nature biotechnology

RNA synthetic biology

Farren J Isaacs[1], Daniel J Dwyer[2,3] & James J Collins[3]

RNA molecules play important and diverse regulatory roles in the cell by virtue of their interaction with other nucleic acids, proteins and small molecules. Inspired by this natural versatility, researchers have engineered RNA molecules with new biological functions. In the last two years efforts in synthetic biology have produced novel, synthetic RNA components capable of regulating gene expression *in vivo* largely in bacteria and yeast, setting the stage for scalable and programmable cellular behavior. Immediate challenges for this emerging field include determining how computational and directed-evolution techniques can be implemented to increase the complexity of engineered RNA systems, as well as determining how such systems can be broadly extended to mammalian systems. Further challenges include designing RNA molecules to be sensors of intracellular and environmental stimuli, probes to explore the behavior of biological networks and components of engineered cellular control systems.

Advances in molecular cellular biology and biotechnology have led to insights into the structure and functional properties of the biomolecules and genetic components underlying the complex processes that govern cellular physiology. Building on these advances, synthetic biology seeks both to employ nonnatural molecules to mimic biological behavior and to assemble well-characterized biomolecular components into circuits that perform prescribed functions[1–6]. This emerging field offers the potential to expand our understanding of biomolecular networks and enhance our ability to engineer novel cellular behavior.

Most synthetic biology efforts have focused on engineering gene circuits that rely on protein-DNA interactions to control transcription[1–6]. Recent advances in RNA biology (**Box 1** and **Fig. 1**) and nucleic acid engineering, however, are inspiring the use of RNA components in the construction of synthetic biological systems[7–19]. RNA molecules, which derive sophisticated behavior from an ability to adopt complex structures, can be generated from potentially all possible sequence combinations, leading to diverse secondary structure and function. These structures can exist in the form of modular domains that confer specific and unique functionality. The causal relationship between sequence, structure and function significantly affects the interaction of RNA molecules with proteins, metabolites and other nucleic acids, making RNA a malleable and attractive molecule to drive programmable function.

In this review, we focus on recent studies where RNA molecules have been engineered to regulate gene expression. We first discuss several engineered RNA systems that modulate expression at the level of translation and then describe designer RNAs that function as transcriptional regulators. We highlight efforts in which functional RNAs have been designed to respond to other nucleic acids, proteins or small molecules to precisely control gene regulation. We present this work in the context of synthetic biology, and discuss some of the immediate challenges, future directions and applications of engineered RNA-based systems.

[1]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Program in Molecular Biology, Cell Biology, and Biochemistry, [3]Center for BioDynamics and Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA. Correspondence should be addressed to F.J.I. (farren@genetics.med.harvard.edu).

Published online 5 May 2006; doi:10.1038/nbt1208

Engineered riboregulators

Protein synthesis requires a series of catalytic and regulatory steps involving key cellular machinery. In bacteria, several factors affect translation initiation, including ribosomal recognition of the mRNA ribosome-binding site (RBS)[20,21] and the start codon[22]. Recognizing the importance of RNA interactions between the ribosome and RBS, and based on work on endogenous riboregulators[23–27] (see **Box 1**), we sought to regulate bacterial gene expression by interfering with ribosomal docking at the RBS[13]. From the outset, our objective was to create a modular post-transcriptional regulation system that could be integrated into biological networks and implemented with any promoter or gene. This is in contrast to endogenous riboregulators, which are limited to specific transcriptional and regulatory elements.

In previous engineered schemes for prokaryotic post-transcriptional regulation, repression was attained through antisense RNA or *trans*-acting ribozymes[28]. Conversely, our approach obtains repression of a target gene by forming a hairpin structure in the 5′ untranslated region (5′-UTR) of the mRNA, sequestering the RBS and preventing it from serving as an accessible substrate for ribosome binding (**Fig. 2**). Because this transcript was designed to inhibit translation initiation through an intramolecular reaction, we refer to it as *cis*-repressed mRNA (crRNA). We designed a noncoding RNA (ncRNA), produced in *trans* from a second promoter to activate expression of the *cis*-repressed gene by targeting the stem-loop structure of the crRNA. This ncRNA, referred to as *trans*-activating RNA (taRNA), forms an RNA duplex that causes a conformational change in the crRNA, which unfolds the stem-loop, exposes the RBS and permits translation (**Fig. 2**).

To determine which secondary structures within the *cis* element confer effective repression, we compared the activity of four crRNA variants with various degrees of stem sequence complementarity in the 5′-UTR of the reporter *gfp* gene (green fluorescent protein, GFP). Assaying GFP expression showed that the *cis*-repressed motifs with hairpin structures composed of either pure duplexes, inner loops or dispersed bulges conferred ~98% repression, whereas a fourth variant with less complementarity in the hairpin exhibited significantly less repression[13].

Activation of gene expression requires an efficient intermolecular interaction in which taRNAs target the crRNAs to relieve repression.

Successful intermolecular base-pairing depends on diffusion, sequence composition, higher-ordered structural motifs and RNA dynamics during binding[29]. Given the complexity of these factors, we imitated a mode of intermolecular RNA interactions—linear-loop pairings—seen in natural riboregulator systems, including the *hok/sok* system of plasmid R1 (ref. 26). Interestingly, the variants that showed activation contain dispersed bulges in the crRNA stem-loops, suggesting that these bulges make crRNA molecules susceptible to open complex formation in the presence of a cognate taRNA.

Our system displayed two attractive features: a rapid response time and tunable gene expression activation. We detected a response within the first 5 min of taRNA induction. By assaying steady-state GFP levels at

Box 1 Natural regulatory RNAs

Noncoding RNA (ncRNA) molecules (that is, ones that do not encode proteins) are responsible for various cellular functions[66], including splicing and editing RNA, modifying rRNA, catalyzing biochemical reactions and regulating gene expression[66–69]. These versatile roles arise from RNA's ability to form complex structures that can interact with other RNA molecules, DNA, proteins and small molecules.

Antisense RNAs, for example, can silence gene expression by targeting specific mRNA sequences[29,70] (**Fig. 1**). In prokaryotes, small regulatory RNAs (sRNA) both activate and repress bacterial gene expression in *trans*, by base pairing with one or more target RNAs[23,24,27]. Many of these sRNAs, or riboregulators, use the RNA chaperone Hfq, which is thought to mediate binding between sRNAs and their target mRNAs, and protect sRNAs from RNase degradation[24,71]. These riboregulators (**Fig. 1**) are expressed in response to a variety of conditions and stimuli, including oxidative stress, the presence of toxins, changes in temperature[24] and fluctuations in the concentrations of other RNA molecules or RNA-binding proteins.

Ribozymes (that is, RNA enzymes) share general mechanistic similarities with protein enzymes in the catalysis of biochemical reactions. They perform a diverse array of integral cellular functions, including nucleotide splicing and phosphodiester bond cleavage and formation[67–69] (**Fig. 1**). For example, the ability of ribozymes to mediate the formation of peptide bonds at the ribosome during protein synthesis is a critical step in gene expression[72–75]. To accomplish these biochemical tasks, ribozymes operate in a sequence-specific fashion and often use metal cofactors[69,76,77].

Recently, riboswitches (that is, RNAs that bind small molecules to regulate metabolism and gene regulation) have been discovered and characterized across numerous prokaryotes and eukaryotes[35,37,38,78]. Riboswitches contain aptamer domain sites[36], comprising highly specific pockets in the 5′ untranslated region (UTR) of the mRNAs that bind small molecules or ligands. Once a ligand selectively binds an aptamer site, a conformational change in the RNA structure leads to a change in gene expression (**Fig. 1**). Such changes in expression can originate from riboswitch-mediated control of translation initiation[37], transcriptional termination[37,78] or cleavage of mRNA[35,38].

Small interfering RNAs (siRNA) are a class of 21- to 22-nucleotide double-stranded RNAs (dsRNA) that stimulate post-transcriptional gene silencing through the RNA interference (RNAi) pathway in higher eukaryotes[79–81] (**Fig. 1**). When a cell encounters a long dsRNA, the Dicer enzyme cleaves it into siRNAs. An RNA-induced silencing complex (RISC) degrades the sense strand of the siRNA and uses the antisense strand to target genes for silencing, destroying the target mRNA[81].

Researchers have also discovered a class of 19- to 25-nucleotide, single-stranded RNAs that are encoded in the genome of various multicellular organisms[82–84]. These microRNAs (miRNA) can be developmentally regulated, and are probably formed from cleavage of hairpin RNAs, single-stranded molecules that form self-complementary, stem-loop structures. Like siRNAs, miRNAs are also processed by Dicer and regulate expression by inhibiting protein translation through base-pairing with target mRNAs (**Fig. 1**). Animal miRNAs exhibit partial complementarity to their mRNA targets and act as translational repressors. Plant miRNAs, like siRNAs and rare metazoan miRNAs, typically display near-perfect complementarity and thus cause mRNA degradation[85,86]. An mRNA can contain many binding sites at the 3′-UTR where more than one miRNA can target the same region; the degree of translational repression probably depends on the number of miRNAs bound to a particular mRNA[80,81].

