ACS SyntheticBiology
Cite This: ACS Synth. Biol. 2019, 8, 2558−2565
Research Article
pubs.acs.org/synthbio

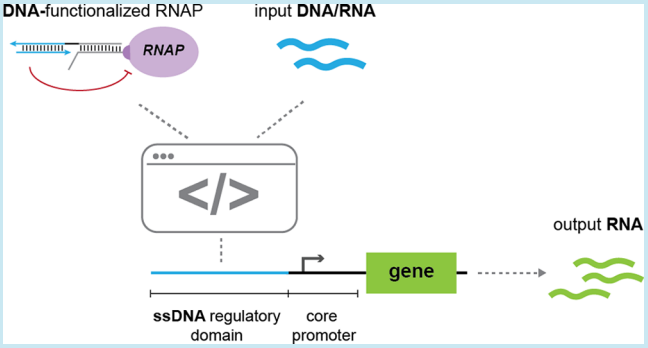
In Vitro Transcriptional Regulation *via* Nucleic-Acid-Based Transcription Factors

Leo Y. T. Chou*,†,‡,§ and William M. Shih*,†,‡

†Dana-Farber Cancer Institute, Boston, Massachusetts 02215, United States
‡Wyss Institute of Biologically Inspired Engineering at Harvard University, Boston, Massachusetts 02215, United States

Supporting Information

ABSTRACT: Cells execute complex transcriptional programs by deploying distinct protein regulatory assemblies that interact with cis-regulatory elements throughout the genome. Using concepts from DNA nanotechnology, we synthetically recapitulated this feature in *in vitro* gene networks actuated by T7 RNA polymerase (RNAP). Our approach involves engineering nucleic acid hybridization interactions between a T7 RNAP site-specifically functionalized with single-stranded DNA (ssDNA), templates displaying cis-regulatory ssDNA domains, and auxiliary nucleic acid assemblies acting as artificial transcription factors (TFs). By relying on nucleic acid hybridization, *de novo* regulatory assemblies can be computationally designed to emulate features of protein-based TFs, such as cooperativity and combinatorial binding, while offering unique advantages such as programmability, chemical stability, and scalability. We illustrate the use of nucleic acid TFs to implement transcriptional logic, cascading, feedback, and multiplexing. This framework will enable rapid prototyping of increasingly complex *in vitro* genetic devices for applications such as portable diagnostics, bioanalysis, and the design of adaptive materials.



KEYWORDS: *DNA nanotechnology, cell-free synthetic biology, transcription factors, gene circuits*

Living cells use information encoded in biochemical circuits to make complex decisions and perform sophisticated tasks. Inspired by their rich functionality, synthetic gene circuits are currently being developed to model biology and engineer organisms for various applications.[1,2] Recently, there has also been increasing interest to create gene circuits that operate *in vitro* using reconstituted molecular components. Compared to cellular devices, these cell-free ones are more portable, accessible, and robust. These advantages are now being explored for applications such as point-of-care diagnostics,[3] artificial cells,[4−6] expression of toxic products,[7] screening,[8] and even for educational purposes.[9] Yet as with cellular devices, scaling up the complexity of synthetic gene circuits requires a large toolbox of regulatory elements that can wire up genetic elements without introducing cross-talk. In living cells, this circuit wiring is achieved *via* interactions between TFs with cis-regulatory elements distributed throughout the genome. The molecular properties of TFs enable sophisticated self-assembly mediated regulatory behaviors, including recognition of specific promoters,[10] recruitment of coregulatory units, signal integration *via* multicomponent assembly,[11] and even physical alteration of genome structure.[12] Engineering these regulatory behaviors has been a rate-limiting step in the design of synthetic gene circuits.

In contrast to proteins, nucleic-acid-based regulatory elements offer a solution for programmable gene regulation by relying on Watson−Crick hybridization for predictable self-assembly, and by taking advantage of the sophisticated software tools that are available to predict nucleic acid interactions. Recently, several synthetic, programmable RNA-based regulatory devices have been developed, such as small transcriptional activating RNAs (STARs)[13] and toehold switches,[14] which regulate transcription elongation or the translation of mRNA into proteins, respectively. In contrast, fewer programmable mechanisms exist for regulation at the level of transcription initiation. Pioneering efforts to engineer *in vitro* transcriptional networks using nucleic acids have largely focused on manipulating the interactions of T7 RNAP with its promoter by either making the promoter single-stranded (*i.e.*, inactive) or double-stranded (*i.e.*, active).[15] While this strategy has enabled the construction of *in vitro* circuits with interesting dynamics,[16,17] its scalability is restricted by the T7 promoter sequence. Here we describe a new regulatory architecture for *in vitro* transcriptional regulation that alleviates this constraint. This architecture supports the use of arbitrary sequences of DNA or RNA as inputs to produce arbitrary RNA-based outputs, making the transcriptional network modular and composable (Figure 1), which lends itself to standardization, abstraction, and scaling. In this study, we experimentally test

