Supplementary Note 1: FACS-Seq analysis of a ribozyme library establishes loop sequence flexibility and consensus

The general applicability of our tertiary interaction switching architecture relies on the ability to restore activity to a ribozyme that has one loop sequence modified by integration of an aptamer. Restoration of cleavage activity is accomplished through the selection of an appropriate opposite loop sequence that restores tertiary interactions and geometries conducive to self-cleavage. The results for the theophylline, tetracycline, and neomycin aptamers demonstrate the feasibility of this mechanism for these particular loop sequences. To further investigate whether this strategy holds more generally, we built a library of HHRzs with loops I and II, each containing four random nucleotides (**Supplementary fig. 5a)**. The activities of the N4/N4 library members (in terms of the GFP/mCherry ratio (μ)) were measured using the described FACS-Seq strategy to provide a genotype-phenotype map for these sequences.

The data allowed us to examine our original hypothesis that it is possible to obtain significant activity in a ribozyme with an arbitrary sequence on one loop by varying the opposite loop sequence in the case of loops four nucleotides long. For each of the 256 possible sequences on each loop, we identified the sequence on the opposite loop that produced the lowest μ (**Supplementary fig. 6)**. The data indicate that for 80% of the possible loop I sequences there is a corresponding loop II sequence that results in a μ of ≤ 0.36. Similarly, for 80% of the possible loop II sequences, there exists a sequence for loop I that results in a μ of ≤ 0.34. For comparison, the inactive control ribozyme (sTRSVctl; **Supplementary table 2**) results in a μ of 1.5 and the average μ over the N4/N4 library is 0.50. Increasing the loop library to contain longer loops up to eight nucleotides long, as used with our aptamer libraries, provides significantly greater flexibility and will have a "best" μ value (at least as low as observed here, since the larger library is a superset of the N4/N4 library). The data support the hypothesis that the HHRz structure is

sufficiently flexible in its sequence space to allow loop replacement with arbitrary sequences, provided that the opposite loop sequence can be optimized.

We further examined the N4/N4 dataset to determine whether there are clusters in the sequence space that lead to higher ribozyme activity. Such clustering information may be useful in reducing the size of libraries needed for functional screening or may ultimately allow fully rational design. We analyzed these data for effects on μ due to either single nucleotide or interaction effects due to the identities of two nucleotides (**Supplementary fig. 7**). For each nucleotide position or pair of nucleotides, we computed the 10^{th}-percentile value of μ over all the measured sequences with those nucleotide position(s) as given (see **Online Methods**). The analysis indicates a strong effect of UG as the 5' nucleotides of loop I and A as the 3' nucleotide of loop II, together resulting in the greatest reduction in expression averaged over all other nucleotide pairs. The consensus sequences of UGNN for loop I and NNNA for loop II are consistent with prior work[1] that observed that natural type III HHRzs similar to sTRSV exhibited loop motifs with a loop I of the form UN_mYN and loop II of the form RN_mA. (where N_m is one or more unspecified nucleotides; Y is C or U; R is A or G).

Our library characterization data also indicate that the N4/N4 sequences cover the entire range of gene expression. A new toolkit for modulating gene expression levels was developed by characterizing library members that span a wide range of expression levels. From the FACS-Seq NGS data for the N4/N4 library we estimated the mean GFP/mCherry ratio (μ) for 20,744 sequences. A subset of these sequences were individually reconstructed and validated through flow cytometry analysis as previously described. The μ values obtained through flow cytometry analysis of the reconstructed sequences are tightly correlated with those obtained through the FACS-Seq analysis (**Supplementary fig. 5b;** R^2=0.98**).** We then examined the range and

coverage of the full set of sequences as an expression control library. The sequences span a wide range of μ with consistent coverage (**Supplementary fig. 8)**; the highest activity (lowest μ) is close to that of the wild-type sTRSV ribozyme and the lowest activity is similar to that of the inactive ribozyme control (sTRSVctl). This graded ribozyme library thus provides a new genetic tool for modulating gene expression levels over a 77-fold range through choice of the particular ribozyme sequence. A subset of validated sequences, which uniformly span the range, are provided in **Supplementary table 4.**

Supplementary Note 2: Identified best practices for FACS-Seq related methods

In the implementation of the described methods we encountered several issues, which required additional experimentation and controls to resolve. With attention to these issues, below we propose some "best practices" for similar experiments.

Chimeras. In preparing samples for NGS, it is important to avoid over-amplification during PCR steps. Excessive amplification results in depletion of primers followed by imperfect annealing of template strands that differ in the random library regions. Since our libraries have variability only on loops I and II of the hammerhead ribozyme, different library sequences have sequence homology at both ends and in the middle. These homologies allow various forms of partial hybridization. We observed chimera sequences composed of loop I sequences from one library entry with loop II sequences from a different library entry. The most salient of these chimeras consists of sequences with different aptamers on each stem, a case that was not present in the initial libraries, and thus readily identified. However, chimeras that are also in the initial libraries are more insidious since they cannot be readily identified and yet are counted as if they had been sorted in the same way as the library constituents, resulting in assigning incorrect values to these sequences. In addition, over-amplification can result in daisy chains of strands that result in incorrect measurement of sample molarities and consequent incorrect loading for NGS. All of these issues are mitigated by stopping PCR amplifications well before primer depletion, which is most easily achieved by using an intercalating dye and monitoring PCR progress on a quantitative PCR machine. We used the Kapa Real Time PCR Kit (KAPA Biosystems, Wilmington, MA, USA**)** for this purpose.

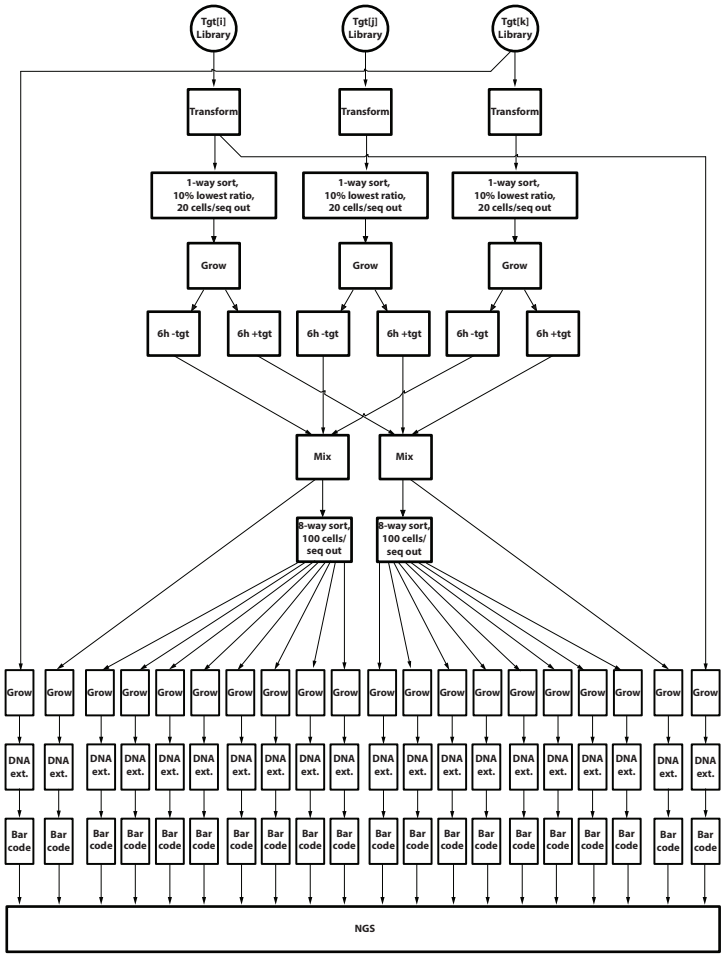
Contamination control. As with any procedure involving low numbers of DNA molecules, contamination by prior PCR products is an ever-present concern. This is primarily an issue at the

first stage of amplification from the plasmid preparations of each bin. We encountered this issue for many sequences, since we had previously run a pilot experiment of a similar library resulting in small, but significant counts of contaminant molecules in the laboratory. We track these contaminations by preparing a plasmid extract of cells grown in parallel to our samples that did not contain a switch insert and carried out the same preparations on these through to sequencing. Reads for this sample can thus only result from cross-contamination. To reduce this contamination issue to an acceptable level, we used nested PCR to amplify the sequences of interest from the plasmids. The outer PCR was run for 10 cycles using primers flanking the area of interest, which amplifies each plasmid molecule well above any contaminant background, while keeping the concentration low enough so that it does not become a significant source of contamination for future experiments that reuse the outer PCR primers. The product from the outer PCR was used as a template in a subsequent PCR to amplify the precise region of interest to nanomolar concentrations without depleting the primers, as described above.

Diversity control. With NGS, it is relatively easy to obtain 100's of millions of sequencing reads, which can lead to high precision of measurements of Poisson processes such as counting the number of occurrences of a particular sequence in a bin. However, if there are significant constrictions on the total number of molecules present at any stage in the process, the Poisson statistics will overestimate the precision since the repeated reads of the same parent molecule amount to little more than repeated observations of the same measurement. The most obvious constriction is the cell sorting itself – the number of viable cells captured during sorting sets an upper bound on the measurement precision. Additional constrictions of diversity at later stages in the process can further undermine the power of the statistics without any obvious indication. To

monitor these cases, we spiked into each sample a diversity control, or a sample of DNA encoding a sufficiently long random sequence such that the likelihood of two molecules having the same sequence is vanishingly small (we used an N17 region for this purpose). The random region of this diversity control is flanked by the same set of nested primers that are used for amplification of the sample of interest. In this way, the occurrence of multiple NGS reads of the same diversity control sequence can be used to measure the actual diversity at the point of addition of this control. We added this diversity control to the plasmid preparations prior to PCR amplification and confirmed that the number of molecules of interest was significantly greater than the number of cells sorted for that sample.

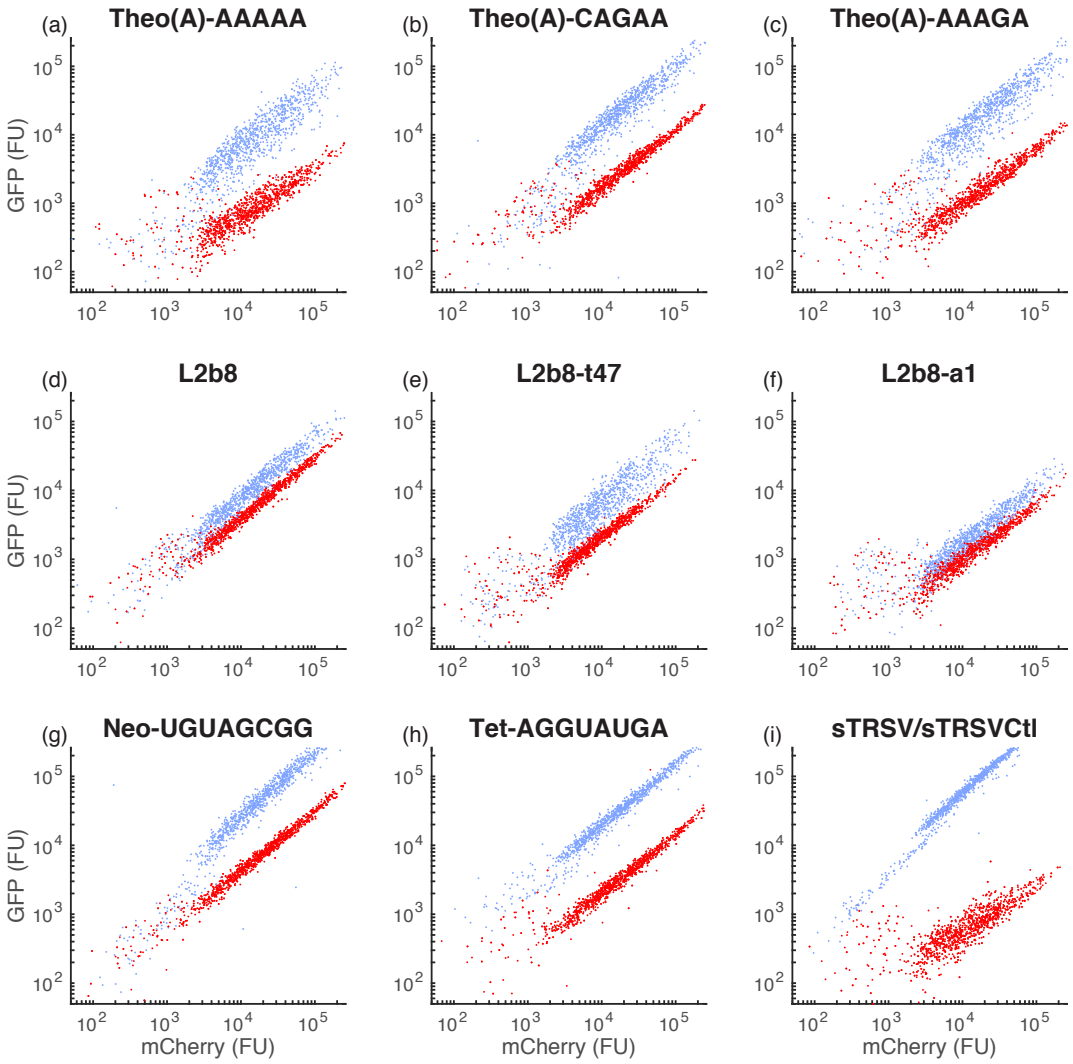
Supplementary fig. 1 Overall process flowchart of the FACS-Seq pipeline.



