Cell-free systems enable mechanistic characterization of genetically encoded RNA strand exchange circuits for programmable protein expression

Fernanda Piorino[1,2], Eugenia Romantseva[1], and Samuel W. Schaffter[1]*

[1]National Institute of Standards and Technology; Gaithersburg, MD, 20899, USA.

[2]Johns Hopkins University; Baltimore, MD, 21218, USA.

*Correspondence: samuel.schaffter@nist.gov

ABSTRACT

Nucleic acid circuits are powerful tools for programming biology, but the design principles for operating these circuits in complex cellular environments remain poorly understood compared to simple *in vitro* settings. Cell-free expression systems (CFES) are uniquely suited to address this challenge, as they can integrate measurements ideal for either *in vitro* or *in cyto* settings. Further, because CFES are open systems, they enable a level of control over component concentration unattainable in cells. Here, we use cell-free systems to characterize genetically encoded RNA circuits that operate *via* toehold-mediated strand exchange (TMSE). These circuits have been extensively characterized *in vitro* and were recently deployed as translational riboregulators in *E. coli*, revealing key differences between the two environments. By systematically modulating the design parameters of these RNA circuits in both purified protein and lysate-based CFES, we elucidate the mechanisms linking TMSE to protein expression. Further, we combine measurements and alterations in CFES composition that are infeasible in cells to investigate interactions between cellular components and RNA circuit components, identifying a potential interaction with ribosomes that informs circuit design. Our results establish a unified set of principles for designing and operating genetically encoded TMSE circuits across *in vitro*, CFE, and bacterial environments, which should catalyze the widespread adoption of this platform for new applications in molecular programming, synthetic biology, and biotechnology.

INTRODUCTION

Nucleic acid-based biotechnologies and molecular circuits have many applications in synthetic biology, spanning *in vitro* diagnostics to cellular computation. The programmable nature of nucleic acid base pairing interactions enables rational design of biosensors and diagnostics[1–4], therapeutics[5], riboregulators[6–8], and molecular computing networks[9–11]. Further, because base pairing only requires positive ions, nucleic acid technologies have the potential to be deployed across application environments and organisms[12,13].

Nucleic acid circuits based on toehold-mediated strand exchange[14] (TMSE) hold great promise for synthetic biology, particularly for molecular computing applications. TMSE circuits consist of an input strand that interacts with a complementary sequence in a partially double-stranded gate *via* a single-stranded toehold, prompting the release of an output strand that can facilitate new strand exchange reactions (Figure 1A). The modular design of the double-stranded gates enables construction of networks of hundreds of interacting components that can be programmed to execute information processing tasks[15,16]. These programmable TMSE networks have been developed into *in vitro* diagnostics that classify complex patterns of infection and disease[17,18] or precisely discriminate nucleic acid biomarkers[19]. These *in vitro* successes were facilitated by mechanistic understanding of the thermodynamics and kinetics of TMSE reactions in well-defined solutions[14,20].

Cotranscriptionally encoded RNA strand displacement (ctRSD) systems, which operate analogously to TMSE circuits and can be genetically encoded, have been developed to expand the application space of TMSE circuits and characterized in *in vitro* transcription (IVT) environments[21–23]. Recently, ctRSD circuits were successfully deployed in bacteria using a toehold exchange (THE) riboregulator that couples RNA strand exchange to fluorescent protein expression[24] (Figure 1B). However, ctRSD circuits in bacteria exhibited some behaviors that contrasted with the expectations from *in vitro* studies, raising questions about the role of the cellular environment in TMSE. Developing a detailed mechanistic understanding of TMSE reactions in cells remains challenging, in part due to the difficulty in conducting kinetic measurements and the lack of suitable tools for directly measuring strand exchange in cells. Further, the precise control of relative ratios of nucleic acid components necessary for mechanistic characterization of TMSE reactions is difficult to achieve in cells. Thus, there is a gap in our understanding of how the complex cellular environment influences the properties of TMSE reactions and the extent to which design principles elucidated *in vitro* translate to cells.

Cell-free expression systems[25,26] (CFES) are uniquely positioned as platform for characterizing genetic circuits because they combine features of *in vitro* and cellular systems (Figure 1C). Like simple, buffered *in vitro* systems, CFES are easy to operate and customize, enabling direct manipulation of reaction composition. For example, users can add any number of DNA templates encoding genetic circuit components, each at a defined concentration, a flexibility unfeasible in cellular systems. Further, because they are non-living, non-replicating systems, CFES support kinetic measurements with tools unsuitable for cells[3,27]. In contrast to simple *in vitro* environments used to characterize TMSE reactions, such as buffered salt solutions[14] or transcription reactions[21],

CFES can execute both transcription and translation. Further, lysate-based CFES can closely mimic the cellular environment, containing residual nucleic acids, small molecules, and proteins native to the host organism. Together, these characteristics make CFES an appealing platform for rapid prototyping of genetic circuits, both as a testbed for circuits designed for implementation in cells and as the main expression platform[28–31]. In addition, several studies have leveraged CFES to prototype RNA circuits and identified potential characteristics of CFES confounding prototyping efforts and responsible for diverging behaviors across CFES and cellular systems[29,30,32].

Here, we use CFES to mechanistically characterize the properties of genetically encoded RNA strand exchange circuits that regulate protein expression. By combining measurement tools previously used in either *in vitro* systems or cells — but incompatible with both systems —, we simultaneously measure transcription, RNA strand exchange, and translation dynamics of > 25 different circuit designs using sequences from the ctRSD toolkit[21,22]. Using these measurements, we unambiguously characterize how changes in input and gate design influence strand exchange and protein expression. We conduct our measurements in *E. coli*-based CFES in both purified protein and lysate formats, which allows us to elucidate phenomena that appear to be related to cellular environment rather than component design. Additionally, we leverage the open environment of CFES to change their composition and identify a potential interaction between ribosomes and non-coding RNAs that interferes with strand exchange. Together, these results provide a greater understanding of TMSE reactions in the cellular environment and highlight mechanisms for using these reactions to tune protein expression. By bridging the gap between *in vitro*[21] and cellular environments[24], these results also lay the foundation for new applications of TMSE and for future studies exploring interactions between cellular components and genetic circuits.

RESULTS AND DISCUSSION

Study and measurement overview

Cotranscriptionally encoded RNA strand displacement (ctRSD) circuits are composed of single-stranded input RNAs and partially double-stranded gate RNAs that are transcribed from DNA templates (Figure 1B). Gates are initially transcribed as a single RNA strand that folds into a hairpin. An internal self-cleaving ribozyme induces the hairpin transcript to cut, producing a partially doubled-stranded gate suitable for TMSE. After transcription, folding, and ribozyme cleavage, a gate presents a single-stranded input toehold (u') that can recruit a complementary input RNA to bind and initiate strand exchange. Successful strand exchange releases a single-stranded output strand from the gate, which can initiate other strand exchange reactions. A toehold exchange (THE) riboregulator represents a slight modification of a ctRSD gate that couples TMSE to protein expression[24].

The modularity and composability of TMSE gates is enabled by a process of toehold exchange[14], whereby binding of the input toehold on the gate is exchanged for dissociation of the output toehold on the gate (Figure 1A). The desired measurement of a TMSE reaction is the rate of, or extent of, output toehold dissociation from a gate when different input sequences or concentrations are

supplied. In *in vitro* environments, such as buffered solutions or transcription reactions[14,21], a fluorescently labeled molecular beacon (MB) complex is added to measure TMSE; the dissociation of the output toehold from a gate initiates the displacement of a fluorescent strand on the MB (Figure 1B). In bacteria, THE riboregulators enable measurement of TMSE reactions by coupling dissociation of the output toehold to expression of a fluorescent protein[24]; in this scheme, the output toehold of a gate is designed to be a sequestered ribosome binding sequence (*rbs*) such that TMSE initiates translation (Figure 1B). However, it remains untested how well the MB and fluorescent protein measurements correlate for elucidating differences in TMSE reactions. CFES allow us to conduct both measurements simultaneously on the same system; we can easily add the MB to the reactions, which also execute transcription and translation necessary for measuring TMSE with THE riboregulators. Additionally, we can include a fluorogenic RNA aptamer construct downstream of the fluorescent protein coding sequence on THE riboregulators[33–35] to explore the role of transcription kinetics across designs and environments.

We characterized THE riboregulators with the above measurements in both a reconstituted system (PURExpress) and a lysate-based system (extract from BL21 Star (DE3) *E. coli*) (Figure 1C). PURExpress, subsequently termed PURE, is a commercially available mixture of purified proteins that enable transcription and translation, and provides an accessible, well-defined platform for rapid genetic circuit prototyping. PURE has minimal nuclease activity, which limits the role of RNA degradation and enables use of linear DNA templates for expression. Lysates from BL21 Star (DE3), a common host strain for protein expression[36] and CFE[37], retain nucleases and other cellular components after cell lysis, and thus presumably resemble the cellular environment more closely. This strain was also used to test THE riboregulators in *E. coli*[24]. We selected plasmids rather than linear DNA to enable comparison between the two CFES and between CFES and cells[24], with input and gate sequences encoded on separate plasmids for independent control of concentrations. We encoded RNA strand exchange components on high-copy plasmid DNA under constitutive expression from a T7 RNAP promoter, and optimized reaction conditions for our measurements in both CFES (Supplementary Section 2).

