ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

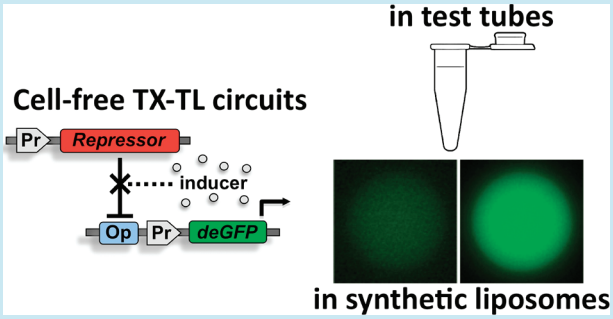
An *E. coli* Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells

Jonghyeon Shin and Vincent Noireaux*

School of Physics and Astronomy, University of Minnesota, 116 Church Street SE, Minneapolis, Minnesota 55455, United States

Supporting Information

ABSTRACT: Cell-free protein synthesis is becoming a powerful technique to construct and to study complex informational processes *in vitro*. Engineering synthetic gene circuits in a test tube, however, is seriously limited by the transcription repertoire of modern cell-free systems, composed of only a few bacteriophage regulatory elements. Here, we report the construction and the phenomenological characterization of synthetic gene circuits engineered with a cell-free expression toolbox that works with the seven *E. coli* sigma factors. The *E. coli* endogenous holoenzyme E_{70} is used as the primary transcription machinery. Elementary circuit motifs, such as multiple stage cascades, AND gate and negative feedback loops are constructed with the six other sigma factors, two bacteriophage RNA polymerases, and a set of repressors. The circuit dynamics reveal the importance of the global mRNA turnover rate and of passive competition-induced transcriptional regulation. Cell-free reactions can be carried out over long periods of time with a small-scale dialysis reactor or in phospholipid vesicles, an artificial cell system. This toolbox is a unique platform to study complex transcription/translation-based biochemical systems *in vitro*.



KEYWORDS: *in vitro transcription-translation, cell-free gene circuits, design principles, phospholipid vesicles*

Cell-free synthetic biology provides an experimental framework for broadening our knowledge of the molecular repertoire of biology through the construction of complex biochemical systems *in vitro*. Synthetic biology performed in the test tube offers a means to investigate informational or metabolic processes in isolation[1−4] and to construct artificial systems that would be impossible to develop *in vivo*.[5−9] The cell-free approach to complex molecular systems is a research area where the living cell does not necessarily stand as an absolute reference.

A particular goal of cell-free synthetic biology is the design and quantitative characterization of molecular networks composed of informational biopolymers, namely, DNA, RNA, or proteins. The development of enzyme-free biochemical circuits has been so far the most active area in that field. Models of information processing and molecular computation are tested through the synthesis of cell-free nucleic acids circuits in buffered aqueous solutions.[2,5,10] This step-by-step constructive approach is a learning method to design predictable large-scale biochemical systems.[11−14] Based on powerful reaction primitives, consisting mostly of nucleotide sequence recognition, strand displacement and branch migration, enzyme-free DNA reaction systems can theoretically reproduce the dynamical behaviors of arbitrary systems of coupled chemical reactions.[15] Furthermore, the algorithmic self-assembly of nucleic acid nanostructures has been achieved.[16−18] Certain classes of chemical behaviors, however, such as pattern formation and the regulation and time control of self-assembly and disassembly, seem far more challenging to construct with DNA only. More elaborate circuit functions and molecular devices can be constructed using a reduced set of enzymes.[7−9,19] Yet, the lack of control of enzyme synthesis and degradation limits the variety of biochemical behaviors that can be engineered *in vitro*. A cell-free system with a controlled synthesis and degradation of proteins would significantly extend the range of biochemical systems that can be constructed.

Transcription/translation cell-free systems could be used, in principle, as a technology to develop complex informational and active self-organization processes that are not accessible by the other reductionist cell-free approaches. In addition, the transcription/translation machinery is the only set of molecules that can manufacture all of its own components. This unique feature allows envisioning the synthesis of artificial self-reproducing entities. The modern cell-free expression systems, however, are optimized for biotechnology purposes,[20,21] rather than for the development of complex biochemical systems *in vitro*. The transcription of the current available cell-free systems is performed by a bacteriophage RNA polymerase added to a crude cytoplasmic extract, which provides the translation machinery. Consequently, the poor repertoire of transcription regulation offered by these hybrid systems is a considerable limitation for the development of gene circuits *in vitro*. Synthetic gene circuits, pattern formation, and artificial cell

Received: October 29, 2011
Published: December 27, 2011

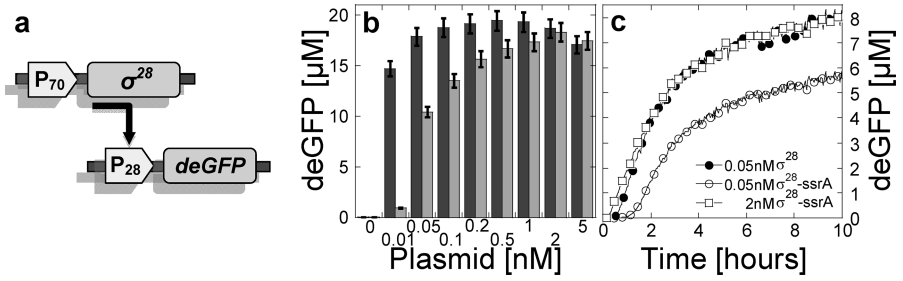


Figure 1. A two-stage transcriptional activation cascade using the *E. coli* σ^{28}. (a) Schematic of the cascade. σ^{28}, cloned under a promoter P_{70} specific to σ^{70}, activates the expression of deGFP cloned under a promoter P_{28} specific to σ^{28} (DNA part list in Supplementary Table S1). (b) End-point deGFP production as a function of the concentration of plasmid encoding the sigma factor (symbol: light gray for σ^{28}-SsrA and dark gray for σ^{28}, σ^{70} salt conditions, see Table 1). The concentration of plasmid P_{28}-*deGFP* was fixed to 5 nM. The same characterization was performed for σ^{19}, σ^{24}, σ^{32}, σ^{38}, T7, and T3 (Supplementary Figure 1). (c) Kinetics of deGFP expression (5 nM P_{28}-*deGFP*).

systems[1,22−24] have emphasized these limitations as well as the potential of cell-free systems to study complex biochemical systems *in vitro*. The PURE system, also based on bacteriophage transcription, presents the same limitations.[25] Moreover, the protein production with the PURE system is less efficient, and long-lived continuous-exchange PURE reactions have not yet been proven to work.

The recent preparation of an endogenous *E. coli* cell-free expression system, as efficient as the conventional bacteriophage systems, is a first step to transform cell-free expression as an effective platform for constructing complex transcription/translation processes.[26] This endogenous system integrates two methods to adjust the global mRNA inactivation rate and the degradation of the synthesized proteins.[27] Protein synthesis and degradation with this system has also been recently modeled.[28] The repertoire of regulatory elements provided by σ^{70} specific promoters is much larger than bacteriophage systems. Yet, the transcription modularity with one sigma factor only is restrictive. The construction of interesting gene circuits, composed of DNA parts that cannot be repeated, requires a larger repertoire of transcriptional regulatory elements.

In this work, we report the development and the phenomenological characterization of synthetic gene circuits constructed with a cell-free expression toolbox that works with the seven *E. coli* sigma factors as well as with the T7 and T3 bacteriophage RNA polymerases. The basic transcription activation layout of the toolbox is constructed so as to work like the main *E. coli* transcription scheme: the core RNA polymerase and the σ^{70} (holoenzyme E_{70}), present in the *E. coli* crude extract, are used as the primary transcription machinery. The six other sigma factors and the two bacteriophage RNA polymerases are expressed to engineer elementary gene circuits, such as transcriptional activation cascades and an AND gate. The σ^{19}, σ^{24}, σ^{28}, σ^{32}, and σ^{38}, tagged with specific AAA+ proteolytic degrons, can be expressed as degradable transcription factors. We show that the competition of different sigma factors for the core RNA polymerase can lead to a strong passive transcriptional repression, which can be used in circuits as an autoregulation of gene expression. The construction of a five-stage transcriptional activation cascade highlights the importance of the global mRNA inactivation rate, which turns out to be critical for the circuit output signal's specificity by decreasing the amplification of crosstalks between activation units. Negative feedback loops are constructed from a set of σ^{70} transcriptional repression units. The time course of gene expression, limited to 4 h in batch mode, can be extended to 16 h with a small-scale dialysis reactor. Alternatively, cell-free expression can also be carried out inside phospholipid vesicles, a synthetic cell system. The toolbox presented in this work demonstrates that cell-free expression systems can be developed as quantitative platforms to construct and to study complex transcription/translation-based biochemical systems *in vitro*.

RESULTS AND DISCUSSION

Repertoire of Cell-Free Transcriptional Activation Units. The construction of transcription/translation circuits is limited by the transcription repertoire of modern cell-free systems, which are composed of only a few bacteriophage RNA polymerases and promoters. We used an *E. coli* endogenous cell-free expression system, based on the housekeeping transcription machinery,[26] to develop a repertoire of alternative transcriptional activation units. We tested the main transcription space of *E. coli*, composed of six sigma factors in addition to the primary σ^{70}.[29] In this work, we define a transcriptional activation unit as a transcriptional activation protein and its specific promoter. A transcriptional activation cascade is the serial assembly of two or more transcriptional activation units.

