Supplementary Information

CRISPR-Hybrid: A CRISPR-Mediated Intracellular Directed Evolution Platform for RNA Aptamers

Qiwen Su-Tobon[1], Jiayi Fan[1], Michael Goldstein[1], Kevin Feeney[1], Hongyuan Ren[1], Patrick Autissier[2], Peiyi Wang[1], Yingzi Huang[1], Udayan Mohanty[1] and Jia Niu[1]*

[1]Department of Chemistry, Boston College, Chestnut Hill, MA, USA.
[2]Department of Biology, Boston College, Chestnut Hill, MA, USA.

Table of Contents

Supplementary Tables

Supplementary Figures

Supplementary Table 1. gRNA sequences used in E. coli cells.

Name	Sequence (5’ to 3’)
G0	GUGAACCGCAUCGAGCUGAA
G1	AACACGCACGGUGUUACAUU
G2	UAGGGACCCUACAACACGCA
G3	CGUUUUGCUGAGGAGACUUA
G4	CUGGGCCUUUCGUUUUGCUG
G5	AAAGGCUCAGUCGAAAGACU
G6	GUAACACCGUGCGUGUUGUA
G7'	AGUCUCCUCAGCAAAACGAA
G8'	AGCCUUUCGUUUUAUUUGAC
G9'	AUGUCCUCUUGACAGUAAGA
G10'	AUCCGGUCAAAUAAAACGAA
G11'	CUUACUGUCAAGAGGACAUC
G12'	UUACCCGUCUUACUGUCAAG
1.1	AACACGCACGGUGUUACAUU
1.2	CGGAUCUUCCACAACACGCA
1.3	UUAUCACCUUGGCUGCAGGC
1.4	AUGUAACACCGUGCGUGUUG
1.5	CGUGCGUGUUGUGGAAGAUC
1.6	GAAGAUCCGGCCUGCAGCCA

Supplementary Table 2. gRNA sequences used in HEK293T cells.

Name	**Sequence (5' to 3')**
Gluc	GUGAAUGAUGAUAAUACGAUG
ASCL1-1	GCUGGGUGUCCCAUUGAAA
ASCL1-2	GCAGCCGCUCGCUGCAGCAG
ASCL1-3	GUGGAGAGUUUGCAAGGAGC
ASCL1-4	GUUUAUUCAGCCGGGAGUC
XIST1-1	GGCAGCGCUUUAAGAACUGAA
XIST1-2	GGACUGAAGAUCUCUCUGCACUU
XIST1-3	GGCCAUAUUUCUUACUCUCUCG

Supplementary Table 3. qPCR primers used in this work.

Target	Forward primer	Reverse primer
ASCL1	GGAGCTTCTCGACTTCACCA	AACGCCACTGACAAGAAAGC
XIST1	AGGTCAGGCAGAGGAAGTCA	CTGCCTCCCGATACAACAAT
GAPDH	TTCGACAGTCAGCCGCATCTTCTT	GCCCAATACGACCAAATCCGTTGA

Supplementary Table 4. Amplification primers used to generate the aptamer libraries.

Oligo	Sequence (5' to 3')	Description
QSL2	GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGCGCACATNNNNNNNNNNNNATGTGCTTTTTTTGAAGCTTGGGCCCGAACAAA	Forward primer for piece A of N11 library
QSL5	TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGCGCNNNNTAGCTAAGCTANNNNGCTTTTTTTGAAGCTTGGGCCCGAACAAAA	Forward primer for piece A of N8 library
QS94	ACTCGAGTAAGGATCCAGTTCACCGAC	Reverse primer for piece A
QS91	GAAAAGTGCCACCTGACGTCTAAGAAACCATTATT	Forward primer for piece B
QS92	TTTTCAAGTTGATAACGGACTAGCCTTATT	Reverse primer for piece B
QS95	GTCGGTGAACTGGATCCTTACTCGAGT	Forward primer for piece C
QS96	AATAATGGTTTCTTAGACGTCAGGTGGCACTTTTC	Reverse primer for piece C

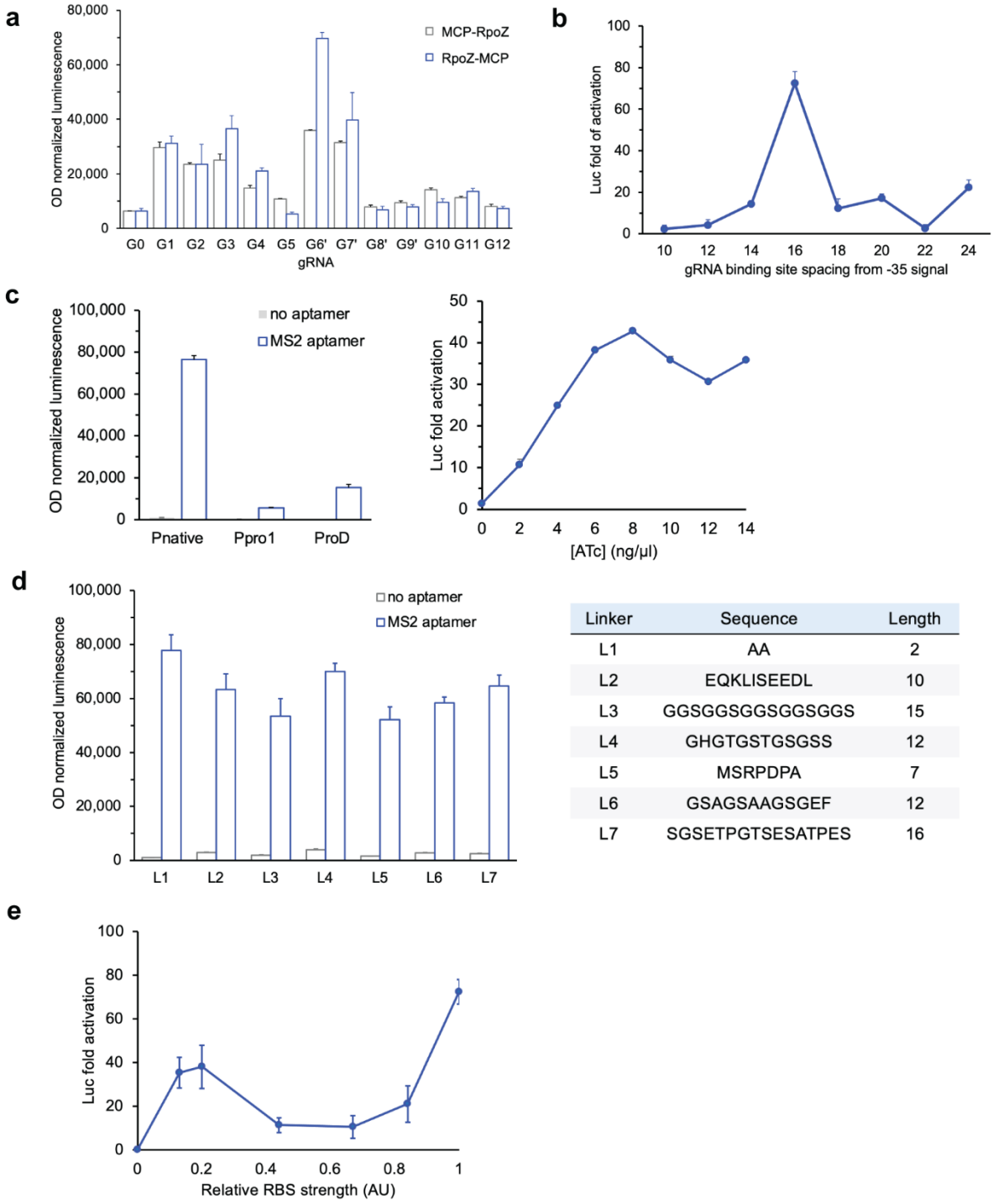
Supplementary Table 5. Oligonucleotides used for NGS.

Oligo	Sequence (5' to 3')
QS_seq18	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGACAGCTAGCTCAGTC
QS_seq19	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTTTTTGTTCGGGCCCAAG

Sequencing adapters are highlighted in yellow

Supplementary Table 6. Aptamers used in this work.

Aptamer	Sequence (5’ to 3’)
MS2	ACAUGAGGAUCACCCAUGU
The Qβ RNA	AUGCAUGUCUAAGACAGCAU
A1	ACAUUAGCUAAGCUAAUGU
A2	ACAUGUACUAAGUACAUGU
A3	ACAUAGGCUAAGCCUAUGU
A4	GGAUAGCUAAGCUACC
A5	UGCUAGCUAAGCUAACA
A6	UGUUAGCUAAGCUAUCA
A7	AGUUAGCUAAGCUAAUU
A8	UUUAGCUAAGCUAGGG
A9	CCUUAGCUAAGCUAAUGG
A10	UGAUAGCUAAGCUACCA
A11	UUCUAGCUAAGCUAGGA
A12	GAGUAGCUAAGCUACCC
A13	UUAUAGCUAAGCUAUAG
A14	CCAUUAGCUAAGCUAAUGG
A15	CCGUUAGCUAAGCUAAUGG
A16	CCUUUAGCUAAGCUAAUGG
A17	CCUCUAGCUAAGCUAUAGG
A18	CCCUUAGCUAAGCUAAGUG
A19	CCCUUAGCUAAGCUAAGGU
A20	CCUUAGCUAAGCUAAUC
A21	UUUCAUAACCGUCCAGAGG

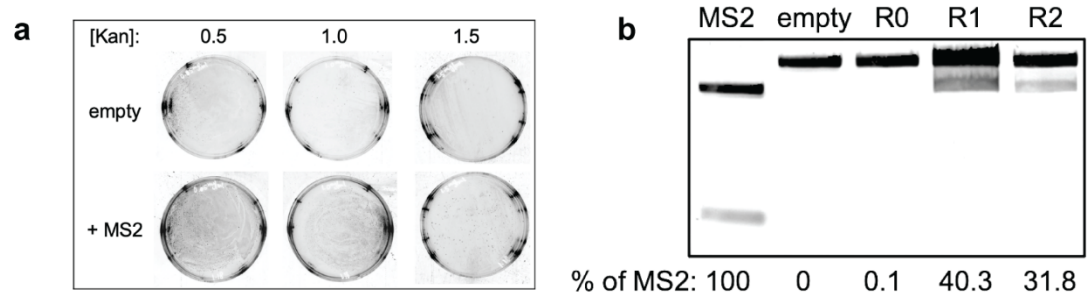


Linker	Sequence	Length
L1	AA	2
L2	EQKLISEEDL	10
L3	GGSGGSGGSGGSGGS	15
L4	GHGTGSTGSGSS	12
L5	MSRPDPA	7
L6	GSAGSAAGSGEF	12
L7	SGSETPGTSESATPES	16

Supplemental Figure 1: CRISPR-hybrid optimizations.