Received: June 4, 2019
Published: October 1, 2019

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

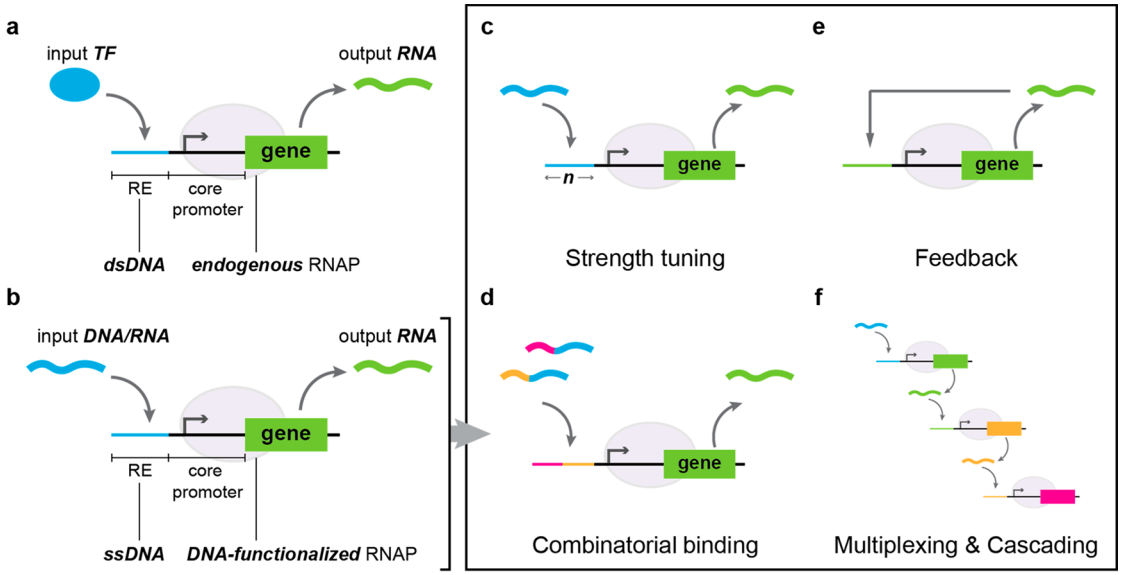


Figure 1. Overview of nucleic-acid-regulated transcription. (a) General schematic of endogenous regulation of transcription initiation. Protein-based transcription factors (TFs) bind to regulatory element (RE) upstream of the core promoter region to either enhance or suppress gene expression. (b) Nucleic-acid-based regulatory architecture developed in this study. Instead of using protein-based transcription factors, DNA/RNA regulatory assemblies are engineered to interact with ssDNA *cis*-regulatory elements *via* sequence-programmable hybridization for local enhancement or suppression of the activity of a DNA-functionalized T7 RNAP. Since both input and output of the gene are in the form of nucleic acids, and there are no sequence constraints, this mechanism of gene regulation allows for the rational design of a number of behaviors including the following: (c) tunable transcriptional strengths *via* the length of the RE domain, *n*, (d) combinatorial and cooperative activation, (e) feedback, and (f) multiplexing and cascading.

and validate this strategy *in vitro* using defined transcriptional buffers, as well as discuss how this gene-regulatory strategy might be deployed to biochemically more complex systems such as cell-free extracts in future studies.

RESULTS

Programmable Transcription Using a "Caged" DNA-Functionalized T7 RNAP. The wild-type T7 RNAP lacks regulatory mechanisms beyond its ability to recognize and bind to its 17-bp promoter. To expand on its regulatory capacity, we created a ssDNA-functionalized T7 RNAP by covalently coupling a 21-nt single-stranded DNA to a recombinant T7 RNAP bearing an N-terminal SNAP-tag (Figure S1). Hybridization of a synthetic duplex to this ssDNA-tag yielded a "caged" T7 RNAP whose activity could be controlled in programmable fashion *via* DNA strand displacement (Figure 2a).[18] The cage duplex encodes a T7 promoter lacking a downstream gene (labeled as $P_{T7,\,\Delta GGG}$ in Figure 2b). Initial experiments showed that this duplex does not produce any transcript on its own but strongly inhibits the transcription of other DNA templates in an *in vitro* transcription reaction (Figure S2). For example, the addition of this cage duplex at a 1:1 molar ratio relative to the DNA template reduced the amount of transcript produced from the reaction by >90% (Figure S2b,c). A comparison between six cage designs (see Figure S2a) revealed that this strong level of inhibition is maintained for cages encoding 0−3 nucleotide transcript downstream of the promoter but diminishes for longer transcripts (Figure S2a−c). Previous structural[19] and single-molecule studies[20] have shown that the T7 RNAP remains bound to its promoter throughout the transcription initiation phase, and that promoter escape and eventual transition to the fully processive elongation phase is mediated *via* the synthesis of an 8−12 nucleotide transcript. These prior studies, together with our results, suggest that a "gene-less" promoter functions as a T7 RNAP cage by trapping the polymerase in its early initiation phase and preventing its association with other DNA templates in solution. In addition, a 14-nt "RNAP-tether" domain on the 5′ end of the cage (colored gray in Figure 2b) enabled its hybridization to the ssDNA-functionalized RNAP, which further enhanced its efficiency of suppressing RNAP activity. Recovery of RNAP activity is achieved *via* a simple strand-displacement operation. For example, a template can be programed to display a complementary ssDNA "operator" domain that invades and displaces the cage duplex from the RNAP-cage complex (Figure 2c). This operation concurrently anchors the uncaged RNAP onto the operator domain upstream of the promoter of the template, thereby priming transcription initiation of the downstream gene.

When bound to the RNAP *via* its ssDNA-tag, the cage is highly efficient by virtue of its close proximity to the RNAP, whereas the transition from RNAP-cage complex to RNAP-template complex abolishes this situation and concurrently places the RNAP in close proximity to the template. We expected this to result in a large change in transcriptional velocity between the caged (*i.e.*, OFF) and the uncaged (*i.e.*, ON) states (Figure 2d). We identified parameters that affect the dynamic range between these two states by monitoring the production of a fluorescent RNA aptamer (*e.g.*, Broccoli) under both conditions. We found that Broccoli expression in the ON state scaled with the length of the ssDNA domain on the template by stabilizing RNAP binding (Figure S3). On the other hand, Broccoli expression in the OFF state was determined by caging efficiency, which varied as a function of cage sequence, concentration, tether stability, and buffer ionic strength (Figure S4). Optimizing these conditions reduced OFF-state expression to near background levels, resulting in 336-fold gene activation on a linear template (Figure 2e). As a

