A Synthetic Biology Framework for Programming Eukaryotic Transcription Functions

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1016/j.cell.2012.05.045">http://dx.doi.org/10.1016/j.cell.2012.05.045</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Elsevier</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://hdl.handle.net/1721.1/91623">http://hdl.handle.net/1721.1/91623</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
</tbody>
</table>
A Synthetic Biology Framework for Programming Eukaryotic Transcription Functions

Ahmad S. Khalil,1,7 Timothy K. Lu,2,7,* Caleb J. Bashor,1,7 Cherie L. Ramirez,3,4 Nora C. Pyenson,1 J. Keith Joung,3,5 and James J. Collins1,6

1Howard Hughes Medical Institute, Department of Biomedical Engineering, and Center for BioDynamics, Boston University, Boston, MA 02215, USA
2Synthetic Biology Group, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
3Molecular Pathology Unit, Center for Cancer Research and Center for Computational and Integrative Biology, Massachusetts General Hospital, Charlestown, MA 02129, USA
4Biological and Biomedical Sciences Program, Division of Medical Sciences
5Department of Pathology
Harvard Medical School, Boston, MA 02115, USA
6Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA
7These authors contributed equally to this work
*Correspondence: timlu@mit.edu
http://dx.doi.org/10.1016/j.cell.2012.05.045

SUMMARY

Eukaryotic transcription factors (TFs) perform complex and combinatorial functions within transcriptional networks. Here, we present a synthetic framework for systematically constructing eukaryotic transcription functions using artificial zinc fingers, modular DNA-binding domains found within many eukaryotic TFs. Utilizing this platform, we construct a library of orthogonal synthetic transcription factors (sTFs) and use these to wire synthetic transcriptional circuits in yeast. We engineer complex functions, such as tunable output strength and transcriptional cooperativity, by rationally adjusting a decomposed set of key component properties, e.g., DNA specificity, affinity, promoter design, protein-protein interactions. We show that subtle perturbations to these properties can transform an individual sTF between distinct roles (activator, cooperative factor, inhibitory factor) within a transcriptional complex, thus drastically altering the signal processing behavior of multi-input systems. This platform provides new genetic components for synthetic biology and enables bottom-up approaches to understanding the design principles of eukaryotic transcriptional complexes and networks.

INTRODUCTION

The genetic program of a living cell is governed by the faithful execution of a number of fundamental molecular functions by transcription factors (TFs). These include wiring specific connections to promoter regulatory elements, modulating the transcriptional output of a gene, tuning molecular noise, recruiting coactivator/repressor complexes and basal transcriptional machinery, cooperating with other TFs to regulate a gene, integrating an array of environmental signals, and even physically manipulating the geometrical configuration of chromosomes (Hahn and Young, 2011; Pedraza and van Oudenaarden, 2005; Ptashne, 1986, 1988; Rosenfeld et al., 2005). A tremendous amount of progress has been made toward understanding eukaryotic transcription regulation. Yet, there is still much to be learned about how the molecular properties of TFs give rise to the complex behavior of transcriptional networks. A synthetic approach, whereby minimal and insulated components and circuitry can be constructed to recapitulate eukaryotic transcription function, would be valuable for studying how transcriptional regulatory complexes are assembled and how TFs wire together transcriptional networks.

A framework for eukaryotic transcription regulation would also be broadly valuable to synthetic biology efforts, which seek to uncover the design principles of gene regulatory networks and program novel biological functions for a range of biotechnological and industrial applications (Andrianantoandro et al., 2006; Bashor et al., 2010; Khalil and Collins, 2010; Mukherji and van Oudenaarden, 2009; Nandagopal and Elowitz, 2011; Smolke and Silver, 2011). Engineering synthetic transcriptional networks has been a major focus of the field, and a variety of circuit behaviors have been implemented, including memory, oscillations, logic operations, filtering, and noise propagation (Basu et al., 2005; Becskei and Serrano, 2000; Elowitz and Leibler, 2000; Friedland et al., 2009; Gardner et al., 2000; Guet et al., 2002; Pedraza and van Oudenaarden, 2005; Rosenfeld et al., 2005). In these and virtually all other studies of synthetic transcriptional networks, circuitry has been constructed using...
Transcriptional functions performed by eukaryotic TFs

- Operate cooperatively with other TFs
- Wire connections to promoter REs
- Trans. initiation complex
- Pol II
- OUTPUT

Constructing transcriptional functions using synthetic TFs

- Tunable cooperativity with other TFs
- Customizable wiring between sTFs and REs
- Programmable specificity
- Trans. initiation complex
- Pol II
- OUTPUT

