NAR Breakthrough Article

Automated physics-based design of synthetic riboswitches from diverse RNA aptamers

Amin Espah Borujeni[1], Dennis M. Mishler[2], Jingzhi Wang[2], Walker Huso[1] and Howard M. Salis[1,3,*]

[1]Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, USA, [2]Department of Chemistry, Emory University, Atlanta, GA 30322, USA and [3]Department of Biological Engineering, The Pennsylvania State University, University Park, PA 16802, USA

ABSTRACT

Riboswitches are shape-changing regulatory RNAs that bind chemicals and regulate gene expression, directly coupling sensing to cellular actuation. However, it remains unclear how their sequence controls the physics of riboswitch switching and activation, particularly when changing the ligand-binding aptamer domain. We report the development of a statistical thermodynamic model that predicts the sequence-structure-function relationship for translation-regulating riboswitches that activate gene expression, characterized inside cells and within cell-free transcription–translation assays. Using the model, we carried out automated computational design of 62 synthetic riboswitches that used six different RNA aptamers to sense diverse chemicals (theophylline, tetramethylrosamine, fluoride, dopamine, thyroxine, 2,4-dinitrotoluene) and activated gene expression by up to 383-fold. The model explains how aptamer structure, ligand affinity, switching free energy and macromolecular crowding collectively control riboswitch activation. Our model-based approach for engineering riboswitches quantitatively confirms several physical mechanisms governing ligand-induced RNA shape-change and enables the development of cell-free and bacterial sensors for diverse applications.

INTRODUCTION

Riboswitches are RNA-based sensors that use an aptamer domain to bind a chemical ligand, change shape and alter gene expression level; typically, by modulating translation initiation or transcriptional termination of mRNAs (1–3). While natural riboswitches have evolved as exquisite sensors, it remains difficult to engineer non-natural riboswitches that utilize different aptamers to detect and respond to ligands of interest, even though their application as biosensors or medical diagnostics would be transformative (4,5). To date, hundreds of aptamers have been discovered, or generated by SELEX, to bind specifically to cellular metabolites (cofactors, nucleotides, amino acids), chemicals of interest (antibiotics, explosives, pesticides, toxins) and proteins indicative of infection or disease (HIV envelope protein, human RUNX1 tumor suppressor) (6–11). However, most aptamers have not been converted into functional riboswitches as existing approaches have relied on qualitative design by experts, combinatorial library generation and high-throughput screening, which have limited the breadth of riboswitch-based applications (12–20). In particular, high-throughput screening for functional riboswitches cannot be carried out when using ligands that are cytotoxic, insoluble, highly reactive or impermeable to cell membranes. Successfully implementing these 'second and third steps' in a reliable riboswitch engineering pipeline have stymied the community for a decade (5).

Computational RNA design offers a scalable and repeatable approach to solve these challenges. For example, small RNAs have been rationally engineered to control gene expression through quantification of their RNA-RNA interactions and activation barriers (21–24). Computational design approaches have been developed for un-

derstanding the function of ligand-binding ribozymes (25), translation-regulating riboswitches (26) and termination-regulating riboswitches (27). Using computational design, specific theophylline-binding riboswitches have been engineered to regulate transcriptional termination (28) and translation initiation (26). Such RNA devices, including riboswitches, ribozymes and aptazymes, can be modeled and readily combined to tune protein expression levels, engineer signal-responsive genetic circuits and dynamically control metabolic pathways (29). However, a model has yet to be developed that can predict a riboswitch's function from its sequence, using in principle any RNA aptamer. Creating such a model from thermodynamic first principles would critically test our understanding of the physics of ligand-induced RNA conformational changes, while enabling the design and optimization of a toolbox of riboswitch-based RNA devices that can sense different ligands.

We therefore created a sequence-to-function physics-based model and automated optimization algorithm that converts RNA aptamers into synthetic translation-regulating riboswitches. To do this, we leveraged a recently developed statistical thermodynamic model of ribosome–RNA interactions that can accurately predict the translation initiation rates of mono-cistronic and multi-cistronic bacterial mRNAs (30–34). We validated our computational approach by characterizing 77 synthetic riboswitches with highly diverse RNA aptamers, sequences and structures, performing both *in vivo* and *in vitro* cell-free measurements to quantify their function. To illustrate the versatility of our approach, we selected RNA aptamers that bind specifically to both ideal and non-ideal ligands, including theophylline, tetramethylrosamine (TMR), fluoride, dopamine, thyroxine and 2,4-dinitrotoluene (DNT). Our automated approach generated functional riboswitches with high activation ratios (ARs), up to 383-fold. The validated model provides a comprehensive, quantitative and falsifiable physical mechanism for riboswitch function, including how changing the aptamer's structure and affinity, surrounding mRNA sequence, switching free energy, ligand and mRNA concentrations, and macromolecular crowding collectively control riboswitch activation.

MATERIALS AND METHODS

Strains and plasmids

All riboswitch and promoter sequences are listed in Supplementary Data. Theophylline, fluoride and DNT riboswitches were constructed and inserted into an mRFP1 fluorescent protein expression vector, derived from plasmid pFTV1 (ColE1 origin, Cm^R) (31). Three theophylline riboswitches (Theo-40, Theo-41 and Theo-45) also controlled the translation of a fusion mRFP1 protein. To create the fusion protein, four or five non-rare codons were introduced between the start codon and SacI restriction site within mRFP1 coding section. All the riboswitches were constructed using standard molecular cloning. Briefly, DNA fragments were computationally designed, synthesized and assembled using either annealing of oligonucleotides, polymerase chain reaction (PCR) assembly of oligonucleotides, or PCR amplification of gBLOCK DNA fragments (Integrated DNA Technologies). DNA fragments were then

digested by XbaI and SacI restriction enzymes, followed by ligation with digested plasmid, transformation, plating on selective media and verification of purified plasmid by sequencing. Similarly, promoter replacements were performed by annealing designed pairs of oligonucleotides, followed by digestion with BamHI/XbaI restriction enzymes, ligation, transformation, selective plating and verification by sequencing.

The promoters AEB-3, J23100 and LmrA were selected or designed to significantly vary riboswitch transcription rates (Supplementary Figure S17). The promoter AEB-3 is a result of mutating the -10 and -35 hexamers of promoter J23100, resulting in 10-fold lower transcription rate. The promoter LmrA is a near-consensus promoter with a 5-fold higher transcription rate. Unless noted otherwise in the text, all theophylline and fluoride riboswitches use the J23100 promoter, while all DNT riboswitches used the AEB-3 promoter.

Theophylline and TMR riboswitches were then constructed in plasmids expressing the luciferase reporter protein using standard molecular cloning. The plasmid is derived from the pBESTluc vector (Promega) initially using a pUC19 origin, where the riboswitch-reporter mRNA is transcribed by a Ptac promoter. As described in the text, plasmid origins were replaced with either pBAC, p15A or pFTV1 by PCR amplifying the expression cassette (promoter to transcriptional terminator) and digesting it with BamHI and SpeI, followed by ligation to the corresponding digested vectors, transformation, selective plating and verification by sequencing.

Dopamine and thyroxine riboswitches were constructed in pFTV1-derived plasmids containing the luciferase expression cassette, where pFTV1 was previously modified to insert AatII and HindIII restriction sites after the Ptac promoter and after the start codon, respectively. Riboswitch-encoding DNA fragments were PCR-amplified from designed gBLOCKs, followed by digestion with AatII and HindIII, ligation to digested plasmids, transformation, selective plating and verification by sequencing.

Ligand, media and buffer conditions

Theophylline, TMR, sodium fluoride, dopamine, thyroxine, 2,4-DNT, DMSO, ascorbic acid and sodium chloride were purchased from Sigma-Aldrich. Luria broth Miller (LB) media was purchased from VWR. All ligand conditions are provided in the Supplementary Data.

DNT has a solubility of 1.48 mM (270 μg/ml) in water. To increase DNT solubility, 1% (v/v) DMSO was added to the media when characterizing DNT riboswitches and no-aptamer controls using the 0 and 1 mM DNT conditions. Dopamine is enzymatically oxidized by *Escherichia coli* DH10B cells, and spontaneously oxidized by molecular oxygen. To minimize dopamine oxidation, 5 mM ascorbic acid was added to cell-free transcription–translation assays to characterize dopamine riboswitches and no-aptamer controls using the 0 and 1 mM dopamine conditions. Thyroxine has a solubility of 135 nM (0.105 μg/ml) in neutral water. To increase thyroxine solubility, 10 mM NaOH was added to create a 1 mM thyroxine working stock solution, followed by adding 2 μl of stock solution to 11 μl

cell-free transcription–translation assays (154 μM thyroxine). A total of 100 mM NaCl was then added to cell-free transcription–translation assays to maintain physiological RNA folding. Characterization of thyroxine riboswitches and no-aptamer controls took place in the same salt and pH conditions, using either 0 and 154 μM thyroxine.

Growth, fluorescence and luminescence measurements

All *in vivo* fluorescence measurements were carried out using the strain *E. coli* DH10B during long-time cultures that maintain cells in the exponential growth phase. Cells harboring cloned plasmids expressing the mRFP1 reporter were grown overnight in 700 μl LB media and 50 μg/ml Cm antibiotic in a 96 deep-well plate at 37°C and 200 rpm orbital shaking. Cultures were then diluted to an OD_{600} of 0.01 into 200 μl LB media, 50 ug/ml Cm and pre-defined concentrations of ligand and buffer solution (see Supplementary Data for all ligand concentrations). OD_{600} and fluorescence values were recorded by a M1000 spectrophotometer (TECAN) every 10 min until the OD_{600} reached about 0.15. Cells were then diluted to 0.01 into fresh media with the same composition, and grown until reaching mid-exponential phase. A third dilution was carried out in the same way. Time course fluorescence measurements per cell (FLPC) were analyzed by evaluating the equation:

$$FLPC = \frac{F - F_M}{OD_{600} - OD_{600,M}} - \frac{F_{NC} - F_M}{OD_{600,NC} - OD_{600,M}} \quad (1)$$

where F is the fluorescence of cells, F_M is the fluorescence of media, F_{NC} is the auto-fluorescence of non-transformed cells, OD_{600} is the optical density of cells and $OD_{600,M}$ is the optimal density of media. Supplementary Figure S8 shows examples of FLPC time series data.