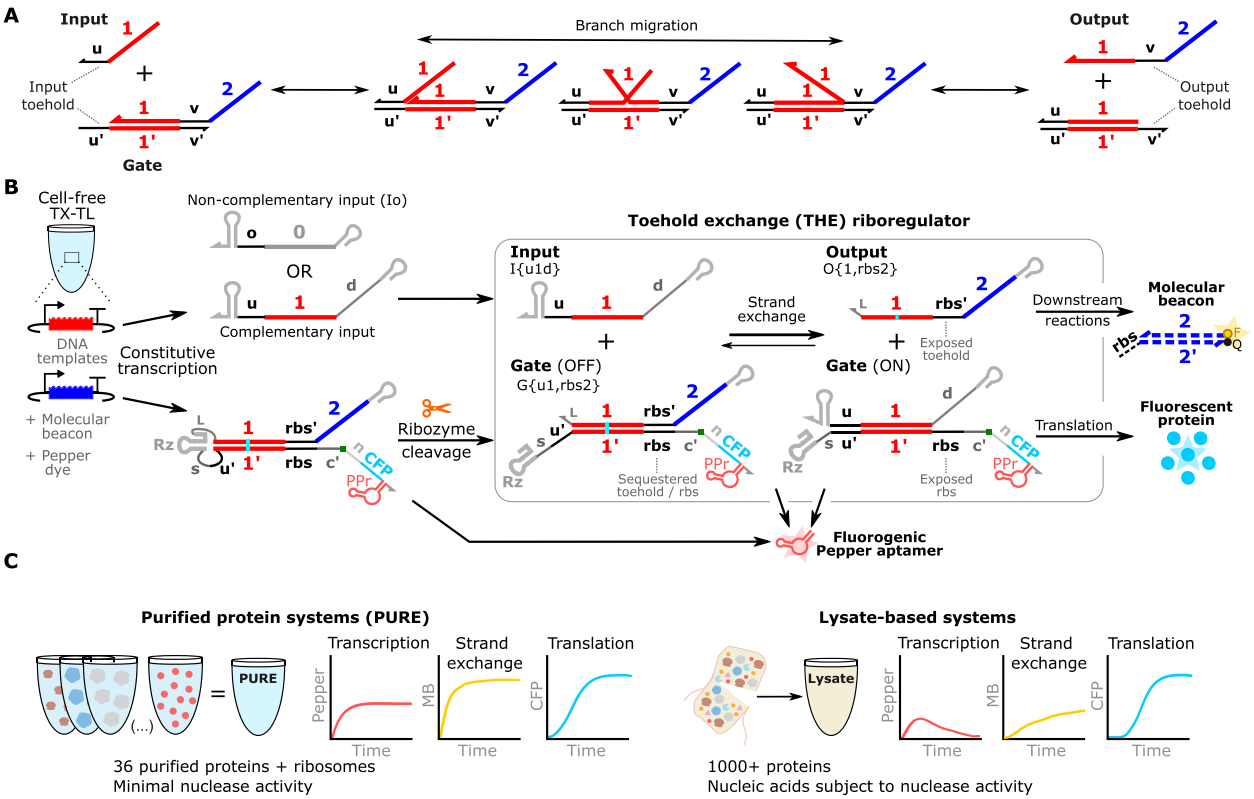

FIGURE 1: Overview of toehold exchange (THE) riboregulators and associated measurements in cell-free expression systems. (**A**) Schematic of a TMSE reaction. (**B**) Schematic of THE riboregulator production, operation, and measurement. The THE riboregulator (gate) and a complementary input are encoded on different plasmids added to a cell-free reaction that supports transcription and translation (TX-TL). Upon transcription, a ribozyme encoded on the gate cleaves the transcript, resulting in a single-stranded region (*u′*) complementary to the toehold on the input strand (*u*), and a double-stranded region including a branch migration domain (*1*) and a sequestered ribosome binding site (*rbs*), thus preventing translation of a downstream cyan fluorescent protein (CFP). The input toehold binds to the gate to initiate strand exchange. The output strand, once dissociated from the gate, can release a fluorophore strand (F) from a previously quenched (Q) molecular beacon complex. Strand exchange also exposes the *rbs*, allowing translation of CFP. The gate also encodes a fluorogenic Pepper RNA aptamer (PPr), which generates a fluorescence signal upon binding to an exogenously added dye. Gate transcripts in the uncleaved, OFF, and ON states contribute to the fluorogenic aptamer signal. A template encoding a non-complementary input, Io, can be added to the reaction to measure leak in the absence of a complementary input, and to maintain the same global transcription levels in reactions containing different concentrations of DNA templates. Additional details on circuit domains and nomenclature are shown in Supplementary Figures 1 and 2. Plasmid schematics are in Supplementary Figure 3. (**C**) Overview of reconstituted (PURE) versus lysate-based cell-free expression systems with expected kinetic trajectories for each measurement.

Validating individual measurements in CFES

Before measuring TMSE, we first conducted control experiments to ensure that each of our intended measurements worked independently in both CFES. To validate MB measurements, we titrated a plasmid expressing the output strand of a THE riboregulator, which should react directly with the MB (Figure 2A). In these and subsequent control experiments, we kept the total DNA concentration constant at 20 nmol/L across samples by adding a template expressing a non-complementary input RNA (Io). A constant transcription load allows for comparison across samples as the overall transcription rate depends on the total DNA template concentration (Supplementary Figures 8-10). As expected, the MB signal increased faster at increasing output plasmid concentrations in both CFES (Figure 2B). In lysate, the rate of MB signal increase was > 2-fold slower than in PURE and output plasmid concentrations below 5 nmol/L were unable to saturate the MB, which could be due to degradation of the output RNA.

We next tested the Pepper and cyan fluorescent protein (CFP) measurements using plasmids encoding hairpin (HP) equivalents of an OFF and ON THE riboregulator (Figure 2C). These constructs serve as best case scenarios for OFF and ON protein expression for a THE riboregulator as HP OFF strongly sequesters the *rbs* and HP ON translates directly after folding. As expected, CFP production was low for HP OFF plasmids and high for HP ON plasmids in both CFES (Figure 2D). Interestingly, the final protein concentration was inversely proportional to HP ON plasmid concentration, perhaps due to faster depletion of energy as the rate of mRNA production increases. In both CFES, the Pepper signal increased with increasing plasmid concentration for both HP OFF and HP ON, although in PURE the differences between the 5 nmol/L and 10 nmol/L HP plasmid concentrations were minimal. In lysate, the Pepper signal decreased steadily after reaching a maximum value, reflecting ribonuclease-mediated degradation of transcripts. In these control experiments, both MB and Pepper are indicators of transcription rates, but Pepper took ~2 times longer than MB to reach 0.5 µmol/L. This lag could be due to Pepper's location at the end of the transcript, slow aptamer folding, or slow dye binding (Figure 2B vs. 2D).

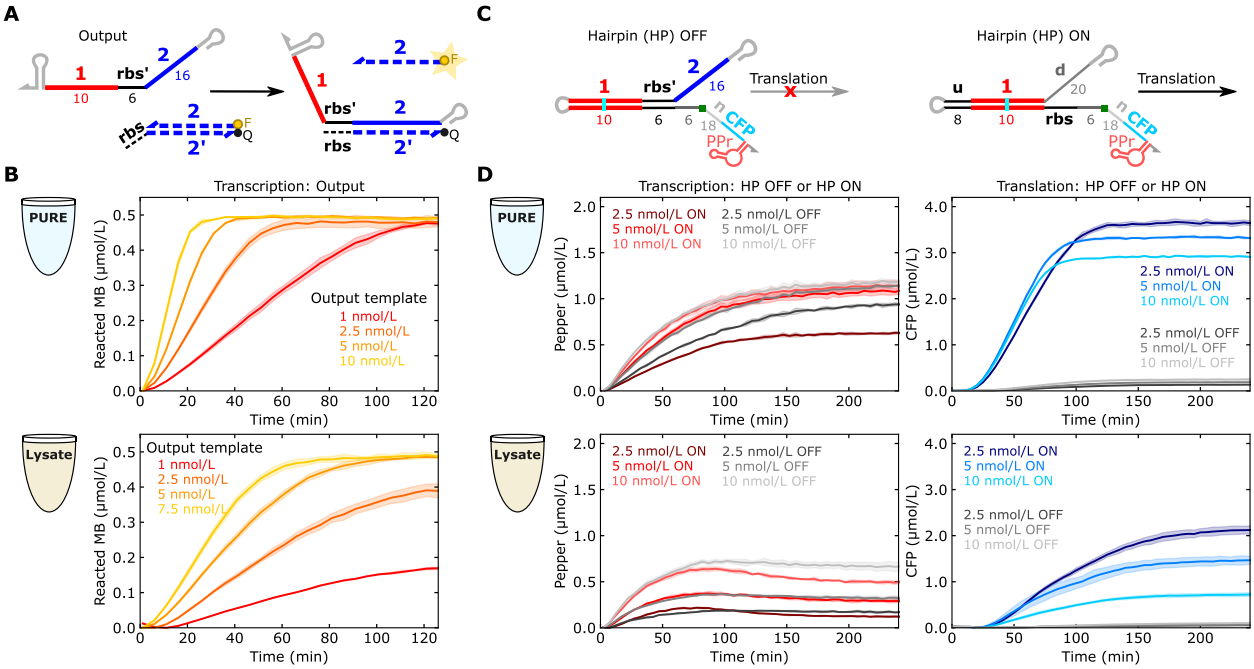

FIGURE 2: Characterizing measurements with DNA templates expressing control constructs. (**A**) Schematic of a control for molecular beacon measurements in CFE systems. An output RNA that reacts with the molecular beacon was constitutively transcribed alongside 500 nmol/L of molecular beacon complex. (**B**) Reacted molecular beacon kinetics for varying concentrations of output template DNA in PURE (top) or lysate (bottom). (**C**) Schematics of translation controls for fluorogenic aptamer and fluorescent protein measurements. HP OFF and HP ON were transcribed separately to determine minimum and maximum translation levels for a THE riboregulator. Domain lengths (non-bold numbers below schematics) were chosen to be consistent with previous ctRSD gate designs[21,22], while also maximizing the predicted translation rate for the ON state[24]. (**D**) Pepper and CFP signals at different template concentrations of HP OFF and HP ON in PURE (top) or lysate (bottom). In (C) and (E), the total DNA template concentration was 20 nmol/L, with Io making up the difference from the concentrations shown in the plots, and the shaded regions indicate the standard deviation of three technical replicates. Additional experiments at lower total template concentrations are shown in Supplementary Figures 9 and 10.

Tuning protein expression by modulating transcription or translation strength

Having validated each individual measurement in both CFES, we next applied these measurements to characterize the full RNA strand exchange process. Specifically, we tested different equimolar concentrations of input and gate templates (Figure 3A) using either a complementary input (I{u1d}, ON state) or a non-complementary input, (Io, OFF state). As expected, we observed that increasing concentrations of input and gate templates increased the rate of production across all measurements (Figure 3B). In PURE, MB and CFP measurements showed similar kinetics and CFP levels followed a similar trend to the HP ON control, with increasing gate concentrations resulting in faster CFP production but lower final CFP concentration. However, in lysate the MB signal was indistinguishable between ON and OFF states, and the CFP trends contrasted the HP ON control trends, with increasing gate concentrations resulting in higher CFP production levels. These results suggest that RNA strand exchange may be inefficient in lysate, resulting in a low concentration of gate in the ON state and thus weak protein production. Notably, CFP production in lysate was greater for a THE riboregulator in the OFF state than HP OFF, which could indicate incorrect folding of the gate that increases leak and inhibits RNA strand exchange. We did not measure ribozyme cleavage efficiency directly but gates with a mutated ribozyme incapable of cleavage produced low protein concentration in the ON and OFF states in both CFES (Supplementary Figure 11). Additionally, we observed similar performance in both CFES for THE riboregulators with an alternative input toehold sequence (*v*), an alternative ribozyme sequence (*Rh*), and an alternative input branch migration sequences (domain *4*) (Supplementary Figure 12). Together, these results demonstrate tuning protein expression levels with RNA strand exchange by modulating the transcription rates of components (Figure 3B, Supplementary Figure 13C).

We next explored tuning protein expression by modulating the translation strength of a THE riboregulator at a constant transcription level. Translation strength is typically tuned by changing the complementarity of the *rbs* to *E. coli* 16S rRNA[38]. However, changing the sequence of an mRNA has disadvantages, as minor changes can alter RNA folding[39] to influence protein expression in difficult-to-predict ways. THE riboregulators offer a method to modulate translation strength without changing the sequence of the protein-encoding mRNA by instead changing the length of a single-stranded domain, termed *d*, upstream of the branch migration sequence in the input (Figure 3C). The *d* domain is hypothesized to serve as a standby site for ribosomes, thereby influencing ribosome binding and translation initiation[40–42]. In both PURE and lysate, decreasing the *d* domain length from 20 bases to 1 base decreased the rate of protein production (Figure 3D), consistent with results in *E. coli*[24]. Pepper measurements indicated global transcription rates were similar across samples and MB measurements in PURE indicated RNA strand exchange was unchanged for inputs with different *d* domain lengths (Figure 3D), suggesting protein expression differences were primarily due to changes in translation strength of the ON gate.

