

# Characterizing and Prototyping Genetic Networks with Cell-Free Transcription-Translation Reactions

Melissa K. Takahashi<sup>1#</sup>, Clarmyra A. Hayes<sup>2#</sup>, James Chappell<sup>1</sup>, Zachary Z. Sun<sup>2</sup>, Richard M. Murray<sup>2</sup>, Vincent Noireaux<sup>3\*</sup>, Julius B. Lucks<sup>1\*</sup>

1 - School of Chemical and Biomolecular Engineering, Cornell University, Ithaca NY 14850

2 - Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

3 - School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455

# - Joint First Authors

\* - Corresponding Authors

Vincent Noireaux

612-624-6589, noireaux@umn.edu, 115 Union Street SE, Minneapolis, MN, 55455

Julius B. Lucks

607-379-0865, jblucks@cornell.edu, 120 Olin Hall, Ithaca NY, 14850

## Abstract

A central goal of synthetic biology is to engineer cellular behavior by engineering synthetic gene networks for a variety of biotechnology and medical applications. The process of engineering gene networks often involves an iterative ‘design-build-test’ cycle, whereby the parts and connections that make up the network are built, characterized and varied until the desired network function is reached. Many advances have been made in the design and build portions of this cycle. However, the slow process of *in vivo* characterization of network function often limits the timescale of the testing step. Cell-free transcription-translation (TX-TL) systems offer a simple and fast alternative to performing these characterizations in cells. Here we provide an overview of a cell-free TX-TL system that utilizes the native *Escherichia coli* TX-TL machinery, thereby allowing a large repertoire of parts and networks to be characterized. As a way to demonstrate the utility of cell-free TX-TL, we illustrate the characterization of two genetic networks: an RNA transcriptional cascade and a protein regulated incoherent feed-forward loop. We also provide guidelines for designing TX-TL experiments to characterize new genetic networks. We end with a discussion of current and emerging applications of cell free systems.

**Keywords:** cell-free systems, transcription-translation (TX-TL), rapid prototyping, genetic networks, synthetic biology

**Abbreviations:** TX-TL (transcription-translation), I1-FFL (incoherent feed-forward loop type 1), GFP (green fluorescent protein)

## 1. Introduction

Cells have a remarkable ability to sense and process information about their external environment and their internal physiological state. They use this ability to adapt to constantly changing environments by making genetic decisions that control their behavior. These decisions range from selectively expressing metabolic enzymes that can produce a wide array of products, to altering motility patterns and differentiating their cell state. Harnessing and manipulating this diversity of cell behavior is thus a core aspect of many biotechnologies. These biotechnologies include engineering cells to

make products from renewable feedstocks [1,2], using cells as new therapeutic agents [3], and many others.

The natural ability of cells to control their behavior is encoded in genetic networks – interconnected webs of regulatory molecules that control gene expression in defined patterns. These genetic networks, or circuits, take internal and external signals as inputs, and shape the flow of genetic information through gene expression. In this way, genetic networks ultimately act as one of the core information processing units of the cell [4]. Because of this, our ability to engineer cellular behavior is directly related to our ability to engineer genetic networks, which is a central goal of synthetic biology [5-7].

There has been a great deal of interest in developing systematic and efficient methods for engineering genetic networks with predictable behaviors [5,8]. Because of the influences from other engineering disciplines, synthetic biologists often think of engineering genetic networks in terms of iterative ‘design-build-test’ cycles. A design-build-test cycle typically entails: designing DNA sequences that encode genetic regulatory ‘parts’ and their interconnections, building these DNA sequences in expression constructs, and testing the performance of the parts and networks in cells by monitoring the expression of detectable outputs. Recently, there has been rapid progress in both the design and build aspects of this cycle. Specifically, there are a number of new computational tools that can facilitate the design of genetic regulators [9-13]. There are also a plethora of high-throughput and high-fidelity DNA synthesis and assembly techniques that can be used to build, or even commercially order, DNA that encodes whole genetic networks [14-21]. However, progress has lagged in establishing reliable and efficient methods for testing these networks, making the testing step the current bottleneck in engineering and optimizing gene networks.

Two aspects of testing genetic networks in cells make this process slow and complex. First, DNA elements must be configured in compatible formats, such as different plasmid systems, which imposes constraints on the physical assembly and relative expression levels of each network part. Second, the process of incorporating this DNA into cells, for example by transformation, selection, and subculturing, takes several days, which delays characterization. Cell-free systems directly address these

limitations, and offer an alternative approach for characterizing outputs of genetic networks in a simplified *in vitro* environment that closely mimics the cell [22-28].

Cell-free systems typically consist of a cell lysate or purified transcription/translation machinery and a buffer/energy mix optimized to express genes from template DNA. The utility of cell-free systems was first realized in the 1960's when lysates were used to translate defined synthetic RNAs into proteins leading to the elucidation of the genetic code [29]. Since then, protein production has continued to be the major use of *E. coli* cell-free systems, though early applications were limited due to short-lived reactions and low protein yield [30-32]. This motivated a number of optimizations in energy source and energy regeneration [30,31,33-37] as well as the development of new *E. coli* strains that were engineered to stabilize amino acids [38] and improve protein expression from PCR products [39,40]. At the same time, preparation of crude extract was simplified [31,36,41], and alternate cell-free expression systems were developed by reconstituting *in vitro* transcription and translation from purified components [42]. These advancements have not only improved protein production, but have also allowed new applications such as the production of proteins with unnatural amino acids [43-46].

Building off of this rich history, researchers have now begun to leverage the flexibility of cell-free systems to express entire genetic networks for their functional characterization. A major advantage of cell-free systems for network characterization lies in their cell-free nature: testing cycles can be decoupled from the DNA formatting and transformation/cell growth issues that have hampered traditional network characterization. This has the effect of removing complications associated with ensuring that the DNA encoding the networks are on compatible plasmid origin and antibiotic resistance constructs. In fact, recent innovations have made it possible to test genetic networks constructed on linear PCR products, enabling the testing cycle to be directly coupled to high-throughput design and construction techniques [47].

In addition, cell-free systems decouple the experiment from cell growth, which allows an order of magnitude decrease in testing times, going from a typical three day experiment for testing in bacterial cells to a mere three hour experiment with cell-free systems [22]. Cell-free systems are also convenient because they are open reactions.

This allows flexibility in experimental design as well as control of the biochemical and biophysical components of the reactions. In addition, cell-free systems are opening the door to new types of applications, such as new molecular diagnostics which use TX-TL reactions to detect the presence of analytes in solutions [48]. Finally, because of their ease-of-use and rapid turnaround times, cell-free systems are finding new uses as teaching tools for synthetic biology through hands-on experiments [28].

The advantages and flexibility of cell-free systems have prompted recent efforts to develop a cell-free system that closely mimics the characteristics of the live cell environment. Such a system would enable the rapid prototyping of genetic networks for eventual deployment in the cell. Because the original focus of cell-free systems was on protein expression and maximizing protein yields, bacteriophage polymerases and promoters, such as T7, were used due to their high levels of transcription and specificity. Recently, an all *E. coli* cell-free transcription-translation (TX-TL) expression system was shown to be as efficient as the bacteriophage systems [23]. This system has all of the benefits of bacteriophage hybrid cell-free systems, but instead it only uses the native *E. coli* TX-TL machinery and recapitulates the seven-*E. coli* sigma factor transcription scheme. Consequently, the repertoire of regulatory elements that can be used is expanded to hundreds of parts rather than just variants of a single bacteriophage promoter. Multiple stage cascades, logic gates, negative feedback loops and other networks have been characterized with this system [49], which can be used in test tube reactions, microfluidics and liposomes [50]. In addition, methods to tune mRNA and protein degradation rates have been devised for this system [51], enabling even finer grained control over genetic network performance. Thus the *E. coli* cell-free TX-TL platform holds great promise as a toolbox for rapid testing and optimization of a large array of regulatory networks.

In this article we focus on the all *E. coli* cell-free TX-TL system [49]. We explain its versatility, its current capabilities and limitations. We start by providing general guidelines for executing TX-TL reactions by describing two examples where TX-TL reactions were used to characterize genetic parts and networks. We then discuss important considerations for a new user designing their own TX-TL experiments. Finally,

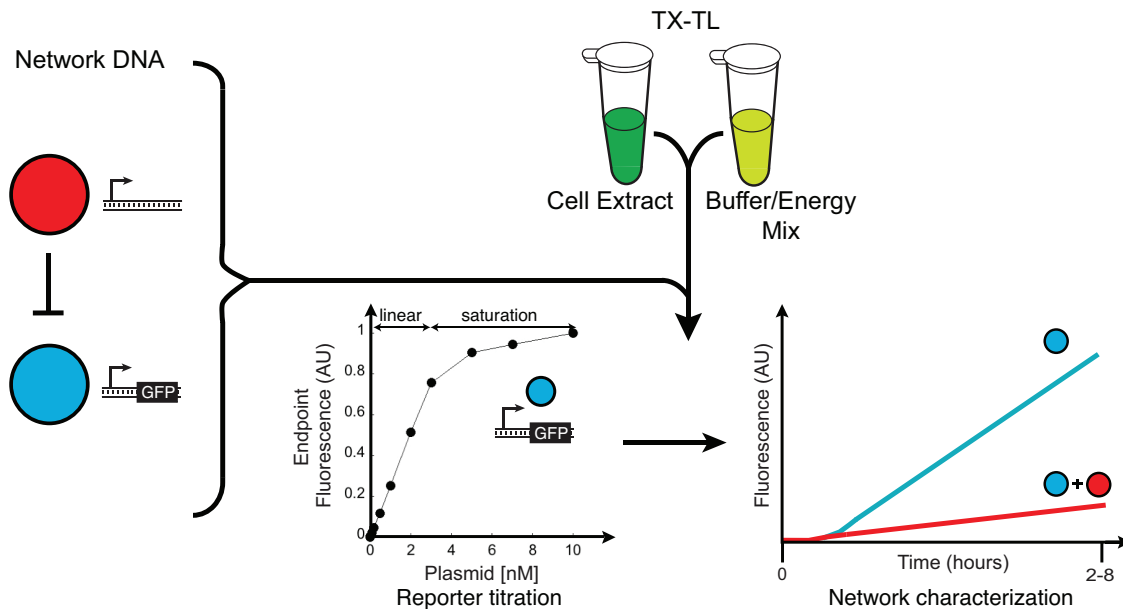
we end by discussing current applications and potential avenues for cell-free TX-TL systems.

## 2. Basic TX-TL preparation and experiment

The all *E. coli* cell-free TX-TL expression system was first developed by modifying existing *E. coli* S30 extract protocols [31,41] to create an *in vitro* gene expression system optimized for examining the dynamics of genetic networks driven by *E. coli* promoters [49]. The system combines crude *E. coli* lysate with a buffer mixture containing resources necessary for transcription and translation. The lysate is generated by bead-beating cell resuspensions from BL21 Rosetta2 cell cultures, however cell strain and lysis method can be optimized to meet individual needs [52]. The buffer contains protein and RNA building blocks (twenty natural amino acids, NTPs), tRNAs, an energy regeneration system composed of a phosphate donor such as phosphoenolpyruvate (PEP), as well as multiple other small molecules. Additional molecules can be added for specific effects. For instance, the protein GamS can be added to prevent the degradation of linear double-stranded DNA templates, while the protein complex ClpXP can be added to increase the degradation of *ssrA*-tagged proteins [47,53,54]. A detailed protocol for preparing both crude extract and buffer, as well as for setting up and running a TX-TL experiment, is described by Sun et al. [22].