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

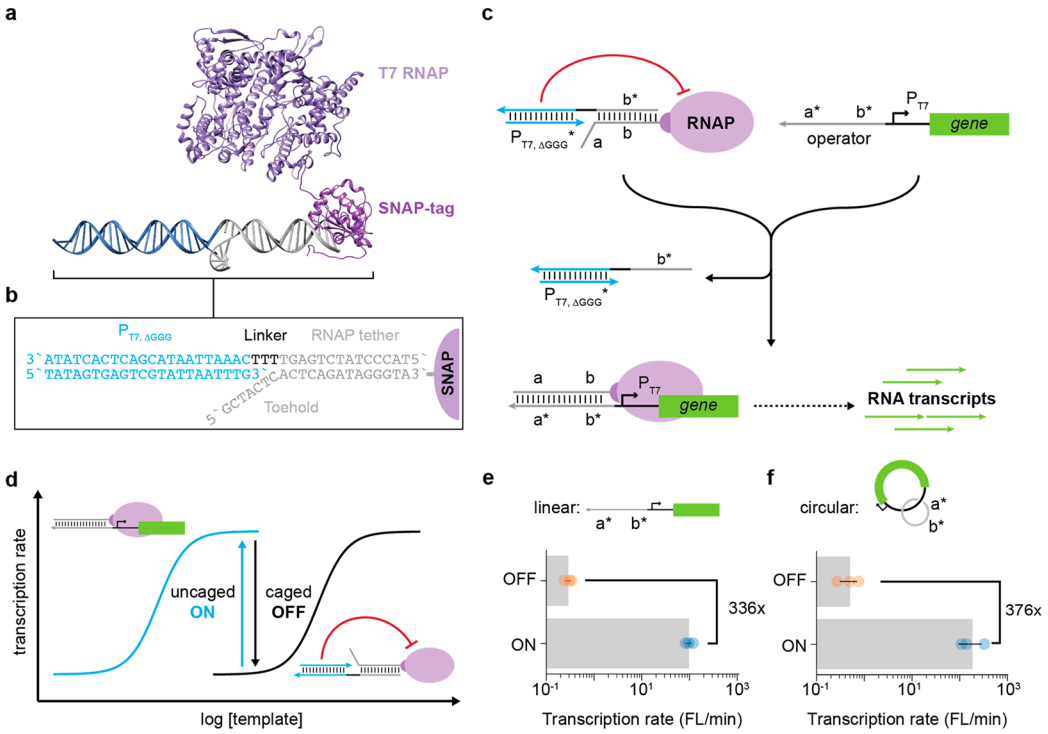


Figure 2. Transcriptional activation for DNA-functionalized T7 RNAP system. (a) Schematic of the "caged" T7 RNAP used in this study. Recombinantly expressed T7 RNAP bearing N-terminal SNAP-tag is covalently conjugated to a 21-nt ssDNA (gray strand). Hybridization of a "cage" duplex to this ssDNA-tag yields an RNAP whose activity is activatable *via* programmed DNA strand-displacement. (b) The cage duplex encodes a truncated T7 promoter (blue domain), which fails to induce transcription but nevertheless occupies the active site of the RNAP and prevents its association with other DNA templates. The duplex can be removed *via* strand displacement mediated by a 7-nt toehold positioned at the 5′ end of the ssDNA-tag. (c) Example schematic of a strand-displacement reaction for RNAP uncaging and anchoring onto a gene-of-interest. (d) Schematic depicting the expected relationship between enzyme kinetics and cage state. Hybridization of cage duplex to RNAP introduces a locally bound competitor for template binding, resulting in a large shift in RNAP activity between the caged *vs* uncaged states at most concentrations of template. (e,f) Transcription velocity of the RNAP, here monitored as the rate of production of a fluorescent RNA aptamer per unit time, in the caged (*i.e.*, OFF) *vs* uncaged (*i.e.*, ON) states, measured on either linear (e) or circular templates (f). Error bars represent standard deviation from three independent replicates.

more stringent test, we repeated this experiment by synthesizing a circular DNA template catenated with another circular ssDNA to which RNAP can localize and produce RNA continuously *via* rolling-circle transcription (Figure S5). Comparing transcription between the ON and OFF states on this circular template also showed >300-fold increase in transcription rate (Figure 2f, see also Figure S6). Together, these results illustrate programmable activation of ssDNA-functionalized T7 RNAP using nucleic acid hybridization.

Nucleic-Acid Structures as Synthetic Transcription Factors. To introduce additional regulatory mechanisms, we programmed the RNAP to colocalize with its template *via* auxiliary nucleic acid structures serving as artificial TFs. As two examples, we designed nucleic acid repressors that respectively emulate the inducible (*e.g.*, *lac*) and repressible (*e.g.*, *trp*) gene systems in *E. coli* (Figure 3a,d). These systems were picked because they have been the workhorses of synthetic gene networks, and therefore, the construction of functional nucleic acid analogs may similarly provide the basis for building more complex *in vitro* circuits. In the *lac* system, the repressor protein binds to the operator domain of the gene to block the RNAP from engaging with the promoter. Repression is allosterically alleviated by effector molecules (*e.g.*, allolactose) binding to the repressor (Figure 3a). Our mimic of the *lac* repressor consists of a linear strand that blocks the DNA-functionalized T7 RNAP from binding to the ssDNA operator domain on the template. Derepression occurs when an effector strand removes the repressor strand from the template *via* toehold-mediated strand displacement, allowing template-mediated uncaging of RNAP (Figure 3b, see also Figure S7 for sequence-level schematic). Figure 3c shows the dose−response of this scheme as a function of effector-strand concentration. The response is much sharper compared to typical *lac*-allolactose systems, reflecting the stronger binding energetics between complementary nucleic acid strands. Notably, the dynamic range of the dose−response is systematically tunable *via* sequence design. For example, the inset in Figure 3c shows how the response for a given effector concentration decreases as the length of its hybridizing domain decreases (see also Figure S8). In contrast to the lac system, the *trp* repressor consists of effector and repressor molecules associating cooperatively to suppress gene expression, which is the logical equivalent of a digital NAND gate (Figure 3d, see also Figure S9 for sequence-level schematic). We recapitulated this logic by designing two nucleic acid strands that associate to form a four-way junction (4WJ) with the operator domain on the template (Figure 3e). Figure 3f shows the dose−response curve of this design as a function of the effector-to-repressor ratio. Inset in Figure 3f shows how the repression changes as a function of deletions in the hybridization domain between the effector and repressor strands. We also tested replacing the 4WJ motif in the *trp* mimic with a three-way junction (3WJ) motif, which resulted in a more graded dose−response (Figure S10),

