ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

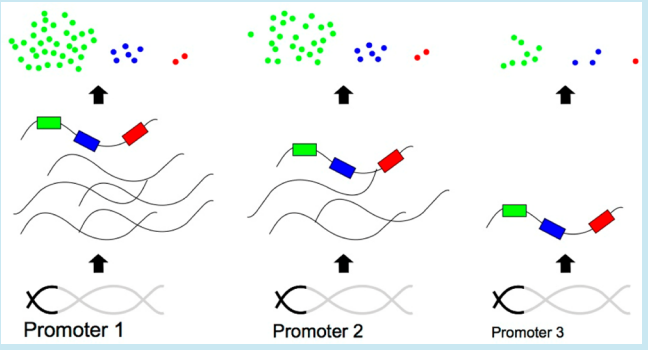
Gene Position More Strongly Influences Cell-Free Protein Expression from Operons than T7 Transcriptional Promoter Strength

Fabio Chizzolini, Michele Forlin, Dario Cecchi, and Sheref S. Mansy*

CIBIO, University of Trento, via delle Regole 101, 38123 Mattarello, Italy

Supporting Information

ABSTRACT: The cell-free transcription−translation of multiple proteins typically exploits genes placed behind strong transcriptional promoters that reside on separate pieces of DNA so that protein levels can be easily controlled by changing DNA template concentration. However, such systems are not amenable to the construction of artificial cells with a synthetic genome. Herein, we evaluated the activity of a series of T7 transcriptional promoters by monitoring the fluorescence arising from a genetically encoded Spinach aptamer. Subsequently the influences of transcriptional promoter strength on fluorescent protein synthesis from one, two, and three gene operons were assessed. It was found that transcriptional promoter strength was more effective at controlling RNA synthesis than protein synthesis *in vitro* with the PURE system. Conversely, the gene position within the operon strongly influenced protein synthesis but not RNA synthesis.



KEYWORDS: *cell-free, Spinach, transcription-translation, T7 promoter, metabolic load, PURE system*

Cell-free gene expression has served as a powerful platform from which to advance basic research and biotechnology. The triplet code for peptide synthesis was deciphered by using cell-free protein synthesis,[1] and difficult to express and toxic proteins are oftentimes produced with cell-free systems.[2] More recently, synthetic biologists have exploited cell-free systems for the optimization of genetic circuitry for later placement in living cells[3,4] and as a basis for the construction of artificial cells.[5−9] In terms of methodology, what most of these studies share is a reliance on multiple plasmids or linear DNA fragments encoding the biological parts needed for desired activity. Additionally, since protein yield is largely dependent upon template DNA concentration,[10] pathway optimization often follows the screening of relative DNA concentrations of each construct. However, these methodologies are not amenable to efforts in building artificial cells. First, it is difficult to efficiently encapsulate multiple pieces of DNA inside of a single vesicle. Since a minimal cell that is dependent upon protein activity is hypothesized to require 200 genes[11] or more, many genes will need to be on the same piece of DNA to support artificial cellular life. Second, the expression of gene clusters as operons, which requires that the genes are on the same piece of DNA, facilitates the construction of genetic cascades by decreasing the number of needed biological parts.

While recent work has helped to advance cell-free research, many questions remain. Computational modeling suggests that metabolic load plays an important role in protein synthesis,[12] which is consistent with cell-free optimization methods that exploit dialysis[13] or permeable vesicles.[14] More specifically, the supply of two components (ATP and GTP) that are involved in both transcription and translation was identified as the primary constraint on protein synthesis.[12] In other words, according to this model, increased transcription could lead to decreased protein production. Conversely, a combined modeling and experimental study by Stögbauer et al. observed a direct correlation between DNA template concentration and mRNA and protein production for monocistronic genetic constructs,[15] meaning that more mRNA resulted in more protein. Instead of metabolic load, ribosome inactivation appeared to be responsible for the premature cessation of protein synthesis.[15] Assuming that both metabolic load and ribosome stability are important determinants of gene expression, then perhaps the metabolic demands of expressing a single gene is not sufficient to exhaust the chemical resources before ribosome activity degrades. If true, then the cell-free expression of polycistronic operons may show signs of metabolic load on expression levels. It should be noted that although metabolic load is important for living cells,[16] the effects are not likely to be the same *in vitro* due to the lack of efficient regeneration systems in cell-free systems.[17]

Many artificial cell experiments exploit a minimal, fully defined transcription−translation mix consisting of T7 RNA polymerase and *E. coli* translation machinery called the PURE

Special Issue: Cell-Free Synthetic Biology

Received: July 31, 2013
Published: November 27, 2013

   dx.doi.org/10.1021/sb4000977 | *ACS Synth. Biol.* 2014, 3, 363−371

system[18] that efficiently produces protein *in vitro*[19,20] and inside of vesicles.[21,22] Although environmental sensing artificial cellular mimics have been constructed by exploiting the activity of the PURE system,[23,24] little is known about the influence of genetic organization on *in vitro* protein expression. Previously, we characterized the influences of bicistronic spacing and different fluorescent proteins on protein expression and detection, respectively.[25] In light of recent modeling and monocistronic gene expression work,[12,15] we sought to further characterize *in vitro* polycistronic expression. More specifically, the influence of T7 transcriptional promoter strength was evaluated in terms of mRNA and protein production of one-gene, two-gene, and three-gene operons. RNA levels were quantified through the incorporation of a Spinach aptamer domain[26] within the 3′ untranslated region (UTR) of the transcript. Simultaneous protein and RNA synthesis were monitored with genetic constructs encoding both Spinach and a fluorescent protein.

RESULTS AND DISCUSSION

Real-Time *In Vitro* Transcription Can Be Monitored with Spinach. The Spinach aptamer[26] is an RNA sequence that binds a small molecule mimic of the chromophore of green fluorescent protein. In aqueous solution the fluorescence of the small molecule 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) is low but upon binding to Spinach the quantum yield of DFHBI greatly increases. To determine if Spinach can be exploited for the real-time monitoring of transcription, the Spinach aptamer sequence was encoded within DNA behind a T7 transcriptional promoter and the reaction initiated by the addition of T7 RNA polymerase in the presence of DFHBI at 37 °C (Figure 1a). An increase of fluorescence was observed for approximately 2 h that subsequently leveled off. The transcription rate was 1.6 nucleotides/s, which was consistent with previously determined values.[15] Since different versions of the aptamer were described earlier,[26,27] including a truncated version (Spinach-min) and a full-length version (Spinach) that was used above, the truncated version was also examined for compatibility with *in vitro* transcription. For stability reasons, both sequences are often embedded within a tRNA scaffold,[28] and so, Spinach and Spinach-min were additionally tested within a tRNA scaffold. Such scaffolding sequences are thought to stabilize the three-dimensional fold of the RNA and to protect against nuclease activity *in vivo*. After 2 h of expression, both Spinach-min constructs failed to give detectable fluorescence under the conditions employed (Figure 1b), whereas Spinach embedded within a tRNA scaffold gave a fluorescence signal that was 39% ± 4 lower than Spinach without the scaffold. To confirm that the fluorescence signal was due to Spinach, an antisense oligonucleotide designed to disrupt the DFHBI binding pocket was added after 2 h of transcription. Fluorescence decreased in proportion to the oligonucleotide concentration, consistent with a linear relationship between Spinach fluorescence and RNA concentration (Supporting Information Figure S1a). Finally, the influence of DNA template concentration on the fluorescence signal was evaluated. The data fit a logarithmic function and reached a plateau after 30 nM of DNA template (Supporting Information Figure S1b).

