ARTICLE
PUBLISHED ONLINE: 2 FEBRUARY 2015 | DOI: 10.1038/NCHEMBIO.1737
nature chemical biology

Creating small transcription activating RNAs

James Chappell, Melissa K Takahashi & Julius B Lucks*

We expanded the mechanistic capability of small RNAs by creating an entirely synthetic mode of regulation: small transcription activating RNAs (STARs). Using two strategies, we engineered synthetic STAR regulators to disrupt the formation of an intrinsic transcription terminator placed upstream of a gene in *Escherichia coli*. This resulted in a group of four highly orthogonal STARs that had up to 94-fold activation. By systematically modifying sequence features of this group, we derived design principles for STAR function, which we then used to forward engineer a STAR that targets a terminator found in the *Escherichia coli* genome. Finally, we showed that STARs could be combined in tandem to create previously unattainable RNA-only transcriptional logic gates. STARs provide a new mechanism of regulation that will expand our ability to use small RNAs to construct synthetic gene networks that precisely control gene expression.

RNA regulators have become an important component of the synthetic biology toolbox for precisely controlling gene expression and constructing synthetic gene networks[1]. They are increasingly attractive substrates owing to their mechanistic diversity[1] and to the emergence of computational[2–5] and experimental tools[6,7] that predict and characterize RNA structures, ultimately informing their functional design[1]. Of the many types of regulatory RNAs that have been engineered, bacterial small RNAs (sRNAs) have proven to be particularly versatile. These *trans*-acting antisense sRNAs are abundant in nature[8] and exert regulation via direct RNA-RNA interactions with a sense RNA sequence contained within a target mRNA. This interaction results in structural rearrangements of the target that regulate gene expression, typically by occluding or exposing regulatory features, such as ribosome binding sites (RBSs) in the case of translation or intrinsic terminator hairpins in the case of transcription[8].

sRNAs that activate or repress translation are found throughout nature[8] and have been engineered to tune gene expression in metabolic pathways[9], to silence endogenous genes in *E. coli*[10], and to act as key components of genetic circuits that perform cellular computations, including genetic switchboards[11] and counters[12,13]. Moreover, sRNAs that repress transcription have been engineered to create orthogonal and composable regulators that can be used to construct RNA-only transcriptional networks[14,15]. These versatile sRNA transcriptional repressors, called attenuators, have been used to construct RNA-only networks that can act as genetic logic gates, propagate information in transcriptional cascades and control the timing of expression of multiple genes[14,16]. Furthermore, because these networks propagate signals directly as RNA species, they operate on the fast timescales set by RNA degradation rates[16].

Although diverse mechanistically, to our knowledge there are no known naturally occurring bacterial sRNA mechanisms that directly activate transcription[1,8,17], representing an opportunity to expand the mechanistic capability of this important class of gene regulators for synthetic biology. To address this gap, we pursued two mechanistic strategies to creating sRNA transcriptional activators, which we refer to as STARs. Using these strategies, we created four different STARs that display up to 94-fold transcriptional activation of their targets. We then systematically modified sequence features of this group to derive design principles for STAR function. By characterizing these STAR variants, we were able to derive a kinetic gene expression model that captures the essence of the sequence-function relationship of STAR activity. We demonstrated the utility of this model by using it to forward engineer a STAR that targets a terminator found in the *E. coli* genome. To test whether STARs could be used as components of higher-order RNA regulatory circuitry, we next confirmed that STARs were orthogonal to themselves as well as to a panel of RNA transcriptional repressors previously used to make RNA-only genetic circuits. Finally, we used the expanded functionality of STARs to construct previously unattainable RNA-only transcriptional logic gates, demonstrating the composability and utility of this new class of RNA regulator.

RESULTS

Engineering STARs with an anti-anti-terminator mechanism

As a starting point, we chose to re-engineer the pT181 transcriptional attenuator[18,19], which has been used in a number of synthetic biology contexts[14–16,20,21]. The pT181 attenuator is a sense RNA sequence that regulates transcription elongation through RNA structural rearrangements that either enable or inhibit the formation of an intrinsic transcription terminator hairpin upstream of a coding region (**Fig. 1a**)[18]. In the absence of a second RNA called the antisense sRNA, the attenuator folds so that an anti-terminator sequence sequesters the 5′ side of the intrinsic terminator hairpin, thereby inhibiting terminator formation and allowing transcription elongation (**Fig. 1a**). In its presence, the antisense sRNA interacts with the attenuator region containing the anti-terminator, which enables terminator formation that causes RNA polymerase (RNAP) to abort transcription near the beginning of the mRNA. The pT181 attenuator thus structurally encodes its repressive regulation (**Fig. 1a**), which we hypothesized we could invert through RNA structural engineering.

We initially adopted a previously reported anti-anti-terminator strategy used to design transcriptional activators that function *in vitro*[22]. In this strategy, an anti-anti-terminator sequence is fused upstream of the attenuator to sequester the anti-terminator itself, adding an additional layer of structural repression that inverts the overall attenuator function from repression to activation (**Fig. 1b**). To construct such a mechanism that functions *in vivo* in *E. coli* cells, we fused designed anti-anti-terminator sequences to the 5′ end of the pT181 attenuator sequence within the sense target RNA. These anti-anti-terminator sequences consisted of a region complementary to the anti-terminator followed by a sRNA recognition sequence (**Fig. 1b**) taken from modular RNA-RNA interaction domains that we have previously used to construct chimeric transcriptional attenuators[15]. Four variants of the sense target RNA were designed using the anti-anti-terminator mechanism and four different sRNA recognition sequences. For each of these, STAR antisense sRNA

School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York, USA *e-mail: jblucks@cornell.edu

