Design rules of synthetic non-coding RNAs in bacteria

Young Je Lee and Tae Seok Moon*

Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, USA

* To whom correspondence should be addressed.

Tae Seok Moon

One Brookings Dr., Box 1180

St. Louis, MO 63130, USA

Tel: +1 (314) 935-5026

Fax: +1 (314) 935-7211

Email: tsmoon@wustl.edu

Abstract

One of the long-term goals of synthetic biology is to develop designable genetic parts with predictable behaviors that can be utilized to implement diverse cellular functions. The discovery of non-coding RNAs and their importance in cellular processing have rapidly attracted researchers' attention towards designing functional non-coding RNA molecules. These synthetic non-coding RNAs have simple design principles governed by Watson-Crick base pairing, but exhibit increasingly complex functions. Importantly, due to their specific and modular behaviors, synthetic non-coding RNAs have been widely adopted to modulate transcription and translation of target genes. In this review, we summarize various design rules and strategies employed to engineer synthetic non-coding RNAs. Specifically, we discuss how RNA molecules can be transformed into powerful regulators and utilized to control target gene expression. With the establishment of generalizable non-coding RNA design rules, the research community will shift its focus to RNA regulators from protein regulators.

Keywords

Non-coding RNA, antisense RNA, STAR, toehold switch, CRISPR, aptazyme

Contents

1. Introduction

For many years, RNA molecules were thought to be passive carriers of genetic information from DNA to protein. However, since ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) were discovered in the 1950s, there has been an explosion in the identification of non-coding RNAs (ncRNAs) in prokaryotic and eukaryotic systems. These versatile molecules regulate a broad spectrum of cellular processes [1, 2], modulate protein activities [3], and catalyze biochemical reactions [4]. Natural ncRNAs are now considered as key regulators of gene expression, along with proteins. There has been a great deal of effort to repurpose and re-engineer natural ncRNAs to control gene expression in synthetic biology. Recently, attention has been given to understanding of the relationship between the structure and function of ncRNAs, which will ultimately allow for rational design and construction of synthetic ncRNAs to accomplish a variety of biological functions.

Remarkable progress has been made in developing synthetic ncRNAs based on natural RNA elements, to precisely regulate target gene expression. Natural ncRNAs adopt not only Watson-Crick base paring, but also complex interactions such as non-canonical base pairing and pseudoknot formation [5]. However, synthetic ncRNAs have been designed predominantly by Watson-Crick base paring with their target DNA, mRNA, or other ncRNAs. This simple but elegant mechanism facilitates the rational design and construction of synthetic ncRNAs with a high degree of orthogonality, which is essential in regulating expression of multiple genes. Similar to protein regulators that fold into functional 3-dimensional structures, synthetic ncRNAs can also fold via intramolecular or intermolecular interactions. These interactions are often quantified as Gibbs free energy changes (ΔG, often referred to simply as 'Gibbs free energy' in literature) and the difference (ΔΔG, often as ΔG in literature) between ΔG values, each of which describes the free energy related to the different mechanistic or folding states of RNA molecules. The Gibbs free energy of a given sequence in a given secondary structure is estimated mainly based on the nearest neighbor model, which assumes that the stability of a base pair depends on the adjacent base pair [6, 7]. There are a number of RNA secondary structure prediction programs, including NUPACK and Mfold (with different constraints), which estimate the Gibbs free energy of RNAs and thus predict their possible secondary structures [8, 9], facilitating forward-engineering of ncRNAs. In this review, ΔG values are used to indicate such model-predicted values, not experimentally measured ones.

Various classes of synthetic ncRNAs have been rationally engineered to regulate gene expression. For example, the small transcription activating RNA (STAR) system [10], the attenuator system [11], CRISPR interference (CRISPRi) [12], and CRISPR activation (CRISPRa) [13] were developed to regulate the transcription of target genes. Furthermore, synthetic antisense

RNA (asRNA) [14], the toehold switch [15], translation-regulating riboswitch [16, 17], and aptazyme [18, 19] were developed to regulate gene expression at the post-transcriptional or translational level. In addition to the regulatory function, some of these synthetic ncRNAs demonstrate sensing and enzymatic functions. Such widespread uses of synthetic ncRNAs have prompted efforts to understand the relationship between RNA structures and functions in the cell. These efforts have accelerated with the advent of high-throughput RNA structure characterization technologies [20-22] and software tools [8, 9] that can experimentally and computationally determine RNA structures, respectively. As we better understand the structure-function relationship of ncRNAs, our ability to engineer biological systems is continuously expanding.

In this review, we summarize recent progress in design and construction of synthetic ncRNAs for predictable gene expression control, mainly in bacteria. In particular, we focus on ncRNA design principles that have been determined by investigating the effect of ncRNA secondary structures and various design parameters on regulatory functions of ncRNAs. Specifically, five major examples of synthetic ncRNAs (asRNA, STAR/attenuator, toehold switch, CRISPRi/a, and riboswitch/aptazyme) are discussed regarding their design parameters, including secondary structure features, free energy (ΔG), target binding lengths, and target locations. We also introduce an ncRNA design-build-test cycle to efficiently build and characterize candidate ncRNAs that can regulate the target gene expression. Because this short review focuses primarily on basic, synthetic ncRNA design rules for researchers who are interested in using ncRNAs as their new tools, a lot of excellent reports are omitted in this methodology-focused review. For many studies that are not covered in this paper, we direct the readers to other reviews [23-28].

2. Synthetic non-coding RNA

2.1. Antisense RNA

Natural ncRNAs have been identified in bacteria by computational and experimental approaches [29-31]. In bacteria, regulatory ncRNAs are usually 50 - 250 nucleotides in length and are thus known as small RNAs (sRNAs). Most sRNAs characterized to date modulate target gene expression at the post-transcriptional or translational level by base-pairing with their target mRNA. Because these sRNAs contain a sequence complementary to that of their target mRNA, they are often called antisense RNAs (asRNAs). Upon base-pairing, asRNAs usually prevent translation by altering mRNA accessibility to the translation machinery and/or increasing the rate of transcript degradation [32-34].

Our understanding of native asRNA systems and their regulatory mechanisms enables design of synthetic asRNAs in bacteria [14, 35-38]. Most synthetic asRNAs repress the target gene expression by binding to the 5' untranslated region (UTR) or a coding region of the target mRNA (Fig. 1A). The repressing ability of synthetic asRNAs can also be significantly improved by introducing an Hfq binding site on the 3'-end of target binding region (TBR) of a synthetic asRNA. In this review, TBR is defined as a designed sequence that is complementary to the target mRNA. In nature, base pairing through multiple regions often occurs (e.g., sRNA Spot 42 [39] and sRNA DsrA [40]), while one region (or TBR) is often designed for synthetic asRNAs. The Hfq binding site allows for binding of the Hfq protein, which is a hexameric chaperone protein that is native to diverse bacteria. This RNA-binding protein has been proposed to protect native asRNAs from ribonuclease cleavage and facilitate the annealing of asRNA-mRNA complex by unfolding both molecules [41-44]. While synthetic biologists consider asRNA systems to be modular, native asRNA systems show structural and mechanistic complexity that makes asRNAs difficult to engineer. For example, multiple roles of Hfq have been found in native asRNA systems as

mentioned above, but there is no mechanistic evidence that binding between synthetic asRNA and mRNA is facilitated by Hfq. Additionally, asRNAs that target multiple mRNAs are often found in nature [45, 46], but most synthetic asRNAs have been designed to target only one mRNA sequence. Recently, Lahiry *et al*. showed that sRNAs can be designed and engineered to bind multiple mRNAs, enabling construction of multi-targeting sRNAs [47].

A challenge in developing synthetic asRNA design rules is a wide range of parameters that need to be considered: an Hfq binding site, thermodynamics, the asRNA length, the number of mismatches, and the binding location. Studying one parameter in isolation can be difficult because of the interdependence of each of these parameters. For example, changing the asRNA length or the number of mismatches can affect the thermodynamics of asRNA-mRNA interaction. Furthermore, the mechanisms of base pairing between an asRNA and its target create another challenge in developing asRNA design rules. Two general mechanisms have been proposed based on base pairing processes of native sRNAs: one-step and multi-step mechanism. In the one-step mechanism, the initial interaction between an sRNA and its target mRNA leads to direct and complete duplex formation [48, 49]. In the multi-step mechanism, however, the initial interaction is transient, and the complete duplex formation is achieved only when the initial interaction is stabilized by a protein or the formation of additional base pairs [50]. Whether asRNAs base pair with its target mRNA via one-step, multi-step, or a combination of both, and how these mechanisms affect the stability of asRNA-mRNA interaction remain to be determined.

The first design rule is to introduce an Hfq binding site on the 3'-end of TBR of a synthetic asRNA. Several studies have tested different Hfq binding sites in an attempt to identify the best performing site [14, 38, 51-53]. In most cases, these studies have shown that adding the Hfq binding site improved asRNA's silencing capability. For example, five different TBRs were

combinatorially fused to five Hfq binding sites (Spot42, MicC, MicF, MicF M7.4, or SgrS), and the resultant 25 asRNAs were tested for their repression efficiency. All Hfq binding sites except SgrS led to significantly higher overall average repression (67-77%) than the no-Hfq binding site control (48%) (Fig. 1B) [53]. Furthermore, all tested Hfq binding sites showed low off-target effects except for Spot42. It is worth noting that non-native target mRNAs may lack A-rich sequences (ARN motifs) or UA-binding motifs that are often found in native target mRNAs and that bind to the distal face or the rim of Hfq, respectively [54, 55]. In this case, the effect of Hfq binding sites on repression might be due to increased stability of asRNAs through binding to Hfq [41], and addition of ARN motifs or UA-binding motifs to non-native target mRNAs can be considered for further optimization.

The second design rule is to consider thermodynamics of asRNA-mRNA interaction. This is quantified as ΔG of the asRNA-mRNA complex (ΔΔG to be accurate, while ΔG is often used in literature). Several studies have identified ΔG as an important parameter by showing a strong correlation between the ΔG of the asRNA-mRNA complex and repression [14, 53, 56-58]. For instance, Na *et al*. showed a quantitative relationship between asRNA binding energy and repression efficiency [14]. Hao *et al*. used RyhB-*sodB* mutant pairs to show a correlation between the free energy of duplex formation (ΔE) and fold repression, in which ΔE is defined as the difference between the free energy of the RyhB-*sodB* duplex and the sum of the self-folding free energy of RyhB and the *sodB* control region [58]. Similarly, Hoynes-O'Connor and Moon also found that there is a negative correlation between ΔG Complex Formation and percent repression (Spearman Rank Correlation Coefficient [SRCC] = -0.322) [53]. ΔG Complex Formation is defined as the difference between the ΔG Complex (i.e., ΔG of asRNA-mRNA complex) and ΔG TBR. In the same study, asRNAs that had ΔG Complex Formation values less than -40 kcal/mol

185 demonstrated significantly higher repression than those with higher (or less negative) ΔG values
186 (Fig. 1C). Recently, Vazquez-Anderson *et al*. developed the inTherAcc model (*in vivo*-optimized
187 thermodynamic accessibility-adjusted model) by incorporating a number of thermodynamic terms
188 that consider intramolecular folding energy, intermolecular binding, and the target region
189 availability [59]. This model will assist in effective selection of target RNA binding sites in
190 bacteria.