Single-cell fluorescence levels were recorded by a Fortessa flow cytometer (BD Biosciences) and used to calculate ARs. During growth, and before each dilution, 10 μl samples were taken and added to 200 μl phosphate buffered saline with 2 mM kanamycin antibiotic for measurement. All distributions were unimodal. Auto-fluorescences of non-transformed cells were subtracted. Averages and standard deviations calculated from at least two independent measurements. The number of measurements for each riboswitch are listed in Supplementary Data.

The *in vivo* luminescence measurements for theophylline and TMR riboswitches were conducted by transforming *E. coli* DH10B using plasmids expressing the luciferase reporter and measuring luminescence in a Biotek Synergy HT plate reader, as described in Ref 14. Briefly, cells were diluted 1:500 in fresh LB/AMP from an overnight culture and grown to mid-exponential phase. Cultures were then prepared for measurement using dry ice and 1× CCLR (Promega) following the manufacturer's procedure. The luminescence data are normalized relative to OD_{600} values at the time of harvesting.

Cell-free transcription–translation assays and luminescence measurements

Cell-free transcription–translation assays were conducted using the S30 *E. coli* extract (Promega, L1020). TMR riboswitches were characterized by adding 20 ng/μl plasmid DNA to cell-free transcription–translation assays and incubating at 37°C for 30 min. Quenching was carried out by placing the reaction on ice and adding dilution buffer. To measure luciferase activity, equal volumes of cell-free product and luciferase reagent (Promega) were combined in black, opaque 96-well plates. Luminescence was recorded using a Biotek Synergy HT plate reader with a sensitivity of 150, and divided by amount of added plasmid. The same procedure was applied when adding increasing plasmid concentrations for the three TMR riboswitches. Dopamine and thyroxine riboswitches were characterized similarly, except using white, opaque 96-well plates, a 60-min incubation period and a SpectraMax M5 plate reader with an integration time of 1500 ms.

Transcription–translation decoupling experiment for TMR riboswitches

Decoupling experiments were conducted as previously described (14). Briefly, 40 ng/μl of DNA template was used, as for a normal cell-free reaction. Amino acids were omitted during the initial 5 min incubation. Rifampicin was then added at 250 μg/ml final concentration in the reaction to stop transcription. After a 1 min incubation, amino acids were added to start translation. A total of 40 μM TMR was added at time 0 or 6 min. The mixtures were incubated for up to 10 min after translation has started before quenching and measuring luminescence.

RESULTS

A biophysical model of translation-regulating riboswitches

The biophysical model considers a riboswitch as a long mRNA whose RNA–RNA, RNA–ribosome and RNA–ligand interactions control its translation initiation rate, and thereby the output protein's expression level (Figure 1A). As a single mRNA is transcribed, its folding trajectory is dictated by contact with ligand (Figure 1B). In the absence of ligand, the mRNA freely folds toward its minimum free energy structure (state **1**) and binds to the ribosome with free energy $\Delta G_{total,OFF}$. We calculate the ribosome's binding free energy to the mRNA in state **1** ($\Delta G_{total,OFF}$) using the RBS Calculator v2.0 model (31,32) (Figure 2A), which includes the energy needed to unfold inhibitory mRNA structures (ΔG_{mRNA}); the energy released when a canonical or non-canonical Shine-Dalgarno (SD) sequence binds to the last 9 nt of the 16S rRNA ($\Delta G_{mRNA-rRNA}$); an energetic penalty when the ribosome is stretched or compressed, due to a non-optimal distance between the SD and start codon ($\Delta G_{spacing}$); the energy released when the $tRNA^{fMet}$ binds to the start codon (ΔG_{start}); and an energetic penalty controlled by the single-stranded surface area of upstream standby sites ($\Delta G_{standby}$) (32). RNA–RNA interactions are calculated using established RNA free energy models that account for both secondary structures (35) and pseudoknots (36,37).

In the absence of ligand, the mRNA's translation initiation rate is predicted according to $r_{OFF} = \exp(-\beta \Delta G_{total,OFF})$ on a 100 000-fold proportional scale. When the ligand is available, a newly transcribed mRNA
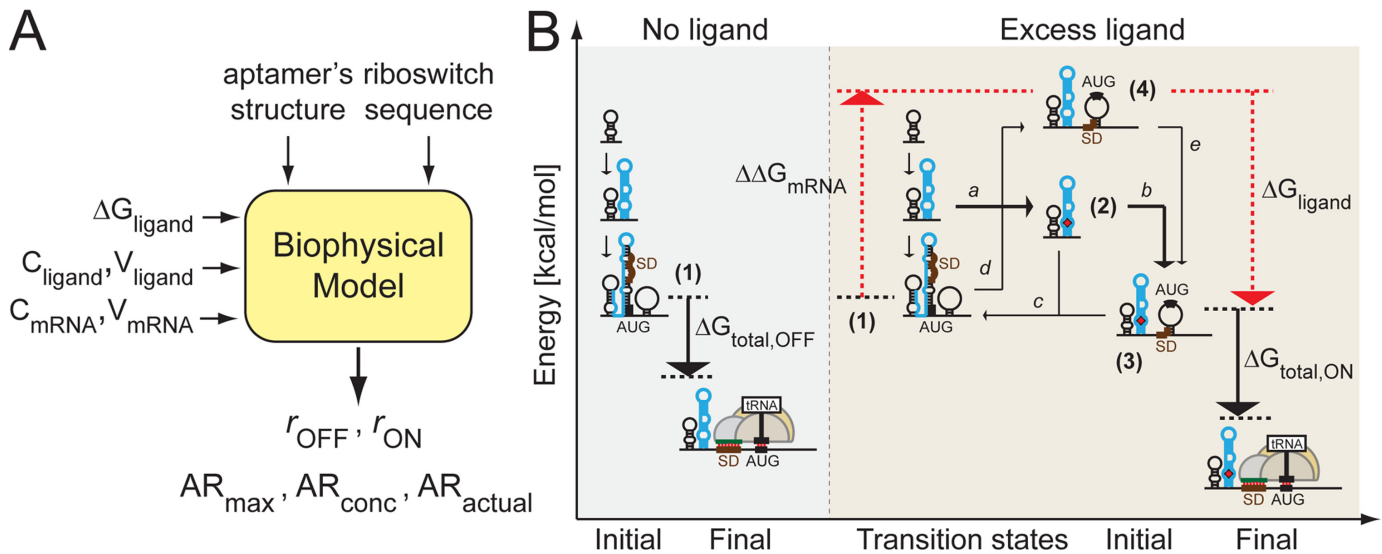
Figure 1. A biophysical model of riboswitch regulation. (**A**) A biophysical model was developed to predict the function of translation-regulating riboswitches. The model uses the riboswitch's mRNA sequence, the aptamer's ligand-bound structure, the concentrations of the mRNA and ligand, and their molar volumes to calculate the riboswitch's translation initiation rates in the ON and OFF states (r_{ON}, r_{OFF}) as well as its maximum, concentration-limited and actual activation ratios (AR_{max}, AR_{conc}, AR_{actual}). (**B**) A reaction coordinate diagram showing the states, energies, and transition paths during mRNA co-transcriptional folding, ligand-binding and translation initiation. Without ligand, the ribosome binds to a folded mRNA (state **1**) with free energy $\Delta G_{total,OFF}$. When excess ligand is present, the co-transcriptional folding of mRNA and ligand through paths a and b ends with a mRNA–ligand complex (state **3**) that binds the ribosome with free energy $\Delta G_{total,ON}$. The stability of the mRNA–ligand complex is controlled by the switching free energy ($\Delta\Delta G_{mRNA} + \Delta G_{ligand}$).

can rapidly bind to it via co-transcriptional folding (path *a*), reaching an intermediate structure that contains the ligand-bound aptamer (state **2**), and continuing to fold (path *b*) toward an alternate minimum free energy structure (state **3**) that now can bind more tightly to the ribosome with free energy $\Delta G_{total,ON}$ (Figure 1B). In the mRNA's state **3**, the RNA aptamer's structure is now locked into a ligand-bound state, while the remaining portions of mRNA are allowed to freely fold toward their minimum free energy structure (Figure 2B).

We then use the same free energy model to calculate the ribosome's binding free energy to mRNA's state **3**, $\Delta G_{total,ON}$ and its translation initiation rate $r_{ON} = \exp(-\beta\Delta G_{total,ON})$, where the differences between $\Delta G_{total,ON}$ and $\Delta G_{total,OFF}$ arise from ligand-induced changes in $\Delta G_{mRNA-rRNA}$, ΔG_{mRNA} and $\Delta G_{standby}$. As the model's first generalized sequence-to-function prediction, the maximum increase in a riboswitch's translation initiation rate (maximum activation ratio, AR_{max}) is determined by the difference in the ribosome's binding free energy, according to $AR_{max} = r_{ON}/r_{OFF} = \exp(-\beta [\Delta G_{total,ON} - \Delta G_{total,OFF}])$. The maximum possible activation is only obtained when co-transcriptional folding is present, when excess amounts of ligand are added and when the mRNA–ligand complex (state **3**) is stable.