Here, we present an alternative approach to engineering transcriptional regulation in eukaryotes using synthetic transcription factors (sTFs) constructed from Cys2-His2 zinc finger (ZF) domains (Figure 1). The sTFs feature a modular design in which separate protein domains carry out individual molecular functions: ZF domains enable binding to DNA at user-specified sequences embedded within an engineered promoter, the transcriptional output for that promoter is driven by an activation domain that recruits basal transcription machinery, and a protein-protein interaction domain allows cooperative interactions with adjacent TFs. This decomposed design permits the facile tuning of individual sTF component properties. ZF domains were selected to carry out sTF DNA binding function because of their potential for engineered sequence specificity. ZFs are small (~30 amino acid) domains that bind to ~3 bps of DNA (Elrod-Erickson et al., 1998; Pavletich and Pabo, 1991). ZFs are utilized in natural transcriptional networks in virtually all eukaryotic taxa to solve the combinatorial problem of DNA recognition by binding to promoter sequences in tandem arrays (Pabo et al., 2001). Recent advances have made it possible to purposefully re-engineer the DNA-binding specificity of individual ZFs to bind to a wide variety of 3 bp sequences, and then covalently link them together into artificial, multifinger arrays capable of recognizing longer DNA sequences with a high degree of specificity (Beerli and Barbas, 2002; Maeder et al., 2008, 2009; Pabo et al., 2001; Sander et al., 2011). Notably, with oligomerized pool engineering (OPEN) (Maeder et al., 2008) and other “context-dependent” engineering methods (Sander et al., 2011), multifinger arrays with defined specificities have been generated to design ZF nucleases (ZFNS) for targeted gene and genome modification (Foley et al., 2009; Maeder et al., 2008; Sebastiano et al., 2011; Townsend et al., 2009; Zou et al., 2009).
Using the OPEN platform, we construct a library of specific and orthogonal sTF-promoter pairs, and demonstrate that these pairs can be used to wire synthetic transcriptional cascades in Saccharomyces cerevisiae. We then use these circuits as a testbed system for exploring the relationship between circuit output and sTF function. We find that a few, key properties, e.g., DNA specificity, DNA affinity, promoter-operator design, and protein interactions, can be rationally and independently adjusted to tune transcriptional behavior. For example, we demonstrate the tuning of transcriptional output through the perturbation of ZF binding affinity and operator number. Additionally, we engineer cooperative transcriptional systems by multimerizing weakly-activating sTF monomers using modular protein-protein interaction domains. Finally, in order to synthetically explore transcriptional signal integration, we construct a set of simple two-input promoters that recruit two individual sTFs. By systematically altering the architecture of the complex through subtle changes to the component properties of the sTFs, we can assign entirely different transcriptional roles to an individual sTF and thus dramatically alter the signal processing of the system.

RESULTS

Wiring Specific and Orthogonal Transcriptional Connections with a Library of Synthetic TF-Promoter Pairs

Transcriptional networks, natural and synthetic, are wired together with sequence-specific protein-DNA interactions. We sought to program DNA-binding specificity, via artificial ZF proteins, in order to wire specific and orthogonal transcriptional connections in the eukaryote, S. cerevisiae. To do so, we first devised a platform by which ZF-based sTFs could be readily constructed and customized. The platform consists of a cassette, into which artificial three-finger arrays with engineered specificities are inserted to generate sTF species. The sTF cassette is paired with a synthetic promoter bearing ZF binding sequences that act as operators for the sTFs (Figure 2A).

Transcriptional activation is one of the most common mechanisms for the control of gene regulation and appears to be a universally conserved process in all eukaryotes, from fungi to metazoans (Fischer et al., 1988; Ma et al., 1988; Webster et al., 1988). We utilized the principle of activation by recruitment (Ptashne, 1988; Ptashne and Gann, 1997) to test our sTFs as minimal transcriptional activators. In our design, the engineered ZF array recapitulates the TF function of binding to a specific DNA site, in this case, to its cognate 9 bp operator in a synthetic promoter. The ZF protein is fused to a VP16 minimal activation domain (AD), which autonomously facilitates recruitment of the RNA polymerase II machinery for mRNA initiation (Ptashne, 1988). This scheme provides a decoupled, modular approach to transcriptional activation, whereby TFs and the initiation machinery can be synthetically recruited in combinatorial fashion. From these components, we constructed a synthetic transcriptional cascade and used it as a test bed for rationally customizing the properties of our transcriptional components to program in vivo behaviors (Figure 2A). Within the circuit, sTF activators are first transcribed from a previously described TetR-controlled GAL1 promoter (Ellis et al., 2009; Murphy et al., 2007), which is induced by anhydrotetracycline (ATc). Addition of ATc activates flux through the circuit to produce sTF activators, which in turn activate downstream transcription from cognate synthetic promoters to produce yEGFP expression (Figure 2B; Figure S1 available online). The resulting gene regulatory transfer function, which combines the effects of the TetR expression system and the operation of sTFs on their synthetic promoters, exhibits monotonic, dose-dependent production of yEGFP (Figure 2B). These results suggest that desired synthetic transcriptional connections can be made based on the specificity of engineered ZF proteins to their target sites.

With the OPEN selection system, we have the ability to rapidly alter the ZF-DNA interaction specificity to create a large library of interaction partners (i.e., engineered ZF proteins and corresponding target sites). We used artificial ZF arrays constructed by OPEN to generate a library of sTF-promoter pairs. In particular, we identified 19 three-finger arrays with binding specificities predicted to be orthogonal to one another (we predominately chose OPEN ZF arrays that had been engineered to bind sequences in orthologous genes found in plants, insects, and metazoans) (Figure 3A). The artificial arrays and cognate binding sequences were inserted into our framework, and the resulting
library of sTF-promoter pairs were tested for activation by triggering our synthetic circuits. We found that the sTFs activated yEGFP expression from cognate promoters by factors of 1.3–6.6 (compared to uninduced cells) (Figure 3B), showing that we could indeed make sequence-specific transcriptional connections with artificially designed ZF arrays. Notably, yEGFP expression levels in uninduced cells were mostly found to be similar to the basal expression levels of cells harboring only synthetic promoters (Figure S2). Thus, in general, a signal was produced only when we induced expression of an sTF in the presence of a cognate promoter.