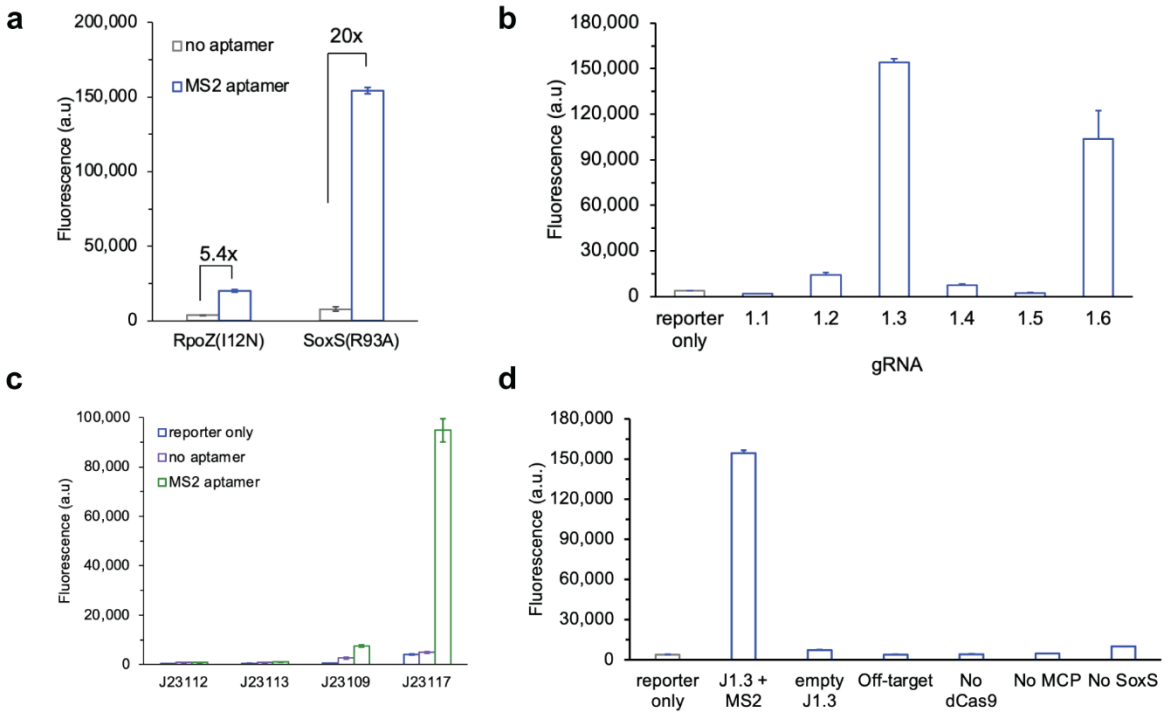
Luciferase expression in *E. coli* was used as reporter gene. **a**, gRNAs targeting the luciferase reporter, as well as a scrambled guide RNA negative control (G0) were tested with MCP-RpoZ and RpoZ-MCP. The highest levels of transcriptional activation were achieved with G6' guide

RNA and RpoZ-MCP. **b**, Guide RNA-promoter binding site spacing strongly affects transcriptional activation levels; G6' binding at 16 base pairs upstream of the -35box in the J23107 promoter resulted in the most robust activation. **c**, Expression of dCas9 was driven by different constitutive promoters (left) and anhydrotetracycline inducible promoter (right). dCas9 expressed from the *S.pyogenes* promoter (Pnative) led to the strongest luciferase transcriptional activation by MS2 aptamer recruitment of RpoZ-MCP. **d**, Tests of seven different linkers between MCP and RpoZ. Linker 1 resulted with the highest reporter activation. **e**, Ribosome-binding site (RBS) modification enables robust modulation of the reporter activation levels. Data represent mean of three independent experiments ± s.d.



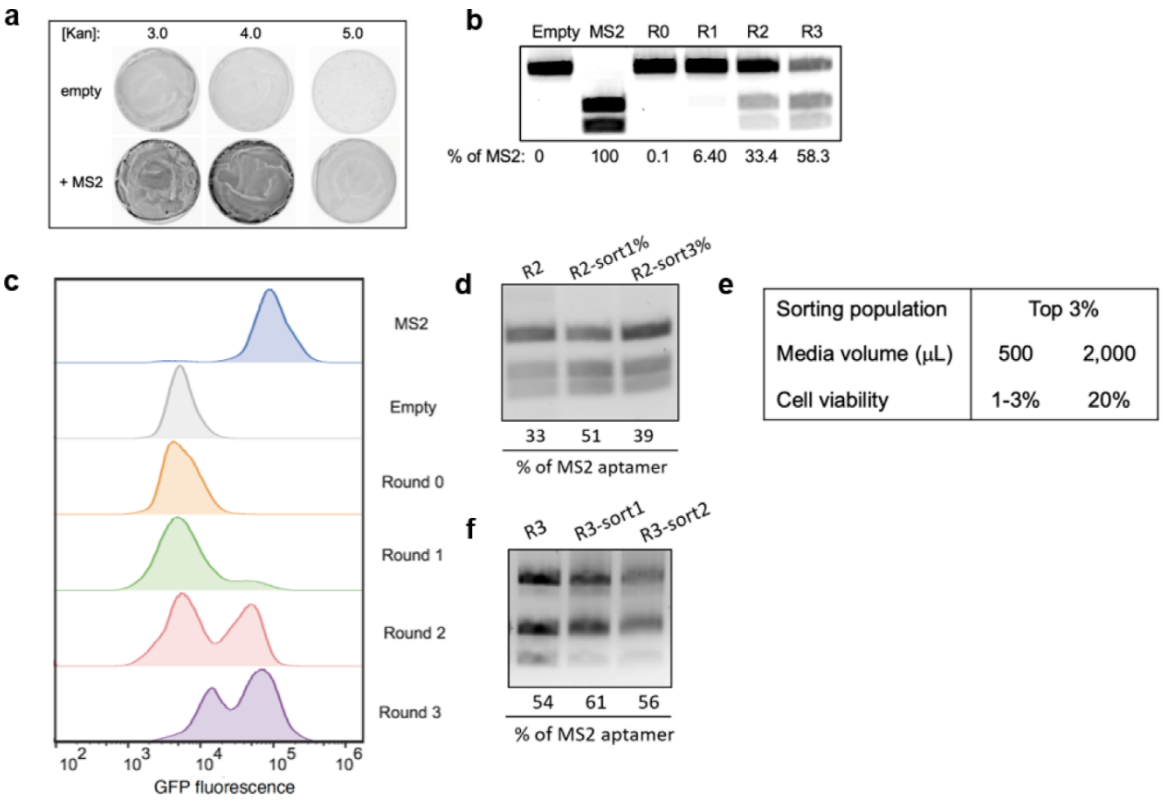
Supplemental Figure 2: Antibiotic mock selection with RpoZ activator.

a, MS2 can recruit RpoZ-MCP to activate the expression of antibiotic resistance gene. e compared with *ScaI* and *AleI* digestion. Survival of cells harboring MS2 aptamer or no aptamer was measured on solid medium supplemented with increasing concentration of kanamycin. **b**, Gene compositions before and after mock selection were compared with *ScaI* and *AleI* digestion.



Supplemental Figure 3: Optimizations with $SoxS_{R93A}$ transcriptional activator.

