

A modular and extensible RNA-based gene-regulatory platform for engineering cellular function

Maung Nyan Win and Christina D. Smolke*

Division of Chemistry and Chemical Engineering, 1200 East California Boulevard, MC 210-41, California Institute of Technology, Pasadena, CA 91125

Edited by Arthur D. Riggs, Beckman Research Institute, City of Hope, Duarte, CA, and approved July 12, 2007 (received for review May 1, 2007)

Engineered biological systems hold promise in addressing pressing human needs in chemical processing, energy production, materials construction, and maintenance and enhancement of human health and the environment. However, significant advancements in our ability to engineer biological systems have been limited by the foundational tools available for reporting on, responding to, and controlling intracellular components in living systems. Portable and scalable platforms are needed for the reliable construction of such communication and control systems across diverse organisms. We report an extensible RNA-based framework for engineering ligand-controlled gene-regulatory systems, called ribozyme switches, that exhibits tunable regulation, design modularity, and target specificity. These switch platforms contain a sensor domain, comprised of an aptamer sequence, and an actuator domain, comprised of a hammerhead ribozyme sequence. We examined two modes of standardized information transmission between these domains and demonstrate a mechanism that allows for the reliable and modular assembly of functioning synthetic RNA switches and regulation of ribozyme activity in response to various effectors. In addition to demonstrating examples of small molecule-responsive, *in vivo* functional, allosteric hammerhead ribozymes, this work describes a general approach for the construction of portable and scalable gene-regulatory systems. We demonstrate the versatility of the platform in implementing application-specific control systems for small molecule-mediated regulation of cell growth and noninvasive *in vivo* sensing of metabolite production.

aptamer | regulatory systems | ribozyme | RNA switches | synthetic biology

Basic and applied biological research and biotechnology are limited by our ability to get information into and from living systems and to act on information inside living systems (1–3). For example, there are only a small number of inducible promoter systems available to provide control over gene expression in response to exogenous molecules (4, 5). Many of the molecular inputs to these systems are not ideal for broad implementation, because they can be expensive and introduce undesired pleiotropic effects. In addition, broadly applicable methods for getting information from cells noninvasively have been limited to strategies that rely on protein and promoter fusions to fluorescent proteins, which enable researchers to monitor protein levels and localization and transcriptional outputs of networks, leaving a significant amount of the cellular information content currently inaccessible.

To address these challenges, scalable platforms are needed for reporting on, responding to, and controlling any intracellular component in a living system. A striking example of a biological communication and control system is the class of RNA-regulatory elements called riboswitches, comprised of distinct sensor and actuation (gene-regulatory) functions, that control gene expression in response to specific ligand concentrations (6). Building on these natural examples, engineered riboswitch elements have been developed for use as synthetic ligand-controlled gene-regulatory systems (7–10). However, these early examples of riboswitch engineering do not address the challenges posed above because they

lack portability across organisms and systems, and their designs and construction do not support modularity and component reuse.

We set out to develop a universal and extensible RNA-based platform that will provide a framework for the reliable design and construction of gene-regulatory systems that can control the expression of specific target genes in response to various effector molecules.[†] We implemented five engineering design principles (DPs) in addressing this challenge: DP1, scalability (a sensing platform enabling *de novo* generation of ligand-binding elements for implementation within the sensor domain); DP2, portability (a regulatory element that is independent of cell-specific machinery or regulatory mechanisms for implementation within the actuator domain); DP3, utility (a mechanism through which to modularly couple the control system to functional level components); DP4, composability (a mechanism by which to modularly couple the actuator and sensor domains without disrupting the activities of these individual elements); and DP5, reliability (a mechanism through which to standardize the transmission of information from the sensor domain to the actuator domain).