We next investigated whether the transcriptional connections made within our library of sTFs were indeed specific only to their cognate synthetic promoters. We selected a subset of six sTFs from our library that exhibited robust activation (>2.5-fold) (Figure 3B, red stars), and crossed them with each of the other non-cognate promoters. Upon triggering the circuit, we observed no cross-activation in the subset of tested sTFs (Figure 3C) with one notable exception: the effect of sTF43-8 on Promoter21-16. Examination of the sequence just downstream of the ZF operator for Promoter21-16 revealed the fortuitous creation of a sequence possessing significant similarity to the binding sequence of 43-8 (at 8 out of 9 bps) (Figure 3A, blue boxes). Thus, we attribute the observed cross-activation to the presence of this binding sequence within the noncognate promoter. Overall, these results show that synthetic transcriptional connections can be designed to be orthogonal to one another by using the OPEN method to engineer the DNA-binding specificities of ZF arrays.

In the design of synthetic elements and gene circuitry, a further “orthogonality” criterion is the degree to which the synthetic system interacts with pathways and machinery native to the cellular host. Ideally, insulated networks would interact with host pathways only at desired nodes and otherwise function independently. Using our synthetic yeast platform, we investigated one potential and rapid method for assessing sTF-host interactions. Specifically, we measured the growth of cells with and without the induction of sTFs, under the assumption that unwanted interactions with the host genome would impose a fitness cost on the cells. We observed no adverse or modest effects on growth in the great majority of sTFs from our library (Figure 3D and Figure S3). Our scheme may thus represent
a starting point for designing and screening sTFs with optimal functionality and orthogonality within a desired host.

These results show that engineered ZF arrays are effective building blocks for minimal sTF activators, and that DNA interaction specificity is a component property that can be programmed to mediate the construction of specific and orthogonal synthetic transcriptional connections in yeast. Moreover, largely through this ability to engineer DNA specificity for many interaction partners, our platform is able to make meaningful predictions about orthogonality (among synthetic components and with host machinery), which remains a major unaddressed issue in synthetic biology.

**Tuning Transcriptional Output**

ZF s are well-studied structural motifs with crystallographic information providing blueprints for harnessing their structure-function relationship to program more complex transcriptional behaviors. We investigated how we could rationally engineer various component properties pertaining to the ZF-DNA interaction to tune transcriptional outputs in our synthetic eukaryotic system. For these studies, we focused on the sTF pair 42-10 and 43-8 because they activate transcription robustly to similar levels but show orthogonal activities to one another. In addition, these two activators show some distinct properties, e.g., 42-10 seemed to impose a fitness cost on the yeast host, whereas 43-8 did not.

To tune up the level of transcriptional activation, we focused on alterations to the promoter architecture. We multimerized ZF binding sequences to create promoters with repeat operators that would correspondingly recruit greater numbers of sTF interactions and thus ADs. With promoters harboring one, two, and eight tandem operators, we observed a corresponding increase in the transcriptional output of the system, confirming that we could tune up the level of activation (Figure 4A).

**Figure 4. Tuning Transcriptional Outputs by Rationally Adjusting Multiple Component Properties**

(A) Tuning up output strength by increasing ZF operator number in synthetic promoter (sTF43-8).

(B) Integrating two distinct sTFs at a single synthetic promoter. sTF43-8 and sTF42-10 were expressed independently from ATc- and IPTG-inducible GAL1 promoters.

(C) Schematic representation of the canonical Cys2-His2 ZF protein (top). Each finger is composed of two β strands and a recognition helix, which makes sequence-specific contacts to three DNA bps. Four arginine residues in the ZF framework that mediate nonspecific interactions with the phosphate backbone were targeted for mutation to alanine residues (gray boxes and highlighted in red) in order to alter the affinity of the ZF for its cognate binding sequence.

(D) Tuning down activation output by engineering ZF affinity variants in sTF42-10 (3x: R2A/R39A/R67A, 4x: R2A/R11A/R39A/R67A). Horizontal axis begins at basal (promoter-only) fluorescence level (B and D).

(E) Phosphate backbone mutants of 42-10 rescue the fitness cost of sTF42-10 on host cell growth. Error bars represent SD of three experiments. See also Figures S4 and S5.
Importantly, no cross-activation was observed between these sTFs and any of the tandem synthetic promoters (Figure S4). Eukaryotic promoters are known to integrate multiple inputs by binding to distinct TFs. In fact, transcriptional networks may even act as logic gates through such regulation schemes. Our synthetic promoters can similarly be designed to recruit distinct sTFs through architectures that include distinct operators. We constructed a two-input promoter with operators specific for sTF43-8 and sTF42-10-4x. We then directed the expression of sTF43-8 and sTF42-10 under the independent and respective control of TetR- and LacI-controlled GAL1 promoters, which could in turn be induced by the chemical inputs, ATc and iso-propyl-β-D-thiogalactopyranoside (IPTG). Upon induction of either or both of the sTF species, we observed transcriptional activation over the uninduced case (Figure 4B), confirming that our promoter design can indeed integrate distinct transcriptional signals in Boolean OR-like fashion.

