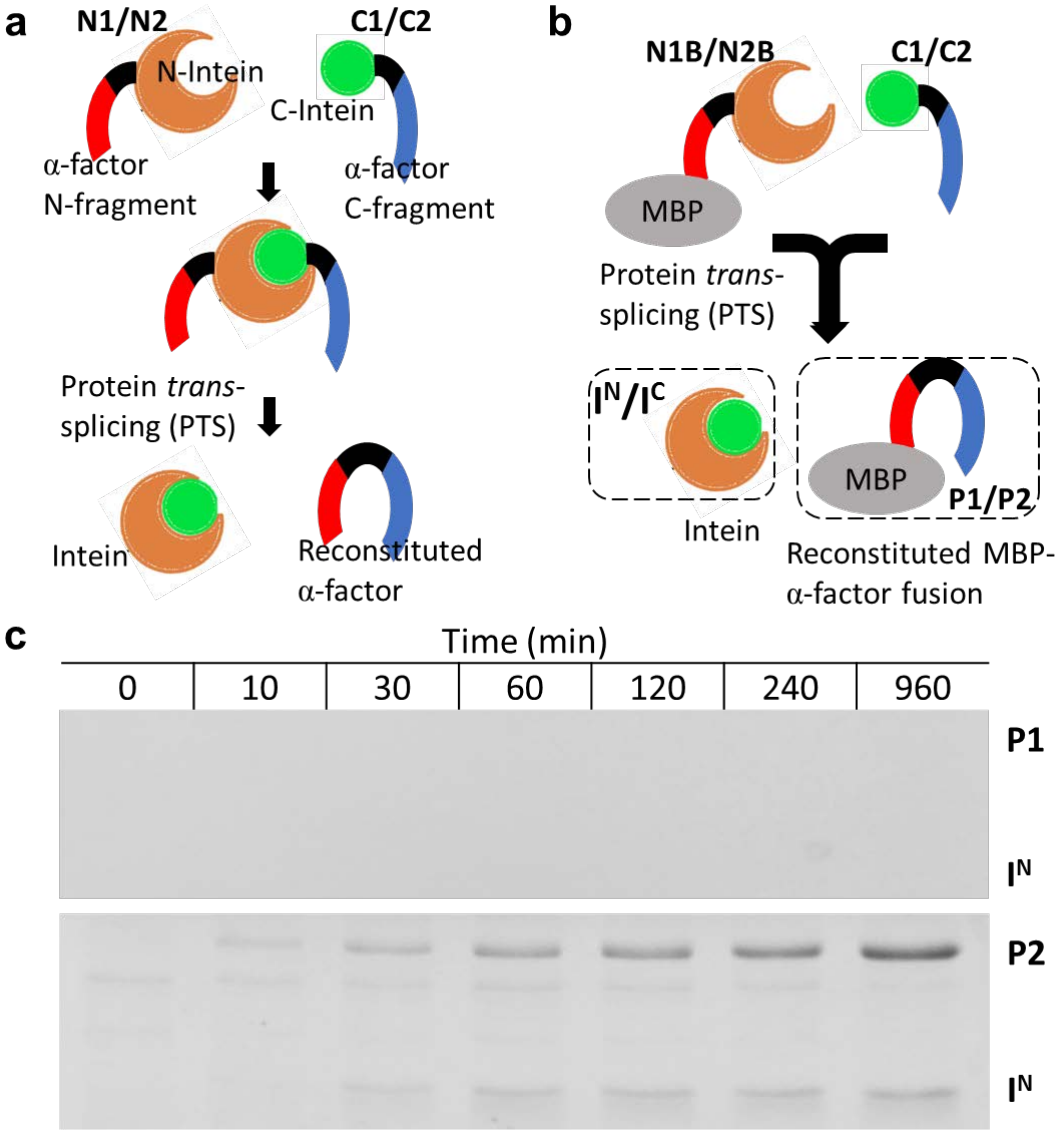


Figure 2.7 Testing PTS to reconstitute α-factor analogue *in vitro*. a) Peptide analogues 1 and 2 were genetically split and fused on the termini of the naturally split *Npu* DnaE intein for PTS. b) An additional maltose-binding protein (MBP) was added to the N-terminus of the N fragment of the α-factor to ease analysis of the reaction. c) Comparison of the two reactions using the two split analogue sequences (~1μM of each protein) showed that only the N2B/C2 mixture resulted in the accumulation of the expected spliced product (P2) over the course of 16 hours.

Once we confirmed that the split fragments are inactive, we next sought to determine the efficiency of intein ligation. To demonstrate the splicing of our split synthetic analogues into a single peptide, we first fused a maltose binding protein (MBP) to the N-terminus of the N-fragments to create constructs N1B and N2B for easy visualization of ligation by SDS-PAGE (Figure 2.7b) as it has been shown that additional sequences on the exteins beyond the first few residues do not affect PTS significantly (Shah et al. 2013).

The corresponding purified N1B/C1 and N2B/C2 fusion proteins were mixed together and the formation of new products were monitored and analyzed by SDS PAGE (Figure 2.7c). Strikingly, the difference of a single residue at the +2 position (Gly in C1 vs. Ala in C2) resulted in the abolishment of any observable *trans*-splicing reaction (Figure 2.7c and Figure 2.8, P1 vs P2), confirming previous studies on the effects of the +2 Gly residue on reaction kinetics using *Npu* split inteins (Amitai et al. 2009; Shah et al. 2012). Although the estimated $t_{1/2}$ for reconstitution of N2B/C2 is approximately 180 min, well short of the rate reported for the wild-type C+2 Phe ($t_{1/2}$ ~30s) (Shah et al. 2013), this represented a compromise between PTS efficiency and biological activity for the split α-factor analogue; the insertion of a bulky Phe residue into the loop domain of the signaling peptide would likely prohibit both termini from interacting properly with the Ste2p receptor, as suggested by the lack of any bulky residues in the flexible domain from previous mutational screening studies (Naider and Becker 1986). This was further confirmed by our initial plate-based screen (Figure 2.5c; Table 2.1, variant 3).

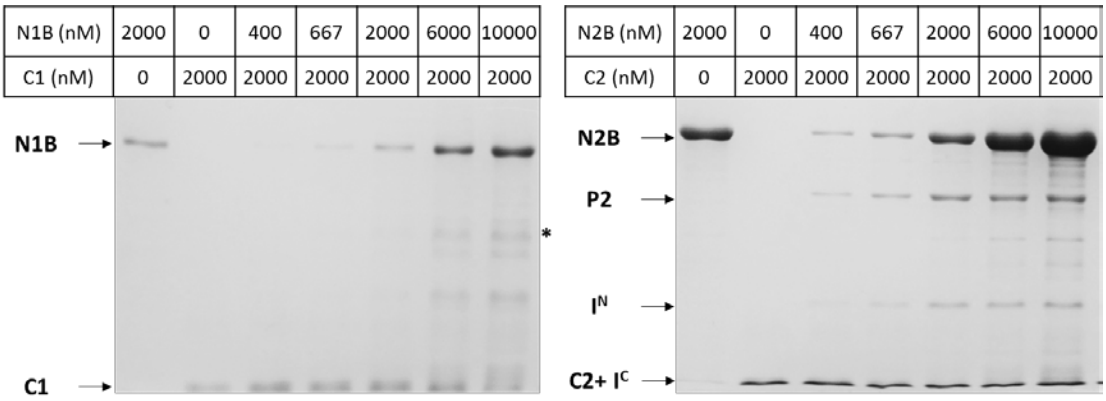
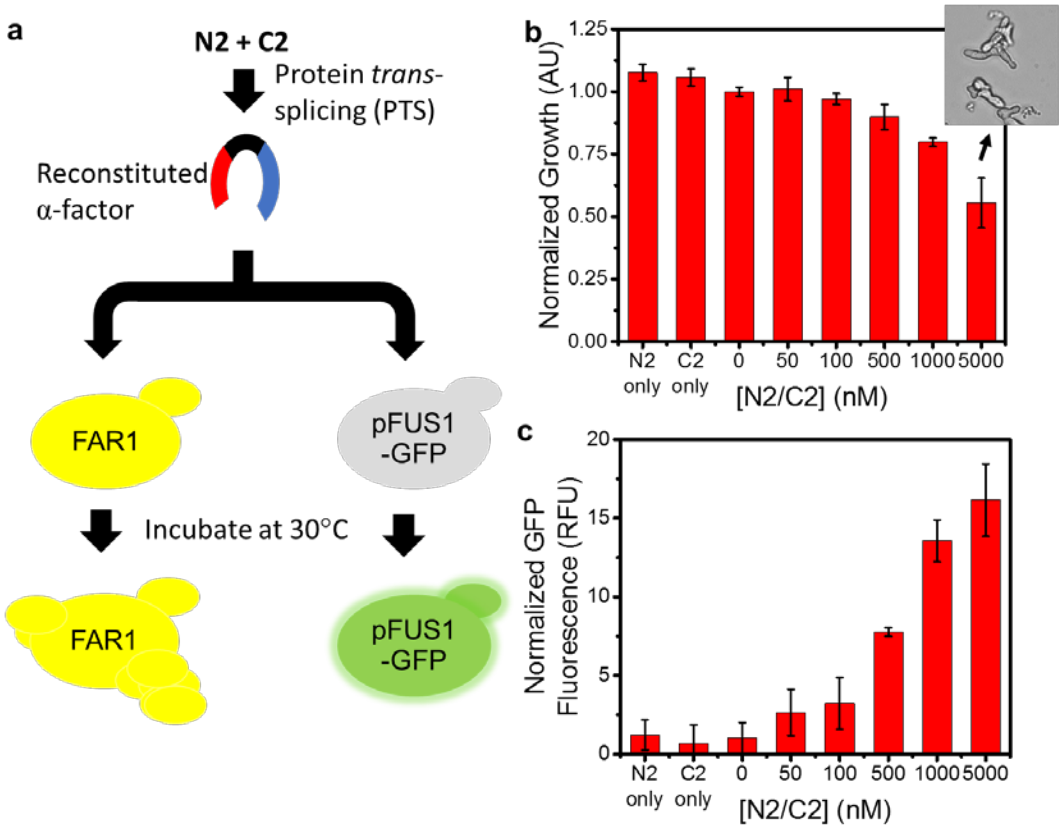


Figure 2.8 Extent of protein *trans*-splicing (PTS) reactions by N1B/C1 vs N2B/C2. The extent of PTS for reconstitution of peptide analogues 1 or 2 was compared using increasing amounts of N1B/N2B in reaction mixtures. No notable ligation products were observed with N1B+C1 even at the highest concentrations tested while products were readily seen with N2B+C2. (*) denotes truncated N1B fusion protein.

2.4.3 Reconstituted α-factor is Biologically Active

Using the 240 min window, we proceeded to ligate N2/C2 and test the resulting ligation products for the ability to trigger cell cycle arrest in liquid cultures (Figure 2.9a). The addition of these reaction mixtures produced a dosage-dependent growth inhibition that strongly suggested the inducible response was a direct result of adding increasing amounts of our reconstituted α-factor analogue (Figure 2.9b). This is further supported by detecting the signature change in cell morphology associated with activation of mating response under a microscope (Figure 2.9b, inset). However, the level of growth inhibition was rather modest even at the highest reactant concentrations of 5μM, suggesting either a lower ligation efficiency than expected or an artifact of the sensitivity of our growth assay. Since the ultimate goal of our strategy is to induce transcriptional upregulation, we further assessed the reaction products to induce GFP expression from the pFus1 promoter. Again, a dosage-

dependent GFP induction was observed through fluorescence measurements, flow cytometry, and immunoblotting (Figure 2.9c). Interestingly, the transcriptional response is far more sensitive than the growth arrest assay with ~7-fold GFP induction even at 500nM concentration. This result suggests that growth inhibition assay may be biased by the growth conditions and even a low level of α- factor is sufficient to induce the required expression phenotype. Collectively, these results confirm our ability to reconstitute functional α-factor analogues for transcriptional activation, albeit at a slightly lower sensitivity than the wild-type version.


a
N2 + C2
Protein trans-splicing (PTS)
Reconstituted α-factor
FAR1
pFUS1-GFP
Incubate at 30°C
FAR1
pFUS1-GFP
b
Normalized Growth (AU)
1.25
1.00
0.75
0.50
0.25
0.00
N2 only
C2 only
0
50
100
500
1000
5000
[N2/C2] (nM)
c
Normalized GFP Fluorescence (RFU)
20
15
10
5
0
N2 only
C2 only
0
50
100
500
1000
5000
[N2/C2] (nM)

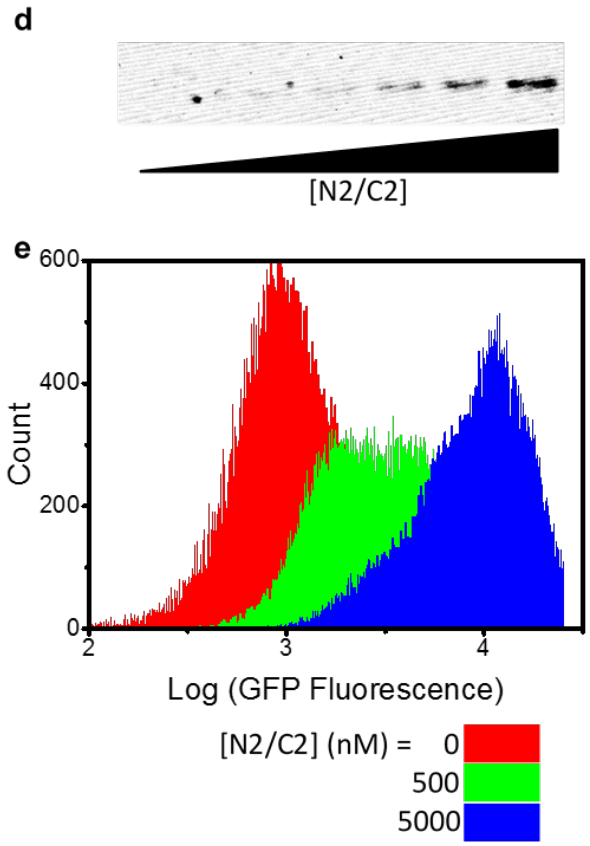
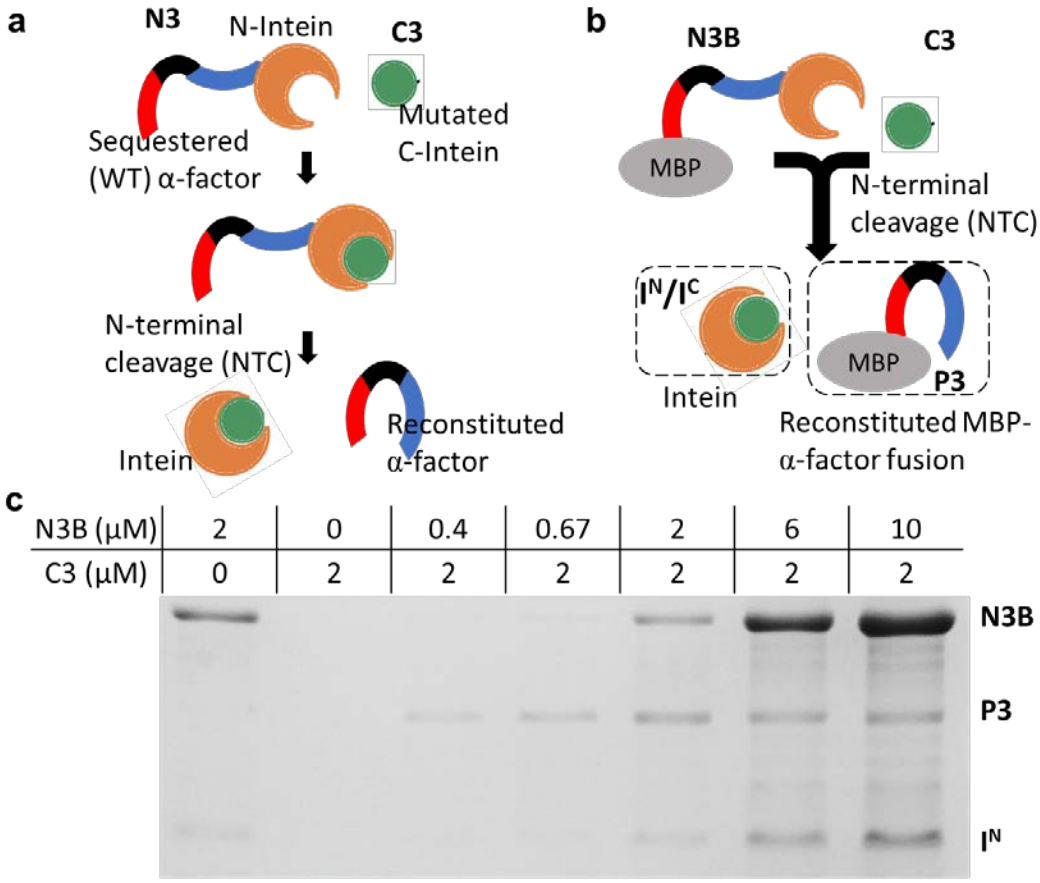


Figure 2.9 Protein *trans*-splicing reaction for reconstitution of functional synthetic α-factor analogue. a) Schematics of functional assays for reconstituted peptides after PTS or NTC reactions. b) Addition of reaction mixtures induce cell cycle arrest in a dosage-dependent manner. Inset: Signature change in cell morphology observed in yeast cultures exposed to reaction mixtures containing 5μM of N2/C2 proteins. Activation of mating response signaling pathway exploited to drive expression of reporter GFP was confirmed and characterized by c) microplate assay, d) immunoblotting, and e) flow cytometry. Values in bar graphs are mean ± s.e.m. with n=3 independent experiments

2.4.4 N-terminal Cleavage Reaction as an Alternative Scheme to Protein *Trans*-splicing

One key reason why the sensitivity for the protein *trans*-splicing approach is lower is the fact that the α-factor analogue itself is more than tenfold less active

compared to the wild type (Figure 2.5 and Figure 2.6; Table 2.1). One way to bypass this hurdle is to devise an alternative strategy for reconstitution using the wild-type α-factor sequence. Almost all reaction intermediates of PTS are susceptible to nucleophile-induced cleavage (Figure 2.2) (Amitai et al. 2009; Shah et al. 2013). Although these side reactions are typically undesirable as they lower the yields of the final ligated products from PTS, they can also be exploited as an alternative reaction pathway to produce a cleaved product (Wood et al. 1999). It may be possible to modify our strategy to take advantage of these side reactions in the form of N-terminal cleavage (NTC) in place of PTS to reconstitute functional signaling peptides (Figure 2.1 and Figure 2.10a). This cleavage scheme is promiscuous to the sequence of the excised N-extein, allowing the use of wild-type α-factor instead of a synthetic analogue with reduced biological activity.


a
N3
N-Intein
C3
Mutated
C-Intein
Sequestered
(WT) α-factor
N-terminal
cleavage (NTC)
Intein
Reconstituted
α-factor
b
N3B
C3
MBP
N-terminal
cleavage (NTC)
IN/IC
MBP
P3
Intein
Reconstituted MBP-
α-factor fusion
c
N3B (μM) 2 0 0.4 0.67 2 6 10
C3 (μM) 0 2 2 2 2 2 2
N3B
P3
IN

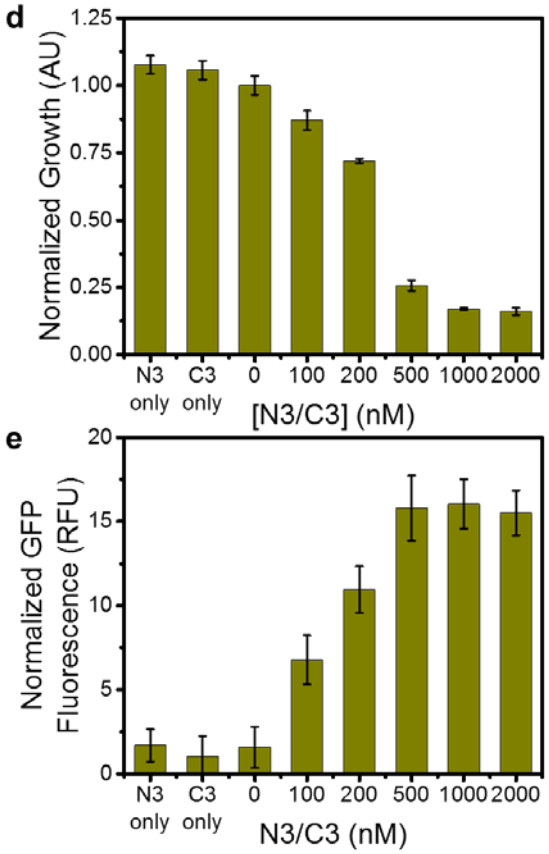


Figure 2.10 N-terminal cleavage reactions to reconstitute α-factor. a) The entire wild-type α-factor was fused to *Npu*N intein, thereby eliminating the signal peptide's ability to bind onto its receptor until NTC reactions cleave it from the fusion protein. b) An additional maltose-binding protein (MBP) was added to the N-terminus of the α-factor to ease analysis of the reaction. b) The expected cleavage product (P3) was observed to accumulate in the reaction mixtures after 1 hour of incubation at 37°C. Cleaved α-factor remained biologically active and was able to induce d) growth arrest and e) transcriptional activation of pFUS1-GFP. Values are mean ± s.e.m. with n=3 independent experiments.

However, it was uncertain whether fusion of the N-intein to the C-terminus of the α-factor is sufficient to block binding to the Ste2p receptor. To investigate this possibility, a new fusion protein N3 was purified and was shown to be incapable of

eliciting any observable growth arrest (Figure 2.10d). This result indicates that N-intein is effective in blocking Ste2p binding, rendering N3 inactive in mating activation. To further probe the kinetics of NTC, MBP was similarly fused to the N-terminus of N-extein (i.e. the α-factor peptide) to create N3B for easy visualization (Figure 2.10b). The formation of new products after mixing N3B/C3 were analyzed by SDS PAGE (Figure 2.10c), and the expected MBP-α-factor product with yield >50% over the course of an hour was detected. The significantly higher efficiency of NTC relative to PTS was likely because NTC avoided the rate-limiting step in PTS pathway (Figure 2.2), which was exacerbated by the substitution of Ala at the C+2 extein residue.

Having proven the ability to effectively excise N-exteins using NTC, we proceeded to test whether the cleaved products are biologically active and can provide functional signaling in live yeast cultures. The addition of the N3/C3 reaction mixtures resulted in significant growth inhibition even at 500nM, a concentration 10-fold lower than that observed for the ligation reactions (Figure 2.10d). A similar enhancement in response was also detected with the GFP induction assay (Figure 2.10e). This higher sensitivity can be attributed to both the improved reaction kinetics and the signaling efficiency of the wild-type α-factor sequence.

2.4.5 *In situ* N-terminal Cleavage to Control Yeast Phenotypes

Because of these improved properties, the N3/C3 system was further evaluated for the ability to elicit signaling response in an *in vivo* culture setting. Since yeast cultures are grown at pH 6.6 as compared to the optimum reaction pH of 7.4, the ability to carry out the NTC reaction under these lower pH conditions was first investigated. The ligation rate remained fairly constant for all pH values higher than

5.7, indicating that this approach will likely work even for growing yeast cultures (Figure 2.11). This was verified by observing the ability to induce GFP synthesis by the in situ release of functional α-factors by NTC. Purified N3 and C3 proteins (1μM each) were added directly into two different yeast cultures grown in either rich or synthetic medium and up to a 6-fold increase in whole-cell fluorescence was detected within 4 h (Figure 2.12). A similar level of response was observed by using an *in vitro* reaction mixture at the same N3/C3 concentration (Figure 2.10e), signifying robustness of the ligation mechanism to reactivate functional signaling peptides in a culture setting and even in the presence of live cells.

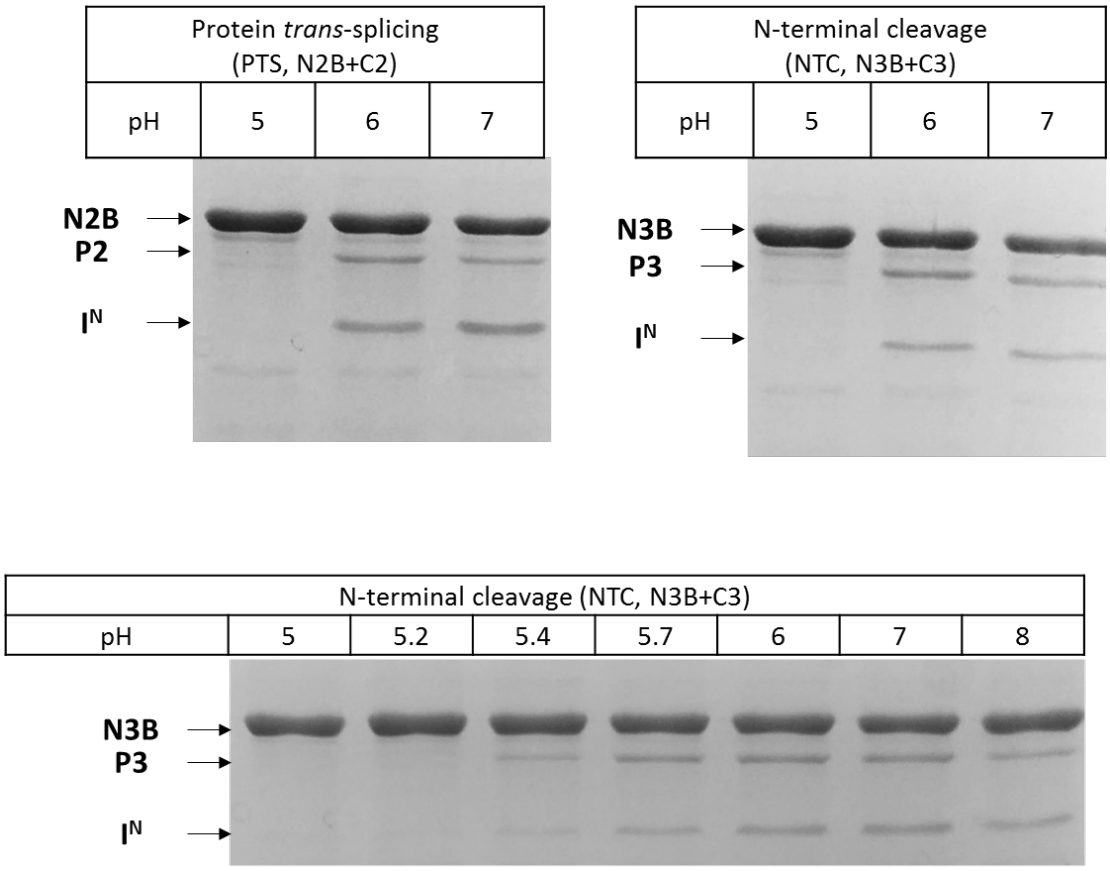


Figure 2.11 Effects of pH on protein *trans*-splicing (PTS) and N-terminal cleavage (NTC) by *Npu* split intein. Both PTS and NTC were tested in different pH to determine whether the split inteins remain active in yeast culture conditions. Activities were only diminished below pH 5.7 and thus a typical yeast culture buffered at pH 6 or above would allow these intein-mediated reactions to occur.

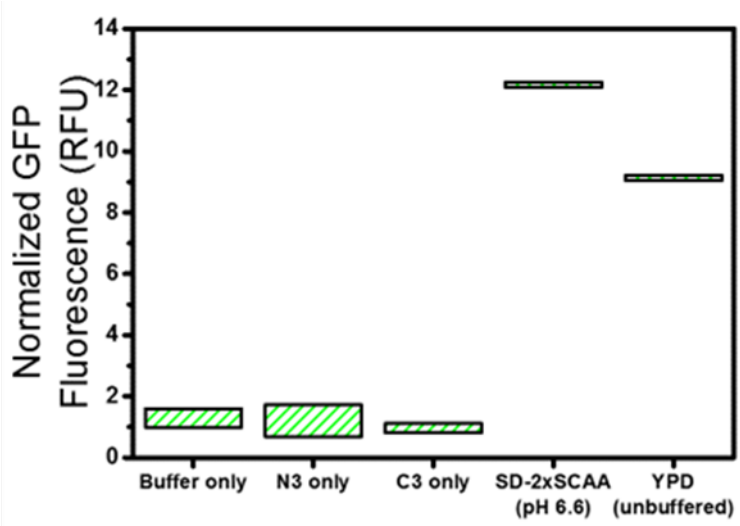


Figure 2.12 Reconstitution of α-factor peptide by NTC in various media with growing yeast cultures. Addition of 1μM of N3/C3 induced GFP expression in either synthetic dropout medium (SD-2xSCAA) or enriched complex medium (YPD), albeit at slightly lower efficiency relative to *in vitro* reactions (Fig 2.10e). Each bar represents the high and low values obtained from two independent experiments performed on separate days.

2.4.6 Prospects of Split Intein-mediated Reactions as Foundation for Synthetic Sensing Circuits

In principle, any GPCRs associated with a peptide signaling molecule, such as the human somatostatin, can be similarly activated by this method. Importantly, the ability to perform *in situ* activation in live cell cultures opens up the possibility of creating orthogonal yeast strains secreting either N3 or C3 under the control of different external inputs in order to create a wide range of AND or OR gate logic yeast co-culture systems. More complex logic circuits can be designed by adapting conditional protein splicing (CPS), in which ligation is initiated by the presence of a small molecule such as rapamycin (Mootz et al. 2003) (Figure 2.13). Other hetero-dimerization domains for proteins or metabolites can be similarly exploited for increased flexibility.