A two-stage transcriptional activation cascade was constructed for each of the *E. coli* σ^{19}, σ^{24}, σ^{28}, σ^{32}, and σ^{38} transcription factors (Figure 1a, Supplementary Figure S1a−e, DNA part list in Supplementary Table S1). The σ^{54} unit, which requires the gene *ntrC* for the transcriptional co-activation of genes, was studied separately. Each transcription factor was cloned under the same promoter P_{70}. To create a set of degradable factors, the five sigma factors were also tagged with AAA+ specific proteolytic degrons. σ^{19}, σ^{28}, and σ^{32} were tagged with the C-terminal SsrA degron.[30] σ^{24} and σ^{38}, not active with the SsrA degron (data not shown), were tagged with the N-terminal OmpA degron.[30] The reporter gene *deGFP* was cloned under a promoter specific to each sigma factor (Supplementary Table S1). We chose strong promoters previously described in the literature. For comparison, we also constructed a T7 and a T3 transcriptional activation units (Supplementary Figure S1f and S1g). Degradable versions of these two units with AAA+ specific degrons were not tested. The transcriptional activation units were first validated as usable units on the basis of the maximum deGFP production in batch mode reaction, also defined as the output signal of the cascade. The requirement for the magnitude of the output signal was set to 500 nM, the average cytoplasmic protein abundance in *E. coli* cells.[31] We used this concentration as the relevant threshold output signal for all of the circuits constructed in this work.

dx.doi.org/10.1021/sb200016s | *ACS Synth. Biol.* 2012, 1, 29−41

Table 1. Optimum Magnesium Glutamate, Potassium Glutamate, and Plasmid Concentrations for 14 Different Transcription Factors and RNA Polymerases (End-Point Measurements)[a]

					deGFP[c]	
transcription factor	Mg glutamate [mM]	K glutamate [mM]	plasmid encoding the transcription factor [nM]	reporter plasmid [nM]	[μM]	(mg/mL)
σ^{70} [b]	3	60	NA	10	25	(0.63)
σ^{19}	3	20	2	10	7	(0.18)
σ^{19}-SsrA	3	20	5	10	7	(0.18)
σ^{24}	6	30	2	10	11	(0.28)
OmpA-σ^{24}	6	30	2	10	8	(0.20)
σ^{28}	3	60	0.2	5	21	(0.53)
σ^{28}-SsrA	3	60	2	5	18	(0.46)
σ^{32}	5	70	0.5	5	19	(0.48)
σ^{32}-SsrA	5	70	2	5	14	(0.36)
σ^{38}	5	100	0.5	15	13	(0.33)
OmpA-σ^{38}	5	100	0.5	15	10	(0.25)
σ^{54}/NtrC	5	30	1 (each)	10	5	(0.13)
T3 RNAP	2	80	0.2	5	27	(0.69)
T7 RNAP	2	80	0.2	5	29	(0.74)

[a]Except for the endogenous σ^{70}, deGFP was synthesized through a two-stage transcriptional activation cascade as shown in Figure 1a for σ^{28}, in Supplementary Figure 1 for the other transcription factors, and in Figure 2a for σ^{54} and NtrC (DNA part list in Supplementary Table S1). [b]Endogenous sigma factor. All the other alternative sigma factors are expressed using endogenous sigma factor 70 (Figure 1a, Supplementary Figure S1). [c]1 mg/mL deGFP = 39.4 μM.

We determined the magnesium and the potassium glutamate concentrations to get the maximum deGFP production for each cascade, before determining the optimum plasmid concentrations in each stage (Table 1). The magnesium and the potassium glutamate concentrations span a range of 4 mM and 80 mM, respectively, σ^{19} and σ^{38} being the most distant. These rather small differences, within what has been measured in *in vitro* transcription assays, are expected since promoter selectivity is based on cellular conditions.[32] We found that all of the activation units have an output signal largely above the reference concentration of 500 nM. As estimated before,[26] the maximum amount of active deGFP produced with the endogenous transcription machinery only (25 μM) is comparable to the amount produced with a bacteriophage RNA polymerase expressed through a two-stage activation cascade (27−29 μM).

σ^{28}, used *in vivo* for motility functions, is the strongest sigma factor transcriptional activation unit. In the best conditions, 21 μM of active reporter protein are produced (Figure 1b). As expected, the reporter protein production at low concentration of plasmid encoding σ^{28}, a regime where the activity of the AAA + proteases is not saturated,[27,28] is much higher than the degradable version (σ^{28}-SsrA). This trend, also observed for σ^{19}-SsrA and σ^{32}-SsrA, is not observed for OmpA-σ^{24} and OmpA-σ^{38} (Supplementary Figure S1), presumably due to the weakness or the inaccessibility of the OmpA tag. As opposed to the OmpA tag, the degradation with the SsrA degron is enhanced by the specific SspB factor present in the extract, which may also explain the difference observed between the two degrons.[28,33] The Michaelis constant of the SspB-SsrA degradation pathway is inferior to 10 nM in our cell-free system.[28] At high concentration of plasmid encoding sigma factors, the production of reporter protein with the non-degradable and the degradable versions of sigma factors is similar (Figure 1b and c, Supplementary Figure S1). In this regime of plasmid concentrations, the rate of protein degradation by the AAA+ proteases is negligible compared to the rate of protein synthesis. This trend is observed for all of the sigma factors (Supplementary Figure S1).

In *E. coli*, the sigma factor family competes for the same core RNA polymerase.[29,34−36] In our experiments, the change in the magnitude of the output signal for each unit with respect to the concentration of plasmid encoding the alternative sigma factor is an indirect measurement of the competition between each alternative sigma factor and the primary one. The change in the magnitude of the output signal depends on the relative binding affinity for each sigma factor for the core RNA polymerase. The magnitude of the output signal depends also on the strength of the promoter. Two trends are observed. For σ^{28} and σ^{32} (Figure 1b, Supplementary Figure S1d), the output signal is high even at low plasmid concentration. In the case of σ^{28}, the output signal increases by only 25% when the plasmid concentration is increased by a factor of 100 in the linear regime of plasmid concentrations (0.01 nM to 1 nM). For the σ^{19}, σ^{24}, and σ^{38} units, a continuous increase of the output signal is observed when the plasmid concentration is increased in the linear regime of plasmid concentration (Supplementary Figure S1a, S1b, and S1e). These observations globally agree with the relative binding affinities of the seven *E. coli* sigma factors for the core RNA polymerase.[34] As expected, the strong σ^{28} and σ^{32} take over the core RNA polymerase even at low plasmid concentration. The transmission of information through these two units is highly efficient compared to the σ^{19}, σ^{24}, and σ^{38} units. The switch-like input-output function of the σ^{28} and σ^{32} units may be a problem for circuit constructions as one may want a linear easily adjustable output signal. The input-output function of these two units could be linearized by adjusting the strength of the different regulatory parts (promoters and untranslated regions) and/or by using degradable sigma factors, as shown in Figure 1b and Supplementary Figure S1d. A net decrease of the output signal is observed for the σ^{38} and T7 units at high plasmid concentration (Supplementary Figure S1e and S1g), for which no explanation can be provided. The sharp response of the T7 and the T3 units to the change of concentration of the plasmid encoding the RNA polymerase is attributed to the high efficiency and specificity of bacteriophage transcriptions (Supplementary Figure S1f and S1g).

dx.doi.org/10.1021/sb200016s | *ACS Synth. Biol.* 2012, 1, 29−41

Table 2. Crosstalk between Transcriptional Activation Units Measured in the Linear Regime of Plasmid Concentration[a]

deGFP [μM]		Transcription factor expressed							
		$(\sigma^{70})^*$	σ^{19}	σ^{24}	σ^{28}	σ^{32}	σ^{38}	T3RNAP	T7RNAP
Promoters	P_{70}	6.40	5.89	5.45	3.96	4.74	6.01	6.06	5.76
	P_{19}	< 0.01	0.89	0.01	0.02	< 0.01	0.01	< 0.01	< 0.01
	P_{24}	0.01	0.03	0.74	< 0.01	0.01	0.03	0.01	0.01
	P_{28}	< 0.01	0.02	< 0.01	8.06	< 0.01	< 0.01	< 0.01	< 0.01
	P_{32}	0.20	0.21	0.20	0.12	5.44	0.20	0.20	0.22
	P_{38}	< 0.01	0.13	0.03	0.01	< 0.01	0.83	0.01	< 0.01
	P_{T3}	0.02	0.02	0.03	< 0.01	0.02	0.03	13.44	0.02
	P_{T7}	0.02	0.03	0.02	< 0.01	< 0.01	0.02	0.02	17.91
Rate of deGFP synthesis [nM/min]		Transcription factor expressed							
		$(\sigma^{70})^*$	σ^{19}	σ^{24}	σ^{28}	σ^{32}	σ^{38}	T3RNAP	T7RNAP
Promoters	P_{70}	23.91	23.53	25.68	20.16	23.08	26.98	23.15	22.98
	P_{19}	< 0.01	3.67	0.04	0.01	< 0.01	0.05	< 0.01	< 0.01
	P_{24}	0.02	0.02	2.35	< 0.01	< 0.01	0.08	0.03	0.03
	P_{28}	< 0.01	< 0.01	< 0.01	35.28	< 0.01	< 0.01	< 0.01	< 0.01
	P_{32}	0.53	0.55	0.76	0.43	24.30	0.81	0.73	0.78
	P_{38}	< 0.01	0.27	0.08	0.02	< 0.01	5.38	0.04	< 0.01
	P_{T3}	0.07	0.05	0.10	< 0.01	0.18	0.1	75.33	0.16
	P_{T7}	0.04	0.06	0.05	< 0.01	< 0.01	0.08	0.83	93.87

[a]0.1 nM plasmid encoding the sigma factor and 1 nM reporter plasmid, σ^{70} salt conditions, see Table 1). The end-point deGFP productions (upper table) and the maximum deGFP synthesis rates (lower table) for the seven *E. coli* sigma factors (non-degradable versions) and for the T7 and the T3 RNA polymerases were measured with respect to each of the specific promoters (shaded in grey) and against each of the other non-specific promoters. Except for the endogenous σ^{70} present in the reaction, the transcription factors were expressed as shown in Figure 1a. Expression through the promoter P_{70} is high in all of the cases because the endogenous σ^{70} is present in the extract. *: endogenous sigma factor. All of the others are expressed using the endogenous sigma factor 70 (Figure 1a, Supplementary Figure S1).