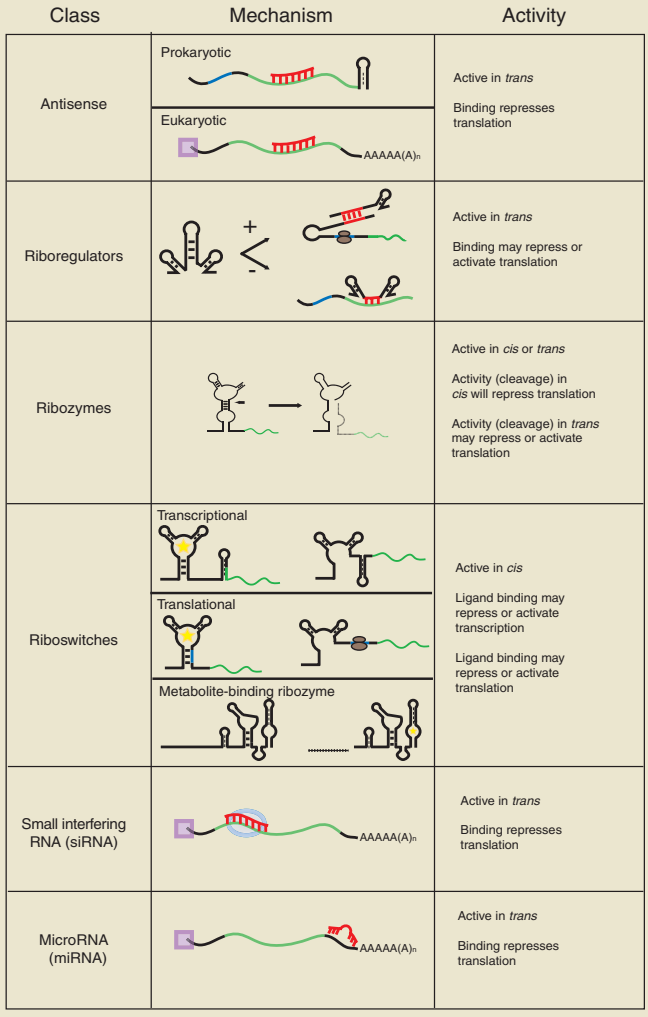

Figure 1 Endogenous RNAs that regulate gene expression. Where applicable, gene coding regions are shown in green, ribosomes in brown and ligands in yellow. mRNA elements highlighted include the prokaryotic ribosome binding site (blue) and eukaryotic 5′ cap (purple) and 3′ poly-A tail.

controlled induction levels of taRNA, we measured a linear dependence between taRNA concentration and GFP expression.

We introduced a number of rational changes, including altering the GC content and size of the *cis*-repressed stem, varying the number of base pairs that participate in intermolecular pairings and incorporating an RNA stability domain on the taRNA. We found that increasing the GC content in the crRNA stem and having a larger number of base pairs participating in the taRNA-crRNA intermolecular interaction improved activation—eightfold for a 24-base-pair design to 19-fold for a 25-base-pair design—from the crRNA repressed state.

Based on the design principles described above, we independently designed four taRNA-crRNA riboregulator pairs. To determine their mutual independence, or orthogonality[30,31], we tested all 16 taRNA-crRNA combinations, which consisted of pairing all 4 cognate (e.g., taRNAa-crRNAa) and 12 noncognate (e.g., taRNAa-crRNAb) taRNA-crRNA combinations. We detected activation only from cognate pairs that contain dispersed bulges, indicating that taRNA-crRNA interactions that expose the RBS require highly specific cognate RNA pairings.

We found that we could achieve comparable levels of *cis*-repression and *trans*-activation with different promoters driving expression of the taRNA and crRNA transcripts, thereby demonstrating the modular nature of the system. We thus think these engineered riboregulators can be used with any prokaryotic promoter or gene, making them well suited to be used as control switches in the construction of synthetic gene networks and in the rewiring of natural networks.

Future extensions of engineered riboregulators face two immediate technological challenges. One is to determine how to successfully integrate rational design and evolution-based techniques to generate new and enhanced (e.g., ligand-modulated) riboregulation. The other is to determine how the modular design principles used in constructing prokaryotic riboregulators can be applied to create related systems that operate in eukaryotes, including mammalian cells. Initiation of translation in eukaryotic cells is an intricate process that involves a series of tightly regulated events and specific mechanisms (e.g., cap-dependent scanning and internal ribosome entry[32,33]), allowing the recruitment of ribosomal subunits to mRNA. Designing mammalian riboregulators may be achieved by interfering with critical eukaryotic initiation factors that direct ribosomal subunits to the mRNA, similar in spirit to the engineered prokaryotic version.

Work on synthetic riboregulation may also further motivate the search for similar, endogenous riboregulators in prokaryotes and eukaryotes. The task of identifying additional native riboregulators will require advanced bioinformatics approaches integrated with biological assays.

Engineered ribosome-mRNA pairs

Rackham and Chin took a second approach to regulating translation via synthetic RNA, combining elements of rational design with directed evolution techniques in engineering the specificities of ribosomes and mRNAs to control gene expression[18]. They tailored the molecular interaction of *Escherichia coli* ribosome-mRNA pairs such that an engineered ribosome could translate only its engineered mRNA partner and not any endogenous mRNA; similarly, a native *E. coli* ribosome would not be able to initiate translation on an engineered mRNA. The engineered ribosome-mRNA pairs function in an orthogonal[30,31] fashion with respect to the native ribosome and mRNAs.

A key goal in the development of synthetic orthogonal ribosome-mRNA pairs is to eliminate any pleiotropic effects. From the outset, Rackham and Chin pursued a strategy to reduce interference with ribosome assembly, rRNA processing and cellular viability. To do so, they developed a positive and negative selection strategy to evolve orthogonal ribosome-orthogonal mRNA (O-ribosome-O-mRNA) pairs that permit robust translation.

Rackham and Chin used a positive selection step, involving the gene encoding chloramphenicol acetyltransferase (CAT) and the antibiotic chloramphenicol, in conjunction with a negative selection strategy involving uracil phosphoribosyltransferase (UPRT)[34]. By fusing the genes encoding CAT (*cat*) and UPRT (*upp*) downstream of a constitutive promoter and the RBS, the authors constructed a single transcript that affords either positive or negative selection, determined by the type and dose of the small molecule added to the medium (**Fig. 3**).

To generate the O-ribosome-O-mRNA pairs, Rackham and Chin first synthesized a library of all possible RBSs and another of all possible 16S rRNA anti-RBS sequences, allowing them to explore a matrix of more

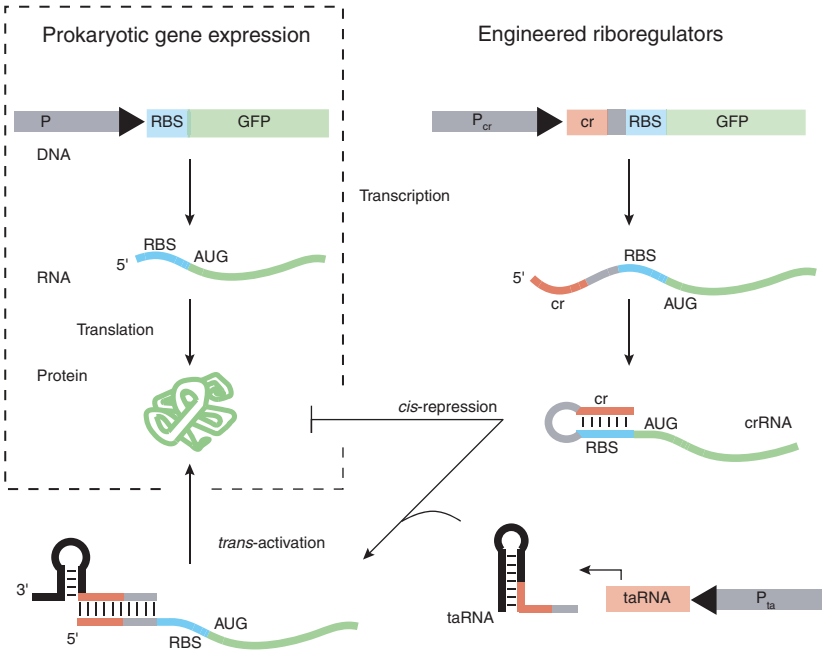

Figure 2 Engineered riboregulators. An engineered riboregulation system employs negative and positive regulation of prokaryotic gene expression at the post-transcriptional level[13]. The engineered riboregulation system contains a short sequence (*cis*-repressed, cr, red) inserted downstream of a promoter (Pcr) and upstream of the ribosome binding site (RBS, blue). After transcription, the *cis*-repressed sequence, which is complementary to the RBS, drives the formation of a stem-loop in the 5′-UTR that prevents ribosome docking and represses translation (*cis*-repression). The resulting mRNA is referred to as *cis*-repressed RNA (crRNA). A second independent promoter (Pta) is responsible for the transcription of a small, noncoding RNA (*trans*-activating RNA, taRNA), which targets its cognate crRNA with high specificity. The subsequent RNA-RNA linear-loop interaction promotes structural rearrangement of the crRNA, thus exposing the obstructed RBS and enabling translation. Control of GFP expression is shown, but this regulation system is modular and can be implemented with potentially any endogenous or synthetic gene.

