SUPPORTING INFORMATION

Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution

Grigory S. Filonov, Jared D. Moon, Nina Svensen, Samie R. Jaffrey

SUPPLEMENTARY FIGURES

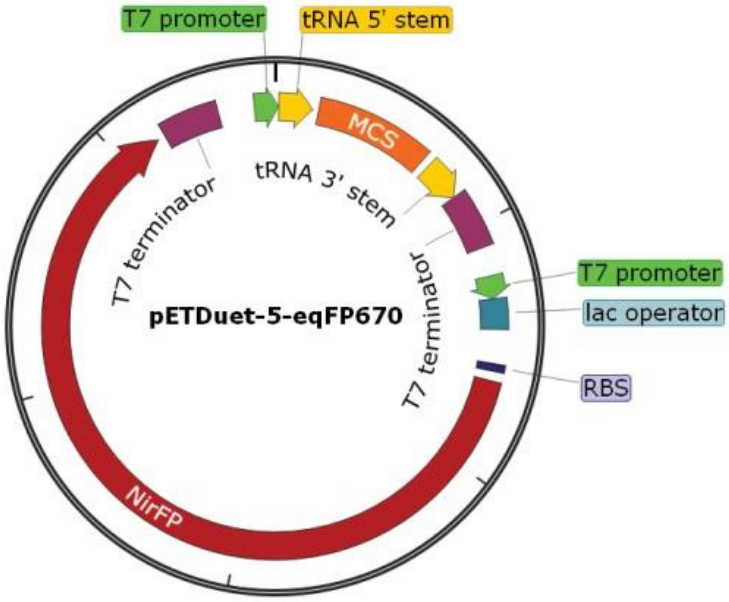


Figure S1. **Simplified scheme of the plasmid used for expressing RNA libraries in bacteria.** Only relevant parts are presented. pETDuet-1 plasmid was extensively modified to be suitable for RNA library expression in *E. coli* (see Methods). The lac operator after the first T7 promoter was removed. Additionally a T7 terminator was introduced into the plasmid at the end of the first expression cassette to ensure proper aptamer transcription termination. The plasmid contains a separate promoter for the transcription of eqFP670, a far-red fluorescent protein.[1] This protein's fluorescence was used to normalize for expression level during FACS and for expression level and colony size during colony imaging on agar plates.

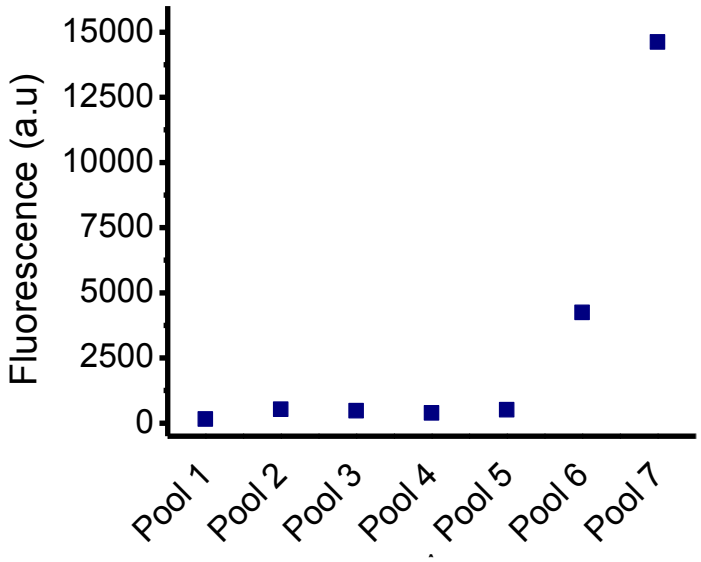


Figure S2. Monitoring of the fluorescence of RNA pools after each round of SELEX guides the transition from SELEX to FACS-based screening of aptamers in *E. coli*. In each experiment, 10 µM of the RNA pool after each round of SELEX was mixed with 20 µM of DFHBI and fluorescence signal was measured (ex = 460 nm, em = 500 nm). The fluorescence was first detectable in the RNA pool after the sixth round of SELEX.

a

```
                           FIXED                                        FIXED                                          FIXED
Library   GAGACGCAACTGAATGAANNNNNNNNNNNNNNNNNNNNNNNNNN      CTGCTTCGGCAGNNNNNNNNNNNNNNNNNNNNNNNNNNTCCGTAACTAGTCGCGTCAC
29-1      GAGACGCAACTGAATGAAATGGTGAAGGAGACGGTCGGGTCCAGGCACAAAAATGTTGCC         TGTTGAGTAGAGTGTGGGCTCCGTAACTAGTCGCGTCAC
29-2      GAGACGCAACTGAATGAAATGGTGAAGGAGACGGTCGGGTCCAGGCACAAAAAAGTTTCC         TGTTGAGTAGAGTGTGGGCTCCGTAACTAGTCGCGTCAC
29-3      GAGACGCAACTGAATGAAATGGTGAAGGAGACGGTCGGGTCCAGGCACAAAAACGTTTCC         TGTCGAGTAGAGTGTGGGCTCCGTAACTAGTCGCGTCAC
Spinach   GAGACGCAACTGAATGAAATGGTGAAGGA      CGGGTCCAGGTGTGGCTGCTTCGGCAGTGCAGCTTGTTGAGTAGAGTGTGAGCTCCGTAACTAGTCGCGTCAC
30-1      GAGACGCAACTGAATGAATATAGTCTTGGAGAGAAGGGCGGTCT      CTGCTTCGGCAGTAGGGGGCGAGACGGTATTGAAAGGTTCCGTAACTAGTCGCGTCAC
31-1      GAGACGCAACTGAATGACTGGAAGGAGGTACGGTTCAGTGTCT       CTGCTTCGGCAGAGGGAGAGCGCAAGATCAGGAGGTGATCCGTAACTAGTCGCGTCAC
```