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

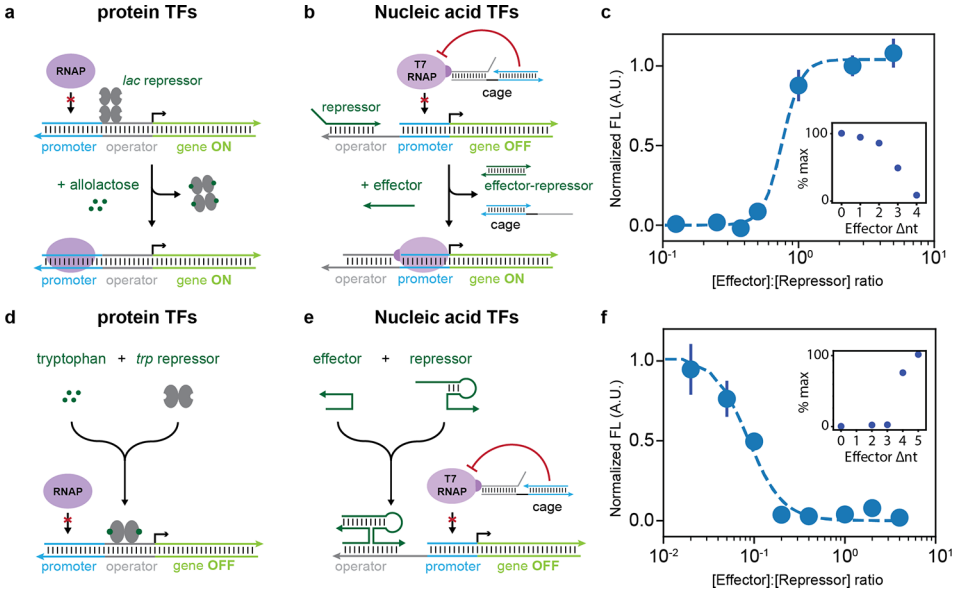


Figure 3. Synthetic recapitulation of endogenous gene-regulatory architectures. (a) Schematic of a *lac*-inducible gene in *E. coli*. Binding of the *lac*-repressor protein to the operator on the DNA template prevents RNAP binding to the promoter, resulting in gene repression. Repression is alleviated upon allosteric binding of allolactose to the *lac*-repressor protein, which turns on gene expression. (b) Nucleic-acid mimic of the *lac* system. A linear ssDNA acts as the repressor by binding to the operator region of the DNA template, masking the toehold required for template-mediated RNAP uncaging. An effector strand displaces the repressor from the template to trigger gene expression. (c) Dose−response of the nucleic acid *lac* mimic as a function of the effector-to-repressor ratio. Inset shows how the level of activation changes as a function of deletions in the hybridization domain between the effector and repressor strands from 0 to 4 nt. (d) The *trp* repressible gene. Binding of tryptophan to the *trp* repressor allosterically strengthens its affinity for the operator, resulting in gene suppression. In the absence of the tryptophan, the *trp* repressor is unable to suppress gene expression. (e) Nucleic-acid mimic of the *trp* system. The effector and repressor strands assemble to form a four-way junction (4WJ) on the template, thereby preventing template-mediated RNAP uncaging. (f) Dose−response of the nucleic acid *trp* mimic as a function of the effector-to-repressor ratio. Inset shows how the level of repression changes as a function of deletions in the hybridization domain between the effector and repressor strands from 0 to 5 nt. Error bars represent standard deviation from three independent replicates.

reflecting the known weaker binding energetics of 3WJ compared to 4WJs.[21] Together, these examples highlight how DNA nanotechnology design principles can be used to engineer different gene expression profiles.

Feedback, Cascading, and Transcriptional Multiplexing. In addition to using DNA as transcriptional regulators, we asked whether our system can also be regulated using RNA, such as using nascent transcripts to execute feedback and/or cascading. To test this, we created an autoinhibitory circuit by constructing a gene that encodes its own repressible effector molecule (Figure 4a). Expression of this gene produced RNA molecules that combine with free-floating DNA repressors to cooperatively inhibit its own transcription. Figure 4b shows how the fluorescence signal changes over time for various concentrations of repressors initially present in the reaction. Autoinhibition was demonstrated by plateauing of the fluorescence signal, and the degree to which this occurred increased with the amount of DNA repressor in solution, consistent with the notion that nascent RNA effectors combine with DNA repressors to cooperatively suppress gene expression. As another example, we constructed a two-step cascade with autocatalytic feedback (Figure 4c). The first step in the cascade is a constitutively active template that produces effectors to activate the second template. The second template is autocatalytic because it also produces its own effectors, but it is initially inhibited by excess DNA repressors. The response of the system is an exponential increase in gene expression, in this case of a fluorescent RNA aptamer, at different time points that is determined by the initial concentration of the DNA repressor (Figure 4d and S11). Together, these results demonstrate how both DNA and RNA can be used as TFs to execute transcriptional logic and feedback.

Another unique advantage of nucleic-acid-based TFs is their scalability. Because DNA hybridization relies on Watson−Crick base pairing, many instances of the same molecular motif can be created by assigning unique sequence choices for each logical domain. To test this hypothesis, we designed a new regulatory motif in which a gene can be activated upon docking of a pair of nucleic acid TFs, denoted TF_A and TF_B, equivalent to a digital AND gate (Figure 5a). Distinct from the motifs shown in Figures 2 to 4, this regulatory motif allowed us to test multiple instances of the same architecture using the same ssDNA-tethered RNAP species. We designed a set of 12 orthogonal templates based on this architecture, each encoding for a different RNA transcript "barcode", and each regulated by an orthogonal pair of TFs (Table S9). To test multiplexed gene activation, we combined the templates into a pool and performed 12 independent *in vitro* transcription reactions using this template pool, each activated using one pair of TFs. We then assayed for the identity of the RNA barcode transcribed using a set of molecular beacons each specific for one RNA barcode (Figure 5b). As a first-stage verification, we visualized the transcription reactions *via* denaturing PAGE (Figure 5c). Here we observed RNA production for all 12 TF pairs added to the template pool ("ALL", Figure 5b). With the exception of TF pairs 4 and 11, the expression levels varied by less than 2-fold across all designs (Figure 5c, bottom graph). It is yet unclear to us why TFs 4 and 11 did not operate as designed; predictions of their secondary structures and binding equilibria using the NUPACK[22] software yielded no conclusive

