nature portfolio

Peer Review File

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

Corresponding Author: Professor Jia Niu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript “CRISPR-Hybrid: A CRISPR-Mediated Intracellular Directed Evolution Platform for RNA Aptamers” described:
(1) a new experimental strategy for discovering functional and specific aptamers. This strategy is an in vivo intracellular aptamer selection process, which differ from the conventional in vitro “SELEX” approach. The main advantage of the in vivo approach is the high functionality of the identified aptamers.
(2) the CRISPR-Hybrid approach allowed for both transcriptional activation and repression at respective targeted genes in one single procedure. This multiplexed genome editing is an original approach different from previously described CRISPRi+a approach (doi 10.1101/gr.275607.121).
Optimization procedures are well documented and I recommend its publishing after revision.
Please some comments below:

- The authors wrote that the conventional SELEX approach is time-consuming and labor-intensive. Could the author bring an estimate of time for their aptamers selection process?
- Supplemental Figure 5b, why the authors used PE-A signal instead of FSC or SSC signal?
- Do the authors control cell viability during FACS analysis for selection process?
- Figure 3d, if possible, could the authors provide results with control/scrambled aptamer in addition to A9 aptamers to control unspecific binding affinity?
- Figure 4c, same comment, when possible, could the author provide results with a scrambled aptamer in addition to A9?
- Figure 4c, if possible, could the author confirm that the transcriptional activation and repression of ASCL1 and XIST1 is functional using an additional approach, like Western blot analysis for example?
- In the discussion, the authors suggested that this approach may be suitable for recruiting endogenous effector protein. Could the author provide an example showing that this approach may work with endogenous RBP?

Reviewer #2

(Remarks to the Author)
The manuscript by Su-Tobon et al. “CRISPR-Hybrid: A CRISPR-Mediated Intracellular Directed Evolution Platform for RNA Aptamers” reports an in-vivo bacterial RNA aptamer selection system. This system uses CRISPR activation as a functional readout for aptamer-mediated RNA binding protein recruitment. The authors show that their screening approach can enrich MS2 aptamer hairpin variants that recruit the RNA binding protein MCP as effectively as the wild-type, unmodified MS2 aptamer. The authors show that this system can be used to select functional library variants that can recruit a second RNA binding protein, QCP. Through a second round of selection, they are able to engineer higher specificity for QCP over MCP. They show that their best performing hairpin has better selectivity and in vivo functionality than the endogenous RNA hairpin for QCP. Lastly, the authors show simultaneous activation and repression of two independent gene targets in mammalian cells with RBPs fused to different effectors (where one of these is recruited using the newly selected hairpin).

The manuscript addresses a timely need: the field as a whole has lacked the ability to expand the number and variety of orthogonal CRISPR systems because of a lack of tools for effector protein recruitment. The development of new orthogonal RNA aptamer-RBP pairs could help address this need. Unfortunately, although the authors outline an approach for

generating such pairs, the data presented in the manuscript leave doubts about the generalizability, and thus impact, of the method. The concern is that the authors may have already reached the limitations of what can be achieved with this approach. At present, the manuscript does not appear to be suitable for publication in Nature Communications. Regardless of whether it is published here, or elsewhere, we hope the authors will find the following suggestions useful for improving the manuscript.

Major Issues

1. One of the central issues for in vivo aptamer selections has always been the limited sequence diversity that can be sampled compared to purely in vitro approaches. Unless the sequence diversity is generated within the cell itself, the diversity of the library is limited by to the transformation efficiency (10^8-10^9 in E. coli). Because the diversity of the library is directly related to the likelihood of generating functional aptamers from a selection, this means that the affinities and specificities of the aptamers generated this way will also be limited. This limitation is further exacerbated by the need to position the CRISPR-recruited activator protein such that effective transcriptional activation and thus selection of the aptamer can be achieved. Given these constraints, it is perhaps not surprising that the newly selected hairpins were so similar to the original hairpins in terms of sequences and structures – surely this was one of the reasons that overall there was so much cross-talk in binding. As presented, it is difficult to see that this approach could be generalized to produce truly orthogonal aptamers to the variety of potential targets that would be of interest. Ideally, results from additional selections to RBP targets more distinct than the ones chosen here would be presented. At a minimum, these considerations deserve much more attention in the discussion.

2. Continuing point #1 above, in the discussion, the authors explain that “Beyond bacteriophage coat proteins, the intracellular selection system developed in this work can be used to discover aptamers capable of recruiting endogenous effector proteins, such as transcription factors, epigenetic editors and readers, kinases and phosphatases, DNA/RNA repair enzymes, translocation regulators, etc.” Given the theoretical and practical challenges of generating aptamers through this kind of in vivo selection, such claims are difficult to support in the absence of specific evidence that this is possible. Taken together with point #1, enthusiasm for the approach and manuscript is significantly diminished. As above, the authors develop a single (albeit orthogonal) aptamer-RBP pair through this approach, and additional discussion of how this methodology could be generalized to a platform for other aptamer-RBP pairs is needed.

3. The authors state that there is a limited set of mutually orthogonal apamer-RBP pairs, but fail to cite key literature in this area, e.g., Fukunaga-2022 https://doi.org/10.1093/nar/gkab527 and Kirkpatrick-2020 https://doi.org/10.1021/acssynbio.0c00012.

4. The field has struggled with finding orthogonal RBP / RNA hairpin pairs for multiplexed CRISRPa/i in bacteria as well as in mammalian systems. The authors claim that their RNA aptamer-RBP pairs would be orthogonal in both mammalian cells and bacteria but only test orthogonality in mammalian cells. The authors selected the aptamers with an E. coli system but did not test their recruitment abilities in bacteria. It has been previously reported that RBP recruitment in bacteria is more challenging, so testing their RNA aptamer-RBP pairs would be necessary to support this claim and increase the potential impact of the paper.

5. The authors state that “cells containing MS2 exhibited higher levels of fluorescence, while cells lacking the aptamer exhibited only background levels of fluorescence” for their antibiotic-sfGFP fusion in supplementary figure 4c. This claim would be on firmer ground if they actually sorted and ran their diagnostic digestion or sequenced the positive population. To this point, it is interesting that the antibiotic selection method was not successful; additional explanation about this would be welcomed.

6. In general, the writing needs to be clearer. e.g. “Twenty colonies were sequenced in each round and the MS2 gene was observed in the second population.” What is the second population? What is R1-picked in Figure 1b?

7. The manuscript would be greatly improved with clearer diagrams and labels in the figures. Currently it is very difficult to understand the figures without careful reading of the manuscript text. Ex. flow plots should be labeled with sfGFP not FITC and it should be indicated somewhere on the figure what aptamer is being used for the selection shown.

Minor Issues

1. For the flow plots, the authors should include positive or negative gates on the flow diagrams. For example, for the first round of library selection in figure 2 it doesn’t look like the percent of positive cells changes at all which is surprising given that you’re sorting on the top 7 out of 100,000.