Results

Component Specification for a Scalable and Portable Gene-Regulatory System. To satisfy the engineering DP of scalability, DP1, we chose RNA aptamers (11), nucleic acid ligand-binding molecules, as the sensing platform for the universal control system. Our choice of sensing platform was driven by the proven versatility of RNA aptamers. Standard *in vitro* selection strategies or systematic evolution of ligands by exponential enrichment (SELEX) (12, 13) have been used to generate RNA aptamers *de novo* to a wide variety of ligands, including small molecules, peptides, and proteins (14). In addition, the specificity and affinity of an aptamer can be tuned through the selection process to meet the specific performance requirements of a given application. The continued selection of new aptamers to appropriate cellular molecules that function under *in vivo* conditions will enable these elements to be implemented as sensors in RNA-based control systems.

To satisfy the engineering DP of portability, DP2, we chose the hammerhead ribozyme, a catalytic RNA, as the regulatory element in the universal control system. Our choice of regulatory element was driven by the ability of the hammerhead ribozyme to exhibit self-cleavage activity across various organisms and its demonstrated potential in biomedical and biotechnological applications owing to its small size, relative ease of design, and rapid kinetics (15). The utility of hammerhead ribozymes as gene-regulatory elements has

Author contributions: M.N.W. and C.D.S. designed research; M.N.W. performed research; M.N.W. and C.D.S. analyzed data; and M.N.W. and C.D.S. wrote the paper.

Conflict of interest statement: The authors declare competing financial interests in the form of a pending patent application whose value may be affected by the publication of this manuscript.

This article is a PNAS Direct Submission.

Abbreviations: DP, design principle; sTRSV, satellite RNA of tobacco ringspot virus.

*To whom correspondence may be addressed. E-mail: smolke@cheme.caltech.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0703961104/DC1.

[†]A glossary of terms is available in [supporting information \(SI\) Text](#).

© 2007 by The National Academy of Sciences of the USA

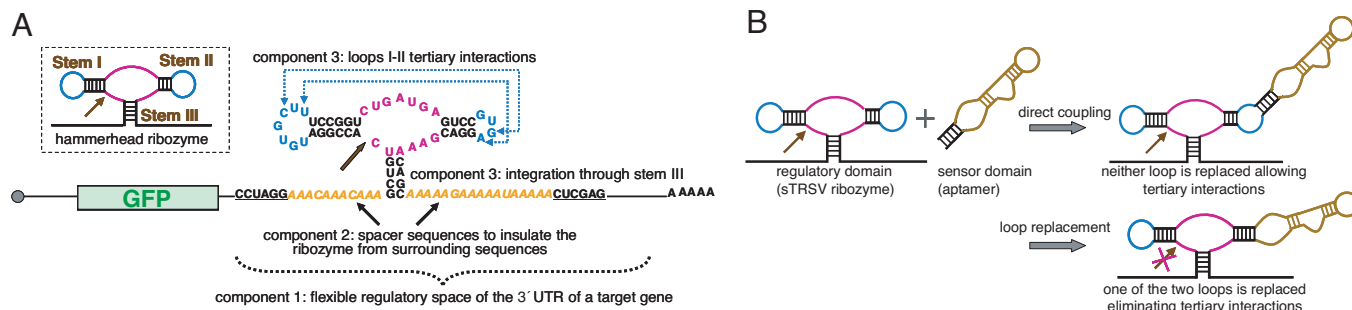


Fig. 1. General design strategy for engineering ribozyme switches. The color scheme is as follows: catalytic core, purple; aptamer sequences, brown; loop sequences, blue; spacer sequences, yellow; brown arrow, cleavage site. (A) General compositional framework and design strategy for engineering cis-acting hammerhead ribozyme-based regulatory systems. Restriction enzyme sites are underlined. (B) Modular coupling strategies of the sensor and regulatory domains to maintain *in vivo* activity of the individual domains.

been demonstrated in various systems (16–18). In addition, several research groups have engineered a special class of synthetic hammerhead ribozymes referred to as allosteric hammerhead ribozymes that contain separate catalytic and ligand-binding domains, which interact in a ligand-dependent manner to control the activity of the ribozyme (19–22). Although this class of ribozymes enables a better control system because of the presence of the integrated ligand-binding domain, there has been no success in translating them to *in vivo* environments.