Initial validation of the biophysical model

The theophylline RNA aptamer has been frequently utilized to engineer riboswitches, as theophylline binds tightly to its aptamer and can be added to excess without toxicity (3,26,29). We therefore initially tested the model's predictions for AR_{max} by characterizing 15 previously designed

theophylline-binding riboswitch variants that regulate luciferase expression (14) (Figure 2C), measuring their steady-state *in vivo* luminescences after adding either 0 or 2 mM theophylline. The calculated ribosome binding free energies, $\Delta G_{total,OFF}$ and $\Delta G_{total,ON}$, correctly predicted the changes in translation rate and luminescence ($R^2 = 0.68$, $P = 2 \times 10^{-8}$), according to the expected Boltzmann log-linear relationship (Figure 2D). The apparent slope ($\beta = 0.42 \pm 0.03$ mol/kcal) is similar to our previous translation rate measurements (31,32). The riboswitches' ARs varied from 1.9 to 326.6-fold and were well-predicted by the model's calculated AR_{max} ($R^2 = 0.68$, $P = 1.5 \times 10^{-4}$) (Figure 2E). These results suggested that the model's thermodynamic calculations can adequately predict riboswitch activation when a mRNA stably binds to a ligand, changes its structure and thereby alters the strengths of the ribosome–mRNA interactions controlling translation rate. We then expanded the biophysical model to predict riboswitch activation when the ligand concentration is limiting, when the mRNA–ligand complex is not stable, and when macromolecular crowding causes the amount of free volume to be limiting.

Riboswitch stability and concentration-dependence

We determine how changing ligand and mRNA concentrations affect riboswitch activation by calculating the ensemble partitioning of riboswitch mRNA into states **1** and **3** (shown in Figure 1B), within a water-filled, constant-volume system where the total amount of available volume is low to mimic the crowded environment found inside cells. The concentrations of mRNA in states **1** and **3** ($C_{mRNA,1}$ and $C_{mRNA,3}$) are solved by minimizing the system's total free energy, constrained by the energies of forma-
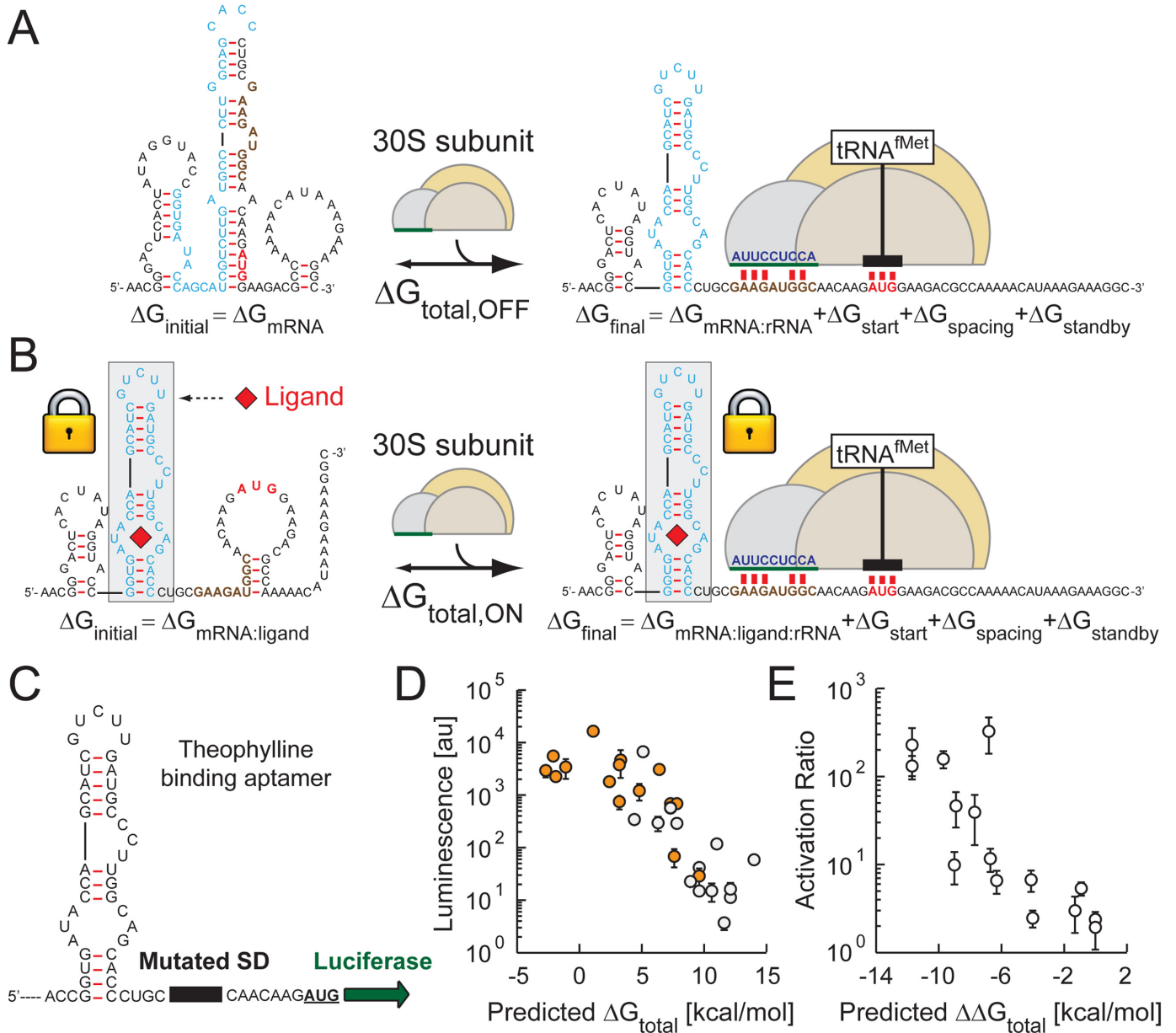
Figure 2. Biophysical model calculations. (**A**) Without ligand, when a ribosome binds to a folded mRNA, it undergoes several structural changes with corresponding free energy changes in the mRNA–ribosome interactions. Here, the ribosome binds poorly to the mRNA's initial state because it has an inaccessible standby site and sequestered (brown) Shine–Dalgarno (SD) sequence. In the final mRNA state, the ribosome's 16S rRNA and tRNA[fMet] hybridize to the SD sequence and start codon, respectively, with significant refolding of the mRNA. The model uses the mRNA sequence to calculate the total free energy change. (**B**) When the ligand is present, the same free energy calculation is performed, while constraining the (blue) aptamer sequence to its ligand-bound structure. (**C**) Fifteen synthetic theophylline riboswitches from Ref. (14) were characterized to test the model's AR_{max} prediction. (**D**) Predicted ΔG_{total} on states are well-correlated to the measured luminescences according to the expected log-linear relationship ($R^2 = 0.68$, $P = 2 \times 10^{-8}$). (**E**) The calculated free energy differences $\Delta\Delta G_{total}$ are well-correlated to the measured ARs according to the expected log-linear relationship ($R^2 = 0.68$, $P = 1.5 \times 10^{-4}$). Each data point and bar represents the mean and s.d. of three measurements.

tion for states **1** and **3**, the total mRNA ($C_{mRNA,total}$), and the total ligand concentration ($C_{ligand,total}$) as well as their molar volumes (V_{mRNA} and V_{ligand}) (Supplementary Figure S1 and Supplementary Information). At a specified total mRNA and ligand concentration, the riboswitch's AR will be $AR_{conc} = (C_{mRNA,1}/C_{mRNA,total}) + (C_{mRNA,3}/C_{mRNA,total})$ AR_{max}.

By converting concentrations to volume fractions and equivalently substituting ($C_{ligand,total} - C_{ligand,free}$) for $C_{mRNA,3}$, we rearranged this equation to explicitly show how the volume fractions and molar volumes of the mRNA and ligand control riboswitch activation (Supplementary Information):

$$AR_{conc} = \frac{v_{mRNA,1}}{v_{mRNA,total}} + \frac{V_{mRNA}}{V_{ligand}} \left[\frac{v_{ligand,total} - v_{ligand,free}}{v_{mRNA,total}} \right] AR_{max} \quad (2)$$

We next quantify the switching free energy to calculate a riboswitch's AR when its mRNA–ligand complex is not stable. We first consider an alternative structure (state **4**) that is not bound by ligand, but is otherwise structurally identical to state **3** (Figure 1B). The mRNA in state **4** is not stable because it has a higher free energy than the mRNA in state **1**. The difference in free energy between states 1 and 4 is always positive ($\Delta\Delta G_{mRNA} > 0$). Therefore, for the mRNA's ligand-bound state **3** to become thermodynamically favorable, the ligand must bind sufficiently tight to stabilize the mRNA–ligand complex; the switching free energy must be negative ($\Delta\Delta G_{mRNA} + \Delta G_{ligand} < 0$). However, if the switching free energy is positive, then the ligand could disassociate from the mRNA–ligand complex prior to its degradation (path *c*), leading to state **1**, and it will be unfavorable for the ligand to bind again without the presence of co-transcriptional folding (paths *d* and *e*). We refer to riboswitches with positive (negative) switching free energies as metastable (stable). Assuming equilibrium between mRNA states **1** and **3**, the fraction of mRNA that remains stably bound to the ligand can be readily derived (Supplementary Information), leading us to the following sequence-to-function expression for the riboswitch's actual activation ratio, AR$_{actual}$, whenever the switching free energy ($\Delta\Delta G_{mRNA} + \Delta G_{ligand}$) is positive:

$$AR_{actual} = 1 + \frac{\exp(-\beta(\Delta\Delta G_{mRNA} + \Delta G_{ligand}))}{1 + \exp(-\beta(\Delta\Delta G_{mRNA} + \Delta G_{ligand}))} (AR_{conc} - 1) \quad (3)$$

Altogether, the inputs into the biophysical model are the riboswitch's mRNA sequence, the structure of the RNA aptamer's ligand-bound state, the protein coding sequence, and the concentrations of the mRNA and ligand. There are no additional free or fit parameters. The model first calculates the translation initiation rates of a riboswitch in its ON and OFF states, r$_{ON}$ and r$_{OFF}$, according to the ribosome's binding free energies in the ligand-bound ($\Delta G_{total,ON}$) and ligand-free state ($\Delta G_{total,OFF}$). The model then uses Equation (2) to decrease r$_{ON}$ according to the specified mRNA and ligand concentrations. The riboswitches' switching free energy is then calculated and, if it's positive, Equation (3) is applied to further reduce r$_{ON}$ to account for metastability of the mRNA–ligand complex. The predicted AR is the ratio between r$_{ON}$ and r$_{OFF}$. The model assumes that the mRNA, ribosome and ligand reach chemical equilibrium faster than the mRNA's half-life, which requires two key conditions: first, the ligand must rapidly bind to the mRNA, for example, via co-transcriptional folding; and second, the mRNA must rapidly, but reversibly, change shape from its OFF to ON state after the ligand has bound. The model's accuracy also depends on the accuracy of the inputs; for example, the RNA aptamer's ligand-bound structure must be correctly specified. In addition, the model does not account for potential changes in mRNA stability, for example, when the translation rate of the mRNA is so low that it becomes a target for RNase activity.