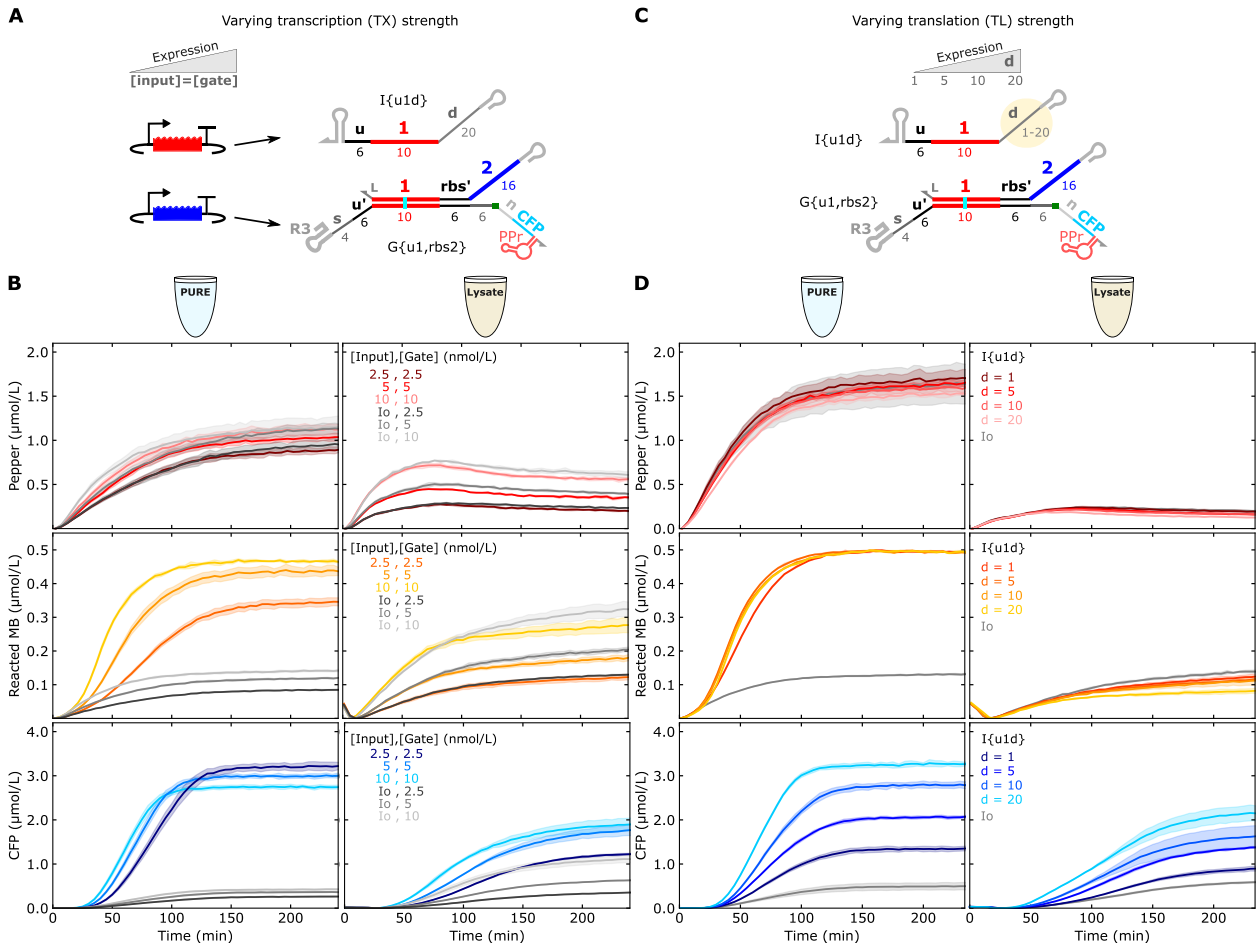

FIGURE 3: Tuning protein expression by modulating transcription or translation strength. (**A**) Transcription strength was modulated by changing the concentrations of input and gate templates. (**B**) Kinetics of the system in (A) in PURE (left) and lysate (right). Input and gate DNA template concentrations are labeled in each plot with Io added to keep a constant 20 nmol/L total DNA template across samples. Additional conditions with unequal template concentrations are shown in Supplementary Figure 13. (**C**) Translation strength was varied by changing the length of the single stranded *d* domain on the input. This domain serves as an unstructured ribosome standby site that enhances translation initiation[41], allowing translation strength to be controlled without changing the *rbs* sequence. (**D**) Kinetics of the system in (C) in PURE (left) and lysate (right). In PURE, the input and riboregulator template concentrations were both 5 nmol/L. In lysate, the input template concentration was 5 nmol/L and the riboregulator template concentration was 2.5 nmol/L. In (B) and (D), the shaded regions indicate the standard deviation of three technical replicates.

Tuning protein expression by modulating RNA strand exchange rates

TMSE offers kinetic control of reactions by altering the lengths of input and output toeholds. This capability is crucial for engineering circuits that execute logic, subtraction, and timing[15,16,43,44]. TMSE can be viewed as a three-step process whereby an input strand can reversibly bind to a gate and undergo a random walk of branch migration. Once the input strand nears the end of the branch migration domain, the output strand can reversibly unbind from the gate[20] (Figure 4A). By modulating the relative lengths of the input and output toeholds of a strand exchange reaction, the rate of output production can be controlled[14,20]. We asked whether modulating toehold lengths could be used to tune protein expression from THE riboregulators. We started by increasing the degree of complementarity (*via* the *s′* domain) between the input strand toehold and the gate, which should decrease the input strand dissociation rate and result in faster strand exchange (Figure 4B, left). In PURE, output RNA production, as measured by the MB, increased with increasing input toehold complementarity (Figure 4C, left). As expected, the rates of protein production increased with increasing input toehold complementarity in both CFES (Figure 4C,D, left). In addition to changing the length of complementarity with the input toehold, the length of the spacer (*s*) domain on the gates should modulate strand exchange kinetics (Figure 4B, right). In IVT experiments, decreasing the length of the spacer from 4 bases to 0 bases decreased the strand exchange rate constant by nearly two orders of magnitude[21,22], likely due to steric hindrance[45,46]. Decreasing spacer length on a THE riboregulator greatly reduced the rate of protein production in both CFES (Figure 4C,D, right).

We next explored how increasing the length of the output toehold influenced strand exchange and protein expression kinetics. We designed a gate with a double-stranded branch migration domain extended by 6 base pairs (*e:e′* domains). If used with an input strand without an *e* domain, a total of 12 bases (*e+rbs′*) would have to dissociate after branch migration for the output to be released and protein translation to initiate. By designing input strands with increasing *e* lengths, the length of the gate output toehold can be modulated between 6 bases and 12 bases (Figure 4E, left).

Modulating the length of the output toehold had different effects on protein expression kinetics in PURE and lysate. In PURE, a gate with a 12-base output toehold showed minimal output release and protein expression, and decreasing the length of the output toehold increased the rates of strand exchange and protein expression (Figure 4F, left). In lysate, however, the extended gate had high protein expression even with a non-complementary input (Io), which masked any trend in protein expression with increasing output toehold length (Figure 4G, left). This high level of leak was not observed in *E. coli*[24], which used an extended gate with a different 3′ untranslated region (UTR) that lacked the Pepper aptamer and used a less efficient terminator[47]. Testing the extended gate with this alternative 3′ UTR (Figure 4F, right) reduced leak in lysate, but, in contrast to PURE, protein expression in the ON state did not vary with output toehold length (Figure 4G, right). This phenomenon was not specific to this gate sequence as variants with a shorter *e* domain, an alternative *e* sequence, or an alternative input domain had the same trend (Supplementary Figures 15, 16). With the alternative 3′ UTR, this unexpected trend now matched the trend observed for THE riboregulators in *E. coli*[24], suggesting cellular components not involved in transcription or translation may play a role.