b

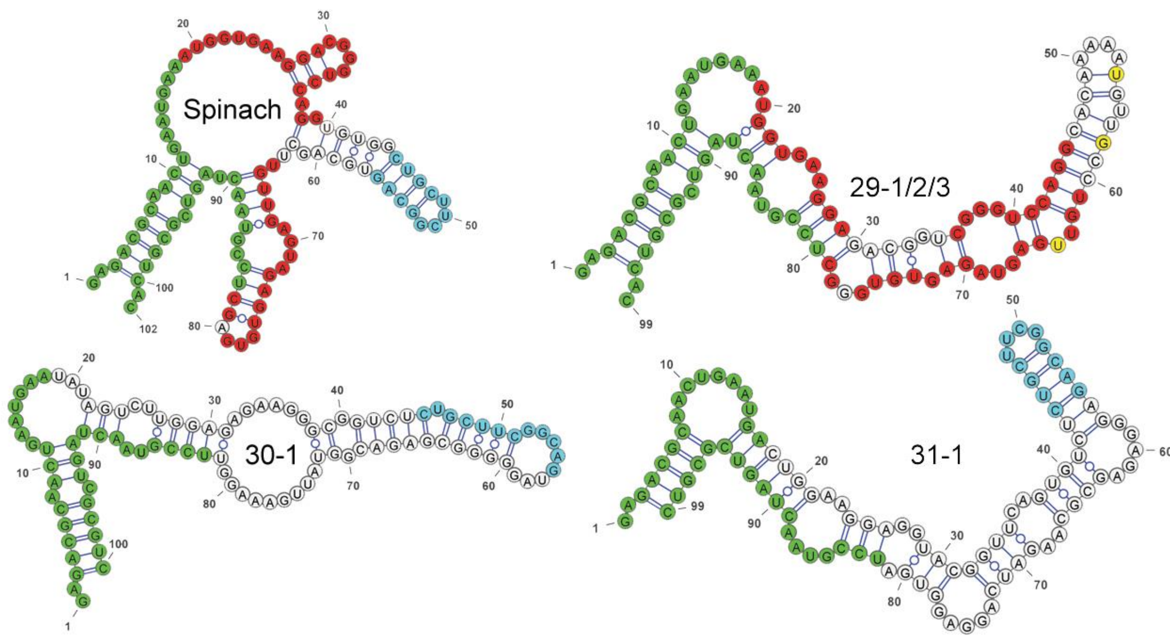


Figure S3. SELEX-FACS identifies diverse "light up" aptamers. (a). Alignment of three mutants from 29-n family (29-1, 29-2, 29-3), Spinach and two unrelated sequences which were also recovered by FACS and which also exhibited fluorescence activation (30-1, 31-1). Above is the sequence of the original library used in SELEX. Green indicates constant regions used for PCR amplification. These regions become part of all the aptamers made during *in vitro* transcription. The blue indicates a fixed region which is predicted to form a stable tetraloop and which was previously shown to facilitate the formation of structured aptamers during SELEX.[2] This region was lost in all the 29-n clones. Red indicates identical regions in Spinach and 29-n aptamers. Yellow indicates nucleotides which were variable in the three 29 family aptamers. Neither 30-1 and 31-1 show sequence similarity with 29-1 nor Spinach (b). All structures shown were predicted using the mFold online software[3]. Folds with the minimal energy were chosen for presentation. The mFold-predicted structure of Spinach was taken from Paige *et al.*, Science, 2011 (ref. 4).

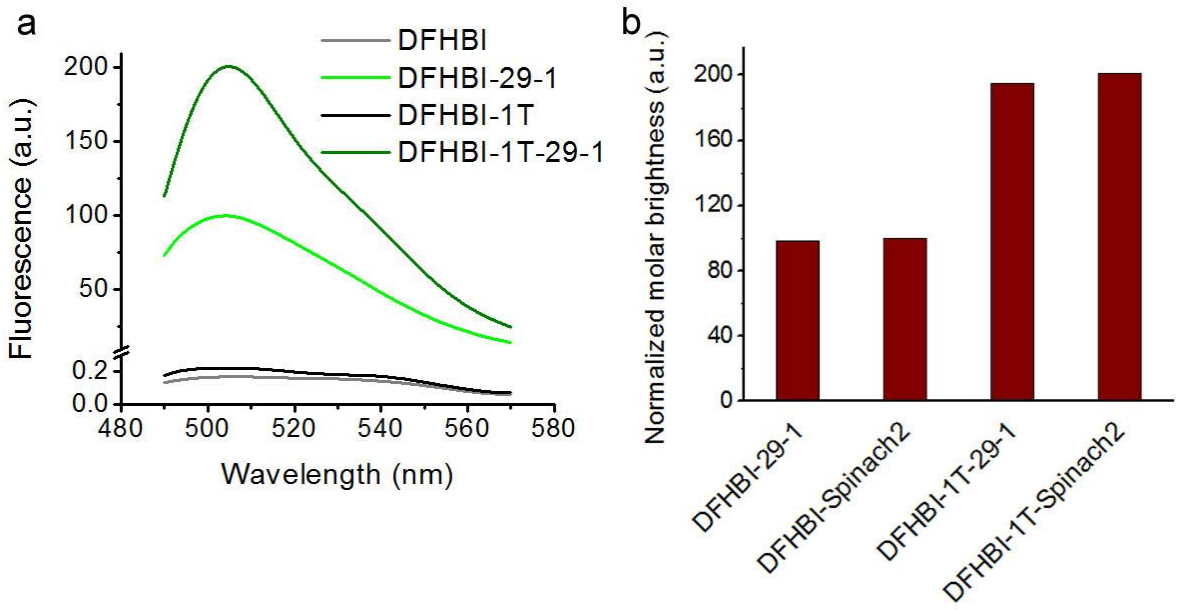


Figure S4. Characterization of the fluorescence properties of 29-1.

(a) We sought to biochemically confirm that 29-1 induces the fluorescence of DFHBI and DFHBI-1T *in vitro*. In these experiments, we incubated 2 µM DFHBI or DFHBI-1T with 20 µM 29-1. We typically use excess RNA to ensure that all the DFHBI is bound by fluorophore. In these experiments 29-1 induces a ~800-fold increase of the fluorophore fluorescence above its background fluorescence.

(b) The 29-1-DFHBI and 29-1-DFHBI-1T complexes exhibit similar molar brightness as Spinach2. To calculate molar brightness, we incubated 2 µM DFHBI with 20 µM 29-1 or Spinach2 to form 2 µM RNA-DFHBI complex. The fluorescence was measured at the respective excitation and emission maxima and was essentially identical. RNA-DFHBI-1T complexes are brighter than the complexes with DFHBI because of the higher extinction coefficient of DFHBI-1T.[5]

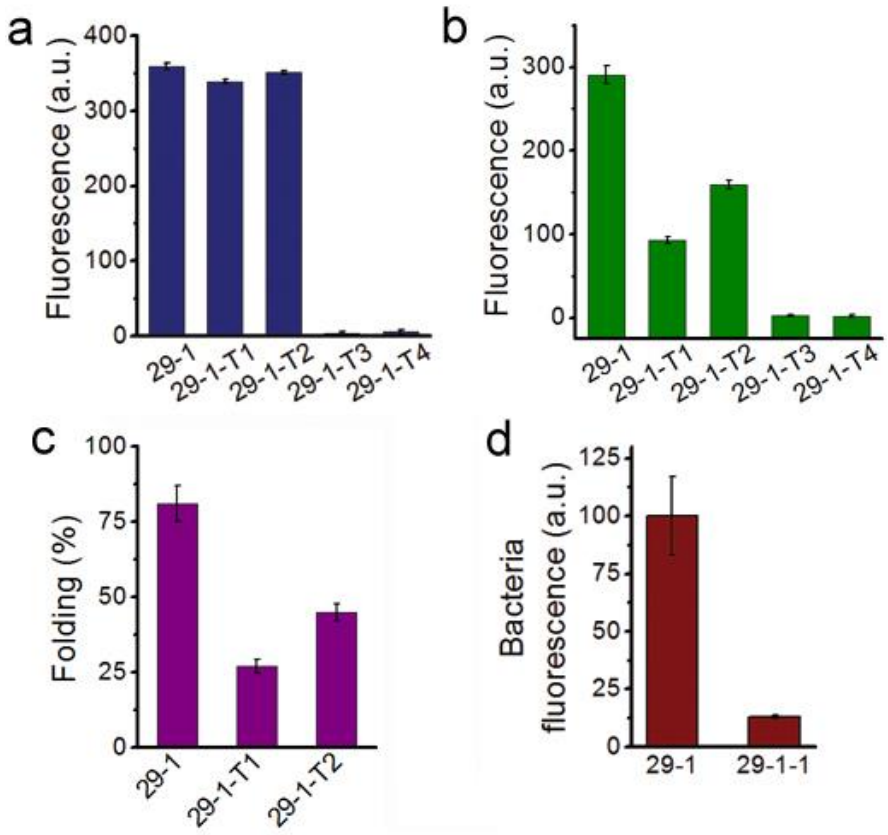


Figure S5. The *in vitro* and *in vivo* performance of the 29-1 truncation mutants.

(a) Fluorescence measurements of the different 29-1 truncation mutants. In these experiments, the fluorescence was measured using 20 µM of the indicated aptamer and 2 µM DFHBI-1T. Thus, the fluorescence shown is the fluorescence of the indicated aptamer-DFHBI-1T complex at 2 µM. This experiment identifies the minimal domain of 29-1 necessary to induce the fluorescence of DFHBI-1T. Additionally, this experiment shows that those truncations mutants which remained fluorescent have almost identical molar brightness.