Promoter architecture can be designed to alter the number of sTFs recruited and thus tune transcriptional output strength. An alternative approach is to regulate the ZF-DNA interaction through structure-guided modification of the ZF backbone to alter nonspecific DNA affinity. Along these lines, we targeted four arginine residues outside of the DNA recognition helices that are known, based on structural studies, to mediate nonspecific interactions of a three-finger array with the DNA phosphate backbone (Elrod-Erickson et al., 1996; Pavletich and Pabo, 1991). The first arginine residue (position 2) is located upstream of the first β strand of the amino-terminal finger, whereas the remaining three (positions 11, 39, 67) are found within the β sheets of each of the three fingers, immediately upstream of each recognition helix (Figure 4C). The arginine residues mediate nonspecific interactions, in part, through their positive charge; thus, we altered each of these to alanine residues.

We screened the DNA-binding activities of ZF arrays possessing various combinations of these four phosphate backbone mutations using a previously described bacterial-two-hybrid (B2H) system (Maeder et al., 2008; Wright et al., 2006). Single residue alanine substitution mutations yielded modest effects, whereas sets of mutations revealed a step-wise decrease in DNA-binding activity (Figure S5). We next incorporated the phosphate backbone mutations into sTF42-10-4x and tested the transcriptional activity of the resulting variants in our synthetic eukaryotic system. We found that transcriptional output could be analogously tuned down as the number of mutations was increased from zero (sTF42-10) to three (sTF42-10-3A) to four (sTF42-10-4A), in effect creating weaker-activating sTF variants from the lower-affinity variants (Figure 4D). We additionally investigated the effects the weaker-activating sTFs have on the yeast host. In the fitness assay, we found that the phosphate backbone mutations were able to systematically rescue the growth inhibition observed with sTF42-10 (Figure 4D and Figure S3). We presume that this effect occurs because the mutant ZF proteins are less able to mediate off-target DNA interactions that may be at the root of the fitness cost.

Taken together, these results demonstrate that the rational engineering of ZF binding sites in the promoter architecture and the ZF-DNA binding interaction—two component properties of our synthetic system—provide effective strategies for tuning transcriptional output.

**Engineering Cooperative Transcriptional Systems from Weakly-interacting Components**

The assembly of TFs into multimeric complexes is a mechanism for achieving cooperativity and shaping input-output responses to regulate transcription. Inspired by natural cooperative systems, we next sought to assemble sTFs into multimeric complexes that could achieve synergistic transcriptional behaviors (Ptashne and Gann, 2002). To do so, we harnessed PDZ interaction domains from metazoan cells. These domains are naturally responsible for organizing intracellular signaling complexes, so we explored whether they could be utilized to assemble and organize our synthetic factors in transcriptional applications. Because these domains are modular, they provide an additional tunable component property to our framework and allow for generalizable designs for multimerization. Furthermore, canonical PDZ domains are extremely rare in nonmetazoans (Harris and Lim, 2001), and are therefore unlikely to interact with endogenous yeast machinery.

The weakly-activating sTFs represent ideal components with which to demonstrate multimerization and the synthetic construction of transcriptional cooperativity. These sTFs were built from ZF mutants with lower nonspecific DNA binding activity. We therefore investigated whether their assembly could stabilize a protein-DNA complex to better initiate transcription, presumably by slowing the off-rate of each component from the bound promoter. The well-studied PDZ domain from the mammalian protein α1-syntrophin (Craven and Bredt, 1998; Harris et al., 2001) was fused to the C terminus of 43-8-4x, and its cognate peptide ligand (GSGS-VKESLV), to which it binds with low micromolar affinity, was fused to the C terminus of 42-10-4x. For these studies, only sTF43-8-4x, carried an AD, making it a single locus for the recruitment of transcription initiation machinery. We did not attach an AD to the dimeric partner (sTF42-10-4A) in order to test whether this additional (nonactivating) factor could stabilize the complex through dimerization and thus aid in the cooperative activation of transcription by sTF43-8-4x. In addition, we generated a nonbinding partner by fusing a mutated form of the cognate ligand (GSGS-VKEAAA) to 42-10-4x. Expression of each “monomeric” sTF was driven by independently inducible GAL1 promoters. Upon induction, the sTFs operate on a synthetic “dimeric” promoter to activate downstream transcription (Figure 5).

We titrated expression of the PDZ-harboring sTF43-8-4x, both in the presence and absence of its ligand-carrying partner. In the resulting dose responses, we observed a significant synergistic effect on the transcriptional output of sTF43-8-4x when the partner factor was present, as compared to when it was not present (Figure 5A). Critically, we observed no cooperative effect on the transcriptional output of the system in the analogous dose response with the nonbinding ligand partner (Figure 5B). These results suggest that the monomeric sTFs, aided by engineered protein-protein interactions, cooperate to recruit and stabilize one another at the synthetic promoter, thereby increasing the promoter occupancy and the resulting transcriptional activity of the complex.
The PDZ-mediated sTF dimerization therefore serves as a key component property for enabling this type of synergistic two-input behavior. By simply modifying the ligand to abolish the binding interaction (i.e., mutating it to the noncognate GSGS-VKEAAA), we rendered the distal monomer transcriptionally noncontributory in the proximal activator scenario, and subsequently engineered a different two-input behavior: one that shows equal output levels in the presence of either both inputs or the input directing the proximal activator (Figure 6D). In other words, we created a two-input system with a “null” effect when both inputs are present.