Overview process flowchart of the FACS-Seq pipeline used to generate tertiary interaction RNA devices to three targets: theophylline (tgt [i]), tetracycline (tgt[j]), and neomycin (tgt[k]). Device libraries developed to existing aptamer-target pairs are gap-repaired into yeast, and populations harboring the library constructs are separately sorted to enrich for devices exhibiting the highest gene-silencing activities. One-way sorts are performed to collect at least 15 cells per library sequence for members exhibiting the lowest 10% GFP/mCherry ratio (μ). For each library, the collected prescreened cells are re-grown, back-diluted 20:1, and then assayed. Assay conditions

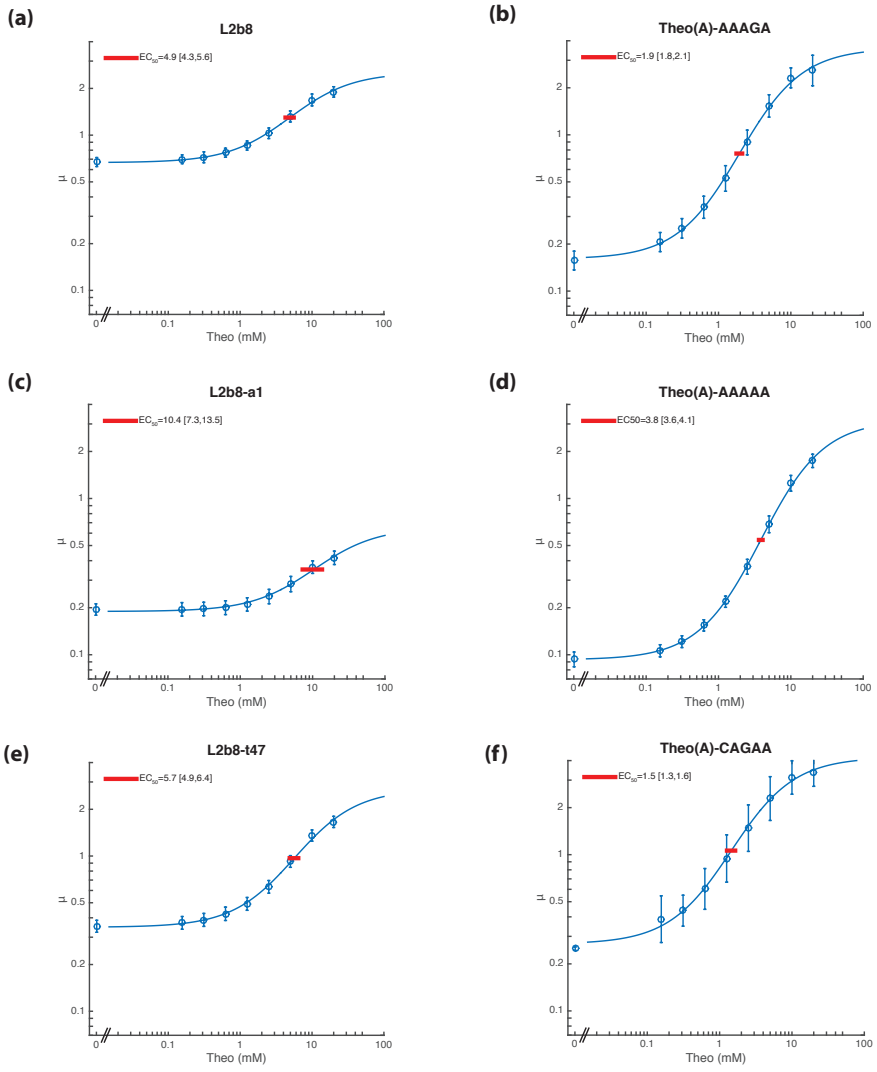
are standardized at 6 hrs of growth in the presence (+tgt) and absence (-tgt) of their respective target. The three –tgt and +tgt samples are combined prior to FACS bin sorting (‘mix’ step). FACS is carried out on the combined –tgt and +tgt samples separately, collecting 5.9M-7.6M cells into eight bins based on μ values during each sort. The DNA from each bin is extracted (ext.) and amplified with unique sequence barcodes. DNA from each ‘mix’ (prior to the 8-way sort) is also isolated and barcoded with the bin samples. The entire process is performed in replicate (not shown in figure) with orthogonal barcodes, and DNA from both replicates is submitted in the same NGS sample. The following libraries are also submitted in the NGS sample: DNA from the library-transformed cells (prior to any sorting), DNA library constructs (used in the transformation), and DNA from non-transformed cells (not depicted in flowchart).

Supplementary fig. 2 Flow cytometry data for representative cultures.



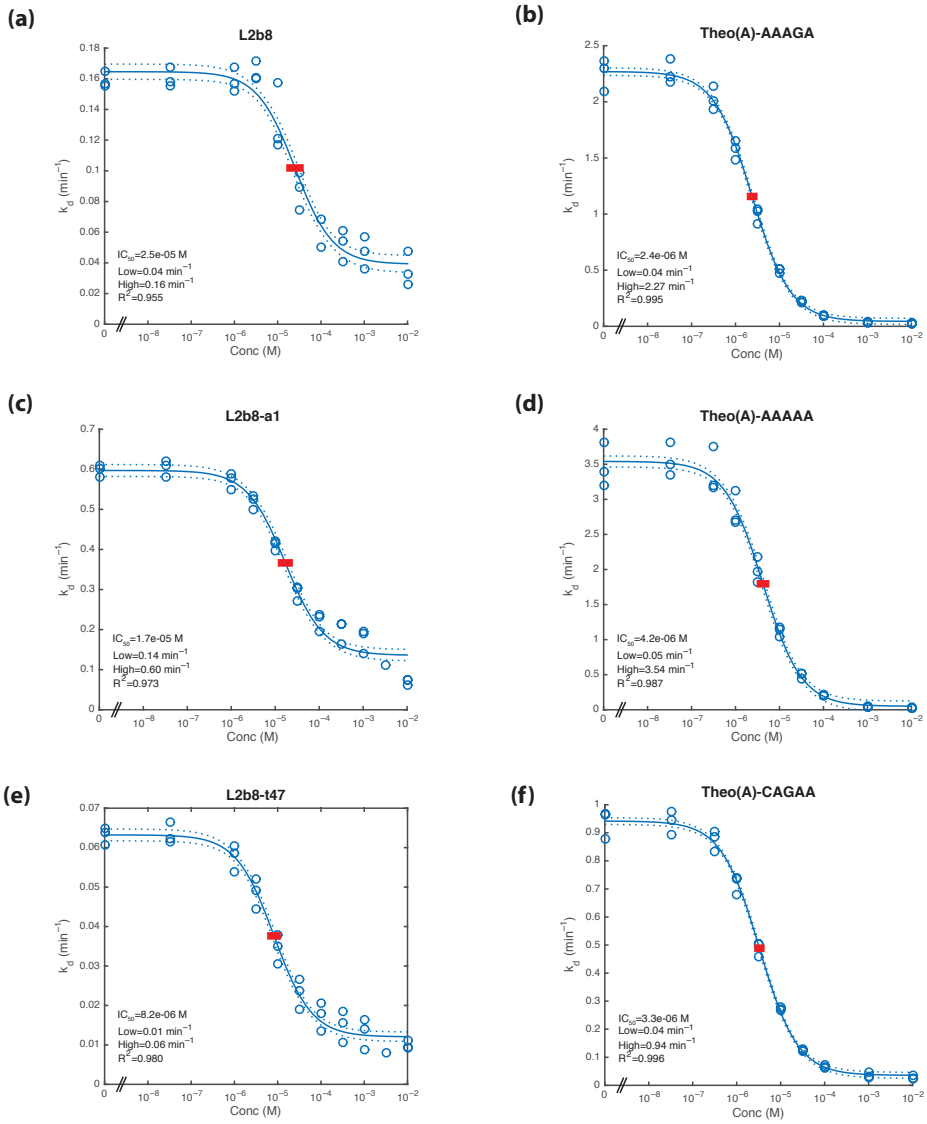
Each panel shows ungated mCherry vs. GFP fluorescence for 1,000 events for two cultures from a common pre-assay culture. For panels (a)-(h), the red points show a culture that was assayed without addition of target and the blue points show the same culture, assayed with target. Switches represented are (a) Theo(A)-AAAAA, (b) Theo(A)-CAGAA, (c) Theo(A)-AAAGA, (d) L2b8, (e) L2b8-t47, (f) L2b8-a1, (g) Neo-UGUAGCGG, (h) Tet-AGGUAUGA. Panel (i) shows the native hammerhead ribozyme (sTRSV) in blue and the scrambled-core version (sTRSVctl) in red.

Supplementary fig. 3 Dose-response curves from flow cytometry for six theophylline-responsive switches.



GFP/mCherry ratios are reported for cells harboring one of six devices grown at the indicated theophylline concentrations for L2b8 (a), Theo(A)-AAAGA (b), L2b8-a1 (c), Theo(A)-AAAAA (d), L2b8-t47 (e), and Theo(A)-CAGAA (f). Data show the average (circles) and standard deviation (error bars) of measurements from at least four biological replicate experiments for each condition. The data were fit to a 4-parameter logistic model (solid lines) with the Hill slope fixed at 1.0. The EC_{50} is shown as a red bar covering the 80% confidence interval.

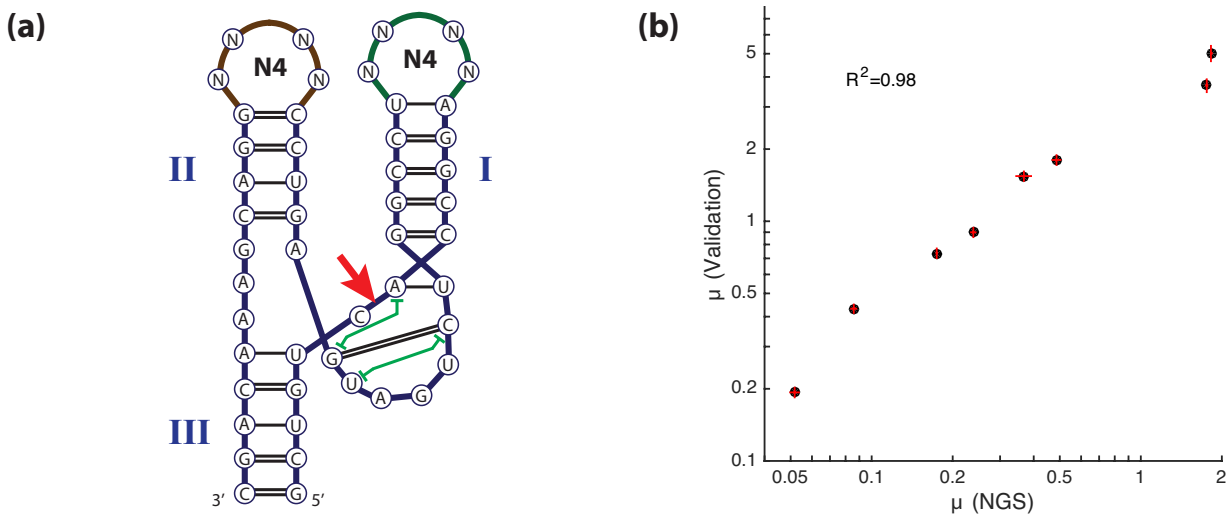
Supplementary fig. 4 Relationship between cleavage kinetics and theophylline levels.