2. Supplementary Figure 7b, it seems like MCP has some specificity to A1 also in the loop region? What to make of that? Does A1 share similar loop motifs to MS2?

3. Not all CRISPR systems consist solely of a Cas9 and sgRNA. It would be helpful for the authors to specify the CRISPR-Cas9 system in the first paragraph of the introduction.

4. Other linker lengths have been demonstrated to work in CRISPR-hybrid systems. Can authors elaborate on why these particular linkers were chosen over previously demonstrated linkers.

5. The authors should give comments as to why they optimize their MCP-MS2 with RpoZ instead of using other RBP recruitment methods that have previously demonstrated higher transcriptional activation.

6. There are minor grammatical, spelling and typographical errors throughout the manuscript. For example at the beginning of the “initial intracellular selection” section, there are periods instead of commas.

7. Figure 4 where is the comparison with QB and the original aptamer? Perhaps the comparison is there, but it is challenging to find.

Reviewer #3

(Remarks to the Author)
This paper reports on the selection of RNA aptamers against RBPs that can be fused to sdRNA of CRISPR to improve gene regulations. As such the subject of the paper is interesting but I found it very difficult to understand from my perspective (I am more an expert on aptamers). In the case of aptamers selection the paper falls short in their full analytical characterization. The only actual quantitative data regarding the selected aptamers is reported in suppl. Fig. 9 with SPR experiments. However, as this paper is mostly about aptamer selection I would have liked to see a more thorough characterization of the aptamers not only in terms of their affinity but also specificity, cross-reactivity, possible binding mechanism, kinetics, etc. This is missing in the paper. I understand the authors are more interested in the "final" effect of CRISPR activity but I think the characterization of aptamers should be more focused. Also, authors should explain how this approach (directed evolution) gives different results compared to a conventional aptamer selection protocol. How the aptamers selected are different from the aptamers that would have been selected (or have been selected) using SELEX or other more standard selection approaches? At the end the specificity/orthogonality that was searched by the authors is not really that high, so I wonder if a more conventional approach would have given better results.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors made illustrated and convincing answers to my comments. I do not have any more comments regarding this manuscript.

Reviewer #3

(Remarks to the Author)
Unfortunately the authors did not take my comments in serious consideration so I cannot really have any positive remarks about this revision. I remain convinced that the characterization of the aptamers behaviour should have been largerly improved. Too bad the authors were not interested.

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)
The authors this time put more efforts in characterizing their aptamer (only A9). I remain quite skeptical about their overall conclusions. This is particularly evident in Figure 4 results where the authors conclude that the the affinity and specificity of the selected aptamer is due to its specific sequence. However, the mutated aptamers (especially A18 and A20) do not really perform that bad in comparison in terms of affinity. Also, no results are shown for A15 and A19, why?
This said, in view of the others' comments, I am incline to suggest publication of the paper in the present form.

Our response is in blue.

Reviewers' comments:

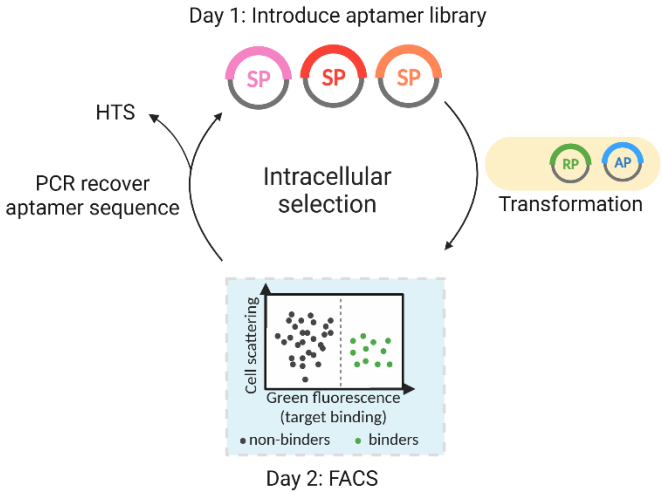
Reviewer #1 (R1):

"The manuscript "CRISPR-Hybrid: A CRISPR-Mediated Intracellular Directed Evolution Platform for RNA Aptamers" described:
(1) a new experimental strategy for discovering functional and specific aptamers. This strategy is an in vivo intracellular aptamer selection process, which differ from the conventional in vitro "SELEX" approach. The main advantage of the in vivo approach is the high functionality of the identified aptamers.
(2) the CRISPR-Hybrid approach allowed for both transcriptional activation and repression at respective targeted genes in one single procedure. This multiplexed genome editing is an original approach different from previously described CRISPRi+a approach (doi 10.1101/gr.275607.121).
Optimization procedures are well documented and I recommend its publishing after revision."

We thank R1 for the supportive comments.

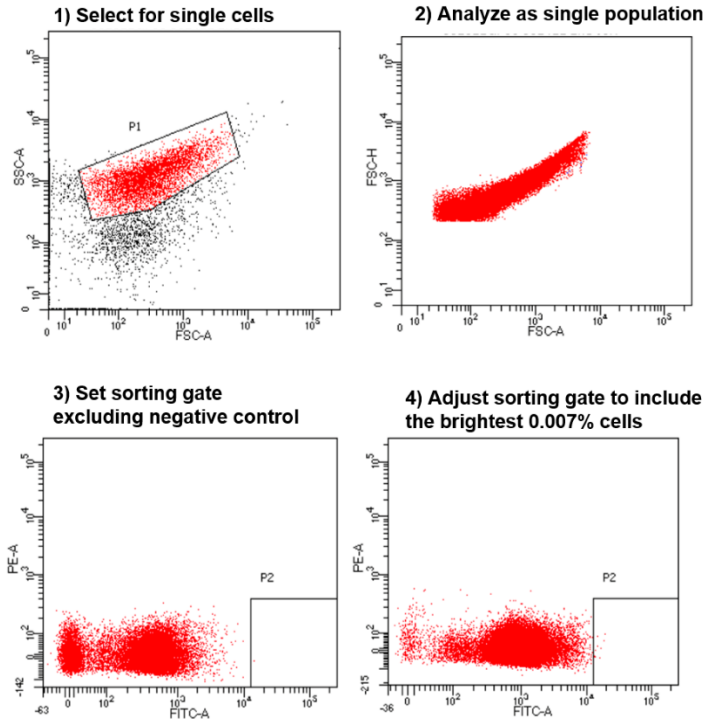
"The authors wrote that the conventional SELEX approach is time-consuming and labor-intensive. Could the author bring an estimate of time for their aptamers selection process?"

We thank R1 for the question. Our selection process takes two days to complete one round. An overview of CRISPR-hybrid selection process coupled with FACS is demonstrated in Suppl. Fig. 5c.



"Supplemental Figure 5b, why the authors used PE-A signal instead of FSC or SSC signal?"

We thank R1 for this question. Both FSC vs SSC and PE-A vs FITC-A signals are utilized in our sorting experiments. Initially, a P1 gate was generated in the FSC vs SSC plot to identify the cell population to be analyzed. Next, the PE-A vs FITC-A plot was used to adjust color compensation, and a P2 gate was generated in this plot to sort the most fluorescent cells. Suppl. Fig. 6 is updated to include detailed FACS plots.