than 10^9 unique mRNA-rRNA combinations (**Fig. 3**). The authors next used a two-step process that combined their positive-negative selection strategy with their mRNA and rRNA libraries. They first selected for mRNA sequences that are not substrates for endogenous ribosomes, by introducing the mRNA library into *E. coli* cells and growing the cells in the presence of 5-fluorouracil (5-FU), which selects against any mRNAs that permit translation of UPRT. Viable cells possessed orthogonal mRNAs that do not interact with endogenous ribosomes. These cells were subsequently transformed with the library of mutant ribosomes and grown in medium containing chloramphenicol; thus, only active ribosomes that translate orthogonal mRNA pairs were selected for under these conditions. From among the 10^{11} clones selected by this two-step process, the authors found four distinct O-mRNAs and ten distinct O-rRNA sequences, suggesting that the positive selection step led to significant convergence of the mRNA sequences.

In a follow-up study[19], Rackham and Chin combined several orthogonal pairs in a single cell, to see whether simultaneous expression of multiple distinct ribosomes can be achieved. They constructed a set of logical AND/OR gates; in the case of the AND gate, for example, the authors separately cloned the genes for two fragments—α and ω—of *lacZ* onto distinct O-mRNAs such that the expression of both genes is required for β-galactosidase activity (*lacZ* expression). They found that a β-galactosidase signal is detected only when O-mRNAs containing α and ω are coexpressed with each of their O-ribosomes, demonstrating that multiple orthogonal ribosomes can simultaneously function and generate logic operations in the same cell.

This engineered ribosome-mRNA system is well-suited for the creation of synthetic, orthogonal cellular pathways, and could be used to probe or modify the translational dynamics of natural networks, providing a powerful tool for the study of isolated network components. Moreover, this system affords the ability both to tune the flux through cellular pathways and to generate translationally based reversible knockouts. Having the ability to modify translational flux could be of tremendous utility in metabolic engineering and biofactory design applications.

Before large-scale implementation of orthogonal networks, two critical issues need to be addressed. First, rigorous optimization of O-ribosome-O-mRNA pairs is required to eliminate crosstalk between orthogonal pairs. Second, ways of channeling the appropriate cellular resources (e.g., ATP) to orthogonal pathway expression that do not debilitate basic cellular function need to be developed.

Engineered ligand-controlled riboregulators

Recent studies have shown that many essential genes in bacteria rely on RNA-based mechanisms such as riboswitches for post-transcriptional control of gene expression[24,35]. Riboswitches[35] are a class of aptamer-like[36] (target-binding nucleic acids, where the target can be various molecular species), *cis*-acting elements localized in the 5′-UTR of the mRNA encoding the regulated gene (**Box 1** and **Fig. 1**). These motifs are composed, in part, of highly specific binding pockets for target metabolites and respond by conformationally altering the mRNA, resulting in modulation of gene expression.

The first natural riboswitch was discovered by Breaker and colleagues who showed that mRNAs encoding enzymes involved in thiamine (vitamin B_1) biosynthesis in *E. coli* can bind thiamine in their 5′-UTRs[37]. The mRNA-ligand complex adopts a distinct structure that sequesters the RBS and leads to a reduction in gene expression. The fabrication of synthetic RNAs that undergo conformational transitions after exposure to specific small molecules has become possible through a framework of rational design and selection methods. For example, in recent work Suess *et al.* integrated a theophylline-binding aptamer into the 5′-UTR of a reporter gene such that binding of the molecule to the reporter mRNA results in repression of gene expression via helix slippage and subsequent RBS occlusion[14].

Researchers in synthetic biology have assembled ncRNAs that couple riboswitch-like aptamer domains with antisense effector domains to create gene expression tools. Bayer and Smolke recently designed a set of *trans*-acting riboregulators, called antiswitches, that function in *Saccharomyces cerevisiae*[16]. When a small molecule binds to the aptamer domain, it causes

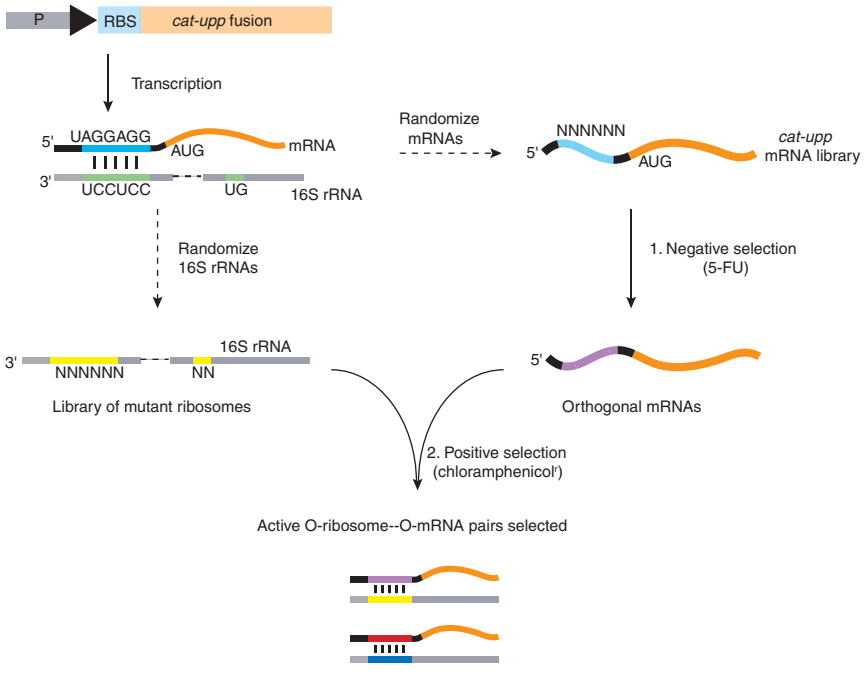

Figure 3 Engineered ribosome-mRNA pairs. An engineered prokaryotic gene expression system was designed that utilizes synthetic ribosomes and mRNA RBS sequences to control target protein expression[18,19]. The engineered expression system modifies translation initiation by changing the sequence of the 16S rRNA anti-RBS (aRBS) component of the ribosome (gray and green) and two downstream rRNA bases, which geometrically assist the ribosome-mRNA interaction[21], with the ribosome binding site (RBS, light blue) on the mRNA (only the classic RBS-aRBS interaction is illustrated). As illustrated, a chloramphenicol acetyltransferase (*cat*)-uracil phosphoribosyltransferase (*upp*) fusion gene (orange) was constructed, and a *cat-upp* mRNA library was fabricated via randomized evolution of the RBS domain. *E. coli* cells were transformed with the synthesized library, and the *cat-upp* mRNAs, which contained RBS sequences that prohibited translation by endogenous ribosomes (purple), were negatively selected for in the presence of 5-fluorouracil (5-FU). Cells that survived selection were subsequently transformed with a library of mutagenized ribosomes (gray and yellow). Those harboring orthogonal ribosome-mRNA pairs (O-ribosome-O-mRNA pairs, yellow-purple and blue-red pairs) were positively selected for in the presence of the antibiotic, chloramphenicol.