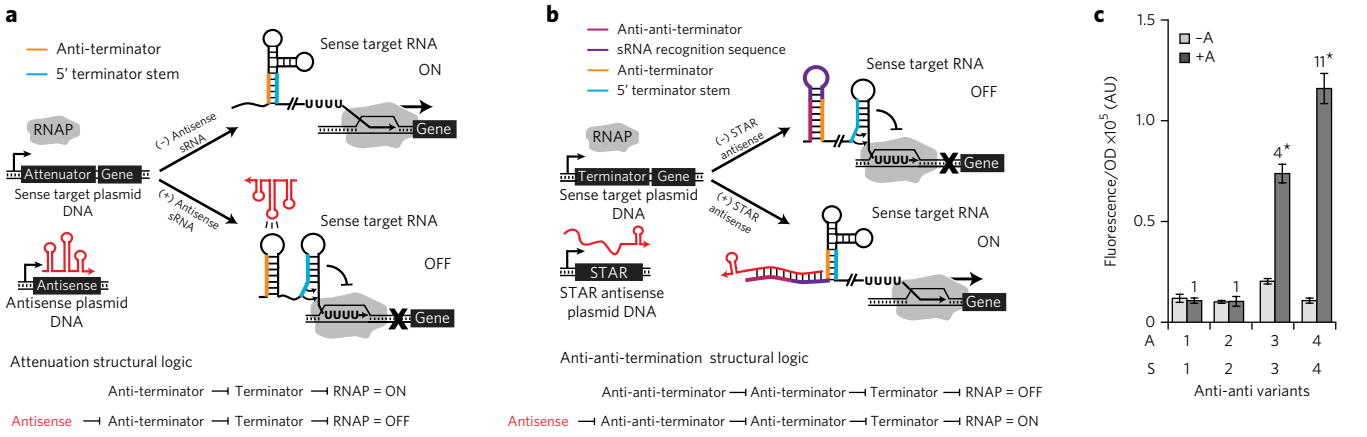

Figure 1 | Design and characterization of the anti-anti-terminator STAR mechanism. (**a**) A schematic of the wild-type pT181 attenuator mechanism[18,19]. The attenuator system is composed of two DNA templates, the sense plasmid and the antisense plasmid, both of which are transcribed by RNAP. In the absence of antisense sRNA, the anti-terminator sequence sequesters the 5′ terminator stem, preventing terminator formation and allowing downstream transcription (ON). When present, the antisense sRNA sequesters the anti-terminator of the sense RNA, allowing terminator formation, and preventing downstream transcription (OFF). The structural logic of the attenuation mechanism, depicting how the structural elements interact to give the overall regulatory function, is shown at the bottom. Blunt end arrows indicate repression. (**b**) A schematic of the engineered T181 anti-anti-terminator STAR. This mechanism inverts the structural logic of the attenuator mechanism through the addition of an anti-anti-terminator sequence. This anti-anti-terminator acts to add an additional structural repression connection in the structural logic diagram shown at the bottom. (**c**) Fluorescence characterization was performed (measured in units of fluorescence/optical density (OD) at 600 nm) on different anti-anti-terminator sense targets (S) in the absence of STAR antisense (−A) and presence of STAR antisense (+A). Data represents mean values of $n = 9$ biological replicas ± s.d. Fold activations are labeled above each A/S pair tested. Welch's t-test was performed on each −A/+A pair; $^*P < 0.05$, indicating conditions where the FL/OD for the +A condition was statistically significant from the −A condition. AU, arbitrary units.

sequences were designed to bind the sRNA recognition sequence and sequester the anti-anti-terminator so that transcription anti-termination (activation) was achieved (**Fig. 1b**).

To characterize transcriptional activation, plasmids were constructed whereby each sense target RNA was placed downstream of a constitutive promoter and upstream of a superfolder GFP (SFGFP)[23] coding sequence with its own RBS (**Supplementary Results**, **Supplementary Fig. 1**). STAR antisense sRNAs were expressed on a separate plasmid from a constitutive promoter and were followed by their own transcription terminators (t500 terminator; **Supplementary Fig. 1**). A no-antisense control plasmid consisting of the constitutive promoter followed directly by a transcription terminator (rrnB terminator (TrrnB); (**Supplementary Fig. 1**) was also constructed. For each sense target (S) plasmid tested, *E. coli* cells were transformed together with either the STAR antisense-expressing plasmid (A) or the no-antisense control plasmid, and SFGFP fluorescence (485 nm excitation and 520 nm emission) and optical density (600 nm) were measured for each culture. Of the four designs tested, two showed significant ($P < 0.05$) activation of gene expression in the presence of the STAR antisense, with STAR antisense and sense target pairs 3 and 4 (anti-anti A3/S3 and A4/S4) showing 3.6-fold (± 0.3) and 10.8-fold (± 1.5) activation, respectively (mean ± s.d.; **Fig. 1c**). These results demonstrated that we could successfully reengineer the structural logic of a sRNA transcriptional repression mechanism to convert it into a transcriptional activator.

Engineering STARs through direct anti-termination

Although we were successful, we realized that an alternative and simpler strategy for creating STARs would be to design STARs that directly act as anti-terminators (**Fig. 2a**). In this approach, STARs that contain an anti-terminator sequence are designed to prevent the formation of terminator hairpins placed upstream of coding regions within the target RNA. This has the effect of removing a layer of structural repression, which also accomplished our goal of inverting the overall attenuator function from repression to activation (**Fig. 2a**). To implement this design, we began by fusing the sequence encoding only the pT181 terminator hairpin upstream of the RBS-SFGFP region (**Fig. 2a**). In this configuration, the transcription terminator should form by default, preventing downstream transcription of the target RNA. STAR antisenses were designed to contain sequences complementary to the 5′ side of the terminator hairpin, so that when present, they would bind the nascent 5′ terminator region as *trans*-acting anti-terminators that allow transcription elongation. We initially created a series of STAR antisense and sense target variants that varied in length and sequence composition to achieve up to 12.4-fold (± 2.0) activation (T181 A4/S5; **Fig. 2b** and **Supplementary Fig. 2**).

To optimize further, we added additional complementary RNA sequences to both the STAR antisense and sense targets present upstream of the natural pT181 terminator to increase the potential interaction region between STAR and target. By adding this sequence in ~10-nt increments, we created six new STAR-target pairs, with T181 A6/S7 showing the strongest activation (18-fold (± 3.2); **Fig. 2b** and **Supplementary Fig. 3**). Notably, increasing the length of the interaction sequence between the STAR antisense and sense target only improved activation up to a point (**Supplementary Fig. 3**), which we hypothesized was because of a trade-off between increasing the binding strength of the intermolecular STAR antisense-sense target interaction and increasing the potential interference of intramolecular secondary structures of the individual strands.

To test whether this approach could be generalized to create additional STAR regulators, we applied this strategy to create STARs that target terminator hairpins from other transcriptional attenuators and transcriptional riboswitches (**Fig. 2c**,**d** and **Supplementary Fig. 4**). For transcriptional attenuator mechanisms, we focused on targeting the terminators from the pIP501, pCF10 and pAD1 plasmid attenuation systems[17]. Of these systems, the AD1 A1/S1 pair was the most promising, showing 7.5-fold (± 0.8) activation (**Supplementary Fig. 4**). To further optimize this system, we lengthened the STAR antisense-sense target interaction region by adding an additional sequence upstream of the natural terminator to this pair in ~10-nt increments, as before (**Fig. 2c** and