Automated design of synthetic riboswitches

Using the complete model, we developed an automated optimization algorithm to computationally convert any RNA aptamer into a synthetic riboswitch. The inputs are the RNA aptamer's sequence, its ligand-bound structure and the output protein's coding sequence. The algorithm then identifies an optimal pre-aptamer sequence and post-aptamer sequence to create a non-natural riboswitch sequence (Figure 3A). Starting with a population of randomized pre- and post-aptamer sequences, the genetic algorithm performs rounds of *in silico* random mutation, prediction, selection and recombination to efficiently search for pre- and post-aptamer sequences that maximize a selected objective function, such as maximizing AR$_{actual}$, minimizing the switching free energy or selecting specific translation rates in the OFF and ON states (Supplementary Figure S2). The computational search space contains about 10^{36} sequences, which greatly surpasses the library generation and measurement capacities of high-throughput approaches (Supplementary Figure S3).

We then employed automated computational optimization to design 52 synthetic riboswitches using three RNA aptamers with the objective of testing the model's ability to correctly predict how different pre- and post-aptamer sequences controlled riboswitch function. Specifically, we selected the theophylline TCT8–4 [38], TMR [39] and fluoride 78-Psy(6) aptamers (Figure 3B), because their experimentally measured ligand binding energies, ligand sizes and ligand toxicities vary considerably (ΔG_{ligand} = -9, -10.2 and -6 kcal/mol; V_{ligand}/V_{water} = 10, 20 and 1; $C_{toxicity}$ = 5 mM, 20 μM and 150 mM for theophylline, TMR and fluoride, respectively). Our choice of RNA aptamers provided appropriate test cases for sensing ligands at low concentrations (TMR) and sensing ligands that bind weakly to their RNA aptamer (fluoride) together with a commonly engineered RNA aptamer that serves as a baseline for comparisons (theophylline). The fluoride RNA aptamer also has a highly pseudoknotted tertiary structure [40], creating a more complex structural constraint. Initially, our optimization algorithm's objective function was tailored to design both stable and metastable riboswitches by varying the targeted switching free energies from -8.4 to 11.52 kcal/mol. As a result, 37 of the designed riboswitches have positive switching free energies (Supplementary Figure S4). The designed riboswitches have dissimilar pre-aptamer and post-aptamer sequences, which do not resemble any natural sequences or structures (sequence entropy 63%, Supplementary Figure S5). All riboswitch structures were also predicted to change significantly when the ligand is bound (examples shown in Supplementary Figures S6 and S7). Three of the previously designed theophylline riboswitches also have positive switching free energies.

Validating model-predicted requirements for riboswitch activation

To test the model's predictions, we measured reporter expression levels in the riboswitches' ON and OFF states by using spectrophotometry and flow cytometry to record mRFP1 fluorescence and luciferase luminescence levels, with and without added ligand, inside *E. coli* DH10B cells maintained in exponential growth for 12 h ('Materials and Methods' section and Supplementary Figure S8 for time series data). Either 2 mM theophylline, 20 μM TMR or 150 mM fluoride were added throughout the long-
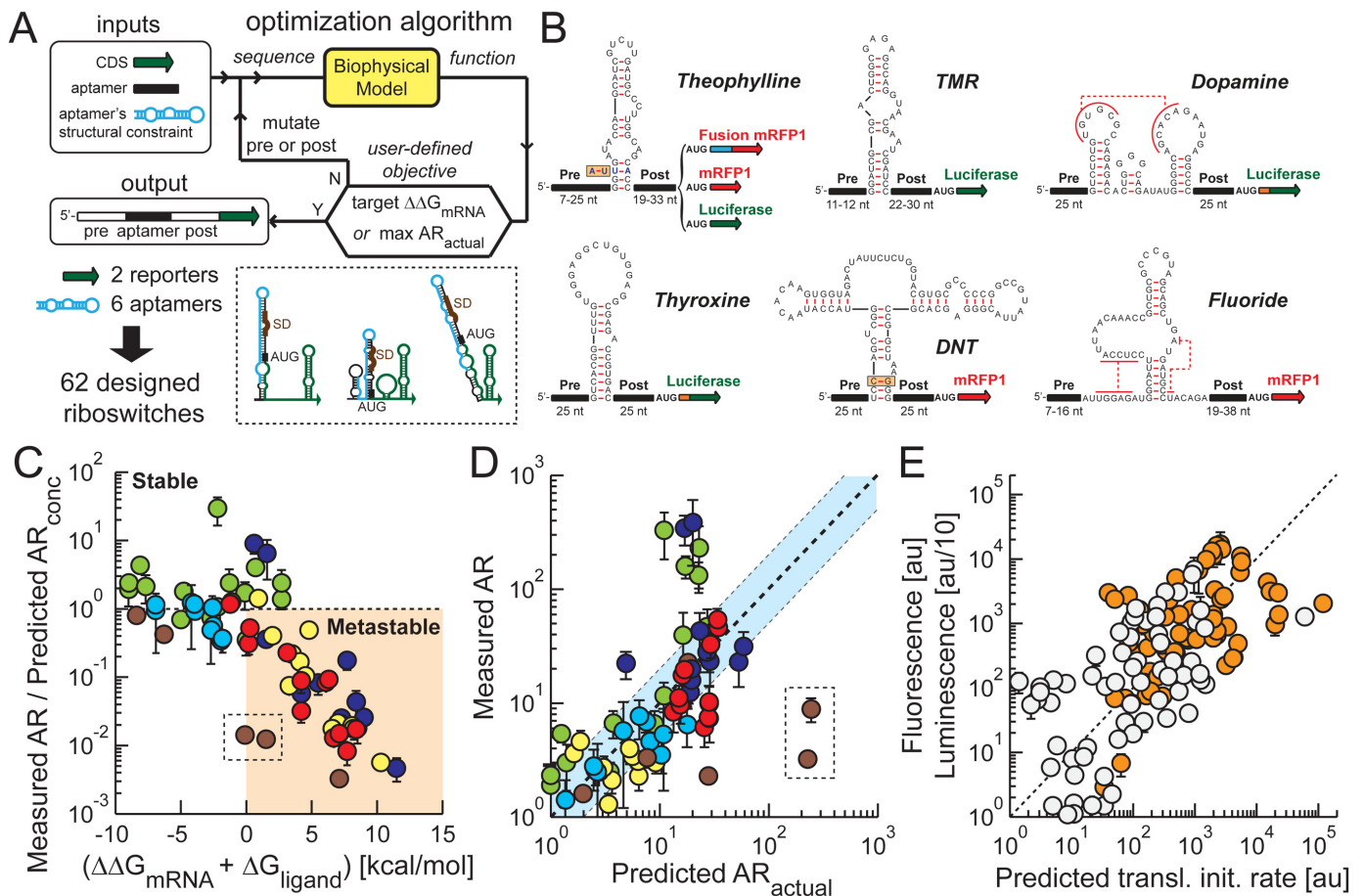
Figure 3. Automated design of synthetic riboswitches. (**A**) An optimization algorithm converts RNA aptamers into synthetic riboswitch sequences through rounds of mutation, prediction, selection and recombination. (**B**) The ligand-bound structures of the theophylline, TMR, fluoride, dopamine (dopa1.3/c.3), thyroxine (ApT4-J-min) and 2,4-dinitrotoluene aptamers are shown (see also Supplementary Figure S16). Pseudoknotted base pairs are indicated by dotted lines. Pre- and post-aptamer sequences varied after optimization. (**C**) The error in the model's calculated AR_{conc} is compared to the calculated switching free energy, showing that the result of Equation 2 loses accuracy as the switching free energy grows. Riboswitches are colored according to their aptamer: (dark blue) 12 designed theophylline riboswitches; (red) 12 theophylline riboswitches with mutated aptamer; (brown) 6 theophylline riboswitches with different mRFP1 coding sequences; (green) 15 previously engineered theophylline riboswitches (14); (yellow) 12 designed fluoride riboswitches; (light blue) 10 TMR riboswitches. The boxed outliers are riboswitches Theo-42 and Theo-44, suggesting a systematic malfunction in their function. (**D**) Model predictions using Equation 3 (AR_{actual}) are in good agreement with the measured ARs for 59–67 riboswitches (Spearman $R = 0.69$, $P = 1.2 \times 10^{-10}$, $N = 67$; Pearson $R^2 = 0.61$, $P = 2.6 \times 10^{-13}$, $N = 59$). (Shaded blue) The model predicted the ARs of 37 riboswitches to within 2-fold. (**E**) The predicted translation initiation rates in the riboswitches' (white circles) OFF and (orange circle) ON states are compared to their measured fluorescence or luminescence levels (Spearman $R = 0.68$, $P = 3.7 \times 10^{-19}$, $N = 134$; Pearson $R^2 = 0.44$, $P = 4.3 \times 10^{-18}$, $N = 134$). Each data point and bar represents the mean and s.d. of 2–4 measurements.

time cultures to measure the riboswitches' ON state expression levels. We performed the same measurements on a set of no-aptamer controls to measure the ligands' non-specific effects on reporter expression levels under identical *E. coli* growth conditions. For example, 2 mM theophylline and 150 mM fluoride both increased the no-aptamer control's reporter expression by 1.4- and 1.3-fold, while 20 μM TMR decreased it by 2.6-fold. We removed these non-specific effects by dividing the riboswitch's measured ON/OFF expression ratio by the no-aptamer control's measured ON/OFF expression ratio, measured under identical growth and ligand conditions. The resulting measurement is the riboswitch's specific AR. Altogether, we validated the model's predictions by performing these measurements on 67 riboswitches, including the 15 previously designed theophylline riboswitches. All sequences, model calculations, ri-

boswitch measurements and no-aptamer control measurements are provided in Supplementary Data.