(b) Fluorescence measurements using an "excess fluorophore" conditions. In this experiment, fluorescence was measured using 2 µM of the indicated aptamer and 20 µM DFHBI-1T. The fluorescence that is seen is a reflection of the proportion of the 2 µM of the aptamer that folds properly. This experiment shows that 29-1-T1 and 29-1-T2 are not as well folded as 29-1. The percent folded is directly measured in panel c.

(c) Folding was calculated as a ratio of the values from panel b to values form the panel a. This folding assay was originally described in Strack *et al.*, Nature Methods, 2013 (ref. 6). These data confirm that the difference between 29-1 and 29-1-T1 and 29-1-T2 is that the truncation mutants do not fold as well.

(d) Performance of 29-1-T2 in *E. coli*. We next wanted to understand if 29-1-T2, which does not fold as well as 29-1, also exhibits poor performance in cells. 29-1-T2 was chosen since it is both the shortest functional mutant, and it shows the highest fluorescence relative to 29-1. To assess

the performance of 29-1-T2 *in vivo* we expressed it in bacterial cells and measured the fluorescent signal of bacterial colonies growing on an agar dish supplemented with 40 μM DFHBI-1T. 29-1-expressing colonies and empty plasmid-transformed colonies were used as the positive and negative control respectively. The latter was used as a background signal and subtracted. Even though both 29-1 and 29-1-T2 are fused to the tRNA scaffold which should facilitate their folding, their performance is very different. This indicates that shorter sequence of 29-1-T2 impairs its function in cells, presumably due to its poor folding. Error bars for all panels indicate standard deviations (n=3).

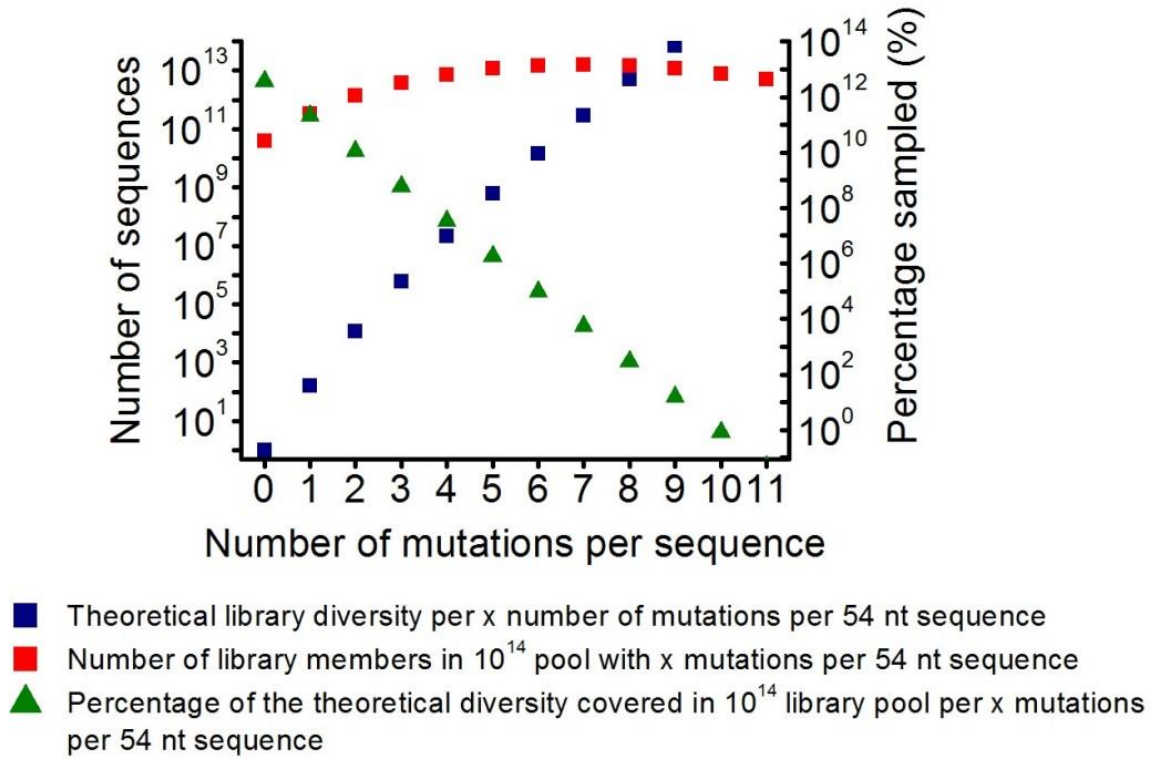


Figure S6. **Theoretical calculations guide the design of the doped library.** The doped library is synthesized as a mixture of ssDNA so that each member resembles the parental sequence. The degree of similarity of the library members to the parental sequence depends on the doping (or mutagenesis) level. The doping level indicates the probability that a nucleotide in the parental sequence is converted to one of the other three nucleotides. Technically it is achieved by using not one nucleoside phosphoramidite during the library synthesis but a mixture of nucleoside phosphoramidites. A doping level of 14% is achieved by using 86% of the parental nucleoside phosphoramidite and 14% of a mixture of the other three phosphoramidites (see Methods for more technical details). Combinatorics equations allow predicting the composition of this library.[7]

The graph presents theoretical calculations for a library with a 54-nt long mutagenized region doped at 14% mutagenesis level.

Blue dots are the number of all theoretically possible nucleotide sequences that are different from the parental sequence by the number of mutations indicated on the *x* axis. These numbers (left *y* axis) do not depend on the doping level and are governed only by the mutagenized sequence length. As can be seen, a library that contains all possible sequences different from the parent by 9 mutations approaches the maximal possible library size that is feasible to work with in a SELEX experiment (~10^{15} library members).

Red dots are the predicted actual number of library members that have x mutations per sequence in a library with 10^{14} total members. These numbers (left *y* axis) follow a normal distribution with the peak value governed by both doping level and mutagenized sequence length (14% and 54 nt, respectively, in this case).

Finally, the percentage of the sampled theoretical diversity is shown in green. These numbers (right *y* axis) shows how well the library covers all the possible diversity of the mutants with 1, 2, 3, etc., mutations per parental sequence. 100% coverage or more ensures that all the possible sequences with x number of mutations are present in the library.

The plot demonstrates that the optimal doping level for the 10^{14} sequences library with 54 nt mutagenized region is around 14%. In this library all possible sequences with 1, 2, 3, 4, 5, 6, 7 and 8 mutations should be present. Decreasing the doping level will result in poor sampling of the sequences with 7 and 8 mutations and in unnecessary oversampling of the library members with lower number of mutations. Increasing the doping level will not substantially improve sampling of the library members with more than eight mutations due to combinatorial explosion of the theoretical diversity of these sequences. Further increase of the doping percentage will also result in poor sampling of mutants with 7 and 8 substitutions.