A basic TX-TL experiment consists of adding DNA encoding a genetic regulator or network suspended in water, into a mixture of extract and buffer (Figure 1) [22]. Both plasmid DNA and linear PCR products can be used, though linear DNA will be degraded by endogenous exonucleases present in the extract unless the exonuclease inhibitor GamS is added (see Section 5.2). The volume of a typical TX-TL reaction is 10  $\mu$ L, with DNA added in concentrations ranging from 0.01 to 20 nM depending on expression strength. The network being tested can consist of one or multiple DNA constructs, with at least one regulatory output designed to produce a measurable signal, such as a fluorescent reporter protein (e.g. GFP). After DNA is added to the buffer and extract, the mixture is typically incubated either at 29°C or 37°C, and the fluorescence output measured over time. For endpoint measurements, the mixture is often placed in a microcentrifuge tube and kept in an incubator for 1-8 hours before being transferred

into a microplate and measured on a plate reader. Most experiments described in this article use a 10  $\mu$ L total reaction volume, placed in a 384-well microplate and measured for 1.5-8 hours on a Biotek plate reader. While we use 10  $\mu$ L for convenience, TX-TL reactions can be run at larger volumes as long as they are properly oxygenated.



**Figure 1.** Characterizing genetic networks in TX-TL reactions. Networks or parts are tested in TX-TL reactions by mixing the DNA for a given network with TX-TL extract and energy mix. Characterization typically begins with a titration of the DNA that controls the reporter protein expression. A plot of measured endpoint fluorescence versus the input DNA concentration is useful to determine the expression capacity of the TX-TL system and typically shows two regimes: a linear response up to a few nM of DNA template, and then saturation due to TX-TL molecular machinery depletion. Once reporter DNA concentration is chosen, the remaining parts of the network can be titrated and characterized.

### 3. How to characterize a genetic network in TX-TL

While the basic setup of a single TX-TL reaction is straightforward, there are several steps that we have found to be useful when characterizing a new regulatory part or network in the TX-TL system. To start, we have found that network characterization is most easily and rapidly carried out when fluorescent proteins are used to report on the gene expression outputs of the network. When choosing fluorescent reporter proteins it is important to consider maturation times and it is best to choose fluorescent proteins with faster and comparable maturation times. Other techniques, such as electrophoresis (SDS PAGE) to measure protein output, or FRET-based probes [55], can be used for quantification, but they are either more time consuming, or require the use of

specialized TX-TL systems [56]. After deciding on a characterization method, the next step is to construct the DNA that encodes the network elements. We have found that it is easiest to encode each part of the network on separate plasmids or DNA elements (see Section 5.2 for guidelines). This allows both ease of assembly and additional flexibility for optimizing network function.

Once the DNA parts have been assembled, we typically start the network characterization process by titrating the concentration of the DNA that controls the reporter protein expression, and characterizing its output. A plot of the measured endpoint fluorescence versus the input DNA concentration is useful to determine the expression capacity of the TX-TL system and typically shows two regimes: a linear response up to a few nM of DNA template, and then saturation due to TX-TL molecular machinery saturation (Figure 1). This curve also allows the user to choose a template concentration that balances signal to background, while minimizing the DNA input to allow the addition of other network elements (see Section 5.3 for resource limitations). From here, we titrate other network elements to test the effect of varying the concentration ratios of these elements on overall network function. Fluorescence versus time plots are collected in order to evaluate the dynamics of gene expression for the network. We also use knowledge of previously tested *in vivo* elements, such as relative plasmid copy number, to choose starting points for DNA concentrations. Results are typically presented as either production rates from raw fluorescence data [25,28] (Section 3.1) or as protein concentrations using a calibration curve generated from purified fluorescent protein [23,47,49] (Section 3.2). We now demonstrate this approach in the context of two example genetic networks – an RNA transcriptional cascade [28,57] and a protein regulated feed-forward loop.

### 3.1 RNA network example

As an example of how to use the all *E. coli* TX-TL system to test RNA genetic networks, we will describe the steps taken to characterize an RNA transcriptional cascade [28,57]. The central element of this cascade is an RNA-mediated transcriptional repressor, engineered from the pT181 transcriptional attenuator [58,59]. The attenuator lies in the 5' untranslated region of the gene it regulates, and functions



like a transcriptional repressor, either allowing (ON) or preventing (OFF) elongation of RNA polymerase through the transcript [58,59]. The OFF state is induced by an RNA-RNA interaction between the attenuator and a complementary antisense RNA, which is expressed in *trans* in this synthetic context (Supplementary Figure S1). By transcriptionally fusing the pT181 repressor to the coding sequence for super folder green fluorescent protein (SFGFP) [60], the function of the repressor can be characterized by measuring SFGFP fluorescence with and without antisense RNA present. A second, mutated version of the repressor, engineered to be orthogonal to the wild type repressor, enables the configuration of a double-repression transcriptional cascade that functions in *E. coli* [57].

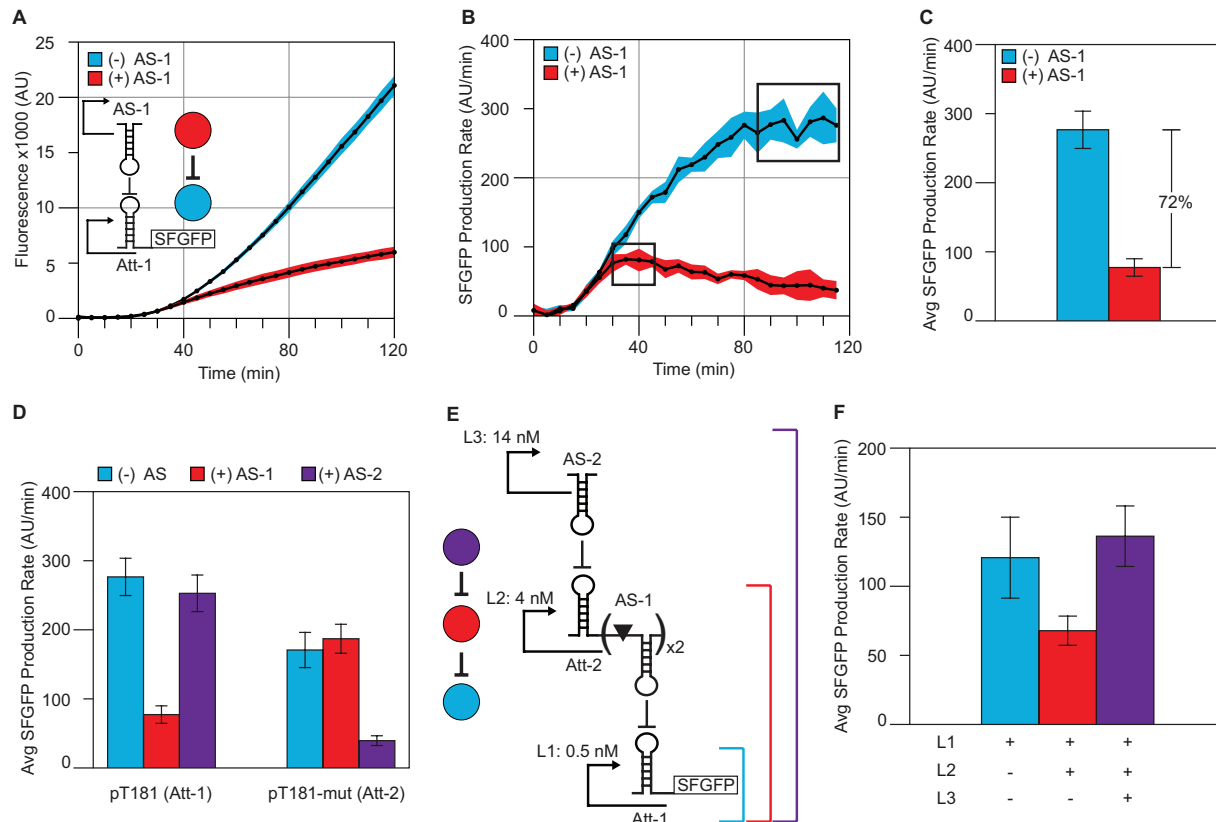
In order to characterize this RNA transcriptional cascade with TX-TL reactions, the first step was to test the function of the wild type pT181 repressor. To do this, the basic repression system was configured on two separate plasmids – one containing the expression cassette for the attenuator target region fused to SFGFP (reporter level), and the other containing the expression cassette for the antisense RNA (Figure 2A, Supplementary Figure S2). From here, the concentrations of plasmid DNA containing the attenuator (Att-1) and antisense (AS-1) RNAs were titrated in different TX-TL reactions to achieve a repression in TX-TL reactions comparable to that found *in vivo* [28] (Figure 2A-C). Attenuator plasmid concentrations were varied between 0.25 nM and 1 nM, while antisense plasmid concentrations were varied between 4 nM and 16 nM (Supplementary Figure S3). Since SFGFP is very stable and not degraded in TX-TL reactions, an increase in fluorescence over time is always observed due to imperfect repression even in the (+) antisense conditions. Because of this, and resource depletion effects that occur in TX-TL reactions [22], end point fluorescence measurements can be misleading [25]. Therefore network performance was characterized by quantifying SFGFP production rates that are calculated as the slope between consecutive time points in the fluorescence time course data [25]. In particular, these calculations are typically done at the beginning of the time courses (up to 2 hours) when nutrients (nucleotides, amino acids) are not limiting. From these rates, windows of constant maximum protein production for each trajectory were found, averaged over several replicate experiments, and compared to assess overall gene repression in the TX-TL

reactions (Figure 2B-C) [28]. We note that the region of maximum protein production occurs at different times between the with and without antisense RNA reactions. Furthermore, the with antisense RNA production rate decreases ~ 40 min after the maxima is reached. One reason for this decrease could be due to resource depletion caused by the RNA-RNA interactions triggering additional RNA degradation pathways. Alternatively, this could be a result of the slow degradation of Att-1-SFGFP transcripts that escape attenuation at the start of the reaction. Independent of a specific cause, we use the region of maximum production rate as a conservative estimate of attenuator repression [28]. These titrations led to a final concentration of 0.5 nM attenuator and 8 nM antisense plasmids.

Next, the function and orthogonality of the mutated version of the pT181 attenuator [57] (Att-2, AS-2) was tested in the TX-TL system. Orthogonality was assessed by comparing average SFGFP production rates for reactions of each attenuator with a no-antisense control (used as a resource utilization control, see Section 5.3), its cognate antisense, or the non-cognate antisense using DNA plasmid concentrations found from the experiments described above. For each attenuator, cognate antisense RNAs caused repression while non-cognate antisense RNAs resulted in SFGFP production rates that were within error of the no-antisense control, thus confirming orthogonality in TX-TL (Figure 2D) [28].

These parts then allowed for the characterization of the full RNA transcription cascade that combined these two elements together (Figure 2E) [28,57]. In this configuration, the bottom level of the cascade (L1) contains the wild type pT181 attenuator (Att-1), which regulates the expression of SFGFP expression via its interaction with its cognate antisense (AS-1). The production of AS-1 is in turn regulated by the mutant pT181 attenuator (Att-2) on the second level of the cascade (L2) through its interaction with AS-2, which is transcribed on level three of the cascade (L3) (Figure 2E). DNA concentrations for the three elements of the cascade were titrated for optimal performance [28], with the final result shown in Figure 2F. The presence of just L1 results in high SFGFP production. A combination of L1 and L2 leads to repression of Att-1, thus lower SFGFP production. Finally, a combination of L1, L2, and L3 results in a double inversion leading to high SFGFP production (Figure 2F) and confirmation of a

functional RNA transcription cascade in the TX-TL system. The characterization of the three level cascade, once DNA was prepared, required five TX-TL experiments, each completed in three hours.