The kinetics of deGFP expression for each unit is similar to the typical protein synthesis dynamics in cell-free systems. An accumulation of reporter protein is observed for a few hours before the reactions stops (Figure 1c, Supplementary Figure S1). A 5−10 min delay of expression is observed for each transcriptional activation cascade compared to the expression of deGFP from a promoter P_{70}, a typical time for such two-stage cascades.[1] This delay is much larger for the degradable version of the σ^{28} and σ^{32} transcriptional activation units at low plasmid concentration due to the high efficiency of proteolysis with the SsrA tag. The kinetics of deGFP expression through the σ^{19} unit is two times shorter, which could be due to an increased instability of this sigma factor.[37]

Crosstalks between Transcriptional Activation Units. A quantitative estimation of nonspecific signals generated in elementary gene circuits is essential to construct complex informational systems with predictable behaviors. We next studied the crosstalks between the elements (sigma factors and promoters) of all of the units. To determine the complete crosstalk space of the activation unit set (σ^{19}, σ^{24}, σ^{28}, σ^{32}, σ^{38}, σ^{70}, T3, and T7), we measured in individual assays the leak of deGFP expression through each promoter against the other sigma factors and the two bacteriophage RNA polymerases. The specific and nonspecific end-point production of active deGFP and the maximum rate of deGFP synthesis were first measured in the linear regime of plasmid concentrations (Table 2, scaled values in Supplementary Table S2). The generation of undesirable signals due to the presence of alternative endogenous sigma factors in the extract was also examined. The magnesium and the potassium glutamate conditions were fixed to 3 and 60 mM respectively, the optimum salt conditions for σ^{70} (Table 1).

The nonspecific gene expression generated through the promoter P_{70} by the alternative sigma factors could not be measured because the σ^{70} is present in the reaction. The decrease of gene expression from the promoter P_{70} when one alternative sigma factor is expressed in the reaction agrees with the relative binding affinities of each sigma factor for the core enzyme measured *in vitro*,[34] σ^{28} and σ^{32} being the most competitive (Table 2). Protein synthesis through the promoter P_{70} is not sensitive to the expression of the T3 or the T7 RNA polymerases, which do not compete for the core RNA polymerase.

Several features emerge from the crosstalk table (Table 2). First, σ^{32} is the most leaky unit. The nonspecific expression induced by σ^{32} through the other promoters (P_{70} excluded) is rather small, less than 1% in the worst case. The nonspecific expression through the promoter P_{32}, however, is large even in the presence of the primary sigma factor only. This could be due either to the leak of the primary sigma factor on the promoter P_{32} or to the presence of endogenous σ^{32} in the extract. On the basis of the intracellular concentration of σ^{32} (<10 nM, Supplementary Table S5) and taking into consideration the dilution factor during extract preparation (20−30 times dilution), we hypothesize that the leak observed on the promoter P_{32} is due to the primary sigma factor. Other well-characterized σ^{32} dependent promoters could be tested to determine whether a more specific unit can be constructed with this sigma factor.

Another important feature is the high specificity of σ^{28} and its promoter. The leak through the promoter P_{28} is almost systematically below the detection limit. The nonspecific expression generated by σ^{28} through the other promoters (P_{70} excluded), less than 0.2% in the worst case, is smaller than the leak generated by the other sigma factors. The σ^{28} unit is the most efficient and the most specific sigma factor unit tested in this work. The low level of crosstalk for the two bacteriophage units, well-known for their high specificity, was expected. Overall, we can also conclude that all of the alternative sigma factors are either not present in the extract or are present at insufficient concentrations to generate undesirable signals.

The last feature in the crosstalk table is the high level of nonspecific expression induced by the σ^{19} through the promoter P_{38}. Whether this crosstalk can be decreased with a more specific σ^{38}-dependent promoter has yet to be tested. The entire crosstalk table was also determined for the saturation regime of plasmid concentration (Supplementary Table S3,

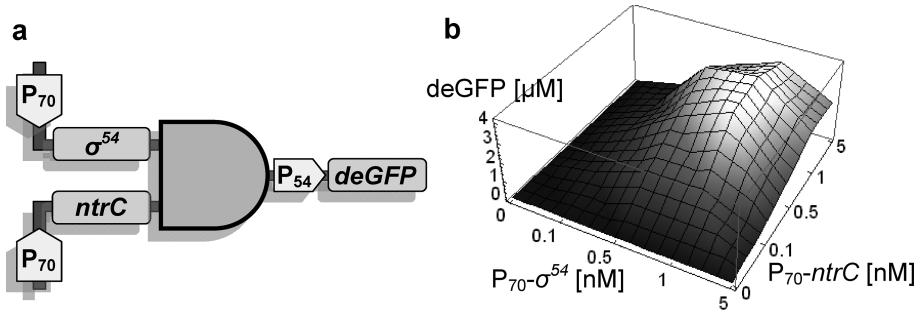


Figure 2. Co-activation of deGFP synthesis by σ^{54} and NtrC. (a) Schematic of the co-activation, shown as an AND gate circuit. The genes σ^{54} and *ntrC* were cloned under a promoter P_{70} in separate plasmids and expressed concurrently to transcribe the gene *deGFP* cloned under a promoter P_{54} (DNA part list in Supplementary Table S1). (b) 3D plot of deGFP production (end-point measurements) as a function of the concentration of the two plasmids (5 nM P_{54}-*deGFP*).

scaled values in Supplementary Table S4). As anticipated, the crosstalk level in this regime is much higher, but no new features are observed.

Transcription Co-activation by σ^{54} and NtrC: An AND Gate System. In *E. coli*, the transcription through σ^{54}-dependent promoters is co-activated by the enhancer protein NtrC (Nitrogen regulatory protein C). The *E. coli* glnAp2 regulatory DNA part, a well-characterized σ^{54} specific promoter,[38] contains a set of operator sites specific to NtrC located upstream of the promoter region. Transcription through the glnAp2 promoter works literally as an AND gate. Phosphorylation of the co-activator NtrC, also required for activation, is carried out either by the kinase/phosphatase NtrB (Nitrogen regulatory protein B) or by autophosphorylation with specific chemical substrates.[39]

To construct a synthetic AND gate, we cloned the σ^{54} and the *ntrC* genes under the same promoter P_{70} in separate plasmids and the *deGFP* gene under the glnAp2 promoter (Figure 2a). Tagging NtrC and σ^{54} with AAA+ specific degrons was not tested. The AND gate was studied with a slightly modified extract. For this transcriptional activation unit only, we noticed that the magnitude of the output signal was much higher when the crude extract was prepared with the S30 buffer A adjusted to pH 8.2 rather than 7.7. The activity of the other transcriptional activation units is decreased by only 20% with this extract. In the optimal conditions, the maximun deGFP production is largely above the reference concentration of 500 nM (Table 1). The output signal as a function of each input is similar (Figure 2b, Supplementary Figure S2c and S2d). The orthogonality of the glnAp2 promoter with respect to the primary σ^{70} was determined by measuring the synthesis of deGFP in the presence of one of the inputs. The protein production (0.04 μM) and the rate of protein synthesis (0.25 nM/min) decrease by a factor of 100 when only one input is used (Supplementary Figure S2c and S2d).

The reaction conditions and the buffer 3-PGA are sufficient for the phosphorylation of NtrC. The output signal is twice as great when the high-energy substrate carbamyl phosphate, specific for the phosphorylation of NtrC,[39] is added to the reaction (Supplementary Figure S2b). No increase of deGFP synthesis is observed when NtrB is also expressed in the reaction (data not shown).

Competition-Induced Transcription Regulation. *In vivo*, the competition between different sigma factors for the core RNA polymerase is responsible for passive control of promoter selectivity and gene regulation.[35,36] To determine the passive transcription regulations that result from such enzymatic contests, a competition assay was performed as follows: the expression of deGFP through each of the σ^{19}, σ^{24}, σ^{28}, σ^{32}, σ^{38}, and σ^{70} single transcriptional activation units was carried out with and without all of the other sigma factors in the reaction. The assay was performed in the linear and the saturation regime of plasmid concentrations (Figure 3a and b), using the σ^{70} salt conditions (Table 1).

The σ^{28}, σ^{32}, and σ^{70} units resist competition the most, a result in agreement with *in vitro* measurements of binding affinity between individual sigma factors and the core RNA polymerase.[34] The σ^{28} unit is mostly insensitive to competition, even at high plasmid concentration. A 10-fold increase of sensitivity is observed for σ^{19} between the two regimes of plasmid concentrations. Cell-free expression with σ^{24} or σ^{38} is strongly inhibited by the presence of the other sigma factors in both regimes.

To better characterize passive transcription regulation by competition, we tested two subsets of sigma factors. First, we studied gene expression with σ^{28} and σ^{38}, the most distant sigma subunits in competition sensitivity. Whereas deGFP synthesis through the promoter P_{28} is insensitive to the presence of σ^{38} at low and high plasmid concentrations, gene expression through the promoter P_{38} is decreased by a factor of 20 when σ^{28} is expressed (Figure 3c and d). The magnitude of the passive repression induced by σ^{28} was also measured for σ^{19}. At low plasmid concentration, both units are insensitive to competition (Supplementary Figure S3a). At high plasmid concentration, gene expression through the promoter P_{19} is decreased by a factor of 4 when σ^{28} is expressed (Supplementary Figure S3b). The passive transcriptional repression created by σ^{28} illustrates the additional gene regulations that emerge when transcriptional activation units are used in the same circuit (Figure 3e). In this particular case, σ^{28} can be used to down regulate the upstream part of a circuit, which increases the efficiency of the information flow and reduces the crosstalks. We shall see later that the sigma factor competition does not prevent the construction of parallel circuits.