Design Strategies for Engineering Portability, Utility, and Composability into a Biological Control System. To support a framework for engineering ligand-controlled gene-regulatory systems, we specified a design strategy that is in accordance with our engineering DPs, stated above (Fig. 1). This strategy is comprised of three components that address mechanisms for the portability, utility, and composability (DP2–DP4, respectively) of the control system and are critical to the development of a general ribozyme switch platform. First, the cis-acting hammerhead ribozyme constructs are integrated into the flexible regulatory space of the 3' UTR (Fig. 1A). We chose to locate the synthetic ribozymes within the 3' UTR of their target gene as opposed to the 5' UTR to isolate their specific cleavage effects on transcript levels from their nonspecific structural effects on translation initiation, because secondary structures have been demonstrated to repress efficient translation when placed in the 5' UTR (ref. 23 and K. Hawkins and C.D.S., unpublished observations). In addition, cleavage within the 3' UTR is a universal mechanism for transcript destabilization in eukaryotic and prokaryotic organisms. Second, each ribozyme construct is insulated from surrounding sequences, which may disrupt its structure and therefore its activity, by incorporating spacer sequences immediately 5' and 3' of stem III (Fig. 1A). By implementing these two components, we ensure that these control systems will be portable across organisms and modular to coupling with different coding regions (Y. Chen and C.D.S., unpublished work). The third component was necessitated by the fact that previously engineered *in vitro* allosteric ribozyme systems, which replace stem loops I or II with part of the aptamer domain (Fig. 1B, lower right), do not function *in vivo*. From previous studies on the satellite RNA of tobacco ringspot virus (sTRSV) hammerhead ribozyme (17), we suspect that this lack of *in vivo* functionality in earlier designs results from removal of stem loop sequences that may play a critical role in tertiary interactions that stabilize the catalytically active conformation under physiological Mg^{2+} concentrations. To develop ribozyme switches that function *in vivo*, we chose to integrate the hammerhead ribozyme into the target transcript through stem III and couple the sensor domain directly to the ribozyme through stem loops I or II to maintain these potentially essential sequence elements (Fig. 1B, upper right). Construction and characterization

of the regulatory activity of a series of ribozyme control constructs in the eukaryotic model organism *Saccharomyces cerevisiae* (Fig. 1A) indicate that maintenance of loop I and II sequences and thus integration through stem III are essential for their *in vivo* functionality (SI Text and SI Fig. 7).

Engineering Mechanisms for Information Transmission Between the Modular Switch Domains. The final design challenge in building a universal switch platform is to develop a standardized means of transmitting information (encoded within an information transmission domain) from the sensor (aptamer) domain to the regulatory (ribozyme) domain (i.e., DP5). There are two different strategies for transmitting information between the aptamer and ribozyme domains: strand displacement and helix slipping. We constructed and characterized ribozyme switch platforms based on both mechanisms.

The first information transmission domain that we developed is based on a strand-displacement mechanism, which involves the rational design of two sequences that compete for binding to a general transmission region (the base stem of the aptamer) (Fig. 2A and B). We used this mechanism in engineering a ribozyme switch platform that enables both up- and down-regulation of gene expression in response to increasing effector concentrations (“ON” and “OFF” switches, respectively). An initial ribozyme switch, L2bulge1, was constructed to up-regulate gene expression through the corresponding base platform, L2Theo (SI Fig. 7C), by incorporating a competing strand following the 3' end of the theophylline aptamer (24) (Fig. 2A). This competing strand is perfectly complementary to the base stem of the aptamer at the 5' end. Using the same DPs, we engineered another ribozyme switch, L2bulgeOff1 (Fig. 2B), for down-regulating gene expression. Our strand displacement strategy is based on the conformational dynamics characteristic of RNA molecules that enables them to distribute between at least two different conformations at equilibrium: one conformation in which the competing strand is not base-paired or is base-paired such that the ligand-binding pocket is not formed, and the other conformation in which the competing strand is base-paired with the aptamer base stem, displacing the switching strand and thus allowing the formation of the ligand-binding pocket. Strand displacement results in the disruption (L2bulge1) or restoration (L2bulgeOff1) of the ribozyme's catalytic core. Binding of theophylline to the latter conformation shifts the equilibrium distribution to favor the aptamer-bound form as a function of increasing theophylline concentration. An ≈ 25 -fold increase in target expression levels at 5 mM theophylline, relative to those in the absence of effector, was observed in L2bulge1 (Fig. 2C and SI Fig. 8). In contrast, an ≈ 18 -fold reduction in expression levels at 5 mM theophylline, relative to those in the absence of effector, was observed in L2bulgeOff1 (Fig. 2D and SI Fig. 8). Through our