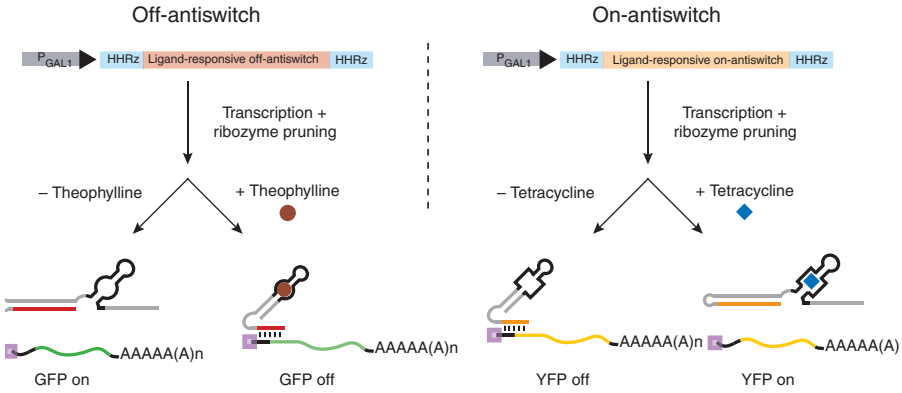

Figure 4 Engineered ligand-controlled riboregulators. Ligand-controlled riboregulators, called antiswitches, were engineered to control translation in *S. cerevisiae*[16]. These antiswitches, expressed from the GAL1 promoter, contain an aptamer domain that responds to the presence of specific ligands. Ligand binding causes a conformational rearrangement that affects the availability of the antisense domain. In the default state of the off-antiswitch, the antisense domain (red) is sequestered in a 5′ stem-loop (gray) and GFP expression proceeds. Binding of theophylline (brown circle) induces antiswitch structural rearrangement that exposes the antisense domain for interaction with its mRNA target. Specifically, the antiswitch binds to a portion of both the mRNA 5′-UTR (black) and coding region (green), effectively repressing translation. In the case of the on-antiswitch, gene expression is repressed (yellow) in the default state by the exposed antisense domain. Following tetracycline (blue diamond) binding, the antisense domain is sequestered and target mRNA translation is permitted (yellow). Two strategically placed hammerhead ribozyme sequences[108] (HHRz, light blue) provide highly efficient cleavage at the 5′- and 3′ -boundaries to ensure formation of the desired antiswitch structure, which is free of interfering ribonucleotides that could disrupt structure and binding. mRNA 5′-UTR and 3′-UTR (black) sequences are depicted, as are the 5′ cap (purple) and 3′ poly-A tail.

a conformational change that affects the ability of the antisense domain to bind its target mRNA and modulate its translation. The authors initially chose the small molecule theophylline and its corresponding aptamer domain, and designed their antisense domain to bind a region on the mRNA that spans a fraction of the 5′-UTR and nucleotides in the open reading frame (ORF) of the reporter *gfp* gene (**Fig. 4**).

Bayer and Smolke constructed two types of ligand-responsive riboregulators: an off-antiswitch and an on-antiswitch (**Fig. 4**). In the absence of theophylline, the antisense domain of the off-antiswitch is sequestered in a stem-loop, preventing target mRNA repression. When theophylline binds to the aptamer domain, the off-antiswitch undergoes a conformational change that exposes the antisense domain and represses translation of the mRNA. In the absence of theophylline, expression of the off-antiswitch decreased reporter gene expression by ~30%, relative to controls, indicating that antiswitch structural instability causes moderate translational attenuation even when the antisense domain is predicted to be bound in the stem loop. The authors showed that, after the addition of theophylline, the antiswitch molecules quickly suppress reporter gene translation and reduce expression to background levels, through a switch-like response, indicating effective repression and a rapid temporal response.

For the on-antiswitch variant, in the absence of the ligand, an antisense domain is free to bind a target mRNA (**Fig. 4**). Upon ligand binding at the aptamer domain, the on-antiswitch undergoes an allosteric transition that causes the antisense domain to form a stem-loop, rendering it inactive. As a result, target gene expression is upregulated.

By introducing targeted mutations, Bayer and Smolke changed the thermodynamic properties of their off-antiswitches, creating variants with both lower (less stable stem structures) and higher (more stable stem structures) theophylline-switching thresholds. Adding nucleotides to the antisense stem reduced background translational inhibition from 30% to 10%. Through these mutations, they also isolated antiswitch variants that exhibit a trend to a more graded response, where transitions occur over various levels of theophylline concentration[16]. These experiments show that targeted mutations to important structural domains in RNA molecules can lead to phenotypic changes in system response.

To demonstrate the modular ligand-binding capability of their off-antiswitches, Bayer and Smolke replaced the aptamer domain for theophylline in one case with a theophylline domain with a tenfold higher K_d and in another case with an aptamer domain for tetracycline. As expected, the modified theophylline aptamer switched at a higher ligand concentration, setting the stage for customized switches that elicit a response at different levels of input ligand. The tetracycline aptamer variant, similar to the theophylline antiswitch, responded in a switch-like manner to tetracycline. The authors also constructed two antiswitches where they combined different antisense domains, one for GFP and the other for yellow fluorescent protein (YFP), with theophylline and tetracycline domains, respectively, and demonstrated that each switch can respond independently and regulate different mRNA targets in the same cell (**Fig. 4**).

These antiswitches provide a novel framework combining multiple domains that, together, modulate gene expression in eukaryotes post-transcriptionally. By assembling various combinations of antisense and aptamer domains, researchers can now repress or activate the expression of presumably any target gene in response to a nontoxic and membrane-permeable small molecule. In addition, engineered aptamer-antisense RNA constructs could serve as ligand-dependent *in vivo* sensors of mRNA levels and metabolite-sensitive tools for rewiring regulatory networks. Constructing a genome-wide set of functional antiswitches represents a future challenge because the antisense domain targets an mRNA sequence that spans part of the 5′-UTR and ORF (**Fig. 4**). This may require revised design strategies that account for potential secondary structure formations at the 5′-UTR-ORF junctions, which could prevent antiswitch binding to its target mRNA.

Engineered ribozyme-mediated RNA switches

Research on riboswitches has shown that direct interaction of intracellular metabolites and RNA sequences can affect translation and transcription through RNA self-cleavage[35,38]. For example, the mRNA that encodes the *glmS* gene in *Bacillus subtilis* and other Gram-positive bacteria contains a ribozyme that cleaves the transcript in response to increasing levels of glucosamine-6-phosphate (GlcN6P), the metabolic product of *glmS*[38]. As the basis for the design of an engineered RNA gene expression platform in mammalian systems, Yen *et al.* employed the properties of a ligand-responsive ribozyme to regulate mRNA translation[15].

The strategy used by Yen *et al.*[15] relies on a specific self-cleaving ribozyme that efficiently cleaves an mRNA molecule at a particular location and a signal (e.g., a small molecule or nucleic acid) that inhibits self-cleavage of the *cis*-acting ribozyme. The authors explored a large set of candidate ribozyme sequences capable of efficient cleavage and

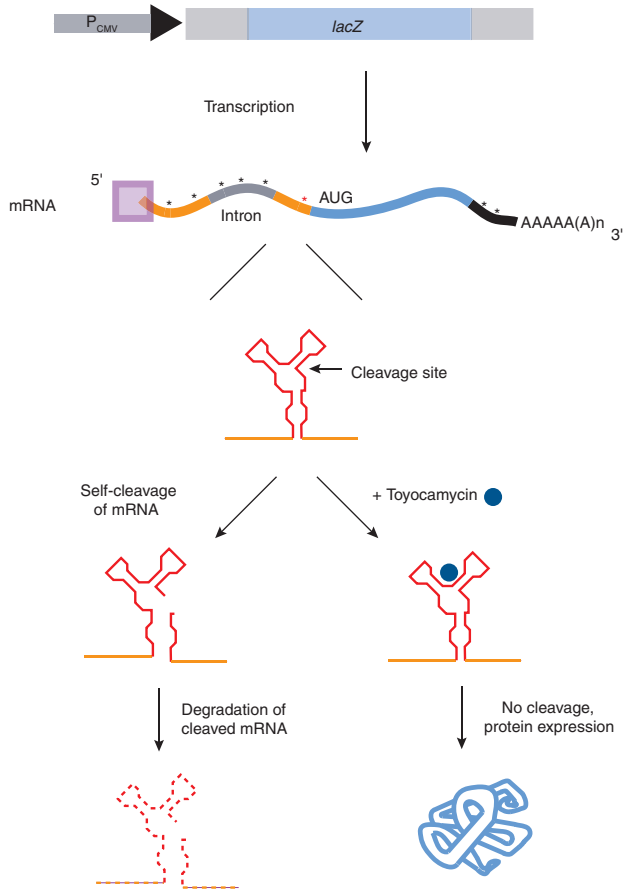

Figure 5 Engineered ribozyme-mediated gene expression. A ribozyme-mediated gene expression system was developed to control translation in mammalian systems[15]. In this expression system, a cytomegalovirus (CMV) promoter drives the transcription of the engineered ribozyme mRNA, which encodes *lacZ* (blue). *Cis*-acting ribozymes were constructed, using the *Schistosoma mansoni* Sm1 hammerhead ribozyme as a foundation. Through high-throughput screening studies, ligands that inhibit the catalytic activity of the Sm1 ribozyme were identified. The best performing ligand-binding ribozyme was inserted at various locations (noted by asterisks) in the 5′-UTR, 3′-UTR and intron region of the *lacZ* reporter gene. In the absence of ligand (toyocamycin, blue circle), the ribozyme exacts efficient and constitutive autocleavage of the *lacZ* mRNA. In the presence of toyocamycin, catalytic activity is inhibited, mRNA degradation is prevented, and protein expression occurs. The red asterisk denotes the aptamer-ribozyme location that conferred the greatest *lacZ* expression. mRNA 5′-UTR (orange), 3′-UTR (black) and intron (gray) sequences are depicted, as are the 5′ cap (purple) and 3′ poly-A tail.

introduced these sequences into various locations in a *lacZ* vector construct (**Fig. 5**).