We first evaluated the original biophysical model that only accounts for the ligand-induced changes in mRNA–ribosome interactions, using our formula for AR_{max}. We found that a riboswitch must have a high model-calculated AR_{max} for it to successfully activate translation in response to ligand binding. For example, five of our characterized riboswitches have $AR_{max} < 3.0$; when they bind theophylline, all five activate translation by 3-fold or less. However, by itself, the model-calculated AR_{max} is not a quantitative predictor of a riboswitch's AR as quantified by a log–log comparison across the 67 riboswitches (Pearson $R^2 = 0.002$, $P = 0.71$, $N = 67$) (Supplementary Figure S9). Designing a riboswitch with a high model-calculated AR_{max} is necessary, but not sufficient, for achieving high activation.

Instead, we found that incorporating concentration-dependence into the biophysical model enabled it to accurately predict riboswitch activation, but only for stable riboswitches (Figure 3C). Using the formula for AR_{conc} (Equation 2), 25 out of the 27 total riboswitches with negative switching free energies had well-predicted ARs with two outliers (Supplementary Figure S10A) (Spearman $R = 0.72$, $P = 2.5 \times 10^{-5}$, $N = 27$; Pearson $R^2 = 0.62$, $P = 3.5 \times 10^{-6}$, $N = 25$). The ARs of all 10 TMR riboswitches were well-predicted, due to their high-affinity RNA aptamers and modestly high structural free energy changes ($\Delta\Delta G_{mRNA}$ <8.4 kcal/mol) (Supplementary Figure S10A). In contrast, without accounting for the instability of the mRNA–ligand complex, the model-calculated AR_{conc} could not predict the ARs of the remaining 40 metastable riboswitches, which included 28 theophylline riboswitches that undergo energetically costly switching ($\Delta\Delta G_{mRNA}$ > 9 kcal/mol) and all 12 low-affinity fluoride-binding riboswitches (Spearman $R = 0.27$, $P = 0.09$, $N = 40$; Pearson $R^2 = 0.04$, $P = 0.22$, $N = 40$) (Supplementary Figure S10B).

Finally, when we account for the instability of the mRNA–ligand complex, the biophysical model could predict the ARs of both stable and metastable riboswitches (Figure 3D). Using the formula for AR_{actual} (Equation 3), the calculated ARs for 59 out of the 67 riboswitches agreed with their measured ARs across a 105-fold range (Spearman $R = 0.69$, $P = 1.2 \times 10^{-10}$, $N = 67$; Pearson $R^2 = 0.61$, $P = 2.6 \times 10^{-13}$, $N = 59$). The eight outliers were theophylline riboswitches, and five of them had higher-than-predicted ARs. The model errors in AR_{actual}, r_{ON} and r_{OFF} were log-normally distributed and 55% of all riboswitches had their ARs correctly predicted to within 2-fold (Supplementary Figure S11). Moreover, the model's accuracy did not depend on the RNA aptamer employed to engineer the riboswitch; the apparent error in the model's free energy calculations, quantified by the metric |log(AR/AR_{actual})|, was 0.64, 0.41 and 0.59 for the theophylline, TMR and fluoride riboswitches, respectively.

Interestingly, the riboswitches' ARs were more accurately calculated than their translation rates in their ON and OFF states (Spearman $R = 0.68$, 3.7×10^{-19}, $N = 134$; Pearson $R^2 = 0.44$, $P = 4.3 \times 10^{-18}$, $N = 134$) (Figure 3E and Supplementary Figure S12) as the error in predicting r_{ON} and r_{OFF} for the same riboswitch had a high degree of correlation (Spearman $R = 0.78$, $P < 10^{-30}$, $N = 67$; Pearson $R^2 = 0.87$, $P = 3.6 \times 10^{-27}$, $N = 59$). This shared uncertainty factor was eliminated when calculating the r_{ON}/r_{OFF} ratio (Supplementary Figure S13). This source of error could arise from any interaction, not included within our model, that affects riboswitch expression similarly in both ON and OFF states. By analyzing the physical interactions distinguishing the riboswitch inliers from outliers, we also found that riboswitches with lower-than-predicted ARs bound much more tightly to the ribosome's 16S rRNA (highly negative $\Delta G_{SD-antiSD}$) and had a high energetic cost to switching (high $\Delta\Delta G_{mRNA}$) (Supplementary Figure S14), which could be responsible for their higher-than-predicted r_{OFF} and lower-than-predicted r_{ON}, respectively.

Overall, by developing a sequence-to-function biophysical model for riboswitches and coupling it to an automated optimization algorithm, we successfully converted the theo-phylline, TMR and fluoride aptamers into highly active riboswitches, while performing a small number of measurements (Figure 4A, Supplementary Figure S15). The model explains why some RNA aptamers are more conducive to riboswitch engineering, compared to others; for example, the highest possible AR_{actual} for a fluoride riboswitch is 22-fold lower than for a theophylline riboswitch, due to their differences in ligand affinity and ligand-bound structure. Further, because TMR is toxic and limits its maximum concentration, the highest possible AR_{actual} for a TMR riboswitch is 62-fold lower than for a theophylline riboswitch.

Design of dopamine, thyroxine and DNT-binding riboswitches

We next tested whether our computational design approach could convert uncharacterized RNA aptamers that bind non-ideal ligands into functional riboswitches. As a demonstration, we selected the dopamine (dopa2/C.1 and dopa1.3/C.3) (41), thyroxine (ApT4-J-min) (42) and 2,4-DNT (17,43) aptamers and ligand targets (Figure 3B and Supplementary Figure S16) for their potential to diagnose disease and detect explosives. Dopamine is rapidly oxidized by bacterial membrane-bound monoamide oxidases (44) and thyroxine is highly insoluble in water, motivating the characterization of such riboswitches in cell-free transcription–translation assays. A DNT-binding riboswitch with 10-fold apparent activation was found by elegantly screening a 10^{18} sequence library (17). However, there are no published reports of the dopamine-binding or thyroxine-binding RNA aptamers being utilized in any aptazyme or riboswitch.

Accordingly, we applied our biophysical model and automated optimization algorithm to design five dopamine, two thyroxine and three DNT riboswitches. We measured the ARs of the dopamine and thyroxine riboswitches using cell-free transcription–translation assays and characterized the activation of the DNT riboswitches inside *E. coli* DH10B cells ('Materials and Methods' section). We also measured the non-specific effects of these ligands by recording the no-aptamer control's reporter expression levels under identical conditions; 1 mM DNT increased the no-aptamer control's expression by 4.8-fold, while 1 mM dopamine and 150 μM thyroxine decreased it by 1.67- and 2.5-fold, respectively. As before, the ligands' non-specific effects on gene expression were removed from our definition of the riboswitches' measured ARs. The best DNT-binding riboswitch achieved an AR of 11.1-fold. However, if the non-specific effects of DNT on gene expression were not removed from the apparent AR, then the DNT-binding riboswitch's apparent AR would be 53.3-fold. The best thyroxine- and dopamine-binding riboswitches achieved ARs of 2.4- and 2.0-fold, respectively, making these the first riboswitches that can regulate gene expression using these RNA aptamers (Figure 4A and Supplementary Figure S15). In contrast to high-throughput screening, these prototypical examples demonstrate how physics-based design can engineer functional riboswitches even when the ligand is reactive, insoluble or impermeable to cell membranes.