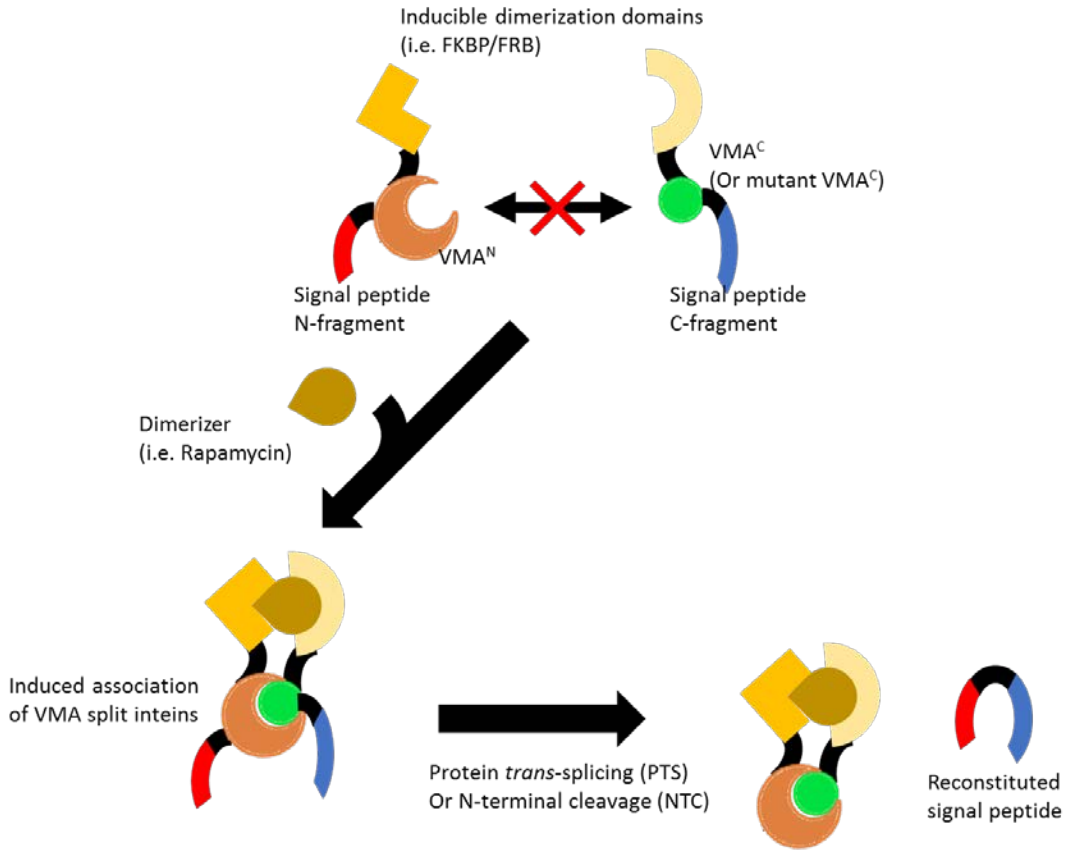


Figure 2.13 Conditional split inteins, such as VMA from *Saccharomyces cerevisiae*, can be used in place of the *Npu* inteins used in the study. These split inteins have very little native affinity and must rely on chemical induced dimerizers (i.e. rapamycin) to bring each VMA intein fragments into close proximity to trigger PTS or NTC. This allows more precise spatial and temporal controls to be implemented over the reconstitution of functional peptides.

Because both the identities of the reconstituted signal peptide and the induced dimerization domains are flexible, this design can serve as an extracellular sensor that is able to recognize any target analyte of choice and to activate any signal pathway once PTS or NTC reconstitutes the peptide. This design, thus, represents a basic sensory element in extracellular sensing circuits.

2.5 Acknowledgements

This work was supported by NSF (CBET1263719). We thank Prof. Zhilei Chen at TAMU and Prof. Wendell Lim at UCSF for their gracious gifts of a few noted plasmids and strains used in this study.

Chapter 3

ENGINEERING SGRNA AS TOEHOLD SWITCHABLE RIBOREGULATORS FOR PROGRAMMABLE CONTROL OF CRISPR/CAS9 FUNCTION

This Chapter is adapted with permission from Siu, Ka Hei, and Wilfred Chen. 2018. "Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function." *Nature Chemical Biology.* doi: 10.1038/s41589-018-0186-1. (Appendix B.2)

3.1 Summary

Predictable control over gene expression is essential to elicit desired synthetic cellular phenotypes. Although CRISPR interference (CRISPRi) offers a simple RNA-guided method for targeted transcription gene silencing, it lacks the ability to integrate endogenous cellular information for efficient signal processing. Here, we present a new class of riboregulators termed toehold-gated gRNA (thgRNA) by integrating toehold riboswitches into sgRNA scaffolds, and demonstrate their programmability for multiplexed regulation in *E. coli* with minimal cross-talk. The versatility of the design provides a highly flexible platform to guide cellular behaviors in a variety of contexts.

3.2 Introduction

The ability to provide orthogonal control over gene expression is critical for the creation of complex biological circuits that can reliably redirect cellular functions into new programmable phenotypes (Khalil and Collins 2010; Lim 2010; Taylor et al. 2015). Previous efforts relied on the use of ligand-responsive transcription factors

(LRTFs), which must be customized for each target of interest (Taylor et al. 2015; S. Y. Tang and Cirino 2011). Their limited number and orthogonality further hinder the construction of more complex circuits that can operate robustly in living cells. An alternative is to use RNA-based regulators, which contain a specific pocket for target binding and an allosteric cis-RNA element for gene regulation (Martini et al. 2015; Serganov and Nudler 2013; Chappell et al. 2015). However, the small number of well-characterized RNA sensor-actuator pairs and the rather modest dynamic range limit their wide-spread utility (Callura, Cantor, and Collins 2012; Mutalik et al. 2012). An ideal design is to create new hybrid protein-RNA devices combining the unique advantages of both systems for orthogonal gene regulation.

The recently discovered CRISPR/Cas9 system offers a unique RNA-guided approach for DNA targeting (Horvath and Barrangou 2010; Wiedenheft, Sternberg, and Doudna 2012), and nuclease-null Cas9 (dCas9) has been repurposed as transcriptional regulators for both gene repression and activation (Qi et al. 2013). Switchable guide RNA (gRNA) motifs that are modulated by specific stimuli have been created to enable tunable gene regulation and genome editing (W. Tang, Hu, and Liu 2017; Liu et al. 2016). Conditional activation of quiescent gRNAs was achieved by using either a ligand-induced conformation switch or a ligand-controlled RNA-cleaving unit to uncage the spacer-blocking region. While these strategies enable programmable activation of Cas9 functions by externally delivered ligands or selected intracellular protein triggers (Liu et al. 2016; Davis et al. 2015), they lack the ability to implement autonomous control of gene expression based on endogenous cellular information. Inspired by the simplicity of toehold-mediated strand displacement (D. Y. Zhang and Winfree 2009) and the success of new de-novo-designed regulators termed

"toehold switches" for translational activation (Green et al. 2014), we designed a new class of conditional gRNA structures termed toehold-gated gRNA (thgRNA) (Figure 3.1) and demonstrated their utility to provide orthogonal gene regulation using both synthetic and endogenous RNA triggers.

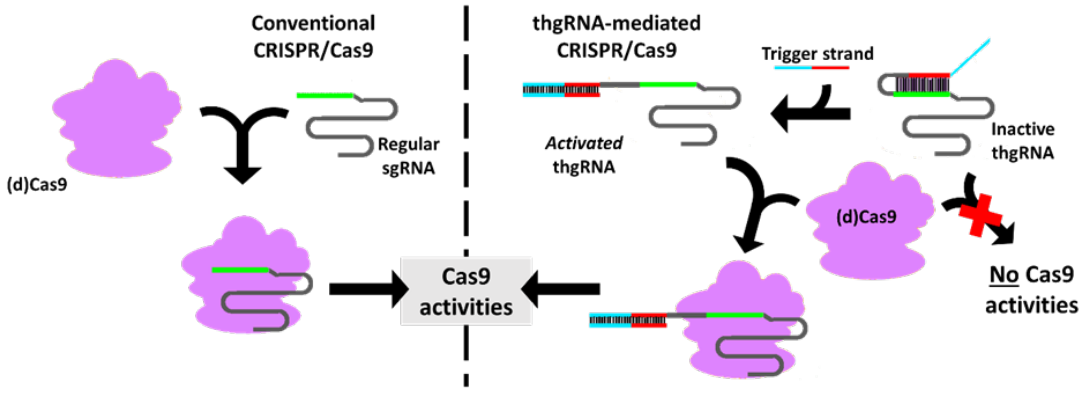


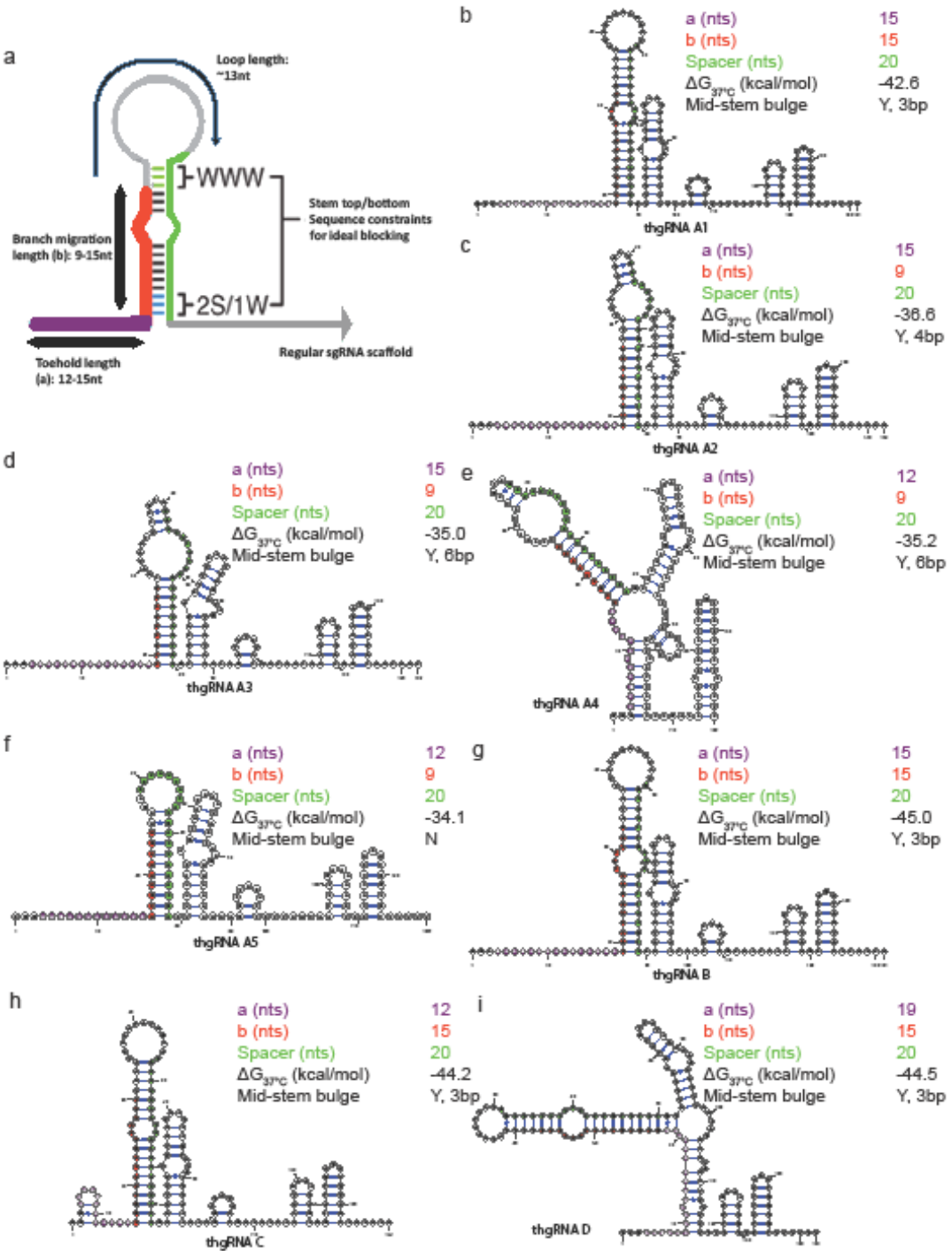
Figure 3.1 Design of toehold-gated guide RNAs (thgRNAs). Schematic representation of thgRNA-based activation of CRISPR/Cas9 functions. Binding of Cas9 to specific DNA targets relies on hybridization of the spacer region of sgRNA to the unwound DNA strand. With its spacer region sequestered, thgRNA is unable to hybridize with its target DNA until a cognate trigger strand is added to initiate a toehold strand displacement reaction to open the hairpin structure and unblock the spacer region to freely hybridize to the specific DNA target and activate CRISPR/Cas9 functions.

3.3 Materials and Methods

3.3.1 Design and Computational Screening of thgRNA Variants

NUPACK algorithm (Zadeh et al. 2011) was used to model all thgRNA variants prior to any experimental work (Figure 3.2). Except for thgRNAs **A2-5**, which were designed and screened to test the limits of thgRNA structures (length of toehold, stem-loop, etc.), the toehold, branch migration, and spacer regions of all other

thgRNAs were screened and modified to follow the general design of thgRNA **A1**, emulating rules outlined by Green et al. (Green et al. 2014) (Figure 3.2a), wherever possible. Stable secondary structures in the toehold region were avoided in all instances where artificial trigger sequences were used. Pair-wise complexes between thgRNAs and trigger strands were also modeled by the same algorithm to predict formation of hybridized dimers, as intended or otherwise. The full sequences of all thgRNAs, trigger strands, and DNA targets tested are reported in Table A.2.1.


a
Loop length: ~13nt
WWW
2S/1W
Stem top/bottom Sequence constraints for ideal blocking
Branch migration length (b): 9-15nt
Toehold length (a): 12-15nt
Regular sgRNA scaffold
b
a (nts) 15
b (nts) 15
Spacer (nts) 20
ΔG37°C (kcal/mol) -42.6
Mid-stem bulge Y, 3bp
thgRNA A1
c
a (nts) 15
b (nts) 9
Spacer (nts) 20
ΔG37°C (kcal/mol) -36.6
Mid-stem bulge Y, 4bp
thgRNA A2
d
a (nts) 15
b (nts) 9
Spacer (nts) 20
ΔG37°C (kcal/mol) -35.0
Mid-stem bulge Y, 6bp
thgRNA A3
e
a (nts) 12
b (nts) 9
Spacer (nts) 20
ΔG37°C (kcal/mol) -35.2
Mid-stem bulge Y, 6bp
thgRNA A4
f
a (nts) 12
b (nts) 9
Spacer (nts) 20
ΔG37°C (kcal/mol) -34.1
Mid-stem bulge N
thgRNA A5
g
a (nts) 15
b (nts) 15
Spacer (nts) 20
ΔG37°C (kcal/mol) -45.0
Mid-stem bulge Y, 3bp
thgRNA B
h
a (nts) 12
b (nts) 15
Spacer (nts) 20
ΔG37°C (kcal/mol) -44.2
Mid-stem bulge Y, 3bp
thgRNA C
i
a (nts) 19
b (nts) 15
Spacer (nts) 20
ΔG37°C (kcal/mol) -44.5
Mid-stem bulge Y, 3bp
thgRNA D

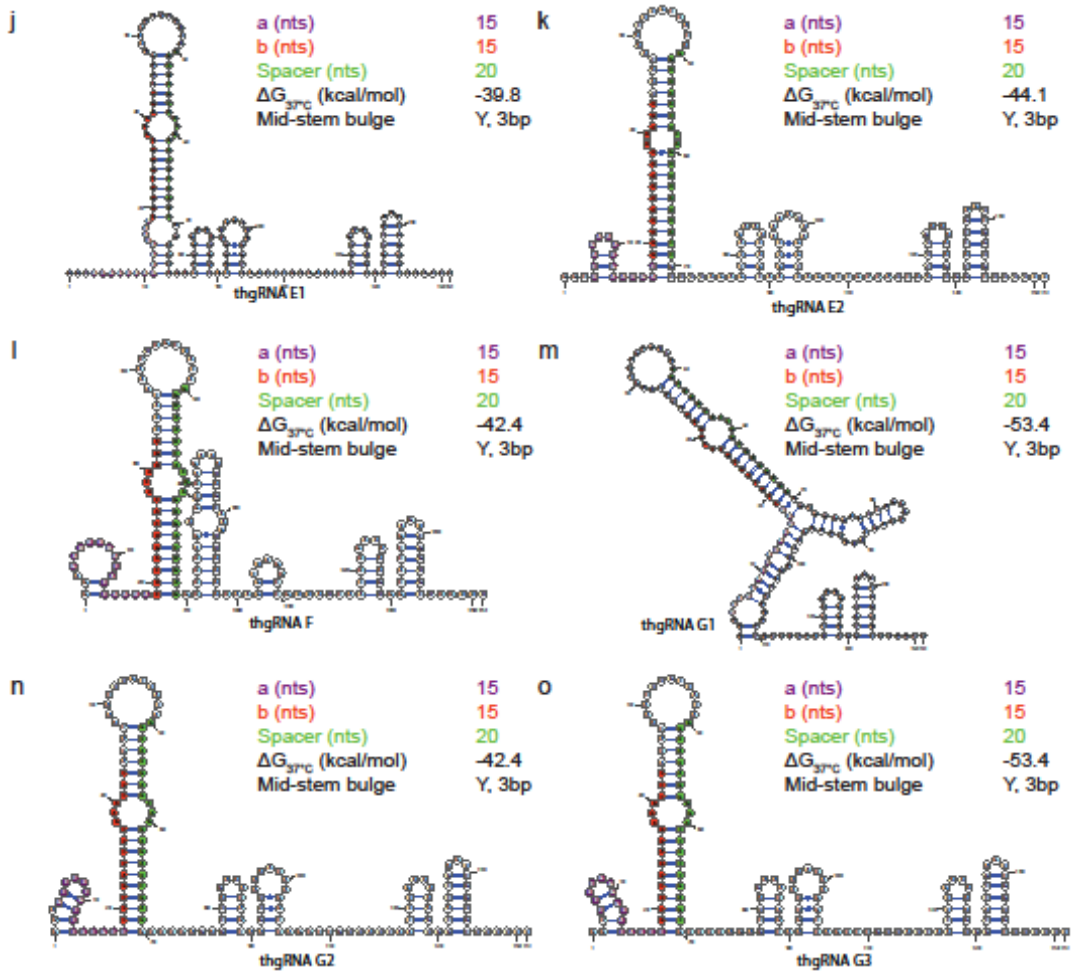


Figure 3.2 Design and schematics of thgRNAs. a) General design schematics of thgRNA. Length of the toehold region (a) and branch migration region (b) can be somewhat flexible if the trigger sequence impose certain constraints, though the ideal lengths are 15nts for both regions. As discussed in Green et al.1, the top of the stem should ideally be limited to A-U base pairing while the bottom should include 2 C-G and 1 A-U pairs for optimal sequestration. b) to o) Minimum free energy (MFE) structures predicted by NUPACK for all thgRNA variants tested. The toehold (a), branch migration (b), and spacer regions are colored purple, orange, and green respectively. ΔG is calculated at 37°C with 1M Na^{+} and 0M Mg^{2+}. If imperfect hybridization was designed between the branch migration and the spacer region, a mid-stem bulge of 3-6bp would be predicted to form.

When endogenous sRNAs were intended to be used as trigger strands, previous literature (Altuvia et al. 1998; Massé and Gottesman 2002) were referenced wherever available to check the computationally predicted structures modeled by NUPACK. The artificial trigger strands **C*** and **D*** were designed to mimic single-stranded regions of the sRNA or regions where hybridization with the native targets were experimentally observed by previous works. Both the artificial triggers and endogenous sequences were further modeled for hybridization with the thgRNAs. All thgRNAs **A**, **B**, **C**, and **D** and any variants characterized in subsequent experimental studies were predicted to form stable complexes with their cognate trigger strands.

3.3.2 DNA Cleavage and Förster Resonance Energy Transfer (FRET)-based Beacon Assays

Plasmid containing DNA target **A** was constructed as described in Chapter 3.3.3 below (Table A.2.2). Target DNA was harvested from NEB®5α cells transformed with the relevant plasmids using Zyppy® miniprep kit according to vendor's instructions (Zymo Research, Irvine, CA, USA) and linearized by restriction digests using NotI enzyme prior to cleavage assays. Cas9 proteins were expressed from *E. coli* and purified as described by a previous study (Gagnon et al. 2014). All RNAs were transcribed *in vitro* using HiScribe™ T7 Quick High Yield RNA Synthesis Kit (New England BioLabs®, Inc., Ipswich, MA, USA) and purified by standard phenol-chloroform extraction and ethanol precipitation. RNA purity and quality was analyzed by spectrophotometry on a NanoDrop™ 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific™, Inc., Waltham, MA, USA) and denaturing urea PAGE electrophoresis using 5% polyacrylamide gels containing 8M urea (Summer, Grämer, and Dröge 2009). Unmodified and fluorophore- or quencher-

labeled oligonucleotides were synthesized and purified by a commercial vendor (Integrated DNA Technologies®, Inc., Coralville, IA, USA).

For the DNA cleavage assay, thgRNA and DNA mimetic trigger strand or unmodified sgRNA were first incubated at 37°C for 15min. Reactions containing linearized plasmid targets, purified Cas9 protein, and pre-incubated thgRNA/trigger strand or sgRNA were then mixed with final concentrations as denoted into cleavage reaction buffer (100mM NaCl, 50mM Tris-HCl, 10mM $MgCl_2$, 100μg/ml BSA, pH 7.9), and incubated at 37°C for 1 hour. Reaction products were then analyzed by electrophoresis on a 1% agarose gel with ethidium bromide stain. Percent of cleavage was estimated by densitometry using ImageJ (NIH).

For the FRET-based beacon assay (Figure 3.6), beacon complexes containing either target **A**, **B**, **C**, or **D** were assembled by mixing the 5'-fluorophore-labeled target strand, the PAM-containing strand, and the 3'-quncher labeled strand to a final concentration of 2μM in nuclease-free water, heated to 90°C, and cooled to room temperature at a rate of 0.1°C/second using a S1000™ Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA). All fluorescence measurements were carried out in binding buffer (20mM Tris-HCl, 120mM NaCl, 5%v/v glycerol, 0.1mM DTT, 1mM MgCl2, 0.02%v/v Tween 20, pH 7.9) (Mekler et al. 2016) and measured using a Synergy™ H4 Hybrid microplate reader (BioTek Instruments, Inc., Winooski, VT, USA) at 25°C. Final assay mixtures contained 5nM assembled beacon complex, 50nM Cas9 protein, 50nM sgRNA or thgRNA, and varying concentrations of trigger strands as indicated. Excitation and emission wavelengths were tailored to the fluorophore used for each target (**A**: FAM, Ex: 498nm, Em: 520nm; **B**: TYE™563, Ex: 545nm, Em: 565nm; **C**: TEX615, Ex: 595nm, Em: 615nm; **D**: TYE™665, Ex: 645nm, Em:

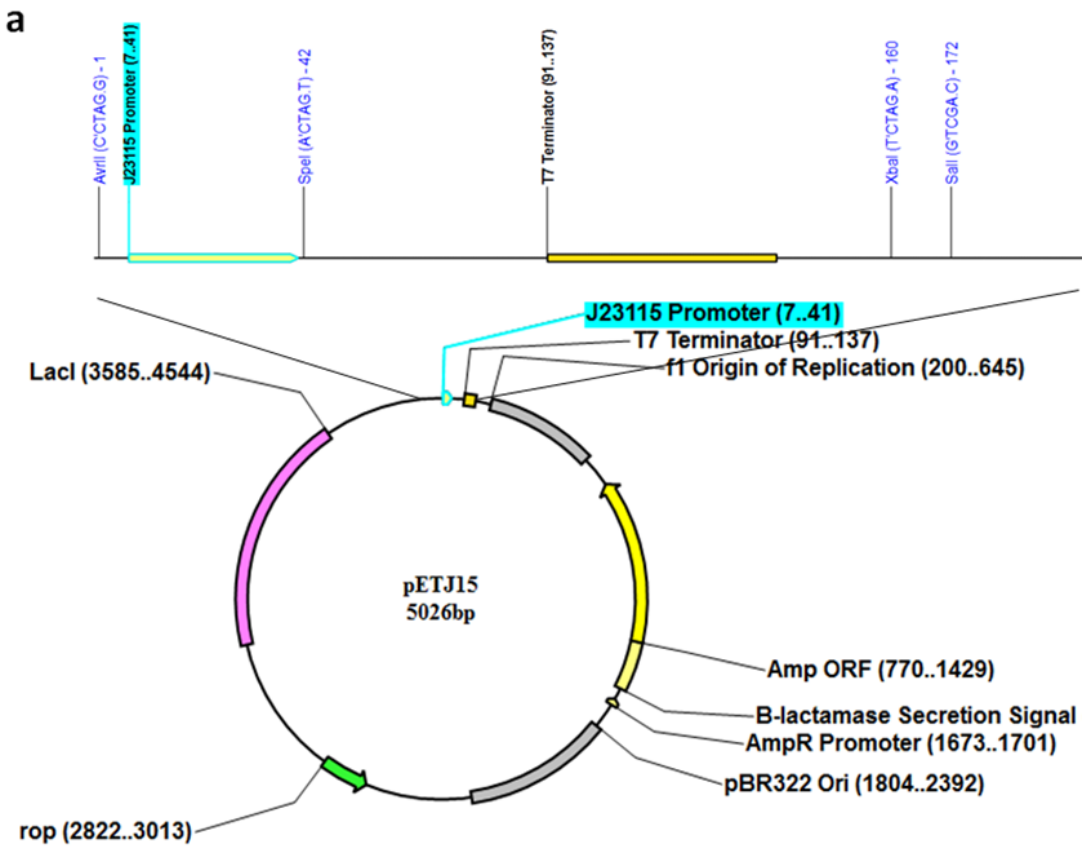
665nm). Pre-incubated thgRNA/trigger strand or sgRNA were added at t = 0s and measurements were taken every 6-30 seconds for the initial 600s and every 30 seconds for the remainder of the assays.

3.3.3 Strains Used and Plasmids Construction

All strains and plasmids used in this chapter are listed in Table A.2.2.

Plasmid used for *in vitro* DNA cleavage assays containing target **A** was constructed by standard subcloning techniques based on high-copy number backbone pUC19 and transformed into NEB®5α *E. coli* strain (New England BioLabs®, Inc., Ipswich, MA, USA) to ensure high plasmid yields. Expression construct for Cas9 protein was obtained as a gift from Dr. David Liu (Addgene plasmid # 62374) and transformed into BL21-Gold (DE3) cells for expression. dCas9 gene was amplified from pHAGE-TO-dCas9-3XmCherry (Addgene plasmid # 64108), a gift from Thoru Pederson.

The unmodified ePathBrick19 vector, pETM6, was a generous gift from Prof. Mattheos Koffas and was used as the initial backbone for our intracellular expression constructs. Briefly, we created a new BioBrick-compatible vector capable of constitutive expression by substituting the synthetic constitutive promoter J23115p in place of the lac-inducible T7-lacO promoter and constructed a set of expression constructs for each sgRNA/thgRNA, target and downstream reporter, and Cas9 or dCas9; this allowed us to rapidly combine sets of expression cassettes into a single plasmid that can be easily co-transformed into *E. coli* with a trigger plasmid (See Figure 3.3 for detailed scheme). The trigger plasmids were constructed by Gibson assembly and standard subcloning techniques.


a
AvrII (C'CTAG.G) - 1
J23115 Promoter (7..41)
SpeI (A'CTAG.T) - 42
T7 Terminator (91..137)
XbaI (T'CTAG.A) - 160
SalI (G'TCGA.C) - 172
J23115 Promoter (7..41)
T7 Terminator (91..137)
f1 Origin of Replication (200..645)
LacI (3585..4544)
pETJ15
5026bp
Amp ORF (770..1429)
B-lactamase Secretion Signal
AmpR Promoter (1673..1701)
pBR322 Ori (1804..2392)
rop (2822..3013)

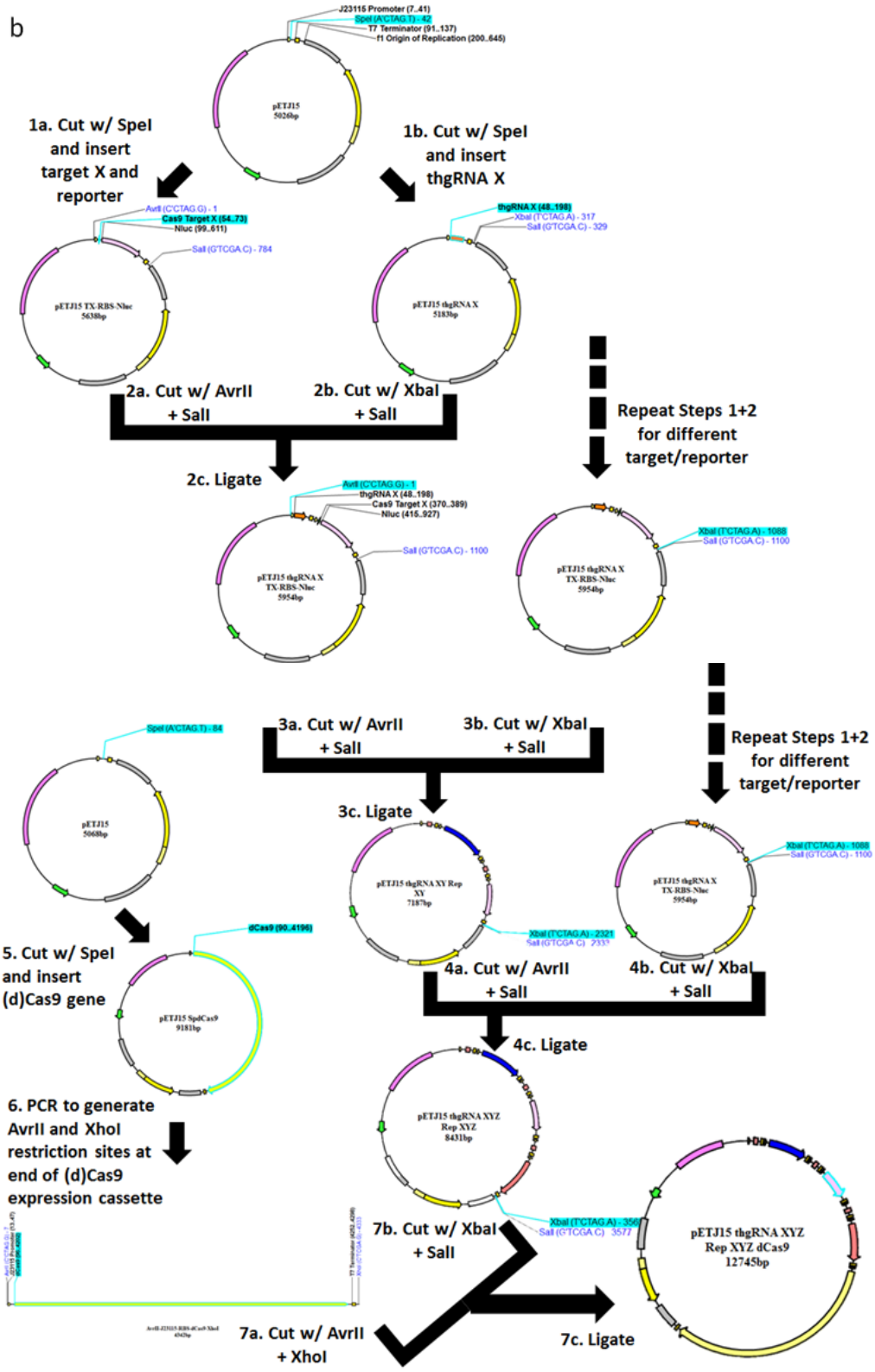

b
J23115 Promoter (7..41)
SpeI (A'CTAG.T) - 42
T7 Terminator (91..137)
f1 Origin of Replication (200..645)
pETJ15
5026bp
1a. Cut w/ SpeI and insert target X and reporter
1b. Cut w/ SpeI and insert thgRNA X
AvrII (C'CTAG.G) - 1
Cas9 Target X (54..73)
Nluc (99..611)
SalI (G'TCGA.C) - 784
pETJ15 TX-RBS-Nluc
5638bp
thgRNA X (48..198)
XbaI (T'CTAG.A) - 317
SalI (G'TCGA.C) - 329
pETJ15 thgRNA X
5183bp
2a. Cut w/ AvrII + SalI
2b. Cut w/ XbaI + SalI
2c. Ligate
Repeat Steps 1+2 for different target/reporter
AvrII (C'CTAG.G) - 1
thgRNA X (48..198)
Cas9 Target X (370..389)
Nluc (415..927)
SalI (G'TCGA.C) - 1100
pETJ15 thgRNA X TX-RBS-Nluc
5954bp
XbaI (T'CTAG.A) - 1088
SalI (G'TCGA.C) - 1100
pETJ15 thgRNA X TX-RBS-Nluc
5954bp
3a. Cut w/ AvrII + SalI
3b. Cut w/ XbaI + SalI
3c. Ligate
Repeat Steps 1+2 for different target/reporter
SpeI (A'CTAG.T) - 84
pETJ15
5068bp
pETJ15 thgRNA XY Rep XY
7187bp
XbaI (T'CTAG.A) - 2321
SalI (G'TCGA.C) 2333
pETJ15 thgRNA X TX-RBS-Nluc
5954bp
XbaI (T'CTAG.A) - 1088
SalI (G'TCGA.C) - 1100
5. Cut w/ SpeI and insert (d)Cas9 gene
dCas9 (90..4196)
pETJ15 SpdCas9
9181bp
4a. Cut w/ AvrII + SalI
4b. Cut w/ XbaI + SalI
4c. Ligate
6. PCR to generate AvrII and XhoI restriction sites at end of (d)Cas9 expression cassette
pETJ15 thgRNA XYZ Rep XYZ
8431bp
XbaI (T'CTAG.A) - 3567
SalI (G'TCGA.C) 3577
7b. Cut w/ XbaI + SalI
pETJ15 thgRNA XYZ Rep XYZ dCas9
12745bp
7a. Cut w/ AvrII + XhoI
7c. Ligate

Figure 3.3 Map and schematic of the base BioBrick compatible plasmid pETJ15 used for constitutive expression of all thgRNA, DNA target with downstream reporter, and dCas9. a) Plasmid map of pETJ15, with the BioBrick gene expression cassette and restriction sites displayed. b) Schematic of the iterative cloning process to create plasmids containing all thgRNA, DNA targets with downstream reporters, and dCas9 for simultaneous constitutive expression in transcriptional repression experiments. See Table A.2.2 for all plasmids constructed and used in the study.