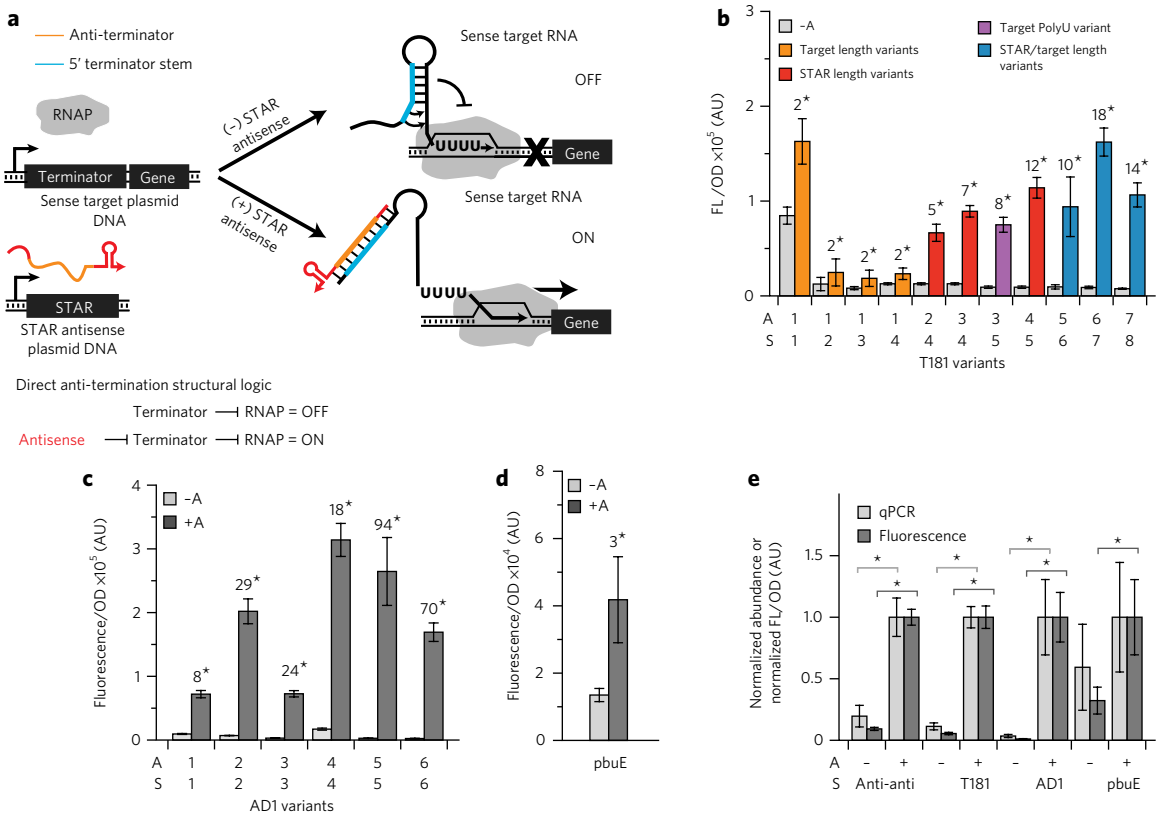

Figure 2 | Design and characterization of the direct anti-terminator STAR mechanism. (**a**) Schematic of the mechanism. In the absence of a STAR antisense, an intrinsic terminator is formed in the sense target RNA preventing transcription elongation (OFF). In the presence of the STAR antisense, the 5′ intrinsic terminator stem is sequestered by the STAR antisense, allowing downstream transcription by RNAP. This mechanism removes a structural repression connection from the attenuation mechanism (**Fig. 1a**) inverting the function from repression to activation, as shown at the bottom. (**b**–**d**) Fluorescence characterization was performed (measured in units of fluorescence (FL)/optical density (OD) at 600 nm) on STAR sense targets (S) in the absence of STAR antisense (–A) and presence of STAR antisense (+A) for the T181 (**b**), AD1 (**c**) and pbuE (**d**) systems. Fold activations are labeled above each A/S pair tested. In **b**, +A variants are color-coded according to sequence optimizations. Data represent mean values of $n = 9$ biological replicas ± s.d. (**e**) Comparison of qPCR and fluorescence characterization of the best STAR-target variants. Fluorescence data are from panels **b**–**d**. The ON condition for the qPCR and FL/OD data were normalized to 1 within each system. qPCR data represent mean values of $n = 3$ biological replicas ± s.d. For both qPCR and FL/OD data, a Welch's *t*-test was performed on each –A/+A pair; $^*P < 0.05$, indicating conditions where the FL/OD for the +A condition was statistically significant from that of the –A condition. AU, arbitrary units.

Supplementary Fig. 5). In this way, we were able to find a pair (AD1 A5/S5) that displayed 94-fold (± 26) activation.

For conversion of transcriptional riboswitches into STAR-target pairs, we focused on creating STARs from the terminator hairpins of three well-characterized riboswitches shown to have a high degree of modularity: metH[24], xpt-pbuX[25] and pbuE[24]. Of these, the pbuE STAR showed 3.1-fold (± 1.0) activation (**Fig. 2d** and **Supplementary Fig. 4**). Optimizations were attempted as before, though no greater fold change in activation was achieved by increasing the STAR-target interaction sequence (**Supplementary Fig. 6**).

To corroborate that these systems regulate expression through transcriptional activation, we used quantitative PCR (qPCR) to determine the steady-state level of SFGFP mRNA in the presence and absence of STAR antisense expression for the best activators (**Fig. 2e**). For clarity, the STAR-target RNAs for these systems are denoted anti-anti (Anti-anti.A4/S4), T181 (T181.A6/S7), AD1 (AD1.A5/S5) and pbuE (pbuE.A1/S1). For the anti-anti, T181 and AD1 STAR-target pairs, we observed a statistically significant ($P < 0.05$) increase in the steady-state abundance of SFGFP mRNA in the presence of their STAR antisenses, thus corroborating that these systems operated through transcriptional activation. We note that for these systems, we observed small discrepancies between qPCR quantifications of SFGFP mRNA and the measured SFGFP fluorescence that we attribute to the mass normalization of qPCR samples to total RNA concentrations, which can vary depending on the overall gene expression in each condition tested. The pbuE system showed an overall increase in SFGFP mRNA abundance in the presence of its STAR antisense sequence in the qPCR experiments, though it was not statistically significant ($P > 0.05$). This is most likely because of the low fold activation of this system and the inherent noise of the qPCR experiment.

To further demonstrate that the observed *in vivo* transcriptional activation of the STAR-target systems is not due to an off-target or nonspecific gene expression response in the cell, we tested their function using *in vitro* transcription and translation (TX-TL) reactions. TX-TL reactions contain all of the necessary cellular machinery for gene expression but contain no endogenous genomic DNA templates, and so they provide a reduced gene expression system independent of other host genes[26–29]. Thus STARs are only expected to activate gene expression in TX-TL reactions if their function is not dependent on other genomic targets. We observed activation for the AD1 and T181 direct anti-terminator STARs in TX-TL reactions (**Supplementary Fig. 7**), though the levels of activation differed from those in *in vivo* experiments, most likely because of the known differences between *in vivo* and TX-TL systems[16].

Collectively, these results demonstrated that we could create sRNA-mediated transcriptional activators through a *trans*-acting anti-terminator mechanism. Although mechanistically simple, to

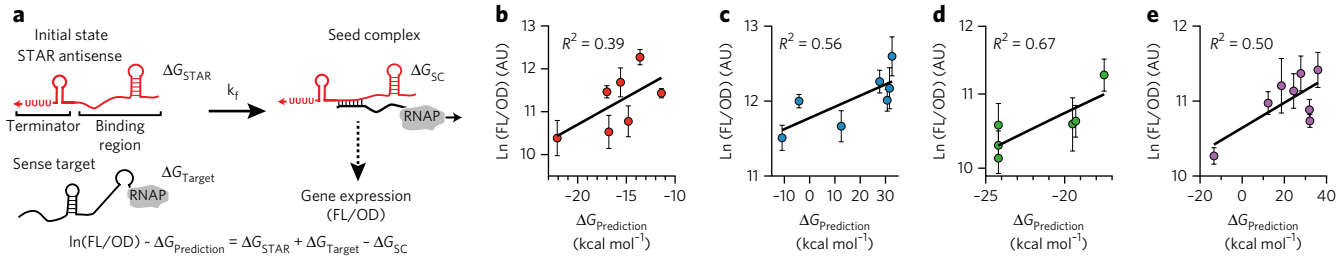

Figure 3 | STAR design principles. (**a**) A kinetic model of STAR anti-termination showing the hypothesized interactions between the STAR antisense and the sense target region. Our model considers an initial state and a SEED complex (SC). The initial state consists of a fully transcribed STAR antisense, with free energy ΔG_{STAR}, and the upstream portion of the sense target that is transcribed before the transcription elongation decision has been made, with free energy ΔG_{Target}. These interact with a forward rate k_f to form the SC with free energy ΔG_{SC}. Under the hypothesis that the formation of the SC is sufficient to allow transcription elongation and downstream gene expression, the natural log of observed gene expression (fluorescence (FL)/optical density (OD)) is linearly related to $\Delta G_{Prediction}$, which is the difference in free energies between the initial state and SC (**Supplementary Note**). (**b**–**e**) Observed correlations between fluorescence characterization (measured in units of natural log FL/OD at 600 nm) and $\Delta G_{Prediction}$ of different length STARs against the optimally functioning target region from the T181 (**b**), AD1 (**c**), pbuE (**d**) systems (shown in **Fig. 2e**) and the intrinsic terminator of the *E. coli* ribA gene (**e**). Data represent mean values of $n = 9$ biological replicas ± s.d. The R^2 correlation coefficient between ln(FL/OD) and $\Delta G_{Prediction}$ is shown in the upper left of each plot. AU, arbitrary units.

our knowledge no naturally occurring sRNA has been shown to act at this level of regulation by such a mechanism.