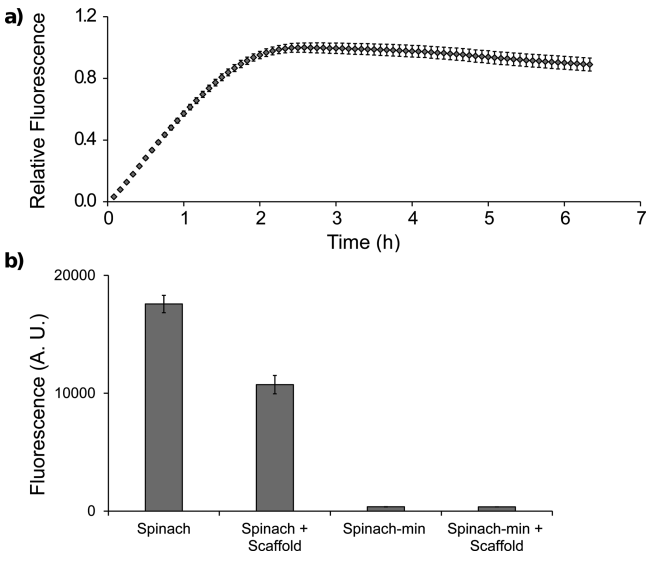


Figure 1. Real time *in vitro* transcription monitored through Spinach fluorescence. (a) Transcription with T7 RNA polymerase at 37 °C of a DNA template encoding full length Spinach without a tRNA scaffold. (b) Comparison of the fluorescence arising from transcription reactions of DNA templates encoding different versions of the Spinach aptamer after 2 h at 37 °C. Here, Scaffold indicates that the aptamer sequence is embedded within a tRNA sequence. More sequence information can be found in Supporting Information Table S1.

The *in vitro* expression of the Spinach aptamer was used to evaluate the effect of the DNA template sequence on RNA yield. It was desirable to use linear, PCR product DNA template rather than plasmid DNA so as to facilitate the construction and rapid testing of many constructs. Since the sequence context of the promoter can influence transcription, regions 5′ and 3′ to a standard T7 promoter were first evaluated in terms of their influences on RNA production. A standard T7 promoter sequence was used without additional 5′ residues and with two different 5′ extensions that were 5 bp in length (Supporting Information Figure S2a). The inclusion of either prepromoter sequence resulted in over 80% more RNA produced in comparison to the promoter lacking additional 5′ residues (Supporting Information Figure S2b). Four different 10 bp 5′ extensions were then evaluated. None of the sequences improved or diminished RNA yield when compared to the 5 bp extension sequences (Supporting Information Figure S2b). Therefore, a 5′ extension of CCGGT was used for the remainder of the experiments. Next, the sequences 3′ to the promoter were assessed by individually mutating each position from +1 to +6. The data were consistent with previous reports showing a strong dependence on the presence of guanosine residues for high RNA production (Supporting Information Figure S3). The region 3′ to the promoter was chosen to be GGGAGA for the remaining experiments.

A set of 21 T7 transcriptional promoters was generated by mutating promoter positions within constructs maintaining constant regions 5′ and 3′ to the promoter. Fifteen of the constructs were mutated at single positions, whereas double, triple, and quadruple mutants were represented by two, three, and one construct(s). The mutation positions were based on previous work[29] and sequences found naturally in T7 bacteriophage. Promoter activities were gauged by the total amount of RNA produced after 2 h of transcription at 37 °C. The set of 21 promoters showed activity from 5% to over 110% of the standard T7 promoter (Figure 2). The data were confirmed by qPCR (Supporting Information Figure S4), and

dx.doi.org/10.1021/sb4000977 | ACS Synth. Biol. 2014, 3, 363−371

a)

FC122	TAA**G**ACGACTCACTATA
FC109	**C**AA**CC**CGACTCACTATA
FC113	TAA**C**ACGACTCACTATA
FC111	TAAT**TGA**ACTCACTA**A**A
FC108	TAATACG**T**CTCACTATA
FC095	TAA**C**ACGACTCACTA**A**A
FC097	**C**AAT**C**CGACTCACTA**A**A
FC096	TAATACGACTC**T**CTATA
FC186	TAATACGACTCA**G**TATA
FC114	TAATACGACTCACTAT**T**
FC094	**C**TAT**T**CGACTCACTATA
FC115	TA**T**TACGACTCACTA**A**A
FC107	TAA**A**ACGACTCACTATA
FC098	TA**T**TACGACTCACTATA
FC091	TAATACGACTCACTA**A**A
FC089	**C**AATACGACTCACTATA
FC092	TAATA**T**GACTCACTATA
FC112	TAATAC**A**ACTCACTATA
FC093	**G**AATACGACTCACTATA
FC074	TAATACGACTCACTATA
FC090	**A**AATACGACTCACTATA

b)