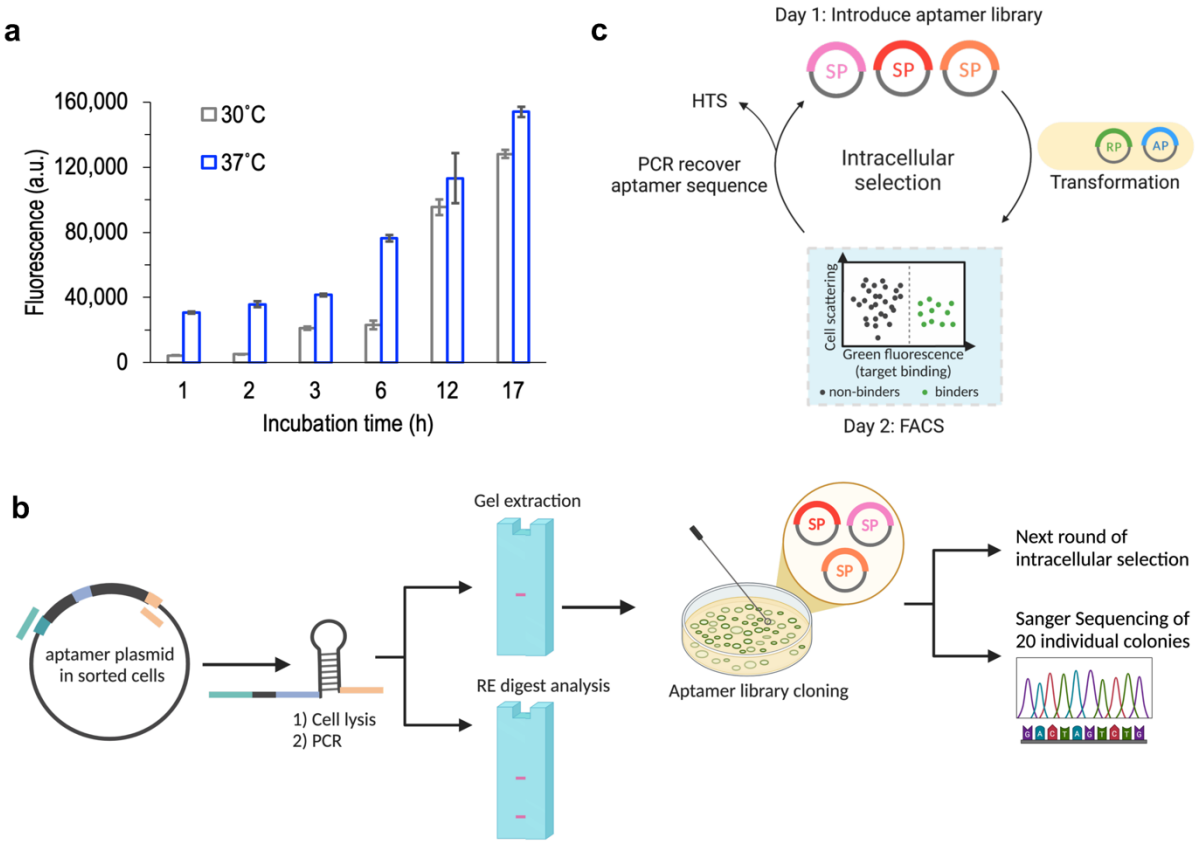
sfGFP was used as reporter gene. **a**, Transcriptional activator $SoxS_{R93A}$ showed 20-fold increase in reporter gene expression. **b**, Guide RNA 1.3 binding at 80 base pairs upstream of transcriptional start site yielded a 56-fold increase in GFP expression, compared to expression from reporter plasmid only. **c**, RNAP-minimal promoter affinity was evaluated with a series of bacterial J23 promoters, and J23117 resulted in the largest reporter fold activation by CRISPRa. **d**, Reporter activation was observed only when all components of CRISPRa are included. Data represent mean of three independent experiments ± s.d.



Sorting population	Top 3%	
Media volume (μL)	500	2,000
Cell viability	1-3%	20%

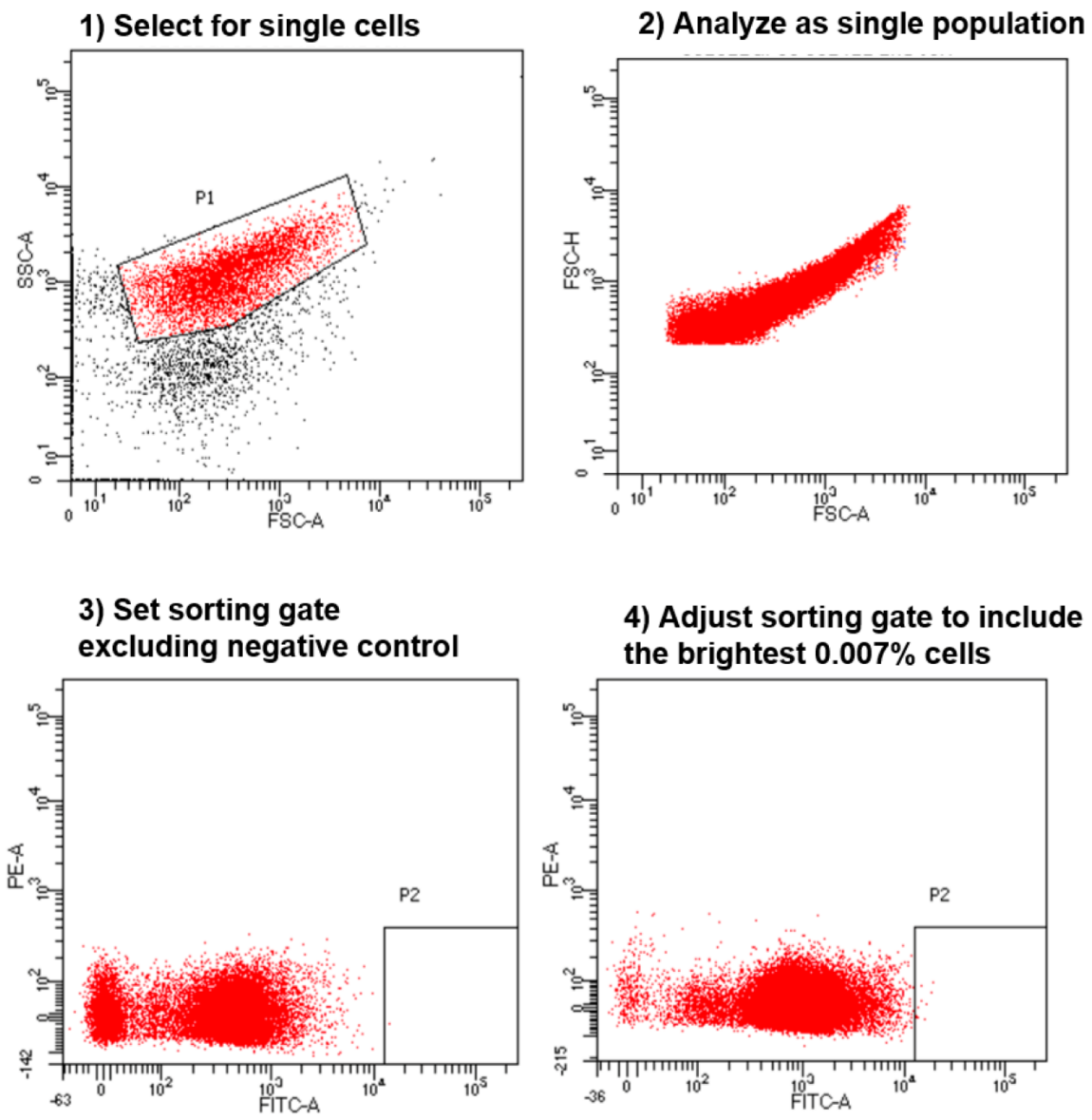
Supplemental Figure 4: Antibiotic mock selection with $SoxS_{R93A}$ activator.

a, Dynamic range of [Kan-GFP] reporter. Survival of cells harboring MS2 aptamer or randomized aptamer was measured on solid medium supplemented with increasing concentration of kanamycin. Cells with MS2 recruiting MCP-$SoxS_{R93A}$ showed robust expression of Kan-GFP resistance gene for cell survival of up to 5.0mg/ml. **b**, Aptamer compositions before and after mock selection were compared with *ScaI* and *AleI* digestion. **c**, Fluorescence histogram overlays showed cell populations after antibiotic selection are distinguishable on flow cytometry. **d**, Sorting most fluorescent 1% cells from round 2 of antibiotic selection yielded higher recovery of MS2 aptamer than sorting the top 3% cells. **e**, Increasing collection media volume improved cell viability post FACS. **f**, Following optimized conditions to sort the most fluorescent 1% cells from round 3 of antibiotic selection resulted in slight higher enrichment of MS2 aptamer.



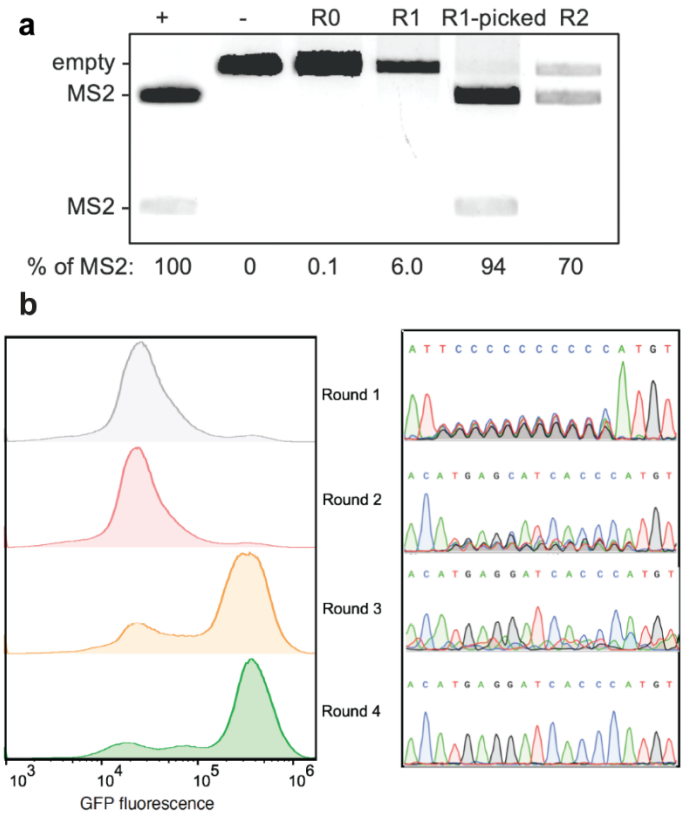
Supplemental Figure 5: FACS optimizations.

a, Transcriptional activation of *sfGFP* reporter by CRISPR-hybrid with MS2 was measured at different *E. coli* growth temperature and time. Maximal *sfGFP* activation was observed from cells grown at 37°C for 17 h. Data represent mean of three independent experiments ± s.d. **b**, Schematic of the Rescue-PCR strategy for maximal recovery of sorted aptamer sequences. **c**, Overview of CRISPR-hybrid selection process coupled with FACS.