Development of a sequence-function model of STARs

One of the advantages of RNA regulators over their protein counterparts is the wealth of available computational structure prediction tools that can serve as a starting point for model-guided RNA regulator design[1]. Recently, these tools have been combined with mechanistic models of RNA regulation to rationally design and optimize a range of systems that control translation, including RBSs[30], sRNAs[3,31,32] and riboswitches[4]. We sought to use this approach for the direct anti-terminator STAR mechanism to develop a kinetic model that could explain the range in activation we observed as a function of STAR antisense and target RNA sequence. To develop this model, we first considered the different RNA structural states formed as the STAR antisense interacts with the sense target RNA (**Fig. 3a**). On the basis of previous work characterizing the RNA-RNA binding pathways involved in kissing hairpin translational regulators[33], we hypothesized the presence of three structural states in the STAR mechanism: the initial state (IS), consisting of the individually folded STAR antisense and sense target; an extended duplex that consists of perfect base pairing between STAR and target; and a seed complex (SC) in which STAR-target interactions are initiated that serves as an intermediate state between the initial state and the extended duplex (**Fig. 3a** and **Supplementary Note**). Because the STAR-mediated transcriptional regulatory decision must happen during transcription elongation by RNAP, we hypothesized that seed complex formation is much faster than extended duplex formation, which has been observed in the pT181 transcriptional repression system[18]. We further hypothesized that seed complex formation is sufficient to prevent the formation of the terminator hairpin and enact the regulatory decision, and thus the rate of overall gene expression is proportional to the rate of seed complex formation (**Supplementary Note**).

An analysis of this model predicted that the observed STAR-mediated gene expression was a function of the free energies of the different RNA structural states (**Supplementary Note** and **Supplementary Fig. 8**). Specifically, this analysis predicted that the natural log of the observed gene expression (fluorescence (FL)/optical density (OD)) is linearly related to the difference in free energy between the initial state and the seed complex: $\ln(FL/OD) \sim \Delta G_{STAR} + \Delta G_{Target} - \Delta G_{SC}$. This free energy difference naturally reflects the competing effects of intramolecular base pairs within the STAR and target that need to be broken before the formation of intermolecular base pairs that lead to the seed complex and, ultimately, transcription activation.

The only thing that had to be addressed before using this model to explain the variation in observed STAR activation was determining which RNA sequences to use in computational folding algorithms to approximate the ΔG terms. This amounted to choosing the length of the sense target strand that comprised the initial state and the region of interaction that characterized the seed complex. To investigate this, we characterized the fluorescence observed from a combination matrix of different-length STAR antisense and sense target variants by challenging each different-length target with each different-length STAR antisense for the T181, AD1 and pbuE

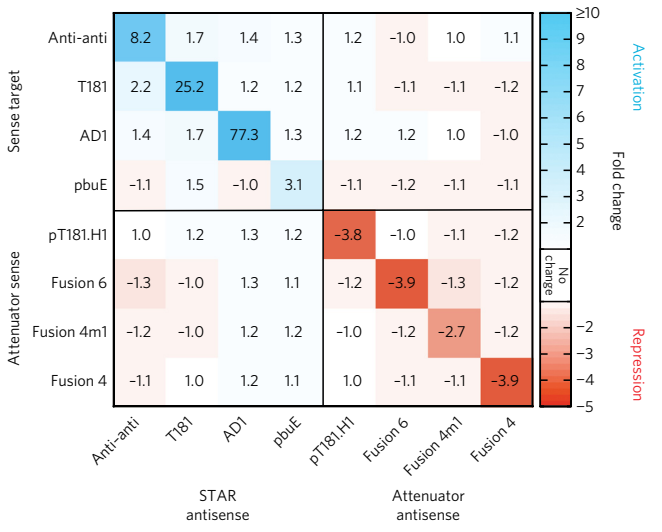

Figure 4 | Determining the orthogonality of STAR regulators and transcriptional attenuators. Characterization of an 8 × 8 orthogonality matrix of four different STAR regulators created in this work and four transcriptional repressors engineered in previous studies[14,15]. Each element of the matrix represents the fold change of gene expression for the indicated antisense/sense target plasmid combination compared to a no-antisense/sense target plasmid condition. Fold changes for different combinations are written within each of the elements of the matrix. Fluorescence characterization (measured in units of fluorescence/optical density at 600 nm) was used to calculate fold change, which is represented by a color scale in which values greater than or equal to tenfold are blue (activation), singlefold is white (no activation or repression) and negative fivefold is red (repression). FL/OD plots for each combination are shown in **Supplementary Figure 12**. Data represents mean values of $n = 9$ biological replicas ± s.d.

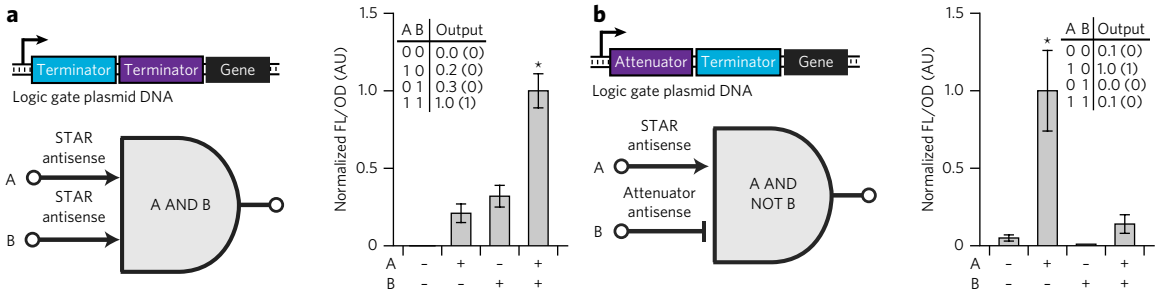

Figure 5 | Characterization of novel RNA-only transcriptional logic gates. (**a**,**b**) DNA template (upper left), logic schematic (lower left) and fluorescence characterization (measured in units of fluorescence (FL)/optical density (OD) at 600 nm) (right) of the A AND B logic gate (**a**) and the A AND NOT B logic gate (**b**). Fluorescence data were normalized to 1 for the ON condition in the presence of both antisenses A and B in **a** or in the presence of only antisense A in **b**. Insets show the measured output (normalized FL/OD measurements) performances of the logic gates and the expected values for a perfect digital logic gate (parentheses). Data represent mean values of $n = 9$ biological replicas ± s.d. Welch's t-test was performed on each ON/OFF condition; $^*P < 0.05$, indicating conditions where the FL/OD for the ON condition was statistically significant from all OFF conditions. AU, arbitrary units.

systems. As predicted by our model, there was an optimum STAR antisense length above which no further increase in fluorescence was observed for each sense target variant (**Supplementary Fig. 9** and **Supplementary Note**). We then used this data to determine the specific sequences needed to calculate the $\Delta G_{\text{Prediction}}$ terms of our model. For each system, the initial state consisted of the entire STAR antisense sequence and the 5′ region of the sense target, including the 5′ half of the terminator hairpin (**Fig. 3a**), whereas the seed complex region was found by testing predicted correlations between ln(FL/OD) and the $\Delta G_{\text{Prediction}}$ term using different regions. We observed a consistent positive correlation between the measured STAR-mediated gene expression and the $\Delta G_{\text{Prediction}}$ term (**Fig. 3b–d** and **Supplementary Fig. 10**). For the T181, AD1 and pbuE systems used in **Figure 2e**, R^2 was 0.39, 0.56 and 0.67, respectively, and we observed equally strong correlations for predictions when applied to different-length sense target variations (**Supplementary Fig. 10**). These results indicated that our kinetic model captured the essence of the STAR direct anti-terminator mechanism.