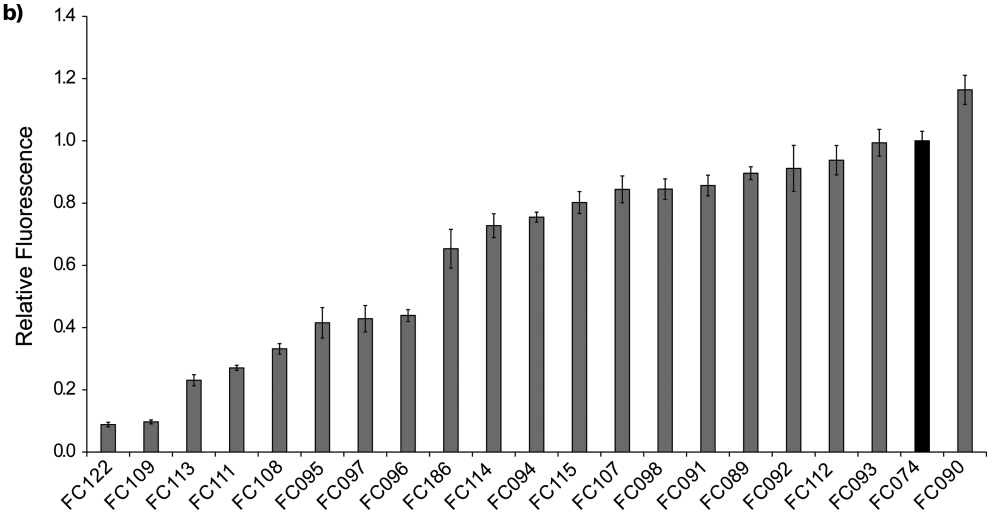


Figure 2. Library of T7 promoters characterized for *in vitro* transcription. (a) The promoter sequences are shown with residues in bold indicating positions different from the standard T7 promoter labeled here as FC074. Several of the promoter sequences are analogous to those found naturally within the T7 bacteriophage, including FC97, FC186, FC114, FC94, FC91, FC89, FC112, which correspond to natural T7 promoters phi4C, phi1.3, phi2.5, phi4.7, phi4.3, phi1.1A, and phiOL, respectively. All sequences additionally contained on the 5′-end CCGGT and the 3′-end GGGAGA followed by a sequencing encoding Spinach without a scaffold. (b) The fluorescence arising from transcription reactions after 2 h at 37 °C with each of the promoters listed in panel a is shown.

the quantification of band intensities of ethidium bromide stained agarose gels (Supporting Information Figure S5). Taken together, Spinach can be used to monitor transcription in real-time, and a set of T7 promoters was characterized for subsequent use in cell-free transcription-translation reactions.

Spinach and Fluorescent Proteins Allow for the Simultaneous Monitoring of *In Vitro* Transcription and Translation in Real-Time. A Spinach coding sequence was placed within the 3′-UTR of a gene encoding the red fluorescent protein (RFP) mRFP1 so that both RNA and protein production could be monitored in real-time by fluorescence spectroscopy. Although Spinach in the absence of a tRNA scaffold gave more intense fluorescence readings for transcription reactions alone, both constructs with and without the scaffold were tested for activity in transcription-translation reactions. In this case, the inclusion of the tRNA scaffold increased fluorescence over 10-fold with respect to the Spinach construct lacking the scaffold (Figure 3). Importantly, RNA and protein could be monitored in real-time simultaneously. As was expected,[15] RNA production was observed prior to protein synthesis, and the protein signal followed a sigmoidal distribution. The time needed to reach half maximal fluorescence[25] was 36 min ± 1 for RNA and 66 min ± 0.3 for protein synthesis. The measured translation rate of 0.03 amino acids/s was the same as previously reported for the PURE system.[15] PCR product DNA template lead to double the RNA and 12% ± 3 more protein than plasmid template when the same number of molecules of DNA were used (Supporting Information Figure S6). The DNA template concentration more strongly influenced the amount of RNA produced. A 4-fold difference in DNA concentration resulted in a 2-fold difference in final RNA concentration, whereas the same range of DNA template concentration resulted in less than a 25% difference in the final protein concentration (Supporting Information Figure S7).

T7 Transcriptional Promoter Strength Is More Effective in Controlling RNA Synthesis *In Vitro* than Protein Synthesis. DNA sequences encoding either RFP or

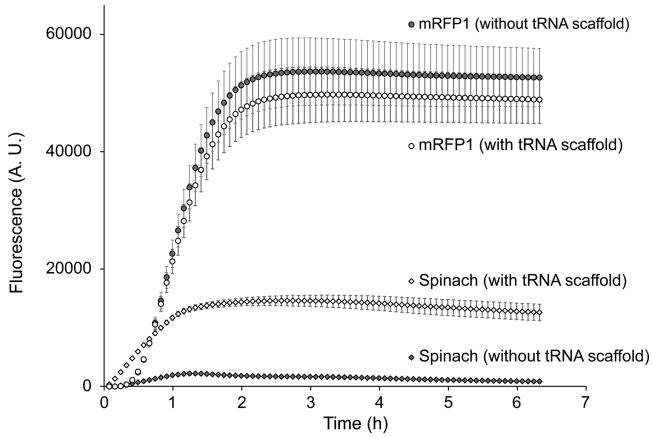


Figure 3. Real-time detection of transcription and translation by fluorescence spectroscopy. The DNA templates encoded the red fluorescent protein RFP. Additionally, either Spinach or Spinach with a tRNA scaffold was placed within the 3′-UTR of each sequence. Transcription−translation exploited the PURE system at 37 °C. Diamonds represent the signal from Spinach (bottom two curves), and the circles represent the signal from RFP (top two curves). Gray filled symbols are from constructs that contain unscaffolded Spinach and empty symbols are of data from Spinach within a tRNA scaffold.

the green fluorescent protein (GFP) GFPmut3b and Spinach within a tRNA scaffold were placed behind the set of T7 transcriptional promoters described above. For both RFP and GFP containing constructs, the fluorescence arising from Spinach, and thus the amount of RNA produced, gave consistent results in the sense that the relative promoter strengths were the same during transcription-translation as for transcription alone (Supporting Information Figure S8) and were the same for both RFP and GFP containing constructs (Figure 4). A plot of the concentration of RNA and protein produced for each promoter from RFP and GFP encoding templates showed a linear relationship (Supporting Information Figure S9) suggesting that the relative promoter strength was not greatly influenced by the folding or sequence composition