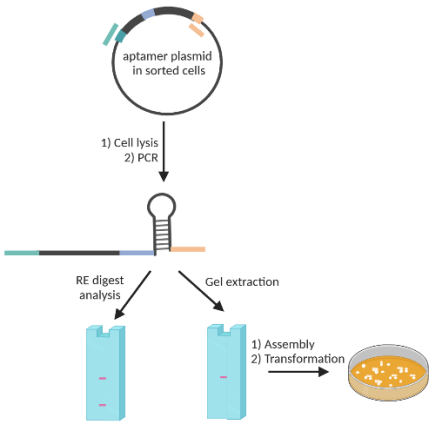


"Do the authors control cell viability during FACS analysis for selection process?"

Cell viability was controlled by collecting sorted cells in different volume of media (Suppl. Fig. 4e).

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

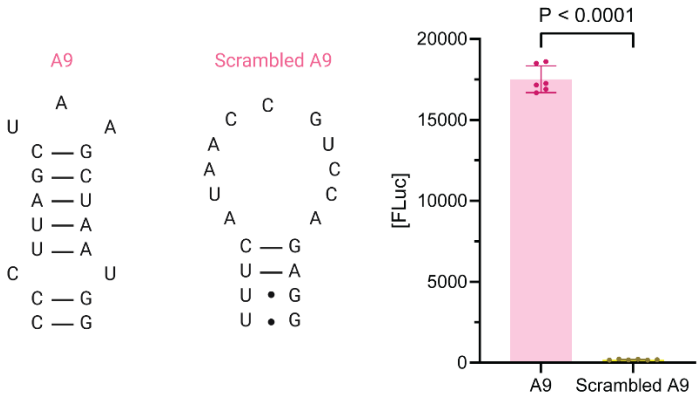
To maximize cell viability, post-sorting cells were not directly amplified; instead, they were lysed after sorting and the DNA fragments encoding the enriched sgRNA-aptamer chimera were PCR amplified, followed by assembly with selection plasmid backbone and reintroduced into selection strain carrying appropriate accessory and reporter plasmids for subsequent rounds, Suppl. Fig. 5b, also copied below, illustrates the workflow of the Rescue-PCR strategy for maximal recovery of sorted aptamer sequences.



"Figure 3d, if possible, could the authors provide results with control/scrambled aptamer in addition to A9 aptamers to control unspecific binding affinity?"

"Figure 4c, same comment, when possible, could the author provide results with a scrambled aptamer in addition to A9?"

We thank R1 for this important and constructive comment. In light of R1's question, we tested the intracellular activity of a scrambled aptamer, scrambled A9, in HEK293T mammalian cell using the [Fluc] transcriptional activation assay described in Fig. 4a. As shown in Suppl. Fig. 14, also copied below, an A9-scrambled sequence was introduced into HEK293T mammalian cells to replace A9 in luciferase reporter activation. The scrambled A9 showed negligible target binding in [FLuc] activation, whereas A9 exhibited significant activity.



"Figure 4c, if possible, could the author confirm that the transcriptional activation and repression of ASCL1 and XIST1 is functional using an additional approach, like Western blot analysis for example?"

We thank R1 for this question. We hope to point out that the RT-qPCR assay described in Fig. 4c has been widely used to determine transcriptional activation and repression of the target genes. For example, please see: *Mol Syst Biol* **16**, e9427 (2020); *Nat. Methods* **15**, 611–616 (2018).

"In the discussion, the authors suggested that this approach may be suitable for recruiting endogenous effector protein. Could the author provide an example showing that this approach may work with endogenous RBP?"

We thank R1 for this question. We believe the CRISPR-hybrid platform is widely applicable to discover active aptamers targeting intracellular proteins, such as transcriptional modulators, or epigenetic effectors. An *in vitro* pre-enrichment process can be implemented to allow an initial library with high diversity to be used, thereby allowing the CRISPR-hybrid system to be used to select larger aptamers. As examples, we are currently applying the CRISPR-hybrid with an *in vitro* pre-enrichment protocol to select aptamers targeting O-GlcNAc transferase (OGT) and L7Ae, a naturally occuring RNA-binding protein that is frequently used in synthetic biology. These experiments and results are beyond the scope of this Communication and will be reported in our future papers.

Reviewer #2 (R2):
"The manuscript by Su-Tobon et al. "CRISPR-Hybrid: A CRISPR-Mediated Intracellular Directed Evolution Platform for RNA Aptamers" reports an in-vivo bacterial RNA aptamer selection system. This system uses CRISPR activation as a functional readout for aptamer-mediated RNA binding protein recruitment. The authors show that their screening approach can enrich MS2 aptamer hairpin variants that recruit the RNA binding protein MCP as effectively as the wild-type, unmodified MS2 aptamer. The authors show that this system can be used to select functional library variants that can recruit a second RNA binding protein, QCP. Through a second round of selection, they are able to engineer higher specificity for QCP over MCP. They show that their best performing hairpin has better selectivity and in vivo functionality than the endogenous RNA hairpin for QCP. Lastly, the authors show simultaneous activation and repression of two independent gene targets in mammalian cells with RBPs fused to different effectors (where one of these is recruited using the newly selected hairpin).

The manuscript addresses a timely need: the field as a whole has lacked the ability to expand the number and variety of orthogonal CRISPR systems because of a lack of tools for effector protein recruitment. The development of new orthogonal RNA aptamer-RBP pairs could help address this need. Unfortunately, although the authors outline an approach for generating such pairs, the data presented in the manuscript leave doubts about the generalizability, and thus impact, of the method. The concern is that the authors may have already reached the limitations of what can be achieved with this approach. At present, the manuscript does not appear to be suitable for publication in Nature Communications. Regardless of whether it is published here, or elsewhere, we hope the authors will find the following suggestions useful for improving the manuscript."

We thank R2 for the detailed and insightful comments. We agree with R2 that this work addressed a timely need by showing that orthogonal aptamer-RNA pairs could be generated using the CRISPR-hybrid system. However, we respectfully disagree with R2

that this system has reached its limit. A main argument of R2 for such a negative conclusion is the concern over the sequence space that our method can be used to explore. This challenge, however, can be overcome by the introduction of an in vitro pre-enrichment procedure prior to intracellular selection. Please see below for detailed explanation.

"One of the central issues for in vivo aptamer selections has always been the limited sequence diversity that can be sampled compared to purely in vitro approaches. Unless the sequence diversity is generated within the cell itself, the diversity of the library is limited by to the transformation efficiency (10^8-10^9 in E. coli). Because the diversity of the library is directly related to the likelihood of generating functional aptamers from a selection, this means that the affinities and specificities of the aptamers generated this way will also be limited. This limitation is further exacerbated by the need to position the CRISPR-recruited activator protein such that effective transcriptional activation and thus selection of the aptamer can be achieved. Given these constraints, it is perhaps not surprising that the newly selected hairpins were so similar to the original hairpins in terms of sequences and structures – surely this was one of the reasons that overall there was so much cross-talk in binding. As presented, it is difficult to see that this approach could be generalized to produce truly orthogonal aptamers to the variety of potential targets that would be of interest. Ideally, results from additional selections to RBP targets more distinct than the ones chosen here would be presented. At a minimum, these considerations deserve much more attention in the discussion."