Whereas most ribozymes in the library did not affect gene expression, two—the hammerhead Pst-3 ribozyme[39] and Sm1 ribozyme[40]—exhibited detectable changes in expression; Pst-3 and Sm1 showed 13-fold and 19-fold differences in expression, respectively. Yen *et al.* found that placement of the Sm1 ribozyme directly upstream of the translation start site yielded the most dramatic decrease in gene expression (**Fig. 5**), and that inserting a second adjacent ribozyme further suppressed expression. In addition, they demonstrated the modularity of their system by successfully incorporating alternative promoter and reporter gene elements in the design. They showed that their Sm1 ribozyme-mediated expression system can function efficiently in several commonly used mammalian cell lines transfected with their ribozyme-engineered *lacZ* construct.

To achieve ribozyme-mediated regulation, the researchers identified nontoxic and membrane permeable ligands that specifically interact with the ribozyme target. They found that synthetic morpholino oligonucleotides[41] repressed ribozyme cleavage, leading to increased *lacZ* expression. Using a high-throughput screening method, the authors likewise found the nucleoside analog toyocamycin[42] to be a strong inhibitor of ribozyme cleavage of mRNA.

Yen *et al.* used adeno-associated virus vectors incorporating the toyocamycin-responsive Sm1 ribozyme, to examine whether their ribozyme-mediated gene expression system could function *in vivo*[15]. In the absence of toyocamycin, reporter gene expression from the ribozyme construct was undetectable in the retinas of nude mice injected with the vector. After toyocamycin treatment and incorporation into the mRNA, repressed ribozyme activity resulted in readily detectable levels of reporter gene expression.

The immediate challenge facing this novel platform is expanding the number of ligands that exogenously modulate translation. This can be achieved by screening for additional small molecules that inhibit the Sm1 ribozyme. Moreover, it may be possible to implement rational design strategies to customize the inducible switch by incorporating small-molecule aptamer domains.

Such design modifications could bring the prospects of sophisticated RNA-based function within reach and yield potential therapeutic applications. Key challenges for such applications include the development of safe and effective means for delivering the construct to a sufficient number of cells, and incorporating this RNA-based expression system into synthetic networks with memory[43–47], which would require only transient introduction of a small molecule for long-term, stable functionality.

Engineered RNA transcriptional activators

Eukaryotic transcription factors consist of two modular protein domains: a DNA-binding domain and an activating domain. A typical transcriptional activator binds to specific sites on the DNA through its DNA-binding domain, localizing proteins responsible for transcriptional initiation to a specific promoter. The activating domain is primarily responsible for modulating specific interactions with protein machinery that direct transcription[48]. Because eukaryotic transcription factors separate their functionality into modular DNA-binding and regulatory domains, they represent excellent components for an engineered regulatory switch. Recognizing this, Saha *et al.*[10] and Buskirk *et al.*[11] independently constructed RNA-based transcription factors that activate transcription, a role traditionally thought to belong to proteins. Both studies employed a similar design strategy, inspired by previous work of Sengupta *et al.*[49], where a set of naturally occurring RNA sequences found in *Saccharomyces cerevisiae* activates transcription in the absence of a protein transcriptional activator.

The Saha[10] and Buskirk[11] studies each used a yeast three-hybrid strain[49,50], which combines auxotrophic markers, reporter genes and a method for localizing RNA to these elements. Specifically, these studies[10,11] used a yeast strain containing a LexA-MS2 fusion protein[51,52], together with a construct incorporating the selectable *HIS3* gene and a *lacZ* reporter driven by promoters with LexA binding sites (**Fig. 6**). Thus, RNAs containing an MS2 hairpin can be localized to the promoters driving the expression of *HIS3* and *lacZ*.

Libraries of synthetic transcriptional activator RNAs were created in both studies[10,11], such that a variable region was designed as the RNA-based transcriptional activation domain and two MS2 hairpin regions serve to localize the synthetic RNA molecule to the LexA-MS2 fusion

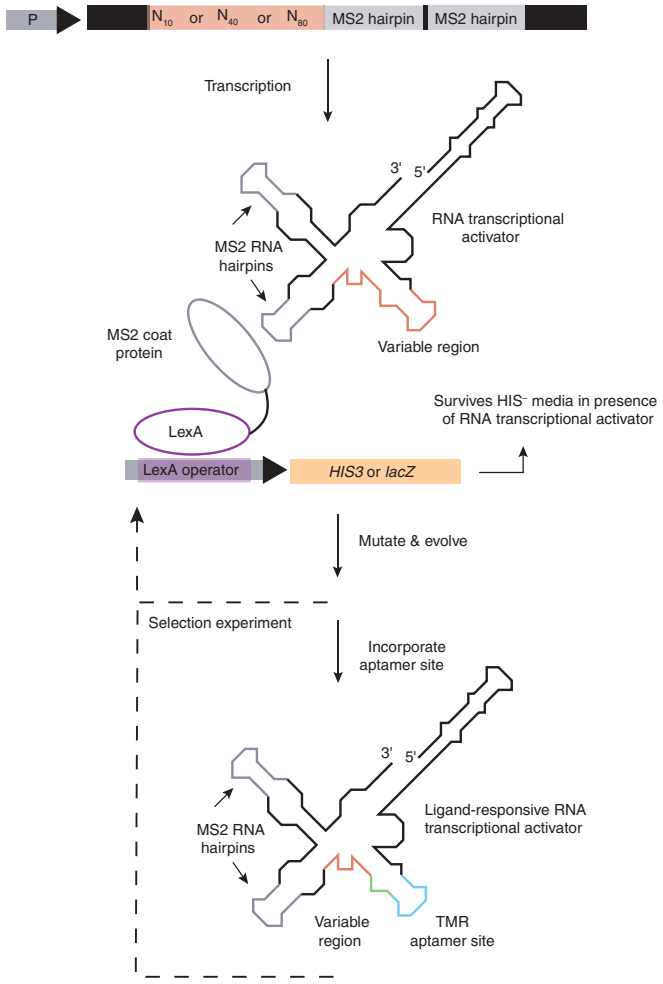

Figure 6 Engineered RNA transcriptional activators. The schematic outlines the development of engineered RNA transcriptional activators[10–12] that contain two main domains: an MS2 RNA hairpin domain (gray) and an activating domain that contains a randomized sequence of variable length (orange). The two MS2 RNA hairpins serve to localize the engineered RNA to a LexA-MS2 fusion protein with high affinity (purple) that is bound to the LexA operator site in a promoter. Three randomized libraries of 10, 40 and 80 nucleotide activation regions were synthesized[10–12]. HIS3 auxotrophic selection of transcriptionally active engineered RNAs was achieved by growth in HIS⁻ media. Transcription activation levels of selected RNAs were directly quantified by *lacZ* reporter gene expression measurements. A tetramethylrosamine (TMR) aptamer site (blue) was integrated into a highly expressing RNA transcriptional activator sequence (red). TMR binding allosterically modulates the transcriptional activation domain, permitting exogenous control of the RNA transcriptional activator. The randomized linker region is green.

protein (**Fig. 6**). The length of the variable regions differed between the studies—Saha *et al.* synthesized RNA activator libraries with randomized domains of 10 ribonucleotides (N_{10}), whereas Buskirk *et al.* created two libraries with variable regions of 40 (N_{40}) and 80 (N_{80}) randomized ribonucleotides, respectively (**Fig. 6**).

Saha *et al.* found several RNA molecules capable of activating transcription, with the highest variant exhibiting expression levels 400-fold above background. Compared to protein transcription factor fusions, they found that their engineered RNA activators exhibited tenfold greater expression than LexA-TBP (TATA box-binding protein) and ~15-fold lower expression than the strong LexA-Gal11 and LexA-Gal4 activators in yeast. Buskirk *et al.* found several N_{40} RNA molecules that activated transcription at levels comparable to a Gal4 protein control[11], which is significantly higher than the reported activation from the N_{10} library of Saha *et al.*[10].

Saha *et al.* also found that their system worked well with other engineered promoters, demonstrating that RNA molecules serve as broad transcriptional activators in yeast. Of note, all of the N_{10} functional RNA variants from the Saha *et al.* study[10] contained a six-nucleotide consensus region, and upon searching the yeast genome database, Saha *et al.* found 27 transcribed regions that encode a ten-nucleotide region identical to one of their RNA transcriptional activators[10].

