An Integrated CRISPR-Cas Toolkit for Engineering Human Cells

by

Samuel David Perli

B.Tech., Indian Institute of Technology Madras (2008)
S.M., Massachusetts Institute of Technology (2010)

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

Author .....
Department of Electrical Engineering and Computer Science
May 19, 2015

Certified by .....
Timothy K. Lu
Associate Professor
Thesis Supervisor

Accepted by .....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Theses

An Integrated CRISPR-Cas Toolkit for Engineering Human Cells

by

Samuel David Perli

Submitted to the Department of Electrical Engineering and Computer Science
on May 19, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Natively functioning Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) system is a prokaryotic adaptive immune system that confers resistance to foreign genetic elements including plasmids and phages. Very recently, a two-component CRISPR-Cas technology from *Streptococcus Pyogenes* comprising of the RNA-guided DNA endonuclease Cas9 and the guide RNA (gRNA) has been demonstrated to enable unprecedented genome editing efficiency across all domains of life. Current applications however, employ CRISPR/Cas technology in a stand-alone fashion, isolated from the rich biological machinery of the host environment in which it is applied.

In this thesis, I present a toolkit designed by integrating CRISPR/Cas technology with a wide array of mammalian molecular components, thereby enabling altogether novel applications while enhancing the efficiency of current applications. By integrating a catalytically dead version of the CRISPR/Cas protein Cas9 (dCas9) with mammalian transcriptional activator VP64 and mammalian transcriptional repressor KRAB, we build and characterize tunable, multifunctional and orthogonal CRISPR/Cas transcription factors (CRISPR-TFs) in human cells. By integrating CRISPR-TFs and Cas6/Csy4 based RNA processing with multiple mammalian RNA regulatory strategies including RNA Polymerase II (RNAP II) promoters, RNA-triple-helix structures, introns, microRNAs and ribozymes, we demonstrate efficient modulation of endogenous promoters and the implementation of tunable synthetic circuits such as multistage cascades and RNA-dependent networks that can be rewired with Csy4. In summary, our integrated toolkit enables efficient and multiplexed modulation of endogenous gene networks, construction of highly scalable and tunable synthetic gene circuits. Our toolkit can be used for perturbing endogenous networks towards developmental, therapeutic and synthetic biology applications.

Thesis Supervisor: Timothy K. Lu
Title: Associate Professor

Acknowledgments

I give thanks to the Lord, my God, for granting me wisdom, spirit of understanding, strength, faith and perseverance to shepherd me through this journey of life. I am thankful for this wonderful gift of nature that I can study and engineer. I thank my parents for their exemplary lives, deep love, prayers and support. My father, especially for his encouragement and counsel. My mother, for her loving care and uplifting spirit.

I thank my PhD mentor Prof. Timothy Lu. He has been a great support for me all through the past four and a half years of being in his lab. He took me in as his first graduate student when I had very little knowledge of biology. I am grateful for the valuable time he spent with me and teaching me experimental biological skills starting from the very basics of pouring gels and cloning plasmids. I appreciate the freedom he has given me in choosing across from different projects and for the encouragement he gave me in incorporating computational approaches while doing biological research. I also thank him for nominating me for numerous awards and fellowships.

I thank Prof. Doug Lauffenburger for his insightful comments and guidance throughout my PhD thesis work. From him, I have learned the importance of being methodical in both approach and practice while doing biological research. I am also grateful for his availability given the very demanding administrative roles he has been engaged with. I thank Prof. Bruce Tidor for always challenging me to abstract biological systems and to think critically about my design principles. I am grateful to have Prof. Lauffenburger and Prof. Tidor on my thesis committee.

I am thankful for the interdisciplinary lab that Tim has set up. Lior Nissim has been a good friend and helped me pick up the ropes with tissue culture. Cheryl Cui has been an awesome partner in the lab. Fahim, Mark, Jake, Rob, Nate, Alan, Gigi, Adina, Sasha, Barbara, Isaak, Eleanore, ZJ, Cong, Oli, Sebastian, Sarah, Hiroki, Bijan, Kevin, Mario, Gianluca, Ming-Ru, Ying-Chou and Ramez, you all have provided me a fun environment to work in. Ky has been a tremendously gifted lab manager. Olga and Jen, you made everything run in the lab, you guys are awesome!

I thank my small group friends, for their heartfelt fellowship, prayers and awesome fun activities including board games and potlucks. Steph has been an awesome friend and I love the conversations we had discussing ministry, science and life. Michelle, Heather, Dan, Anya, Zach, Zech and Amanda have been great friends. I thank Albert, Courtney, Andrew and Anthony for encouraging me and giving me valuable counsel when I transitioned to do research in biology. I am thankful for my awesome roommates - Marcus, George, Yong, Adam, Keith and Dixia. Ming, Yukkee, David and John were great accountability partners. I'm also very grateful for Sunday night GIG (Group Investigating God) dinners, discussions and prayers with Neel, XiaoYu, Dixia, Yang-ting, Soyoung and others. My time with the Graduate Christian Fellowship (GCF) has been deeply formative in my character and faith. I am grateful for all the wonderful relationships I built with each person in the fellowship, especially, Kunle, Emily, Eric and Victoria, Emily, Po-Ru and Charis, Peng and Annie, Gerald, Sam, Grace, Nigel, Frank, Alice, Lyndon, Ken, Mark, John, Dawsen, Kevin, Mary and Marianne. I have been very blessed by my time at Park Street Church, for the very intellectually enlightening sermons by Gordon Hugenberger, the faith filled community and the delicious after service lunches and discussions.

I thank my former advisor, Prof. Dina Katabi, for her insights and guidance. It has been a wonderful experience working along with her. I am grateful for being the recipient of multiple fellowships over the years that have instilled in me a sense of assurance while navigating across the fields of computer science and biology. I thank my undergrad mentors, Prof. David Koilpillai and Prof. Anurag Kumar, for the indelible impression they left on me. Finally, I thank my brother, who has also been a great friend to me, my sister and friends near and far, who constantly remember me in their thoughts and prayers.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Advances in the ability to synthesize and modify DNA have fueled the burgeoning discipline of Synthetic Biology. Enzymes (including polymerases, ligases, and restriction endonucleases), oligonucleotides and polymerase chain reaction (PCR) have provided ways to isolate genes and gene fragments, as well as to introduce small-scale mutations into genes *in-vitro*, in cells, and in model organisms [134, 141, 121, 137]. Custom designed nucleases such as Zinc Finger Nucleases (ZFNs) and Transcription Activator-Like (TAL) Effector Nucleases (TALENs) marked a newer generation of tools in the genome engineering toolkit following oligonucleotides and restriction enzymes [11, 27]. TALENs were easier than ZFNs to build and characterize but the requirement of individual protein synthesis for each separate DNA target sequence remained a barrier for their widespread adoption. The RNA guided DNA endonuclease Cas9 from the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) prokaryotic adaptive immune system is the newest tool in the genome engineering toolkit. Unlike ZFNs or TALENs the same Cas9 nuclease can be directed to target multiple different DNA sequences guided by sequence complementarity specified by the complexing guide-RNA (gRNA). The ease of use and high efficiency of this system has led to its widespread adoption and is currently the state-of-the-art tool for genome engineering [35].

1.1 Motivations for building an integrated toolkit

Although the CRISPR-Cas system has satisfactory performance for current genome engineering applications, we believe it holds greater promise towards engineering human cells at large. Current applications employ CRISPR-Cas technology in a standalone fashion, isolated from the rich biology of the human host cell. Biological components found in nature, however, are often multifunctional and heavily employ the concept of reuse and integration to achieve complex cellular behavior. To mention a few examples - signaling modules such as sonic hedgehog and Wnt are reused at different time points during differentiation and development in the same organism [50], disparate cell surface receptors [22] use the same downstream signaling cascades such as MAPK, and Nuclear Receptor (NR) transcription factors effect drastically different transcriptional programs by integrating with and recruiting many different types of effectors complexes depending on the specific cellular context [45]. Along similar lines, we were inspired by the idea that integrating CRISPR-Cas technology with molecular components involving human RNA Polymerase II (RNAP II) promoters, endogenous RNA processing elements and proteins with application-specific functions can potentially enhance the breadth of applications that can be made possible with this technology.

Our motivations for building an integrated CRISPR-Cas toolkit is threefold:

1.1.1 Repurposing Cas9's DNA binding ability

In addition to genome engineering, the ability to bind a specific sequence of DNA can be repurposed for enabling different applications such as gene regulation and chromatin modification. Cas9's ability to specifically recognize and cleave double stranded DNA can be theoretically decoupled in to two separate biological functions - 1) the binding of a specific sequence of DNA and 2) the catalytic introduction of a double stranded break on to the bound DNA. If the active site residues with in the nuclease domains of Cas9 were mutated to result in a catalytically dead version of Cas9 (dCas9), then such a protein could function purely as a DNA binding protein

with out any associated endonuclease activity. If in addition, dCas9 were to be fused to specific human transcription factors, one could envision modulating gene expression or affecting chromatin dynamics in a programmable fashion.

1.1.2 gRNA expression from RNAP II promoters

One of the hallmarks of multicellular life is tissue specific gene expression. This feature is enabled in higher organisms by evolving highly sophisticated RNAP II promoters that integrate multiple different cellular signals at the transcriptional level. By encoding gRNA transcription from RNAP II promoters, CRISPR-Cas activity can be similarly integrated with a plethora of cellular signals. Moreover, a single, wild-type human RNAP II transcript often encodes distinct RNA species including microRNAs, non-coding RNAs and multiple different iso-forms of a given protein, all transcribed in a concomitant fashion. Leveraging such an architecture for gRNA production can provide us with multiplexing gains associated with expressing multiple different gRNAs from a single transcript, targeting correspondingly different DNA sequences.

1.1.3 Enabling novel applications

The native CRISPR-Cas system is an adaptive immune system originally fashioned by nature to function in a prokaryotic host environment. Once a double stranded break is introduced by Cas9 in the plasmid/phage DNA, the replication of targeted DNA is stalled and ultimately the targeted DNA is degraded or diluted away due to the absence of a Non-homologous DNA repair pathway. However, human cells contain the Non-Homologous End Joining (NHEJ) DNA repair pathway that can repair DNA double stranded breaks in the absence of a repair template. NHEJ is a highly active, albeit, error-prone DNA repair pathway present in all eukaryotes. If Cas9 were to repeatedly cleave a particular locus of DNA with in human cells, such mechanism, aided by the error-prone NHEJ pathway, could theoretically result in the programmable and continuous generation of exponentially increasing genomic

diversity at a specific locus. The generated genomic diversity can serve as a molecular record of the intensity or duration of Cas9 activity which in itself can be coupled to cellular signals via RNAP II promoters as discussed above. When activated in proliferating cells, the generated genomic diversity at the particular locus can serve as a traceable marker for a phylogenetic reconstruction of lineage maps of each individual cell. Finally, when operated under a directed evolution paradigm, the continuous generation of genomic diversity can be exploited for protein engineering applications.

1.2 Challenges towards integration

Multiple challenges have to be kept in mind while designing an integrated toolkit, some of which are discussed below:

1.2.1 Loading and retroactivity

A traditional challenge while expressing non-native molecular species such as Cas9 or the gRNA within a host cell is the effect of "loading" [20, 112, 56, 86]. The loading effect is triggered by the potential sequestration of limited shared resources such as DNA and RNA polymerases, mRNA transporters, chaperone proteins, etc which are needed for the maintenance and expression of the exogenous molecular species as well as the molecular components endogenous to the host cell. The effect of loading can range from slower growth phenotype and altered energy metabolism to cell-cycle arrest, severe cytotoxicity and apoptosis of the host cell. Although such effects can never be absolutely eliminated, approaches that allow only a transient introduction of the non-native molecular species or those that pose minimal metabolic burden on the host cell need to be employed. A related effect of sequestration of proteins across different modules in a circuit termed retroactivity [31, 66, 18] has been reported. Although natural systems function despite the presence of retroactivity, sometimes even exploiting it [18, 81], synthetic biologists have developed fast acting buffering devices that can alleviate the effects of retroactivity [101].

1.2.2 Crosstalk

As discussed earlier as one of the efforts of integration, we seek to endow DNA binding proteins with multifunctional abilities by recruiting specific host transcription factors. It is important to note that such recruitment of transcription factors can result in undesirable modulation of gene expression elsewhere in the genome. This is especially a cause of concern while recruiting transcription factors such as Histone Deacetylases (HDACs) which can affect chromatin condensation over several kbps of DNA [78, 79] resulting in unanticipated repression of gene expression several kbps away from the targeted region. In addition, according the principle of reversibility, just as the transcription factors are recruited by the DNA binding protein to the targeted locus, the DNA binding protein is recruited to the native sites of function of the host transcription factors as well. This can result in undesirable transcriptional effects with in the host cell, especially if the DNA binding protein is huge in size such as Cas9, resulting in a steric hindrance of access for other complexing proteins at the native sites.

1.2.3 Matching the expression levels

While building an integrated toolkit with multiple components involved, it is important to match the expression levels of all individual components to function well together as an effective complex. In addition to the stoichiometries of each of the molecular candidates within functional complexes, different affinities of the components to each other must be considered. The need to match expression levels is further underscored if the components involved in the toolkit belong to different molecular species such as RNA and protein as in our system. Typically, a single mRNA transcript can be translated to produce about 100 to 1000 copies of protein [128] however, each Cas9 protein complexes with only one gRNA. If gRNAs were to be generated from an RNAP II promoter then it follows that much stronger promoters be used for expressing Cas9 compared to those used for expressing gRNAs, assuming similar rates of RNA decay, etc.

The rest of the thesis is organized as follows:

Chapter 2 introduces ZF technology and will describe ZF based transcription factor design and their characterization in *Saccharomyces cervisiae*.

Chapter 3 introduces the native CRISPR-Cas system and will describe the design and characterization of CRISPR-Cas transcription factors in Human cells.

Chapter 4 covers the integration of CRISPR-Cas with multiple RNA regulatory strategies including gRNA expression from RNAP II promoters, RNA-triple-helix structures, introns, microRNAs and ribozymes.

We will conclude in Chapter 5

Chapter 2

Synthetic Zinc Finger transcription factors

Zinc Finger (ZF) proteins are amongst the most abundant proteins encoded by the eukaryotic genome [85]. Originally discovered in *Xenopus* transcription factor IIIA [83], ZF domains are now found to be present in many other eukaryotic transcription factors and in lipid binding proteins, chaperone proteins and RNA packaging proteins [85]. Unlike the DNA binding proteins known prior to the discovery of ZFs, which bind DNA as symmetric dimers making use of the 2-fold symmetry of the DNA helix backbone and/or palindromic sequences with in DNA, ZFs were found to be modular, containing a similar, structural framework tandemly linked together in a linear fashion. The variation in binding specificity is achieved by varying key amino acid residues that make contact with specific sets of nucleotide bases with in the DNA [64]. The classic Cys_2His_2 is the most abundant zinc finger motif comprising of ~ 30 amino acid residues arranged in a simple $\beta\beta\alpha$ fold which is stabilized by hydrophobic interactions and a Zinc ion [85]. The interaction of the α-helix domain with the major grove of DNA results in sequence-specific base contacts, with each Cys_2His_2 finger recognizing 3 bp of DNA [39, 106].

2.1 Custom designed Zinc Finger proteins

Ever since the structural basis for ZF DNA recognition has been uncovered [39, 106], efforts were underway to build custom designed ZF DNA binding proteins that can specifically recognize and bind any desired sequence of DNA. Enhancing the specificity of DNA recognition was potentially possible by arraying multiple structurally similar but distinct Cys_2His_2 ZF domains that recognize different sets of 3 bp target sequences in tandem [53, 122]. However, arraying individual Cys_2His_2 ZF domains that recognize specific sets of 3 bp DNA did not prove to be thoroughly modular as predicted and alternative, screening based strategies were developed to build arrayed ZF proteins that specifically bind 9, 12 and 18 bps of DNA [70, 62, 61]. In parallel, chimeric fusions of transcriptional activation and repression domains with ZF DNA binding domains were built for controlling gene expression at will [8]. Similarly, ZF Nucleases were built by chimeric fusions of the nuclease domain of a type IIs restriction enzyme *FokI* with arrayed ZF proteins [12, 21, 92]. Because of the dimerization requirement for *FokI* nuclease activity, ZF nucleases were designed to function in a dimeric fashion, recognizing a doublet of 9 bp target sites located in close proximity.

2.2 Synthetic Zinc Finger transcription factor circuits

Synthetic gene circuits built previously such as pulse generators [7], oscillators [38, 2] and counters [44] have used natural DNA binding domains from bacteria and yeast such as TetR, LacI, LexA and Gal4. Limited availability of well characterized and orthogonal transcription factors remained a barrier for engineering and modulating complex transcriptional circuits. Custom designed Zinc Finger transcription factors ushered in a new set of tools for building complex synthetic gene circuits. The Joung lab has graciously shared with us some synthetic 9 bp targeting ZFs they have built and characterized using a bacterial two-hybrid (B2H) selection system [92] which are listed in Table 2.1. Using these ZFs, we set out to build and characterize ZF based

synthetic gene circuits in *Saccharomyces cerevisiae*. Building a well characterized set of synthetic transcriptional activators and repressors is a simple first step along this endeavor.

ZF	B2H (Act)	Amino acid residues			Target Site	GI #
		Finger #1	Finger #2	Finger #3		
42-10	7.19	TGQILDR	VAHSLKR	DPSNLRR	aGACGCTGCTc	78
43-8	11.6	RQDRLDR	QKEHLAG	RRDNLNR	aGAGTGAGGAc	52
21-16	3.5	RNFILQR	QGGNLVR	QQTGLNV	aTTAGAAGTGa	153
97-4	3.7	RQSNLSR	RNEHLVL	QKTGLRV	aTTATGGGAGa	83
150-4	7.39	KGERLVR	RMDNLST	RKDALNR	gGTGTAGGGGt	17
151-1	7.22	IPNHLAR	QSAHLKR	QDVSLVR	tGCAGGAGGTg	52
158-2	9.74	DKTKLRV	VRHNLTR	QSTSLQR	tGTAGATGGAg	86
159-3	9.43	SAQALAR	QGGNLTR	QHPNLTR	tGAAGAAGCTg	302
13-6	5.43	TNQKLEV	VRHNLQR	QHPNLTR	aGAAGATGGTg	292
14-3	5.57	APSKLDR	LGENLRR	DGGNLGR	gGACGACGGCa	32
62-1	3.8	TGQRLRI	QNQNLAR	DKSVLAR	gGCCGAAGATa	119
63-4	4.9	KKDHLHR	QRPHLTN	VGASLKR	aGCTGGAGGGt	37
54-8	14.5	NKTDLGR	RRDMLRR	RMDHLAG	aTGGGTGGCAt	52
55-1	8.15	DESTLRR	MKHHLGR	RSDHLSL	cTGGGGTGCCc	38
36-4	6.9	GRQALDR	DKANLTR	QRNNLGR	cGAAGACGCTg	124
37-12	9.7	RNFILQR	DRANLRR	RHDQLTR	tGAGGACGTGt	27
92-1	12.4	DSPTLRR	QRSSLVR	ERGNLTR	aGATGTAGCCt	61
93-10	5.6	APSKLKR	HKSSLTR	QRNALSG	cTTTGTTGGCa	214
128-2	7.3	TMAVLRR	RREVLEN	QTVNLDR	cGAAGTGGTCc	83
129-3	4.2	TAAVLTR	DRANLTR	RIDKLGD	cGGGGACGTCa	21
172-5	7.2	MKNTLTR	RQEHLVR	QKPHLSR	aGGAGGGGCTc	27
173-3	6	SAQALAR	QQTNLAR	VGSNLTR	aGATGAAGCTg	232

Table 2.1: List of Zinc Fingers built and characterized using a bacterial two-hybrid (B2H) system by the Joung lab [92]. For each of the ZF listed, the B2H fold activation is listed along side, as well as the number of *Saccharomyces cerevisiae* genomic instances (GI #) of each of the target sites is listed in the far right column.

2.2.1 Zinc Finger activators

We built a fusion of ZF43-8 (Table 2.1) with a commonly used eukaryotic transcriptional activator VP64 [122] as depicted in Figure 2-1-A.

A reporter gene was built using the minimal Cyc1 promoter from *S. cerevisiae*

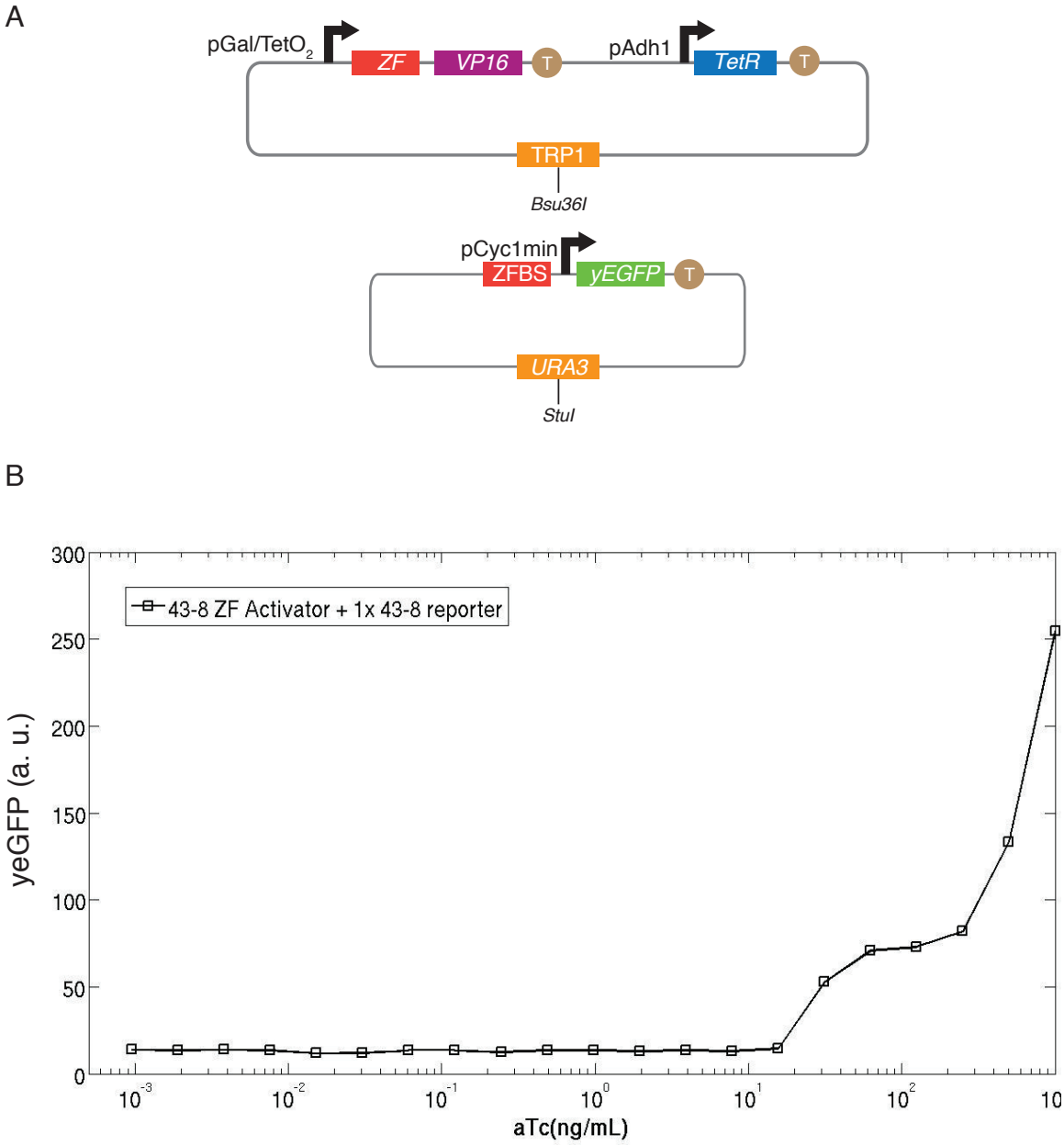


Figure 2-1: Tetracycline inducible activation of a synthetic reporter gene using a custom designed Zinc Finger transcription factor in *S. cerevisiae*. The mean of the measured yEGFP levels is plotted as a function of aTc concentration.

with a single (1x) ZF 43-8 binding site placed immediately upstream of the promoter. The expression of ZF43-8_VP16 was designed to be regulated by a hybrid Galactose/Tetracycline inducible promoter (pGal/$TetO_2$) [36], with TetR constitutively expressed by the Alcohol dehydrogenase promoter (pAdh) from *S. cerevisiae*. The plasmid constructs were linearized with unique cutting sites and were sequentially integrated in to the *S. cerevisiae* YPH500 genome. YPH500 colonies containing successful genomic integrations of the plasmid constructs are selected by auxotrophic selections. Refer to Section 2.4 for more details on the methods.

In Figure 2-1-B, we plot yEGFP fluorescence measured as a function of the concentration of the inducer, anhydrotetracycline (aTc). We observe a steep dose response with $>$ 10-fold activation of *yEGFP* expression with aTc concentrations approaching 1 ug/mL.

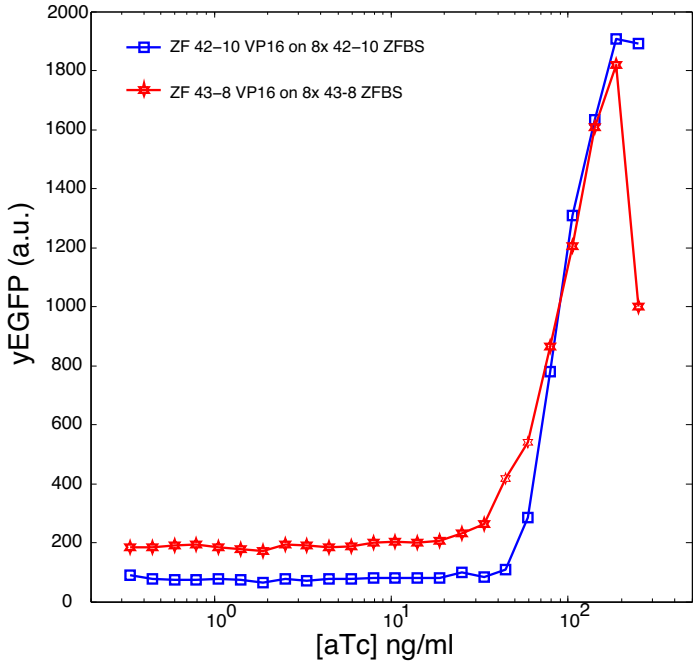


Figure 2-2: Tetracycline inducible activation of synthetic reporter genes containing 8x ZF binding sites. The mean of the measured yEGFP levels is plotted as a function of aTc concentration.

In an attempt to tune the dose response curve observed in Figure 2-1-B, we modified the reporter architecture by increasing the number of ZF binding sites from 1x to 8x. In addition, we also tried another ZF, ZF42-10 which has a different bind-

ing affinity to its target sequence (Table 2.1). The resulting dose response curves for ZF42-10 and ZF43-8 are plotted in Figure 2-2. Owing to positive co-operativity resulting from multiple transcription factor binding sites, we observe a much steeper dose-response response compared to the scenario with 1x ZF binding site. We observe no appreciable difference in the maximal levels of *yEGFP* expression for the two different ZFs although they have potentially different binding affinities (Table 2.1). This implies that ZF binding is not a rate limiting determinant for gene activation at high inducer concentration, suggesting that ZF-VP16 activation follows a Michaelis-Menten like kinetics. Given the steepness of the dose-response in Figure 2-2, it is hard to directly infer the difference in binding affinities for ZF42-10 and ZF43-8. We also note that different binding site sequences in the promoter regions of ZF42-10 and ZF43-8 reporters results in different basal levels of *yEGFP* expression.

2.2.2 Zinc Finger repressors

Given that we had transcriptional activators working in a predictable fashion, we turned our focus to building synthetic gene repressors in *S. cerevisiae*.

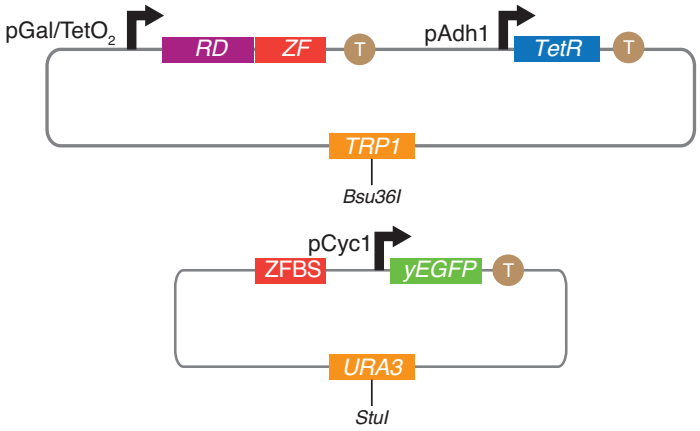


Figure 2-3: Tetracycline inducible repression of synthetic reporter genes containing ZF binding sites.

SSN6 is a well characterized, general repressor of transcription in yeast [77, 130]. SSN6 works as a chromatin remodeling transcription factor that regulates transcription of RNA polymerase II (RNAP II) promoters by binding and recruiting histone

deacetylases (HDACs) [116, 77, 76]. SSN6 is functions as a transcriptional repressor complex with Tup1p. In addition to SSN6, we have also tested UME6 [71, 118, 72] and SIN3 [142, 125] which are both subunits of the Rpd3L HDAC complex and effect transcriptional repression by chromatin remodeling.

The plasmid constructs for characterizing inducible repression are similar to those designed for characterizing inducible transcriptional activation. However, instead of a minimal Cyc1 promoter, we use the full-length Cyc1 promoter whose activity is targeted for repression by placing a ZF binding sequence (ZFBS) upstream of the promoter driving *yEGFP* expression (Ref. Figure 2-3). We fuse a repression domain (RD) to the N terminus of the ZF, whose expression is regulated by the hybrid Galactose/Tetracycline inducible promoter (pGal/$TetO_2$).

Table 2.2: Tetracycline induced SSN6-ZF fusion results

5'-SSN6-ZF-3' repressors integrated into 1x ZFBS-UAS-TATA-pCyc1-*yEGFP* clones				
Clone	Description (ZF/ZFBS)	w/ aTc	w/o aTc	Fold repression
SPY111_1	SSN6-ZF42-10 + 1x 43-8 ZFBS	56.74	201.69	3.55
SPY111_2	SSN6-ZF43-8 + 1x 43-8 ZFBS	113.42	355.45	3.13
SPY111_3	SSN6-ZF43-8 + 1x 43-8 ZFBS	54.25	130.97	2.41
SPY111_4	SSN6-ZF43-8 + 1x 43-8 ZFBS	74.99	191.1	2.54
SPY112_1	SSN6-ZF42-10 + 1x 43-8 ZFBS	72.34	286.44	3.95
SPY112_2	SSN6-ZF42-10 + 1x 43-8 ZFBS	71.05	254.83	3.58
SPY112_3	SSN6-ZF42-10 + 1x 43-8 ZFBS	66.71	220.67	3.30
SPY112_4	SSN6-ZF42-10 + 1x 43-8 ZFBS	100.9	252.55	2.50
SPY113_1	SSN6-ZF43-8 + 1x 42-10 ZFBS	532.8	547.37	1.02
SPY113_2	SSN6-ZF43-8 + 1x 42-10 ZFBS	283.87	243.62	0.85
SPY113_3	SSN6-ZF43-8 + 1x 42-10 ZFBS	216.74	210.97	0.97
SPY114_4	SSN6-ZF43-8 + 1x 42-10 ZFBS	289.03	278.81	0.96
SPY114_1	SSN6-ZF42-10 + 1x 42-10 ZFBS	187.69	777.37	4.14
SPY114_2	SSN6-ZF42-10 + 1x 42-10 ZFBS	165.48	567.42	3.42
SPY114_3	SSN6-ZF42-10 + 1x 42-10 ZFBS	152.61	532.8	3.49
SPY114_4	SSN6-ZF42-10 + 1x 42-10 ZFBS	138.24	557.31	4.03

We observed a very characteristic and pronounced slow-growth phenotype while testing for repression using each of the candidate repressors. Moreover, the repression results were not reproducible when multiple clones were assayed for each of the

Table 2.3: Tetracycline induced SIN3-ZF fusion based repression results

5‘-SIN3-ZF-3’ repressors integrated into 1x ZFBS-UAS-TATA-pCyc1-*yEGFP* clones				
Clone	Description	w/ aTc	w/o aTc	Fold repression
SPY103_1	SIN3-ZF43-8 + 1x 43-8 ZFBS	209.08	189.38	0.90
SPY103_2	SIN3-ZF43-8 + 1x 43-8 ZFBS	41.05	78.09	1.90
SPY104_1	SIN3-ZF42-10 + 1x 43-8 ZFBS	144.6	151.25	1.04
SPY104_2	SIN3-ZF42-10 + 1x 43-8 ZFBS	22.47	113.42	5.04
SPY105_1	SIN3-ZF43-8 + 1x 42-10 ZFBS	649.38	604.3	0.93
SPY105_2	SIN3-ZF43-8 + 1x 42-10 ZFBS	577.72	626.43	1.08
SPY106_1	SIN3-ZF42-10 + 1x 42-10 ZFBS	96.03	469.76	4.89
SPY106_2	SIN3-ZF42-10 + 1x 42-10 ZFBS	286.44	399.54	1.39

construct combinations. In Tables 2.2 2.3 and 2.4, we summarize the results obtained for inducible repression via SSN6, SIN3 and UME6 respectively.

We speculated one of the reasons for characteristic and pronounced slow-growth phenotype and non-reproducibility of inducible repression might be due to off-target effects, given multiple occurences of each of the ZF target site present in the *S. Cerevisiae* genome. To address this issue, we tried an inducible synthetic repressor construct using the bacterial LexA DNA binding domain which specifically binds a 16 bp palindromic sequence 5‘-CTGTATATATATACAG-3’. Given that there are no instances of the LexA target sequences present in the *S. Cerevisiae* genome, we expected to see an ameliorated performance. However, we observed a similar slow-growth phenotype and an erratic dose response with LexA-SSN6 repressor as shown in Figure 2-4. This observation rules out the role of potential off-target effects mediated by ZF-SSN6 fusion proteins and points to a possibly unanticipated effect of repression via chromatin modification. We believe that when ZF-SSN6, ZF-SIN3 or ZF-UME6 are recruited to the synthetic target promoter to effect transcriptional repression, they also effect transcriptional repression at endogenous genes located in the vicinity of the synthetic target promoter, resulting in the associated slower-growth phenotype.

Table 2.4: Tetracycline induced UME6-ZF fusion based repression results

5'-UME6-ZF-3' repressors integrated into 1x ZFBS-UAS-TATA-pCyc1-*yEGFP* clones				
Clone	Description (ZF/ZFBS)	w/ aTc	w/o aTc	Fold repression
SPY107_1	UME6-ZF43-8 + 1x 43-8 ZFBS	102.27	106.02	1.03
SPY107_2	UME6-ZF43-8 + 1x 43-8 ZFBS	42.55	149.89	3.52
SPY107_3	UME6-ZF43-8 + 1x 43-8 ZFBS	19.99	140.75	7.04
SPY107_4	UME6-ZF43-8 + 1x 43-8 ZFBS	42.94	162.53	3.78
SPY107_5	UME6-ZF43-8 + 1x 43-8 ZFBS	16.65	158.2	9.50
SPY107_6	UME6-ZF43-8 + 1x 43-8 ZFBS	16.4	152.61	9.30
SPY108_1	UME6-ZF42-10 + 1x 43-8 ZFBS	60.98	189.38	3.10
SPY108_2	UME6-ZF42-10 + 1x 43-8 ZFBS	56.74	153.99	2.71
SPY108_3	UME6-ZF42-10 + 1x 43-8 ZFBS	52.8	198.1	3.75
SPY108_4	UME6-ZF42-10 + 1x 43-8 ZFBS	45.32	212.88	4.69
SPY108_5	UME6-ZF42-10 + 1x 43-8 ZFBS	54.74	259.46	4.73
SPY108_6	UME6-ZF42-10 + 1x 43-8 ZFBS	56.23	162.53	2.89
SPY109_1	UME6-ZF43-8 + 1x 42-10 ZFBS	60.98	504.81	8.27
SPY109_2	UME6-ZF43-8 + 1x 42-10 ZFBS	873.79	991.05	1.13
SPY109_3	UME6-ZF43-8 + 1x 42-10 ZFBS	64.94	542.47	8.35
SPY109_4	UME6-ZF43-8 + 1x 42-10 ZFBS	60.98	465.55	7.63
SPY109_5	UME6-ZF43-8 + 1x 42-10 ZFBS	64.36	582.94	9.05
SPY109_6	UME6-ZF43-8 + 1x 42-10 ZFBS	62.08	598.89	9.64
SPY110_1	UME6-ZF42-10 + 1x 42-10 ZFBS	143.3	1055.45	7.36
SPY110_2	UME6-ZF42-10 + 1x 42-10 ZFBS	94.75	842.9	8.89
SPY110_3	UME6-ZF42-10 + 1x 42-10 ZFBS	77.74	572.55	7.36
SPY110_4	UME6-ZF42-10 + 1x 42-10 ZFBS	107.46	691.58	6.43
SPY110_5	UME6-ZF42-10 + 1x 42-10 ZFBS	842.91	827.88	0.98
SPY110_6	UME6-ZF42-10 + 1x 42-10 ZFBS	126.35	577.72	4.57

2.3 A transcription factor ratiometer

As building inducible synthetic repressors in *S. cerevisiae* proved harder than initially anticipated, we turned to using stand-alone synthetic ZFs as repressors that can effect transcriptional repression via steric hinderance. Along these lines, we built synthetic gene circuits that exploit the principle of competition across similar ZF transcription factors for a given target sequence. The design is depicted in Figure 2-5-A.

Our intuition behind the design of the transcription factor ratio meter circuit in Figure 2-5-A is as follows. For different intracellular levels of competing transcription

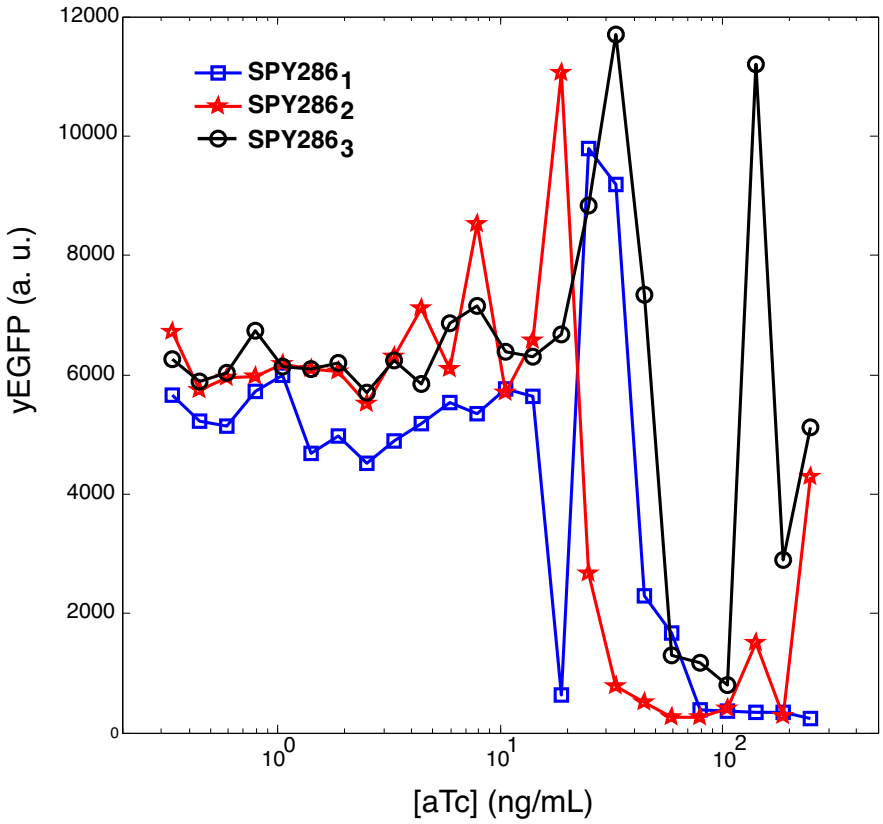


Figure 2-4: Dose response of synthetic LexA-SSN6 based repressor in *S. cerevisiae*. The mean of the measured yEGFP levels is plotted as a function of aTc concentration. Potential chromatin modification via LexA-SSN6 results in slow-growth phenotype and unpredictable *yEGFP* expression.

factor ZF, proportionately higher intracellular levels of activating transcription factor ZF-VP16 would be needed to achieve a given level of *BFP2* expression. Thus, the expression level of *BFP2* would be a function of the ratio of intracellular concentration of ZF-VP16 and ZF. In Figure 2-5-A, we modified the ZF activator design in Figure 2-1 by adding an additional module that expresses the same ZF under the control of a hybrid Galactose/Isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible promoter (pGal/LacO$_2$) [36], with LacI being constitutively expressed by the Alcohol dehydrogenase promoter (pAdh) from *S. cerevisiae*. In this design, the expression levels of transcription factors ZF-VP16 and ZF can be tuned independently by varying the levels of aTc and IPTG respectively.