First of all, we would like to point out that the aptamer identified in this work, A9, is truly orthogonal to existing aptamer-RBP pairs. We have presented results in both *E. coli* (Suppl. Fig. 12) and mammalian cells (Fig. 4). The identification of A9 in this work suggest that the intracellular aptamer selection system – CRISPR-Hybrid – is a powerful platform for identifying orthogonal aptamers to intracellular protein targets.

Second, we would like to point out that higher library diversity does not necessarily guarantee better aptamers; rather, the optimal length of the random region of the aptamer library is target dependent[3]. There has been no evidence in literature to suggest that 10^8 to 10^9 library diversity is insufficient to generate functional aptamers with high affinity and specificity in intracellular selections. To the contrary, both literature and our results have shown the opposite. For example, Thiel et al. have shown that functional aptamers were abundant from a library consisting of a 20-nt random region[4]. Our results also showed that the CRISPR-Hybrid method has enabled the discovery of A9 with low nM affinity and high specificity binding to the target RBP QCP in both bacterial and mammalian cells, from a library where in total 19 nucleotides have been randomized in two selections (see Fig. 3d and Fig. 4). In contrast, although a library of higher diversity (N30, 10^18) was used in Hirao et al.[5] in SELEX to bind QCP, we found that selected aptamer showed had low activity and low specificity in cells (see Fig. 4b of our paper).

Third, R2's argument that the intracellular selection systems are limited by the transformation efficiency can be straightforwardly addressed by the introduction of an in vitro pre-enrichment process of the starting library. Such an *in vitro* pre-enrichment procedure is analogous to the traditional in vitro SELEX process using a N30 or N40 library and a biotinylated/immobilized purified protein target. The pre-enrichment protocol would allow an initial library of high diversity (10^12–10^14) to be reduced to 10^8–10^9 in 2~3 rounds, assuming the enrichment factor of each round is ~100. Therefore, combining in vitro pre-enrichment with intracellular selection by CRISPR-hybrid, a library of high diversity could be selected targeting intracellular proteins. As examples, we are currently applying the CRISPR-hybrid with an *in vitro* pre-enrichment protocol to identify aptamers targeting O-GlcNAc transferase (OGT) and L7Ae, a naturally occurring RNA-binding protein that is frequently used in synthetic biology. These experiments and results are beyond the scope of this Communication and will be reported in our future papers.

Finally, we would like to point out that the significant amount of work described in this paper, *i.e.*, the optimization of the CRISPR-hybrid system, the mock selections, the selections against MCP, the selections against QCP, and the validation of the orthogonality of the aptamers in both bacterial and mammalian cells, has already constitutes a self-contained, complete story that validates the principle of intracellular aptamer selection targeting intracellular proteins. While we agree with R2 that additional selections are desirable, and have already started working to apply CRISPR-hybrid to additional selections, these new experiments and results have exceeded the scope of this Communication and will be reported in our future papers. Requesting all of these new directions to be squeezed into one paper is not only unfair, but also complicates the central message of the paper that makes it hard to read.

In light of R2's comment, we have added the following paragraph in the Discussion section: "Despite these advantages, an intracellular aptamer selection system also has its own intrinsic limitations. As an example, unless high-frequency mutation to the evolving aptamer library is generated within the cell itself, the diversity of the library is limited by the transformation efficiency (10^8–10^9 in *E. coli*). While such a library diversity may be sufficient for some protein targets (*e.g.*, QCP), targeting large intracellular proteins may require a greater number of randomized bases and higher library diversity. In such cases, in vitro pre-enrichment of the initial library using purified proteins via SELEX may become necessary to predispose the library for the desired target and productively reduce library diversity prior to the intracellular selection procedure."

"Continuing point #1 above, in the discussion, the authors explain that "Beyond bacteriophage coat proteins, the intracellular selection system developed in this work can be used to discover aptamers capable of recruiting endogenous effector proteins, such as transcription factors, epigenetic editors and readers, kinases and phosphatases, DNA/RNA repair enzymes, translocation regulators, etc." Given the theoretical and practical challenges of generating aptamers through this kind of in vivo selection, such claims are difficult to support in the absence of specific evidence that this is possible. Taken together with point #1, enthusiasm for the approach

and manuscript is significantly diminished. As above, the authors develop a single (albeit orthogonal) aptamer-RBP pair through this approach, and additional discussion of how this methodology could be generalized to a platform for other aptamer-RBP pairs is needed."

In the discussion, we pointed out that the CRISPR-Hybrid as an intracellular aptamer selection system have the promising potential to be used to discover new aptamers targeting a wide variety of intracellular proteins. As responded above, the challenge of limited sequence space to be explored can be addressed by the introduction of an *in vitro* pre-enrichment protocol prior to the intracellular selection process. Such an *in vitro* pre-enrichment protocol is essentially identical to the well-established SELEX procedure, and can be easily implemented.

"The authors state that there is a limited set of mutually orthogonal apamer-RBP pairs, but fail to cite key literature in this area, e.g., Fukunaga-2022 https://doi.org/10.1093/nar/gkab527 and Kirkpatrick-2020 https://doi.org/10.1021/acssynbio.0c00012."

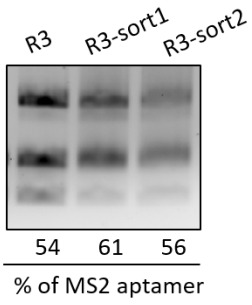
We thank R2 for pointing out these additional references and will cite these additional papers. But we would like to note that the work described in these papers are distinct from our work. Fukunaga-2022 used the *in vitro* SELEX strategy, while Kirkpatrick-2020 is a rationally designed system without selection. Therefore, these published reports does not affect the novelty or significance of this work.

"The field has struggled with finding orthogonal RBP / RNA hairpin pairs for multiplexed CRISRPa/i in bacteria as well as in mammalian systems. The authors claim that their RNA aptamer-RBP pairs would be orthogonal in both mammalian cells and bacteria but only test orthogonality in mammalian cells. The authors selected the aptamers with an E. coli system but did not test their recruitment abilities in bacteria. It has been previously reported that RBP recruitment in bacteria is more challenging, so testing their RNA aptamer-RBP pairs would be necessary to support this claim and increase the potential impact of the paper."

We thank R2 for this important question. We further confirmed the orthogonality of RNA aptamer-RBP pairs in *E. coli* using a dual-color reporter system. We designed the sgRNA-A9 chimera to activate an *sfGFP* reporter, while sgRNA-MS2 chimera activated an *RFP* reporter. In the presence of both MCP and QCP in bacteria, the *sfGFP* reporter is activated by A9-QCP recruitment, while the *RFP* reporter is activated by MS2-MCP recruitment, respectively (Suppl. Fig. 12, also incorporated below).