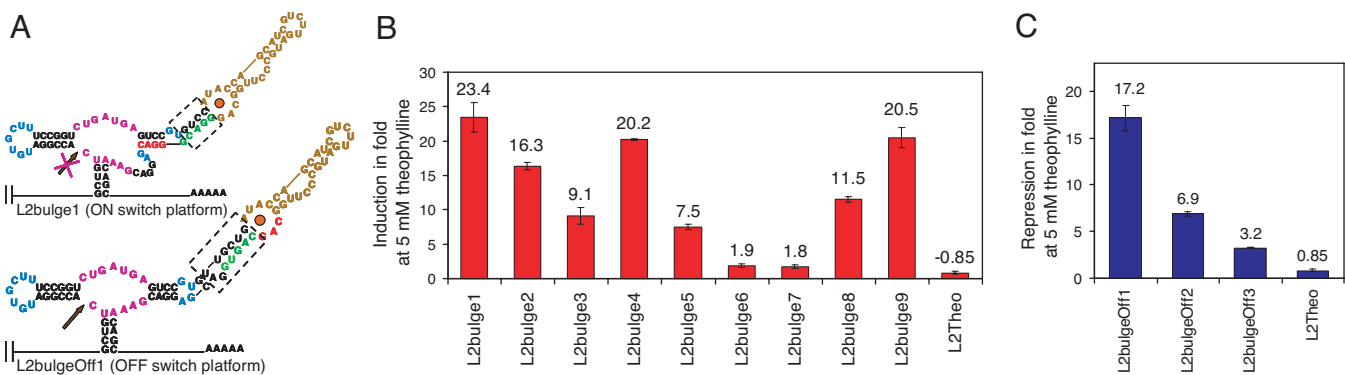


Fig. 4. Tunability of the strand-displacement-based ribozyme switches. (A) Sequences targeted by the rational tuning strategies are indicated in the dashed boxes on the effector-bound conformations of L2bulge1 (ribozyme-inactive) and L2bulgeOff1 (ribozyme-active). (B and C) Regulatory activities of tuned strand-displacement-based ON (B) and OFF (C) ribozyme switches. Gene-regulatory effects of these switches at 5 mM theophylline are reported in fold induction for ON switches and fold repression for OFF switches relative to the expression levels in the absence of theophylline as described in Fig. 2.

steady-state transcript levels in the absence and presence of effector are consistent with corresponding fluorescent protein levels (SI Table 1), indicating that cleavage in the 3' UTR results in rapid decay and inactivation of the target transcript. In addition, we demonstrated that changes in expression levels are induced shortly after effector addition (SI Fig. 11), indicating that the response of the regulatory elements to changes in effector levels is relatively rapid.

Rational Tuning Strategies Enable Programming of Switch-Regulatory Response.