In Figure 2-5-B, we plot the levels of yEGFP as a function of the concentration of

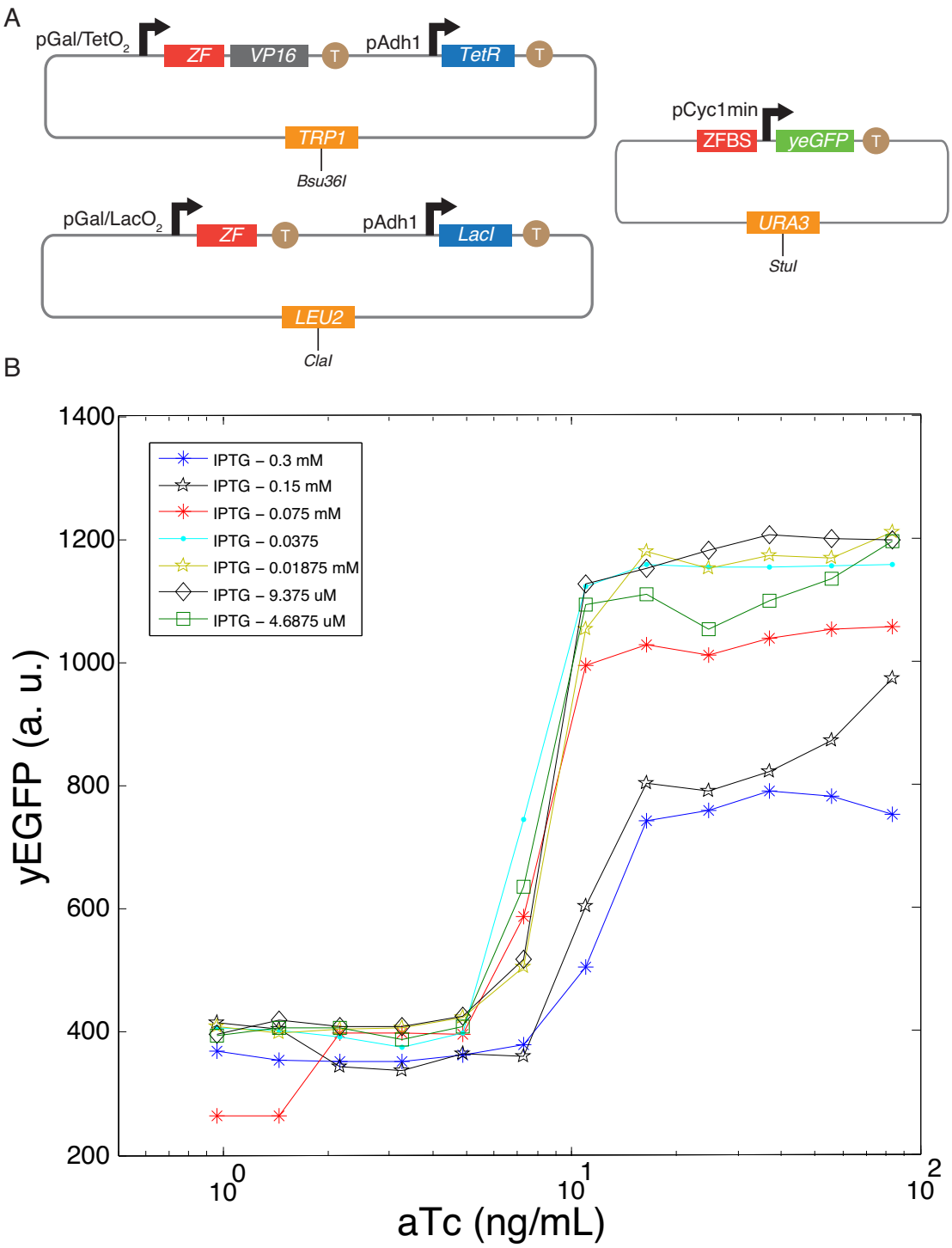


Figure 2-5: ZF ratio meter in *S. cerevisiae*. A) The reporter gene *yEGFP*'s expression level is designed to be a function of the ratio of the levels of transcription factors ZF-VP16 and ZF. B) The mean of yEGFP levels are plotted as a function of the concentration of aTc for different concentrations of IPTG.

aTc for different concentrations of IPTG. We observe that for a given concentration of aTc, as the concentration of IPTG increases, the levels of the output protein yEGFP decrease as was predicted. We also notice that the shape of the aTc dose response curves remains the same for different levels of IPTG, once again demonstrating a Michaelis-Menten like kinetics observed earlier with ZF activators in Section 2.2.1.

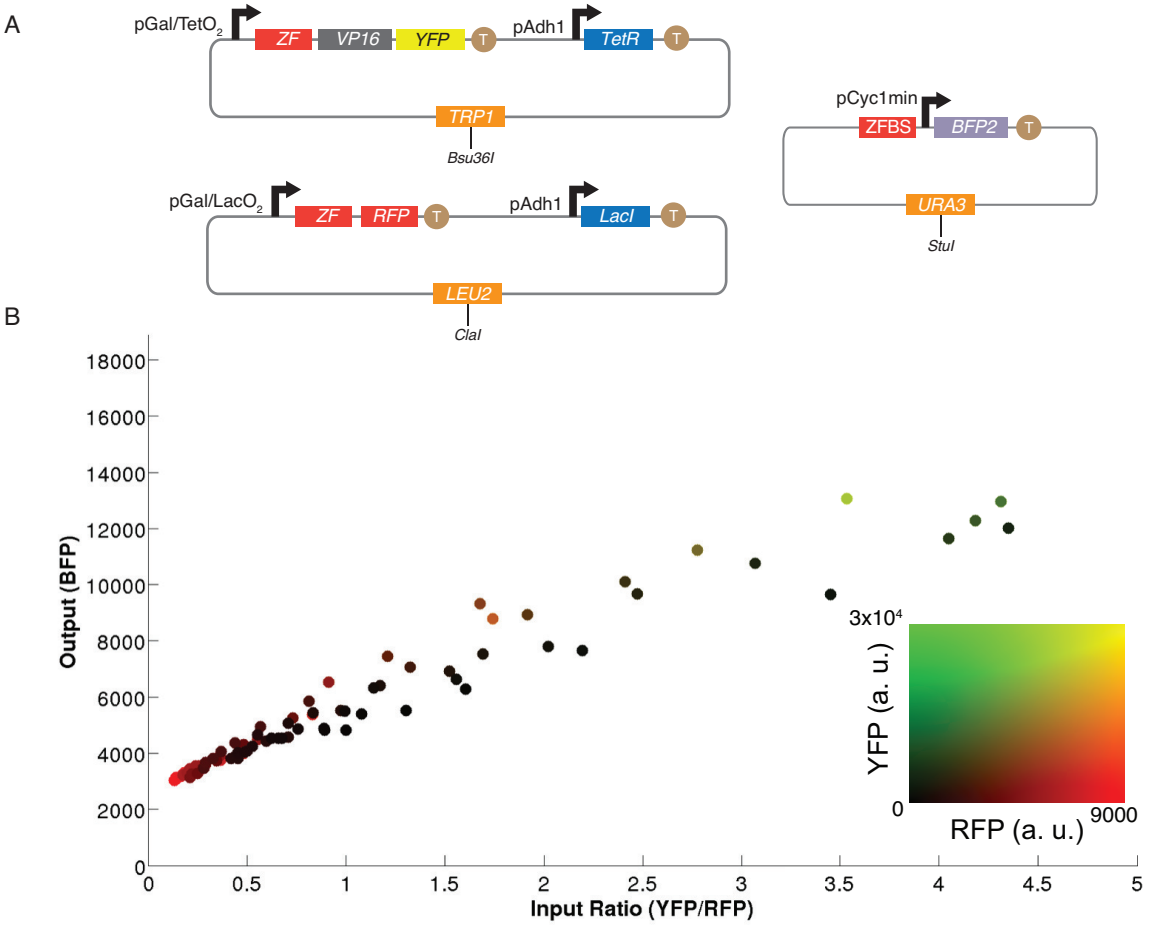


Figure 2-6: ZF ratio meter characterization in *S. cerevisiae*. A) The levels of output protein BFP2 is designed to be a function of the ratio of the levels of transcription factors ZF-VP16-YFP and ZF-RFP. B) The mean BFP levels measured are plotted against the ratio of the mean levels of YFP and RFP for different sets of inducer concentrations. Each data point plotted is color coded to capture the relative levels of YFP and RFP as indicated in the inset.

In order to accurately measure the intracellular levels of activating and competing transcription factors, we modified the circuit design to incorporate fusions of fluorescent proteins to the transcription factors as depicted in Figure 2-6-A. The output

protein used is BFP2 while the ZF-VP16-YFP and ZF-RFP constitute the input transcription factors, all of whose levels can be measured by multicolor flow cytometry. In Figure 2-6-B, we plot the levels of BFP2 as the ratio of YFP and RFP per cell, measured by varying the concentration of aTc and IPTG. We observe an impressive linear correlation of the output being a measure of the ratio of input transcription factors. Moreover, we observe that for any given ratio, there is at least a 10-fold wide dynamic range of inputs over which the output is a ratio of the inputs.

2.3.1 Modeling the transcription factor ratio meter

The dynamics of the transcription factor ratio meter can be described by the following affinity equations:

$$D + TF \underset{k_{d-}}{\overset{k_{d+}}{\rightleftharpoons}} D\text{-}TF \xrightarrow{k_p} P$$
$$D + I \underset{k_{i-}}{\overset{k_{i+}}{\rightleftharpoons}} D\text{-}I$$

where TF is the activating transcription factor, I is the competing transcription factor, D is the DNA binding site upstream of the reporter gene and P is the protein expressed. $k_{d+}, k_{d-}, k_{i+}, k_{i-}$ and k_p are the associated rate constants as shown above.

According to the law of mass action, we have

$$\frac{d}{dt}[D\text{-}TF] = k_{d+}[D][TF] - k_{d-}[D\text{-}TF] - k_p[D\text{-}TF]$$
$$\frac{d}{dt}[D\text{-}I] = k_{i+}[D][I] - k_{i-}[D\text{-}TF]$$

In addition, since the total amount of reporter genes per cell, D_{Tot} is constant, $(= 1)$ in our case, we have

$$[D_{Tot}] = [D\text{-}TF] + [D] + [D\text{-}I] \tag{2.1}$$

At the steady state, we set $\frac{d}{dt}[D\text{-}TF] = 0$ and $\frac{d}{dt}[D\text{-}I] = 0$. This results in

$$[D\text{-}TF] = \frac{k_{d+}}{(k_{d-} + k_p)}[D][TF] = \frac{[D][TF]}{K_d} \qquad \left(K_d = \frac{k_{d-} + k_p}{k_{d+}}\right) \tag{2.2}$$

$$[D\text{-}I] = \frac{k_{i+}}{ki-}[D][I] = \frac{[D][I]}{Ki} \qquad \left(K_i = \frac{k_{i-}}{k_{i+}}\right) \tag{2.3}$$

From equation 2.1, we have

$$\begin{aligned}
[D_{Tot}] &= [D\text{-}TF] + K_d\frac{[D\text{-}TF]}{TF} + \frac{K_d}{K_i}\frac{[I][D\text{-}TF]}{[TF]} \\
\implies [D_{Tot}] &= [D\text{-}TF]\left(1 + \frac{K_d}{[TF]} + \frac{K_d}{K_I}\frac{[I]}{[TF]}\right) \\
\implies [D\text{-}TF] &= \frac{[D_{Tot}][TF]}{\left([TF] + K_d + \frac{K_d}{K_i}[I]\right)}
\end{aligned}$$

The rate of expression of the reporter gene, v is given by $k_p[D\text{-}TF]$

$$\begin{aligned}
v &= \frac{k_p[D_{Tot}][TF]}{\left([TF] + K_d + \frac{K_d}{K_i}[I]\right)} \\
&= \frac{v_{max}[TF]}{\left([TF] + K_d + \frac{K_d}{K_i}[I]\right)} \\
&= \frac{v_{max}\rho}{\left(\rho + \frac{K_d}{K_i} + \frac{K_d}{[I]}\right)}
\end{aligned}$$

where $v_{max} = k_p \times D_{Tot}$ is the maximum possible rate of expression of the reporter gene and $\rho = \frac{[TF]}{[I]}$ is the ratio of in-vivo concentration of the activating transcription factor to the in-vivo concentration of the competing DNA binding protein.

In Figure 2-7, we plot the input-output relationship for the transcription factor ratio meter as predicted by the above model. Assuming $K_d \simeq K_i$ (the activating transcription factor and the competing transcription factor are comprised of the same

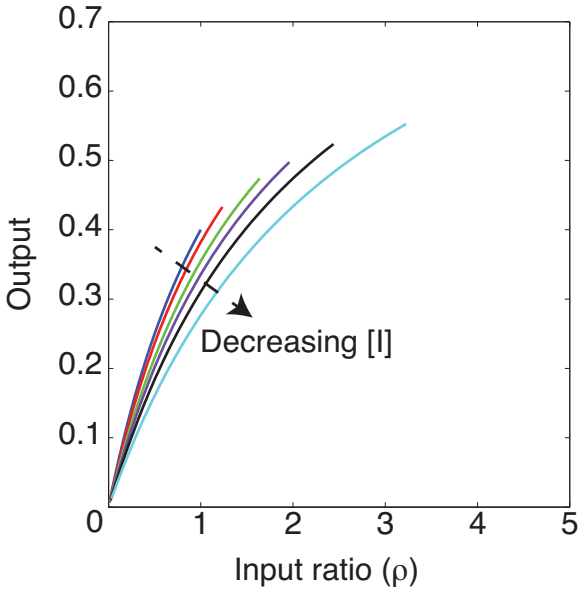


Figure 2-7: Transcription factor ratio meter model. The rate of production of the output protein is plotted as a function of the input transcription factor ratio for different concentrations of the competing transcription factor.

zinc finger), we observe that the rate of production of the output protein $v \propto \rho$ when $K_d << [I]$ i.e, for transcription factors that have very high affinity to their target DNA.

2.4 Methods

2.4.1 Strain and Plasmid construction

The *Saccharomyces cerevisiae* strain YPH500 ($MAT\alpha$ *ura*3-52 *lys*2-801_*amber* *ade*2-101_*ochre* *trp*1-Δ63 *his*3-Δ200 *leu*2-Δ1) was used for all of the above experiments. The Zinc Finger transcription factors were cloned in to pRS404 and pRS405 backbones using pTPGI and pLOGI promoters described in [36]. The reporter plasmids were cloned in the pRS406 backbone. The respective sequences for all the relevant constructs can be found in Section A.1. The designed plasmids were linearized with the restriction sites StuI, Bsu36I and ClaI and integrated via homologous recombination into the URA3, TRP1 and LEU2 loci of YPH500.

2.4.2 Yeast transformation

Log phase YPH500 cells cultured in 250 mL Erlenmeyer flasks at 30C, 300 rpm were harvested by spinning down at 3000g, 5 min after washing twice. A 360 uL transformation mix comprising 240 uL of 50% Poly Ethylene Glycol (PEG), 36 uL of 1mM Lithium Acetate, 50 uL of 2 mg/mL single stranded salmon sperm carrier DNA and 1 ug of linearized DNA is added to the harvested cells, resuspended and incubated at 42C for 45 min. Following incubation, yeast cells are spun down at 13000g for 30 seconds and the supernatant was removed. The cells were resuspended in 100 uL sterile double de-ionized water and plated on to CSM-dropout plates for prototrophic selection and incubated at 30C for two to three days. Colonies of yeast containing successful integrants are picked and cultured in dropout media for further analysis.

2.4.3 Fluorescence assays

To assess expression of the reporter constructs, yeast cells expressing different ZF transcription factors were grown overnight (900 rpm, 30C) in 96-deep-well plates in yeast minimal media supplemented with galactose with appropriate selection (multiple independent colonies for each sample). Ten microliters of these cultures were then transferred into fresh media supplemented with galactose and specified concentrations of anhydrotetracycline (aTc) and/or IPTG and grown for 20 h (900 rpm, 30C) in deepwell plates before analysis by flow cytometry. An LSR Fortessa II flow cytometer equipped with 405 nm, 488 nm, and 561 nm lasers was used for all the experiments. yEGFP/YFP, BFP2, and RFP levels were detected using 488/FITC, 405/Pacific-Blue, and 561/TX-red laser/filter sets respectively. All samples were uniformly gated by forward and side scatter.

Chapter 3

Synthetic CRISPR-Cas transcription factors

3.1 History and Biology of CRISPR-Cas systems

The first *bona-fide* CRISPRs (clustered regularly interspaced palindromic repeats) were described by Japanese researchers in 1987 as a series of short direct repeats interspersed with short sequences in the *Escherichia Coli* genome [63]. Such interspersed repeats were later found in numerous other bacteria and archaea [102] although, their specific function remained unclear. Soon after, in 2002, CRISPR sequences were found to be represented in the non-coding RNA population of the archaeal transcriptome [136]. In 2005, three different research groups made the observation that the spacer sequences in CRISPRs derive from associated plasmid and phage DNA sequences [110, 103, 16]. With the prediction that various cas (CRISPR-associated genes) [65, 54, 110, 16] encode proteins with putative nuclease and helicase domains, researchers in [94] proposed that CRISPR-Cas is an adaptive immune system in bacteria and archaea, analogous to the eukaryotic RNAi, that uses spacer sequences as memory signatures of past infections. The first experimental evidence of CRISPR-Cas mediated adaptive immunity came in 2007 when researchers in [5] conducted infection experiments of the lactic acid bacterium *Streptococcus Theromophilus* with lytic phages. In 2008, mature CRISPR RNAs (crRNAs) were shown to complex with

Cas proteins and serve as guides to interfere with phage proliferation in *Escherichia Coli* [19]. In the same year, the CRISPR-Cas system was reported to exhibit DNA targeting activity in the pathogen *Staphylococcus Epidermidis* [98].

Functional CRISPR-Cas loci comprise a CRISPR array of identical repeats interspersed with short spacers derived from plasmid or phage DNA that encode crRNAs and an operon of *cas* genes that encode the Cas proteins. Adaptive immunity in bacteria and archaea that contain the CRISPR/Cas system occurs in the following three stages [6, 139]:

1. **Acquisition:** Integration of a short target sequence of plasmid or phage DNA as a spacer sequence in to the CRISPR array

2. **crRNA biogenesis:** Transcription of the CRISPR array and maturation of the precursor-crRNAs to generate mature crRNAs

3. **Targeting:** crRNA guided cleavage of the target nucleic acid by Cas proteins at sites complementary to the crRNA spacer sequences

Three types of CRISPR-Cas systems exist (I, II and III) that use distinct molecular mechanisms to achieve nucleic acid recognition and cleavage with in the overall theme above [93, 95]. For type I and type II systems, the protospacer adjacent motif (PAM), a short sequence motif adjacent to the crRNA-targeted sequence on the target DNA plays an essential role in the stages of acquisition and targeting [123]. The type I and type III systems use a large complex of Cas proteins for crRNA-guided targeting [57, 117, 59, 55] while the type II system uses a single protein Cas9 (also previously known as COG3513, Csx12, Cas5, or Csn1) for RNA-guided DNA recognition and cleavage [68, 48].

Cas9 is a large, multifunctional protein with two putative nuclease domains, HNH [16, 54, 94] and RuvC-like [94]. In 2012, the Doudna lab showed that Cas9 from *Streptococcus Pyogenes* is a dual-RNA-guided DNA endonuclease that uses the tracrRNA:crRNA duplex to effect DNA recognition and cleavage [68]. Mutating either the HNH or the RuvC-like domain in Cas9 results in a protein (nickase) with

single stranded DNA cleavage activity whereas mutating both domains (D10A and H840A) results in an RNA-guided DNA binding protein (dCas9) that does not cleave the target DNA. Target DNA recognition requires both, crRNA/target-DNA base pairing and the presence of a PAM adjacent to the target DNA [68, 35]. The dual tracrRNA:crRNA was then engineered as a single guide RNA (gRNA) that retains the two critical features: 1) the 20-nucleotide sequence at the 5'-end of the gRNA that determines the DNA target sequence by Watson-crick base pairing and 3) the stem loop and hairpin structures at the 3'-end that bind to Cas9 [68, 35].

3.2 CRISPR-Cas transcription factors (crisprTFs)

The observation that Cas9 when mutated to silence the endonuclease activity (Cas9:D10A, H840A) can be repurposed to function as an RNA-guided DNA binding domain paved way to building novel synthetic transcription factors based on Cas9. Recently Qi et. al., [114] showed that dCas9 can be used for gene silencing in *E. Coli*. When targeted to a gene of interest, they demonstrated that dCas9 can block the progression of RNA polymerase by forming d-loops and hence, silence the expression of the targeted gene. Along similar lines, researchers in [13] fused the omega subunit of RNA polymerase to dCas9 to achieve transcriptional activation in *E. Coli*.

We have taken a similar approach and have built and characterized synthetic transcription factors based on dCas9 in Human cells. In addition to modulating gene expression in a programmable fashion, we also demonstrate the potential of dCas9 in uncovering the regulatory maps of mammalian promoters. To this end, we fused a human-optimized dCas9 to three copies of SV40 nuclear localization sequence (NLS) and four tandem copies of VP16 (VP64) on the C-terminus to build a CRISPR transcription factor (crisprTF) as shown in Figure 3-1-A. The expression of the crisprTF cassette was set to be driven by the cytomegalovirus immediate-early promoter (pCMV). The gRNAs targeting different DNA sequences were constitutively expressed from a separate plasmid by the RNA polymerase III (RNAP III)-dependent U6 promoter (pU6) as previously described [84, 97]. We investigated the regulatory

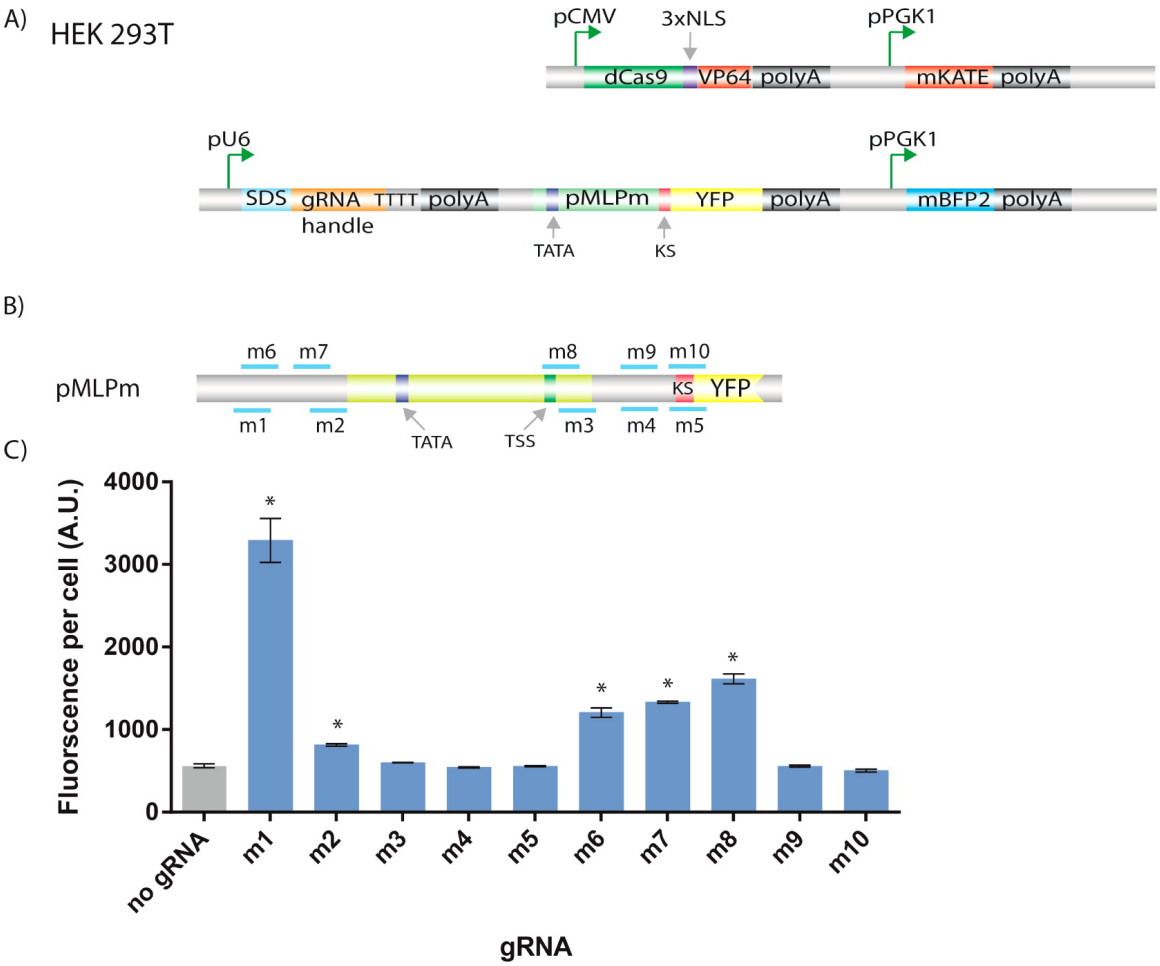


Figure 3-1: Regulation of *yfp* expression from a minimal MLP promoter (pMLPm) by crisprTFs in HEK293T cells. (A) dCas9_VP64 is expressed in HEK293T cells by the pCMV promoter and directed to target sequences in pMLPm. The mKATE (red) and mBFP2 (blue) fluorophores act as flow-cytometry gating controls for successful plasmid transfections. (B) Map of pMLPm illustrating the relative positions of known regulatory elements. Blue lines indicate target sites for each gRNA. (C) Regulation of *yfp* expression from pMLPm by crisprTFs based on the gRNAs shown in (B). HEK293T cells were co-transfected with the plasmids shown in (A), with specific gRNAs labeled as shown in the x-axis. Targeting crisprTFs to sequences upstream of the TATA box (by m1, m2, m6, and m7 gRNAs) resulted in higher *yfp* expression compared with the no gRNA control. Error bars indicate the standard error of the mean for three independent biological replicates. Asterisks on each bar indicate statistically significant changes in *yfp* expression relative to the no gRNA control (based on the two-sided Welch's t test, p-value < 0.05)

architecture of the minimal adenovirus major late promoter (pMLPm) by targeting crisprTFs to different positions across this promoter [120] as show in Figure 3-1-B. YFP as measured by flow cytometry was used as a readout for the pMLPm promoter

activity. We observe that crisprTFs activated gene expression when targeted to sequences upstream of the pMLPm TATA box (using m1, m2, m6, or m7 gRNA) or downstream of the transcription start site (m8 gRNA) (Figure 3-1-B and C).

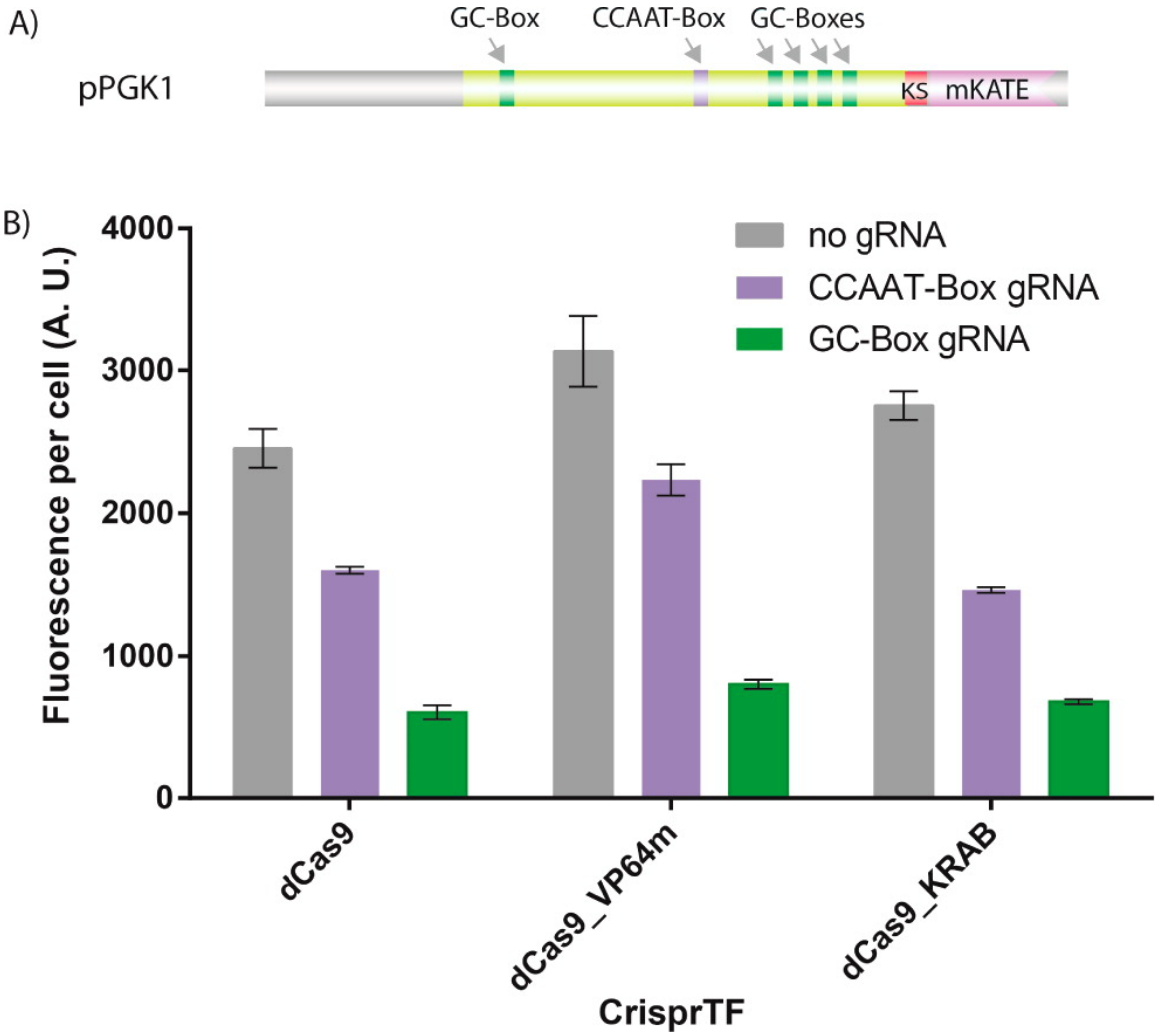


Figure 3-2: CrisprTF-mediated repression of the constitutive pPGK1 promoter in HEK293T cells. (A) Map of the pPGK1 promoter illustrating the relative positions of known regulatory elements. (B) CrisprTF-based targeted repression of the constitutive pPGK1 promoter. Constructs expressing different dCas9-based proteins (dCas9, dCas9-VP64, and dCas9-KRAB) were co-transfected with plasmids containing pPGK1_mKATE and constructs expressing no gRNAs or gRNAs targeting the CCAAT box or the GC-box gRNA. Significant repression of the pPGK1 promoter relative to the no gRNA control was observed with all of the three different dCas9 constructs (dCas9, dCas9-VP64, and dCas9-KRAB). Error bars indicate the standard error of the mean for three independent biological replicates.

Since the basal expression level of the pMLPm promoter is low, it was challeng-

ing to detect significant repression from this promoter. Thus, to demonstrate that crisprTFs can function as transcriptional repressors in mammalian cells, we placed mKATE under the control of a constitutive mammalian promoter, phosphoglycerate kinase 1 (pPGK1) [1], and targeted crisprTFs to this promoter (Figure 3A). pPGK1 is a strong, constitutive, TATA-less promoter that contains a CCAAT box [34] and five GC-boxes [14] These sites are the binding sites for the endogenous human transcription factors CBP and SP1, respectively. Targeting dCas9 alone, dCas9 fused to VP64 domain, or dCas9 fused to KRAB domain to the CCAAT box or the GC-boxes resulted in significant repression of the reporter gene (Figure 3B), presumably by preventing endogenous transcription factors from binding to specific DNA recognition elements [34, 14] within the pPGK1 promoter. Our finding that one can activate or repress the expression of a gene of interest by directing a single protein to different positions of a promoter is advantageous for the efficient design of synthetic transcriptional networks or rewiring natural ones. This property obviates the need for using separate orthogonal Cas9 protein fusions as activators and repressors.

3.2.1 Tunability

We further sought to explore the tunability of crisprTFs in the context of synthetic promoters. We observed synergistic activation in HEK293T cells when multiple gRNA operator sites were placed upstream of pMLPm (ref Figure 3-3), with up to 56× activation attained with 3× gRNA operator sites. The level of activation that is achieved by crisprTFs in human cells is comparable to the levels of activation reported for ZF- and TALE-activators [80, 88, 108] where higher activation levels can be achieved by increasing the number of operator sites or by targeting multiple synthetic transcription factors to the same locus [108, 91]. These results demonstrate that crisprTFs can be used to build synthetic promoters with tunable strengths by the straightforward engineering of gRNA-binding sites.

In addition, we noticed a strand bias associated with the orientation of the operator sites inserted upstream of pMLPm. We inserted 1× and 2× synthetic operator sites upstream of pMLPm in both the forward and the reverse orientation, result-

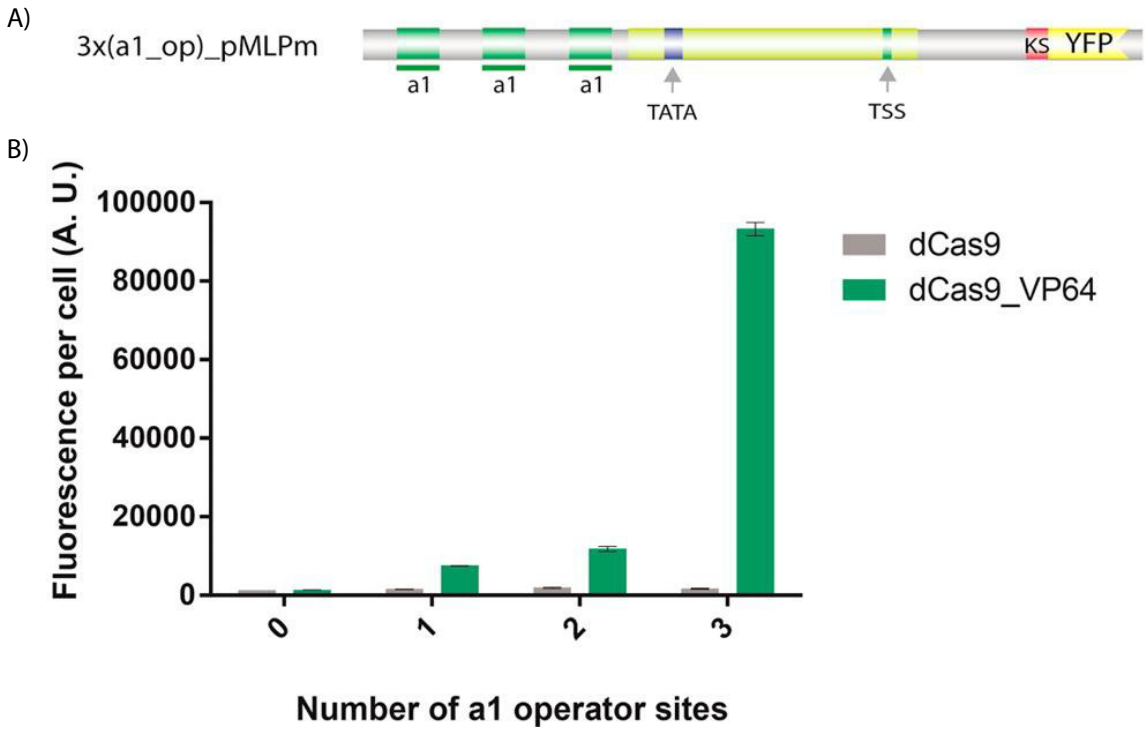


Figure 3-3: Synergistic and tunable activation of synthetic promoters with arrayed operator sites upstream of pMLPm in HEK293T cells using crisprTFs. (A) A schematic view of the pMLPm synthetic promoter with three a1_gRNA operator sites arrayed upstream of pMLPm, thus named 3×(a1_op)_pMLPm. (B) Increasing the number of arrayed a1_gRNA operator sites upstream of pMLPm resulted in higher *yfp* expression in HEK293T cells when co-transfected with a1_gRNA and dCas9-VP64 versus when co-transfected with a1_gRNA and dCas9. Error bars indicate the standard error of the mean for three independent biological replicates.

ing in crisprTFs targeting either the template or the non-template strand of DNA. In Figure 3-4, we observe that CrisprTFs exhibit greater activation potential while targeting the non-template strand DNA as opposed to the template strand. The observed strand bias gives us an additional degree of tunability in building synthetic transcriptional networks and while activating endogenous genes for basic biology and therapeutic applications.

3.2.2 Inducibility

In many applications, control of the activity of a transcription factor by an inducer (e.g., a small molecule) is desired. With crisprTFs, one viable strategy is to constitu-

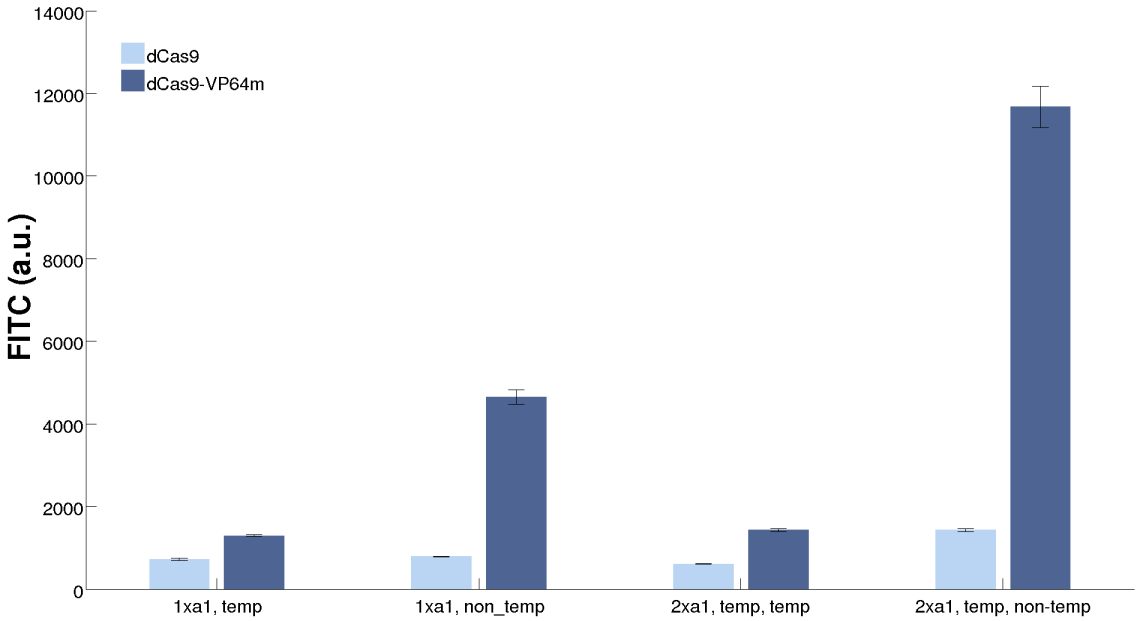


Figure 3-4: Strand bias in activation by crisprTFs. Different combinations of forward and reverse oriented 1× and 2× a1_gRNA operator sites were introduced upstream of pMLPm resulting in crisprTFs targeting the template or the non-template DNA strand. When co-transfected with crisprTFs, we observe a higher activation potential for the non-template strand operators compared with the template ones. Errors bars indicate the standard error of the mean for three independent biological replicates.

tively express the protein component of the system (i.e., dCas9) and then modulate the amount of gRNA available for binding to dCas9 and thus, the activity achieved at the target DNA. To test this strategy, we constructed the following two inducible gRNA promoters as shown in Figure 3-5-A:

1. Doxycycline (Dox) inducible RNAP III H1 promoter built by placing 2×TetR operator sites in the H1 promoter and constitutively expressing TetR in the system as described previously [74].

2. Isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible U6 promoter built by placing 3× LacI operator sites in the U6 promoter and constitutively expressing LacI in the system as described previously [82].

We co-transfected HEK293T cells with the 3×(a1_op)_pMLPm_*yfp* reporter plasmid along with plasmids expressing crisprTF, TetR and a1_gRNA under the inducible H1 promoter and plated the cells in the absence or presence of Dox. In

Figure 3-5-B, we plotted the mean fluorescence values measured for YFP expression in different conditions and we observe a $>$ 20-fold Dox inducible activation of *yfp* expression. Similarly, when we co-transfect 3×(a1_op)_pMLPm_*yfp* reporter plasmid along with plasmids expressing crisprTF, LacI and a1_gRNA under the inducible U6 promoter, we observe $\sim$ 10-fold activation of *yfp* expression. We demonstrate that sophisticated regulatory motifs, such as small-molecule responsive modules can be built for crisprTFs, thus enabling external control of crisprTF-based transcriptional circuits. Such synthetic modules can be interfaced with other regulatory elements to achieve more complex regulation for synthetic biology.

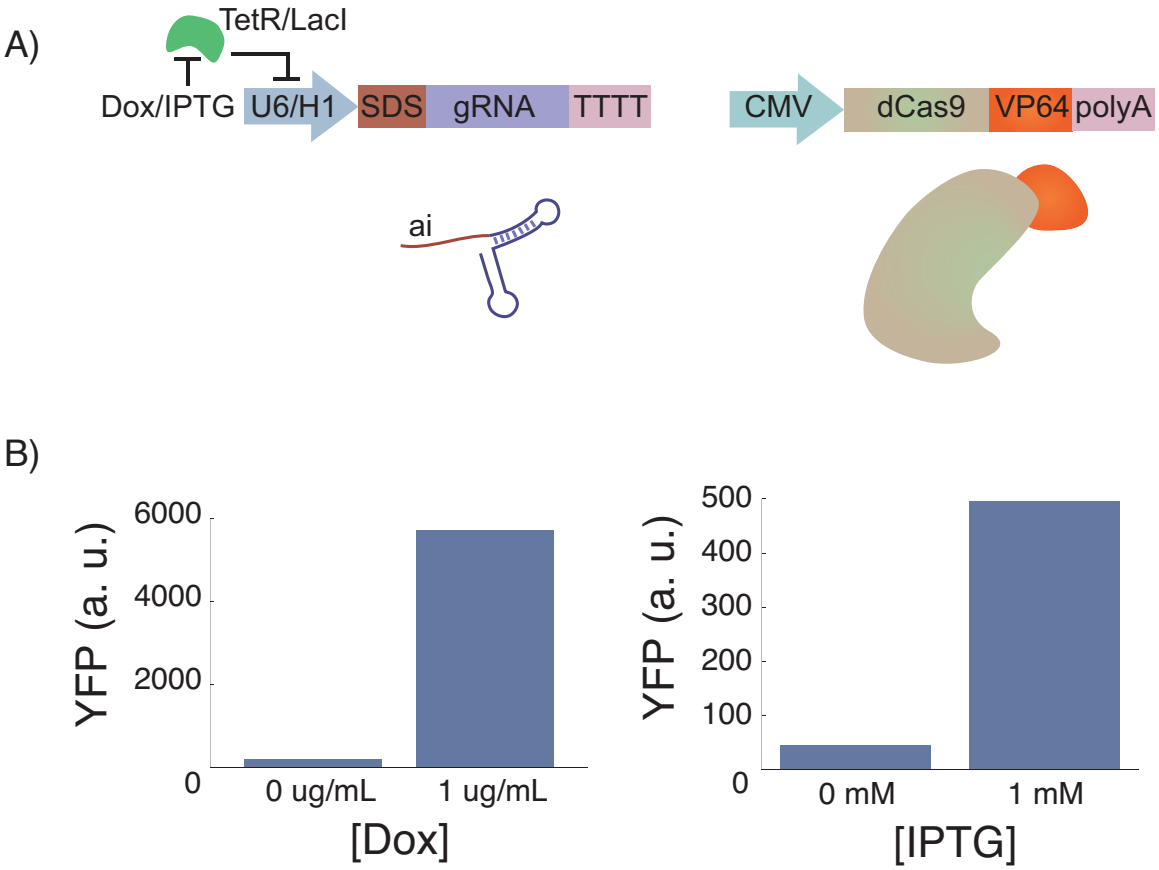


Figure 3-5: Inducible activation of target genes by crisprTFs in HEK293T cells. A) Doxycycline and IPTG inducible modified H1 and U6 promoter architecture is shown respectively. CrisprTF is constitutively expressed under CMVp. B) When the 3×(a1_op)_pMLPm_*yfp* plasmid is co-transfected with the plasmids expressing crisprTF, a1_gRNA and TetR or LacI, we observe an inducible activation of *yfp* expression. The bars represent the mean of the YFP levels of 100,000 individual cells.