To further validate our model, we sought to use it in designing new STARs that target alternative sources of intrinsic terminators. As all of our current STAR antisense sequences target terminators present in existing transcriptional switches, we chose to focus on targeting intrinsic terminators present in the *E. coli* genome at the ends of genes to test whether STARs could target this class of terminator. We placed intrinsic terminators upstream of a strong RBS and SFGFP in our two-plasmid system and constructed STAR antisense sequences to target their 5′ halves. Initial screening identified several functional variants, and the strongest activation was a 2.3-fold (± 0.37) increase, observed for the intrinsic terminator of the GTP cyclohydrolase II ribA gene (**Supplementary Fig. 11a**). Using our mechanistic model of STAR activation, we designed seven more STAR antisense variants that we predicted would cover a range of activation levels (**Supplementary Fig. 11b**). A comparison of the observed fluorescence caused by these STAR antisense variants with the $\Delta G_{\text{Prediction}}$ term for this system revealed a good correlation (R^2 = 0.50), consistent with the results from our previous model (**Fig. 3e**). Furthermore, the largest $\Delta G_{\text{Prediction}}$ term successfully identified our optimal STAR antisense, which had a 7.2-fold (± 1.9) activation of expression (ribA A6 from **Supplementary Fig. 11b**). It should be noted that because the sense target RNA sequence was constant in this series, the ΔG_{Target} term was removed from our model for these predictions, which amounts to a shift in all $\Delta G_{\text{Prediction}}$ terms. These results demonstrated that we could successfully apply our model to aid in the design of new STARs and that STARs could be designed to target naturally occurring intrinsic terminators not derived from RNA regulators.

Determining the orthogonality of STARs

We next sought to test whether STARs could be used as components of higher-order RNA regulatory circuitry. A prerequisite to such utility is their orthogonality to each other, i.e., the ability of a STAR antisense to only activate its cognate target without cross-talk to other targets. To determine STAR orthogonality, we measured the fold activation of all possible STAR-target pairs among the best direct anti-terminator activators from the T181, AD1 and pbuE systems and the best anti-anti-terminator mechanism (**Fig. 4** and **Supplementary Fig. 12**). Noting that a singlefold change is no activation, we observed a high degree of orthogonality between these activators. The one exception was for the pbuE sense target, which was activated 1.5-fold by the T181 STAR antisense (compared to its cognate, which had 3.1-fold activation), although this result was most likely biased by the overall low fold activation of this activator. To our surprise, we observed orthogonality between the two pT181-derived activators, the T181 (direct anti-terminator mechanism) and anti-anti (anti-anti-terminator mechanism) activators, suggesting that our changes to the STAR antisense and sense target pairs were substantial enough to allow independent regulation. This unanticipated result led us to hypothesize that these activators may be orthogonal to a series of previously published transcriptional repressors (attenuators) engineered from the pT181 attenuator[14,15]. To test this, an 8 × 8 orthogonality matrix of attenuator and activator sense and antisense sRNAs were tested as before, and fold change relative to the no-antisense control was determined (**Fig. 4** and **Supplementary Fig. 12**). We observed a high level of orthogonality between the noncognate pairs of sense and antisense from the attenuator and activator systems. Although we observed some small levels of cross-talk, most fold changes were within error of the no-antisense control (**Supplementary Fig. 12**). These results demonstrate that the STARs are highly orthogonal to themselves and to the existing sRNA transcriptional attenuator libraries, suggesting that these activators in fact expand the versatility of the RNA transcriptional regulatory toolbox required for engineering RNA-only genetic networks[14].

Another type of orthogonality that is only beginning to be studied in synthetic biology is orthogonality to the host cell[1]. To determine these effects, we performed RNA-seq on total RNA isolated from *E. coli* cells transformed with either one of the four STARs or the no-antisense control plasmid. It should be noted that *E. coli* strain K12 MG1655 was used, in which we showed our STAR antisense sRNAs to be functional (**Supplementary Fig. 13**). Differential gene expression analysis between a specific STAR antisense condition and our no-antisense control showed that there were global changes in gene expression due to STAR antisense expression, although the majority of genes are unaffected (**Supplementary Fig. 14**).

We also found this to be true for the best ribA STAR antisense (**Supplementary Fig. 14**). As each STAR antisense seems to behave similarly, these observed changes in gene expression could be due to a general response to the presence of a highly expressed RNA.

Applying STARs to construct novel RNA-only logic gates

Finally, inspired by the orthogonality of STARs to each other and RNA transcriptional repressors (attenuators), we aimed to construct new RNA-only transcriptional logic gates that were previously unattainable owing to the lack of sRNA activators. Genetic logic gates are necessary network elements for constructing circuits that integrate signals and process information to control cellular behavior[34] and have been used in a number of biotechnological applications[35,36]. However, the only synthetic RNA-mediated transcriptional logic gates that have been demonstrated are NOR gates, which only allow gene expression when none of the gate inputs are present[14,37]. We therefore sought to construct two new RNA-only transcriptional logic gates that combined both RNA transcriptional activators and attenuators: an A AND B gate (**Fig. 5a**) and an A AND NOT B gate (**Fig. 5b**). These logic gates were constructed by transcriptionally fusing STAR target sense and attenuator sequences in series and were tested against all possible input combinations of antisense sRNAs (**Fig. 5a**,**b** and **Supplementary Fig. 1**). This characterization revealed that both the A AND B and A AND NOT B gates were functional; the AND gate was only ON when both inputs were present, whereas the A AND NOT B was in the ON state when only the A input was present. These results provided further evidence that STARs act on the transcriptional level and demonstrated that STARs can be used within more sophisticated RNA genetic circuitry devices.

DISCUSSION

Our results showed for what is to our knowledge the first time that sRNAs can directly activate transcription *in vivo*. As such, this represents both an expansion of the known mechanistic capabilities of sRNA regulation and a valuable addition to the rapidly growing RNA regulatory toolbox[1,38]. As STARs are highly orthogonal to each other and to a family of RNA transcriptional repressors, they enabled construction of new RNA-only transcriptional logic gates that were previously unattainable. These RNA devices will provide new tools for engineering cells to perform signal processing and integration and ultimately higher-order genetic regulatory functions. As such, we believe the work described here will enhance a broad array of synthetic biology and biotechnological applications.

Our mechanistic model was successfully used to understand the sequence-function relationships of STAR performance and will aid future development of new STAR-target pairs. Although this model explains the trend observed in STAR performance as a function of STAR and target sequence, it does have several limitations. First, the STAR target RNAs are being actively transcribed during the regulatory decision, and thus they exist in a highly nonequilibrium folding regime. Our underlying hypotheses modeled this by considering intermediate lengths of the target RNA sequence to be present in the initial state and seed complex and further used thermodynamic minimum free energy calculations to estimate the free energy of these states. Both of these are approximations that could be refined by incorporating more detailed models of cotranscriptional folding that incorporate the kinetics of RNA folding and the nonuniform dynamics of RNAP transcription. Second, our model uses free energy calculations that use parameters derived from *in vitro* conditions to model *in vivo* gene expression processes[39]. Third, STARs represent a completely new mode of RNA regulation, and our model represents a coarse-grained and heuristic understanding of this mechanism. Further studies are required to gain a better mechanistic understanding of this regulation, which could enable a more fine-grained model of how the STAR antisense and sense target RNAs interact that could ultimately yield more accurate predictions. Nevertheless, this model was able to be successfully applied to develop a new STAR antisense that targets an intrinsic terminator derived from the *E. coli* genome, and it accurately captures the overall trend in STAR performance as a function of STAR and target sequence. Although we believe further work is needed to make this model more accurate, we believe we have captured the essence of the design principles of STAR function, which adds to the increasing body of work highlighting the ability of RNA regulators of gene expression to be engineered and designed[1].