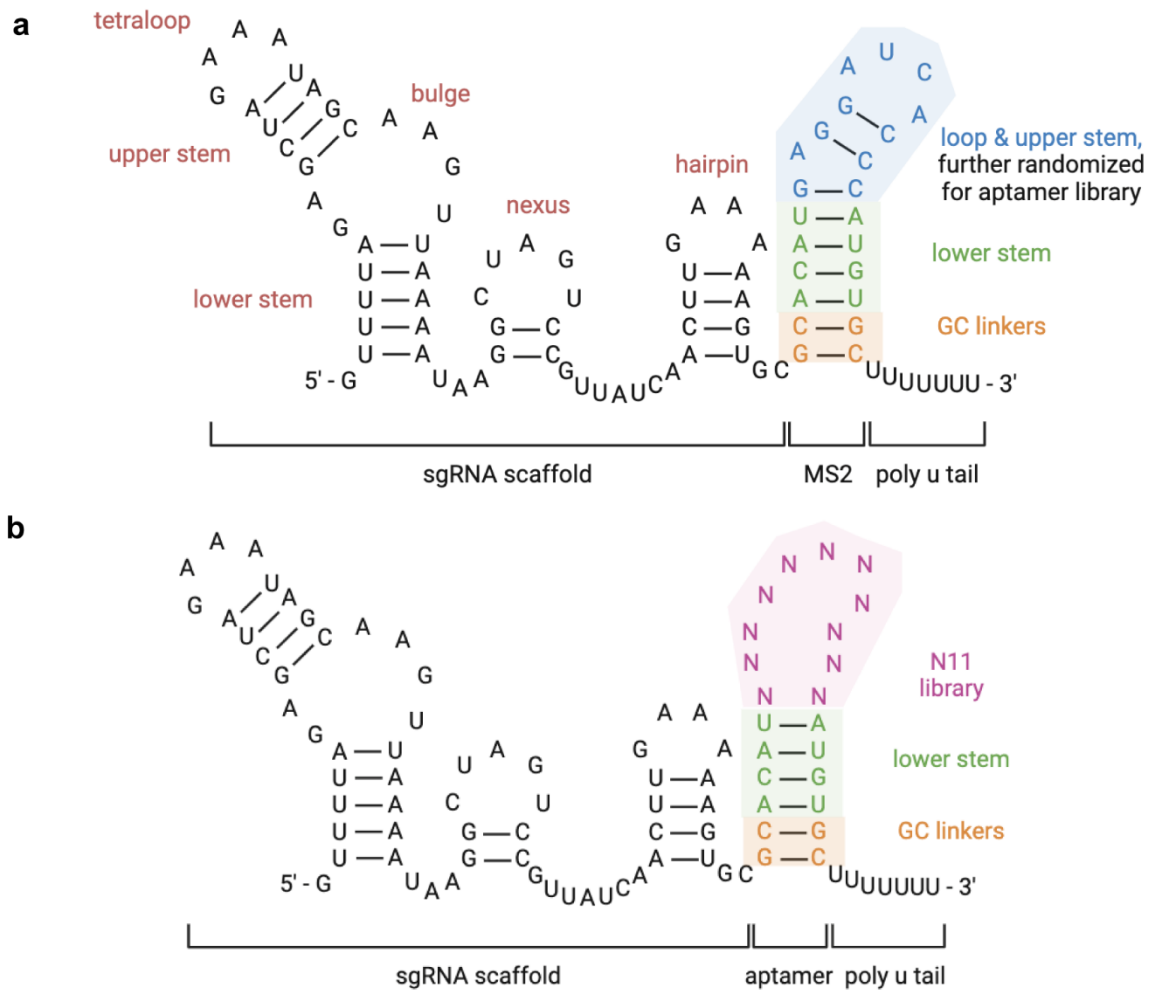
Supplemental Figure 6: FACS gating strategy.

Single bacterial cells were selected **(1)** by applying a gate (P1) on the FSC-A vs. SSC-A plot. These cells were analyzed **(2)** as single population on the FSC-A vs. FSC-H plot. The GFP fluorescence of the single cells were analyzed in the FITC channel and plotted as FITC-A vs. PE-A. Color compensation was applied to correct potential GFP fluorescence spilled over to the PE channel. A sorting gate (P2) was applied **(3)** by excluding cells carrying empty sgRNA (no aptamer), which served as negative control. Then, the sorting gate was adjusted **(4)** to the right and only the brightest 0.007% of cells were collected.



Supplemental Figure 7: Mock Selection.

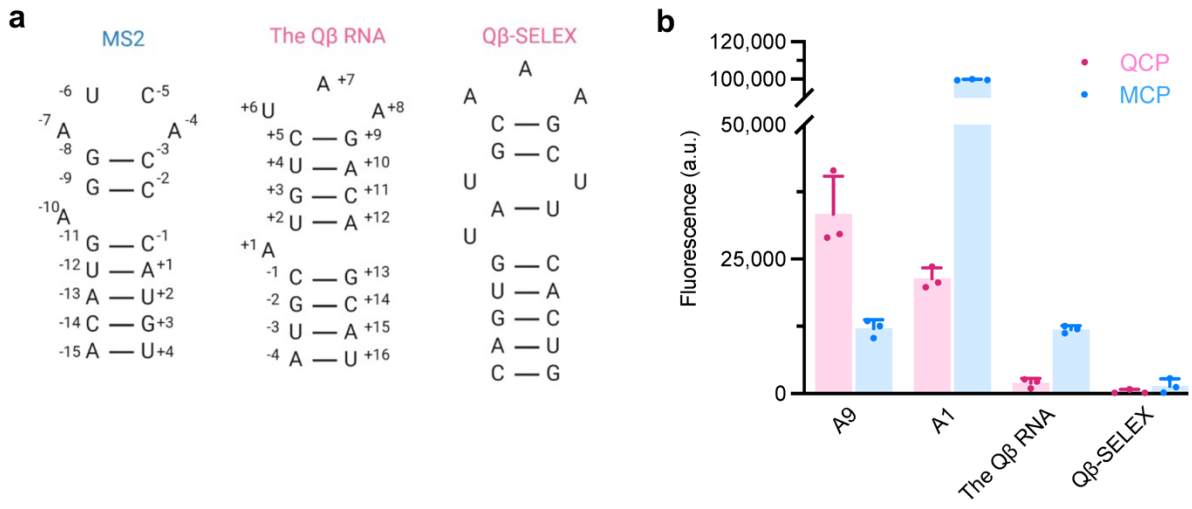
a, Gene compositions before and after mock selection were compared with *ScaI* and *AleI* digestion, showing ≥700-fold enrichment of MS2. R1-picked: R2 done by hand-picking colonies from R1. **b**, FACS histogram overlays of four rounds selection of the aptamer library in cells expressing MCP-SoxS.



Supplemental Figure 8: Design of aptamer library.

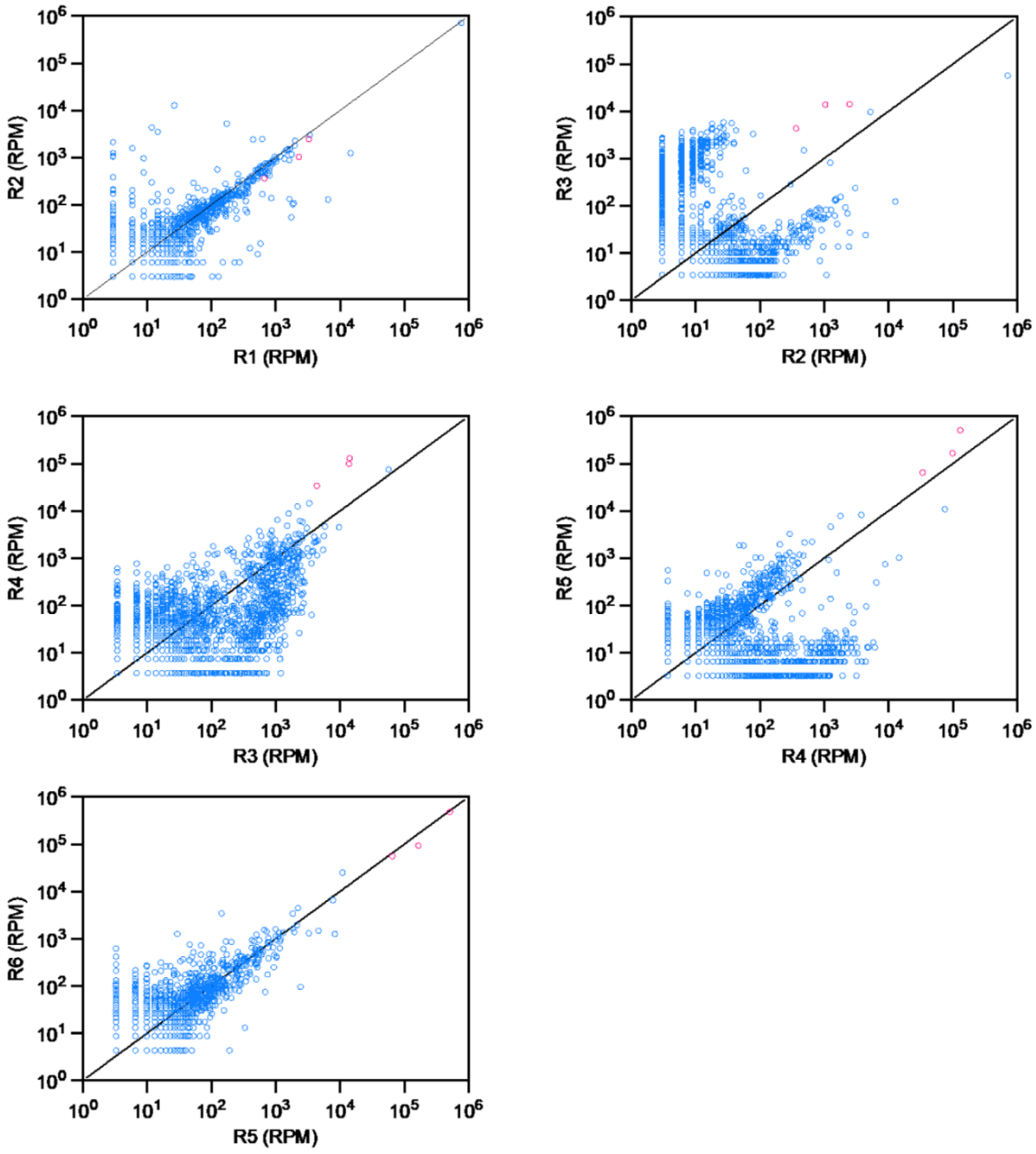
a, Schematic of sgRNA scaffold extended with MS2 aptamer on the 3'-end with two GC linkers.

b, The loop and upper stem region of MS2 aptamer, 11 nucleotides total, is randomized for the N11 library.