3.3.4 Induced Transcriptional Repression by CRISPRi using Artificial Trigger Strands

NEB®5α cells were transformed with a plasmid containing the constitutively expressed reporter, thgRNA, and dCas9 cassettes and a trigger plasmid containing either **A***, **B***, **C***, **D***, or combinations thereof by heat shock. Successful transformants were picked from agar plates and grown in LB medium (10g/L tryptone, 5g/L yeast extract, 10g/L NaCl) supplemented with 100µg/mL carbenicillin and 50µg/mL kanamycin for ~10-12hrs at 37°C. The resulting cultures were used to inoculate subcultures at an initial OD_{600} ~0.03 and grown to OD_{600} ~1. Expression of the artificial trigger strands were then induced by inoculating fresh LB media containing the corresponding inducers (500mM IPTG, 10ng/mL ATc, or both) at an initial OD_{600} ~0.03. These cultures were incubated at 37°C for ~4-5hrs, at which point mid-late exponential phase (OD_{600} ~1) would be reached and samples were taken.

Cell samples were pelleted by centrifugation at 3000*g* for 5 minutes, washed twice with PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4), and resuspended to OD_{600} ~2 in PBS prior to measurements. Whole cell luminescence were measured according to the NanoGlo® vendor's instructions (Promega Corporation, Madison, WI, USA) using a Synergy™ H4 Hybrid microplate reader. Whole cell fluorescence was also measured with the microplate reader where

needed. Analysis of luminescence and fluorescence data were completed using Origin software (OriginLab Corporation, Wellesley Hills, MA, USA).

Total RNA was extracted from samples using Direct-zol™ miniprep kit according to the vendor's instructions (Zymo Research, Irvine, CA, USA). Transcript expression was then quantified using qPCR using Luna® Universal qPCR Master Mix (New England BioLabs®, Inc., Ipswich, MA, USA) according to the manufacturer's protocol. All samples were run in technical duplicate on a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA). All PCR primers were verified as being specific on the basis of melting curve analysis and were as follows: Nluc: 5'-GGTGTCCGTAACTCCGATCC-3' and 5'-ATCCACAGGGTACACCACCT-3'; *ssrA*: 5'-TTAGGACGGGGATCAAGAGA-3' and 5'-GCGTCCGAAATTCCTACATC-3'. Transcript levels of Nluc were calculated by subtracting housekeeping control (*ssrA*) cycle threshold (C_t) values from Nluc C_t values to normalize for total input, yielding ΔC_t levels. Relative transcript levels were computed as $2^{-\Delta Ct}$.

3.3.5 sRNA and mRNA Induced Repression

NEB®5α cells were transformed with a plasmid containing the expression cassettes for the reporter, thgRNA, and dCas9 and grown as described above. Full-length OxyS, MicF, SgrS sRNA and mCherry mRNA replaced artificial trigger sequence to activate their respective thgRNA. Induction of RyhB sRNA was achieved by addition of varying concentrations of 2'2-bipridyl as noted in Figure 3.16 instead of IPTG or ATc as described above for induction under pL*lac*O-1 or pL*tet*O-1 promoters.

3.3.6 Induced Plasmid Loss

NEB®5α cells were transformed with a plasmid containing the expression cassettes for the reporter and thgRNA **B** as well as a compatible plasmid containing *tet*-inducible trigger **B*** and nuclease active Cas9 protein under constitutive promoter J23115. Transformed cells were grown, induced, and collected as described above. 2-fold serial dilutions of collected cells resuspended to initial OD_{600}~10^{-3} were prepared in PBS and 20μL of each dilution spotted onto LB agar plates containing only kanamycin or both ampicillin and kanamycin as noted in Figure 3.19. Colonies forming units were visually counted and compared between the two plates to estimate the loss of ampicillin resistance, indicating loss of the corresponding plasmid containing Cas9 target and AmpR marker.

3.4 Results and Discussions

3.4.1 Initial Screening and Validation of thgRNA Structural Variants

The thgRNA is initially trapped in a sequestered state, with a stem-loop structure that renders the spacer region unavailable for target binding (Figure 3.1). A structural change is induced by binding the trigger strand to the toehold region preceding the 5' end of the stem-loop, initiating branch migration inside the stem-loop region, and exposing the spacer for target binding (Figure 3.1). We first screened our designs computationally by using NUPACK (Zadeh et al. 2011) to minimize unintended secondary structures and to maximize interactions between the trigger and thgRNA. We selected five candidates that target the same target sequence **A** and *in vitro* synthesized thgRNA variants (**A1**-**5**) for further testing (Figure 3.2b-f and Figure 3.4, Table A.2.1).

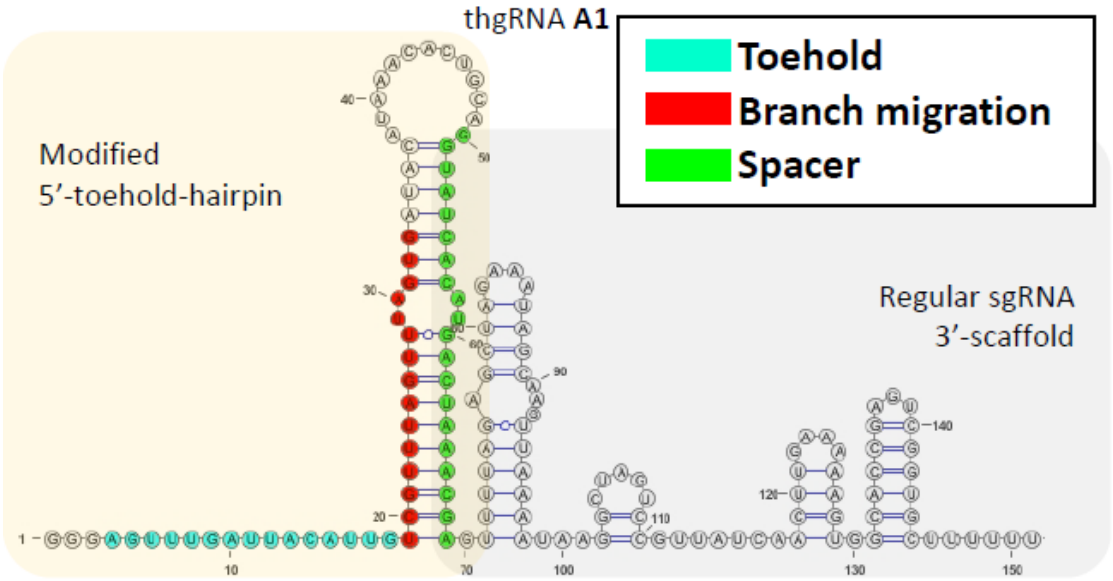


Figure 3.4 The minimum free energy (MFE) structure of thgRNA **A1** at 37°C as predicted by the NUPACK algorithm. The toehold (a), branch migration (b), and spacer regions are colored cyan, red, and green respectively.

The initial characterization of thgRNA candidates exploited the native nuclease activity of Cas9 to cut its cognate DNA target sequence (Figure 3.5a). A successful design would result in cleavage only in the presence of the trigger strand **A***. Of the five variants, most exhibited positive correlations between trigger concentrations and cleavage activities (Figure 3.5b), which suggest that these thgRNA variants were activated through binding with the trigger strand **A***. Only variant A5 displayed a substantial background activity without the trigger. Because the stem region of **A5** is the shortest (11bp, Figure 3.2f), we speculate that the hairpin structure on **A5** may be unstable and thus spontaneously unwind to free up the spacer region for DNA binding.

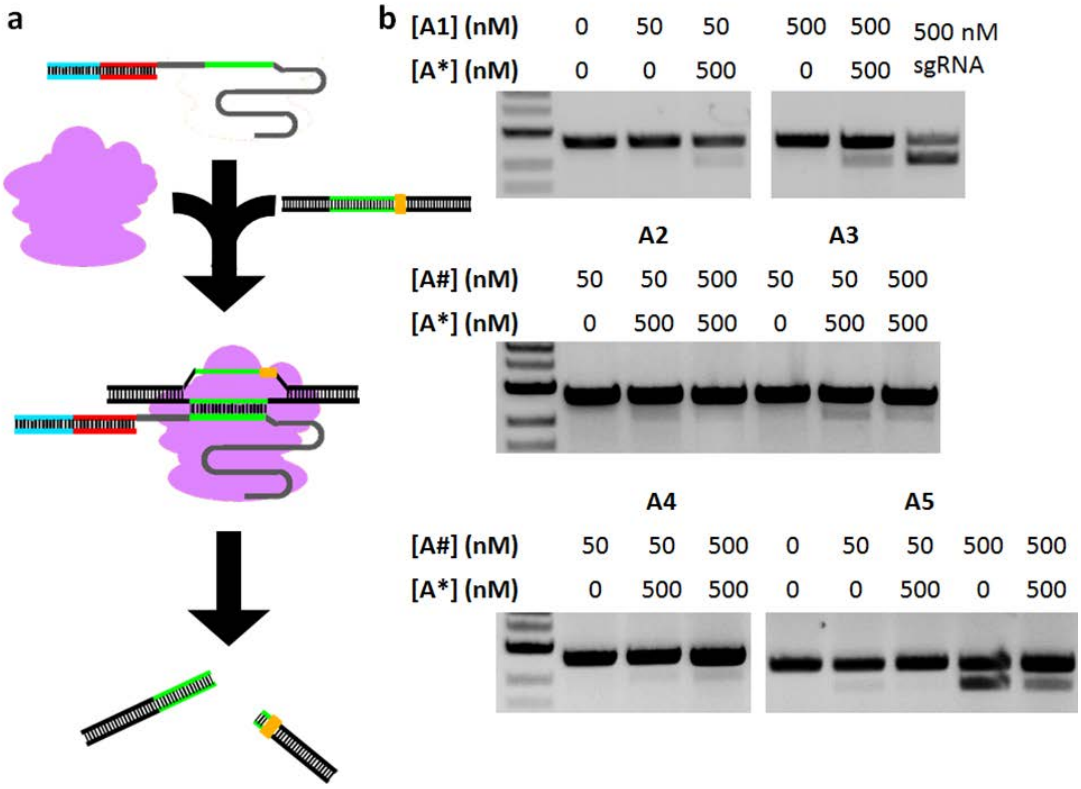


Figure 3.5 DNA cleavage assays for thgRNA **A** variants. a) Schematics of DNA cleavage assay for *in vitro* screening of thgRNA variants. b) Gel electrophoresis showed the extent of DNA cleavage using different concentrations of thgRNA variants and trigger strands. Note the differences in apparent efficiencies. Gels are representative results of three independent experiments.

To evaluate the kinetics of activation, we employed a newly developed FRET-based beacon assay which detects dissociation of a fluorescently labeled strand from a quencher strand upon Cas9 binding (Figure 3.6) (Mekler et al. 2016). The high temporal resolution and ease of parallel measurements by the assay greatly facilitate more quantitative, real-time, and higher throughput screening and kinetics characterization of multiple thgRNA variants. Results confirmed that target binding by Cas9-thgRNA is dependent on the presence of the corresponding trigger strand (Figure

3.7). Background fluorescence remained low for all designs except for **A5**, consistent with the DNA cleavage assay. Binding of Cas9 to the DNA target was restored upon addition of the trigger stand **A***. Except for **A4**, the rise in fluorescence in the presence of **A*** was comparable to that of Cas9-unmodified sgRNA (Figure 3.7), suggesting that neither extension at the 5' end of the spacer region nor hybridization of the trigger strand appreciably altered target binding capabilities of the Cas9 complex. Together with the DNA cleavage assay, we have successfully demonstrated a new generalizable framework for conditional activation of Cas9 activity using switchable gRNA structures that are modulated by toehold-mediated strand displacement.

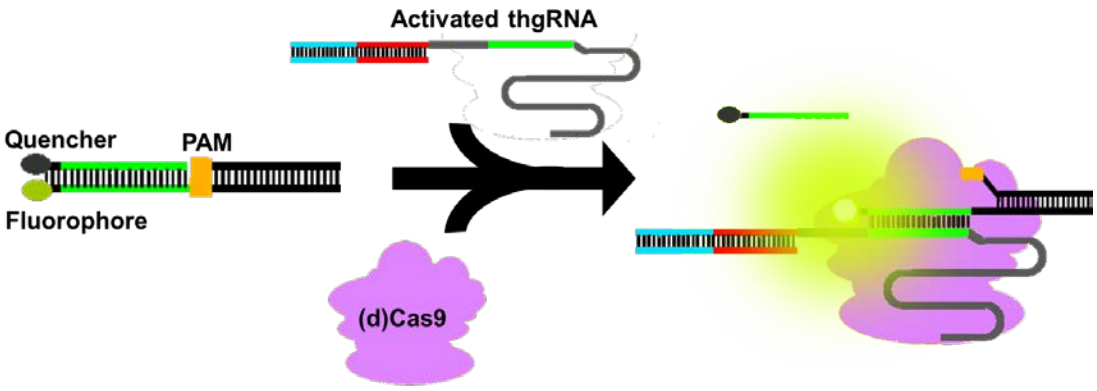
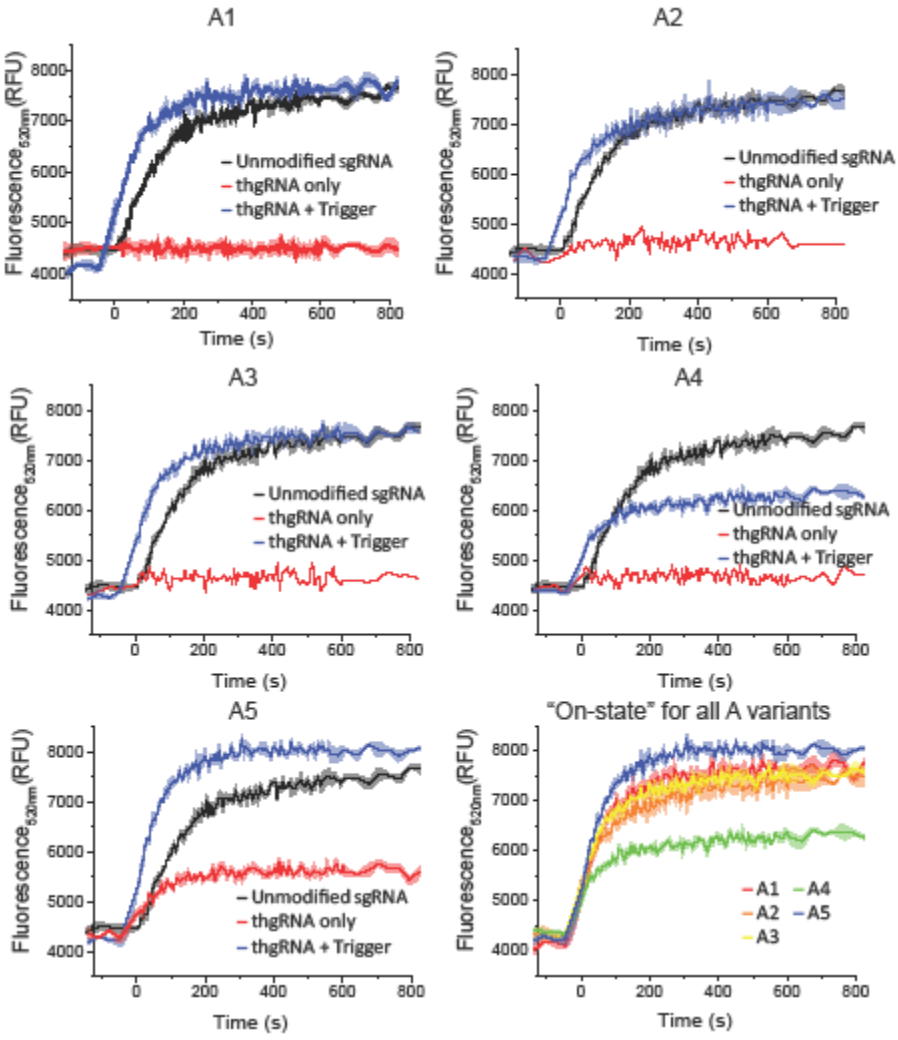


Figure 3.6 FRET-based Cas9 beacon assays for screening thgRNA variants. The assembled beacon consisting of a 5' quencher-labeled strand, a 3' fluorophore-labeled strand, and a PAM-containing strand exhibits low fluorescence as the quencher is placed in close proximity to the fluorophore. Addition of Cas9 and thgRNA/trigger or sgRNA unwinds the beacon and displaces the quencher, allowing the fluorophore to fluoresce.


A1
A2
A3
A4
A5
"On-state" for all A variants
Fluorescence520nm (RFU)
Time (s)
Unmodified sgRNA
thgRNA only
thgRNA + Trigger
A1
A2
A3
A4
A5

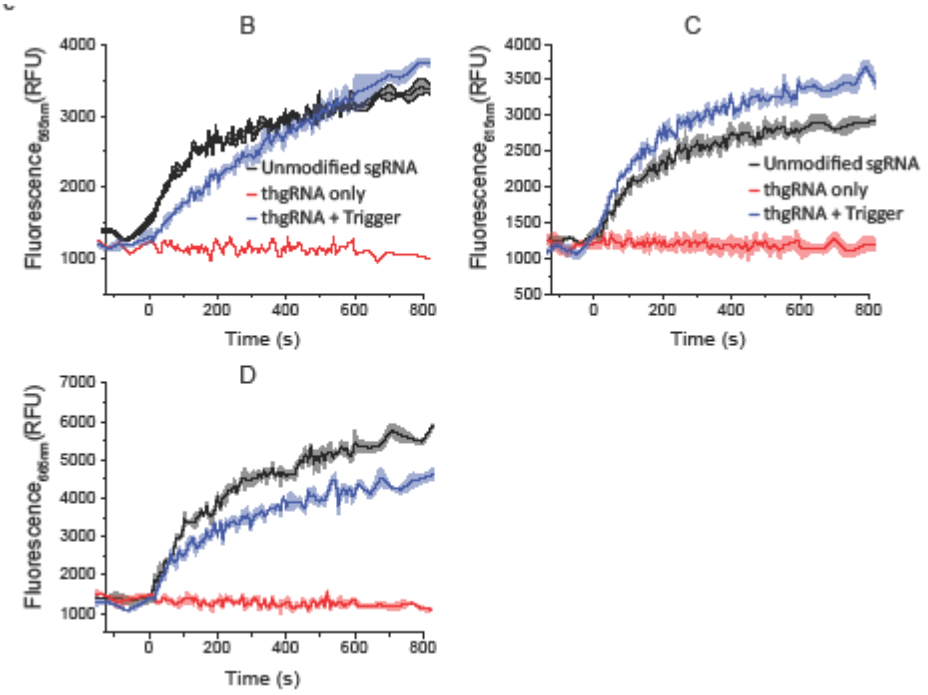


Figure 3.7 Fluorescence data for all thgRNAs **A1**-**5**, **B**, **C**, and **D**. Fluorescence measurements enabled real-time characterization of thgRNA activation of Cas9 binding in the presence or absence of the trigger strand. thgRNA and cognate trigger strands were added to the mixture at 0s. Fluorescence traces represent mean ± s.e.m. with n = 3 independent experiments.

3.4.2 Generalization of thgRNA Design to Multiple Targets

Using the design of **A1** as template, we generated three additional thgRNAs targeting three separate orthogonal DNA targets, **B**, **C**, and **D** (Figure 3.2), and experimentally characterized each using the beacon assay (Figure 3.7). Of the four designs, the toehold and branch migration sequences of **A** and **B** were entirely artificial while those of **C** and **D** were designed to hybridize with two endogenous small RNA (sRNA) sequences (OxyS and RyhB). The need to pair with native sequences imposed more constraints on the designs, resulting in a more sequestered conformation in the predicted structure of thgRNA **D** (Figure 3.2); this may have

caused the observed decreased kinetics and lowered fluorescence signal in the beacon assay (Figure 3.7), similar to thgRNA **A4**, which also exhibited diminished DNA cleavage activities as well as lowered fluorescence activation (Figure 3.5 and Figure 3.7). Nonetheless, thgRNA **A1**, **B**, **C**, and **D** all displayed negligible background activities and >75% increase in fluorescence signals relative to the unmodified sgRNA, suggesting that the basic design of thgRNA can be readily adapted for many sequences, both native or synthetic. More importantly, all four thgRNAs exhibited excellent orthogonality and provided selective activation of the corresponding beacon even when mixed together in a multiplexed setting (Figure 3.8).

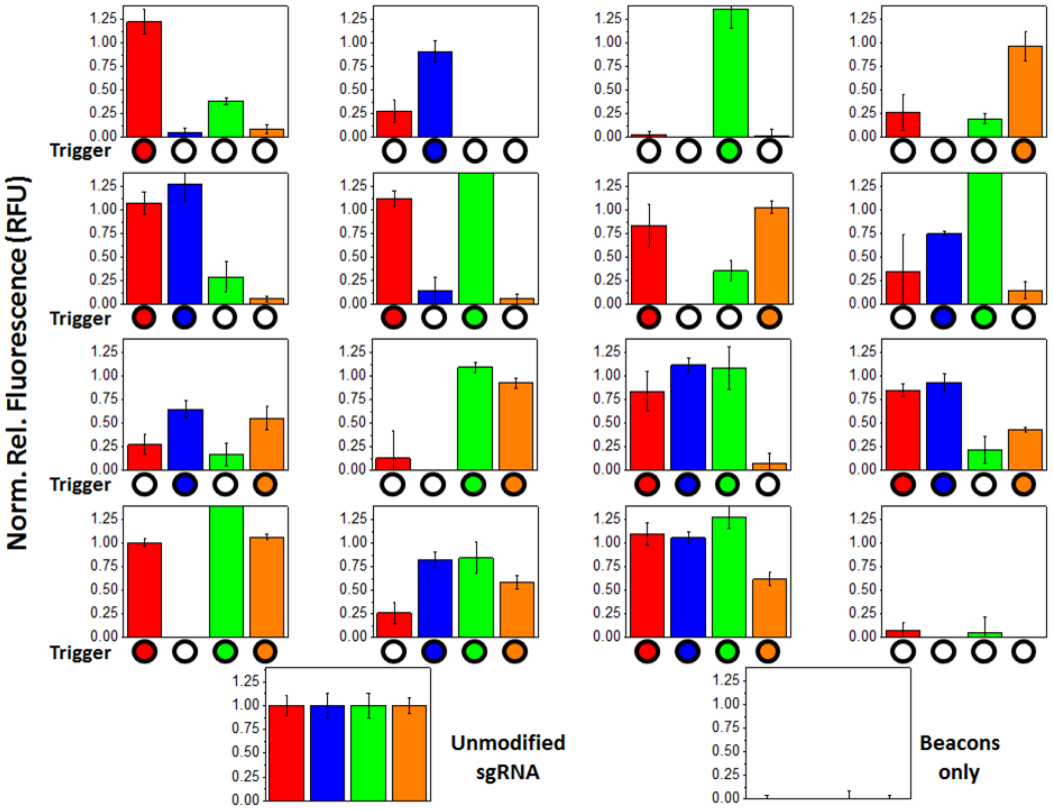


Figure 3.8 *In vitro* multiplexed beacon assay. All four beacons, thgRNA (**A1**, **B**, **C** & **D**), and Cas9 proteins are present in all samples. Fluorescence was measured at time = 600s after the addition of thgRNA/trigger strand mixtures and normalized to the average fluorescence observed with samples containing all unmodified sgRNA **A**, **B**, **C**, and **D** as shown in the bottom left panel. Filled circle(s) on the x-axis denotes the presence of the corresponding trigger (**A*** = red, **B*** = blue, **C*** = green, and **D*** = orange). Note that large increase in fluorescence is observed when and only when the corresponding trigger is present, irrespective of the presence of other non-cognate trigger strands. Values are mean ± s.e.m. with n=3 independent experiments.

3.4.3 Implementation and Characterization of thgRNAs as Intracellular Riboregulators

We next tested the use of thgRNA as an intracellular RNA-responsive switch to regulate CRISPRi-mediated Nanoluciferase (Nluc) repression in *E. coli* (Figure

3.9). Co-expression of dCas9 and the corresponding sgRNA resulted in complete Nluc repression (Figure 3.10). In contrast, co-expression of dCas9 and thgRNA had little impact on Nluc expression, indicating *in vivo* blocking of dCas9 binding. Induction of trigger RNA expression by IPTG reduced the Nluc level by >10-fold for samples expressing both thgRNA and dCas9, but had no impact for the control expressing dCas9 alone (Figure 3.11, **A1**). Although a similar ΔthgRNA without the flanking toehold sequence was also effective in blocking dCas9 activation, the addition of IPTG had little effects on Nluc expression (Figure 3.11, **A1**). This result highlights the importance of the toehold sequence and confirms opening of the blocking strand by displacement. Collectively, these results demonstrate the successful execution of toehold-gated dCas9 regulators and a good correlation between *in vitro* and *in vivo* results.

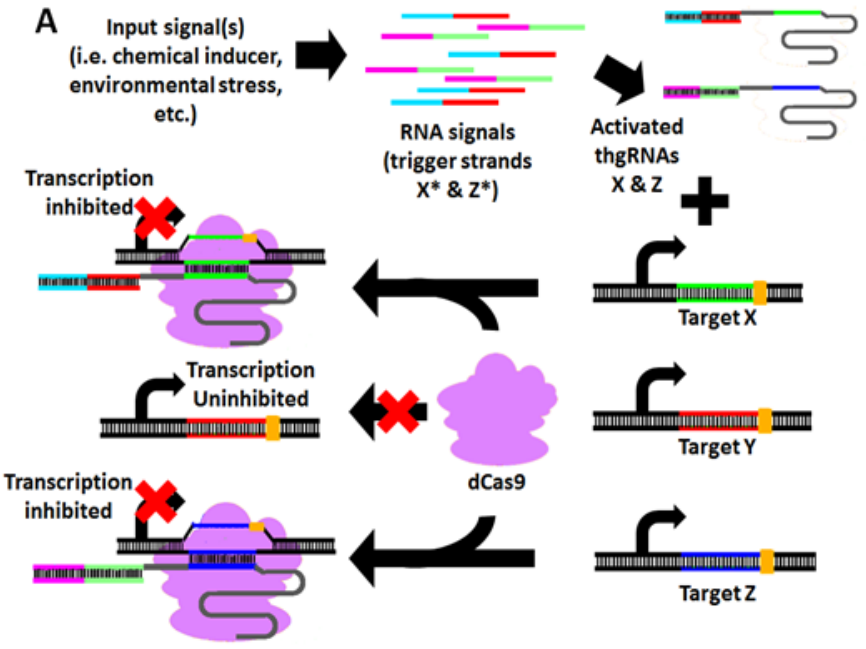


Figure 3.9 thgRNAs can be selectively activated intracellularly by induced expression of cognate trigger RNAs. RNA-responsive thgRNA could be applied to restore Cas9 functions to specific targets inside the cells, regardless of the origin of the RNA trigger sequence.