191 The asRNA-mRNA interaction is noncatalytic in nature because an asRNA molecule is
192 likely to be degraded with its target mRNA, instead of being reused [60]. Thus, the transcription
193 rates of mRNAs and asRNAs (i.e., a 'threshold-linear' response [61]), as well as their interaction
194 energetics, should be considered when asRNA-mediated gene expression control is pursued. It is
195 expected (in an mRNA transcription rate-expression level curve) that above a threshold set by the
196 asRNA transcription rate, the target gene expression increases linearly with the difference between
197 the transcription rates of mRNAs and asRNAs, while complete silencing occurs below the
198 threshold [61]. To ensure predictable and successful gene repression, the asRNA system can be
199 designed such that asRNAs would be in stoichiometric excess of the target mRNAs (e.g., use of a
200 higher copy number plasmid or a stronger promoter for asRNA expression) [53].

201 The third design rule is to consider the TBR length of a synthetic asRNA and the maximum
202 length of continuous dsRNA present in the asRNA-mRNA complex. The maximum length of the
203 dsRNA region is defined as the longest uninterrupted interstrand length of dsRNA in the asRNA-
204 mRNA complex. It was found that an increase in the TBR length or an increase in the maximum
205 length of dsRNA resulted in a higher repression (SRCC = 0.284 for TBR length; SRCC = 0.360
206 for dsRNA length) [53]. The same study showed that asRNAs with at least 15 nucleotides of
207 dsRNA had significantly higher repression efficiency than those with lengths shorter than 15

nucleotides of dsRNA (Fig. 1D). In other studies, a target-binding sequence of 20-30 nucleotides is recommended for reliable repression [14, 38]. It is also important to understand that the length of TBR, the maximum length of continuous dsRNA present in the asRNA-mRNA complex, and ΔG of the asRNA-mRNA complex are interdependent parameters. For example, varying the length of TBR can alter both the maximum length and ΔG of the asRNA-mRNA complex. Therefore, these parameters should be simultaneously considered when designing an asRNA.

The fourth design rule is to minimize the number of mismatched nucleotides in the TBR sequence of a synthetic asRNA (Fig. 1E). Many natural sRNAs bind imperfectly to their target, and it is not clear whether this imperfect pairing facilitates gene repression (via mRNA degradation or inhibition of translation) in bacteria [62]. However, hindered gene silencing was observed when artificial sRNAs contained mismatched (forming internal loops) or unmatched (forming bulges) regions in the middle of the target-recognition sequences [63]. Similarly, it has been shown that an increase in the number of mismatched nucleotides decreases percent repression by synthetic asRNAs in bacteria [53]. In other words, there was a negative correlation between the percent mismatch (defined in Fig. 1E) and target gene repression by synthetic asRNAs (SRCC = -0.345). It was also found that asRNAs with less than 15% mismatch had significantly higher repression than asRNAs with more than 15% mismatch [53]. It is worth noting that additional base-pairing regions could increase repression. As Beisel *et al*. demonstrated, Hfq-binding sRNAs that base pair through multiple unstructured regions can be designed to enable high repression [39], while these sRNAs would have high percent mismatch.

The last design rule is to minimize off-target effects. Studies have demonstrated that synthetic asRNAs are most effective when they bind to the translation initiation region [14, 64, 65]. These findings are not surprising since many natural sRNAs also regulate target gene

expression by binding to the same region [66, 67]. However, the translation initiation region that includes the Shine-Dalgarno sequence and the start codon is conserved in *E. coli* [68, 69]. Therefore, asRNAs that target the translation initiation region may result in off-target repression, as demonstrated using three mRNA-asRNA pairs [14]. To extensively explore off-target repression, Hoynes-O'Connor and Moon designed 96 different TBRs: a third of the asRNAs targeted the *rfp* mRNA, a third targeted the *gfp* mRNA, and the final third targeted the *cfp* mRNA [53]. From this experiment, a positive correlation between sequence identity and off-target repression was observed ($R^2 = 0.496$), indicating that it is important to select asRNA target regions whose sequences are dissimilar to non-target mRNAs (Fig. 1F) [53].

2.2. Small transcription activating RNA and attenuator

ncRNAs that can activate the translation of target genes are found in nature [70, 71]. However, sRNAs that can directly activate transcription have not been found in nature. Recently, an sRNA-mediated, transcription activation system was developed and referred to as STAR (small transcription activating RNA) [10]. The STAR system is composed of two different RNA elements, an intrinsic transcription terminator (e.g., the pT181 terminator [72, 73] and the terminator from the pAD1 plasmid attenuation system [74]) and an sRNA transcriptional activator (*trans*-acting anti-terminator). The intrinsic transcription terminator is placed upstream of a coding region of the target gene (between a promoter and ribosome binding site (RBS)), and represses the downstream transcription of the target gene by default. The sRNA transcriptional activator, or STAR regulator, contains an anti-terminator sequence, which prevents the terminator hairpin formation via RNA-RNA interaction, and thus allows for the downstream transcription of the target gene (Fig. 2A). Using this strategy, Chappell *et al*. demonstrated 94-fold activation of target gene expression [10]. The same group found that a tight control of gene expression in the default

state is crucial in achieving a high dynamic range of the target gene expression. Two design rules were described to generate high performing STAR systems [75]. First, the formation of secondary structures within the linear region of the intrinsic transcription terminator should be avoided. It was shown that the formation of secondary structures negatively affected transcription termination efficiency. Second, the presence of base stacking interactions within the target RNA's linear region decreased transcription termination efficiency, and thus should be avoided. Following these design rules, Chappell *et al*. achieved up to ~9000-fold activation of target gene expression [75].

An intrinsic transcription terminator is a central feature of the STAR system, and thus it is important to determine how the terminator sequence contributes to its strength. *E. coli* intrinsic transcription terminators consist of a short RNA hairpin and a U-rich sequence (U-tract). The RNA hairpin and U-tract cause a rapid dissociation of the ternary elongation complex (the RNA polymerase, DNA template, and RNA), resulting in transcription termination [76, 77]. Chen *et al*. characterized a large library of natural and synthetic *E. coli* intrinsic terminators to understand the sequence-function relationship of the terminators [78]. Each terminator sequence was divided into an A-rich sequence (A-tract), stem, loop, and U-tract (Fig. 2B). In general, insertion of a perfect A-tract or U-tract improved the terminator strength. Furthermore, the best loop was GAAA, and the terminators were strongest when they contained the longest 8-bp stem. Similarly, Cambray *et al*. characterized a collection of natural and synthetic terminators that had a wide range of termination efficiencies [79]. Both reports provide researchers with well-characterized terminator libraries, facilitating engineering of synthetic genetic systems.

A model was developed to describe the gene activation by the STAR system [10]. The model predicts that the natural log of gene expression levels is linearly correlated to the difference in free energy between the initial state of the STAR system and the seed complex ($\Delta G_{Prediction}$ in

Fig. 2C; $R^2 = 0.39 \sim 0.67$ for four different STAR systems tested). The free energy of the initial state (ΔG_{IS}) is the sum of the free energy of the intrinsic transcription terminator (ΔG_{Target}) and the STAR regulator (ΔG_{STAR}). The free energy of the seed complex (ΔG_{SC}) describes the binding energy of the STAR-terminator complex. In the future, this model can be utilized to design new STAR systems.

ncRNAs that repress transcription have been engineered to construct RNA-based circuits [11]. This attenuator system, which was engineered based on the pT181 transcriptional attenuator [72, 73], is composed of two different RNA elements: an attenuator and an asRNA. The attenuator resides in the 5' UTR of mRNA and regulates the transcription of the downstream gene via RNA structural changes. In the absence of the asRNA, which contains a sequence that is complementary to the attenuator, the formation of the terminator hairpin is prevented (Fig. 2D). However, base pairing of this attenuator with the asRNA promotes the formation of the terminator hairpin, which halts the transcription. While engineered attenuator systems using natural attenuators have been successfully demonstrated to control gene expression, only a limited number of orthogonal attenuators exist in nature. This has presented a challenge to utilizing attenuators as genetic tools to construct complex RNA-based networks. One strategy to overcome this challenge is to utilize translational regulators to create orthogonal transcriptional attenuators [80, 81]. For example, Takahashi and Lucks replaced the interacting sequence of the pT181 attenuator with sequences from loop-loop translational regulators (TransSysR and TransSysC) to create chimeric attenuators with high orthogonality (Fig. 2E). Subsequently, the number of chimeric attenuators was greatly expanded via mutagenesis [81]. Furthermore, using in-cell SHAPE-Seq experiments [20] and molecular dynamics simulations [82], Takahashi *et al*. showed that interior loops in the attenuator hairpin confer structural flexibility, establishing interior loops as important structural elements

[83]. These studies have expanded capabilities to use transcriptional attenuators as orthogonal regulatory tools.

2.3. Toehold switch

Nature has a wide number of ncRNAs that act directly on mRNA to regulate translation [84, 85]. Inspired by these natural systems, researchers have developed a number of engineered regulatory ncRNAs, including riboregulators that respond to their cognate *trans*-activating RNAs (taRNAs) [86-89]. For example, Rodrigo *et al*. developed a computational algorithm for the rational design of novel riboregulators by considering free energy of complex formation and activation energy of complex formation [89]. In conventional riboregulators, repression is achieved by intramolecular base-pairing to the RBS region, while activation occurs by displacing the repressing sequence through intermolecular base-pairing and thus exposing the RBS region. This mechanism imposes sequence constraints on *cis*-repressed RNA because intramolecular base-pairing to the RBS region should be present for the repressed state. To address this limitation, Green *et al.* developed *de novo* designed and programmable riboregulators, the so-called toehold switches, to regulate translation through toehold-mediated RNA strand displacement [15] (Fig. 3A). The toehold switch consists of a switch RNA and a trigger RNA (trRNA). The switch RNA forms a *cis*-repressing RNA hairpin, which represses translation by sequestering the region around the RBS and start codon. The binding of trRNA to a toehold sequence at the 5'-end of the hairpin initiates a branch migration process with the hairpin to expose the RBS and start codon, thus allowing ribosomes to bind and initiate the translation of the target mRNA [90, 91].

Unlike conventional riboregulators, the start codon is placed and left unpaired within the stem of the switch RNA, introducing a 3-nt bulge near the midpoint of the 18 nucleotide stem (Fig. 3B). This sequesters the start codon within RNA duplexes. Furthermore, the RBS sequence is

inserted and left unpaired within the loop of the hairpin. This structural configuration of the switch RNA relieves sequence constraints. In addition, a common 21 nucleotide linker sequence, coding for low molecular-weight amino acids, is inserted between the switch RNA and the coding region to act as a buffer region. Based on these design rules, Green *et al.* constructed toehold switches that demonstrated a high level of orthogonality and high fold activation of target gene expression (up to 400-fold) in *E. coli* [15]. Furthermore, they found that the free energy of the sequence from the RBS region to the end of the linker ($\Delta G_{RBS\text{-}linker}$) is positively correlated with the fold activation ($R^2 = 0.40$ for the first-generation library; $R^2 = 0.79$ for the forward-engineered library). $\Delta G_{RBS\text{-}linker}$ describes the energy required by the ribosome to unwind the RBS and linker sequences as it begins translation.