dx.doi.org/10.1021/sb4000977 | *ACS Synth. Biol.* 2014, 3, 363−371

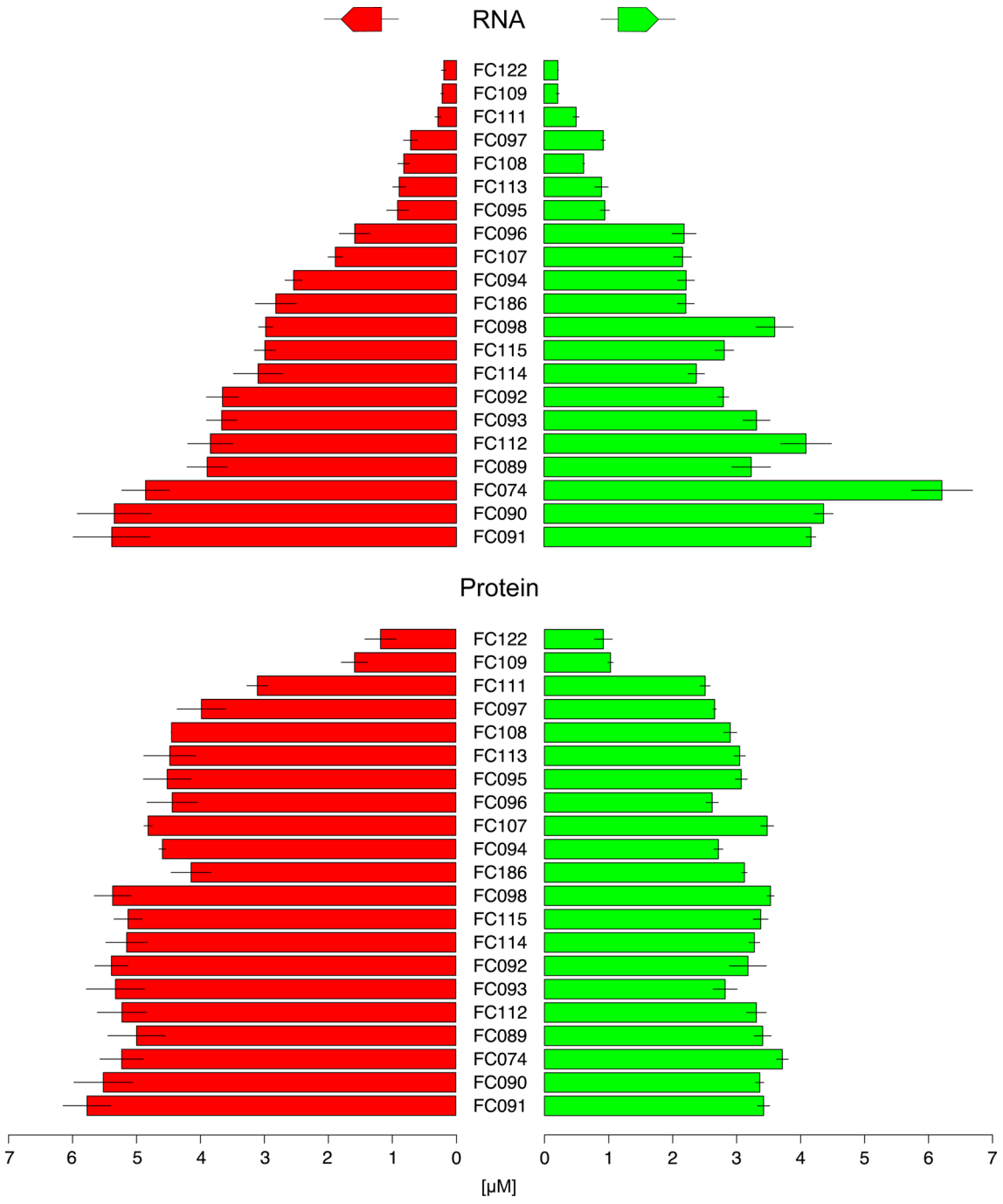


Figure 4. Influence of transcriptional promoter strength on RNA and protein synthesis from monocistronic constructs. Twenty-one PURE system reactions were incubated for 6 h at 37 °C with monocistronic DNA constructs encoding either RFP or GFP. Both constructs additionally contained Spinach with a tRNA scaffold in the 3′-UTR. RNA and protein concentrations are shown as red or green bars.

of the transcript. Although protein levels were also controlled by transcriptional promoter strength (Figure 4), the final protein concentrations were clustered between low and high without many transcriptional promoters resulting in intermediate protein concentrations. For example, weak transcriptional promoters gave final protein concentrations of approximately 1 μM, whereas the stronger transcriptional promoters typically resulted in 3 μM and greater than 4 μM GFP and RFP, respectively. Only protein capable of fluorescing was detected by this assay, and so, differences between GFP and RFP levels may reflect differences in the formation of the chromophore or protein folding in addition to the amount of protein expressed. Since both RNA and protein production were monitored simultaneously, the influence of RNA template concentration on protein synthesis was evaluated. Protein synthesis showed a much more pronounced dependence on RNA template concentration than RNA synthesis had on either DNA template concentration or transcriptional promoter strength (Supporting Information Figure S10). Therefore, small changes in the amount of RNA resulted in greater changes in the final protein output, particularly when RNA concentrations were low. Indeed, the ratio of protein to RNA concentration was 8 for weak transcriptional promoters, such as FC122 and FC109. As the transcriptional promoter strength