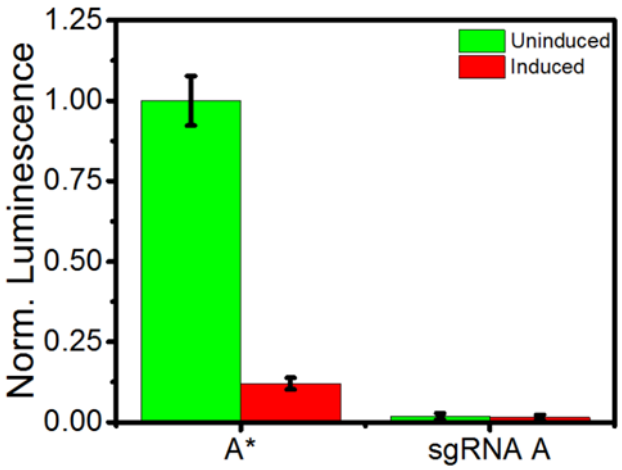


Figure 3.10 Induced repression of the Nluc reporter by trigger **A***. Luminescence is normalized to the mean value observed without **A*** induction. Values are mean ± s.e.m. with n=3 independent experiments.

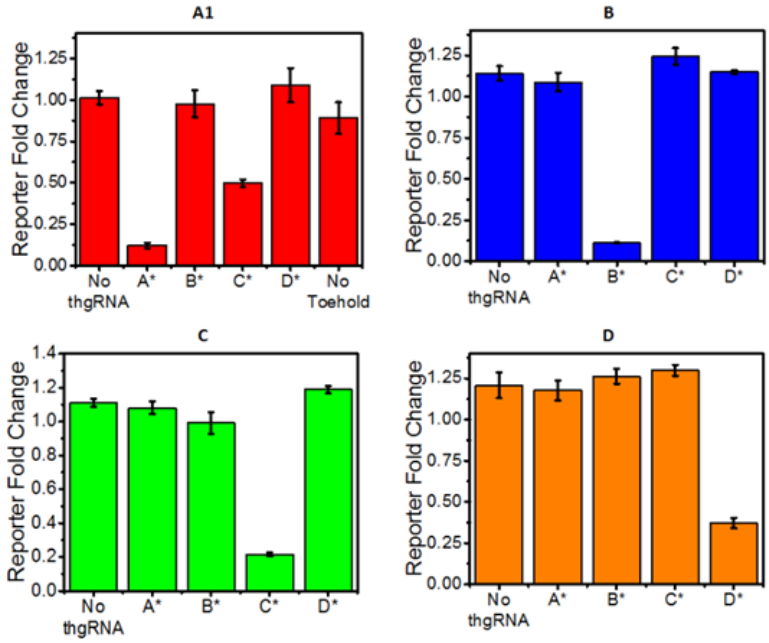


Figure 3.11 Induced repression of the Nluc reporter by different combinations of thgRNAs and triggers. Reporter fold change is defined as the Nluc luminescence observed with induction of the indicated trigger on the x-axis over the Nluc luminescence observed without induction of trigger. Values are mean ± s.e.m. with n=3 independent experiments.

To test *in vivo* orthogonality, we assembled the expression cassettes for dCas9, thgRNAs, and Nluc reporter into a single plasmid (Peng Xu et al. 2012) (Figure 3.3, Table A.2.2) and placed the different trigger strands under the control of either an inducible pL*lac*O-1 or pL*tet*O-1 promoter (Lutz and Bujard 1997) using a separate plasmid. Addition of either IPTG or ATc induces expression of the trigger strand to unblock the thgRNA, resulting in transcription repression. Although the extent of Nluc repression (4 to 12-fold) varied among the four thgRNA designs, CRISPRi was indeed activated by the cognate trigger strands in all cases (Figure 3.11). More importantly, very limited cross-talks were observed except for the ~50% repression between

thgRNA **A1** and trigger **C***. This is likely caused by the unintended interaction between thgRNA **A1** and **C*,** also observed in the beacon assay. Comparing to regular sgRNA, there is some reduction in dCas9 repression even after thgRNA is activated by the trigger and the degree of reduction varied with the different thgRNA designs. This is likely caused by the differential binding kinetics of thgRNA toward dCas9 and the trigger, leading to less than 100% unblocking and dCas9 activities. Previous studies on similar toehold-based RNA switches have yielded a wide range of dynamic ON/OFF behaviors that correlates well with the predicted thermodynamic changes upon trigger binding (Green et al. 2014). It is thus likely possible to manipulate the blocking hairpin structures of our thgRNA to further increase the thermodynamic driving force for strand displacement and Cas9 activation to mitigate these potential limitations.

Using the thgRNA **B**-trigger **B*** pair that exhibited the best repression ratio, we further evaluated the kinetics and sensitivity of the CRISPRi activation. Repression was observed as early as 1h post-induction, with a maximum repression of ~10-fold detected after 4 h (Figure 3.12). A similar trend was detected at the transcript level with a ~10-fold reduction in Nluc mRNA within 4 h. In addition, repression could be modulated by different dosages of ATc with maximum repression detected at a modest 1 ng/mL, a concentration 10-fold lower than the maximum expression achieved using the pL*tet*O-1 promoter system (Figure 3.13) (Lutz and Bujard 1997). Because the number of DNA targets are limited by the plasmid copy number, this high sensitivity is unsurprising as relatively small number of trigger strands is needed to activate Cas9 binding to all available DNA binding sites. These results demonstrate that the

switchable activation of thgRNAs by their trigger strands is not only fast and sensitive, but also provides good specificity and orthogonality.

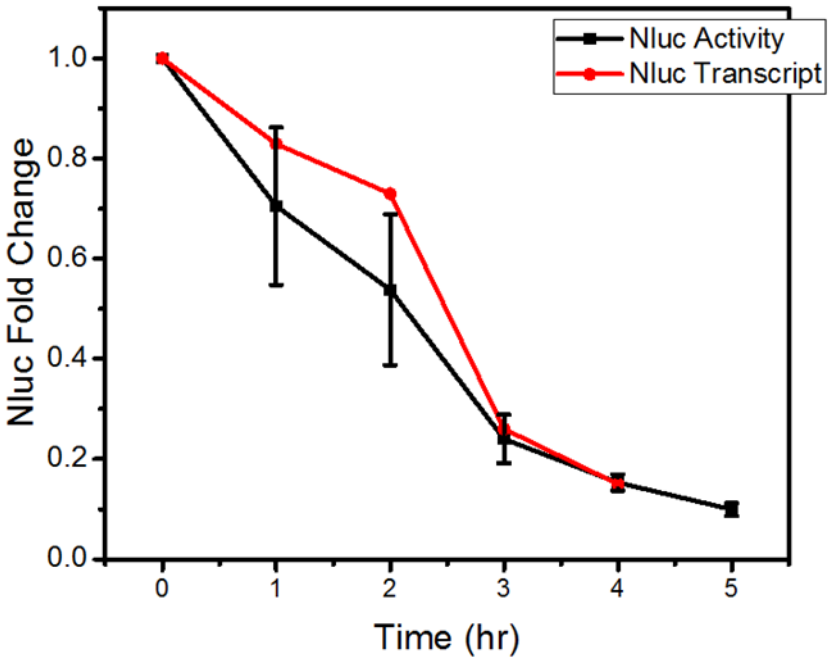


Figure 3.12 Time course for induced repression of Nanoluciferase reporter signal after induction of trigger **B***. Values are mean ± s.e.m. with n=3 independent experiments.

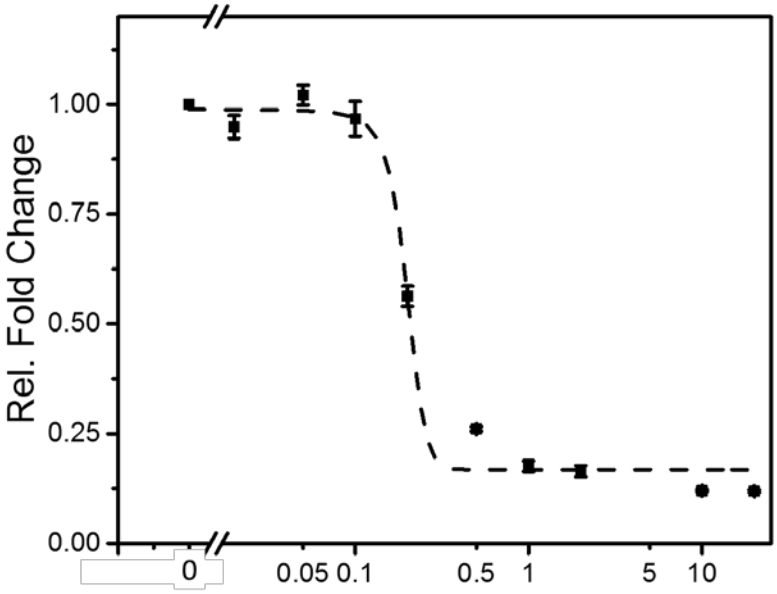


Figure 3.13 Dosage response curve of thgRNA-activated CRISPRi of Nluc repression fitted to data obtained by induction using varying concentrations of ATc. Values are mean ± s.e.m. with n=3 independent experiments.

3.4.4 Multiplexing thgRNAs for Control of Multiple Genes

To establish the possibility of deploying different thgRNAs for multiplexed regulation, we replaced Nluc with mCherry and BFP, for targets **B** and **D** to minimize any spectral overlap. We constructed a plasmid expressing the three reporters, thgRNAs, and dCas9, and a second set of plasmids expressing different trigger strands using separate inducible promoters (Table A.2.2). Significant repression was detected only when the appropriate triggers were induced with no significant cross-talks observed in all cases (Figure 3.14). Given the high selectivity of using thgRNAs for multiplexed regulation, we believe this new strategy can be easily layered to construct even more complex genetic circuit designs.

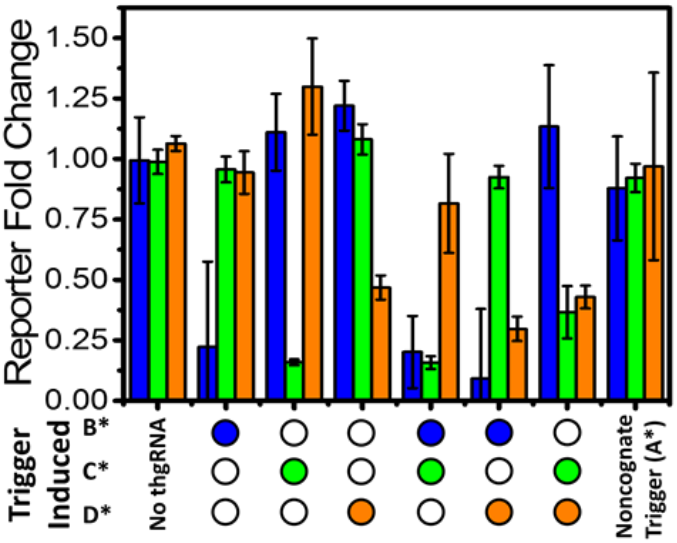


Figure 3.14 Multiplexed activation of thgRNA-dCas9-mediated repression by induced expression of specific trigger strands indicated on the x-axis. Normalized fold change is defined as the reporter signal observed with induction of the indicated trigger(s) over the signal observed without induction of trigger(s). Values are mean ± s.e.m. with n=3 independent experiments.

3.4.5 Integration of Endogenous Cellular Information for thgRNA Regulation

Another important feature is to incorporate endogenous RNAs as the trigger (Green et al. 2014; Groves et al. 2015). To demonstrate this possibility, we designed thgRNAs **C** and **D** to be responsive to two small RNAs (sRNA), OxyS and RyhB, respectively (Altuvia et al. 1998; Massé, Escorcia, and Gottesman 2003) (Figure 3.15). These sRNA were chosen for their relative stability and abundance after different stress stimuli. Expression of RyhB is induced by iron deficiency, which can be experimentally simulated by addition of 2'2-bipridyl, while OxyS is expressed upon exposure to oxidizing agents such as H_2O_2 or paraquat. Addition of 2'2-bipridyl resulted in a dose-dependent repression in Nluc expression only by paring thgRNA **D** with the corresponding reporter cassette (Figure 3.16). Similar experiments using

H_2O_2 or paraquat were less successful as cell growth and protein expression were significantly inhibited. To allow more direct assessment, expression of full-length OxyS sRNA was induced artificially and a significant decrease in Nluc level, similar to that observed with the synthetic trigger **C***, was detected (Figure 3.17).

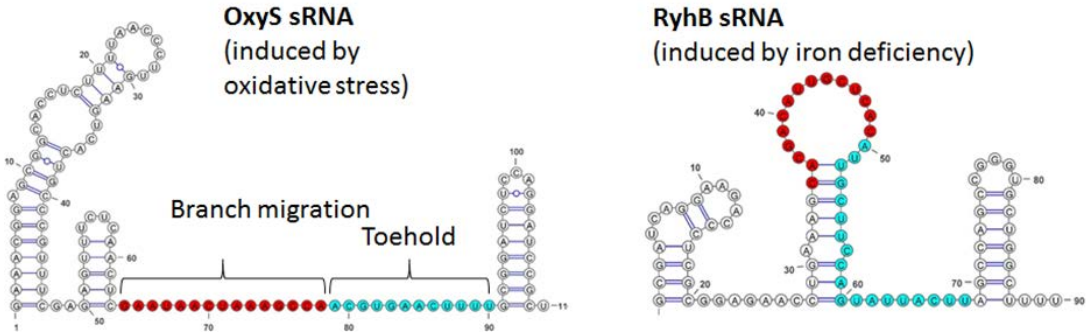


Figure 3.15 Endogenous sRNA can be used as trigger strand to activate CRISPRi through thgRNA. Specific sRNA can be induced by exposure to different environmental stresses, such as OxyS by oxidative stress and RyhB by iron-deficiency. thgRNA **C** and **D** were designed to hybridize with these sRNAs, respectively.

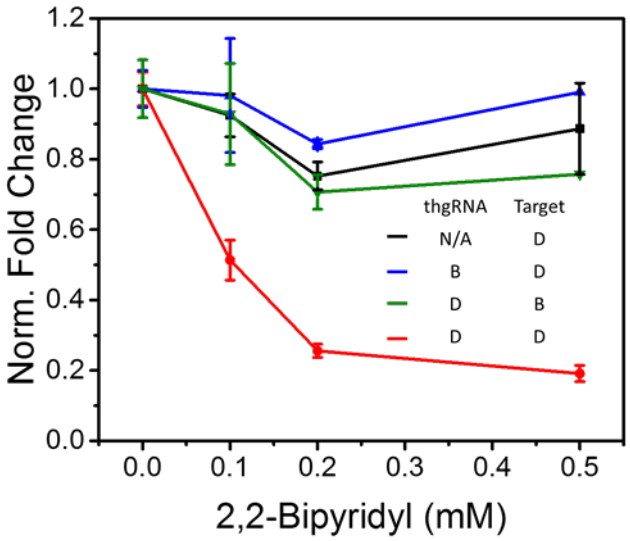


Figure 3.16 Induced repressed of the Nluc reporter using the native iron-responsive RyhB sRNA as the trigger. Repression of Nluc by addition of increasing concentrations of 2,2'-bipyridyl is observed only when both thgRNA D and target D are both present (red). Values are mean ± s.e.m. with n=3 independent experiments.

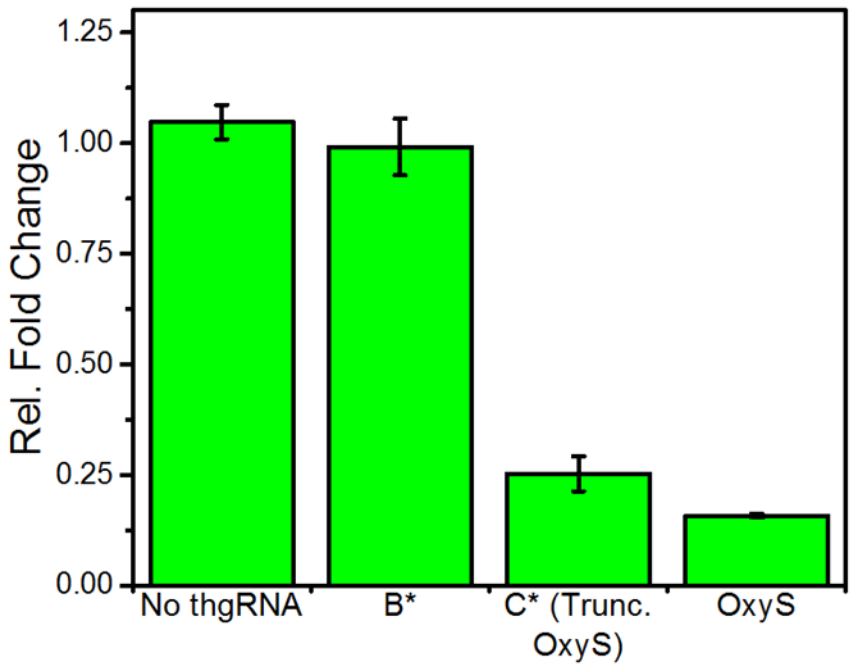


Figure 3.17 Induced repression by full-length OxyS sRNA. Values are mean ± s.e.m. with n=3 independent experiments.

We further demonstrated the generality of the approach for different endogenous sequences by designing six additional thgRNAs targeting two native sRNAs (97bp MicF and 227bp SgrS) and three regions of the full-length mCherry mRNA (Figure 3.2, Table A.2.1). While not all the new thgRNAs were as effective as **C** and **D**, most exhibited a reasonable level of intracellular responses and even the thgRNAs designed for the longer SgrS and the full-length mCherry transcripts (**F** and **G2**) elicited >60% repression (Figure 3.18). Importantly, no notable decrease in mCherry fluorescence was observed for any samples that exploited mCherry mRNA as the trigger, indicating that translation of the mRNA trigger is not inhibited

significantly even after binding to the thgRNA (Figure 3.19). These results highlight the feasibility of designing thgRNAs that are responsive to a range of native RNA sequences, including full-length mRNA, without compromising the cellular functions of the endogenous strands.

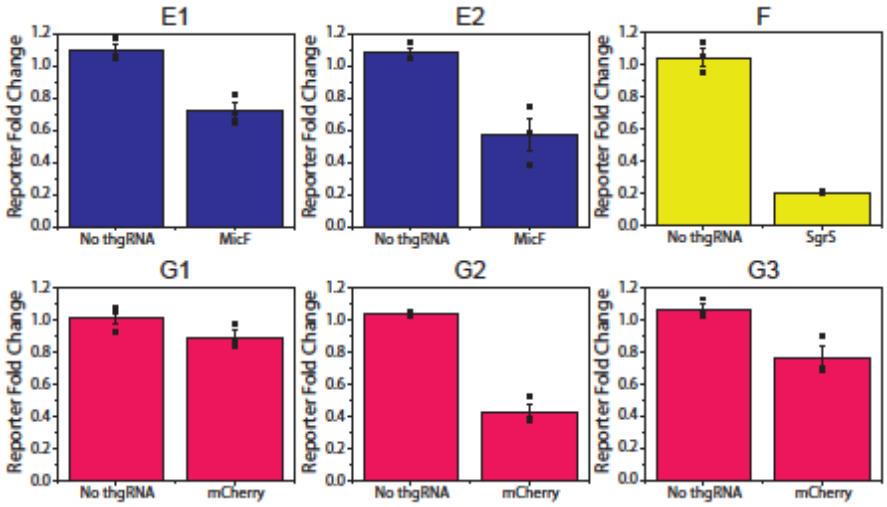


Figure 3.18 Induced repression by additional full-length endogenous sequences. Six additional thgRNA that designed to be responsive to three arbitrarily chosen endogenous trigger sequences: two (**E1** and **E2**) for MicF sRNA, one (F) for SgrS sRNA, and three (**G1-3**) for mCherry mRNA. All additional thgRNA displayed some level of activities, though only two out of six of these additional thgRNA led to ≥40% repression, despite the lack of more extensive *in vitro* characterization and refinement. Including thgRNA **C** and **D**, these brings the total of four out of eight thgRNAs (50%) that were able to be activated by endogenous sequences robustly. Importantly, the only endogenous sequence out of five studied that failed to activate any thgRNA tested and induce >50% repression was MicF sRNA, suggesting that there is a high probability of successful thgRNA design for any given sequence. Values are mean ± s.e.m. with n=3 independent experiments.

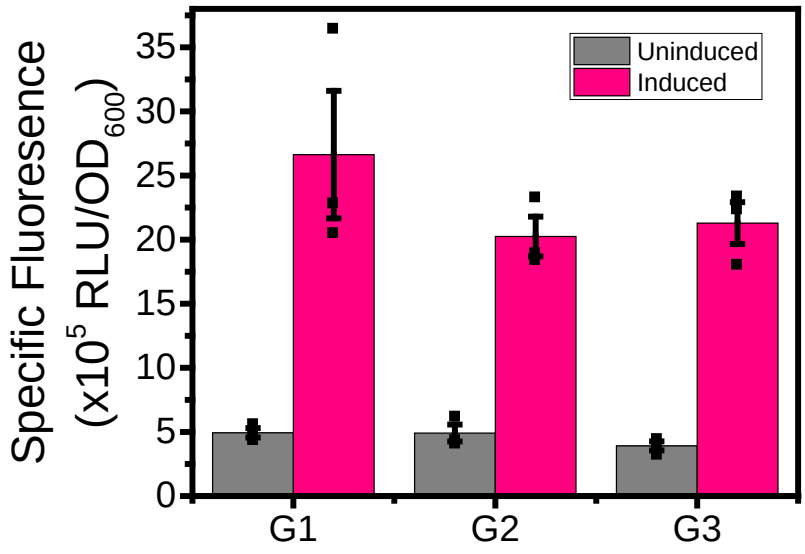
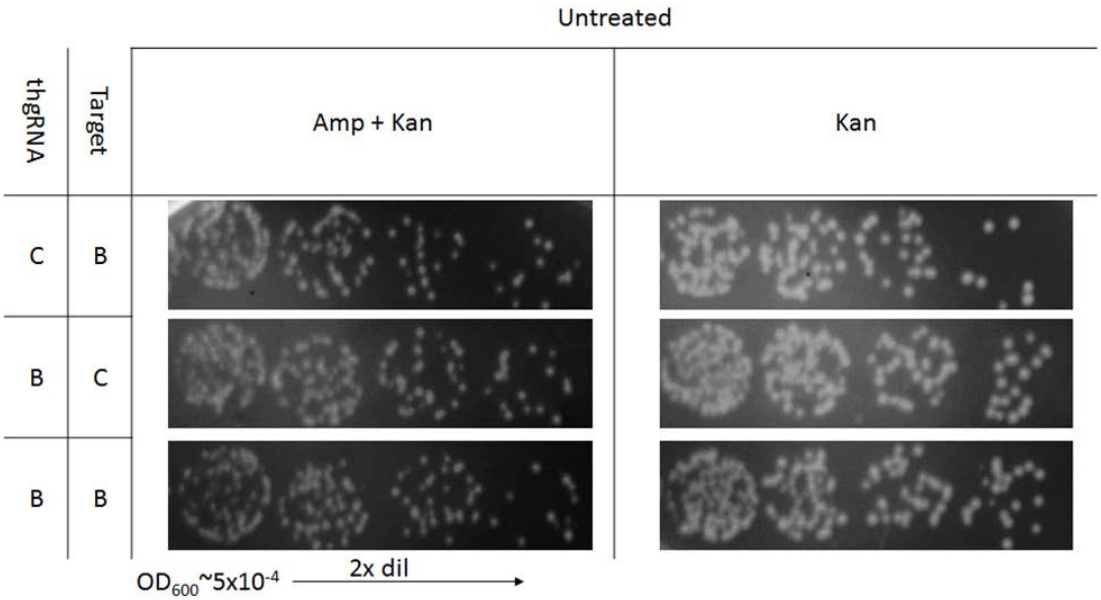


Figure 3.19 Observed mCherry fluorescence exhibited no correlations with degree of induced repression by activated thgRNA. All samples induced to express mCherry mRNA to be used to activate thgRNA showed similar levels of fluorescence, suggesting that the use of the mRNA as trigger strand does not significantly inhibit its ability to be translated by cellular translational machinery. Values are mean ± s.e.m. with n = 3 independent experiments.

3.4.6 Inducible Gene Knockout through Activation of thgRNA

To expand the use of thgRNAs as a tool for inducible gene knockout (Ran, Hsu, Wright, et al. 2013), we investigated whether the same thgRNA/trigger pair **B** can be used to induce plasmid loss by nuclease active Cas9. We introduced two plasmids (containing either AmpR or KanR selection marker) into *E. coli* and induced the expression of trigger **B*** by ATc, which selectively cleaves the plasmid containing AmpR. After 4 h induction, we plated cells onto different selective agar plates to estimate the loss of the AmpR plasmid (Figure 3.20). We observed a 60% reduction in

the number of colonies by the induced sample relative to the un-induced sample, while there were no observable differences for the control samples carrying mismatched thgRNA and trigger pairs. The ability to provide conditional gene knockout in *E. coli* paves the way for inducible genome editing in eukaryotic systems based on differential RNA expression.


Untreated
thgRNA
Target
Amp + Kan
Kan
C
B
B
C
B
B
OD600~5x10-4
2x dil

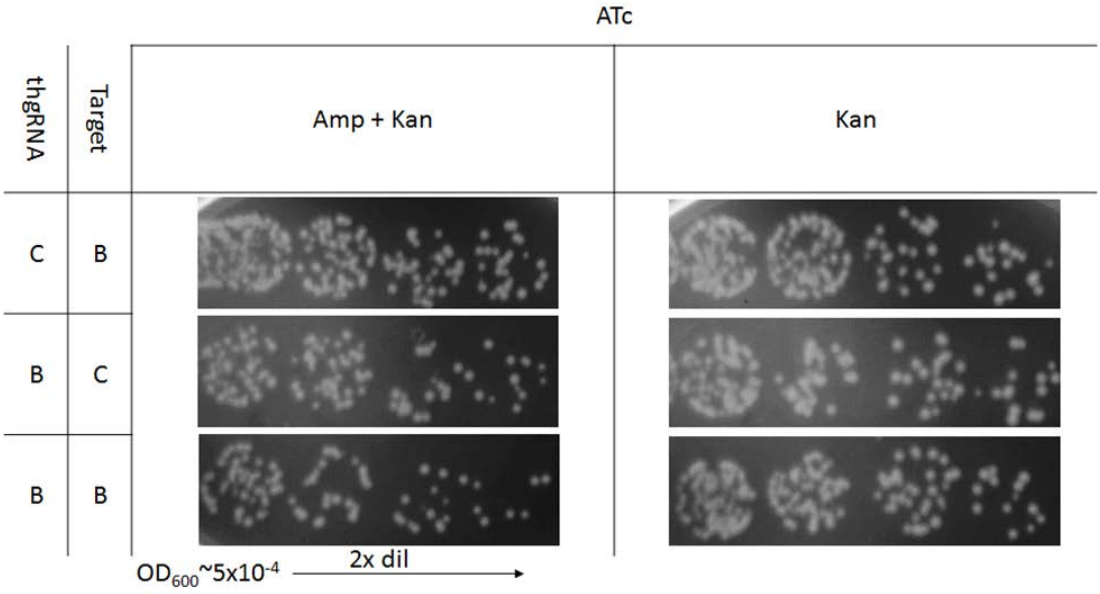
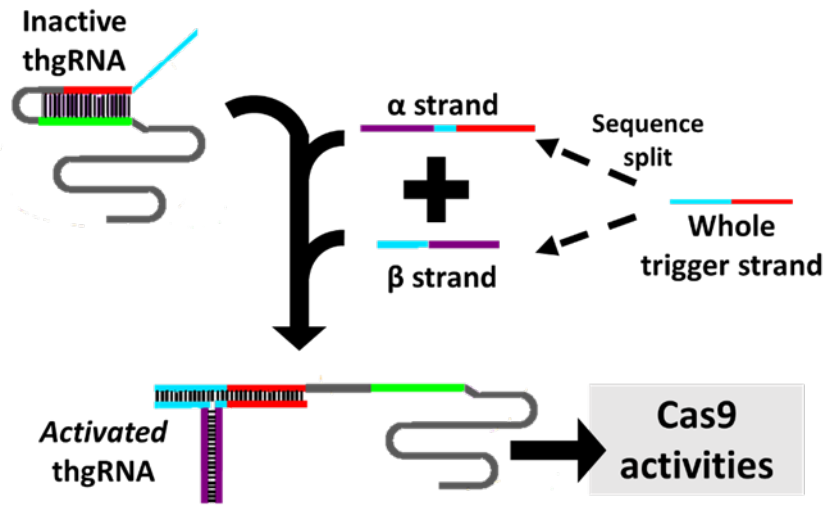

ATc
thgRNA
Target
Amp + Kan
Kan
C
B
B
C
B
B
OD600~5x10-4
2x dil

Figure 3.20 Spot plates for induced plasmid loss. The loss of ampicillin resistance and thus its associated plasmid can be assessed by the ability to form colonies. By comparing the number of colonies in each dilution series between the two selective plates, we can estimate the fraction of transformed cells that have lost the AmpR and the Cas9-thgRNA target-bearing plasmid. Note that induced plasmid loss were only observed when both thgRNA **B** and target **B** were present, with ~40% of cells still retaining functional ampicillin resistance after 4 hours of trigger induction (last row).