3.2.3 Orthogonality

To further demonstrate the potential of crisprTFs toward constructing synthetic promoters and gRNAs that are orthogonal with respect to each other, we tested three randomly designed gRNAs (a1, a2 and a3 gRNAs) for their ability to activate each other's target sequences. As shown in Figure 3-6, each of the gRNAs exhibited high activity at their cognate target sequences but low activity at non-cognate sequences.

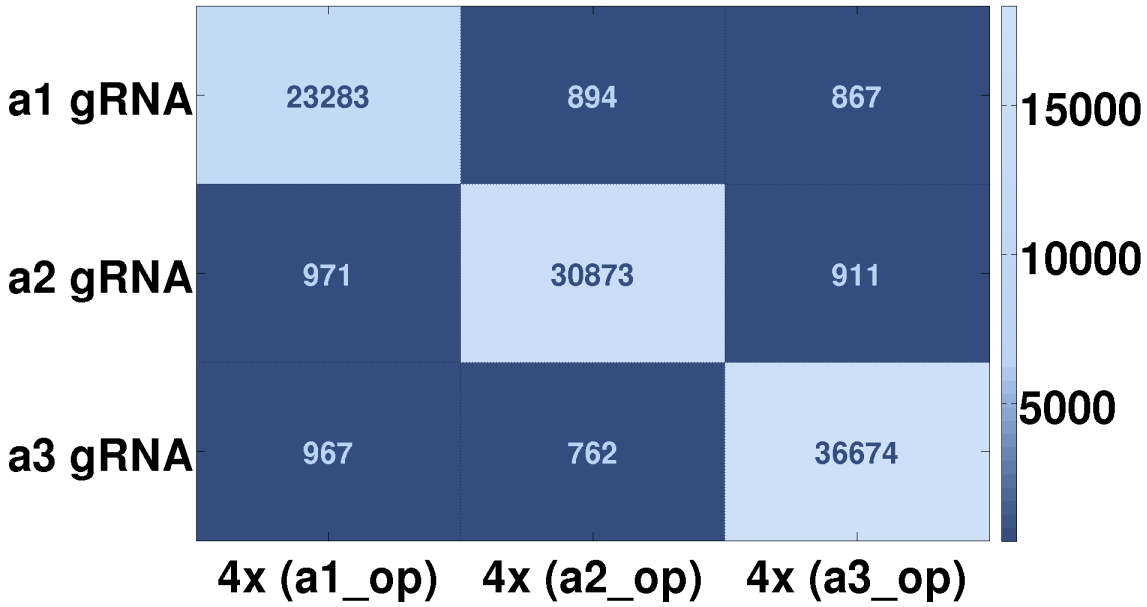


Figure 3-6: Constructing Orthogonal crisprTF responsive promoters. Heat map illustrating the orthogonality of crisprTFs in human cells. Plasmids encoding three orthogonal gRNAs (a1, a2, and a3 gRNAs) were co-transfected into HEK293T cells along with one of the three reporter plasmids (each encoding 4× operator sites for a given gRNA) upstream of pMLPm promoter driving *yfp* expression. Only cognate interactions between gRNAs and target binding sites resulted in significant activation. The standard error of the mean of YFP fluorescence for three independent biological replicates is indicated in each cell of the heat-map plot.

These results suggest that one can construct synthetic promoters and gRNAs that are orthogonal with respect to each other and to the host genome. The ease of design and expression of customized gRNAs in comparison to ZFs and TALEs make CRISPR-based transcription factors appealing as synthetic TFs for modulating endogenous gene expression as well as for synthetic biology. The ability to customize the target site of dCas9 via the expression of short gRNAs obviates the need to engineer multiple

orthogonal DBDs in order to construct complex transcriptional circuits. This could potentially reduce the overall metabolic burden on cells and enable the integration of more complex synthetic computation and logic within living cells.

3.3 CrisprTF based synthetic gene circuits

More complex regulatory and logic circuits, such as cascades and complex digital logics gates can be built by layering crisprTFs we have built so far. The possibility of integrating multiple inputs at a single promoter expands the regulatory potential and provides us with increased flexibility that can be leveraged while designing synthetic transcriptional networks or rewiring endogenous pathways. Furthermore, since both activation and repression functions can be achieved with crisprTFs, the crisprTF platform may be advantageous compared to RNAi-based gene regulatory platforms where only repression can be achieved.

3.3.1 'A' and 'B' boolean logic in single cells

To demonstrate the multiplexing ability of crisprTFs, we designed a synthetic gene circuit that employs multiple gRNAs in single cells. In Figure 3-7, we built and characterized a circuit that uses two different gRNAs, a3 and g5, to independently activate two different promoters and drive the expression of *yfp* and *mKate* respectively. The reporter plasmids are comprised of following architectures: 4×(a3_op)_pMLPm_*yfp* and 4×(g5_op)_pMLPm_*mKate*, which are similar to the ones we used earlier. However, instead of co-transfecting a plasmid encoding dCas9-VP64, we used a stable cell line derived from HEK293T cells that constitutively expresses dCas9-VP64 (HEK293T-taCas9). For each different scenario in Figure 3-7-B, we co-transfected both the reporter plasmids and one of four two-wise combinations of a non-specific gRNA, a3 gRNA and g5 gRNA as indicated.

Although the circuits have a robust performance as indicated by flow cytometry data in Figure 3-7-B, the microscopy data reveals that not all individual cells express YFP and mKate simultaneously when transfected with both a3 and g5 gRNAs. We

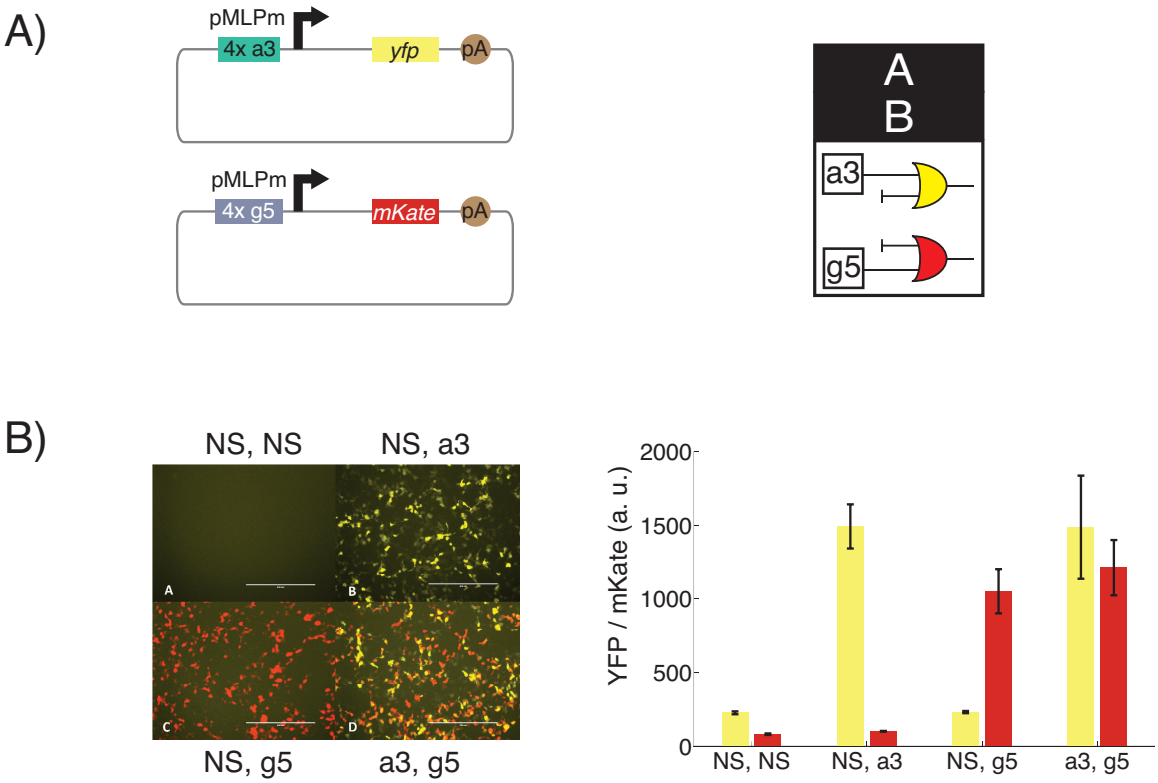


Figure 3-7: Orthogonal crisprTF responsive promoters implementing two 2-input logic gates 'A' and 'B' in single cells. Two-wise combinations of plasmids encoding orthogonal gRNAs (a3 and g5 gRNAs) and a non-specific (NS) gRNA were co-transfected into HEK293T-taCas9 cell line that constitutively expresses dCas9-VP64. Two reporter plasmids were also co-transfected - The first reporter plasmid encoded 4× operator sites for a3 gRNA upstream of pMLPm promoter driving *yfp* expression and the second reporter plasmid encoded 4× operator sites for g5 gRNA upstream of pMLPm promoter driving *mKate* expression. The circuit schematic implements two 2-input logic gates 'A' and 'B' in single cells as demonstrated by fluorescent microscopy and multicolor flow cytometry. Error bars indicate the standard error of mean of YFP fluorescence and mKate fluorescence respectively for three independent biological replicates.

believe this is because of sub-optimal transfection efficiencies while co-transfecting four different plasmids, resulting in some cells not receiving all four plasmids.

3.3.2 'NOT A' and 'B' boolean logic in single cells

For our next circuit, we decided to take advantage of the fact that crisprTFs can affect both activation and repression of gene expression depending on the location of the DNA target site with respect to the promoter. As depicted in Figure 3-8-A, we

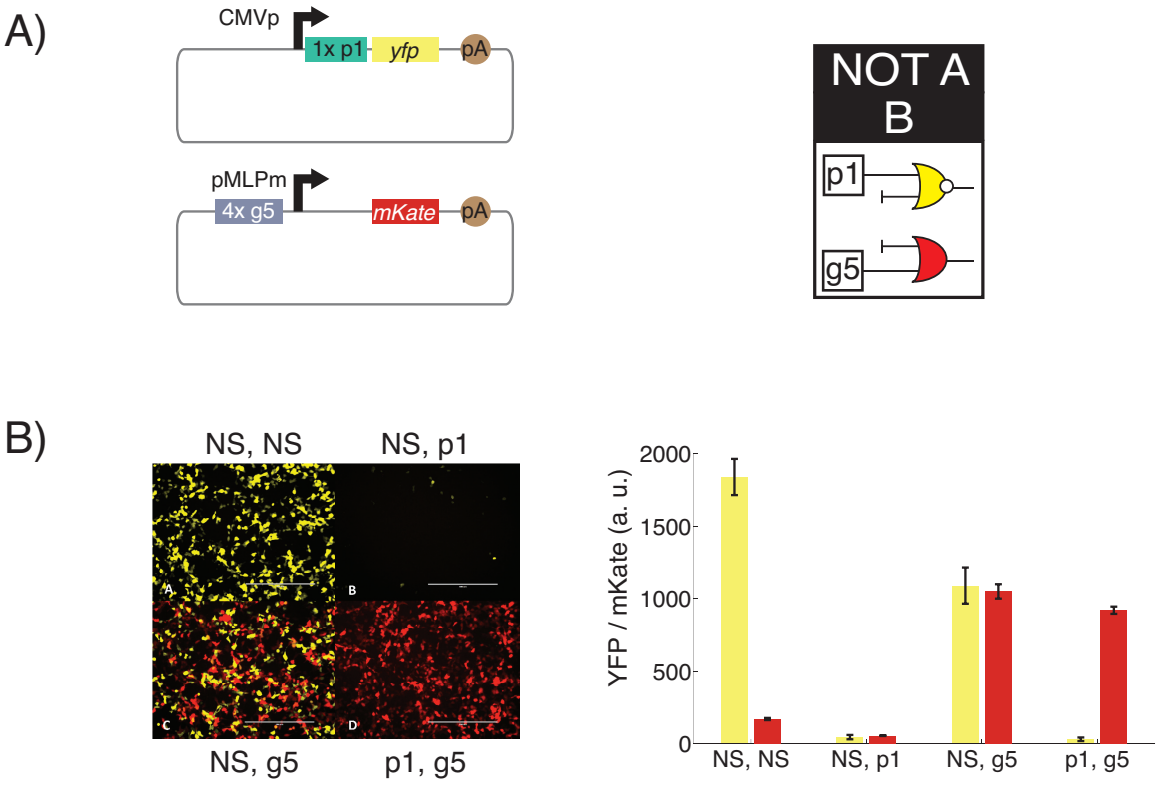


Figure 3-8: Orthogonal crisprTF responsive promoters implementing two 2-input logic gates 'NOT A' and 'B' in single cells. Two-wise combinations of plasmids encoding orthogonal gRNAs (a3 and g5 gRNAs) and a non-specific (NS) gRNA were co-transfected into HEK293T-taCas9 cell line that constitutively expresses dCas9-VP64. Two reporter plasmids were also co-transfected - The first reporter plasmid encoded an operator site for p1 gRNA downstream CMV promoter driving *yfp* expression and the second reporter plasmid encoded 4× operator sites for g5 gRNA upstream of pMLPm promoter driving *mKate* expression. The circuit schematic implements two 2-input logic gates 'NOT A' and 'B' in single cells as demonstrated by fluorescent microscopy and multicolor flow cytometry. Error bars indicate the standard error of mean of YFP fluorescence and mKate fluorescence respectively for three independent biological replicates.

designed a p1 gRNA that targets downstream of the native CMV promoter (CMVp) that drives the expression of *yfp*. Although CMVp constitutively expresses YFP, when co-transfected with p1 gRNA in to HEK293-taCas9 cells, we found that CMVp activity is thoroughly silenced (> 50-fold) by p1 gRNA guided crisprTF. By co-transfecting with a second reporter plasmid encoding 4× operator sites for g5 gRNA upstream of pMLPm promoter driving *mKate* expression and with plasmids encoding respective gRNAs, we were able to demonstrate the simultaneous implementation of

two different logic gates 'NOT A' and 'B' in single cells. The circuits also have a robust performance with > 10-fold difference between the ON and the OFF states as measured by multicolor flow cytometry ref., Figure 3-8-B.

3.3.3 'A XOR B' boolean logic in single cells

Given its complexity, 'A XOR B' logic is considered to be one of the most challenging logics to implement in single cells [127, 17]. We leverage crisprTF's ability to simultaneously active and repress gene expression and design a compact architecture that encodes XOR logic. As depicted in Figure 3-9-A, we build two symmetric promoters with each of the gRNAs a2 and a3 activating one of the promoters while repressing the other. To realize such a design, we introduce 4× operator sites for a2 gRNA upstream and a 1× operator site for a3 gRNA downstream of pMLPm promoter driving the expression of *yfp*. Similar but inverted design is used for the second promoter in which we introduce 4× operator sites for a3 gRNA upstream and a 1× operator site for a2 gRNA downstream of pMLPm promoter. Both of the promoters are designed to drive *yfp* expression.

Our design results in a simple but efficient implementation of the XOR logic in single cells. Although the circuit does have > 10-fold difference between the ON and the OFF states, the ON state when a3 gRNA is expressed is lower than the one when a2 gRNA is expressed. Further tuning of individual gRNAs to match their affinities to the respective target DNA sequences and/or increasing the number of upstream gRNA operator sites should result in better performance. Given the inefficiencies associated with DNA delivery inherent to transient transfections, moving on to a Lentivirus based, stable integration platform should also provide additional boosts in logic circuit performance.

Unlike ZFs or TALENs, a single dCas9-VP64 protein shared across multiple different gRNAs, each targeting a different DNA sequence can be used as a versatile transcription factor. Moreover, the complexity of the circuits measured in terms of the unique number of parts employed can be much smaller when both the activating and the repressing ability of crisprTFs is leveraged, as is evident with our XOR design

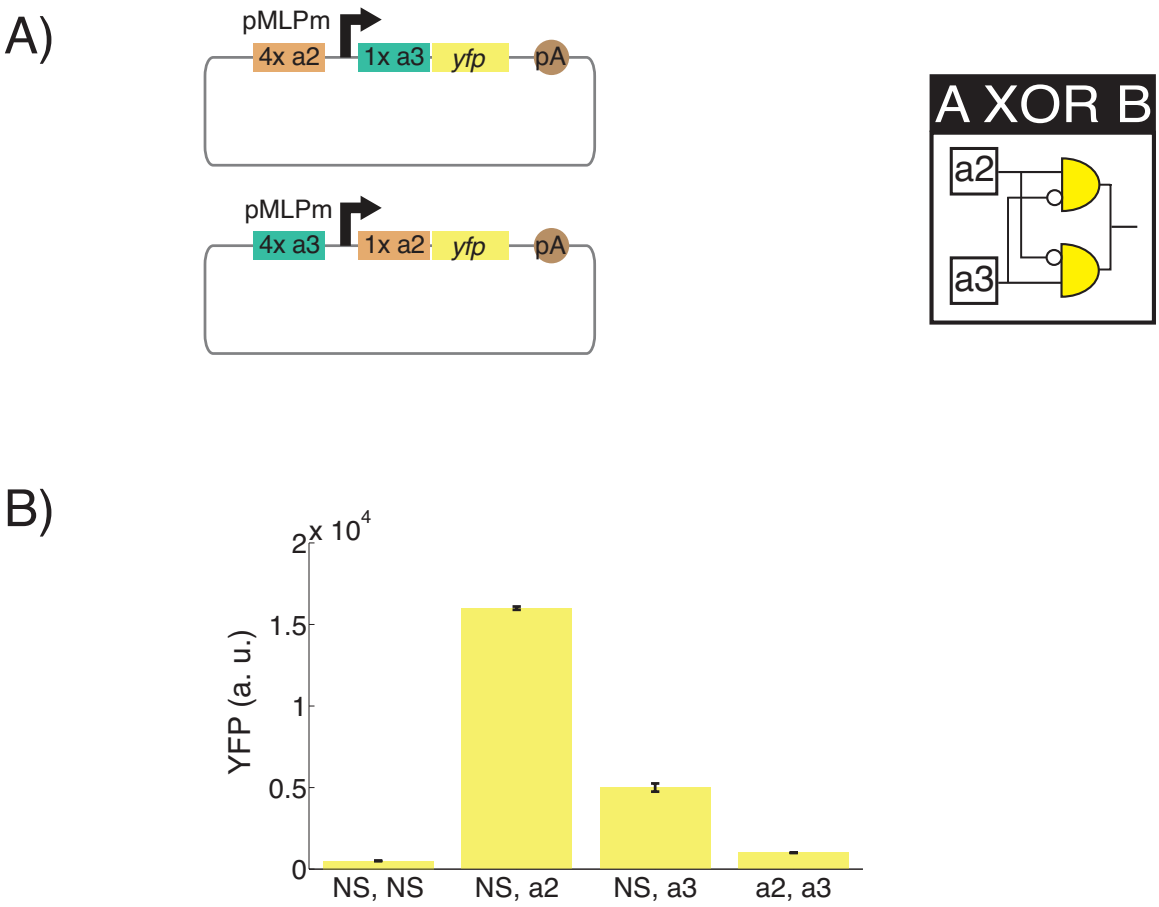


Figure 3-9: Orthogonal crisprTF responsive promoters implementing 'A XOR B' logic in single cells. Two-wise combinations of plasmids encoding orthogonal gRNAs (a2 and g3 gRNAs) and a non-specific (NS) gRNA were co-transfected into HEK293T-taCas9 cell line that constitutively expresses dCas9-VP64. Two reporter plasmids were also co-transfected. The first reporter plasmid encoded a 4× operator sites for a2 gRNA upstream and 1× operator site for a3 gRNA downstream of pMLPm promoter driving *yfp* expression. The second reporter plasmid encoded 4× operator sites for a3 gRNA upstream and 1× operator site for a2 gRNA downstream of pMLPm promoter driving *mKate* expression. The circuit schematic implements a 2-input logic gate 'A XOR B' in single cells as demonstrated by multicolor flow cytometry. Error bars indicate the standard error of mean of YFP fluorescence and mKate fluorescence respectively for three independent biological replicates.

implemented above. The ease with which each of the above boolean logic gates could be implemented demonstrates the potential crisprTFs have towards enabling complex synthetic transcriptional network design and for probing endogenous networks.

3.4 Methods

3.4.1 Plasmid construction

To construct the mammalian dCas9-VP64 expressing plasmid, we first introduced D10A and H841A mutations into hCas9 [97] (Addgene, Plasmid #41815). Then, three repeats of SV40 Nuclear Localization Signal (3xNLS) were fused to the C-terminus of the mutated hCas9 using a PCR-based assembly protocol. Using a multipart Gibson assembly protocol, the immediate-early promoter of cytomegalovirus (pCMV), dCas9-3xNLS, VP64, and SV40 polyA terminator were cloned into the NotI site of the pG5-Luc plasmid (Promega). To monitor successfully transfected cells by flow cytometry, we replaced the original luciferase gene in pG5-Luc with mKATE (Evrogen). The resulting pPGK1-mKATE cassette served as a constitutive fluorescent protein control that was used to gate for the presence of the crisprTF-expressing plasmid with flow cytometry in Figure 3-1.

The gRNA expression plasmids were constructed by cloning the 138 bp human U6 promoter (an RNA-polymerase-III-dependent promoter [84]), along with the gRNA handle and terminator into a plasmid containing pPGK1-eBFP2 flanked by the SV40 polyA terminator (a gift from Lior Nissim, Lu lab). A SacI site was placed at the 3'-end of the U6 promoter to enable the cloning of different specificity determining sequences for each gRNA. The reporters were assembled into the gRNA-expressing plasmid through a one-step Gibson assembly reaction, where the upstream polyadenylation (pA) signal and transcriptional pause site from pG5-Luc, along with a 41 bp, minimal adenovirus type 2 major late promoter (pMLPm), mYFP, and HSV polyA signal were cloned into the AatII site of the gRNA-expressing plasmids.

For the synthetic promoter experiments, additional gRNA operator sites were cloned in the NheI site upstream of the pMLPm promoter (see Table A.2). For the repression experiments, dCas9-KRAB was constructed by cloning a 366 bp KRAB domain to the C-terminus of dCas9. GCCACC was used as the Kozak sequence for the expression of dCas9-VP64, mYFP, eBFP2, and mKATE. Unless directly targeted by gRNAs for repression assays or used a reporter in building logic gate circuits,

the mKATE fluorescent protein on the crisprTF-expression plasmid and the eBFP2 fluorescent protein on the reporter/gRNA plasmid served as our gating controls for flow cytometry analysis.

More information on the plasmid DNA sequences can be found in Table A.2 in the appendix.

3.4.2 Cell lines

To build the HEK293T-taCas9 cell line, a lentiviral (LV) plasmid encoding dCas9-VP64-p2a-puroR was constructed by cloning in amplified dCas9-VP64 and p2a-puroR in to the BamHI and EcoRI sites of the 3rd generation lentiviral plasmid - FUGw [89] (Addgene #14883) via Gibson assembly. LV particles were produced by transfecting 200,000 HEK293T cells with 1 ug of lentivial plasmid encoding dCas9-VP64-p2a-puroR and 0.5 ug of pCMV-VSV-G (Addgene #8454) and 0.5 ug of pCMV-dR8.2 (Addgene #8455). The cell culture supernatant containing LV particles was collected 48 hrs post transfection, filtered with a 0.2 uM Cellulose acetate filter and was used to infect HEK293T cells supplemented with 8 ug/mL polybrene. Successfully transduced cells were obtained by selection with puromycin at 3 ug/mL for four days.

3.4.3 Cell culture

HEK293T kidney epithelial cells and the derived cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 1% glutamine, and Non-essential amino acids. Cells were grown under 5% CO2 at 37C. Continuous culture of HEK293T and derived cells was maintained by passaging cells at 80-100% confluency. 0.05% Trypsin (diluted 1/5 in PBS from the standard 0.25% Trypsin) was used for passaging HEK293T and derived cells. Cells were transfected with Fugene-HD transfection reagent (Promega) using the protocol below and assayed for gene expression with flow cytometry or fluorescent microscopy at 48 h post transfection.

3.4.4 Transfections

Promega's Fugene-HD was used as our transfection reagent. Fresh growth medium was added to cells already growing in logarithmic phase, four to six hours prior to transfection. The cells (in a 10 cm dish) to be transfected were ensured to be ~ 80 to 100% confluent at the time of transfection. At the time of transfection, cells were washed with warm PBS, trypsinized with 0.05% trypsin and pipetted up and down using a 10 mL stripette to dissociate cells from cell-cell adhesions. The cells were then collected in a 50 mL falcon tube, spun down at 2000 rpm for 5 min and resuspend in ~ 8 mL of warm, fresh medium to achieve a density of $1,000,000$ cells/mL.

Once the DNA, opti-MEM and transfection reagent were allowed to arrive at room temperature, the following mixture was prepared in a 1.5 mL tube for a six well transfection: 80 uL opti-MEM + 20 uL total DNA (The DNA is normalized to be at 100 ng/uL). 10 uL of Fugene-HD at room temperature was added to the above mixture and the mixture was vortexed, briefly spun down and left undisturbed at room temperature for 15 to 20 min to allow DNA/lipid complex formation. After 15 to 20 min of complex formation, 200 uL of cells (200,000 cells from a 1,000,000 cells/mL suspension) were added to the complexed mixture, pipetted up and down and let sit for two minutes. All of the cell + complex mixture was plated in a well (of a six well plate) and incubated at 37C, 5% CO_2. The transfection reagent + DNA mixture was never left undisturbed for more than 20 minutes as this would result in larger complexes and will lower the transfection efficiency.

3.5 Acknowledgements and Copyright

This work was supported by the Defense Advanced Research Projects Agency, the National Institutes of Health New Innovator Award (1DP2OD008435), and the National Science Foundation (1124247).

The above work has been reprinted from **Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas**, Fahim Farzadfard, Samuel D. Perli, and Timothy K. Lu, ACS Synthetic Biology, 2013, 2 (10), pp 604-613, with permission from ACS.

Chapter 4

Integrating CRISPR-Cas with Mammalian RNA biology

In Chapter 3, we discussed how the Cas9 protein and the guide RNA (gRNA) from type II CRISPR/Cas system can be adapted to achieve programmable DNA binding without requiring complex protein engineering. In this system, the sequence specificity of the Cas9 DNA-binding protein is determined by guide RNAs (gRNAs) that have base-pairing complementarity to DNA target sites. This enables simple and highly flexible programing of Cas9 binding. Cas9's nuclease activity has been adapted for precise and efficient genome editing in prokaryotic and eukaryotic cells [29, 67, 68, 69, 97]. A mutant derivative of this protein (dCas9), which lacks nuclease activity, was modified to enable programmable transcriptional regulation of both ectopic and native promoters to create CRISPR-based transcription factors (CRISPR-TFs) in mammalian cells [25, 42, 51, 90, 96, 107]. Type III CRISPR/Cas systems (RNA-targeting Cas proteins) have also been adapted for synthetic biology applications [113]. For example, the type III CRISPR/Cas-associated Csy4 protein from *Pseudomonas aeruginosa* has been used in bacteria to achieve predictable regulation of multigene operons by cleaving precursor mRNAs. The functionality of Csy4 has also been demonstrated in bacteria, archaea, and eukaryotes [113].

CRISPR-TFs can enable the construction of large-scale synthetic gene circuits and the rewiring of natural regulatory networks. This is due to the ease of defining

new, orthogonal transcriptional regulators by designing artificial gRNAs. However, up until now, gRNAs in human cells have only been expressed from RNA polymerase III (RNAP III) promoters, presumably since RNAs expressed from most RNA polymerase II (RNAP II) promoters are expected to be exported to the cytoplasm, while gRNAs and Cas9 need to interact with DNA in the nucleus. Because RNAP III promoters comprise only a small portion of cellular promoters and are mostly constitutively active [105], this is an important limitation for programming CRISPR/Cas activity for conditional gene regulation and genome engineering. For example, conditional regulatory systems in which gRNA production is linked to tissue-specific [24], temporally controllable [49] and/or inducible expression systems [75] cannot be readily constructed with RNAP III promoters, whereas many such systems that utilize RNAP II promoters exist. Another limitation of existing CRISPR/Cas regulatory schemes is that multiple gRNAs are typically needed to efficiently activate endogenous promoters [25, 90, 97, 107], but strategies for multiplexed gRNA production from single transcripts have not yet been described. As a result, multiple gRNA expression constructs are currently needed to perturb natural transcriptional networks, thus limiting scalability.

In addition to transcriptional regulation, natural circuits leverage RNA-based posttranscriptional regulation to achieve complex behaviors [3, 23, 87, 99, 144]. Short hairpin RNAs (shRNAs) and microRNAs (miRNAs) have been expressed from both RNAP II and RNAP III promoters and can be embedded in 3' UTRs or as introns [52, 124, 133]. Moreover, multiple miRNAs can be expressed concomitantly from the same transcript [124], and synthetic circuits that use shRNA- and miRNA-based regulation in mammalian systems have been built [52, 146]. Thus, gene regulatory strategies based on combining RNA engineering with transcriptional and posttranscriptional regulation would be useful in studying and modeling natural systems or implementing artificial behaviors.

In this Chapter, we integrate mammalian and bacterial RNA-based regulatory mechanisms to create complex synthetic circuit topologies and to regulate endogenous promoters. We combine multiple mammalian RNA processing strategies, including 3'

RNA triple helixes (triplexes) [145], introns, and ribozymes with mammalian miRNA regulation, bacteria-derived CRISPR-TFs, and the Csy4 RNA-modifying protein. We use these tools to generate functional gRNAs from RNAP-II-regulated mRNAs in human cells for activating both synthetic and endogenous promoters, while rendering the concomitant translation of the harboring mRNAs tunable. Furthermore, we develop strategies for multiplexed gRNA production that enable compact encoding of proteins and multiple gRNAs in single transcripts. To demonstrate the utility of these tools, we implemented multistage transcriptional cascades and combined mammalian miRNA-based regulation with CRISPR-TFs to create multicomponent genetic circuits whose feedback loops, interconnections, and behaviors could be rewired by Csy4-based RNA processing. Thus, integration of CRISPR-TFs with mammalian RNA regulatory architectures can be used to construct complex, synchronized, and switchable networks using synthetic transcriptional regulation and RNA-processing mechanisms.

4.1 Strategies to express gRNAs from RNAP II promoters

An important first step to enabling complex CRISPR-TF-based circuits is to generate functional gRNAs from RNAP II promoters in human cells, which would allow for the coupling of gRNA production to specific regulatory signals. For example, the activation of gRNA-dependent circuits could be conditionally initiated in defined cell types or states or tuned in response to external inputs with specific RNAP II promoters. Furthermore, the ability to simultaneously express gRNAs along with proteins from a single transcript would allow multiple outputs, including effector proteins and regulatory links, to be produced from a concise genetic architecture. Thus, we sought to simultaneously generate functional gRNAs and proteins from RNAP II promoters.

4.1.1 Functional gRNA Generation with an RNA Triple Helix and Csy4

We first utilized the RNA-binding and RNA endonuclease capabilities of Csy4 [58, 132] to release gRNAs from transcripts that also encode functional protein sequences, generated by RNAP II promoters. Csy4 recognizes a 28 nt RNA sequence (hereafter referred to as the "28" sequence), cleaves the RNA, and remains bound to the upstream RNA fragment [58]. Specifically, we used the CMV promoter (CMVp) to express a gRNA (gRNA1), flanked by two Csy4 binding sites, downstream of an *mKate2* coding region as shown in Figure 4-1-A. In this architecture, RNA cleavage by Csy4 releases a functional gRNA but also removes the poly-(A) tail from the upstream mRNA, which is known to result in impaired translation of most eukaryotic mRNAs [111].

To complement the Csy4-mediated loss of the poly(A) tail, we cloned a 110 bp fragment derived from the 3' end of the mouse MALAT1 locus [145] downstream of *mKate2* and immediately upstream of the gRNA sequence flanked by Csy4 recognition sites. This sequence forms a highly conserved 3' triple helical structure known to stabilize transcripts lacking a poly-(A) tail and enables their translation [145]. Thus, our final "triplex/Csy" architecture was a CMVp-driven *mKate2* transcript with a 3' triplex sequence followed by a 28-gRNA-28 sequence (CMVp-*mK*-Tr-28-gRNA-28) 4-1.

To characterize gRNA activity, we co-transfected HEK293T cells with the CMVp-*mK*-Tr-28-gRNA1-28 expression plasmid, along with a plasmid encoding a synthetic P1 promoter that is specifically activated by the gRNA1/taCas9 complex to express EYFP (see taCas9 definition below) [42]. In this experiment and all the following ones (unless otherwise indicated), we also co-transfected a plasmid expressing a transcriptionally active dCas9-NLS-VP64 protein (taCas9), which consists of dCas9 fused to a SV40 nuclear localization signal (NLS) [73] and the viral VP64 transcription activation domain, which efficiently recruits RNA Pol II to initiate transcription [9]. We also transfected HEK293T cells with 0–400 ng of a Csy4-producing plasmid (where Csy4 was produced by the murine PGK1 promoter) along with 1 μg of the other

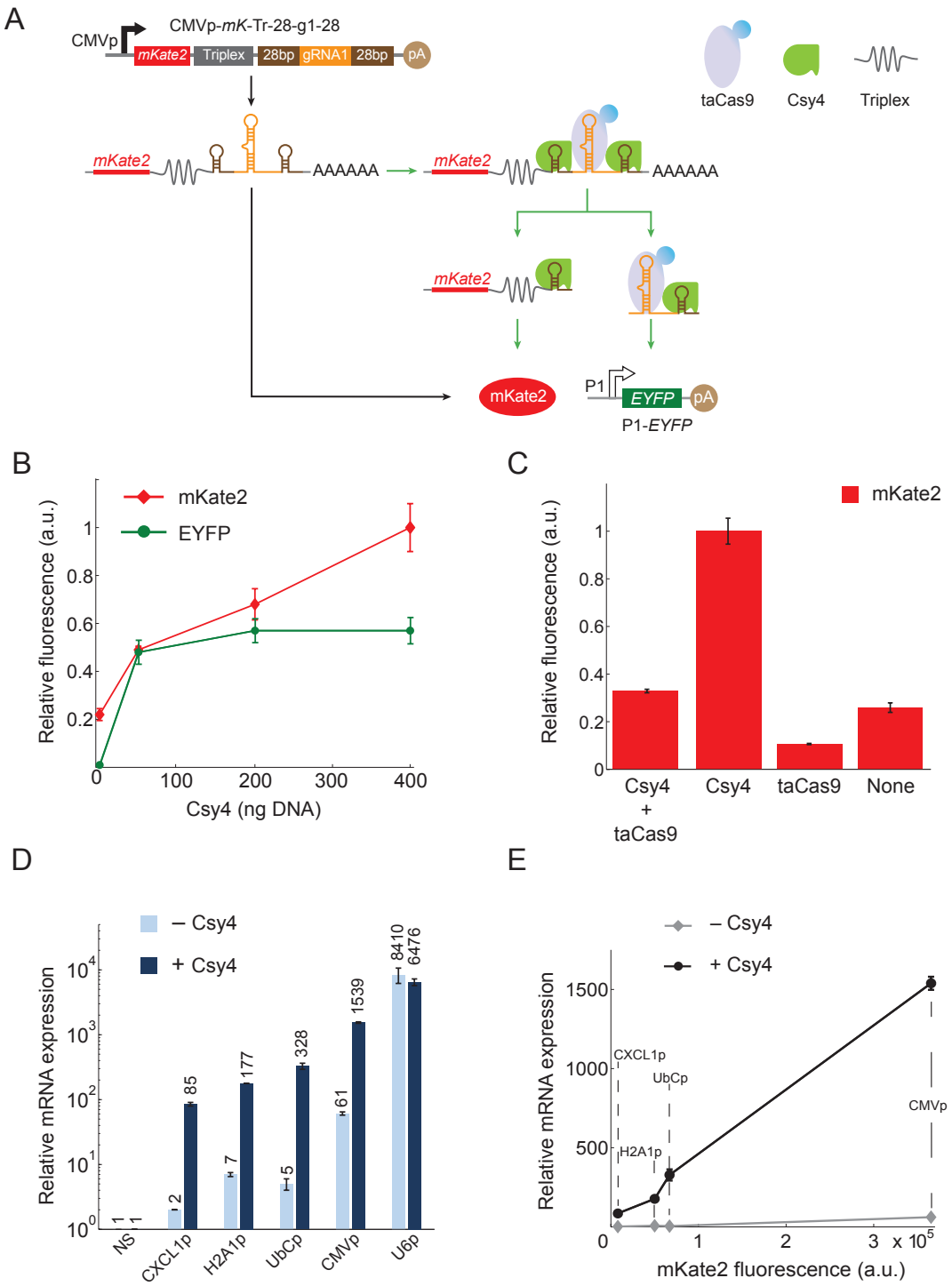


Figure 4-1: The "Triplex/Csy4" Architecture (CMVp-*mK*-Tr-28-g1-28) Produces Functional gRNAs from RNAP II Promoters while Maintaining Expression of the Harboring Gene.

Figure 4-1: (A) gRNA1 was flanked by two Csy4 recognition sites ("28"), placed downstream of an *mKate2* gene followed by an RNA triplex, and produced by CMVp. Csy4 generates gRNAs that can be incorporated into a transcriptionally active dCas9-VP64 (taCas9) to activate a synthetic promoter (P1) driving EYFP (P1-*EYFP*).

(B) The presence of Csy4 enabled a 60-fold increase in EYFP levels, validating the generation of functional gRNAs. Fluorescence values were normalized to the maximum respective fluorescence between the data in this figure and Figures 2B-2D to enable cross-comparisons between the "triplex/Csy4" and "intron/Csy4" architectures.

(C) Csy4 and taCas9 have opposite effects on mKate2 fluorescence generated by CMVp-*mK*-Tr-28-g1-28. The mKate2 fluorescence levels were normalized to the maximum mKate2 value observed (Csy4 only) across the four conditions tested here.

(D) The human RNAP II promoters CXCL1, H2A1, and UbC and the viral CMVp were used to drive expression of four different gRNAs (gRNA3-gRNA6, Table A.3) previously shown to activate the IL1RN promoter [107] from the "triplex/Csy4" construct. These results were compared to the RNAP III promoter U6p driving direct expression of the same gRNAs. Four different plasmids, each containing one of the indicated promoters and gRNAs 3-6, were co-transfected along with a plasmid encoding taCas9 and with or without a plasmid expressing Csy4. Relative IL1RN mRNA expression, compared to a control construct with nonspecific gRNA (NS, CMVp-*mK*-Tr-28-g1-28), was monitored using qRT-PCR.

(E) The input-output transfer curve for the activation of the endogenous IL1RN loci by the "triplex/Csy4" construct was determined by plotting the mKate2 levels (as a proxy for the input) versus the relative IL1RN mRNA expression levels (as the output). Tunable modulation of endogenous loci can be achieved with RNAP II promoters of different strengths, with the presence of Csy4 greatly increasing activation compared with the absence of Csy4. The IL1RN data is the same as shown in (D). Data are represented as mean ± SEM. See also Figures B-1 and B-2

plasmids (Figures 4-1-B and B-1-A, for raw data).

Increasing Csy4 plasmid concentrations increased mKate2 levels by up to 5-fold (Figure 4-1-B). Furthermore, functional gRNAs generated from this construct induced EYFP expression by up to 60-fold from the P1 promoter. While mKate2 expression continued to increase with the concentration of the Csy4-producing plasmid, EYFP activation plateaued after 50 ng of the Csy4-producing plasmid. Examination of

cell cultures with microscopy showed visual evidence of Csy4-mediated cytotoxicity (roughly ∼20% cell death) at 400 ng Csy4 plasmid concentrations while minimal cytotoxicity was observed at 0–200 ng Csy4 plasmid concentrations. Thus, we used 100–200 ng of the Csy4 plasmid in subsequent experiments (except where otherwise noted), although this reduced the number of Csy4-positive cells after transfection. This issue could be addressed in future work by using weaker promoters to regulate Csy4 expression or by generating stable Csy4-producing cell lines.

We characterized the relative effects of Csy4 and taCas9 in this architecture on the levels of the gRNA-harboring protein by measuring mKate2 fluorescence in the presence of Csy4 and taCas9, Csy4 alone, taCas9 alone, or neither protein (Figures 4-1-C and B-2-A). The lowest mKate2 fluorescence resulted from the taCas9-only condition. Since we used a taCas9 with a strong NLS, this effect could have been mediated by taCas9 binding to the gRNA within the mRNA and localizing the transcript to the nucleus, thus inhibiting translation of the harboring mKate2. This hypothesis is supported by data demonstrating that endogenous promoters can be activated by gRNAs produced from the "triplex/Csy4"-based architecture even in the absence of Csy4 (see below and Figures 4-1-D and E). The highest mKate2 levels were obtained with Csy4 alone, suggesting that Csy4 processing could enhance mKate2 levels. The levels of mKate2 in the absence of both Csy4 and taCas9, as well as in the presence of both Csy4 and taCas9, were similar and were reduced by 3- to 4-fold compared with Csy4 only. Thus, titrating Csy4 and taCas9 levels can tune the input-output relationship of CRISPR/Cas-based circuits.

To validate the robustness of the "triplex/Csy4" architecture, we adapted it to regulate the expression of a native genomic target in human cells. We targeted the endogenous IL1RN locus for gene activation via the coexpression of four distinct previously described gRNAs: gRNA3, gRNA4, gRNA5, and gRNA6 (Table A.3) [107]. The IL1RN gene cluster encodes the expression of IL-1Ra protein, which is a modulator of the immune response and has been shown to be beneficial for treating inflammatory diseases [33].

We designed each of the four gRNAs to be expressed concomitantly with mKate2,

each from a separate plasmid. Each set of four gRNAs was regulated by one of the following promoters: the CMVp, human Ubiquitin C (UbCp), human Histone H2A1 (H2A1p) [115], and human inflammatory chemokine CXCL1 (CXCL1p) promoters [140]. H2A1p and CXCL1p have been shown to be deregulated in a malignant transformation model [100]. As a control, we also used the RNAP III promoter U6 (U6p) to drive expression of the four gRNAs. For each promoter tested, four plasmids encoding the four different gRNAs were cotransfected along with plasmids expressing taCas9 and Csy4. As a negative control, we substituted the IL1RN-targeting gRNA expression plasmids with plasmids that expressed gRNA1, which is nonspecific for the IL1RN promoter (Figure 4-1-D, "NS").