The programmability of the toehold switch, originally shown in *E. coli*, was also demonstrated *in vitro*. Pardee *et al*. developed a low-cost virus detection kit in which a paper-based, cell-free expression platform was combined with the toehold switch technology, successfully demonstrating strain-specific Ebola virus detection as well as extremely sensitive Zika virus sensing [92, 93]. For example, extracted Zika RNA from a sample was isothermally amplified by nucleic acid sequence-based amplification (NASBA). The amplified Zika RNA functions as trRNA, and the binding of trRNA to the 5'-end of the switch RNA initiates the translation of the downstream reporter gene. The level of the reporter protein can be measured and can indicate the presence or absence of Zika virus in a sample at clinically relevant concentrations of Zika virus RNA. Importantly, when combined with a CRISPR-Cas9 module, the toehold switch RNA sensors were able to discriminate between Zika genotypes with single-base resolution. Recently, Green *et al*. constructed ribocomputing systems in *E. coli*, using multiple toehold switches to enable complex intracellular computations such as four-input AND and six-input OR

346 logic gates [94]. These logic gates were built in a single circuit layer by co-localizing multiple
347 toehold switches within a single transcript, which can enhance signal propagation to the output
348 and decrease metabolic costs.

349 **2.4. CRISPR interference and activation**

350 ncRNAs that interact with DNA have been found in nature. In many bacteria and archaea,
351 the CRISPR-Cas (clustered regularly interspaced short palindromic repeats-CRISPR-associated
352 proteins) system provides acquired immunity against bacteriophages and plasmids by using
353 ncRNAs and Cas proteins to target and cleave foreign DNA, respectively, in a sequence-specific
354 manner [95]. Among different types of CRISPR-Cas systems that have been identified, the type II
355 CRISPR-Cas system from *Streptococcus pyogenes* was originally repurposed and has been widely
356 used for gene regulation [12]. The repurposed CRISPR-Cas system is called CRISPR interference
357 (CRISPRi) system, which represses target gene expression in a programmable manner. Other
358 CRISPRi systems, including the CRISPR-Cas9 system from *Streptococcus thermophilus* [96] and
359 the type I CRISPR-Cas system in *E. coli* [97], have also been used for programmable gene
360 repression.

361 To develop the CRISPRi system, a catalytically dead Cas9 mutant (dCas9) was used, which
362 has mutations in the RuvC1 and HNH nuclease domains of Cas9 [12, 98]. The dCas9 protein is
363 defective in DNA cleavage but still functions as a DNA-binding protein. Additionally, CRISPR
364 RNA (crRNA) and *trans*-acting crRNA (tracrRNA) were combined into a chimeric single guide
365 RNA (sgRNA), which simplifies the experimental procedure by bypassing the crRNA maturation
366 step [98]. sgRNA contains ~20 nucleotides of a DNA binding region (complementary to a DNA
367 target sequence) and 42 nucleotides of a dCas9 binding handle. When sgRNA-guided dCas9 binds
368 to a specific DNA target, the sgRNA-dCas9 riboprotein complex can interfere with RNA

polymerase binding or transcriptional elongation, repressing transcription of the target gene in eukaryotes and prokaryotes (Fig. 4A) [99-101].

Repression is affected by multiple factors in the CRISPRi system. dCas9 binding specificity is determined by both sgRNA-DNA base pairing and a short Protospacer Adjacent Motif (PAM) sequence present in proximity to the target DNA region (NGG for dCas9 from *S. pyogenes*, where N is any nucleotide and G is the nucleotide containing guanine). It was found that the target location can affect fold repression [12, 13, 102]. For example, while sgRNAs binding to the nontemplate DNA strand of the coding region resulted in effective silencing (10- to 300-fold repression), sgRNAs binding to the template DNA strand of the coding region showed little effect (Fig. 4B) [12]. Effective gene silencing was observed when sgRNAs targeted the promoter region. However, unlike targeting of the coding region, targeting of both the template and nontemplate DNA strands of the promoter region yielded silencing effects.

dCas9 can be converted into a transcriptional activator by fusing an effector, generating the CRISPR activation (CRISPRa) system. Several effectors, including transcription activator VP16 or p65 activation domain, have been fused to dCas9 to activate the target gene expression in eukaryotes [99]. However, only one effector (the ω subunit of RNA polymerase) fused to dCas9 has been shown to activate gene expression in bacteria (Fig. 4C). The RNA polymerase ω subunit stabilizes the binding of RNA polymerase to a promoter [103]. Bikard *et al*. demonstrated up to 23-fold activation using this fusion protein in a host with a deletion of *rpoZ*, which encodes for the ω subunit of RNA polymerase [13]. The fold-activation was dependent on the target location of the sgRNA-dCas9-ω riboprotein complex, and its optimal distance from the promoter should be determined to induce maximum gene expression.

One exciting use of the CRISPRi system is to dynamically and tunably regulate gene

expression (e.g., repression and derepression, as opposed to gene knockout which means irreversible 100% repression) within a single cell. For example, a gene target that is repressed by the CRISPRi system can be derepressed by expressing a synthetic asRNA, which sequesters and prevents an sgRNA from binding to the target DNA (Fig. 4D). Lee *et al.* found that the hybridization ΔG of the sgRNA-asRNA complex negatively correlated with the derepression efficiency (i.e., the more negative the ΔG was, the higher derepression efficiency was observed; $R^2 = 0.62$) [104]. In addition, instead of targeting the DNA binding site of sgRNA, asRNA was designed to target an artificial linker, which was introduced between the dCas9 binding site and the transcription terminator of sgRNA [104]. This strategy allows for flexible sequence selection of asRNA because its binding sequence can be designed based on the inserted linker region of sgRNA, as opposed to the DNA binding site of sgRNA. This design flexibility enabled asRNAs with low ΔG (e.g., -115 kcal/mol) to be created, resulting in a high derepression efficiency (up to 95%).

2.5. Aptamer, riboswitch, and aptazyme

Aptamers are molecules that specifically bind to their target ligand with high affinity. Aptamers can be generated by *in vitro* selection termed systematic evolution of ligands by exponential enrichment (SELEX) [105, 106], which uses a randomly-generated oligonucleotide library to find oligonucleotide molecules that have a high affinity for a specific target ligand. Due to their small size (20-100 nt; 6-30 kDa), high affinity and specificity, and wide range of potential targets, aptamers have been used for many applications, such as developing diagnostic kits [107], drug delivery systems [108], biosensors for hazard detection [109], metabolite sensors [110-113], and analytical reagents [114]. Furthermore, natural RNA aptamers have been found in prokaryotic riboswitches that control gene expression at the transcriptional or translational levels [115].

Translation-regulating riboswitches are often found within the 5' UTR of bacterial mRNA and contain an aptamer domain and an expression platform that includes RBS. The binding of the target ligand to the aptamer domain induces a conformational change in the expression platform, thereby regulating gene expression. For example, the thiamine pyrophosphate (TPP)-dependent riboswitch represses the expression of thiamine biosynthetic proteins when TPP is present in bacteria [16]. Upon TPP binding, the riboswitch adopts a conformation that occludes RBS and inhibits translation. In addition, a riboswitch that has two different modes of action also exists [116]. Caron *et al*. showed that the *lysC* riboswitch, upon lysine binding, folds into a conformation that not only inhibits translation initiation but also enables RNA degradation by exposing RNase E cleavage sites. By modifying either the aptamer or expression platform of natural riboswitches in bacteria, researchers have engineered riboswitches that exhibit different regulatory functions. For example, the TPP-dependent riboswitch was engineered to activate, rather than repress, gene expression by randomizing the sequence of the expression platform [117].

To construct functional riboswitches, most approaches rely on screening of combinatorial libraries. Recently, Espah Borujeni *et al*. developed a statistical thermodynamic model that can be used to design synthetic translation-regulating riboswitches [118]. This biophysical model predicts the sequence-structure-function relationship by considering the effects of diverse factors on riboswitch activation: the riboswitch's mRNA sequence, aptamer structure, ligand affinity, switching free energy, ligand and mRNA concentrations, and macromolecular crowding. Their computational approach was validated by experimentally characterizing synthetic riboswitches and comparing predicted activation ratios with measured activation ratios. The model-predicted actual activation ratios were in good agreement with the measured activation ratios for 59 synthetic riboswitches (Pearson $R^2 = 0.61$, $P = 2.6\times10^{-13}$). Importantly, they created an automated

optimization algorithm that enabled generation of functional riboswitches with high activation ratios (up to 383-fold), and its user-friendly online version, called the Riboswitch Calculator, is available at http://salislab.net/software [118].

Transcription-regulating riboswitches are also found in nature [115]. These riboswitches usually contain an intrinsic transcription terminator, and upon ligand binding, this terminator structure is either formed (transcription attenuation) or disrupted (transcription activation). Recently, Wachsmuth *et al*. successfully designed and constructed synthetic riboswitches that can control gene expression at the transcriptional level [119]. They developed an *in silico* pipeline for riboswitch design and demonstrated theophylline-dependent regulation of gene expression in *E. coli*. In the follow-up study, Wachsmuth *et al*. identified two important design parameters for the functionality of transcription-regulating riboswitches [120]. First, the terminator hairpin stability should be within a relatively narrow range (e.g., for the theophylline-binding aptamer tested, it should be between -29.0 and -18.1 kcal/mol). If the terminator is too stable, the riboswitch will not be turned on even in the presence of ligands. If the hairpin stability is too low, the riboswitch will be in a permanent 'on' state. Second, possible folding traps should be considered. If the energy barrier for the terminator hairpin formation is too high (i.e., there are folding traps), the riboswitch will not be functional. Further studies on design parameters using other aptamers would facilitate the construction of a variety of synthetic riboswitches that control transcription in response to diverse ligands.

Aptamers can be further engineered by fusing them to a catalytic ribozyme sequence, forming an aptazyme. In aptazymes, the self-cleavage activity of ribozyme is regulated by the ligand binding to the aptamer domain. This inducible cleavage can be used to control gene expression through various strategies, including release of an anti-RBS sequence, which allows

for ribosome access and translation initiation (Fig. 5A), or transcript destabilization and inactivation if the cleavage occurs within the 3' UTR. The hammerhead ribozyme has been widely used with ligand binding aptamers [18, 121-123]. For example, Soukup *et al.* developed an aptazyme by integrating the theophylline aptamer with the hammerhead ribozyme [124]. This theophylline-dependent aptazyme is composed of three domains: a theophylline binding aptamer, a communication module, and a hammerhead ribozyme. The communication module is a unique modular bridge sequence that can be used to rapidly engineer a new aptazyme by replacing the aptamer or ribozyme with different aptamers or ribozymes (Fig. 5B). The authors demonstrated more than 3000-fold activation in cleavage activity *in vitro* [124]. Win and Smolke fused the theophylline aptamer to the hammerhead ribozyme using a communication module sequence to control gene expression *in vivo* [18]. This aptazyme was inserted on the 3' UTR of a target mRNA, as opposed to the 5' UTR, to avoid nonspecific structural effects on translation initiation. This report demonstrated that ncRNAs can act as a sensor with catalytic activity to regulate gene expression in yeast through control of mRNA stability. Furthermore, a recent report showed that differential gene expression can be achieved by using aptazymes that sequester or expose RBS in response to theophylline or TPP in *E. coli* [125]. A better understanding of nucleic acid folding, ribozyme mechanisms, and the binding between nucleic acid and ligand will enable rational design and construction of new aptazymes.