A Five-Stage Transcriptional Activation Cascade: An Example of Series Circuit. Using the repertoire of transcriptional activation units and the preliminary circuit design rules, we next constructed a five-stage transcriptional activation cascade that creates a 1-h delay with a specific and biologically relevant end-point output signal (i.e., with a magnitude of at least 500 nM). Our objective was to understand how the crosstalks alter the signal's specificity for a circuit composed of a large number of activation units placed in series. The

dx.doi.org/10.1021/sb200016s | *ACS Synth. Biol.* 2012, 1, 29–41

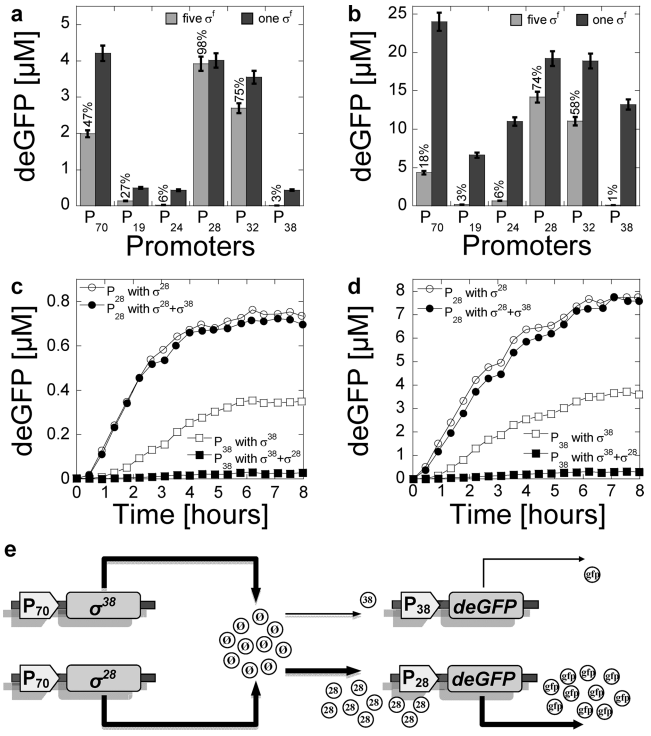


Figure 3. Passive transcription regulation by competition between sigma factors. (a) End-point measurements of deGFP synthesis with one or five sigma factors expressed in the reactions (linear regime of plasmid concentration: 0.1 nM of plasmid encoding each of the sigma factors and 1 nM reporter plasmid, σ^{70} salt conditions, see Table 1). In the first set of reactions, deGFP was synthesized through each individual sigma transcriptional activation cascade (one σ^f). In the second set of reactions, σ^{19}, σ^{24}, σ^{28}, σ^{32}, and σ^{38} cloned under the same promoter P_{70} were expressed simultaneously, whereas only one reporter plasmid was added to the reaction (five σ^f). The percentage indicates the level of expression compared to the one σ^f case. (b) The same experiment as in panel a in the saturation regime of plasmid concentration (0.5 nM plasmid encoding each of the sigma factors, 5 nM reporter plasmid). (c) Passive transcription regulation by competition with σ^{28} and σ^{38} (linear regime of plasmid concentration, each plasmid at 0.2 nM). The expression of deGFP through the σ^{28} transcriptional activation unit is not sensitive to the co-expression of σ^{38} in the reaction. The expression of deGFP through the σ^{38} transcriptional activation unit decreases by a factor of 20 when σ^{28} is also expressed in the reaction. (d) The same experiment as in panel c in the saturation regime of plasmid concentration (1 nM plasmid encoding each of the sigma factors and 4 nM reporter plasmid). (e) Schematic of the competition between σ^{28} and σ^{38}. The two sigma factors, expressed from the same promoter P_{70}, compete for the free core RNA polymerase (symbol Ø). The free core RNA polymerase interacts predominantly with σ^{28} (symbol 28) rather than with σ^{38} (symbol 38). As a result of the enzymatic contest, the expression of deGFP (symbol gfp) through the promoter P_{38} is decreased by a factor of 20.

optimization of the output signal was also used to test the adjustable parameters of the toolbox.

We chose to order the units by strength from the weakest to the strongest to obtain a favorable cascading of the amplification factors (Figure 4a). The position of σ^{19} in the second stage just after σ^{38} takes advantage of the leak of σ^{19} through P_{38} to create a positive feedback loop. The position of σ^{28} in the third position creates a negative feedback loop, by competition-induced passive regulation, that turns off the second stage of the cascade (P_{38}) and partially inhibits

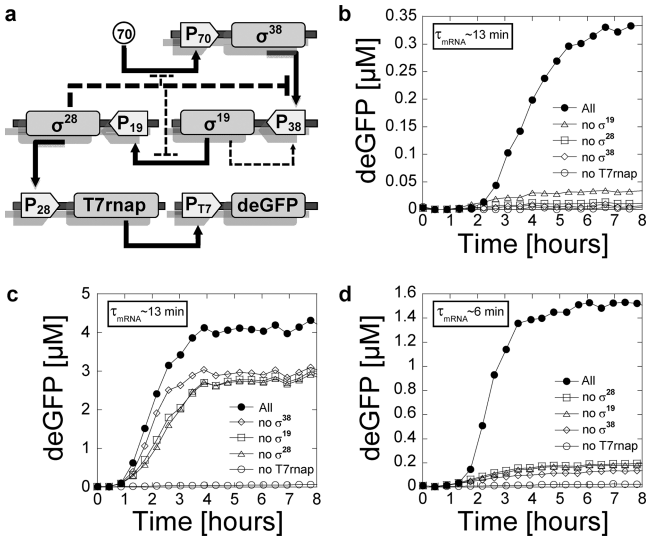


Figure 4. A five-stage transcriptional activation cascade. (a) Schematic of the cascade. Nonspecific interactions are shown as dotted lines (inhibition of transcription due to the competition of σ^{28} with the other sigma factors, leak of σ^{19} on P_{38}). (b) Kinetics of deGFP expression in the linear regime of plasmid concentration (each plasmid at 0.2 nM, σ^{70} salt conditions, see Table 1) with a global mRNA mean lifetime of 13 min. The specificity of the output signal is confirmed by the four negative controls, which consist of removing one of the four sigma factor stages (no σ^{19} means no plasmid P_{38}-σ^{19}). The output signal, however, barely reaches the relevant concentration fixed to 500 nM. (c) Kinetics of deGFP expression in the saturation regime of plasmid concentration (each plasmid at 1 nM) with a global mRNA mean lifetime of 13 min. The negative controls indicate that the output signal of the cascade is not specific. (d) Kinetics of deGFP expression in the saturation regime of plasmid concentration (each plasmid at 1 nM) with a global mRNA mean lifetime of 6 min. The specificity of the output signal is confirmed by the four negative controls, and the output signal is largely above the biologically relevant concentration of 500 nM.

transcription of the first and third stages (promoter P_{70} and P_{19}). This creates an autoregulation of the first part of the cascade, which improves both the transmission of the information along the series circuit and the bookkeeping of resources. σ^{28} turns off the first stages of the cascade, which decreases unnecessary protein synthesis along the cascade. Furthermore, because it is the most orthogonal and the most competitive sigma factor, σ^{28} attenuates the leaks from the other sigma factors. The T7 unit, used as the last amplification stage, is not sensitive to competition and it is the strongest amplifier. The reporter gene is placed in the last stage of the circuit to monitor the output signal. The T3, σ^{24}, σ^{32}, and AND gate units were not used in this example of circuit construction.

The concentration of each gene was first fixed to 0.2 nM to study the circuit in the linear regime of plasmid concentration (1 nM total plasmid concentration, σ^{70} salt conditions, Table 1). In these conditions, the output signal barely reaches the 500 nM reference level (Figure 4b) with a relatively low rate of deGFP production (~1.5 nM/min). The output signal, however, is specific with a signal to leak ratio of 10. An amplification factor of 10 was calculated as the ratio between the input signal (σ^{70}, 35 nM, Supplementary Table S5) and the output signal (deGFP, 350 nM). To get a greater amplification and faster response, the gain of each stage was increased by adjusting the concentration of each plasmid to 1 nM (saturation regime of plasmid concentration), which corre-

dx.doi.org/10.1021/sb200016s | ACS Synth. Biol. 2012, 1, 29–41

sponds to the concentration of genes in *E. coli*. Whereas a much larger output signal and a faster response are observed, the circuit loses its specificity (Figure 4c). Except for the T7 stage, large output signals are observed when one of the stages is removed from the series circuit. The crosstalks between units lead to a large amplification of nonspecific signals when the concentration of each plasmid is increased from 0.2 nM (linear regime of plasmid concentration) to 1 nM (saturation regime of plasmid concentration). The different steps of construction of the circuit at 1 nM of each plasmid show that the amplification of the leak is already significant when the circuit is composed of only two stages (Supplementary Figure S4a-d).

Rather than looking for a better range of plasmid concentration to optimize the output signal, we tested the behavior of the circuit as a function of the global mRNA turnover. The concentration of each plasmid was held at 1 nM. Using a method described previously,[27] the global mRNA mean lifetime was decreased from 13 to 6 min, a lifetime in the range of the mRNA half-life measured in *E. coli*.[40] In this case, the magnitude of the output signal (1.5 μM) is largely above the reference mark (Figure 4d). The output signal is specific, with a signal to leak ratio of 10. An amplification factor of 40 is measured. Furthermore, the rate of deGFP synthesis is increased by a factor of 10 (~15 nM/min), and the onset of reporter synthesis is observed after 1 h of incubation, rather than 2 h in the case of low plasmid concentration (0.2 nM) and slow mRNA turnover (13 min). The acceleration of the mRNA inactivation rate becomes critical at the five-stage level. With a smaller number of stages, both systems (1 nM plasmid and 6 min mean lifetime, 0.2 nM plasmid and 13 min mean lifetime) have specific output signals with similar magnitudes (Supplementary Figure S4e and S4f). This circuit demonstrates how critical the mRNA turnover is for the specificity of the signal. Nonspecific signals are filtered out only by a 2-fold acceleration of the global mRNA turnover.

Parallel Circuits. Gene circuits usually consist of series and parallel connections. The transcription regulation induced by competition between sigma factors, although useful for some circuit configurations (Figure 4a), can be also a serious limitation to design series or parallel circuits. To show that strong and weak sigma factors can be used simultaneously, multiple transcriptional activation units were carried out in parallel. First, the σ^{28} and the σ^{38} cascades were performed simultaneously (Figure 5a). σ^{28} was tagged with the YbaQ AAA + specific degron[27,30] to attenuate its domination. By simply adjusting the gene concentrations, the end point output signal of each cascade can be chosen over a wide range of magnitudes all above the 500 nM reference concentration (Figure 5b). The same observations were made with three transcriptional activation units placed in parallel (Figure 5c). When the σ^{19} cascade is added to the previous circuit, the output signals of each unit can be adjusted over a wide range of magnitudes (Figure 5d). Parameters can be chosen so as to get output signals with same magnitude in each branch of the parallel circuits. The freedom to adjust the gene concentrations, the global mRNA turnover, and the protein degradation rate allows constructing various series and parallel circuits without even tuning the strength of the other regulatory parts (promoter, untranslated region, ribosome binding site).