We used qRT-PCR to quantify the mRNA levels of the endogenous IL1RN gene, with the results normalized to the nonspecific control. With the four gRNAs regulated by U6p, IL1RN activation levels were increased by 8,410-fold in the absence of Csy4 and 6,476-fold with 100 ng of the Csy4-producing plasmid (Figure 4-1-D, "U6p"). IL1RN activation with gRNAs expressed from the CMVp was substantial (Figure 4-1-D, "CMVp"), with 61-fold enhancement in the absence of Csy4 and 1,539-fold enhancement in the presence of Csy4. The human RNAP II promoters generated $\sim$ 2- to 7-fold activation in the absence of Csy4 and $\sim$ 85- to 328-fold activation in the presence of Csy4 (Figure 4-1-D; "CXCL1p,"H2A1p," and "UbCp"). Strikingly, IL1RN activation observed in the absence of Csy4 suggests that taCas9 can utilize gRNAs encoded in long nonprocessed RNA transcripts, albeit with much lower efficiency than Csy4-processed gRNAs.

To characterize the input-output transfer function for endogenous gene regulation, we used mKate2 fluorescence generated by each promoter as a marker of input promoter activity for the various RNAP II promoters (Figure 4-1-E). These data indicate that IL1RN activation was not saturated in the conditions tested and that a large dynamic range of endogenous gene regulation can be achieved with human RNAP II promoters. Thus, tunable modulation of native genes can be achieved using CRISPR-TFs with gRNAs expressed from the "triplex/Csy4" architecture.

4.1.2 Functional gRNA Generation from Introns with Csy4

Complementary to the "triplex/Csy4" architecture, we developed an alternative strategy for generating functional gRNAs from RNAP II promoters by encoding gRNAs within an intronic sequence of a gene. Such a strategy has been used previously to generate synthetic siRNAs in mammalian cells [52]. Specifically, we encoded a gRNA as an intron within the mKate2 coding sequence using previously described "consensus" acceptor, donor, and branching sequences [129, 135]. We expected that, once spliced, the gRNA would associate with taCas9 to activate a cognate synthetic promoter regulating EYFP. However, this simple architecture resulted in undetectable EYFP levels (FigureB-2-B, bottom panel). These data are consistent with previous studies highlighting that the lifetime of most intronic RNAs is typically only a few minutes [3, 28]. Thus, we concluded that without any stabilization, intronic gRNAs would be expected to be rapidly degraded.

We tried two different approaches to stabilize intronic gRNAs. First, we used intronic sequences that have been reported to produce long-lived introns. This included constructs such as the HSV-1 latency-associated intron, which forms a stable circular intron [15], and the sno-lncRNA2 (snoRNA2) intron, which is processed on both ends by the snoRNA machinery, which protects it from degradation [148]. However, these approaches also resulted in undetectable activation of the target promoter (data not shown).

As an alternative strategy, we sought to stabilize intronic gRNAs by flanking the gRNA cassette with two Csy4 recognition sites (Figure 4-2). In this model, spliced gRNA-containing introns should be bound by Csy4, which should release functional gRNAs. However, Csy4 can also potentially bind and digest the pre-mRNA before splicing occurs. In this case, functional gRNA would be produced, but the mKate2-containing pre-mRNA would be degraded in the process (Figure 4-2-A). Thus, increased Csy4 concentrations are expected to result in decreased mKate2 levels but higher levels of functional gRNA. The decrease in mKate2 levels and the increase in functional gRNA levels with Csy4 concentrations could be expected to

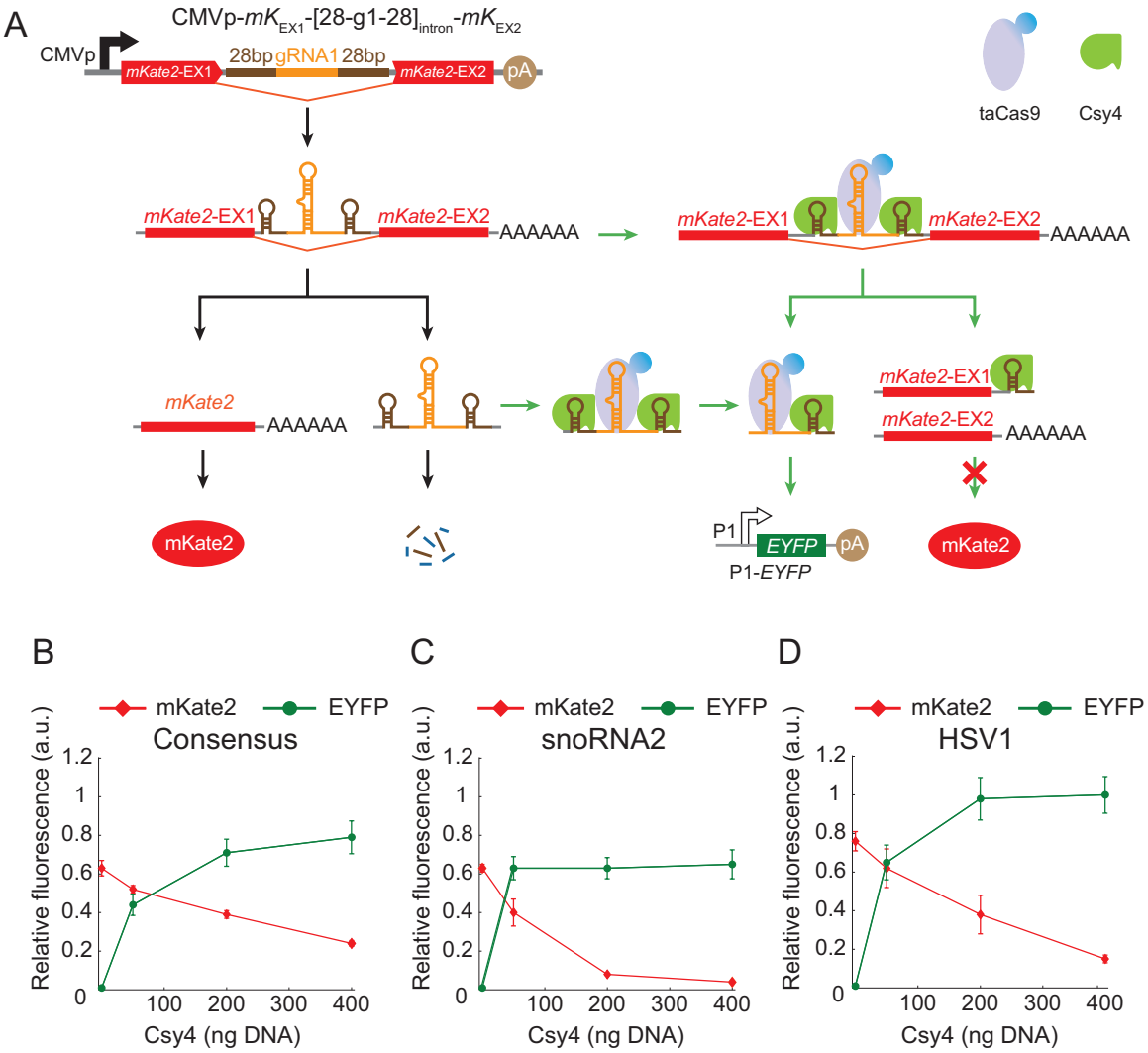


Figure 4-2: The "Intron/Csy4" Architecture (CMVp-mK_{EX1}-[28-g1-28]$_{\text{intron}}$-mK_{EX2}) Generates Functional gRNAs from Introns in RNAP-II-Expressed Transcripts

(A) gRNA1 is flanked by Csy4 recognition sites and encoded within an intron, leading to functional gRNA1 generation with Csy4 and activation of a downstream P1-*EYFP* construct. In contrast to the "triplex/Csy4" construct in Figure 4-1, the "intron/Csy4" architecture results in decreased expression of the harboring gene with increased Csy4 levels, which may be due to cleavage of pre-mRNA prior to splicing.

(B–D) Three introns – a consensus intron (B), snoRNA2 intron (C), and an HSV1 intron (D) – combined with Csy4 resulted in functional gRNAs as assessed by EYFP expression. Fluorescence values were normalized to the maximum fluorescence between these data and Figure 4-1-B. Data are represented as mean ±SEM. See also FiguresB-1-B-3.

depend on several factors, which are illustrated in Figure 4-2 (black arrows, Csy4-independent processes; green arrows, Csy4-mediated processes). These competing factors include the rate at which Csy4 binds to its target sites and cleaves the RNA, the rate of splicing, and the rate of spliced gRNA degradation in the absence of Csy4. To examine the behavior of the "intron/Csy4" architecture, we used CMVp to drive expression of mKate2 harboring HSV1, snoRNA, and consensus introns containing gRNA1 flanked by two Csy4 binding sites (CMVp-mK_{EX1}-[28-g1-28]$_{\text{intron}}$-mK_{EX2}) along with a synthetic P1 promoter regulating the expression of EYFP (Figure 4-2-A).

4.1.3 Functional gRNA Generation with Cis-Acting Ribozymes

Very recently, the generation of gRNAs from RNAP II promoters for genome editing was demonstrated in wheat [138] and yeast [47] with cis-acting ribozymes. In addition to the "triplex/Csy4" and "intron/Csy4" based mechanisms described above, we employed self-cleaving ribozymes to enable gene regulation in human cells via gRNAs generated from RNAP II promoters. Specifically, our gRNAs were engineered to contain a hammerhead (HH) ribozyme [109] on their 5' end and a HDV ribozyme [43] on their 3' end (Figure 4-3).

We tested ribozymes in three different architectures for their ability to generate gRNA1 to activate EYFP from P1-*EYFP* and maintain mKate2 expression: (1) an mKate2 transcript followed by a triplex and a HH-gRNA1-HDV sequence (CMVp-*mK*-Tr-HH-g1-HDV, Figure 4-3-A); (2) an mKate2 transcript followed a HH-gRNA1-HDV sequence (CMVp-*mK*-HH-g1-HDV, Figure 4-3-B); and (3) the sequence HH-gRNA1-HDV with no associated protein-coding gene (CMVp-HH-g1-HDV, Figure 4-3-C). We compared gRNAs generated from these architectures with gRNAs produced by U6p and the "triplex/Cys4" architecture (with 200 ng of the Csy4 plasmid).

All constructs that contained mKate2 exhibited detectable mKate2 fluorescence (Figures 4-3-D and B-4). Surprisingly, this included CMVp-*mK*-HH-g1-HDV, which did not have a triplex sequence and was thus expected to have low mKate2 levels due to removal of the poly-(A) tail. Our observations could result from inefficient

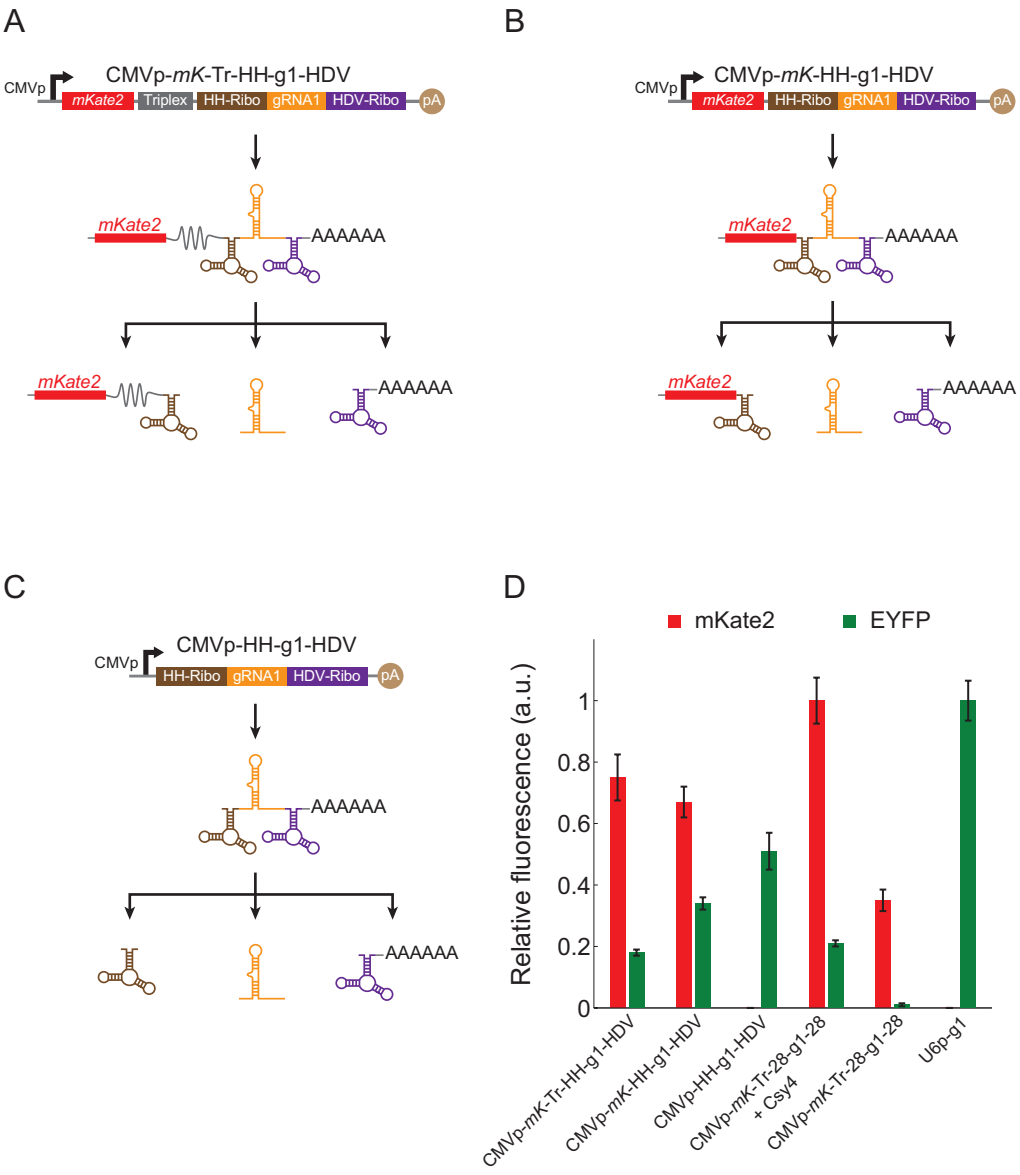


Figure 4-3: Ribozyme Architectures Expressed from CMVp Can Produce Active gRNAs

(A) gRNA1 was flanked with hammerhead (HH) and HDV ribozymes and encoded downstream of mKate2 with an RNA triplex (CMVp-*mK*-Tr-HH-g1-HDV).

(B) gRNA1 was flanked with HH and HDV ribozymes and encoded downstream of mKate2 with no RNA triplex (mK-HH-g1-HDV).

(C) gRNA1 was flanked with HH and HDV ribozymes (HH-g1-HDV).

(D) The three architectures efficiently generated gRNA1 to activate P1-*EYFP*. The "triplex/Csy4" construct (CMVp-*mK*-Tr-28-g1-28), with and without Csy4, and the RNAP III promoter U6p driving gRNA1 (U6p-g1) are shown for comparison. Data are represented as mean ±SEM. See also FigureB-4.

ribozyme cleavage [26] that allows non-processed transcripts to be transported to the cytoplasm and translated, protection of the mKate2 transcript by the residual 3' ribozyme sequence, or other mechanisms yet to be determined. In terms of target EYFP activation, the highest EYFP fluorescence level was generated from gRNAs expressed by U6p, followed by the CMVp-HH-g1-HDV and CMVp-*mK*-HH-g1-HDV constructs (Figure 4-3-D). The CMVp-*mK*-Tr-HH-gRNA1-HDV and "triplex/Csy4" architectures resulted in similar EYFP levels. Thus, cis-acting ribozymes can mediate functional gRNA expression from RNAP II promoters.

4.2 Multiplexed gRNA Expression from Single RNA Transcripts

A major challenge in constructing CRISPR-TF-based circuits in human cells, especially ones that interface with endogenous promoters, is that multiple gRNAs are necessary to achieve desired activation levels, since single gRNAs do not typically achieve significant activation [25, 90, 96, 107]. Current techniques rely on multiple gRNA expression cassettes, each with their own promoter and terminator. Our toolkit can generate multiple functional gRNAs from a single compact transcript.

We expressed two independent gRNAs from a single RNA transcript to activate two independent downstream promoters. In the first architecture ("intron-triplex"), we encoded gRNA1 within an HSV1 intron flanked by two Csy4 binding sites within the coding sequence of mKate2. Furthermore, we encoded gRNA2 enclosed by two Csy4 binding sites downstream of the mKate2-triplex sequence (Figure 4A, "Input A"). In the second architecture ("triplex-tandem"), we surrounded both gRNA1 and gRNA2 with Csy4 binding sites and placed them in tandem, downstream of the mKate2-triplex sequence (Figure 4-4-B, "Input B"). In both architectures, gRNA1 and gRNA2 targeted the synthetic promoters P1-*EYFP* and P2-*ECFP*, respectively.

Both strategies for multiplexed gRNA expression enable functional taCas9 activity at multiple downstream targets and can be tuned for desired applications (Figures 4-

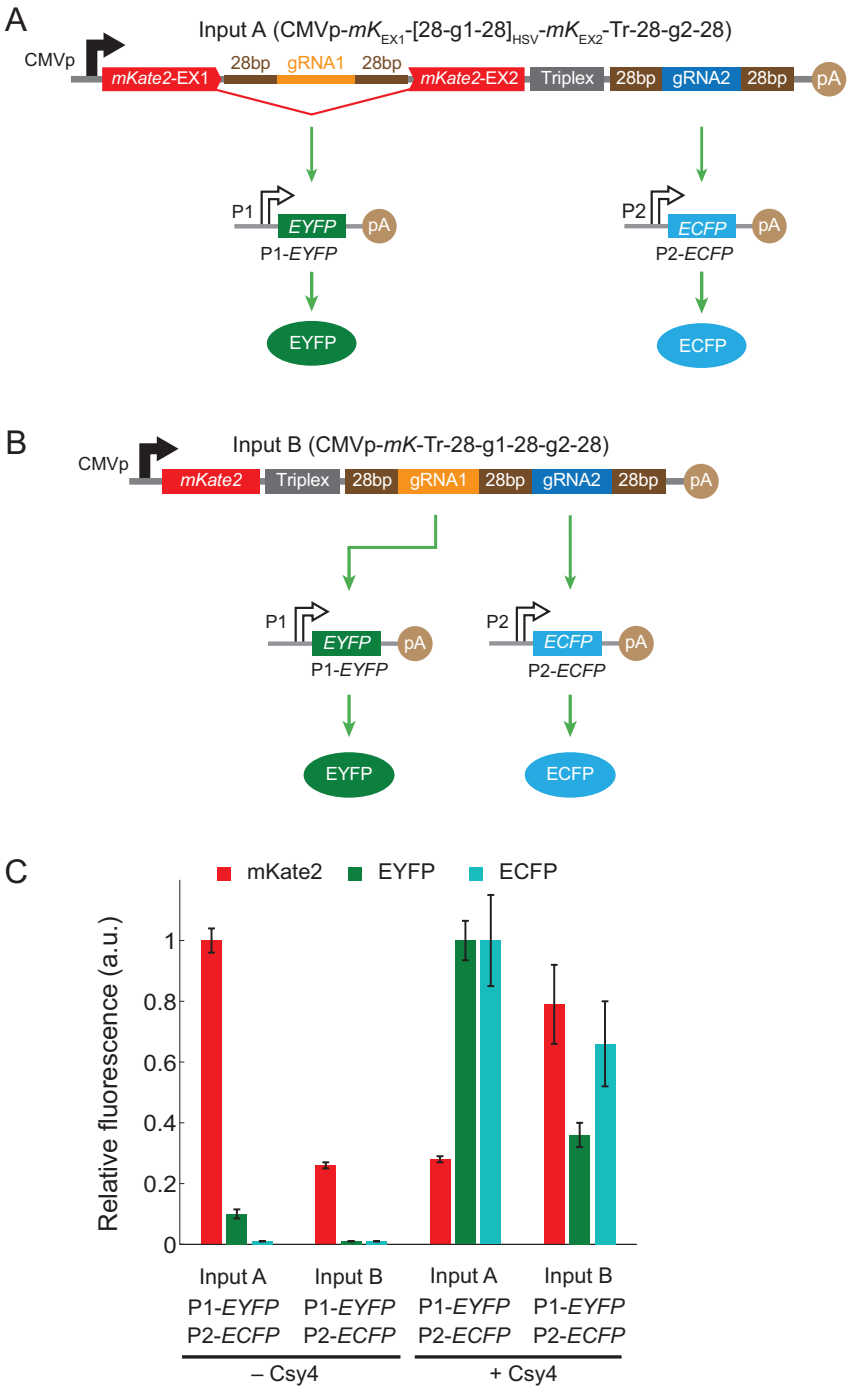


Figure 4-4: The "Triplex/Csy4" and "Intron/Csy4" Architectures Enable Multiplexed gRNA Expression from a Single Transcript and Compact Encoding of Synthetic Circuits with Multiple Outputs

Figure 4-4: (A) In the first design (Input A, "intron-triplex"), we encoded gRNA1 within a HSV1 intron and gRNA2 after an RNA triplex. Both gRNAs were flanked by Csy4 recognition sites. Functional gRNA expression was assessed by activation of a gRNA1-specific P1-*EYFP* construct and a gRNA2-specific P2-*ECFP* construct.

(B) In the second design (Input B, "triplex-tandem"), we encoded both gRNA1 and gRNA2 in tandem, with intervening and flanking Csy4 recognition sites, downstream of mKate2 and an RNA triplex. Functional gRNA expression was assessed by activation of a gRNA1-specific P1-*EYFP* construct and a gRNA2-specific P2-*ECFP* construct.

(C) Both multiplexed gRNA expression constructs efficiently activated EYFP and ECFP expression in the presence of Csy4, thus demonstrating the generation of multiple active gRNAs from a single transcript. Data are represented as mean ±SEM. See also Figure B-5.

4-C and B-5). Specifically, the "intron-triplex" construct exhibited a 3-fold decrease in mKate2, a 10-fold increase in EYFP, and a 100-fold increase in ECFP in the presence of 200 ng of the Csy4 plasmid compared to no Csy4. In the "triplex-tandem" architecture, mKate2, EYFP, and ECFP expression increased by 3-fold, 36-fold, and 66-fold, respectively, in the presence of 200 ng of the Csy4 plasmid when compared to no Csy4. The "intron-triplex" architecture had higher EYFP and ECFP levels compared with the "triplex-tandem" architecture. Thus, Csy4 enables the generation of multiple gRNAs from a single mRNA transcript.

To demonstrate the utility of our multiplexing toolkit in targeting endogenous loci, we encoded multiple gRNAs targeting the IL1RN promoter into a single transcript. Previously described gRNAs for IL1RN activation, separated by Csy4 binding sites, were cloned in tandem, downstream of an mKate2-triplex sequence on a single transcript (Figure 4-5). We measured IL1RN activation by the multiplexed single-transcript construct as well as an architecture where the four different gRNAs were expressed from four different plasmids (Figure 4-5-B, "Multiplexed" versus "Non multiplexed", respectively). In the presence of 100 ng of the Csy4 plasmid, the multiplexed architecture resulted in a ∼1,111-fold activation over nonspecific gRNA1 ("NS") compared with ∼461-fold for the non-multiplexed set of single-gRNA-expressing plasmids.

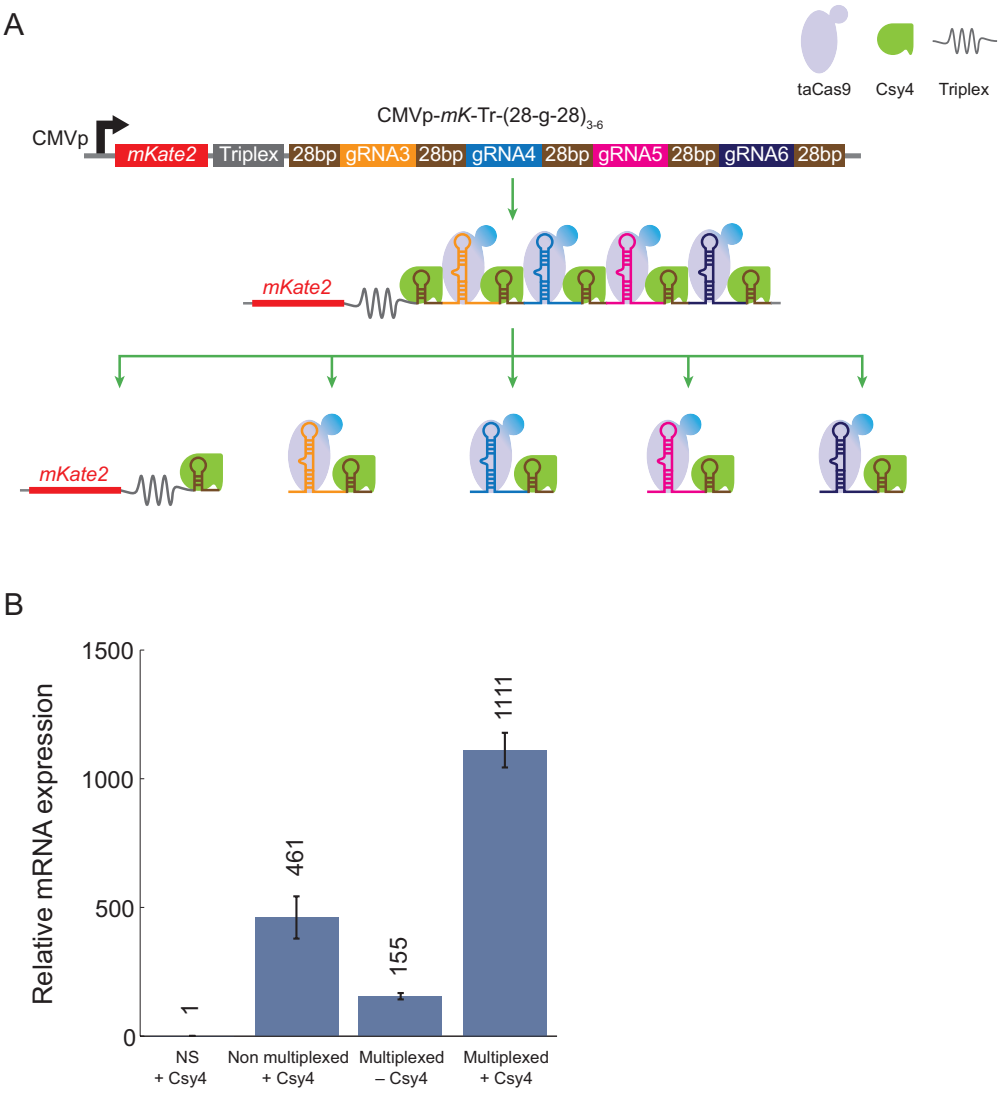


Figure 4-5: Multiplexed gRNA Expression from a Single, Compact Transcript Enables Efficient Activation of Endogenous Loci.
(A) Four different gRNAs (gRNA3–gRNA6) were multiplex-encoded in tandem, with intervening and flanking Csy4 recognition sites, downstream of *mKate2* and an RNA triplex (CMVp-*mK*-Tr-(28-g-28)$_{3-6}$).
(B) The multiplexed *mK*-Tr-(28-g-28)$_{3-6}$ construct exhibited high-level activation of IL1RN expression in the presence of Csy4 compared with the same construct in the absence of Csy4. Relative *IL1RN* mRNA expression was determined based on a control construct with nonspecific gRNA1 (NS, CMVp-*mK*-Tr-28-g1-28) expressed via the "triplex/Csy4" architecture. For comparison, a non-multiplexed set of plasmids containing the same gRNAs (gRNA3–gRNA6), each produced from separate, individual plasmids (CMVp-*mK*-Tr-28-gRNA-28) with the "triplex/Csy4" architecture is shown. Data are represented as mean ±SEM.

Moreover, ∼155-fold IL1RN activation was detected with the multiplexed architecture even in the absence of Csy4, further supporting our observation that taCas9 can utilize gRNAs encoded in long transcripts, albeit with significantly lower efficiency. Together with Figure 4-4, these results show that it is possible to generate multiple functional gRNAs for multiplexed expression from a single, concise RNA transcript. These architectures thus enable compact programming of Cas9 function for implementing multi-output synthetic gene circuits, for modulating endogenous genes, and potentially for achieving conditional multiplexed genome editing.

4.3 Synthetic Transcriptional Cascades with RNA-Guided Regulation

To demonstrate the utility of our RNA-dependent regulatory toolkit, we used it to create the first CRISPR-TF-based transcriptional cascades. Many complex gene circuits require the ability to implement cascades, in which signals integrated at one stage are transmitted to downstream stages for processing and actuation [36, 60]. Furthermore, transcriptional cascades are important in natural regulatory systems, such as those that control segmentation, sexual commitment, and development [32, 126]. However, despite the potential utility of CRISPR-TFs for artificial gene circuits, CRISPR-TF-based cascades have not been built to date.

We integrated the "triplex/Csy4" and "intron/Csy4" strategies to build two different, three-stage CRISPR-TF-mediated transcriptional cascades (Figure 4-6). In the first design, CMVp-driven expression of gRNA1 from an "intron/Csy4" construct generated gRNA1 from an HSV1 intron, which activated a synthetic promoter P1 to produce gRNA2 from a "triplex/Csy4" architecture, which then activated a downstream synthetic promoter P2 regulating *ECFP* (Figure 4-6). In the second design, the intronic gRNA expression cassette in the first stage of the cascade was replaced by a "triplex/Csy4" architecture for expressing gRNA1 (Figure 4-6-B). We tested these two designs in the presence of 200 ng of the Csy4 plasmid (Figures 4-6-C, D and B-6).

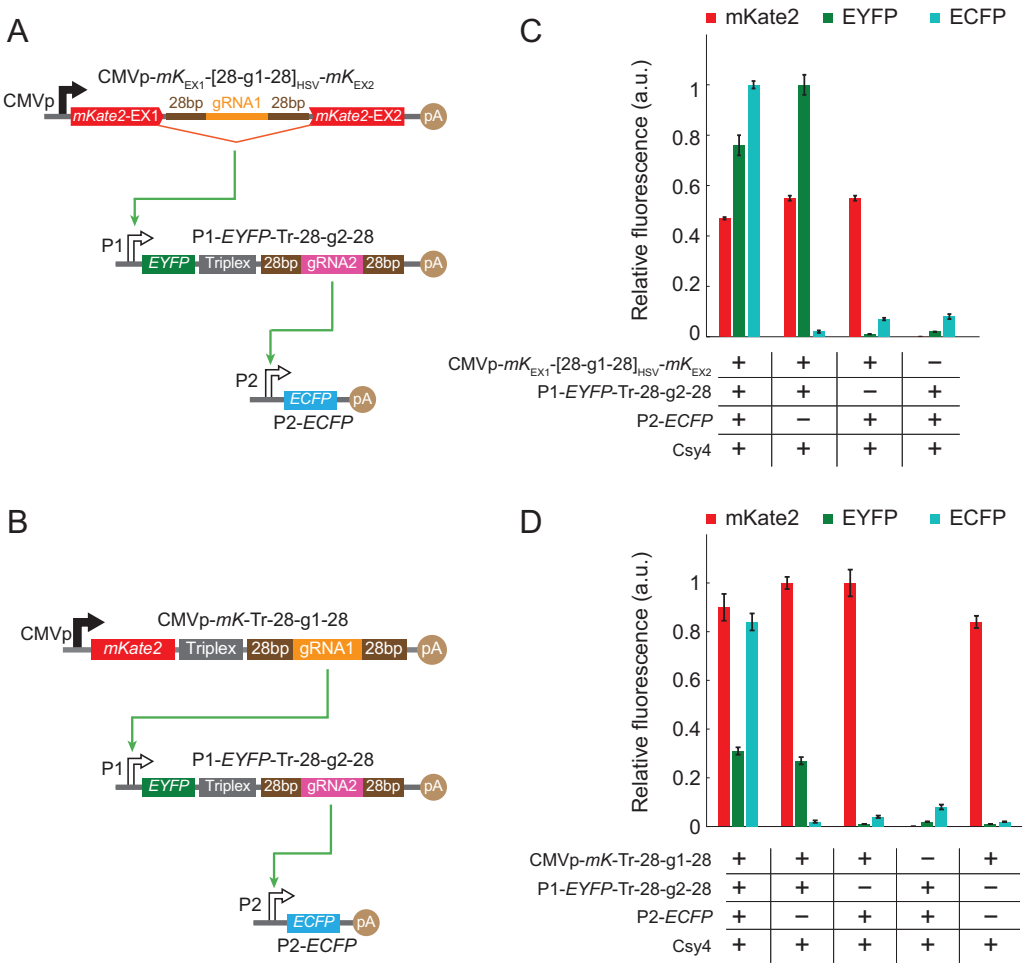


Figure 4-6: Multistage Transcriptional Cascades Can Be Implemented with Our CRISPR-TF Architectures

(A) A three-stage transcriptional cascade was implemented by using intronic gRNA1 (CMVp-mK_{EX1}-[28-g1-28]$_{HSV}$-mK_{EX2}) as the first stage. gRNA1 specifically targeted the P1 promoter to express gRNA2 (P1-*EYFP*-Tr-28-g2-28), which then activated expression of ECFP from the P2 promoter (P2-*ECFP*).

(B) A three-stage cascade was implemented by using a "triplex/Csy4" architecture to express gRNA1 (CMVp-*mK*-Tr-28-g1-28). gRNA1 specifically targeted the P1 promoter to express gRNA2 (P1-*EYFP*-Tr-28-g2-28), which then activated expression of *ECFP* from P2 (P2-*ECFP*).

(C and D) The complete three-stage cascade from (A) and (B), respectively, exhibited expression of all three fluorescent proteins. The removal of one of each of the three stages in the cascade resulted in the expected loss of fluorescence of the specific stage and dependent downstream stages. The control condition in column 4 in (C) and (D) are duplicated, since the two circuits in (A) and (B) were tested in the same experiment. Data are represented as mean ±SEM. See also Figure B-6.

In the first cascade design, a 76-fold increase in EYFP and a 13-fold increase in ECFP were observed compared to a control in which the second stage of the cascade (P1-*EYFP*-Tr-28-g2-28) was replaced by an empty plasmid (Figure 4-6-C). In the second cascade design, a 31-fold increase in EYFP and a 21-fold increase in ECFP were observed compared to a control in which the second stage of the cascade (P1-*EYFP*-Tr-28-g2-28) was replaced by an empty plasmid (Figure 4-6-D). These results demonstrate that there is minimal nonspecific activation of promoter P2 by gRNA1 (see also FigureB-2-C), which is essential for the scalability and reliability of transcriptional cascades. Furthermore, the activation of each stage in the cascade was dependent on the presence of all upstream nodes, which is expected in properly functioning transcriptional cascades (Figures 4-6-C and D).

4.4 Rewiring RNA-Dependent Synthetic Regulatory Circuits

We integrated CRISPR-TF regulation with mammalian RNAi to implement more sophisticated circuit topologies and to show how network motifs could be rewired by Csy4-based RNA processing. Specifically, we incorporated miRNA regulation with CRISPR-TFs and used Csy4 to disrupt miRNA inhibition of target RNAs by removing cognate miRNA binding sites. We first built a single RNA transcript that was capable of multiplexed gene regulation by expressing both a functional miRNA [52, 146] and a functional gRNA. This was achieved by encoding a mammalian miRNA inside the consensus intron within the mKate2 gene, followed by a triplex sequence and a gRNA1 sequence flanked by Csy4 recognition sites (Figure 4-7-A, CMVp-mK_{EX1}-[miR]-mK_{EX2}-Tr-28-g1-28). The first output was a constitutively expressed *ECFP* gene followed by a triplex sequence, a Csy4 recognition site, 8× miRNA binding sites (8× miRNA-BS), and another Csy4 recognition site (Figure 4-7-A). The second output was the P1 promoter regulating *EYFP* (Figure 4-7-A).

In the absence of Csy4, ECFP and EYFP levels were low because the miRNA

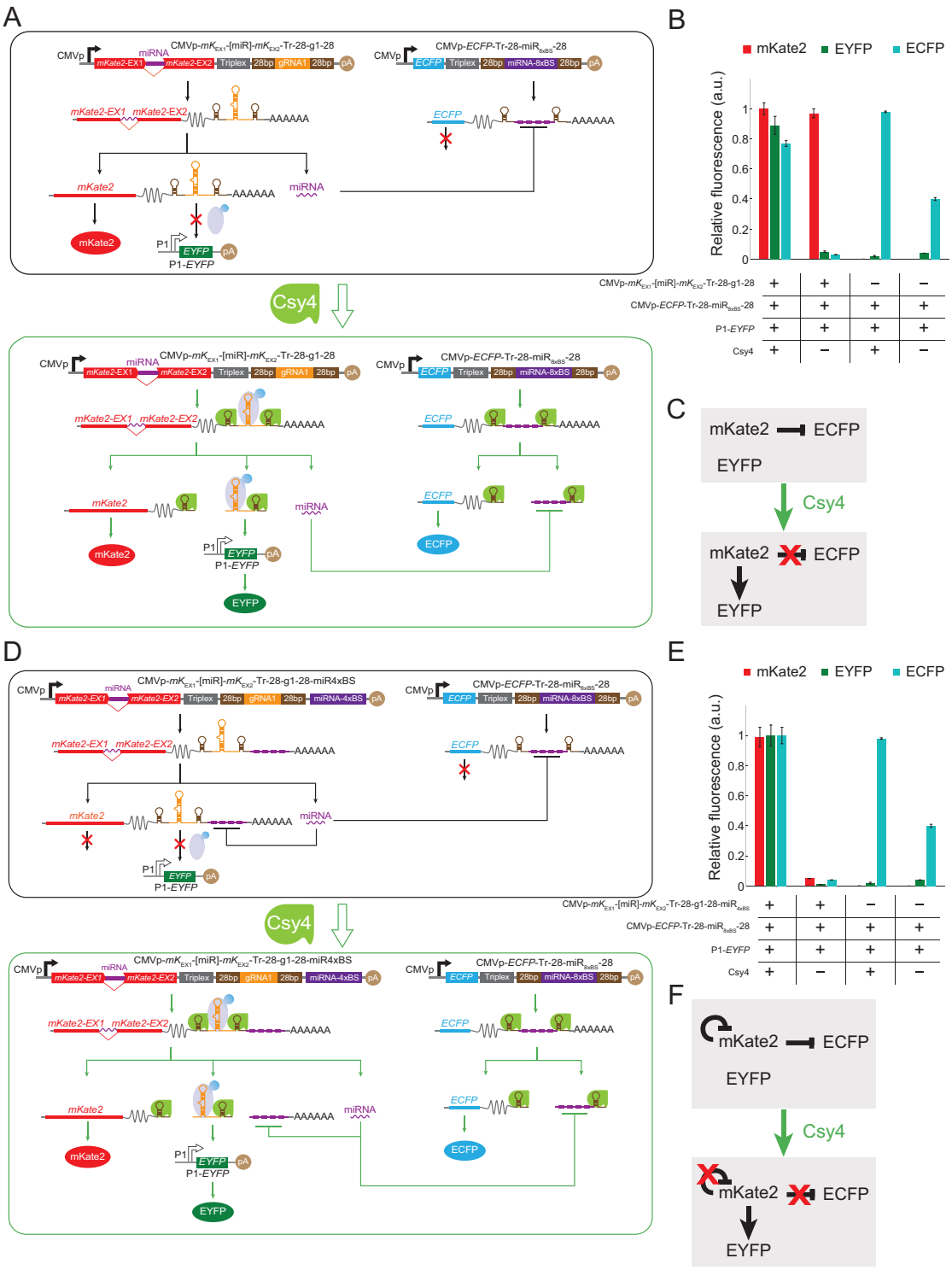


Figure 4-7: CRISPR-Based Transcriptional Regulation Can Be Integrated with Mammalian miRNAs and RNA Processing Mechanisms as well as with Csy4-Dependent RNA Processing to Implement Feedback Loops and Multioutput Circuits that Can Be Rewired at the RNA Level

Figure 4-7: (A) We created a single transcript that encoded both miRNA and CRISPR-TF regulators by expressing a miRNA from an intron within *mKate2* and gRNA1 from a "triplex/Csy4" architecture (CMVp-mK_{EX1}-[miR]-mK_{EX2}-Tr-28-g1-28). In the presence of taCas9, but in the absence of Csy4, this circuit did not activate a downstream gRNA1-specific P1-*EYFP* construct and did repress a downstream ECFP transcript with 8× miRNA binding sites flanked by Csy4 recognition sites (CMVp-*ECFP*-Tr-28-miR8×BS). In the presence of both taCas9 and Csy4, this circuit was rewired by activating gRNA1 production and subsequent EYFP expression, as well as by separating the ECFP transcript from the 8× miRNA binding sites, thus ablating miRNA inhibition of ECFP expression.

(B and C) Csy4 production changes the behavior of the circuit in (A) by rewiring circuit interconnections.

(D) We incorporated an autoregulatory feedback loop into the network topology of the circuit described in (A) by encoding 4× miRNA binding sites at the 3' end of the input transcript (CMVp-mK_{EX1}-[miR]-mK_{EX2}-Tr-28-g1-28-miR4×BS). This negative feedback suppressed mKate2 expression in the absence of Csy4. However, in the presence of Csy4, the 4× miRNA binding sites were separated from the *mKate2* mRNA, thus leading to *mKate2* expression.

E and F) Csy4 production changes the behavior of the circuit in (D) by rewiring circuit interconnections. In contrast to the circuit in (A), *mKate2* was suppressed in the absence of Csy4 but was highly expressed in the presence of Csy4 due to elimination of the miRNA-based autoregulatory negative feedback. Each of the mKate2, EYFP, and ECFP levels in (B) and (E) were normalized to the respective maximal fluorescence among all tested scenarios. The controls in column 3 and 4 in (B) and (E) are duplicated, since the two circuits in (A) and (D) were tested in the same experiment with the same controls. Data are represented as mean ±SEM. See also FigureB-8.

suppressed *ECFP* expression and no functional gRNA1 was generated (Figures 4-7-B and B-8, "Mechanism 1"). In the presence of Csy4, ECFP expression increased by 30-fold compared to the no Csy4 condition, which we attributed to Csy4-induced separation of the 8× miRNA-BS from the *ECFP* transcript (Figure 4-7-B). Also, the presence of Csy4 generated functional gRNA1, leading to 17-fold increased *EYFP* expression compared to the no Csy4 condition (Figure 4-7-B). The mKate2 fluorescence levels were high in both the Csy4-positive and Csy4-negative conditions. Thus, Csy4 catalyzed RNA-based rewiring of circuit connections between the input node and its

two outputs by simultaneously inactivating a repressive output link and turning on an activating output link (Figure 4-7-C).