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

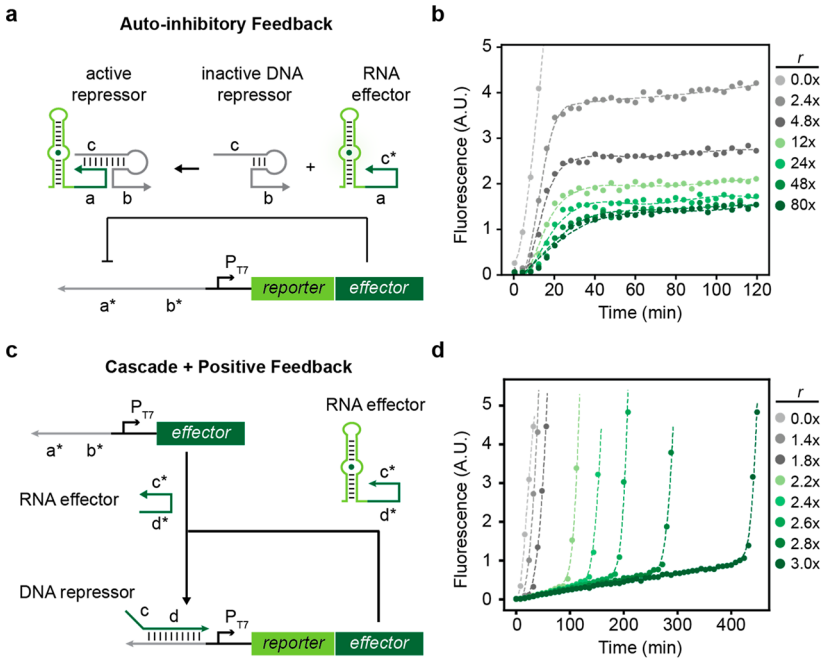


Figure 4. Implementation of negative and positive feedback. (a) Schematic of an autoinhibitory feedback loop. The DNA template encodes a fluorescent RNA aptamer (reporter) attached at its 3′-end to an effector sequence. Upon transcription, this RNA molecule assembles with free DNA repressors in solution for cooperative suppression of its own production. (b) Kinetics of the autoinhibitory response as a function of the initial repressor-to-template ratio, *r*. (c) Schematic of a two-step cascade driving an autocatalytic feedback loop. The first template in the cascade is constitutively active and produces an effector sequence that removes the DNA repressor initially bound to the second template in the cascade. The second template encodes an autocatalytic RNA molecule consisting of a fluorescent RNA aptamer (reporter) attached at its 3′-end to the same effector which, when transcribed, alleviates its own inhibition by the repressors. (d) Kinetics of the autocatalytic response as a function of the initial repressor-to-template ratio, *r*. Data shown in b and d are representative kinetic traces from a single experiment.

explanation for their performance discrepancy (Table S11). On the other hand, repeating this experiment with the target template removed from the pool resulted in no detectable gene expression across all 12 reactions, suggesting that the RNA production we trigger using the TF pair is specific to its cognate template ("LOO" panel, Figure 5b). To further validate the identities of the RNA barcodes transcribed from each reaction, we designed a set of 12 molecular beacons each specific for an RNA transcript (Figure S12). To test the ability of the RNA transcripts to activate their respective beacons, we first normalized the concentrations of the RNA transcripts across all 12 transcription reactions, and then added them separately to each of the molecular beacons. The normalization ensured that differences in the amount of activation induced by the RNA is dependent on their sequence and not due to differences in concentration. Here we observed that each RNA transcript was able to switch on their cognate molecular beacons with minimal activation of other beacons, consistent with orthogonal transcription of their respective genes in a multiplexed format (Figure 5d, raw values of the heat map in Table S12). These results demonstrate the possibility of prototyping *de novo* regulatory elements for multiplexed operation.

DISCUSSION

The utility of T7 RNAP for biotechnology applications have motivated numerous synthetic biology efforts to regulate T7 RNAP transcription. These include approaches involving RNAP engineering, such as (i) RNAP mutants with orthogonal promoter specificities,[23] (ii) split or fragmented RNAPs whose expressions are separately inducible,[24−27] and (iii) RNAP fusion with DNA-binding proteins that mediate polymerase binding to a weakened T7 promoter.[28] Alternatively, allosteric TFs have been adapted for regulating T7 RNAP transcription.[29,30] This strategy involves inserting an "operator" domain downstream of the T7 promoter to which an allosteric TF binds and prevents polymerase from switching from its initiation to the elongation phase, whereas transcription activation occurs upon ligand-induced TF de-repression. This approach has the advantage that it circumvents polymerase engineering, and that different TF-operator pairs can be used to create DNA templates responsive to different ligands. However, the scalability and programmability of this approach depends on the development of allosteric TFs with different ligand-specificities and ligand-induced binding properties. In this study, we circumvented the need to develop *de novo* protein−protein and protein-DNA binding interactions by developing an entirely nucleic-acid-based gene-regulatory framework for T7 RNAP *in vitro* transcription. This framework uses synthetic nucleic acid assemblies to program the association and subsequent activity of an ssDNA-functionalized T7 RNAP with its DNA templates, analogous to how endogenous TFs interact with cis-regulatory elements to recruit RNAP to transcriptional start sites. By replacing protein-DNA interactions with nucleic acid hybridization, we were able to rapidly prototype orthogonal gene-regulatory elements and different gene-regulatory logics. While nucleic acids have been previously used to control T7 RNAP activity,[15,31] these studies have largely focused on switching the state of the promoter between ssDNA *versus* dsDNA, which poses inherent sequence and structural constraints. By alleviating these constraints, our framework further improves the scalability of nucleic-acid-based gene-regulatory networks. Another unique feature of our