**Figure 2.** Characterizing an RNA transcriptional cascade in TX-TL reactions. (A) Fluorescence time courses of TX-TL reactions containing the pT181 attenuator reporter plasmid at 0.5 nM, with 8 nM antisense plasmid (+) or 8 nM no-antisense control plasmid (-). Colored circles represent the different plasmids in the system; the antisense (red) represses (blunt end line) the attenuator-SFGFP target (blue). (B) SFGFP production rates were calculated from the data in (A) by calculating the slope between consecutive time points. Boxes represent regions of constant SFGFP production. Blue and red shaded regions in (A) and (B) represent standard deviations of at least seven independent reactions performed over multiple days calculated at each time point. (C) Average SFGFP production rates were calculated from the data in boxed regions in (B). Error bars represent standard deviations of those averages. The (+) antisense condition shows 72% repression compared to the (-) antisense condition in TX-TL reactions. (D) Orthogonality of the pT181 attenuator (Att-1) to a pT181 mutant attenuator (Att-2). Average SFGFP production rates were calculated as in (C). Bars represent each attenuator at 0.5 nM with 8 nM of no-antisense control plasmid (blue), pT181 antisense plasmid (AS-1, red), or pT181 mutant antisense plasmid (AS-2, purple). (E) Schematic of an RNA transcriptional cascade. L1 is the same pT181 attenuator (Att-1) reporter plasmid used in (A) – (D). In the plasmid for L2, the pT181-mut attenuator (Att-2) regulates two copies of the pT181 antisense (AS-1), each separated by a ribozyme (triangle) [57]. The L3 plasmid transcribes the pT181-mut antisense (AS-2). Colored circles represent the different plasmids in the system with blunt end lines showing the repressive connections of the cascade. (F) Average SFGFP production rates for the three combinations of the transcription cascade levels depicted in (E). L1 alone (blue bar) leads to high SFGFP production. L1+L2 (red bar) results in AS-1 repressing Att-1, thus lower SFGFP production. L1+L2+L3 (purple bar) results in a double inversion leading to high SFGFP production. Total DNA concentration in each reaction was held constant at 18.5 nM. In (D) and (F) error

bars represent standard deviations from at least seven independent reactions performed over multiple days. Figure from Takahashi et al. ACS Synth. Biol., 4 (2015) 503-515 [28].

### 3.2 Protein network example

The steps taken to test protein-mediated networks in the TX-TL system are similar to those for RNA-mediated networks, with a few protein-specific additions. For instance, while RNAs are readily degradable by RNases within the extract, there is often not enough protein degradation machinery native to the extract to significantly degrade expressed protein. Protein degradation rate can be greatly increased by the addition of the protease ClpXP, either on DNA or as a purified protein. Any proteins that require degradation must be expressed with a degradation tag specific to ClpXP [61] (Supplementary Figure S4). Another difference between RNA and protein-mediated networks is that proteins take a longer time to express than RNAs because of the additional steps of translation and protein folding. The amount of time needed is protein specific and generally on the order of minutes. Consequently, proteins take longer than RNAs to reach the concentration thresholds needed to be active in the specific network tested [62].

As an example of how to test protein regulated networks in the TX-TL system, we will explain how we prototyped an incoherent type-1 feed-forward loop (I1-FFL) [63]. Our version of the I1-FFL consists of three “transcription units” (TU, promoter through terminator), each on its own plasmid (Figure 3A). The first unit (TU1) encodes the gene for sigma-28 (rpoF), constitutively expressed by an *E. coli* promoter specific to sigma-70. *In vivo*, sigma-28 is related to flagellum formation [64]; however in TX-TL, sigma-28 acts as an activator for any gene downstream of the sigma-28 dependent promoter  $P_{tar}$  [49]. The second unit (TU2) encodes the protein transcriptional repressor TetR, downstream of  $P_{tar}$ . TetR sterically represses transcription of promoters containing the tetO operator [65]. The third unit (TU3) encodes the fluorescent reporter GFP (green fluorescent protein) downstream of a rationally designed  $P_{tar}$ -tetO hybrid promoter. This promoter is activated by sigma-28 and repressed by TetR.

The I1-FFL functions as follows. At time zero, all promoters are off except for the constitutive promoter, which starts transcription of sigma-28. After translation, sigma-28 together with RNA polymerase binds to and initiates transcription of both  $P_{tar}$  promoters

on units TU2 and TU3. This results in the simultaneous transcription and then translation of TetR and GFP, and an initial increase in measured fluorescence. However, over time TetR accumulates in the system, and eventually reaches a concentration sufficient to shut down GFP production from TU3. In the TX-TL system with minimal native protein degradation, this results in the fluorescent signal plateauing at a constant value. If ClpXP is added to TX-TL reactions, the output is a pulse of measured fluorescence over time (Figure 3A).

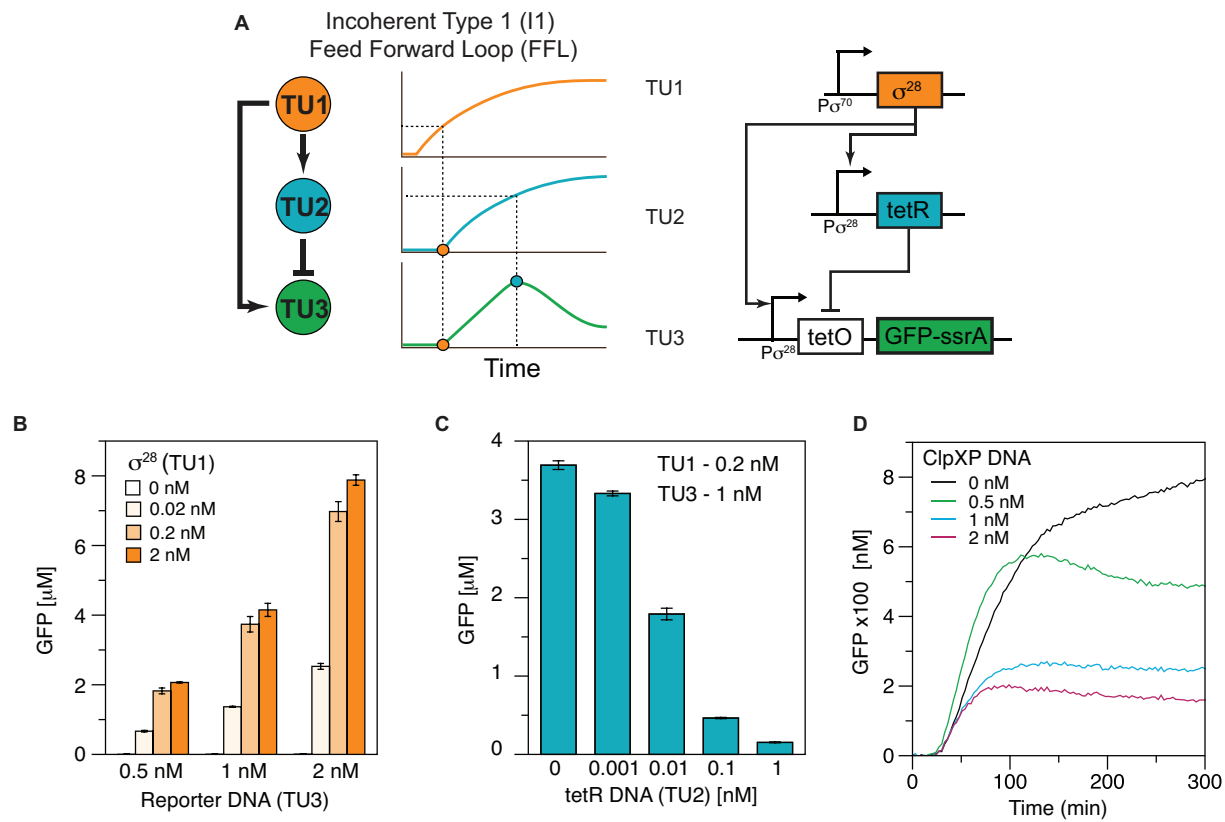
To prototype the I1-FFL in the TX-TL system, the first step was to determine optimal sigma-28 DNA (TU1) and reporter DNA (TU3) concentrations for *in vitro* testing. The concentration of reporter DNA was varied between 0.5 and 2 nM, and the concentration of sigma-28 DNA was varied from 0 to 2 nM (Figure 3B). To obtain maximum signal, fluorescence was measured at 8 hours, when the reactions had plateaued. The results show that i) the previously untested  $P_{tar-tetO}$  hybrid promoter successfully produces GFP in the presence of sigma-28, ii) the hybrid promoter's leakiness is negligible, with very little GFP produced in the absence of sigma-28, and iii) that while all non-zero concentrations of sigma-28 and reporter DNA tested are sufficient to produce fluorescent signal, both 0.2 and 2 nM of sigma-28 DNA result in significantly higher signal than 0.02 nM. (Figure 3B). To avoid saturating the system with unnecessary DNA, we chose to use 0.2 nM of sigma-28 DNA for future experiments. In addition, the results show that adding 2 nM of reporter DNA gives less than twice the signal of 1 nM reporter DNA, indicating that the TX-TL molecular machinery is starting to become saturated at 2 nM of reporter DNA. We therefore chose to use 1 nM of reporter DNA, as it gives relatively high signal and does not saturate the TX-TL machinery.

The second step was to explore the effect of TetR DNA concentration on GFP output. Setting reporter DNA (TU3) concentration constant at 1 nM and sigma-28 DNA (TU1) at 0.2 nM, we varied TetR DNA (TU2) from 0 to 1 nM (Figure 3C). The results show that i) the novel  $P_{tar-tetO}$  hybrid promoter is repressed by TetR and thus functions as designed, ii) the TU2  $P_{tar}$  promoter produces sufficient TetR to repress GFP production, and iii) that even a very small amount of TetR DNA (0.001 nM) is sufficient to decrease signal. Since our aim was to prototype an I1-FFL, we wanted to choose a

TetR concentration that significantly represses signal, but not so strongly that the signal falls into the noise range of our instruments and becomes difficult to detect. Therefore, we chose to use 0.01 nM of TetR DNA, which decreases signal more than 2-fold but still gives us more than 1  $\mu$ M of fluorescent protein.

The third step was to determine if the network would exhibit a pulse-like signal in TX-TL when protein degradation machinery was added. To test this, reporter (TU3), sigma-28 (TU1), and TetR (TU2) DNA concentrations were set at 1, 0.2 and 0.01 nM, respectively. DNA encoding ClpXP was added in concentrations ranging from 0 to 2 nM (Figure 3D). The results showed that i) ClpXP successfully degraded the degradation-tagged GFP protein, ii) the addition of 0.5 nM ClpXP resulted in a small pulse, characteristic of the I1-FFL, and iii) that adding ClpXP DNA to the system significantly decreases maximum signal.

Overall, characterization of the I1-FFL in the TX-TL system required only three experiments, each completed in less than a day. We were able to rapidly test different concentrations of the circuit transcriptional units to determine what ratios result in successful I1-FFL pulse-like behavior, a test which *in vivo* would have required varying plasmid copy numbers or swapping in different strength promoters or ribosomal binding sites. Additionally, because TX-TL reactions are entirely *in vitro*, we were able to test a non-traditional “activator,” sigma factor 28, without potential *in vivo* side effects such as flagellum activation and toxicity.