To demonstrate facile circuit programming using RNA-dependent mechanisms, we extended the design in Figure 4-7-A by incorporating an additional 4× miRNA-BS at the 3' end of the *mKate2*-containing transcript (Figure 4-7-D, CMVp-mK_{EX1}-[miR]-mK_{EX2}-Tr-28-g1-28-miR4×BS). In the absence of Csy4, this resulted in autoregulatory, negative-feedback suppression of *mKate2* expression by the miRNA generated within the *mKate2* intron (Figures 4-7-E and B-8, "Mechanism 2"). In addition, both ECFP and EYFP levels remained low due to repression of *ECFP* by the miRNA and the lack of functional gRNA1 generation. However, in the presence of Csy4, mKate2 levels increased by 21-fold due to Csy4-mediated separation of the 4× miRNA-BS from the *mKate2* transcript. Furthermore, *ECFP* inhibition by the miRNA was relieved in a similar fashion, resulting in a 27-fold increase in ECFP levels. Finally, functional gRNA1 was generated, leading to a 50-fold increase in EYFP levels (Figure 4-7-E). Thus, Csy4 catalyzed RNA-based rewiring of circuit connections between the input node and its two outputs by simultaneously inactivating a repressive output link, turning on an activating output link, and inactivating an autoregulatory feedback loop (Figure 4-7-F).

4.5 Discussion

Synthetic biology provides tools for studying basic biology by disrupting, rewiring, and mimicking natural network motifs [37]. In addition, synthetic circuits have been used to link exogenous signals to endogenous gene regulation [46, 147], to address biomedical applications [104, 143], and to perform cellular computation [10, 30]. Although many synthetic gene circuits have been based on transcriptional regulation, RNA-based regulation has also been used to construct a variety of artificial gene circuits [4, 10, 119, 146]. However, previous efforts have not yet integrated RNA-based regulation with CRISPR-TFs, which are both promising strategies for implementing scalable genetic circuits given their programmability and potential for multiplexing.

In this work, we created a rich toolkit for engineering artificial gene circuits and endogenous gene regulation in human cells. We developed multiple complementary approaches to generate functional gRNAs from transcripts of translated genes regulated by RNAP II promoters. These architectures, based either on Csy4 or ribozymes, enable the encoded RNA and protein levels to be tuned. Choosing between these architectures depends on the specific application. When gRNA expression does not need to be specifically timed or synchronized in a complex fashion, a constitutively active ribozyme-based system is beneficial, since it does not need the expression of an additional protein (Csy4). Conditional gRNA production via ribozyme-based architectures can be achieved by using conditional RNAP II promoters or ligand-dependent cleavage [131] to trigger gRNA release using exogenous control. Applications that require more complicated regulation, synchronization or rewiring of multiple genes 4-7, and/or tunable input-output relationships can benefit from the additional control afforded by Csy4. For example, functional gRNAs are produced with the expression of the harboring RNA transcript and the presence of Csy4. The expression of both of these components could be linked to regulated or conditional promoters for more specific spatial or temporal control of CRISPR-TF circuits.

Complementary to synthetic circuits, we showed that this toolkit can be used to activate endogenous promoters from multiple different endogenous human RNAP II promoters, as well as viral CMVp. We also described useful strategies for multiplexed gRNA expression from compact single transcripts to modulate both synthetic and native promoters. This feature is beneficial because it can be used to regulate multiple nodes from a single concise one, thus enabling sophisticated circuits with a large number of parallel "fan-outs" (i.e., outgoing interconnections from a given node) and networks with dense interconnections. Moreover, the ability to target endogenous loci with several gRNAs in a condensed fashion is critical, since multiple gRNAs are needed for substantial modulation of native promoters. Thus, our tools can be used to build efficient artificial gene networks and to perturb native regulatory networks.

The native CRISPR RNA context can be used for multiplexed genome engineering when expressed from RNAP III promoters in mammalian cells [29]. However, it

remains unclear whether CRISPR/Cas-based systems can be multiplexed when expressed from RNAP II promoters using this approach and what cellular factors mediate this process. In addition to transcriptional regulation, if a nuclease-proficient Cas9 was used with our platform instead of taCas9, then in vivo multiplexed genome editing activity could be conditionally linked to cellular signals via regulation of gRNA expression. In addition to genetic studies, this capability could be potentially used to create in vivo DNA-based "ticker tapes" that link cellular events to mutations.

This framework integrates mammalian RNA regulatory mechanisms with the RNA-dependent protein, dCas9, and the RNA-processing protein, Csy4, from bacteria. These architectures lay a foundation for sophisticated and compact synthetic gene circuits in human cells, such as multistage, multi-input/multi-output networks and feedback circuits capable of logic, computing, and interfacing with endogenous systems. Theoretically, since the specificity of regulatory interconnections with these tools is determined only by RNA sequences, scalable circuits with almost any network topology could be constructed. We demonstrated highly specific and effective three-layer transcriptional cascades with two different architectures that incorporated RNA triplexes, introns, Csy4, and CRISPR-TFs. The absence of undesired crosstalk between different stages of the cascade underscores the orthogonality and scalability of RNA-dependent regulatory schemes.

We also combined RNA regulatory parts, CRISPR-TFs, and RNAi to create various circuit topologies that can be rewired via conditional RNA processing. Since both positive and negative regulation are possible with the same taCas9 protein [42] and miRNAs enact tunable negative regulation, many important multifunctional and multicomponent network topologies can be implemented using this set of regulatory parts. In addition, Csy4 can be used to catalyze changes in gene expression by modifying RNA transcripts and rewiring network topologies. For example, functional gRNAs were liberated for transcriptional modulation, and miRNA binding sites were removed from RNA transcripts to eliminate miRNA-based links. This mechanism demonstrates the potential of adapting bacterial proteins to modulate mammalian gene regulation at the RNA level. This feature could be used to minimize unwanted

leakage in positive feedback loops and to dynamically switch circuits between different states. By linking Csy4 expression to endogenous promoters, interconnections between circuits and network behavior could also be conditionally linked to specific tissues, events (e.g., cell cycle phase and mutations), or environmental conditions. With genome mining or directed mutagenesis, orthogonal Cas9 and Csy4 variants could be discovered and used for more complicated regulatory and RNA processing schemes [41].

In summary, this work provides a diverse set of tools for constructing scalable regulatory gene circuits, tuning them, modifying connections between circuit components, and synchronizing the expression of multiple genes in a network. Furthermore, these regulatory parts could be used to interface synthetic gene circuits with endogenous systems as well as to rewire endogenous networks. Importantly, the promoters, proteins, and miRNAs used in these architectures are not limited to synthetic ones. For many applications, it will be useful to utilize endogenous cellular components, such as tissue- or cell-phase-specific promoters and miRNAs, in order to interface engineered systems with native networks. Similarly, the outputs for these architectures are not limited to reporter genes, but can be effector genes, multiplexed gRNAs that target endogenous promoters, or any other encodable gene. We envision that integrating RNA-dependent regulatory mechanisms with RNA processing will enable sophisticated transcriptional and post-transcriptional regulation, accelerate synthetic biology, and facilitate the study of basic biology in human cells.

4.6 Methods

4.6.1 Plasmid Construction

The **CMVp-dCas9-**3×**NLS-VP64** (taCas9, Construct 1, Table A.3) plasmid was built as described previously [42]. The csy4 gene from *P. aeruginosa* strain UCBPP-PA14 [113] was codon optimized for expression in human cells and cloned downstream of the PGK1 promoter in a PGK1p-EBFP2 plasmid [42] to create **PGK1p-Csy4-**

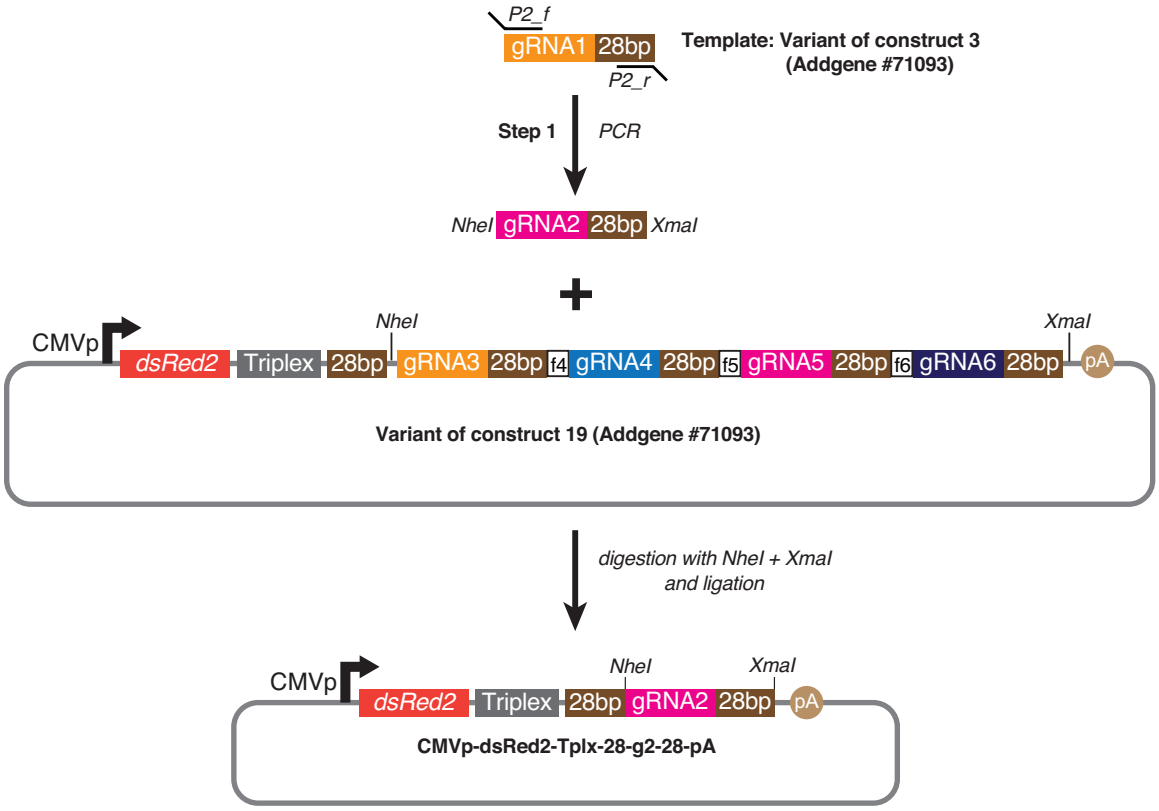
pA (Construct 2, Table A.3). The various gRNA expression constructs were built using conventional restriction enzyme cloning, Golden Gate assembly, and/or Gibson assembly as described below and in Table A.3.

The plasmid **CMVp-*mKate2*-Triplex-28-gRNA1-28-pA** (Construct 3, Table A.3) was built using Gibson Assembly from three parts amplified with appropriate homology overhangs: 1) the full length coding sequence of *mKate2*; 2) the first 110 bp of the mouse MALAT1 3' triple helix [145]; and 3) gRNA1 containing a 20 bp Specificity Determining Sequence (SDS) and a *S. pyogenes* gRNA scaffold along with 28nt Csy4 recognition sites. For cloning in a desired gRNA with a triplex, please refer to Figure 4-8

The reporter plasmids **P1-*EYFP*-pA** (Construct 5, Table A.3), **P2-ECFP-pA** (Construct 6, Table A.3) and **P2-EYFP-pA** (Construct 27, Table A.3) were built by cloning in repeated gRNA1 binding sites and repeated gRNA2 binding sites into the NheI site of pG5-Luc (Promega) via annealing complementary oligonucleotides. Then, EYFP and ECFP were cloned into the NcoI/FseI sites, respectively.

The plasmid **CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-pA** (Construct 4, Table A.3) was built by Gibson Assembly of the following parts with appropriate homology overhangs: 1) the *mKate2*_EX1 (a.a. 1-90) of *mKate2*; 2) *mKate2*_EX2 (a.a. 91-237) of *mKate2*; and 3) gRNA1 containing a 20 bp SDS followed by the *S. pyogenes* gRNA scaffold flanked by Csy4 recognition sites and the HSV1 acceptor, donor and branching sequences. Variations of the **CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-pA** plasmid containing consensus and snoRNA2 acceptor, donor, and branching sequences and with and without the Csy4 recognition sequences (Constructs 8-11, Table A.3) were built in a similar fashion.

The ribozyme-expressing plasmids **CMVp-*mKate2*-Triplex-HHRibo-gRNA1-HDVRibo-pA** and **CMVp-*mKate2*-HHRibo-gRNA1-HDVRibo-pA** plasmids

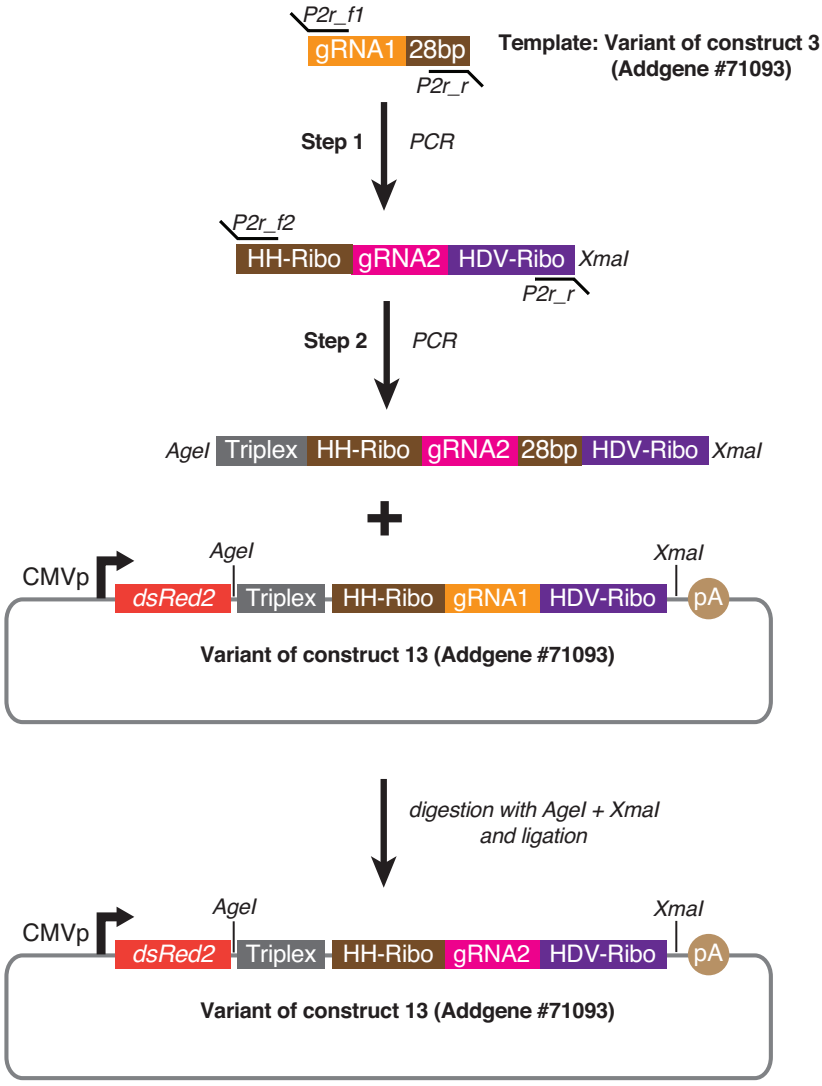


Oligo	*Sequence*
P2_f	ATA*GCTAGC*GTAAGTCGGAGTACTGTCCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P2_r	ATA*CCCGGG*TTTCTTAGCTGCCTATACGGCAGTGAACGAAAAAAAGCACCG

Figure 4-8: Instructions for cloning in desired gRNAs with triplex - **CMVp-*mKate2*-Triplex-28-gRNA1-28-pA** (Construct 3, Table A.3)

(Constructs 13 and 14, respectively, Table A.3) were built by Gibson Assembly of XmaI-digested CMVp-*mKate2*, and PCR-extended amplicons of gRNA1 (with and without the triplex and containing HHRibo [47] on the 5' end and HDVRibo [47] on the 3' end). For cloning in a desired gRNA with ribozymes, please refer to Figure 4-9

The plasmid **CMVp-HHRibo-gRNA1-HDVRibo-pA** (Construct 15, Table A.3) was built similarly by Gibson Assembly of SacI-digested **CMVp-*mKate2*** and a PCR-extended amplicon of gRNA1 containing HHRibo on the 5' end and HDVRibo on the 3' end.



Oligo	*Sequence*
P2r_f1	GGCCTTTCCCTAGCTTTAAAAAAAAAAAAGCAAAACGACTACTGATGAGTCCGTGAGGACGAAACG AGTAAGCTCGTC*GTAAGTCGGAGTACTGTCCT*GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P2r_f2	ATA*ACCGGT*GATTCGTCAGTAGGGTTGTAAAGGTTTTTCTTTTCCTGAGAAAACAACCTT TTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTCCCTAGCTTTAAAAAAAAAAAAGCAAAA
P2r_r	ATA*CCCGGG*GTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGC GAGGAGGCTGGGACCATGCCGGCCAAAAGCACCGACTCGGTGCCACTTTTTC

Figure 4-9: Instructions for cloning in desired gRNAs with ribozymes - **CMVp-*mKate2*-Triplex-HHRibo-gRNA1-HDVRibo-pA** (Construct 13, Table A.3)

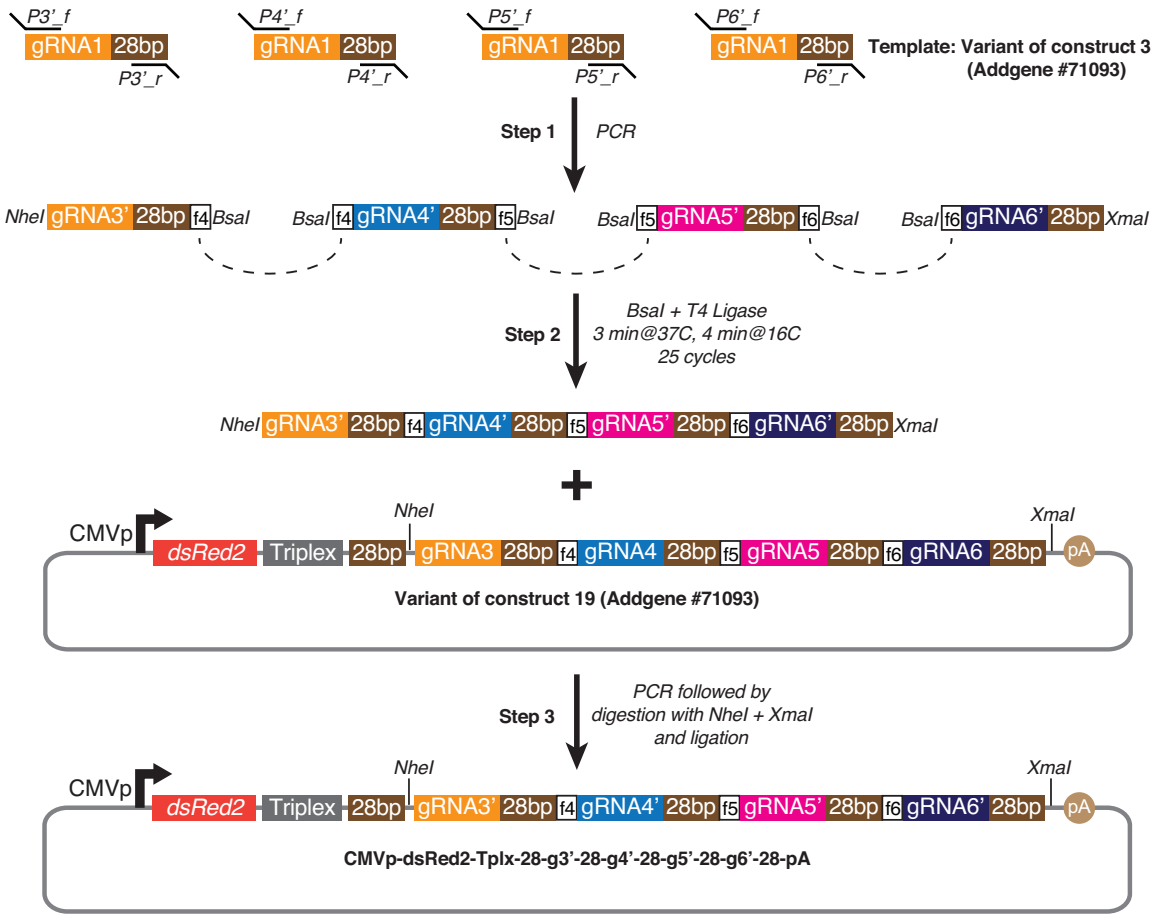
The plasmid **CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-Triplex-28-gRNA2-28-pA** (Construct 16, Table A.3) was built by Gibson Assembly of the following parts using appropriate homologies: 1) XmaI-digested **CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-pA** (Construct 4, Table A.3) and 2) PCR amplified **Triplex-28-gRNA2-28** from **CMVp-*mKate2*-Triplex-28-gRNA1-28-pA** (Construct 3, Table A.3).

The plasmid **CMVp-*mKate2*-Triplex-28-gRNA1-28-gRNA2-28-pA** (Construct 17, Table A.3) was built by Gibson Assembly with the following parts using appropriate homologies: 1) XmaI-digested **CMVp-*mKate2*-Triplex-28-gRNA1-28-pA** (Construct 3, Table A.3) and 2) PCR amplified **28-gRNA2-28**.

The plasmid **CMVp-*mKate2*-Triplex-28-gRNA3-28-gRNA4-28-gRNA5-28-gRNA6-28** (Construct 19, Table A.3) was constructed using a Golden Gate approach using the Type IIs restriction enzyme, BsaI [40], as shown in Figure 4-10

1. **Step 1** Each individual gRNA of interest (e.g., gRNA3, gRNA4, etc.) containing a 20 bp SDS followed by the S. pyogenes gRNA scaffold and a Csy4 '28' sequence was PCR amplified to be flanked on both the 5' and 3' ends by the BsaI type IIs restriction enzyme site. Short, 5 base pair sequences (e.g., f3, f4, etc.) were introduced on the 5' and 3' ends immediately downstream and upstream of the BsaI sites, respectively. These 5 bp sequences were designed such that each adjoining gRNA to be cloned would contain the same 5 bp sequence.

2. **Step 2** The amplified products containing the BsaI sites were mixed in a single reaction mixture containing the enzyme BsaI and T4 DNA ligase. The reaction mixture was subjected to 25 repeats of 3 minute digestions followed by 4 minute ligations at 37C and 16C, respectively.



Oligo	*Sequence*
P3'_f	ATA*GCTAGC*GTGTACTCTCTGAGGTGCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P3'_r	ATA*GGTCTC*AACCTTTTCTTAGCTGCCTATACGGCAGTGAACGAAAAAAAGCACCG
P4'_f	ATA*GGTCTC*AAGGT*GACGCAGATAAGAACCAGTT*GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P4'_r	ATA*GGTCTC*ACCTGTTTCTTAGCTGCCTATACGGCAGTGAACGAAAAAAAGCACCG
P5'_f	ATA*GGTCTC*ACAGG*GCATCAAGTCAGCCATCAGC*GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P5'_r	ATA*GGTCTC*AGACTTTTCTTAGCTGCCTATACGGCAGTGAACGAAAAAAAGCACCG
P6'_f	ATA*GGTCTC*AAGTC*GGAGTCACCCTCCTGGAAAC*GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P6'_r	ATA*CCCGGG*TTTCTTAGCTGCCTATACGGCAGTGAACGAAAAAAAGCACCG

Figure 4-10: Instructions for cloning four multiplexed gRNAs on to a single transcript; **CMVp-*mKate2*-Triplex-28-gRNA3-28-gRNA4-28-gRNA5-28-gRNA6-28** (Construct 19, Table A.3)

3. **Step 3** The assembled product containing all the gRNAs of interest from Step 2 was further PCR amplified to produce a larger amount of the assembled gRNA construct. This product was then digested with NheI-HF and XmaI and cloned

into the **CMVp-*mKate2*-Triplex-28-gRNA1-28** plasmid (a variant of Construct 3, Table A.3).

The **CMVp-*mKate2*_EX1-[miRNA]-*mKate2*_EX2-pA** plasmid containing an intronic FF4 (a synthetic miRNA) was received as a gift from Lila Wroblewska. The synthetic FF4 miRNA was cloned into an intron with consensus acceptor, donor and branching sequences between a.a. 90 and 91 of mKate2 to create **CMVp-*mKate2*_EX1-[miRNA]-*mKate2*_EX2-Triplex-28-gRNA1-28-pA** (Construct 20, Table A.3) and **CMVp-*mKate2*_EX1-[miRNA]-*mKate2*_EX2-Triplex-28-gRNA1-28-**4×**miRNA-BS-pA** (Construct 21, Table A.3).

The plasmid **CMVp-*ECFP*-Triplex-28-**8×**miRNA-BS-28-pA** (Construct 22, Table A.3) was cloned via Gibson Assembly with the following parts: 1) full length coding sequence of ECFP and 2) 110 nt of the MALAT1 3' triple helix sequence amplified via PCR extension with oligonucleotides containing eight FF4 miRNA binding sites and Csy4 recognition sequences on both ends.

Further information on construct sequences can also be found at the MIT RLE website - `http://www.rle.mit.edu/sbg/resources/`

4.6.2 Cell Culture and Transfections

Low-passage HEK293T cells were obtained from ATCC; freshly thawed cells were used in this study and were replaced with a fresh batch every 2 months. They were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS, 1% penicillin-streptomycin, 1% GlutaMAX, and non-essential amino acids at 37C with 5% CO2. HEK293T cells were transfected with FuGENE-HD Transfection Reagent (Promega) according to the manufacturer's instructions. Each transfection was made using 200,000 cells/well in a 6-well plate. As a control, with 2 μg of a single plasmid in which a CMVp regulated *mKate2*, transfection efficiencies were routinely

higher than 90% (determined by flow cytometry). Unless otherwise indicated, each plasmid was transfected at 1 μg/sample. All samples were transfected with taCas9, unless specifically indicated. Cells were processed for flow cytometry or qRT-PCR analysis 72 hrs after transfection.

4.6.3 Quantitative RT-PCR

The RT-PCR procedure was developed in [108] and described here. Cells were harvested 72 hrs post-transfection. Total RNA was isolated by using the RNeasy Plus RNA isolation kit (Qiagen). cDNA synthesis was performed by using qScript cDNA SuperMix (Quanta Biosciences). Real-time PCR using PerfeCTa SYBR Green FastMix (Quanta Biosciences) was performed with the Mastercycler ep realplex real-time PCR system (Eppendorf) with oligonucleotide primers, which are as follows:

$IL1RN$:forward GGAATCCATGGAGGGAAGAT
reverse TGTTCTCGCTCAGGTCAGTG
$GAPDH$:forward CAATGACCCCTTCATTGACC
reverse TTGATTTTGGAGGGATCTCG

The primers were designed using Primer3Plus software and purchased from IDT. Melting curve analysis was used to confirm primer specificity. To ensure linearity of the standard curve, reaction efficiencies over the appropriate dynamic range were calculated. Using the ddCt method, we calculated fold-increases in the mRNA expression of the gene of interest normalized to GAPDH expression. We then normalized the mRNA levels to the non-specific gRNA1 control condition. Reported values are the means of three independent biological replicates with technical duplicates that were averaged for each experiment. Error bars represent standard error of the mean (SEM).

Table 4.1: Compensation setup for flow cytometry

Fluorochrome	-% Fluorochrome	Spectral Overlap
PE-Tx-Red-YG	FITC	0%
Pacific Blue	FITC	0.2%
FITC	PE-Tx-Red-YG	21.1%
Pacific Blue	PE-Tx-Red-YG	1%
FITC	Pacific Blue	7.5%

4.6.4 Flow Cytometry

Cells were washed with DMEM and 1× PBS, resuspended in 1× PBS, and immediately assayed with a Becton Dickinson LSRII Fortessa flow cytometer. At least 50,000 cells were recorded per sample in each data set. The results of each experiment represent at least three biological replicates. Error bars are SEM on the weighted median fluorescence values.

Compensation controls

Compensation controls were strict and designed to remove false-positive cells even at the cost of removing true-positive cells. Compensation was done with BD FACSDiva (version no. 6.1.3; BD Biosciences) as detailed in Table 4.1:

Flow cytometry analysis

Compensated flow cytometry results were analyzed using FlowJo software (vX.0.7r2). Calculations were performed as described below:

1. All samples were gated to exclude cell clumps and debris (population P1)

2. Histograms of P1 cells were analyzed according to the following gates, which were determined according to the auto-fluorescence of non-transfected cells in the same acquisition conditions such that the proportion of false-positive cells would be lower than 0.1%

(a) mKate2: 'mKate2 positive' cells were defined as cells above a fluorescence threshold of 100 a.u.

(b) EYFP: 'EYFP positive' cells were defined as cells above a fluorescence threshold of 300 a.u.

(c) ECFP: 'ECFP positive' cells were defined as cells above a fluorescence threshold of 400 a.u.

3. The percent of positive cells (% positive) and the median fluorescence for each 'positive cell' population were calculated. The % positive cells was multiplied by the median fluorescence, resulting in a weighted median fluorescence expression level that correlated fluorescence intensity with cell numbers. This measurement strategy is consistent with several previous studies [4, 146].

4. The weighted median fluorescence was determined for each sample. The mean of the weighted median fluorescence of biological triplicates was calculated. These are the data presented in the paper. The standard error of the mean (SEM) was also computed and presented as error bars.

5. To facilitate comparisons between various constructs and to account for variations in the brightness of different fluorescent proteins, the weighted median fluorescence for each experimental condition was divided by the maximum weighted median fluorescence for the same fluorophore among all conditions tested in the same set of experiments.

Flow cytometry data plots shown in the Supplemental Information are representative compensated data from a single experiment. As noted above, cells were gated to exclude cell clumps and debris (population P1), and the entire gated population of viable cells are presented in each figure. The threshold for each sub-population Q1-Q4 was set according to the thresholds described above. The percentage of cells in each sub-population is indicated in the plots. Black crosses in the plots indicate the median fluorescence for a specific sub-population.

4.7 Acknowledgements and Copyright

We thank members of the Lu lab, especially Ramiz Daniel and Ky Lowenhaupt, for helpful discussions. We thank Lila Wroblewska for an *mKate2* plasmid expressing intronic miRNA and Jeremy Wilusz and Courtney JnBaptiste for the cGFP_MALAT1_3' plasmid. The above work was supported by the Defense Advanced Research Projects Agency and the National Institutes of Health (DP2 OD008435 and P50 GM098792).

Chapter 5

Conclusion

In this thesis, I presented a toolkit designed by integrating the two component Cas9-gRNA CRISPR technology with a wide array of mammalian molecular components which have not been previously demonstrated to work in concert.

First, by integrating a catalytically dead version of the CRISPR/Cas protein Cas9 (dCas9) with mammalian transcriptional activator VP64 and mammalian transcriptional repressor KRAB, we built and characterized tunable, multifunctional and orthogonal CRISPR/Cas transcription factors (CRISPR-TFs) in human cells. We leveraged the observation that CRISPRTFs can simultaneously activate and repress gene expression depending upon the location of the target sequence relative to the promoter and built highly scalable synthetic logic circuits in human cells. We have also characterized the efficacy of our toolkit in enhancing current applications including transcriptional modulation of endogenous genes and have demonstrated unparalleled multiplexing gains. Furthermore, unlike the traditional knock-out/knock down and over expression approaches to phenotype screening, our work demonstrates that taCas9 can provide great advantages because of its ability to both activate and repress gene expression simultaneously.

Second, we have integrated CRISPR-TFs and Cas6/Csy4 based RNA processing with multiple mammalian RNA regulatory strategies including RNA Polymerase II (RNAP II) promoters, RNA-triple-helix structures, introns, microRNAs and ribozymes. By simultaneously employing CRISPR-TFs, Csy4 and microRNAs in an

integrated framework, we were able to build some of the most complex while programmable gene networks. Our work also ushers in a new set of novel tools which can be indispensable for building RNA-only devices towards engineering mammalian cells.

Finally, by integrating Cas9 nuclease activity with the Non Homologous End Joining (NHEJ) DNA repair pathway, we enable altogether novel applications including DNA based memory in single cells, and continuous generation of genomic diversity that can be exploited for protein engineering applications.

Given the rapid pace at which CRISPR/Cas technology is employed in current research and practice, especially in the mammalian context, our work is a much welcomed development. We have also made all of our constructs and cell lines quickly accessible to the wider research community via our lab's website, Addgene and Genebank. We believe our toolkit enhances the breadth and scope of technologies that define mammalian synthetic biology research for the near future.

Appendix A

Tables

Table A.1: Sequences of DNA constructs used in building Synthetic ZF transcription circuits

Construct	Sequence
ZF43-8-VP16	ATGGCTCCTAAGAAAAAGCGCAAAGTCCGGCCGGCATCTAGACCC GGGGAGCGCCCCTTCCAGTGTCGCATTTGCATGCGGAACTTTTCG CGCCAGGACAGGCTTGACAGGCATACCCGTACTCATACCGGTGAA AAACCGTTTCAGTGTCGGATCTGTATGCGAAATTTCTCCCAGAAG GAGCACTTGGCGGGGCATCTACGTACGCACACCGGCGAGAAGCCA TTCCAATGCCGAATATGCATGCGCAACTTCAGTCGCCGCGACAACC TGAACCGGCACCTAAAAACCCACCTGAGAGGATCCGGAGACGTCG CCGTCGACGCCCCACCAACCGATGTTTCTTTAGGTGATGAGTTACA TTTGGACGGTGAGGACGTAGCTATGGCACATGCAGATGCCTTAGA CGATTTTGACTTGGATATGTTAGGAGATGGAGATTCACCAGGACC TGGTTTCACACCACACGATTCAGCTCCATACGGTGCTTTGGACATG GCCGATTTCGAATTTGAACAGATGTTTACTGATGCCTTGGGTATA GACGAATACGGTGGTGGAGCTAGCGCCAGATCTGCTCCATGGTAA
1x ZFBS43-8 in pCyc1min	CCCGGGAGAGTGAGGACGGATCCCAGATCCGCCAGGCGTGTATAT ATAGCGTGGATGGCCAGGCAACTTTAGTGCTGACACATACAGGCA TATATATATGTGTGCGACGACACATGATCATATGGCATGCATGTG CTCTGTATGTATATAAAACTCTTGTTTTCTTCTTTTCTCTAAATAT TCTTTCCTTATACATTAGGACCTTTGCAGCATAAATTACTATACTT CTATAGACACACAAACACAAATACACACACTAATCTAGATATTAAA ATG
8x ZFBS43-8 in pCyc1min	GTCGACAGAGTGAGGACTCGAAAATATTAATAGAGTGAGGACTCG AAAATATTAATAGAGTGAGGACTCGAAAATATTAATAGAGTGAGG ACTCGAAAATATTAATATCGATAGAGTGAGGACTCGAAAATATTA ATAGAGTGAGGACTCGAAAATATTAATGAATTCAGAGTGAGGACT CGAAAATATTAATAGAGTGAGGACTCGAAAATATTAATGGATCCC AGATCCGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAACTT TAGTGCTGACACATACAGGCATATATATATGTGTGCGACGACACA TGATCATATGGCATGCATGTGCTCTGTATGTATATAAAACTCTTGT TTTCTTCTTTTCTCTAAATATTCTTTCCTTATACATTAGGACCTTT GCAGCATAAATTACTATACTTCTATAGACACACAAACACAAATACA CACACTAATCTAGATATTAAAATG
SSN6-ZF43-8	ATGGCTCCTAAGAAAAAGCGCAAAGTCGGTATCCATGGCGTTCCC GTCGACAATCCGGGCGGTGAACAAACAATAATGGAACAACCCGCT CAACAGCAACAACAACAGCAACAACAACAGCAGCAACAGCAACAG CAGGCAGCAGTTCCTCAGCAGCCACTCGACCCATTAACACAATCAA CTGCGGAAACTTGGCTCTCCATTGCTTCTTTGGCAGAAACCCTTGG TGATGGCGACAGGGCCGCAATGGCATATGACGCCACTTTACAGTT CAATCCCTCATCTGCAAAGGCTTTAACATCTTTGGCTCACTTGTAC CGTTCCAGAGACATGTTCCAAAGAGCTGCAGAATTATATGAAAGA GCACTTTTGGTAAATCCCGAACTATCAGATGTGTGGGCTACTTTA GGTCATTGTTATCTGATGCTGGATGATCTGCAAAGAGCTTACAAT GCCTATCAACAGGCTCTCTACCACCTCAGTAATCCCAACGTACCGA AATTATGGCATGGAATCGGCATTCTTTATGACAGATATGGTTCGC TCGACTATGCCGAAGAAGCTTTTGCCAAAGTTTTGGAATTGGACC CTCATTTTGAAAAGGCAAACGAAATTTACTTCAGACTAGGTATTAT TTATAAACATCAGGGTAAATGGTCTCAAGCTTTGGAATGCTTCAG ATACATTCTCCCTCAACCTCCTGCTCCCTTGCAGGAGTGGGACATA TGGTTTCAGTTGGGTAGTGTTTTGGAGAGTATGGGAGAGTGGCAA GGTGCGAAGGAAGCCTACGAGCATGTCTTGGCTCAAAATCAACAT

SSN6-ZF43-8 (contd . . .)	CATGCCAAAGTATTACAACAATTAGGTTGTCTTTACGGTATGAGT AACGTACAATTTTATGACCCTCAAAAGGCATTGGATTATCTTCTAA AGTCGTTAGAAGCAGATCCCTCCGATGCCACTACATGGTATCATCT CGGTAGAGTGCATATGATTAGAACAGATTATACTGCCGCATATGA TGCTTTCCAACAAGCTGTTAATAGAGATTCAAGAAACCCTATCTTT TGGTGCTCAATCGGTGTTTTATATTACCAAATTTCTCAATACAGAG ACGCCTTAGACGCGTACACAAGAGCCATAAGATTAAATCCTTATAT TAGTGAAGTTTGGTACGATCTAGGTACTCTTTACGAAACTTGTAA CAACCAATTATCTGACGCCCTTGATGCGTATAAGCAAGCTGCAAG ACTGGACGTAAATAATGTTCACATAAGAGAAAGATTAGAAGCTTT AACAAAGCAGTTAGAAAACCCAGGCAATATAAACAAATCGAACGG TGCGCCAACGAATGCCTCTCCTGCCCCACCTCCTGTGATTTTACAA CCTACCTTACAACCTAATGATCAAGGAAATCCTTTGAACACTAGAA TTTCAGCCCAATCTGCCAATGCTACTGCTTCAATGGTACAACAACA GCATCCTGCTCAACAAACGCCTATTAACTCTTCTGCAACAATGTAC AGTAATGGAGCTTCCCCTCAATTACAAGCTCAAGCTCAAGCTCAAG CTCAAGCACAAGCTCAAGCACAAGCACAAGCTCAAGCACAAGCAC AAGCACAAGCGCAAGCACAAGCACAAGCACAGGCGCAAGCACAGG CACAAGCACAAGCACAAGCACATGCACAAGCGCAAGCACAAGCAC AAGCACAGGCACAAGCACAAGCACAGGCGCAGGCACAACAACAAC AACAACAACAGCAACAACAACAACAACAACAACAACAACAACAACA ACAACAACAACAACAACAACAACAACAGCAGCAGCAATTACAGCCC CTACCAAGACAACAGCTGCAGCAAAAGGGAGTTTCTGTGCAAATG TTAAATCCTCAACAAGGGCAACCATATATCACACAGCCAACAGTCA TACAAGCTCACCAACTGCAACCATTTTCTACACAAGCTATGGAACA TCCGCAAAGCTCTCAACTGCCACCTCAACAGCAACAACTACAATCT GTTCAACATCCACAACAACTTCAAGGCCAGCCTCAAGCCCAAGCTC CCCAACCTTTAATCCAGCATAACGTGGAACAGAACGTTTTACCTCA AAAGAGATACATGGAAGGTGCAATCCACACTTTAGTAGATGCCGC CGTATCCAGTAGCACCCACACAGAGAATAACACAAAGTCTCCTCGT CAACCAACCCATGCCATTCCAACGCAAGCTCCCGCAACAGGAATAA CGAACGCTGAACCACAGGTAAAGAAGCAAAAGTTGAACTCTCCAA ATTCAAACATCAACAAATTAGTAAATACTGCTACTTCCATTGAAGA AAATGCAAAATCTGAGGTGAGCAACCAATCGCCAGCAGTAGTGGA GTCTAATACCAATAATACTTCACAAGAAGAAAAACCTGTAAAAGC AAACTCAATACCTTCAGTAATTGGCGCACAGGAACCTCCACAGGA AGCTAGTCCTGCTGAAGAAGCTACCAAAGCAGCTTCTGTTTCTCCT TCTACAAAACCGCTTAATACGGAACCAGAGTCATCTAGTGTCCAAC CAACTGTATCATCAGAAAGTTCAACAACAAAAGCAAATGACCAAA GCACTGCTGAGACCATAGAACTTTCTACTGCTACTGTTCCTGCAGA AGCAAGCCCTGTAGAAGACGAAGTAAGACAGCATTCTAAAGAGGA AAACGGCACAACTGAAGCATCTGCACCTTCTACTGAAGAGGCGGA GCCAGCAGCTTCCAGAGATGCTGAAAAACAACAAGATGAAACCGC TGCTACAACGATAACTGTAATCAAACCTACTTTGGAAACAATGGA AACAGTGAAAGAGGAGGCCAAAATGCGTGAGGAAGAGCAAACATC TCAAGAAAAATCCCCACAGGAGAACACACTTCCAAGAGAAAATGT AGTAAGGCAAGTGGAAGAAGATGAAAACTACGACGACGCTAGACC CGGGGAGCGCCCCTTCCAGTGTCGCATTTGCATGCGGAACTTTTC GCGCCAGGACAGGCTTGACAGGCATACCCGTACTCATACCGGTGA AAAACCGTTTCAGTGTCGGATCTGTATGCGAAATTTCTCCCAGAA GGAGCACTTGGCGGGGCATCTACGTACGCACACCGGCGAGAAGCC ATTCCAATGCCGAATATGCATGCGCAACTTCAGTCGCCGCGACAAC CTGAACCGGCACCTAAAAACCCACCTGAGATAA