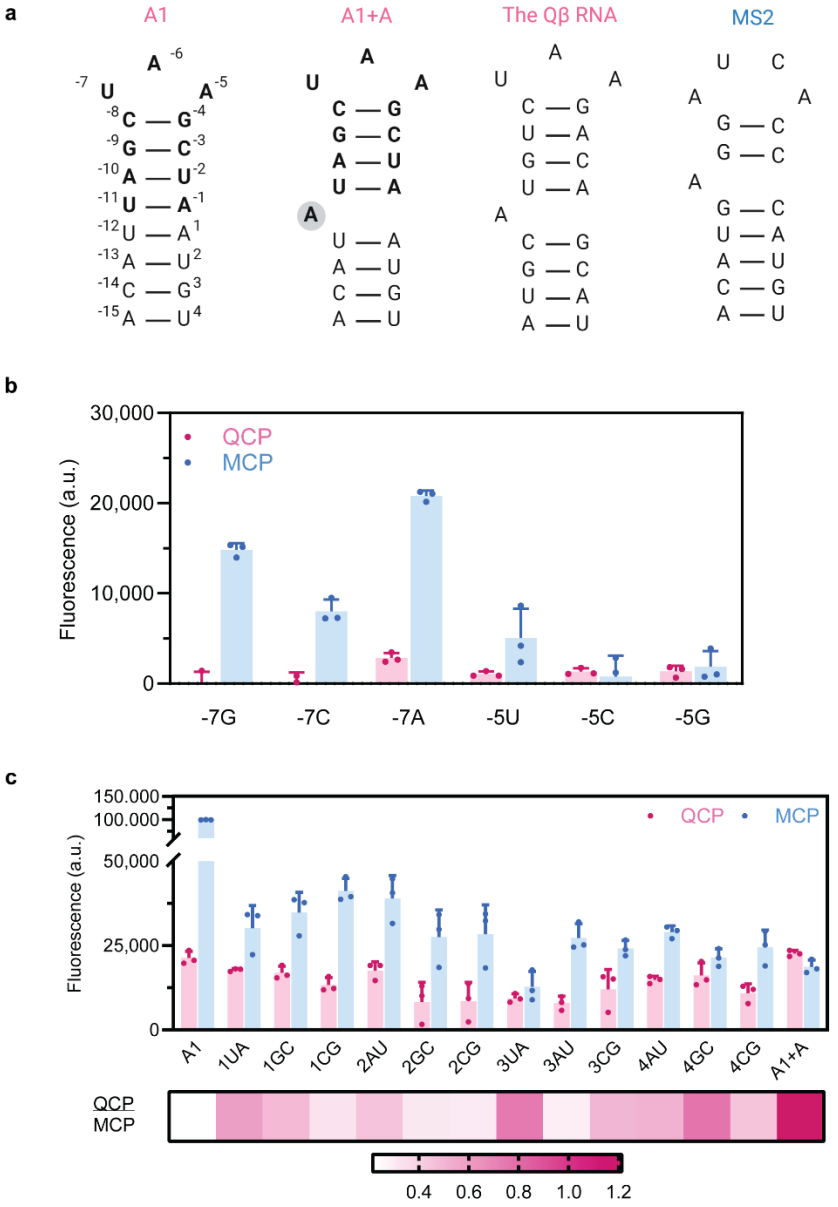
Supplemental Figure 9: Binding activity and selectivity of QCP aptamers.

a, Sequences of consensus aptamers MS2, the Qβ RNA, and SELEX-selected aptamer (Qβ-SELEX). **b**, A9 and A1 showed strong binding activity to QCP, while A9 exhibited reduced affinity for MCP. The Qβ RNA and Qβ-SELEX showed minimum to no binding activity to QCP *in vivo,* but exhibited enhanced nonspecific binding activity to MCP. Binding activity was measured by recruitment of RBP fused to $SoxS_{R93A}$ to activate downstream GFP reporter. Data represent mean of three independent experiments ± s.d.



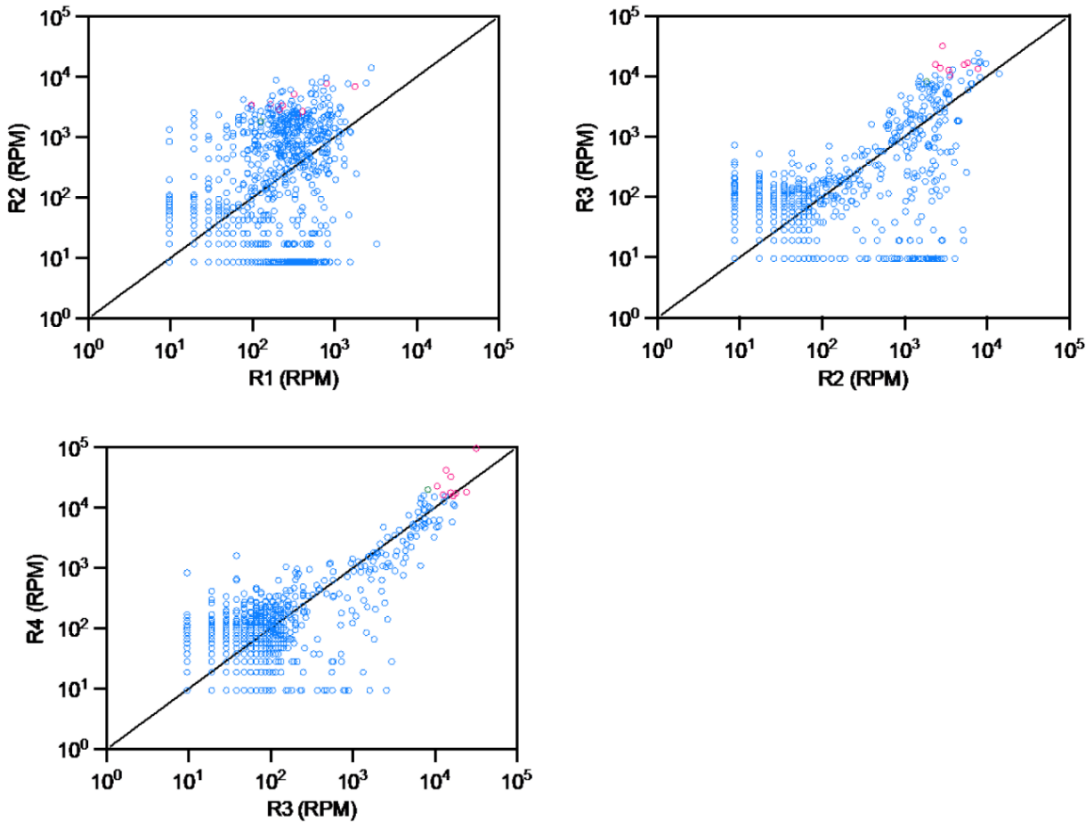
Supplemental Figure 10: Analysis of HTS data from QCP initial intracellular aptamer selection.

The abundances (reads per million, RPM) of sequences from two sequential rounds of library selection are shown as scatter plots. Enrichment patterns observed are consistent with overlay histograms in Fig. 2c, and R5-R6 shows a high degree of correlation. The top three aptamers in R6 are colored in pink.

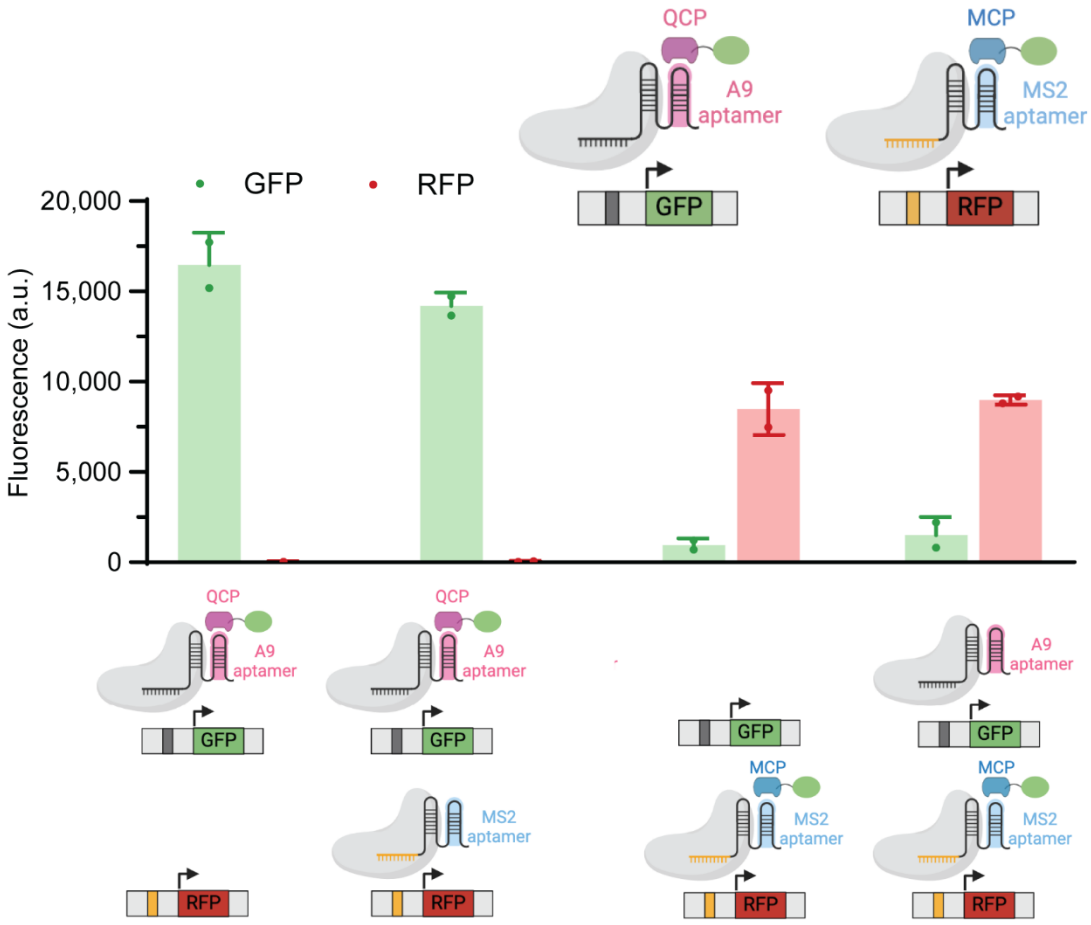


Supplemental Figure 11: A1 mutants binding activity and specificity.