3.4.7 Use of thgRNAs as Riboregulators in Context

Unlike previous approaches, which rely on distinct conformational change induced either by ligand binding or by RNA cleavage to unblock the spacer region (W. Tang, Hu, and Liu 2017; Liu et al. 2016), our thgRNA strategy offers a simple 'plug and play' design for conditional activation of CRISPR-based systems by a virtually unlimited set of RNA triggers using the highly predictable toehold-mediated strand displacement reactions. Because activation is enabled by sequence-specific unblocking of the spacer region using a specific RNA trigger, this design offers a high degree of orthogonality with very low system cross-talk with other unrelated endogenous information. These advantages enabled us to create a small panel of riboregulators each responsive to different input trigger sequences with relative ease. Even with the added restrictions targeting endogenous sRNA and mRNAs, we were successful in designing several new thgRNAs with only minor refinements, suggesting a high degree of sequence versatility for the trigger. The flexibility to exploit endogenous RNAs to regulate gene expression provides a simple interface between native signals and synthetic transcriptional outputs, bypassing the often tedious process of screening large libraries currently needed to create specific LRTFs or riboregulators.

Theoretically, our thgRNA design allows us to execute complex multi-input logic operations by stacking several RNAs into a single trigger. Preliminary results of such a control scheme using multiple trigger strands to activate thgRNA is briefly discussed in Figure 3.21. Because of the potential suggested by these results and the similar success that has recently been demonstrated using optimized toehold switch design by stacking up to five different RNA transcripts into a single trigger (Green et al. 2017), we envision these thgRNA can be layered into multiplexed control for more complex, self-regulating genetic circuits for both gene repression and activation as well as inducible gene knockouts. However, we also anticipate the need for further optimizations of thgRNA and trigger designs to account for additional secondary structures and interactions with Cas9 protein. These complications may introduce kinetic and thermodynamic barriers that may prevent efficient activation of thgRNAs. Investigations on multi-input designs are underway.


Inactive
thgRNA
α strand
Sequence
split
Whole
trigger strand
β strand
Activated
thgRNA
Cas9
activities

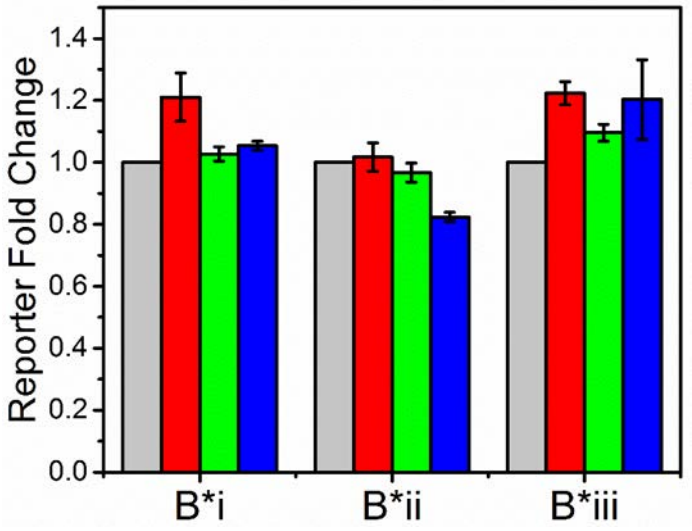

Reporter Fold Change
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
B*i
B*ii
B*iii

Figure 3.21 Schematics and preliminary results of induced repression using split trigger strands. Schematics of split trigger strands for activation of thgRNA. **α** and **β** strands contain the two fragments of the whole sequence used in one-input activation of thgRNA. Preliminary results using split trigger versions **B*i-iii** to activate thgRNA **B** (uninduced = gray, strand α induced by ATc = red, strand **β** induced by IPTG = green, both strands induced = blue). A modest (~20%) but statistically significant ($P<0.05$; one-sided student's *t*-test) repression of the Nanoluciferase reporter was observed when and only when both α and β strands of split trigger **B*ii** were induced, suggesting the potential to adapt this basic design for multi-input control schemes with optimization. Values are mean ± s.e.m. with n=3 independent experiments.

3.5 Acknowledgements

This work was supported by grants from NSF (CBET1510817 and MCB1615731).

Chapter 4

COILED-COIL STRAND DISPLACMENT FOR DYNAMIC CONTROL OF INTRACELLULAR PROTEIN INTERACTIONS AND FUNCTIONS

4.1 Summary

Intracellular protein interactions found in nature have evolved to be sensitive to changes in environmental, intercellular, and intracellular conditions. Ideally, synthetic biomolecular interactions used to construct control circuits should also exhibit the same trait; however, even the current state-of-the-art synthetic modules often fall short of mimicking nature, particularly in implementing dynamic responses to react to the ever-changing environment. This disparity is partly caused by the inability to reliably predict protein-protein interactions *a pirori*, limit unspecific associations with both synthetic and endogenous components within the cell, and reconfigure protein assemblies in response to stimuli. To address these deficiencies, we have employed a class of protein structural motif known as coiled-coils as synthetic actuators and exploited their capacity to partake in strand displacement reaction to impose dynamic controls over protein interactions.

4.2 Introduction

The α-helical coiled-coil motif is currently the most well-characterized protein folding motif that has been successfully used in bioinspired systems involving self-assembly. The success of this motif is driven by the straightforward sequence-to-structure relationships, specifically the repeating heptad pattern of hydrophobic and

polar residues (Gradišar and Jerala 2011), which substantially facilitate design of individual coils and reliably predict interactions between different coils *a priori*. The ability to design molecular association *de novo* has enabled construction of sizable libraries of orthogonal coiled-coils (Reinke, Grant, and Keating 2010) and applications such as biomaterials (Banwell et al. 2009) and protein labelling (Meyenberg et al. 2011).

While extensive works have been devoted to constructing supramolecular complexes using the existing coiled-coils, comparatively few studies have applied coiled-coils for dynamic control of protein-protein interactions. Recently, Gröger et al. demonstrated reversible coiled-coil association using pairs of heterodimeric coiled-coil based on the Hodges EK peptides by modulating the lengths of the individual coils (Gröger, Gavins, and Seitz 2017). The varying lengths of the coiled-coils resulted in a range of dissociation constants of μM to sub-nM, allowing longer coils to dynamically displace shorter coils akin to strand displacement for nucleic acids (D. Y. Zhang and Winfree 2009). Displacement of these heterodimeric coiled-coils was used to create a two-state switching system to control interaction of attached moieties such as fluorescent dyes and protein kinase substrates. Thus far, this application represents the only reported example of dynamic coiled-coil systems using peptides composed of only native proteinogenic amino acids. Inspired by the potential of this dynamic control system to direct protein-protein interactions, we seek to expand the utility of coiled-coil peptide-based strand displacement. The ease to directly express coiled-coil motifs recombinantly as peptide tags makes them ideal synthetic actuators to coordinate functions of multiple proteins, particularly for intracellular applications such as reversible transcriptional switch and enzyme complex rearrangement.

4.3 Materials and Methods

4.3.1 Strains used and Plasmids Construction

All strains and plasmids used in this chapter are listed in Table A.3.2.

Split Nanoluciferase-coiled-coil fusions are constructed by standard subcloning techniques using restriction digests and ligation into pMAL™-c5x vector (New England BioLabs®, Inc., Ipswich, MA, USA). Yeast expression constructs were generated by method described in details by Lee et al (Lee et al. 2015). Briefly, MoClo-YTK "parts" containing the genes of interests were generated via BsmBI Golden Gate assembly. These parts were then combined with other parts from the MoClo-YTK toolkit in a BsaI Golden Gate reaction and transformed into NEB®5α cells. Each cassette was then assembled via another BsmBI Golden Gate reaction into a multi-gene plasmid for transformation and expression into yeast.

4.3.2 Luminescence by Reconstituted Split Nanoluciferase Fusions

Plasmids containing split Nanoluciferase-coiled-coil fusions are transformed into NEB® Express cells for expression. Successful transformants were picked from agar plates and grown in LB medium (10g/L tryptone, 5g/L yeast extract, 10g/L NaCl) supplemented with 100µg/mL ampicillin overnight at 37°C. The resulting cultures were used to inoculate 25mL LB medium supplemented with 100µg/mL ampicillin at an initial OD_{600} ~0.05 and grown for 2hrs to OD_{600} ~0.7. Expression of the split Nanoluciferase-coiled-coil fusions were then induced by addition of 300µM IPTG. These cultures were then incubated at 37°C for 2hrs, after which the cultures were harvested and centrifuged into cell pellets. The harvested cell pellets were subsequently resuspended in PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4) to OD_{600} ~20 and lysed using sonication. Lysed samples

were centrifuged at 13000*g* for 20min and the soluble fractions were collected as the supernatants. The resulting samples were analyzed by standard SDS-PAGE and densitometry analysis to estimate the purity and quantity of split Nanoluciferase-coiled-coil fusion proteins within the collected fractions (Figure 4.1).

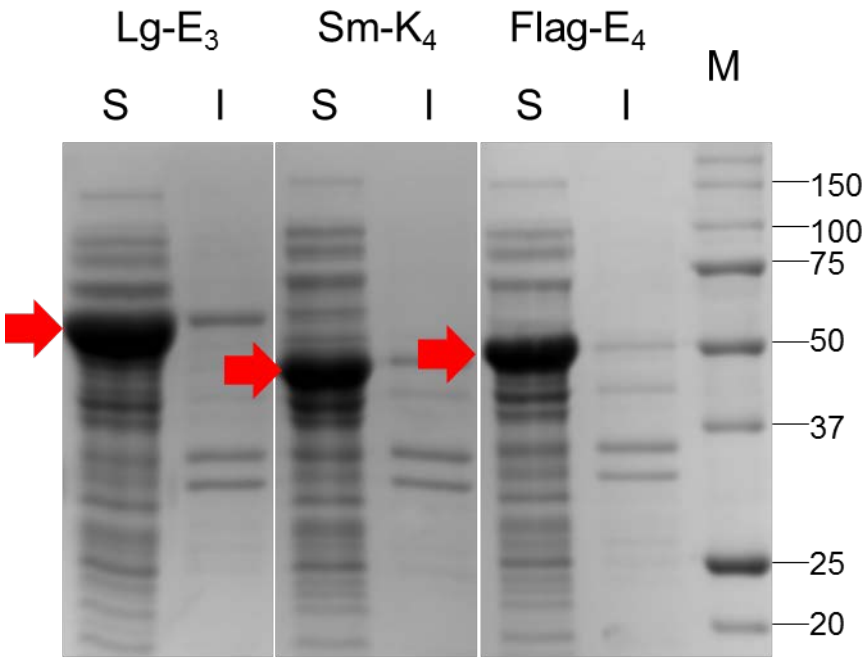


Figure 4.1 Expression of split Nanoluciferase-coiled-coil fusions in NEB® Express cells. Cell lysates were analyzed by standard SDS-PAGE. All fusions contain an N-terminal Maltose-binding protein (MBP) tag.

Reaction mixtures containing varying amounts of split Nanoluciferase-coiled-coil fusions were incubated at room temperature for 45min prior to luminescence measurement. Luminescence assays were performed according to NanoGlo® vendor's instructions (Promega Corporation, Madison, WI, USA) using a Synergy™ H4 Hybrid microplate reader.

4.3.3 Förster Resonance Energy Transfer (FRET) in Yeast

Plasmids with expression cassettes of recombinant fluorescent protein-coiled-coil fusions were transformed into BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) cells. Successful transformants were picked from synthetic dropout medium (SD-2xSCAA; 6.7 g/L yeast nitrogen base w/o amino acids, 20g/L raffinose, 0.2g/L dextrose, 10.19g/L Na_2HPO_4; and 8.56 g/L NaH_2PO_4, amino acid supplements, pH 6.6) (Wittrup and Benig 1994) agar plates with the appropriate selective dropout marker and grown overnight at 30°C. Expression of the displacer coiled-coil fusion was then induced by inoculating fresh SD-2xSCAA medium containing 20g/L raffinose and 0.2g/L galactose in place of dextrose at an initial OD600~0.1 with the grown overnight cultures. These subcultures were incubated at 30°C and samples were taken at 0, 2, 4, and 20hrs post-induction.

Collected samples were pelleted by centrifugation at 3000*g* for 5 minutes, washed twice with PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4), and resuspended to OD_{600} ~1 in PBS prior to measurements. Whole cell fluorescence were measured using a Synergy^{TM} H4 Hybrid microplate reader. Analysis of fluorescence data were completed using Origin software (OriginLab Corporation, Wellesley Hills, MA, USA).

4.3.4 Yeast Violacein Production and Analysis

Plasmids containing genes for violacein biosynthesis were constructed as described above. The parent yeast strain for these experiments was BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). The complete 5-gene cassette for violacein pathway production was integrated at the leu2 locus to create strain designated as ySK001 (Table A.3.2). To introduce the different fluorescent protein-coiled-coil fusions, multi-

gene plasmids with CEN2/ARS6 origin with LEU marker were transformed into ySK001. Transformants were selected as colonies on SD-2xSCAA-Ura-Leu agar plates after incubation of ~48hrs at 30°C.

Successful transformants were then grown in SD-2xSCAA-Ura-Leu supplemented with 2% raffinose + 0.2% dextrose overnight at 30°C. The resulting cultures were then used to inoculate fresh SD-2xSCAA medium supplemented only with 2% raffinose at an initial OD_{600}~0.1 as subcultures and grown for additional 4hr at 30°C. Expression of the displacer coiled-coil fusions were then induced by addition of 0.2g/L galactose and/or 50µM Cu_2SO_4 into the subcultures. The cultures were then grown for 36hr at 30°C and harvested as pellets by centrifugation at 3000*g* for 10 min.

Pelleted cells were suspended in 300µL methanol and boiled at 95°C for 10 minutes, vortexing twice during the incubation. Solutions were then centrifuged twice at 13000*g* for 10min to remove cell debris, and the supernatant (extract) was analyzed by HPLC/MS on a Waters AutoPurification HPLC/MS System with a SunFire™ C18 column. Absorbance of the eluent at 560nm and 600nm were measured for violacein/deoxyviolacein and proviolacein/prodeoxyviolacein respectively, while mass was measured with the attached SQD2 MS system to confirm the identity of the isolated molecule corresponding to each eluent peak (Figure 4.2). The product yields can be estimated as the integrated peak areas from the chromatograms by comparison to standards of violacein (MilliporeSigma, St. Louis, MO, USA).

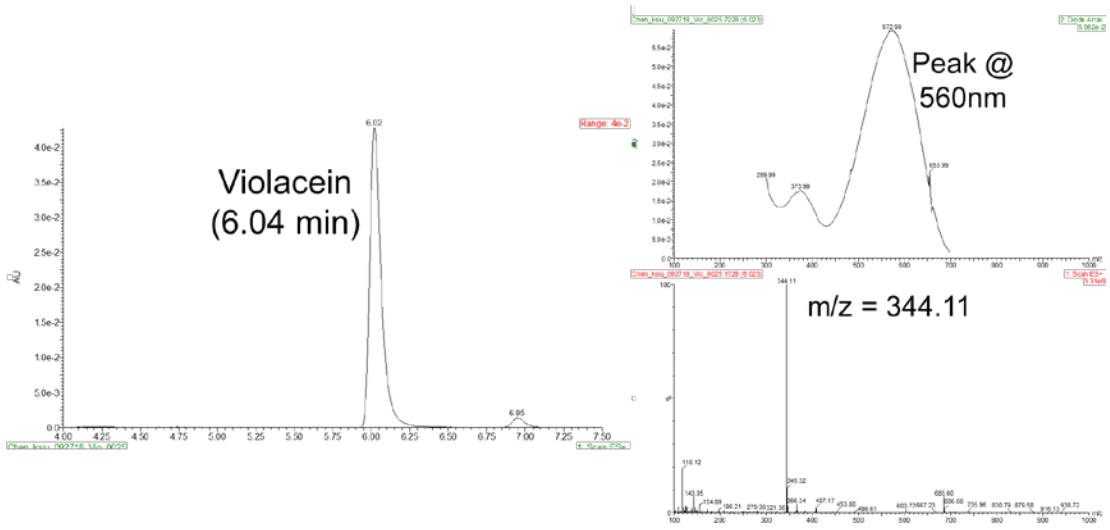


Figure 4.2 Representative chromatogram, photo spectrum, and mass spectrum of product obtained from yeast violacein biosynthesis highlighting the detection of violacein from mixtures.

4.4 Results

4.4.1 Strand Displacement using Heterodimeric Coiled-coils can Control Protein Interaction

To demonstrate the use of coiled-coils that can be recombinantly expressed as peptide tags to control protein-protein interactions, we first constructed pairs of EK coiled-coils fused to MBP and split Nanoluciferase reporters (Figure 4.3). Split Nanoluciferase was developed by splitting NanoLuc luminescent reporter into two subunits, LgBit and SmBit, and was shown to exhibit low affinity between the two split fragments (K_D~190μM), making them ideal for use in protein complementation assays to investigate of protein interactions both in and out of the cell (Dixon et al. 2016). The recombinant fusions were then expressed in *E. coli* and harvested (Figure 4.1). The resultant cell lysates were then mixed in concentrations indicated in Figure 4.3 below. Each reaction mixture was further subjected to NanoLuc® luminescence

assays according to manufacturer's instructions (Promega Corporation, Fitchburg, WI, USA). As expected, the addition of an equimolar amount of the longer E_4 coil fused to MBP and a Flag peptide led to a dramatic decrease (>25-fold) in luminescence, suggesting that the LgBit-E_3 fusion was displaced from the SmBit-K_4 fusion by Flag-E_4. This decrease in luminescence was driven lower by another order of magnitude by adding slight excess of the longer Flag-E_4 protein to ensure stoichiometric pairing of E_4 and K_4 coils. Further increase in the amount of Flag-E_4 added did not yield lower luminescence (data not shown), indicating that the low luminescence signal originated from spontaneous reconstitution of LgBit and SmBit caused by their low affinity at these concentrations. These results demonstrated the ability to recombinantly express coiled-coils motifs as peptide tags to modulate protein-protein interactions.

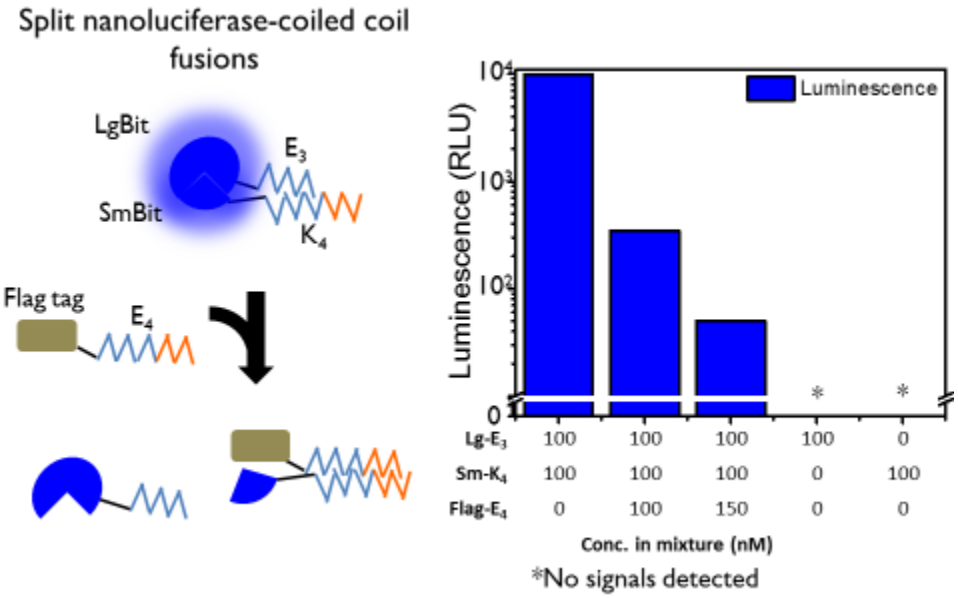


Figure 4.3 EK coiled-coils recombinantly fused and expressed as peptide tags onto proteins of interests. Each fragment of the split Nanoluciferase reporter and a Flag peptide tag is fused to either an E or K coil and an N-terminal MBP tag (MBP not shown for simplicity). The fusions were then expressed in *E. coli* and harvested. Cell lysates containing the different fusions were then mixed and tested for luminescence. Luminescence assays for reaction mixtures containing different combinations of coiled-coil fusions suggested that the fragments of the split Nanoluciferase reporter were displaced from close proximity by strand displacement of the E_3 coil by the longer E_4 coil.

4.4.2 Coiled-coil Strand Displacement is Functional Intracellularly

After confirming strand displacement functions of coiled-coil motifs fused to proteins *in vitro*, we next designed a three-color FRET system in yeast using the EK coils and three fluorescent proteins to demonstrate the use of coiled-coil displacement as a dynamic mechanism to control intracellular protein interactions (Figure 4.4). The fluorescent proteins are cloned and expressed as N-terminal fusions of the EK coils using the MoClo-YTK toolkit (Lee et al. 2015) (Table A.3.2). In this scheme, mTurquoise2-E_3 and Venus-K_4 are constitutively expressed and self-assemble into

heterodimers that place mTurquoise2 into close proximity of Venus. Upon induction of mRuby2-E_4 by galactose, mTurquoise2-E_3 is displaced from Venus-K_4, thereby gradually switching from the initial cyan-to-yellow FRET signal to yellow-to-red signal and increasing mTurquoise2 emission.

The FRET signal changes observed following galactose induction suggested that coiled-coil strand displacement can indeed control protein interactions in an intracellular setting (Figure 4.5). Induction of mRuby2-E_4 resulted in the restoration of mTurquiose2 emission spectrum and the reduction of Venus emission at 530nm, suggesting that mRuby2 replaced mTurquoise2 in the heterodimers. Importantly, the rate of displacement is notably enhanced by inducing the expression of mRuby2-E_4 relatively to inducing expression of mRuby2-E_3 (Figure 4.6). Compared to mRuby2-E_4, mRuby2-E3 is unable to displace mTurquoise2-E3 as quickly or as completely, suggesting that the extra heptad repeat is acting as a "toehold" region for initiation of the strand displacement reaction akin to DNA strand displacement (D. Y. Zhang and Seelig 2011) described in Chapter 1.3.2. Lastly, yellow-to-red FRET ratio rise slowly compared to the decrease in blue-to-yellow FRET ratio, likely because mRuby2 needs time to mature ($t_{1/2}$~150min) to become fully fluorescent. These results suggests that higher thermodynamic drive to form the E_4/K_4 heterodimeric complex endowed by the toehold region was necessary for the dynamic rearrangement of the fluorescent protein complexes.

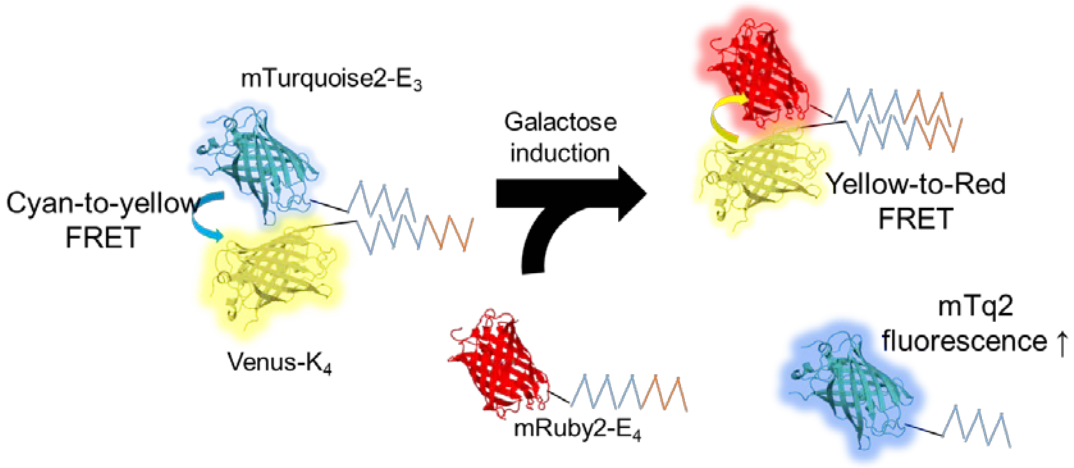


Figure 4.4 Detection of intracellular coiled-coil strand displacement using three-color Förster resonance energy transfer system. The initially dominant cyan-to-yellow FRET is disrupted as Venus-K_4 is displaced away from mTurquoise2-E_3 and instead complexes with mRuby2-E_4, shifting the FRET to yellow-to-red as mRuby2 fluorescent protein matures and allowing mTurquoise2 to fluoresce.

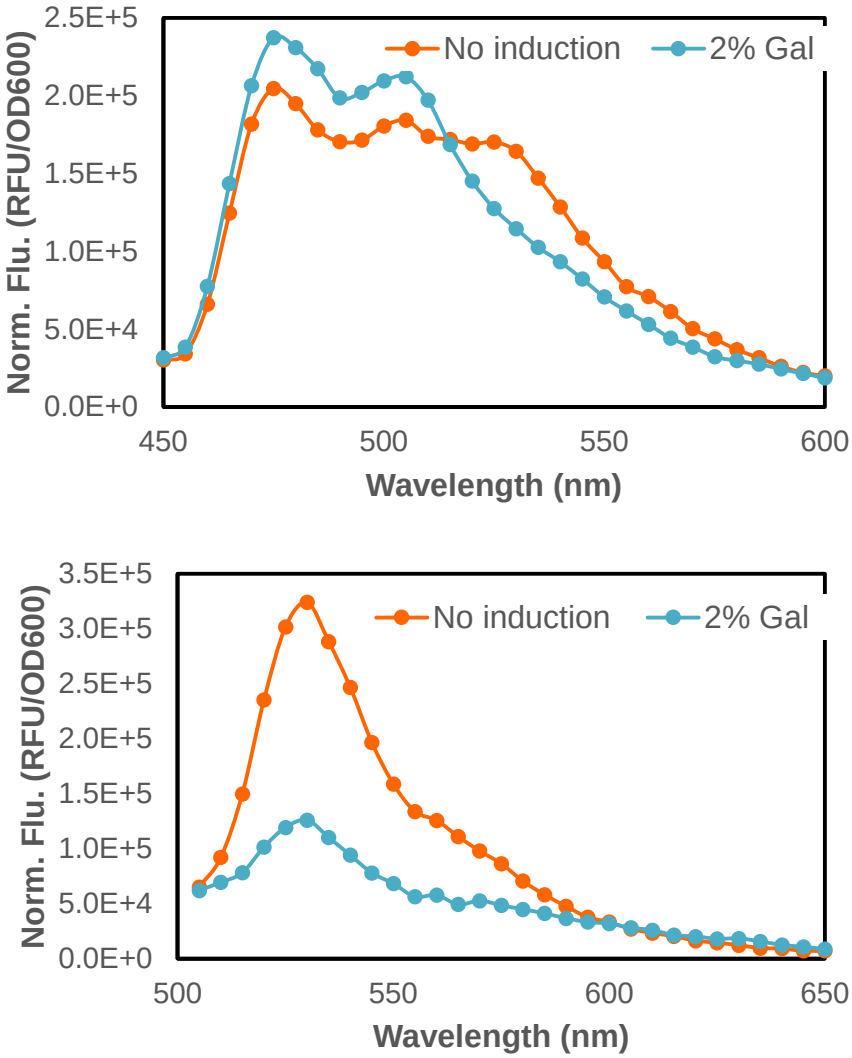


Figure 4.5 Emission spectra change after induced expression of the displacing mRuby2-E_4 fusion at 20hrs after galactose induction. The rise in fluorescence at 470nm with 430nm excitation and the drop in fluorescence at 530nm with 475nm excitation observed in the induced sample relative to the uninduced sample suggested that the proximities of the fluorescent proteins have been changed.

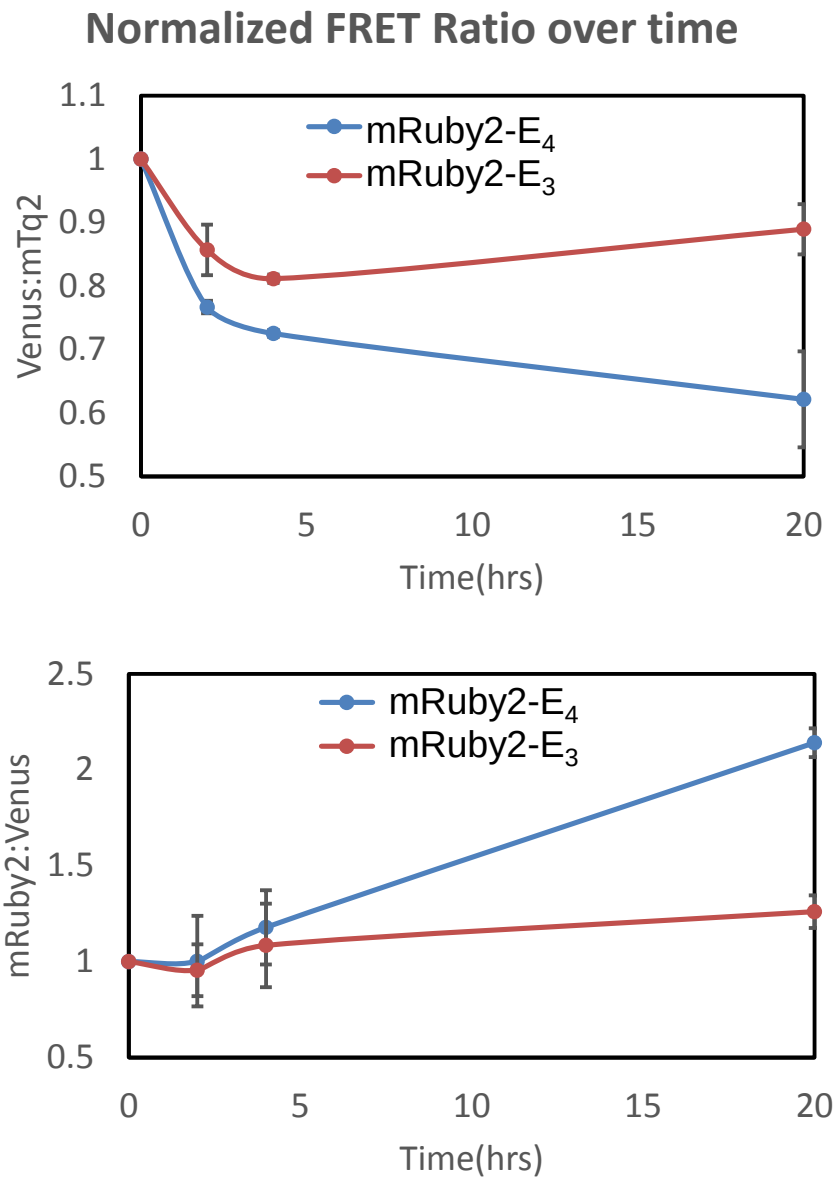


Figure 4.6 Toehold is necessary for efficient coiled-coil strand displacement. Compared to mRuby2-E_4, mRuby2-E3 is unable to displace mTurquoise2-E3 as quickly. Because mRuby2 needs time to mature ($t_{1/2}$~150min) to become fully fluorescent, yellow-to-red FRET ratio rise slowly compared to the decrease in blue-to-yellow FRET ratio. Values are mean ±s.e.m. for n=3 independent experiments.