The ability to program the regulatory response of a universal switch platform is an important property in tuning the platform performance to comply with the design specifications for a particular application. We demonstrate that our strand-displacement-based switch platform incorporates an information transmission mechanism that is amenable to rational tuning strategies for programming response properties. Programming of new regulatory information is achieved by sequence alteration, resulting in a change in the molecule's structural stability, which may affect its switching dynamics if the molecule can adopt multiple conformations. These rational sequence modification tuning strategies are not applicable to communication module-based switches because of an inability to predict their activities. A more complete description of our tuning strategies is provided in SI Text, SI Fig. 12, and SI Table 2. Briefly, our rational tuning strategies target alteration of the nucleotide composition of the base stem of the aptamer domain to affect the stabilities of individual constructs and the energies required for the construct to switch between two adoptable conformations. Using these strategies, we rationally engineered a series of tuned ON and OFF switches from L2bulge1 and L2bulgeOff1, respectively (Fig. 4A). These tuned switches exhibit different regulatory ranges in accordance with our rational energetic tuning strategies (Fig. 4B and C and SI Figs. 13 and 14).

The Ribozyme Switch Platform Exhibits Component Modularity and Specificity.

In implementing a standardized mechanism through which to transmit information between the domains of a switch platform (DP5), we needed to confirm that the modular coupling between the aptamer and ribozyme domains is maintained (DP4). We performed modularity studies on our strand-displacement-based ribozyme switch platform, in which aptamers possessing sequence flexibility in their base stems can be swapped into the sensor domain. To begin to demonstrate that ribozyme switch activity may be controlled by different effector molecules, we replaced the theophylline aptamer of L2bulge1 with a tetracycline miniaptamer (25) to construct a tetracycline-responsive ON switch (L2bulge1tc) (Fig. 5A). Despite similar aptamer ligand affinities

(24, 25), the extent of up-regulation with L2bulge1tc was greater than that with L2bulge1 at the same extracellular concentration of their respective ligands (Fig. 5B). This is likely because of the high cell permeability of tetracycline (26) compared with theophylline (27). These results demonstrate that our strand-displacement-based switch platform maintains modularity between the aptamer and ribozyme domains. We also performed similar modularity studies on the helix-slipping-based switch platform by replacing the theophylline aptamer of L1cm10, L2cm4, and L2cm5 with the tetracycline miniaptamer (L1cm10tc, L2cm4tc, and L2cm5tc, respectively). These constructs do not exhibit effector-mediated gene-regulatory effects (data not shown). We also demonstrated that the aptamer sequences (theophylline and tetracycline) incorporated into our ribozyme switch platforms maintain highly specific target recognition capabilities *in vivo*, similar to their *in vitro* specificities generated during the selection process against corresponding molecular analogues (caffeine and doxycycline, respectively) (24, 25) (Fig. 5B). This is an important property in implementing these platforms in cellular engineering applications that involve complex environments where molecular species similar to the target ligand may be present.

Component Modularity Enables Implementation of Ribozyme Switches as Regulatory Systems in Diverse Applications.

To demonstrate the scalability and utility of these switch platforms as application-specific control systems, we demonstrate the implementation of ribozyme switches in two distinct cellular engineering application areas. First, utility (DP3) and the ability to respond to and control cellular information are demonstrated by the application of ribozyme switches to small molecule-mediated regulation of cell growth. Second, scalability (DP1) and the ability to respond to and report on cellular information are demonstrated by the implementation of ribozyme switches as noninvasive *in vivo* sensors of metabolite production.

The first system explores the application of our ribozyme switches to the regulation of a survival gene, where modification of expression levels is expected to produce an observable and titratable phenotypic effect on cell growth. The reporter gene within the original constructs was replaced with a growth-associated gene, *his5*, responsible for the biosynthesis of histidine in yeast (28) (Fig. 6A). We performed growth-regulation assays across various effector concentrations using representative switch constructs and demonstrated that these switches mediate cell growth in a highly effector-dependent manner (Fig. 6B and SI Fig. 15). This application demonstrates the utility (DP3) of our switch platform, in which the control system exhibits modularity to the functional level components in the regulatory system.

feeding in correlation with HPLC data (Fig. 6D and SI Fig. 16), illustrating the noninvasive metabolite-sensing capabilities of these switches through transmitting changes in metabolite accumulation to changes in reporter expression levels. This application demonstrates the scalability (DP1) of our switch platform, in which the unique properties of the sensing platform used in this control system enable broad implementation in diverse applications not generally accessible by other regulatory systems.