a, Structure of A1 with its bases numbered, A1 variant A1+1, and two consensus aptamer the Qβ RNA and MS2. **b**, Fluorescence measurements of A1 variants with single base mutations in the loop region. Interactions to QCP was completely abolished and nonspecific binding to MCP was greatly reduced. **c**, Fluorescence measurements of A1 variants with single base pair mutations in the lower stem region (top panel). Binding preferences are plotted as ratios of QCP over MCP (bottom panel). Addition of a bulge nucleotide, A1+A variant, showed highest binding activity and specificity to QCP. Data represent mean of three independent experiments ± s.d.



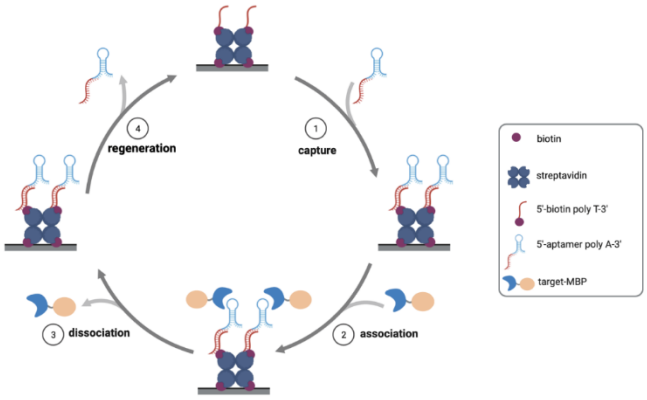
Supplemental Figure 12: Analysis of HTS data from QCP second intracellular selection. The abundances (reads per million, RPM) of sequences from two sequential rounds of the second intracellular selection for QCP are shown as scatter plots. Enrichment patterns observed are consistent with overlay histograms in Fig. 3a, and R3-R4 shows a high degree of correlation. The top ten aptamers in R4 are colored in pink with the most active aptamer, A9, colored in green.



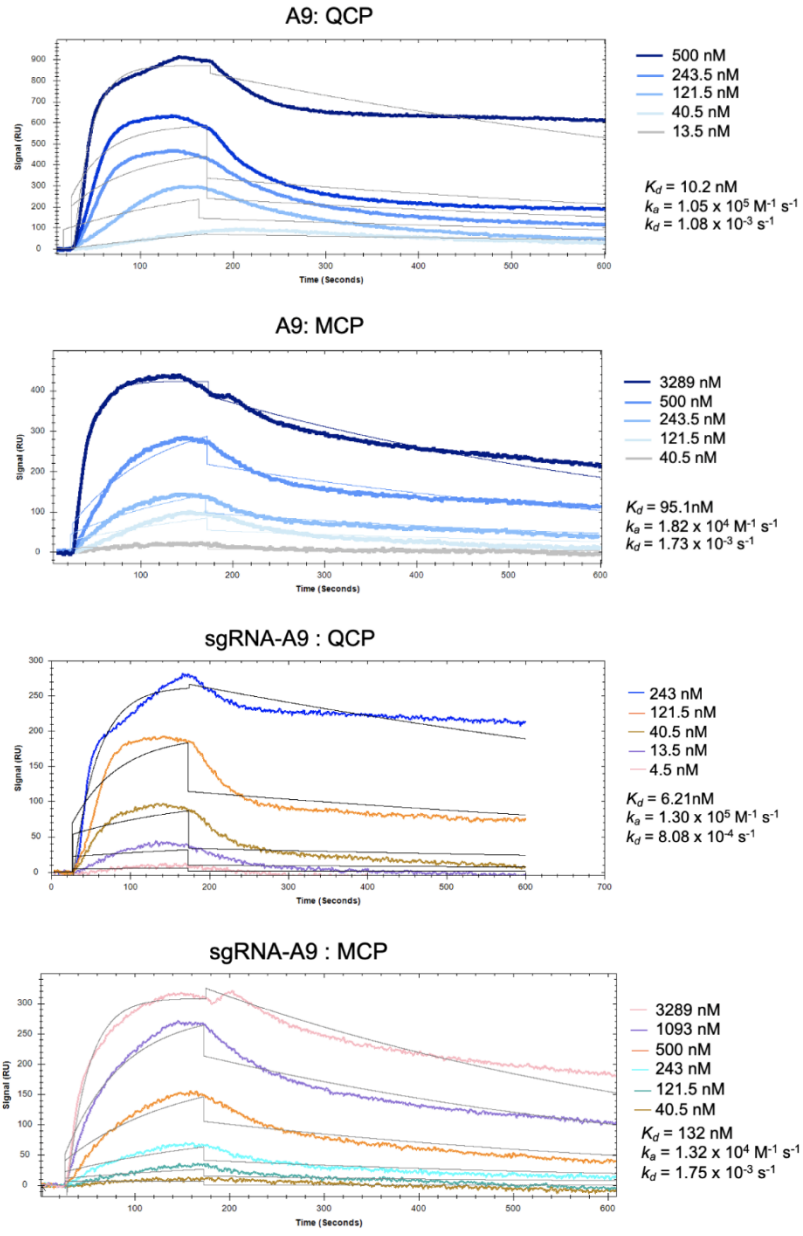
Supplemental Figure 13: Bacteria CRISPR activation of dual-fluorescent reporter with dual aptamers.

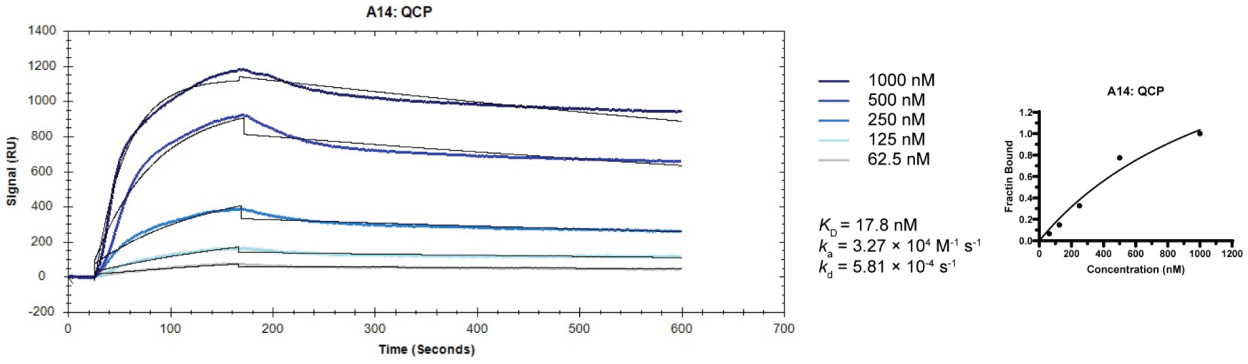
Two CRISPR constructs with programmed sgRNA-aptamer chimeras recruit specific RBP-transcriptional activator fusion proteins to target green or red fluorescent reporter (top panel). Fluorescence measurements of the designed dual CRISPR activation constructs with varied components (bottom panel). Values are mean ± s.d. ($n = 2$ independent replicates).