Within the field of RNA synthetic biology, one of the most recent advances in RNA regulation has been the repurposing of the bacterial clustered regularly interspaced short palindromic repeat (CRISPR) defense system as a transcriptional regulatory system. CRISPR interference (CRISPRi) relies upon the use of CRISPR small guide RNAs, in combination with a dead catalytic mutant of the CRISPR Cas9 protein (dCas9), to target specific DNA sequences for transcriptional repression or activation in a variety of organisms[40–44]. As such, STARs and CRISPRi are comparable in several aspects. As they both regulate transcription via RNA-guided targeting, they both have been shown to be designable because of the simplicity of Watson-Crick base pairing and the availability of RNA secondary structure prediction tools to aid rational design. However, there are several distinctions between the mechanisms. First, in *E. coli*, transcriptional activation by CRISPRi requires both a fusion of the dCas9 protein to the ω-fragment of RNAP and a strain containing a knockout of the endogenous ω-fragment of RNAP that currently shows low fold activation on strong promoters, potentially limiting applications[40]. Second, in bacterial systems, CRISPRi regulators have not yet been shown to be able to be used to construct logic gates or networks, thus limiting their application in synthetic circuit design. Although many of these issues have been resolved in higher eukaryotes[41–44], much work is required to address these limitations for bacterial systems. In contrast, CRISPRi has clear application for strain engineering because of its ability to effectively repress or activate genes located on the host genome[41]. Although we were able to make a STAR that targets a natural terminator, we could not detect a change in expression of genes surrounding this terminator on the genome with RNA-seq (**Supplementary Fig. 14**). It should be noted that our original intent in creating STARs was to expand the types of RNA-only logics and networks that can be constructed, an application they appear to be ideally suited to. We thus view STARs and CRISPRi as complementary elements of the RNA synthetic biology toolbox for bacterial systems.

In summary, this work represents a major advance in our ability to engineer RNA regulators by effectively reengineering the structural logic of a natural RNA transcriptional repressor as well as the ability to convert intrinsic *E. coli* terminators into STAR regulators. Because of their mechanistic novelty, STARs effectively increase the parameter space that can be explored in applications ranging from sRNA-based metabolic pathway optimization[9] all the way to engineering RNA-only genetic networks that control the timing of gene expression inside cells[16]. Finally, the relative simplicity by which this was achieved suggests that sRNA transcriptional activation may be a natural mechanism of gene regulation yet to be discovered.

Received 15 May 2014; accepted 25 November 2014;
published online 2 February 2015

METHODS

Methods and any associated references are available in the online version of the paper.

References

1. Chappell, J. *et al.* The centrality of RNA for engineering gene expression. *Biotechnol. J.* **8**, 1379–1395 (2013).

2. Carothers, J.M., Goler, J.A., Juminaga, D. & Keasling, J.D. Model-driven engineering of RNA devices to quantitatively program gene expression. *Science* **334**, 1716–1719 (2011).
3. Rodrigo, G., Landrain, T.E. & Jaramillo, A. *De novo* automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proc. Natl. Acad. Sci. USA* **109**, 15271–15276 (2012).
4. Wachsmuth, M., Findeiss, S., Weissheimer, N., Stadler, P.F. & Morl, M. *De novo* design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res.* **41**, 2541–2551 (2013).
5. Xayaphoummine, A., Viasnoff, V., Harlepp, S. & Isambert, H. Encoding folding paths of RNA switches. *Nucleic Acids Res.* **35**, 614–622 (2007).
6. Lucks, J.B. *et al.* Multiplexed RNA structure characterization with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc. Natl. Acad. Sci. USA* **108**, 11063–11068 (2011).
7. Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. & Weissman, J.S. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures *in vivo*. *Nature* **505**, 701–705 (2014).
8. Storz, G., Vogel, J. & Wassarman, K.M. Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell* **43**, 880–891 (2011).
9. Na, D. *et al.* Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* **31**, 170–174 (2013).
10. Sharma, V., Yamamura, A. & Yokobayashi, Y. Engineering artificial small RNAs for conditional gene silencing in *Escherichia coli*. *ACS Synth. Biol.* **1**, 6–13 (2012).
11. Callura, J.M., Cantor, C.R. & Collins, J.J. Genetic switchboard for synthetic biology applications. *Proc. Natl. Acad. Sci. USA* **109**, 5850–5855 (2012).
12. Friedland, A.E. *et al.* Synthetic gene networks that count. *Science* **324**, 1199–1202 (2009).
13. Isaacs, F.J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* **22**, 841–847 (2004).
14. Lucks, J.B., Qi, L., Mutalik, V.K., Wang, D. & Arkin, A.P. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *Proc. Natl. Acad. Sci. USA* **108**, 8617–8622 (2011).
15. Takahashi, M.K. & Lucks, J.B. A modular strategy for engineering orthogonal chimeric RNA transcription regulators. *Nucleic Acids Res.* **41**, 7577–7588 (2013).
16. 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.* doi:10.1021/sb400206c (12 March 2014).
17. Brantl, S. Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr. Opin. Microbiol.* **10**, 102–109 (2007).
18. Brantl, S. & Wagner, E.G. Antisense RNA-mediated transcriptional attenuation: an *in vitro* study of plasmid pT181. *Mol. Microbiol.* **35**, 1469–1482 (2000).
19. Kumar, C.C. & Novick, R.P. Plasmid pT181 replication is regulated by two countertranscripts. *Proc. Natl. Acad. Sci. USA* **82**, 638–642 (1985).
20. Qi, L., Lucks, J.B., Liu, C.C., Mutalik, V.K. & Arkin, A.P. Engineering naturally occurring *trans*-acting non-coding RNAs to sense molecular signals. *Nucleic Acids Res.* **40**, 5775–5786 (2012).
21. Sakai, Y. *et al.* Improving the gene-regulation ability of small RNAs by scaffold engineering in *Escherichia coli*. *ACS Synth. Biol.* **3**, 152–162 (2014).
22. Dawid, A., Cayrol, B. & Isambert, H. RNA synthetic biology inspired from bacteria: construction of transcription attenuators under antisense regulation. *Phys. Biol.* **6**, 025007 (2009).
23. Pédelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C. & Waldo, G.S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–88 (2006).
24. Ceres, P., Trausch, J.J. & Batey, R.T. Engineering modular 'ON' RNA switches using biological components. *Nucleic Acids Res.* **41**, 10449–10461 (2013).
25. Ceres, P., Garst, A.D., Marcano-Velazquez, J.G. & Batey, R.T. Modularity of select riboswitch expression platforms enables facile engineering of novel genetic regulatory devices. *ACS Synth. Biol.* **2**, 463–472 (2013).
26. Sun, Z.Z. *et al.* Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp.* **2013**, 16 e50762 (2013).
27. 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).
28. 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).
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 Res.* **41**, 3471–3481 (2013).
30. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
31. Mutalik, V.K., Qi, L., Guimaraes, J.C., Lucks, J.B. & Arkin, A.P. Rationally designed families of orthogonal RNA regulators of translation. *Nat. Chem. Biol.* **8**, 447–454 (2012).
32. 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).
33. Kolb, F.A. *et al.* Progression of a loop-loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. *EMBO J.* **19**, 5905–5915 (2000).
34. Anderson, J.C., Voigt, C.A. & Arkin, A.P. Environmental signal integration by a modular AND gate. *Mol. Syst. Biol.* **3**, 133 (2007).
35. Ellis, T., Wang, X. & Collins, J.J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol.* **27**, 465–471 (2009).
36. Kotula, J.W. *et al.* Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci. USA* **111**, 4838–4843 (2014).
37. Liu, C.C. *et al.* An adaptor from translational to transcriptional control enables predictable assembly of complex regulation. *Nat. Methods* **9**, 1088–1094 (2012).
38. Qi, L.S. & Arkin, A.P. A versatile framework for microbial engineering using synthetic non-coding RNAs. *Nat. Rev. Microbiol.* **12**, 341–354 (2014).
39. Mathews, D.H. & Turner, D.H. Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.* **16**, 270–278 (2006).
40. Bikard, D. *et al.* Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41**, 7429–7437 (2013).
41. Qi, L.S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
42. Farzadfard, F., Perli, S.D. & Lu, T.K. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth. Biol.* **2**, 604–613 (2013).
43. Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P. & Lu, T.K. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell* **54**, 698–710 (2014).
44. Kiani, S. *et al.* CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods* **11**, 723–726 (2014).