"The authors state that "cells containing MS2 exhibited higher levels of fluorescence, while cells lacking the aptamer exhibited only background levels of fluorescence" for their antibiotic-sfGFP fusion in supplementary figure 4c. This claim would be on firmer ground if they actually sorted and ran their diagnostic digestion or sequenced the positive population. To this point, it is interesting that the antibiotic selection method was not successful; additional explanation about this would be welcomed."

We thank R2 for this question. We did sort and completed diagnostic digestion of the positive population, and data shows that MS2 aptamer was enriched from 54% in round 3 (R3) of antibiotic selection to 61% after one round of sorting (R3-sort1). However, further enrichment was not observed in subsequent round (R3-sort 2). Please see the attached Suppl. Fig. 4f below. We hypothesized that adequate sorting optimizations including cell viability and sorting gate would improve enrichment, hence we proceeded to implement FACS in our selection strategy for faster turn-around time and completed optimizations from round 0 rather than using a pre-enriched population from antibiotic selection.



“In general, the writing needs to be clearer. e.g. “Twenty colonies were sequenced in each round and the MS2 gene was observed in the second population.” What is the second population? What is R1-picked in Figure 1b?”

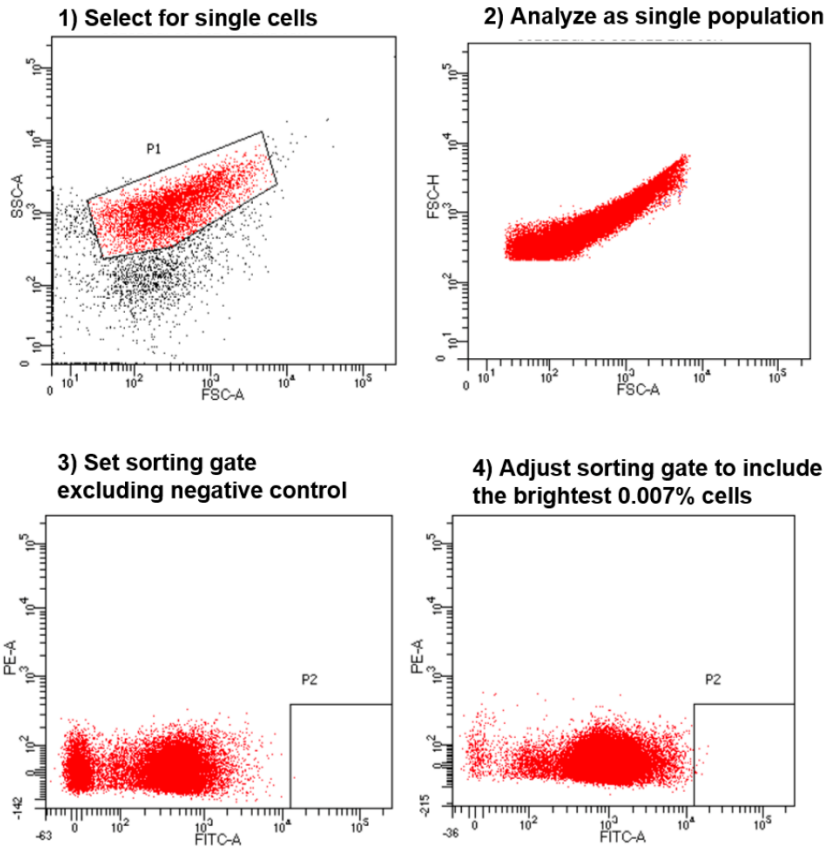
We thank R2 for this question. The second population indicates the fluorescence-increased cell population in the FACS histogram (Fig. 1d left). R1-picked is the R2 selection done by hand-picking colonies from R1. We have re-written the corresponding language to make it easier to follow.

“The manuscript would be greatly improved with clearer diagrams and labels in the figures. Currently it is very difficult to understand the figures without careful reading of the manuscript text. Ex. flow plots should be labeled with sfGFP not FITC and it should be indicated somewhere on the figure what aptamer is being used for the selection shown.”

We thank R2 for this suggestion. FITC is the channel used to measure GFP fluorescence. We have updated the flow plots accordingly. We have also added schemes to each figure in the manuscript to improve their readability.

“For the flow plots, the authors should include positive or negative gates on the flow diagrams. For example, for the first round of library selection in figure 2 it doesn’t look like the percent of positive cells changes at all which is surprising given that you’re sorting on the top 7 out of 100,000.”

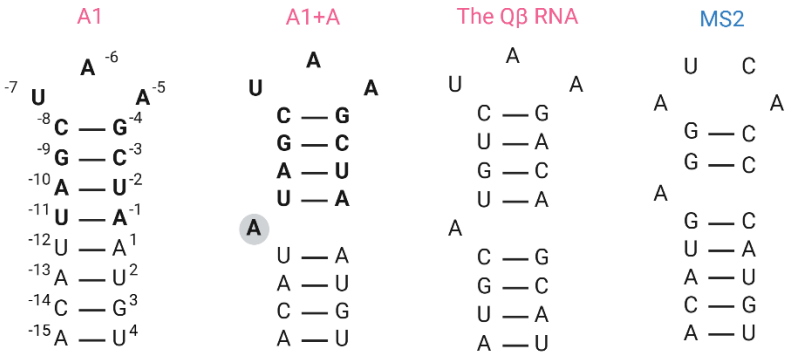
We thank R2 for this suggestion, and have included flow plots showing positive and negative gates in Suppl. Fig. 6, also shown below.



Although we are sorting the top 7 out of 100,000 cells, we do not expect to see a significant enrichment after one round of sorting. While cells containing genes leading to higher finesses are enriched, not all cells express the same amount of proteins in the subsequent culturing. In fact, even cells grown from a single colony could demonstrate heterogeneity in protein expression. Furthermore, due to their small size, FACS sorting of *E. coli* cells often lead to multiple cells being sorted at the same time, causing imperfect enrichment. These negative cells would be de-enriched in the subsequent rounds of sorting.

"Supplementary Figure 7b, it seems like MCP has some specificity to A1 also in the loop region? What to make of that? Does A1 share similar loop motifs to MS2?"

A1 does not share similar loop motifs to MS2. A1 consists of a three-nucleotide loop UAA while MS2 consists of a four-nucleotide loop AUCA. The secondary structure of MS2 and A1 are included in Suppl. Fig. 10a for easier comparison.



"Not all CRISPR systems consist solely of a Cas9 and sgRNA. It would be helpful for the authors to specify the CRISPR-Cas9 system in the first paragraph of the introduction."

We thank R2 for the helpful suggestion and have included the suggested specification in the introduction.

"Other linker lengths have been demonstrated to work in CRISPR-hybrid systems. Can authors elaborate on why these particular linkers were chosen over previously demonstrated linkers."