We next sought to reverse the monomer roles by simply switching the placement of the AD. We loaded the AD onto ligand-carrying sTF$_{42-10-3x}$ and removed it from PDZ-carrying sTF$_{43-8-4x}$, while directing the two monomers at the same dimeric promoter (Figure 6, “distal activator” architecture). Strikingly, we did not observe transcriptional output synergy. Instead, we observed an inhibition of the output signal in the presence of both inputs (Figure 6E). Furthermore, the inhibitory behavior was conserved in both PDZ-binding and nonbinding cases (Figure 6F). These results suggest that, with this particular combination of components in the distal activator scenario, the proximal monomers take on an inhibitory as opposed to a cooperative role. If this were the case, then we reasoned that we should be able to further strengthen the inhibition by increasing the ZF affinity of the inhibitory monomer (43-8-4x) to its operator DNA operating specificity of 18 bps rather than 9 bps. Furthermore, the two-component complex achieved transcriptional output levels similar to those of the single WT activator (Figure 6A), but only through the addition of both inputs and a total DNA operating specificity of 18 bps rather than 9 bps. Furthermore, we found that we could boost this effect by adding an AD onto the distal monomer, thus engineering a two-activator system and providing another source for transcriptional machinery recruitment at the promoter (Figure 6C).

Figure 5. Transcriptional Cooperative Systems Can Be Engineered from Weakly-Activating sTF “monomers” that Are Dimerized with a PDZ Interaction Domain

(A) The dimerization interaction promotes cooperative behavior in transcriptional activation. Syntrophin PDZ domain (dark gray) was fused to the C-terminal of ZF affinity mutant 43-8-4x, and the resulting AD-carrying sTF “monomer” was expressed from ATc-inducible pGAL1. The heterologous ligand (light gray) was fused to the C-terminal of 42-10-4x, and the resulting AD-less factor was expressed from IPTG-inducible pGAL1. The factors assemble at a synthetic “dimeric” promoter to cooperatively activate downstream transcription (“IPTG+” = full induction with 20 mM IPTG). The PDZ-mediated sTF dimerization therefore serves as a key component property for enabling this type of synergistic two-input behavior. By simply modifying the ligand to abolish the binding interaction (i.e., mutating it to the noncognate GSGS-VKEAAA), we rendered the distal monomer transcriptionally noncontributory in the proximal activator scenario, and subsequently engineered a different two-input behavior: one that shows equal output levels in the presence of either both inputs or the input directing the proximal activator (Figure 6D). In other words, we created a two-input system with a “null” effect when both inputs are present.

We next sought to reverse the monomer roles by simply switching the placement of the AD. We loaded the AD onto ligand-carrying sTF$_{42-10-3x}$ and removed it from PDZ-carrying sTF$_{43-8-4x}$, while directing the two monomers at the same dimeric promoter (Figure 6, “distal activator” architecture). Strikingly, we did not observe transcriptional output synergy. Instead, we observed an inhibition of the output signal in the presence of both inputs (Figure 6E). Furthermore, the inhibitory behavior was conserved in both PDZ-binding and nonbinding cases (Figure 6F). These results suggest that, with this particular combination of components in the distal activator scenario, the proximal monomers take on an inhibitory as opposed to a cooperative role. If this were the case, then we reasoned that we should be able to further strengthen the inhibition by increasing the ZF affinity of the inhibitory monomer (43-8-4x) to its operator DNA operating specificity of 18 bps rather than 9 bps. Furthermore, we found that we could boost this effect by adding an AD onto the distal monomer, thus engineering a two-activator system and providing another source for transcriptional machinery recruitment at the promoter (Figure 6C).

The PDZ-mediated sTF dimerization therefore serves as a key component property for enabling this type of synergistic two-input behavior. By simply modifying the ligand to abolish the binding interaction (i.e., mutating it to the noncognate GSGS-VKEAAA), we rendered the distal monomer transcriptionally noncontributory in the proximal activator scenario, and subsequently engineered a different two-input behavior: one that shows equal output levels in the presence of either both inputs or the input directing the proximal activator (Figure 6D). In other words, we created a two-input system with a “null” effect when both inputs are present.

We next sought to reverse the monomer roles by simply switching the placement of the AD. We loaded the AD onto ligand-carrying sTF$_{42-10-3x}$ and removed it from PDZ-carrying sTF$_{43-8-4x}$, while directing the two monomers at the same dimeric promoter (Figure 6, “distal activator” architecture). Strikingly, we did not observe transcriptional output synergy. Instead, we observed an inhibition of the output signal in the presence of both inputs (Figure 6E). Furthermore, the inhibitory behavior was conserved in both PDZ-binding and nonbinding cases (Figure 6F). These results suggest that, with this particular combination of components in the distal activator scenario, the proximal monomers take on an inhibitory as opposed to a cooperative role. If this were the case, then we reasoned that we should be able to further strengthen the inhibition by increasing the ZF affinity of the inhibitory monomer (43-8-4x) to its operator DNA operating specificity of 18 bps rather than 9 bps. Furthermore, we found that we could boost this effect by adding an AD onto the distal monomer, thus engineering a two-activator system and providing another source for transcriptional machinery recruitment at the promoter (Figure 6C).