Buskirk *et al.* isolated the strongest activating clone (ten times greater than the Gal4 protein control) from their N_{40} library and subjected its 40-nucleotide variable region to directed mutagenesis. After exposing this mutagenized library to more stringent selective conditions to evolve activators with greater dynamic range (**Fig. 6**), the authors isolated a clone with expression levels 53-fold greater than the Gal4 activation domain. These findings show that a mutagenesis–high stringency selection strategy can be used to evolve transcriptional activating RNAs with improved expression, comparable to activation by protein transcription factors. Furthermore, this work[11], together with the work of Saha *et al.*[10], suggests that variable regions of intermediate length (~40 nucleotides) provide the strongest set of RNA transcriptional activators. One likely interpretation of these results, as proposed by Buskirk *et al.*[11], is that the secondary structural diversity derived from longer RNAs (40 versus 10 nucleotides) may be required to achieve activation on par with its protein counterparts, and at longer lengths (80 nucleotides), unstructured single-stranded regions are prone to degradation and thus cause RNA instability.

In subsequent work[12], Buskirk *et al.* constructed a ligand-dependent RNA transcriptional activator such that transcription can be modulated exogenously by a small molecule. To achieve this mode of control, the authors designed into their gene expression system an aptamer domain, comprising a highly specific binding pocket for target metabolites (**Fig. 6**). The authors replaced several nonessential bases of their RNA transcriptional activator with a well-characterized RNA aptamer that binds tetramethylrosamine (TMR)[53], a molecule that is membrane-permeable and nontoxic to yeast at low concentrations. The authors designed the ligand-dependent RNA activator such that the TMR aptamer domain destabilizes the loop, leaving the RNA nonfunctional. Upon TMR binding, the loop stabilizes, returning function to the RNA transcriptional activator. For this RNA-aptamer system, the authors found a linearly dependent relationship between transcriptional activation and TMR concentration. This result supports the notion that TMR increases levels of transcriptional activation by specifically binding to its RNA aptamer, thereby regulating the transcriptional domain.

Regulation of transcription with engineered RNA molecules provides a novel tool to control eukaryotic gene expression. The current design contains specific RNA motifs (MS2) that localize the synthetic RNA transcriptional activators to a target promoter, which contains the corresponding MS2 protein fused to a DNA-binding protein (LexA). It will be challenging to develop a set of modular components to generalize the system, given that the current design requires specific domains (e.g., MS2 directs RNA-protein interaction and LexA directs protein-DNA interaction) to achieve functionality. For example, additional domains that direct RNA-protein interactions would be required to simultaneously use more than one synthetic transcriptional activator *in vivo*. Buskirk *et al.* made a significant advance to address this by incorporating aptamer domains that sensitively respond to a specific small molecule[12]. Such developments will enable simultaneous control of multiple genes from large numbers of synthetic RNA transcriptional activators.

The design and selection of allosteric RNA transcription factors containing ligand-responsive aptamer domains adds an additional layer of gene expression control, and increasing the number of aptamer domain variants represents another potential hurdle for this system. One can readily imagine designing a library of modular RNA transcriptional regulators, with a wide array of ligand aptamer sites, which would permit precise transcriptional control in response to potentially any nontoxic, membrane-permeable small molecule. This work may also motivate the development of RNA transcriptional activators that operate in mammalian cells. Along these lines, it may be possible to couple RNA transcriptional activators to endogenous signals or molecules of interest to create cell-based biosensors than can measure and respond to physiological signals of interest (e.g., changes in blood glucose levels).

Future directions and applications

All of these engineered RNA-based systems employ an element of rational design in which functionality is derived in a modular fashion through domains of RNA sequence that confer a specific activity (e.g., RNA-RNA binding, ligand-RNA binding, ribozyme activity). These studies show that engineering sequence-specific changes (e.g., changing the degree of complementarity in an RNA stem-loop) in these domains can result in significant changes in RNA structure and function. Such changes affect the thermodynamic properties of the RNA molecules, as well as their response to environmental stimuli and their ability to bind to other molecules.

Designer RNA switches, which possess quick response times to biological and environmental stimuli, can be employed as cellular sensors to detect fluctuations in biological signals, metabolites or other small

Box 2 *In vitro* nucleic acid systems

A large body of *in vitro* work has led to the development of designer nucleic acid molecules and RNA- and DNA-based devices, including nucleic acid aptamers, which respond to chemical or biological inputs[36,87–89]; ribozymes[90–97] and deoxyribozymes[98–100], which catalyze a variety of chemical transformations; molecular computers[101] and DNA- and RNA-based structures[102]. In addition, modular rational design approaches and *in vitro* selection techniques have been used to create *in vitro* RNA switches and logic gates with tailored responses[91,93,97,103,104] (**Fig. 7**).

Researchers have also developed nucleic acid-based autonomous machines, or molecular automata, that perform sophisticated tasks in response to their environment (**Fig. 7**). In one study, a molecular automaton comprising a Boolean network of deoxyribozyme-based logic gates encoded a version of the game tic-tac-toe[105]. The logic gates were constructed from a deoxyribozyme module that cleaves a fluorogenic substrate and a stem-loop module that undergoes a conformational change upon binding of a complementary oligonucleotide. Complementary input oligonucleotides and loop pairs were selected to generate a set of logic gates, which were combined to create a deoxyribozyme-based molecular automaton. The system was set up to execute a linear program based on the rules of tic-tac-toe, releasing fluorescence signals in response to input oligonucleotides. In impressive fashion, the automaton performed as programmed, executing several successful games of tic-tac-toe.

In a related study, a molecular automaton was constructed from DNA molecules and enzymes[106], and used to 'diagnose' mRNA of disease-related genes in an *in vitro* model[107]. This system contains an input module that recognizes specific mRNA levels, a computation module that implements a stochastic molecular automaton and an output module that releases a short single-stranded DNA molecule or antisense drug. Detection of each mRNA is achieved through two automata, one for a positive diagnosis and one for a negative diagnosis. The presence of a particular mRNA increases the probability of a positive transition and decreases the probability of a negative transition (vice versa in the absence of the mRNA). Changing the ratio between the positive and negative transitions for a particular mRNA provides fine control over the sensitivity of diagnosis. The positive diagnosis automaton releases a drug antisense molecule, and the negative diagnosis automaton releases a drug suppressor molecule. Together, the two automata provide fine control of the final active drug concentration by precisely determining the ratio between drug antisense and drug suppressor molecules.

Future *in vitro* goals include increasing the precision, number and functional complexity of molecular switches and automata. The key *in vivo* challenge, given the success of current efforts of engineered RNA systems, lies in developing methods to integrate these *in vitro* systems into cellular environments, while eliminating pleiotropic effects and maintaining the desired behavior of the synthetic network.

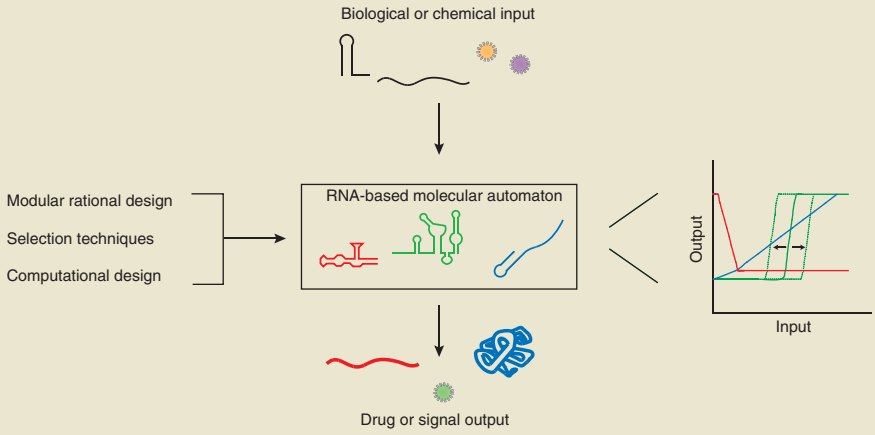

Figure 7 *In vitro* nucleic acid systems. A variety of engineered nucleic acid molecules have been constructed using *in vitro* synthesis and selection techniques. These *in vitro* DNA and RNA molecules produce a specified output (biological molecule or signal) in response to a biological or chemical input. The system response is determined by a single nucleic acid or interacting network of nucleic acids (e.g., molecular automaton) that perform sophisticated tasks in response to their environment. The ability to generate varied responses (colored curves in the input-output plot) and finely tune a particular response (green curves in the input-output plot) is obtained from a combination of rational design, computational design and selection techniques. Inputs consist of nucleic acids, proteins or signals; networks of nucleic acids comprise the molecular automaton; outputs can exist as nucleic acids (red), signals (green) and protein (blue).

molecules. For example, Sando *et al.*[54] constructed a luminescence-linked riboregulator detector for genotyping, and used this system to sensitively distinguish between different input nucleic acid alleles. It should also be possible to interface engineered RNA switches with natural regulatory networks and signaling pathways in a 'plug-and-play' fashion to create programmable cells[46] and whole-cell biosensors.

The tight regulation and highly specific base-pair interactions of RNA switches also afford immediate applications for regulatory RNAs in functional genomics. Highly specific RNA probes could be engineered and used to selectively perturb natural biological systems to reveal functional properties of large-scale networks and specific proteins and genes.