To expand our repertoire coiled-coil pairs that can be used for further applications, we also tested SYNZIP5/SYNZIP6 coils (Reinke, Grant, and Keating 2010) elongated with P1/P2 coils (Gradišar and Jerala 2011) functioning as toeholds to initiate strand displacement reactions. These synthetic chimera can be used a second set of peptide tags to execute strand displacement that are orthogonal to the EK coils. The same three-color FRET system in yeast was used for testing their efficacy and similar results as EK coils were obtained using the SYNZIP5/SYNZIP6 coiled-coils (Figure 4.7), albeit slightly less in magnitudes; this is likely a result of tighter binding between SYNZIP5-SYNZIP6$(P1)_4$ relative to E3-K4 and differences in expression levels of the fusion proteins. All amino acid sequences of coiled-coil motifs used in this study are listed in Table A.3.1.

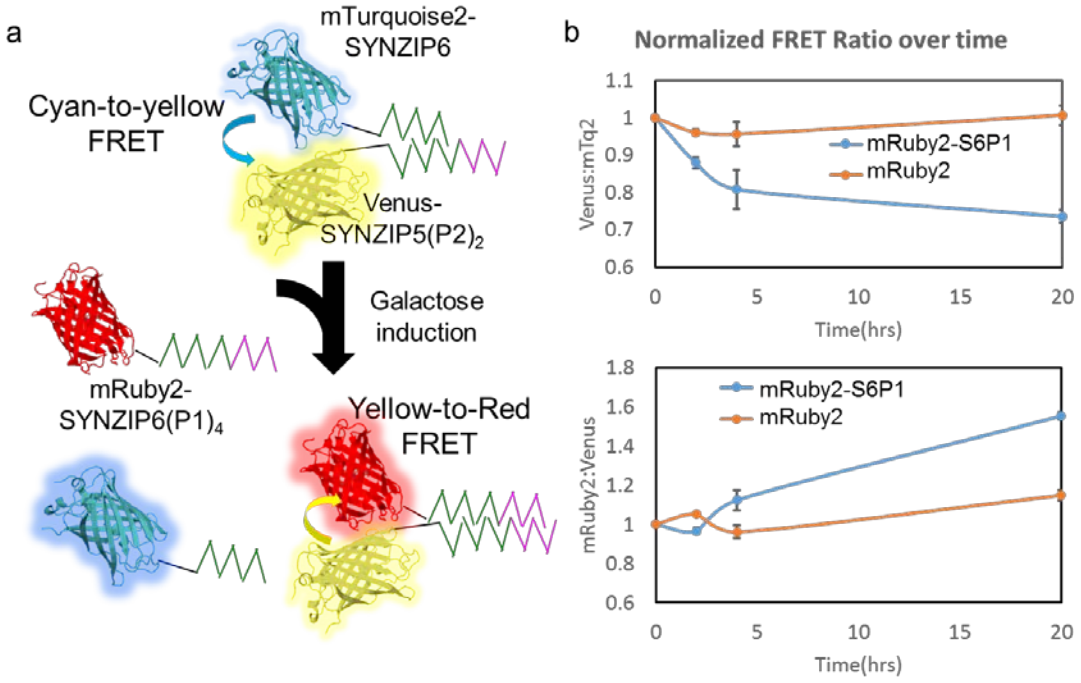


Figure 4.7 Second pair of coiled-coils can also execute strand displacement intracellularly. a) Three-color FRET coiled-coil displacement system using SYNZIP5/6-P2/1 coils in place of EK coils. Analogous to the additional EK coil repeats, P2/1 coils act as the toehold to initiate strand displacement. b) The changes in FRET ratios using these second set of coiled-coils are similar to the EK coils. Values are mean ±s.e.m. for 3 independent experiments.

4.4.3 Directing Violacein Pathway Fluxes by Coiled-coil Displacement

To demonstrate the versatility of coiled-coil strand displacement to dynamically rearrange protein complexes, we created an artificial supramolecular complex consisting of VioC, VioD, and VioE enzymes from the violacein pathway (Lee et al. 2013) (Figure 4.8a) scaffolded together with the EK and SYNZIP5/SYNZIP6 coiled-coils (Figure 4.8b). Scaffolding allow substrate channeling of the intermediates between VioC, VioD, and VioE enzymes, thereby increasing apparent reaction rates and shifting the ratios of the four different products in

accordance with the arrangement of enzymes in the complex. Because the four products exhibits strikingly different colors, this bifurcating pathway provides an easily measurable outcome for us to determine the success of our dynamic scaffolding scheme. Different enzyme complex arrangements are formed using fluorescent protein-coiled-coil fusions from the previously described three-color FRET system to displace one or both of VioC and VioD from VioE (Figure 4.8b). Once displaced, the enzymes are no longer in proximity to the others and thus loses the enhanced reaction rates enabled by substrate channeling, ultimately leading to changes in yields and distributions of the four colored products.

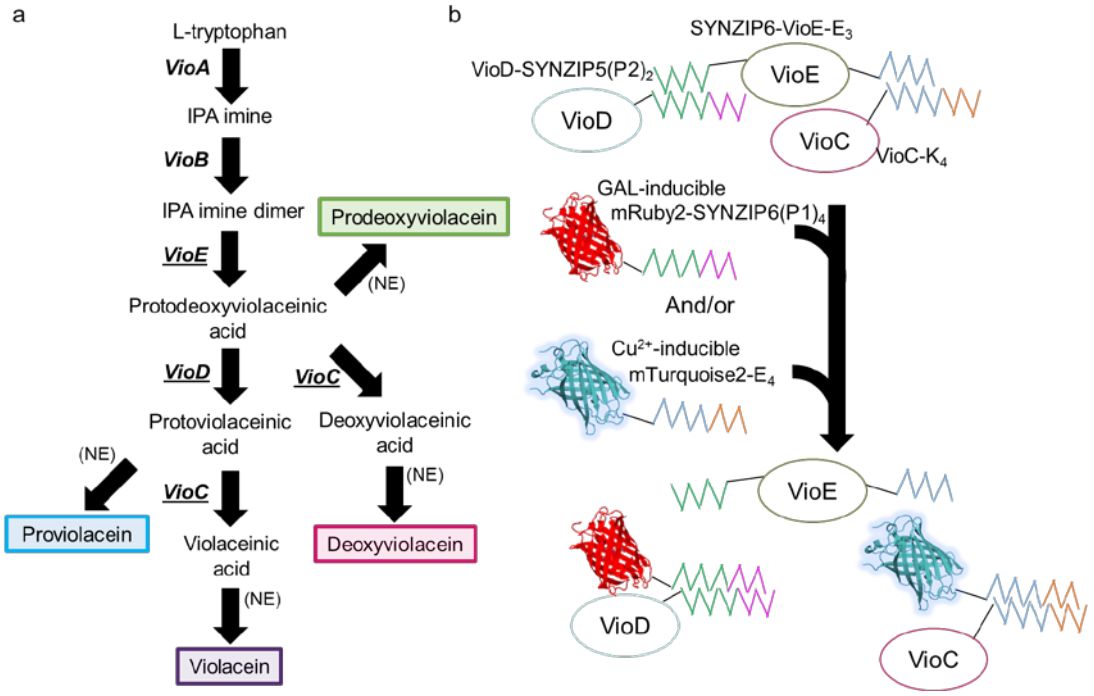


Figure 4.8 Violacein biosynthesis in yeast controlled by coiled-coil-mediated dynamic scaffolding. a) Five enzymatic reactions, along a non-enzymatic reaction, produce the primary purple-colored product, violacein. Each reaction is labelled with the italicized name of the associated enzyme or "(NE)" for non-enzymatic reactions. The synthetic dynamic enzyme complex is made of the underlined enzymes, VioC, VioD, and VioE. b) By displacing one or both of VioC or VioD enzymes from the supramolecular enzyme complex, fluxes through the pathway can be redirected as a result of substrate channeling, leading to changes in yields and compositions of the final product mixtures.

Significant changes in the compositions of the products are visually observable as colorimetric changes (Figure 4.9a). Quantitatively, the shifts in yields and compositions can be observed from chromatograms obtained from HPLC analysis of the extracted products from different induction conditions (Figure 4.9b). Inducing expressions of fluorescent protein-coiled-coil fusions to displace VioC and/or VioD enzyme from the artificial enzyme complex resulted in substantial changes in product

yields, while only modest changes were observed by inducing expressions of non-displacing fluorescent proteins alone. These results indicated that not only does substrate channeling play a significant role for our scaffolding scheme, but the supramolecular enzyme complex is also altered by the expression of the displacement coils. In particular, the >2-fold reduction in yields of violacein when VioD-SYNZIP5$(P2)_2$ was displaced away from SYNZIP6-VioE-E_3 by the induced expression of mRuby2-SYNZIP6$(P1)_4$ suggested that channeling of protodeoxyviolacenic acid to either VioD or VioC from VioE significantly altered the overall productivity and product quality of the heterologous pathway. This effect, in turn, was dynamically modulated by inducing enzyme complex rearrangement using coiled-coil displacement.

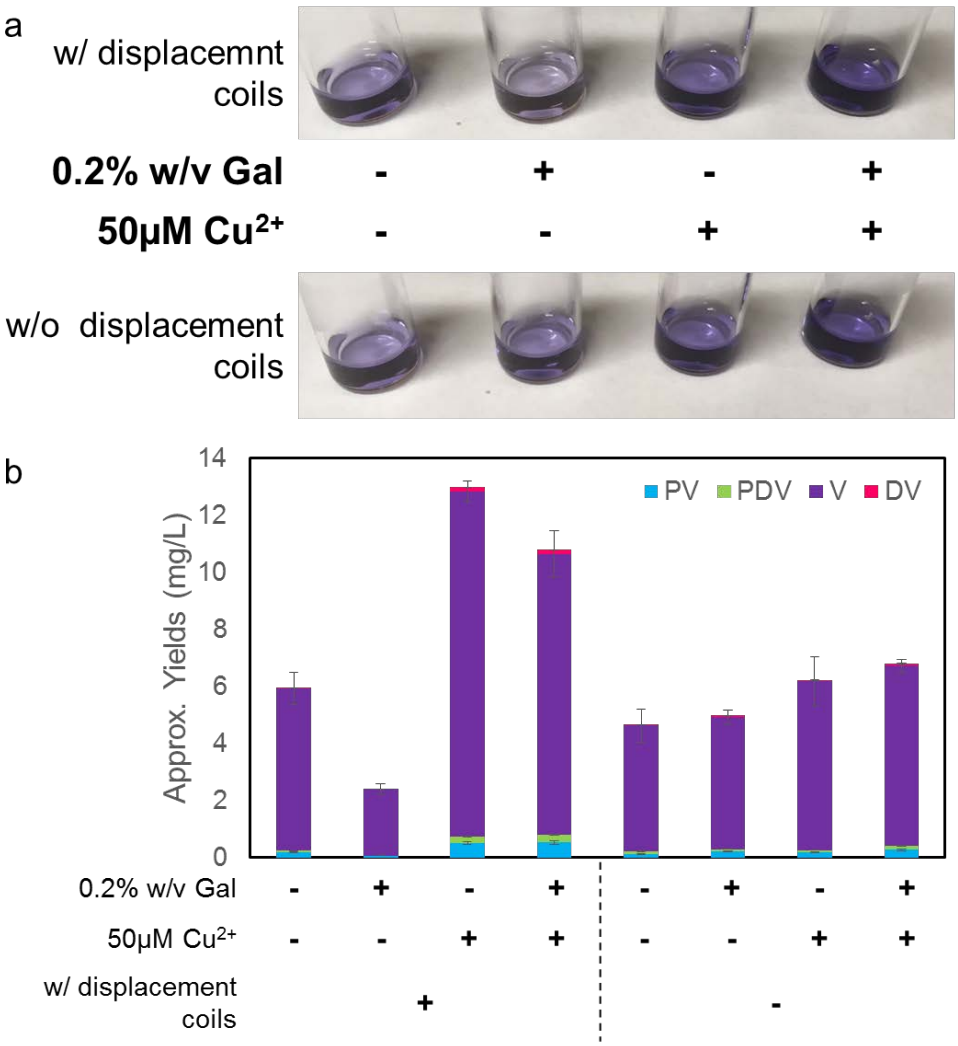


Figure 4.9 Different enzyme complex arrangements using coiled-coils lead to different product yields and compositions from violacein biosynthesis. a) Intensities of the dominant purple color visually changed under different induction conditions, especially apparent when induced fluorescent proteins are fused to displacement coils. b) HPLC analysis of the extracts from different induction conditions. Values are mean ±s.e.m for 3 independent experiments.

4.5 Discussions

4.5.1 Strand Displacement using Heterodimeric Coiled-coils can Control Protein Interaction

The ability of coiled-coil motifs to organize functional protein supramolecular complexes has proven useful in the bottom-up design and construction of biomolecular systems. By exploiting specific interactions between coiled-coils of varying lengths, dynamic systems based on strand displacement can be readily designed (Gröger, Gavins, and Seitz 2017). As this work and previous studies show, this mechanism exhibits fast kinetics, often reaching completion within minutes (Groth et al. 2018). This fast kinetics enable protein rearrangement to occur quickly, often limited only by protein synthesis and turnover rates for intracellular applications.

Similar to the nucleic acid-based counterpart (R. P. Chen et al. 2018), coiled-coil strand displacement reaction is able to not only organize any moieties attached to the peptide motifs, but also change the configurations of the resulting supramolecular complexes dynamically. In contrast to nucleic acid-based schemes, however, coiled-coil motifs can be genetically fused onto any protein of interests. Direct recombinant expression of these peptides bypasses the need to deliver multiple exogenous components and thus expedites applications that require control of protein organization inside the cell. As this work demonstrated, this peptide-based strand displacement strategy can be readily applied in a variety of context. This versatility is further enhanced by the expanding repertoire of orthogonal coiled-coils that enables multiplexing where multiple interactions need to be controlled independently of one another (Reinke, Grant, and Keating 2010; Negron and Keating 2014).

4.5.2 Coiled-coil Strand Displacement can Direct Dynamic Reassembly of Enzyme Complexes

The ability to direct protein interactions by coiled-coils can be used to construct and rearrange artificial enzyme complexes. While artificial enzyme complexes have been shown to enhance the productivity of metabolic pathways in many contexts (Dueber et al. 2009; Moon et al. 2009), a dynamic complex consisting solely of exogenous enzymes that can be assembled and disassembled on demand has yet to be reported. This rearrangement process needs to disrupt complexes formed in one particular state and reconfigure the pathway to another state in response to metabolic cues. Hence, using induced expression of coiled-coil-enzyme fusions controlled by metabolic changes can exploit the displacement mechanism to redirect fluxes through enzyme pathways. Gaining dynamic control over fluxes can potentially alleviate excessive metabolic burdens, especially as heterologous pathways become increasingly complex and intertwined with native components (Nielsen and Keasling 2016; P. Xu et al. 2014).

For biosynthesis that involve multiple enzymatic steps and bifurcating reaction pathways such as the violacein pathway tested in this study, controlling spatial proximities of the relevant enzymes through dynamic reassembly may provide a potentially effective and convenient way to achieve high yields and quality of specific products that is independent of the more conventional transcriptional control of enzyme expression. Because coiled-coil strand displacement is a relatively fast process, enzyme complex reassembly using coiled-coils can potentially be a more responsive, rapid control strategy, as conventional expression control schemes rely not only on switching gene expression, but also protein turnover rates.

4.6 Acknowledgements

This work was supported by National Science Foundation (MCB1817675).

Chapter 5

CONCLUSIONS

5.1 Engineering Biological Control Systems: Opportunities in Synthetic Biology

This work illustrates multiple avenues to manipulate regulatory mechanisms of cellular functions through the creation of synthetic biomolecular constructs. As discussed in Chapter 1, these methods not only reflect the great diversity of ways that nature has evolved to respond to changes, but also represent the ample opportunities for synthetic biologists to design control elements and circuits not found in nature. In that vein, the successes of our designs are largely defined by our own motivations and aspirations. Nonetheless, we hope our studies have highlighted a few lessons of value for the synthetic biology field at large:

First, our parallel schemes of split signal peptide reconstitution using intein-mediated reactions in Chapter 2 demonstrated that even "undesirable" side reactions (i.e. N-terminal cleavage) that diverge from a canonical biochemical reaction pathway (i.e. protein *trans*-splicing) can be adapted to serve a useful function for synthetic applications. While nature may have driven the evolution of a given biochemical pathway to increase fitness for the host organism, this selective pressure does not necessarily apply in our synthetic context. The study in fact underscored the potential benefits of exploiting NTC in place of the canonical PTS for our particular purposes, because of the higher reaction efficiencies upon additional engineering and higher signaling activity of the reconstituted wild-type peptide (Figure 2.9 and Figure 2.10). The discrepancy between native and synthetic objectives thus should not only be

viewed as a source of complications but also an opportunity to be freed from constraints imposed by natural selection, a recurring theme that is frequently encountered in metabolic (Nielsen and Keasling 2016) and protein engineering (Nobeli, Favia, and Thornton 2009) studies.

Second, the interface established between endogenous transcripts and synthetic riboregulators in the form of our thgRNAs in Chapter 3 provided a glimpse of the potential rewards of creating an autonomous control element capable of assessing cellular states. If implemented successfully, such a device may be particularly useful in bioprocessing applications to prevent the biosynthesis of unnecessary RNAs/proteins/metabolites and increase the efficiencies of energy and material usage. For instance, our dynamic control system can promote or repress expression of pathway enzymes based on growth conditions or stress signals reflected in concentrations of cellular sRNA and mRNA, thereby decoupling cell growth and production phases such that resources can be efficiently allocated on one major task at a time (P. Xu et al. 2014; F. Zhang, Carothers, and Keasling 2012). These thgRNAs therefore represent an early attempt toward creating autonomous synthetic regulatory systems. Future progress in this efforts will be indispensable for making biotechnology competitive in industry.

Third, the successful application of entirely synthetic modules presented in Chapter 4 involving coiled-coil motifs that are designed *de novo* and their associated strand displacement validates the efficacy of modeling and computational methods as powerful tools that can aid future designs of synthetic control elements and engineered biological systems in general. As discussed in Chapter 1.4.1, years of theoretical and experimental work on coiled-coils culminated in a uniquely deep understanding of

their sequence-to-structure relationship. Predictive algorithms based on this understanding enabled the precise design of tailor-made nanostructures. In addition, much like the toehold-mediated strand displacement reaction for nucleic acids discussed in Chapter 1.3.2, coiled-coil strand displacement is also thermodynamically driven; hence, it is likely that similar computational models can be made for coiled-coil displacement to predict the rates and efficiencies of any particular reaction, besides the kinetic model fitting studies already underway (Groth et al. 2018). The robustness of the coiled-coil complex reassembly process facilitates the use of the structural motifs in and out of the cellular environment. These synthetic modules can thus be straightforwardly used to construct more complex control networks even inside the crowded cellular environment, given a sufficiently thorough understanding of their characteristics.

Lastly, an obligatory word of caution: despite our growing knowledge of biological components and their interactions, surprising outcomes such as slow kinetics observed for protein *trans*-splicing in Chapter 2 or activation of certain sequestered thgRNAs in Chapter 3 remain commonplace throughout our studies. These hurdles, while unforeseen, are not necessarily unexpected, as we are only beginning to understand some of the complex processes present in living systems. We only overcame these complications by experimentally testing many variants and refining the design in molecular details over several iterations. While time-consuming and, at times, tedious, this iterative process was nonetheless humbling and valuable as an affirmation of a well-known proverb:

What I cannot create, I do not understand. – Richard Feynman

5.2 Ongoing Studies and Future Perspectives on Synthetic Biomolecular Control Circuits

All results presented in this dissertation represent work in progress. It is, therefore, appropriate to elaborate on their prospects, for both the immediate and long term future:

The control of yeast cellular behaviors by extracellularly reconstituted α-factor signaling peptide detailed in Chapter 2 is an important step toward the eventual objective of creating a synthetic extracellular sensing circuit. As suggested in Figure 2.13, the next step is to append the actual sensing domains onto our chimeric sensor units. The sensors' modular design enabled by the promiscuity of intein-mediated reactions allows us to change the specific sensing and signal peptide domains depending on the application. Beyond extensions on the molecular architecture, we attempted to secrete our sensors from *S. cerevisiae* using both native and engineered secretory pathways, with unsatisfactory levels of success (data not shown). A potential workaround for this practical issue may be to screen for secretory peptide signals that maximize the secretion of our sensors (Hashimoto, Koyabu, and Imoto 1998; Kjeldsen, Frost Pettersson, and Hach 1999; Deschuyteneer and Garcia 2010). Alternatively, a host with higher secretory capacities such as *Pichia pastoris* (Ahmad et al. 2014) can be adapted to secrete our extracellular sensors and be co-cultured with engineered *S. cerevisiae*.

The thgRNA design described in Chapter 3 proves to be highly flexible in regards to both sequences and types of RNA that could be used as trigger strands. Despite this valuable trait, some sequence constraints remain in place. Namely, the DNA target for inducible Cas9 protein is coupled by base-pairing complementarity to the sequence of the spacer, which in turn is coupled to the branch migration region in

the hairpin structure (Figure 3.2a). This means that while virtually any trigger sequence can be used to activate Cas9 functions, the DNA target inevitably shares sequence identity with parts of the branch migration region. In other words, Cas9 proteins can only be induced to act on the very specific subset of DNA sequences that are somewhat complementary to the trigger sequence. The utility of this version of our thgRNA is thus limited to only synthetic targets. The next development must free the design from target sequence constraints for this class of synthetic control elements to seamlessly interface with native components. Likely, this will involve either changing the basic schematic of the riboregulators, perhaps to a molecular beacon-like structure (Zheng et al. 2015). Another avenue is to rely on multi-layered strand displacement reactions to produce a trigger strand that differs from the desired transcript signal(s). Recent successes of implementing similar series of displacement reactions give early indications that such a scheme can be adapted for our purposes (Wang et al. 2009; L. Li et al. 2016).

The ability to modulate protein-protein interactions by using coiled-coil strand displacement shown in Chapter 4 presents a unique opportunity to expand the utility of synthetic modules in dynamic control circuits. Beyond transcriptional switching and enzyme complex rearrangement, we anticipate applications in engineering cell signaling (Wei et al. 2012), protein translocation (Guntas et al. 2015; Niopek et al. 2014), and targeted protein degradation (Nishimura et al. 2009). Additionally, coiled-coil strand displacement schemes that parallel those devised for DNA strand displacement such as displacement reaction cascades (D. Y. Zhang and Seelig 2011) would allow us to construct different logic gates that can further diversify their potential applications in controlling protein-protein interactions.

The fundamental questions addressed in Chapters 2, 3, and 4 revolve around the idea that dynamic control at the cellular level is not only beneficial, but also essential for the application of synthetic biological systems in biotechnology and human health. To that end, our results demonstrated the vast potential of biomolecules as valuable and malleable parts for engineering synthetic control circuits in biological systems. The incredible diversity provided by naturally evolved molecules represents an almost infinitely large space from which these engineered components can be generated. Furthermore, as computational power and methods continue to advance, biomolecular modelling algorithms will undoubtedly be improved in their accuracies and speed. As such, we expect the use of synthetic modules like the thgRNAs in Chapter 3 and the coiled-coil motifs in Chapter 4 to take on increasing importance.

The synthetic control elements designed and constructed in this work provide the basic toolkits that can be integrated as high-level decision-making regulatory mechanisms in engineered organisms capable of performing many tasks simultaneously. While the devices and platforms created in this work remain early in the development stage, the core conceptual principles that drove their designs— modularity, connectivity, transferability— should endure as the devices themselves become more refined in synthetic biologists' quest to establish robust, adaptable platforms with autonomous control to advance bioprocessing and human health applications.

REFERENCES

Ahmad, Mudassar, Melanie Hirz, Harald Pichler, and Helmut Schwab. 2014. "Protein Expression in Pichia Pastoris: Recent Achievements and Perspectives for Heterologous Protein Production." *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-014-5732-5.

Altuvia, Shoshy, Aixia Zhang, Liron Argaman, Anita Tiwari, and Gisela Storz. 1998. "The Escherichia Coli OxyS Regulatory RNA Represses FhlA Translation by Blocking Ribosome Binding." *EMBO Journal* 17 (20): 6069–75. doi:10.1093/emboj/17.20.6069.

Amitai, Gil, Brian P Callahan, Matt J Stanger, Georges Belfort, and Marlene Belfort. 2009. "Modulation of Intein Activity by Its Neighboring Extein Substrates." *Proceedings of the National Academy of Sciences of the United States of America* 106 (27): 11005–10. doi:10.1073/pnas.0904366106.

Andersen, Jens Bo, Claus Sternberg, Lars Kongsbak Poulsen, Sara Petersen Bjørn, Michael Givskov, and Søren Molin. 1998. "New Unstable Variants of Green Fluorescent Protein for Studies of Transient Gene Expression in Bacteria." *Applied and Environmental Microbiology* 64 (6): 2240–46.

Anderson, J Christopher, Christopher a Voigt, and Adam P Arkin. 2007. "Environmental Signal Integration by a Modular AND Gate." *Molecular Systems Biology* 3 (133): 133. doi:10.1038/msb4100173.

Armbruster, Blaine N, Xiang Li, Mark H Pausch, Stefan Herlitze, and Bryan L Roth. 2007. "Evolving the Lock to Fit the Key to Create a Family of G Protein-Coupled Receptors Potently Activated by an Inert Ligand." *Proceedings of the National Academy of Sciences of the United States of America* 104 (12): 5163–68. doi:10.1073/pnas.0700293104.

Atkinson, Mariette R., Michael A. Savageau, Jesse T. Myers, and Alexander J. Ninfa. 2003. "Development of Genetic Circuitry Exhibiting Toggle Switch or Oscillatory Behavior in Escherichia Coli." *Cell*. doi:10.1016/S0092-8674(03)00346-5.

Ballister, Edward R., Chanat Aonbangkhen, Alyssa M. Mayo, Michael A. Lampson,

and David M. Chenoweth. 2014. “Localized Light-Induced Protein Dimerization in Living Cells Using a Photocaged Dimerizer.” *Nature Communications*. doi:10.1038/ncomms6475.

Banwell, Eleanor F., Edgardo S. Abelardo, Dave J. Adams, Martin A. Birchall, Adam Corrigan, Athene M. Donald, Mark Kirkland, Louise C. Serpell, Michael F. Butler, and Derek N. Woolfson. 2009. “Rational Design and Application of Responsive α-Helical Peptide Hydrogels.” *Nature Materials*. doi:10.1038/nmat2479.

Bashor, Caleb J, Noah C Helman, Shude Yan, and Wendell A Lim. 2008. “Using Engineered Scaffold Interactions to Reshape MAP Kinase Pathway Signaling Dynamics.” *Science (New York, N.Y.)* 319 (5869): 1539–43. doi:10.1126/science.1151153.

Broadwater, D. W.Bo, and Harold D. Kim. 2016. “The Effect of Basepair Mismatch on DNA Strand Displacement.” *Biophysical Journal*. doi:10.1016/j.bpj.2016.02.027.

Burkhard, Peter, Sergei Ivaninskii, and Ariel Lustig. 2002. “Improving Coiled-Coil Stability by Optimizing Ionic Interactions.” *Journal of Molecular Biology*. doi:10.1016/S0022-2836(02)00114-6.

Callura, J. M., C. R. Cantor, and J. J. Collins. 2012. “Genetic Switchboard for Synthetic Biology Applications.” *Proceedings of the National Academy of Sciences* 109 (15): 5850–55. doi:10.1073/pnas.1203808109.

Cameron, D. Ewen, Caleb J. Bashor, and James J. Collins. 2014. “A Brief History of Synthetic Biology.” *Nature Reviews Microbiology*. doi:10.1038/nrmicro3239.

Chappell, James, Kyle E. Watters, Melissa K. Takahashi, and Julius B. Lucks. 2015. “A Renaissance in RNA Synthetic Biology: New Mechanisms, Applications and Tools for the Future.” *Current Opinion in Chemical Biology*. doi:10.1016/j.cbpa.2015.05.018.

Chen, Ming-Tang, and Ron Weiss. 2005. “Artificial Cell-Cell Communication in Yeast Saccharomyces Cerevisiae Using Signaling Elements from Arabidopsis Thaliana.” *Nature Biotechnology* 23 (12): 1551–55. doi:10.1038/nbt1162.

Chen, Rebecca, Qi Chen, Heejae Kim, Ka-Hei Siu, Qing Sun, Shen-Long Tsai, and Wilfred Chen. 2014. “Biomolecular Scaffolds for Enhanced Signaling and Catalytic Efficiency.” *Current Opinion in Biotechnology* 28C (August). Elsevier Ltd: 59–68. doi:10.1016/j.copbio.2013.11.007.

Chen, Rebecca P., Daniel Blackstock, Qing Sun, and Wilfred Chen. 2018. “Dynamic Protein Assembly by Programmable DNA Strand Displacement.” *Nature Chemistry*. doi:10.1038/s41557-018-0016-9.

Collins, Cynthia H., Jared R. Leadbetter, and Frances H. Arnold. 2006. “Dual Selection Enhances the Signaling Specificity of a Variant of the Quorum-Sensing Transcriptional Activator LuxR.” *Nature Biotechnology* 24 (6): 708–12. doi:10.1111/j.1467-9655.2006.00359_31.x.

Conklin, BR, EC Hsiao, and Sylvie Claeysen. 2008. “Engineering GPCR Signaling Pathways with RASSLs.” *Nature …* 5 (8): 673–78. doi:10.1038/nmeth.1232.Engineering.

Crick, F. H.C. 1952. “Is α-Keratin a Coiled Coil?” *Nature*. doi:10.1038/170882b0.

Daringer, Nichole M., Rachel M. Dudek, Kelly A. Schwarz, and Joshua N. Leonard. 2014. “Modular Extracellular Sensor Architecture for Engineering Mammalian Cell-Based Devices.” *ACS Synthetic Biology*. doi:10.1021/sb400128g.