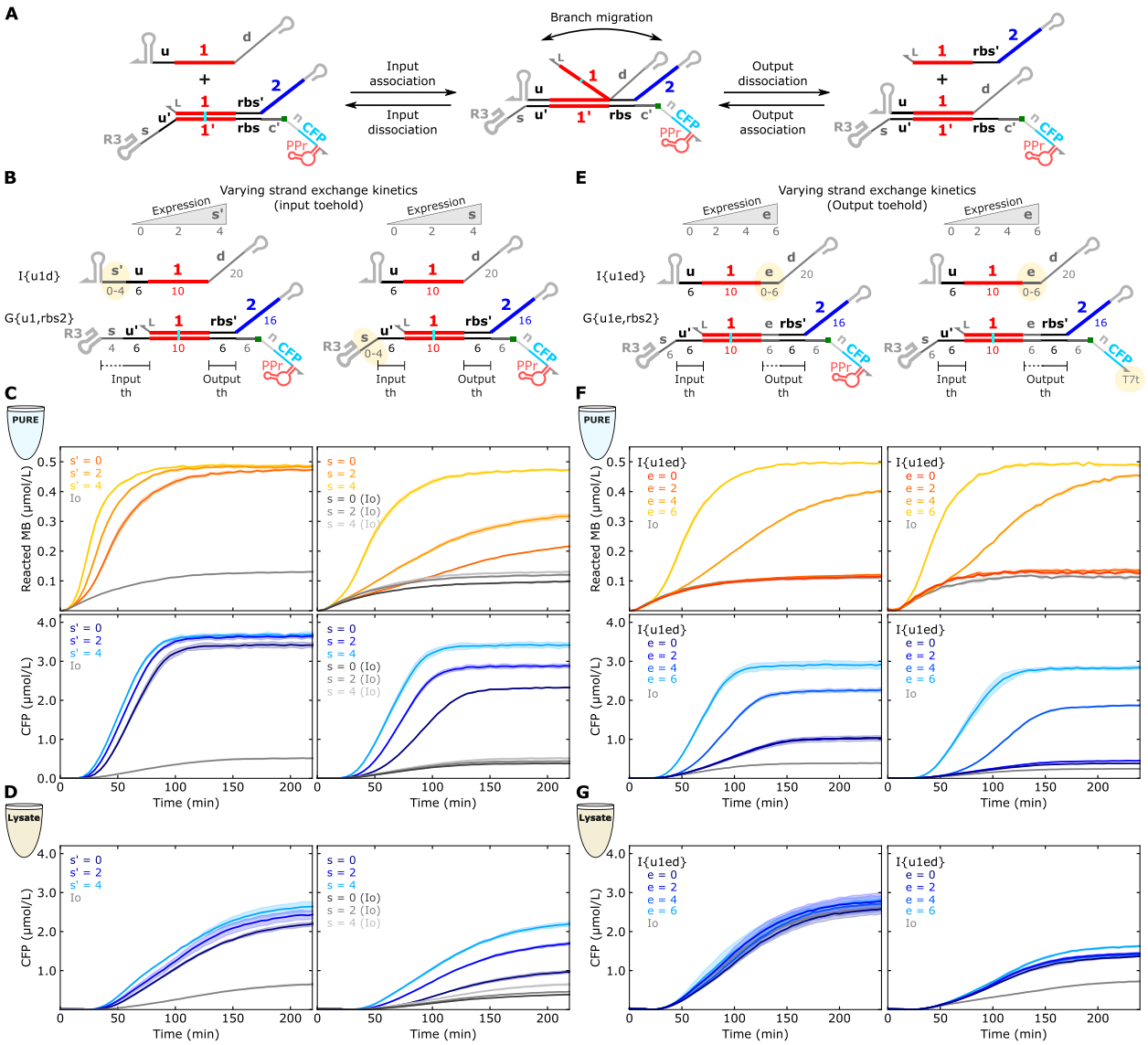

FIGURE 4: Tuning protein expression by modulating RNA strand exchange parameters. (**A**) An expanded schematic of toehold mediated strand exchange (TMSE). With respect to base pairing thermodynamics, TMSE is a reversible process composed of three steps: input/output association, branch migration, and input/output dissociation[20]. In the canonical THE riboregulator design, the *u* and *rbs* toeholds are both six bases so TMSE is not heavily biased in either the forward or reverse direction. (**B**) TMSE can be biased in a forward direction by increasing the length of the toehold on the input strand to disfavor input dissociation (left) or by changing the length of a single-stranded spacer (*s*) between the input toehold and the ribozyme on the THE riboregulator (right). Longer spacers disfavor input dissociation. (**C**,**D**) Kinetics of the systems in (B) in PURE (C) or lysate (D). Pepper measurements and results for additional combinations of *s*:*s′* lengths are in Supplementary Figure 14. (**E**) TMSE can be biased in the reverse direction by increasing the length of the output toehold that has to dissociate after branch migration, designated by *e* on the riboregulator (left). Longer output toeholds can be partially overcome by increasing the length of the input toehold (right). (**F**,**G**) Kinetics of the systems in (E) in PURE (F) or lysate (G). In PURE, the input and riboregulator template concentrations were both 5 nmol/L. In lysate, the input template concentration was 5 nmol/L and the riboregulator template concentration was 2.5 nmol/L. Pepper measurements and results for a gate with a 4-base *e* domain are in Supplementary Figure 15. In (C), (D), (F) and (G), the shaded regions indicate the standard deviation of three technical replicates.

Rapidly prototyping multi-layer RNA strand exchange cascades in PURE

RNA strand exchange circuits that process and transduce information require multiple layers of TMSE reactions whereby outputs from one layer serve as inputs to a downstream layer[15,16,43,44]. Such cascades have been demonstrated for ctRSD circuits in IVT and used to execute multi-input logic[21,22]. Further, up to 3-layer ctRSD circuits have been measured using THE riboregulators in *E. coli*[24]. In *E. coli*, ctRSD gates upstream of a THE riboregulator generally required a 3′ hairpin on the output strand to transduce a measurable signal[24], a constraint not necessary in IVT[21,22]. We hypothesized that these 3′ hairpins were necessary in *E. coli* to confer stability to the output strands so that they would not degrade before being able to react downstream.

To explore this hypothesis, we tested multi-layer cascades with different upstream gate designs in PURE, which lacks ribonucleases that degrade RNA. We started with 2-layer cascades testing upstream gates with and without 3′ hairpins on their output strands (*L2* domain on O{4,u1d}, Figure 5A). Even in PURE, the 2-layer cascade required the 3′ hairpin (hp) to produce signal above the OFF state (2-layer, Figure 5B), suggesting the results in *E. coli* may not be related to RNA stability alone. We confirmed this result was not sequence specific, as the same trend held for an upstream gate with an alternative input domain (domain *5*), an alternative ribozyme (*R3*), an alternative output domain (domain *3*), and an alternative 3′ hairpin (Supplementary Figure 17A-C). Direct transcription of a 2nd layer output (*e.g.*, O{4,u1d}) also required a 3′ hairpin to induce CFP production above the OFF state (Supplementary Figure 17D), indicating this phenomenon is related to the free output strand rather than the dsRNA gate. This phenomenon likely extends to input RNAs, but all input RNAs used in this study had terminators with 3′ hairpins by design.

To investigate whether these results were confined to the THE riboregulator layer, we tested cascades with more than 2 layers. We designed 3- and 4-layer cascades using gates with and without 3′ hairpins in the highest layer while maintaining 3′ hairpins in all downstream layers (Figure 5A). These multi-layer cascades also required 3′ hairpins to transduce signal above the OFF state (Figure 5B, top vs. bottom). Again, this was not specific to the chosen sequence as the same trend held for a 3-layer cascade with a different combination of gate sequences (Supplementary Figure 18).

We hypothesized that ribosomes, which bind unstructured regions of RNA promiscuously[42,48,49], could be interacting with ssRNA output strands and interfering with strand exchange, with the structure of a 3′ hairpin disfavoring such interactions. To test this hypothesis, we implemented a 1-layer circuit triggered by O{5,u1d} with and without a 3′ hairpin in a PURE reaction lacking ribosomes (Figure 6A). In the absence of ribosomes, O{5,u1d} with and without a 3′ hairpin had the same MB kinetics (Figure 6B, left). Adding ribosomes to a concentration equivalent to previous PURE experiments (Supplementary Figure 20) substantially damped the MB and CFP signals from O{5,u1d} without a 3′ hairpin (Figure 6B, right). The same trend was observed for a 2-layer cascade using an upstream gate with and without a 3′ hairpin on the output strand (Supplementary Figure 21). We confirmed these results did not arise from differences in ionic conditions in the reactions with and without ribosomes (Supplementary Figure 22). Together, these results suggest ribosomes play a role in the requirement for a 3′ hairpin on ssRNAs that serve as

inputs for strand exchange reactions. It is still possible that contaminants in the purified ribosome solution that co-isolate with ribosomes are responsible for our observations, which warrants future exploration.

Given that removing ribosomes from PURE resolved the discrepancy between output strands with and without 3′ hairpins, we next asked whether inhibiting translation with different antibiotics would produce similar results. With a 1-layer strand exchange reaction (Figure 6A), we tested antibiotics[50] that inhibit either translation initiation, elongation, or termination (Figure 6C, Supplementary Figure 23). At the antibiotic concentrations used, we observed minimal CFP expression, indicating successful inhibition of translation (Figure 6D). However, the antibiotics did not affect the MB signal when expressing O{5,u1d} without a 3′ hairpin. O{5,u1d} with a 3′ hairpin performed similarly to experiments in PURE with active translation, although inhibiting translation initiation slowed down MB kinetics. Similar trends were observed for a 2-layer cascade using a gate with and without a 3′ hairpin on the output strand (Supplementary Figure 24). These results suggest that binding between output strands without 3′ hairpins and ribosomes interferes with RNA strand exchange, rather than reactions associated with active translation.

Regardless of the mechanism underlying the requirement for a 3′ hairpin on the output, we demonstrated that multi-layer cascades can be reliably constructed in CFES. As observed in IVT and in *E. coli*, CFP production decreased and leak generally increased with the number of layers (Figure 5C, D). In an effort to reduce leak, we tested 4-layer cascades using upstream gates with an inverted transcription order, *i.e.*, gates that transcribe the bottom strand of the gate before the output strand. However, as in IVT studies[22], leak reduction was inconsistent across inverted gate sequence combinations (Supplementary Figure 19).

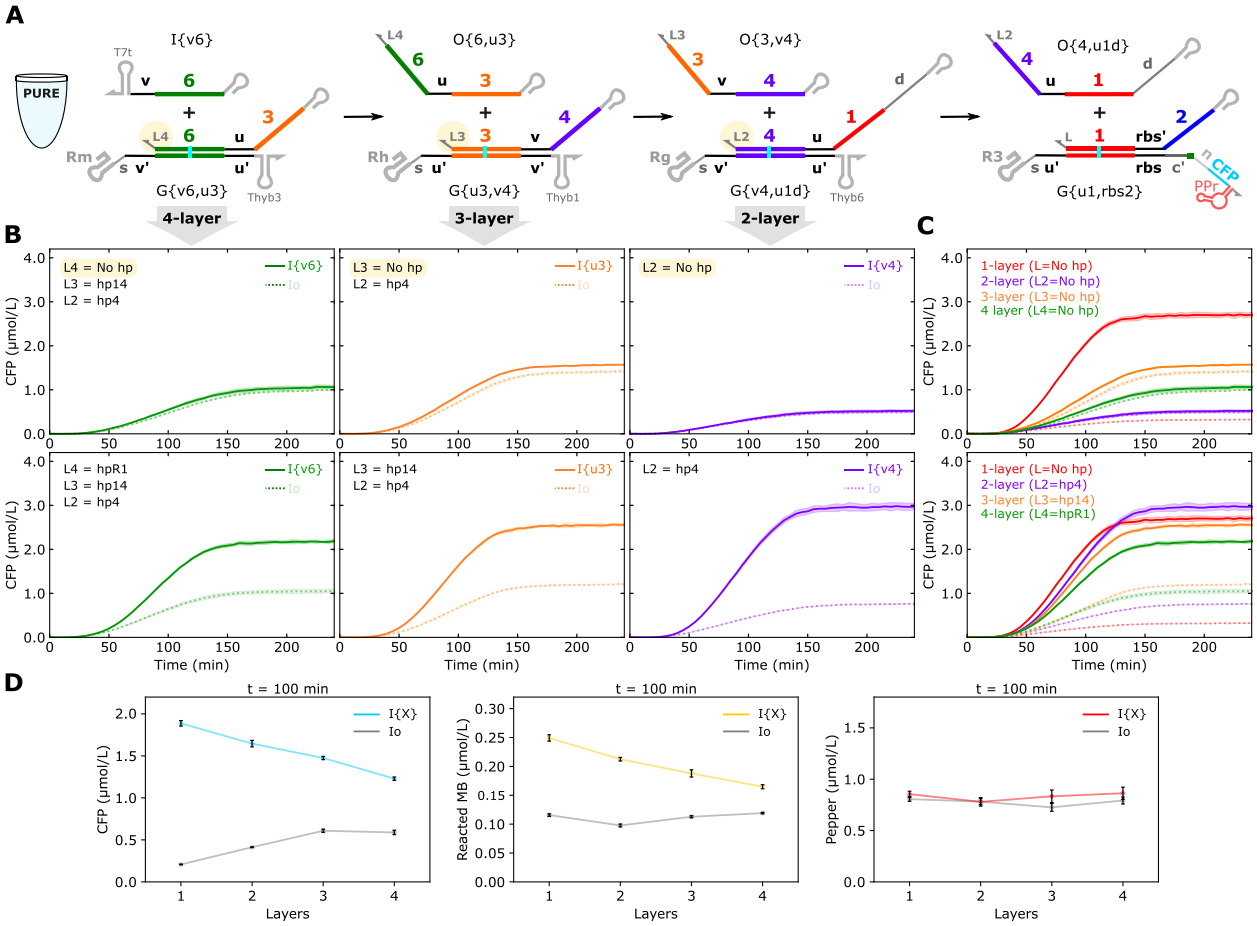

FIGURE 5: Characterizing multi-layer RNA strand exchange reactions in PURE. (**A**) Schematic of a 4-layer RNA strand exchange cascade. Linker domains *L2*, *L3*, and *L4* are designed either to be a short 3-base linker (no hp) or a hairpin-forming linker (hpX) as specified in the plots. (**B**) CFP kinetics for multi-layer cascades without (top) and with (bottom) hairpins at the 3′ end of the output strands. Dashed lines indicate cascades expressed with a non-complementary input (Io). (**C**) Overlay of multi-layer cascade results from (B) with 1-layer results shown for comparison. (**D**) CFP, MB, and Pepper signals at 100 min as a function of cascade layers for cascades of gates with hairpins at the 3′ end of the output strands. ON/OFF ratios decrease for CFP and MB while the Pepper signal remains constant, indicating similar transcription levels for each cascade. Linear DNA templates were used for upstream gates. Io and the THE riboregulator were on plasmid templates. Input and gate DNA template concentrations were each 4 nmol/L with Io added to keep a constant 20 nmol/L total DNA template across samples. In (B) and (C), the shaded regions indicate the standard deviation of three technical replicates. In (D), the error bars indicate the standard deviation of three technical replicates.

