

Supplementary material

Riboswitching with ciprofloxacin – Development and characterization of a novel RNA regulator

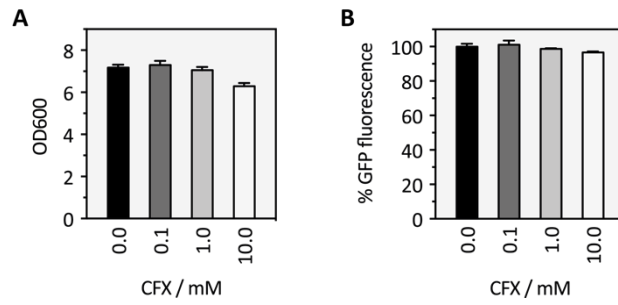
Florian Groher ¹, Cristina Bofill-Bosch ¹, Christopher Schneider ¹, Johannes Braun ¹, Sven Jager ², Katharina Geißler ¹, Kay Hamacher ^{2,3}, Beatrix Suess ^{1,*}

¹ *Synthetic Genetic Circuits, Dept. of Biology, TU Darmstadt, Darmstadt, Germany*

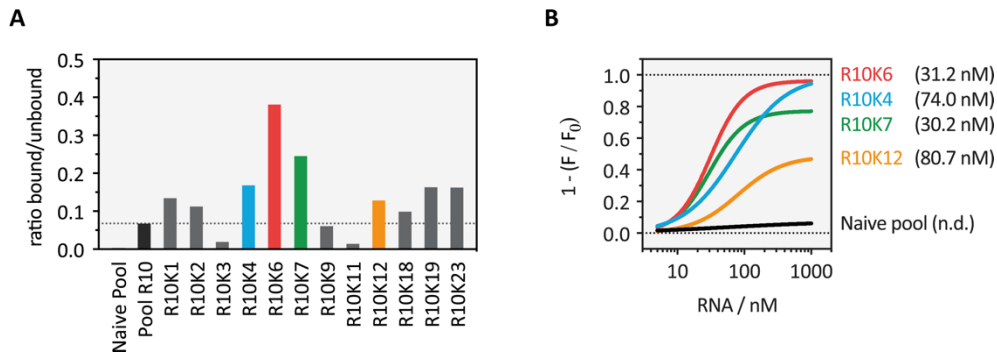
² *Computational Biology and Simulation, Dept. of Biology, TU Darmstadt, Darmstadt, Germany*

³ *Dept. of Physics, Dept. of Computer Science, TU Darmstadt, Darmstadt, Germany*

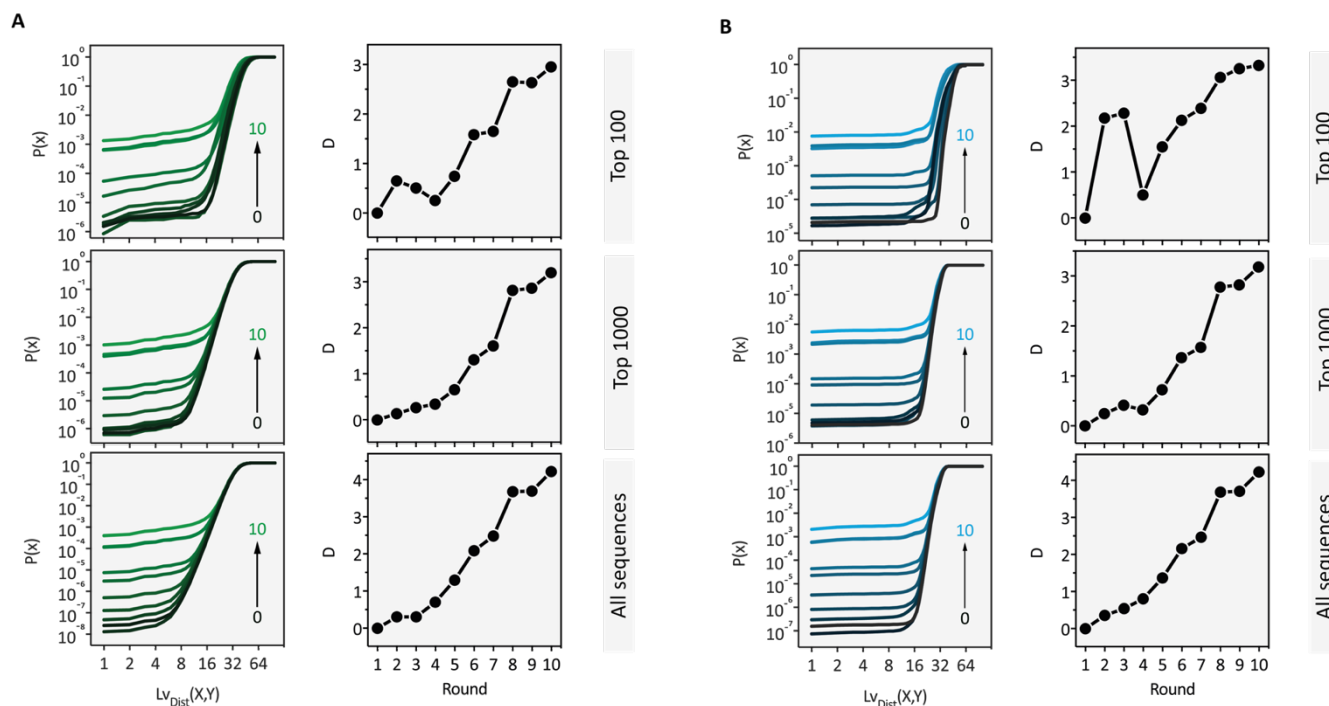
* To whom correspondence should be addressed. Tel: +49 6151 1622000; Fax: +49 6151 1622003; Email: bsuess@bio.tu-darmstadt.de