Notably, the ligands' pleiotropic effects on *in vitro* and *in vivo* gene expression are particularly important because,
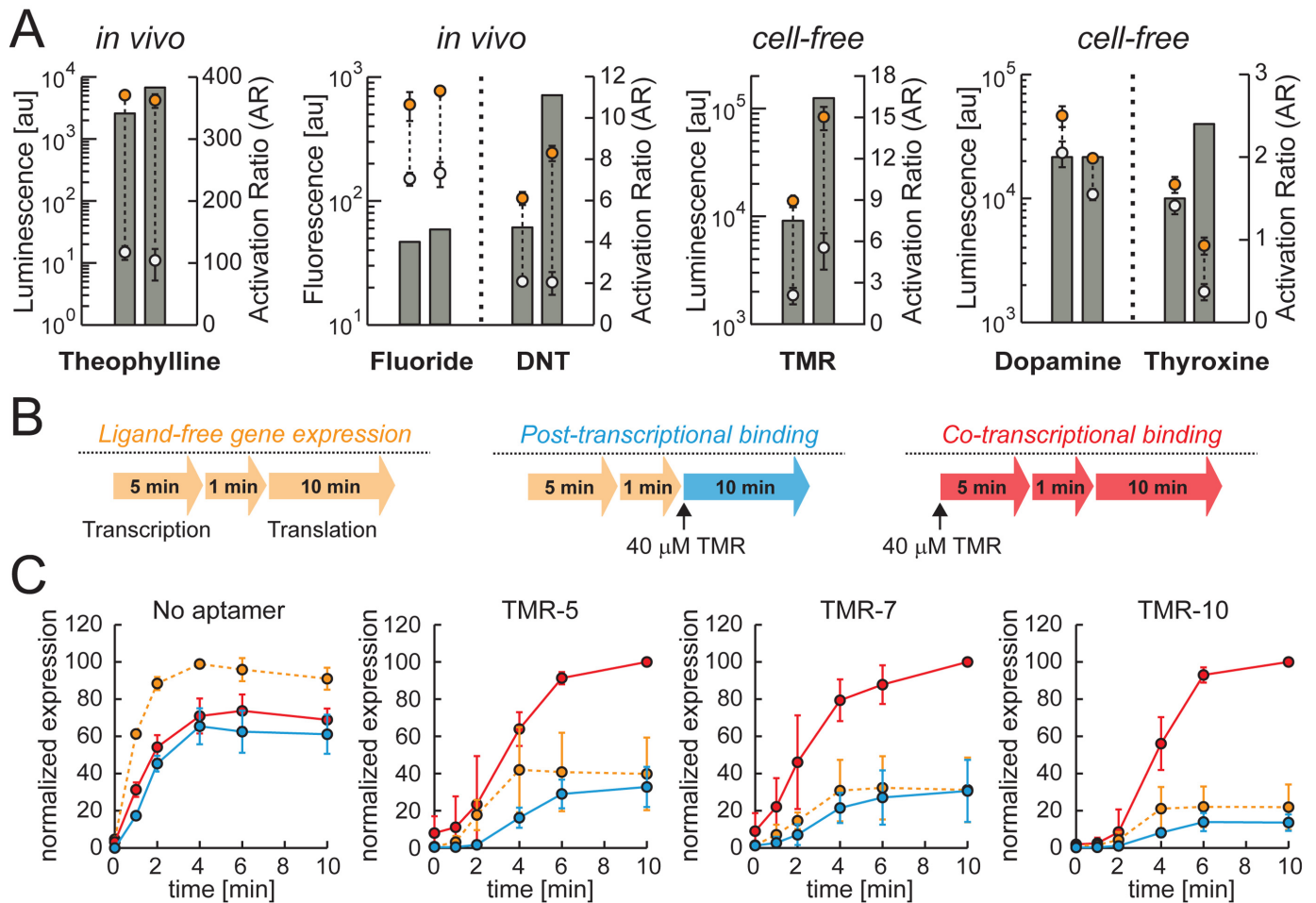
Figure 4. *In vivo* and cell-free characterization of designed riboswitches. (**A**) For each aptamer, the (circles) measured expression levels in the ON and OFF states and (bars) the ARs for the two best riboswitch variants are shown. To characterize their ON states, added ligand concentrations were 2 mM theophylline, 30 μM TMR, 150 mM fluoride, 1 mM DNT, 1 mM dopamine supplemented with 5 mM ascorbic acid and 150 μM thyroxine supplemented with 100 mM NaCl and 1.5 mM NaOH (see Supplementary Figure S15). Each data point and bar represents the mean and standard deviation of 2–4 measurements. (**B**) Cell-free transcription–translation assays are used to characterize riboswitch function during (yellow) the absence of ligand, (blue) post-transcriptional ligand binding and (red) co-transcriptional ligand binding. (**C**) Using cell-free transcription–translation assays, the expression levels of three TMR riboswitches and a no-aptamer control were measured to determine the role of co-transcriptional folding. As described in Ref. (14), expression levels are normalized to the highest luminescence level during the 10 min assay for each individual construct. Each data point and bar represents the mean and s.d. of 3–4 measurements.

depending on the sensing application, they will increase or decrease the apparent dynamic range of regulation independent of riboswitch function. These pleiotropic effects do depend on the environmental factors beyond the ligand itself, including additives to the buffer, temperature and pH changes, and salt concentration. Here, to validate the biophysical model's predictions and to elucidate riboswitch-specific mechanisms, we focused on measuring riboswitch-specific ARs. However, when incorporating riboswitch assays into sensing devices to measure the concentration of a ligand-of-interest, it will be equally important to carry out the same measurements using no-aptamer controls and to use those measurements for calibration.

Co-transcriptional folding is required for riboswitch activation

With additional use of cell-free transcription–translation assays, we next applied our model to evaluate how co-transcriptional folding and macromolecular crowding affects riboswitch function inside cells. First, we characterized the activation of three TMR riboswitches in cell-free assays either without ligand, without co-transcriptional ligand binding or with co-transcriptional ligand binding (Figure 4B and 'Materials and Methods' section). When ligand was added after transcription had been stopped, the riboswitches were inactive and behave similarly as the no-aptamer control (Figure 4C). The riboswitches only activated reporter expression when ligand was added during transcription. With co-transcriptional ligand-binding, all 10 TMR riboswitches activated similarly in both cell-free and *in vivo* environments (Supplementary Figure S15EF).

These results show that co-transcriptional folding is required for reaching a fast mRNA–ligand equilibrium, a key model assumption, which occurs either inside cells or within cell-free transcription–translation assays. Our results are consistent with our earlier work showing that co-transcriptional folding is essential even when ligand is present in excess (14).

Macromolecular crowding controls riboswitch activation

An intriguing and counter-intuitive model prediction is that higher riboswitch mRNA levels will actually decrease riboswitch ARs inside cells, but not within a relatively dilute cell-free assay. Inside cells, high concentrations of native RNA and protein occupy a significant portion of the intracellular volume, called macromolecular crowding (45,46). As a result of volume exclusion, there is significantly less free volume available to a bulky mRNA. According to the model, when more mRNA is expressed inside the cell, the amount of free volume will further decrease and could eventually lead to the exclusion of the ligand from the same volumetric space as the mRNA with a correspondingly lower amount of mRNA–ligand complex. We calculate the free mRNA, free ligand and ligand-bound mRNA volume fractions by minimizing the thermodynamic free energy of the constant volume system, constrained by the total mRNA and ligand concentrations and their molar volumes (Supplementary Information). In Figure 5A, we show how these volume fractions vary as more riboswitch mRNA is expressed. Using these volume fractions and Equation 2, the model qualitatively predicts that expressing more riboswitch mRNA will lead to higher translation rates in both the ON and OFF states (higher r_{ON} and r_{OFF}), however the riboswitch's AR will actually decrease because r_{ON} increases less than r_{OFF}. In contrast, in a dilute system, there is enough free volume that adding more riboswitch mRNA should not exclude ligand from the system. Riboswitch activation will increase if ligand-binding follows second order kinetics.

We tested these predictions by adding increasing amounts of riboswitch DNA template to cell-free transcription–translation assay, and by increasing riboswitch mRNA levels inside *E. coli* cells using modified promoters and plasmid copy numbers (Supplementary Figure S17). We found that increasing the mRNA levels of the TMR-10 riboswitch by about 1000-fold in cell-free assays increased its reporter expression levels, while also enhancing its AR from 3.9- to 14.8-fold (Figure 5B). However, when we increased the *in vivo* mRNA levels for the same TMR riboswitch, by replacing the single-copy BAC origin with either low-copy or high-copy origins of replication (p15A, ColE1 or pUC19), the measured AR substantially dropped from 12.3- to 5.7-fold (Figure 5C and Supplementary Figure S17A). The same differences in cell-free versus *in vivo* activation were observed for three different TMR riboswitches. We then verified that changing mRNA levels, and not plasmid burden, were responsible for modulating riboswitch activation by introducing three different promoters to vary the transcription rates of three fluoride and two theophylline riboswitches. A 50-fold increase in promoter transcription rate (Supplementary Figure S17B) dropped the F-11 riboswitch's AR from 4.0- to 1.6-fold and the Theo-44 riboswitch's AR from 12.6- to 6.3-fold (Figure 5E). The same trend was observed for all five riboswitches without finding a reduction in the hosts' growth rates, compared to wild-type or no-aptamer controls (Supplementary Figure S18A). We also found that the reporter expression levels from the no-aptamer controls always exceeded the riboswitches' expression levels, indicating that the host's maximum translational capacity had not been reached even as mRNA levels were increased (Supplementary Figure S18B). These results are qualitatively consistent with our model-calculated ARs (Figure 5D) and the expectation that volume exclusion will have a more pronounced effect on larger ligands (theophylline, TMR), compared to tiny fluoride. Additional experiments will be needed to confirm how macromolecular crowding affects riboswitch function and to quantify the factors that control its extent. Regardless of the mechanism, as a general engineering design principle, lowering a riboswitch's mRNA level will increase its AR inside cells.

Model-predicted limits of riboswitch-based sensors

Finally, we illustrate the potential limits of riboswitch-based sensors by designing an idealized 'perfect' riboswitch that uses a hypothetical aptamer and highly optimized pre, post and coding sequences for maximum possible detection ($AR_{max} = 16\,000$, $\Delta\Delta G_{mRNA} = 15.2$ kcal/mol) (Figure 6A, see Supplementary Information for more details). When a very high-affinity aptamer is available ($K_D \sim 10$ pM) and its ligand is plentiful, such a riboswitch could activate with a high dynamic range ($AR_{actual} > 1000$-fold). However, when faced with the most stringent sensing application, for example, sensing nanomolar ligand levels, even the most ideal riboswitch will only activate expression by 20-fold using a high-affinity aptamer and 3-fold using a low-affinity aptamer (Figure 6B). These limits provide an approximate best-case for engineering novel riboswitches, regardless of aptamer sequence and structure, and help determine when downstream signal amplification will be necessary.