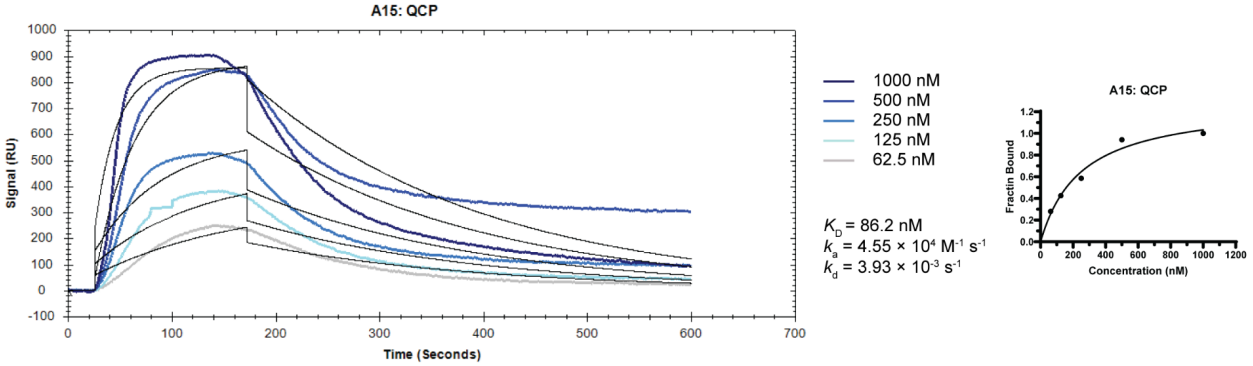
a

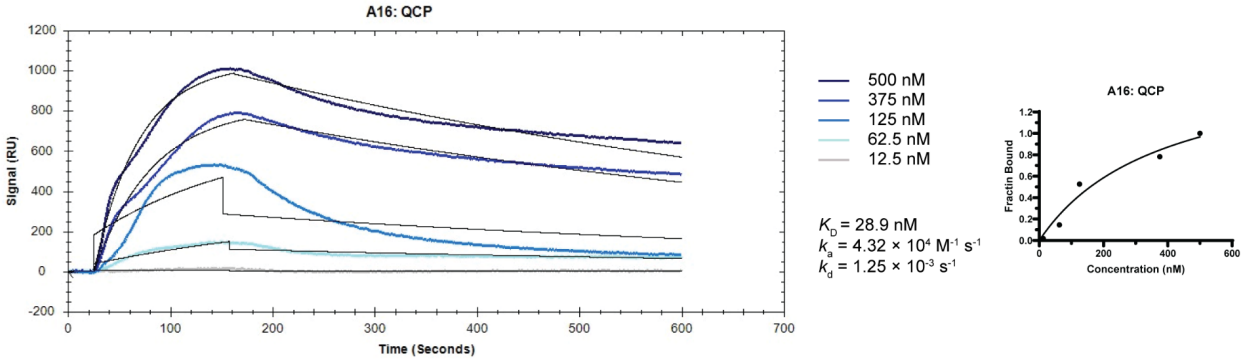

1
capture
2
association
3
dissociation
4
regeneration
biotin
streptavidin
5'-biotin poly T-3'
5'-aptamer poly A-3'
target-MBP

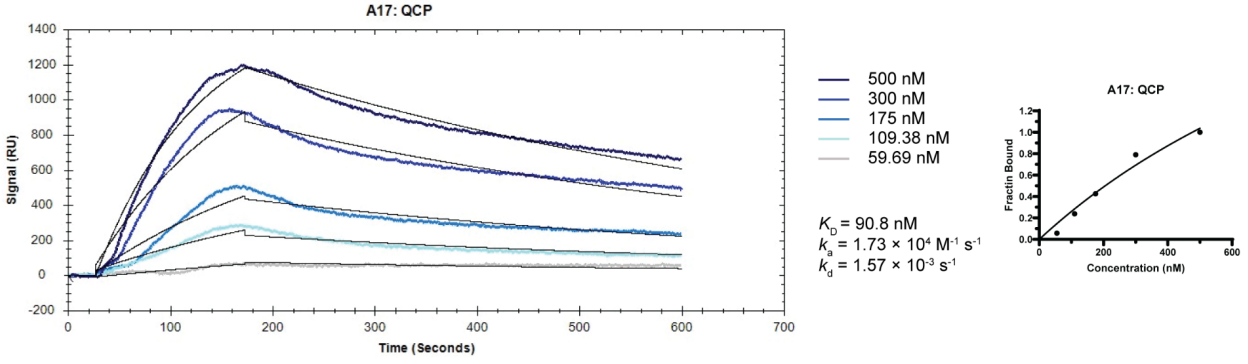
b

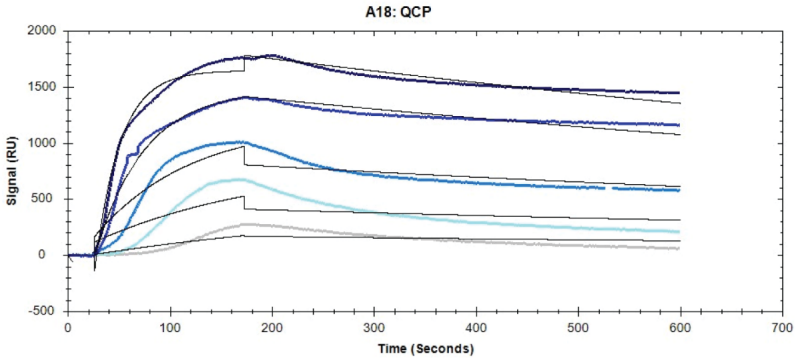

A9: QCP
Signal (RU)
Time (Seconds)
500 nM
243.5 nM
121.5 nM
40.5 nM
13.5 nM
K_d = 10.2 nM
k_a = 1.05 x 10^5 M^{-1} s^{-1}
k_d = 1.08 x 10^{-3} s^{-1}
A9: MCP
3289 nM
500 nM
243.5 nM
121.5 nM
40.5 nM
K_d = 95.1nM
k_a = 1.82 x 10^4 M^{-1} s^{-1}
k_d = 1.73 x 10^{-3} s^{-1}
sgRNA-A9 : QCP
243 nM
121.5 nM
40.5 nM
13.5 nM
4.5 nM
K_d = 6.21nM
k_a = 1.30 x 10^5 M^{-1} s^{-1}
k_d = 8.08 x 10^{-4} s^{-1}
sgRNA-A9 : MCP
3289 nM
1093 nM
500 nM
243 nM
121.5 nM
40.5 nM
K_d = 132 nM
k_a = 1.32 x 10^4 M^{-1} s^{-1}
k_d = 1.75 x 10^{-3} s^{-1}


A14: QCP
Signal (RU)
Time (Seconds)
1000 nM
500 nM
250 nM
125 nM
62.5 nM
K_D = 17.8 nM
k_a = 3.27 × 10^4 M^{-1} s^{-1}
k_d = 5.81 × 10^{-4} s^{-1}
A14: QCP
Fractin Bound
Concentration (nM)

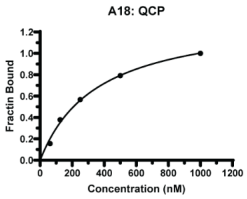

A15: QCP
Signal (RU)
Time (Seconds)
1000 nM
500 nM
250 nM
125 nM
62.5 nM
K_D = 86.2 nM
k_a = 4.55 × 10^4 M^{-1} s^{-1}
k_d = 3.93 × 10^{-3} s^{-1}
A15: QCP
Fractin Bound
Concentration (nM)

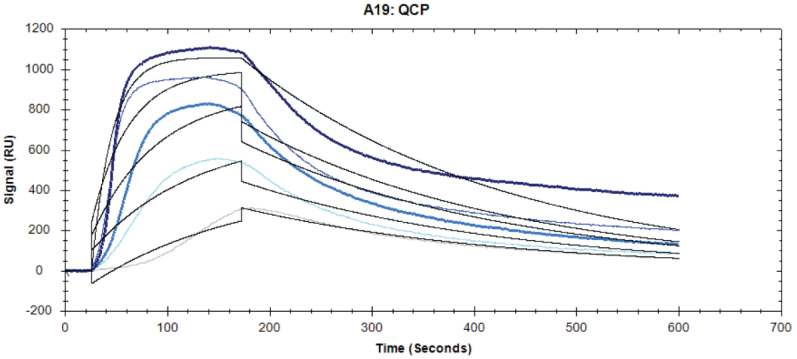

A16: QCP
Signal (RU)
Time (Seconds)
500 nM
375 nM
125 nM
62.5 nM
12.5 nM
K_D = 28.9 nM
k_a = 4.32 × 10^4 M^{-1} s^{-1}
k_d = 1.25 × 10^{-3} s^{-1}
A16: QCP
Fractin Bound
Concentration (nM)


A17: QCP
Signal (RU)
Time (Seconds)
500 nM
300 nM
175 nM
109.38 nM
59.69 nM
K_D = 90.8 nM
k_a = 1.73 × 10^4 M^{-1} s^{-1}
k_d = 1.57 × 10^{-3} s^{-1}
A17: QCP
Fractin Bound
Concentration (nM)

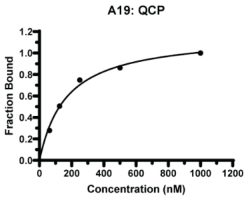


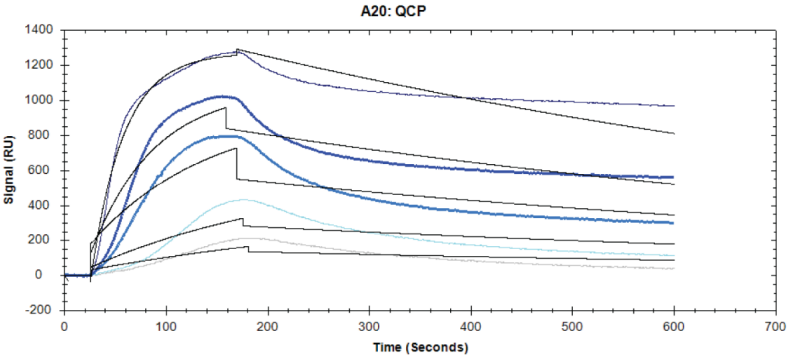
K_D = 16 nM
k_a = 4.01 × 10^4 M^{-1} s^{-1}
k_d = 6.43 × 10^{-4} s^{-1}



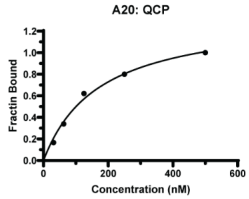


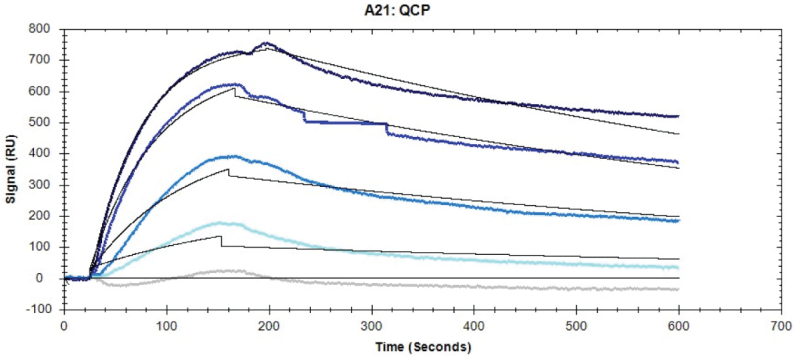
K_D = 80.4 nM
k_a = 4.78 × 10^4 M^{-1} s^{-1}
k_d = 3.85 × 10^{-3} s^{-1}