Discussion

A key component in the development of an RNA-based framework for engineering ligand-controlled gene-regulatory systems is captured within DP5: a mechanism through which to reliably transmit information between distinct domains of the molecule. The strand-displacement and helix-slipping mechanisms demonstrate different strengths and weaknesses as standardized means of transmitting information from the aptamer domain to the ribozyme domain. Only 7 of the 26 tested communication modules exhibited regulatory activity in our system. In addition, all of the functional communication module sequences demonstrated OFF activity in our *in vivo* system, whereas one of these sequences (cmd) exhibited ON activity in an *in vitro* system (20). These results indicate that *in vitro* functionality of these elements is selectively translated to *in vivo* activity due to their sensitivity to surrounding sequences. Furthermore, modularity studies performed on this platform indicate that the helix-slipping mechanism is not amenable to modular domain-swapping strategies. In contrast, we have demonstrated that strand displacement exhibits greater reliability as an information transmission mechanism in our platform and is characterized by engineering properties such as modular assembly, rational *de novo* design, flexible induction and repression profiles, and response programmability. Although not preferred for the rational design strategies presented here, our helix-slipping platform can be used for the effective generation of new ribozyme switches by *in vivo* screening for helix-slipping elements that function with new aptamer sequences, different regulatory ranges, and flexible regulatory profiles. In addition, screening strategies may represent a powerful alternative when rational design strategies fail. For example, we were unable to successfully apply our rational design strategies to the construction of strand-displacement-based ribozyme switches that modulate cleavage through stem I (L1bulge1–L1bulge6 in SI Table 3). These results indicate that screening strategies may be more effective in generating switches that modulate activity through stem I.

We have developed and demonstrated universal RNA-based regulatory platforms called ribozyme switches by using engineering DPs. This work describes a framework for the reliable *de novo* construction of modular, portable, and scalable control systems that can be used to achieve flexible regulatory properties, such as up- and down-regulation of target expression levels and tuning of regulatory response to fit application-specific performance requirements, thereby expanding the utility of our platforms to a broader range of applications. For example, these switch platforms may be

applied to the construction of transgenic regulatory control systems that are responsive to cell-permeable, exogenous molecules of interest for a given network. In regulating sets of functional proteins, these switches can act to rewire information flow through cellular networks and reprogram cellular behavior in response to changes in the cellular environment. In regulating reporter proteins, ribozyme switches can serve as synthetic cellular sensors to monitor temporal and spatial fluctuations in the levels of diverse input molecules. The switch platforms described here represent powerful tools for constructing ligand-controlled gene-regulatory systems tailored to respond to specific effector molecules and enable regulation of target genes in various living systems; due to their general applicability, our platforms offer broad utility for applications in synthetic biology, biotechnology, and health and medicine.

Materials and Methods

Plasmid and Switch Construction. By using standard molecular biology techniques (30), a characterization plasmid, pRzS, harboring the yeast-enhanced GFP (yEGFP) (31) under control of a GAL1-10 promoter, was constructed. For the ribozyme switch-mediated growth studies, the *yegfp* gene was replaced with the *his5* gene (28). See SI Text for details.

RNA Secondary Structure Prediction and Free Energy Calculation. RNAstructure 4.2 (<http://rna.urmc.rochester.edu/rnastructure.html>) was used to predict the secondary structures of all switch constructs and their thermodynamic properties. RNA sequences that are predicted to adopt at least two stable equilibrium conformations (ribozyme-inactive and -active) were constructed and examined for functional activity.

Ribozyme Characterization, Cell Growth Regulation, and Metabolite Sensing Assays. See SI Text for details. Briefly, cells harboring appropriate plasmids were grown in the absence or presence of corresponding ligands or substrates and characterized for ligand-regulated ribozyme switch activity, cell growth, and metabolite sensing.