Supplementary Figure S1. Influence of CFX on yeast growth and GFP expression. **A** OD600 of yeast cultures grown overnight in media supplemented with the respective CFX concentration. **B** Relative GFP fluorescence of yeast cultures supplemented with the respective CFX concentration compared to untreated cells. Measurements were repeated three times with technical replicates.



Supplementary Figure S2. Analysis of single clone binding from round 10. **A** Ratios of bound vs. unbound RNA for different clones from selection round 10 are displayed. As references, the naive pool and the pool from round 10 are depicted. According to the SELEX procedure, RNA was transcribed and 500 kcpm were loaded onto the CFX-derivatized column. After 10 wash steps with 1 CV binding buffer each, the RNA was eluted by 4 wash steps with 1 mM CFX in solution. Afterwards each fraction was measured on a scintillation counter. Measured radioactivity in the fractions flow through and wash steps were summed up (unbound) and also for elution fractions (bound). The ratio of bound to unbound gives a direct qualitative feedback of the binding capacity of the tested clones. **B** Determination of binding affinity of the selected aptamer candidates by fluorescence titration spectroscopy. Measurements were repeated at least twice. Standard deviations and individual data points were omitted for clarity. K_D values are written in brackets.



Supplementary Figure S3. Next generation sequencing analysis Displayed are the cumulative distribution function (CDF) and Kolmogorv Smirnov's ks test (D) for Top100, Top1000 and all sequences. **A** Results based on calculated minimal free energy (MFE) secondary structure. **B** Results based on sequence. The CDF for each round based on calculated Levenshtein distances on MFE structures is plotted for each round (left in A and B), resulting in an increased $P(x)$ over the selection experiment. Based on CDF, D was derived and its logarithm is plotted against the selection rounds for Top100, Top1000 and all sequences (right panel in A and B). Here, D is computed between the first round and all remaining.

One major drawback is the computational time that it takes to compute a $Lv_{Dist}(X,Y)$ distribution where we compare every sequence with every other (often 10^{12} single computations). Due to this, advanced computational resources as well as efficient software and memory management is required. However, the data suggests that calculating all Levenshtein distances for each sequence and each round is not necessary and it is sufficient to look at the Top1000 enriched sequences to draw conclusions (at least in this SELEX experiment). This fact reduced the calculation efforts required by several orders of magnitude. We can conclude that comparing Top1000 vs all sequences by its Levenshtein distance can improve the process of SELEX round selection for future work. Additionally, using only the Top1000 made the computation feasible on a desktop computer by reducing the computational time by several orders of magnitude.