Relationship between cleavage kinetics and theophylline levels for selected theophylline-responsive tertiary interaction devices (Theo(A)-AAAGA, Theo(A)-AAAAA, Theo(A)-CAGAA) and secondary-structure switching devices (L2b8, L2b8-a1, L2b8-t47). Cleavage rates were measured as RNA dissociation rate constants (k_d) using a previously described SPR-based cleavage assay[2] for L2b8 (a), Theo(A)-AAAGA (b), L2b8-a1 (c), Theo(A)-AAAAA (d), L2b8-t47 (e), and Theo(A)-CAGAA (f) with varying levels of theophylline. Three separate runs were

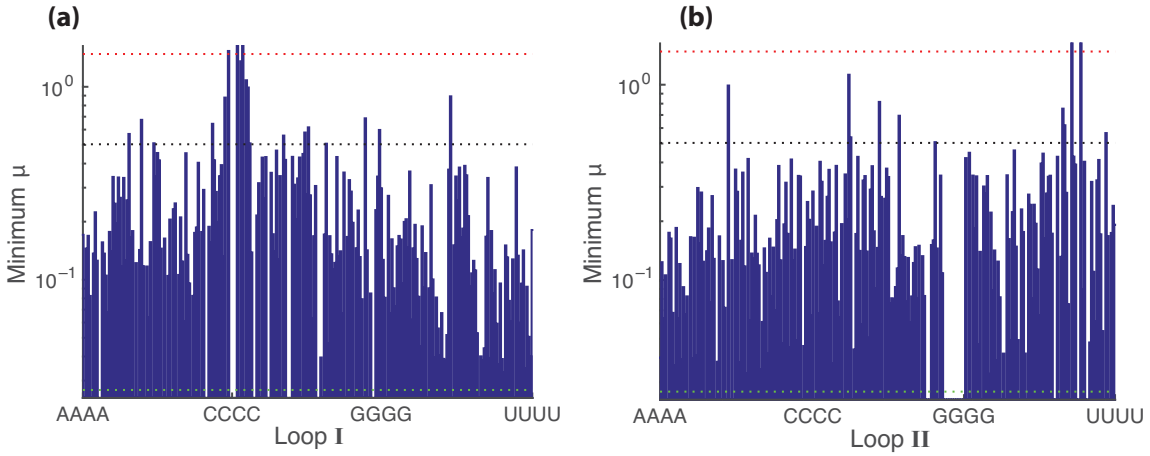
performed for each device starting with fresh buffer for each run. During a run, measurements of k_d were acquired at each concentration of theophylline, with each individual measurement shown as a separate circle. Dose response curves of cleavage activity inhibition were fit independently for each switch with MATLAB using a 4-parameter logistic model with the Hill slope set to 1.0 with the best fit shown as solid lines and 80% confidence intervals for the fits shown as dotted lines. The 80% confidence interval for the IC_{50} is shown as a red bar. Cleavage reactions were carried out at 25^{o}C, in 500 μM $MgCl_2$, 150 mM NaCl, 10 mM HEPES (pH 7.4) and the indicated theophylline concentrations.

Supplementary fig. 5 Validation of an N4/N4 ribozyme library.



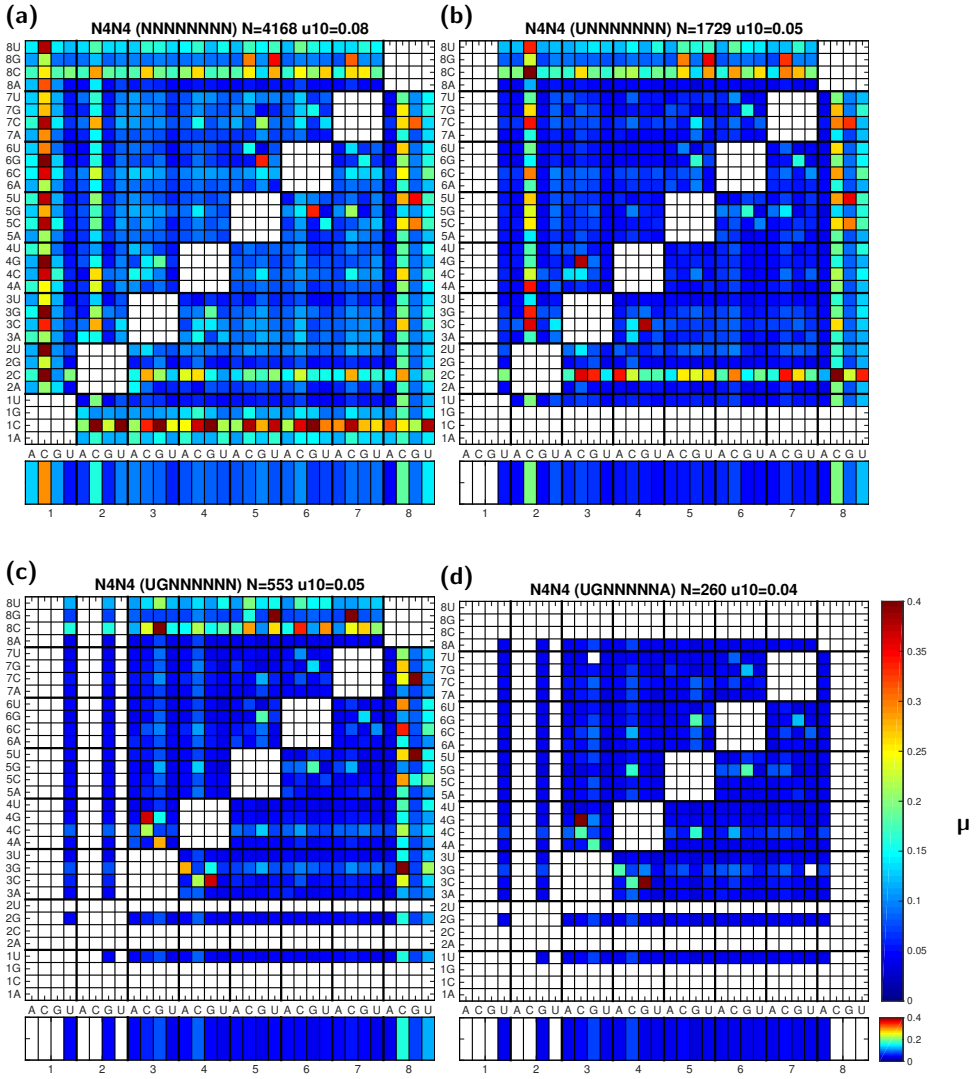
(a) Library design for N4/N4 ribozyme tertiary interactions. Shown are loop libraries (N4/N4; green/brown) grafted onto stems I and II of the sTRSV hammerhead ribozyme (blue). Red arrow indicates the ribozyme cleavage site in the backbone. (b) Comparison of NGS- and flow cytometry-based GFP/mCherry activity measurements for individual N4/N4 library members. Each point is a single library sequence that was identified from the NGS analysis. Error bars for the flow cytometry validated measurements represent the standard error of the mean over at least three biological replicates. Error bars for measurements from the NGS analysis represent the range over the two biological replicates.

Supplementary fig. 6 GFP/mCherry activities (μ) of N4/N4 ribozyme library members for fixed values of one loop.



Minimum μ values observed over all sequences as a function of the sequence on loop I (a) and loop II (b) are shown. In each case, a total of 2,174 sequences (for which at least 100 cells were measured) were considered. Activities for sTRSV (green; lower), sTRSVctl (red; upper), and N4/N4 library mean (black; central) are shown as dotted lines. Note that coverage at the level of cell counts was low, with ~5-10 distinct opposite loop sequences available from which to select the minimum for each loop. Thus, the minimum reported values are an upper bound on the minimum attainable if all 256 opposite loops were considered.

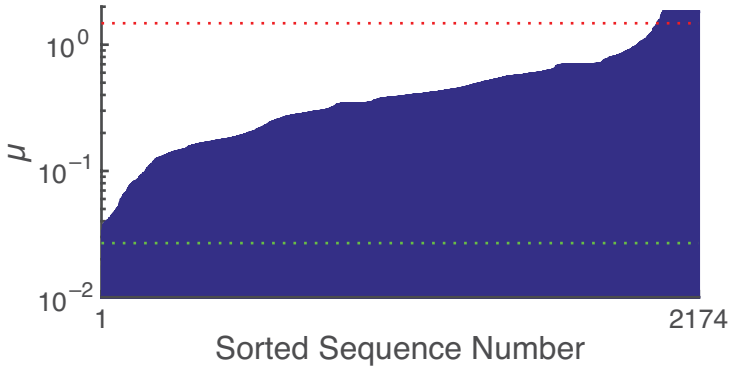
Supplementary fig. 7 Effect of loop sequence on activity of the N4/N4 ribozymes.



Each panel shows 10th-percentile μ values, as a pairwise function of the identity of the four randomized nucleotides on each stem of a ribozyme, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. Sequence positions are labeled 5' to 3' starting with the four nucleotides on loop I followed by the four nucleotides on loop II. (a) Sequence analysis based on iterating over all possible values of two nucleotides and averaging over all possibilities for the remaining six

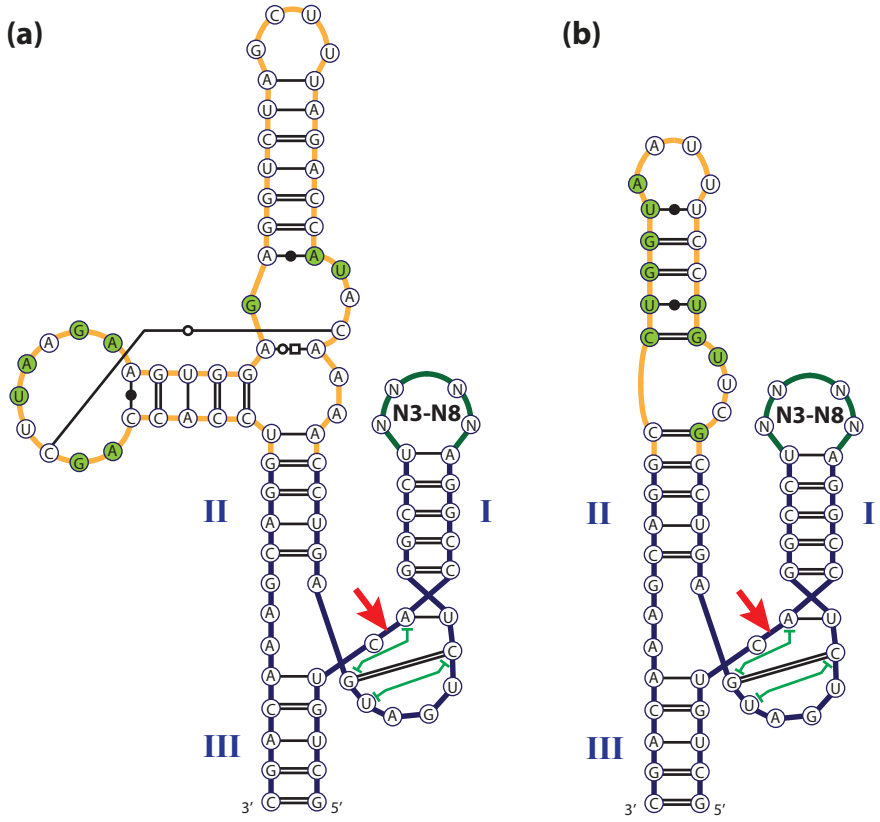
nucleotides in the loops. The strongest single nucleotide positive effect on catalytic activity arises when position 1 as a U. There is also a strong negative effect of a C in position 1. (b) Sequence analysis based on fixing position 1 as a U and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 2 is a G. (c) Sequence analysis based on fixing position 1 as a U and position 2 as a G and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 8 is an A. (d) Sequence analysis based on fixing position 1 as a U, position 2 as a G, and position 8 as an A and averaging over the remaining three nucleotides for each entry. With the consensus sequences set to UGNNNNNNA, the 10^{th}-percentile μ is 0.04.

Supplementary fig. 8 GFP/mCherry activities (μ) of N4/N4 ribozyme library members.



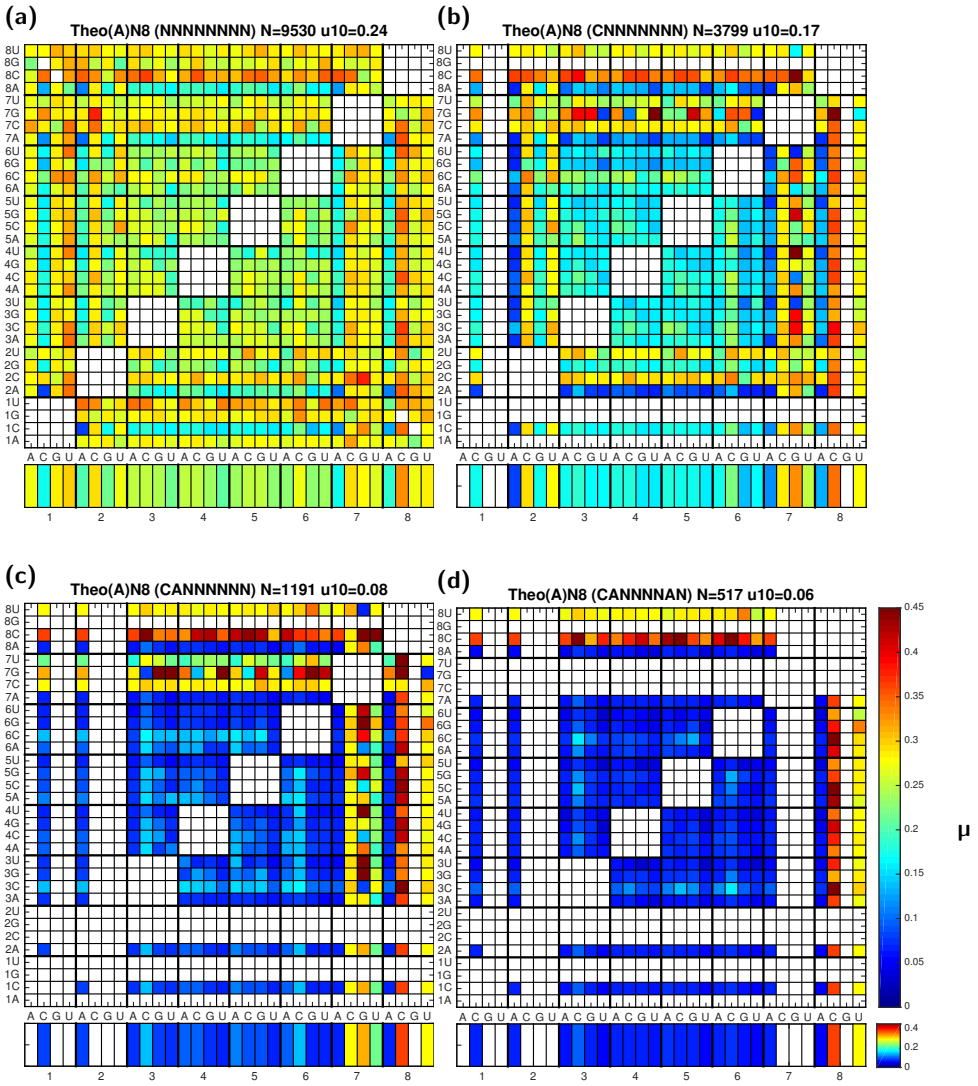
Activities are shown only for those sequences that had at least 100 cells sorted. Activities for sTRSV (green; lower) and sTRSVctl (red; upper) are shown as dotted lines.