The PDZ-mediated sTF dimerization therefore serves as a key component property for enabling this type of synergistic two-input behavior. By simply modifying the ligand to abolish the binding interaction (i.e., mutating it to the noncognate GSGS-VKEAAA), we rendered the distal monomer transcriptionally noncontributory in the proximal activator scenario, and subsequently engineered a different two-input behavior: one that shows equal output levels in the presence of either both inputs or the input directing the proximal activator (Figure 6D). In other words, we created a two-input system with a “null” effect when both inputs are present.

We next sought to reverse the monomer roles by simply switching the placement of the AD. We loaded the AD onto ligand-carrying sTF$_{42-10-3x}$ and removed it from PDZ-carrying sTF$_{43-8-4x}$, while directing the two monomers at the same dimeric promoter (Figure 6, “distal activator” architecture). Strikingly, we did not observe transcriptional output synergy. Instead, we observed an inhibition of the output signal in the presence of both inputs (Figure 6E). Furthermore, the inhibitory behavior was conserved in both PDZ-binding and nonbinding cases (Figure 6F). These results suggest that, with this particular combination of components in the distal activator scenario, the proximal monomers take on an inhibitory as opposed to a cooperative role. If this were the case, then we reasoned that we should be able to further strengthen the inhibition by increasing the ZF affinity of the inhibitory monomer (43-8-4x) to its operator DNA operating specificity of 18 bps rather than 9 bps. Furthermore, we found that we could boost this effect by adding an AD onto the distal monomer, thus engineering a two-activator system and providing another source for transcriptional machinery recruitment at the promoter (Figure 6C).

The PDZ-mediated sTF dimerization therefore serves as a key component property for enabling this type of synergistic two-input behavior. By simply modifying the ligand to abolish the binding interaction (i.e., mutating it to the noncognate GSGS-VKEAAA), we rendered the distal monomer transcriptionally noncontributory in the proximal activator scenario, and subsequently engineered a different two-input behavior: one that shows equal output levels in the presence of either both inputs or the input directing the proximal activator (Figure 6D). In other words, we created a two-input system with a “null” effect when both inputs are present.
Figure 6. Synthetic ZF-Based Transcription Framework Can Be Used to Engineer Diverse Two-Input Behaviors

(A) The transcriptional operation of a single sTF43-8 (carrying a PDZ domain) at the proximal position of a two-input promoter.

(B) Cooperative two-input synergy engineered with PDZ-carrying sTF43-8 as the proximal activator and cognate ligand-carrying sTF42-10 as the distal partner.

(C) Cooperative two-input synergy further enhanced by the addition of an AD onto the distal partner to create a two-activator system.

(D) A “null” two-input system engineered by abolishing the dimerization interaction with a PDZ nonbinding ligand on the distal partner, thus rendering it noncontributory.

(E and F) Inhibitory two-input behavior engineered by reversing the activator location (from proximal to distal) and using either PDZ binding (E) or nonbinding ligands (F).

(G and H) Inhibition by the proximal monomer can be further increased by increasing the proximal ZF affinity to DNA (43-8-4x to 43-8-3x) and decreasing the distal ZF affinity to DNA (42-10-3x to 42-10-4x) in both PDZ binding (G) and nonbinding cases (H).

(I) By reversing the orientation of the operators, sTF43-8 is converted from an inhibitor to a cooperative factor to, once again, obtain cooperative transcriptional synergy in the two-input behavior. All sTFs were expressed from either ATc- or IPTG-inducible pGAL1 (500 ng/ml ATc and/or 20 mM IPTG). Horizontal axes correspond to “mean fluorescence intensity per cell (AU)” and begin at basal (promoter-only) fluorescence level. Error bars represent SD of three experiments.

the promoter (Ptashne, 1986).
Engineered transcription activator-like (TAL) effectors have recently emerged as an important alternative to customized ZFs for programming DNA specificity. Naturally occurring TAL effectors encoded by Xanthomonas bacteria bind to target DNA sequences using arrays of highly conserved 33–35 amino acid repeat domains. A single TAL effector repeat binds to one nucleotide of DNA with specificity of binding associated with the identities of amino acids at two positions known as repeat variable di-residues (RVDs). TAL effector repeats bearing different RVDs have been described for specific binding to each of the four possible DNA nucleotides, and these repeats can be simply joined together to create arrays capable of binding to extended target DNA sequences (Bogdanove and Voytas, 2011). The simplicity and modularity of the TAL effector repeats suggests that nearly any DNA sequence should be targetable, and recent work has demonstrated that engineered TAL effector nuclease (TALEN) fusions can be robustly generated for a wide variety of different sequences (Reyon et al., 2012). Nonetheless, engineered ZFs have several important advantages, including their considerably smaller size, their less repetitive coding sequence (potentially important for packaging into viral vectors), and a greater understanding of their biochemical properties, structure, and function, which is important for creating variations in affinity, specificity, and tunability.