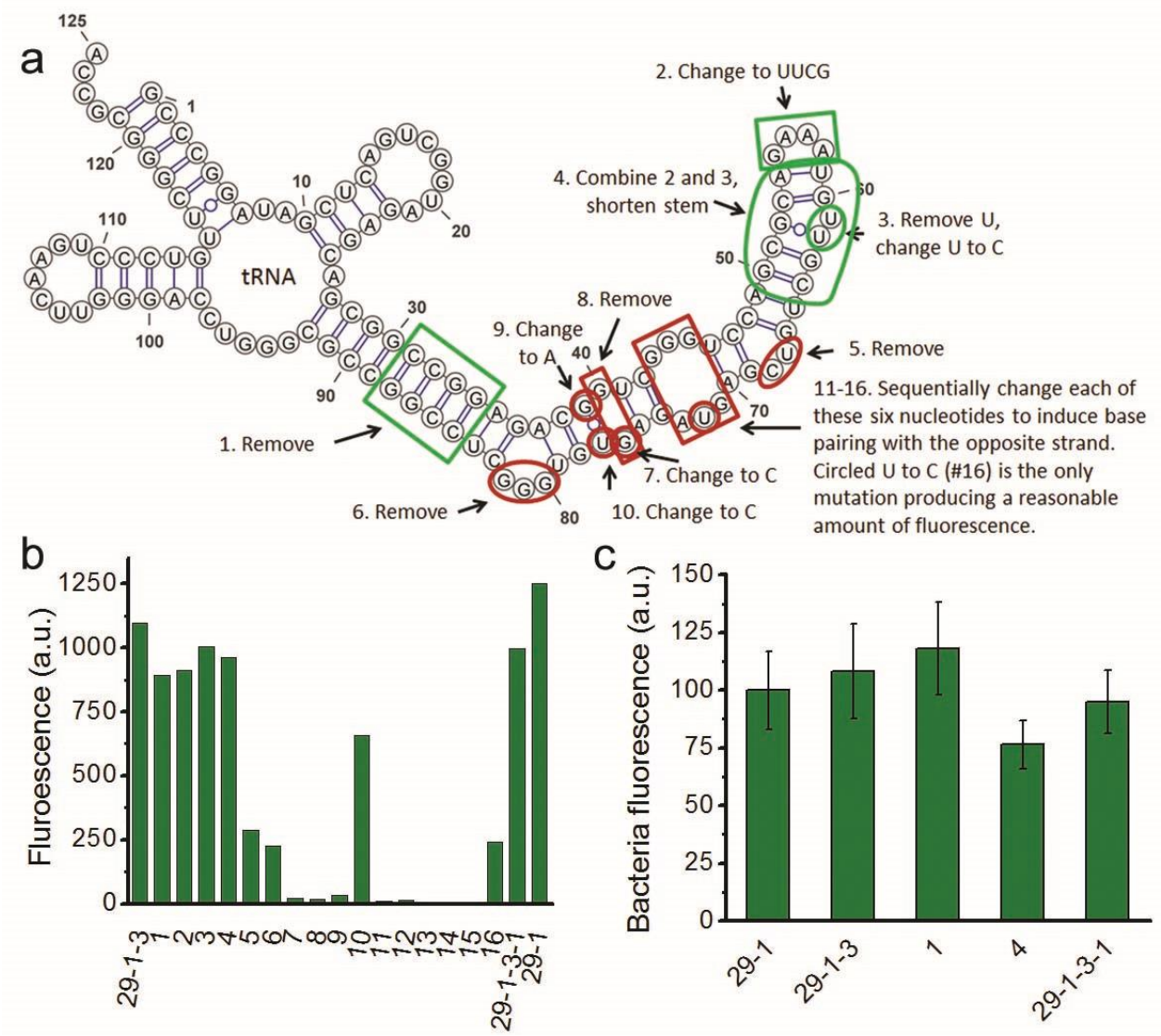


Figure S7. Characterization of the structure of 29-1-3.

(a) Mutagenesis of 29-1-3 validates predictions of mutation intolerant domains made by comparing aptamers produced during directed evolution. Shown is the mFold prediction of the secondary structure of 29-1-3. The tested regions are indicated. The red color indicates the changes that reduced or blocked fluorescence. Green indicates fluorescence-preserving mutations.

(b) *In vitro* brightness of the mutants of 29-1-3 shown in panel a. The same numbering system used in panel a is used in b. For comparison, 29-1-3 and 29-1 are included. The 29-1-3-1 aptamer (which was designated Broccoli) was engineered as a combination of mutations and deletions 1 and 4 from the panel a. In these experiments, we *in vitro* transcribed RNA from the mutant DNA templates and measured their fluorescence on a plate reader in complex with excess of DFHBI-1T. All RNAs contained the tRNA scaffold.

(c) Measurement of aptamer fluorescence in *E. coli*. The fluorescent of the best aptamers in b were measured in *E. coli* colonies grown on a DFHBI-1T-supplemented LB-agar plate. 29-1-3-1 shows almost as strong a signal as the parental 29-1 aptamer while being ~50 nt shorter. All aptamers tested were expressed in the tRNA scaffold. Error bars indicate standard deviations (n=3).

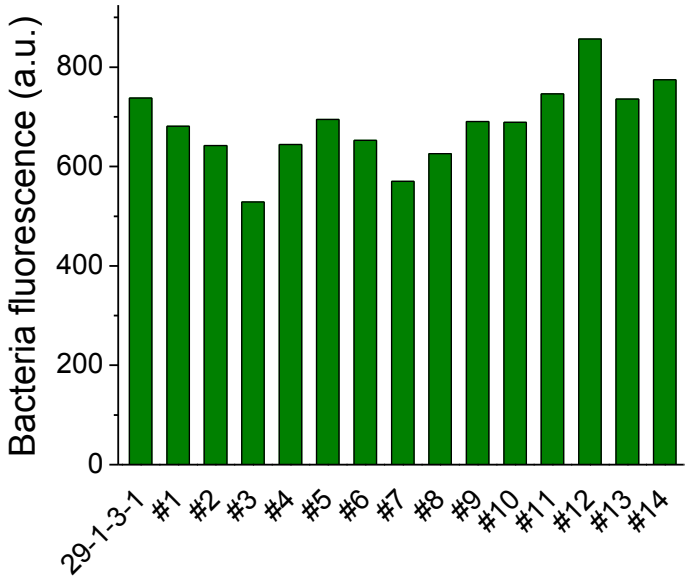


Figure S8. An additional round of directed evolution does not improve the *in vivo* brightness of 29-1-3-1. 29-1-3-1 was selected for further directed evolution. A library was made using the doping strategy described in the text and Methods. However, after FACS, the brightest selected colonies exhibited similar or only slightly (~15%) higher fluorescence intensity compared to the parental 29-1-3-1 when measured in bacterial colonies. Thus, 29-1-3-1 was considered the "final" aptamer and was designated Broccoli.

a. tBroccoli

GCCCGGATAGCTCAGTCGGTAGAGCAGCGGAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTA
GAGTGTGGGCTCCGCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCA

b. tRNA-Broccoli-c-diGMP sensor

GCCCGGATAGCTCAGTCGGTAGAGCAGCGGAGACGGTCGGGTACGCACAGGGCAAACCATTCGAA
AGAGTGGGACGCAAAGCCTCCGGCCTAAACCAGAAGACATGGTAGGTAGCGGGGTTACCGATAGTA
GAGTGTGGGCTCCGCGGGTCCAGGGTTCAAGTCCCTGTTCGGGCGCCA

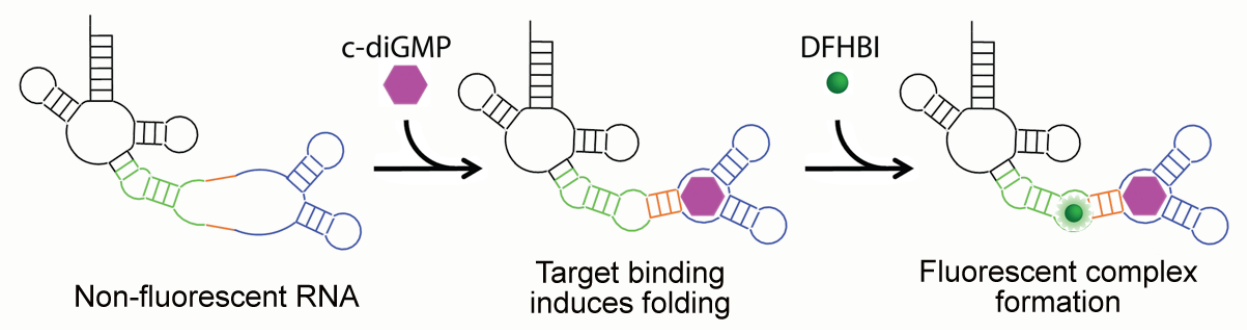
c. tdBroccoli

GCCCGGATAGCTCAGTCGGTAGAGCAGCGGAGACGGTCGGGTCCATCTGAGACGGTCGGGTCCAG
ATATTCGTATCTGTCGAGTAGAGTGTGGGCTCAGATGTCGAGTAGAGTGTGGGCTCCGCGGGTCCA
GGGTTCAAGTCCCTGTTCGGGCGCCA

Figure S9. Sequences of Broccoli and its modifications.