Transcriptional Repression Units. In gene regulation, active transcription repression is an essential counterpart to transcription activation. To show that transcriptional repression could be effectively implemented with our cell-free system, we

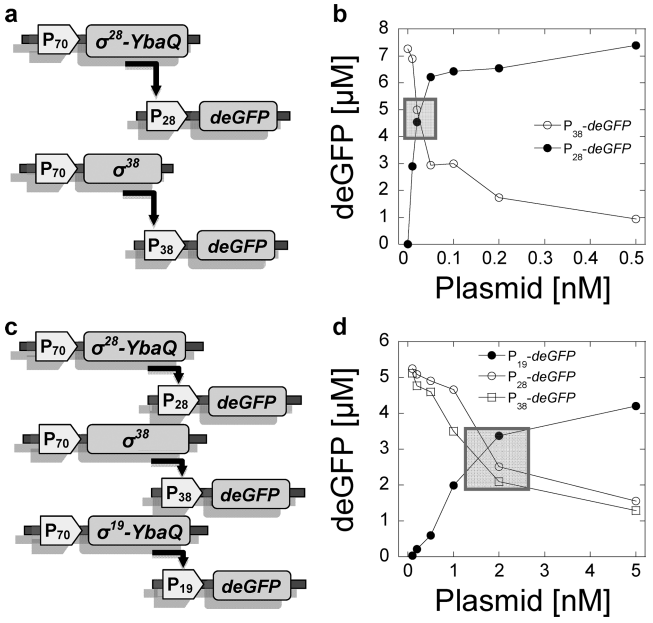


Figure 5. Parallel circuits using transcriptional activation cascades. (a) Schematic of two transcriptional activation units carried out in parallel. σ^{28} is tagged with the degron YbaQ (Supplementary Table S1). (b) End-point deGFP production as a function of the concentration of plasmid P_{70}-σ^{28}-*YbaQ* (0.5 nM P_{70}-σ^{38}, 10 nM P_{38}-*deGFP*, and 2 nM P_{28}-*deGFP*, σ^{70} salt conditions, see Table 1). The output signals of each cascade have the same magnitude for a P_{70}-σ^{28}-*YbaQ* plasmid concentration of about 0.025 nM (gray frame). (c) Schematic of three transcriptional activation units carried out in parallel. Both σ^{19} and σ^{28} are tagged with the YbaQ AAA+ specific degron. (d) End-point deGFP production as a function of the concentration of plasmid P_{70}-σ^{19}-*YbaQ* (0.5 nM P_{70}-σ^{38}, 10 nM P_{38}-*deGFP*, 0.025 nM P_{70}-σ^{28}-*YbaQ*, 2 nM P_{28}-*deGFP*, and 10 nM P_{19}-*deGFP*). The output signals of each cascade have the same magnitude for a P_{70}-σ^{19}-*YbaQ* plasmid concentration of about 2 nM (gray frame).

studied four σ^{70} transcriptional repression units (Figure 6a, Supplementary Figure S5).

First, we tested the lactose system with the synthetic regulatory element $P_{LlacO-1}$, composed of a strong promoter specific to σ^{70} and two lac operators.[41] The *lacI* repressor gene was cloned under the $P_{LlacO-1}$ element to make a negative feedback loop. The *deGFP* reporter gene was cloned under the same element in a separate plasmid. The repression of deGFP expression is observed after 30 min of incubation (Figure 6b). The reporter protein is fully expressed when 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) is added to the reaction. The range of IPTG concentration required to inhibit the repression is similar to the amount of IPTG used for *in vivo* induction (Supplementary Figure S5a). A much higher concentration of lactose, also comparable to *in vivo* experiments, is required to induce the expression of deGFP (Supplementary Figure S5b). At low plasmid concentration, however, these observations are slightly biased by the presence of endogenous lac repressor in the extract. A 50% repression is observed when only the reporter plasmid is used at a concentration of 0.5 nM (Supplementary Figure S5c). At higher plasmid concentration, the repression due to the presence of endogenous lac repressor protein in the system is negligible (Supplementary Figure S5d).

Next, we tested the arabinose inducible system with the plasmid pBAD, which contains the araBAD regulatory element

dx.doi.org/10.1021/sb200016s | *ACS Synth. Biol.* 2012, 1, 29–41

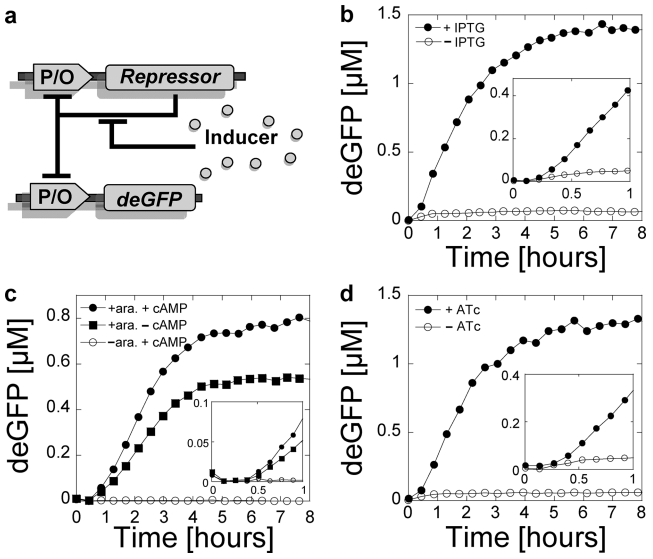


Figure 6. Lactose, arabinose, and tetracycline inducible transcriptional repressions. (a) Schematic of the circuit (P/O: promoter/operator). (b) The lactose system. The *E. coli lacI* repressor gene and the *deGFP* gene were cloned under the $P_{LlacO-1}$ regulatory element (1 nM $P_{LlacO-1}$-*lacI* and 0.5 nM $P_{LlacO-1}$-*deGFP*, σ^{70} salt conditions, see Table 1). The expression of deGFP is fully derepressed when 0.5 mM IPTG is added to the reaction. Inset: a blow-up of the first hour of expression. (c) The arabinose system. The *deGFP* gene was cloned under the araBAD promoter/operator into the plasmid pBAD bearing the arabinose repressor gene araC (5 nM pBAD-*deGFP*, σ^{70} salt conditions, see Table 1). The expression of deGFP is derepressed when a concentration of 10 mM arabinose is used in the reaction. A slight increase of deGFP synthesis is observed when cAMP is added to the reaction. Inset: a blow-up of the first hour of expression. (d) The tetracycline system. The *E. coli tetR* repressor gene and the *deGFP* gene were cloned under the $P_{LtetO-1}$ regulatory element (2 nM $P_{LtetO-1}$-*tetR* and 1 nM $P_{LtetO-1}$-*deGFP*, σ^{70} salt conditions, see Table 1). The expression of deGFP is derepressed when 10 μM ATc is added to the reaction. Inset: a blow-up of the first hour of expression.

and the *araC* repressor gene. Transcription through the araBAD element, a σ^{70} specific promoter with two operator sites, is repressed by the protein AraC in the absence of arabinose and activated by AraC in the presence of arabinose.[42] Transcription through the araBAD promoter is also stimulated by the cAMP catabolic activator protein Crp via a CAP operator site.[43] The synthesis of deGFP, cloned under the araBAD promoter, is repressed in the absence of arabinose, whether cAMP is present in the solution or not (Figure 6c). When a concentration of 10 mM of arabinose is used in the reaction (0.2% w/v), a typical concentration used for induction in *E. coli* cells, deGFP is fully expressed. A 2-fold maximum increase of gene expression is observed upon the addition of 0.75 mM of cAMP to the reaction (Supplementary Figure S5e), a stimulation smaller than the 5- to 6-fold increase observed *in vivo*.[43] The expression of the *crp* gene in the reaction (cloned under a promoter P_{70}) does not have any effect on the expression of deGFP (data not shown), presumably because the concentration of endogenous CRP protein present in the extract is high enough to co-activate the expression. Compared to the reference plasmid P_{70}-*deGFP*, the deGFP production in the open state (~1 μM) and the synthesis rate (~5 nM/min) are smaller. This is due to the promoter and to the native untranslated region, not as strong as the promoter P_{70} and the UTR1 untranslated region used in this study. The synthesis of deGFP in the repressed state is at or below the background level (0.1 < nM/min). Consequently, the efficiency of repression, determined by the ratio of the rate of protein synthesis in the open state and the repressed state, is larger than 1000 (detection limit). These observations are not biased by the presence of endogenous AraC repressor in the extract. The expression of deGFP is identical in the presence and in the absence of arabinose when the gene *araC* is knocked out of the plasmid construction (Supplementary Figure S5f).

Although not systematic, leftovers of repressor proteins in the extract can slightly bias the function of synthetic circuits at low plasmid concentrations. An approach to get around this limitation is to use repressors not present in the *E. coli* extract. The tetracycline system, constructed as a negative feedback loop, was tested with the synthetic regulatory element $P_{LtetO-1}$, composed of a strong promoter specific to σ^{70} and two tet operators.[41] The tetracycline repressor gene *tetR* and the *deGFP* gene were cloned under $P_{LtetO-1}$ in two separate plasmids. As expected, no repression is observed when only the reporter gene is used (Supplementary Figure S5g). The repression of deGFP expression is observed after 30 min of incubation (Figure 6d). The reporter protein is fully expressed when a concentration of 10 μM (5 μg/mL) of anhydrotetracycline (ATc) is used in the reaction. The high concentration range of ATc required for full expression (Supplementary Figure S5h), 50 times greater than the amount used for *in vivo* induction,[41] is due to the high level of TetR repressor expressed in the open state, on the order of a few micromolars. The rate of deGFP synthesis decreases 200-fold between the open and closed states. The concentration of TetR repressor monomers required for repression, estimated from the kinetics of deGFP expression to be between 10 and 50 nM, agrees with previous characterizations of the tetracycline operon (dimer dissociation constant $K_d \approx 10^{-7}$ to 10^{-8} M and repressor dimer-operator dissociation constant $K_d \approx 10^{-12}$ to 10^{-13} M[44,45]). Coliphages also provide a large number of repression units. The lambda repressor, for example, is fully active in our system. A 100-fold repression is also observed between the closed and the open states (Supplementary Figure S5i and S5j).