Supplementary fig. 9 Library design for tetracycline- and neomycin-responsive tertiary interaction switches.



Shown are loop libraries (N3-N8; green) grafted onto stem II of the sTRSV hammerhead ribozyme (blue), with the tetracycline aptamer (a, gold) and neomycin (b, gold) aptamer on the opposing loop. The libraries also contain the aptamers grafted onto stem I (not shown). Red arrow indicates the ribozyme cleavage site in the backbone. Notation of interactions in the ribozyme and aptamers follow Leontis-Westhof notation[3] with the addition of green "I-beams" showing non-adjacent base-stacking interactions. Nucleotides in contact with the ligand are indicated in green. Interactions were based on 3D structures from the RCSB PDB[4] entries 2QUS[5] (ribozyme), 2KXM[6] (neomycin aptamer), and 3EGZ[7] (tetracycline aptamer). Base pair interactions were extracted from the PDB entries using FR3D[8].

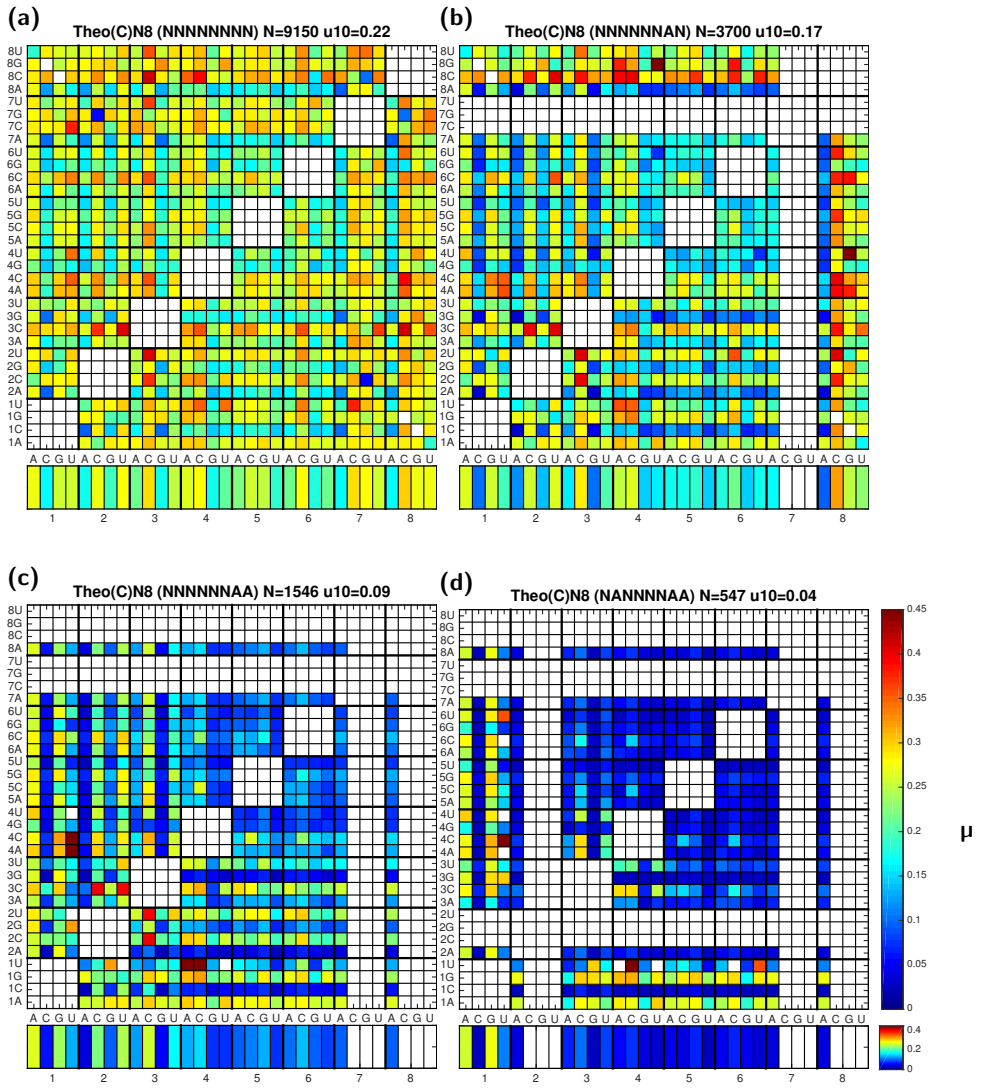
Supplementary fig. 10 Effect of loop II sequence on activity of theophylline-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile μ values as a pairwise function of the identity of two of the randomized nucleotides on stem II of a ribozyme containing a theophylline aptamer (AAG-variant) on the opposing stem I, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on varying two nucleotides and averaging over all possibilities for the remaining

six nucleotides in the loop. The strongest single nucleotide positive effect on catalytic activity arises when position 1 is a C. Interaction effects between any of nucleotides 2, 7, or 8 being an A are also evident. (b) Sequence analysis based on fixing position 1 as a C and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 2 is an A and there is a negative effect observed when position 8 is a C or position 7 is a G. (c) Sequence analysis based on fixing position 1 as a C and position 2 as an A and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 7 is an A. (d) Sequence analysis based on fixing position 1 as a C, position 2 as an A, and position 7 as an A and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to CANNNNAN, the 10^{th}-percentile μ is 0.06.

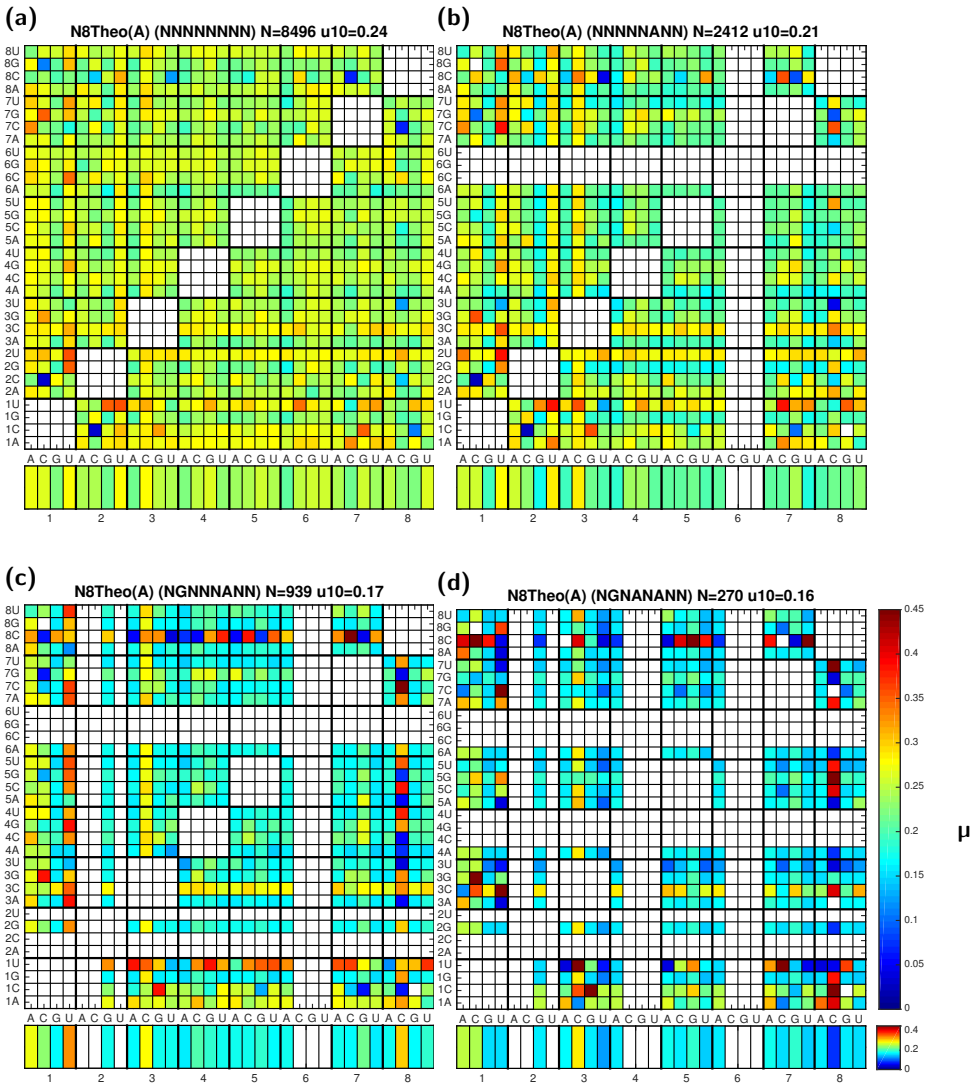
Supplementary fig. 11 Effect of loop II sequence on activity of theophylline-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem II of a ribozyme containing a theophylline aptamer (CAG-variant) on the opposing stem I, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on averaging over all possibilities for the remaining six nucleotides in

the loop. The strongest single nucleotide positive effect on catalytic activity arises when position 7 is an A. (b) Sequence analysis based on fixing position 7 as an A and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 8 is an A. (c) Sequence analysis based on fixing position 7 as an A and position 8 as an A and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 2 is an A. (d) Sequence analysis based on fixing position 2 as an A, position 7 as an A, and position 8 as an A and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to NANNNNAA, the 10^{th}-percentile μ is 0.04.

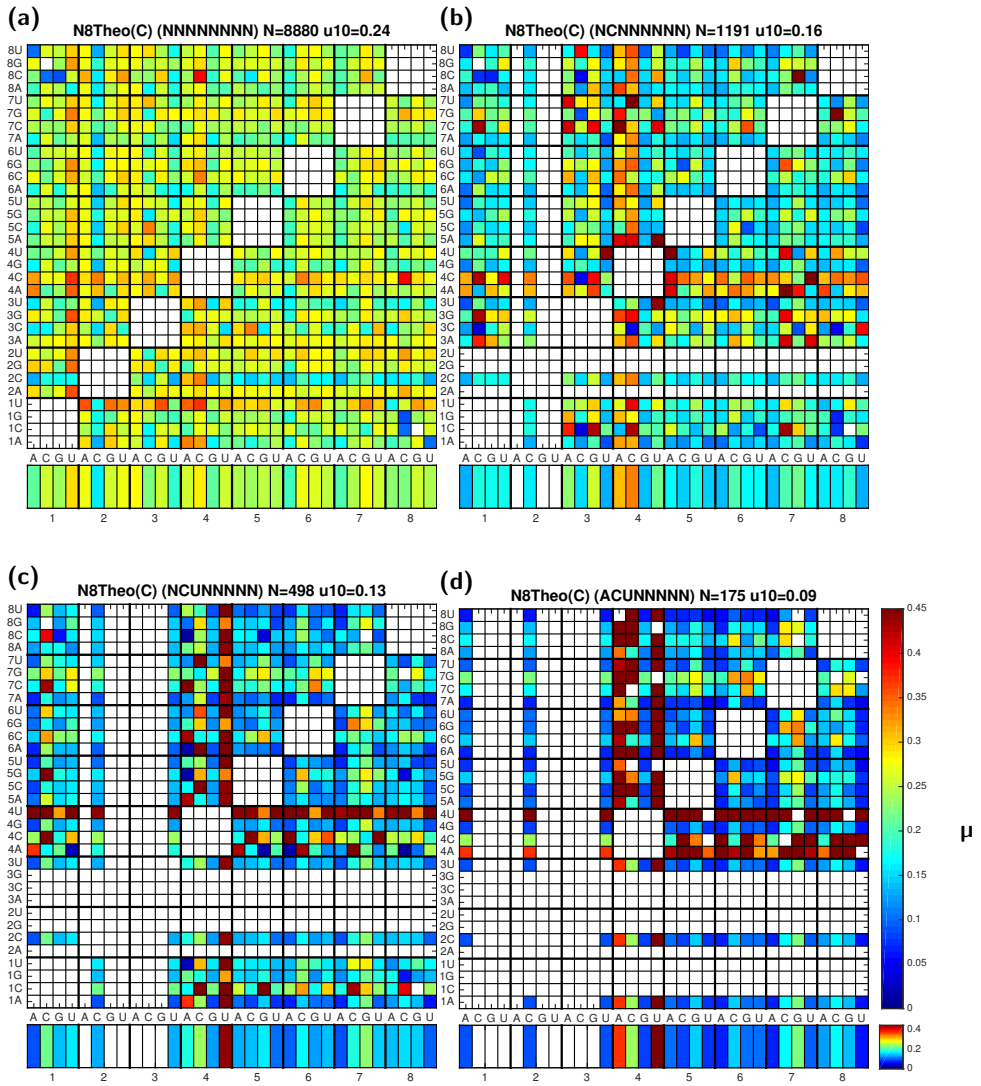
Supplementary fig. 12 Effect of loop I sequence on activity of theophylline-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem I of a ribozyme containing a theophylline aptamer (AAG-variant) on the opposing stem II, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on averaging over all possibilities for the remaining six nucleotides

in the loop. The strongest single nucleotide positive effect on catalytic activity arises when position 6 is an A. (b) Sequence analysis based on fixing position 6 as an A and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 2 is a G. (c) Sequence analysis based on fixing position 2 as a G and position 6 as an A and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 4 is an A. (d) Sequence analysis based on fixing position 2 as a G, position 4 as an A, and position 6 as an A and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to NGNANANN, the 10^{th}-percentile μ is 0.16.