DISCUSSION

We have developed a parsimonious statistical thermodynamic model that predicts the sequence-structure-function relationship for translation-regulating riboswitches, accounting for the interactions between the ribosome, mRNA and ligand. We experimentally validated the model by designing and characterizing 62 synthetic riboswitches, using six different RNA aptamers, that bind to their respective ligands (theophylline, TMR, fluoride, dopamine, thyroxine or 2,4-DNT) and activated translation rate by up to 383-fold. We used both cell-free *in vitro* transcription–translation assays as well as *in vivo* bacterial expression assays to measure riboswitch function. Importantly, we also performed extensive measurements of no-aptamer controls to ensure that the non-specific effects of the ligands on gene expression levels are correctly subtracted. For 55% of these riboswitches, the model was able to correctly calculate the ligand-induced changes in the ribosome-mRNA binding free energy to within 1.5 kcal/mol.
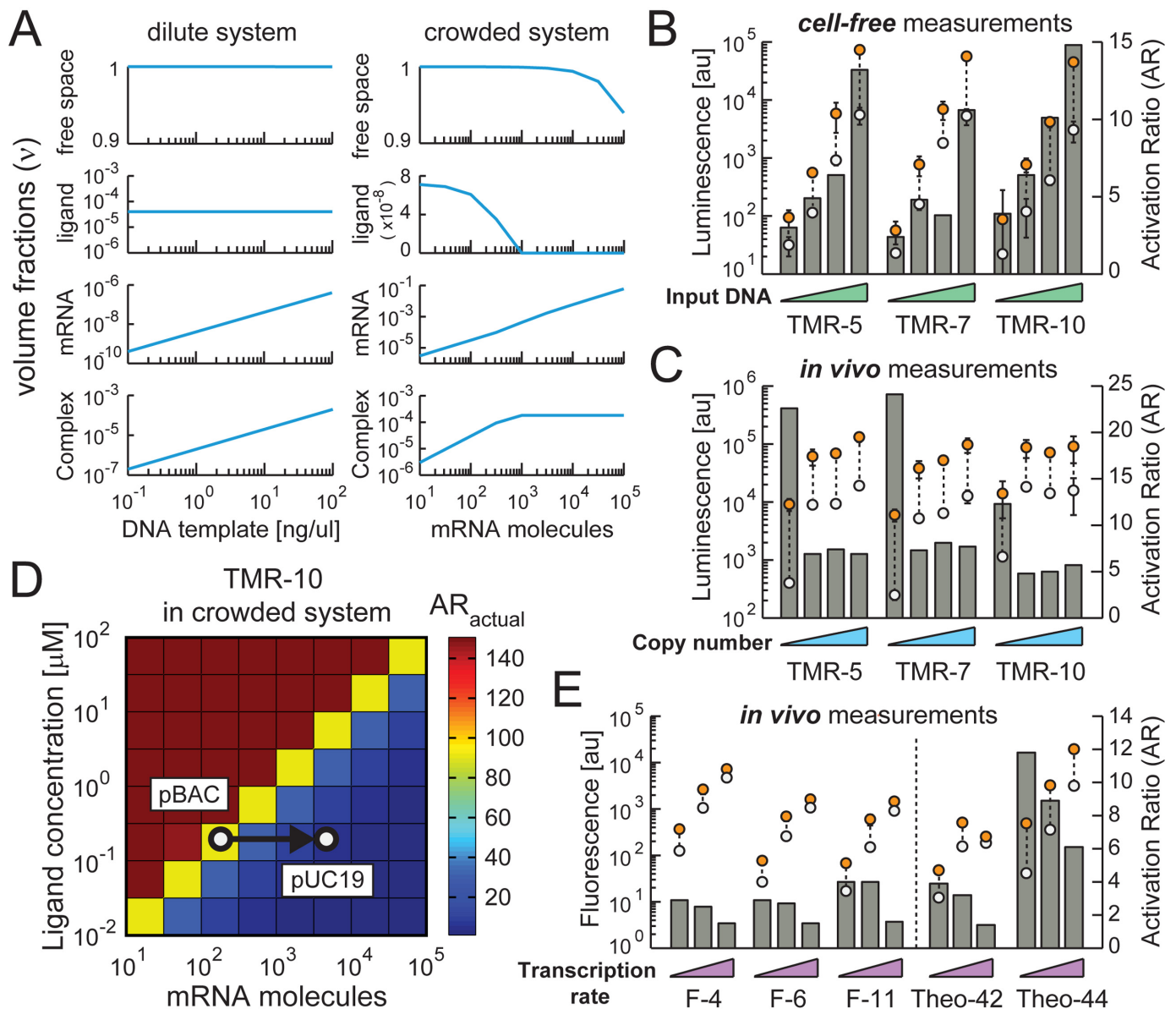
Figure 5. Macromolecular crowding controls riboswitch activation inside cells. (**A**) The relationship between riboswitch mRNA level and the volume fractions (ν) of free mRNA, free ligand, mRNA–ligand complex and free volume in either dilute or crowded systems were calculated according to Equation (2) for the TMR-10 riboswitch. (**B**) Cell-free transcription–translation assays were performed for three TMR riboswitches using increasing DNA template concentrations (0.1, 1, 10 and 100 ng/μl). Colors, symbols and ligand concentrations are the same as Figure 4. (**C**) The TMR riboswitches' *in vivo* expression levels were measured at zero and 20 μM TMR after increasing the plasmid copy numbers, using the pBAC, p15A, pColE1 and pUC19 replication origins. (**D**) Model-predicted AR_{actual} for the TMR-10 riboswitch when varying the total concentrations of mRNA and ligand. The effect of changing the plasmid's copy number on mRNA level is shown. (**E**) The *in vivo* expression levels and ARs for theophylline and fluoride riboswitches were measured after increasing riboswitch mRNA levels, using three promoters with steadily increasing transcription rates. Colors, symbols and ligand concentrations are the same as Figure 4. Each data point and bar represents the mean and s.d. of 2–4 measurements.

Using our thermodynamic model, we experimentally confirmed several mechanisms that control riboswitch function. First, some RNA aptamers are more conducive to riboswitch engineering than others, though for previously unknown reasons. We demonstrated that an aptamer's ligand binding free energy places an upper limit on the possible RNA conformational changes that may occur, according to its switching free energy, which then places a maximum limit on the riboswitch's AR. The same thermody-

namic limit will govern the ligand-induced conformational changes of all RNA regulators. Second, the effect of volume exclusion on biomolecular interactions inside crowded cells has remained poorly understood, though it is particularly important when expressing bulky mRNAs. A key model prediction is that higher mRNA concentrations will lower riboswitch ARs by reducing the amount of free volume. We systematically increased intracellular mRNA concentrations to show that, indeed, riboswitch ARs drop by
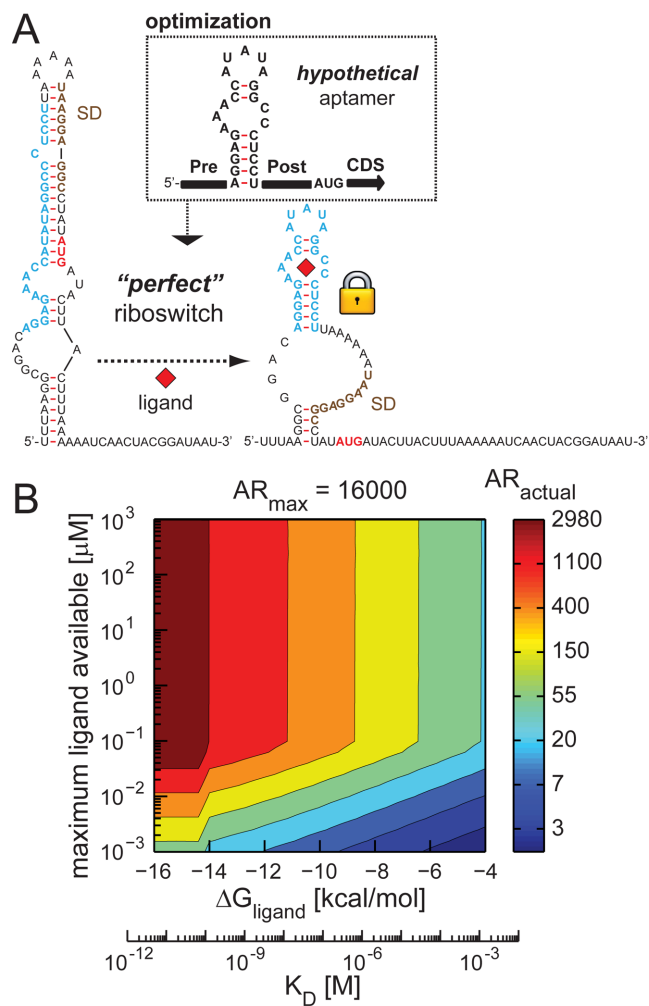
Figure 6. A 'perfect' riboswitch tests the limits of sensing. (**A**) Optimization was used to design a 'perfect' riboswitch using a hypothetical aptamer, which is shown in its ligand-free and ligand-bound states. (**B**) Model calculations show the effects of the aptamer's binding free energy (affinity) and maximum ligand concentration on the 'perfect' riboswitch's actual AR to illustrate the best possible expression changes under potential sensing scenarios.

considerable amounts. This effect was not observed in several *in vivo* no-aptamer controls and was found to depend on the ligand's molar volume, as expected according to the theory. Importantly, we repeated these measurements in the 20-fold more dilute cell-free assay and expectedly found the opposite behavior. The principles that we experimentally validated to develop this model are not specific to riboswitches and encompass RNA molecules whose ligand-induced conformational changes affect their interactions with proteins, other RNAs, or supramolecular complexes (e.g. the ribosome).