We constructed a synthetic transcriptional cascade in yeast that was used as a test bed for systematically tuning stF component properties. For instance, we tuned the strength of transcriptional outputs through modifications of ZF binding sites in the promoter architecture as well as through structure-guided modifications to the ZF protein to alter ZF–DNA interactions. Additionally, our framework provides the ability to engineer and tune transcriptional cooperativity. To date, there exists no simple way of building cooperative transcriptional systems, even though their importance is well-documented in both natural and synthetic gene regulation. As a result, in most synthetic studies, researchers have used TFs with integrated, cooperative properties. In contrast, our approach establishes a modular framework for constructing cooperative transcriptional activation schemes de novo, through the multimerization of weakly-activating ZF-TFs using low-affinity protein interaction domains (i.e., PDZ domains). This has important consequences for constructing higher-order complexes that more accurately mimic eukaryotic transcription regulation schemes, lead to sharper switch-like responses, and modulate cooperativity within circuits. Indeed, multimerization and cooperativity are ubiquitous molecular regulation schemes that underlie complex gating and decision-making in cells. For example, the yeast GAL1 promoter is able to integrate coactivator proteins in specific temporal order by utilizing the cooperativity of certain interactions to gate subsequent recruitment events (Bryant and Ptashne, 2003). It is of great interest to understand how activators function cooperatively to assemble specific initiation complexes and regulate transcription. Our bottom-up and modular approach to transcriptional cooperativity could be used to synthetically recapitulate such phenomena so as to study these fundamental mechanisms of regulation. This type of approach has been used to understand transcriptional synergy in prokaryotes (Joung et al., 1993, 1994), and our platform should now enable this strategy to be used to model more combinatorially complex eukaryotic promoters. Furthermore, our extensible and modular framework for multimeric and cooperative transcriptional systems may allow for the implementation of expanded computational operations in eukaryotes, such as logic devices with more input possibilities.

Previous reports have described various frameworks for creating dimeric ZF proteins. In all of these studies, elements derived from naturally occurring TFs (Pomerantz et al., 1998; Wolfe et al., 2000) or ones selected from combinatorial peptide libraries (Wang et al., 2001; Wang and Pabo, 1999) were used to dimerize two-finger units. A disadvantage of this strategy is that a dimerized two-finger complex would have a maximum specificity of 12 bps (assuming that each of the four fingers in the dimer specifies 3 bps). Our approach differs by utilizing three-finger monomers that have had their binding activities reduced by mutagenesis of nonspecific phosphate-contacting residues. This strategy creates dimeric proteins that can have specificities as high as 18 bps, a sequence long enough to be potentially unique in a mammalian genome. Furthermore, the use of modular protein-protein interaction domains for multimerization is advantageous for various reasons. For example, the interaction is orthogonal and tunable, allowing us to “match” affinities of all the component interactions making up the stF complex, and it provides extensibility for further expanding the complex architecture and dynamically increasing or decreasing the DNA binding specificity of the complexes.

We also showed that complex signal processing behaviors can arise when control of TF cooperativity is combined with the ability to engineer promoters of multimerized ZF binding sites. Cellular signal processing is a mechanism by which environmental and other signals are integrated to modulate transcription and thus control cellular processes, such as growth and stress responses. We constructed a simple, synthetic two-input transcriptional system that allowed us to decompose contributions from the stF component properties to the system’s processing behavior. We showed that, with the same two core TFs and promoter operator sites, a cell could process and integrate signals in a variety of ways. For example, subtle changes, such as reversing promoter operators and disrupting protein-protein interactions, can have striking effects on the output of the system. This led to the construction of not just varied digital logic behavior, but a range of analog tunability. In an inhibitory system (Figures 6E–6H), we arrived at an interesting Boolean logic gate that produced a positive signal only in the presence of a single input: A > B (A does not imply B). A broad observation from our studies was that specific perturbations to an stF’s component properties (DNA affinity, multimerization with other species, location of operator, etc.) could allow it to interconvert between different transcriptional roles within the complex, such as activator, cooperative factor, noncontributory, and inhibitor. This synthetic approach could be utilized to explore the diversity of behaviors that can be programmed by even just a few transcriptional components; furthermore, our findings provide simple strategies for reprogramming the signal processing behavior of a cell. Similar strategies are undoubtedly employed naturally, where there are many examples of individual proteins that can take on either activating or repressing roles
depending on the cellular and environmental states (Ma and Ptashne, 1988; Maxon and Herskowitz, 2001; Rubin-Bejerano et al., 1996).

Given that TFs containing ZFs play a central role in eukaryotic promoter regulation (Pabo et al., 2001), our system represents a promising means for synthetic recapitulation of eukaryotic promoter function, and thus will significantly enhance our ability to construct synthetic gene networks in mammalian cells, an area of tremendous potential in synthetic biology (Weber and Fussenegger, 2009). Indeed, yeast may serve as a well-characterized testbed for the design and validation of synthetic gene circuits that can be subsequently ported to higher organisms. Synthetic transcriptional regulators based on the sTFs described here can be used to create classifier circuits to identify cell state (Nissim and Bar-Ziv, 2010; Xie et al., 2011), memory devices to record cellular events, and logic gates for cellular signal processing (Kramer et al., 2004), which can aid in the study and control of stem cell differentiation, therapeutics, and complex human diseases. Additionally, ZF-based proteins have been shown to be powerful targeting elements of endogenous genomic loci in many mammalian cells, including cancer, immune, and stem cells. These proteins include ZF nuclease (Zou et al., 2009), which are being tested in therapeutic applications for modifying/disrupting disease-causing alleles and genes, and artificial TFs (Blancafort et al., 2005), which can be used to modify the expression of native genes for reprogramming purposes. Together with these ZF technologies, our work may lead to the construction of integrated gene circuits—artificial circuitry that seamlessly and specifically integrates into endogenous gene networks and/or leads to the modification of endogenous genes—for more dynamical and sophisticated genetic control in cell-based therapeutic applications.