dx.doi.org/10.1021/sb4000977 | *ACS Synth. Biol.* 2014, 3, 363−371

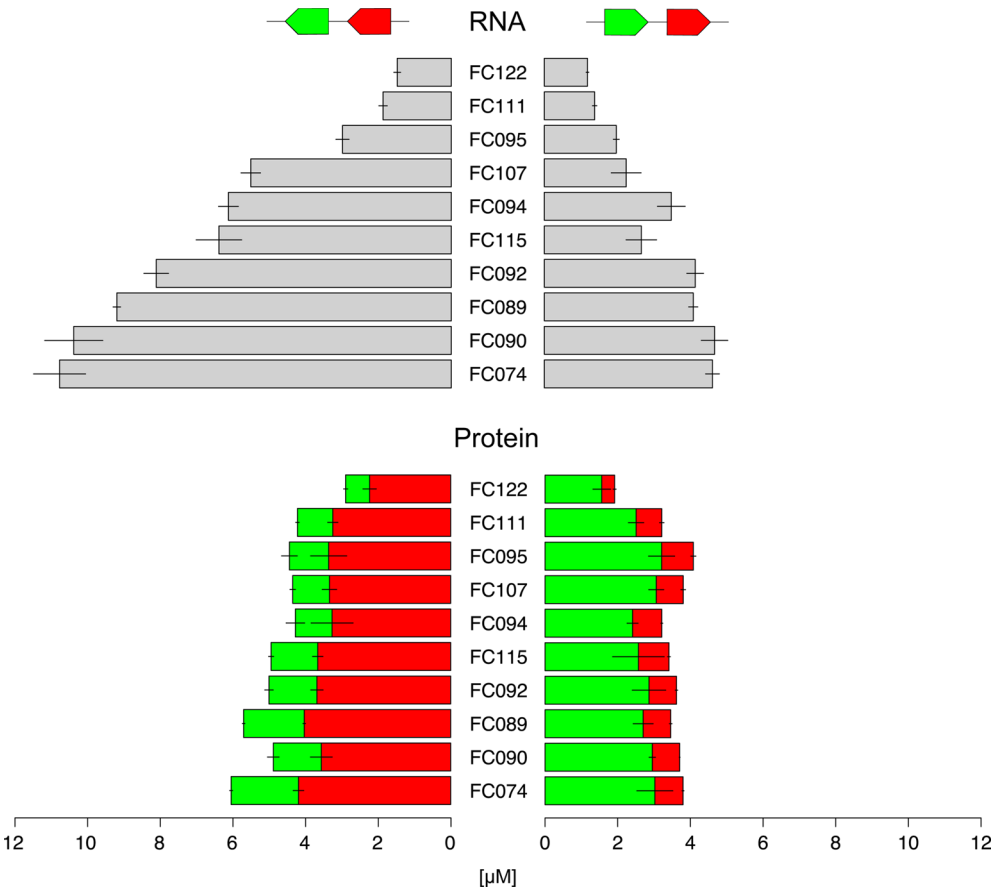


Figure 5. Bicistronic transcription−translation with the PURE system. The DNA templates encoded RFP and GFP with Spinach in a tRNA scaffold within the 3′-UTR. Reactions were incubated for 6 h at 37 °C. Although the RNA yield was less when GFP was in the first gene position, the relative strength of the promoters were largely constant.

increased, the protein to RNA concentration ratio decreased down to 1 for the strongest promoters tested. It should be noted that for all reactions, protein synthesis stopped before 3 h, consistent with previous reports with the PURE system.[15] The addition of more nucleotides or purified ribosomes did not extend the synthesis time (Supporting Information Figure S11).

To determine the influence of transcriptional promoter strength on two-gene operons, genetic constructs were assembled in which RFP was placed upstream of GFP and *vice versa*. All constructs contained Spinach within a tRNA scaffold at the 3′-end and 10 different transcriptional promoters at the 5′-end. Due to spectral overlap between Spinach and GFP, versions of both two-gene operons were additionally constructed with a mutated, nonfluorescent GFP in place of the fluorescently active GFP. With these four constructs, the protein and RNA concentrations could be determined. The amount of full-length transcript decreased by 45% ± 14 when GFP was in the first and RFP in the second gene position than the opposite arrangement, suggesting that the differing positioning of the genes affected the folding of the mRNA.[30] The relative promoter strengths were largely consistent regardless of the arrangement of genes or whether the construct encoded one or two genes (Figure 5). The two-gene operon data were also similar to the one-gene constructs in that the resulting protein levels showed a narrower distribution (the highest protein concentration was 2 to 3-fold greater than the lowest) as a result of transcriptional promoter strength than observed for RNA (the highest RNA concentration was 5 to 10-fold greater than the lowest).

Finally, the influence of transcriptional promoter strength on the expression of three-gene operons was evaluated. A three-gene construct was assembled with the first and second genes coding for nonfluorescent versions of GFP and the cyan fluorescent protein (CFP) mCerulean, respectively. The third gene encoded RFP and the 3′-UTR contained a scaffolded spinach sequence. Ten transcriptional promoters were tested. The relative RNA concentrations resulting from the *in vitro* transcription-translation reactions were consistent with the promoter strengths determined from the one-gene and two-gene constructs. Similarly, RFP expression from the third position of the three-gene operon gave a small range of protein concentration (here, between 0.1 μM and 0.3 μM). To ensure that the promoter strength continued to affect protein expression from the first two genes of the three-gene operon, an analogous three-gene operon was constructed with all three genes encoding functional fluorescent protein, and three different transcriptional promoters were tested. Promoter strength clearly influenced protein expression from all three gene positions (Figure 6).

Gene Position within Multigene Operons Strongly Affects Protein Expression. The most dramatic effect on protein expression observed did not stem from the influence of transcriptional promoter strength, but rather the position of the gene within the operon. For example, protein yield decreased 7-fold when monicistronic and bicistronic RFP constructs were

dx.doi.org/10.1021/sb4000977 | *ACS Synth. Biol.* 2014, 3, 363−371

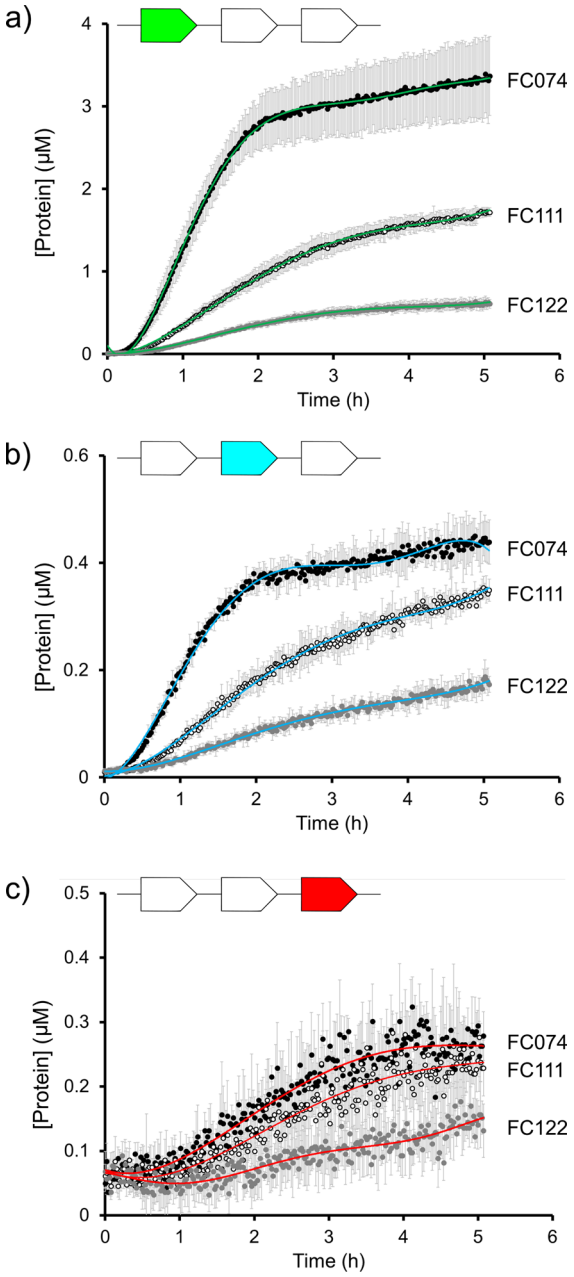


Figure 6. *In vitro* expression of a three gene operon with three different transcriptional promoters at 37 °C with the PURE system. The first, second, and third genes coded for GFP (a), CFP (b), and RFP (c), respectively.

compared (RFP was encoded by the position furthest from the 5′-end). The last position of the three-gene operon decreased RFP expression an additional 4-fold when compared with the two-gene operon (Supporting Information Figure S12), consist with previous reports on polycistronic expression.[31] The influence of transcript length on protein expression was not the result of variable mRNA concentration since the amount of RNA produced from each construct varied by less than 20% and tended to increase rather than decrease in concentration as the number of genes in the operon increased (Supporting Information Figure S12). Even in cases in which the transcript length was maintained, the switching of gene order within the bicistronic operon greatly influenced the amount of protein produced. RFP and GFP protein concentrations dropped 6-fold and 2-fold, respectively, when encoded by the second instead of the first gene of the operon (Figure 5). To gain further insight into the effect of gene position on expression levels, three-gene operons were constructed in which the encoded transcript length was maintained but the order of the genes was changed. The amount of each protein decreased as the gene coding for the protein moved from the first, to the second, and the third gene-position of the operon (Figure 7). On average, the decrease in protein concentration was 6-fold and 12-fold, respectively, in comparison to expression from the first gene position.