Supplementary Table S1. Plasmids used in this study

Name	Description	Reference
pWHE601	2 μ plasmid with constitutively expression of <i>gfp+</i> from an <i>adh</i> promoter	(1)
pWHE601*	Derived from pWHE601 with deletion of AUG in <i>gfp+</i> / AflIII --> Agel	(2)
10A	Active riboswitch found in initial <i>in vivo</i> screening	This work
Δ AUG	Deletion of AUG within the sequence of 10A	This work
GOF	Introduction of 7 point mutations in Δ AUG	This work
G1U	Investigation of the named point mutation within Δ AUG	This work
A11C	Investigation of the named point mutation within Δ AUG	This work
A25C	Investigation of the named point mutation within Δ AUG	This work
U47C	Investigation of the named point mutation within Δ AUG	This work
C51U	Investigation of the named point mutation within Δ AUG	This work
A56C	Investigation of the named point mutation within Δ AUG	This work
U61G	Investigation of the named point mutation within Δ AUG	This work
A35G	Investigation of the named point mutation within Δ AUG	This work
U41G	Investigation of the named point mutation within Δ AUG	This work
A50G	Investigation of the named point mutation within Δ AUG	This work
U92G	Investigation of the named point mutation within Δ AUG	This work
A102G	Investigation of the named point mutation within Δ AUG	This work
M1	Mutation of the C31 and G32 to G and C within GOF, respectively	This work
M1R	Compensatory point mutations for M1 to restore function	This work
U37A	Investigation of the named point mutation within GOF	This work
G72C	Investigation of the named point mutation within GOF	This work
M2	Mutation of GUU75 to CAA within GOF	This work
M2R	Compensatory mutations for M2 to restore pseudoknot and function	This work
M3	Mutation of G75C and C79G within GOF	This work
M3R	Compensatory mutations for M3 to restore pseudoknot and function	This work
COMP	Complementary sequence of GOF for investigation of basal expression	This work

Corresponding oligonucleotides for cloning are listed in Supplementary Table S2.

Supplementary Table S2. Oligonucleotides used for cloning

Name	Sequence (5'->3')
10A_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGCAC
10A_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAG
ΔAUG_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGCAC
ΔAUG_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAG
GOF_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGCACG
GOF_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATCTGACGCCTAATTTGGGGAG
G1U_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGCAC
G1U_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAG
A11C_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGC
A11C_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAG
A25C_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATACGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGC
A25C_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAG
U47C_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCAAATTAGGAGTCATATAGCGGC
U47C_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAG
C51U_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAAATTAGGAGTCATATAGCGGC
C51U_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAATTTAGGGAGATAG
A56C_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGCGTCATATAGCGGC
A56C_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACGCCTAGTTTAGGGAGATAG
U61G_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCAGATAGCGGC
U61G_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATCTGACTCCTAGTTTAGGGAGATAG
A35G_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGCTCTATCTCCCTAACTAGGAGTCATATAGCGGCAC
A35G_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAG
U41G_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTAGCTCCCTAACTAGGAGTCATATAGCGGCAC
U41G_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGCTAGAG
A50G_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAAGCTAGGAGTCATATAGCGGCAC
A50G_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGCTTAGGGAGATAGAG
U92G_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGCAC
U92G_rev	GGCCGCTAGCCATTTTGTGACGCGACTCGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAG
A102G_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAACGCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGCAC
A102G_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAG
M1_fwd	CGCGACCGGTTGGAGACGCAACTGAATCAACATAAGTGAAGCCGACTCTATCTCCCTAACTAGGAGTCATATAGCGGC
M1_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGG
M1R_fwd	- identical to M1_fwd -
M1R_rev	GGCCGCTAGCCATTTTGTGAGCCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAGTCG
U37A_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGC
U37A_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATCTGACGCCTAATTTGGGGAGATAGTG
G72C_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGCACCG
G72C_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTAACTCCGTGCCGCTATCTGACGCCTAATTTGGGGAGATAGATC
M2_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGC
M2_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTGTGTTGTCGGTGCCGCTATCTGACGCCTAATTTGGGGAGATAG
M2R_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGCACGG
M2R_rev	GGCCGCTAGCCATTTTGTGACGCGACTACAAACGGATCGTGTGTTGTCGGTGCCGCTATCTGACGCCTAATTTGGGGAG
M3_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGC
M3_rev	GGCCGCTAGCCATTTTGTGACGCGACTAGTTACGGATCGTCTAAGTCCGTGCCGCTATCTGACGCCTAATTTGGGGAGATAG
M3R_fwd	CGCGACCGGTTGGAGACGCACCTGAATCAACATACGTGAACGCGACTCTATCTCCCAAATTAGGCGTCAGATAGCGGCACG
M3R_rev	GGCCGCTAGCCATTTTGTGACGCGACTACTTAGGGATCGTCTAAGTCCGTGCCGCTATCTGACGCCTAATTTGGGGAGATAG
COMP_fwd	CGCGACCGGTACCTCTGCGTGGACTTAGTTGTATGCCTGCGCTGAGATAGAGGGTTAATCCGAGTCTATCCCGGTG
COMP_rev	GGCCGCTAGCCATTTTCACTGCGCTGATCAATGCCTAGCACATTGAGGCACGGCGATAGACTGCGGATTAACCCCTCTATCTC

Supplementary Table S3. Oligonucleotides used for cloning of doped pools for *in vivo* screening