Synthetic biology is helping us to understand how organisms behave and develop through the forward engineering of molecular circuitry with well-understood genetic components (Elowitz and Lim, 2010). The present work expands the synthetic biology toolkit with new genetic components, beyond re-purposed bacterial transcriptional components, to program eukaryotic cells. Additionally, it provides a bottom-up framework for exploring the complexity of eukaryotic promoters and their combinatorial regulation by TF complexes and circuitry. This framework can be a starting point for determining the transcriptional components, modules, and circuitry needed to implement the sophisticated behaviors that control the development and function of eukaryotic cells.

**EXPERIMENTAL PROCEDURES**

**Strains and Media**

*S. cerevisiae* YPH500 (*α*, ura3-52, lys2-801, ade2-101, trp1Δ63, his3Δ200, leu2Δ1) (Stratagene) was used as the host strain in all yeast experiments. Culturing, genetic transformation, and verification of transformation were done as previously described (Murphy et al., 2007), using either the URA3, TRP1, or LEU2 genes as selectable markers.

**Plasmid Construction**

Synthetic promoter plasmids were constructed from integrative plasmid pRS406 (Stratagene) by cloning ZF binding sequences (BS) directly upstream of the CYC1 minimal promoter TATA box. The corresponding ZF-activated promoter drives the expression of a yeast enhanced green fluorescent protein (yEGFP) (Cormack et al., 1997), which is preceded by a Kozak consensus sequence.

sTF circuit plasmids were constructed from the previously described yeast integrative plasmids pTPG1 (TX: TetR-regulated control promoter) and pLOG1 (LX: LacI-regulated control promoter) (Ellis et al., 2009). Briefly, these plasmids consist of a GAL1 upstream activation signal (UAS) followed by either a TetR- or LacI-regulated wild-type GAL1 promoter, which drives the expression of our sTF cassettes. The strong constitutive TEF1 promoter also directs the expression of yeast codon-optimized versions of TetR (Tn10.B tetracycline repressor) and the Escherichia coli LacI inhibitor (LacI). Constitutive expression of the repressors ensures low basal levels of expression of our sTF cassettes from the engineered GAL1 promoter, which can be relieved by the respective addition of the chemical inputs, ATP and IPTG, to the medium.

The sTF cassette, which was synthesized (DNA2.0) and cloned as a KpnI/Xhol fragment into pTPG1 and pLOG1, consists of an open cloning site for engineered ZF arrays. Upon insertion of a ZF gene, the resulting minimal sTF becomes (N- to C-terminal): 3xFLAG – NLS – VP16 AD – ZF array. All ZF genes were codon-optimized, individually synthesized (IDT), and cloned as XbaI/BamHI fragments into the cassette. Protein-protein interaction domains, namely, syntrophin PDZ domain and peptide ligands, were added as C-terminal fusions to the sTF, separated by a GSGS linker, and cloned from synthesized, codon-optimized gene fragments (DNA2.0).

**Induction Experiments**

Single yeast colonies for each strain were picked and used to inoculate 500 μl of SD-Glu (synthetic drop-out media containing 2% glucose with selectable amino acid mixtures) in Costar 96-well assay blocks (V-bottom; 2 ml max volume; Fisher Scientific). The cultures were grown at 30°C with 900 rpm shaking for 24–48 hr. A triplicate set of 500 μl YEP-Gal (yeast extract peptone media containing 2% galactose) cultures, with and without inducers, were inoculated to an OD<sub>600</sub> of ~0.08–0.1 and grown at 30°C with 900 rpm shaking for ~14–16 hr. Cells were then treated with cycloheximide to inhibit protein synthesis, and then assayed for yEGFP expression by flow cytometry.

**Flow Cytometry and Data Analysis**

For all data, we acquired 5,000–10,000 events using a BD LSFortessa equipped with a High Throughput Sampler (BD Biosciences). Events were gated by forward and side scatter. The geometric means of the fluorescence distributions were calculated. The autofluorescence value of *S. cerevisiae* YPH500 cells harboring no genomic integrations was subtracted from these values to give the fluorescence values reported in this study. “Fold activation” values were calculated as the ratio of fluorescence values from induced cells to those from uninduced cells.

**Growth Assays**

Growth assays were performed similarly to induction experiments, except that experimental cultures were inoculated to an OD<sub>600</sub> of ~0.03–0.05 and grown at 30°C with 900 rpm shaking for 30 hr. OD<sub>600</sub> measurements were taken using a SpectraMax M5 fluorescence microplate reader (Molecular Devices) using culture volumes of 100 μl. A “no ZF” control—a strain engineered with synthetic promoter and sTF cassette lacking a ZF array—was assayed in parallel.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Extended Experimental Procedures and five figures and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2012.05.045.

**ACKNOWLEDGMENTS**

We thank members of the Collins Lab for helpful discussions, Carl O. Pabo for helpful suggestions on reducing nonspecific affinity of zinc fingers, and Katie M. Flynn for help with artwork. This work was supported by the Howard
REFERENCES