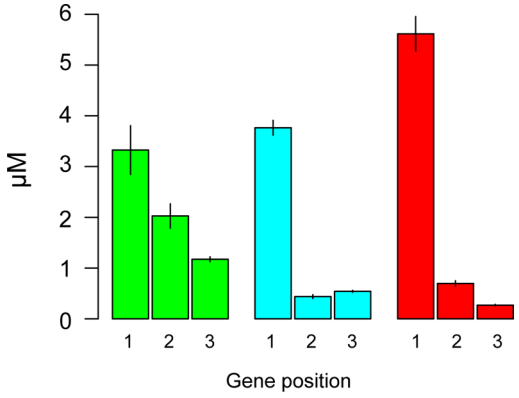


Figure 7. *In vitro* expression of three different three-gene operons with the PURE system. Genes coding for GFP, RFP, and CFP were placed in operons so that each gene could be evaluated in each of the three possible positions. The constructs encoded GFP-CFP-RFP, RFP-GFP-CFP, and CFP-RPF-GFP. Reactions were incubated for 6 h at 37 °C.

The reason for the differences in protein expression from each gene in the operon is not clear. The fact that protein expression generally decreases as the length of the transcript increases is consistent with the greater possibility of longer RNA molecules attaining more stable tertiary structures that interfere with translation. For example, RNA folding can occlude ribosome binding sites.[32] In our system, we observed that the expression of RFP from the first gene position of two-gene and three-gene operons gave 25% ± 9 and 7% ± 10, respectively, less protein than moncistronic expression. The data are in contrast to similar *in vivo* measurements in *E. coli* by Lim et al. where longer transcripts were associated with increased protein expression.[33] It may be that the increased accessibility of the 5′-end of mRNA simply facilitates recognition of the ribosome binding site by the ribosome resulting in preferential expression. It should be noted, however, that transcription and translation are not coupled in the PURE system, in the sense that RNA is produced much faster with T7 RNA polymerase than protein is synthesized with *E. coli* ribosomes. Therefore, the influences of transcriptional promoter strength with the PURE system are not expected to closely resemble the behavior observed *in vivo* or with *E. coli*-based extract systems.[3] The optimization of genetic devices for *in vivo* applications are likely better pursued with cell extract systems. Although a metabolic load effect was not seen when transcriptional promoter strength or DNA template concentration (Supporting Information Figure S13) was decreased, most of the constructs tested resulted in the synthesis of the same amount of total protein (approximately 6 μM).

dx.doi.org/10.1021/sb4000977 | *ACS Synth. Biol.* 2014, 3, 363−371

The simplicity of T7-based expression is attractive for the construction of cellular mimics. Weaker transcriptional promoters effectively reduce the amount of RNA made with the PURE system but do not greatly influence total protein output nor the relative expression of protein from each gene position of an operon. Conversely, the position of a gene within an operon strongly influences protein expression while having almost no effect on RNA concentration. Therefore, proteins needed at higher concentrations should be placed at the beginning of the operon, and protein levels are likely better tuned with ribosome binding sites than with transcriptional promoters. Nevertheless, the fact that halving the RNA concentration often has no effect on protein levels demonstrates how inefficient transcription-translation can be. Such inefficiencies may become more important as the complexity of *in vitro* genetic devices increases.

METHODS

Genetic Constructs. Genes encoding the fluorescent proteins GFPmut3b (BBa_E0040) and mRFP1 (BBa_E1010) were from the registry of standard biological parts (http://partsregistry.org). The monomeric version of the cyan fluorescent protein Cerulean (mCerulean) was from Lentini et al.[25] The different versions[26−28] of Spinach were synthesized by Genscript. Nonfluorescent versions of GFPmut3b and mCerulean (i.e. Y66F G67A GFPmut3b and W68F G69A mCerulean) were generated by site-directed mutagenesis with Phusion DNA polymerase (Finnzymes). All genes were subcloned into pET21b by isothermal Gibson assembly.[34] All constructs were confirmed by sequencing by Genechron. More details on the genetic constructs can be found in Supporting Information Table S1.

***In Vitro* Transcription.** Plasmids were amplified in *E. coli* DH5α and purified with the Wizard Plus SV Minipreps DNA Purification System (Promega). Subsequently, the DNA was phenol-chloroform extracted, ethanol precipitated, and resuspended in deionized and diethylpyrocarbonate (DEPC) treated water. Genetic constructs were amplified by PCR with Phusion DNA polymerase (Finnzymes) and purified from agarose gels with the Wizard Plus SV PCR cleanup system (Promega). Unless otherwise indicated, a 30 nM final concentration of DNA was used for the transcription reactions with 50 units T7 RNA polymerase (New England Biolabs), 2 mM of each nucleotide (New England Biolabs), 20 units of human placenta RNase Inhibitor (New England Biolabs), and 60 μM DHFBI (Lucerna). The final reaction volume was 50 μL. Reaction components were assembled on ice and then fluorescence kinetic data were acquired at 37 °C for 2 h with a Tecan Infinite M200 plate reader with excitation and emission wavelengths of 469 and 501 nm, respectively. Each reaction was repeated at least three times on three different days for a total of nine measurements.

RT-qPCR. An aliquot (3 μL) was taken from each transcription reaction and reverse transcribed with Superscript II Reverse Transcriptase (Invitrogen) following the protocol provided by the manufacturer. Each sample was then 100-fold diluted, and 1 μL of this 100-fold diluted sample was mixed with 5 μL SSo Advanced SYBR Green Supermix (Bio-Rad), 200 nM forward primer (5′-GACGCAACTGAATGAAATGGTG-3′), and 200 nM reverse primer (5′-GACGCGACTAGTTACGGAGC-3′). The volume was adjusted to 10 μL with deionized, sterile water. Real-time PCR was performed with a CFX96 Real-Time PCR machine (Bio-Rad) using an optimized annealing temperature of 59.4 °C. Reactions were performed in triplicate.