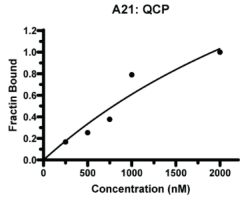


K_D = 18 nM
k_a = 6.01 × 10^4 M^{-1} s^{-1}
k_d = 1.08 × 10^{-3} s^{-1}



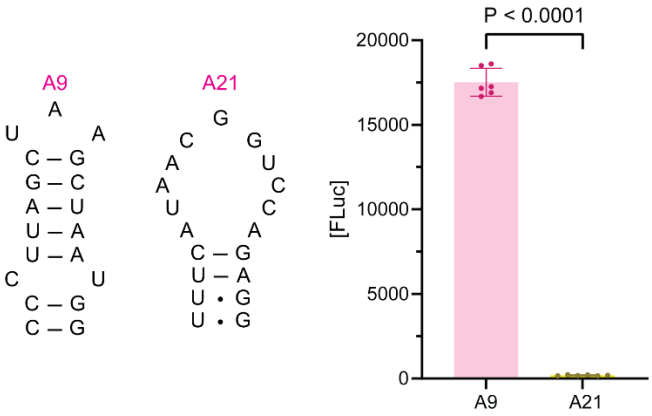


K_D = 119 nM
k_a = 9.8 × 10^3 M^{-1} s^{-1}
k_d = 1.16 × 10^{-3} s^{-1}



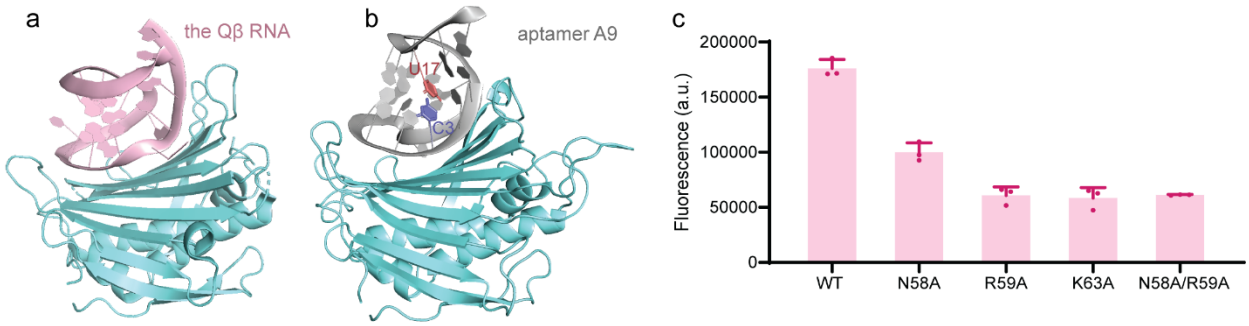
Supplemental Figure 14: SPR kinetic measurements.

a, Illustration of SPR design. Biotinylated poly-T ssDNA is immobilized to streptavidin binding on biotin sensor. RNA aptamer extended with poly-A tail is captured onto the sensor by hybridizing with poly-T. Binding affinity of RNA aptamer with specific target protein is measured by recording the association and dissociation kinetics. Sensor is regenerated by removal of poly-A RNA aptamer for next cycle. **b**, Binding curves of A9, sgRNA-A9 and A9 variants with QCP and MCP, respectively.



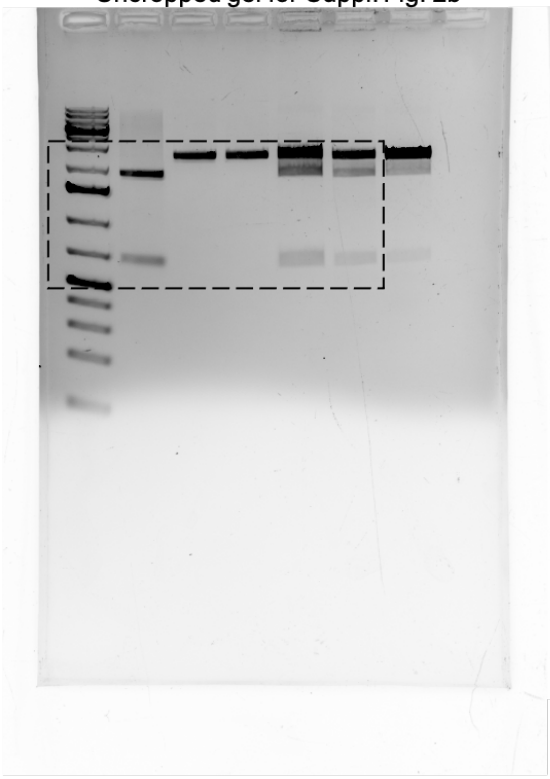
Supplemental Figure 15: Scrambled A9 binding activity in HEK293T cells.

A9 shows high specificity for its cognate target QCP compared to an A9-scrambled sequence. Binding activities were measured using luciferase reporter, ($n = 6$ technical replicates). Values are mean ± s.d., *P* values were determined by two-tailed unpaired Student's *t*-test.

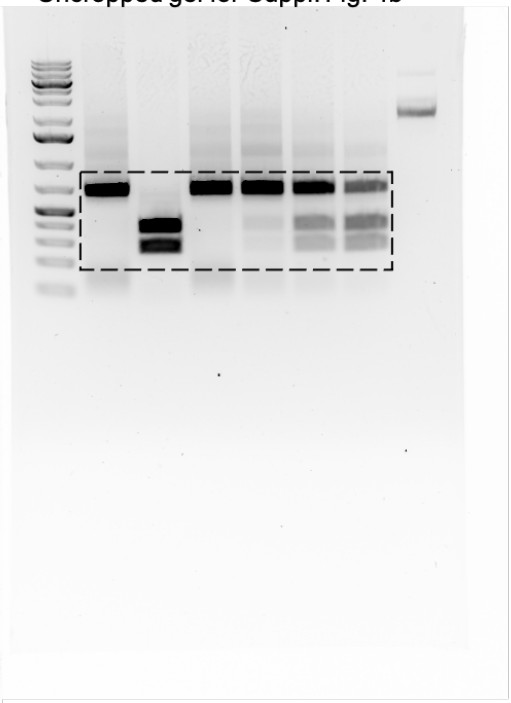


Supplemental Figure 16: QCP-A9 complex structure and binding activity of QCP mutants. **a**, Structure of the Qβ RNA (left, from PDB ID 4L8H) binding to the QCP. **b**, Structure of A9 aptamer binding to the QCP. C3 and U17 are represented in blue and red. The overall binding pose of A9 is similar to that of the Qβ RNA. **c**, Fluorescence measurement of *E. coli* cells expressing A9 and mutants of QCP. Mutations to N58, R59, and K63 of QCP all led to significantly reduced fluorescence, indicating that these residues are essential for the A9-QCP interaction. Data represent mean of three independent experiments ± s.d.

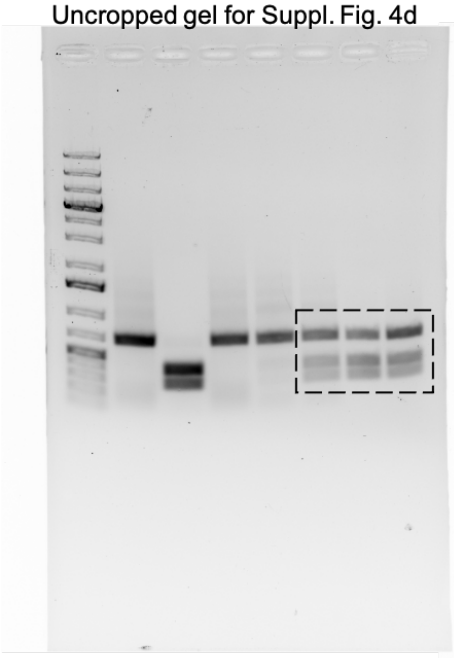
Uncropped gel for Suppl. Fig. 2b



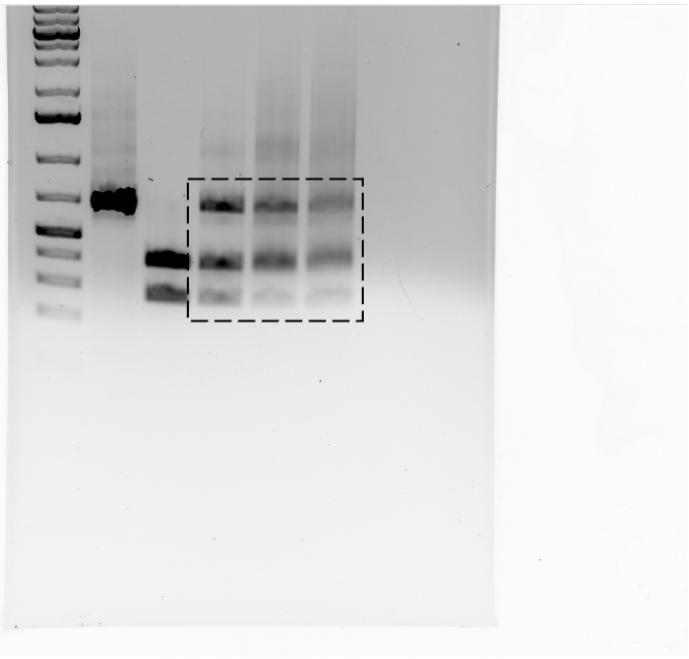
Uncropped gel for Suppl. Fig. 4b



Uncropped gel for Suppl. Fig. 4d



Uncropped gel for Suppl. Fig. 4f



Uncropped gel for Suppl. Fig. 7a