We used the lactose and the tetracycline repressors to make an inducible circuit with two outputs and four possible states (Figure 7a). Each branch of the circuit was first tested individually (Supplementary Figure S6a–d). The concentration of the regulatory part $P_{LlacO-1}$ was kept high so that the action of endogenous lac repressor is negligible. Without any inducer in the reaction both sides of the circuit are closed because both promoters have comparable strength and both repressions occur at comparable concentrations of repressor dimers (Figure 7b and c). As expected, addition of one inducer opens only one side of the circuit, both sides are open in the presence of IPTG (50 μM) and ATc (10 μM).

Long-Lived Cell-Free Expression. The availability of resources is an issue for batch mode cell-free reactions, especially for relatively slow processes requiring a significant amount of energy and building blocks like transcription and translation. Typically, batch mode cell-free expression reactions last 4 to 5 h in the best conditions. In batch mode, gene expression is independent from the resources only for a short period of time (~1–2 h for conventional systems and our system). The decrease of the energy charge, the degradation of some amino acids and the change of pH during protein synthesis rapidly alter the kinetics of gene expression.[21] The accumulation of waste products is also a concern for pure

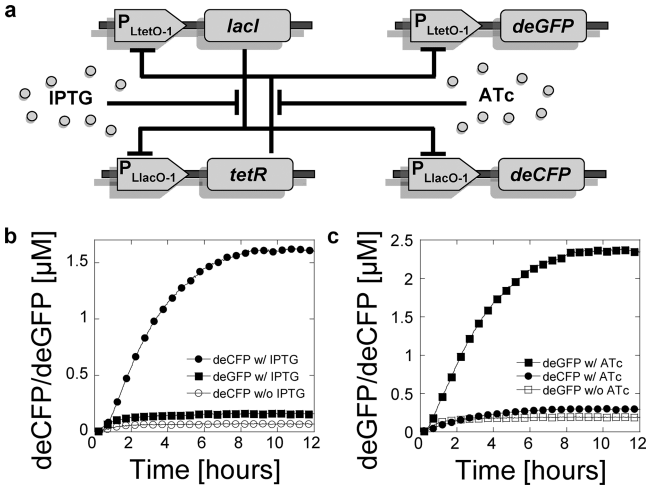


Figure 7. An inducible circuit constructed from two transcriptional repression cascades. (a) Schematic of the circuit composed of the lactose and tetracycline inducible repressions and the reporter genes *deCFP* and *deGFP*. The concentration of the four plasmids was fixed (1 nM $P_{LtetO-1}$-*lacI*, 2 nM $P_{LlacO-1}$-*tetR*, 4 nM $P_{LlacO-1}$-*deCFP*, and 2 nM $P_{LtetO-1}$-*deGFP*, σ^{70} salt conditions, see Table 1). deCFP channel: Ex 434, Em 485. deGFP channel: Ex 485, Em 528. (b) Kinetics of deCFP and deGFP expression in the closed state (0 μM IPTG) and the open state (50 μM IPTG). deCFP is expressed in the presence of IPTG and repressed without IPTG. deGFP is repressed by the TetR repressor. (c) Kinetics of deCFP and deGFP expression in the closed state (0 μM ATc) and the open state (10 μM ATc). deGFP is expressed in the presence of ATc and repressed without ATc. deCFP is repressed by the LacI repressor.

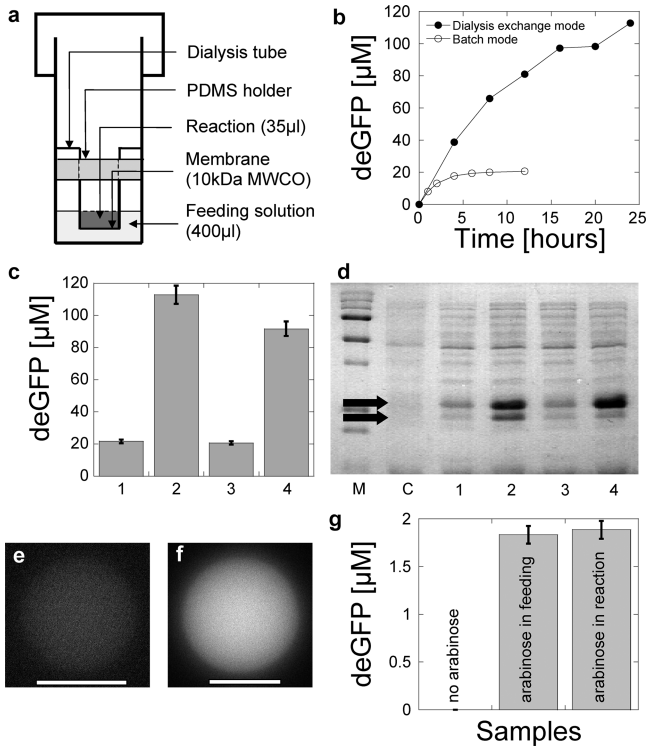


Figure 8. Long-lived cell-free reaction system and synthetic phospholipid vesicles. (a) Schematic of the exchange dialysis system. A dialysis tube (MWCO 10 kDa, 35 μL of cell-free reaction) is held in a larger tube (7 mL) containing 400 μL of feeding solution. (b) Kinetics of deGFP expression (5 nM P_{70}-*deGFP*) in batch mode and dialysis mode (using the σ^{70} salt conditions, see Table 1). (c) End-point deGFP production in batch mode (columns 1 and 3) and dialysis mode (columns 2 and 4). The expression of deGFP was carried out with the plasmid P_{70}-*deGFP* (columns 1 and 2) and through the two-stage cascade shown in Figure 1a (0.2 nM P_{70}-σ^{28} and 5 nM P_{28}-*deGFP*). (d) SDS-PAGE 13% of cell-free reactions. The top arrow indicates the deGFP band, and the bottom arrow indicates the ampicillin resistance protein band. M, marker; C, cell-free reaction with no plasmid; 1, expression of deGFP in batch mode (P_{70}-*deGFP*); 2, expression of deGFP in dialysis mode (P_{70}-*deGFP*); 3, expression of deGFP in batch mode (cascade, 0.2 nM P_{70}-σ^{28} and 5 nM P_{28}-*deGFP*); 4, expression of deGFP in dialysis mode (cascade, 0.2 nM P_{70}-σ^{28} and 5 nM P_{28}-*deGFP*). (e) Fluorescence microscopy image of deGFP expression using the arabinose system inside a phospholipid vesicle with no arabinose added to the reaction or to the feeding solution (5 nM pBAD-*deGFP*, scale bar: 10 μm). (f) Same as in panel e with 10 mM arabinose added to the feeding solution (5 nM pBAD-*deGFP*, scale bar: 10 μm). (g) End-point measurements of deGFP expression inside phospholipid vesicles (average fluorescence intensity of 20 large phospholipid vesicles). In the absence of arabinose, no deGFP can be measured, whereas 1.5–2 μM deGFP is produced inside the vesicles when 10 mM arabinose is added to the reaction or to the feeding solution.

synthetic biochemical systems.[7] The development of cell-free reactions stable over long periods of time is necessary to test and to model larger circuits.

Long-lived cell-free gene expression reactions that produce proteins in high yield have been already engineered with conventional hybrid systems.[46] Typical continuous-exchange bacteriophage systems are prepared for 1 mL of reaction and 10 mL of feeding solution. To test our endogenous system in continuous-exchange mode, we devised a dialysis reactor that works with 35 μL of reaction and 400 μL of feeding solution (Figure 8a). Two different reactions were tested. In the first reaction, deGFP is expressed through the promoter P_{70} (plasmid P_{70}-*deGFP*). We observe a time extension of cell-free expression by a factor of 4, from 4 h in batch mode to 16 h in dialysis mode (Figure 8b). In the second reaction, deGFP is expressed through the transcriptional activation cascade P_{70}-σ^{28} → P_{28}-*deGFP*. We also observe a time extension of cell-free expression by a factor of 4, from 4 h in batch mode to 16 h in dialysis mode (data not shown). In both cases, the amount of active deGFP produced is also increased by a factor of 5–6 (Figure 8c, 0.55 mg/mL in batch mode, 2.5 to 3 mg/mL in dialysis mode). In batch mode, a total concentration of 0.7 mg/mL deGFP (0.55 mg/mL active deGFP based on fluorescence) is measured on SDS-PAGE for both reactions (Figure 8d). In dialysis mode, a total concentration of 4 mg/mL deGFP is measured for both reactions. Our endogenous cell-free toolbox works as well as conventional hybrid systems in both batch and dialysis modes.