3. Non-coding RNA design cycle

The power of engineering synthetic ncRNAs lies in our ability to design, build, and characterize RNA elements. This ability continues to increase as we better understand the mechanism of natural ncRNAs and develop more advanced engineering strategies. The design-

build-test cycle of synthetic ncRNAs are described as follows (Fig. 6). First, the selection of an organism of interest (Step 1) is followed by the identification of a target gene of interest (Step 2). An organism related to the purpose of the research is selected. For example, *Synechocystis* sp. PCC 6803 (hereafter, *Synechocystis*) can be selected as a microbial cell factory capable of using CO_2 and light [126]. Target genes for repression can be genes in the competing pathways that divert intermediate compounds from heterologous pathways, which have been introduced into *Synechocystis* to produce value-added chemicals from CO_2 and light. A synthetic ncRNA is then selected based on its function (e.g., transcriptional regulation vs. post-transcriptional regulation; activation vs. repression; Step 3). Background research is helpful for Step 3. For example, it has been shown that the *Synechocystis* Hfq is unable to recover the function of the *E. coli* Hfq in an Hfq knockout strain of *E. coli* [127]. Thus, it is unlikely that the native *Synechocystis* Hfq protein will be able to bind to *E. coli* Hfq binding sites. Researchers may express the *E. coli* Hfq to use *E. coli* Hfq binding sites, or may search for RNA sequences that associate with the *Synechocystis* Hfq. Alternatively, the CRISPRi system developed for this strain can be used [128]. Such literature search is accompanied by identifying whether the target gene is under the control of a single promoter with a cluster of other non-target genes (as an operon). If synthetic ncRNAs regulate gene expression at the transcriptional level (e.g., CRISPRi), utilization of these ncRNAs will affect the expression of non-target genes in the same operon.

The iterative design-build-test cycle consists of Steps 4-6. The synthetic ncRNA is designed following the design rules and strategies discussed in the previous section. The structure of ncRNAs is typically predicted using software tools [8, 9] (Step 4), while it can be experimentally verified later using high-throughput RNA structure characterization technologies [20-22]. A special caution should be taken in predicting the structure of ncRNAs because it can determine

ncRNAs' ability to regulate expression of target genes. Next, the synthetic ncRNA is expressed mostly using a plasmid in bacteria, or possibly by genome integration of the ncRNA system (Step 5). The plasmid typically consists of a promoter to transcribe the ncRNA, an origin of replication, and an antibiotic resistance gene. Usually, preliminary tests are performed in Step 5 to see whether the selected ncRNA works in the organism of interest (prior to extensive characterization in Step 6). Importantly, the transcription start site should be identified to avoid transcribing unwanted RNA sequence on the 5'-end of ncRNA, which alters the sequence of a seed region. The seed region is crucial for RNA-DNA and RNA-RNA interactions [98, 129, 130]. Lastly, the ncRNA functions (e.g., dynamic range of gene expression and orthogonality of ncRNAs) are tested in the organism of interest (Step 6). Reporter genes are usually used in the iterative cycle as target genes (Steps 4-6), and once selected, ncRNAs can be tested to regulate the target gene of interest identified in Step 2.

4. Conclusion

The study of ncRNA has been one of the most exciting areas of research in biology. The combination of scientific and engineering approaches has transformed how we utilize ncRNAs to program biological networks and systems. The discoveries in the field of RNA biology continue to demonstrate the versatility of ncRNAs. The recent advances of synthetic biology have influenced and inspired us to engineer diverse types of synthetic ncRNAs. Importantly, the design rules and strategies discussed in this review can be applied to integrate synthetic ncRNAs as functional components in biological networks and systems to process complex information in bacteria. However, to further advance our ability to design and construct a synthetic ncRNA with a predictable function, we still need to better understand the RNA-DNA, RNA-RNA, and RNA-

protein interactions. Furthermore, improved RNA structure prediction tools, along with experimental methods for quick and accurate RNA structure determination *in vivo*, are required to better elucidate the structure-function relationship. We envision that the research community will shift its focus to ncRNAs from protein regulators, and this review contributes to this movement by discussing ncRNA design principles and strategies developed so far.

Funding

This work was supported by the National Science Foundation (CBET-1350498 and MCB-1714352).

Acknowledgements

We thank James Ballard, Drew DeLorenzo, and Kalpana Bodavula for helpful comments on the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

[1] F.J. Mojica, C. Diez-Villasenor, J. Garcia-Martinez, E. Soria, Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements, J. Mol. Evol. 60 (2005) 174-182.
[2] H.L. Jin, Endogenous small RNAs and antibacterial immunity in plants, FEBS Lett. 582 (2008) 2679-2684.
[3] M.M. Yusupov, G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H.D. Cate, H.F. Noller, Crystal structure of the ribosome at 5.5 angstrom resolution, Science 292 (2001) 883-896.
[4] M.J. Fedor, J.R. Williamson, The catalytic diversity of RNAS, Nat. Rev. Mol. Cell Biol. 6 (2005) 399-412.
[5] A. Serganov, E. Nudler, A Decade of Riboswitches, Cell 152(1) (2013) 17-24.
[6] T.B. Xia, J. SantaLucia, M.E. Burkard, R. Kierzek, S.J. Schroeder, X.Q. Jiao, C. Cox, D.H. Turner, Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs, Biochemistry-Us 37(42) (1998) 14719-14735.
[7] D.H. Mathews, J. Sabina, M. Zuker, D.H. Turner, Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, J. Mol. Biol. 288(5) (1999) 911-940.
[8] J.N. Zadeh, C.D. Steenberg, J.S. Bois, B.R. Wolfe, M.B. Pierce, A.R. Khan, R.M. Dirks, N.A. Pierce, NUPACK: Analysis and design of nucleic acid systems, J. Comput. Chem. 32 (2011) 170-173.
[9] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. 31

(2003) 3406-3415.
[10] J. Chappell, M.K. Takahashi, J.B. Lucks, Creating small transcription activating RNAs, Nat. Chem. Biol. 11 (2015) 214-220.
[11] J.B. Lucks, L. Qi, V.K. Mutalik, D. Wang, A.P. Arkin, Versatile RNA-sensing transcriptional regulators for engineering genetic networks, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 8617-8622.
[12] L.S. Qi, M.H. Larson, L.A. Gilbert, J.A. Doudna, J.S. Weissman, A.P. Arkin, W.A. Lim, Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression, Cell 152 (2013) 1173-1183.
[13] D. Bikard, W.Y. Jiang, P. Samai, A. Hochschild, F. Zhang, L.A. Marraffini, Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system, Nucleic Acids Res. 41 (2013) 7429-7437.
[14] D. Na, S.M. Yoo, H. Chung, H. Park, J.H. Park, S.Y. Lee, Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs, Nat. Biotechnol. 31 (2013) 170-174.
[15] A.A. Green, P.A. Silver, J.J. Collins, P. Yin, Toehold Switches: De-Novo-Designed Regulators of Gene Expression, Cell 159 (2014) 925-939.
[16] W. Winkler, A. Nahvi, R.R. Breaker, Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression, Nature 419 (2002) 952-956.
[17] S. Jang, S. Jang, J. Yang, S.W. Seo, G.Y. Jung, RNA-based dynamic genetic controllers: development strategies and applications, Current opinion in biotechnology 53(Supplement C) (2018) 1-11.
[18] M.N. Win, C.D. Smolke, A modular and extensible RNA-based gene-regulatory platform for engineering cellular function, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 14283-14288.
[19] J.M. Carothers, J.A. Goler, D. Juminaga, J.D. Keasling, Model-driven engineering of RNA devices to quantitatively program gene expression, Science (New York, N.Y 334(6063) (2011) 1716-9.
[20] K.E. Watters, T.R. Abbott, J.B. Lucks, Simultaneous characterization of cellular RNA structure and function with in-cell SHAPE-Seq, Nucleic Acids Res. 44 (2016) e12.
[21] S.W. Sowa, J. Vazquez-Anderson, C.A. Clark, R. De La Pena, K. Dunn, E.K. Fung, M.J. Khoury, L.M. Contreras, Exploiting post-transcriptional regulation to probe RNA structures in vivo via fluorescence, Nucleic Acids Res. 43 (2015) e13.
[22] P. Tijerina, S. Mohr, R. Russell, DMS footprinting of structured RNAs and RNA-protein complexes, Nature protocols 2(10) (2007) 2608-23.
[23] T.R. Cech, J.A. Steitz, The Noncoding RNA Revolution-Trashing Old Rules to Forge New Ones, Cell 157 (2014) 77-94.
[24] L.S. Qi, A.P. Arkin, A versatile framework for microbial engineering using synthetic non-coding RNAs, Nat. Rev. Microbiol. 12 (2014) 341-354.
[25] M.N. Win, J.C. Liang, C.D. Smolke, Frameworks for Programming Biological Function through RNA Parts and Devices, Chem. Biol. 16 (2009) 298-310.
[26] S. Gottesman, G. Storz, Bacterial small RNA regulators: versatile roles and rapidly evolving variations, Cold Spring Harbor perspectives in biology 3(12) (2011).
[27] G. Rodrigo, T.E. Landrain, S. Shen, A. Jaramillo, A new frontier in synthetic biology: automated design of small RNA devices in bacteria, Trends in genetics : TIG 29(9) (2013) 529-36.
[28] E. Levine, T. Hwa, Small RNAs establish gene expression thresholds, Current opinion in microbiology 11(6) (2008) 574-9.
[29] R. Hershberg, S. Altuvia, H. Margalit, A survey of small RNA-encoding genes in Escherichia coli, Nucleic Acids Res. 31 (2003) 1813-1820.
[30] L. Argaman, R. Hershberg, J. Vogel, G. Bejerano, E.G. Wagner, H. Margalit, S. Altuvia, Novel small RNA-encoding genes in the intergenic regions of Escherichia coli, Curr. Biol. 11 (2001) 941-950.
[31] K.M. Wassarman, F. Repoila, C. Rosenow, G. Storz, S. Gottesman, Identification of novel small RNAs using comparative genomics and microarrays, Genes Dev. 15 (2001) 1637-1651.