**Figure 3.** Characterizing a protein-mediated incoherent feed-forward loop. (A) Schematic of an incoherent type-1 feed-forward loop (I1-FFL). Transcription unit 1 (TU1) constitutively expresses sigma-28 ( $\sigma^{28}$ ), which activates the expression of both transcription unit 2 (TU2) and transcription unit 3 (TU3). TU2 produces the TetR repressor protein, which represses the production of GFP from TU3. The simultaneous activation of TU2 and TU3 results in a pulse of GFP when the protease ClpXP is present to degrade the GFP signal. (B) Average measured GFP concentration for TX-TL reactions with varying TU1 (0–2 nM) for three different TU3 concentrations (0.5, 1, 2 nM) at 8 hours of reaction time. Error bars represent standard deviations from three independent reactions. The bars corresponding to the 0 nM TU1 reactions are not visible on this scale. Average concentrations were 3, 5, and 8 nM GFP for 0.5, 1, and 2 nM TU3, respectively. (C) Average measured GFP concentration for 1 nM TU3, 0.2 nM TU1, and varying TU2 DNA from 0–1 nM at 8 hours of reaction time. Error bars represent standard deviations from three independent reactions. (D) Full I1-FFL with added ClpXP DNA. Plot of measured GFP concentration over time with 1 nM TU3, 0.2 nM TU1, 0.01 TU2, and varying ClpXP DNA from 0–2 nM. Each result represents the average of two independent TX-TL reactions. The TX-TL experiments for figure 3D were run with a different extract batch from figures 3B and 3C, so the protein concentrations are not directly comparable.

#### 4. Using spike experiments to measure responses to network perturbations

The openness of TX-TL reactions enables a great deal of flexibility in experimental design. In particular, the lack of a cell membrane enables the addition of new DNA constructs, small molecules such as inducers, and macromolecules such as transcription factors at any point during an experiment. These ‘spike’ additions, coupled



with frequent monitoring of network output on a microplate reader, allows for the determination of network response times to the addition of a new species and more broadly dynamic perturbations of network conditions. In fact, ligand response times measured using TX-TL reactions have been shown to be relevant to similar response times measured *in vivo* (Section 5.7). This demonstrates that TX-TL spike experiments can be used to rapidly test network perturbations for *in vivo* use.

DNA spike experiments were recently used to measure the response time of the RNA transcription cascade described in Section 3.1 (Figure 2E) to the addition of the L3 antisense RNA [28]. The L3 antisense RNA (AS-2) triggers the double inversion of the cascade. Therefore, its addition to a reaction only containing the bottom two levels of the cascade should cause an increase in the fluorescence trajectory for the reaction. Fluorescence trajectories for reactions with and without the addition of L3 were compared to determine the time at which the trajectories diverge. This gives a measure of the network response time to the addition of L3. To measure this, TX-TL reactions with 0.5 nM L1 and 4 nM L2 were setup and allowed to proceed for 20 min in the microplate, at which point 14 nM of either L3 or a no-antisense control plasmid was spiked into the reactions on the microplate using a multi-channel pipette ( $t=0$ ). Reactions were placed back on the microplate reader and fluorescence monitored every minute. The fluorescence trajectories showed that the L3 spike differed from the control approximately 15 min after the spike (Supplementary Figure S5). A Welch's t-test was used to determine the point at which the two trajectories were significantly different across multiple experimental replicates ( $14.6 \pm 4.8$  min – Supplementary Figure S5) [28]. This represents the response time of the cascade to the addition of L3 DNA. Similar response times can be measured in this way for other protein or RNA networks, inducible promoters, aptamer constructs, etc.

There are a few important considerations when designing/performing spike experiments: (i) An appropriate control reaction should be designed to maintain equal volumes and equal total amounts of DNA in reactions that will be compared. In the RNA cascade example the total DNA concentration in each reaction was especially important (see section 5.4), therefore a no-antisense control was added in parallel. (ii) When performing the actual spike addition, it is important to avoid addition of bubbles to the



reaction wells. If bubbles are present after the spike, the data should be discarded and additional replicates performed. (iii) If the response time for the experiment is expected to be fast, it is best to use a multichannel pipette to add elements to multiple reactions at the same time.

## 5. Important considerations

### 5.1 Batch to batch variation

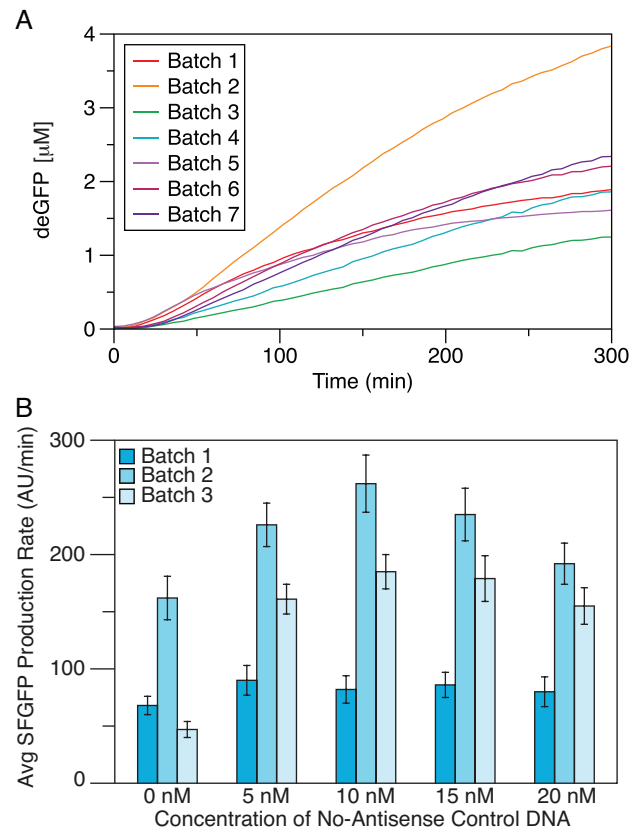
As described earlier, the TX-TL system is a combination of S30-like cell free lysate, buffer solution, and DNA. We define batch-to-batch variation as variability that results when the lysis method and buffer solution is held constant to a set protocol. In lieu of characterizing individual components in the extract, we broadly determine this variability using a plasmid that strongly expresses GFP. While all extracts express the plasmid, they can differ in the dynamics and strength of expression (Figure 4A).

Theoretically, lysates made using the same strain and following the same preparation protocol will have the same expression dynamics. However, in practice there can be significant variability depending on the person preparing the extracts and factors intrinsic to the process but hard to control for, such as the exact optical density (OD) the cells are captured in at mid-log phase, time cells are on ice, or lysis efficiency. Additionally, due to the buffer being composed of at least 25 ingredients, different buffer preparations may have slight variations in reagent concentrations. Therefore, we define a combination of a similarly prepared lysate and buffer as a “batch” and conduct all experiments within this batch. Due to batch-to-batch variation, experiments from different batches are generally not directly comparable, though results can often be correlated.

Our typical extract uses a commercial strain of BL21-Rosetta2. However, lysates have been made from multiple strains of *E. coli* such as BL21-type, Rosetta2-type, JM109-type, MG1655-type, DE3-expressing and custom variants [52]. To date, all of the strains have been usable for prototyping. However, protein-expressing ability seems to be partially dependent on strain.

Variability can also result from the processes used to prepare the materials. There are many lysis methods that can be employed, such as mechanical lysis with

bead-beating [22,23], microfluidizing [35], and sonication [52,66]. There are also different types of buffering solutions, each using independent energy sources [30,33-35,37,67]. There is limited work, however, on evaluating the ability of these extract variants to conduct gene network prototyping.



**Figure 4.** Batch-to-batch variation. (A) Batch-to-batch variation demonstrated by the expression of 1nM of a strong deGFP-expressing plasmid using seven disparate batches composed of extracts and buffers prepared by different people on different days. (B) RNase batch-to-batch variation. Average maximum SFGFP production rates from three different buffer and extract preparations using 0.5 nM L1 and 0, 5, 10, 15, and 20 nM no-antisense control DNA from the RNA transcriptional cascade (Figure 2). Error bars represent standard deviations from at least 11 independent reactions. Batches in (B) were not the same as in (A). Part (B) adapted from Takahashi et al. ACS Synth. Biol., 4 (2015) 503-515 [28].

## 5.2 DNA preparation

DNA inputs for cell-free reactions can be plasmid DNA and linear or circularized [25] PCR products, though there are specific considerations for each type of input. For plasmid DNA, a pure and well-supercoiled solution in a low-salt buffer is ideal. As mentioned previously, there are no limitations on compatibility of plasmid origins or antibiotic resistance.

An advantage to using linear DNA to prototype new parts and networks is that it can be constructed via a rapid DNA assembly method such as Golden Gate assembly [14]. In this way, a new network construct can be assembled and PCR amplified for use in a TX-TL reaction without going through *in vivo* selection and purification [47]. TX-TL extract contains endogenous DNA exonucleases, making it important to include a truncated version of the bacteriophage Gam protein, GamS, (3.5  $\mu$ M in the final reaction) in the reactions to protect linear DNA from degradation [47,53]. In addition, adding 250 bases of steric protection (non-coding DNA sequence) to each end of the linear construct can further slow degradation.

It is important to note that transcriptional promoters can function differently in TX-TL reactions depending on whether linear or plasmid DNA templates are used [25,47]. Such variation has been hypothesized to be due to differences in preparation of DNA templates [47] as well as conformational differences such as supercoiling between linear and plasmid DNA [25,47]. Two methods for correcting this difference have been suggested. First, linear DNA templates can be circularized using a USER-ligase reaction. The circularized templates have been shown to be comparable to plasmid templates in TX-TL reactions [25]. Limitations are extra processing time and the relatively low efficiency of the circularization reaction. One alternative is to characterize the difference between the expression profiles of a given regulator on circular and linear DNA templates, and calibrate characterization data accordingly [47].

### 5.3 Resource limitations/usage

Most cell-free TX-TL reactions are performed in batch mode, i.e. in a fixed volume much larger than a cell (typically 10  $\mu$ l), using a diluted cytoplasmic extract. A direct consequence of this setup is resource limitations at two levels: (i) fixed concentrations of nutrient resources (NTPs, amino acids, etc.), which limit the lifetime of reactions, and (ii) fixed concentrations of the core TX-TL machinery, which limit the rate of RNA and protein synthesis.

In terms of nutrient resources, translation places some of the highest demands on chemical energy usage since two ATP and two GTP are consumed per peptide bond formed. In modern cell-free TX-TL systems, batch mode reactions terminate because of

chemical energy limitations and accumulation of byproducts. Therefore, most of the effort to increase cell-free protein expression is spent on developing new metabolisms to energize translation, which consists of devising methods to i) maintain ATP and GTP concentrations at their initial concentrations (about 1-2 mM), ii) regenerate ADP and AMP, and iii) recycle reaction byproducts such as inorganic phosphate, which is a strong inhibitor of protein synthesis. An ATP regeneration system composed of a phosphate donor (for example, creatine phosphate) and a kinase is also added to the energy mixture. Such a system extends gene expression up to 3-4 hours [26].

Recently, it was shown that adding a carbon source (maltose or maltodextrin) extends protein synthesis up to 10 hours by activation of the glycolytic pathway [67,68]. In this method, the carbon source acts as a substrate for both ATP regeneration and recycling of inorganic phosphate. In effect, the phosphate donor and the carbon source are used to keep the adenylate energy charge at a sufficient level until protein synthesis ceases. This is important, since adding fresh ATP to the solution after a few hours of incubation does not work, unless the amount of fresh ATP is large, because the batch reaction is inevitably loaded with ADP and AMP. In addition, adding fresh ATP to the reaction is complicated since magnesium also has to be added due to the addition of negatively charged ATP. The ultimate solution to this problem is to carry out continuous buffer exchange through a dialysis membrane against a feeding solution containing the nutrients [49,69], or using microfluidics devices either based on diffusion exchange or continuous dilution [50,56]. With those setups, cell-free TX-TL reactions can be extended from a few hours to at least one day until the TX-TL machineries lose their functions.