Name	Sequence (5'->3')
Agel_doped_fwd	GCATACAATCAACTCCAAGCTAGATCTACCGGT
NheI_[3.0/4.5/9.0/30.0]_doped_rev	CGAGCTAGCCATTTT[GTGACGCGACTAGTTACGGATCGTGTA ACTCCGTGCCGCTATATGACTCCTAGTTTAGGGAGATAGAGTCGCGTTCACTTATGTTGATTGAGTTGCGTCTCCC]ACCGGTAGATCTAGCTTGAGTTGATTGTATGC

For all cloning steps Agel_doped_fwd was used for PCR. For generating different degrees of randomization, the part in brackets of NheI_ATG_Kozac_doped_rev was synthesized with mixed phosphoramidites for 3.0%, 4.5%, 9.0% and 30.0% incorporation of the other three bases.

Supplementary Table S4. Oligonucleotides and barcodes used for Illumina sequencing

Name	Round	Barcode	Sequence (5'->3')
Seq_IL_fwd	-	-	GGGAGACGCAACTGAATGAA
Seq_IL_rev0	0	GTGT	ACACGTGACGCGACTAGTTACGGA
Seq_IL_rev1	1	ACAC	GTGTGTGACGCGACTAGTTACGGA
Seq_IL_rev2	2	ATAT	ATATGTGACGCGACTAGTTACGGA
Seq_IL_rev3	3	AGAG	CTCTGTGACGCGACTAGTTACGGA
Seq_IL_rev4	4	TATA	TATAGTGACGCGACTAGTTACGGA
Seq_IL_rev5	5	TCTC	GAGAGTGACGCGACTAGTTACGGA
Seq_IL_rev6	6	TGTG	CACAGTGACGCGACTAGTTACGGA
Seq_IL_rev7	7	CACA	TGTGGTGACGCGACTAGTTACGGA
Seq_IL_rev8	8	CGCG	CGCGGTGACGCGACTAGTTACGGA
Seq_IL_rev9	9	CTCT	AGAGGTGACGCGACTAGTTACGGA
Seq_IL_rev10	10	GAGA	TCTCGTGACGCGACTAGTTACGGA

Supplementary Table S5. Oligonucleotides for template generation for *in vitro* transcription

Name	Sequence (5'→3')
10A_T7_fwd	CCAAGTAATACGACTCACTATAGGGAGACGCAACTGAATGAACATAAGTGAAC GCGACTCTATCTCCCTAACTAGG
10A_T7_rev	GTGACGCGACTAGTTACGGATCGTGTA ACTCCGTGCCGCTATATGACTCCTAGTT TAGGGAGATAGAGTCGCGTTC
ΔAUG_T7_fwd	CCAAGTAATACGACTCACTATAGGGAGACGCAACTGAATCAACATAAGTGAACGC GACTCTATCTCCCTAACTAGG
ΔAUG_T7_rev	- identical to 10A_T7_rev -
GOF_T7_fwd	CCAAGTAATACGACTCACTATAGGGAGACGCACCTGAATCAACATACGTGAACGC GACTCTATCTCCCAAATTAGGCGTCAG
GOF_T7_rev	GTGACGCGACTAGTTACGGATCGTGTA ACTCCGTGCCGCTATCTGACGCCTAATT TGGGGAGATAGAGTCGCGTTCACG
U37A_T7_fwd	CCAAGTAATACGACTCACTATAGGGAGACGCACCTGAATCAACATACGTGAACGC GACTATCTCCCAAATTAGGCG
U37A_T7_rev	GTGACGCGACTAGTTACGGATCGTGTA ACTCCGTGCCGCTATCTGACGCCTAATT TGGGGAGATAGTGTGCGGTTACG
G72C_T7_fwd	CCAAGTAATACGACTCACTATAGGGAGACGCACCTGAATCAACATACGTGAACGC GACTCTATCTCCCAAATTAGGCG
G72C_T7_rev	GTGACGCGACTAGTTACGGATCGTGTA ACTCGGTGCCGCTATCTGACGCCTAATT TGGGGAGATAGAGTCGCGTTCACG
GOF_CAA4_T7_fwd	- identical to GOF_T7_fwd -
GOF_CAA4_T7_rev	TTGTTGTTGTTGTGACGCGACTAGTTACGGATCGTGTA ACTCCGTGCCGCTATCT GACGCCTAATTTGGGGAGATAGAGTCGCGTTCACG