By combining certain features of these engineered RNA systems, one can envision creating more complicated RNA switches with sophisticated functionality. For example, engineered RNA switches that combine multiple functional domains could be designed and used to sense different stimuli and generate stimulus-specific functional responses. Along such lines, interesting *in vitro* systems have been recently developed (**Box 2**) that could be translated into *in vivo* systems. Novel nucleic acid engineering strategies are needed to link functional aspects of such networks with regulatory, catalytic and structural components that control endogenous cellular processes. As these systems become more complex and the need to generate enhanced functionality grows, computational design and experimental validation approaches[55] for allosteric RNAs will become increasingly valuable.

The utility of designer RNA molecules may be further realized by linking engineered RNA regulatory components with other existing or emerging biotechnologies. Saha *et al.*[10], Buskirk *et al.*[11,12] and Rackham and Chin[18,19] highlight how rational design and evolutionary schemes can be used to increase the number and optimize the functionality of RNA molecular switches. These integrated design-evolution strategies may also be useful for constructing other, more complicated RNA systems, as well as synthetic gene networks that incorporate RNA components. In this regard, there is a need to develop selection strategies for evolving input-ouput responses[56] (**Fig. 7**) and dynamic functions, such as oscillatory behavior. Integrating libraries of RNA control modules with enhanced DNA synthesis[57] and recombination technologies[58] could also lead to synthetic genomes[59,60] that integrate novel RNA-based sensors, probes and regulators.

Synthetic biology is still establishing its identity. This emerging field has the potential to create new biological components and organisms, and improve the behavior of existing ones. Future benefits from synthetic biology, in addition to those mentioned above, may include new classes of therapeutics, novel proteins[61,62], engineered regulatory[2,6] and metabolic pathways[63], and the development of microbial fuel cells[64]. An immediate challenge for the field is the creation of an infrastructure, comprising modular and scalable functional components[65]. Versatile RNA molecules are well-suited to meet this challenge. Future advances in synthetic biology will likely involve a growing number of engineered networks that contain diverse sets of RNA molecules. The expanding catalog of designer RNAs could lead to novel modes of cellular control that are robust, scalable and applicable to a wide range of organisms.

ACKNOWLEDGMENTS

We thank Jason Chin for providing access to unpublished results in the form of a preprint manuscript and George Church for insightful and relevant discussions. We also thank Nick Reppas, Duhee Bang and the anonymous reviewers for providing valuable suggestions to improve the paper. This work was supported by the National Institutes of Health and the National Science Foundation.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Biotechnology* website for details).

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Hasty, J., McMillen, D., Isaacs, F. & Collins, J.J. Computational studies of gene regulatory networks: in numero molecular biology. *Nat. Rev. Genet.* **2**, 268–279 (2001).
2. Hasty, J., McMillen, D. & Collins, J.J. Engineered gene circuits. *Nature* **420**, 224–230 (2002).
3. Wall, M.E., Hlavacek, W.S. & Savageau, M.A. Design of gene circuits: lessons from bacteria. *Nat. Rev. Genet.* **5**, 34–42 (2004).
4. Benner, S.A. & Sismour, A.M. Synthetic biology. *Nat. Rev. Genet.* **6**, 533–543 (2005).
5. McDaniel, R. & Weiss, R. Advances in synthetic biology: on the path from prototypes to applications. *Curr. Opin. Biotechnol.* **16**, 476–483 (2005).
6. Sprinzak, D. & Elowitz, M.B. Reconstruction of genetic circuits. *Nature* **438**, 443–448 (2005).
7. Thompson, K.M., Syrett, H.A., Knudsen, S.M. & Ellington, A.D. Group I aptazymes as genetic regulatory switches. *BMC Biotechnol.* **2**, 21 (2002).
8. Werstuck, G. & Green, M.R. Controlling gene expression in living cells through small molecule-RNA interactions. *Science* **282**, 296–298 (1998).
9. Grate, D. & Wilson, C. Inducible regulation of the *S. cerevisiae* cell cycle mediated by an RNA aptamer-ligand complex. *Bioorg. Med. Chem.* **9**, 2565–2570 (2001).
10. Saha, S., Ansari, A.Z., Jarrell, K.A. & Ptashne, M. RNA sequences that work as transcriptional activating regions. *Nucleic Acids Res.* **31**, 1565–1570 (2003).
11. Buskirk, A.R., Kehayova, P.D., Landrigan, A. & Liu, D.R. *In vivo* evolution of an RNA-based transcriptional activator. *Chem. Biol.* **10**, 533–540 (2003).
12. Buskirk, A.R., Landrigan, A. & Liu, D.R. Engineering a ligand-dependent RNA transcriptional activator. *Chem. Biol.* **11**, 1157–1163 (2004).
13. Isaacs, F.J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* **22**, 841–847 (2004).
14. Suess, B., Fink, B., Berens, C., Stentz, R. & Hillen, W. A theophylline responsive riboswitch based on helix slipping controls gene expression *in vivo*. *Nucleic Acids Res.* **32**, 1610–1614 (2004).
15. Yen, L. *et al.* Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* **431**, 471–476 (2004).
16. Bayer, T.S. & Smolke, C.D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat. Biotechnol.* **23**, 337–343 (2005).
17. Desai, S.K. & Gallivan, J.P. Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.* **126**, 13247–13254 (2004).
18. Rackham, O. & Chin, J.W. A network of orthogonal ribosome-mRNA pairs. *Nat. Chem. Biol.* **1**, 159–166 (2005).
19. Rackham, O. & Chin, J.W. Cellular logic with orthogonal ribosomes. *J. Am. Chem. Soc.* **127**, 17584–17585 (2005).
20. Shine, J. & Dalgarno, L. Identical 3′-terminal octanucleotide sequence in 18S ribosomal ribonucleic acid from different eukaryotes. A proposed role for this sequence in the recognition of terminator codons. *Biochem. J.* **141**, 609–615 (1974).
21. Yusupova, G.Z., Yusupov, M.M., Cate, J.H. & Noller, H.F. The path of messenger RNA through the ribosome. *Cell* **106**, 233–241 (2001).
22. Wikstrom, P.M., Lind, L.K., Berg, D.E. & Bjork, G.R. Importance of mRNA folding and start codon accessibility in the expression of genes in a ribosomal protein operon of *Escherichia coli*. *J. Mol. Biol.* **224**, 949–966 (1992).
23. Lease, R.A. & Belfort, M. A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc. Natl. Acad. Sci. USA* **97**, 9919–9924 (2000).
24. Majdalani, N., Vanderpool, C.K. & Gottesman, S. Bacterial small RNA regulators. *Crit. Rev. Biochem. Mol. Biol.* **40**, 93–113 (2005).
25. Gottesman, S. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu. Rev. Microbiol.* **58**, 303–328 (2004).
26. Franch, T., Petersen, M., Wagner, E.G., Jacobsen, J.P. & Gerdes, K. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *J. Mol. Biol.* **294**, 1115–1125 (1999).
27. Majdalani, N., Hernandez, D. & Gottesman, S. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol. Microbiol.* **46**, 813–826 (2002).
28. Wagner, E.G. & Flardh, K. Antisense RNAs everywhere? *Trends Genet.* **18**, 223–226 (2002).
29. Good, L. Translation repression by antisense sequences. *Cell. Mol. Life Sci.* **60**, 854–861 (2003).
30. Ting, A.Y. *et al.* Phage-display evolution of tyrosine kinases with altered nucleotide specificity. *Biopolymers* **60**, 220–228 (2001).
31. Shi, Y. & Koh, J.T. Selective regulation of gene expression by an orthogonal estrogen receptor-ligand pair created by polar-group exchange. *Chem. Biol.* **8**, 501–510 (2001).
32. Pestova, T.V. *et al.* Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. USA* **98**, 7029–7036 (2001).