These linkers were previously tested for CRISPR-dCas9 systems with RpoZ and SoxS transcriptional activators[6,7], and both activators were used in our CRISPR-Hybrid system.

"The authors should give comments as to why they optimize their MCP-MS2 with RpoZ instead of using other RBP recruitment methods that have previously demonstrated higher transcriptional activation."

RpoZ was previously used an *in vivo* CRISPR selection strategy[6], while other transcriptional activators have not been previously used in CRISPR selections.

"There are minor grammatical, spelling and typographical errors throughout the manuscript. For example at the beginning of the "initial intracellular selection" section, there are periods instead of commas."

We thank R2 for the helpful suggestion and have corrected these minor errors.

"Figure 4 where is the comparison with QB and the original aptamer? Perhaps the comparison is there, but it is challenging to find."

Qβ is the original RNA sequence found at the beginning of the replicase cistron of the bacteriophage coat protein Qβ[8]. We revised the nomenclature of this sequence to be "the Qβ RNA" to avoid this confusion.

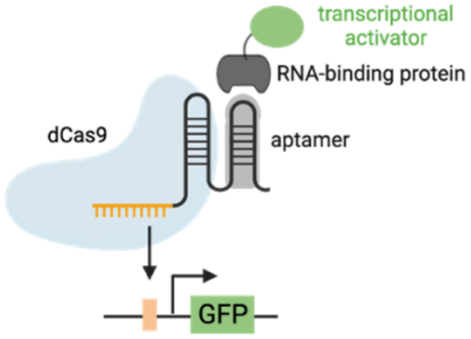
Reviewer #3 (R3):

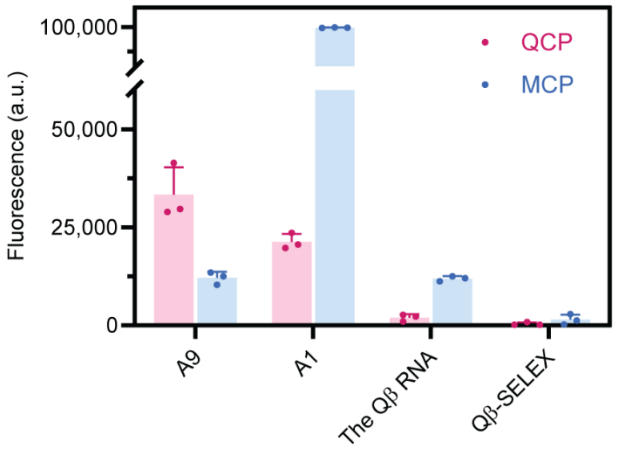
"This paper reports on the selection of RNA aptamers against RBPs that can be fused to sdRNA of CRISPR to improve gene regulations. As such the subject of the paper is interesting but I found it very difficult to understand from my perspective (I am more an expert on aptamers). In the case of aptamers selection the paper falls short in their full analytical characterization. The only actual quantitative data regarding the selected aptamers is reported in suppl. Fig. 9 with SPR experiments. However, as this paper is mostly about aptamer selection I would have liked to see a more thorough characterization of the aptamers not only in terms of their affinity but also specificity, cross-reactivity, possible binding mechanism, kinetics, etc. This is missing in the paper. I understand the authors are more interested in the "final" effect of CRISPR activity but I think the characterization of aptamers should be more focused. Also, authors should explain how this approach (directed evolution) gives different results compared to a conventional aptamer selection protocol. How the aptamers selected are different from the aptamers that would have been selected (or have been selected) using SELEX or other more standard selection approaches? At the end the specificity/orthogonality that was searched by the authors is not really that high, so I wonder if a more conventional approach would have given better results."

We respectfully disagree with R3 that the "characterization of the aptamers not only in terms of their affinity but also specificity, cross-reactivity, possible binding mechanism, kinetics, etc." are missing. In fact, we have performed characterizations of the specificity and orthogonality of our selected aptamer A9 in bacterial cells (Suppl. Fig. 12) and mammalian cells (Fig. 4), as well as in *in vitro* SPR experiments (Fig. 3f), which confirmed the K_d and specificity of our reported aptamer A9. The k_{off} rate of the aptamers were characterized by the SPR experiment shown in Suppl. Fig. 13. We further conducted systematic mutational studies to elucidate the binding mechanism of the aptamers A1 (Fig. 2d and 2e, Suppl. Fig. 10) and A9 (Fig. 3d and 3e). All of these results have specifically addressed the affinity, specificity, orthogonality, kinetics, and binding mechanisms of the selected aptamers *in vitro* and in cells.

Our *in vivo*-selected aptamers are completely different from the aptamers selected using SELEX[5]. The loop motifs, stem, bulged nucleotide sequence and position are all different than reported *in vitro* selected aptamers. Fig 4. in our manuscript shows that our selected aptamer A9 displays significant orthogonality *in vivo*.

In addition, we have tested the conventional SELEX-selected aptamer for QCP, known as Qβ-SELEX, in both *E. coli* cells (Suppl. Fig. 8) and mammalian cells (Fig. 4B). Data show that Qβ-SELEX exhibits no activity *in vivo* at all, while our *in vivo*-selected aptamers A9 exhibit 17-and 11-fold stronger transcriptional activation activity in mammalian cells (see the attached figure below).





Reference

1. Ho, H., Fang, J. R., Cheung, J. & Wang, H. H. Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol Syst Biol* **16**, e9427 (2020).
2. Yeo, N. C. *et al.* An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat Methods* **15**, 611–616 (2018).
3. Komarova, N. & Kuznetsov, A. Inside the Black Box: What Makes SELEX Better? *Molecules* **24**, (2019).
4. Thiel, W. H. *et al.* Nucleotide Bias Observed with a Short SELEX RNA Aptamer Library. *Nucleic Acid Ther* **21**, 253–263 (2011).
5. Hirao, I., Spingola, M., Peabody, D. & Ellington, A. D. The limits of specificity: An experimental analysis with RNA aptamers to MS2 coat protein variants. *Mol Divers* **4**, 75–89 (1998).
6. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).
7. Dong, C., Fontana, J., Patel, A., Carothers, J. M. & Zalatan, J. G. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* **9**, 2489 (2018).
8. Weber, H. The binding site for coat protein on bacteriophage Qβ RNA. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **418**, 175–183 (1976).