SIN3-ZF43-8	ATGGCTCCTAAGAAAAAGCGCAAAGTCGGTATCCATGGCGTTCCC GTCGACTCACAGGTTTGGCATAATTCGAATTCGCAATCAAACGAT GTGGCTACTTCAAATGACGCTACGGGTTCCAACGAAAGAAACGAA AAAGAACCGTCCCTCCAGGGAAATAAGCCCGGTTTTGTTCAACAG CAGCAACGGATTACTTTACCCTCGCTATCTGCCTTGAGTACTAAGG AGGAAGATAGAAGAGATTCCAATGGCCAACAGGCTCTAACTTCTC ATGCTGCTCACATATTAGGTTATCCTCCCCCACATTCAAATGCTAT GCCCTCAATTGCAACTGATTCAGCATTGAAACAGCCCCACGAGTAT CACCCTCGCCCTAAATCTTCGTCCTCTTCTCCCTCTATAAACGCTT CGCTTATGAATGCTGGTCCAGCTCCCCTCCCCACAGTGGGAGCCG CCAGTTTTTCTTTGTCGAGATTTGACAATCCATTACCGATAAAAGC TCCTGTTCATACAGAGGAACCAAAAAGTTATAATGGTCTTCAGGA AGAAGAAAAGGCGACGCAACGGCCTCAAGATTGCAAGGAAGTTCC CGCTGGTGTGCAGCCTGCTGATGCCCCTGACCCTAGCAGCAACCA TGCAGATGCTAACGATGACAATAATAACAACGAAAATTCTCACGA TGAAGATGCTGACTACAGACCTCTAAACGTGAAGGATGCCCTATC TTACCTCGAACAGGTCAAATTTCAATTTAGTTCGCGTCCGGATATC TATAATTTATTTTTAGATATTATGAAGGACTTTAAATCTCAGGCAA TAGACACACCGGGCGTTATTGAAAGAGTATCCACTTTGTTCAGAG GTTATCCAATTTTGATTCAAGGGTTCAATACTTTTCTACCCCAAGG CTATAGAATCGAATGCTCCTCTAATCCGGACGACCCCATTAGAGTT ACTACACCAATGGGTACTACGACAGTAAACAATAACATCAGTCCA TCTGGTAGAGGTACAACGGATGCACAGGAACTTGGTTCTTTTCCA GAAAGCGATGGAAATGGTGTTCAACAGCCCTCCAATGTGCCAATG GTGCCTTCGAGTGTGTATCAATCGGAACAAAACCAAGACCAACAA CAATCTTTGCCTCTTTTAGCTACTTCTTCTGGTTTACCTTCAATTC AACAACCTGAAATGCCTGCACATCGCCAAATCCCACAAAGTCAATC TTTAGTGCCTCAAGAAGATGCTAAGAAAAACGTTGATGTCGAATT TAGTCAAGCCATAAGCTACGTTAATAAAATTAAAACTAGATTTGCC GACCAACCTGATATTTACAAGCATTTTCTGGAAATACTACAAACTT ATCAGCGAGAGCAAAAGCCAATAAACGAAGTCTACGCACAAGTGA CGCATCTTTTCCAAAATGCTCCTGATTTACTAGAAGATTTCAAGAA ATTCTTGCCGGACTCTTCAGCTTCTGCCAATCAGCAGGTGCAACAT GCTCAGCAACATGCTCAACAACAACATGAGGCCCAAATGCATGCA CAGGCACAAGCTCAGGCTCAGGCACAGGCACAGGTGGAACAACAG AAGCAGCAACAGCAATTCTTGTATCCAGCTTCGGGCTATTACGGC CATCCTTCTAACCGAGGTATTCCACAGCAAAACTTGCCTCCTATAG GAAGCTTCTCGCCTCCAACAAACGGTTCGACCGTACATGAAGCCTA TCAGGATCAACAGCACATGCAACCACCCCACTTCATGCCCTTACCA TCGATCGTTCAACACGGGCCAAATATGGTACATCAAGGGATAGCG AATGAAAATCCACCTCTATCAGATCTAAGAACGTCTCTCACTGAAC AGTACGCTCCTTCTAGTATACAACATCAGCAGCAGCACCCGCAAAG TATTAGTCCTATAGCAAATACGCAGTATGGTGATATCCCTGTTAG ACCGGAAATTGATTTAGATCCTAGCATTGTGCCTGTGGTCCCTGA ACCCACTGAGCCCATCGAAAACAATATATCGCTTAATGAGGAAGT CACTTTCTTCGAAAAGGCCAAGAGGTATATCGGCAATAAACATTT ATACACTGAGTTTTTGAAAATTTTAAATTTGTACTCTCAAGATATA CTTGATCTTGACGATTTAGTGGAAAAGGTAGATTTCTACTTGGGT TCCAATAAAGAACTATTTACGTGGTTCAAAAACTTTGTTGGCTACC AAGAAAAAACCAAATGTATCGAGAATATTGTTCATGAAAAACATA GATTGGATTTAGATTTATGTGAGGCATTTGGCCCAAGTTACAAGA GGCTACCAAAAAGTGACACTTTCATGCCATGCTCAGGTAGGGATG ATATGTGTTGGGAAGTCTTGAACGATGAATGGGTTGGACATCCTG TATGGGCTTCCGAAGATTCGGGATTTATTGCTCATCGTAAAAACC AGTATGAGGAAACACTATTCAAGATCGAAGAGGAAAGACATGAGT

SIN3-ZF43-8 (contd. . .)	ATGATTTTTACATTGAATCAAATTTAAGAACTATTCAATGTTTGGA AACAATTGTAAATAAGATCGAGAACATGACTGAGAATGAAAAAGC CAATTTTAAACTGCCTCCAGGTCTTGGCCATACTTCAATGACTATT TATAAAAAAGTGATAAGGAAAGTTTATGATAAGGAAAGGGGGTTC GAGATTATTGATGCTTTGCATGAGCACCCTGCAGTGACAGCCCCA GTCGTTCTGAAAAGGTTAAAGCAAAAAGACGAAGAATGGAGAAGA GCTCAACGTGAATGGAATAAAGTTTGGAGGGAGTTAGAACAGAAG GTTTTTTTCAAGTCATTAGATCATTTAGGCTTAACATTTAAACAGG CTGACAAGAAATTATTAACTACAAAGCAGTTGATATCAGAGATTA GCAGCATCAAAGTTGATCAAACAAACAAAAAAATTCACTGGTTAA CTCCTAAACCAAAGAGCCAGTTAGATTTTGATTTCCCTGATAAAAA CATTTTCTATGATATCTTGTGTTTGGCTGACACTTTTATAACCCAT ACCACAGCCTATTCTAATCCCGATAAAGAAAGATTGAAAGATTTAC TAAAATACTTCATCTCTTTGTTTTTTTCTATTTCTTTCGAAAAAAT CGAAGAATCGTTGTACTCCCATAAGCAAAACGTGTCAGAATCTAG CGGATCTGACGATGGCAGTTCTATTGCATCAAGGAAGAGGCCCTA TCAACAAGAAATGAGTTTACTTGATATTTTACACAGGAGCAGATA TCAAAAGCTAAAGCGTTCTAATGATGAAGATGGCAAAGTTCCCCA GCTCTCTGAACCACCCGAAGAAGAACCTAATACCATTGAGGAGGA AGAGCTCATCGATGAAGAAGCTAAAAATCCGTGGCTAACTGGGAA TTTAGTGGAGGAAGCAAACTCTCAGGGTATTATCCAAAATCGCAG TATTTTTAATCTATTCGCTAATACGAATATTTACATATTTTTCCGT CATTGGACAACGATTTATGAGCGGCTTTTGGAAATTAAGCAAATG AATGAAAGGGTCACAAAGGAAATCAACACAAGGTCGACAGTTACT TTTGCCAAAGATCTAGATTTATTATCGAGTCAACTTTCCGAAATGG GACTAGATTTTGTTGGTGAAGACGCGTACAAACAGGTTTTAAGAC TGAGTAGAAGGTTGATTAATGGGGATCTTGAACATCAGTGGTTTG AAGAGAGTTTGCGGCAAGCTTACAATAACAAAGCGTTTAAACTCT ATACAATTGATAAAGTCACCCAATCGTTGGTAAAGCATGCTCATA CCTTGATGACTGACGCTAAAACTGCGGAAATAATGGCTTTGTTCG TTAAAGATAGAAATGCCTCCACCACGAGTGCGAAGGACCAAATTA TCTATCGCTTGCAGGTGCGCTCACATATGTCCAACACAGAAAATAT GTTTAGAATAGAGTTTGATAAAAGAACTCTGCATGTTTCCATTCA ATATATTGCACTTGATGATTTGACACTAAAGGAACCAAAGGCAGA CGAAGATAAATGGAAATATTATGTAACATCGTATGCTCTGCCACA TCCAACAGAAGGCATTCCGCACGAGAAACTGAAGATACCATTTTT GGAAAGGCTCATCGAATTTGGACAAGATATTGATGGAACAGAGGT AGATGAGGAGTTTTCCCCCGAAGGCATCTCTGTATCGACATTGAA AATTAAGATCCAGCCTATAACCTACCAACTGCATATTGAAAACGG GTCTTACGATGTTTTTACCCGTAAGGCTACTAACAAATATCCTACT ATTGCTAATGATAATACTCAAAAAGGAATGGTTTCTCAGAAGAAG GAGTTGATATCAAAATTTCTAGACTGTGCAGTTGGTTTGAGAAAT AATTTAGATGAGGCTCAAAAATTAAGTATGCAAAAAAAATGGGAG AACCTAAAAGACAGCATAGCAAAGACGAGTGCTGGAAACCAGGGA ATAGAGAGTGAAACCGAAAAAGGTAAAATTACGAAACAGGAGCAG TCGGATAATTTGGACAGCTCTACTGCAAGTGTACTGCCTGCATCC ATTACTACTGTACCACAAGATGATAATATAGAAACGACTGGGAAT ACTGAATCTTCAGACAAGGGGGCTAAGATTCAACCCGGGGAGCGC CCCTTCCAGTGTCGCATTTGCATGCGGAACTTTTCGCGCCAGGACA GGCTTGACAGGCATACCCGTACTCATACCGGTGAAAAACCGTTTC AGTGTCGGATCTGTATGCGAAATTTCTCCCAGAAGGAGCACTTGG CGGGGCATCTACGTACGCACACCGGCGAGAAGCCATTCCAATGCC GAATATGCATGCGCAACTTCAGTCGCCGCGACAACCTGAACCGGC ACCTAAAAACCCACCTGAGATAA

UME6-ZF43-8	ATGGCTCCTAAGAAAAAGCGCAAAGTCGGTATCCATGGCGTTCCC GTCGACCTAGACAAGGCGCGCTCTCAAAGCAAACACATGGACGAA TCTAATGCGGCTGCCTCTCTGCTTTCGATGGAAACAACCGCCAACA ATCATCACTATTTGCACAATAAAACATCTCGTGCCACGCTGATGAA TAGCAGCCAAGACGGCAAAAAACATGCAGAAGATGAAGTTAGTGA TGGAGCTAACTCCCGCCATCCTACAATTTCCAGTGCTAGCATCGAA TCTCTCAAGACAACCTACGATGAAAACCCTTTGCTTTCTATTATGA AATCGACATGTGCGCCCAACAACACTCCCGTGCATACTCCGTCTGG TTCGCCGAGTTTGAAAGTCCAAAGTGGCGGAGATATCAAAGACGA TCCTAAGGAAAACGATACTACTACTACGACCAATACTACACTTCAA GATCGCCGTGACAGCGATAACGCTGTGCATGCTGCCGCCAGCCCA CTCGCGCCTTCCAATACACCTTCGGATCCTAAGTCATTGTGCAATG GCCATGTCGCACAGGCTACAGACCCACAAATTTCCGGTGCTATTCA GCCGCAGTATACTGCAACCAACGAGGATGTTTTCCCTTACTCCTCC ACCTCCACTAATAGTAACACTGCCACTACTACTATCGTCGCCGGCG CCAAAAAAAAAAATACATTTGCCGCCACCACAAGCTCCAGCGGTTTC TTCCCCCGGTACCACCGCAGCAGGCTCGGGCGCGGGCACGGGCTC GGGCATCCGTTCCCGCACAGGATCGGATTTGCCGCTCATCATTACC AGCGCCAACAAGAACAACGGTAAGACTACCAATTCGCCTATGTCG ATACTGAGCAGAAACAACAGTACCAACAACAACGACAATAATTCG ATACAAAGCTCAGATTCGAGAGAATCCTCTAATAACAATGAGATT GGCGGGTATTTGCGCGGCGGAACTAAGCGCGGCGGCAGTCCATCT AACGACTCTCAGGTCCAGCATAATGTGCATGATGACCAATGTGCC GTGGGCGTGGCGCCCAGGAACTTCTATTTCAACAAGGATAGAGAG ATAACAGACCCAAATGTAAAACTGGACGAGAACGAATCAAAAATC AACATATCGTTCTGGCTAAATTCGAAATACAGAGATGAGGCTTAT TCTTTGAATGAATCATCCTCCAACAATGCTAGTTCAAACACGGATA CGCCTACAAACTCTCGACATGCGAACACCAGCTCTTCCATTACCAG CAGAAACAATTTCCAGCATTTTAGGTTCAACCAAATACCTTCTCAA CCTCCAACTTCCGCATCTTCGTTTACAAGCACAAACAACAACAACC CTCAACGGAACAATATCAATCGCGGTGAAGACCCGTTTGCCACTTC GTCAAGACCTTCTACTGGGTTTTTTTACGGCGATTTGCCGAATCGT AACAATAGAAATAGTCCCTTCCATACAAATGAACAATACATCCCAC CACCTCCACCGAAATACATCAATTCTAAATTGGATGGATTGAGATC AAGATTATTGCTCGGTCCGAATTCTGCATCTTCATCTACCAAACTA GACGACGACTTGGGTACAGCAGCAGCAGTGCTATCAAACATGAGA TCATCCCCATATAGAACTCATGATAAACCCATTTCCAATGTCAATG ACATGAATAACACAAATGCGCTCGGTGTGCCGGCTAGTAGGCCTC ATTCGTCATCTTTTCCATCAAAGGGTGTCTTAAGACCAATTCTGTT ACGTATCCATAATTCCGAACAACAACCCATTTTCGAAAGCAACAAT TCTACAGCTGTTTTTGATGAAGACCAGGACCAAAATCAAGACTTG TCTCCTTACCATTTAAATCTAAACTCTAAAAAGGTTTTAGATCCCA CTTTTGAGTCAAGGACAAGGCAAGTTACTTGGAATAAGAATGGTA AGCGAATAGACAGACGCCTTTCTGCTCCAGAACAACAACAGCAAC TGGAAGTTCCACCATTGAAAAAATCGAGAAGGTCAGTTGGAAACG CAAGAGTAGCAAGCCAAACCAATAGCGATTATAATTCTCTTGGCG AATCCTCAACTTCGTCAGCTCCATCGTCTCCATCTTTGAAGGCTTC TTCTGGCTTGGCATATACCGCTGATTATCCTAACGCTACTTCGCCG GATTTCGCTAAATCTAAAGGAAAAAATGTCAAGCCTAAGGCAAAA TCAAAGGCGAAACAGTCATCAAAGAAAAGACCAAATAATACTACT TCGAAATCAAAAGCAAACAATTCTCAAGAATCGAATAATGCTACTT CCTCAACGTCTCAAGGTACAAGGTCCCGTACTGGTTGCTGGATTT GTAGATTAAGGAAAAAGAAGTGTACCGAGGAAAGACCGCACTGTT TCAACTGTGAAAGGTTGAAATTGGACTGTCACTATGACGCGTTCA AACCAGATTTTGTATCTGATCCAAAGAAAAAACAGATGAAACTGG

UME6-ZF43-8 (contd. . .)	AGGAAATCAAGAAAAAAACAAAAGAGGCCAAAAGAAGAGCAATGA AAAAAAAACCCGGGGAGCGCCCCTTCCAGTGTCGCATTTGCATGC GGAACTTTTCGCGCCAGGACAGGCTTGACAGGCATACCCGTACTC ATACCGGTGAAAAACCGTTTCAGTGTCGGATCTGTATGCGAAATT TCTCCCAGAAGGAGCACTTGGCGGGGCATCTACGTACGCACACCG GCGAGAAGCCATTCCAATGCCGAATATGCATGCGCAACTTCAGTC GCCGCGACAACCTGAACCGGCACCTAAAAACCCACCTGAGATAA
LexA-VP16	ATGGCTCCTAAGAAAAAGCGCAAAGTCCGGCCGGCATCTAGAAAG GCTTTGACTGCTAGACAACAAGAAGTTTTCGACTTGATCAGAGAC CACATCTCTCAAACTGGTATGCCACCAACTAGAGCTGAAATCGCTC AACGTCTCGGCTTCAGATCGCCAAATGCTGCTGAAGAACACTTGA AGGCTTTGGCTAGAAAGGGTGTTATCGAAATCGTTTCTGGTGCTT CTCGTGGTATCAGATTGTTGCAAGAAGAAGAAGAAGGTTTGCCAT TGGTTGGTAGAGTTGCTGCTGGTGAACCAGGTGGTGGTTCTGTCG ACGCCCCACCAACCGATGTTTCTTTAGGTGATGAGTTACATTTGG ACGGTGAGGACGTAGCTATGGCACATGCAGATGCCTTAGACGATT TTGACTTGGATATGTTAGGAGATGGAGATTCACCAGGACCTGGTT TCACACCACACGATTCAGCTCCATACGGTGCTTTGGACATGGCCG ATTTCGAATTTGAACAGATGTTTACTGATGCCTTGGGTATAGACG AATACGGTGGTGGAGCTAGCTAA

Table A.2: Sequences of DNA constructs used in building Synthetic CRISPR-Cas transcription factors

Construct	Sequence
dCas9-3xNLS-VP64	GCCACCATGGACAAGAAGTACTCCATTGGGCTCGCCATCGGCACA AACAGCGTCGGCTGGGCCGTCATTACGGACGAGTACAAGGTGCCG AGCAAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATA AAGAAGAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACG GCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACC CGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAAT GAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAG TCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATC TTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCA ACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAG GCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAA TTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAAC AGCGATGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATC AGCTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCA AAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAA ACCTCATCGCACAGCTCCCTGGGGAGAAGAAGAACGGCCTGTTTG GTAATCTTATCGCCCTGTCACTCGGGCTGACCCCCAACTTTAAATC TAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGA CACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGA CCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGC CATTCTGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAA AGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCA CCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCC TGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTA CGCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAA ATTTATTAAGCCCATCTTGGAAAAAATGGACGGCACCGAGGAGCT GCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCAC TTTCGACAATGGAAGCATCCCCCACCAGATTCACCTGGGCGAACT GCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAA AGATAACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACC CTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTCGCGTG GATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGA GGAAGTCGTGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAG GATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCC TAAACACTCTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTC ACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTC CTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAG ACGAACCGGAAAGTTACCGTGAAACAGCTCAAAGAAGACTATTTC AAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAG GATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAA TCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACA TTCTTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGG AGATGATTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACG ACAAAGTCATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGG GGCGGCTGTCAAGAAAACTGATCAATGGGATCCGAGACAAGCAGA GTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCA ACCGGAACTTCATGCAGTTGATCCATGATGACTCTCTCACCTTTAA GGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCT TCACGAGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAA

dCas9-3xNLS-VP64 (contd. . .)	GGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGT AATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCG AGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAG GATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAAT CCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAA GCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGA TCAGGAACTGGACATCAATCGGCTCTCCGACTACGACGTGGATGC CATCGTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAA GTGTTGACAAGATCCGATAAAAATAGAGGGAAGAGTGATAACGTC CCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAG CTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATAATCTG ACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGC TTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCAC GTGGCCCAAATTCTCGATTCACGCATGAACACCAAGTACGATGAA AATGACAAACTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCT AAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTG AGAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAAT GCAGTGGTAGGCACTGCACTTATCAAAAAATATCCCAAGCTTGAA TCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAA ATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAG TACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTA CACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAA ACGGAGAAACAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCG CGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTA AAAAGACCGAAGTACAGACCGGAGGCTTCTCCAAGGAAAGTATCC TCCCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATT GGGACCCCAAGAAATACGGCGGATTCGATTCTCCTACAGTCGCTT ACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAA AACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGC GATCAAGCTTCGAAAAAAACCCCATCGACTTTCTCGAGGCGAAAG GATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGT ACTCTCTCTTTGAGCTTGAAAACGGCCGGAAACGAATGCTCGCTA GTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTA AATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAA AGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACA ACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATT CTCCAAAAGAGTGATCCTCGCCGACGCTAACCTCGATAAGGTGCT TTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGC AGAAAACATTATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCT GCAGCCTTCAAGTACTTCGACACCACCATAGACAGAAAGCGGTAC ACCTCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCA ATTACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCTCGGT GGAGACAGCAGGGCTGACGGGCCCTCACTGGGTTCAGGGTCACCC AAGAAGAAACGCAAAGTCGAGGATCCAAAGAAGAAAAGGAAGGTT GAAGACCCCAAGAAAAAGAGGAAGGTGGATGGGATCGGCTCAGGC AGCAACGGCGGTGGAGGTTCAGACGCTTTGGACGATTTCGATCTC GATATGCTCGGTTCTGACGCCCTGGATGATTTCGATCTGGATATG CTCGGCAGCGACGCTCTCGACGATTTCGACCTCGACATGCTCGGG TCAGATGCCTTGGATGATTTTGACCTGGATATGCTCTCATGATGA
KRAB	GATGCTAAGTCACTAACTGCCTGGTCCCGGACACTGGTGACCTTC AAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTG GACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGAGAAC TATAAGAACCTGGTTTCCTTGGGTTATCAGCTTACTAAGCCAGAT GTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCTGGCTGGTGGAG AGAGAAATTCACCAAGAGACCCATCCTGATTCAGAGACTGCATTT

KRAB (contd. . .)	GAAATCAAATCATCAGTTTCCAGCAGGAGCATTTTTAAAGATAAG CAATCCTGTGACATTAAAATGGAAGGAATGGCAAGGAATGATCTC TGG
PGK1	AATTCTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGGAG CATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTG GCCTCTGGCCTCGCACAAATTCTACCGGGTAGGGGAGGCGCTTTT CCCAAGGCAGTCTGGAGCATGCGCTTTAGCAGCCCCGCTGGGCAC TTGGCGCTACACAAGTGGCCTCTGGCCTCGCACACATTCCACATCC ACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTCGCG CCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCGCCCCGC AGCTCGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTCTCA CTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGG TAGGCCTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTCGCTT TCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGCGGGCTC AGGGGCGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTCCTCCGGA GGCCCGGCATTCTGCACGCTTCAAAAGCGCACGTCTGCCGCGCTG TTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCTGCATCCCAAGCT TGCCACCATG
pMLPm	GGGGGGCTATAAAAGGGGGTGGGGGCGTTCGTCCTCACTCTAGAT CTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCC ACCATGGCGCCACCATG
0xa1_pMLPm	ATCGATAGCTAGCGGGGGGCTATAAAAGGGGGTGGGGGCGTTCGT CCTCACTCTAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACT GTTGGTAAAGCCACCATGGCGCCACCATG
1xa1_pMLPm	ATCGATAGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGT AGGCTAGCGGGGGGCTATAAAAGGGGGTGGGGGCGTTCGTCCTCA CTCTAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTG GTAAAGCCACCATGGCGCCACCATG
2xa1_pMLPm	ATCGATAGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGT AGGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGTAGGC TAGCGGGGGGCTATAAAAGGGGGTGGGGGCGTTCGTCCTCACTCT AGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGTAA AGCCACCATGGCGCCACCATG
3xa1_pMLPm	ATCGATAGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGT AGGCTAGCCTGAAAATATTAATTACTGTATGTACATACAGTAGGC TAGCCTGAAAATATTAATTACTGTATGTACATACAGTAGGCTAGC GGGGGGCTATAAAAGGGGGTGGGGGCGTTCGTCCTCACTCTAGAT CTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCC ACCATGGCGCCACCATG
pU6_gRNA	TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGA TCCGGTACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATT CCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATT AGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACG TGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAA TTATGTTTTAAAATGGACTATCATTGCTTACCGTAACTTGAAAGTA TTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC GNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTT AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT CGGTGCTTTTTTT
	where NNNNNNNNNNNNNNNNNNN above is one of the following
m1 gRNA	CTGTCCCCAGTGCAAGTGC
m2 gRNA	TTCTCTATCGATAGCTAGC
m3 gRNA	TCTGCGATCTAAGTAAGCT

m4 gRNA	ACTGTTGGTAAAGCCACCA
m5 gRNA	ATGGCGCCACCATGAGCAG
m6 gRNA	GGCACCTGCACTTGCACTG
m7 gRNA	TATCGATAGAGAAATGTTC
m8 gRNA	AGATCGCAGATCTAGAGTG
m9 gRNA	GTGGCTTTACCAACAGTAC
m10 gRNA	GCTCATGGTGGCGCCATGG
CCAAT gRNA	GAAGGAGCAAAGCTGCTAT
GC-box gRNA	GGCGGGCTCAGGGGCGGGG
a1 gRNA	TACTGTATGTACATACAGT
a2 gRNA	AGTCGCGTGTAGCGAAGCAT
a3 gRNA	CAACGCGACGCTAGATAGCA
g5 gRNA	TAAGTCGGAGTACTGTCCT
p1 gRNA	CTGAGTCCGGTAGCGCTAG

Table A.3: Construct names, designs, and abbreviations used in Chapter 4

Construct	Sequence
Construct 1	**CMVp-dCas9-**3×**NLS-VP64-3'LTR**
Abbreviation	taCas9
Construct 2	**PGK1p-Csy4-pA**
Abbreviation	Csy4
Construct 3	**CMVp-*mKate2*-Triplex-28-gRNA1-28-pA**
Abbreviation	CMVp-*mK*-Tr-28-g1-28
Construct 4	**CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-pA**
Abbreviation	CMVp-*mK*EX1-[28-g1-28]HSV1-*mK*EX2
Construct 5	**P1-*EYFP*-pA**
Abbreviation	P1-EYFP
Construct 6	**P2-ECFP-pA**
Abbreviation	P2-ECFP
Construct 7	**U6p-gRNA1-TTTTT**
Abbreviation	U6p-g1
Construct 8	**CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{consensus}}$-*mKate2*_EX2-pA**
Abbreviation	CMVp-*mK*EX1-[28-g1-28]cons-*mK*EX2
Construct 9	**CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{snoRNA2}}$-*mKate2*_EX2-pA**
Abbreviation	CMVp-*mK*EX1-[28-g1-28]sno-*mK*EX2
Construct 10	**CMVp-*mKate2*_EX1-[28-gRNA1]$_{\text{HSV1}}$-*mKate2*_EX2-pA**
Abbreviation	CMVp-*mK*EX1-[28-g1]HSV1-*mK*EX2
Construct 11	**CMVp-*mKate2*_EX1-[gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-pA**
Abbreviation	CMVp-*mK*EX1-[g1-28]HSV1-*mK*EX2
Construct 12	**CMVp-*mKate2*_EX1-[gRNA2]$_{\text{consensus}}$-*mKate2*_EX2-pA**
Abbreviation	CMVp-*mK*EX1-[g2]cons-*mK*EX2
Construct 13	**CMVp-*mKate2*-Triplex-HHRibo-gRNA1-HDVRibo-pA**
Abbreviation	CMVp-*mK*-Tr-HH-g1-HDV
Construct 14	**CMVp-*mKate2*-HHRibo-gRNA1-HDVRibo-pA**
Abbreviation	CMVp-*mK*-HH-g1-HDV
Construct 15	**CMVp-HHRibo-gRNA1-HDVRibo-pA**
Abbreviation	CMVp-HH-g1-HDV

Construct	Sequence
Construct 16	**CMVp-*mKate2*_EX1-[28-gRNA1-28]$_{\text{HSV1}}$-*mKate2*_EX2-Triplex-28-gRNA2-28-pA**
Abbreviation	CMVp-*mK*EX1-[28-g1-28]HSV1-*mK*EX2-Tr-28-g2-28
Construct 17	**CMVp-*mKate2*-Triplex-28-gRNA1-28-gRNA2-28-pA**
Abbreviation	CMVp-*mK*-Tr-28-g1-28-g2-28
Construct 18	**P1-*EYFP*-Triplex-28-gRNA2-28-pA**
Abbreviation	P1-EYFP-Tr-28-g2-28
Construct 19	**CMVp-*mKate2*-Triplex-28-gRNA3-28-gRNA4-28-gRNA5-28-gRNA6-28**
Abbreviation	CMVp-*mK*-Tr-(28-g-28)3-6
Construct 20	**CMVp-*mKate2*_EX1-[miRNA]-*mKate2*_EX2-Triplex-28-gRNA1-28-pA**
Abbreviation	CMVp-*mK*EX1-[miR]-*mK*EX2-Tr-28-g1-28
Construct 21	**CMVp-*mKate2*_EX1-[miRNA]-*mKate2*_EX2-Triplex-28-gRNA1-28-**4×**miRNA-BS-pA**
Abbreviation	CMVp-*mK*EX1-[miR]-*mK*EX2-Tr-28-g1-28-miR4xBS
Construct 22	**CMVp-*ECFP*-Triplex-28-**8×**miRNA-BS-28-pA**
Abbreviation	CMVp-ECFP-Tr-28-miR8xBS-28
Construct 23	**CMVp-*mKate2*-Triplex-28-gRNA3-28**
Abbreviation	CMVp-*mKate2*-Tr-28-g3-28
Construct 24	**CMVp-*mKate2*-Triplex-28-gRNA4-28**
Abbreviation	CMVp-*mKate2*-Tr-28-g4-28
Construct 25	**CMVp-*mKate2*-Triplex-28-gRNA5-28**
Abbreviation	CMVp-*mKate2*-Tr-28-g5-28
Construct 26	**CMVp-*mKate2*-Triplex-28-gRNA6-28**
Abbreviation	CMVp-*mKate2*-Tr-28-g6-28
Construct 27	**P2-*EYFP*-pA**
Abbreviation	P2-*EYFP*

Table A.4: Sequences of DNA constructs used in building Integrated CRISPR/Cas circuits

Construct	Sequence (gene sequences start with Kozak following the start codon)
dCas9-3xNLS-VP64-3'LTR (Construct 1)	GCCACCATGGACAAGAAGTACTCCATTGGGCTCGCCATCGGCACA AACAGCGTCGGCTGGGCCGTCATTACGGACGAGTACAAGGTGCCG AGCAAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATA AAGAAGAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACG GCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACC CGCAGAAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAAT GAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAG TCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATC TTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCA ACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAG GCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAA TTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAAC AGCGATGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATC AGCTTTTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCA AAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAA ACCTCATCGCACAGCTCCCTGGGGAGAAGAAGAACGGCCTGTTTG GTAATCTTATCGCCCTGTCACTCGGGCTGACCCCCAACTTTAAATC TAACTTCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGA CACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGA CCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGC CATTCTGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAA AGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCA CCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCC TGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTA CGCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAA ATTTATTAAGCCCATCTTGGAAAAAATGGACGGCACCGAGGAGCT GCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCAC TTTCGACAATGGAAGCATCCCCCACCAGATTCACCTGGGCGAACT GCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAA AGATAACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACC CTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTCGCGTG GATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGA GGAAGTCGTGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAG GATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCC TAAACACTCTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTC ACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTC CTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAG ACGAACCGGAAAGTTACCGTGAAACAGCTCAAAGAAGACTATTTC AAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAG GATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAA TCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACA TTCTTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGG AGATGATTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACG ACAAAGTCATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGG GGCGGCTGTCAAGAAAACTGATCAATGGGATCCGAGACAAGCAGA GTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCA ACCGGAACTTCATGCAGTTGATCCATGATGACTCTCTCACCTTTAA GGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCT

dCas9-3xNLS-VP64-3'LTR (Construct 1) (contd. . .)	TCACGAGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAA GGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGT AATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCG AGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAG GATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAAT CCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAA GCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGCCAC CATGGACAAGAAGTACTCCATTGGGCTCGCCATCGGCACAAACAG CGTCGGCTGGGCCGTCATTACGGACGAGTACAAGGTGCCGAGCAA AAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAA GAACCTCATTGGCGCCCTCCTGTTCGACTCCGGGGAGACGGCCGA AGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAG AAAGAATCGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGAT GGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAGTCCTT TTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGG CAATATCGTGGACGAGGTGGCGTACCATGAAAAGTACCCAACCAT ATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGA CTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAATTTCG GGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGA TGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTT TTCGAAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCA ATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTC ATCGCACAGCTCCCTGGGGAGAAGAAGAACGGCCTGTTTGGTAAT CTTATCGCCCTGTCACTCGGGCTGACCCCCAACTTTAAATCTAACT TCGACCTGGCCGAAGATGCCAAGCTTCAACTGAGCAAAGACACCT ACGATGATGATCTCGACAATCTGCTGGCCCAGATCGGCGACCAGT ACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTC TGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAAAGCTC CGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCACCAAG ACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGA AGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTACGCCG GATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTTA TTAAGCCCATCTTGGAAAAAATGGACGGCACCGAGGAGCTGCTGG TAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCG ACAATGGAAGCATCCCCCACCAGATTCACCTGGGCGAACTGCACG CTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAAAGATAA CAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTACTA TGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTCGCGTGGATGAC TCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGT CGTGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAGGATGAC TAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACAC TCTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGG TCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTCCTGTCTG GAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACC GGAAAGTTACCGTGAAACAGCTCAAAGAAGACTATTTCAAAAAGA TTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGATCGCT TCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAATCATTAA AGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGA GGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGAT TGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTC ATGAAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTG TCAAGAAAACTGATCAATGGGATCCGAGACAAGCAGAGTGGAAAG ACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACT TCATGCAGTTGATCCATGATGACTCTCTCACCTTTAAGGAGGACAT CCAGAAAGCACAAGTTTCTGGCCAGGGGGACAGTCTTCACGAG

dCas9-3xNLS-VP64-3'LTR (Construct 1) (contd. . .)	CACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGGGAATA CTGCAGACCGTTAAGGTCGTGGATGAACTCGTCAAAGTAATGGGA AGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAAC CAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAG AGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCTTAAG GAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTAC CTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAA CTGGACATCAATCGGCTCTCCGACTACGACGTGGATGCCATCGTG CCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGA CAAGATCCGATAAAAATAGAGGGAAGAGTGATAACGTCCCCTCAG AAGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGA ACGCCAAACTGATCACACAACGGAAGTTCGATAATCTGACTAAGG CTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCA AAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCACGTGGCCC AAATTCTCGATTCACGCATGAACACCAAGTACGATGAAAATGACA AACTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGG TCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTGAGAGAGA TCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGG TAGGCACTGCACTTATCAAAAAATATCCCAAGCTTGAATCTGAATT TGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGC AAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTT TTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGGCC AATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAA ACAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTC CGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACC GAAGTACAGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAA AGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCC AAGAAATACGGCGGATTCGATTCTCCTACAGTCGCTTACAGTGTA CTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAAAACTCAAA AGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGC TTCGAAAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAA GAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCT TTGAGCTTGAAAACGGCCGGAAACGAATGCTCGCTAGTGCGGGCG AGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATACGTTA ATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCC CGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAACACAAACA CTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAG AGTGATCCTCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTAC AATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATT ATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCA AGTACTTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAA AGGAGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGC TCTATGAAACAAGAATCGACCTCTCTCAGCTCGGTGGAGACAGCA GGGCTGACGGGCCCTCACTGGGTTCAGGGTCACCCAAGAAGAAAC GCAAAGTCGAGGATCCAAAGAAGAAAAGGAAGGTTGAAGACCCCA AGAAAAAGAGGAAGGTGGATGGGATCGGCTCAGGCAGCAACGGCG GTGGAGGTTCAGACGCTTTGGACGATTTCGATCTCGATATGCTCG GTTCTGACGCCCTGGATGATTTCGATCTGGATATGCTCGGCAGCG ACGCTCTCGACGATTTCGACCTCGACATGCTCGGGTCAGATGCCTT GGATGATTTTGACCTGGATATGCTCTCATGATGA

PGK1p-Csy4-pA (Construct 2)	GCCACCATGAAATCTTCTCACCATCACCATCACCATGAAAACCTGT ACTTCCAATCCAATGCAGCTAGCGACCACTATCTGGACATCAGACT GAGGCCCGATCCTGAGTTCCCTCCCGCCCAGCTGATGAGCGTGCT GTTTGGCAAGCTGCATCAGGCTCTGGTCGCCCAAGGCGGAGACAG AATCGGCGTGTCCTTCCCCGACCTGGACGAGTCCCGGAGTCGCCT GGGCGAGCGGCTGAGAATCCACGCCAGCGCAGACGATCTGCGCGC CCTGCTGGCCCGGCCTTGGCTGGAGGGCCTGCGGGATCATCTGCA GTTTGGCGAGCCCGCCGTGGTGCCACACCCAACACCCTACCGCCA GGTGAGCCGCGTGCAGGCCAAGTCAAATCCCGAGAGACTGCGGCG GAGGCTGATGAGGCGACATGATCTGAGCGAGGAGGAGGCCAGAAA GAGAATCCCCGACACAGTGGCCAGAGCCCTGGATCTGCCATTTGT GACCCTGCGGAGCCAGAGCACTGGCCAGCATTTCAGACTGTTCAT CAGACACGGGCCCCTGCAGGTGACAGCCGAGGAGGGCGGATTTAC ATGCTATGGCCTGTCTAAAGGCGGCTTCGTGCCCTGGTTCTGA
mKate2-Triplex-28-gRNA1-28-pA (Construct 3)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGG CGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCT CATCTACAACGTCAAGATCAGAGGGGTGAACTTCCCATCCAACGG CCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGA GATGCTGTACCCCGCTGACGGCGGCCTGGAAGGCAGAAGCGACAT GGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACTTGAA GACCACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCC CGGCGTCTACTATGTGGACAGAAGACTGGAAAGAATCAAGGAGGC CGACAAAGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAG ATACTGCGACCTCCCTAGCAAACTGGGGCACAAACTTAATTGATA AACCGGTGATTCGTCAGTAGGGTTGTAAAGGTTTTTCTTTTCCTG AGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTCC CTAGCTTTAAAAAAAAAAAAGCAAAACTCACCGAGGCAGTTCCAT AGGATGGCAAGATCCTGGTATTGGTCTGCGAGTTCACTGCCGTAT AGGCAGCTAAGAAATAGTCGCGTGTAGCGAAGCAGTTTTAGAGCT AGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAA AAGTGGCACCGAGTCGGTGCTTTTTTTCGTTCACTGCCGTATAGG CAGCTAAGAAACAAACAGGAATCGAATGCAACCGGCGCAGGAACA CTGCCAGCGCATCAACCCCGGG
mKate2_EX1-[28-gRNA1-28]$_{HSV1}$-mKate2_EX2-pA (Construct 4)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGTAAGTGTTCACTGCCGTATAGGCAGCTAAGAAATAGTCGC GTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCT TTTTTTCGTTCACTGCCGTATAGGCAGCTAAGAAAGAGGGAGTCG AGTCTTCTTTTTTTTTTTCACAGGGCTTCACATGGGAGAGAGTCAC CACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGGACACCAG CCTCCAGGACGGCTGCCTCATCTACAACGTCAAGATCAGAGGGGT GAACTTCCCATCCAACGGCCCTGTGATGCAGAAGAAAACACTCGG CTGGGAGGCCTCCACCGAGATGCTGTACCCCGCTGACGGCGGCCT GGAAGGCAGAAGCGACATGGCCCTGAAGCTCGTGGGCGGGGGCCA

(Construct 4) (contd. . .)	CCTGATCTGCAACTTGAAGACCACATACAGATCCAAGAAACCCGCT AAGAACCTCAAGATGCCCGGCGTCTACTATGTGGACAGAAGACTG GAAAGAATCAAGGAGGCCGACAAAGAGACCTACGTCGAGCAGCAC GAGGTGGCTGTGGCCAGATACTGCGACCTCCCTAGCAAACTGGGG CACAAACTTAATTGA
P1-EYFP-pA (Construct 5)	GCTAGCCATGCTTCGCTACACGCGACTATTAATATTTTCAGGCTAG CCATGCTTCGCTACACGCGACTATTAATATTTTCAGGCTAGCCATG CTTCGCTACACGCGACTATTAATATTTTCAGGCTAGCCATGCTTCG CTACACGCGACTATTAATATTTTCAGGCTAGCCATGCTTCGCTACA CGCGACTATTAATATTTTCAGGCTAGCCATGCTTCGCTACACGCGA CTATTAATATTTTCAGGCTAGCCATGCTTCGCTACACGCGACTATT AATATTTTCAGGCTAGCCATGCTTCGCTACACGCGACTATTAATAT TTTCAGGCTAGCCATGCTTCGCTACACGCGACTATTAATATTTTCA GGCTAGCGGGGGGCTATAAAAGGGGGTGGGGGCGTTCGTCCTGCT ATCTAGCGTCGCGTTGACCATGGCGCCACCATGAGCAGCGGCGCC CTGCTGTTCCACGGCAAGATCCCCTACGTGGTGGAGATGGAGGGC GATGTGGATGGCCACACCTTCAGCATCCGCGGTAAGGGCTACGGC GATGCCAGCGTGGGCAAGGTGGATGCCCAGTTCATCTGCACCACC GGCGATGTGCCCGTGCCCTGGAGCACCCTGGTGACCACCCTGACC TACGGCGCCCAGTGCTTCGCCAAGTACGGCCCCGAGCTGAAGGAT TTCTACAAGAGCTGCATGCCCGATGGCTACGTGCAGGAGCGCACC ATCACCTTCGAGGGCGATGGCAATTTCAAGACCCGCGCCGAGGTG ACCTTCGAGAATGGCAGCGTGTACAATCGCGTGAAGCTGAATGGC CAGGGCTTCAAGAAGGATGGCCACGTGCTGGGCAAGAATCTGGAG TTCAATTTCACCCCCCACTGCCTGTACATCTGGGGCGATCAGGCCA ATCACGGCCTGAAGAGCGCCTTCAAGATCTGCCACGAGATCGCCG GCAGCAAGGGCGATTTCATCGTGGCCGATCACACCCAGATGAATA CCCCCATCGGCGGCGGCCCCGTGCACGTGCCCGAGTACCACCACA TGAGCTACCACGTGAAGCTGAGCAAGGATGTGACCGATCACCGCG ATAATATGAGCCTGACGGAGACCGTGCGCGCCGTGGATTGCCGCA AGACCTACCTGTAA
P2-ECFP-pA (Construct 6)	GCTAGCCCAGGACAGTACTCCGACTTACTTAATATTTTCAGGCTAG CCCAGGACAGTACTCCGACTTACTTAATATTTTCAGGCTAGCCCAG GACAGTACTCCGACTTACTTAATATTTTCAGGCTAGCCCAGGACAG TACTCCGACTTACTTAATATTTTCAGGCTAGCCCAGGACAGTACTC CGACTTACTTAATATTTTCAGGCTAGCCCAGGACAGTACTCCGACT TACTTAATATTTTCAGGCTAGCCCAGGACAGTACTCCGACTTACTT AATATTTTCAGGCTAGCCCAGGACAGTACTCCGACTTACTTAATAT TTTCAGGCTAGCGGGGGGCTATAAAAGGGGGTGGGGGCGTTCGTC CTGCTATCTAGCGTCGCGTTGACCATGGTGAGCAAGGGCGAGGAG CTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGAC GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGC AAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTGGG GCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACG ACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGG TGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGG GCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG AGTACAACGCCATCAGCGACAACGTCTATATCACCGCCGACAAGC AGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCG AGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCC CCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGA GCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC

(Construct 6) (contd. . .)	ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCG GCATGGACGAGCTGTACAAGTAA
mKate2_EX1-[28-gRNA1-28]$_{consensus}$-mKate2_EX2-pA (Construct 8)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGTAAGTGTTCACTGCCGTATAGGCAGCTAAGAAATAGTCGC GTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCT TTTTTTCGTTCACTGCCGTATAGGCAGCTAAGAAATACTAACTTCG AGTCTTCTTTTTTTTTTTCACAGGGCTTCACATGGGAGAGAGTCAC CACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGGACACCAG CCTCCAGGACGGCTGCCTCATCTACAACGTCAAGATCAGAGGGGT GAACTTCCCATCCAACGGCCCTGTGATGCAGAAGAAAACACTCGG CTGGGAGGCCTCCACCGAGATGCTGTACCCCGCTGACGGCGGCCT GGAAGGCAGAAGCGACATGGCCCTGAAGCTCGTGGGCGGGGGCCA CCTGATCTGCAACTTGAAGACCACATACAGATCCAAGAAACCCGCT AAGAACCTCAAGATGCCCGGCGTCTACTATGTGGACAGAAGACTG GAAAGAATCAAGGAGGCCGACAAAGAGACCTACGTCGAGCAGCAC GAGGTGGCTGTGGCCAGATACTGCGACCTCCCTAGCAAACTGGGG CACAAACTTAATTGACCCGGG
mKate2_EX1-[28-gRNA1-28]$_{snoRNA2}$-mKate2_EX2-pA (Construct 9)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGTAAGTGTTCATTTCTCAAAAGACCCTAATGTTCTTCCTTTA CAGGAATGAATACTGTGCATGGACCAATGATGACTTCCATACATG CATTCCTTGGAAAGCTGAACAAAATGAGTGGGAACTCTGTACTAT CATCTTAGTTGAACTGAGGTCCGGATCCGTTCACTGCCGTATAGG CAGCTAAGAAATAGTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGCTTTTTTTCAGATCTGTTCACTGCCGTATA GGCAGCTAAGAAATCTAGATGGATCGATGATGACTTCCATATATA CATTCCTTGGAAAGCTGAACAAAATGAGTGAAAACTCTATACCGT CATTCTCGTCGAACTGAGGTCCAACCGGTGCACATTACTCCAACAG GGGCTAGACAGAGAGGGCCAACATTGATTCGTTGACATGGGTGGC TGCAGTACTAACTTCGAGTCTTCTTTTTTTTTTTCACAGGGCTTCA CATGGGAGAGAGTCACCACATACGAAGACGGGGGCGTGCTGACCG CTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGT CAAGATCAGAGGGGTGAACTTCCCATCCAACGGCCCTGTGATGCA GAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGATGCTGTACCC CGCTGACGGCGGCCTGGAAGGCAGAAGCGACATGGCCCTGAAGCT CGTGGGCGGGGGCCACCTGATCTGCAACTTGAAGACCACATACAG ATCCAAGAAACCCGCTAAGAACCTCAAGATGCCCGGCGTCTACTAT GTGGACAGAAGACTGGAAAGAATCAAGGAGGCCGACAAAGAGACC TACGTCGAGCAGCACGAGGTGGCTGTGGCCAGATACTGCGACCTC CCTAGCAAACTGGGGCACAAACTTAATTGA

mKate2-Triplex-HHRibo-gRNA1-HDVRibo-pA (Construct 13)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGG CGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCT CATCTACAACGTCAAGATCAGAGGGGTGAACTTCCCATCCAACGG CCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGA GATGCTGTACCCCGCTGACGGCGGCCTGGAAGGCAGAAGCGACAT GGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACTTGAA GACCACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCC CGGCGTCTACTATGTGGACAGAAGACTGGAAAGAATCAAGGAGGC CGACAAAGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAG ATACTGCGACCTCCCTAGCAAACTGGGGCACAAACTTAATTGATA AACCGGTGATTCGTCAGTAGGGTTGTAAAGGTTTTTCTTTTCCTG AGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTCC CTAGCTTTAAAAAAAAAAAAGCAAAACGACTACTGATGAGTCCGT GAGGACGAAACGAGTAAGCTCGTCTAGTCGCGTGTAGCGAAGCAG TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTGGCCGGCATGG TCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCAT GGCGAATGGGACCCCGGG
mKate2-HHRibo-gRNA1-HDVRibo-pA (Construct 14)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGG CGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCT CATCTACAACGTCAAGATCAGAGGGGTGAACTTCCCATCCAACGG CCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGA GATGCTGTACCCCGCTGACGGCGGCCTGGAAGGCAGAAGCGACAT GGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACTTGAA GACCACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCC CGGCGTCTACTATGTGGACAGAAGACTGGAAAGAATCAAGGAGGC CGACAAAGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAG ATACTGCGACCTCCCTAGCAAACTGGGGCACAAACTTAATTGATA AACCGGTCGACTACTGATGAGTCCGTGAGGACGAAACGAGTAAGC TCGTCTAGTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCAC CGAGTCGGTGCTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGC GCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACCCCGGG
HHRibo-gRNA1-HDVRibo-pA (Construct 15)	CGACTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCTA GTCGCGTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAA AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG GTGCTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCT GGGCAACATGCTTCGGCATGGCGAATGGGACCCCGGG

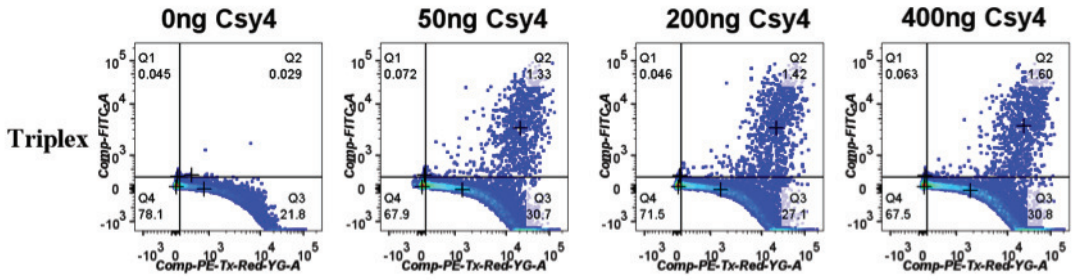
mKate2-Triplex-28-gRNA3-28-gRNA4-28-gRNA5-28-gRNA6-28-pA (Construct 19)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGG CGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCT CATCTACAACGTCAAGATCAGAGGGGTGAACTTCCCATCCAACGG CCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGA GATGCTGTACCCCGCTGACGGCGGCCTGGAAGGCAGAAGCGACAT GGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACTTGAA GACCACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCC CGGCGTCTACTATGTGGACAGAAGACTGGAAAGAATCAAGGAGGC CGACAAAGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAG ATACTGCGACCTCCCTAGCAAACTGGGGCACAAACTTAATTGATA AACCGGTGATTCGTCAGTAGGGTTGTAAAGGTTTTTCTTTTCCTG AGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTCC CTAGCTTTAAAAAAAAAAAAGCAAAACTCACCGAGGCAGTTCCAT AGGATGGCAAGATCCTGGTATCGGTCTGCGAGTTCACTGCCGTAT AGGCAGCTAAGAAAGCTAGCGTGTACTCTCTGAGGTGCTCGTTTT AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCGTTCACTGCCG TATAGGCAGCTAAGAAAAGGTGACGCAGATAAGAACCAGTTGTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAA CTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCGTTCACTGCC GTATAGGCAGCTAAGAAACAGGGCATCAAGTCAGCCATCAGCGTT TTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA ACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCGTTCACTGC CGTATAGGCAGCTAAGAAAAGTCGGGAGTCACCCTCCTGGAAACG TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCGTTCACT GCCGTATAGGCAGCTAAGAAACCCGGG
CMVp-mK$_{EX1}$-[miR]-mK$_{EX2}$-Tr-28-g1-28 (Construct 20)	GCCACCATGGTGTCTAAGGGCGAAGAGCTGATTAAGGAGAACATG CACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTC AAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAG ACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCC TTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCA TCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCC TGAGGTAAGTGTGCTCGCTTCGGCAGCACATATACTATGTTGAAT GAGGCTTCAGTACTTTACAGAATCGTTGCCTGCACATCTTGGAAA CACTTGCTGGGATTACTTCTTCAGGTTAACCCAACAGAAGGCTCG AGTGCTGTTGACAGTGAGCGCCGCTTGAAGTCTTTAATTAAATAG TGAAGCCACAGATGTATTTAATTAAAGACTTCAAGCGGTGCCTAC TGCCTCGGAGAATTCAAGGGGCTACTTTAGGAGCAATTATCTTGT TTACTAAAACTGAATACCTTGCTATCTCTTTGATACATTTTTACAA AGCTGAATTAAAATGGTATAAATTAAATCACTTTTTTCAATTGTAC TAACTTCGAGTCTTCTTTTTTTTTTTCACAGGGCTTCACATGGGAG AGAGTCACCACATACGAAGACGGGGGGCGTGCTGACCGCTACCCAG GACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAGATCA GAGGGGTGAACTTCCCATCCAACGGCCCTGTGATGCAGAAGAAAA CACTCGGCTGGGAGGCCTCCACCGAGATGCTGTACCCCGCTGACG GCGGCCTGGAAGGCAGAAGCGACATGGCCCTGAAGCTCGTGGGCG GGGGCCACCTGATCTGCAACTTGAAGACCACATACAGATCCAAGA AACCCGCTAAGAACCTCAAGATGCCCGGCGTCTACTATGTGGACA

(Construct 20) (contd. . .)	GAAGACTGGAAAGAATCAAGGAGGCCGACAAAGAGACCTACGTCG AGCAGCACGAGGTGGCTGTGGCCAGATACTGCGACCTCCCTAGCA AACTGGGGCACAAACTTAATTGATAAACCGGTGATTCGTCAGTAG GGTTGTAAAGGTTTTTCTTTTCCTGAGAAAACAACCTTTTGTTTTC TCAGGTTTTGCTTTTTGGCCTTTCCCTAGCTTTAAAAAAAAAAAAG CAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTAT CGGTCTGCGAGTTCACTGCCGTATAGGCAGCTAAGAAATAGTCGC GTGTAGCGAAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCT TTTTTTCCCGCTTGAAGTCTTTAATTAAACCGCTTGAAGTCTTTAA TTAAACCGCTTGAAGTCTTTAATTAAACCGCTTGAAGTCTTTAATT AAAGTTCACTGCCGTATAGGCAGCTAAGAAACCCGGG
ECFP-Triplex-28-8×miRNA-BS-28-pA (Construct 22)	GCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTC AGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGG CCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTGCTTCGCCC GCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACAC CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA CGGCAACATCCTGGGGCACAAGCTGGAGTACAACGCCATCAGCGA CAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGC CAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAG CAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA GTAAACCGGTGATTCGTCAGTAGGGTTGTAAAGGTTTTTCTTTTC CTGAGAAAACAACCTTTTGTTTTCTCAGGTTTTGCTTTTTGGCCTT TCCCTAGCTTTAAAAAAAAAAAAGCAAAACTCACCGAGGCAGTTC CATAGGATGGCAAGATCCTGGTATCGGTCTGCGAGTTCACTGCCG TATAGGCAGCTAAGAAACCGCTTGAAGTCTTTAATTAAACCGCTT GAAGTCTTTAATTAAACCGCTTGAAGTCTTTAATTAAACCGCTTGA AGTCTTTAATTAAACCTCTGGCCACATCGGTTCCTGCTCCGCTTGA AGTCTTTAATTAAACCGCTTGAAGTCTTTAATTAAACCGCTTGAAG TCTTTAATTAAACCGCTTGAAGTCTTTAATTAAAGTTCACTGCCGT ATAGGCAGCTAAGAAACCCGGG
gRNAs	where NNNNNNNNNNNNNNNNNNN is one of the following:
gRNA1	GAGTCGCGTGTAGCGAAGCA
gRNA2	GTAAGTCGGAGTACTGTCCT
gRNA3	GTGTACTCTCTGAGGTGCTC
gRNA4	GACGCAGATAAGAACCAGTT
gRNA5	GCATCAAGTCAGCCATCAGC
gRNA6	GGAGTCACCCTCCTGGAAAC

Appendix B

Figures

A. Triplex/gRNA expression mechanism



B. Intron/Csy4 expression mechanism

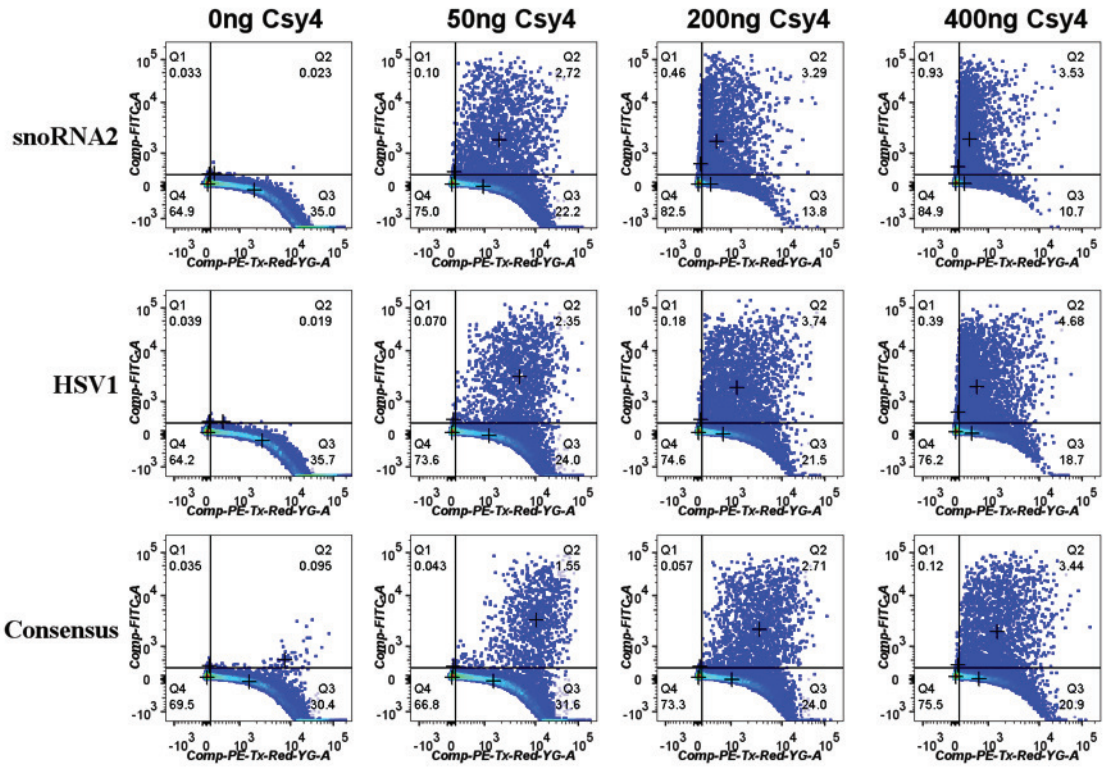


Figure B-1: Flow cytometry data corresponding to: (A) the 'triplex/Csy4' strategy (Figure 4-1) and (B) the 'intron/Csy4' (Figure 4-2) strategy for generating functional gRNAs from RNAP II transcripts.
Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP).
Triplex: Construct #3 (CMVp-mK-Tr-28-g1-28, 1 μg).
Consensus, snoRNA2, and HSV1: Constructs #8-10, respectively (CMVp-mK_{EX1}-[28-g1-28]$_{\text{'intron type'}}$-mK_{EX2} with the corresponding intron sequences flanking the gRNA and Csy4 recognition sites ('28')). These plasmids were transfected at 1 μg.
In addition, the amount of the Csy4-expressing plasmid (Construct #2) transfected in each sample is indicated. Other plasmids transfected include Construct #1 (taCas9, 1 μg) and #5 (P1-EYFP, 1 μg).

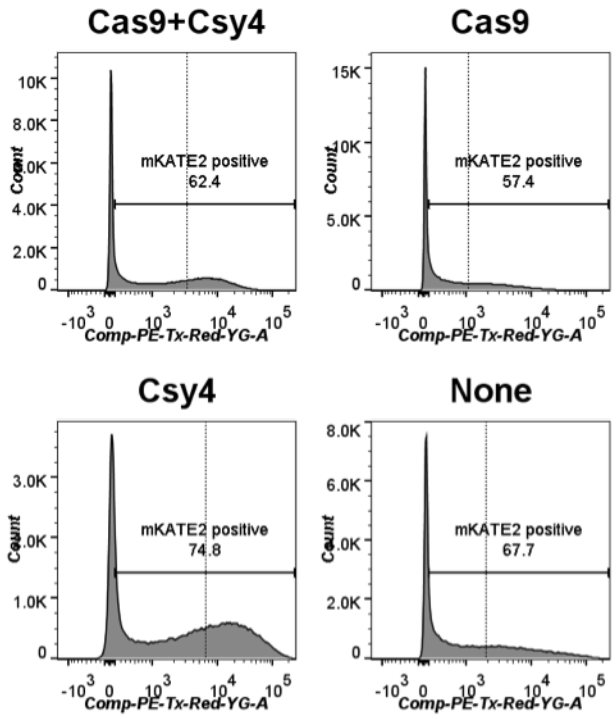


Figure B-2: Flow cytometry data corresponding to Figure 4-1-B to analyze how various combinations of Csy4 and taCas9 affect expression of the harboring *mKate2* gene for the CMVp-mK-Tr-28-g1-28 architecture.

Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2).

All samples contained Construct #3 (CMVp-mK-Tr-28-g1-28, 1 μg). Construct #1 (taCas9, 1 μg) and Construct #2 (Csy4, 100 ng) were applied as indicated.

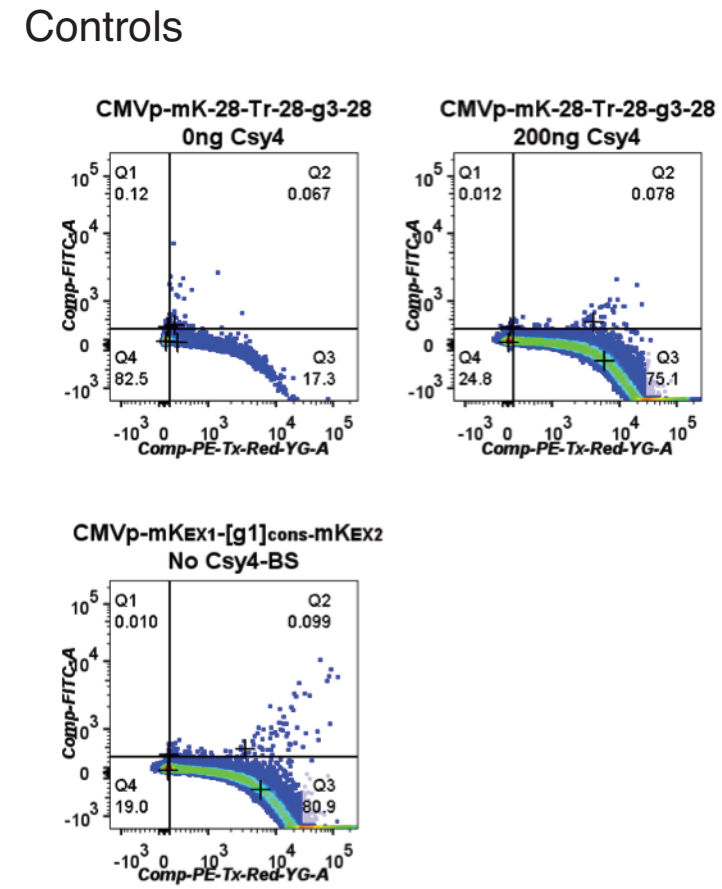


Figure B-3: Flow cytometry data providing various controls to demonstrate minimal non-specific activation of the P1 promoter by gRNA3 (top two panels) and minimal *EYFP* activation from the promoter P1 with intronic gRNA1 without Csy4 binding sites (bottom panel).

Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP).

The amount of Csy4 DNA transfected in each sample in the top two panels is indicated in the figure. The lower panel (CMVp-mK$_{EX1}$-[g1]$_{cons}$-mK$_{EX2}$) was tested in the absence of Csy4. Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg) and Construct #5 (P1-EYFP, 1 μg).

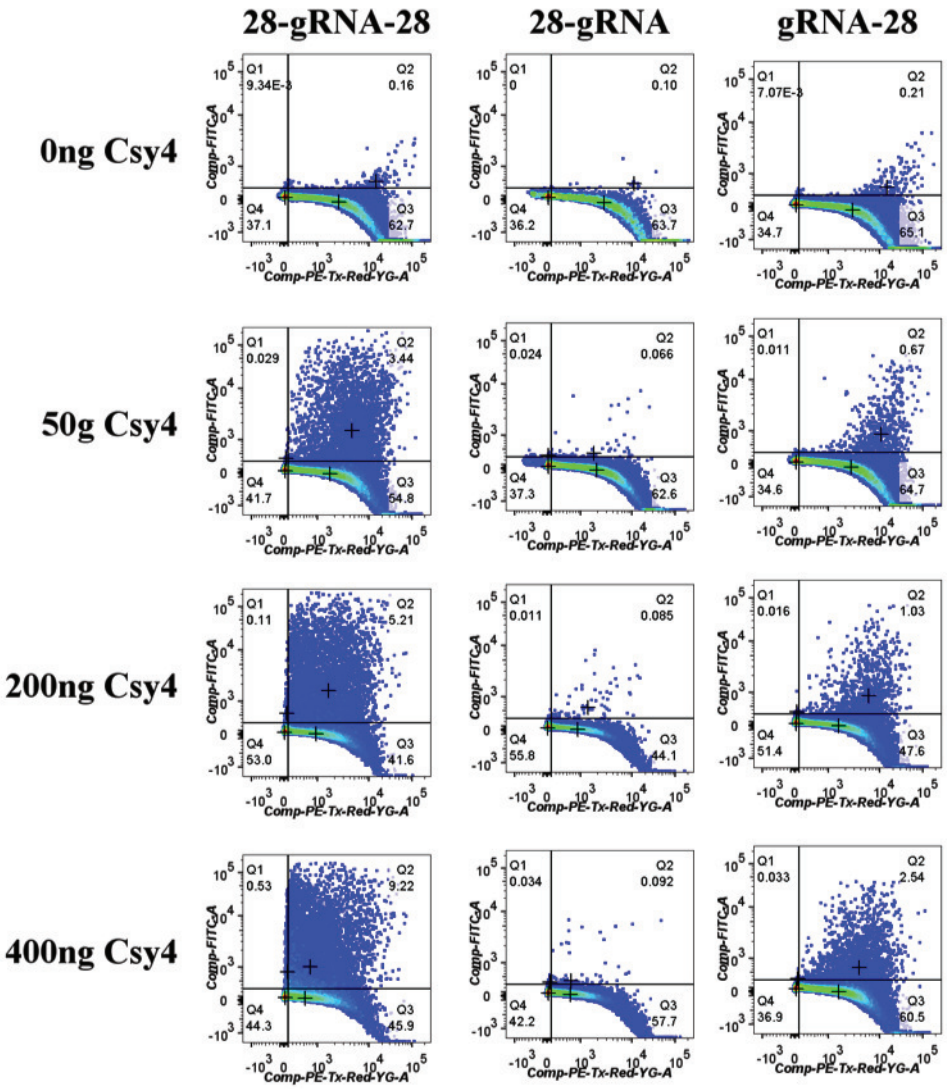


Figure B-4: Flow cytometry data corresponding to Figures 4-2-E, F to analyze how various configurations of Csy4 recognition sites flanking the gRNA within an intron affect CRISPR-TF activity.
Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP). '28-gRNA-28' is HSV1 intronic gRNA flanked by two Csy4 recognition sites (Construct #4, CMVp-mK$_{EX1}$-[28-g1-28]$_{HSV1}$-mK$_{EX2}$)
'28-gRNA' is HSV1 intronic gRNA with a 5' Csy4 recognition site only (construct #10, CMVp-mK$_{EX1}$-[28-g1]$_{HSV1}$-mK$_{EX2}$)
'gRNA-28' is HSV1 intronic gRNA with a 3' Csy4 recognition site only (construct #11, CMVp-mK$_{EX1}$-[g1-28]$_{HSV1}$-mK$_{EX2}$).
In addition, the amount of the Csy4-expressing plasmid transfected in each sample is indicated with each figure. Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg) and Construct #5 (P1-EYFP 1 μg).

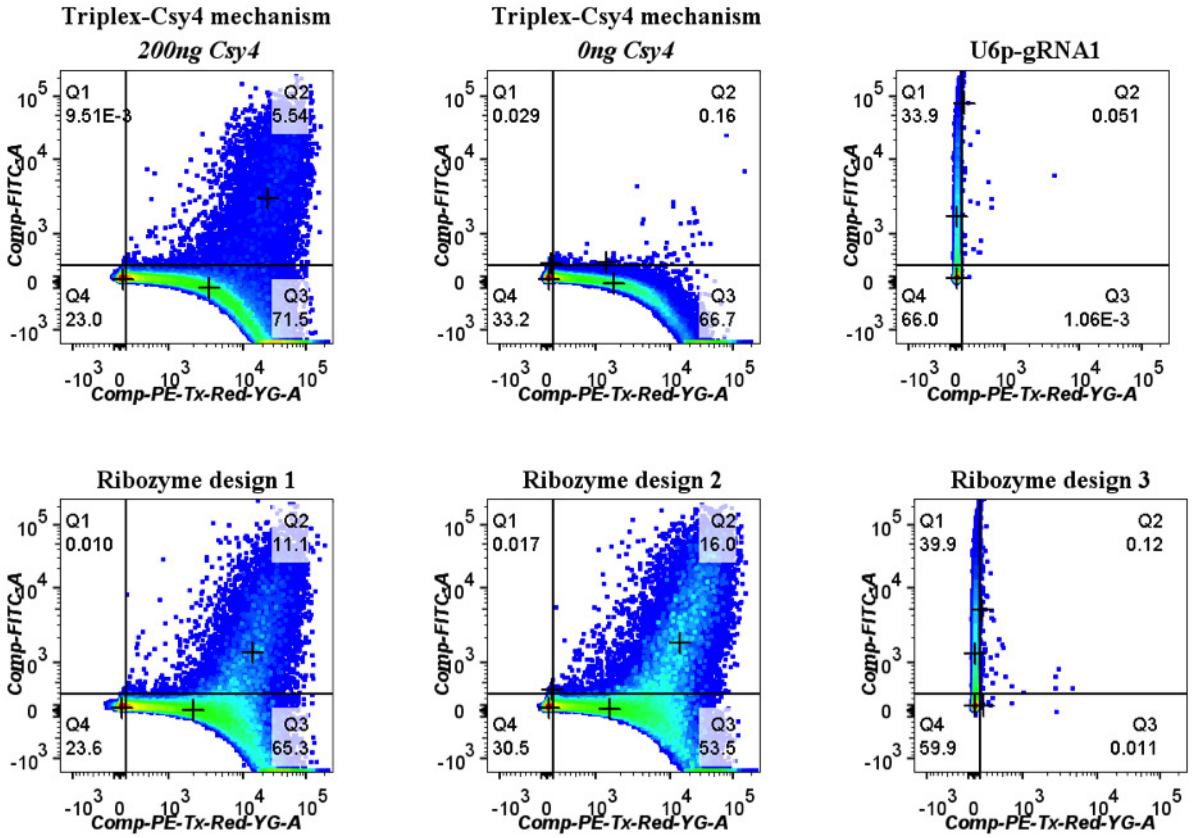


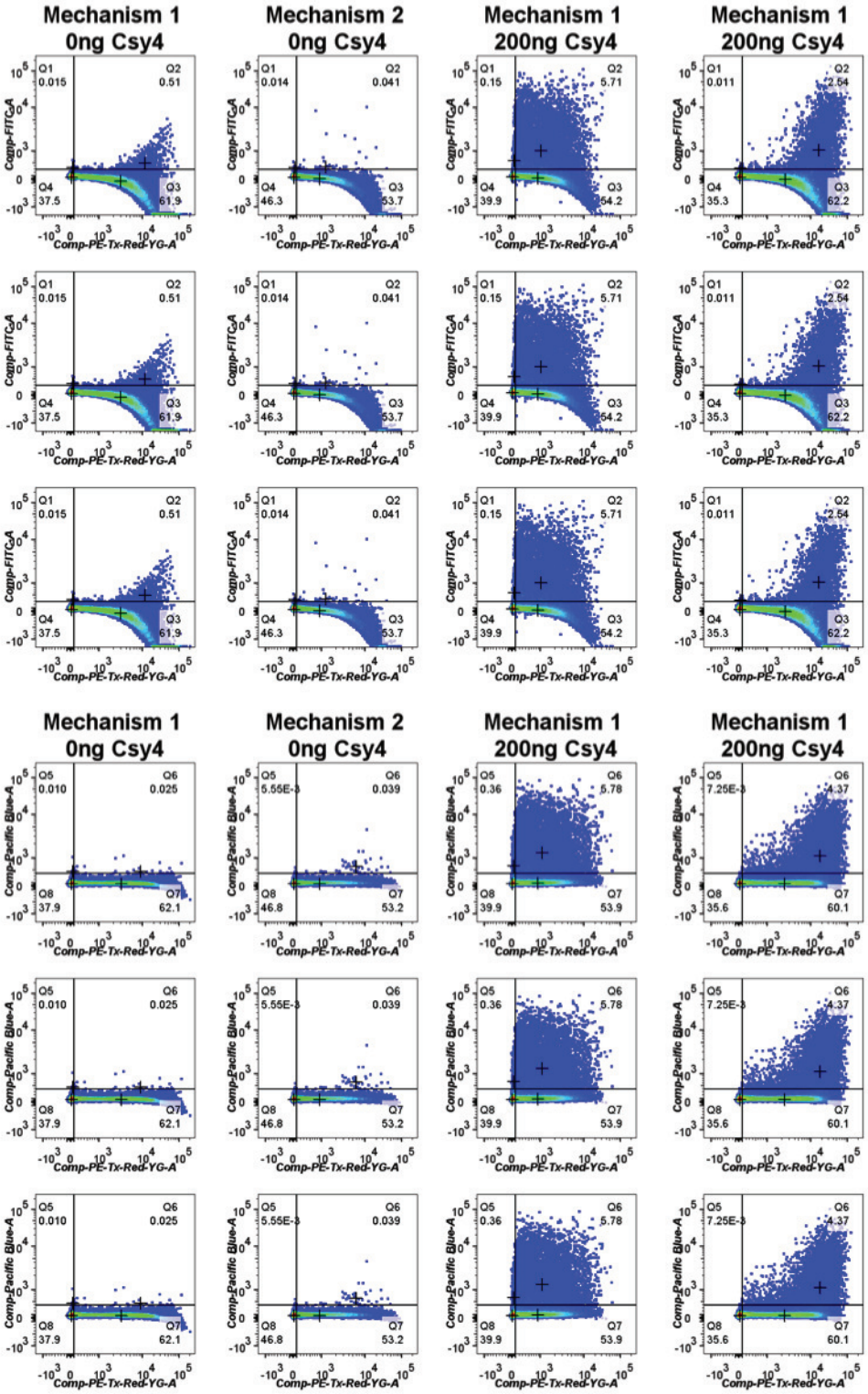
Figure B-5: Flow cytometry data corresponding to Figure 4-3.

Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP).

'Triplex-Csy4' mechanism contains Construct #3 (CMVp-mK-Tr-28-g1-28). Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg); Construct #5 (P1-EYFP); Construct #2 (Csy4, concentrations indicated).

'Ribozyme design 1' contains Construct #13 (CMVp-mK-Tr-HH-g1-HDV). Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg); Construct #5 (P1-EYFP, 1 μg).
'Ribozyme design 2' contains Construct #14 (CMVp-mK-HH-g1-HD). Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg); Construct #5 (P1-EYFP, 1 μg).
'Ribozyme design 3' contains Construct #15 (CMVp-HH-g1-HDV). Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg); Construct #5 (P1-EYFP, μg).
'U6p-gRNA1' contains Construct#7 (U6p-g1, 1 μg). Other plasmids transfected in this experiment include Construct #1 (taCas9, 1 μg).

Multiplexed expression of 2x gRNAs from a single transcript



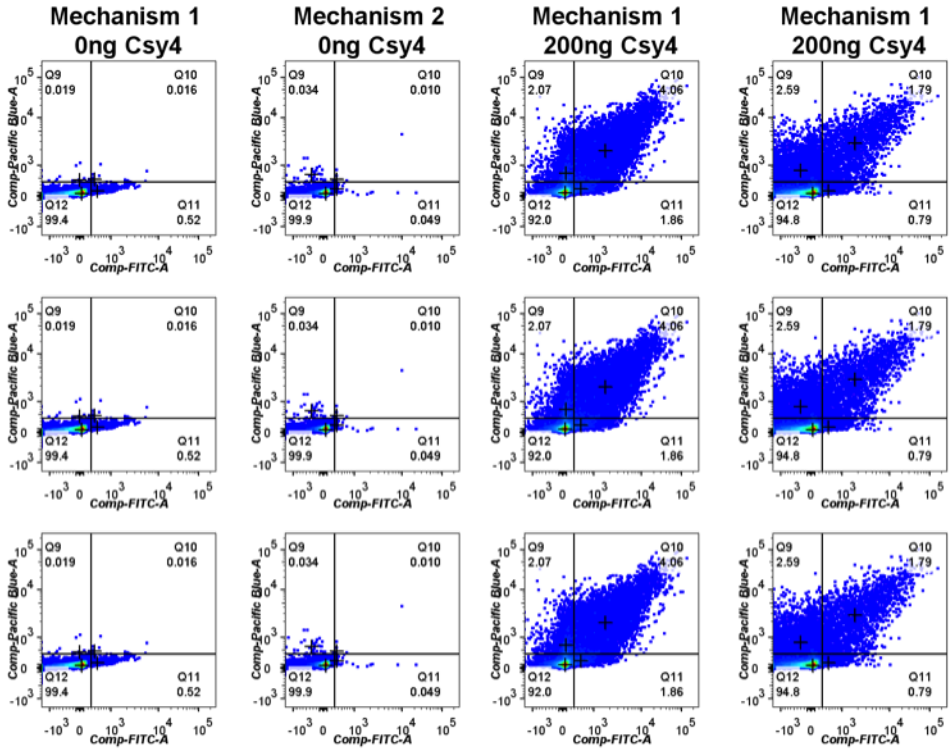


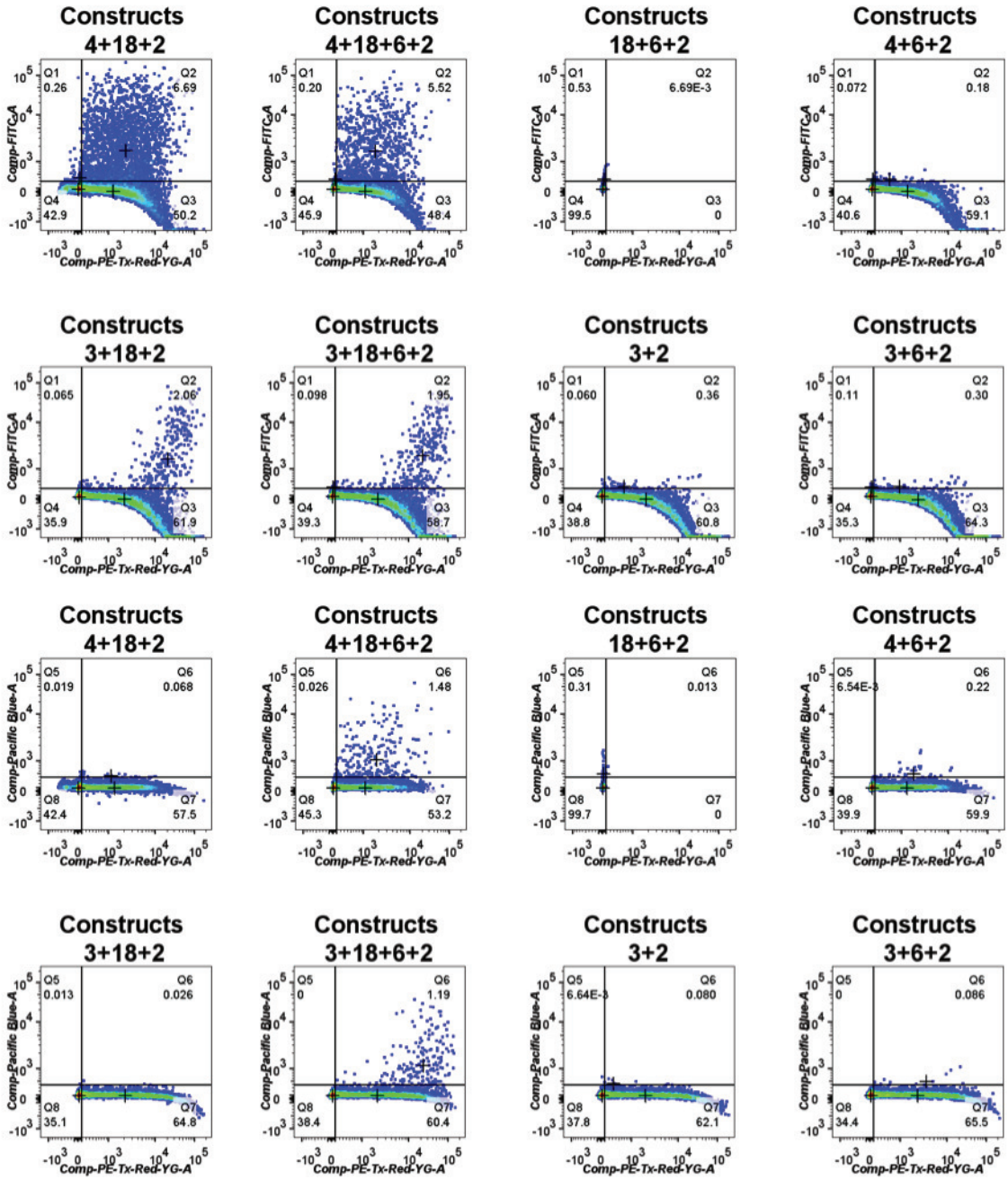
Figure B-6: Flow cytometry data corresponding to Figure 4-4.

Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP); *Comp-Pacific Blue-A* (ECFP)

'Mechanism 1' refers to the 'intron-triplex' architecture and contains Constructs #16 (CMVp-mK$_{EX1}$-[28-g1-28]$_{HSV1}$-mK$_{EX2}$-Tr-28-g2-28, 1 μg); #5 (P1-EYFP, 1 μg); #6 (P2-ECFP, 1 μg); and #1 (taCas9, 1 μg).

'Mechanism 2' refers to the 'tandem-triplex' architecture and contains Constructs #17 (CMVp-mK-Tr-28-g1-28-g2-28, 1 μg); #5 (P1-EYFP, 1 μg) and #6 (P2-ECFP, 1 μg); and #1 (taCas9, 1 μg).

In addition, the amount of Csy4-expressing plasmid DNA (Construct #2) transfected in each sample is indicated above each plot.


Constructs 4+18+2
Q1 0.26
Q2 6.69
Q4 42.9
Q3 50.2
Constructs 4+18+6+2
Q1 0.20
Q2 5.52
Q4 45.9
Q3 48.4
Constructs 18+6+2
Q1 0.53
Q2 6.69E-3
Q4 99.5
Q3 0
Constructs 4+6+2
Q1 0.072
Q2 0.18
Q4 40.6
Q3 59.1
Constructs 3+18+2
Q1 0.065
Q2 2.06
Q4 35.9
Q3 61.9
Constructs 3+18+6+2
Q1 0.098
Q2 1.95
Q4 39.3
Q3 58.7
Constructs 3+2
Q1 0.060
Q2 0.36
Q4 38.8
Q3 60.8
Constructs 3+6+2
Q1 0.11
Q2 0.30
Q4 35.3
Q3 64.3
Constructs 4+18+2
Q5 0.019
Q6 0.068
Q8 42.4
Q7 57.5
Constructs 4+18+6+2
Q5 0.026
Q6 1.48
Q8 45.3
Q7 53.2
Constructs 18+6+2
Q5 0.31
Q6 0.013
Q8 99.7
Q7 0
Constructs 4+6+2
Q5 6.54E-3
Q6 0.22
Q8 39.9
Q7 59.9
Constructs 3+18+2
Q5 0.013
Q6 0.026
Q8 35.1
Q7 64.8
Constructs 3+18+6+2
Q5 0
Q6 1.19
Q8 38.4
Q7 60.4
Constructs 3+2
Q5 6.64E-3
Q6 0.080
Q8 37.8
Q7 62.1
Constructs 3+6+2
Q5 0
Q6 0.086
Q8 34.4
Q7 65.5
Comp-FITC-A
Comp-Pacific Blue-A
Comp-PE-Tx-Red-YG-A

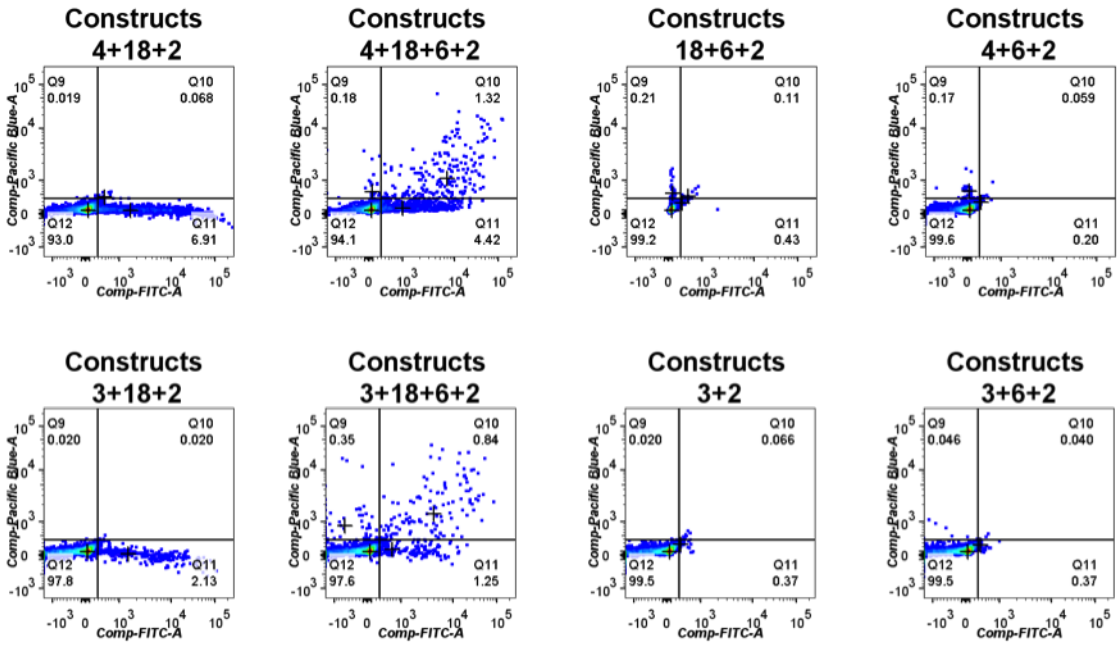


Figure B-7: Flow cytometry data corresponding to Figure 4-6.

Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP); *Comp-Pacific Blue-A* (ECFP)

All samples were transfected with the constructs listed in each plot title (1 μg each, Table A.3) and 200 ng of the Csy4-expressing plasmid (Construct #2).

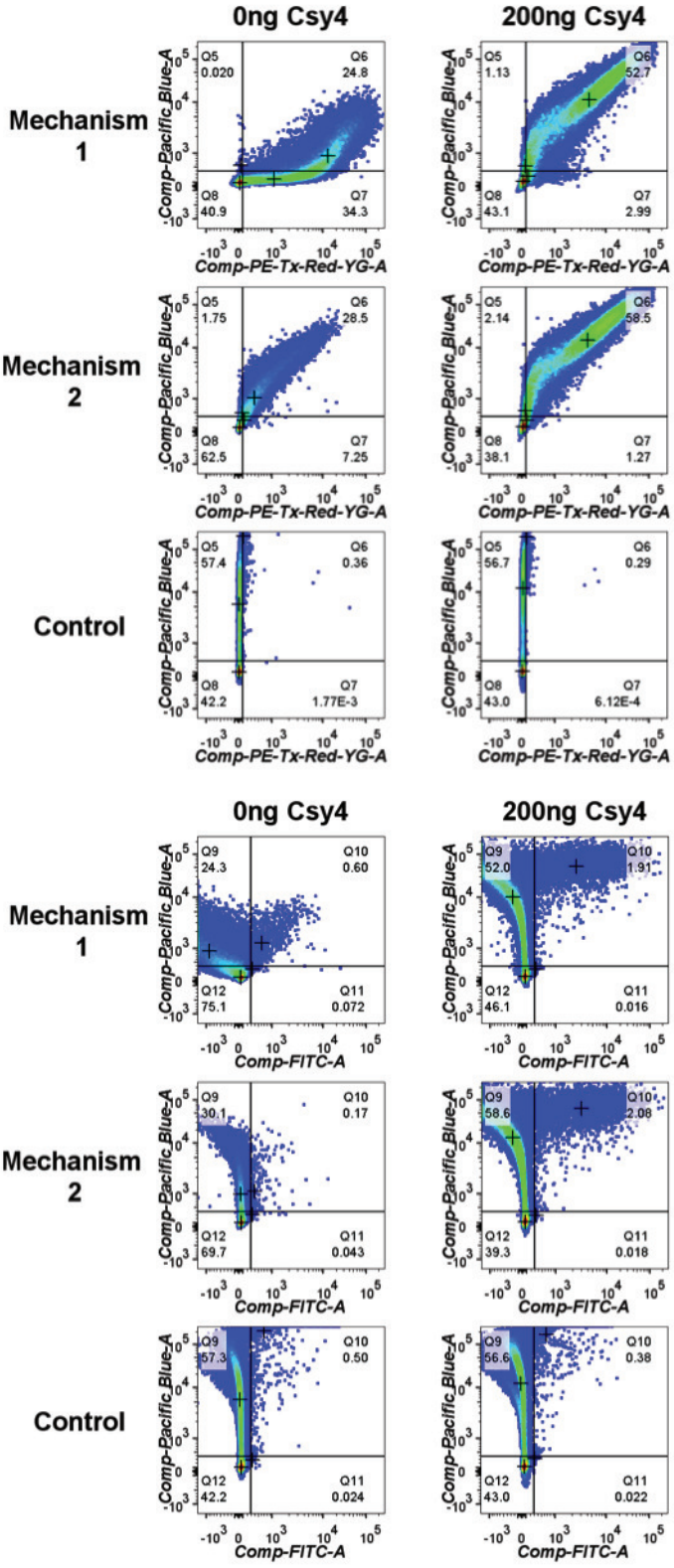

0ng Csy4
200ng Csy4
Mechanism 1
Mechanism 2
Control
Comp-Pacific Blue-A
Comp-PE-Tx-Red-YG-A
Comp-FITC-A
Q5 0.020
Q6 24.8
Q8 40.9
Q7 34.3
Q5 1.13
Q6 52.7
Q8 43.1
Q7 2.99
Q5 1.75
Q6 28.5
Q8 62.5
Q7 7.25
Q5 2.14
Q6 58.5
Q8 38.1
Q7 1.27
Q5 57.4
Q6 0.36
Q8 42.2
Q7 1.77E-3
Q5 56.7
Q6 0.29
Q8 43.0
Q7 6.12E-4
Q9 24.3
Q10 0.60
Q12 75.1
Q11 0.072
Q9 52.0
Q10 1.91
Q12 46.1
Q11 0.016
Q9 30.1
Q10 0.17
Q12 69.7
Q11 0.043
Q9 58.6
Q10 2.08
Q12 39.3
Q11 0.018
Q9 57.3
Q10 0.50
Q12 42.2
Q11 0.024
Q9 56.6
Q10 0.38
Q12 43.0
Q11 0.022

Figure B-8: Flow cytometry data corresponding to Figure 4-7.

Abbreviations: *Comp-PE-Tx-Red-YG-A* (mKate2); *Comp-FITC-A* (EYFP); *Comp-Pacific Blue-A* (ECFP).

'Mechanism 1' contains the following Constructs: #20 (CMVp-mK$_{Ex1}$-[miR]-mK$_{Ex2}$-Tr-28-g1-28); #22 (CMVp-ECFP-Tr-28-miR-8×BS-28); and #5 (P1-EYFP). These plasmids were transfected at a concentration of 1 μg each. This mechanism corresponds to the circuit diagram in Figure 4-7-A.

'Mechanism 2' contains the following Constructs: #21 (CMVp-mK$_{Ex1}$-[miR]-mK$_{Ex2}$-Tr-28-g1-28-miR-4×BS); #22 (CMVp-ECFP-Tr-28-miR-8×BS-28); and #5 (P1-EYFP). These plasmids were transfected at a concentration of 1 μg each. This mechanism corresponds to the circuit diagram in Figure 4-7-D.

'Control' samples contain Constructs #22 (CMVp-ECFP-Tr-28-miR-8×BS-28) and #5 (P1-EYFP) only. These plasmids were transfected at a concentration of 1 μg each.

In addition, the amount of Csy4-expressing plasmid (Construct #2) transfected in each sample is indicated above each plot.

B.1 Copyright

We thank members of the Lu lab, especially Ramiz Daniel and Ky Lowenhaupt, for helpful discussions. We thank Lila Wroblewska for an *mKate2* plasmid expressing intronic miRNA and Jeremy Wilusz and Courtney JnBaptiste for the cGFP_MALAT1_3' plasmid. The above work was supported by the Defense Advanced Research Projects Agency and the National Institutes of Health (DP2 OD008435 and P50 GM098792).

Bibliography

[1] Chaker N. Adra, Poppo H. Boer, and Michael W. McBurney. Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene*, 60(1):65 – 74, 1987.

[2] Mariette R. Atkinson, Michael A. Savageau, Jesse T. Myers, and Alexander J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in escherichia coli. *Cell*, 113(5):597 – 607, 2003.

[3] A. Audibert, D. Weil, and F. Dautry. In vivo kinetics of mrna splicing and transport in mammalian cells. *Molecular and Cellular Biology*, 22(19):6706–6718, 2002.

[4] S. Auslander, D. Auslander, M. Muller, M. Wieland, and M. Fussenegger. Programmable single-cell mammalian biocomputers. *Nature*, 487(7405):123–127, Jul 2012.

[5] Rodolphe Barrangou, Christophe Fremaux, HÃľlÃĺne Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A. Romero, and Philippe Horvath. Crispr provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819):1709–1712, 2007.

[6] Rodolphe Barrangou and Luciano A. Marraffini. Crispr-cas systems: Prokaryotes upgrade to adaptive immunity. *Molecular Cell*, 54(2):234 – 244, 2014.

[7] Subhayu Basu, Rishabh Mehreja, Stephan Thiberge, Ming-Tang Chen, and Ron Weiss. Spatiotemporal control of gene expression with pulse-generating networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(17):6355–6360, 2004.

[8] Roger R. Beerli, Birgit Dreier, and Carlos F. Barbas. Positive and negative regulation of endogenous genes by designed transcription factors. *Proceedings of the National Academy of Sciences*, 97(4):1495–1500, 2000.

[9] Roger R. Beerli, David J. Segal, Birgit Dreier, and Carlos F. Barbas. Toward controlling gene expression at will: Specific regulation of the erbb-2/her-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences*, 95(25):14628–14633, 1998.

[10] Y. Benenson. Biomolecular computing systems: principles, progress and potential. *Nat. Rev. Genet.*, 13(7):455–468, Jul 2012.

[11] Marina Bibikova, Kelly Beumer, Jonathan K. Trautman, and Dana Carroll. Enhancing gene targeting with designed zinc finger nucleases. *Science*, 300(5620):764, 2003.

[12] Marina Bibikova, Dana Carroll, David J. Segal, Jonathan K. Trautman, Jeff Smith, Yang-Gyun Kim, and Srinivasan Chandrasegaran. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Molecular and Cellular Biology*, 21(1):289–297, 2001.

[13] David Bikard, Wenyan Jiang, Poulami Samai, Ann Hochschild, Feng Zhang, and Luciano A. Marraffini. Programmable repression and activation of bacterial gene expression using an engineered crispr-cas system. *Nucleic Acids Research*, 41(15):7429–7437, 2013.

[14] M. C. Blake, R. C. Jambou, A. G. Swick, J. W. Kahn, and J. C. Azizkhan. Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. *Mol. Cell. Biol.*, 10(12):6632–6641, Dec 1990.

[15] Timothy M Block and James M Hill. The latency associated transcripts (lat) of herpes simplex virus: Still no end in sight. *Journal of Neurovirology*, 3(5):313–321, 1997.

[16] Alexander Bolotin, Benoit Quinquis, Alexei Sorokin, and S. Dusko Ehrlich. Clustered regularly interspaced short palindrome repeats (crisprs) have spacers of extrachromosomal origin. *Microbiology*, 151(8):2551–2561, 2005.

[17] Jerome Bonnet, Peter Yin, Monica E. Ortiz, Pakpoom Subsoontorn, and Drew Endy. Amplifying genetic logic gates. *Science*, 340(6132):599–603, 2013.

[18] Robert C. Brewster, Franz M. Weinert, Hernan G. Garcia, Dan Song, Mattias Rydenfelt, and Rob Phillips. The transcription factor titration effect dictates level of gene expression. *Cell*, 156(6):1312 – 1323, 2014.

[19] Stan J. J. Brouns, Matthijs M. Jore, Magnus Lundgren, Edze R. Westra, Rik J. H. Slijkhuis, Ambrosius P. L. Snijders, Mark J. Dickman, Kira S. Makarova, Eugene V. Koonin, and John van der Oost. Small crispr rnas guide antiviral defense in prokaryotes. *Science*, 321(5891):960–964, 2008.

[20] Stefano Cardinale and Adam Paul Arkin. Contextualizing context for synthetic biology – identifying causes of failure of synthetic biological systems. *Biotechnology Journal*, 7(7):856–866, 2012.

[21] Dana Carroll, J. Jason Morton, Kelly J. Beumer, and David J. Segal. Design, construction and in vitro testing of zinc finger nucleases. *Nat. Protocols*, 1(3):1329–1341, Oct 2006. Protocol.

[22] Lufen Chang and Michael Karin. Mammalian map kinase signalling cascades. *Nature*, 410(6824):37–40, Mar 2001.

[23] Mo Chen and James L. Manley. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol*, 10(11):741–754, Nov 2009.

[24] Xin Chen, Jian-min Wu, Klaus Hornischer, Alexander Kel, and Edgar Wingender. Tiprod: the tissue-specific promoter database. *Nucleic Acids Research*, 34(suppl 1):D104–D107, 2006.

[25] Albert W. Cheng, Haoyi Wang, Hui Yang, Linyu Shi, Yarden Katz, Thorold W. Theunissen, Sudharshan Rangarajan, Chikdu S. Shivalila, Daniel B. Dadon, and Rudolf Jaenisch. Multiplexed activation of endogenous genes by crispron, an rna-guided transcriptional activator system. *Cell Res*, 23(10):1163–1171, Oct 2013. Supplementary information available for this article at http://www.nature.com/cr/journal/v23/n10/suppinfo/cr2013122s1.html.

[26] B. M. Chowrira, P. A. Pavco, and J. A. McSwiggen. In vitro and in vivo comparison of hammerhead, hairpin, and hepatitis delta virus self-processing ribozyme cassettes. *J. Biol. Chem.*, 269(41):25856–25864, Oct 1994.

[27] Michelle Christian, Tomas Cermak, Erin L. Doyle, Clarice Schmidt, Feng Zhang, Aaron Hummel, Adam J. Bogdanove, and Daniel F. Voytas. Targeting dna double-strand breaks with tal effector nucleases. *Genetics*, 186(2):757–761, 2010.

[28] Jade Q. Clement, Lian Qian, Nickalus Kaplinsky, and Miles F. Wilkinson. The stability and fate of a spliced intron from vertebrate cells. *RNA*, 5:206–220, 2 1999.

[29] Le Cong, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, Xuebing Wu, Wenyan Jiang, Luciano A. Marraffini, and Feng Zhang. Multiplex genome engineering using crispr/cas systems. *Science*, 339(6121):819–823, 2013.

[30] R. Daniel, J. R. Rubens, R. Sarpeshkar, and T. K. Lu. Synthetic analog computation in living cells. *Nature*, 497(7451):619–623, May 2013.

[31] Domitilla Del Vecchio, Alexander J Ninfa, and Eduardo D Sontag. Modular cell biology: retroactivity and insulation. *Molecular Systems Biology*, 4(1), 2008.

[32] Mary-Lee Dequeant and Olivier Pourquie. Segmental patterning of the vertebrate embryonic axis. *Nat Rev Genet*, 9(5):370–382, May 2008.

[33] Charles A. Dinarello. Immunological and inflammatory functions of the interleukin-1 family. *Annual Review of Immunology*, 27(1):519–550, 2009. PMID: 19302047.

[34] Arnulf Dorn, Jacques Bollekens, Adrian Staub, Christophe Benoist, and Diane Mathis. A multiplicity of ccaat box-binding proteins. *Cell*, 50(6):863 – 872, 1987.

[35] Jennifer A. Doudna and Emmanuelle Charpentier. The new frontier of genome engineering with crispr-cas9. *Science*, 346(6213), 2014.

[36] Tom Ellis, Xiao Wang, and James J. Collins. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol*, 27(5):465–471, May 2009. 19377462[pmid].

[37] Michael Elowitz and Wendell A. Lim. Build life to understand it. *Nature*, 468(7326):889–890, Dec 2010.

[38] Michael B. Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, Jan 2000.

[39] Monicia Elrod-Erickson, Mark A Rould, Lena Nekludova, and Carl O Pabo. Zif268 protein–dna complex refined at 1.6a: a model system for understanding zinc finger–dna interactions. *Structure*, 4(10):1171 – 1180, 1996.

[40] C. Engler and S. Marillonnet. Combinatorial DNA assembly using Golden Gate cloning. *Methods Mol. Biol.*, 1073:141–156, 2013.

[41] K. M. Esvelt, P. Mali, J. L. Braff, M. Moosburner, S. J. Yaung, and G. M. Church. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods*, 10(11):1116–1121, Nov 2013.

[42] Fahim Farzadfard, Samuel D. Perli, and Timothy K. Lu. Tunable and multifunctional eukaryotic transcription factors based on crispr/cas. *ACS Synthetic Biology*, 2(10):604–613, 2013. PMID: 23977949.

[43] A. R. Ferre-D'Amare, K. Zhou, and J. A. Doudna. Crystal structure of a hepatitis delta virus ribozyme. *Nature*, 395(6702):567–574, Oct 1998.

[44] Ari E. Friedland, Timothy K. Lu, Xiao Wang, David Shi, George Church, and James J. Collins. Synthetic gene networks that count. *Science*, 324(5931):1199–1202, 2009.

[45] Christopher J. Fry and Peggy J. Farnham. Context-dependent transcriptional regulation. *Journal of Biological Chemistry*, 274(42):29583–29586, 1999.

[46] Martin Fussenegger, Rowan P. Morris, Cornelia Fux, Markus Rimann, Beryl von Stockar, Charles J. Thompson, and James E. Bailey. Streptogramin-based gene regulation systems for mammalian cells. *Nat Biotech*, 18(11):1203–1208, Nov 2000.

[47] Yangbin Gao and Yunde Zhao. Self-processing of ribozyme-flanked rnas into guide rnas in vitro and in vivo for crispr-mediated genome editing. *Journal of Integrative Plant Biology*, 56(4):343–349, 2014.

[48] Giedrius Gasiunas, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys. Cas9–crrna ribonucleoprotein complex mediates specific dna cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, 109(39):E2579âĂŞE2586, 2012.

[49] Nicholas Paul Gauthier, Lars Juhl Jensen, Rasmus Wernersson, SÃÿren Brunak, and Thomas S. Jensen. Cyclebase.org: version 2.0, an updated comprehensive, multi-species repository of cell cycle experiments and derived analysis results. *Nucleic Acids Research*, 38(suppl 1):D699–D702, 2010.

[50] John Gerhart, Marc Kirschner, and Eileen Starr Moderbacher. *Cells, embryos, and evolution: Toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability.* Blackwell Science Malden, 1997.

[51] Luke A. Gilbert, Matthew H. Larson, Leonardo Morsut, Zairan Liu, Gloria A. Brar, Sandra E. Torres, Noam Stern-Ginossar, Onn Brandman, Evan H. Whitehead, Jennifer A. Doudna, Wendell A. Lim, Jonathan S. Weissman, and Lei S. Qi. Crispr-mediated modular rna-guided regulation of transcription in eukaryotes. *Cell*, 154(2):442–451, 2013.

[52] David Greber, Marie Daoud El-Baba, and Martin Fussenegger. Intronically encoded sirnas improve dynamic range of mammalian gene regulation systems and toggle switch. *Nucleic Acids Research*, 36(16):e101, 2008.

[53] Harvey A. Greisman and Carl O. Pabo. A general strategy for selecting high-affinity zinc finger proteins for diverse dna target sites. *Science*, 275(5300):657–661, 1997.

[54] D. H. Haft, J. Selengut, E. F. Mongodin, and K. E. Nelson. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.*, 1(6):e60, Nov 2005.

[55] Caryn R. Hale, Peng Zhao, Sara Olson, Michael O. Duff, Brenton R. Graveley, Lance Wells, Rebecca M. Terns, and Michael P. Terns. Rna-guided rna cleavage by a crispr rna-cas protein complex. *Cell*, 139(5):945 – 956, 2009.

[56] Leland H. Hartwell, John J. Hopfield, Stanislas Leibler, and Andrew W. Murray. From molecular to modular cell biology. *Nature*, 402:C47–C52, 1999.

[57] Asma Hatoum-Aslan, Inbal Maniv, and Luciano A. Marraffini. Mature clustered, regularly interspaced, short palindromic repeats rna (crrna) length is measured by a ruler mechanism anchored at the precursor processing site. *Proceedings of the National Academy of Sciences*, 108(52):21218–21222, 2011.

[58] R. E. Haurwitz, S. H. Sternberg, and J. A. Doudna. Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. *EMBO J.*, 31(12):2824–2832, Jun 2012.

[59] Rachel E. Haurwitz, Martin Jinek, Blake Wiedenheft, Kaihong Zhou, and Jennifer A. Doudna. Sequence- and structure-specific rna processing by a crispr endonuclease. *Science*, 329(5997):1355–1358, 2010.

[60] Sara Hooshangi, Stephan Thiberge, and Ron Weiss. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 102(10):3581–3586, 2005.

[61] Jessica A. Hurt, Stacey A. Thibodeau, Andrew S. Hirsh, Carl O. Pabo, and J. Keith Joung. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proceedings of the National Academy of Sciences*, 100(21):12271–12276, 2003.

[62] Mark Isalan, Aaron Klug, and Yen Choo. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the hiv-1 promoter. *Nat Biotechnol*, 19(7):656–660, Jul 2001. 11433278[pmid].

[63] Y Ishino, H Shinagawa, K Makino, M Amemura, and A Nakata. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in escherichia coli, and identification of the gene product. *Journal of Bacteriology*, 169(12):5429–5433, 1987.

[64] Shiro Iuchi and Natalie (Eds.) Kuldell. *Zinc Finger Proteins: From Atomic Contact to Cellular Function*. Springer US, 2005.

[65] Ruud. Jansen, Jan. D. A. van Embden, Wim. Gaastra, and Leo. M. Schouls. Identification of genes that are associated with dna repeats in prokaryotes. *Molecular Microbiology*, 43(6):1565–1575, 2002.

[66] Peng Jiang, Alejandra C. Ventura, Eduardo D. Sontag, Sofia D. Merajver, Alexander J. Ninfa, and Domitilla Del Vecchio. Load-induced modulation of signal transduction networks. *Science Signaling*, 4(194):ra67–ra67, 2011.

[67] Wenyan Jiang, David Bikard, David Cox, Feng Zhang, and Luciano A. Marraffini. Rna-guided editing of bacterial genomes using crispr-cas systems. *Nat Biotech*, 31(3):233–239, Mar 2013. Research.

[68] Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. A programmable dual-rnaâĂŞguided dna endonuclease in adaptive bacterial immunity. *Science*, 337(6096):816–821, 2012.

[69] Martin Jinek, Alexandra East, Aaron Cheng, Steven Lin, Enbo Ma, and Jennifer Doudna. Rna-programmed genome editing in human cells. *eLife*, 2, 2013.

[70] J. Keith Joung, Elizabeth I. Ramm, and Carl O. Pabo. A bacterial two-hybrid selection system for studying proteinâĂŞdna and proteinâĂŞprotein interactions. *Proceedings of the National Academy of Sciences*, 97(13):7382–7387, 2000.

[71] David Kadosh and Kevin Struhl. Repression by ume6 involves recruitment of a complex containing sin3 corepressor and rpd3 histone deacetylase to target promoters. *Cell*, 89(3):365 – 371, 1997.

[72] David Kadosh and Kevin Struhl. Targeted recruitment of the sin3-rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. *Molecular and Cellular Biology*, 18(9):5121–5127, 1998.

[73] Daniel Kalderon, Bruce L. Roberts, William D. Richardson, and Alan E. Smith. A short amino acid sequence able to specify nuclear location. *Cell*, 39(3, Part 2):499 – 509, 1984.

[74] Sven Kappel, Yves Matthess, Brigitte Zimmer, Manfred Kaufmann, and Klaus Strebhardt. Tumor inhibition by genomically integrated inducible rnai-cassettes. *Nucleic Acids Research*, 34(16):4527–4536, 2006.

[75] Maria Karlsson, Wilfried Weber, and Martin Fussenegger. Design and construction of synthetic gene networks in mammalian cells. In Wilfried Weber and Martin Fussenegger, editors, *Synthetic Gene Networks*, volume 813 of *Methods in Molecular Biology*, pages 359–376. Humana Press, 2012.

[76] Cynthia A. Keleher, Caroline Goutte, and Alexander D. Johnson. The yeast cell-type-specific repressor $\alpha 2$ acts cooperatively with a non-cell-type-specific protein. *Cell*, 53(6):927 – 936, 1988.

[77] Cynthia A. Keleher, Michael J. Redd, Janet Schultz, Marian Carlson, and Alexander D. Johnson. Ssn6-tup1 is a general repressor of transcription in yeast. *Cell*, 68(4):709 – 719, 1992.

[78] Albert J. Keung, Caleb J. Bashor, Szilvia Kiriakov, James J. Collins, and Ahmad S. Khalil. Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell*, 158(1):110–120, 2015/04/20 2014.

[79] Albert J. Keung, J. Keith Joung, Ahmad S. Khalil, and James J. Collins. Chromatin regulation at the frontier of synthetic biology. *Nat Rev Genet*, 16(3):159–171, Mar 2015. Review.

[80] Ahmad S. Khalil, Timothy K. Lu, Caleb J. Bashor, Cherie L. Ramirez, Nora C. Pyenson, J. Keith Joung, and James J. Collins. A synthetic biology framework for programming eukaryotic transcription functions. *Cell*, 150(3):647 – 658, 2012.

[81] Yoosik Kim, Maria Jose Andreu, Bomyi Lim, Kwanghun Chung, Mark Terayama, Gerardo Jimenez, Celeste A. Berg, Hang Lu, and Stanislav Y. Shvartsman. Gene regulation by mapk substrate competition. *Developmental Cell*, 20(6):880 – 887, 2011.

[82] K. Kiss, A. Brozik, N. Kucsma, A. Toth, M. Gera, L. Berry, A. Vallentin, H. Vial, M. Vidal, and G. Szakacs. Shifting the paradigm: the putative mitochondrial protein ABCB6 resides in the lysosomes of cells and in the plasma membrane of erythrocytes. *PLoS ONE*, 7(5):e37378, 2012.

[83] Aaron Klug. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual Review of Biochemistry*, 79(1):213–231, 2010. PMID: 20192761.

[84] Young-Don Kwak, Hiroko Koike, and Kiminobu Sugaya. Rna interference with small hairpin rnas transcribed from a human u6 promoter-driven dna vector. *Journal of Pharmacological Sciences*, 93(2):214–217, 2003.

[85] John H Laity, Brian M Lee, and Peter E Wright. Zinc finger proteins: new insights into structural and functional diversity. *Current Opinion in Structural Biology*, 11(1):39 – 46, 2001.

[86] Douglas A. Lauffenburger. Cell signaling pathways as control modules: Complexity for simplicity? *Proceedings of the National Academy of Sciences*, 97(10):5031–5033, 2000.

[87] R. Lin, S. Maeda, C. Liu, M. Karin, and T. S. Edgington. A large noncoding rna is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene*, 26(6):851–858, Jul 2006.

[88] Jason J. Lohmueller, Thomas Z. Armel, and Pamela A. Silver. A tunable zinc finger-based framework for boolean logic computation in mammalian cells. *Nucleic Acids Research*, 40(11):5180–5187, 2012.

[89] Carlos Lois, Elizabeth J. Hong, Shirley Pease, Eric J. Brown, and David Baltimore. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*, 295(5556):868–872, 2002.

[90] Morgan L. Maeder, Samantha J. Linder, Vincent M. Cascio, Yanfang Fu, Quan H. Ho, and J. Keith Joung. Crispr rna-guided activation of endogenous human genes. *Nat Meth*, 10(10):977–979, Oct 2013. Brief Communication.

[91] Morgan L. Maeder, Samantha J. Linder, Deepak Reyon, James F. Angstman, Yanfang Fu, Jeffry D. Sander, and J. Keith Joung. Robust, synergistic regulation of human gene expression using tale activators. *Nat Meth*, 10(3):243–245, Mar 2013.

[92] Morgan L. Maeder, Stacey Thibodeau-Beganny, Anna Osiak, David A. Wright, Reshma M. Anthony, Magdalena Eichtinger, Tao Jiang, Jonathan E. Foley, Ronnie J. Winfrey, Jeffrey A. Townsend, Erica Unger-Wallace, Jeffry D. Sander, Felix MÃijller-Lerch, Fengli Fu, Joseph Pearlberg, Carl GÃűbel, JustinÂăP. Dassie, Shondra M. Pruett-Miller, Matthew H. Porteus, Dennis C. Sgroi, A. John Iafrate, Drena Dobbs, Paul B. McCray Jr., Toni Cathomen, Daniel F.

Voytas, and J. Keith Joung. Rapid âĂIJopen-sourceâĂİ engineering of customized zinc-finger nucleases for highly efficient gene modification. *Molecular Cell*, 31(2):294 – 301, 2008.

[93] K. S. Makarova, L. Aravind, Y. I. Wolf, and E. V. Koonin. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct*, 6:38, 2011.

[94] K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, and E. V. Koonin. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct*, 1:7, 2006.

[95] Kira S. Makarova, Daniel H. Haft, Rodolphe Barrangou, Stan J. J. Brouns, Emmanuelle Charpentier, Philippe Horvath, Sylvain Moineau, Francisco J. M. Mojica, Yuri I. Wolf, Alexander F. Yakunin, John van der Oost, and Eugene V. Koonin. Evolution and classification of the crispr–cas systems. *Nat Rev Micro*, 9(6):467–477, Jun 2011.

[96] Prashant Mali, John Aach, P. Benjamin Stranges, Kevin M. Esvelt, Mark Moosburner, Sriram Kosuri, Luhan Yang, and George M. Church. Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotech*, 31(9):833–838, Sep 2013. Research.

[97] Prashant Mali, Luhan Yang, Kevin M. Esvelt, John Aach, Marc Guell, James E. DiCarlo, Julie E. Norville, and George M. Church. Rna-guided human genome engineering via cas9. *Science*, 339(6121):823–826, 2013.

[98] Luciano A. Marraffini and Erik J. Sontheimer. Crispr interference limits horizontal gene transfer in staphylococci by targeting dna. *Science*, 322(5909):1843–1845, 2008.

[99] Tim R. Mercer, Marcel E. Dinger, and John S. Mattick. Long non-coding rnas: insights into functions. *Nat Rev Genet*, 10(3):155–159, Mar 2009.

[100] Michael Milyavsky, Yuval Tabach, Igor Shats, Neta Erez, Yehudit Cohen, Xiaohu Tang, Marina Kalis, Ira Kogan, Yosef Buganim, Naomi Goldfinger, Doron Ginsberg, Curtis C. Harris, Eytan Domany, and Varda Rotter. Transcriptional programs following genetic alterations in p53, ink4a, and h-ras genes along defined stages of malignant transformation. *Cancer Research*, 65(11):4530–4543, 2005.

[101] Deepak Mishra, Phillip M. Rivera, Allen Lin, Domitilla Del Vecchio, and Ron Weiss. A load driver device for engineering modularity in biological networks. *Nat Biotech*, 32(12):1268–1275, Dec 2014. Research.

[102] Francisco J. M. Mojica, Cesar Diez-Villasenor, Elena Soria, and Guadalupe Juez. Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria. *Molecular Microbiology*, 36(1):244–246, 2000.

[103] Francisco J.M. Mojica, Cesar Diez-Villasenor, Jesus Garcia-Martinez, and Elena Soria. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, 60(2):174–182, 2005.

[104] L. Nissim and R. H. Bar-Ziv. A tunable dual-promoter integrator for targeting of cancer cells. *Mol. Syst. Biol.*, 6:444, Dec 2010.

[105] Andrea Orioli, Chiara Pascali, Aldo Pagano, Martin Teichmann, and Giorgio Dieci. {RNA} polymerase {III} transcription control elements: Themes and variations. *Gene*, 493(2):185 – 194, 2012.

[106] NP Pavletich and CO Pabo. Zinc finger-dna recognition: crystal structure of a zif268-dna complex at 2.1 a. *Science*, 252(5007):809–817, 1991.

[107] Pablo Perez-Pinera, D. Dewran Kocak, Christopher M. Vockley, Andrew F. Adler, Ami M. Kabadi, Lauren R. Polstein, Pratiksha I. Thakore, Katherine A. Glass, David G. Ousterout, Kam W. Leong, Farshid Guilak, Gregory E. Crawford, Timothy E. Reddy, and Charles A. Gersbach. Rna-guided gene activation by crispr-cas9-based transcription factors. *Nat Meth*, 10(10):973–976, Oct 2013. Brief Communication.

[108] Pablo Perez-Pinera, David G. Ousterout, Jonathan M. Brunger, Alicia M. Farin, Katherine A. Glass, Farshid Guilak, Gregory E. Crawford, Alexander J. Hartemink, and Charles A. Gersbach. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat Meth*, 10(3):239–242, Mar 2013.

[109] Heinz W. Pley, Kevin M. Flaherty, and David B. McKay. Three-dimensional structure of a hammerhead ribozyme. *Nature*, 372(6501):68–74, Nov 1994.

[110] C. Pourcel, G. Salvignol, and G. Vergnaud. Crispr elements in yersinia pestis acquire new repeats by preferential uptake of bacteriophage dna, and provide additional tools for evolutionary studies. *Microbiology*, 151(3):653–663, 2005.

[111] N. J. Proudfoot. Ending the message: poly(A) signals then and now. *Genes Dev.*, 25(17):1770–1782, Sep 2011.

[112] Priscilla E. M. Purnick and Ron Weiss. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol*, 10(6):410–422, Jun 2009.

[113] Lei Qi, Rachel E. Haurwitz, Wenjun Shao, Jennifer A. Doudna, and Adam P. Arkin. Rna processing enables predictable programming of gene expression. *Nat Biotech*, 30(10):1002–1006, Oct 2012. Research.

[114] Lei S. Qi, Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin, and Wendell A. Lim. Repurposing crispr as an rna-guided platform for sequence-specific control of gene expression. *Cell*, 152(5):1173 – 1183, 2013.

[115] Emmy P. Rogakou, Duane R. Pilch, Ann H. Orr, Vessela S. Ivanova, and William M. Bonner. Dna double-stranded breaks induce histone h2ax phosphorylation on serine 139. *Journal of Biological Chemistry*, 273(10):5858–5868, 1998.

[116] Sharon Y Roth. Chromatin-mediated transcription repression in yeast. *Current Opinion in Genetics & Development*, 5(2):168 – 173, 1995.

[117] Christophe Rouillon, Min Zhou, Jing Zhang, Argyris Politis, Victoria Beilsten-Edmands, Giuseppe Cannone, Shirley Graham, CarolÂăV. Robinson, Laura Spagnolo, and MalcolmÂăF. White. Structure of the crispr interference complex csm reveals key similarities with cascade. *Molecular Cell*, 52(1):124 – 134, 2013.

[118] Stephen E. Rundlett, Andrew A. Carmen, Noriyuki Suka, Bryan M. Turner, and Michael Grunstein. Transcriptional repression by ume6 involves deacetylation of lysine 5 of histone h4 by rpd3. *Nature*, 392(6678):831–835, Apr 1998.

[119] H. Saito, T. Kobayashi, T. Hara, Y. Fujita, K. Hayashi, R. Furushima, and T. Inoue. Synthetic translational regulation by an L7Ae-kink-turn RNP switch. *Nat. Chem. Biol.*, 6(1):71–78, Jan 2010.

[120] MichÃĺle Sawadogo and Robert G. Roeder. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the {TATA} box region. *Cell*, 43(1):165 – 175, 1985.

[121] S Scherer and R W Davis. Replacement of chromosome segments with altered dna sequences constructed in vitro. *Proceedings of the National Academy of Sciences*, 76(10):4951–4955, 1979.

[122] David J. Segal, Birgit Dreier, Roger R. Beerli, and Carlos F. Barbas. Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5'-gnn-3' dna target sequences. *Proceedings of the National Academy of Sciences*, 96(6):2758–2763, 1999.

[123] Shiraz A. Shah, Susanne Erdmann, Francisco J.M. Mojica, and Roger A. Garrett. Protospacer recognition motifs. *RNA Biology*, 10(5):891–899, 2013.

[124] Kum-Joo Shin, Estelle A. Wall, Joelle R. Zavzavadjian, Leah A. Santat, Jamie Liu, Jong-Ik Hwang, Robert Rebres, Tamara Roach, William Seaman, Melvin I. Simon, and Iain D. C. Fraser. A single lentiviral vector platform for microrna-based conditional rna interference and coordinated transgene expression. *Proceedings of the National Academy of Sciences*, 103(37):13759–13764, 2006.

[125] Rebecca A. Silverstein and Karl Ekwall. Sin3: a flexible regulator of global gene expression and genome stability. *Current Genetics*, 47(1):1–17, 2005.

[126] Abhinav Sinha, Katie R. Hughes, Katarzyna K. Modrzynska, Thomas D. Otto, Claudia Pfander, Nicholas J. Dickens, Agnieszka A. Religa, Ellen Bushell, Anne L. Graham, Rachael Cameron, Bjorn F. C. Kafsack, April E. Williams, Manuel Llinas, Matthew Berriman, Oliver Billker, and Andrew P. Waters. A cascade of dna-binding proteins for sexual commitment and development in plasmodium. *Nature*, 507(7491):253–257, Mar 2014. Letter.

[127] Piro Siuti, John Yazbek, and Timothy K. Lu. Synthetic circuits integrating logic and memory in living cells. *Nat Biotech*, 31(5):448–452, May 2013. Research.

[128] Marlena Siwiak and Piotr Zielenkiewicz. A comprehensive, quantitative, and genome-wide model of translation. *PLoS Comput Biol*, 6(7):e1000865, Jul 2010. 20686685[pmid].

[129] Christopher W. J. Smith, Elena B. Porro, James G. Patton, and Bernardo Nadal-Ginard. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature*, 342(6247):243–247, Nov 1989.

[130] Rebecca L Smith and Alexander D Johnson. Turning genes off by ssn6âĂŞtup1: a conserved system of transcriptional repression in eukaryotes. *Trends in Biochemical Sciences*, 25(7):325 – 330, 2000.

[131] G. A. Soukup and R. R. Breaker. Design of allosteric hammerhead ribozymes activated by ligand-induced structure stabilization. *Structure*, 7(7):783–791, Jul 1999.

[132] S. H. Sternberg, R. E. Haurwitz, and J. A. Doudna. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA*, 18(4):661–672, Apr 2012.

[133] S. A. Stewart, D. M. Dykxhoorn, D. Palliser, H. Mizuno, E. Y. Yu, D. S. An, D. M. Sabatini, I. S. Chen, W. C. Hahn, P. A. Sharp, R. A. Weinberg, and C. D. Novina. Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA*, 9(4):493–501, Apr 2003.

[134] Scott A. Strobel and Peter B. Dervan. Single-site enzymatic cleavage of yeast genomic dna mediated by triple helix formation. *Nature*, 350(6314):172–174, Mar 1991.

[135] Allison J. Taggart, Alec M. DeSimone, Janice S. Shih, Madeleine E. Filloux, and William G. Fairbrother. Large-scale mapping of branchpoints in human pre-mrna transcripts in vivo. *Nat Struct Mol Biol*, 19(7):719–721, Jul 2012.

[136] Thean-Hock Tang, Jean-Pierre Bachellerie, Timofey Rozhdestvensky, Marie-Line Bortolin, Harald Huber, Mario Drungowski, Thorsten Elge, Jurgen Brosius, and Alexander Huttenhofer. Identification of 86 candidates for small non-messenger rnas from the archaeon archaeoglobus fulgidus. *Proceedings of the National Academy of Sciences*, 99(11):7536–7541, 2002.

[137] Kirk R. Thomas, Kim R. Folger, and Mario R. Capecchi. High frequency targeting of genes to specific sites in the mammalian genome. *Cell*, 44(3):419 – 428, 1986.

[138] Santosh Kumar Upadhyay, Jitesh Kumar, Anshu Alok, and Rakesh Tuli. Rna-guided genome editing for target gene mutations in wheat. *G3: Genes/Genomes/Genetics*, 3(12):2233–2238, 2013.

[139] John van der Oost, Edze R. Westra, Ryan N. Jackson, and Blake Wiedenheft. Unravelling the structural and mechanistic basis of crispr-cas systems. *Nat Rev Micro*, 12(7):479–492, Jul 2014. Review.

[140] Dingzhi Wang, Haibin Wang, Joanne Brown, Takiko Daikoku, Wei Ning, Qiong Shi, Ann Richmond, Robert Strieter, Sudhansu K. Dey, and Raymond N. DuBois. Cxcl1 induced by prostaglandin e2 promotes angiogenesis in colorectal cancer. *The Journal of Experimental Medicine*, 203(4):941–951, 2006.

[141] G Wang, D D Levy, M M Seidman, and P M Glazer. Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation. *Molecular and Cellular Biology*, 15(3):1759–68, 1995.

[142] H Wang and D J Stillman. Transcriptional repression in saccharomyces cerevisiae by a sin3-lexa fusion protein. *Molecular and Cellular Biology*, 13(3):1805–1814, 1993.

[143] W. Weber and M. Fussenegger. Emerging biomedical applications of synthetic biology. *Nat. Rev. Genet.*, 13(1):21–35, Jan 2012.

[144] Ross C. Wilson and Jennifer A. Doudna. Molecular mechanisms of rna interference. *Annual Review of Biophysics*, 42(1):217–239, 2013. PMID: 23654304.

[145] J. E. Wilusz, C. K. JnBaptiste, L. Y. Lu, C. D. Kuhn, L. Joshua-Tor, and P. A. Sharp. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.*, 26(21):2392–2407, Nov 2012.

[146] Zhen Xie, Liliana Wroblewska, Laura Prochazka, Ron Weiss, and Yaakov Benenson. Multi-input rnai-based logic circuit for identification of specific cancer cells. *Science*, 333(6047):1307–1311, 2011.

[147] Haifeng Ye, Marie Daoud-El Baba, Ren-Wang Peng, and Martin Fussenegger. A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science*, 332(6037):1565–1568, 2011.

[148] Qing-Fei Yin, Li Yang, Yang Zhang, Jian-Feng Xiang, Yue-Wei Wu, Gordon G. Carmichael, and Ling-Ling Chen. Long noncoding rnas with snorna ends. *Molecular Cell*, 48(2):219 – 230, 2012.