[32] K. Prevost, G. Desnoyers, J.F. Jacques, F. Lavoie, E. Masse, Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage, Genes & development 25(4) (2011) 385-96.
[33] T. Mizuno, M.Y. Chou, M. Inouye, A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA), Proceedings of the National Academy of Sciences of the United States of America 81(7) (1984) 1966-70.
[34] M. Bouvier, C.M. Sharma, F. Mika, K.H. Nierhaus, J. Vogel, Small RNA binding to 5' mRNA coding region inhibits translational initiation, Molecular cell 32(6) (2008) 827-37.
[35] K.V. Solomon, T.M. Sanders, K.L. Prather, A dynamic metabolite valve for the control of central carbon metabolism, Metab. Eng. 14 (2012) 661-671.
[36] S.B. Tummala, N.E. Welker, E.T. Papoutsakis, Design of antisense RNA constructs for downregulation of the acetone formation pathway of Clostridium acetobutylicum, J. Bacteriol. 185 (2003) 1923-1934.
[37] Y.P. Yang, Y.H. Lin, L.Y. Li, R.J. Linhardt, Y.J. Yan, Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products, Metab. Eng. 29 (2015) 217-226.
[38] S. Man, R. Cheng, C. Miao, Q. Gong, Y. Gu, X. Lu, F. Han, W. Yu, Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria, Nucleic Acids Res. 39 (2011) e50.
[39] C.L. Beisel, T.B. Updegrove, B.J. Janson, G. Storz, Multiple factors dictate target selection by Hfq-binding small RNAs, EMBO J 31(8) (2012) 1961-74.
[40] R.A. Lease, M. Belfort, A trans-acting RNA as a control switch in Escherichia coli: DsrA modulates function by forming alternative structures, Proceedings of the National Academy of Sciences of the United States of America 97(18) (2000) 9919-24.
[41] J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, Nat. Rev. Microbiol. 9 (2011) 578-589.
[42] T. Moller, T. Franch, P. Hojrup, D.R. Keene, H.P. Bachinger, R.G. Brennan, P. Valentin-Hansen, Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction, Mol. Cell 9 (2002) 23-30.
[43] S.C. Pulvermacher, L.T. Stauffer, G.V. Stauffer, Role of the Escherichia coli Hfq protein in GcvB regulation of oppA and dppA mRNAs, Microbiology 155 (2009) 115-123.
[44] S. Sagawa, J.E. Shin, R. Hussein, H.N. Lim, Paradoxical suppression of small RNA activity at high Hfq concentrations due to random-order binding, Nucleic Acids Res. 43 (2015) 8502-8515.
[45] E.G. Wagner, P. Romby, Small RNAs in bacteria and archaea: who they are, what they do, and how they do it, Advances in genetics 90 (2015) 133-208.
[46] C.M. Sharma, F. Darfeuille, T.H. Plantinga, J. Vogel, A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites, Genes & development 21(21) (2007) 2804-17.
[47] A. Lahiry, S.D. Stimple, D.W. Wood, R.A. Lease, Retargeting a Dual-Acting sRNA for Multiple mRNA Transcript Regulation, ACS Synth Biol 6(4) (2017) 648-658.
[48] T. Franch, A.P. Gultyaev, K. Gerdes, Programmed cell death by hok/sok of plasmid R1: processing at the hok mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding, J. Mol. Biol. 273 (1997) 38-51.
[49] T. Franch, M. Petersen, E.G. Wagner, J.P. Jacobsen, K. Gerdes, Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure, J. Mol. Biol. 294 (1999) 1115-1125.
[50] P. Sulc, T.E. Ouldridge, F. Romano, J.P. Doye, A.A. Louis, Modelling toehold-mediated RNA strand displacement, Biophys. J. 108 (2015) 1238-1247.
[51] Y. Sakai, K. Abe, S. Nakashima, W. Yoshida, S. Ferri, K. Sode, K. Ikebukuro, Improving the Gene-Regulation Ability of Small RNAs by Scaffold Engineering in Escherichia coli, ACS Synth. Biol. 3 (2014) 152-162.
[52] V. Sharma, A. Yamamura, Y. Yokobayashi, Engineering Artificial Small RNAs for Conditional Gene Silencing in Escherichia coli, ACS Synth. Biol. 1 (2012) 6-13.

[53] A. Hoynes-O'Connor, T.S. Moon, Development of Design Rules for Reliable Antisense RNA Behavior in E. coli, ACS Synth. Biol. 5 (2016) 1441-1454.
[54] D.J. Schu, A. Zhang, S. Gottesman, G. Storz, Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition, EMBO J 34(20) (2015) 2557-73.
[55] T.M. Link, P. Valentin-Hansen, R.G. Brennan, Structure of Escherichia coli Hfq bound to polyriboadenylate RNA, Proceedings of the National Academy of Sciences of the United States of America 106(46) (2009) 19292-7.
[56] S.M. Yoo, D. Na, S.Y. Lee, Design and use of synthetic regulatory small RNAs to control gene expression in Escherichia coli, Nat. Protoc. 8 (2013) 1694-1707.
[57] V.K. Mutalik, L. Qi, J.C. Guimaraes, J.B. Lucks, A.P. Arkin, Rationally designed families of orthogonal RNA regulators of translation, Nat. Chem. Biol. 8 (2012) 447-454.
[58] Y. Hao, Z.J. Zhang, D.W. Erickson, M. Huang, Y. Huang, J. Li, T. Hwa, H. Shi, Quantifying the sequence-function relation in gene silencing by bacterial small RNAs, Proceedings of the National Academy of Sciences of the United States of America 108(30) (2011) 12473-8.
[59] J. Vazquez-Anderson, M.K. Mihailovic, K.C. Baldridge, K.G. Reyes, K. Haning, S.H. Cho, P. Amador, W.B. Powell, L.M. Contreras, Optimization of a novel biophysical model using large scale in vivo antisense hybridization data displays improved prediction capabilities of structurally accessible RNA regions, Nucleic acids research 45(9) (2017) 5523-5538.
[60] E. Masse, F.E. Escorcia, S. Gottesman, Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli, Genes & development 17(19) (2003) 2374-83.
[61] E. Levine, Z. Zhang, T. Kuhlman, T. Hwa, Quantitative characteristics of gene regulation by small RNA, PLoS biology 5(9) (2007) e229.
[62] S. Gottesman, Micros for microbes: non-coding regulatory RNAs in bacteria, Trends Genet. 21 (2005) 399-404.
[63] H. Park, G. Bak, S.C. Kim, Y. Lee, Exploring sRNA-mediated gene silencing mechanisms using artificial small RNAs derived from a natural RNA scaffold in Escherichia coli, Nucleic Acids Res 41(6) (2013) 3787-3804.
[64] J. Coleman, P.J. Green, M. Inouye, The Use of Rnas Complementary to Specific Messenger-Rnas to Regulate the Expression of Individual Bacterial Genes, Cell 37 (1984) 429-436.
[65] H. Park, Y. Yoon, S. Suk, J.Y. Lee, Y. Lee, Effects of different target sites on antisense RNA-mediated regulation of gene expression, BMB Rep. 47 (2014) 619-624.
[66] A.A. Rasmussen, J. Johansen, J.S. Nielsen, M. Overgaard, B. Kallipolitis, P. Valentin-Hansen, A conserved small RNA promotes silencing of the outer membrane protein YbfM, Mol. Microbiol. 72 (2009) 566-577.
[67] A.A. Rasmussen, M. Eriksen, K. Gilany, C. Udesen, T. Franch, C. Petersen, P. Valentin-Hansen, Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control, Mol. Microbiol. 58 (2005) 1421-1429.
[68] K. Yamagishi, T. Oshima, Y. Masuda, T. Ara, S. Kanaya, H. Mori, Conservation of translation initiation sites based on dinucleotide frequency and codon usage in Escherichia coli K-12 (W3110): Non-random distribution of A/T-rich sequences immediately upstream of the translation initiation codon, DNA Res. 9 (2002) 19-24.
[69] V. Vimberg, A. Tats, M. Remm, T. Tenson, Translation initiation region sequence preferences in Escherichia coli, BMC Mol. Biol. 8 (2007) 100.
[70] G. Storz, J. Vogel, K.M. Wassarman, Regulation by Small RNAs in Bacteria: Expanding Frontiers, Mol. Cell 43 (2011) 880-891.
[71] G. Desnoyers, M.P. Bouchard, E. Masse, New insights into small RNA-dependent translational regulation in prokaryotes, Trends Genet. 29 (2013) 92-98.
[72] S. Brantl, E.G.H. Wagner, Antisense RNA-mediated transcriptional attenuation: an in vitro study of

plasmid pT181, Mol. Microbiol. 35 (2000) 1469-1482.
[73] C.C. Kumar, R.P. Novick, Plasmid Pt181 Replication Is Regulated by 2 Countertranscripts, Proc. Natl. Acad. Sci. U.S.A. 82 (1985) 638-642.
[74] S. Brantl, Regulatory mechanisms employed by cis-encoded antisense RNAs, Current opinion in microbiology 10(2) (2007) 102-9.
[75] J. Chappell, A. Westbrook, M. Verosloff, J.B. Lucks, Computational design of small transcription activating RNAs for versatile and dynamic gene regulation, Nature communications 8(1) (2017) 1051.
[76] I. Gusarov, E. Nudler, The mechanism of intrinsic transcription termination, Mol. Cell 3 (1999) 495-504.
[77] L.E. Macdonald, Y. Zhou, W.T. McAllister, Termination and slippage by bacteriophage T7 RNA polymerase, J. Mol. Biol. 232 (1993) 1030-1047.
[78] Y.J. Chen, P. Liu, A.A. Nielsen, J.A. Brophy, K. Clancy, T. Peterson, C.A. Voigt, Characterization of 582 natural and synthetic terminators and quantification of their design constraints, Nat. Methods 10 (2013) 659-664.
[79] G. Cambray, J.C. Guimaraes, V.K. Mutalik, C. Lam, Q.A. Mai, T. Thimmaiah, J.M. Carothers, A.P. Arkin, D. Endy, Measurement and modeling of intrinsic transcription terminators, Nucleic acids research 41(9) (2013) 5139-48.
[80] C.C. Liu, L. Qi, J.B. Lucks, T.H. Segall-Shapiro, D. Wang, V.K. Mutalik, A.P. Arkin, An adaptor from translational to transcriptional control enables predictable assembly of complex regulation, Nat. Methods 9 (2012) 1088-1094.
[81] M.K. Takahashi, J.B. Lucks, A modular strategy for engineering orthogonal chimeric RNA transcription regulators, Nucleic Acids Res. 41 (2013) 7577-7588.
[82] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29(7) (2013) 845-54.
[83] M.K. Takahashi, K.E. Watters, P.M. Gasper, T.R. Abbott, P.D. Carlson, A.A. Chen, J.B. Lucks, Using in-cell SHAPE-Seq and simulations to probe structure-function design principles of RNA transcriptional regulators, Rna 22(6) (2016) 920-33.
[84] A.P. Gultyaev, T. Franch, K. Gerdes, Programmed cell death by hok/sok of plasmid R1: Coupled nucleotide covariations reveal a phylogenetically conserved folding pathway in the hok family of mRNAs, J. Mol. Biol. 273 (1997) 26-37.
[85] J. Praszkier, A.J. Pittard, Pseudoknot-dependent translational coupling in repBA genes of the IncB plasmid pMU720 involves reinitiation, J. Bacteriol. 184 (2002) 5772-5780.
[86] F.J. Isaacs, D.J. Dwyer, C. Ding, D.D. Pervouchine, C.R. Cantor, J.J. Collins, Engineered riboregulators enable post-transcriptional control of gene expression, Nat. Biotechnol. 22 (2004) 841-847.
[87] J.M. Callura, C.R. Cantor, J.J. Collins, Genetic switchboard for synthetic biology applications, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 5850-5855.
[88] J.M. Callura, D.J. Dwyer, F.J. Isaacs, C.R. Cantor, J.J. Collins, Tracking, tuning, and terminating microbial physiology using synthetic riboregulators, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 15898-15903.
[89] G. Rodrigo, T.E. Landrain, A. Jaramillo, De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells, Proceedings of the National Academy of Sciences of the United States of America 109(38) (2012) 15271-6.
[90] P. Sulc, T.E. Ouldridge, F. Romano, J.P.K. Doye, A.A. Louis, Modelling Toehold-Mediated RNA Strand Displacement, Biophys. J. 108 (2015) 1238-1247.
[91] N. Srinivas, T.E. Ouldridge, P. Sulc, J.M. Schaeffer, B. Yurke, A.A. Louis, J.P.K. Doye, E. Winfree, On the biophysics and kinetics of toehold-mediated DNA strand displacement, Nucleic Acids Res. 41 (2013) 10641-10658.
[92] K. Pardee, A.A. Green, T. Ferrante, D.E. Cameron, A. DaleyKeyser, P. Yin, J.J. Collins, Paper-Based