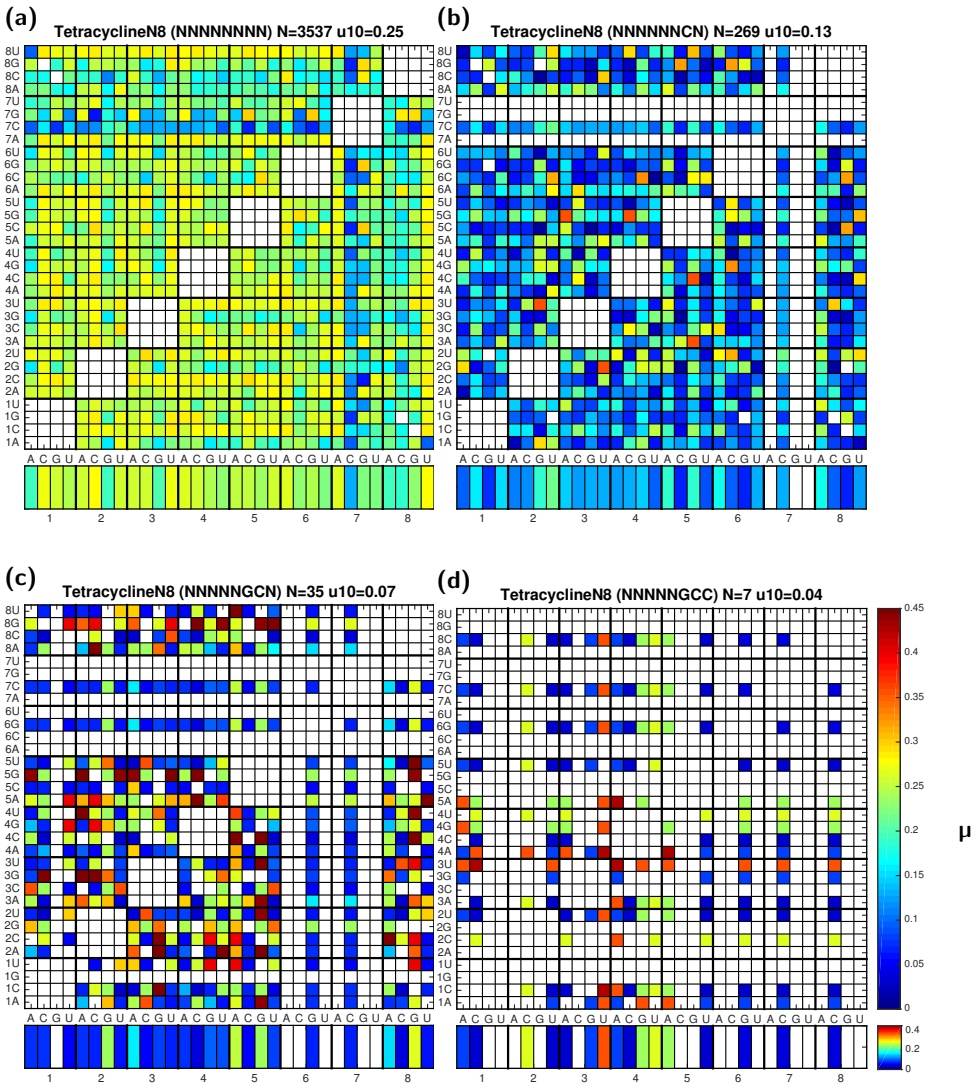
Supplementary fig. 13 Effect of loop I sequence on activity of theophylline-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem I of a ribozyme containing a theophylline aptamer (CAG-variant) on the opposing stem II, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on averaging over all possibilities for the remaining six nucleotides

in the loop. The strongest single nucleotide positive effect on catalytic activity arises when position 2 is a C. (b) Sequence analysis based on fixing position 2 as a C and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 3 is a U. (c) Sequence analysis based on fixing position 2 as a C and position 3 as a U and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 1 is an A. (d) Sequence analysis based on fixing position 1 as an A, position 2 as a C, and position 3 as a U and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to ACUNNNNN, the 10^{th}-percentile μ is 0.09.

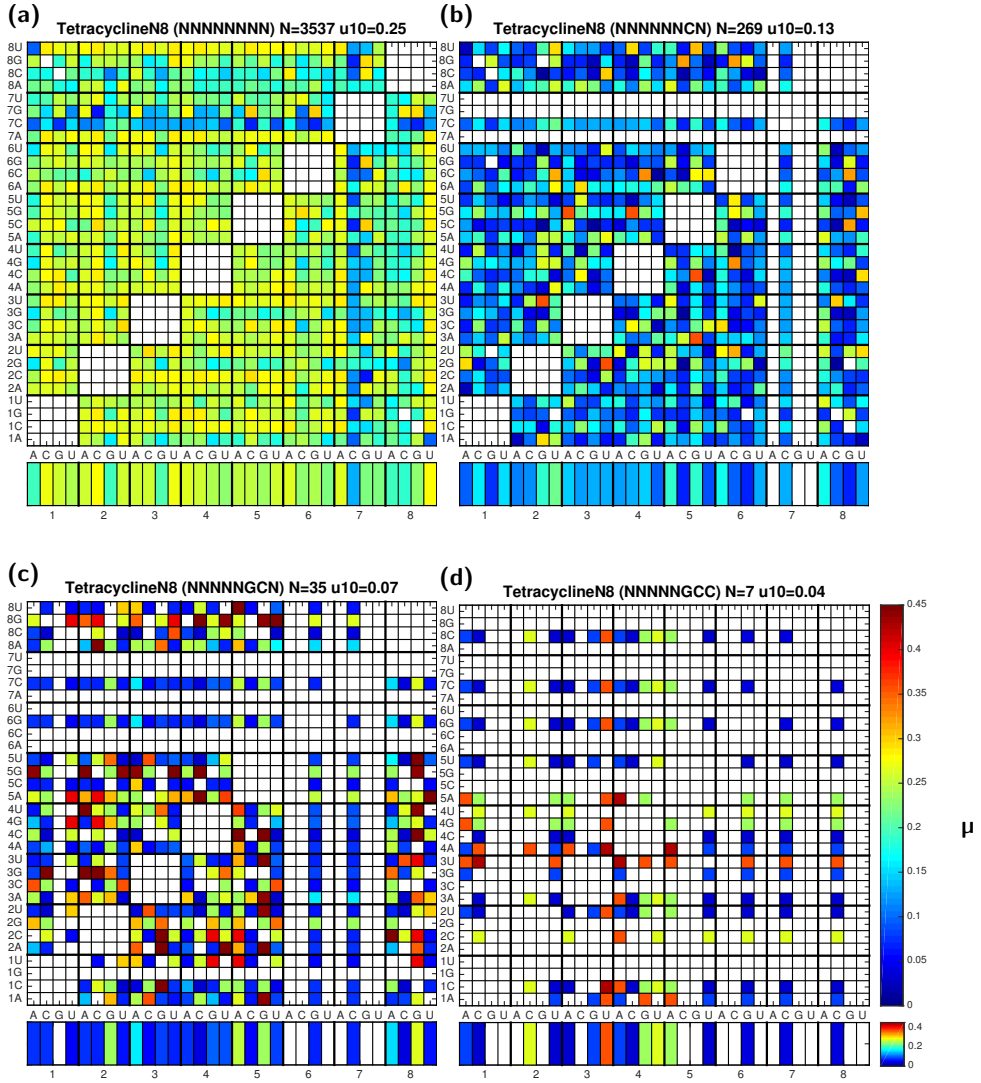
Supplementary fig. 14 Effect of loop II sequence on activity of tetracycline-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem I of a ribozyme containing a tetracycline aptamer on the opposing stem II, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on averaging over all possibilities for the remaining six nucleotides in the loop.

The strongest single nucleotide positive effect on catalytic activity arises when position 7 is a C. (b) Sequence analysis based on fixing position 7 as a C and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 6 is a G. (c) Sequence analysis based on fixing position 6 as a G and position 7 as a C and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 8 is a C. (d) Sequence analysis based on fixing position 6 as a G, position 7 as a C, and position 8 as a C and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to NNNNNGCC, the 10^{th}-percentile μ is 0.04.

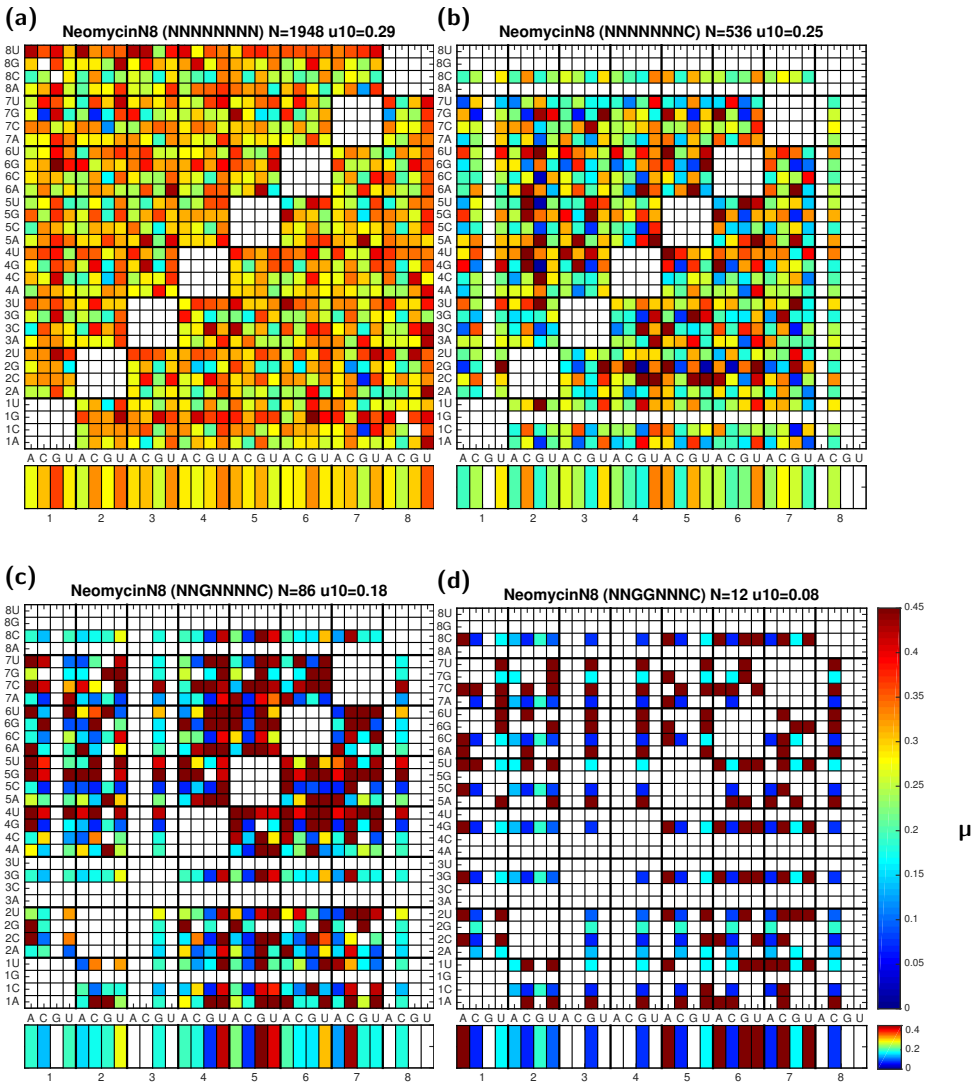
Supplementary fig. 15 Effect of loop I sequence on activity of tetracycline-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem II of a ribozyme containing a tetracycline aptamer on the opposing stem I, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on averaging over all possibilities for the remaining six nucleotides in the loop.

The strongest single nucleotide positive effect on catalytic activity arises when position 1 is a G. (b) Sequence analysis based on fixing position 1 as a G and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 4 is a U. (c) Sequence analysis based on fixing position 1 as a G and position 4 as a U and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 8 is an A. (d) Sequence analysis based on fixing position 1 as a G, position 4 as a U, and position 8 as an A and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to GNNUNNNA, the 10^{th}-percentile μ is 0.02.