(a) Sequence of tBroccoli. (b) Sequence of Broccoli-based c-diGMP sensor in the tRNA scaffold. (c) Sequence of tdBroccoli. Green – Broccoli sequence. Orange – transducer domain (in sensor) or connector domain and terminal stem-loop (in monomer and dimer). Blue – c-diGMP sensor sequence. tRNA scaffold is not colored.

a.



b.

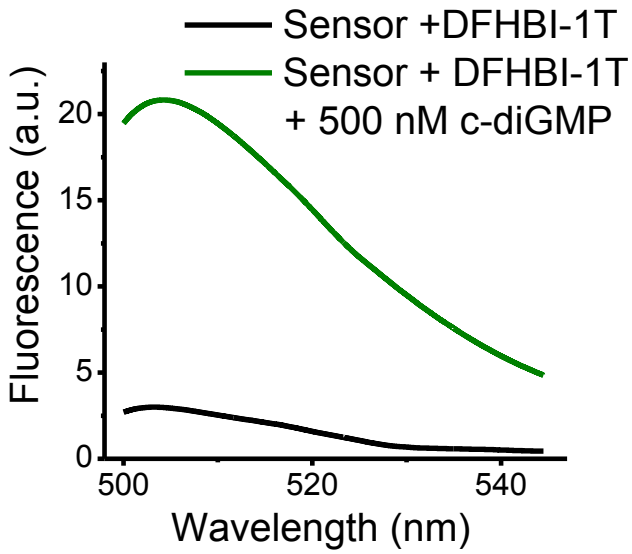


Figure S10. Broccoli can be modified into a small molecule sensor.

(a) Schematic depiction of the sensor structure and its activation mechanism. The Broccoli-based sensor becomes folded and capable of forming a stable complex with the fluorophore only upon analyte binding.

(b) Broccoli-based sensor shows ~7 fold activation upon treatment with c-diGMP. In these experiments, 1 µM of *in vitro* transcribed c-diGMP Broccoli-based sensor was premixed with 10 µM of DFHBI-1T and then treated with 500 nM of c-diGMP. The fluorescence signal increase was recorded on the fluorometer.

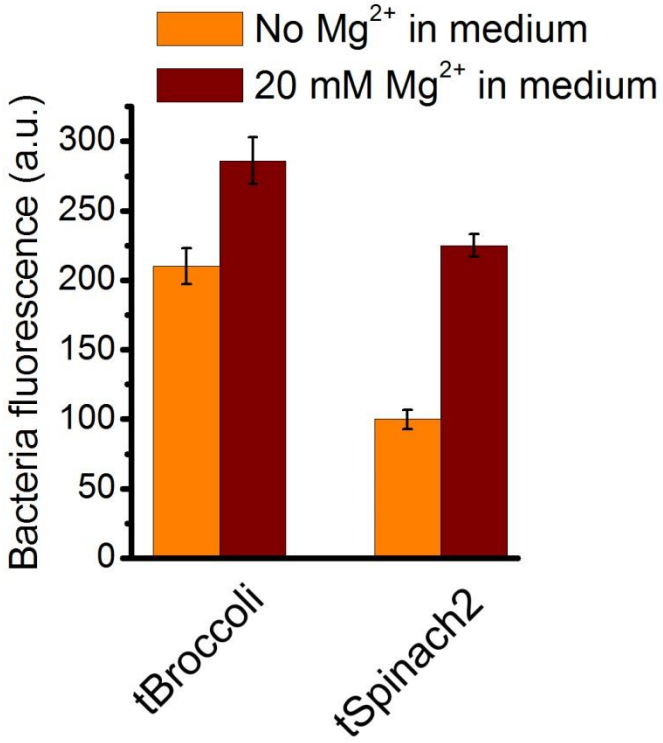


Figure S11. Supplementation of bacteria growth media with magnesium substantially increases brightness of tSpinach2 expressing cells. The main *in vitro* differences between Broccoli and Spinach were their magnesium and temperature dependence. Since all bacterial experiments at room temperature showed higher fluorescence associated with Broccoli expression compared to Spinach2, temperature is not the sole factor accounting for improved performance of Broccoli over Spinach2. Thus, the requirement of Spunach2 for Mg2+ is more likely to be the difference that explains their different performance in cells. In order to test this idea, we incubated *E. coli* expressing either tBroccoli or tSpinach2 with either 200 µM DFHBI-1T, or 200 µM DFHBI-1T and 20 mM $MgCl_2$ for one hour. Next, we measured the fluorescence of the cellular suspension on a plate reader. The result is presented as a bar graph. The tSpinach2 signal without magnesium was set as 100. The fluorescent signal from the negative cells was subtracted as a background. tSpinach2 fluorescence is highly sensitive to magnesium in the media, while tBroccoli is less sensitive to added magnesium. Since our imaging experiments were performed in media lacking added magnesium, this likely accounts for the difference in cellular behavior of Spinach2 vs. Broccoli. Error bars indicate standard deviations (n=3).

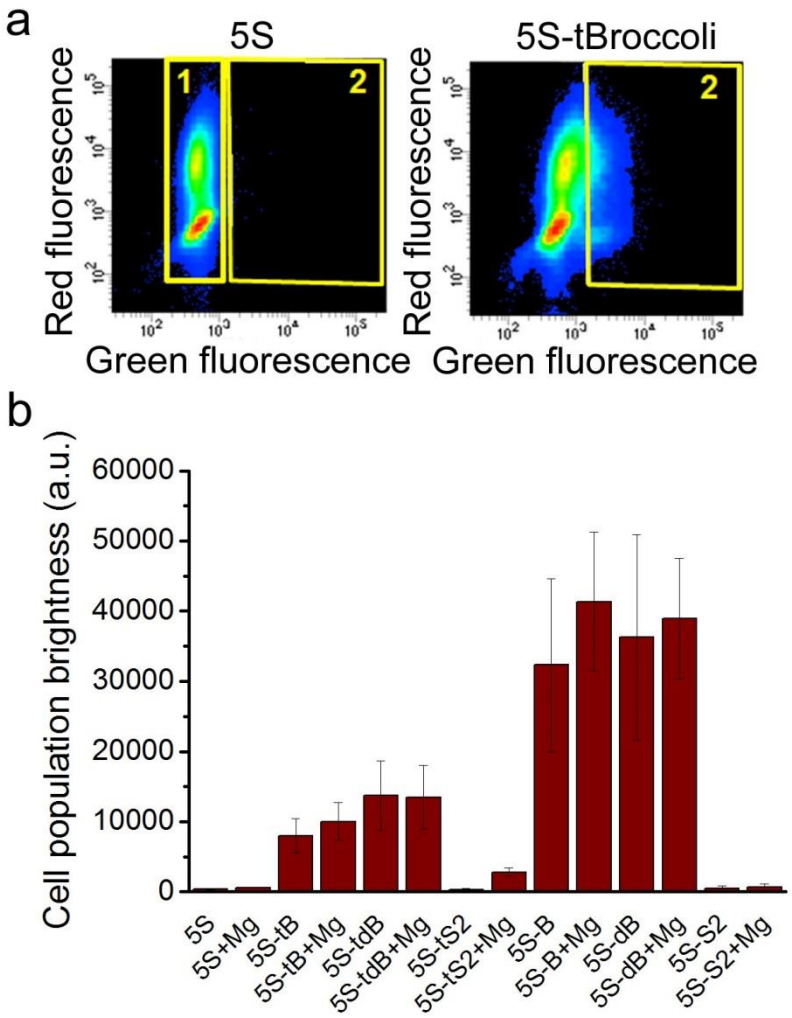


Figure S12. Quantification of the flow cytometry analysis of aptamer expression in mammalian cells. To compare aptamer brightness in mammalian cells, we transfected HEK293T with two plasmids: one expressing the aptamer (with or without tRNA, indicated with a "t") fused to 5S rRNA, and another expressing mCherry. Cells were then treated with 40 µM DFHBI-1T and analyzed on a FACSAria III instrument in two channels: green (ex=488 nm, em= 525±50 nm) and red (ex=561 nm, em=610±20).