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

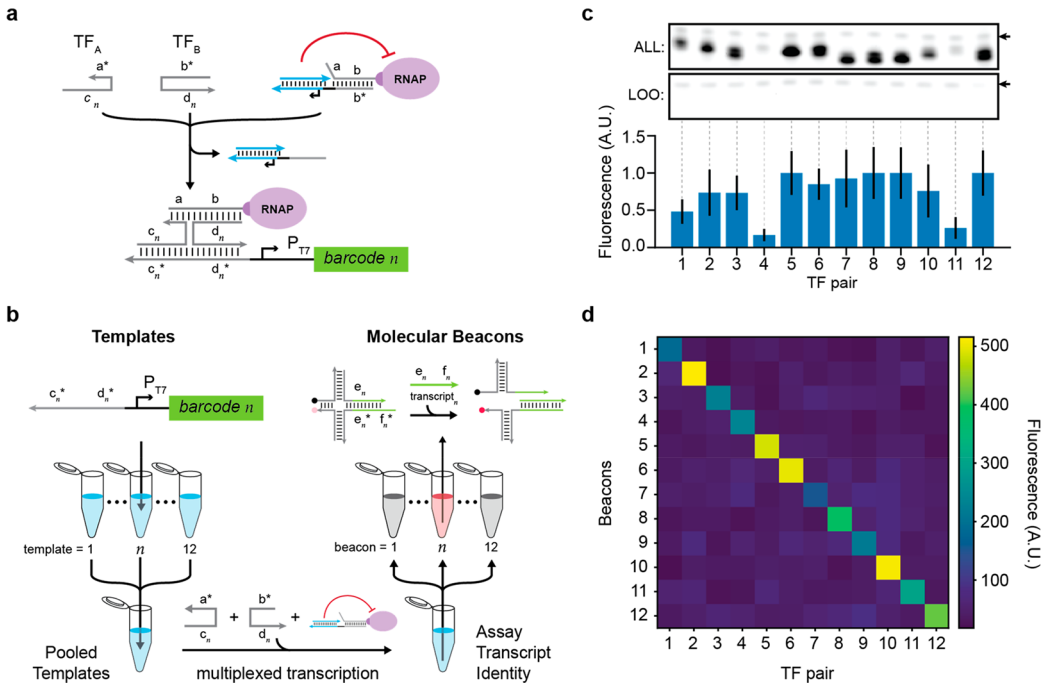


Figure 5. Implementation of transcriptional multiplexing. (a) Schematic of the regulatory architecture of the genetic templates used for testing multiplexed transcription. Each template *n* encodes a unique RNA barcode *n* whose transcription is activated upon the binding of a pair of nucleic acids, denoted TF_A and TF_B, to the template operator. (b) Schematic of the experiment used for multiplexed transcription and RNA identification. Twelve templates each encoding a different RNA barcode and under the regulation by the architecture shown in (a) are combined into a pool. This pool is used to set up 12 independent *in vitro* transcription reactions each performed using one TF pair. The identity of the RNA transcribed from each reaction is verified using a set of 12 molecular beacons specific for each RNA barcode. (c) TBE-Urea PAGE showing RNA produced from the 12 independent multiplexed *in vitro* transcription reactions. "ALL": *in vitro* transcriptions using a template pool containing all 12 templates. "LOO": *in vitro* transcriptions using a "leave-one-out" pool, in which the cognate template for the TF pair was removed from the pool. Arrow in the gels points to the cage duplex, here used as a loading control to normalize signals across samples. Bottom: Quantification of RNA produced from the transcription reactions from the "ALL" template pool normalized to the highest level. Error bars denote standard deviation from three independent experiments. (d) Activation of molecular beacons by RNA produced from each multiplexed transcription reaction displayed as an activation matrix. Signals across matrix diagonal represent specific activation while off-diagonals indicate nonspecific activation.

framework is that both the regulatory input and output consist of nucleic acids, which should allow for the programmable design of transcriptional cascades. These advantages will enable the construction of more sophisticated *in vitro* genetic circuits for various applications.

A potential concern of our T7 RNAP regulatory mechanism is the possibility for an uncaged T7 RNAP to initiate the transcription of templates that it is not bound to *via* its DNA tether, which would introduce leaks in the gene network. Two strategies were used to minimize this possibility: (1) our design stipulates that polymerase uncaging only occurs locally on the template-of-interest, either *via* direct displacement of the cage duplex by the template operator domain, or indirectly *via* nucleic acid transcription factors preassembled on the template. This stipulation couples the uncaging of the polymerase to its anchoring on the template, thereby priming transcription of the bound template over other templates in solution; (2) we modified the transcription buffer with 1× phosphate buffered saline, which contains chloride ions known to weaken polymerase affinity for its promoter. We found that this modification suppressed transcription of unbound templates by 1 order of magnitude compared to 1.4-fold for bound templates (Figure S4). This differential effect on polymerase activity further reinforced nucleic-acid-mediated transcriptional regulation.

Another important question concerns whether the proposed strategy works for longer linear DNA templates, such as templates encoding mRNAs. Since the uncaging process anchors the polymerase to a specific location on the gene, this immobilization can impede its ability to slide along the gene during the course of transcription. Addressing this question will require further work as the longest RNA transcript tested in this study was 66-nt. One strategy we used to circumvent this concern altogether was to construct circular DNA templates catenated with another ssDNA circle, which we referred to as the leash. The leash encoded binding sites to anchor the polymerase for rolling-circle transcription of the DNA template without introducing strain (Figure 2f). This strategy should be applicable to any circular DNA templates, including plasmids encoding mRNA-length transcripts. To avoid generating concatemers, a T7 terminator sequence can be included in the plasmid sequence.

We note some limitations in the current implementation of this technology. First, the uncaging step is currently irreversible, which limits control over circuit dynamics. Two potential strategies to enable reversibility include: (1) using the catalytic seesaw gate motifs developed by the Qian group,[32] and (2) producing RNA-based TFs that can be degraded using endoribonucleases, such as RNase H. How well these strategies perform as a function of circuit design warrants future studies. Second, the implementation of feedback currently uses free DNA-based repressors in solution, whereas in some reaction cascades it might be preferable to express all the regulatory components dynamically, *i.e.*, using RNA. However, doing so

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

might slow down circuit response time as the concentrations of these components need to accumulate. One strategy to overcome this is to build spatially localized gene clusters that coregulate over short diffusion lengths, such as encapsulating feedback circuits within artificial cells.[33] Whether to use pre-existing, though "inactive" DNA regulatory components that become activated upon assembly with RNA transcripts, or to express all components dynamically as RNA, will likely be dependent on the application. Finally, all of the experiments presented in this study were performed in chemically defined buffers, whereas in some cases, it might be desirable to operate this regulatory mechanism in more complex systems, such as cell-free extracts, such as for cell-free protein synthesis applications. This would require manipulating RNase and DNase concentrations in the extract to prevent premature circuit deterioration due to uncontrolled nucleic acid degradation.

We propose a number of future directions for further advancing this technology. The first is to realize large libraries of standardized regulatory components. In contrast to proteins, such DNA-based regulatory components should be easier to design, tune, and characterize. As proof-of-concept in this study, we developed a panel of 12 orthogonal nucleic acid transcription factors (*e.g.*, Figure 5). The design used in Figure 5a was advantageous over the designs in Figures 2−4 by enabling the regulation of transcriptional multiplexing using a single RNAP species with a universal ssDNA tether, as opposed to 12 different ssDNA-tethered RNAP species that would be required using the designs shown in Figures 2−4. Future studies could expand on the size of this panel, as well as panels of other regulatory motifs, to hundreds of standardized parts, which will support the design of more complex circuits. Nevertheless, for such larger sized networks, we speculate that it might be advantageous to further insulate genes from each other *via* the use of multiple ssDNA-tethered RNAP species, each under the control of an orthogonal cage duplex. Second, we envision future iterations of this technology to interface even more closely with DNA-based computing technologies. Systems of synthetic oligonucleotides have been successfully designed as switches, amplifiers, logic gates, and oscillators.[32,34−36] By programming these circuits to produce specific TF sequences as outputs, they can function as embedded controllers for programming gene-expression dynamics under our framework. This use of nucleic acid computing for the active, on-demand synthesis of functional RNAs could find applications in biological analysis, directed evolution, and molecular information processing. Third, we foresee ample opportunities for synthetic recapitulation of native gene-regulatory mechanisms using DNA nanotechnology. In this study, we created nucleic acid inducible and repressible genes by mimicking the structure of a prokaryotic operon (*e.g.*, Figure 3). In the future, coregulated gene clusters can be envisioned by designing nucleic acid scaffolds that mediate higher-order organization of genes and TFs, analogous to the actions of endogenous long-noncoding RNAs,[37] or by organizing genes into artificial DNA nanostructures that can reconfigure in analogous fashion to chromatin reorganization.[38] These efforts will enable more sophisticated levels of synthetic gene regulation. Finally, yet still more refined gene regulation can be explored by merging our work with those operating at the post-transcriptional levels[13,14] and with methods based on spatial patterning and compartmentalization.[6]