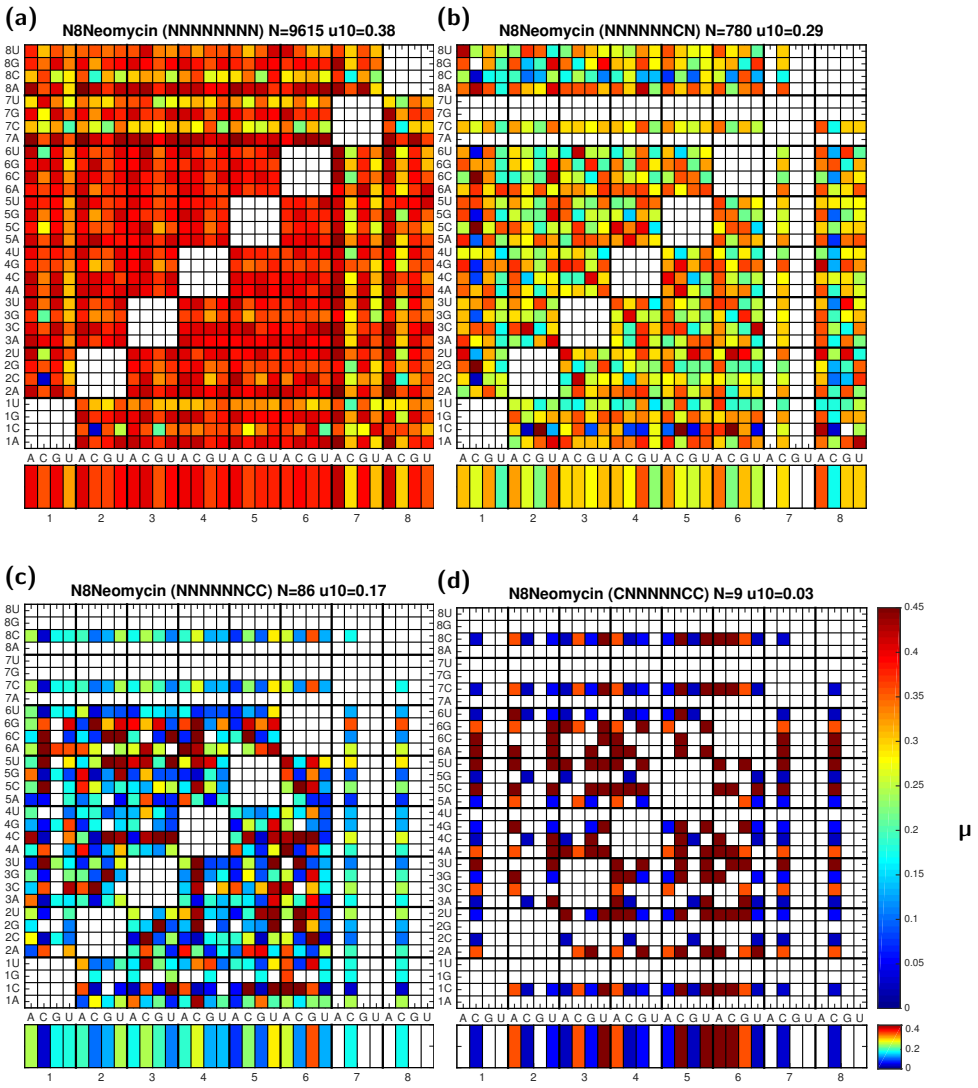
Supplementary fig. 16 Effect of loop II sequence on activity of neomycin-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem II of a ribozyme containing a neomycin aptamer on the opposing stem I, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence analysis based on averaging over all possibilities for the remaining six nucleotides in the loop.

The strongest single nucleotide positive effect on catalytic activity arises when position 8 is a C. (b) Sequence analysis based on fixing position 8 as a C and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 3 is a G. (c) Sequence analysis based on fixing position 3 as a G and position 8 as a C and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 4 is a G. (d) Sequence analysis based on fixing position 3 as a G, position 4 as a G, and position 8 as a C and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to NNGGNNNC, the 10^{th}-percentile μ is 0.08.

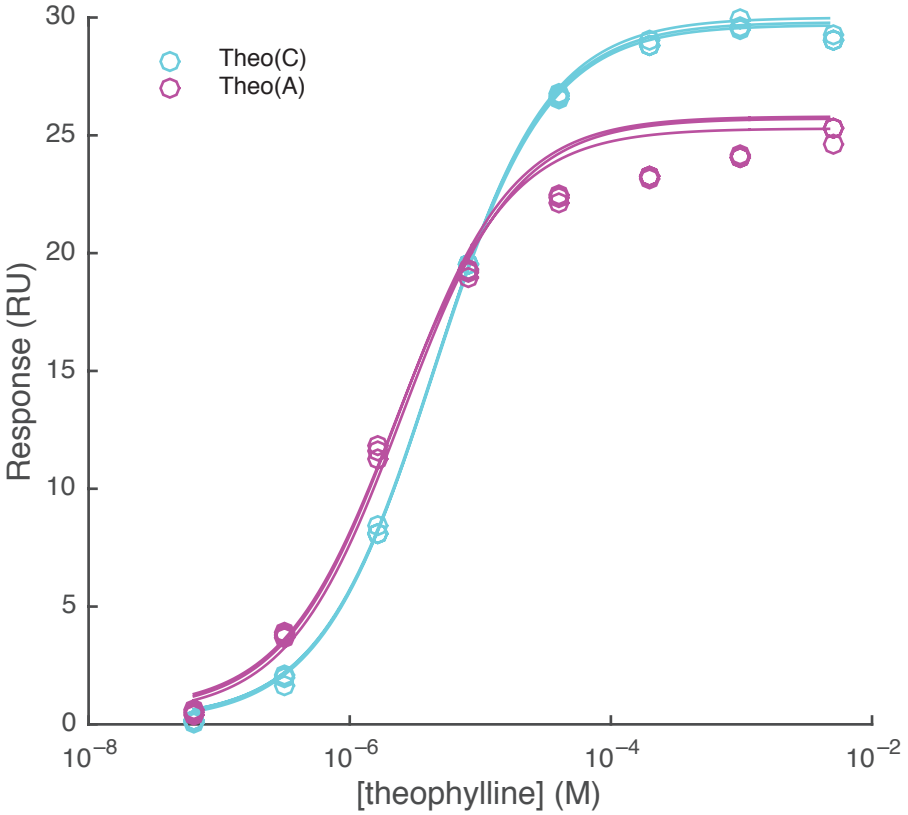
Supplementary fig. 17 Effect of loop I sequence on activity of neomycin-responsive tertiary interaction switches in the absence of target.



Each panel shows the 10th-percentile activity, GFP/mCherry (μ), as a pairwise function of the identity of two of the randomized nucleotides on stem I of a ribozyme containing a neomycin aptamer on the opposing stem II, averaged over all other nucleotide possibilities in the remaining position. The narrow strip along the bottom of each plot shows the effect of each nucleotide by itself. Lower values of μ indicate higher levels of ribozyme catalytic activity. (a) Sequence

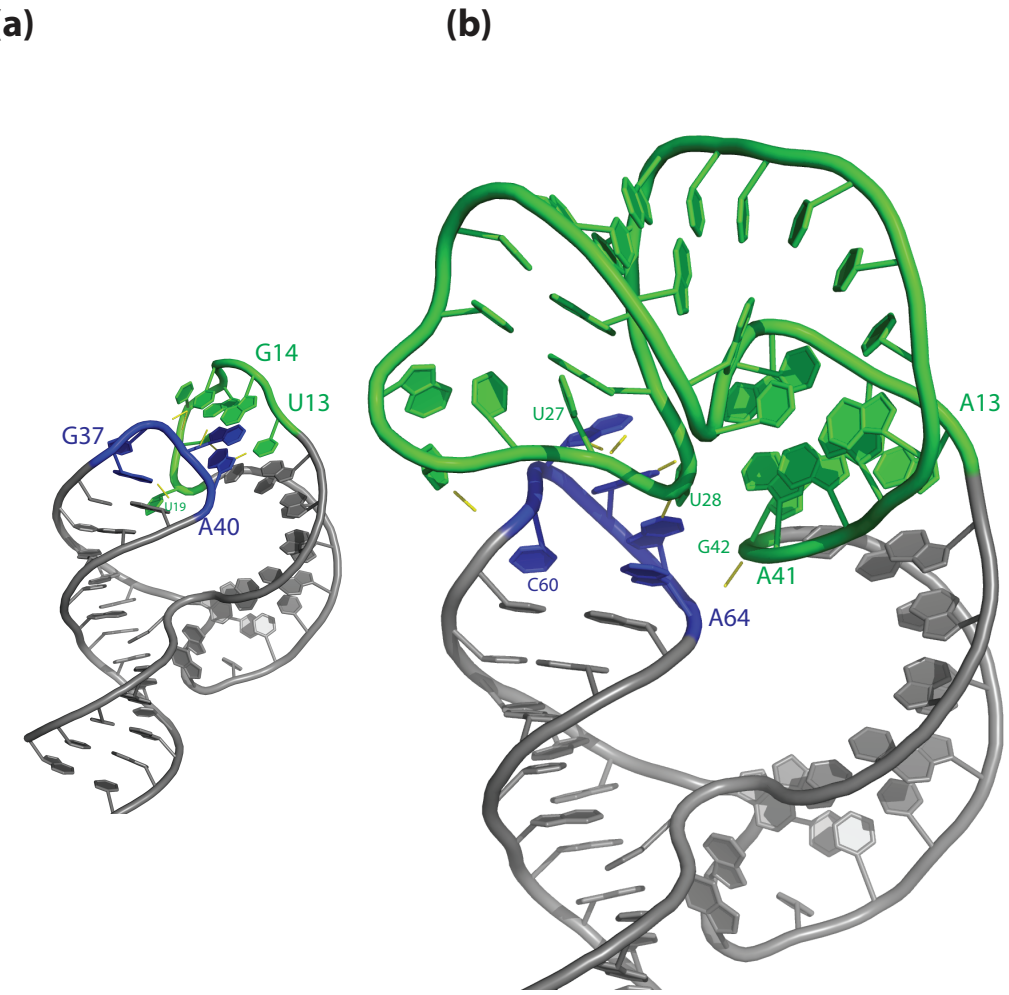
analysis based on averaging over all possibilities for the remaining six nucleotides in the loop. The strongest single nucleotide positive effect on catalytic activity arises when position 7 is a C. (b) Sequence analysis based on fixing position 7 as a C and averaging over the remaining five nucleotides for each entry. The dominant positive effect on catalytic activity is observed when position 8 is a C. (c) Sequence analysis based on fixing position 7 as a C and position 8 as a C and averaging over the remaining four nucleotides for each entry. The strongest effect on catalytic activity is observed when position 1 is a C. (d) Sequence analysis based on fixing position 1 as a C, position 7 as a C, and position 8 as a C and averaging over the remaining three nucleotides for each entry. With the consensus sequence set to CNNNNNCC, the 10^{th}-percentile μ is 0.03.

Supplementary fig. 18 Representative theophylline binding isotherms of the theo-CAG and theo-AAG aptamer variants.



The SPR-based binding assay and analysis was carried out as previously described[9]. SPR responses (RU) due to steady-state theophylline binding were measured in triplicate at the indicated theophylline concentrations at 500 µM $MgCl_2$, 150 mM NaCl, and 10 mM HEPES (pH 7.4). The triplicate data were each fit using MATLAB to a steady-state affinity model to estimate the equilibrium dissociation constants (K_D), and maximum and minimum RU levels; the fitted curves are shown as the solid lines. The replicate data was then pooled to obtain overall K_Ds of theo-CAG (4.4 µM; 95% confidence interval [4.1 µM, 4.8 µM]) and theo-AAG (2.4 µM; 95% confidence interval [1.7 µM, 3.1 µM]).