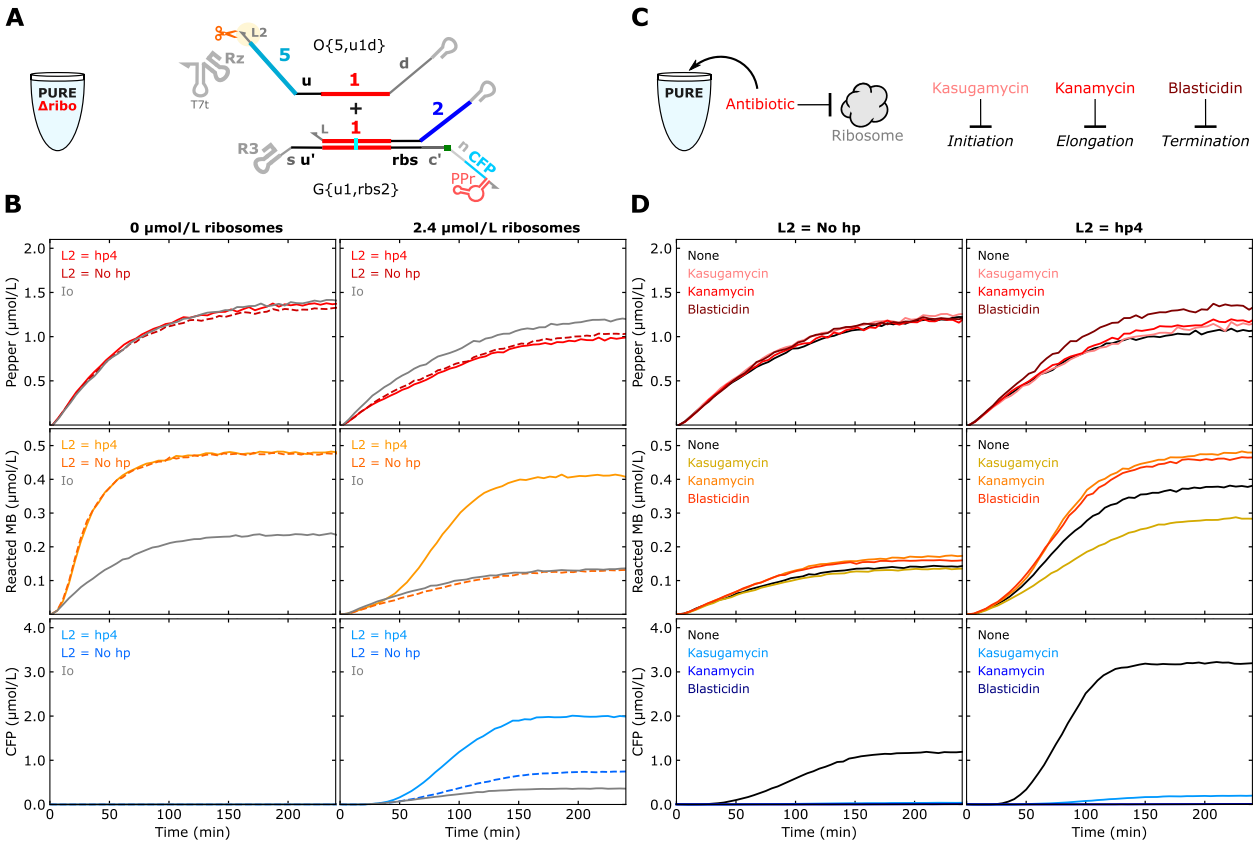

FIGURE 6: Characterizing potential interactions of ribosomes and RNA strand exchange components. (**A**) Schematic of a 1-layer cascade using O{5,u1d} as the input in a PURE reaction missing ribosomes. (**B**) Kinetics of the system in (A) in PURE without (left) and with (right) ribosomes. Dashed lines indicate O{5,u1d} without a 3′ hairpin. Results for the same experiments with a 2-layer cascade are shown in Supplementary Figure 21. (**C**) The system in (A) was characterized in PURE supplemented with different antibiotics that inhibit translation. (**D**) Kinetics of the system in (A) in PURE with different antibiotics (C). Results for O{5,u1d} without or with a 3′ hairpin are shown in the left and right columns, respectively. Kasugamycin, kanamycin, and blasticidin were added to a final concentration of 12.0 mmol/L, 85.8 μmol/L, and 109 μmol/L, respectively. In (B) and (D), output and gate DNA template concentrations were each 4 nmol/L with Io added to keep a constant 20 nmol/L total DNA template across samples, and all reactions had a single replicate. Results for the same experiments with a 2-layer cascade are in Supplementary Figure 24.

CONCLUSIONS

Here we demonstrated how CFES can be leveraged to mechanistically characterize genetically encoded RNA strand exchange circuits. Our characterization was enabled by a combination of measurements that simultaneously probed differences in transcription, translation, and strand exchange. Using these measurements, we demonstrate how different properties of TMSE reactions can be modulated to tune protein expression at the transcriptional (Figure 3B), post-transcriptional (Figure 4), and translational levels (Figure 3D). The ability to manipulate the composition of CFES allowed us to identify potential interactions between RNA strand exchange components and ribosomes, bridging our gap in understanding between previous IVT studies and recent studies in *E. coli*[24]. We expect the framework for prototyping in CFES presented here will be applicable for

characterizing, understanding, and successfully deploying biotechnologies beyond the specific RNA circuits we studied[2,5–7].

A central motivation of this study was to use CFES to complement and supplement previous TMSE characterization in IVT and in cells. Reassuringly, we found the same toehold, input/output, and ribozyme sequences generally functioned well across environments, highlighting interoperability. However, different CFES may serve different purposes when working with RNA circuits. For evaluating initial designs and comparing to IVT, a well-defined system like PURE is ideal. This system removes confounding variables such as RNA degradation, which makes interpretation of results much simpler. However, many trends that agreed in IVT and PURE did not transfer to lysate and cells[24] (Figure 4). Thus, if the primary research question is to explore phenomena observed in cells, then lysate produced more comparable results. Notably, we found that behavior in lysate, like in cells, was more sensitive to changes in 3′ UTRs than in PURE. In PURE, the magnitude of signals varied slightly for different input and gate 3′ UTRs, but the same fundamental trends were observed (Supplementary Section 5). In lysate, however, we observed very different trends in protein expression when the 3′ UTR of the THE riboregulator was changed (Figure 4, Supplementary Figure 16). Interestingly, the trends for different 3′ UTRs of the THE riboregulator in lysate did not match those observed in *E. coli*[24]. Therefore, it is important to keep in mind that adding fluorescent aptamer sequences or changing terminator sequences may change results substantially.

CFES allowed us to study three phenomena that had markedly different results in IVT compared to *E. coli*: the trends of increasing (1) input and (2) output toehold lengths on strand exchange and (3) the need for a 3′ hairpin on the output strands of gates in multi-layer cascades. The trend of increasing strand exchange rate when increasing the input toehold length from 6 bases to 10 bases held across IVT[21], PURE, and lysate, suggesting the opposite trend observed in *E. coli*[24] is unique to the cellular environment. However, we can tune reaction kinetics across all environments by increasing the length of the *s* domain on the gate, which had the same effect in IVT, PURE, lysate, and *E. coli*. The relationship between output toehold length and strand exchange was split between reconstituted systems (IVT and PURE) and whole-cell systems (lysate and cells). The IVT and PURE results matched expectations[14] from a base-pairing perspective, with increasing output toehold length showing decreasing strand exchange (Figure 4). In contrast, lysate and cells showed little dependence on output toehold length, which could implicate RNA degradation pathways[51–53]. The requirement of 3′ hairpins on RNA inputs for strand exchange seemed to be related to interactions with ribosomes (Figures 5 and 6), rather than RNA stability, and will likely be necessary for effective strand exchange in any environment in which translational machinery is present.

Our results only begin to characterize the myriad possible interactions with cellular components and synthetic RNA systems. A class of enzymes to explore further are RNases[51–54], which could play a role in some of the differences between PURE and lysate. Endogenous RNAs absent in reconstituted systems or at low concentrations in lysate systems could also interfere with RNA strand exchange[55,56]. Beyond components that could interfere with RNA strand exchange, some

RNA-binding proteins may enhance RNA:RNA interactions. Many of these proteins have low sequence specificity but preferentially interact with 3′ secondary structures[54,57,58]. Probing the above interactions will likely require new measurements or improvement of the measurements presented here. For example, the molecular beacon measurement was not sensitive enough to characterize strand exchange in lysate (Figure 3), preventing us from assessing strand exchange without translation. Other molecular beacon sequences or designs could resolve this issue, but we found molecular beacons with different toeholds had markedly different ON/OFF ratios across IVT, PURE, and lysate (Supplementary Section 6). Therefore, other methods to measure strand exchange more directly may be necessary.

Despite the differences observed across environments and other yet untested interactions, the work described here lays the foundation for programming protein expression with THE riboregulators, which has many potential applications in CFES and other environments. For example, THE riboregulators could provide the basis for cell-free biosensors that detect small RNAs[59], such as miRNAs, because THE riboregulators require only 16 bases for function. Given that TMSE reactions are reversible, THE riboregulators are well suited for discriminating single-nucleotide polymorphisms[19,60], at least when operated in PURE. Applications involving the expression of multiple proteins could benefit from cascades of RNA strand exchange reactions. With cascades, users could control the timing and order of production of different proteins[43]. Further, different *d* domain lengths could be used to independently control translation strength at each layer while maintaining a constant flux of strand exchange. Such autonomous temporal control could be useful for coordinating expression of multi-protein systems[61] or for operating biosensors that need to coordinate expression of sensing and reporting proteins[62], especially in closed systems like synthetic cells[63]. Predictive kinetic models relating protein expression levels to RNA strand exchange[22,64,65] would further facilitate the above applications. Our results identifying a unified set of design principles for using genetically encoded RNA strand exchange circuits across common synthetic biology environments positions this platform to catalyze development of a wide range of new biotechnologies.