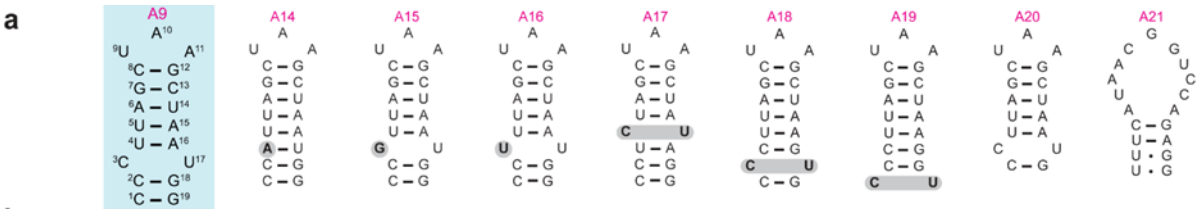
Response to Reviewers

Reviewer #3 (Remarks to the Author):

1. *"This paper reports on the selection of RNA aptamers against RBPs that can be fused to sdRNA of CRISPR to improve gene regulations. As such the subject of the paper is interesting but I found it very difficult to understand from my perspective (I am more an expert on aptamers). In the case of aptamers selection the paper falls short in their full analytical characterization. The only actual quantitative data regarding the selected aptamers is reported in suppl. Fig. 9 with SPR experiments."*

"However, as this paper is mostly about aptamer selection I would have liked to see a more thorough characterization of the aptamers not only in terms of their affinity but also specificity, cross-reactivity, possible binding mechanism, kinetics, etc. This is missing in the paper. I understand the authors are more interested in the "final" effect of CRISPR activity but I think the characterization of aptamers should be more focused."

We appreciate the reviewer's interest in a detailed characterization of our aptamers. In light of R3's comments, we performed additional experiments and computational studies focusing on the affinity, specificity, and kinetic properties of aptamer A9. First, in vitro affinity characterization by SPR showed that the intracellularly selected aptamer A9 alone, without the sgRNA scaffold, has a strong affinity for its target protein QCP with a K_D of 10.2 nM, in contrast to its much lower affinity (95.1 nM) for MCP, confirming its specificity (Fig. 3e). To further probe the structure-activity relationship of A9, we created variants A14 to A20 by systematically mutating the unpaired C-U bulge of A9. A scrambled sequence of A9, A21, was also generated. SPR studies suggested that A14–A20 all have lower affinity to QCP than A9, highlighting the importance of the proper C-U bulge for QCP binding. Alterations of the C-U bulge was found to mainly impact association rate constant (k_{on}), with A9 showing 2–10 folds higher k_{on} than A14–A20 (Fig. 4a, c). It is noteworthy that although a weak in vitro affinity to QCP (K_D =119 nM) by SPR was still observed for A21, it showed no activity in the transcriptional activation assay (Suppl. Fig. 15).
.

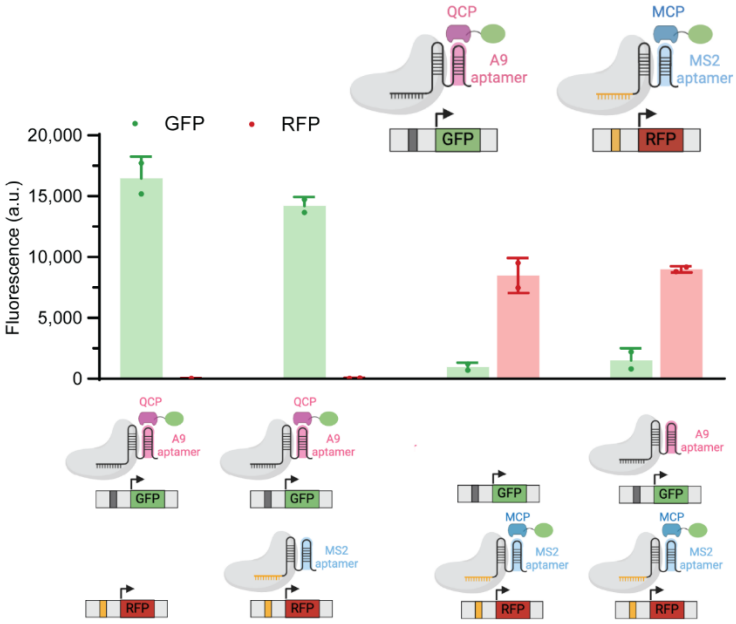


c

	K_D (nM)	k_a (M^{-1} s^{-1})	k_d (s^{-1})
A9	10.2	1.05×10^5	1.08×10^{-3}
A14	17.8	3.27×10^4	5.81×10^{-4}
A16	28.9	4.32×10^4	1.25×10^{-3}
A17	90.8	1.73×10^4	1.57×10^{-3}
A18	16	4.01×10^4	6.43×10^{-4}
A20	18	6.01×10^4	1.08×10^{-3}
A21	119	9.8×10^3	1.16×10^{-3}

Figure 4a, Secondary structures of the A9 and its variants. **c**, Binding properties of A9 variants and QCP by SPR.

We also measured the specificity of A9 and A14–A19 in *E. coli* cells by the recruitment of the on-target QCP-SoxS versus the off-target MCP-SoxS transcriptional activator and the activation of the transcription of a *sfGFP* reporter gene. Consistent with the in vitro SPR assay results, A14–A19 also showed markedly reduced specificity compared to A9 in transcriptional activation assay in *E. coli* (Fig. 4b). The specificity of A9 in *E.* coli cells was further demonstrated in a dual-fluorescent reporter assay: when the cells express QCP-SoxS, the CRISPR activator containing sgRNA-A9 could strongly activate the downstream *sfGFP* gene expression; when the cells express MCP-SoxS, only negligible transcription activation of the *sfGFP* gene was observed (Suppl. Fig. 13).



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. ($n = 3$ independent replicates).

The specificity of A9 was then further validated in mammalian cells through the transcriptional activation of the luciferase gene. Minimal cross-binding was detected between A9 and non-cognate proteins (Fig. 5a)

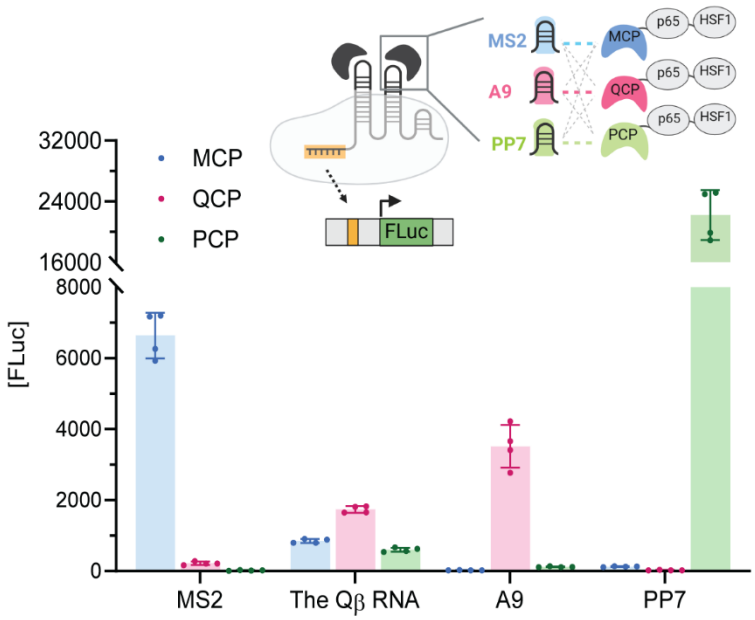


Figure 5a, No significant crossbinding is observed for MS2, PP7 and CRISPR-hybrid aptamer A9 with non-cognate RBPs, while evident crossbinding is observed for the Qβ RNA ($n = 4$ technical replicates).