(a) For quantification, the mean green fluorescent signal of the non-negative population (gate 2, 5S-tBroccoli is shown as an example) was calculated, and then the mean green fluorescent signal of the 5S-expressing cells (gate 1) was subtracted from it to assess the specific fluorescence signal. Finally, this number was multiplied by the percentage of the non-negative population (percentage of the cells in the gate 2) to provide the total fluorescent signal of the non-negative population.

(b) Calculation of the mean fluorescence of the aptamer-expressing cells. Where indicated, cells were also pretreated with 5 mM $MgSO_4$. Error bar indicates standard deviation calculated by the FlowJo program (Tree Star), used for the data analysis. As can be seen, tBroccoli is not substantially enhanced by magnesium supplementation, and tdBroccoli is more fluorescent than tBroccoli. Also, removing the tRNA scaffold altogether resulted in increased Broccoli and dBroccoli fluorescence, but abrogated Spinach2 fluorescence.

METHODS

Reagents and equipment

Unless otherwise stated, all reagents were purchased from Sigma-Aldrich. Commercially available reagents were used without further purification. Fluorophores used in this study were obtained from Lucerna Technologies (New York, NY) or were synthesized as described previously.[4,5] Absorbance spectra were recorded with a Thermo Scientific NanoDrop 2000 spectrophotometer with cuvette capability. Fluorescence excitation and emission spectra were measured with a PerkinElmer LS-55 fluorescence spectrometer. ChemiDoc MP imager (BioRad) was used to record fluorescence in bacterial colonies on agar plates and in gels. Fluorescence also was measured on Safire II or Genios Pro plate readers (Tecan). FACS experiments were performed using FACSAria III instrument (BD Biosciences). Fluorescence imaging experiments were performed using an Eclipse TE2000-E microscope (Nikon).

SELEX procedure

The random library used for our SELEX was generated before and already utilized to isolate Spinach aptamer (for details see Paige *et al.*, Science, 2011 (ref. 4)). Briefly, this library contained two 26-base random stretches separated by a 12-base fixed sequence and flanked from 5′ and 3′ ends with constant regions used for PCR amplification and *in vitro* transcription. This ssDNA library was purified and amplified as described previously.[4]

Doped libraries were created in a way that each encoded aptamer resembles the parent aptamer, except that there are on average seven mutations per sequence. In order to obtain this library every position is chemically synthesized with a phosphoramidite nucleosides mixture that contains primarily the nucleotide that is found at that position in the parent aptamer, but also contains each of the other nucleotides at a lower concentration. For example, if a C is present at a certain position, we use phosphoramidite mixture containing 88.2% C, 3.2% G, 4.8% A, and 3.8% T. Unequal concentrations for the non-parent nucleotides are used to take into account the different chemical activity of the respective phosphoramidites.[7] Using these synthesis conditions, most strands will incorporate a C at that position, but some will incorporate one of the other nucleotides. dsDNA encoding doped libraries were designed with 14% mutagenesis efficiency and were ordered from Protein and Nucleic Acid Facility, Stanford University Medical Center.

The first doped library was generated based on the 29-1-T2 aptamer with removed G46/C60 base pair (numbering as on the Fig. 2). This base pair was found to be unnecessary for the correct folding and fluorescence generation.

Affinity matrix for SELEX (DFHBI-sepharose) was prepared as described previously.[4]

1×10^{14} different sequences of double stranded DNA were transcribed in a 250 µl T7 RNA polymerase transcription reaction using the AmpliScribe T7-Flash Transcription Kit (Epicentre Biotechnologies). After treatment with DNase (Epicentre Biotechnologies) for 1 h, RNAs were purified using RNeasy Mini Kit (Qiagen) following manufacture's recommendations.

For random library SELEX rounds 1-4 RNA was then diluted in selection buffer containing 40 mM HEPES pH 7.4, 100 mM KCl, 1 mM $MgCl_2$, and 0.1% DMSO. Starting round five and

onwards RNA was diluted in the same buffer but containing 0.1 mM $MgCl_2$ to preferentially select for aptamers with low magnesium dependence.

SELEX procedure was conducted essentially as described before.[4] Briefly, during the first step RNA species capable of binding to the sepharose matrix were removed by incubation with "mock" resin consisting of aminohexyl linker bound to sepharose. The resulting RNA solution was then incubated with DFHBI-coupled matrix. RNA bound to DFHBI resin was then washed with 3x0.5 ml of selection buffer during rounds 1-2, 4x0.5 ml during rounds 3-6 and 6x0.5 ml during round 7. Finally, specifically bound RNA was eluted with free DFHBI.

Doped library SELEX was conducted essentially the same way. Magnesium concentration in selection buffer was decreased from 1 mM to 0.1 mM on the second round of SELEX and maintained low during the third round as well.

The eluted RNAs were then ethanol precipitated, reverse-transcribed, PCR amplified and *in vitro* transcribed to yield the pool for the next round. Presence of fluorescent RNA species in each pool was assessed by mixing 20 μM RNA and 10 μM DFHBI and measuring fluorescence emission of this solution on a fluorometer in comparison with the fluorophore alone.

Bacterial expression plasmids

To engineer a pETDuet-1-based vector that allows the insertion of a SELEX library and which bears the gene of a fluorescent protein used for expression normalization, we have done the following: First, the Spinach sequence with flanking human $tRNA_{Lys}$ scaffold regions and downstream T7 terminator was PCR amplified, digested with *Xba*I and *Bsi*WI restriction enzymes and inserted into the pETDuet-1 vector (EMD Biosciences) cut with *Xba*I and *Bsr*GI. This placed Spinach in the tRNA scaffold (tSpinach) after the first T7 promoter, and also introduced a terminator sequence that was not present in the first expression cassette in the original vector. The resulting plasmid was further modified to eliminate *Eag*I site in the vector backbone. That was performed using the QuickChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. The next step was to remove the lac operator sequence after the first T7 promoter so that no additional sequences were present between the promoter and 5' tRNA scaffold sequence. Again, that was done using the QuickChange site-directed mutagenesis kit mentioned above. Finally, eqFP670, a red fluorescent protein[1] gene, was PCR amplified from pNirFP-N plasmid (Evrogen), digested with *Nde*I and *Xho*I restriction enzymes and inserted into the plasmid generated in the previous step cut with the same enzymes. This resulted in pETDuet-5-tSpinach-eqFP670 plasmid which was used as a vector for library insertion via *Sac*II and *Eag*I sites.

To produce RNA aptamers in arabinose-induced LMG194 cells (ATCC), an expression vector was engineered as follows: tSpinach with a T7 terminator on its 5' end was PCR amplified and cut with *Nco*I and *Hind*III and then inserted into pBAD/His A vector (Life Technologies) cut with the same enzymes. Then the region between araBAD promoter and 5' tRNA region was removed by means of QuickChange mutagenesis which finally resulted in the pBAD E tSpinach plasmid. All RNA aptamers studied or sorted in LMG194 cells were later cloned into this vector using *Sac*II and *Eag*I sites. To express final mutants tested with a shortened stem between the tRNA scaffold and the aptamer, the pBAD F plasmid was engineered as follows: the pBAD E plasmid was modified using the QuickChange approach so that a short fragment bearing two *Bbs*I sites was inserted between the araBAD and the T7 terminator sequences. All tRNA

scaffold-based aptamers lacking *Eag*I and *Sac*II sites, including tBroccoli, were cloned into pBAD F plasmid using BbsI sites having different overhanging sequences.