Fluorescence Quantification and Quantification of Cellular Transcript Levels. See SI Text for details. Briefly, total RNA was extracted by employing standard acid phenol methods (32) followed by cDNA synthesis and PCR amplification.

We thank K. Hawkins for assistance with controls and HPLC experiments and data analysis; A. Babiskin (California Institute of Technology) for pRzS and assistance with quantitative RT-PCR assays; K. Dusinger, J. Michener, and J. Liang for assistance with controls; E. Kelsic for assistance with image presentation; and Y. Chen and K. Hoff for critical reading of the manuscript. This work was supported by the Arnold and Mabel Beckman Foundation, the National Institutes of Health, and the Center for Biological Circuit Design at the California Institute of Technology (fellowship to M.N.W.).

- Endy D (2005) *Nature* 438:449–453.
- Voigt CA (2006) *Curr Opin Biotechnol* 17:548–557.
- Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ (2004) *Proc Natl Acad Sci USA* 101:8414–8419.
- Gossen M, Bujard H (1992) *Proc Natl Acad Sci USA* 89:5547–5551.
- Lutz R, Bujard H (1997) *Nucleic Acids Res* 25:1203–1210.
- Mandal M, Breaker RR (2004) *Nat Rev Mol Cell Biol* 5:451–463.
- Kim DS, Gusti V, Pillai SG, Gaur RK (2005) *RNA* 11:1667–1677.
- An CI, Trinh VB, Yokobayashi Y (2006) *RNA* 12:710–716.
- Bayer TS, Smolke CD (2005) *Nat Biotechnol* 23:337–343.
- Isaacs FJ, Dwyer DJ, Collins JJ (2006) *Nat Biotechnol* 24:545–554.
- Bunka DH, Stockley PG (2006) *Nat Rev Microbiol* 4:588–596.
- Tuerk C, Gold L (1990) *Science* 249:505–510.
- Ellington AD, Szostak JW (1990) *Nature* 346:818–822.
- Hermann T, Patel DJ (2000) *Science* 287:820–825.
- Birikh KR, Heaton PA, Eckstein F (1997) *Eur J Biochem* 245:1–16.
- Marschall P, Thomson JB, Eckstein F (1994) *Cell Mol Neurobiol* 14:523–538.
- Khvorova A, Lescoate A, Westhof E, Jayasena SD (2003) *Nat Struct Biol* 10:708–712.
- Yen L, Svendsen J, Lee JS, Gray JT, Magnier M, Baba T, D'Amato RJ, Mulligan RC (2004) *Nature* 431:471–476.
- Koizumi M, Soukup GA, Kerr JN, Breaker RR (1999) *Nat Struct Biol* 6:1062–1071.
- Soukup GA, Breaker RR (1999) *Proc Natl Acad Sci USA* 96:3584–3589.
- Soukup GA, Emilsson GA, Breaker RR (2000) *J Mol Biol* 298:623–632.
- Kertsburg A, Soukup GA (2002) *Nucleic Acids Res* 30:4599–4606.
- Pelletier J, Sonenberg N (1985) *Cell* 40:515–526.
- Jenison RD, Gill SC, Pardi A, Polisky B (1994) *Science* 263:1425–1429.
- Berens C, Thain A, Schroeder R (2001) *Bioorg Med Chem* 9:2549–2556.
- Hanson S, Berthelot K, Fink B, McCarthy JE, Sues B (2003) *Mol Microbiol* 49:1627–1637.
- Koch AL (1956) *J Biol Chem* 219:181–188.
- Nishiwaki K, Hayashi N, Irie S, Chung DH, Harashima S, Oshima Y (1987) *Mol Gen Genet* 208:159–167.
- Ogawa J, Takeda S, Xie SX, Hatanaka H, Ashikari T, Amachi T, Shimizu S (2001) *Appl Environ Microbiol* 67:1783–1787.
- Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY), 3rd Ed.
- Mateu C, Avery SV (2000) *Yeast* 16:1313–1323.
- Caponigro G, Muhrad D, Parker R (1993) *Mol Cell Biol* 13:5141–5148.