753 Synthetic Gene Networks, Cell 159 (2014) 940-954.
754 [93] K. Pardee, A.A. Green, M.K. Takahashi, D. Braff, G. Lambert, J.W. Lee, T. Ferrante, D. Ma, N. Donghia,
755 M. Fan, N.M. Daringer, I. Bosch, D.M. Dudley, D.H. O'Connor, L. Gehrke, J.J. Collins, Rapid, Low-Cost
756 Detection of Zika Virus Using Programmable Biomolecular Components, Cell 165 (2016) 1255-1266.
757 [94] A.A. Green, J.M. Kim, D. Ma, P.A.S. Ilver, J.J. Collins, P. Yin, Complex cellular logic computation using
758 ribocomputing devices, Nature 548 (2017) 117.
759 [95] P. Horvath, R. Barrangou, CRISPR/Cas, the immune system of bacteria and archaea, Science 327
760 (2010) 167-170.
761 [96] J.M. Rock, F.F. Hopkins, A. Chavez, M. Diallo, M.R. Chase, E.R. Gerrick, J.R. Pritchard, G.M. Church, E.J.
762 Rubin, C.M. Sassetti, D. Schnappinger, S.M. Fortune, Programmable transcriptional repression in
763 mycobacteria using an orthogonal CRISPR interference platform, Nature microbiology 2 (2017) 16274.
764 [97] M.L. Luo, A.S. Mullis, R.T. Leenay, C.L. Beisel, Repurposing endogenous type I CRISPR-Cas systems for
765 programmable gene repression, Nucleic acids research 43(1) (2015) 674-81.
766 [98] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-
767 guided DNA endonuclease in adaptive bacterial immunity, Science 337 (2012) 816-821.
768 [99] L.A. Gilbert, M.H. Larson, L. Morsut, Z.R. Liu, G.A. Brar, S.E. Torres, N. Stern-Ginossar, O. Brandman,
769 E.H. Whitehead, J.A. Doudna, W.A. Lim, J.S. Weissman, L.S. Qi, CRISPR-Mediated Modular RNA-Guided
770 Regulation of Transcription in Eukaryotes, Cell 154 (2013) 442-451.
771 [100] E.D. Jensen, R. Ferreira, T. Jakociunas, D. Arsovska, J. Zhang, L. Ding, J.D. Smith, F. David, J. Nielsen,
772 M.K. Jensen, J.D. Keasling, Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA
773 strategies, Microb. Cell Fact. 16 (2017) 46.
774 [101] C.H. Huang, C.R. Shen, H. Li, L.Y. Sung, M.Y. Wu, Y.C. Hu, CRISPR interference (CRISPRi) for gene
775 regulation and succinate production in cyanobacterium S. elongatus PCC 7942, Microb. Cell Fact. 15 (2016)
776 196.
777 [102] A. Radzisheuskaya, D. Shlyueva, I. Muller, K. Helin, Optimizing sgRNA position markedly improves
778 the efficiency of CRISPR/dCas9-mediated transcriptional repression, Nucleic Acids Res. 44 (2016) e141.
779 [103] S.L. Dove, A. Hochschild, Conversion of the omega subunit of Escherichia coli RNA polymerase into
780 a transcriptional activator or an activation target, Genes Dev. 12 (1998) 745-754.
781 [104] Y.J. Lee, A. Hoynes-O'Connor, M.C. Leong, T.S. Moon, Programmable control of bacterial gene
782 expression with the combined CRISPR and antisense RNA system, Nucleic Acids Res. 44 (2016) 2462-2473.
783 [105] A.D. Ellington, J.W. Szostak, Selection in vitro of single-stranded DNA molecules that fold into specific
784 ligand-binding structures, Nature 355 (1992) 850-852.
785 [106] C. Tuerk, L. Gold, Systematic evolution of ligands by exponential enrichment: RNA ligands to
786 bacteriophage T4 DNA polymerase, Science 249 (1990) 505-510.
787 [107] S. Tombelli, M. Minunni, E. Luzi, M. Mascini, Aptamer-based biosensors for the detection of HIV-1
788 Tat protein, Bioelectrochemistry 67 (2005) 135-141.
789 [108] Y.R. Wu, K. Sefah, H.P. Liu, R.W. Wang, W.H. Tan, DNA aptamer-micelle as an efficient
790 detection/delivery vehicle toward cancer cells, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 5-10.
791 [109] L.S. Rotherham, C. Maserumule, K. Dheda, J. Theron, M. Khati, Selection and Application of ssDNA
792 Aptamers to Detect Active TB from Sputum Samples, PLoS ONE 7 (2012) e46862.
793 [110] S. Jang, G.Y. Jung, Systematic optimization of L-tryptophan riboswitches for efficient monitoring of
794 the metabolite in Escherichia coli, Biotechnology and bioengineering 115(1) (2018) 266-271.
795 [111] J. Yang, S.W. Seo, S. Jang, S.-I. Shin, C.H. Lim, T.-Y. Roh, G.Y. Jung, Synthetic RNA devices to expedite
796 the evolution of metabolite-producing microbes, Nature communications 4 (2013) 1413.
797 [112] S. Jang, S. Jang, Y. Xiu, T.J. Kang, S.-H. Lee, M.A.G. Koffas, G.Y. Jung, Development of Artificial
798 Riboswitches for Monitoring of Naringenin In Vivo, ACS Synthetic Biology 6(11) (2017) 2077-2085.
799 [113] C. Hwang, J.M. Carothers, Label-free selection of RNA aptamers for metabolic engineering, Methods
800 106(Supplement C) (2016) 37-41.

[114] S. Shin, I.H. Kim, W. Kang, J.K. Yang, S.S. Hah, An alternative to Western blot analysis using RNA aptamer-functionalized quantum dots, Bioorg. Med. Chem. Lett. 20 (2010) 3322-3325.
[115] W.C. Winkler, R.R. Breaker, Regulation of bacterial gene expression by riboswitches, Annu. Rev. Microbiol. 59 (2005) 487-517.
[116] M.P. Caron, L. Bastet, A. Lussier, M. Simoneau-Roy, E. Masse, D.A. Lafontaine, Dual-acting riboswitch control of translation initiation and mRNA decay, Proceedings of the National Academy of Sciences of the United States of America 109(50) (2012) E3444-53.
[117] Y. Nomura, Y. Yokobayashi, Reengineering a natural riboswitch by dual genetic selection, Journal of the American Chemical Society 129 (2007) 13814-13815.
[118] A. Espah Borujeni, D.M. Mishler, J. Wang, W. Huso, H.M. Salis, Automated physics-based design of synthetic riboswitches from diverse RNA aptamers, Nucleic acids research 44(1) (2016) 1-13.
[119] M. Wachsmuth, S. Findeiss, N. Weissheimer, P.F. Stadler, M. Morl, De novo design of a synthetic riboswitch that regulates transcription termination, Nucleic acids research 41(4) (2013) 2541-51.
[120] M. Wachsmuth, G. Domin, R. Lorenz, R. Serfling, S. Findeiss, P.F. Stadler, M. Morl, Design criteria for synthetic riboswitches acting on transcription, RNA biology 12(2) (2015) 221-31.
[121] J. Tang, R.R. Breaker, Rational design of allosteric ribozymes, Chem. Biol. 4 (1997) 453-459.
[122] M. Araki, Y. Okuno, Y. Hara, Y. Sugiura, Allosteric regulation of a ribozyme activity through ligand-induced conformational change, Nucleic Acids Res. 26 (1998) 3379-3384.
[123] G.A. Soukup, R.R. Breaker, Design of allosteric hammerhead ribozymes activated by ligand-induced structure stabilization, Structure 7 (1999) 783-791.
[124] G.A. Soukup, G.A.M. Emilsson, R.R. Breaker, Altering molecular recognition of RNA aptamers by allosteric selection, J. Mol. Biol. 298 (2000) 623-632.
[125] M. Felletti, J. Stifel, L.A. Wurmthaler, S. Geiger, J.S. Hartig, Twister ribozymes as highly versatile expression platforms for artificial riboswitches, Nat Commun 7 (2016) 12834.
[126] B.M. Berla, R. Saha, C.M. Immethun, C.D. Maranas, T.S. Moon, H.B. Pakrasi, Synthetic biology of cyanobacteria: unique challenges and opportunities, Front Microbiol 4 (2013) 246.
[127] A. Boggild, M. Overgaard, P. Valentin-Hansen, D.E. Brodersen, Cyanobacteria contain a structural homologue of the Hfq protein with altered RNA-binding properties, The FEBS journal 276(14) (2009) 3904-15.
[128] L. Yao, I. Cengic, J. Anfelt, E.P. Hudson, Multiple Gene Repression in Cyanobacteria Using CRISPRi, ACS Synth Biol 5(3) (2016) 207-12.
[129] B. Wiedenheft, G.C. Lander, K. Zhou, M.M. Jore, S.J.J. Brouns, J. van der Oost, J.A. Doudna, E. Nogales, Structures of the RNA-guided surveillance complex from a bacterial immune system, Nature 477 (2011) 486-489.
[130] B.P. Lewis, C.B. Burge, D.P. Bartel, Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets, Cell 120 (2005) 15-20.
[131] A. Ogawa, M. Maeda, Aptazyme-based riboswitches as label-free and detector-free sensors for cofactors, Bioorg. Med. Chem. Lett. 17 (2007) 3156-3160.