For the highest expression level of aptamers in bacterial cells, the pET28c plasmid backbone (EMD Biosciences) was used. Aptamers in the tRNA scaffold were PCR amplified to have a *Bgl*II site and a T7 promoter on 5' end, and a T7 terminator and *Xho*I site on the 3' end. This was then cloned into pET28c-tSpinach2 as described previously[6] using *Bgl*II and *Xho*I restriction enzymes.

Bacterial library generation and FACS sorting

The RNA pool which was prepared by doing six rounds of SELEX on the fully random library, or the RNA pool after three rounds of SELEX on the doped library, were reverse transcribed and PCR amplified. Then these PCR products were cloned into pETDuet-5-tSpinach2-eqFP570 or pBAD E tSpinach using *Eag*I and *Sac*II sites.

The resulting ligation mixtures were purified and electroporated into the Acella strain (Edgebio) or LMG194 (ATCC). LMG194 cells then were grown in LB media overnight in presence of 0.2% arabinose and then collected for sorting. Acella cells were grown overnight with antibiotic and then were diluted in the morning 1 to 10 with fresh LB and with addition of 1 mM IPTG. Then expression was allowed to proceed for four hours followed by collecting these cells for sorting. Typical bacterial libraries contain 10-30x10^6 individual members.

Cells were preincubated with 40 µM DFHBI (FACS of the random library) or DFHBI-1T (subsequent FACS of the doped library) and then sorted on a FACSAria III instrument (BD Biosciences). The sample compartment of the sorter was maintained at 37^{o}C to facilitate sorting of cells expressing the most thermostable aptamers. DFHBI or DFHBI-1T-binding aptamers were excited with the 488 nm laser and their emission was collected using 525±50 emission filter. The eqFP670 fluorescent protein was monitored in the PE-Cy5 channel (561 nm excitation and 660±20 nm emission).

Typically one thousand of the brightest cells were sorted, rescued in 1 ml SOC media at 37^{o}C for 1 h and then plated on LB/agar supplemented with carbenicillin, DFHBI (or DFHBI-1T) and, in the case of LMG194 cells, 0.2% arabinose. When the Acella strain was used 1 mM IPTG was added to the dishes the next morning and cells were allowed to express aptamers for four to six more hours.

Dishes were photographed on a ChemiDoc MP imager (Bio-Rad). Aptamer library fluorescence was collected in a channel with 470±30 nm excitation and 532±28 nm emission. eqFP670 fluorescent protein signal was collected with 630±30 nm excitation and 697±55 nm emission. The same channel was used in case of LMG194 cells to collect autofluorescence signal from bacterial colonies which allowed normalizing for their size. Images were processed and normalized in ImageJ software (NIH).[8]

***In vitro* characterization of aptamers**

dsDNA from the brightest bacterial cells in a library was PCR amplified from eluted plasmids. Truncation, deletion and point mutation mutants were generated from dsDNAs that were PCR amplified from the appropriate ssDNA templates (synthesized at the Protein and Nucleic Acid Facility, Stanford University Medical Center). Primers were designed so the PCR product

contained the T7 promoter sequence. When indicated, dsDNA or ssDNA also contained the tRNA scaffold sequence. PCR products were then purified with PCR purification columns (Qiagen) and *in vitro* transcribed utilizing an AmpliScribe T7-Flash Transcription Kit (Epicenter). RNA was purified using Bio-Spin columns (Bio-Rad), and quantified using both absorbance values and the Quant-iT RiboGreen RNA Assay Kit (Life Technologies).

All *in vitro* RNA properties were measured in 40 mM HEPES pH 7.4, 100 mM KCl, 1 mM $MgCl_2$ buffer unless specified. All values presented are the average of at least three independent experiments. Error bars indicate standard deviations.

Absorption, excitation and emission spectra were measured for solutions using "excess RNA" conditions and limiting amount of fluorophore to ensure that no free fluorophore contributes to the absorbance or fluorescence signal. This approach also allows us to have a fixed concentration of RNA-fluorophore complex which is equal to the concentration of the fluorophore that was initially added. The RNA concentration was 30 μM (fluorescence measurements) and 50 μM (absorption measurements) while DFHBI-1T concentration was 2 μM and 5 μM respectively.

The extinction coefficient was calculated based on the absorbance spectrum and the Beer-Lambert-Bouguer law. For quantum yield calculations, the fluorescence signal of Broccoli-DFHBI-1T complex was compared at different dilutions to that of equally absorbing Spinach2-DFHBI-1T (quantum yield 0.94, Song *et al.*, JACS, 2014 (ref. 5)).

To calculate dissociation constant we performed a titration of 50 nM RNA with increasing concentration of DFHBI-1T and then fitted the resulting data points with the curve based on the Hill equation.

To measure the thermostability RNA-fluorophore complexes, 1 μM of RNA was incubated with 10 μM DFHBI-1T. Then the fluorescence values were recorded in 1°C increments from 20°C to 70°C, with a 2-min incubation at each temperature to allow for equilibration.

Folding measurements were performed essentially as described before.[6] Briefly, the fluorescence intensity of two solutions was compared: one having excess of the fluorophore and limiting amount of RNA, and the other with excess of RNA and limiting amount of the fluorophore. As we described previously, this allows us to calculate the percent of the aptamer that is folded. For these experiments, we used 1 μM of DFHBI-1T (or RNA) and 20 μM of RNA (or DFHBI-1T). The signal from the first condition (limiting RNA) was divided by the signal from the second condition (limiting fluorophore) to determine the fraction folded. To measure the folding in the context of different flanking sequences, RNA was generated from separate dsDNA fragments. These fragments were PCR amplified to have β-actin and 5S RNA sequences on one or both sides of Broccoli. The folding assay then performed as described above.

To measure magnesium dependence, 1 μM RNA was incubated with 10 μM DFHBI-1T in 40 mM HEPES pH 7.4, 100 mM KCl buffer with different concentrations of $MgCl_2$ and then fluorescence emission was measured on a plate reader.

In vitro characterization of tBroccoli-based sensor

c-diGMP sensor based on tBroccoli aptamer was generated basically following the strategy and using the sequences described previously.[4,9] RNA was *in vitro* transcribed from PCR fragments

generated from ssDNA (synthesized at the Protein and Nucleic Acid Facility, Stanford University Medical Center). To test the sensor functionality, 1 μM of c-diGMP Broccoli-based sensor RNA was premixed with 10 μM of DFHBI and then treated with 500 nM of c-diGMP. Fluorescence emission increase was recorded on a fluorometer as described above.

Fluorescence measurements of *E. coli*

Fluorescence measurements of *E. coli* were done in two general ways.

To assess fluorescence brightness of bacterial colonies on a LB-agar plate, Acella or LMG194 strains cells were transferred to a new LB/agar dish with the appropriate antibiotic and allowed to grow overnight at 37ºC. Typically four to six colonies for the same aptamer were grown on the same plate to provide replicates for measuring statistical data. These plates already contained 40 μM DFHBI-1T (or DFHBI) and 0.2% arabinose (in case of LMG194 strain). When the Acella strain was used, 1 mM IPTG was added the next morning and incubation was continued for 4 to 6 hours more at 37ºC. Then bacterial colonies were imaged under ChemiDoc MP imager as described above and images were processed using ImageJ.