While chemical energy is the most important nutrient limitation in run-off reactions due to the high demand for translation, other resources can also be limiting. Some amino acids, for example, are unstable in cell-free reactions [38]. Additionally, as a reaction progresses, the pH decreases (typically from 8 to 6-6.5) due to the accumulation of acid-insoluble species, which impacts the entire reaction [70]. For these reasons, kinetic models of cell-free TX-TL reactions are performed for the first few hours when the reaction is not limited by resources [71-73].

The other resource limitation is the finite concentration of the core TX-TL machinery, which is ultimately determined by cellular machinery concentrations. Part of this resource limitation is due to the 25-30x dilution factor necessary for the extract preparation. The cytoplasmic protein concentration in *E. coli* cells is on the order of 250-300 mg/ml [74] making this concentration on the order of 10 mg/ml in typical *E. coli* cell-free reactions. This happens to correspond to an optimum for protein synthesis [70]. Though no concrete explanations for this particular concentration have been provided, *E. coli* cell-free systems containing more than 10 mg/ml of cytoplasmic extract have been tested and are not more productive [75].

Cell-free reactions therefore contain fixed concentrations of TX-TL molecular machineries, typically 25-30 times less than real *E. coli* cells: a few tens of nanomolar for RNA polymerase, and 1-2  $\mu$ M ribosomes compared to about 50  $\mu$ M in living cells [49]. For transcription, this dilution is not a problem because most of the cell-free systems have T7 RNA polymerase added to the reaction. However, it has been shown that the most efficient cell-free system using the endogenous *E. coli* core RNAP with sigma factor 70 (present in the cell extract) for transcription is as efficient as T7-based systems [67]. For translation, two regimes are typically observed as a function of added DNA template concentration: a linear response and a saturation that occurs at a few nanomolar plasmid [26,47,49] (Figure 1). Both the final amount of protein produced and the constant protein synthesis rate (determined in the first two hours of expression) show those same two regimes. The saturated response, observed using T7 RNA polymerase and T7 promoter as well as *E. coli* transcription with strong promoter-UTR pairs, corresponds to a depletion of the TX-TL machineries onto the genetic templates. As the plasmid concentration increases, either the RNA polymerase pool is entirely sequestered by the input DNA, or the ribosomes pool is entirely sequestered on the produced mRNA [76]. Whether it is the transcription or the translation that is limiting depends on the type of cell-free system used. In either case, protein synthesis reaches a maximum at a specific DNA concentration that is dependent on promoter-UTR strength (usually a few nanomolar for strong promoter-UTR pairs such as T7). Currently, no 'metric' exists with respect to well-defined *E. coli* regulatory parts to compare the transition from linear to saturated regimes between regulatory elements.

Technically, this transition is found by expressing eGFP from the promoter-UTR of interest and by performing a plasmid titration to determine cell-free expression. This allows cell-free users to set plasmids concentrations and know the load on the TX-TL core machinery. In cell-free gene network characterization, accelerating mRNA turnover [51] to endogenous level (5-6 minutes mRNA mean lifetime in *E. coli*) prevents accumulation of mRNA and saturation of the translation machinery [49].

#### 5.4 Specific RNA considerations

The main consideration when testing RNA networks in TX-TL reactions is the presence of RNA degradation machinery (RNases) in the extract. RNA degradation is important for RNA network function since the signals that propagate through the networks are RNA molecules, which need to degrade in order to control the dynamic behavior of the network [62]. Cellular RNases are carried over during extract preparation, and their concentrations can vary from batch to batch. RNase inhibitors, or the *E. coli* interferase MazF [51], can be used to decrease or increase RNA degradation respectively.

It is not necessary to tune RNA degradation for all applications, but it is important to control for the effects of RNases. For example, in the case of the RNA transcriptional repressor (Section 3.1), a no-antisense control plasmid was designed to enable comparisons between reactions with and without the antisense RNA plasmid [28]. The no-antisense control plasmid has the same promoter as the antisense plasmid driving the expression of an RNA transcriptional terminator. A titration of the no-antisense control plasmid with 0.5 nM of the attenuator-SFGFP reporter shows that an addition of 5 nM control plasmid causes an increase in SFGFP production rate (Figure 4B) [28]. We hypothesize that this increase is due to competition effects for RNases. While the control plasmid does not affect the attenuator in a mechanistic way, the RNA terminator can provide a decoy substrate for RNases, in effect stabilizing the attenuator-SFGFP reporter mRNA, and leading to an increase in SFGFP production. For most batches, as the concentration of the no-antisense control plasmid is increased, the SFGFP production rate plateaus once the degradation machinery is saturated (Figure 4B, Batch 1,3) [28]. This saturation point varies due to batch-to-batch differences in RNase

concentration. As a result, it is important to design an appropriate control construct for each network being tested, and to use the control to maintain a constant total DNA concentration in each reaction that will be compared.

### 5.5 Specific protein considerations

When the network being tested in TX-TL reactions expresses proteins, network dynamics will be affected by the timescales of both transcription and translation. Proteins require time to be translated and to fold, and may form dimers or other oligomers. Additionally, protein degradation is dependent on the presence and concentration of the protease ClpXP, and requires the protein to be expressed with a degradation tag [61]. TX-TL resources such as ATP, GTP and amino acids are also consumed by protein production.

Measurements of transcription and translation rates in TX-TL reveal that both are 1-2 orders of magnitude slower than in *E. coli*, due mainly to the 25-30x dilution of TX-TL machinery in the extract as compared with *in vivo* conditions [71]. In the first hour of a reaction, before resource limitations and reaction waste products become relevant, the protein synthesis rate in TX-TL scales linearly with mRNA synthesis dynamics, suggesting that transcription rate, not translation rate, limits protein production rate *in vitro* [71]. Protein folding time is variable and protein-specific; the maturation time in TX-TL of deGFP, a variant of eGFP [23], was measured at 8-8.5 minutes, while that for Luciferase, an alternative reporter protein, was less than 1 minute [51].

There are small amounts of the ClpXP protease that are endogenous to the cell-free extract that can degrade ssrA-tagged proteins [51,71]. To test networks that require significant protein degradation to function, such as oscillators or feed-forward loops, supplemental ClpXP can be added to a reaction through additional DNA templates. However, production of ClpXP will require a delay in degradation ability as well as the use of TX-TL resources to make the subunits. A viable alternative is to add purified synthetically linked hexameric ClpX to the TX-TL reaction [77]. We have found that extracts are generally ClpP saturated, and the addition of ClpX alone can increase degradation rate to a certain cutoff point. The addition of ClpX alone is relatively resource independent [54].



## 5.6 Negative control fluorescence and subtraction

A negative control should be run with each TX-TL experiment. This negative control is a mixture of the extract, buffer, and water (instead of network components). The fluorescence (or other measured output) of the negative control should be measured and subtracted from each time point of the experimental conditions. The negative control does display measureable autofluorescence, and in the green excitation/emission regime (485 nm/520 nm) this fluorescence decreases over the first 20-40 minutes after the buffer and extract are mixed together (Supplementary Figure S6A). This fluorescence decrease is not seen with every excitation/emission combination (Supplementary Figure S6B), and a pre-incubation of the extract and buffer for 20-40 minutes at 37°C before adding the network components can eliminate this decrease (Supplementary Figure S6A). The time necessary for pre-incubation varies with extract batch. A likely hypothesis for this fluorescence decrease is that the autofluorescence observed in the green regime is due to the presence of oxidized flavin-containing molecules, which have a maximum absorbance near 450 nm [78]. The decrease in fluorescence noted above may therefore be caused by the reduction of flavins upon mixing extract and buffer, the latter of which is highly reducing (Supplementary Figure S6C). Besides altering background fluorescence, there is no evidence to indicate that flavin oxidation state affects overall network performance.

## 5.7 Converting TX-TL results to *in vivo* results

An important aspect of TX-TL systems that is beginning to be explored is its use to transition prototyped networks from *in vitro* to *in vivo*. There have already been a handful of demonstrations showing that parts and networks prototyped and optimized in TX-TL function similarly when ported to cells. These include panels of transcriptional and translational units [25,47], the transformation of an extension of the RNA cascade reviewed in Section 3.1 [28], and recent success in prototyping and transitioning novel negative feedback protein oscillators [27].

In addition to characterizing RNA transcription cascade response times (Section 4), we recently used TX-TL reaction spike experiments to test parts required to build a



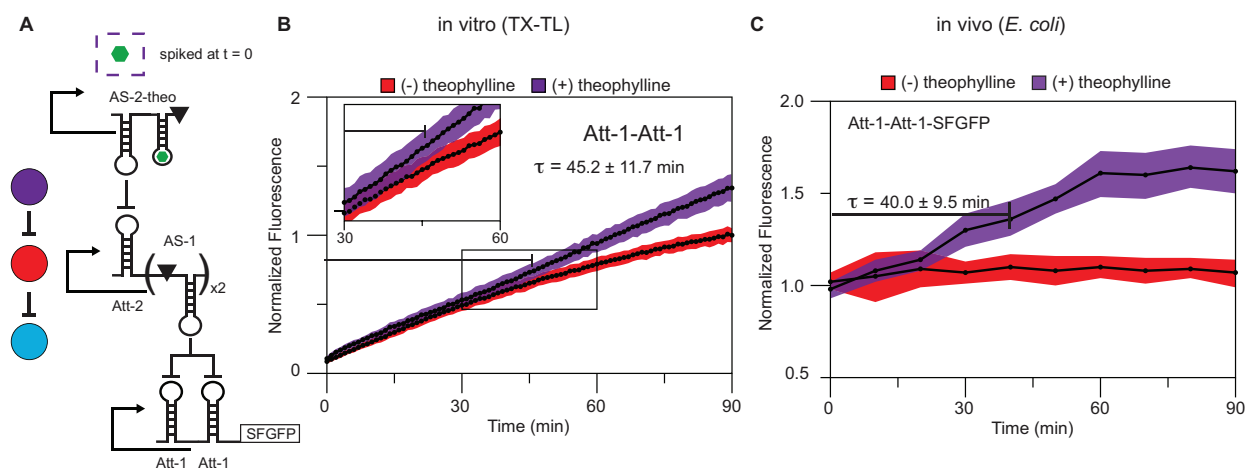
new network called an RNA single input module (SIM). The RNA SIM is an extension of the transcription cascade configured to dynamically stage the expression of two proteins instead of one [28]. Two additional parts were required to build the RNA SIM. The first was a construct that placed two copies of the pT181 attenuator in tandem, upstream of the SFGFP coding sequence. This increases the response time of a cascade by making the bottom level more sensitive to antisense concentration. The second was a construct that would allow for activation of the top level of the cascade with an inducer that could be used for *in vivo* spike experiments. In this case, we used an antisense RNA fused to the theophylline aptamer that was engineered to only be functional in the presence of theophylline [79].

To prototype this new theophylline-activated RNA cascade, we first performed a DNA spike experiment in TX-TL reactions to show that an RNA cascade using the double attenuator construct for L1 indeed had a slower response time (~20 min) than the cascade in Section 4 [28]. We then replaced L3 with the theophylline responsive antisense and performed a theophylline spike experiment in TX-TL reactions to mimic an *in vivo* theophylline spike experiment (Figure 5A,B). These spike experiments showed that an RNA SIM was feasible and the full RNA SIM network was built from these parts. A theophylline spike experiment showed that the complete RNA SIM functioned *in vivo* [28], and a comparison of the response time for the double attenuator construct in the SIM (Figure 5C) was remarkably similar to the response time from the theophylline spike experiment done in TX-TL reactions (Figure 5B). Similar agreement was observed when comparing the oscillation periods of ring oscillators prototyped in TX-TL reactions and further characterized *in vivo* [27].