In the first reaction (plasmid P_{70}-*deGFP*), the ampicillin antibiotic resistance gene, present in the plasmid under a σ^{70} specific promoter, is also significantly expressed in dialysis mode (Figure 8d, sample 2). In the second reaction, the antibiotic resistance gene is almost not expressed due to the competition of σ^{28} with the housekeeping σ^{70}. σ^{28} regulates its own expression by passively repressing the promoter P_{70} through which it is transcribed. We could not confirm this autoregulation mechanism from the SDS-PAGE because the molecular mass of σ^{28} is comparable to the molecular mass of deGFP (25 kDa). We performed the same experiment with the gene *ntrB* encoding the protein NtrB of mass 32 kDa. As expected, the quantity of σ^{28} measured on gel is barely visible (Supplementary Figure S7).

dx.doi.org/10.1021/sb200016s | *ACS Synth. Biol.* 2012, 1, 29–41

The cell-free toolbox presented in this work is also devised to program synthetic phospholipid vesicles. Cell-free gene expression carried out in cell-sized liposomes is a bottom-up approach to constructing an artificial cell that focuses on the informational and self-organization properties of living matters. This idea was demonstrated experimentally[22,24] and discussed recently.[6,47,48] We first verified that the cell-free reactions prepared with the toolbox could be encapsulated in vesicles using the method developed for commercial hybrid systems.[22] A heterogeneous solution of single and aggregates of liposomes with a diameter of 1 μm to a few tens of micrometers is directly produced in a feeding solution (Supplementary Figure S8a). The expression of deGFP inside the liposomes is either as high as in a test tube or slightly greater due to the permeability of the phospholipid membrane.[22] The expression of the fusion protein αHemolysin-eGFP allows extending gene expression inside the vesicles by creating a selective permeability of the membrane through the channel of molecular mass cutoff 2−3 kDa (Supplementary Figure S8b).[22] In addition, we tested the arabinose system, the tightest repression tested in this work. A cell-free reaction containing the plasmid pBAD-*deGFP* was encapsulated inside vesicles. With no arabinose added to the feeding solution or to the reaction, the repression of deGFP expression is as efficient as when the reaction is carried out in a test tube (Figure 8e and f). The expression of deGFP inside the liposomes is identical whether 10 mM arabinose is added to the feeding solution or to the reaction (Figure 8g). Expression of deGFP is a factor of 2 higher than in the test tube due to the feeding of the reaction through the membrane. Adding arabinose only to the feeding solution is enough to completely open the system. The vesicle system adds to the range of applications of the toolbox by providing a cell-like environment to study self-organization processes via the expression of synthetic gene circuits.

Conclusion. The study of complex biochemical systems *in vitro* relies on the development of constructive and quantitative methods. The poor transcription repertoire of conventional cell-free systems has been a major restriction to construct synthetic systems based on the transcription and translation reactions. In this work, we have solved this limitation by developing the most versatile cell-free gene expression toolbox so far reported. This integrated platform provides a repertoire of transcriptional activation and repression units to design synthetic gene circuits. Compared to *in vivo*, cell-free transcription and translation are much less efficient. This does not prevent the development of synthetic circuits with relevant input-output responses, as demonstrated in this work. In addition, important biochemical parameters, often inaccessible *in vivo*, can be adjusted over a wide range using the toolbox. The possibility to adjust the global mRNA degradation rate and the concentration of each part of a circuit, for example, allows investigating the behaviors of gene circuits in a larger parameter space. The optimization of gene circuits using an *in vitro* approach is also potentially much faster than *in vivo*.

The basic working rules of the toolbox are described through a step-by-step construction of elementary gene circuits. The system, however, has not been exploited to its full potentialities. The construction and the characterization of different circuit motifs could certainly reveal a wealth of other dynamical behaviors. For instance, engineering a circuit to get a stable steady state for a given protein concentration would be a clear step toward real dynamical systems.

The platform presented in this article is an open transcription/translation breadboard that will be further developed as more complex circuits are constructed. Increasing the system's repertoire is a matter of testing new DNA parts, such as promoters/operators elements with their respective regulatory proteins.[49] The toolbox is also developed to express, in the long term, minimal genome size programs.[48] The demonstration that long-lived cell-free expression works with this system, in test tubes and in phospholipid vesicles, is a first step toward this goal. The approach one gene-one plasmid used in this work is adequate for circuits composed of a few tens of genes at most. New cloning techniques have to be used to assemble larger circuits. The ability to construct functional circuits involving more than half a dozen genes with the cell-free toolbox described herein is not known. Expressing existing large DNA programs encoding well-described functions and structures, such as coliphages, is a possible approach to scale up the system's capacity. The phospholipid vesicle system is a cell-sized bioreactor to test DNA programs encoding complex self-organized structures at the membrane, such as cell division. The development of an active phospholipid membrane, however, is currently the most serious limitation of the artificial cell system.

Usually considered as a poorly understood technique, cell-free gene expression is becoming more characterized. Kinetics constants, concentrations of the transcription and translation machineries have been determined in recent studies and a model of protein synthesis and degradation has been proposed.[3,28,50] Together with the toolbox presented in this work, this information provides the ingredients to develop mathematical models of cell-free transcription/translation circuits, our current ongoing effort.

METHODS

Crude Extract Preparation and Batch Mode Cell-Free Reaction. The cell-free system used in this work, prepared in our laboratory using a procedure described earlier,[26] is entirely endogenous. mRNA and protein synthesis are performed by the molecular machineries present in the extract, with no addition of external enzymes. Transcription is carried out by the *E. coli* core RNA polymerase and the primary σ^{70} with an estimated concentration in cell-free reactions of 100 nM and 35 nM, respectively (Supplementary Table S5). All gene circuits are booted up with promoters specific to σ^{70}, used as the unique housekeeping transcription factor. Protein synthesis is as efficient as conventional bacteriophage systems, with a batch mode end-point reporter protein production of 0.5 to 1 mg/mL (measured for deGFP and the *firefly* Luciferase[26]). The amount of protein synthesized and the kinetics of expression are reproducible from batch to batch within a 5% error bar.

Cell-free reactions were carried out in a volume of 5−10 μL at 29 °C. The 3-PGA reaction buffer is composed of[26] 50 mM Hepes pH 8, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM coenzyme A, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-phosphoglyceric acid, 1.5 mM each of 20 amino acids, 1 mM DTT, 2% PEG8000. A typical cell-free reaction with our system is composed of 33% (by volume) of *E. coli* crude extract, which corresponds to a protein concentration of 10 mg/mL, an optimum concentration typically used for conventional hybrid systems. The other 66% of the reaction volume are composed of the plasmids and the reaction buffer containing the nutrients. In the experiments, we fix the concentrations of all of the reagents contained in the reaction buffer except for the

magnesium glutamate and the potassium glutamate, two essential ions for transcription/translations reactions and molecular interactions involved in gene circuits. The cell-free expression system is prepared so as to adjust the concentrations of these two ions for any reaction. Cell-free protein synthesis with our system is oxygen dependent. At a scale of 5−10 μL, the oxygen transfer is not a limiting factor, and the reactions do not need to be stirred or shaken.

The preparation of the MazF *E. coli* S30 crude extract, used to accelerate the global mRNA inactivation rate, was described previously.[27] All the crude extracts used in this work are stable for at least 1 year at −80 °C.

DNA Part List and Plasmid Preparation. The DNA parts used in this work are reported in the Supplementary Table S1. Except as otherwise noted, the plasmids contain the highly efficient untranslated region named UTR1. The standard molecular cloning procedures were used to construct the plasmids. A construction noted P_n-σ^m (where *n* and *m* can be 19, 24, 28, 32, 38, 54, or 70) refers to an *E. coli* sigma factor gene *m* cloned under a promoter specific to an *E. coli* sigma factor *n*. This notation was also used for the T7 and T3 RNA polymerases and their respective promoters.

Long-Lived Cell-Free Reaction. The custom-built reactor was composed of a dialysis tube (Slide-A-Lyzer Mini Dialysis Unit, Thermo Scientific, MWCO 10 kDa), held by a PDMS (polydimethylsiloxane) ring into a mini scintillation vial (7 mL, Denville Scientific). The composition of the feeding solution (400 μL added at the bottom of the vial, no plasmid) was the same as the cell-free reaction (35 μL into the dialysis tube) except for the crude extract replaced by a same volume of S30 buffer B (5 mM Tris, 60 mM potassium glutamate, 14 mM magnesium glutamate, 1 mM DTT, pH 8.2). The system was incubated in a shaker incubator (29 °C, 120 rpm).

Phospholipid Vesicle Preparation. The encapsulation of cell-free reactions into large unilamellar phospholipid vesicles was described previously.[22] Briefly, the phospholipid solution was prepared by dissolving egg PC (Avanti Polar lipids) into mineral oil (Sigma-Aldrich) at 2 mg/mL. One microliter of cell-free reaction was added to the phospholipid solution. This solution was vortexed to create an emulsion, and 50 μL of the emulsion was placed on top of 50 μL of feeding solution, the same feeding solution used for the long-lived cell-free reactions. The vesicles were formed by centrifuging the biphasic solution for 20 s at 11,000 rpm in a benchtop centrifuge.

Measurements. Quantitative measurements were carried out with deGFP (1 mg/mL = 39.4 μM), a variant of the reporter eGFP.[26] The *deGFP* reporter gene is up to four times more translatable than eGFP *in vitro*, and the maturation time of the protein (7−8 min) is comparable to the maturation time of eGFP *in vivo*.[27] The excitation and emission spectra of deGFP and eGFP are similar (Supplementary Figure S9). The fluorescence of deGFP produced in batch mode cell-free reaction was measured with a Wallac Victor III plate reader (PerkinElmer, 384-well plate) and with an H1m plate reader (Biotek Instruments, 384-well plate). End-point measurements were carried out after 12 h of incubation. Pure recombinant eGFP (Clonetech) was used for the quantification of deGFP (linear calibration on plate reader). Polyacrylamide gel electrophoresis was carried out according to standard procedures. The gels were stained with SimplyBlue safestain (Invitrogen). The phospholipids vesicles were observed with a CCD camera (QImaging Retiga EXi FAST) mounted on a microscope (Olympus IX-71, inverted) with the proper set of fluorescence filters.

mRNA Turnover and Protein Degradation. The deGFP mRNA turnover rate can be accelerated from the endogenous level (mean lifetime ∼13 min) up to a complete inactivation (mean lifetime ∼0 min) using a method presented recently.[27] The *deGFP* mRNA turnover is used as a reference to set the global mRNA inactivation rate in a cell-free reaction. Specific degradation of the synthesized proteins is carried out by the endogenous AAA+ proteases present in the extract.[27] Degradation by the AAA+ proteases is efficient for proteins with a molecular mass smaller than 50 kDa approximately.[30,51] A quantitative model of deGFP expression and degradation with this system has been reported recently.[28]

Gene Circuits and Regime of Plasmid Concentrations. The behavior of synthetic cell-free gene circuits critically depends on the total concentration of plasmids used in the reaction and on the global mRNA degradation rate. A linear regime and a saturation regime of plasmid concentrations are observed for the output signal of elementary circuits.[1] In our system, with an endogenous mRNA mean lifetime of 13 min, the end-point output signal (total amount of active protein produced) of a circuit is linear with respect to the concentration of plasmid in each stage of the circuit below a total plasmid concentration of 1.5 nM. Above a total plasmid concentration of 1.5 nM, the response is not linear. It curves to saturation when the concentration of plasmid is increased in any stage of the circuit. This behavior is due to a physical saturation of the transcription and the translation machinery, which does not behave as an infinite reservoir. This saturation leads to a competition and a sharing of resources between the different genes expressed in the reaction. At high plasmid concentrations, this saturation results in a poor transmission of the information, as noticed previously,[1] unless the mRNA turnover is accelerated.