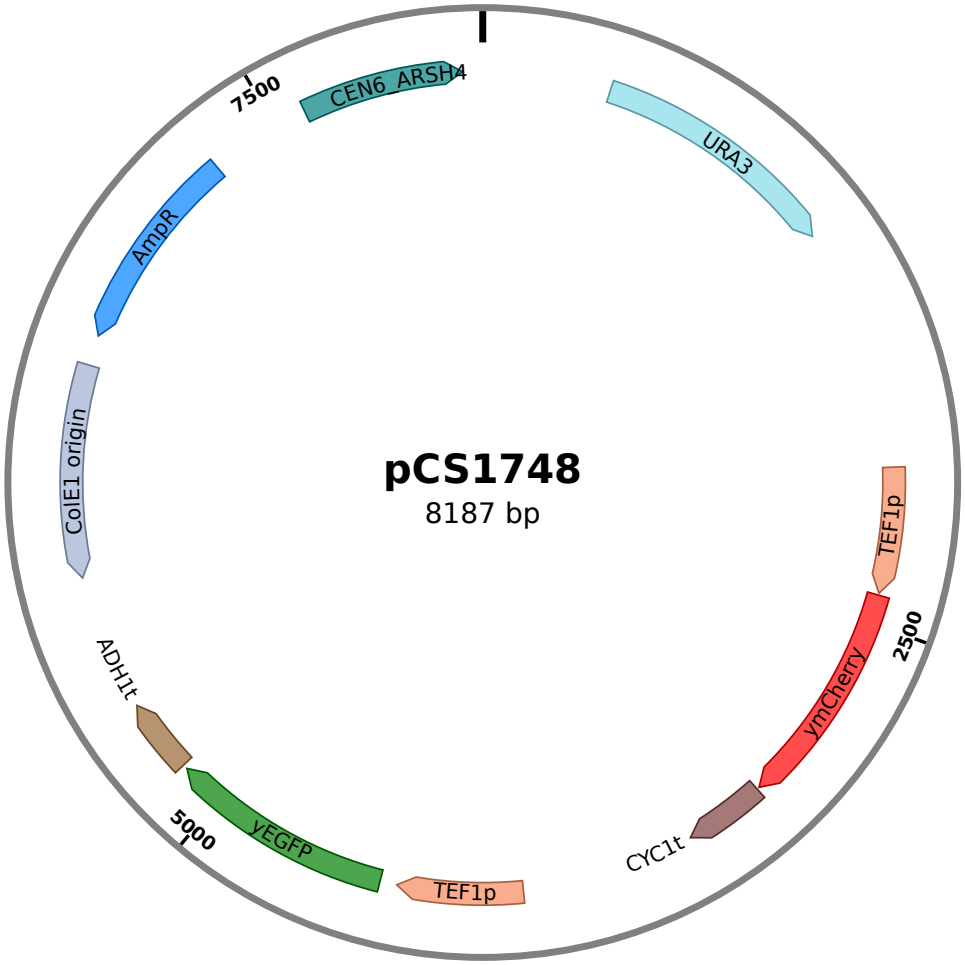
Supplementary fig. 19 Overall three-dimensional structures.



Overall three-dimensional structure of hammerhead ribozyme and ribozyme-based device modeled using Rosetta/FARFAR[10] with full atom-level modeling and rendered using PyMol[11]. (a) Hammerhead ribozyme model derived from 2QUS[12] PDB coordinates. Loop I is shown in green and loop II in blue with polar contacts (hydrogen bonds) between the loops shown as yellow dashes. The G12A mutation used in 2QUS was reverted to a guanine and loop III was replaced with the sequence used in this work. These modified nucleotides were allowed to move freely while the matched residues were fixed during simulation. (b) Theo(A)-CAGAA RNA

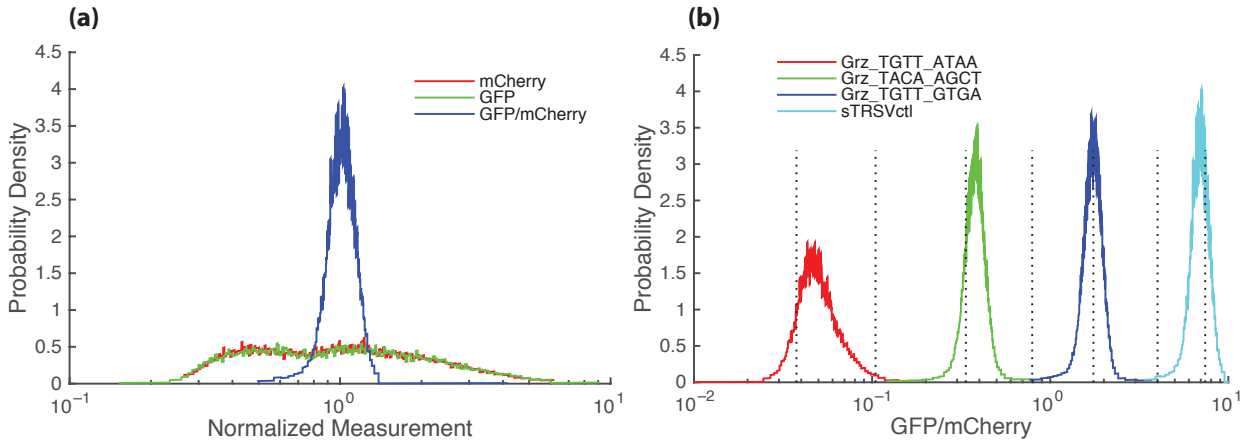
device model. The ribozyme core was derived from 2QUS as in panel (a). The theophylline aptamer structure was based on 1O15[13] with the CAG bulge replaced with AAG and the top loop replaced with the sequence used in this work, and then optimized separately using FARFAR. The full structure of the device was then simulated for the combined structures with the core ribozyme and aptamer coordinates internally fixed with the backbones connected at loop I, but free with respect to each other. The loop II sequence was replaced with CAGAA and also allowed to freely move during simulation. For both panels, 2000 low-resolution structures were generated followed by all-atom optimization of the structures under the Rosetta energy function. The lowest energy ones, based on FARFAR scoring, are shown here in cartoon form.

Supplementary fig. 20 Plasmid Map.



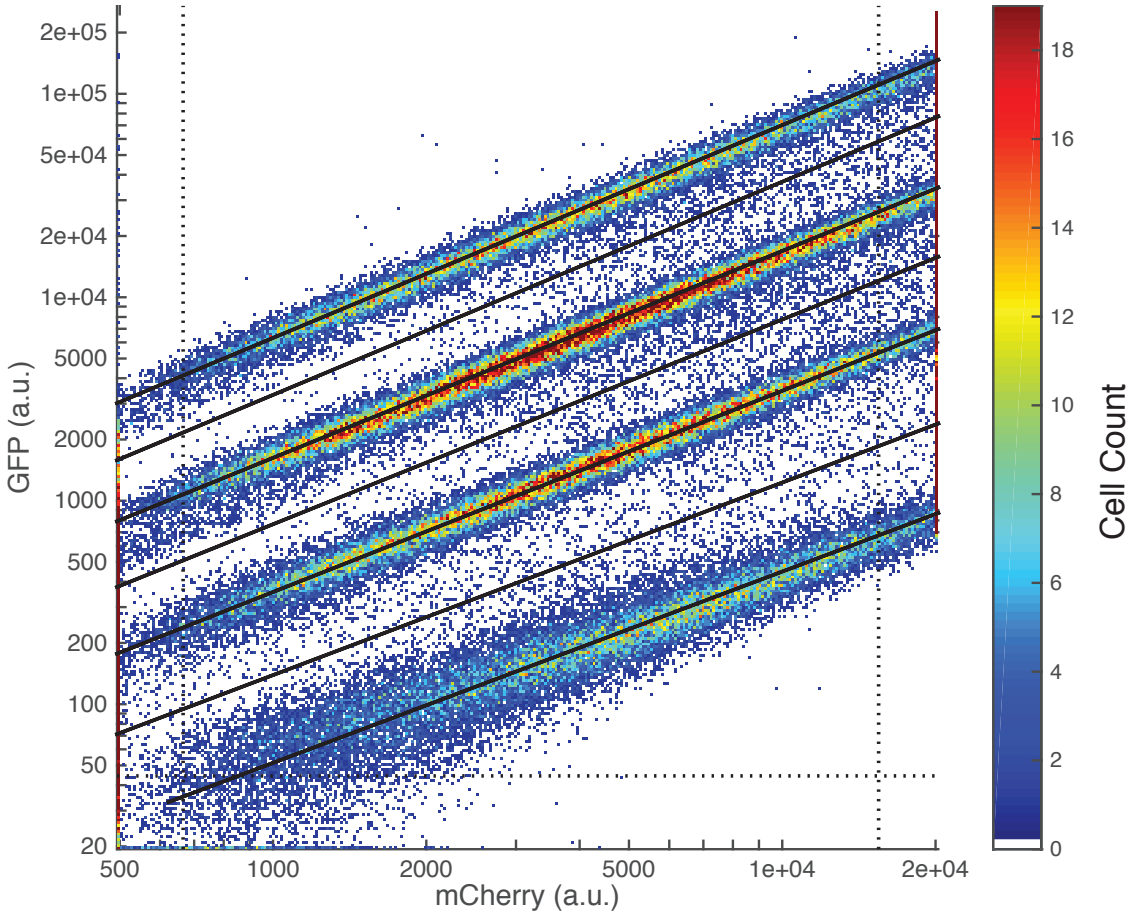
Map of the two-color plasmid used for *in vivo* characterization of all switch, ribozyme, and control sequences. Switch, ribozyme, and control sequences were inserted between the GFP coding region and the terminator.

Supplementary fig. 21 Probability densities from flow cytometry analyses for two-color assay and for binned sorting.



(a) Two-color fluorescence reporter activity distributions observed through flow cytometry analysis. Representative distributions are shown for a cell population harboring a two-color plasmid in which GFP is modulated by one of the graded ribozymes. The GFP and mCherry distributions exhibit much larger spreads individually than the computed GFP/mCherry ratio for the same cell population. (b) GFP/mCherry distribution for cells harboring four graded ribozymes in the two-color construct used to setup the gating for 8-way sorting. The borders of the bins are shown as dashed vertical line.

Supplementary fig. 22 Fluorescent activity for graded ribozymes and gating for 8-way sort.



Gates for the sorting bins were established using the reference set of graded ribozymes, which consists of an equal mixture of cells harboring four different sequences that exhibit GFP/mCherry activities that uniformly span the range of interest on a log scale. The boundary between bins 1&2 was set to follow the first graded ribozyme density with approximately 50% of the cells falling on either side of the gate. Similarly, the boundaries between bins 3&4, 5&6, and 7&8 were specified with reference to the other three graded ribozymes in the reference set. The boundaries between bins 2&3, 4&5, and 6&7 were set between the graded ribozyme peaks, dividing these regions in half. The gating operations were performed on a log-log representation of GFP versus mCherry.

Supplementary table 1. Flow cytometry validation results for switch sequences identified from NGS analyses.

Loop I	Loop II	μ (-tgt)	μ (+tgt)	Activation Ratio
Theo (AAG)	AAAAA	0.056±0.002	0.51±0.07	9.1±1.3
Theo (AAG)	CAGAA	0.158±0.006	1.26±0.09	7.9±0.4
Theo (CAG)	CAGUA	0.165±0.014	1.14±0.08	6.9±0.2
Theo (CAG)	AGGAAA	0.236±0.020	1.60±0.22	7.0±1.5
Theo (CAG)	CAGAGAA	0.159±0.010	0.87±0.11	5.5±0.7
GGAACU	Theo (AAG)	0.164±0.008	0.93±0.04	5.7±0.4
Theo (AAG)	AGAGA	0.107±0.003	1.03±0.07	9.6±0.5
Theo (AAG)	AAAGA	0.088±0.003	1.00±0.08	11.4±0.8
Theo (AAG)	CAAUAA	0.144±0.008	0.98±0.04	6.9±0.5
Theo(CAG)	CAGAUAAA	0.238±0.012	1.23±0.08	5.2±0.2
Theo (CAG)	CACGUGAA	0.232±0.003	1.40±0.20	6.0±0.8
Theo (AAG)	CAUAUAA	0.205±0.006	1.58±0.15	7.8±0.9
Theo (AAG)	CAAGUGAA	0.230±0.007	1.90±0.37	8.2±1.4
Theo (AAG)	CAAUUUAA	0.183±0.011	1.40±0.15	7.6±0.6
Theo (AAG)	CAAUCUAA	0.200±0.011	1.56±0.12	7.8±0.3
Theo (AAG)	CAUGUAAA	0.207±0.010	1.80±0.12	8.8±1.0
ACUUUAA	Theo (AAG)	0.174±0.006	1.27±0.01	7.3±0.2
Tetracycline	AGGUAUGA	0.196±0.006	1.80±0.07	9.1±0.2
GGGGGUGC	Tetracycline	0.154±0.005	1.16±0.04	7.5±0.2
GGGAUUAU	Tetracycline	0.129±0.007	0.92±0.18	7.0±1.0
GGGUAGAU	Tetracycline	0.240±0.012	1.28±0.07	5.3±0.1
AGGGAAAU	Tetracycline	0.173±0.010	1.17±0.10	6.8±0.3
Neomycin	UGUAGCGG	0.376±0.012	2.43±0.11	6.5±0.3
Neomycin	CUGGCCAC	0.406±0.009	2.34±0.07	5.8±0.1
Neomycin	CCGGCCAC	0.316±0.008	1.95±0.10	6.2±0.3
AGGACUAA	Neomycin	0.595±0.024	2.13±0.15	3.6±0.1

The values of μ were determined from flow cytometry analysis of at least 3 independent transformants containing the indicated switch. Error estimates are the standard error of the mean.

Supplementary table 2. Sequences of spacers, ribozymes, aptamers, switches, and primers.