These initial findings indicate that transitioning prototyped networks from *in vitro* to *in vivo* are possible, and that transcriptional and translational units should transition when on plasmid DNA. For further predictability, recently developed tools for standardizing transcriptional and translational strength *in vivo* should also be helpful *in vitro*. These include panels of benchmarked synthetic promoter strengths [80,81] and predictive bi-cistronic device ribosome binding sites [82].

There are several additional considerations that need to be addressed when transitioning networks *in vivo*. To start, one needs to consider putting network units on

plasmids with compatible origins of replication and antibiotic resistances, or integrating networks genomically into the DNA. This limits networks to set copy numbers, which need to be experimentally determined. Additionally, the stability of both the DNA (avoiding hairpins) and of the actual product (toxicity of network components) needs to be considered. For example, while Niederholtmeyer *et al.* found that while TX-TL reactions served as a suitable prototyping environment for complex oscillator networks, cellular toxicity effects were not captured [27]. While high expressing DNA may be best for prototyping *in vitro*, the cellular load induced *in vivo* may be detrimental to cellular health, requiring re-adjustment of transcriptional and translational unit strength [83,84]. Finally, if large timescales are required to test complex networks, devices that emulate cellular steady-state behavior with TX-TL reactions, can be used to more accurately mimic the *in vivo* environment. For example, microfluidic reactors can be used to exchange TX-TL reagents at dilution rates that match the rates of dividing bacteria [27,56].



**Figure 5.** Comparing *in vitro* and *in vivo* response times. (A) Schematic of a modification to the RNA transcriptional cascade in Figure 2E. L1 of the cascade was modified to contain tandem attenuators (Att-1-Att-1) controlling SFGFP production. L3 was replaced with AS-2 fused to a theophylline aptamer (AS-2-theo) from Qi *et al.* [79]. AS-2-theo is only active in the presence of theophylline. (B) TX-TL (*in vitro*) spike experiment. L1 (Att-1- Att-1, 0.5 nM) + L2 (4 nM) + aptamer L3 (14 nM) reaction was setup for 20 min at which point theophylline (final concentration 2 mM, purple curve) or ddH<sub>2</sub>O (red curve) was spiked into the reaction and time reset to 0. Plot shows normalized fluorescence curves combining three separate experiments performed at 37°C with a total of 9 replicates over multiple days. Inset shows that the response time of the circuit to the addition of theophylline ( $\tau = 45.2 \pm 11.7$  min) is slower than the response time from a DNA spike due to the aptamer antisense used for L3 [28]. Shaded regions represent standard deviations calculated at each time point. (C) *In vivo* spike experiment. An extension of the cascade in A where each level of the cascade was encoded on a separate plasmid, with L1 containing both Att-1-RFP and Att-1-Att-1-SFGFP, and co-transformed into *E. coli* TG1 cells. Cultures

were grown to exponential growth, and then split. Theophylline was added to one of the split cultures once in logarithmic growth at which point time was set to zero. Plot shows normalized fluorescence curves with (+) and without (-) theophylline at 2 mM. The response time,  $\tau$ , for Att-1-Att-1-SFGFP was calculated by determining the time at which the (+) and (-) curves were statistically different ( $\tau = 40.0 \pm 9.5$  min). Shaded regions represent standard deviations calculated from 12 biological replicates at each time point. Figure adapted from Takahashi et al. ACS Synth. Biol., 4 (2015) 503-515 [28].

## 6. Catalog of parts tested

A variety of parts and networks have been demonstrated to function in TX-TL, showing the versatility of the platform. The following table is a catalog of parts/networks that have been tested in TX-TL with references for characterization information.

Part	Description	References
Reporter protein	Used to measure gene expression or network output. (eGFP, deGFP, deCFP, Luciferase)	[23,47,49,71,76]
SsrA-tagged reporter protein	SsrA tag marks proteins for degradation by ClpXP (other degradation tags were also tested). ClpXP protein or DNA must be added to TX-TL reactions separately. (eGFP-ssrA)	[23,71]
<i>E. coli</i> sigma factors	Alternative transcriptional activation units. ( $\sigma^{19}$ , $\sigma^{24}$ , $\sigma^{28}$ , $\sigma^{32}$ , $\sigma^{38}$ , $\sigma^{54}$ , NtrC).	[49]
Protein transcriptional repressors and activators	Allows for inducible transcriptional repression or activation. (TetR, LacI, AraC, lambda repressors CI and Cro)	[47,49]
Fusion protein	Provides a fluorescent output to monitor TetR production. (TetR-deGFP)	[22]
Constitutive promoters	Provides constant transcription from the start of the reaction. (Pr, Pr1, Pr2)	[76]
Malachite green aptamer	Fluoresces when the malachite green dye is added to TX-TL reactions. Useful for optically quantifying RNA levels.	[76]

T3, T7 RNA polymerase	Allows for transcription from bacteriophage T3 and T7 promoters.	[49,76]
RNA transcriptional attenuator	Antisense RNA mediated transcriptional repressor.	[28]
Small transcription activating RNAs (STARs)	Small RNAs that activate the transcription of a specific gene regulated by a terminator. (T181, AD1)	[85]

Network	Description	References
Bistable switch	Switches between two states regulated by the lactose and tetracycline regulated promoters.	[49]
Ring oscillator	Three, four, and five-node protein mediated ring oscillators.	[27]
RNA transcription cascade	A three-level RNA transcription cascade.	[28]
RNA Single Input Module	Dynamically controls the expression of two genes with one master regulator.	[28]

## 7. Applications

Cell-free protein synthesis was originally developed to address fundamental questions related to the genetic code and transcription regulation [86,87]. In the last twenty years, this technology has become powerful and versatile enough to be brought to the industrial level [88] for the production of gram quantities of proteins in huge reactions volumes (liters, or hundred of liters). At the same time, cell-free expression has become a popular tool in laboratories where many new applications and assays are created in the research areas of medicine, nanotechnology, cell-free biology, chemical and metabolic engineering [32,89-91]. Cell-free TX-TL reactions can also be carried out

in cell-sized compartments and in microfluidics devices for high-throughput applications. The first major step in this transition was the production of high yield *E. coli*-based *in vitro* protein synthesis systems. Understanding the complex biochemical processes behind TX-TL was critical to achieve enough protein synthesis in batch mode reactions so that useful applications could be envisioned [88]. Beyond TX-TL molecular machineries, endogenous enzymes in *E. coli* cytoplasmic extracts can be used to energize cell-free reactions and reach concentrations of one or a few milligrams per milliliter of soluble proteins [35]. A considerable reduction of the reaction mixture and procedure costs was also achieved, which was essential for commercialization and large-scale production.

This work fostered the development of new methods to energize cell-free reactions. Glycolytic intermediates and carbon sources are used to sustain energy regeneration for longer time periods and help recycle inhibitory byproducts of reactions [37,67,68]. This area of research is advancing rapidly, with new energy mixtures for *E. coli* cell-free TX-TL systems proposed almost every year. Yet, these platforms are limited to cytoplasmic bacterial proteins that do not necessitate specific folding conditions or post-translational modification. Another major step was to demonstrate that complex eukaryotic proteins containing disulfide bonds can be produced in large amounts using *E. coli* extracts [92], leading to the large scale production of antigen and antibodies using cell-free expression. At least two companies have specialized in cell-free protein synthesis for medicine and novel therapeutics, Sutro Biopharma Inc ([www.sutro.bio.com](http://www.sutro.bio.com)) and Exix Bio ([www.exixbio.com](http://www.exixbio.com)). As boundary-free systems, screening of hundreds of protein variants with cell-free systems is another promising avenue to personalized medicine.

In the past decade, cell-free TX-TL has become within the reach of individual research labs. As a result, a myriad of applications have been created that do not require industrial levels of production or large reaction volumes. For example, protein evolution, developed in the 90s, has taken advantage of high yield cell-free systems as well as miniaturization and high-throughput techniques [93]. On the medical side, cell-free TX-TL systems are becoming useful for the synthesis of vaccine candidates, for the identification of drug candidates and for diagnostics [90]. The recent development of

paper-based cell-free TX-TL gene networks to make Ebola sensors and other diagnostics [48] is a particularly intriguing example of this new direction, which promises to create powerful, low-cost diagnostics with high societal impact.

TX-TL systems are also being shown to be able to produce far more than single proteins. A particularly striking example of this is the complete synthesis of functional bacteriophages, such as T7, using only TX-TL reactions and viral genomic DNAs. The T7 genome contains about 60 protein coding genes that encode for viral DNA replication and assembly, and the synthesis of infective T7 virions in TX-TL reactions demonstrates that genome-sized networks and complex self-assembled processes can be achieved outside living cells [94]. In a reverse engineering approach, using robust cell-free gene networks could help construct purely synthetic nanomachines or new materials from natural parts [95,96]. Cell-free TX-TL systems are also expected to expand the scope of applications in chemical and metabolic engineering. For example, cell-free platforms are now used to optimize biosynthetic pathways for the production of therapeutics and fuels [97,98].

### *7.1 Cell-free TX-TL systems as an educational platform*

Recently, the flexibility and ease-of-use of TX-TL systems have been leveraged in an educational setting. Specifically, we brought the TX-TL characterization system to the inaugural Cold Spring Harbor Synthetic Biology Summer Course (CSHL SynBio) to teach aspiring synthetic biologists the principles of genetic network design in the context of addressing real research problems. The results were a resounding success – we were able to teach four students who had little to no experience in wet lab synthetic biology the basics of TX-TL reactions in a matter of days. By the end of the two week course, they had performed many of the preliminary experiments that lead to the results presented in Section 3.1 (Figure 2) [28]. This remarkable success has continued, with an expansion of the use of TX-TL reactions in CSHL SynBio, which is a testament to the robustness of the TX-TL platform. By decoupling genetic network characterization from cell growth, students can rapidly test their own hypotheses and learn about synthetic biology through hands on, immersive design-build-test cycles. We believe that TX-TL systems can also be employed in lower-resource educational settings, since simple

networks can be designed with green fluorescent outputs that are bright enough to be seen with the naked eye under inexpensive blue-light sources.

## 8. Conclusions

Cell-free TX-TL synthetic biology is a rapidly growing research area that spans a wide range of applications, from the development of genetic parts to the construction of complex self-assembled biological systems [99]. Here we have outlined how the simplicity and rapid time scale of TX-TL experiments greatly speeds up the overall design-build-test cycle for engineering genetic networks, thus making it an appealing system for synthetic biology. To help other researchers adopt this powerful platform, we have presented examples and guidelines for using TX-TL reactions to prototype both RNA and protein genetic parts and networks. However, many of the guidelines will be useful for all TX-TL applications. We anticipate an acceleration in the use of TX-TL systems for prototyping and characterizing genetic networks, as well as a whole host of new applications that will emerge from this powerful and flexible technology.

## Funding

This work was supported in part by: the National Science Foundation Graduate Research Fellowship Program [Grant No. DGE-1144153 to MKT]. The Defense Advanced Research Projects Agency Young Faculty Award (DARPA YFA) [N66001-12-1-4254 to JBL]. The Office of Naval Research Young Investigators Program Award (ONR YIP) [N00014-13-1-0531 to JBL]. The Office of Naval Research [N00014-13-1-0074 to VN]. The Defense Advanced Research Projects Agency (DARPA/MTO) Living Foundries program [contract number HR0011-12-C-0065 to CAH, ZZS, RMM]. JBL is an Alfred P. Sloan Research Fellow. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies, either expressly or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

## Acknowledgements

VN thanks Ryan Marshall for providing data in Figure 1. We thank the students of the



2013 Cold Spring Harbor Laboratory Synthetic Biology course (Vipul Singhal, Kevin J. Spring, Shaima Al-Khabouri, and Christopher P. Fall) for carrying out some of the experiments in Figure 2. We also thank David Savage for design and performance of experiments regarding the autofluorescence of the TX-TL extract.