Davis, Kevin M., Vikram Pattanayak, David B. Thompson, John A. Zuris, and David R. Liu. 2015. “Small Molecule-Triggered Cas9 Protein with Improved Genome-Editing Specificity.” *Nature Chemical Biology* 11 (5): 316–18. doi:10.1038/nchembio.1793.

Deschuyteneer, G, and S Garcia. 2010. “Intein-Mediated Cyclization of Randomized Peptides in the Periplasm of Escherichia Coli and Their Extracellular Secretion.” *ACS Chemical …* 5 (7): 691–700.

Dixon, Andrew S., Marie K. Schwinn, Mary P. Hall, Kris Zimmerman, Paul Otto, Thomas H. Lubben, Braeden L. Butler, et al. 2016. “NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells.” *ACS Chemical Biology*. doi:10.1021/acschembio.5b00753.

Doyle, Donald F., David J. Mangelsdorf, and David R. Corey. 2000. “Modifying Ligand Specificity of Gene Regulatory Proteins.” *Current Opinion in Chemical Biology*. doi:10.1016/S1367-5931(99)00052-6.

Dueber, John E., Brian J. Yeh, Kayam Chak, and Wendell A. Lim. 2003. “Reprogramming Control of an Allosteric Signaling Switch through Modular Recombination.” *Science*. doi:10.1126/science.1085945.

Dueber, John E, Gabriel C Wu, G Reza Malmirchegini, Tae Seok Moon, Christopher J Petzold, Adeeti V Ullal, Kristala L J Prather, and Jay D Keasling. 2009. “Synthetic Protein Scaffolds Provide Modular Control over Metabolic Flux.”

Nature Biotechnology 27 (8): 753–59. doi:10.1038/nbt.1557.

Elowitz, Michael B., and Stanislas Leibier. 2000. “A Synthetic Oscillatory Network of Transcriptional Regulators.” *Nature*. doi:10.1038/35002125.

Friedland, Ari E., Timothy K. Lu, Xiao Wang, David Shi, George Church, and James J. Collins. 2009. “Synthetic Gene Networks That Count.” *Science*. doi:10.1126/science.1172005.

Fu, Yanfang, Jeffry D. Sander, Deepak Reyon, Vincent M. Cascio, and J. Keith Joung. 2014. “Improving CRISPR-Cas Nuclease Specificity Using Truncated Guide RNAs.” *Nature Biotechnology*. doi:10.1038/nbt.2808.

Gagnon, James A., Eivind Valen, Summer B. Thyme, Peng Huang, Laila Ahkmetova, Andrea Pauli, Tessa G. Montague, Steven Zimmerman, Constance Richter, and Alexander F. Schier. 2014. “Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs.” *PLoS ONE* 9 (5). doi:10.1371/journal.pone.0098186.

Gardner, Timothy S., Charles R. Cantor, and James J. Collins. 2000. “Construction of a Genetic Toggle Switch in Escherichia Coli.” *Nature*. doi:10.1038/35002131.

Gong, Shanzhong, Helen Hong Yu, Kenneth A. Johnson, and David W. Taylor. 2018. “DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity.” *Cell Reports*. doi:10.1016/j.celrep.2017.12.041.

Gradišar, Helena, and Roman Jerala. 2011. “De Novo Design of Orthogonal Peptide Pairs Forming Parallel Coiled-Coil Heterodimers.” *Journal of Peptide Science* 17 (2): 100–106. doi:10.1002/psc.1331.

Green, Alexander A., Jongmin Kim, Duo Ma, Pamela A. Silver, James J. Collins, and Peng Yin. 2017. “Complex Cellular Logic Computation Using Ribocomputing Devices.” *Nature* 548 (7665): 117–21. doi:10.1038/nature23271.

Green, Alexander A., Pamela A. Silver, James J. Collins, and Peng Yin. 2014. “Toehold Switches: De-Novo-Designed Regulators of Gene Expression.” *Cell* 159 (4): 925–39. doi:10.1016/j.cell.2014.10.002.

Gröger, Katharina, Georgina Gavins, and Oliver Seitz. 2017. “Strand Displacement in Coiled-Coil Structures: Controlled Induction and Reversal of Proximity.” *Angewandte Chemie - International Edition*. doi:10.1002/anie.201705339.

Groth, Mike C., W. Mathis Rink, Nils F. Meyer, and Franziska Thomas. 2018. “Kinetic Studies on Strand Displacement in: De Novo Designed Parallel

Heterodimeric Coiled Coils.” *Chemical Science* 9 (18): 4308–16. doi:10.1039/c7sc05342h.

Groves, Benjamin, Yuan-Jyue Chen, Chiara Zurla, Sergii Pochekailov, Jonathan L. Kirschman, Philip J. Santangelo, and Georg Seelig. 2015. “Computing in Mammalian Cells with Nucleic Acid Strand Exchange.” *Nature Nanotechnology* 11 (3): 287–94. doi:10.1038/nnano.2015.278.

Guntas, Gurkan, Ryan A. Hallett, Seth P. Zimmerman, Tishan Williams, Hayretin Yumerefendi, James E. Bear, and Brian Kuhlman. 2015. “Engineering an Improved Light-Induced Dimer (ILID) for Controlling the Localization and Activity of Signaling Proteins.” *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1417910112.

Harbury, Pehr B., Joseph J. Plecs, Bruce Tidor, Tom Alber, and Peter S. Kim. 1998. “High-Resolution Protein Design with Backbone Freedom.” *Science*. doi:10.1126/science.282.5393.1462.

Hashimoto, Y, N Koyabu, and T Imoto. 1998. “Effects of Signal Sequences on the Secretion of Hen Lysozyme by Yeast: Construction of Four Secretion Cassette Vectors.” *Protein Engineering* 11 (2): 75–77.

Ho, S. N., S. R. Biggar, D. M. Spencer, S. L. Schreiber, and G. R. Crabtree. 1996. “Dimeric Ligands Define a Role for Transcriptional Activation Domains in Reinitiation.” *Nature*. doi:10.1038/382822a0.

Hodges, R. S., A. K. Saund, P. C.S. Chong, S. A. St-Pierre, and R. E. Reid. 1981. “Synthetic Model for Two-Stranded α-Helical Coiled-Coils. Design, Synthesis, and Characterization of an 86-Residue Analog of Tropomyosin.” *Journal of Biological Chemistry*.

Horvath, P., and R. Barrangou. 2010. “CRISPR/Cas, the Immune System of Bacteria and Archaea.” *Science* 327 (5962): 167–70. doi:10.1126/science.1179555.

Hsu, Patrick D., David A. Scott, Joshua A. Weinstein, F. Ann Ran, Silvana Konermann, Vineeta Agarwala, Yinqing Li, et al. 2013. “DNA Targeting Specificity of RNA-Guided Cas9 Nucleases.” *Nature Biotechnology*. doi:10.1038/nbt.2647.

Ishii, Jun, Shizuka Matsumura, Sakurako Kimura, Kenji Tatematsu, Shun’ichi Kuroda, Hideki Fukuda, and Akihiko Kondo. 2006. “Quantitative and Dynamic Analyses of G Protein-Coupled Receptor Signaling in Yeast Using Fus1, Enhanced Green Fluorescence Protein (EGFP), and His3 Fusion Protein.” *Biotechnology Progress* 22 (4): 954–60. doi:10.1021/bp0601387.

Iwai, H, a Lingel, and a Pluckthun. 2001. “Cyclic Green Fluorescent Protein Produced in Vivo Using an Artificially Split PI-PfuI Intein from Pyrococcus Furiosus.” *The Journal of Biological Chemistry* 276 (19): 16548–54. doi:10.1074/jbc.M011639200.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012a. “A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science*. doi:10.1126/science.1225829.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. 2012b. “A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptice Bacterial Immunity.” *Science (New York, N.Y.)* 337 (August): 816–22. doi:10.1126/science.1225829.

Khalil, Ahmad S., and James J. Collins. 2010. “Synthetic Biology: Applications Come of Age.” *Nature Reviews Genetics*. doi:10.1038/nrg2775.

Kjeldsen, Thomas, Annette Frost Pettersson, and Morten Hach. 1999. “The Role of Leaders in Intracellular Transport and Secretion of the Insulin Precursor in the Yeast Saccharomyces Cerevisiae.” *Journal of Biotechnology* 75: 195–208. doi:10.1016/S0168-1656(99)00159-5.

Koh, John T. 2002. “Engineering Selectivity and Discrimination into Ligand-Receptor Interfaces.” *Chemistry and Biology*. doi:10.1016/S1074-5521(02)00087-X.

Kohn, Wayne D., and Robert S. Hodges. 1998. “De Novo Design of α-Helical Coiled Coils and Bundles: Models for the Development of Protein-Design Principles.” *Trends in Biotechnology*. doi:10.1016/S0167-7799(98)01212-8.

Lappano, Rosamaria, and Marcello Maggiolini. 2011. “G Protein-Coupled Receptors: Novel Targets for Drug Discovery in Cancer.” *Nature Reviews. Drug Discovery* 10 (1): 47–60. doi:10.1038/nrd3320.

Lee, Michael E., Anil Aswani, Audrey S. Han, Claire J. Tomlin, and John E. Dueber. 2013. “Expression-Level Optimization of a Multi-Enzyme Pathway in the Absence of a High-Throughput Assay.” *Nucleic Acids Research*. doi:10.1093/nar/gkt809.

Lee, Michael E., William C. DeLoache, Bernardo Cervantes, and John E. Dueber. 2015. “A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly.” *ACS Synthetic Biology*. doi:10.1021/sb500366v.

Li, Juan, Wunchang Sun, Bing Wang, Xiao Xiao, and Xiang-Qin Liu. 2008. “Protein

Trans-Splicing as a Means for Viral Vector-Mediated in Vivo Gene Therapy." *Human Gene Therapy* 19 (9): 958–64. doi:10.1089/hum.2008.009.

Li, Lu, Jie Feng, Haiyun Liu, Qingling Li, Lili Tong, and Bo Tang. 2016. "Two-Color Imaging of MicroRNA with Enzyme-Free Signal Amplification via Hybridization Chain Reactions in Living Cells." *Chemical Science*. doi:10.1039/c5sc03909f.

Lim, Wendell A. 2010. "Designing Customized Cell Signalling Circuits." *Nature Reviews. Molecular Cell Biology* 11 (6). Nature Publishing Group: 393–403. doi:10.1038/nrm2904.

Liu, Yuchen, Yonghao Zhan, Zhicong Chen, Anbang He, Jianfa Li, Hanwei Wu, Li Liu, et al. 2016. "Directing Cellular Information Flow via CRISPR Signal Conductors." *Nature Methods* 13 (11): 938–44. doi:10.1038/nmeth.3994.

Lockless, Steve W, and Tom W Muir. 2009. "Traceless Protein Splicing Utilizing Evolved Split Inteins." *Proceedings of the National Academy of Sciences of the United States of America* 106 (27): 10999–4. doi:10.1073/pnas.0902964106.

Lundstrom, Kenneth, Renaud Wagner, Christoph Reinhart, Aline Desmyter, Nadia Cherouati, Thierry Magnin, Gabrielle Zeder-Lutz, et al. 2006. "Structural Genomics on Membrane Proteins: Comparison of More than 100 GPCRs in 3 Expression Systems." *Journal of Structural and Functional Genomics* 7 (2): 77–91. doi:10.1007/s10969-006-9011-2.

Lutz, Rolf, and Hermann Bujard. 1997. "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* 25 (6): 1203–10. doi:10.1093/nar/25.6.1203.

MacPhee, Cait E., and Derek N. Woolfson. 2004. "Engineered and Designed Peptide-Based Fibrous Biomaterials." In *Current Opinion in Solid State and Materials Science*. doi:10.1016/j.cossms.2004.01.010.

Manfredi, J P, C Klein, J J Herrero, D R Byrd, J Trueheart, W T Wiesler, D M Fowlkes, and J R Broach. 1996. "Yeast Alpha Mating Factor Structure-Activity Relationship Derived from Genetically Selected Peptide Agonists and Antagonists of Ste2p." *Molecular and Cellular Biology* 16 (9): 4700–4709.

Martini, Laura, Adam J. Meyer, Jared W. Ellefson, John N. Milligan, Michele Forlin, Andrew D. Ellington, and Sheref S. Mansy. 2015. "In Vitro Selection for Small-Molecule-Triggered Strand Displacement and Riboswitch Activity." *ACS Synthetic Biology* 4 (10): 1144–50. doi:10.1021/acssynbio.5b00054.

Massé, Eric, Freddy E. Escorcia, and Susan Gottesman. 2003. “Coupled Degradation of a Small Regulatory RNA and Its MRNA Targets in Escherichia Coli.” *Genes and Development* 17 (19): 2374–83. doi:10.1101/gad.1127103.

Massé, Eric, and Susan Gottesman. 2002. “A Small RNA Regulates the Expression of Genes Involved in Iron Metabolism in Escherichia Coli.” *Proceedings of the National Academy of Sciences* 99 (7): 4620–25. doi:10.1073/pnas.032066599.

Maung, Nyan Win, and Christina D. Smolke. 2008. “Higher-Order Cellular Information Processing with Synthetic RNA Devices.” *Science*. doi:10.1126/science.1160311.

Mekler, Vladimir, Leonid Minakhin, Ekaterina Semenova, Konstantin Kuznedelov, and Konstantin Severinov. 2016. “Kinetics of the CRISPR-Cas9 Effector Complex Assembly and the Role of 3′-Terminal Segment of Guide RNA.” *Nucleic Acids Research* 44 (6): 2837–45. doi:10.1093/nar/gkw138.

Meyenberg, Karsten, Antonina S. Lygina, Geert Van Den Bogaart, Reinhard Jahn, and Ulf Diederichsen. 2011. “SNARE Derived Peptide Mimic Inducing Membrane Fusion.” *Chemical Communications*. doi:10.1039/c1cc12879e.

Mitchell, Amir, Gal H G.H. Romano, Bella Groisman, Avihu Yona, Erez Dekel, Martin Kupiec, Orna Dahan, and Yitzhak Pilpel. 2009. “Adaptive Prediction of Environmental Changes by Microorganisms.” *Nature*. doi:10.1038/nature08112.

Miyamoto, Takafumi, Robert DeRose, Allison Suarez, Tasuku Ueno, Melinda Chen, Tai Ping Sun, Michael J. Wolfgang, Chandrani Mukherjee, David J. Meyers, and Takanari Inoue. 2012. “Rapid and Orthogonal Logic Gating with a Gibberellin-Induced Dimerization System.” *Nature Chemical Biology*. doi:10.1038/nchembio.922.

Moon, Tae Seok, Sang Hwal Yoon, Amanda M. Lanza, Joseph D. Roy-Mayhew, and Kristala L. Jones Prather. 2009. “Production of Glucaric Acid from a Synthetic Pathway in Recombinant Escherichia Coli.” *Applied and Environmental Microbiology* 75 (3): 589–95. doi:10.1128/AEM.00973-08.

Mootz, Henning D, Elyse S Blum, Amy B Tyszkiewicz, and Tom W Muir. 2003. “Conditional Protein Splicing: A New Tool to Control Protein Structure and Function in Vitro and in Vivo.” *Journal of the American Chemical Society* 125 (35): 10561–69. doi:10.1021/ja0362813.

Muir, Tom W. 2003. “Semisynthesis of Proteins by Expressed Protein Ligation.” *Annual Review of Biochemistry* 72 (January): 249–89. doi:10.1146/annurev.biochem.72.121801.161900.

Mutalik, Vivek K, Lei Qi, Joao C Guimaraes, Julius B Lucks, and Adam P Arkin. 2012. “Rationally Designed Families of Orthogonal RNA Regulators of Translation.” *Nature Chemical Biology* 8 (5): 447–54. doi:10.1038/nchembio.919.

Naider, Fred, and Jeffrey M Becker. 2004. “The Alpha-Factor Mating Pheromone of Saccharomyces Cerevisiae: A Model for Studying the Interaction of Peptide Hormones and G Protein-Coupled Receptors.” *Peptides* 25 (9): 1441–63. doi:10.1016/j.peptides.2003.11.028.

Naider, Fred, and JM Becker. 1986. “Structure-Activity Relationships of the Yeast α-Factor.” *Critical Reviews in Biochemistry and …* 21 (3): 225–48.

Negron, Christopher, and Amy E. Keating. 2014. “A Set of Computationally Designed Orthogonal Antiparallel Homodimers That Expands the Synthetic Coiled-Coil Toolkit.” *Journal of the American Chemical Society*. doi:10.1021/ja507847t.

Nielsen, Jens, and Jay D. Keasling. 2016. “Engineering Cellular Metabolism.” *Cell*. doi:10.1016/j.cell.2016.02.004.

Nihongaki, Yuta, Fuun Kawano, Takahiro Nakajima, and Moritoshi Sato. 2015. “Photoactivatable CRISPR-Cas9 for Optogenetic Genome Editing.” *Nature Biotechnology*. doi:10.1016/j.cej.2015.10.060.

Niopek, Dominik, Dirk Benzinger, Julia Roensch, Thomas Draebing, Pierre Wehler, Roland Eils, and Barbara Di Ventura. 2014. “Engineering Light-Inducible Nuclear Localization Signals for Precise Spatiotemporal Control of Protein Dynamics in Living Cells.” *Nature Communications*. doi:10.1038/ncomms5404.

Nishimasu, Hiroshi, F. Ann Ran, Patrick D. Hsu, Silvana Konermann, Soraya I. Shehata, Naoshi Dohmae, Ryuichiro Ishitani, Feng Zhang, and Osamu Nureki. 2014. “Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA.” *Cell* 156 (5): 935–49. doi:10.1016/j.cell.2014.02.001.

Nishimura, Kohei, Tatsuo Fukagawa, Haruhiko Takisawa, Tatsuo Kakimoto, and Masato Kanemaki. 2009. “An Auxin-Based Degron System for the Rapid Depletion of Proteins in Nonplant Cells.” *Nature Methods*. doi:10.1038/nmeth.1401.

Nobeli, Irene, Angelo D. Favia, and Janet M. Thornton. 2009. “Protein Promiscuity and Its Implications for Biotechnology.” *Nature Biotechnology*. doi:10.1038/nbt1519.

O’Shea, Erin K., Kevin J. Lumb, and Peter S. Kim. 1993. “Peptide ‘Velcro’: Design of

a Heterodimeric Coiled Coil." *Current Biology*. doi:10.1016/0960-9822(93)90063-T.

Park, Sang-Hyun, Ali Zarrinpar, and Wendell A Lim. 2003. "Rewiring MAP Kinase Pathways Using Alternative Scaffold Assembly Mechanisms." *Science (New York, N.Y.)* 299 (5609): 1061–64. doi:10.1126/science.1076979.

Park, Won Min, Mostafa Bedewy, Karl K. Berggren, and Amy E. Keating. 2017. "Modular Assembly of a Protein Nanotriangle Using Orthogonally Interacting Coiled Coils." *Scientific Reports*. doi:10.1038/s41598-017-10918-6.

Pattanayak, Vikram, Steven Lin, John P. Guilinger, Enbo Ma, Jennifer A. Doudna, and David R. Liu. 2013. "High-Throughput Profiling of off-Target DNA Cleavage Reveals RNA-Programmed Cas9 Nuclease Specificity." *Nature Biotechnology*. doi:10.1038/nbt.2673.

Pauling, Linus, Robert B. Corey, and H. R. Branson. 1951. "The Structure of Proteins; Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain." *Proceedings of the National Academy of Sciences of the United States of America*. doi:10.1073/pnas.37.4.205.

Perler, F B, E O Davis, G E Dean, F S Gimble, W E Jack, N Neff, C J Noren, J Thorner, and M Belfort. 1994. "Protein Splicing Elements: Inteins and Exteins--a Definition of Terms and Recommended Nomenclature." *Nucleic Acids Research* 22 (7): 1125–27.

Qi, Lei S., Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin, and Wendell A. Lim. 2013. "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression." *Cell* 152 (5): 1173–83. doi:10.1016/j.cell.2013.02.022.

Ramirez, Miguel, Dongli Guan, Victor Ugaz, and Zhilei Chen. 2013. "Intein-Triggered Artificial Protein Hydrogels That Support the Immobilization of Bioactive Proteins." *Journal of the American Chemical Society* 135 (14): 5290–93. doi:10.1021/ja401075s.

Ran, F. Ann, Patrick D. Hsu, Chie Yu Lin, Jonathan S. Gootenberg, Silvana Konermann, Alexandro E. Trevino, David A. Scott, et al. 2013. "Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity." *Cell*. doi:10.1016/j.cell.2013.08.021.

Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang. 2013. "Genome Engineering Using the CRISPR-Cas9 System." *Nature Protocols* 8 (11): 2281–2308. doi:10.1038/nprot.2013.143.

Reece, Jane B., Lisa A. Urry, Michael L. Cain, Steven A. Wasserman, Peter V. Minorsky, and Robert B. Jackson. 1989. *Biology*. *Campbell Biology*. doi:10.1109/TPAMI.2012.97.

Reinke, Aaron W., Robert A. Grant, and Amy E. Keating. 2010. "A Synthetic Coiled-Coil Interactome Provides Heterospecific Modules for Molecular Engineering." *Journal of the American Chemical Society*. doi:10.1021/ja907617a.

Rozengurt, Enrique, James Sinnett-Smith, and Krisztina Kisfalvi. 2010. "Crosstalk between Insulin/Insulin-like Growth Factor-1 Receptors and G Protein-Coupled Receptor Signaling Systems: A Novel Target for the Antidiabetic Drug Metformin in Pancreatic Cancer." *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research* 16 (9): 2505–11. doi:10.1158/1078-0432.CCR-09-2229.

Sapranauskas, Rimantas, Giedrius Gasiunas, Christophe Fremaux, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys. 2011. "The Streptococcus Thermophilus CRISPR/Cas System Provides Immunity in Escherichia Coli." *Nucleic Acids Research*. doi:10.1093/nar/gkr606.

Sarramegna, V., I. Muller, A. Milon, and F. Talmont. 2006. "Recombinant G Protein-Coupled Receptors from Expression to Renaturation: A Challenge towards Structure." *Cellular and Molecular Life Sciences*. doi:10.1007/s00018-005-5557-6.

Serganov, Alexander, and Evgeny Nudler. 2013. "A Decade of Riboswitches." *Cell*. doi:10.1016/j.cell.2012.12.024.

Shah, Neel H, Geoffrey P Dann, Miquel Vila-Perelló, Zhihua Liu, and Tom W Muir. 2012. "Ultrafast Protein Splicing Is Common among Cyanobacterial Split Inteins: Implications for Protein Engineering." *Journal of the American Chemical Society* 134 (28): 11338–41. doi:10.1021/ja303226x.

Shah, Neel H, Ertan Eryilmaz, David Cowburn, and Tom W Muir. 2013. "Extein Residues Play an Intimate Role in the Rate-Limiting Step of Protein Trans-Splicing." *Journal of the American Chemical Society* 135 (15): 5839–47. doi:10.1021/ja401015p.

Siu, Ka-Hei, and Wilfred Chen. 2018. "Riboregulated Toehold-Gated GRNA for Programmable CRISPR–Cas9 Function." *Nature Chemical Biology*. doi:10.1038/s41589-018-0186-1.

Siu, Ka Hei, and Wilfred Chen. 2017. "Control of the Yeast Mating Pathway by Reconstitution of Functional α-Factor Using Split Intein-Catalyzed Reactions."

ACS Synthetic Biology. doi:10.1021/acssynbio.7b00078.

Spencer, David M., Thomas J. Wandless, Stuart L. Schreiber, and Gerald R. Crabtree. 1993. "Controlling Signal Transduction with Synthetic Ligands." *Science*. doi:10.1126/science.7694365.

Stricker, Jesse, Scott Cookson, Matthew R. Bennett, William H. Mather, Lev S. Tsimring, and Jeff Hasty. 2008. "A Fast, Robust and Tunable Synthetic Gene Oscillator." *Nature*. doi:10.1038/nature07389.

Summer, Heike, René Grämer, and Peter Dröge. 2009. "Denaturing Urea Polyacrylamide Gel Electrophoresis (Urea PAGE)." *Journal of Visualized Experiments*, no. 32. doi:10.3791/1485.

Tang, Shuang Yan, and Patrick C. Cirino. 2011. "Design and Application of a Mevalonate-Responsive Regulatory Protein." *Angewandte Chemie - International Edition* 50 (5): 1084–86. doi:10.1002/anie.201006083.

Tang, Weixin, Johnny H. Hu, and David R. Liu. 2017. "Aptazyme-Embedded Guide RNAs Enable Ligand-Responsive Genome Editing and Transcriptional Activation." *Nature Communications* 8. doi:10.1038/ncomms15939.

Tavassoli, Ali, and Stephen J Benkovic. 2007. "Split-Intein Mediated Circular Ligation Used in the Synthesis of Cyclic Peptide Libraries in E. Coli." *Nature Protocols* 2 (5): 1126–33. doi:10.1038/nprot.2007.152.

Taylor, Noah D, Alexander S Garruss, Rocco Moretti, Sum Chan, Mark A Arbing, Duilio Cascio, Jameson K Rogers, et al. 2015. "Engineering an Allosteric Transcription Factor to Respond to New Ligands." *Nature Methods* 13 (2): 177–83. doi:10.1038/nmeth.3696.

Tsai, Shengdar Q., Nicolas Wyvekens, Cyd Khayter, Jennifer A. Foden, Vishal Thapar, Deepak Reyon, Mathew J. Goodwin, Martin J. Aryee, and J. Keith Joung. 2014. "Dimeric CRISPR RNA-Guided FokI Nucleases for Highly Specific Genome Editing." *Nature Biotechnology*. doi:10.1038/nbt.2908.

Venkatakrishnan, a J, Xavier Deupi, Guillaume Lebon, Christopher G Tate, Gebhard F Schertler, and M Madan Babu. 2013. "Molecular Signatures of G-Protein-Coupled Receptors." *Nature* 494 (7436): 185–94. doi:10.1038/nature11896.

Volkmann, Gerrit, Wenchang Sun, and Xiang Qin Liu. 2009. "Controllable Protein Cleavages through Intein Fragment Complementation." *Protein Science*. doi:10.1002/pro.249.

Wang, Kemin, Zhiwen Tang, Chaoyong James Yang, Youngmi Kim, Xiaohong Fang, Wei Li, Yanrong Wu, et al. 2009. “Molecular Engineering of DNA: Molecular Beacons.” *Angewandte Chemie - International Edition*. doi:10.1002/anie.200800370.

Wei, Ping, Wilson W Wong, Jason S Park, Ethan E Corcoran, Sergio G Peisajovich, James J Onuffer, Arthur Weiss, and Wendell A Lim. 2012. “Bacterial Virulence Proteins as Tools to Rewire Kinase Pathways in Yeast and Immune Cells.” *Nature* 488 (7411). Nature Publishing Group: 384–88. doi:10.1038/nature11259.

Wiedenheft, Blake, Samuel H. Sternberg, and Jennifer A. Doudna. 2012. “RNA-Guided Genetic Silencing Systems in Bacteria and Archaea.” *Nature* 482 (7385): 331–38. doi:10.1038/nature10886.

Wittrup, K D, and V Benig. 1994. “Optimization of Amino Acid Supplements for Heterologous Protein Secretion In.” *Biotechnol Techniques* 8 (3): 161–66.

Wood, D W, W Wu, G Belfort, V Derbyshire, and M Belfort. 1999. “A Genetic System Yields Self-Cleaving Inteins for Bioseparations.” *Nature Biotechnology* 17 (9): 889–92. doi:10.1038/12879.

Xu, P., L. Li, F. Zhang, G. Stephanopoulos, and M. Koffas. 2014. “Improving Fatty Acids Production by Engineering Dynamic Pathway Regulation and Metabolic Control.” *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1406401111.

Xu, Peng, Amerin Vansiri, Namita Bhan, and Mattheos A G Koffas. 2012. “EPathBrick: A Synthetic Biology Platform for Engineering Metabolic Pathways in E. Coli.” *ACS Synthetic Biology* 1 (7): 256–66. doi:10.1021/sb300016b.

Youk, Hyun, and Wendell A Lim. 2014. “Secreting and Sensing the Same Molecule Allows Cells to Achieve Versatile Social Behaviors.” *Science (New York, N.Y.)* 343 (6171): 1242782. doi:10.1126/science.1242782.

Zadeh, Joseph N., Conrad D. Steenberg, Justin S. Bois, Brian R. Wolfe, Marshall B. Pierce, Asif R. Khan, Robert M. Dirks, and Niles A. Pierce. 2011. “NUPACK: Analysis and Design of Nucleic Acid Systems.” *Journal of Computational Chemistry* 32 (1): 170–73. doi:10.1002/jcc.21596.

Zarrinpar, Ali, Sang Hyun Park, and Wendell A. Lim. 2003. “Optimization of Specificity in a Cellular Protein Interaction Network by Negative Selection.” *Nature*. doi:10.1038/nature02178.

Zettler, Joachim, Vivien Schütz, and Henning D Mootz. 2009. “The Naturally Split

Npu DnaE Intein Exhibits an Extraordinarily High Rate in the Protein Trans-Splicing Reaction." *FEBS Letters* 583 (5). Federation of European Biochemical Societies: 909–14. doi:10.1016/j.febslet.2009.02.003.

Zhang, David Yu, and Georg Seelig. 2011. "Dynamic DNA Nanotechnology Using Strand-Displacement Reactions." *Nature Chemistry*. doi:10.1038/nchem.957.

Zhang, David Yu, and Erik Winfree. 2009. "Control of DNA Strand Displacement Kinetics Using Toehold Exchange." *Journal of the American Chemical Society* 131 (47): 17303–14. doi:10.1021/ja906987s.

Zhang, Fuzhong, James M. Carothers, and Jay D. Keasling. 2012. "Design of a Dynamic Sensor-Regulator System for Production of Chemicals and Fuels Derived from Fatty Acids." *Nature Biotechnology*. doi:10.1038/nbt.2149.

Zheng, Jing, Ronghua Yang, Muling Shi, Cuichen Wu, Xiaohong Fang, Yinhui Li, Jishan Li, and Weihong Tan. 2015. "Rationally Designed Molecular Beacons for Bioanalytical and Biomedical Applications." *Chemical Society Reviews*. doi:10.1039/c5cs00020c.