Acknowledgments

The authors acknowledge J. Roberts, E. Strobel, A. Stroock and the Lucks Lab members for helpful discussions. We would also like to thank C. Trapnell for help with RNA-seq experimental design and analysis. We also thank J. Peters (Department of Microbiology, Cornell University) for providing *E. coli strain* K12 MG1655. Finally, we would like to thank D. Tapias-Rojas for preliminary work on targeting naturally occurring intrinsic terminators. This material is based on work supported by the National Science Foundation Graduate Research Fellowship Program (grant no. DGE-1144153 to M.K.T.), the Defense Advanced Research Projects Agency Young Faculty Award (DARPA YFA; no. N66001-12-1-4254 to J.B.L.) and an Office of Naval Research Young Investigators Program Award (ONR YIP; no. N00014-13-1-0531 to J.B. L.). J.B.L. is an Alfred P. Sloan Research Fellow.

Author contributions

J.C., M.K.T. and J.B.L. conceived the ideas, designed the experiments and wrote the manuscript. J.C. and M.K.T. performed the experiments.

Competing financial interests

The authors declare competing financial interests: details accompany the online version of the paper.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to J.B.L.

ONLINE METHODS

Plasmid assembly. All plasmids used in this study can be found in **Supplementary Table 1** with key sequences provided in **Supplementary Tables 2–4**. All sense and antisense plasmids were constructed using inverse PCR (iPCR). As shown in **Supplementary Figure 1**, all sense plasmids had the p15A origin and chloramphenicol resistance, and all antisense plasmids had the ColE1 origin and ampicillin resistance. All assembled plasmids were verified using DNA sequencing.

Strains, growth media and *in vivo* bulk fluorescence measurements. Fluorescence measurement experiments were performed in *E. coli* strain TG1 except for those in **Supplementary Figure 13**, for which *E. coli* strain K12 MG1655 was used. Experiments were performed for nine biological replicas collected over three separate days unless otherwise stated in the figure legend. For each day of *in vivo* bulk fluorescence measurements, plasmid combinations were transformed into chemically competent *E. coli* TG1 cells and plated on LB + Agar (Difco) plates containing 100 mg/ml carbenicillin and 34 mg/ml chloramphenicol and incubated approximately 17 h overnight at 37 °C. Plates were taken out of the incubator and left at room temperature for approximately 7 h. Three colonies were used to inoculate three cultures of 300 µl of LB containing carbenicillin and chloramphenicol at the concentrations indicated above in a 2-ml 96-well block (Costar), and they were grown for approximately 17 h overnight at 37 °C at 1,000 r.p.m. in a VorTemp 56 (Labnet) bench top shaker. Four microliters of each overnight culture were then added to 196 µl (1:50 dilution) of supplemented M9 minimal medium (1 × M9 minimal salts, 1 mM thiamine hydrochloride, 0.4% glycerol, 0.2% casamino acids, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$) containing the selective antibiotics. Cells were then grown for 4 h for all data except for those in **Figures 4** and **5** and **Supplementary Figure 12**, for which cells were grown for 5 h in the same conditions as the overnight culture. Fifty microliters of this culture were then transferred to a 96-well plate (Costar) containing 50 µl of phosphate-buffered saline (PBS). SFGFP fluorescence (FL; 485 nm excitation, 520 nm emission) and optical density (OD) at 600 nm were then measured using a SynergyH1 plate reader (Biotek).

Bulk fluorescence data analysis. On each 96-well block, there were two sets of controls; a medium blank (M9) and *E. coli* TG1 cells (*E. coli* K12 MG1655 for **Supplementary Fig. 13**) that do not produce SFGFP (transformed with control plasmids JBL001 and JBL002; **Supplementary Tables 3** and **4** and **Supplementary Fig. 1**). The block contained three replicates of each control. OD and FL values for each colony were first corrected by subtracting the corresponding values of the medium blank. The ratio of FL to OD (FL/OD) was then calculated for each well (grown from a single colony), and the mean FL/OD of TG1 cells without SFGFP was subtracted from each colony's FL/OD value. Three biological replicas were collected from one independent transformation, with three colonies characterized per transformation (nine colonies total). Mean FL/OD values were calculated over replicas, and error bars represent s.d. For characterization of orthogonality (**Fig. 4** and **Supplementary Fig. 12**), the fold change (activation or repression) for each pair was determined by dividing the FL/OD of cells containing both the sense and antisense plasmids (ON) by the FL/OD of cells containing the sense plasmid and a no-antisense control plasmid (OFF). If this number was less than 1, indicating repression, the negative reciprocal was taken to give the fold repression, i.e., 0.20 becomes −5-fold repression. A Welch's *t*-test was performed to determine statistical significance ($P < 0.05$) between different conditions; exact comparisons used are in figure legends.

Total RNA extraction for quantitative PCR. For all extraction of total RNA for quantitative PCR (qPCR) experiments, *E. coli* strain TG1 was used. Plasmids were transformed, and subsequent colonies were grown overnight as described for *in vivo* bulk fluorescence measurements. For each biological replica, 20 µl of a single overnight culture was added to three wells containing 980 µl (1:50 dilution) of supplemented M9 minimal medium containing the selective antibiotics and grown for 4 h at the same conditions as the overnight cultures. For each plasmid combination, 500 µl of cells were removed from three wells (grown from one colony) and combined into a 1.6-ml tube and pelleted by centrifugation at 13,000 r.p.m. for 1 min. The supernatant was removed, and the remaining pellet was resuspended in 750 µl of Trizol reagent (Life Technologies), homogenized by repetitive pipetting, incubated at room temperature for 5 min and stored at −80 °C for approximately 17 h. These samples were defrosted on ice, 150 µl of chloroform (Sigma-Aldrich) was added, and the samples were mixed for 15 s and incubated at room temperature for 3 min. Following incubation, the samples were centrifuged for 15 min at 12,000*g* at 4 °C, and 200 µl of the top aqueous layer was removed. One microliter of glycogen (20 µg/µl; Life Technologies) and 375 µl of isopropanol were added to the aqueous phase, and the sample was incubated at room temperature for 10 min and centrifuged for 15 min at 15,000 r.p.m. at 4 °C. Following centrifugation, the isopropanol was carefully removed from the total RNA/glycogen pellets, washed in 600 µl of chilled 70% ethanol (EtOH) and centrifuged for 2 min at 15,000 r.p.m. at 4 °C. EtOH was removed, and tubes were centrifuged for another 2 min at 15,000 r.p.m. at 4 °C to ensure that all of the ethanol was effectively removed. Pellets were resuspended in 20 µl of RNase free double-distilled water (ddH_2O).

DNase treatment of total RNA for qPCR. Purified total RNA samples were quantified by the Qubit Fluorometer (Life Technologies) and were diluted to a concentration of 10 ng/µl in a total of 10 µl RNase free ddH_2O and digested by Turbo DNase (Life Technologies) according to the manufacturer's protocol. After digestion, 150 µl of RNase free ddH_2O and 200 µl phenol/chloroform (Acros Organics) was added, and the sample was vortexed for 10 s and incubated for 3 min at room temperature and centrifuged for 10 min at 15,000 r.p.m. at 4 °C. After centrifugation, 190 µl of the top aqueous layer was carefully removed, 190 µl of chloroform was added, and samples were vortexed for 10 s, incubated for 3 min at room temperature and centrifuged for 10 min at 15,000 r.p.m. at 4 °C. After centrifugation, 170 µl of the top aqueous layer was carefully removed, 170 µl of chloroform was added, and samples were vortexed for 10 s, incubated for 3 min at room temperature and centrifuged for 10 min at 15,000 r.p.m. at 4 °C. After centrifugation, 120 µl of the top aqueous layer was carefully removed and added to 1 µl glycogen, 360 µl of chilled 100% EtOH and 12 µl of 3 M sodium acetate, pH 5. Samples were vortexed for 10 s and stored at −80 °C for 1 h. Samples were then centrifuged for 30 min at 15,000 r.p.m. at 4 °C. Supernatant was removed, and the pellets were washed in 600 µl of chilled 70% EtOH. Samples were then centrifuged for 2 min at 15,000 r.p.m. at 4 °C, and the EtOH was removed. Samples were recentrifuged for 2 min at 15,000 r.p.m. at 4 °C, and residual EtOH was removed, and pellets were air-dried for 10 min and eluted in 10 µl RNase fee ddH_2O.