While all models are imperfect, here we systematically quantified the thermodynamic model's predictive accuracy across a large and diverse dataset, enabling us to distinguish the differences between stable and meta-stable riboswitches and further improve our understanding of the physics of RNA shape-change. Importantly, throughout this study, our model assumes thermodynamic equilibrium between the ribosome, mRNA, and ligand, and does not attempt to predict changes in mRNA stability. We envision that eliminating these assumptions will further improve the model's ability to predict riboswitch function from sequence.

Finally, by automating our physical chemistry calculations, both experts and non-experts alike can perform identical calculations to convert a wide array of RNA aptamers into engineered riboswitches, broadly enabling the rapid prototyping of new biomolecular sensors for use in medical diagnostics, environmental remediation, metabolic engineering and future fields (29). A web-interface to the automated design method, called the Riboswitch Calculator, is accessible at http://salislab.net/software.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Mandal,M. and Breaker,R.R. (2004) Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.*, **5**, 451–463.
2. Montange,R.K. and Batey,R.T. (2008) Riboswitches: emerging themes in RNA structure and function. *Annu. Rev. Biophys.*, **37**, 117–133.
3. Serganov,A. and Nudler,E. (2013) A decade of riboswitches. *Cell*, **152**, 17–24.
4. Michener,J.K., Thodey,K., Liang,J.C. and Smolke,C.D. (2012) Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.*, **14**, 212–222.
5. Berens,C. and Suess,B. (2015) Riboswitch engineering—making the all-important second and third steps. *Curr. Opin. Biotechnol.*, **31**, 10–15.
6. Baker,J.L., Sudarsan,N., Weinberg,Z., Roth,A., Stockbridge,R.B. and Breaker,R.R. (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science*, **335**, 233–235.
7. Meyer,C., Berg,K., Eydeler-Haeder,K., Lorenzen,I., Grötzinger,J., Rose-John,S. and Hahn,U. (2014) Stabilized Interleukin-6 receptor binding RNA aptamers. *RNA Biol.*, **11**, 57–65.
8. Zhou,J., Swiderski,P., Li,H., Zhang,J., Neff,C.P., Akkina,R. and Rossi,J.J. (2009) Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells. *Nucleic Acids Res.*, **37**, 3094–3109.
9. Barton,J.L., Bunka,D.H.J., Knowling,S.E., Lefevre,P., Warren,A.J., Bonifer,C. and Stockley,P.G. (2009) Characterization of RNA aptamers that disrupt the RUNX1-CBFβ/DNA complex. *Nucleic Acids Res.*, **37**, 6818–6830.
10. Cruz-Toledo,J., McKeague,M., Zhang,X.R., Giamberardino,A., McConnell,E., Francis,T., DeRosa,M.C. and Dumontier,M. (2012) Aptamer base: a collaborative knowledge base to describe aptamers and SELEX experiments. *Database (Oxford)*, **2012**, bas006.

11. Klussmann,S.E. (2006) The Aptamer Handbook: Functional Oligonucleotides and Their Applications. John Wiley & Sons, Chicago.

12. Serganov,A. and Patel,D.J. (2012) Metabolite recognition principles and molecular mechanisms underlying riboswitch function. *Annu. Rev. Biophys.*, **41**, 343–370.

13. Robinson,C.J., Vincent,H.A., Wu,M.-C., Lowe,P.T., Dunstan,M.S., Leys,D. and Micklefield,J. (2014) Modular riboswitch toolsets for synthetic genetic control in diverse bacterial species. J. Am. Chem. Soc., **136**, 10615–10624.

14. Mishler,D.M. and Gallivan,J.P. (2014) A family of synthetic riboswitches adopts a kinetic trapping mechanism. *Nucleic Acids Res.*, **42**, 6753–6761.

15. Ketzer,P., Kaufmann,J.K., Engelhardt,S., Bossow,S., von Kalle,C., Hartig,J.S., Ungerechts,G. and Nettelbeck,D.M. (2014) Artificial riboswitches for gene expression and replication control of DNA and RNA viruses. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E554–E562.

16. Soulière,M.F., Altman,R.B., Schwarz,V., Haller,A., Blanchard,S.C. and Micura,R. (2013) Tuning a riboswitch response through structural extension of a pseudoknot. *Proc. Natl. Acad. Sci.*, **110**, E3256–E3264.

17. Davidson,M.E., Harbaugh,S.V., Chushak,Y.G., Stone,M.O. and Kelley-Loughnane,N. (2013) Development of a 2,4-dinitrotoluene-responsive synthetic riboswitch in E. coli cells. *ACS Chem. Biol.*, **8**, 234–241.

18. Wang,J.X., Lee,E.R., Morales,D.R., Lim,J. and Breaker,R.R. (2008) Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling. *Mol. Cell*, **29**, 691–702.

19. Desai,S.K. and Gallivan,J.P. (2004) Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.*, **126**, 13247–13254.

20. Dohno,C., Kohyama,I., Kimura,M., Hagihara,M. and Nakatani,K. (2013) A synthetic riboswitch that operates using a rationally designed ligand–RNA pair. *Angew. Chem. Int. Ed. Engl.*, **52**, 9976–9979.

21. Mutalik,V.K., Qi,L., Guimaraes,J.C., Lucks,J.B. and Arkin,A.P. (2012) Rationally designed families of orthogonal RNA regulators of translation. *Nat. Chem. Biol.*, **8**, 447–454.

22. Rodrigo,G., Landrain,T.E., Shen,S. and Jaramillo,A. (2013) A new frontier in synthetic biology: automated design of small RNA devices in bacteria. *Trends Genet.*, **29**, 529–536.

23. Rodrigo,G., Landrain,T.E. and Jaramillo,A. (2012) De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 15271–15276.

24. Green,A.A., Silver,P.A., Collins,J.J. and Yin,P. (2014) Toehold switches: de-novo-designed regulators of gene expression. *Cell*, **159**, 925–939.

25. Chen,X. and Ellington,A.D. (2009) Design principles for ligand-sensing, conformation-switching ribozymes.PLoS Comput. Biol., **5**, e1000620.

26. Beisel,C.L. and Smolke,C.D. (2009) Design principles for riboswitch function. *PLoS Comput. Biol.*, **5**, e1000363.

27. Wachsmuth,M., Domin,G., Lorenz,R., Serfling,R., Findeiß,S., Stadler,P.F. and Mörl,M. (2015) Design criteria for synthetic riboswitches acting on transcription. *RNA Biol.*, **12**, 221–231.

28. Wachsmuth,M., Findeiß,S., Weissheimer,N., Stadler,P.F. and Mörl,M. (2013) De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res.*, **41**, 2541–2551.

29. Carothers,J.M., Goler,J.A., Juminaga,D. and Keasling,J.D. (2011) Model-driven engineering of RNA devices to quantitatively program gene expression. *Science*, **334**, 1716–1719.

30. Salis,H.M. (2011) The ribosome binding site calculator. *Methods Enzymol.*, **498**, 19–42.

31. Salis,H.M., Mirsky,E.A. and Voigt,C.A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.*, **27**, 946–950.

32. Espah Borujeni,A., Channarasappa,A.S. and Salis,H.M. (2014) Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.*, **42**, 2646–2659.

33. Farasat,I., Kushwaha,M., Collens,J., Easterbrook,M., Guido,M. and Salis,H.M. (2014) Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. Mol. Sys. Bio., **10**, 1–18.

34. Tian,T. and Salis,H.M. (2015) A predictive biophysical model of translational coupling to coordinate and control protein expression in bacterial operons. *Nucleic Acids Res.*, **43**, 7137–7151.

35. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.

36. Cao,S. and Chen,S.-J. (2009) Predicting structures and stabilities for H-type pseudoknots with interhelix loops. *RNA*, **15**, 696–706.

37. Sperschneider,J. and Datta,A. (2010) DotKnot: pseudoknot prediction using the probability dot plot under a refined energy model. *Nucleic Acids Res.*, **38**, e103.

38. Jenison,R.D., Gill,S.C., Pardi,A. and Polisky,B. (1994) High-resolution molecular discrimination by RNA. *Science*, **263**, 1425–1429.

39. Baugh,C., Grate,D. and Wilson,C. (2000) 2.8 Å crystal structure of the malachite green aptamer. *J. Mol. Biol.*, **301**, 117–128.

40. Ren,A., Rajashankar,K.R. and Patel,D.J. (2012) Fluoride ion encapsulation by Mg2 +ions and phosphates in a fluoride riboswitch. *Nature*, **486**, 85–89.

41. Mannironi,C., Di Nardo,A., Fruscoloni,P. and Tocchini-Valentini,G. (1997) In vitro selection of dopamine RNA ligands. *Biochemistry*, **36**, 9726–9734.

42. Lévesque,D., Beaudoin,J.-D., Roy,S. and Perreault,J.-P. (2007) In vitro selection and characterization of RNA aptamers binding thyroxine hormone. *Biochem. J.*, **403**, 129–138.

43. Ehrentreich-Förster,E., Orgel,D., Krause-Griep,A., Cech,B., Erdmann,V.A., Bier,F., Scheller,F. and Rimmele,M. (2008) Biosensor-based on-site explosives detection using aptamers as recognition elements. *Anal. Bioanal. Chem.*, **391**, 1793–1800.

44. Murooka,Y., Doi,N. and Harada,T. (1979) Distribution of membrane-bound monoamine oxidase in bacteria. *Appl. Environ. Microbiol.*, **38**, 565–569.

45. Zhou,H.-X., Rivas,G. and Minton,A.P. (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.*, **37**, 375–397.

46. Ellis,R.J. (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.*, **26**, 597–604.