Applications of cell-free synthetic gene circuits are now beginning to emerge, such as portable diagnostics, distributed biomanufacturing, and therapeutic artificial cells.[3,39,40] Regulating gene-expression dynamics is desirable in these applications in order to focus a finite amount of energy and resources toward manufacturing the right product at the right time. Compared to protein-based gene-regulatory frameworks, the nucleic-acid-based regulatory framework presented in this study offers a number of functional advantages for synthetic gene regulation. First, nucleic acid regulatory elements consume less resources and can potentially accelerate circuit-response time because their production does not involve the translation machinery. Second, for devices designed for portability, nucleic acids have favorable storage and distribution characteristics compared to proteins. Finally, for circuits being developed for point-of-care diagnostics, nucleic acid regulated circuits can directly interface with DNA and RNA molecules extracted from physiological fluids, or else with small molecules and proteins *via* the use of aptamers or DNA-encoded affinity agents. An intrinsic constraint of our approach is the need to synthesize polymerase-DNA conjugates and genes containing ssDNA domains. Nonetheless, we believe these efforts will be scalable[41] and offset by the programmability and gains in performance offered by nucleic-acid-based gene regulation.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.9b00242.

Detailed Materials and Methods, Tables S1−S12, and Figures S1−S12 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: leo.chou@utoronto.ca.

*E-mail: William_Shih@dfci.harvard.edu.

ORCID

William M. Shih: 0000-0002-1395-9267

Present Address

§Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada.

Author Contributions

L.Y.T.C. conceived the project, planned and executed the experiments, and cowrote the manuscript. W.M.S planned experiments, cowrote the manuscript, and supervised the project.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank Dr. Rasmus Sørensen and Dr. Jaeseung Hahn for helpful discussions. L.Y.T.C. acknowledges the Banting Postdoctoral Fellowship for generous support. This work was funded by support from NSF Expeditions CCF-1317291, Army Research Office W911NF-12-1-0420, and the Wyss Institute at Harvard Core Faculty Award to W.M.S.

■ REFERENCES

(1) Bashor, C. J., and Collins, J. J. (2018) Understanding Biological Regulation Through Synthetic Biology. *Annu. Rev. Biophys. 47* (1), 399−423.