33. Stoneley, M. & Willis, A.E. Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. *Oncogene* **23**, 3200–3207 (2004).
34. Rasmussen, U.B., Mygind, B. & Nygaard, P. Purification and some properties of uracil phosphoribosyltransferase from *Escherichia coli* K12. *Biochim. Biophys. Acta* **881**, 268–275 (1986).
35. Winkler, W.C. & Breaker, R.R. Regulation of Bacterial Gene Expression by Riboswitches. *Annu. Rev. Microbiol.* **59**, 487–517 (2005).
36. Hermann, T. & Patel, D.J. Adaptive recognition by nucleic acid aptamers. *Science* **287**, 820–825 (2000).
37. Winkler, W., Nahvi, A. & Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956 (2002).
38. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A. & Breaker, R.R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281–286 (2004).
39. Rojas, A.A. *et al.* Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopoda cave crickets. *Nucleic Acids Res.* **28**, 4037–4043 (2000).
40. Ferbeyre, G., Smith, J.M. & Cedergren, R. Schistosome satellite DNA encodes active hammerhead ribozymes. *Mol. Cell. Biol.* **18**, 3880–3888 (1998).
41. Morcos, P.A. Achieving efficient delivery of morpholino oligos in cultured cells. *Genesis* **30**, 94–102 (2001).
42. Aszalos, A., Lemanski, P., Robison, R., Davis, S. & Berk, B. Identification of antibiotic 1037 as toyocamycin. *J. Antibiot. (Tokyo)* **19**, 285 (1966).
43. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
44. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
45. Isaacs, F.J., Hasty, J., Cantor, C.R. & Collins, J.J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl. Acad. Sci. USA* **100**, 7714–7719 (2003).
46. Kobayashi, H. *et al.* Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci. USA* **101**, 8414–8419 (2004).
47. Kramer, B.P. *et al.* An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.* **22**, 867–870 (2004).
48. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
49. Sengupta, D.J., Wickens, M. & Fields, S. Identification of RNAs that bind to a specific protein using the yeast three-hybrid system. *RNA* **5**, 596–601 (1999).
50. Bernstein, D.S., Buter, N., Stumpf, C. & Wickens, M. Analyzing mRNA-protein complexes using a yeast three-hybrid system. *Methods* **26**, 123–141 (2002).
51. Witherell, G.W., Gott, J.M. & Uhlenbeck, O.C. Specific interaction between RNA phage coat proteins and RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **40**, 185–220 (1991).
52. Stockley, P.G. *et al.* Probing sequence-specific RNA recognition by the bacteriophage MS2 coat protein. *Nucleic Acids Res.* **23**, 2512–2518 (1995).
53. Grate, D. & Wilson, C. Laser-mediated, site-specific inactivation of RNA transcripts. *Proc. Natl. Acad. Sci. USA* **96**, 6131–6136 (1999).
54. Sando, S., Narita, A., Abe, K. & Aoyama, Y. Doubly catalytic sensing of HIV-1-related CCR5 sequence in prokaryotic cell-free translation system using riboregulator-controlled luciferase activity. *J. Am. Chem. Soc.* **127**, 5300–5301 (2005).
55. Penchovsky, R. & Breaker, R.R. Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes. *Nat. Biotechnol.* **23**, 1424–1433 (2005).
56. Yokobayashi, Y., Weiss, R. & Arnold, F.H. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. USA* **99**, 16587–16591 (2002).
57. Tian, J. *et al.* Accurate multiplex gene synthesis from programmable DNA microchips. *Nature* **432**, 1050–1054 (2004).
58. Porteus, M.H. & Carroll, D. Gene targeting using zinc finger nucleases. *Nat. Biotechnol.* **23**, 967–973 (2005).
59. Smith, H.O., Hutchison, C.A., III, Pfannkoch, C. & Venter, J.C. Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci. USA* **100**, 15440–15445 (2003).
60. Chan, L.Y., Kosuri, S. & Endy, D. Refactoring bacteriophage T7. *Mol. Syst. Biol.*, Published online 13 September 2005 (doi:10.1038/msb4100025).
61. Looger, L.L., Dwyer, M.A., Smith, J.J. & Hellinga, H.W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185–190 (2003).
62. Dueber, J.E., Yeh, B.J., Chak, K. & Lim, W.A. Reprogramming control of an allosteric signaling switch through modular recombination. *Science* **301**, 1904–1908 (2003).
63. Martin, V.J., Pitera, D.J., Withers, S.T., Newman, J.D. & Keasling, J.D. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **21**, 796–802 (2003).
64. Rabaey, K. & Verstraete, W. Microbial fuel cells: novel biotechnology for energy generation. *Trends Biotechnol.* **23**, 291–298 (2005).
65. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
66. Eddy, S.R. Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2**, 919–929 (2001).
67. Kruger, K. *et al.* Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**, 147–157 (1982).
68. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857 (1983).
69. Doudna, J.A. & Cech, T.R. The chemical repertoire of natural ribozymes. *Nature* **418**, 222–228 (2002).
70. Wagner, E.G. & Simons, R.W. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* **48**, 713–742 (1994).
71. Masse, E., Escorcia, F.E. & Gottesman, S. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.* **17**, 2374–2383 (2003).
72. Nissen, P., Hansen, J., Ban, N., Moore, P.B. & Steitz, T.A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
73. Steitz, T.A. & Moore, P.B. RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends Biochem. Sci.* **28**, 411–418 (2003).
74. Weinger, J.S., Parnell, K.M., Dorner, S., Green, R. & Strobel, S.A. Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat. Struct. Mol. Biol.* **11**, 1101–1106 (2004).
75. Nahas, M.K. *et al.* Observation of internal cleavage and ligation reactions of a ribozyme. *Nat. Struct. Mol. Biol.* **11**, 1107–1113 (2004).
76. Hampel, A. & Cowan, J.A. A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage. *Chem. Biol.* **4**, 513–517 (1997).
77. Narlikar, G.J. & Herschlag, D. Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. *Annu. Rev. Biochem.* **66**, 19–59 (1997).
78. Winkler, W.C., Cohen-Chalamish, S. & Breaker, R.R. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. USA* **99**, 15908–15913 (2002).
79. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
80. Novina, C.D. & Sharp, P.A. The RNAi revolution. *Nature* **430**, 161–164 (2004).
81. Zamore, P.D. & Haley, B. Ribo-gnome: the big world of small RNAs. *Science* **309**, 1519–1524 (2005).
82. Lau, N.C., Lim, L.P., Weinstein, E.G. & Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862 (2001).
83. Lee, R.C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864 (2001).
84. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858 (2001).
85. Rhoades, M.W. *et al.* Prediction of plant microRNA targets. *Cell* **110**, 513–520 (2002).
86. Yekta, S., Shih, I.H. & Bartel, D.P. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594–596 (2004).
87. Ellington, A.D. & Szostak, J.W. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
88. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
89. Ellington, A.D. & Szostak, J.W. Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **355**, 850–852 (1992).
90. Breaker, R.R. In Vitro Selection of Catalytic Polynucleotides. *Chem. Rev.* **97**, 371–390 (1997).
91. Tang, J. & Breaker, R.R. Rational design of allosteric ribozymes. *Chem. Biol.* **4**, 453–459 (1997).
92. Unrau, P.J. & Bartel, D.P. RNA-catalysed nucleotide synthesis. *Nature* **395**, 260–263 (1998).
93. Araki, M., Okuno, Y., Hara, Y. & Sugiura, Y. Allosteric regulation of a ribozyme activity through ligand-induced conformational change. *Nucleic Acids Res.* **26**, 3379–3384 (1998).
94. Robertson, M.P. & Ellington, A.D. In vitro selection of an allosteric ribozyme that transduces analytes to amplicons. *Nat. Biotechnol.* **17**, 62–66 (1999).
95. Soukup, G.A. & Breaker, R.R. Engineering precision RNA molecular switches. *Proc. Natl. Acad. Sci. USA* **96**, 3584–3589 (1999).
96. Tang, J. & Breaker, R.R. Structural diversity of self-cleaving ribozymes. *Proc. Natl. Acad. Sci. USA* **97**, 5784–5789 (2000).
97. Jose, A.M., Soukup, G.A. & Breaker, R.R. Cooperative binding of effectors by an allosteric ribozyme. *Nucleic Acids Res.* **29**, 1631–1637 (2001).
98. Breaker, R.R. & Joyce, G.F.A. DNA enzyme that cleaves RNA. *Chem. Biol.* **1**, 223–229 (1994).
99. Breaker, R.R. DNA enzymes. *Nat. Biotechnol.* **15**, 427–431 (1997).
100. Silverman, S.K. In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Res.* **33**, 6151–6163 (2005).
101. Ruben, A.J. & Landweber, L.F. The past, present and future of molecular computing. *Nat. Rev. Mol. Cell Biol.* **1**, 69–72 (2000).
102. Seeman, N.C. DNA in a material world. *Nature* **421**, 427–431 (2003).
103. Tang, J. & Breaker, R.R. Examination of the catalytic fitness of the hammerhead ribozyme by in vitro selection. *RNA* **3**, 914–925 (1997).
104. Tang, J. & Breaker, R.R. Mechanism for allosteric inhibition of an ATP-sensitive ribozyme. *Nucleic Acids Res.* **26**, 4214–4221 (1998).
105. Stojanovic, M.N. & Stefanovic, D. A deoxyribozyme-based molecular automaton. *Nat. Biotechnol.* **21**, 1069–1074 (2003).
106. Benenson, Y. *et al.* Programmable and autonomous computing machine made of biomolecules. *Nature* **414**, 430–434 (2001).
107. Benenson, Y., Gil, B., Ben-Dor, U., Adar, R. & Shapiro, E. An autonomous molecular computer for logical control of gene expression. *Nature* **429**, 423–429 (2004).
108. Samarsky, D.A. *et al.* A small nucleolar RNA:ribozyme hybrid cleaves a nucleolar RNA target *in vivo* with near-perfect efficiency. *Proc. Natl. Acad. Sci. USA* **96**, 6609–6614 (1999).