Supplementary Table S6. Detailed summary of the CFX selection process

Round	Negative selection	CFX col. [mM]	# Pre-elution [CV]	# Buffer washes [CV]	Specific elution	Eluent	# Elution steps [CV]	% Input eluted
1	yes	0.6	-	10	-	20 mM EDTA	4	0.2%
2	yes	0.6	-	10	-	20 mM EDTA	4	0.3%
3	yes	0.6	-	10	-	20 mM EDTA	4	0.3%
4	-	0.6	-	10	-	20 mM EDTA	4	4.2%
5	-	0.4	-	20	-	20 mM EDTA	4	2.9%
6	-	0.4	-	20	yes	1 mM CFX	4	8.1%
7	-	0.4	3	20	yes	1 mM CFX	4	4.5% *
8	-	0.4	4	20	yes	1 mM CFX	4	0.4% *
9	-	0.4	-	20	yes	1 mM CFX	4	18.1%
10	-	0.04	-	20	yes	1 mM CFX	4	6.0%

The amount of immobilized CFX was estimated by fluorescence measurement of the derivatized solid support.

CV = column volume

* 23.3% and 4.7% of pre-eluted RNA were discarded in round 7 and 8, respectively

Supplementary Table S7. Randomized regions from clones round 10

Clone	Frequency	Sequence (5'->3')	Length*
R10K1	1	TCAGTGGCATTTCAAACACCAATTTGACGAAAAGAAGACTTAGTGAATACTAAGCGGAATTAAC	104
R10K2	3	AACCAAACAGTTCATCAAGACCTAGGTATCTAGAACTAGCACGTCCGGATATGTCGGTA	101
R10K3	2	ATCAGCATCCCTACAGAGGAAGTACCGCACACTATTGTGGAAAGGCCAGATTC	93
R10K4	5	GAGGTTCCCTATCATTACAGACG <u>CTGCTTCGGCAGT</u> AACTAGAAATGTCCGGCCACTACGTG	102
R10K6	4	AATGTCATTCAAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGG	103
R10K7	1	TTGATTTCCCGTGATGAAAAGAAGACTGCTTCGGCAGCGGAAGGAAAGTTTTCGGACCCTCCA	103
R10K9	1	TGCTGAGGACATTAGTAGCAAGTTCTCTGCTTCGGCAGGCAAATTTGGCAAGTCAGCT	98
R10K11	1	CGCAATTCATTTTCACTAGGTCGTGCTTGAAAAAGTGTGGAGCCAGACTAATTAGCATCAGGG	104
R10K12	1	GTAGGTTCCCTATCATTACAGACG <u>CTGCTTCGGCAGT</u> AACTAGAAATGTCCGGCCACTACGTG	103
R10K13	1	GAGGTTCCCTATCATTACAGACG <u>CTGCTTCGGCGTA</u> AACTAGAAATGTCCGGCCACTACGTG	102
R10K18	1	CGTGGCCGAGCATACATCGTATCGGCCTGCTTCGACCAGGTCGGCCCTGGCG	92
R10K19	1	GACCGTCATTCATGAGTTCTTACGTGCTGCTTCGGCAGGGGGAGAATGGCTCGGACTTAAATGG	104
R10K23	1	CGAACTTCAACTAAACACTCCGATGTAATAACTAGCATCGTAGCCTGTCCCTGCGATAAAGGAG	104

Sequences found in SELEX round 10. Both, 5'- and 3'-regions are removed for clarity.

The reported stem loop (5'-CTGCTTCGGCAG-3') is underlined allowing for one mismatch/mutation.

* including constant regions.

Supplementary Table S9. K_D and regulatory activity of selected fluoroquinolones

Fluoroquinolone	K_D / nM	Activity / x-fold
EFX *	61.3 (1.5)	3.1 (0.1)
CFX	64.2 (1.8)	7.5 (0.3)
DFX	137.1 (10.6)	4.2 (0.7)
NFX	182.6 (22.1)	2.7 (0.2)
EX	236.6 (52.9)	2.8 (0.1)
hCFX	366.7 (67.2)	0.8 (0.0)
dCFX	829.8 (118.0)	1.7 (0.2)
PA	916.1 (195.9)	2.8 (0.2)

For every fluoroquinolone, the dissociation constant (K_D) was determined by fluorescence titration and activity *in vivo* was measured by standard GFP fluorescence assay using the CFX-riboswitch. The standard deviation (\pm SD) is reported in brackets for the titration experiments and regulatory activity, respectively.

* EFX reduced the growth rate of yeast approx. 10-fold [data not shown].

Supplementary references

1. Sues, B., Hanson, S., Berens, C., Fink, B., Schroeder, R. and Hillen, W. (2003) Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res.*, 31, 1853–1858.
2. Schneider, C. and Sues, B. (2016) Identification of RNA aptamers with riboswitching properties. *Methods*, 97, 44–50.