Cell-Free Transcription−Translation. Unless otherwise indicated, 5 μL transcription−translation reactions with the PURExpress *in vitro* protein synthesis kit (New England Biolabs) contained 12.6 nM DNA and 4 units of human placenta RNase inhibitor (New England Biolabs). When needed, DHFBI (Lucerna) was added to a final concentration of 60 μM. The reaction components were assembled on ice, and then, the reaction was initiated by incubation at 37 °C. Reactions were monitored for 6 h with a Bio-Rad CFX96 Real-Time system equipped with a C1000 Touch Thermal Cycler. Channel 1 was used to detect GFPmut3b and Spinach (excitation, 450−490 nm; emission, 515−530 nm), and channel 3 was used to detect mRFP1 (excitation, 560−590 nm; emission, 610−650 nm). Each reaction was repeated at least three times. To test whether transcription−translation could be prolonged, additional reaction components were added to some of the reaction solutions after 6 h incubation at 37 °C. When indicated, the following components were added to separate reactions in a manner that doubled the final reaction volume: 13 mM purified ribosomes (New England BioLabs), a full complement of solution A of the PURE system, a full complement of solution A of the PURE system supplemented with 13 mM purified ribosomes, a full complement of solution B of the PURE system, a full complement of both solutions A and B of the PURE system, 50 U of T7 RNA polymerase (New England BioLabs), 12.6 nM DNA template, the full PURE system supplemented with 12.6 nM DNA template, 0.3 U of inorganic pyrophosphatase (New England BioLabs), and water.

Transcription−translation reactions were also monitored with a Photon Technology International (PTI) QuantaMaster 40 UV−Vis spectrofluorometer equipped with two detectors. The reaction conditions were the same as described above except that 5.8 nM DNA and 20 units of human placenta RNase inhibitor (New England BioLabs) were used in a final reaction volume of 25 μL. The excitation and emission wavelengths were 435 and 478 nm, 506 and 514 nm, and 587 and 608 nm for mCerulean, GFPmut3b, and mRFP1, respectively.

Proteins and RNA Standard Curves. His-Tagged versions of GFPmut3b, mCerulean, and mRFP1 were generated by mutating the stop codon of the constructs CD100A, RL005A, RL008A[25] by phusion site-directed mutagenesis (Finnzymes). The resulting constructs coded for GFPmut3b, mCerulean, or mRFP1 with 24 additional residues that included a carboxy-terminal hexahistidine-tag. *E. coli* BL21(DE3) pLysS (Promega) were transformed with each His-tagged construct and grown in LB supplemented with 50 μg/mL ampicillin and 100 μg/mL chloramphenicol at 37 °C to an optical density of 0.5 at 600 nm before induction with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were harvested 4 h after the addition of IPTG by centrifugation at 5000 rpm for 10 min with a Beckman Coulter Avanti J-E centrifuge with a JLA 9100 rotor. The cell pellets were resuspended in 40 mL buffer R (50 mM NaH_2PO_4, 300 mM NaCl, 5 mM imidazole, pH 8), supplemented with 100 μL protease inhibitor cocktail (Sigma), and sonicated on ice (4 cycles, 10 s each cycle with 1 min cooling on ice between cycles) with a Branson Sonifier 450. Lysed cells were centrifuged at 15 000 rpm for 30 min at 4 °C with a Thermo Scientific Legend X1R centrifuge with a Fiberlite F15-8 × 50 cy rotor. The cleared lysate was loaded on a Ni-NTA column (Qiagen) and successively washed with

dx.doi.org/10.1021/sb4000977 | *ACS Synth. Biol.* 2014, 3, 363−371

buffer R and buffer R supplemented with 20 mM imidazole. Bound protein was eluted with buffer R plus 250 mM imidazole. Eluted protein was dialyzed against 20 mM Tris-HCl, 150 mM NaCl, 5 mM 2-mercaptoethanol, pH 8. Protein concentrations were determined from the extinction coefficients of GFPmut3b ($\varepsilon^{501\ nm}$ = 21,890 M^{-1} cm^{-1}),[35] mCerulean ($\varepsilon^{433\ nm}$ = 43 000 M^{-1} cm^{-1}),[35] and mRFP ($\varepsilon^{584\ nm}$ = 44 000 M^{-1} cm^{-1})[36] with an Agilent 8453 UV−vis.

Spinach mRNA was purified by using acidified phenol extraction followed by an ethanol precipitation. Briefly, transcription reactions were performed as described above under *in vitro transcription* except that the reactions were scaled up to 750 μL. Four different templates were used in four different reactions, including FC013A (monocistronic construct), FC019A (two-gene operon, mRFP1 last), FC022A (two-gene operon, mRFP1 first), and FC023A (three-gene operon, mRFP1 last). Five units of DNase I (New England BioLabs) were added to each sample for 1 h at 37 °C. The samples were then extracted with 5:1 phenol:chloroform (Sigma). The upper aqueous phase was subsequently extracted with 24:1 chloroform:isoamyl alcohol (Sigma) and ethanol precipitated.[37] RNA samples were resuspended in 0.1 mM EDTA and mixed with 1 volume of 2× RNA Loading Dye (Thermo Scientific). Samples were loaded on a 1% agarose gel and compared against a lane containing RiboRuler High Range RNA Ladder (Thermo Scientific) for quantification with ImageJ.[38] Varying amounts of Spinach encoding RNA were then incubated with 60 μM DHFBI to construct a standard curve that relates fluorescence to RNA concentration.

Statistical Analysis. All statistical analyses used R statistical computing software.[39] The RNA transcription rate was identified by fitting the Spinach fluorescence data to

$$\mathrm{RNA}(t) = a - b \cdot \mathrm{e}^{-ct} \tag{1}$$

where c represents the transcription rate. Fitting the protein kinetic data to a logistic model instead identified the translation rate:

$$\mathrm{PROT}(t) = \frac{K}{1 + \mathrm{e}^{-Ct-B}} \tag{2}$$

where C represents the translation rate. In both cases, the parameters were estimated by using a nonlinear least-squares analysis with the Gauss−Newton algorithm. Linear and logarithmic models were estimated using least-squares estimators. The significance of the estimated models (p-values) were based on F-test statistics.

■ ASSOCIATED CONTENT

Supporting Information

Table S1 and Figures S1−S13, as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +39 0461 28 3438. Fax: +39 0461-283091. E-mail: mansy@science.unitn.it.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Armenise-Harvard Foundation, the autonomous province of Trento (Ecomm), and CIBIO for funding.