## References

- [1] J.D. Keasling, Synthetic Biology for Synthetic Chemistry, ACS Chem. Biol. 3 (2008) 64–76. doi:10.1021/cb7002434.
- [2] J.M. Carothers, J.A. Goler, J.D. Keasling, Chemical synthesis using synthetic biology, Curr. Opin. Biotechnol. 20 (2009) 498–503. doi:10.1016/j.copbio.2009.08.001.
- [3] Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, Multi-input RNAi-based logic circuit for identification of specific cancer cells, Science. 333 (2011) 1307–1311. doi:10.1126/science.1205527.
- [4] U. Alon, An Introduction to Systems Biology: Design Principles of Biological Circuits, Chapman & Hall/CRC, 2007.
- [5] P.E.M. Purnick, R. Weiss, The second wave of synthetic biology: from modules to systems, Nat. Rev. Mol. Cell Biol. 10 (2009) 410–422. doi:10.1038/nrm2698.
- [6] D.E. Cameron, C.J. Bashor, J.J. Collins, A brief history of synthetic biology, Nat. Rev. Microbiol. 12 (2014) 381–390. doi:10.1038/nrmicro3239.
- [7] J.C. Way, J.J. Collins, J.D. Keasling, P.A. Silver, Integrating Biological Redesign: Where Synthetic Biology Came From and Where It Needs to Go, Cell. 157 (2014) 151–161. doi:10.1016/j.cell.2014.02.039.
- [8] A.S. Khalil, J.J. Collins, Synthetic biology: applications come of age, Nat Rev Genet. 11 (2010) 367–379. doi:10.1038/nrg2775.
- [9] H.M. Salis, E.A. Mirsky, C.A. Voigt, Automated design of synthetic ribosome binding sites to control protein expression, Nat Biotechnol. 27 (2009) 946–950. doi:10.1038/nbt.1568.
- [10] J.M. Carothers, J.A. Goler, D. Juminaga, J.D. Keasling, Model-Driven Engineering of RNA Devices to Quantitatively Program Gene Expression, Science. 334 (2011) 1716–1719. doi:10.1126/science.1212209.
- [11] M. Wachsmuth, S. Findeiß, N. Weissheimer, P.F. Stadler, M. Mörl, De novo design of a synthetic riboswitch that regulates transcription termination, Nucleic Acids Research. 41 (2013) 2541–2551. doi:10.1093/nar/gks1330.
- [12] G. Rodrigo, A. Jaramillo, RiboMaker: computational design of conformation-based riboregulation, Bioinformatics. 30 (2014) 2508–2510. doi:10.1093/bioinformatics/btu335.
- [13] A.A. Green, P.A. Silver, J.J. Collins, P. Yin, Toehold Switches: De-Novo-Designed Regulators of Gene Expression, Cell. (2014) 1–15. doi:10.1016/j.cell.2014.10.002.
- [14] C. Engler, R. Kandzia, S. Marillonnet, A one pot, one step, precision cloning method with high throughput capability, PLoS ONE. 3 (2008) e3647. doi:10.1371/journal.pone.0003647.



- [15] D.G. Gibson, L. Young, R.-Y. Chuang, J.C. Venter, C.A. Hutchison, H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat. Methods*. 6 (2009) 343–345. doi:10.1038/nmeth.1318.
- [16] S. Kosuri, N. Eroshenko, E.M. LeProust, M. Super, J. Way, J.B. Li, et al., Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips, *Nat Biotechnol*. 28 (2010) 1295–1299. doi:10.1038/nbt.1716.
- [17] E. Weber, C. Engler, R. Gruetzner, S. Werner, PLOS ONE: A Modular Cloning System for Standardized Assembly of Multigene Constructs, *PLoS ONE*. 6 (2011) e16765. doi:10.1371/journal.pone.0016765.g009.
- [18] S. Kosuri, G.M. Church, Large-scale de novo DNA synthesis: technologies and applications, *Nat. Methods*. 11 (2014) 499–507. doi:10.1038/nmeth.2918.
- [19] P. Coussement, J. Maertens, J. Beauprez, W. Van Bellegem, M. De Mey, One step DNA assembly for combinatorial metabolic engineering, *Metabolic Engineering*. 23 (2014) 70–77. doi:10.1016/j.ymben.2014.02.012.
- [20] J.P. Torella, F. Lienert, C.R. Boehm, J.-H. Chen, J.C. Way, P.A. Silver, Unique nucleotide sequence-guided assembly of repetitive DNA parts for synthetic biology applications, *Nat Protoc*. 9 (2014) 2075–2089. doi:10.1038/nprot.2014.145.
- [21] E. Appleton, J. Tao, T. Haddock, D. Densmore, Interactive assembly algorithms for molecular cloning, *Nat. Methods*. 11 (2014) 657–663. doi:10.1038/NMETH.2939.
- [22] Z.Z. Sun, C.A. Hayes, J. Shin, F. Caschera, R.M. Murray, V. Noireaux, Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology, *Journal of Visualized Experiments*. (2013). doi:10.3791/50762.
- [23] J. Shin, V. Noireaux, Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70, *J Biol Eng*. 4 (2010) 8. doi:10.1186/1754-1611-4-8.
- [24] D.K. Karig, S. Iyer, M.L. Simpson, M.J. Doktycz, Expression optimization and synthetic gene networks in cell-free systems, *Nucleic Acids Research*. 40 (2012) 3763–3774. doi:10.1093/nar/gkr1191.
- [25] J. Chappell, K. Jensen, P.S. Freemont, Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology, *Nucleic Acids Research*. 41 (2013) 3471–3481. doi:10.1093/nar/gkt052.
- [26] V. Noireaux, R. Bar-Ziv, A. Libchaber, Principles of cell-free genetic circuit assembly, *Proc Natl Acad Sci USA*. 100 (2003) 12672–12677. doi:10.1073/pnas.2135496100.
- [27] H. Niederholtmeyer, Z. Sun, Y. Hori, E. Yeung, A. Verpoorte, R.M. Murray, et al., A cell-free framework for biological systems engineering, *bioRxiv*, 2015. doi:10.1101/018317.
- [28] M.K. Takahashi, J. Chappell, C.A. Hayes, Z.Z. Sun, J. Kim, V. Singhal, et al., Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription-Translation (TX-TL) Systems, *ACS Synth. Biol*. 4 (2015) 503–515. doi:10.1021/sb400206c.
- [29] M.W. Nirenberg, J.H. Matthaei, O.W. Jones, R.G. Martin, Approximation of genetic code via cell-free protein synthesis directed by template RNA,

- Federation Proceedings Genetic Mechanisms. 22 (1963) 55–61.
- [30] D.-M. Kim, J.R. Swartz, Prolonging cell-free protein synthesis with a novel ATP regeneration system, *Biotechnol Bioeng.* 66 (1999) 180–188. doi:10.1002/(SICI)1097-0290(1999)66:3<180::AID-BIT6>3.0.CO;2-S.
- [31] T. Kigawa, T. Yabuki, Y. Yoshida, M. Tsutsui, Y. Ito, T. Shibata, et al., Cell-free production and stable-isotope labeling of milligram quantities of proteins, *FEBS Lett.* 442 (1999) 15–19.
- [32] F. Katzen, G. Chang, W. Kudlicki, The past, present and future of cell-free protein synthesis, *Trends in Biotechnology.* 23 (2005) 150–156. doi:10.1016/j.tibtech.2005.01.003.
- [33] D.-M. Kim, J.R. Swartz, Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis, *Biotechnol Bioeng.* 74 (2001) 309–316. doi:10.1002/bit.1121.
- [34] T.-W. Kim, I.-S. Oh, J.-W. Keum, Y.-C. Kwon, J.-Y. Byun, K.-H. Lee, et al., Prolonged cell-free protein synthesis using dual energy sources: Combined use of creatine phosphate and glucose for the efficient supply of ATP and retarded accumulation of phosphate, *Biotechnol Bioeng.* 97 (2007) 1510–1515. doi:10.1002/bit.21337.
- [35] M.C. Jewett, K.A. Calhoun, A. Voloshin, J.J. Wu, J.R. Swartz, An integrated cell-free metabolic platform for protein production and synthetic biology, *Mol Syst Biol.* 4 (2008) 4–10. doi:10.1038/msb.2008.57.
- [36] T.-W. Kim, J.-W. Keum, I.-S. Oh, C.-Y. Choi, C.-G. Park, D.-M. Kim, Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system, *Journal of Biotechnology.* 126 (2006) 554–561. doi:10.1016/j.jbiotec.2006.05.014.
- [37] H.C. Kim, T.W. Kim, D.M. Kim, Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source, *Process Biochemistry.* 46 (2011) 1366–1369. doi:10.1016/j.procbio.2011.03.008.
- [38] N. Michel-Reydellet, K. Calhoun, J. Swartz, Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome, *Metabolic Engineering.* 6 (2004) 197–203. doi:10.1016/j.ymben.2004.01.003.
- [39] N. Michel-Reydellet, K. Woodrow, J. Swartz, Increasing PCR fragment stability and protein yields in a cell-free system with genetically modified *Escherichia coli* extracts, *J. Mol. Microbiol. Biotechnol.* 9 (2005) 26–34. doi:10.1159/000088143.
- [40] J.-H. Ahn, H.-S. Chu, T.-W. Kim, I.-S. Oh, C.-Y. Choi, G.-H. Hahn, et al., Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions, *Biochemical and Biophysical Research Communications.* 338 (2005) 1346–1352. doi:10.1016/j.bbrc.2005.10.094.
- [41] D.V. Liu, J.F. Zawada, J.R. Swartz, Streamlining *Escherichia coli* S30 Extract Preparation for Economical Cell-Free Protein Synthesis, *Biotechnol. Prog.* 21 (2008) 460–465. doi:10.1021/bp049789y.
- [42] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, et al., Cell-free translation reconstituted with purified components, *Nat Biotechnol.* 19 (2001) 751–755. doi:10.1038/90802.