Figure Legends

Figure 1. Synthetic antisense RNA design rules. (A) asRNAs modulate the expression of a gene of interest (GOI) at the post-transcriptional or translational level by base-pairing with the 5' UTR or a coding region of the target mRNA. The interaction between the asRNA and its target mRNA

844 can be facilitated by the Hfq protein that binds to the Hfq binding site of the asRNA. RBS stands
845 for ribosome binding site. **(B)** Introduction of an Hfq binding site on the 3'-end of a target binding
846 region (TBR) can improve synthetic asRNA's gene silencing capability. High-performing Hfq
847 binding sites are listed in the box. Spot42, MicC, MicF, and MicF M7.4 all led to high overall
848 repression compared to their control asRNAs without an Hfq-binding site [53]. **(C)**
849 Thermodynamics of asRNA-mRNA interaction is important when designing an asRNA. asRNAs
850 that had ΔG Complex Formation values less than -40 kcal/mol showed significantly higher
851 repression than those with higher (or less negative) ΔG values [14, 53]. ΔG Complex Formation
852 is defined as the difference between the ΔG Complex (i.e., ΔG of asRNA-mRNA complex) and
853 ΔG TBR. **(D)** The maximum length of continuous dsRNA present in the asRNA-mRNA complex
854 is another parameter that can affect asRNA's gene silencing capability. asRNAs with at least 15
855 nucleotides for dsRNA formation showed significantly higher repression than those with lengths
856 shorter than 15 nucleotides for dsRNA formation [53]. **(E)** The percent of the TBR length that is
857 mismatched (% mismatch) must be less than 15% for higher repression. Percent mismatch is
858 calculated by dividing the number of mismatched nucleotides by the TBR length (and multiplying
859 by 100). **(F)** The sequence similarity between the target and the non-target is an important design
860 consideration. A positive correlation between sequence identity and off-target repression was
861 observed [53].
862
863 **Figure 2. STAR and attenuator design rules. (A)** STAR systems modulate the transcription of
864 a gene of interest (GOI) [10, 75]. In the absence of STAR regulator, the terminator forms a hairpin
865 structure that represses the downstream transcription of the GOI. In the presence of STAR
866 regulator, the 5'-end of terminator stem is sequestered by the STAR regulator, allowing

downstream transcription of the GOI by RNA polymerase (RNAP). **(B)** Changing the various components of a terminator has an impact on the terminator strength [78]. Each component is denoted as an A-tract, stem, loop, and U-tract. The sequence that resulted in the highest terminator strength is highlighted in red for each category. **(C)** A model for the STAR system predicts that the natural log of gene expression is linearly correlated with the difference in free energy between the initial state of the terminator/STAR regulator ($\Delta G_{Target} + \Delta G_{STAR}$) and the seed complex ($\Delta G_{SC}$) [10]. **(D)** A transcriptional attenuator in the 5' UTR of mRNA can fold into a structure that allows transcription of the GOI in the absence of asRNA. A kissing hairpin interaction between the attenuator and asRNA results in the formation of a terminator hairpin, halting the transcription of the GOI [11]. **(E)** Chimeric attenuators with high orthogonality were created by fusing the sequences from loop-loop translational regulators (TransSysR and TransSysC) onto the natural pT181 transcriptional attenuator [81]. The chimeric attenuators are denoted as Fusion 1 and 2. The number of chimeric attenuators can be greatly expanded via mutagenesis. Mutation patterns are denoted by stars.

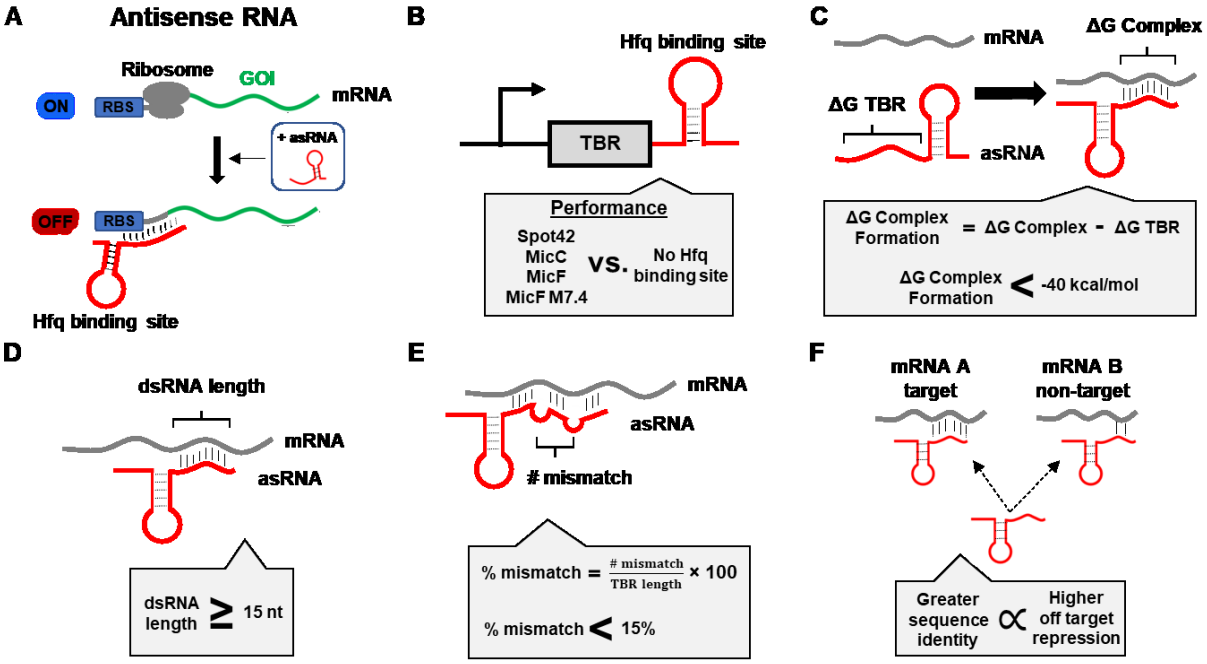
Figure 3. Toehold switch design rules. (A) Switch RNAs repress translation by sequestering the region around the start codon and ribosome binding site (RBS) [15]. The binding of a trigger RNA (trRNA) to the switch RNA through strand displacement exposes the RBS and start codon, thus initiating the translation of the target mRNA. GOI stands for gene of interest. **(B)** The start codon is placed and left unpaired within the stem of switch RNAs, introducing a 3-nt bulge near the midpoint of the 18 nucleotide stem. The RBS sequence is inserted and left unpaired within the loop of the hairpin. A 21 nucleotide linker sequence is inserted between the switch RNA and the coding region to act as a buffer region. The initial binding of domain a of switch RNA to a

complementary domain a' of trRNA leads to the binding of domain b of switch RNA to domain b' of trRNA. The length of these domains and the loop can be systematically varied to increase or decrease the fold-activation of the target gene expression.

Figure 4. single guide RNA (sgRNA) design rules. (A) A CRISPRi system utilizes sgRNA and a catalytically dead Cas9 nuclease (dCas9) to block RNA polymerase (RNAP) binding during transcription initiation and elongation [12]. sgRNA consists of ~20 nucleotides of a base pairing region, 42 nucleotides of a dCas9 binding handle, and a terminator sequence. GOI stands for gene of interest. **(B)** Target location can affect the fold repression. sgRNAs (NT-1,2,3) targeting the nontemplate DNA strand of the coding region resulted in effective silencing [12]. sgRNAs (T-1,2,3) targeting the template DNA strand of the coding region had little effect. Both NT-Promoter and T-Promoter sgRNAs that target the promoter region resulted in effective silencing. RBS stands for ribosome binding site. **(C)** A CRISPRa system utilizes sgRNA and dCas9 fused to the ω subunit of RNAP to activate the target gene expression in *E. coli* with deletion of *rpoZ*, which encodes for the ω subunit of RNAP [13]. The target location of the sgRNA-dCas9-ω riboprotein complex within the promoter region should be optimized to induce maximum gene expression. **(D)** Two independent systems of RNA regulation, CRISPRi and asRNA, were integrated to control target gene expression [104]. The gene target repressed by the CRISPRi system can be derepressed by expressing asRNA, which sequesters sgRNA. The derepression efficiency was improved by increasing the binding affinity between sgRNA and asRNA. Importantly, by introducing an artificial RNA linker between the dCas9 binding site and the transcription terminator of sgRNA, asRNAs can be designed to bind to the RNA linker, instead of the DNA binding site of sgRNA. This strategy allowed for flexible asRNA design, leading to a high derepression efficiency.

Figure 5. Strategies for constructing a synthetic aptazyme. (A) In the absence of the target ligand, an aptazyme can repress translation by sequestering the ribosome binding site (RBS) sequence with the anti-RBS sequence, which prevents the ribosome from binding to the mRNA [131]. The binding of the target ligand to the apatamer domain triggers a conformational change, which results in self-cleavage by an activated aptazyme. The self-cleavage releases the anti-RBS sequence and allows the ribosome to initiate translation. **(B)** The general design strategy for constructing aptazymes is shown. An aptamer and a ribozyme are in red and blue, respectively. A communication module is a unique modular bridge sequence that can be used to rapidly engineer a new aptazyme by simply replacing the aptamer or the ribozyme with different aptamers or ribozymes. The communication module is shown in a dotted black box.

Figure 6. Design and characterization process of non-coding RNAs. An ncRNA design-build-test cycle can include the following steps: choose a species of interest (Step 1); identify a target gene (Step 2); select ncRNA based on its function (Step 3); design the selected ncRNA and computationally predict its structure (Step 4); construct a plasmid to transcribe the ncRNA in the species of interest (Step 5); and characterize the ncRNA functions in the species of interest (Step 6). Steps 4-6 can be repeated.