■ REFERENCES

(1) Khorana, H. G. (1965) Polynucleotide synthesis and the genetic code. *Fed. Proc. 24*, 1473−1487.

(2) Spirin, A. S., and Swartz, J. R. *Cell-free Protein Synthesis*, 1st ed.; John Wiley & Sons, Weinheim, 2008.

(3) Chappell, J., Jensen, K., and Freemont, P. S. (2013) Validation of an entirely *in vitro* approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res. 41*, 3471−3481.

(4) Zoltan, A. T., Singhal, V., Kim, J., and Murray, R. M. An *in silico* modeling toolbox for rapid prototyping of circuits in a biomolecular "Breadboard" system. *Conference on Decision and Control (CDC)*, 2013.

(5) Stano, P., and Luisi, P. L. (2013) Semi-synthetic minimal cells: Origin and recent developments. *Curr. Opin. Biotechnol. 24*, 633−638.

(6) 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.

(7) Forlin, M., Lentini, R., and Mansy, S. S. (2012) Cellular imitations. *Curr. Opin. Chem. Biol. 16*, 586−592.

(8) Forster, A. C., and Church, G. M. (2006) Towards synthesis of a minimal cell. *Mol. Syst. Biol. 2*, 45.

(9) Fritz, B. R., Timmerman, L. E., Daringer, N. M., Leonard, J. N., and Jewett, M. C. (2010) Biology by design: From top to bottom and back. *J. Biomed. Biotechnol. 2010*, 232016.

(10) 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.

(11) Gil, R., Silva, F. J., Pereto, J., and Moya, A. (2004) Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev. 68*, 518−537.

(12) Calviello, L., Stano, P., Mavelli, F., Luisi, P. L., and Marangoni, R. (2013) Quasi-cellular systems: Stochastic simulation analysis at nanoscale range. *BMC Bioinformatics 14*, S7.

(13) Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y., and Alakhov, Y. B. (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science 242*, 1162−1164.

(14) 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.

(15) Stogbauer, T., Windhager, L., Zimmer, R., and Radler, J. O. (2012) Experiment and mathematical modeling of gene expression dynamics in a cell-free system. *Integr. Biol. (Camb) 4*, 494−501.

(16) Bentley, W. E., Mirjalili, N., Andersen, D. C., Davis, R. H., and Kompala, D. S. (2009) Plasmid-encoded protein: The principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnology Bioengineering, 1990. *Biotechnol. Bioeng. 102*, 1284−1297 discussion 1283.

(17) Jewett, M. C., and Swartz, J. R. (2004) Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnol. Bioeng. 87*, 465−472.

(18) 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.

(19) Kita, H., Matsuura, T., Sunami, T., Hosoda, K., Ichihashi, N., Tsukada, K., Urabe, I., and Yomo, T. (2008) Replication of genetic information with self-encoded replicase in liposomes. *ChemBioChem 9*, 2403−2410.

(20) Asahara, H., and Chong, S. (2010) In vitro genetic reconstruction of bacterial transcription initiation by coupled synthesis and detection of RNA polymerase holoenzyme. *Nucleic Acids Res. 38*, e141.

(21) Nishimura, K., Matsuura, T., Nishimura, K., Sunami, T., Suzuki, H., and Yomo, T. (2012) Cell-free protein synthesis inside giant

unilamellar vesicles analyzed by flow cytometry. *Langmuir 28*, 8426−8432.
(22) Stano, P., Kuruma, Y., Souza, T. P., and Luisi, P. L. (2010) Biosynthesis of proteins inside liposomes. *Methods Mol. Biol. 606*, 127−145.
(23) Martini, L., and Mansy, S. S. (2011) Cell-like systems with riboswitch controlled gene expression. *Chem. Commun. (Camb) 47*, 10734−10736.
(24) Kobori, S., Ichihashi, N., Kazuta, Y., and Yomo, T. (2013) A controllable gene expression system in liposomes that includes a positive feedback loop. *Mol. Biosyst. 9*, 1282−1285.
(25) Lentini, R., Forlin, M., Martini, L., Del Bianco, C., Spencer, A. C., Torino, D., and Mansy, S. S. (2013) Fluorescent proteins and *in vitro* genetic organization for cell-free synthetic biology. *ACS Synth. Biol. 2*, 482−489.
(26) Paige, J. S., Wu, K. Y., and Jaffrey, S. R. (2011) RNA mimics of green fluorescent protein. *Science 333*, 642−646.
(27) Paige, J. S., Nguyen-Duc, T., Song, W., and Jaffrey, S. R. (2012) Fluorescence imaging of cellular metabolites with RNA. *Science 335*, 1194.
(28) Ponchon, L., and Dardel, F. (2007) Recombinant RNA technology: The tRNA scaffold. *Nat. Methods 4*, 571−576.
(29) Diaz, G. A., Raskin, C. A., and McAllister, W. T. (1993) Hierarchy of base-pair preference in the binding domain of the bacteriophage T7 promoter. *J. Mol. Biol. 229*, 805−811.
(30) Nagatoishi, S., Ono, R., and Sugimoto, N. (2012) The yields of transcripts for a RNA polymerase regulated by hairpin structures in nascent RNAs. *Chem. Commun. 48*, 5121−5123.
(31) Du, L., Gao, R., and Forster, A. C. (2009) Engineering multigene expression *in vitro* and *in vivo* with small terminators for T7 RNA polymerase. *Biotechnol. Bioeng. 104*, 1189−1196.
(32) Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods 10*, 354−360.
(33) Lim, H. N., Lee, Y., and Hussein, R. (2011) Fundamental relationship between operon organization and gene expression. *Proc. Natl. Acad. Sci. U.S.A. 108*, 10626−10631.
(34) Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343−345.
(35) Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins. *Nat. Methods 2*, 905−909.
(36) Vrzheshch, E. P., Dmitrienko, D. V., Rudanov, G. S., Zagidullin, V. E., Paschenko, V. Z., Razzhivin, A. P., Saletsky, A. M., and Vrzheshch, P. V. (2008) Optical properties of the monomeric red fluorescent protein mRFP1. *Moscow University Biological Sciences Bulletin 63*, 109−112.
(37) Sambrook, J. J., and Russell, D. D. W. (2001) *Molecular cloning: A laboratory manual*. Vol. 2, 2nd ed.; Cold Spring Harbor Laboratory Press, New York.
(38) Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods 9*, 671−675.
(39) R Development Core Team (2007) R: A language and environment for statistical computing, Version 2. 5.0. R Foundation for Statistical Computing, Vienna.