- [43] P. Shrestha, M.T. Smith, B.C. Bundy, Cell-free unnatural amino acid incorporation with alternative energy systems and linear expression templates, *New Biotechnology*. 31 (2014) 28–34. doi:10.1016/j.nbt.2013.09.002.
- [44] S.H. Hong, Y.C. Kwon, M.C. Jewett, Non-standard amino acid incorporation into proteins using *Escherichia coli* cell-free protein synthesis, *Chemical Biology*. 2 (2014) 1–7. doi:10.3389/fchem.2014.00034.
- [45] B.C. Bundy, J.R. Swartz, Site-Specific Incorporation of p-Propargyloxyphenylalanine in a Cell-Free Environment for Direct Protein–Protein Click Conjugation, *Bioconjugate Chemistry*. 21 (2010) 255–263. doi:10.1021/bc9002844.
- [46] A.R. Goerke, J.R. Swartz, High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids, *Biotechnol Bioeng*. 102 (2009) 400–416. doi:10.1002/bit.22070.
- [47] Z.Z. Sun, E. Yeung, C.A. Hayes, V. Noireaux, R.M. Murray, Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* based TX-TL cell-free system, *ACS Synth. Biol*. 3 (2014) 387–397. doi:10.1021/sb400131a.
- [48] K. Pardee, A.A. Green, T. Ferrante, D.E. Cameron, A. DaleyKeyser, P. Yin, et al., Paper-Based Synthetic Gene Networks, *Cell*. 159 (2014) 940–954. doi:10.1016/j.cell.2014.10.004.
- [49] J. Shin, V. Noireaux, An *E. coli* Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells, *ACS Synth. Biol*. 1 (2012) 29–41. doi:10.1021/sb200016s.
- [50] E. Karzbrun, A.M. Tayar, V. Noireaux, R.H. Bar-Ziv, Programmable on-chip DNA compartments as artificial cells, *Science*. 345 (2014) 829–832. doi:10.1126/science.1255550.
- [51] J. Shin, V. Noireaux, Research Study of messenger RNA inactivation and protein degradation in an *Escherichia coli* cell-free expression system, *J Biol Eng*. 4 (2010) 1–9.
- [52] Y.-C. Kwon, M.C. Jewett, High-throughput preparation methods of crude extract for robust cell-free protein synthesis, *Scientific Reports*. 5 (2015) 1–8. doi:10.1038/srep08663.
- [53] K. Sitaraman, D. Esposito, G. Klarmann, S.F. Le Grice, J.L. Hartley, D.K. Chatterjee, A novel cell-free protein synthesis system, *Journal of Biotechnology*. 110 (2004) 257–263. doi:10.1016/j.jbiotec.2004.02.014.
- [54] Z.Z. Sun, J. Kim, V. Singhal, R.M. Murray, Protein degradation in a TX-TL cell free expression system using ClpXP protease, [http://www.cds.caltech.edu/~murray/papers/2014g\\_sksm14-clpxp.html](http://www.cds.caltech.edu/~murray/papers/2014g_sksm14-clpxp.html). (2014).
- [55] H. Niederholtmeyer, L. Xu, S.J. Maerkl, Real-Time mRNA Measurement during an in Vitro Transcription and Translation Reaction Using Binary Probes, *ACS Synth. Biol*. 2 (2013) 411–417. doi:10.1021/sb300104f.
- [56] H. Niederholtmeyer, V. Stepanova, S.J. Maerkl, Implementation of cell-free biological networks at steady state, *Proc Natl Acad Sci USA*. 110 (2013) 15985–15990. doi:10.1073/pnas.1311166110.
- [57] J.B. Lucks, L. Qi, V.K. Mutalik, D. Wang, A.P. Arkin, Versatile RNA-sensing transcriptional regulators for engineering genetic networks, *Proc Natl Acad Sci USA*. 108 (2011) 8617–8622. doi:10.1073/pnas.1015741108.

- [58] R.P. Novick, S. Iordanescu, S.J. Projan, J. Kornblum, I. Edelman, pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator, *Cell*. 59 (1989) 395–404.
- [59] S. Brantl, E.G. Wagner, Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181, *Mol Microbiol*. 35 (2000) 1469–1482.
- [60] J.-D. Pédelacq, S. Cabantous, T. Tran, T.C. Terwilliger, G.S. Waldo, Engineering and characterization of a superfolder green fluorescent protein, *Nat Biotechnol*. 24 (2006) 79–88. doi:10.1038/nbt1172.
- [61] J.M. Flynn, S.B. Neher, Y.I. Kim, R.T. Sauer, T.A. Baker, Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals, *Mol. Cell*. 11 (2003) 671–683.
- [62] N. Rosenfeld, U. Alon, Response delays and the structure of transcription networks, *J Mol Biol*. 329 (2003) 645–654.
- [63] S. Mangan, U. Alon, Structure and function of the feed-forward loop network motif, *Proc Natl Acad Sci USA*. 100 (2003) 11980–11985. doi:10.1073/pnas.2133841100.
- [64] D.N. Arnosti, M.J. Chamberlin, Secondary sigma factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*, *Proc Natl Acad Sci USA*. 86 (1989) 830–834.
- [65] R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements, *Nucleic Acids Research*. 25 (1997) 1203–1210.
- [66] P. Shrestha, T.M. Holland, Streamlined extract preparation for *Escherichia coli*-based cell-free protein synthesis by sonication or bead vortex mixing, *Biotechniques*. (2012) 163–174. doi:10.2144/0000113924.
- [67] F. Caschera, V. Noireaux, Synthesis of 2.3 mg/ml of protein with an all *Escherichia coli* cell-free transcription–translation system, *Biochimie*. (2014) 162–168. doi:10.1016/j.biochi.2013.11.025.
- [68] Y. Wang, Y.-H.P. Zhang, Cell-free protein synthesis energized by slowly-metabolized maltodextrin, *BMC Biotechnology*. 9 (2009) 58. doi:10.1186/1472-6750-9-58.
- [69] A. Spirin, V. Baranov, L. Ryabova, S. Ovodov, Y. Alakhov, A continuous cell-free translation system capable of producing polypeptides in high yield, *Science*. 242 (1988) 1162–1164. doi:10.1126/science.3055301.
- [70] F. Caschera, V. Noireaux, A cost-effective polyphosphate-based metabolism fuels an all *E. coli* cell-free expression system, *Metabolic Engineering*. 27 (2015) 29–37. doi:10.1016/j.ymben.2014.10.007.
- [71] E. Karzbrun, J. Shin, R.H. Bar-Ziv, V. Noireaux, Coarse-Grained Dynamics of Protein Synthesis in a Cell-Free System, *Phys Rev Lett*. 106 (2011) 048104. doi:10.1103/PhysRevLett.106.048104.
- [72] D. Siegal-Gaskins, V. Noireaux, R.M. Murray, Biomolecular resource utilization in elementary cell-free gene circuits, in: 2013 American Control Conference (ACC), IEEE, 2013: pp. 1531–1536. doi:10.1109/ACC.2013.6580053.
- [73] C.Y. Hu, J.D. Varner, J.B. Lucks, Generating effective models and parameters for RNA genetic circuits, *bioRxiv*, 2015. doi:10.1101/018358.
- [74] S.B. Zimmerman, S.O. Trach, Estimation of macromolecule concentrations and

- excluded volume effects for the cytoplasm of *Escherichia coli*, *J Mol Biol.* 222 (1991) 599–620. doi:10.1016/0022-2836(91)90499-V.
- [75] K. Fujiwara, S.-I.M. Nomura, Condensation of an additive-free cell extract to mimic the conditions of live cells, 8 (2013) e54155. doi:10.1371/journal.pone.0054155.
- [76] D. Siegal-Gaskins, Z.A. Tuza, J. Kim, V. Noireaux, R.M. Murray, Gene Circuit Performance Characterization and Resource Usage in a Cell-Free “Breadboard,” *ACS Synth. Biol.* 3 (2014) 416–425. doi:10.1021/sb400203p.
- [77] A. Martin, T.A. Baker, R.T. Sauer, Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines, *Nature.* 437 (2005) 1115–1120. doi:10.1038/nature04031.
- [78] C. Walsh, *Enzymatic reaction mechanisms*, W. H. Freeman, 1979.
- [79] L. Qi, J.B. Lucks, C.C. Liu, V.K. Mutalik, A.P. Arkin, Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals, *Nucleic Acids Research.* 40 (2012) 5775–5786. doi:10.1093/nar/gks168.
- [80] J.C. Anderson, Anderson promoter library registry of standard biological parts, <http://partsregistry.org/Promoters/Catalog/Anderson>. (2006).
- [81] J.H. Davis, A.J. Rubin, R.T. Sauer, Design, construction and characterization of a set of insulated bacterial promoters, *Nucleic Acids Research.* 39 (2011) 1131–1141. doi:10.1093/nar/gkq810.
- [82] V.K. Mutalik, J.C. Guimaraes, G. Cambray, C. Lam, M.J. Christoffersen, Q.-A. Mai, et al., Precise and reliable gene expression via standard transcription and translation initiation elements, *Nat. Methods.* 10 (2013) 354–360. doi:10.1038/nmeth.2404.
- [83] R. Algar, T. Ellis, G.-B. Stan, Modelling the burden caused by gene expression: an in silico investigation into the interactions between synthetic gene circuits and their chassis cell, <http://arxiv.org/abs/1309.7798>. (2013).
- [84] A. Gyorgy, D. Del Vecchio, Limitations and trade-offs in gene expression due to competition for shared cellular resources, *Proc of IEEE Conference on Decision and Control*. <http://web.mit.edu/ddv/www/papers/AndrasCDC14.pdf>, (2014).
- [85] J. Chappell, M.K. Takahashi, J.B. Lucks, Creating small transcription activating RNAs, *Nat Chem Biol.* 11 (2015) 214–220. doi:10.1038/nchembio.1737.
- [86] M. Nirenberg, Historical review: Deciphering the genetic code – a personal account, *Trends in Biochemical Sciences.* 29 (2004) 46–54. doi:10.1016/j.tibs.2003.11.009.
- [87] Y. Cenatiempo, Prokaryotic gene expression in vitro: transcription-translation coupled systems, *Biochimie.* 68 (1986) 505–515. doi:10.1016/S0300-9084(86)80195-X.
- [88] J. Swartz, Developing cell-free biology for industrial applications, *J Ind Microbiol Biotechnol.* 33 (2006) 476–485. doi:10.1007/s10295-006-0127-y.
- [89] M. He, Cell-free protein synthesis: applications in proteomics and biotechnology, *New Biotechnology.* 25 (2008) 126–132. doi:10.1016/j.nbt.2008.08.004.
- [90] E.D. Carlson, R. Gan, C.E. Hodgman, M.C. Jewett, Cell-free protein synthesis: Applications come of age, *Biotechnology Advances.* 30 (2012) 1185–1194. doi:10.1016/j.biotechadv.2011.09.016.
- [91] J.F. Zawada, G. Yin, A.R. Steiner, J. Yang, A. Naresh, S.M. Roy, et al.,



- Microscale to manufacturing scale-up of cell-free cytokine production-a new approach for shortening protein production development timelines, *Biotechnol Bioeng.* 108 (2011) 1570–1578. doi:10.1002/bit.23103.
- [92] K.G. Knapp, A.R. Goerke, J.R. Swartz, Cell-free synthesis of proteins that require disulfide bonds using glucose as an energy source, *Biotechnol Bioeng.* 97 (2006) 901–908. doi:10.1002/bit.21296.
- [93] A. Fallah-Araghi, J.-C. Baret, M. Ryckelynck, A.D. Griffiths, A completely in vitro ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution, *Lab on a Chip.* 12 (2012) 882–891. doi:10.1039/C2LC21035E.
- [94] J. Shin, P. Jardine, V. Noireaux, Genome Replication, Synthesis, and Assembly of the Bacteriophage T7 in a Single Cell-Free Reaction, *ACS Synth. Biol.* 1 (2012) 408–413. doi:10.1021/sb300049p.
- [95] S.S. Daube, T. Arad, R. Bar-Ziv, Cell-Free Co-synthesis of Protein Nanoassemblies: Tubes, Rings, and Doughnuts, *Nano Lett.* 7 (2007) 638–641. doi:10.1021/nl062560n.
- [96] S.S. Daube, R.B. Ziv, Protein nanomachines assembly modes: cell-free expression and biochip perspectives, *WIREs Nanomed Nanobiotechnol.* 5 (2013) 613–628. doi:10.1002/wnan.1234.
- [97] M. Bujara, M. Schümperli, R. Pellaux, M. Heinemann, S. Panke, Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis, *Nat Chem Biol.* 7 (2011) 271–277. doi:10.1038/nchembio.541.
- [98] Q.M. Dudley, A.S. Karim, M.C. Jewett, Cell-free metabolic engineering: Biomanufacturing beyond the cell, *Biotechnology Journal.* 10 (2014) 1–14. doi:10.1002/biot.201400330.
- [99] M.C. Jewett, Cell-Free Synthetic Biology Special Issue, *ACS Synth. Biol.* 3 (2014) 332–332. doi:10.1021/sb5002592.