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565

(2) Weber, W., and Fussenegger, M. (2012) Emerging Biomedical Applications of Synthetic Biology. *Nat. Rev. Genet. 13* (1), 21−35.
(3) Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., DaleyKeyser, A., Yin, P., and Collins, J. J. (2014) Paper-Based Synthetic Gene Networks. *Cell 159* (4), 940−954.
(4) Karzbrun, E., Tayar, A. M., Noireaux, V., and Bar-Ziv, R. H. (2014) Synthetic Biology. Programmable on-Chip DNA Compartments as Artificial Cells. *Science 345* (6198), 829−832.
(5) Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R., and Boyden, E. S. (2017) Engineering Genetic Circuit Interactions within and between Synthetic Minimal Cells. *Nat. Chem. 9* (5), 431−439.
(6) Dupin, A., and Simmel, F. C. (2019) Signalling and Differentiation in Emulsion-Based Multi-Compartmentalized in Vitro Gene Circuits. *Nat. Chem. 11* (1), 32−39.
(7) Salehi, A. S. M., Smith, M. T., Bennett, A. M., Williams, J. B., Pitt, W. G., and Bundy, B. C. (2016) Cell-Free Protein Synthesis of a Cytotoxic Cancer Therapeutic: Onconase Production and a Just-Add-Water Cell-Free System. *Biotechnol. J. 11* (2), 274−281.
(8) Catherine, C., Lee, K.-H., Oh, S.-J., and Kim, D.-M. (2013) Cell-Free Platforms for Flexible Expression and Screening of Enzymes. *Biotechnol. Adv. 31* (6), 797−803.
(9) Huang, A., Nguyen, P. Q., Stark, J. C., Takahashi, M. K., Donghia, N., Ferrante, T., Dy, A. J., Hsu, K. J., Dubner, R. S., Pardee, K., et al. (2018) BioBits Explorer: A Modular Synthetic Biology Education Kit. *Sci. Adv. 4* (8), eaat5105.
(10) Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K. R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013) DNA-Binding Specificities of Human Transcription Factors. *Cell 152* (1−2), 327−339.
(11) Allen, B. L., and Taatjes, D. J. (2015) The Mediator Complex: A Central Integrator of Transcription. *Nat. Rev. Mol. Cell Biol. 16* (3), 155−166.
(12) Cavalli, G., and Misteli, T. (2013) Functional Implications of Genome Topology. *Nat. Struct. Mol. Biol. 20* (3), 290−299.
(13) Chappell, J., Takahashi, M. K., and Lucks, J. B. (2015) Creating Small Transcription Activating RNAs. *Nat. Chem. Biol. 11* (3), 214−220.
(14) Green, A. A., Silver, P. A., Collins, J. J., and Yin, P. (2014) Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell 159* (4), 925−939.
(15) Kim, J., Hopfield, J. J., and Winfree, E. (2004) Neural Network Computation by in Vitro Transcriptional Circuits. *Adv. Neural Inf. Process. Syst. 17*, 681−688.
(16) Kim, J., White, K. S., and Winfree, E. (2006) Construction of an in Vitro Bistable Circuit from Synthetic Transcriptional Switches. *Mol. Syst. Biol. 2* (1), 68.
(17) Kim, J., and Winfree, E. (2011) Synthetic in Vitro Transcriptional Oscillators. *Mol. Syst. Biol. 7* (1), 465.
(18) Zhang, D. Y., and Seelig, G. (2011) Dynamic DNA Nanotechnology Using Strand-Displacement Reactions. *Nat. Chem. 3* (2), 103−113.
(19) Steitz, T. A. (2009) The Structural Changes of T7 RNA Polymerase from Transcription Initiation to Elongation. *Curr. Opin. Struct. Biol. 19* (6), 683−690.
(20) Tang, G.-Q., Roy, R., Bandwar, R. P., Ha, T., and Patel, S. S. (2009) Real-Time Observation of the Transition from Transcription Initiation to Elongation of the RNA Polymerase. *Proc. Natl. Acad. Sci. U. S. A. 106* (52), 22175−22180.
(21) Kadrmas, J. L., Ravin, A. J., and Leontis, N. B. (1995) Relative Stabilities of DNA Three-Way, Four-Way and Five-Way Junctions (Multi-Helix Junction Loops): Unpaired Nucleotides Can Be Stabilizing or Destabilizing. *Nucleic Acids Res. 23* (12), 2212−2222.
(22) Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., and Pierce, N. A. (2011) NUPACK: Analysis and Design of Nucleic Acid Systems. *J. Comput. Chem. 32* (1), 170−173.
(23) Temme, K., Hill, R., Segall-Shapiro, T. H., Moser, F., and Voigt, C. A. (2012) Modular Control of Multiple Pathways Using Engineered Orthogonal T7 Polymerases. *Nucleic Acids Res. 40* (17), 8773−8781.
(24) Han, T., Chen, Q., and Liu, H. (2017) Engineered Photoactivatable Genetic Switches Based on the Bacterium Phage T7 RNA Polymerase. *ACS Synth. Biol. 6* (2), 357−366.
(25) Segall-Shapiro, T. H., Meyer, A. J., Ellington, A. D., Sontag, E. D., and Voigt, C. A. (2014) A 'Resource Allocator' for Transcription Based on a Highly Fragmented T7 RNA Polymerase. *Mol. Syst. Biol. 10* (7), 742.
(26) Shis, D. L., and Bennett, M. R. (2013) Library of Synthetic Transcriptional AND Gates Built with Split T7 RNA Polymerase Mutants. *Proc. Natl. Acad. Sci. U. S. A. 110* (13), 5028−5033.
(27) Schaerli, Y., Gili, M., and Isalan, M. (2014) A Split Intein T7 RNA Polymerase for Transcriptional AND-Logic. *Nucleic Acids Res. 42* (19), 12322−12328.
(28) Hussey, B. J., and McMillen, D. R. (2018) Programmable T7-Based Synthetic Transcription Factors. *Nucleic Acids Res. 46* (18), 9842−9854.
(29) Alam, K. K., Jung, J. K., Verosloff, M. S., Clauer, P. R., Lee, J. W., Capdevila, D. A., Pastén, P. A., Giedroc, D. P., Collins, J. J., and Lucks, J. B. (2019) Rapid, Low-Cost Detection of Water Contaminants Using Regulated In Vitro Transcription. *bioRxiv*, DOI: 10.1101/619296.
(30) Ike, K., Arasawa, Y., Koizumi, S., Mihashi, S., Kawai-Noma, S., Saito, K., and Umeno, D. (2015) Evolutionary Design of Choline-Inducible and -Repressible T7-Based Induction Systems. *ACS Synth. Biol. 4* (12), 1352−1360.
(31) Kar, S., and Ellington, A. D. (2018) *In Vitro* Transcription Networks Based on Hairpin Promoter Switches. *ACS Synth. Biol. 7* (8), 1937−1945.
(32) Qian, L., and Winfree, E. (2011) A Simple DNA Gate Motif for Synthesizing Large-Scale Circuits. *J. R. Soc., Interface 8* (62), 1281−1297.
(33) Joesaar, A., Yang, S., Bögels, B., van der Linden, A., Pieters, P., Kumar, B. V. V. S. P., Dalchau, N., Phillips, A., Mann, S., and de Greef, T. F. A. (2019) DNA-Based Communication in Populations of Synthetic Protocells. *Nat. Nanotechnol. 14* (4), 369−378.
(34) Chen, Y.-J., Dalchau, N., Srinivas, N., Phillips, A., Cardelli, L., Soloveichik, D., and Seelig, G. (2013) Programmable Chemical Controllers Made from DNA. *Nat. Nanotechnol. 8* (10), 755−762.
(35) Srinivas, N., Parkin, J., Seelig, G., Winfree, E., and Soloveichik, D. (2017) Enzyme-Free Nucleic Acid Dynamical Systems. *Science 358* (6369), eaal2052.
(36) Fern, J., Scalise, D., Cangialosi, A., Howie, D., Potters, L., and Schulman, R. (2017) DNA Strand-Displacement Timer Circuits. *ACS Synth. Biol. 6* (2), 190−193.
(37) Rinn, J. L., and Chang, H. Y. (2012) Genome Regulation by Long Noncoding RNAs. *Annu. Rev. Biochem. 81* (1), 145−166.
(38) Li, Y., Hu, M., and Shen, Y. (2018) Gene Regulation in the 3D Genome. *Hum. Mol. Genet. 27* (R2), R228−R233.
(39) Pardee, K., Slomovic, S., Nguyen, P. Q., Lee, J. W., Donghia, N., Burrill, D., Ferrante, T., McSorley, F. R., Furuta, Y., Vernet, A., et al. (2016) Portable, On-Demand Biomolecular Manufacturing. *Cell 167* (1), 248−259.
(40) Krinsky, N., Kaduri, M., Zinger, A., Shainsky-Roitman, J., Goldfeder, M., Benhar, I., Hershkovitz, D., and Schroeder, A. (2018) Synthetic Cells Synthesize Therapeutic Proteins inside Tumors. *Adv. Healthcare Mater. 7* (9), 1701163.
(41) Minev, D., Guerra, R., Kishi, J., Smith, C., Krieg, E., Said, K., Hornick, A., Sasaki, H., Filsinger, G., Beliveau, B., et al. (2019) Rapid and Scalable in Vitro Production of Single-Stranded DNA. *bioRxiv*, DOI: 10.1101/558429.

DOI: 10.1021/acssynbio.9b00242
ACS Synth. Biol. 2019, 8, 2558−2565