MATERIALS AND METHODS:

DNA, bacterial strains, and chemicals

DNA: The majority of the DNA for experiments in the main text was ordered as clonal gene fragments from Twist Biosciences in plasmid backbones provided by the company — either "pTwist Amp High Copy" (high-copy backbone encoding resistance to ampicillin) or "pTwist Kan Medium Copy" (medium-copy backbone encoding resistance to kanamycin). Some DNA sequences encoding inputs and gates were ordered as eBlock gene fragments from Integrated DNA Technologies (IDT), PCR amplified, and subsequently cloned into pTwist backbones or used as linear templates in PURE. See Supplementary Figure 3 for additional details. DNA oligonucleotide primers for PCR were ordered from IDT with standard desalting. DNA sequences, including their origin, are provided in Supplementary File S1. Annotated GenBank files of DNA templates are provided in Supplementary File S2.

RNA: The oligonucleotides of the molecular beacon complex were ordered as 2′-O-methyl RNA (2omR) with standard desalting for the fluorophore-modified (HEX) strand and HPLC purification for the quencher-modified (Iowa Black RQ) strand. 2omR bases were used to prevent degradation of the molecular beacon in lysate, but we also found that DNA-based molecular beacons produced spurious fluorescence in PURE (Supplementary Figure 4). Molecular beacon complexes were prepared by mixing 20 µmol/L of each strand in transcription buffer (Thermo Fisher, cat. no. EP0113) and heating to 90 °C for 5 min followed by cooling to 20 °C at a rate of 1 °C/min. Schematics and nomenclature for RNA species are described in Supplementary Figures 1 and 2.

Protein sequences: The sCFP3A amino acid sequence, referred to throughout as 'CFP' for simplicity, was taken from GenBank (AAZ65848.1). This sequence was then codon optimized for expression in bacteria with IDT's codon optimization tool.

Pepper RNA aptamer: Unless otherwise specified, THE riboregulators used in this work used a dimeric Pepper aptamer grafted onto F30 and tRNA scaffolds[34]. This Pepper aptamer is termed tDF30ppr. A version of this aptamer lacking the tRNA scaffold, termed DF30ppr, was used in control experiments included in the Supplementary Section 5. In both tDF30ppr and DF30ppr, two point mutations were made to the Pepper sequence to reduce sequence complexity and increase the likelihood of successful synthesis.

Strains: E. coli DH10B (Thermo Fisher, cat. no. EC0113) was used for plasmid cloning and extraction. BL21 Star (DE3) (Thermo Fisher, cat. no. C601003) was used to prepare lysates for cell-free gene expression and to purify proteins.

PURE: PURExpress® *In Vitro* Protein Synthesis Kit (cat. no. E6800L), PURExpress® ΔRibosome Kit (cat. no. E3313S), and *E. coli* ribosomes (cat. no. P0763S) were purchased from New England Biolabs.

Chemicals: Supplementary File S4 lists all chemical reagents used in cell-free reactions and all antibiotics used for plasmid selection. Enzyme and kit product numbers are referenced when first introduced below.

Cloning and PCR

PCR: PCRs were conducted with Phusion Flash Master Mix (Thermo Fisher, cat. no. F-548) and 0.5 µmol/L of each DNA primer. Plasmid backbones for cloning were prepared *via* PCR of 0.2 ng/µL plasmid DNA and the following thermocycler protocol: 30 cycles of a denaturing step (30 s, 98 °C), a primer annealing step (30 s, 60 °C), and an extension step (120 s, 72 °C). A 5 min extension at 72 °C was executed at the end of the program. Backbone PCR products were digested by DpnI (New England Biolabs, cat. no. R0176S) to remove methylated plasmid DNA. PCR products were purified using the PureLink PCR purification kit (Invitrogen, cat. no. K310001). To remove non-specific products shorter than 300 bp from plasmid backbone PCRs, a wash step with buffer B3 was included after DNA binding with buffer B2. Purified PCRs were quantified using a Nanodrop spectrophotometer (DeNovix D-11 Series). Linear DNA inserts or transcription templates were prepared similarly except the extension step in the PCR was shortened to 30 s, the DpnI digestion was not included, and the B3 wash step was omitted during purification.

Cloning: Plasmids assembled in house were cloned *via* Gibson assembly or inverse PCR. Gibson reactions were prepared using Gibson assembly master mix (New England Biolabs, cat. no. E2611L) with a 3:1 molar ratio of insert to backbone DNA, and incubated at 50 ˚C for 1 h prior to transformation. For inverse PCRs, DNA primers were first phosphorylated by T4 Polynucleotide Kinase (New England Biolabs, cat. no. M0201S) at 37 ˚C for 30 min. Inverse PCR products were ligated by T4 DNA Ligase (New England Biolabs, cat. no. M0202S) at room temperature for 2 h.

DNA sequence verification

Plasmids underwent whole-plasmid or ZeroPrep sequencing by Plasmidsaurus using Oxford Nanopore Technology with custom analysis and annotation.

DNA plasmid extraction and quantification

Plasmids were transformed into chemically competent DH10B *E. coli*, and transformants were plated on LB agar with appropriate antibiotic and grown overnight at 37 ˚C. 100 mL of LB with appropriate antibiotic was inoculated with one colony from the transformation and allowed to shake at 37 ˚C, 250 rpm, for 16 h in a baffled 250 mL flask. After growth, a glycerol stock was prepared using 200 µL of the culture to enable long-term storage at - 80 ˚C and to bypass the transformation step in future DNA extractions.

Cells were harvested by centrifugation at 4 ˚C, 2700 x g, for 20 min. Plasmid DNA was extracted using the E.Z.N.A Plasmid DNA Midi kit (Omega Biotek, cat. no. D6904-04) and the protocol provided by the manufacturer. Upon being eluted, DNA was precipitated in 100 % isopropanol (Millipore Sigma, cat. no. 34863) and washed in 70 % ethanol (Millipore Sigma, cat. no. E7023) following the kit's protocol. The DNA pellet was allowed to air dry for 30 min, then resuspended in 30 µL of elution buffer. For 'High Copy' plasmids, at this stage the DNA solution typically had a concentration between (2500 and 4000) ng/µL based on spectrophotometric measurements on a Nanodrop instrument.

Each plasmid underwent an additional purification using the PureLink PCR Purification kit (Invitrogen, cat. no. K310002). The 30 µL DNA solution was mixed with 70 µL of water and 400

µL of binding buffer B2. When the mass of DNA exceeded 40 µg — the maximum binding capacity of the column — the mixture of DNA and B2 was divided across more than one column. The rest of the purification was performed following the protocol provided by the manufacturer. DNA was eluted in 35 µL of elution buffer per column; samples divided into multiple columns were eluted into the same container.

Following purification, plasmid DNA was quantified using the Qubit™ 1x dsDNA high-sensitivity assay (Invitrogen, cat. no. Q33231) on a Qubit 4 fluorometer (Invitrogen, cat. no. Q33226). Samples were prepared by mixing 2 µL of a 100-fold dilution of plasmid DNA with 198 µL of Qubit working solution. Concentrations of plasmids measured using the Qubit assay were generally (40 to 50) % lower than those measured using a Nanodrop instrument. All cell-free reactions were prepared based on Qubit-measured concentrations.

Linear DNA templates were prepared *via* PCR as described above, purified using the PureLink PCR Purification kit, and quantified using the Qubit™ 1x dsDNA high-sensitivity assay for comparability to plasmid DNA concentrations. We found Qubit and Nanodrop concentrations were in much closer agreement for linear templates than for plasmids.

The few Medium Copy plasmids tested required two 100 mL LB cultures to be grown and combined onto a single column for plasmid extraction to obtain high enough DNA concentrations for use in multiple experiments. So High Copy plasmids were used for most constructs tested. Additionally, we found that MB and CFP measurements of Medium Copy plasmids encoding a riboregulator deviated more from linear templates in PURE than a High Copy counterpart (Supplementary Figures 5 and 13), even though concentrations were putatively the same across template types. These results could be due to increased carryover of inhibitors from Medium Copy plasmid extraction.

Lysate preparation

The extract preparation protocol was adapted from published protocols[37,66] and divided into a four-day protocol.

Day 1: An *E. coli* BL21 Star (DE3) glycerol stock was streaked onto an LB (10 g/L tryptone, 10 g/L sodium chloride, 5 g/L yeast extract) plate without antibiotic and incubated at 37 ˚C for 16 h.

Day 2: 50 mL of LB without antibiotic was inoculated with one colony from the E. coli BL21 Star (DE3) plate at 37 ˚C and 250 rpm for 16 h in a baffled 250 mL flask.

Day 3: 400 mL of 2xYTP (16 g/L tryptone, 5 g/L sodium chloride, 10 g/L yeast extract, 40 mmol/L potassium phosphate dibasic, 22 mmol/L potassium phosphate monobasic) was inoculated with 25 mL of the overnight culture at 37 ˚C, 250 rpm. About 1.5 h into growth, 0.4 mmol/L isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to the culture to induce expression of T7 RNA polymerase. Cells continued to grow until the OD600 reached (1.6 to 1.7), about (3.5 to 4) h into growth. Upon reaching the target OD600, cells were harvested by centrifugation at 4 ˚C, 2700 x g and washed three times with S30 buffer (10 mmol/L Tris acetate, 14 mmol/L magnesium acetate, 60 mmol/L potassium acetate, pH adjusted to 8.2 with 5 mol/L potassium hydroxide, and 2 mmol/L dithiothreitol), with a centrifugation step between washes. After the last wash step, the mass of the

cell pellet was determined, and cells were stored at - 80 ˚C. Per 400 mL of culture, a cell mass of 2 g to 2.5 g is expected.

Day 4: Cell pellets were thawed on ice, then resuspended in 1 mL of S30 buffer per g of cells. The cellular resuspension was divided into 1 mL aliquots in sterile 1.5 mL microcentrifuge tubes. Cellular resuspensions were then lysed on ice using a Q125 sonicator (Qsonica) with a 3.175 mm diameter probe, 20 kHz frequency, 50 % amplitude, and cycles of 10 s on and 10 s off, delivering (260 ± 10) J after 5 cycles. At the start of each cycle, the tip of the probe was positioned close to the bottom of the tube to prevent bubble formation; approximately 5 times per cycle, the tube was moved down slowly such that the tip reached the 0.5 mL mark of the tube and then reverted to its original position. Immediately after lysis, 3 mmol/L DTT was added to each tube. Lysed cells were centrifuged at 4 ˚C, 12000 x g, for 15 min. The supernatant was consolidated into a single tube, mixed by inversion, then divided into 2 mL aliquots in 14 mL culture tubes and allowed to shake at 37 ˚C, 250 rpm, for 80 min—this is the runoff reaction. The sample was then centrifuged at 4 ˚C, 12000 x g, for 15 min. The supernatant (the final extract to be used in CFE reactions) was consolidated into one tube, mixed by inversion, divided into 200 µL aliquots, and stored at - 80 ˚C for future use.