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +1-612-624-6589. Fax: +1-612-624-4578. Email: noireaux@umn.edu.

Author Contributions

J.S. and V.N. designed and performed the experiments, analyzed the data, and wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Albert Libchaber and Roy Bar-Ziv for comments; to Richard Murray, Jongmin Kim, and Anu Thubagere for useful suggestions; to Catherine Raach for reading and correcting the manuscript; and to Jerome Chalmeau for technical help. This work was supported by UMN startup funds and the National Science Foundation grant PHY-0750133. J.S. was supported by the NSF grant and by a fellowship from the Graduate School Block Grant funds (UMN).

REFERENCES

(1) Noireaux, V., Bar-Ziv, R., and Libchaber, A. (2003) Principles of cell-free genetic circuit assembly. *Proc. Natl. Acad. Sci. U.S.A. 100*, 12672−12677.

(2) Isaacs, F. J., Dwyer, D. J., and Collins, J. J. (2006) RNA synthetic biology. *Nat. Biotechnol. 24*, 545−554.

(3) Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J., and Swartz, J. R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol. 4*, 220.

(4) Hodgman, C. E., and Jewett, M. C. (2011) Cell-free synthetic biology: Thinking outside the cell, *Metab Eng.* E-pub ahead of print, DOI: 10.1016/j.ymben.2011.09.002.

(5) Zhang, D. Y., Turberfield, A. J., Yurke, B., and Winfree, E. (2007) Engineering entropy-driven reactions and networks catalyzed by DNA. *Science 318*, 1121−1125.

(6) Jewett, M. C., and Forster, A. C. (2010) Update on designing and building minimal cells. *Curr. Opin. Biotechnol. 21*, 697−703.

(7) Kim, J., and Winfree, E. (2011) Synthetic in vitro transcriptional oscillators. *Mol. Syst. Biol. 7*, 465.

(8) Montagne, K., Plasson, R., Sakai, Y., Fujii, T., and Rondelez, Y. (2011) Programming an in vitro DNA oscillator using a molecular networking strategy. *Mol. Syst. Biol. 7*, 476.

(9) Kim, J., White, K. S., and Winfree, E. (2006) Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Mol. Syst. Biol. 2*, 68.

(10) Seelig, G., Soloveichik, D., Zhang, D. Y., and Winfree, E. (2006) Enzyme-free nucleic acid logic circuits. *Science 314*, 1585−1588.

(11) Macdonald, J., Li, Y., Sutovic, M., Lederman, H., Pendri, K., Lu, W., Andrews, B. L., Stefanovic, D., and Stojanovic, M. N. (2006) Medium scale integration of molecular logic gates in an automaton. *Nano Lett. 6*, 2598−2603.

(12) Qian, L., and Winfree, E. (2011) A simple DNA gate motif for synthesizing large-scale circuits. *J. R. Soc., Interface 8*, 1281−1297.

(13) Qian, L., and Winfree, E. (2011) Scaling up digital circuit computation with DNA strand displacement cascades. *Science 332*, 1196−1201.

(14) Qian, L. L., Winfree, E., and Bruck, J. (2011) Neural network computation with DNA strand displacement cascades. *Nature 475*, 368−372.

(15) Soloveichik, D., Seelig, G., and Winfree, E. (2010) DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. U.S.A. 107*, 5393−5398.

(16) Rothemund, P. W. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature 440*, 297−302.

(17) Barish, R. D., Schulman, R., Rothemund, P. W., and Winfree, E. (2009) An information-bearing seed for nucleating algorithmic self-assembly. *Proc. Natl. Acad. Sci. U.S.A. 106*, 6054−6059.

(18) Venkataraman, S., Dirks, R. M., Rothemund, P. W., Winfree, E., and Pierce, N. A. (2007) An autonomous polymerization motor powered by DNA hybridization. *Nat. Nanotechnol. 2*, 490−494.

(19) Dittmer, W. U., Kempter, S., Radler, J. O., and Simmel, F. C. (2005) Using gene regulation to program DNA-based molecular devices. *Small 1*, 709−712.

(20) Katzen, F., Chang, G., and Kudlicki, W. (2005) The past, present and future of cell-free protein synthesis. *Trends Biotechnol. 23*, 150−156.

(21) Swartz, J. (2006) Developing cell-free biology for industrial applications. *J. Ind. Microbiol. Biotechnol. 33*, 476−485.

(22) Noireaux, V., and Libchaber, A. (2004) A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. U.S.A. 101*, 17669−17674.

(23) Isalan, M., Lemerle, C., and Serrano, L. (2005) Engineering gene networks to emulate Drosophila embryonic pattern formation. *PLoS Biol. 3*, e64.

(24) Ishikawa, K., Sato, K., Shima, Y., Urabe, I., and Yomo, T. (2004) Expression of a cascading genetic network within liposomes. *FEBS Lett. 576*, 387−390.

(25) Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components. *Nat. Biotechnol. 19*, 751−755.

(26) Shin, J., and Noireaux, V. (2010) Efficient cell-free expression with the endogenous *E. coli* RNA polymerase and sigma factor 70. *J. Biol. Eng. 4*, 8.

(27) Shin, J., and Noireaux, V. (2010) Study of messenger RNA inactivation and protein degradation in an *Escherichia coli* cell-free expression system. *J. Biol. Eng. 4*, 9.

(28) Karzbrun, E., Shin, J., Bar-Ziv, R. H., and Noireaux, V. (2011) Coarse-grained dynamics of protein synthesis in a cell-free system. *Phys. Rev. Lett. 106*, 048104.

(29) Gruber, T. M., and Gross, C. A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol. 57*, 441−466.

(30) Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., and Baker, T. A. (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell 11*, 671−683.

(31) Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J., and Frishman, D. (2008) Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics 9*, 102.

(32) Glaser, B. T., Bergendahl, V., Anthony, L. C., Olson, B., and Burgess, R. R. (2009) Studying the salt dependence of the binding of sigma70 and sigma32 to core RNA polymerase using luminescence resonance energy transfer. *PLoS One 4*, e6490.

(33) Levchenko, I., Seidel, M., Sauer, R. T., and Baker, T. A. (2000) A specificity-enhancing factor for the ClpXP degradation machine. *Science 289*, 2354−2356.

(34) Maeda, H., Fujita, N., and Ishihama, A. (2000) Competition among seven Escherichia coli sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res. 28*, 3497−3503.

(35) Farewell, A., Kvint, K., and Nystrom, T. (1998) Negative regulation by RpoS: a case of sigma factor competition. *Mol. Microbiol. 29*, 1039−1051.

(36) Ishihama, A. (2000) Functional modulation of Escherichia coli RNA polymerase. *Annu. Rev. Microbiol. 54*, 499−518.

(37) Ochs, M., Angerer, A., Enz, S., and Braun, V. (1996) Surface signaling in transcriptional regulation of the ferric citrate transport system of *Escherichia coli*: mutational analysis of the alternative sigma factor FecI supports its essential role in fec transport gene transcription. *Mol Gen Genet 250*, 455−465.

(38) Atkinson, M. R., Pattaramanon, N., and Ninfa, A. J. (2002) Governor of the glnAp2 promoter of *Escherichia coli*. *Mol. Microbiol. 46*, 1247−1257.

(39) Feng, J., Atkinson, M. R., McCleary, W., Stock, J. B., Wanner, B. L., and Ninfa, A. J. (1992) Role of phosphorylated metabolic intermediates in the regulation of glutamine synthetase synthesis in *Escherichia coli*. *J. Bacteriol. 174*, 6061−6070.

(40) Selinger, D. W., Saxena, R. M., Cheung, K. J., Church, G. M., and Rosenow, C. (2003) Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res. 13*, 216−223.

(41) Lutz, R., and Bujard, H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. *Nucleic Acids Res. 25*, 1203−1210.

(42) Lobell, R. B., and Schleif, R. F. (1990) DNA looping and unlooping by AraC protein. *Science 250*, 528−532.

(43) Lichenstein, H. S., Hamilton, E. P., and Lee, N. (1987) Repression and catabolite gene activation in the arabad operon. *J. Bacteriol. 169*, 811−823.

(44) Hillen, W., Gatz, C., Altschmied, L., Schollmeier, K., and Meier, I. (1983) Control of expression of the Tn10-encoded tetracycline resistance genes. Equilibrium and kinetic investigation of the regulatory reactions. *J. Mol. Biol. 169*, 707−721.

(45) Biliouris, K., Daoutidis, P., and Kaznessis, Y. N. (2011) Stochastic simulations of the tetracycline operon. *BMC Syst. Biol. 5*, 9.

(46) Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T., and Yokoyama, S. (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett. 442*, 15−19.
(47) Stano, P., and Luisi, P. L. (2010) Achievements and open questions in the self-reproduction of vesicles and synthetic minimal cells. *Chem. Commun. (Cambridge) 46*, 3639−3653.
(48) Noireaux, V., Maeda, Y. T., and Libchaber, A. (2011) Development of an artificial cell, from self-organization to computation and self-reproduction. *Proc. Natl. Acad. Sci. U.S.A. 108*, 3473−3480.
(49) Voigt, C. A. (2006) Genetic parts to program bacteria. *Curr. Opin. Biotechnol. 17*, 548−557.
(50) Underwood, K. A., Swartz, J. R., and Puglisi, J. D. (2005) Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnol. Bioeng. 91*, 425−435.
(51) Ortega, J., Singh, S. K., Ishikawa, T., Maurizi, M. R., and Steven, A. C. (2000) Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP. *Mol. Cell 6*, 1515−1521.