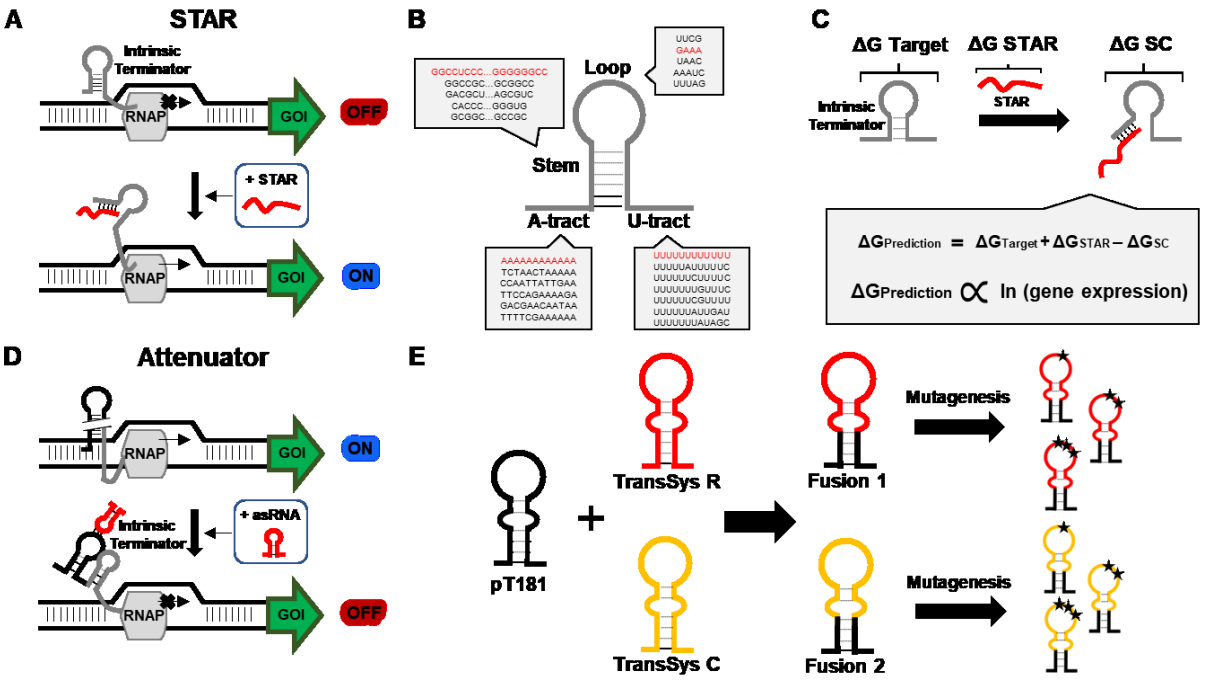
936

937

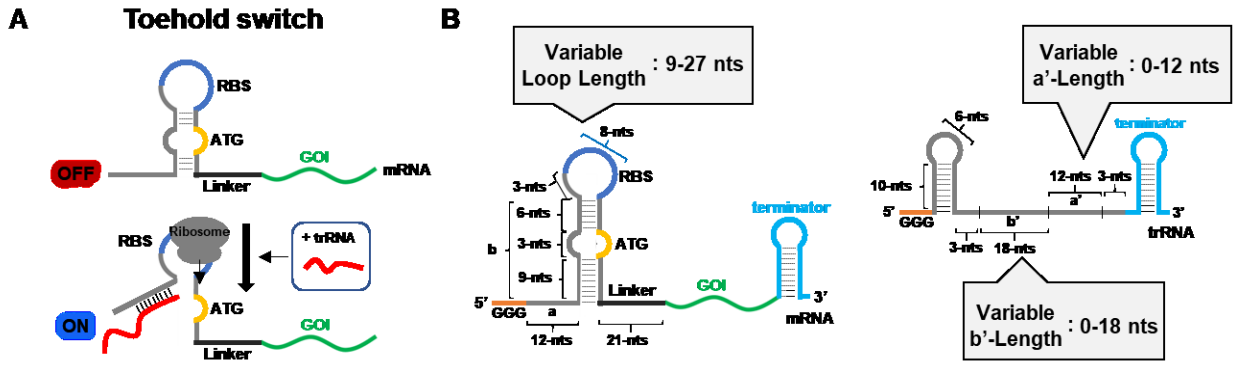
938

939

940



941

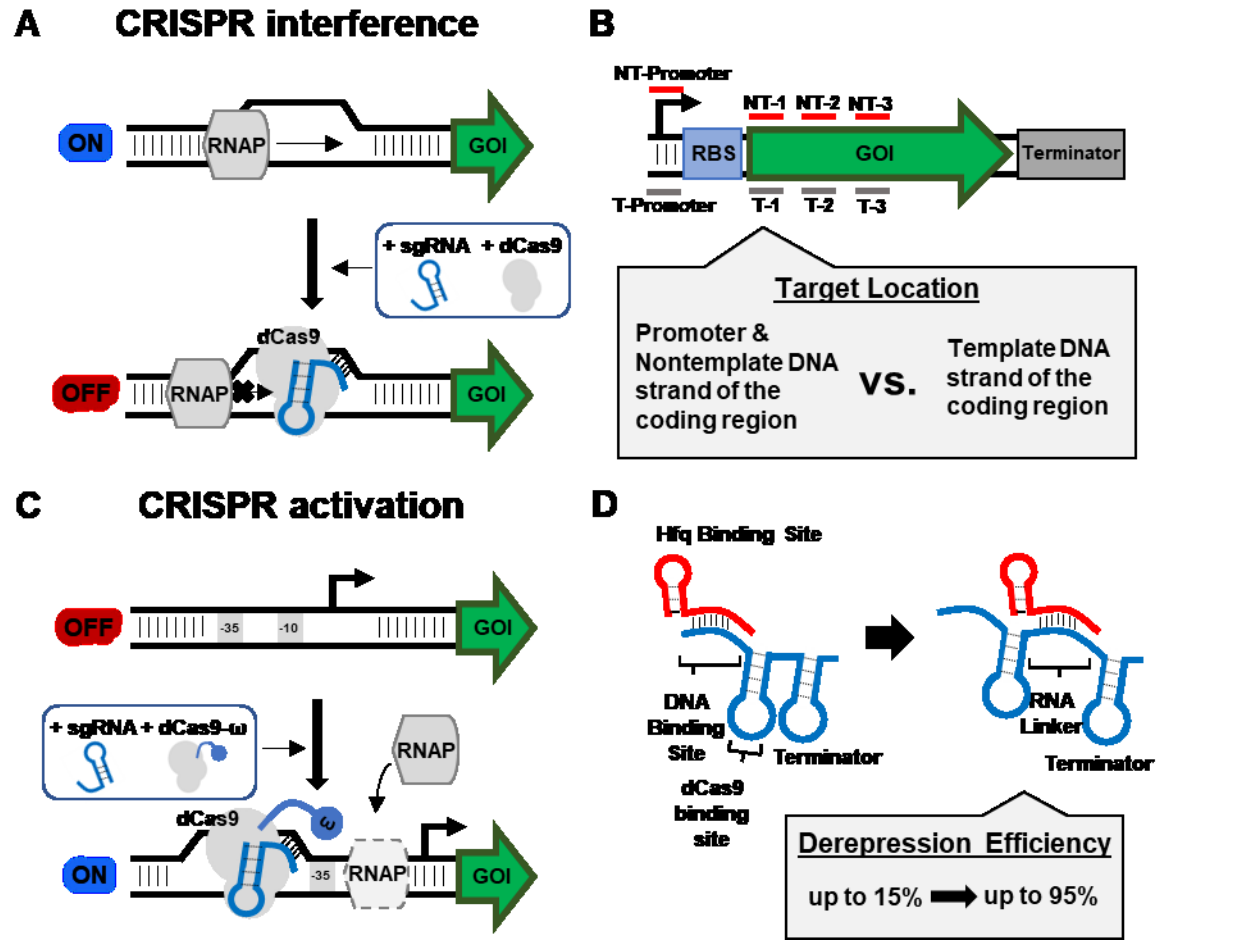


942

943

944

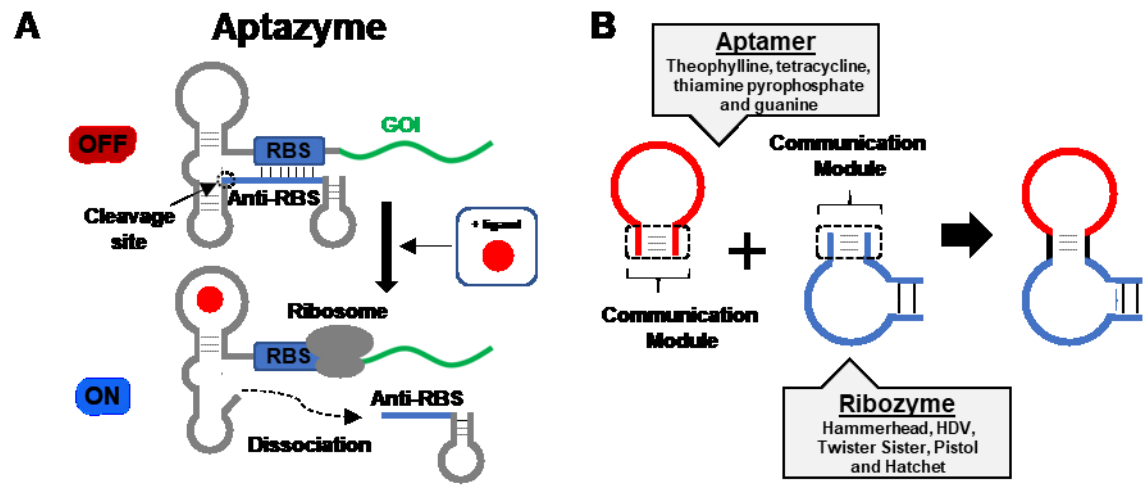
945



946

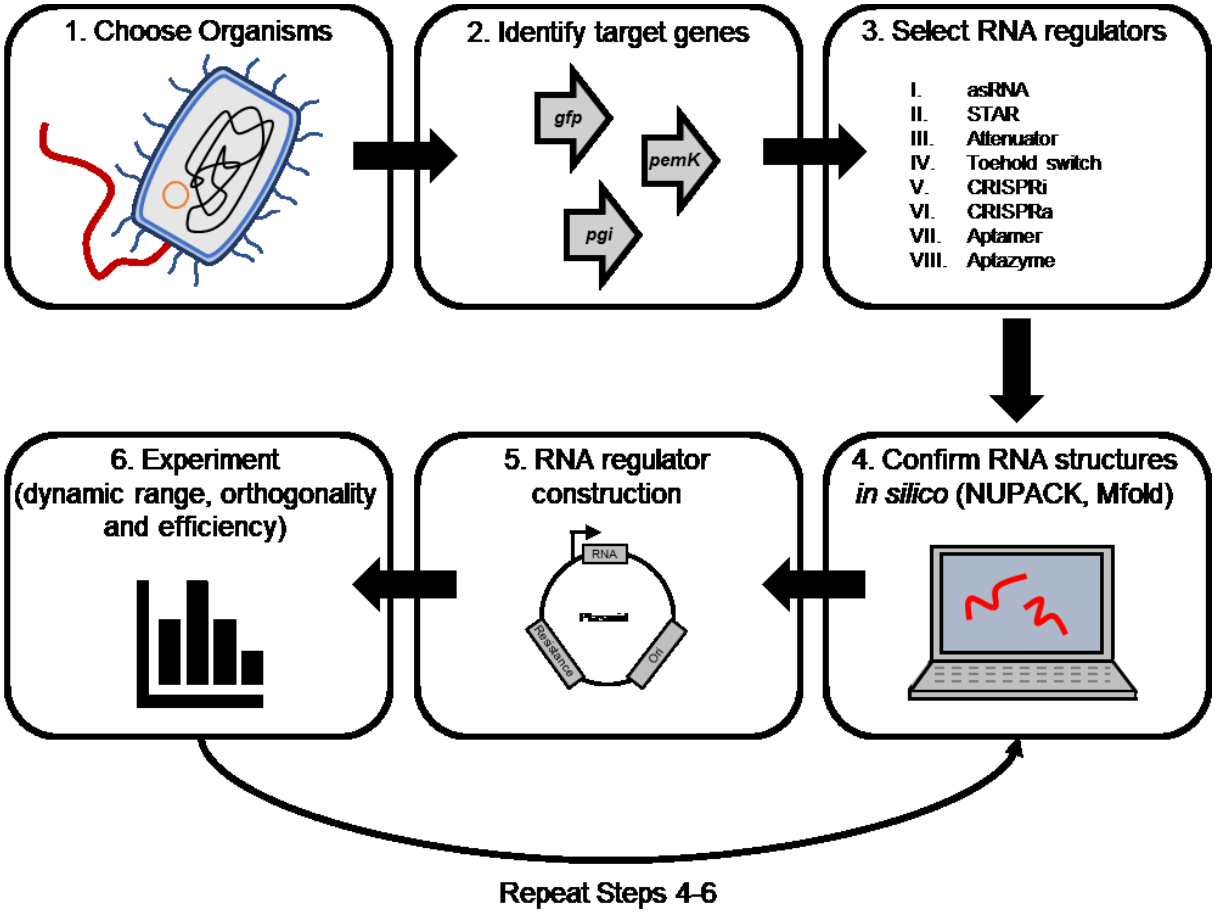
947

948



949

950



951