Plate reader experiments

Unless otherwise specified, all cell-free reactions were run in three 10 µL technical replicates in clear, flat-bottom 384-well plates (Greiner Bio-one, cat. no. 784101) in a BioTek Synergy Neo2 plate reader at 37 ˚C. Measurements of sCFP3A (excitation: 430 nm, emission: 475 nm, bandwidth: 10 nm, gain: 85), MB (excitation: 524 nm, emission: 565 nm, bandwidth: 20 nm, gain: 80), and Pepper-HBC620 complex (excitation: 577 nm, emission: 620 nm, bandwidth: 20, gain: 75) were taken every 5 min from the bottom of the plate. All reactions included 5 µmol/L HBC620 Pepper dye and 500 µmol/L of the molecular beacon complex. DNA templates concentrations for each experiment are specified in the relevant figure captions or figure legends. At the end of each experiment a synthetic trigger for the molecular beacon, composed of 2′-O-methyl RNA bases, was added to each sample (final concentration 2.3 mol/L) to obtain a maximum MB fluorescence signal for concentration calibration. The compositions of the different reaction environments are described below.

Cell-free reactions in lysate: Reactions were prepared as previously described[67]. Unless otherwise specified, reactions contained 17.5 mmol/L magnesium glutamate, 10 mmol/L ammonium glutamate, 133 mmol/L potassium glutamate, 1.2 mmol/L ATP, 0.85 mmol/L GTP, 0.85 mmol/L CTP, 0.85 mmol/L UTP, 0.034 mg/mL folinic acid, 0.171 mg/mL tRNA from E. coli MRE 600, 0.33 mmol/L nicotinamide adenine dinucleotide, 0.267 mmol/L coenzyme A, 4 mmol/L sodium oxalate, 1 mmol/L putrescine, 1.5 mmol/L spermidine, 50 mmol/L HEPES, 2 mmol/L each of the twenty standard amino acids, 0.03 mol/L PEP, 27 % extract by volume. 17.5 mmol/L magnesium glutamate was used because it yielded the highest ON/OFF ratio for CFP of magnesium glutamate concentrations spanning (10 to 30) mmol/L (Supplementary Figure 6).

PURExpress reconstituted system: Reactions were prepared following the protocol provided by the manufacturer. Components other than Solution A and Solution B were mixed on ice, followed

by addition of Solution A (40 % by volume) then Solution B (30 % by volume). All reactions used the same lot of PURExpress (Lot no. 10272724).

In vitro transcription (IVT): Unless otherwise specified, IVT reactions were conducted in transcription buffer (40 mmol/L tris-HCl (pH 7.9), 6 mmol/L $MgCl_2$, 10 mmol/L dithiothreitol, 10 mmol/L NaCl, and 2 mmol/L spermidine) supplemented with 2 mmol/L of each ribonucleotide triphosphate (NTP) (Thermo Fisher, cat. no. R0481). The concentrations of T7 RNAP (Thermo Fisher, cat. no. EP0113) used are specified in the relevant figure captions.

Protein purification

To generate a calibration curve for sCFP3A measurements, sCFP3A was purified in NEBExpress® Ni Spin Columns (New England Biolabs, cat. no. S1427L). Using inverse PCR, 6 histidine amino acid residues (CATCATCACCACCACCAT) were inserted at the C-terminus of sCFP3A in a plasmid encoding T7 RNA polymerase-driven sCFP3A expression and resistance to kanamycin.

BL21 Star (DE3) *E. coli* were transformed with the sCFP3A-6xHis plasmid, and one colony of this transformation was used to inoculate a 5 mL LB culture supplemented with 50 µg/mL kanamycin. This culture was grown at 37 ˚C, 250 rpm for 16 h. A 200 mL LB culture containing 50 µg/mL kanamycin was inoculated with 2 mL of the smaller culture and grown at 37 ˚C and 250 rpm. After 2 h of growth, 0.4 mmol/L IPTG were added to the culture to induce expression of T7 RNA polymerase. The culture was allowed to grow for another 2 h, after which cells were harvested by centrifugation at 4 ˚C, 2700 x g, 15 min. The cell pellet was resuspended in 1 mL of lysis buffer (supplied by the nickel column's manufacturer) per g of cells. The cell resuspension was divided into 1 mL aliquots in 1.5 mL microcentrifuge tubes. Cells were then lysed on ice using a Q125 sonicator (Qsonica) with a 3.175 mm diameter probe, 20 kHz frequency, 50 % amplitude, and cycles of 10 s on and 10 s off, delivering about 300 J after 6 cycles. Lysed cells were centrifuged at 4 ˚C, 12000 xg, for 15 min, and the lysate supernatant was consolidated into one tube. Purification proceeded following the manufacturer's protocol. 1 mL of lysate was added to each of two nickel columns, with 500 µL added at a time followed by binding and centrifugation. Each time, the lysate was allowed to bind to the column for 10 min at 4 ˚C, with gentle mixing by inversion every 3 min. Protein was eluted in 200 µL elution buffer. The purified protein fractions from each column were combined, and purification was verified on an SDS-PAGE gel (Bio-rad, cat. no. 4568096). The histidine tag was not removed.

Purified sCFP3A was quantified with the LabChip® GXII Touch™ protein characterization system (Revvity, CLS138160) and the ProteinEXact™ assay and lab chip (Revvity, CLS150337).

***In vitro* transcription**

To generate a calibration curve for the Pepper aptamer, Pepper RNA transcripts were prepared via *in vitro* transcription (IVT) using a plasmid encoding T7 RNA polymerase-driven Pepper expression. Prior to transcription, this plasmid was linearized via PCR to remove the origin of replication and the antibiotic resistance gene.

The 100 µL IVT reaction contained 1 µg of linear DNA template, 2 mmol/L each of ATP, GTP, CTP and UTP, 0.6 U/µL T7 RNA polymerase (Thermo Scientific, cat. no. EP0113), and the transcription buffer provided by the polymerase's manufacturer. The reaction was run in 100 µL volumes at 37 ˚C for 2 h in a thermocycler. Each reaction was then treated with 0.1 U/µL DNase I for 15 min at room temperature prior to purification and elution in 30 µL nuclease-free water using the RNA Clean & Concentrator-25 kit (Zymo Research, cat. no. R1018).

RNA was quantified using the Qubit™ RNA high-sensitivity assay (Invitrogen, cat. no. Q32855) on a Qubit 4 fluorometer (Invitrogen, cat. no. Q33226).

Calibration curves for Pepper, molecular beacon, and sCFP3A measurements

For each calibration curve, the relevant molecule was added to PURE or lysate reactions without any DNA templates at multiple concentrations capturing fluorescence intensity values above and below those obtained in our experiments. The cell-free reactions were incubated in a plate reader at 37 ˚C and measured for 4 h. Fluorescence values at 90 min were used to create calibration curves for each molecule. Linear regression was used to fit the calibration curves using the polyfit() function of the Numpy package in Python 3.8.18. The slope of these fit lines was used to convert fluorescence intensities to concentrations. To avoid physically unrealistic negative concentrations, the minimum fluorescence value of a given sample was used as the value corresponding to a concentration of 0 mol/L.

Pepper aptamer: Purified Pepper mRNA prepared via IVT was used to generate calibration curves. In PURE, the signal stabilized after 1 h. In lysate, the signal decreased sharply in the first 30 min followed by a slow decay over the remainder of the experiment (Supplementary Figure 30A). Ribonuclease-mediated degradation in lysate confounds concentration calibration; the signal associated with an RNA concentration at a given time point may be used for comparing across samples and conditions, but does not reflect a stable RNA concentration during the reaction.

Molecular Beacon: A synthetic molecular beacon trigger strand, composed of 2omR bases, was added at different concentrations to PURE or lysate reactions containing a fixed concentration of the molecular beacon complex. The resulting calibration curves showed a linear trend between molecular beacon fluorescence and increasing concentrations of trigger RNA added (Supplementary Figure 30B). Thus, to convert molecular beacon signal to concentration the signal in each sample was normalized to be between 0 and 1 based on the minimum and maximum fluorescence values in an experiment and then multiplied by the total concentration of MB added to the reaction:

$$\text{Reacted MB } (\mu\text{mol/L}) = \frac{\text{fluorescence} - \min(\text{fluorescence})}{\max(\text{fluorescence}) - \min(\text{fluorescence})} * [MB]_{added}$$

The synthetic molecular beacon trigger strand was spiked into every sample at the end of each experiment to a concentration in excess of the molecular beacon complex to ensure each sample had an internal measurement of fully reacted molecular beacon signal for normalization.

sCFP3A: Purified sCFP3A was used to generate calibration curves. In both PURExpress and lysate, the protein signal decreased slightly in the first 30 min, likely due to temperature equilibration of the samples in the plate reader, but then remained nearly constant (Supplementary Figure 30C).

AUTHOR CONTRIBUTIONS (CRediT):

FP: Conceptualization (Template and experiment design), Methodology and Investigation (Plasmid cloning and preparation, lysate preparation, lysate experiments, PURE experiments), Project administration, Writing – original draft

ER: Methodology and Investigation (PURE experiments), Writing – review and editing

SWS: Conceptualization, Data curation, Formal analysis, Methodology and Investigation (Plasmid cloning, IVT experiments), Project administration, Supervision, Visualization, Writing – original draft

DATA AND CODE AVAILABILITY:

Fluorescence data for all experiments in this study will be available on Zenodo.

Data analysis and visualization was conducted in Python (3.8.18) in the Spyder IDE (5.4.3). Example analysis and plotting code is available in Supplementary File S3. Additionally, the ctRSD simulator 2.1 package (https://ctrsd-simulator.readthedocs.io/en/latest/SeqCompiler.html) contains a sequence compiling function to stitch together sequences for any combination of sequence domains explored in this study. Scripts for generating the sequences in this study using the sequence compiler are available in Supplementary File S3. A Google CoLab for running the sequence compiler is available at: CoLab link

CONFLICTS OF INTEREST:

SWS is an inventor on two patent applications pertaining to ctRSD circuits (Application number PCT/US2022/053229) and THE riboregulators (Application number PCT/US25/13610). The authors declare no other conflicts.

ACKNOWLEGDMENTS:

The authors thank Svetlana Ikonomova, Zoila Jurado, David Ross, Elizabeth Strychalski, Chad Sundberg, and Molly Wintenberg for insightful discussions throughout the project and feedback on the manuscript. In addition, the authors thank Katharina Yandrofski and Kayla Diaz at the Institute for Bioscience and Biotechnology Research (IBBR) for usage of and assistance with their LabChip® GXII Touch protein characterization instrument. The authors also thank Bright Eyes, whose album *I'm Wide Awake, It's Morning* was inspiration throughout the study.