Zhou, Xin X, Xinzhi Zou, Hokyung K. Chung, Yuchen Gao, Yanxia Liu, Lei S. Qi, and Michael Z Lin. 2017. "A Single-Chain Photoswitchable CRISPR-Cas9 Architecture for Light-Inducible Gene Editing and Transcription." *ACS Chemical Biology*. doi:10.1021/acschembio.7b00603.

Appendix A

SUPPLEMENTARY TABLES AND INFORMATION

A.1 Strains and Plasmids used in Chapter 2

Table A.1 Strains and plasmids used in Chapter 2

Strains		
Escherichia coli		
NEB®5α	fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	
BL21-Gold (DE3)	ompT hsdS(rB- mB-) dcm+ Tetr gal λ(DE3) endA The	
Saccharomyces cerevisiae		
BY4741 bar1 Δ	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 bar1::KanR	
CB009	MATa far1Δ mfa2::pFUS1-GFP bar1::NatR his3 trp1 leu2 ura3	
Plasmids*		
Short Name	Name	Features/Notes
	pI-CutA-L-*Npu*N	Template of Npu^{N} intein fragment
	pI-*Npu*C-S-CutA	Template of Npu^{C} intein fragment
	pET-24a(+)	Empty vector for protein expression
N1	pET24a-α^{N1}-*Npu*N-GB1-his	Expression of **N1**
N1B	pET24a-MBP-α^{N1}-*Npu*N-GB1-his	Expression of **N1B**
C1	pET24a-his-GB1-*Npu*C- α^{C1}	Expression of **C1**
N2	pET24a-α^{N2}-*Npu*N-GB1-his	Expression of **N2**
N2B	pET24a-MBP-α^{N2}-*Npu*N-GB1-his	Expression of **N2B**
C2	pET24a-his-GB1-*Npu*C- α^{C2}	Expression of **C2**
N3	pET24a-α-*Npu*N-GB1-his	Expression of **N3**
N3B	pET24a-MBP-α-*Npu*N-GB1-his	Expression of **N3B**
C3	pET24a-his-GB1-*Npu*C*-his	Expression of **C3**

A.2 Sequences of thgRNA, Trigger Strands, and DNA Targets used in Chapter 3

Table A.2 Sequences of thgRNA, trigger strands, and DNA targets used in Chapter 3

Name	Sequence			
	thgRNA[i]			
	a (Toehold)	b (Branch migration)	Loop	c (Spacer)
A1	AGUUUGAUUACAUUG	UCGUUUAGUUUAGUG	AUACAUAAACACUGCA	GGUAUCACAUGACUAAACGA
A2	AGUUUGAUUACAUUG	UCGUUUAGU	CAUAAACACUGCA	GGUAUCACAUGACUAAACGA
A3	AGUUUGAUUACAUUG	UCGUUUAGU	CUAAAACACUGCA	GGUAUCACAUGACUAAACGA
A4	UUGAUUACAUUG	UCGUUUAGU	CUAAAACACUGCA	GGUAUCACAUGACUAAACGA
A5	UUGAUUACAUUG	UCGUUUAGU	UAGA	GGUAUCACAUGACUAAACGA
B	GAAGAAUAAGCAUGA	AGCUUGGGUCACCGG	AUACAAACUUGCUACA	GGUAUCCGAACACCCAAGCU
C	AAAAGUUCACGU	UGGCUUUAGUUAUUC	AUACAUAAACACUGCA	GGUAUGAACUGCUAAAGCCA
D	AAGUAAUACUGGAAGCAAU	GUGAGCAAUGUCGUG	UAUUAAACACUCA	GGAUACACUCGAUUGCUCAC
E1	ACGGUAAUAAAUAAA	GUUAAUGAUGAUAGC	AUACAAACUUGCUACA	GGUAUGCUUCGAUCAUUAAC

E2	**UCAGAAAUGAAUGAC**	**GGUAAUAAAUAAAGU**	**UAACAAACUUGCUACA**	**GGUUAACUCAGUUUAUUACC**
F	**GAUUUUACACCAAUA**	**CUCAGUCACACAUGA**	**UUACAAACUUGCUACA**	**GGUAAUCACUAGUGACUGAG**
G1	**UGAUGGCCAUGUUAU**	**CCUCCUCGCCCUUGC**	**UUACAAACUUGCUACA**	**GGUAAGCACCUGCGAGGAGG**
G2	**GCCCUCGCCCUCGCC**	**CUCGAUCUCGAACUC**	**AUACAAACUUGCUACA**	**GGUAUGAGCAGGAGAUCGAG**
G3	**CCUCGGGGUACAUCC**	**GCUCGGAGGAGGCCU**	**AUACAAACUUGCUACA**	**GGUAUAGGGGACCUCCGAGC**
Trigger strands[ii]				
	5’	**b* (Branch migration)**	**a* (Toehold)**	**3’**
A*	**GGAUCCACUGACUAUUCUGUGCAAUAGUCAGUAAA**	**CACUAAACUAAACGA**	**CAAUGUAAUCAAACU**	**AGAUCUAGCUAGCAUAACCCCUUGGGGCCUCUAAACGGGUCUUGAGGGGUUUUUU**
B*	**GGAUCCACUGACUAUUCUGUGCAAUAGUCAGUAAA**	**CCGGUGACCCAAGCU**	**UCAUGCUUAUUCUUC**	**AGAUCUAGCUAGCAUAACCCCUUGGGGCCUCUAAACGGGUCUUGAGGGGUUUUUU**
C*	**GGAUCCACUGACUAUUCUGUGCAAUAGUCAGUAAA**	**GAAUAACUAAAG**	**CCAACGUGAACUUUU**	**AGAUCUAGCUAGCAUAACCCCUUGGGGCCUCUAAACGGGUCUUGAGGGGUUUUUU**
D*	**GGAUCCACUGACUAUUCUGUGCAAUAGUCAGUAAA**	**CACGACAUUGCUCAC**	**AUUGCUUCCAGUAUUACUU**	**AGAUCUAGCUAGCAUAACCCCUUGGGGCCUCUAAACGGGUCUUGAGGGGUUUUUU**

OxyS	**GAAACGGAGCGGCACCUCUUUUAACCCUUGAAGUCACUGCCCGUUUCGAGAGUUUCUCAACUC**	**GAAUAACUAAAG**	**CCAACGUGAACUUUU**	**GCGGAUCUCCAGGAUCCGCUG**
RyhB	**GCGAUCAGGAAGACCCUCGCGGAGAACCUGAAAG**	**CACGACAUUGCUCAC**	**AUUGCUUCCAGUAUUACUU**	**AGCCAGCCGGGUGCUGGCUUUU**
MicF (E1)	**(N/A)**	**GCUAUCAUCAUUAAC**	**UUUAUUUAUUACCGU**	**CAUUCAUUUCUGAAUGUCUGUUUACCCCUAUUUCAACCGGAUGCCUCGCAUUCGGUUUUUUUU**
MicF (E2)	**GCUAUCAUCAUUA**	**ACUUUAUUUAUUACC**	**GUCAUUCAUUUCUGA**	**AUGUCUGUUUACCCCUAUUUCAACCGGAUGCCUCGCAUUCGGUUUUUUUU**
SgrS	**GAUGAAGCAAGGGGGUGCCCCAUGCGUCAGUUUUAUCAGCACUAUUUUACCGCGACAGCGAAGUUGUGCUGGUUGCGUUGGUUAAGCGUCCCACAACGAUUAACCAUGCUUGAAGGACUGAUGCAGUGGGAUGACCGCAAUUCUGAAAGUUGACUUGCCUGCA**	**UCAUGUGUGACUGAG**	**UAUUGGUGUAAAAUC**	**ACCCGCCAGCAGAUUAUACCUGCUGGUUUUUUUU**
mCherry (G1)[iii]	**(39)-**	**GCAAGGGCGAGGAGG**	**AUAACAUGGCCAUCA**	**-(722)**

mCherry (G2)[iii]	**(122)-**	**GAGUUC GAGAUC GAG**	**GGCGAGGG CGAGGGC**	**-(639)**
mCherry (G3)[iii]	**(477)-**	**AGGCCU CCUCCG AGC**	**GGAUGUAC CCCGAGG**	**-(284)**
Target (Target strand)[iv]				
A	**TCGTTTAGTCATGTGATACC**			
B	**AGCTTGGGTGTTCGGATACC**			
C	**TGGCTTTAGCAGTTCATACC**			
D	**GTGAGCAATCGAGTGTACC**			
E1	**GGTATGCTTCGATCATTAAC**			
E2	**GGTAATAAACTGAGTTAACC**			
F	**CTCAGTCACTAGTGATTACC**			
G1	**CCTCCTCGCAGGTGCTTACC**			
G2	**CTCGATCTCCTGCTCATACC**			
G3	**GGTATAGGGGACCTCCGAGC**			

[i] 5'-end of thgRNA start with GGN (not shown) prior to the toehold region (a). The canonical sequence(Jinek et al. 2012b) of crRNA repeats (not shown) begins at the 3'-end of the spacer region.

[ii] Artificial trigger strands contain 5' hairpins that enhance stability of the resulting transcript inside the cell(Green et al. 2014) as well as 3' T7 terminator sequence. The endogenous sRNA sequences are shown, with the intended hybridization regions used in toehold-mediated strand displacement highlighted. Only the sequences used for thgRNA activations were shown for mCherry mRNA, with the number of 5' and 3' bases outside of the regions of interest shown in parenthesis.

[iii] Numbers in ( ) indicates the number of bases upstream or downstream of the intended hybridization regions.

[iv] Shown sequences are the intended targets that will hybridize with the spacer region of the gRNA; these are placed on the coding strand downstream of the transcriptional start site and upstream of the RBS of the reporter gene.

A.3 Strains and Plasmids used in Chapter 3

Table A.3 Strains and plasmids used in Chapter 3

Strains			
NEB®5α	fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17		
BL21-Gold (DE3)	ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]$_{K-12}$(λS)		
Plasmids*			
#	Name	Marker	Features/Notes
Reporter^{+}			
CR143	pUC19_J23101p-TA-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23101p promoter and target **A** in a high copy number plasmid; linearized with NotI restriction enzyme prior to use in *in vitro* cleavage assay (Figure 3.5)
CR295	pETJ15_J23115p-TA-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **A** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR277	pETJ15_J23115p-TB-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **B** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR285	pETJ15_J23115p-TC-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **C** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR380	pETJ15_J23115p-TD-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **D** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover

CR479	pETJ15_J23115p-TE1-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **E1** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR480	pETJ15_J23115p-TE2-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **E2** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR482	pETJ15_J23115p-TF-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **F** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR487	pETJ15_J23115p-TG1-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **G1** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR488	pETJ15_J23115p-TG2-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **G2** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR489	pETJ15_J23115p-TG3-Nluc-LAA-T7t	Amp	Nanoluciferase reporter downstream of J23115p promoter and target **G3** in a modified ePathBrick vector; ssrA degradation tag is fused at the C-terminus of Nluc to promote turnover
CR307	pETJ15_J23115p-TB-BFP-T7t	Amp	Blue fluorescent protein reporter downstream of J23115p promoter and target **B** in a modified ePathBrick vector
CR396	pETJ15_J23115p-TD-mCherry-T7t	Amp	mCherry fluorescent protein reporter downstream of J23115p promoter and target **D** in a modified ePathBrick vector
Cas9 protein / gRNA			
62374 (Addgene)	pET-Cas9-6xHis	Amp	Cas9-6xHis expression vector from David Liu via Addgene; transformed in BL21-Gold (DE3) cells for expression and purification

64108 (Addgene)	pHAGE-TO-dCas9-3XmCherry	Amp	dCas9 fused with C-term. 3xmCherry; template of dCas9 gene for all subsequent constructs
CR097	pCDF_ProC-dCas9_J23100p-sgRNA A	Spect	dCas9 protein under constitutive ProC promoter and sgRNA A under constitutive J23100p promoter in a medium copy number plasmid (Figure 3.10)
CR098	pCDF_ProC-dCas9_J23100p-thgRNA A1	Spect	dCas9 protein under constitutive ProC promoter and thgRNA **A1** with toehold under constitutive J23100p promoter in a medium copy number plasmid (Figure 3.10)
CR099	pCDF_ProC-dCas9_J23100p-ΔthgRNA A1	Spect	dCas9 protein under constitutive ProC promoter and thgRNA **A1** without toehold under constitutive J23100p promoter in a medium copy number plasmid (Figure 3.10)
CR294	pETJ15_J23115p-T7RBS-dCas9-T7t	Amp	dCas9 protein under constitutive J23115p promoter in modified ePathBrick plasmid
CR346	pETJ15_J23115p-T7RBS-Cas9-T7t	Amp	Nuclease active Cas9 protein under constitutive J23115p promoter in modified ePathBrick plasmid
CR278/ CR280/ CR282/ CR378/ CR483/ CR484/ CR486/ CR493/ CR494/ CR495	pETJ15_J23115p-thgRNA A1/B/C/D/E1/E2/F/G1/G2/G3-T7t	Amp	thgRNA **A1/B/C/D/E1/E2/F/G1/G2/G3** under constitutive J23115p promoter in modified ePathBrick plasmid

CR279/ CR281/ CR283/ CR379	pETJ15_J23115p-sgRNA A/B/C/D -T7t	Amp	Unmodified sgRNA **A/B/C/D** under constitutive J23115p promoter in modified ePathBrick plasmid
CR296/ CR290/ CR292/ CR397/ CR490/ CR491/ CR492/ CR500/ CR501/ CR502	pETJ15_J23115p-thgRNA A1/B/C/D/E1/E2/F/G1/G2/G3-T7t J23115p-TA/B/C/D/E1/E2/F/G1/G2/G3-Nluc-LAA-T7t	Amp	thgRNA **A1/B/C/D/E1/E2/F/G1/G2/G3** and Nluc reporter under constitutive J23115p promoter in modified ePathBrick plasmid
CR297/ CR291/ CR293/ CR398	pETJ15_J23115p-sgRNA A/B/C/D-T7t J23115p-TA/B/C/D-Nluc-LAA-T7t	Amp	sgRNA **A/B/C/D** and Nluc reporter under constitutive J23115p promoter in modified ePathBrick plasmid
CR304/ CR299/ CR301/ CR397/ CR496/ CR497/ CR498/ CR504/ CR505/ CR506	pETJ15_J23115p- thgRNA A1/B/C/D/E1/E2/F/G1/G2/G3-T7t J23115p-TA/B/C/D/E1/E2/F/G1/G2/G3-Nluc-LAA-T7t J23115p-dCas9-T7t	Amp	thgRNA **A1/B/C/D/E1/E2/F/G1/G2/G3**, Nluc reporter, and dCas9 under constitutive J23115p promoter in modified ePathBrick plasmid (Figures 3.10-19)
CR305/ CR300/ CR302/ CR398	pETJ15_J23115p-sgRNA A/B/C/D-T7t J23115p-TA/B/C/D-Nluc-LAA-T7t J23115p-dCas9-T7t	Amp	sgRNA **A/B/C/D**, Nluc reporter, and dCas9 under constitutive J23115p promoter in modified ePathBrick plasmid (Figure 3.11)
CR303/ CR356/ CR358/	pETJ15_J23115p-TA/B/C/D/E1/E2/F/G1/G2/G3-Nluc-LAA-T7t	Amp	No gRNA control plasmids; Nluc reporter and dCas9 under constitutive

CR399	J23115p-dCas9-T7t		J23115p promoter in modified ePathBrick plasmid (Figures 3.11-19)
CR413	pETJ15_J23115p- thgRNA D-T7t J23115p- TD-mCherry-T7t J23115p- thgRNA B-T7t J23115p- TB-BFP-LAA-T7t J23115p- thgRNA C-T7t J23115p- TC-Nluc-LAA-T7t J23115p-dCas9-T7t	Amp	thgRNA **B**,**C**,& **D**, BFP, mCherry, & Nluc reporters, and dCas9 under constitutive J23115p promoter in modified ePathBrick plasmid (Figure 3.14)
CR412	pETJ15_J23115p- TD-mCherry-T7t J23115p- TB-BFP-LAA-T7t J23115p- TC-Nluc-LAA-T7t J23115p-dCas9-T7t	Amp	No gRNA; BFP, mCherry, & Nluc reporters, and dCas9 under constitutive J23115p promoter in modified ePathBrick plasmid (Figure 3.14)
	pETJ15_J23115p- thgRNA B-T7t J23115p-TB-Nluc-LAA-T7t J23115p-Cas9-T7t	Amp	thgRNA **B**, Nluc reporter with target **B**, and Cas9 under constitutive J23115p promoter in modified ePathBrick plasmid; nuclease active Cas9 used to induce plasmid loss (Figure 3.20)
	pETJ15_J23115p- thgRNA C-T7t J23115p-TB-Nluc-LAA-T7t J23115p-Cas9-T7t	Amp	thgRNA **C**, mismatched with Nluc reporter with target **B**, and Cas9 under constitutive J23115p promoter in modified ePathBrick plasmid; nuclease active Cas9 used to induce plasmid loss (Figure 3.20)
	pETJ15_J23115p- thgRNA B-T7t J23115p-TC-Nluc-LAA-T7t J23115p-Cas9-T7t	Amp	thgRNA **B**, mismatched with Nluc reporter with target **C**, and Cas9 under constitutive J23115p promoter in modified ePathBrick plasmid; nuclease active Cas9 used to induce plasmid loss (Figure 3.20)
Trigger			
CR109	pMB1_pLlacO1-5'hp-A*-T7t	Amp	Trigger **A*** under pLlacO1 promoter with 5' hairpin and 3' T7 terminator in medium copy number plasmid
CR138	pMB1_pLlacO1-5'hp-A*-T7t_ J23101p-TA-Nluc-LAA-T7t	Amp	Trigger **A*** under pLlacO1 promoter with 5' hairpin and 3' T7 terminator and Nanoluciferase reporter downstream of J23101p promoter and target **A** in medium copy number plasmid (Figure 3.10)

CR124	pSC101_pLlacO1-5'hp-A*-T7t	Kan	Trigger **A*** under pLlacO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.11)
CR364	pSC101_pLtetO1-5'hp-B*-T7t	Kan	Trigger **B*** under pLtetO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figures 3.11-13, 3.17, & 3.20)
CR365	pSC101_pLtetO1-5'hp-C*-T7t	Kan	Trigger **C*** under pLtetO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.11)
CR366	pSC101_pLtetO1-OxyS-T7t	Kan	Full-length OxyS sRNA under p*Ltet*O-1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.17)
CR400	pSC101_pLtetO1-5'hp-D*-T7t	Kan	Trigger **D*** under pLtetO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.11)
CR402	pSC101_pLlacO1-5'hp-B*-T7t pLtetO1-5'hp-C*-T7t	Kan	Trigger **B*** under pLlacO1 promoter and Trigger **C*** under pLtetO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.14)
CR403	pSC101_pLlacO1-5'hp-B*-T7t pLtetO1-5'hp-D*-T7t	Kan	Trigger **B*** under pLlacO1 promoter and Trigger **D*** under pLtetO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.14)
CR418	pSC101_pLlacO1-5'hp-C*-T7t pLtetO1-5'hp-D*-T7t	Kan	Trigger **C*** under pLlacO1 promoter and Trigger **D*** under pLtetO1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.14)
CR468	pSC101_pLtetO1-MicF-T7t	Kan	Full-length MicF sRNA under p*Ltet*O-1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.18)
CR469	pSC101_pLtetO1-SgrS-T7t	Kan	Full-length SgrS sRNA under p*Ltet*O-1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.18)
CR470	pSC101_pLtetO1-mCherry-T7t	Kan	Full-length mCherry mRNA under p p*Ltet*O-1 promoter with 5' hairpin and 3' T7 terminator in low copy number plasmid (Figure 3.18)

Miscellaneous			
49795 (Addgene)	pETM6	Amp	ePathBrick parental vector
CR275	pETJ15_J23115p-Empty-T7t	Amp	Modified ePathBrick vector by subsitution of constitutive J23115p promoter in place of T7-lac promoter. No meaningful transcriptional products; parental vector of constructs for induced transcriptional repression assays
CR276	pETJ15_J23115p-T7RBS-Empty-T7t	Amp	RBS inserted after J23115p promoter in CR275; parental vector of Cas9 proteins for induced transcriptional repression assays

*All plasmids listed are constructed for this study unless otherwise noted.

[+]Sequence and characteristics of ssrA tag fused to Nluc is described in Andersen JB et al.(Andersen et al. 1998)

A.4 Amino Acids Sequences of Coiled-coil Motifs used in Chapter 4

Table A.4 Amino acids sequences of coiled-coil motifs used in Chapter 4

Name	**Amino acid sequence***
$E_{3/4}$	EIAALEK EIAALEK EIAALEK / EIAALEK
K_4	KIAALKE KIAALKE KIAALKE KIAALKE
SYNZIP5	N TVKELKN YIQELEE RNAELKN LKEHLKF AKAELEF ELAA
SYNZIP6(15-54)[^]	K ENAYLEN IVARLEN DNANLEK DIANLEK DIANLER DVAR
$(P1)_4$	EIQALEE ENAQLEQ ENAALEE EIAQLEY
$(P2)_2^+$	KIAQLKE KNAALKE

[*]Amino acids group by heptad repeats patterns denoted *gabcdef* in common coiled-coil conventions

[^]Only the portion that binds to SYNZIP5 is used (Reinke, Grant, and Keating 2010)

[+]The first two heptad repeats of P2 coil was used as toehold for initiating displacement with P1 (Gradišar and Jerala 2011; Gröger, Gavins, and Seitz 2017)

A.5 Strains and Plasmids used in Chapter 4

Table A.5 Strains and plasmids used in Chapter 4

Strains			
Escherichia coli			
NEB®5α	fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17		
NEB®Express	fhuA2 [lon] ompT gal sulA11 R(mcr-73::miniTn10--TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 Δ(mcrC-mrr)114::IS10		
Saccharomyces cerevisiae			
BY4741	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0		
ySK001	MATa his3Δ1 met15Δ0 ura3Δ0 leu2::VioABCDE		
Plasmids*			
#	Name	Marker(s)	Features/Notes
***In vitro* assay reporters**			
	pMAL_ptac-MBP-LgBit-E_3	Amp	LgBit fragment of split Nanoluciferase reporter fused to N-terminal MBP and C-terminal E_3 coil (Figure 4.3)
	pMAL_ptac-MBP-SmBit-K_4	Amp	SmBit fragment of split Nanoluciferase reporter fused to N-terminal MBP and C-terminal K_4 coil (Figure 4.3)
	pMAL_ptac-MBP-Flag-E_4	Amp	Flag-tag fused to N-terminal MBP and C-terminal E_4 coil (Figure 4.3)
Added MoClo-YTK parts			
CC004	T4a: E_3	Cam	Type 4a containing E_3 coil
CC005	T4a: K_4	Cam	Type 4a containing K_4 coil
CC008	T4a: E_4	Cam	Type 4a containing E_4 coil
CC042	T4a: SYNZIP5$(P2)_2$	Cam	Type 4a containing SYNZIP5$(P2)_2$ fused coil
CC043	T4a: SYNZIP6$(P1)_4$	Cam	Type 4a containing SYNZIP6$(P1)_4$ fused coil
CC088	T4a: SYNZIP6	Cam	Type 4a containing SYNZIP6 coil
CC057	T3a: SYNZIP6	Cam	Type 3a containing SYNZIP6 coil
CC044	T3: VioA	Cam	Type 3 containing VioA enzyme
CC045	T3: VioB	Cam	Type 3 containing VioB enzyme
CC055	T3: VioC	Cam	Type 3 containing VioC enzyme
CC056	T3:VioD	Cam	Type 3 containing VioD enzyme

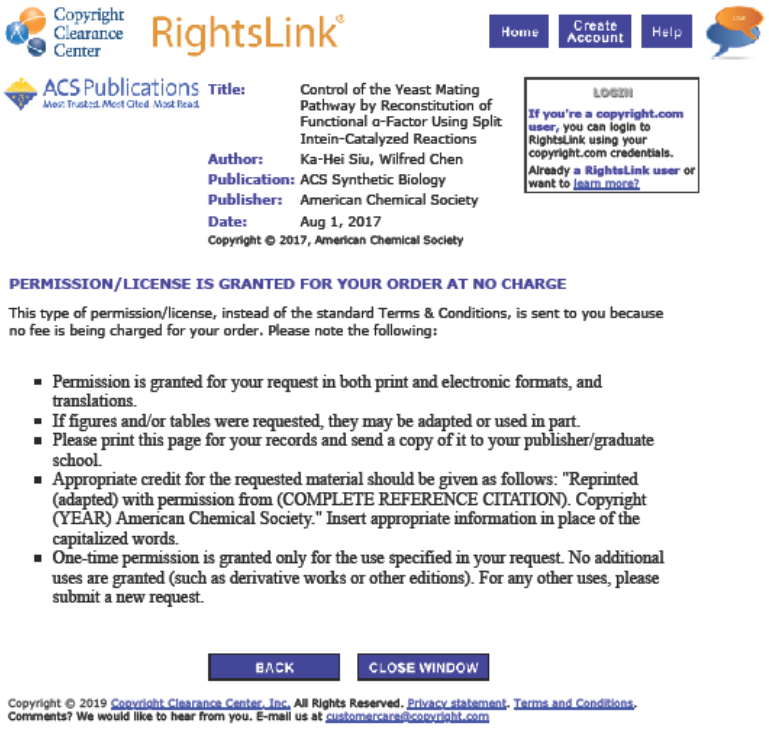
CC060	T3b: VioE	Cam	Type 3b containing VioE enzyme
Assembled MoClo-YTK gene cassessets			
CC006	pHHF2-mTurqoise2-E_3	Amp	mTurqoise2 reporter fused to E_3 coil
CC011	pPGK1-Venus-K_4	Amp	Venus reporter fused to K_4 coil
CC009	pGAL1-mRuby2-E_4	Amp	Galactose inducible mRuby2 reporter fused to E_4 coil
CC046	pGAL1-mRuby2-E_3	Amp	Galactose inducible mRuby2 reporter fused to E_3 coil
CC096	pHHF2-mTurqoise2-SYNZIP6	Amp	mTurqoise2 reporter fused to SYNZIP6 coil
CC097	pPGK1-Venus-SYNZIP5$(P2)_2$	Amp	Venus reporter fused to SYNZIP5$(P2)_2$ coil
CC054	pGAL1-mRuby2-SYNZIP6$(P1)_4$	Amp	Galactose inducible mRuby2 reporter fused to SYNZIP6$(P1)_4$ coil
CC081	pGAL1-mRuby2	Amp	Galactose inducible mRuby2 reporter without any coils
CC053	pCUP1-mTurqoise2-E_4	Amp	Cu_{2+} inducible mTurqoise2 reporter fused to E_4 coil
CC080	pCUP1-mTurqoise2	Amp	Cu_{2+} inducible mTurqoise2 reporter without any coils
CC052	pRNR2-VioA	Amp	VioA enzyme
CC063	pTDH3-VioB	Amp	VioB enzyme
CC058	pTEF2-VioC-K_4	Amp	VioC enzyme fused to K_4 coil
CC059	pHHF2-VioD-SYNZIP5$(P2)_2$	Amp	VioD enzyme fused to SYNZIP5$(P2)_2$ coil
CC065	pPGK1-SYNZIP6-VioE-E_3	Amp	VioE enzyme fused to N-terminal SYNZIP6 and C-terminal E_3 coils
Multi-gene constructs for yeast transformation			
CC012	CEN2/ARS6: CC006+CC011+CC009	Kan/URA	Three-way FRET reporter system for EK coiled-coil strand displacement in yeast (Figures 4.4-6)
CC062	CEN2/ARS6: CC006+CC011+CC046	Kan/URA	Three-way FRET reporter system control without toehold on displacer protein mRuby2 (Figure 4.6)
CC101	CEN2/ARS6: CC096+CC097+CC054	Kan/URA	Three-way FRET reporter system for SYNZIP5/6 coiled-coil strand displacement in yeast (Figure 4.7)

CC102	CEN2/ARS6: CC096+CC097+CC081	Kan/URA	Three-way FRET reporter system control without displacing coil $(P1)_4$ fused to mRuby2 (Figure 4.7)
CC070	LEU2: CC052+CC063+CC058+CC059+CC065	Kan/LEU	Integration plasmid containing all violacein biosynthetic pathway with coiled-coil fusions for constitutive expression of scaffolded complex (Figure 4.8); integrated into leu2 locus to create ySK001
CC077	CEN2/ARS6: CC053+(Spacer)+CC054	Kan/URA	mTurqoise2-SYNZIP6$(P1)_4$ and mRuby2-E_4 displacing reporter fusions for induced rearrangement of violacein scaffolded enzyme complex (Figure 4.9)
CC082	CEN2/ARS6: CC080+(Spacer)+CC081	Kan/URA	mTurqoise2 and mRuby2 non-displacing controls for induced rearrangement of violacein scaffolded enzyme complex (Figure 4.9)

Appendix B

REPRINT PERMISSIONS

B.1 Permission/Rights for Chapter 2



Copyright Clearance Center RightsLink®

Home | Create Account | Help

ACS Publications
Most Trusted. Most Cited. Most Read.

Title: Control of the Yeast Mating Pathway by Reconstitution of Functional α-Factor Using Split Intein-Catalyzed Reactions

Author: Ka-Hei Siu, Wilfred Chen

Publication: ACS Synthetic Biology

Publisher: American Chemical Society

Date: Aug 1, 2017

Copyright © 2017, American Chemical Society

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already a RightsLink user or want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK | CLOSE WINDOW

B.2 Permission/Rights for Chapter 3

Copyright Clearance Center RightsLink®

Home | Create Account | Help

SPRINGER NATURE

Title: Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function
Author: Ka-Hei Siu et al
Publication: Nature Chemical Biology
Publisher: Springer Nature
Date: Dec 10, 2018
Copyright © 2018, Springer Nature

LOGIN
If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.
Already a RightsLink user or want to learn more?

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select no to the question "Are you the Author of this Springer Nature content?".
Ownership of copyright in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:
To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).
The author and any academic institution, where they work, at the time may reproduce the contribution for the purpose of course teaching.
To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.
To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website.
Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read Springer Nature's online author reuse guidelines.

For full paper portion: Authors of original research papers published by Springer Nature are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v1.0

BACK | CLOSE WINDOW