To measure fluorescence of bacterial cell suspensions, we used the BL21 Star (DE3) (Life Technologies) strain and pET28c-based expression plasmid. This combination provided the strongest signal. Colonies were inoculated into LB media with kanamycin and allowed to grow overnight at 37ºC to OD_{600} 0.4. The cells were then induced by addition of 1 mM IPTG for 2-4 h at 37ºC. After induction, cells were normalized for cell density and incubated with 400 μM DFHBI-1T. Cells were then measured for total fluorescence using a Genios Pro plate reader at 465 nm excitation and 535 nm emission. All values were normalized for slightly different excitation and emission spectra of Spinach2 and Broccoli.

Aptamer imaging in bacteria

The BL21 Star (DE3) (Life Technologies) strain was transformed with pET28c-1-based expression vectors encoding RNA aptamers in the tRNA scaffold. Negative cells were transformed with the original pET28c plasmid. Cells were plated, grown overnight and single colonies were picked for inoculation overnight in LB broth with kanamycin. At OD600 = 0.4 1 mM IPTG was added to the media. After 2-4 more hours at 37ºC, the culture was pelleted, resuspended in PBS and transferred to poly-d-lysine coated 3.5 cm dishes (MatTek Corporation). Cells were allowed to attach to the dish for 45 min at 37ºC, then washed with PBS and incubated with 200 μM DFHBI-1T in PBS at 37ºC for another 45 min. Live fluorescence images were taken with a CoolSnap HQ2 CCD camera through a 60X oil objective mounted on a Nikon Eclipse TE2000-E microscope and analyzed with the NIS-Elements software. The filter set used was a sputter coated filter cube with excitation filter 470±40, dichroic mirror 495 (long pass), and emission filter 525±50 (Chroma Technology).

Cloning of the plasmids used for the aptamers imaging in mammalian cells

To engineer plasmids encoding 5S rRNA fusions with tBroccoli or tdBroccoli, we used the previously described pAV-5S-Spinach plasmid.[4] This construct contained Spinach in the context of the $tRNA_{Lys}$ scaffold. The sequence encoding $tRNA_{Lys}$-Spinach was removed from pAV-5S by restriction digest with *Sal*I and *Xba*I. The sequence encoding tBroccoli or tdBroccoli were amplified from pET28c-based plasmids by PCR and then digested with *Xba*I and *Sal*I and finally inserted into pAV-5S.

To fuse Spinach2, Broccoli and dBroccoli to 5S without tRNA scaffold we PCR amplified the aptamers from the respective pET28c-based plasmids and cloned into the above mentioned pAV-5S plasmid using the same *Sal*I and *Xba*I sites.

Cell culture conditions

Cell lines were obtained directly from the American Type Culture Collection (ATCC) for all experiments. HEK-293T (ATCC-CRL-11268) were grown according to ATCC instructions. Cells were screened for mycoplasma contamination before passaging using Hoechst 33258, according to ATCC recommendations.

Imaging 5S-fused aptamers in mammalian cells

Plasmids were transfected into HEK293T cell using FuGENE HD reagent (Promega) following the manufacturer's protocol. pSuperior-mCherry plasmid was also transfected as a transfection efficiency control. Cell imaging was carried out as previously described for 5S-Spinach[4] with DFHBI-1T fluorophore used instead of DFHBI. Briefly, cells were pretreated with 20 µM DFHBI-1T, 5 µg/ml Hoechst 33258 and 0.3 M sucrose 30 min before the imaging. Live fluorescence images were taken with a CoolSnap HQ2 CCD camera through a 60X oil objective mounted on a Nikon Eclipse TE2000-E microscope and analyzed with the NIS-Elements software. The filter set used for cBroccoli detection was a filter cube with excitation filter 470±20 nm, dichroic mirror 495 nm (long pass), and emission filter 525±25 nm. mCherry was detected using 560±20 nm excitation filter, 585 nm (long pass) dichroic mirror and 630±37.5 nm emission filter. Hoechst-stained nuclei were imaged with 350±25 nm excitation filter, 400 nm (long pass) dichroic mirror and 460±25 nm emission filter (all filters are from Chroma Technology). Exposure times: 500 ms for cBroccoli, 200 ms for mCherry and Hoechst. Background signals from cells expressing pAV-5S incubated with DFHBI-1T were subtracted from the corresponding images using NIS-Elements software (Nikon).

In-gel imaging of fluorescent RNAs

Total bacterial or mammalian cell RNA was purified using Trizol LS reagent (Life Technologies) following the manufacturer's protocol. Typically 200-500 ng of total bacterial RNA, 2 µg of mammalian cell RNA or 50-100 ng of *in vitro* transcribed RNA was loaded into a well of precast 6% TBE-Urea Gel (Life Technologies) and ran at 270-300 V in 1x TBE buffer. RiboRuler Low Range RNA Ladder (Thermo Scientific) was used as molecular weight standard.

The in-gel aptamer staining protocol with its abilities and limitations is characterized in a separate manuscript (Filonov *et al.*, manuscript under preparation). Briefly, the protocol is as follows. After the gel was run to completion, the gel was washed 3x5 min with water and then stained for 30 min in 10 µM DFHBI or DFHBI-1T in buffer containing 40 mM HEPES pH 7.4, 100 mM KCl, 1 mM $MgCl_2$. Then gel was imaged using a ChemiDoc MP with 470±30 nm excitation and 532±28 nm emission. Next, to see all the RNA in the sample, the gel was again washed 3x5 min with water followed by staining for 30 min with SYBR Gold (or Green) fluorophore (Life Technologies) diluted 1/10000 in TBE buffer. Then gel was imaged under the same instrument using preset SYBR Gold (or Green) channel (302 nm excitation and 590±110 nm emission). Gel band intensities were quantified using Image Lab 5.0 software (Bio-Rad). The bacterial or mammalian 5S rRNA band was used for loading normalization.

Flow cytometry of mammalian cells

HEK293T cells were transfected with the respective plasmids as described above. Untagged 5S-expressing cells were used as a negative control and mCherry was also expressed from pSuperior-mCherry as a transfection efficiency control. After 48 h the cells were harvested and resuspended in the 4% FBS/1x PBS solution containing 40 µM DFHBI-1T and, where indicated, 5 mM $MgSO_4$ and kept on ice until analysis on the FACSAria III instrument (BD Biosciences). Transfected cells were analyzed in two channels: green (ex=488 nm, em= 525±50 nm) and red (ex=561 nm, em=610±20). Processing and analysis of the data was performed in the FlowJo program (Tree Star)

REFERENCES

(1) Shcherbo, D.; Shemiakina, II; Ryabova, A. V.; Luker, K. E.; Schmidt, B. T.; Souslova, E. A.; Gorodnicheva, T. V.; Strukova, L.; Shidlovskiy, K. M.; Britanova, O. V.; Zaraisky, A. G.; Lukyanov, K. A.; Loschenov, V. B.; Luker, G. D.; Chudakov, D. M. *Nat Methods* **2010**, *7*, 827.

(2) Davis, J. H.; Szostak, J. W. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 11616.

(3) Zuker, M. *Nucleic acids research* **2003**, *31*, 3406.

(4) Paige, J. S.; Wu, K. Y.; Jaffrey, S. R. *Science* **2011**, *333*, 642.

(5) Song, W.; Strack, R. L.; Svensen, N.; Jaffrey, S. R. *J Am Chem Soc* **2014**, *136*, 1198.

(6) Strack, R. L.; Disney, M. D.; Jaffrey, S. R. *Nature Methods* **2013**, *10*, 1219.

(7) Hall, B.; Micheletti, J. M.; Satya, P.; Ogle, K.; Pollard, J.; Ellington, A. D. *Current protocols in molecular biology* **2009**, 39:9.2:9.2.1–9.2.28.

(8) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat Methods* **2012**, *9*, 671.

(9) Kellenberger, C. A.; Wilson, S. C.; Sales-Lee, J.; Hammond, M. C. *J Am Chem Soc* **2013**, *135*, 4906.