Type	Name	Sequence
RNA device	L2b8[14]	AAACAAACAAAGCUGUCACCGGAUGUGCUUUCCGGUC UGAUGAGUCCGUUGUCCAUACCAGCAUCGUCUUGAUG CCCUUGGCAGGGACGGGACGGAGGACGAAACAGCAAA AAGAAAAAUAAAAA
RNA device	L2b8-a1[14]	AAACAAACAAAGCUGUCACCGGAAUCAAGGUCCGGUC UGAUGAGUCCGUUGUCCAUACCAGCAUCGUCUUGAUG CCCUUGGCAGGGACGGGACGGAGGACGAAACAGCAAA AAGAAAAAUAAAAA
RNA device	L2b8-t47[14]	AAACAAACAAAGCUGUCACCGGAUGUGCUUUCCGGUC UGAUGAGUCCGUUGAGUAUACCAGCAUCGUCUUGAU GCCCUUGGCAGACUGUAUACGGAGGACGAAACAGCAA AAAGAAAAAUAAAAA
Spacer	5’ (W)	AAACAAACAAA
Spacer	3’ (X)	AAAAAGAAAAAUAAAAA
Spacer	5’ SPR (SPR-fwd)	GGGAAACAAACAAAGUUGUUUU
Spacer	3’ SPR (SPR-rev)	UUUGUU
Ribozyme	sTRSV	GCUGU C ACCGGA UGUGCUU UCCGGUCUGAUGA GUCC GUGA GGAC GAA ACAGC
Ribozyme	sTRSVctl	GCUGU C ACCGGA UGUGCUU UCCGGUACGUGAG GUCC GUGA GGAC AGA ACAGC
Aptamer	Theophylline (CAG-variant)	AUACCAGCAUCGUCUUGAUGCCCUUGGCAG
Aptamer	Theophylline (AAG-variant)	AUACCAGCAUCGUCUUGAUGCCCUUGGAAG
Aptamer	Neomycin	GCUUGUCCUUUAAUGGUCC
Aptamer	Tetracycline	AAAACAUACCAGAUUUCGAUCUGGAGAGGUGAAGAA UUCGACCACCU
Library	Loop I Integration	AAACAAACAAA GCUGU C ACCGGA <Aptamer> UCCGGUCUGAUGA GUCC <N3:N8> GGAC GAA ACAGC AAAAAGAAAAAUAAAAA
Library	Loop II Integration	AAACAAACAAA GCUGU C ACCGGA <N3:N8> UCCGGUCUGAUGA GUCC <Aptamer> GGAC GAA ACAGC AAAAAGAAAAAUAAAAA
Primer	T7_W_Primer	AATTTAATACGACTCACTATAGGG AAACAAACAAA GCTGTC ACCGGA
Primer	X_Primer	TTTTTATTTTTCTTTTT GCTGT TTC GTCC
Primer	SPR_fwd_primer	TTCTAATACGACTCACTATAGGG
Primer	SPR_rev_primer	AACAAAGCTGTTTCGTCC

SPR template	L2b8_SPR	TTCTAATACGACTCACTATAGGGGGGAAACAAACAAAGT TGTTTTGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGT CCGTTGTCCATACCAGCATCGTCTTGATGCCCTTGGCAGG GACGGGACGGAGGACGAAACAGCTTTGTT
SPR template	L2b8-a1_SPR	TTCTAATACGACTCACTATAGGGGGGAAACAAACAAA GTTGTTTTGCTGTCACCGGAATCAAGGTCCGGTCTGAT GAGTCCGTTGTCCATACCAGCATCGTCTTGATGCCCTTG GCAGGGACGGGACGGAGGACGAAACAGCTTTGTT
SPR template	L2b8-t47_SPR	TTCTAATACGACTCACTATAGGGGGGAAACAAACAAA GTTGTTTTGCTGTCACCGGATGTGCTTTCCGGTCTGATG AGTCCGTTGAGTATACCAGCATCGTCTTGATGCCCTTG GCAGACTGTATACGGAGGACGAAACAGCTTTGTT
SPR template	Theo(A)-AAAGA_SPR	TTCTAATACGACTCACTATAGGGAAACAAACAAAGTTG TTTTGCTGTCACCGGAATACCAGCATCGTCTTGATGCCC TTGGAAGTCCGGTCTGATGAGTCCAAAAAGGACGAAA CAGCTTTGTT
SPR template	Theo(A)-AAAAA_SPR	TTCTAATACGACTCACTATAGGGAAACAAACAAAGTTG TTTTGCTGTCACCGGAATACCAGCATCGTCTTGATGCCC TTGGAAGTCCGGTCTGATGAGTCCAAAAAGGACGAAA CAGCTTTGTT
SPR template	Theo(A)-CAGAA_SPR	TTCTAATACGACTCACTATAGGGAAACAAACAAAGTTG TTTTGCTGTCACCGGAATACCAGCATCGTCTTGATGCCC TTGGAAGTCCGGTCTGATGAGTCCAAAAAGGACGAAA CAGCTTTGTT

Primers are DNA oligonucleotide sequences and SPR templates are dsDNA for transcribing into RNA for the SPR-based experiments[2]. All other sequences are RNA.

Supplementary table 3. Dose-response parameters from flow cytometry for theophylline-responsive RNA devices.

RNA Device	EC_{50} (mM)	μ (0 mM theophylline)	μ (5 mM theophylline)	Activation Ratio
L2b8	4.9 [4.3-5.6]	0.67 [0.63-0.72]	1.32 [1.22-1.43]	2.0 [1.8-2.2]
L2b8-a1	10.4 [7.3-13.5]	0.19 [0.18-0.21]	0.28 [0.25-0.32]	1.5 [1.3-1.7]
L2b8-t47	5.7 [4.9-6.4]	0.35 [0.32-0.39]	0.92 [0.85-1.01]	2.6 [2.3-3.0]
Theo(A)-AAAAA	3.8 [3.6-4.1]	0.09 [0.08-0.10]	0.68 [0.60-0.77]	7.3 [6.2-8.6]
Theo(A)-AAAGA	1.9 [1.8-2.1]	0.16 [0.14-0.18]	1.53 [1.30-1.81]	9.8 [7.9-12.1]
Theo(A)-CAGAA	1.5 [1.3-1.6]	0.25 [0.24-0.26]	2.29 [1.66-3.16]	9.1 [6.6-12.6]

Median GFP/mCherry ratios were computed for at least triplicate independent cultures assayed at theophylline concentrations of 0, 0.16, 0.31, 0.62, 1.25, 2.5, 5, 10, and 20 mM. The data were fit to a 4-parameter logistic model with the Hill slope fixed at 1.0; the 80% confidence interval for the EC_{50} fit is shown. A range of one standard deviation over the replicates is shown for the mean activity levels and activity ratio.

Supplementary table 4. Activity data and sequences for the graded ribozyme set.

Loop I	Loop II	μ (Validation)	N	Note
TGTGCTT	GTGA	0.067 [0.063, 0.071]	5	Native sTRSV
TATG	AGAA	0.067 [0.065, 0.070]	2	
TGTT	ACTA	0.116 [0.104, 0.127]	5	
TTGT	CATA	0.176 [0.157, 0.189]	5	
GGCT	AGCT	0.273 [0.247, 0.298]	5	
TGCA	CGTT	0.497 [0.454, 0.533]	5	
CAGG	AGTT	0.860 [0.833, 0.897]	5	
TGCA	CGCG	1.41 [1.27, 1.54]	5	
TGCT	GTGA	5.16 [4.97, 5.35]	2	
TGTGCTT	GTGA	4.91 [4.70, 5.35]	5	sTRSVctl (scrambled catalytic core not shown)

The set of graded ribozymes span the range of expression control over GFP with GFP/mCherry ratio (μ) varying 77-fold; from close to that of sTRSV (wild-type ribozyme) to higher than that of sTRSVctl (inactive ribozyme). The μ values reported are obtained through flow cytometry analysis of the individual sequences with the number of replicates of each sequence indicated (N). Values in brackets are the range of values observed over the replicates.

Supplementary table 5. GFP/mCherry values for consensus sequences identified for each aptamer tested.

Aptamer	Loop integration	Median μ of library	10^{th}-percentile μ of library	Consensus	10^{th}-percentile μ of consensus
Theophylline (AAG)	I	0.58	0.24	CANNNNAN	0.06
Theophylline (CAG)	I	0.66	0.22	NANNNNNAA	0.04
Theophylline (AAG)	II	0.45	0.24	NGNANANN	0.16
Theophylline (CAG)	II	0.57	0.24	ACUNNNNN	0.09
Neomycin	I	0.54	0.29	NNNNNNNC	0.25
Neomycin	II	0.52	0.38	NNNNNNCN	0.29
Tetracycline	I	0.36	0.25	NNNNNNCN	0.13
Tetracycline	II	0.33	0.06	GNNUNNNA	0.02

Median and 10^{th}-percentile GFP/mCherry (μ) is reported over all measured eight-nucleotide randomized library sequences for the opposing loop. Consensus sequences represent the lowest median μ identified during successive consensus minimization analyses.

Supplementary table 6. Cells sorted and fraction collected for the prescreen sort for each of the six populations.

Replicate	Aptamer	Cells Sorted	Fraction Retained	Library Coverage
1	Theophylline	603,000	10.3%	33.5
1	Neomycin	302,000	11.0%	15.7
1	Tetracycline	327,000	9.2%	20.3
2	Theophylline	814,000	10.2%	45.7
2	Neomycin	602,000	8.8%	39.2
2	Tetracycline	308,000	9.0%	19.6

The prescreen sort was designed to retain ~10% of the cells exhibiting the lowest GFP/mCherry ratio. In each case, an average of at least 15 cells per library entry were sorted, of which ~10% were retained for subsequent steps of the method.

Supplementary table 7. Cells collected for the main sort for each of the four assays.

Bin	Replicate/Condition			
	1-	**1+**	**2-**	**2+**
1	26,675	17,233	17,172	9,307
2	32,995	16,493	25,744	10,170
3	146,144	35,663	111,265	25,678
4	361,365	94,205	248,918	69,424
5	1,438,899	384,375	795,528	265,000
6	4,335,894	2,673,765	2,962,704	1,914,932
7	1,230,248	3,714,895	2,729,291	3,290,934
8	66,881	134,146	186,173	340,159
Total	7,639,101	7,070,775	7,076,795	5,925,604

Each entry is the number of cells sorted into one of the eight bins. The sorts were performed on two biological replicates, each with and without target.

Supplementary table 8. NGS reads per unique plasmid molecule.

Sample	Bin							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1-	1.002	1.002	1.000	1.000	1.009	1.009	1.023	1.000
1+	1.040	1.041	1.006	1.226	1.011	1.009	1.019	1.011
2-	1.037	1.043	1.008	1.019	1.021	1.022	1.161	1.019
2+	1.039	1.114	1.071	1.007	1.249	1.010	1.009	1.004

We attempted to ensure that the plasmid preps of each sample had a sufficient number of molecules such that each molecule would usually result in at most one NGS read, to reduce any statistical biases. Values were obtained based on the statistics of the diversity control spiked into each plasmid preparation prior to PCR amplification and subsequent sequencing.

Supplementary table 9. NGS reads per molecule.

Sample	Bin							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1-	0.13	0.24	0.29	0.27	0.37	0.47	0.38	0.51
1+	0.08	0.07	0.14	0.28	0.30	0.47	0.45	0.53
2-	0.11	0.22	0.51	0.42	0.52	0.76	0.80	1.07
2+	0.05	0.04	0.11	0.43	0.59	0.46	0.49	0.90

Each NGS read corresponds to the number of cells shown. Values were obtained based on the total read counts divided by the number of cells sorted.

Supplementary Information - References

1. Perreault, J. *et al.* Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput. Biol.* **7,** e1002031 (2011).

2. Kennedy, A. B., Liang, J. C. & Smolke, C. D. A versatile cis-blocking and trans-activation strategy for ribozyme characterization. *Nucleic Acids Res.* **41,** e41 (2013).

3. Leontis, N. B., Stombaugh, J. & Westhof, E. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **30,** 3497–3531 (2002).

4. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28,** 235–242 (2000).

5. Chi, Y. I. *et al.* Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS Biol.* **6,** 2060–2068 (2008).

6. Duchardt-Ferner, E. *et al.* Highly modular structure and ligand binding by conformational capture in a minimalistic riboswitch. *Angew. Chemie - Int. Ed.* **49,** 6216–6219 (2010).

7. Xiao, H., Edwards, T. E. & Ferré-D'Amaré, A. R. Structural Basis for Specific, High-Affinity Tetracycline Binding by an In Vitro Evolved Aptamer and Artificial Riboswitch. *Chem. Biol.* **15,** 1125–1137 (2008).

8. Sarver, M., Zirbel, C. L., Stombaugh, J., Mokdad, A. & Leontis, N. B. FR3D: Finding local and composite recurrent structural motifs in RNA 3D structures. *J. Math. Biol.* **56,** 215–252 (2008).

9. Chang, A. L., McKeague, M., Liang, J. C. & Smolke, C. D. Kinetic and equilibrium binding characterization of aptamers to small molecules using a label-free, sensitive, and scalable platform. *Anal. Chem.* **86,** 3273–8 (2014).

10. Cheng, C. Y., Chou, F.-C. & Das, R. Modeling Complex RNA Tertiary Folds with Rosetta. *Methods Enzymol.* **553,** 35–64 (2015).

11. Schrödinger, L. The PyMOL Molecular Graphics System. (2015).

12. Chi, Y. I. *et al.* Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS Biol.* **6,** 2060–2068 (2008).

13. Clore, G. M. & Kuszewski, J. Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J. Am. Chem. Soc.* **125,** 1518–1525 (2003).

14. Liang, J. C., Chang, A. L., Kennedy, A. B. & Smolke, C. D. A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res.* **40,** e154 (2012).