Disclaimer: Certain commercial entities, equipment, or materials may be identified in this document to describe an experimental procedure or concept adequately. Such identification is not

intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose. Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

REFERENCES:

1. Simmel, F. C., Yurke, B. & Singh, H. R. Principles and applications of nucleic acid strand displacement reactions. *Chem. Rev.* **119**, 6326–6369 (2019).
2. Kaminski, M. M., Abudayyeh, O. O., Gootenberg, J. S., Zhang, F. & Collins, J. J. CRISPR-based diagnostics. *Nature Biomedical Engineering* **5**, 643–656 (2021).
3. Jung, J. K., Archuleta, C. M., Alam, K. K. & Lucks, J. B. Programming cell-free biosensors with DNA strand displacement circuits. *Nat. Chem. Biol.* **18**, 385–393 (2022).
4. Koksaldi, I. C. *et al.* RNA-Based Sensor Systems for Affordable Diagnostics in the Age of Pandemics. *ACS Synthetic Biology* **13**, 1026–1037 (2024).
5. Zhu, Y., Zhu, L., Wang, X. & Jin, H. RNA-based therapeutics: an overview and prospectus. *Cell Death & Disease* **13**, 644 (2022).
6. Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell* **159**, 925–939 (2014).
7. Chappell, J., Westbrook, A., Verosloff, M. & Lucks, J. B. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nat. Commun.* **8**, 1051 (2017).
8. Wang, T. & Simmel, F. C. Riboswitch-inspired toehold riboregulators for gene regulation in Escherichia coli. *Nucleic Acids Research* **50**, 4784–4798 (2022).
9. Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).
10. Kim, J. *et al.* De novo-designed translation-repressing riboregulators for multi-input cellular logic. *Nat. Chem. Biol.* **15**, 1173–1182 (2019).
11. Kieffer, C., Genot, A. J., Rondelez, Y. & Gines, G. Molecular Computation for Molecular Classification. *Advanced Biology* **n/a**, 2200203 (2023).
12. Liu, B., Samaniego, C. C., Bennett, M. R., Franco, E. & Chappell, J. A portable regulatory RNA array design enables tunable and complex regulation across diverse bacteria. *Nature Communications* **14**, 5268 (2023).
13. Ameruoso, A., Gambill, L., Liu, B., Villegas Kcam, M. C. & Chappell, J. Brave new 'RNA' world—advances in RNA tools and their application for understanding and engineering biological systems. *Curr. Opin. Syst. Biol.* **14**, 32–40 (2019).
14. Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* **131**, 17303–17314 (2009).
15. Cherry, K. M. & Qian, L. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **559**, 370–376 (2018).
16. Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196–1201 (2011).
17. Zhang, C. *et al.* Cancer diagnosis with DNA molecular computation. *Nat. Nanotechnol.* **15**, 709–715 (2020).
18. Lopez, R., Wang, R. & Seelig, G. A molecular multi-gene classifier for disease diagnostics. *Nat. Chem.* **10**, 746–754 (2018).
19. Zhang, D. Y., Chen, S. X. & Yin, P. Optimizing the specificity of nucleic acid hybridization. *Nature Chemistry* **4**, 208–214 (2012).
20. Srinivas, N. *et al.* On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res* **41**, 10641–10658 (2013).
21. Schaffter, S. W. & Strychalski, E. A. Cotranscriptionally encoded RNA strand displacement circuits. *Sci. Adv.* **8**, eabl4354 (2022).
22. Schaffter, S. W., Wintenberg, M. E., Murphy, T. M. & Strychalski, E. A. Design Approaches to Expand the Toolkit for Building Cotranscriptionally Encoded RNA Strand Displacement Circuits. *ACS Synth. Biol.* **12**, 1546–1561 (2023).
23. Bae, W., Stan, G.-B. V. & Ouldridge, T. E. In situ generation of RNA complexes for synthetic molecular strand-displacement circuits in autonomous systems. *Nano Lett.* **21**, 265–271 (2021).
24. Schaffter, S. W., Vasilyeva, O. B., Wintenberg, M. E., Hurley, J. M. & Alperovich, N. Y. Genetically encoded RNA strand exchange circuits for programmable protein expression and computation in cells. *bioRxiv* (2025).
25. Hunt, A. C. *et al.* Cell-Free Gene Expression: Methods and Applications. *Chem. Rev.* **125**, 91–149 (2025).
26. Garenne, D. *et al.* Cell-free gene expression. *Nature Reviews Methods Primers* **1**, 49 (2021).
27. Niederholtmeyer, H., Xu, L. & Maerkl, S. J. Real-Time mRNA Measurement during an in Vitro Transcription and Translation Reaction Using Binary Probes. *ACS Synth. Biol.* **2**, 411–417 (2013).

28. Shin, J. & Noireaux, V. An E. coli Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. *ACS Synth. Biol.* **1**, 29–41 (2012).
29. Chappell, J., Jensen, K. & Freemont, P. S. Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Research* **41**, 3471–3481 (2013).
30. Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V. & Murray, R. M. Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an Escherichia coli Based TX-TL Cell-Free System. *ACS Synth. Biol.* **3**, 387–397 (2014).
31. Kopniczky, M. B. *et al.* Cell-Free Protein Synthesis as a Prototyping Platform for Mammalian Synthetic Biology. *ACS Synth. Biol.* **9**, 144–156 (2020).
32. Takahashi, M. K. *et al.* Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription–Translation (TX-TL) Systems. *ACS Synth. Biol.* **4**, 503–515 (2015).
33. Chen, X. *et al.* Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nature Biotechnology* **37**, 1287–1293 (2019).
34. Mumbleau, M. M., Chevance, F., Hughes, K. & Hammond, M. C. Investigating the Effect of RNA Scaffolds on the Multicolor Fluorogenic Aptamer Pepper in Different Bacterial Species. *ACS Synth Biol* **13**, 1093–1099 (2024).
35. Piorino, F., Sundberg, C. A., Strychalski, E. A. & Romantseva, E. Characterizing Cell-Free Transcription and Translation Dynamics with Nucleic Acid-Based Assays. *bioRxiv* 2025.09.10.675373 (2025) doi:10.1101/2025.09.10.675373.
36. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113–130 (1986).
37. Kwon, Y.-C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports* **5**, 8663 (2015).
38. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* **27**, 946–950 (2009).
39. Bushhouse, D. Z., Choi, E. K., Hertz, L. M. & Lucks, J. B. How does RNA fold dynamically? *J. Mol. Biol.* **434**, 167665 (2022).
40. de Smit, M. H. & van Duin, J. Translational Standby Sites: How Ribosomes May Deal with the Rapid Folding Kinetics of mRNA. *Journal of Molecular Biology* **331**, 737–743 (2003).
41. Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Research* **42**, 2646–2659 (2014).
42. Sterk, M., Romilly, C. & Wagner, E. G. H. Unstructured 5′-tails act through ribosome standby to override inhibitory structure at ribosome binding sites. *Nucleic Acids Research* **46**, 4188–4199 (2018).
43. Scalise, D. *et al.* Programming the sequential release of DNA. *ACS Synth. Biol.* **9**, 749–755 (2020).
44. Lapteva, A. P., Sarraf, N. & Qian, L. DNA Strand-Displacement Temporal Logic Circuits. *J. Am. Chem. Soc.* **144**, 12443–12449 (2022).
45. Peng, Y., Li, X., Yuan, R. & Xiang, Y. Steric hindrance inhibition of strand displacement for homogeneous and signal-on fluorescence detection of human serum antibodies. *Chem. Commun.* **52**, 12586–12589 (2016).
46. Lysne, D., Hachigian, T., Thachuk, C., Lee, J. & Graugnard, E. Leveraging Steric Moieties for Kinetic Control of DNA Strand Displacement Reactions. *J. Am. Chem. Soc.* **145**, 16691–16703 (2023).
47. Calvopina-Chavez, D. G., Gardner, M. A. & Griffitts, J. S. Engineering efficient termination of bacteriophage T7 RNA polymerase transcription. *G3 Genes|Genomes|Genetics* **12**, jkac070 (2022).
48. Del Campo, C., Bartholomäus, A., Fedyunin, I. & Ignatova, Z. Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. *PLOS Genetics* **11**, e1005613 (2015).
49. Yang, Q., Figueroa-Bossi, N. & Bossi, L. Translation Enhancing ACA Motifs and Their Silencing by a Bacterial Small Regulatory RNA. *PLOS Genetics* **10**, e1004026 (2014).
50. Lin, J., Zhou, D., Steitz, T. A., Polikanov, Y. S. & Gagnon, M. G. Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annual Review of Biochemistry* vol. 87 451–478 (2018).
51. Laalami, S., Zig, L. & Putzer, H. Initiation of mRNA decay in bacteria. *Cell. Mol. Life Sci.* **71**, 1799–1828 (2014).
52. Menendez-Gil, P. & Toledo-Arana, A. Bacterial 3′UTRs: A Useful Resource in Post-transcriptional Regulation. *Frontiers in Molecular Biosciences* **Volume 7-2020**, (2021).

53. Hui, M. P., Foley, P. L. & Belasco, J. G. Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet.* **48**, 537–559 (2014).
54. Quendera, A. P. *et al.* RNA-Binding Proteins Driving the Regulatory Activity of Small Non-coding RNAs in Bacteria. *Frontiers in Molecular Biosciences* **7**, (2020).
55. Mayer, T., Oesinghaus, L. & Simmel, F. C. Toehold-Mediated Strand Displacement in Random Sequence Pools. *J. Am. Chem. Soc.* **145**, 634–644 (2023).
56. Fujiwara, K. *et al.* In vitro transcription–translation using bacterial genome as a template to reconstitute intracellular profile. *Nucleic Acids Research* **45**, 11449–11458 (2017).
57. G. Chaulk, S. *et al.* ProQ Is an RNA Chaperone that Controls ProP Levels in Escherichia coli. *Biochemistry* **50**, 3095–3106 (2011).
58. Kim, H. J. *et al.* Structural basis for recognition of transcriptional terminator structures by ProQ/FinO domain RNA chaperones. *Nature Communications* **13**, 7076 (2022).
59. McSweeney, M. A., Zhang, Y. & Styczynski, M. P. Short Activators and Repressors of RNA Toehold Switches. *ACS Synth. Biol.* **12**, 681–688 (2023).
60. Yan, Z. *et al.* Programmable Fluorescent Aptamer-Based RNA Switches for Rapid Identification of Point Mutations. *medRxiv* 2025.03.07.25323576 (2025) doi:10.1101/2025.03.07.25323576.
61. Nour El-Din, H. *et al.* Cell-free expression system: a promising platform for bacteriophage production and engineering. *Microbial Cell Factories* **24**, 42 (2025).
62. Wang, T. & Lu, Y. Advances, Challenges and Future Trends of Cell-Free Transcription-Translation Biosensors. *Biosensors* **12**, (2022).
63. Rothschild, L. J. *et al.* Building Synthetic Cells─From the Technology Infrastructure to Cellular Entities. *ACS Synth. Biol.* **13**, 974–997 (2024).
64. Poole, W., Pandey, A., Shur, A., Tuza, Z. A. & Murray, R. M. BioCRNpyler: Compiling chemical reaction networks from biomolecular parts in diverse contexts. *PLOS Computational Biology* **18**, e1009987 (2022).
65. Jurado, Z., Pandey, A. & Murray, R. M. A chemical reaction network model of PURE. *bioRxiv* 2023.08.14.553301 (2023) doi:10.1101/2023.08.14.553301.
66. Sun, Z. Z. *et al.* Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *J Vis Exp* 50762 (2013) doi:10.3791/50762.
67. Jewett, M. C. & Swartz, J. R. Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnology and Bioengineering* **86**, 19–26 (2004).