Normalization of total RNA, reverse transcription and qPCR measurements. To enable comparison between different samples, each DNase treated sample was normalized to contain the same total RNA concentration. Each sample was quantified by Qubit Fluorometer, and the sample was diluted to 0.25 ng/µl of total RNA in 20 µl RNase free ddH_2O. One microliter of this total RNA, 1 µl of 2 µM reverse transcription primer (**Supplementary Table 5**), 1 µl of 10 mM of dNTPs (New England BioLabs) and RNase-free ddH_2O (up to 6.5 µl) were incubated for 5 min at 65 °C and cooled on ice for 5 min. 0.25 µl of Superscript III reverse transcriptase (Life Technologies), 1 µl of 100 mM Dithiothreitol (DTT), 1× first-strand buffer (Life Technologies), 0.5 µl RNaseOUT (Life Technologies) and RNase free H_2O up to 3.5 µl were then added, and the solution was incubated at 55 °C for 1 h, 75 °C for 15 min and then stored at −20 °C. qPCR was performed using 5 µl of Maxima SYBR green qPCR master mix (Thermo Scientific), 1 µl of cDNA and 0.5 µl of 2 µM SFGFP qPCR primers (**Supplementary Table 5**) and RNase-free ddH_2O up to 10 µl. A ViiA 7 real-time PCR machine (Applied Biosystems) was used for data collection using the following PCR program: 50 °C for 2 min, 95 °C for 10 min, followed by 30 cycles of 95 °C for 15 s and 60 °C for 1 min. All of the measurements were followed by melting curve analysis. A MicroAmp EnduraPlate Optical 384-well plate (Applied Biosystems) and an Optically Clear seal (Applied Biosystems) were used for all measurements. Results were analyzed using ViiA 7 software (Applied Biosystems) by a relative standard curve. For quantification, a four-point standard curve covering a 1,000-fold range of SFGFP cDNA concentrations was run in parallel and used to determine the relative SFGFP cDNA abundance in each sample. It was shown that the SFGFP qPCR primer set had a primer efficiency between 90–103%. All of the cDNA samples were measured in triplicate, and nontemplate controls run in parallel to control for contamination and nonspecific amplification or primer dimers. In addition, qPCR was performed on total RNA samples to confirm that no DNA plasmid was detected under conditions used. Melting curve analysis was performed to confirm that only a single product was amplified.

Total RNA extraction for RNA-seq. For all extractions of total RNA for RNA-seq experiments, *E. coli* strain K12 MG1655 was used. Antisense or

no-antisense control plasmids were transformed, and subsequent colonies were grown overnight as described for *in vivo* bulk fluorescence measurements except that liquid and solid media contained only 100 mg/ml carbenicillin. For each biological replica, 20 µl of a single overnight culture was added to three wells containing 980 µl (1:50 dilution) of supplemented M9 minimal medium containing 100 mg/ml carbenicillin and grown for 4 h at the same conditions as the overnight cultures. For each plasmid combination, 2–3 ml of cells were removed from three wells (grown from one colony) and two volumes of RNAprotect bacterial reagent (Qiagen) were added and incubated for 5 min at room temperature. Total RNA was then purified using an RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol and eluted in 50 µl of RNase free ddH_2O and stored at −80 °C

DNase treatment of total RNA for RNA-seq. Purified total RNA samples were quantified by Qubit Fluorometer, and 6 µg of total RNA was digested by Turbo DNase according to the manufacturer's protocol and purified as described for DNase treatment of Total RNA for qPCR. The quality of the DNase treated total RNA samples were assessed using a Fragment Analyzer (Advanced Analytical).

rRNA depletion of total RNA for RNA seq. DNase treated total RNA samples were quantified by Qubit Fluorometer, and 3 µg of RNA from each sample was treated with the Ribo-Zero rRNA removal kit (Gram-negative) (Epicentre) to remove rRNA according to the manufacturer's protocol and eluted in 10 µl RNase free ddH_2O. For each sample, the rRNA removal was assessed using a Fragment Analyzer .

RNA-seq library preparation, sequencing and analysis. rRNA-depleted total RNA samples were quantified by Qubit Fluorometer, and 50 ng of RNA was used to prepare RNA-seq libraries using the ScriptSeq v2 RNA-seq Library Preparation Kit (Epicentre) according to the manufacturer's protocol. The quality of the RNA-seq libraries was accessed using a Fragment Analyzer. Samples were sequenced on a MiSeq (Illumina) following the manufacturer's standard cluster generation and sequencing protocols for 50-bp paired-end reads of sequencing. Data are available upon request.

RNA-seq data sets were analyzed using the Tophat/Cufflinks pipeline[45] using Bowtie version 1.1.0, Tophat version 2.0.12 and Cufflinks version 2.2.1. To align against the annotated *E. coli* K-12 MG1655 genome, the ensemble FASTA genomic sequence (.fa) and general feature format (.gff3 file) for the GCA_000005845.2 genome assembly were used. The gff3 annotation file was further manually curated to remove duplicate gene ID entries and then converted to .gtf format using the gffread utility provided in the Cufflinks package. Each set of paired-end sequencing reads for each replicate experiment was aligned to the *E. coli* K-12 MG1655 genome using `tophat` options `--no-novel-juncs` and `--library-type fr-secondstrand`. Differential gene expression was analyzed using `cuffdiff` with the `-u` option that specified the same input .gtf file as used in the `tophat` mapping. Scatter plots were made using CummeRbund version 2.6.1 (ref. 46).

Calculation of free energies for STAR design principles. All ΔG terms were calculated using the command-line version of RNAStructure v5.5 (ref. 47). The `Fold` utility was used to calculate ΔG_{STAR} and ΔG_{Target}, and the `DuplexFold` utility was used to calculate ΔG_{SC}, both using the default options.

Characterization of STARs in TX-TL. Cell extract and reaction buffer were prepared according to ref. 26. Gene expression was optimized via the addition of 5 mM Mg-glutamate and 80 mM K-glutamate. TX-TL buffer and extract tubes were thawed on ice for approximately 20 min. Separate reaction tubes were prepared with combinations of DNA representing a given test condition. Appropriate volumes of DNA, buffer and extract were calculated using a custom spreadsheet developed in ref. 26. A final concentration of 1 nM sense target plasmid DNA and 10 nM STAR antisense plasmid DNA or 10 nM no-antisense control plasmid DNA was run in triplicate. Buffer and extract were mixed together, incubated at 37 °C for 20 min and then added to each tube of DNA according to the previously published protocol[26]. 10 µl of each TX-TL reaction mixture was transferred to a 384-well plate (Nunc), covered with a plate seal (Nunc) and placed on a SynergyH1 plate reader. The temperature was controlled at 37 °C, and SFGFP fluorescence was measured (485 nm excitation, 520 emission) every 5 min.

45. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
46. Goff, L., Trapnell, C. & Kelley, D. cummeRbund: analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. (R package version 2.6.1, 2012).
47. Bellaousov, S., Reuter, J.S., Seetin, M.G. & Mathews, D.H. RNAstructure: Web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res.* **41**, W471–W474 (2013).

doi:10.1038/nchembio.1737