To better understand why A9 was capable of strongly and selectively binding to QCP, we conducted a computational study of the A9-QCP interaction. Our initial hypothesis proposed that the C-U mismatch created a bulge in the RNA secondary structure, allowing the C3 nucleobase to bind to Y89 as an aromatic ring-stacking interaction like the bulged A nucleobase in the Qβ RNA from PDB ID 4L8H. However, this interaction could not fully explain the increased activity and selectivity over the natural Qβ RNA since it already exists in the Qβ RNA, which has low QCP selectivity. Thus, unique structural features of A9 was likely causing additional nucleobase-QCP interactions that contribute to sequence-based selectivity observed in the experiments.

We conducted molecular dynamics (MD) simulations of A9 alone, as well as A9 docked into the QCP. Metadynamics, which was used to better sample configurations along the distance between the mismatched C and U bases, revealed a key difference.

The lowest energy structure of the A9-QCP complex was captured when C3 and U17 of A9 remained stacked within the RNA backbone at a distance of 0.74 nm (Fig. 4d). In this conformation, N58, R59 and K63 interact with A9. Specifically, N58 forms a hydrogen bond with C2 of A9, while R59 and K63 interact with the phosphate group of C2 (Fig. 4e left, Fig. 4f left). However, as the distance between C3 and U17 of A9 increased, QCP-

bound A9 reached another stable conformation with U17 flipped out and a C3-U17 distance of 1.5 nm, after overcoming a modest energy barrier of $\Delta G^{\ddagger}$ = 3.35 kcal/mol (Fig. 4d). In this alternative conformation, U17 flips out of the RNA stem, allowing C3 to base pair with the lower G18, and C2 interacts with G19. This shift leaves C1 exposed, forming a hydrogen bond with N58 in the aforementioned flexible loop in QCP. R59 is now stabilized by interactions with both C2 phosphate and G12 phosphate groups. K63 retains its interaction with C2 phosphate group and remains unchanged between the flip-in and flip-out conformations (Fig. 4e right, Fig. 4f right). To validate the simulation results, we generated three single-point mutants (N58A, R59A, and K63A) and one double mutant (N58A/R59A) of QCP respectively. Functional assay assessing their ability to active the sfGFP reporter in E. coli revealed reduced activity for all mutants, confirming the essential role of these residues in A9 binding (Suppl. Fig. 16c).

In contrast, when A9 is unbound to QCP, the activation energy required for U17 to flip from an "in" to an "out" conformation is significantly higher ($\Delta G^{\ddagger}$ = 6.86 kcal/mol) (Fig. 4d). This decrease in the energy barrier upon A9 binding to QCP highlights the structural flexibility and adaptability of A9 in its interaction with QCP, helping to explain its high k_{on} upon binding to QCP.

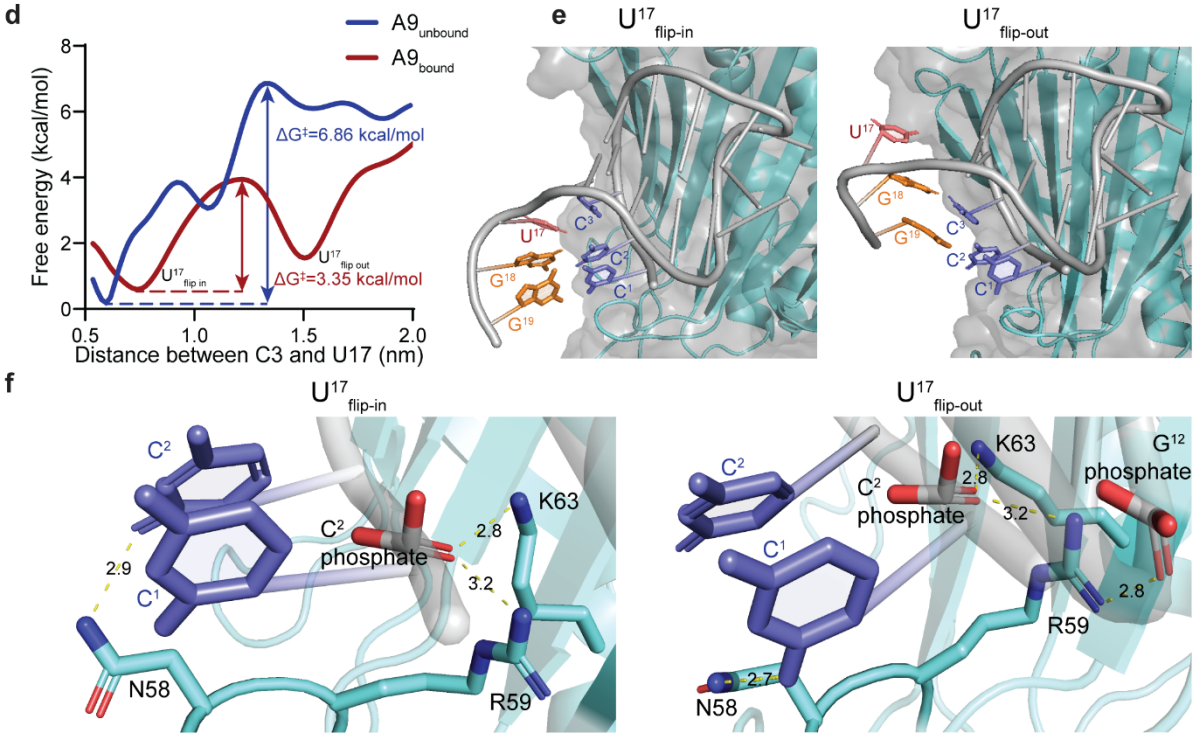
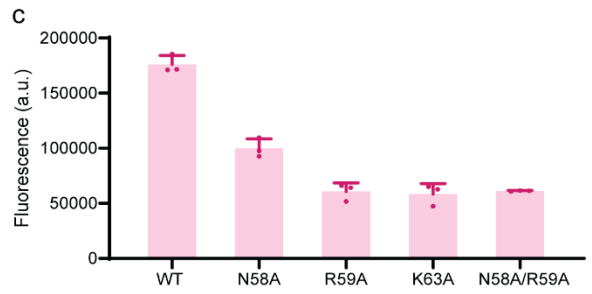


Figure 4d, Free energy profile of A9 U17 nucleobase flipping out of the secondary structure, with $\Delta G^{\ddagger}$ = 6.86 kcal/mol in free solution and $\Delta G^{\ddagger}$ = 3.35 kcal/mol when A9 is bound to the protein. **e**, Selected structures of A9 bound to QCP from metadynamics simulations. Left, structure of A9-QCP complex when U17 flipped in. Right, structure of A9-QCP complex when U17 flipped out. **f**, Residues interactions and distance in selected structures with U17 flipped in and out. Left, when U17 flips in, N58 forms a hydrogen bond with C2, R59 and K63 interact with C2 phosphate. Right, when U17 flips out, N58 forms a hydrogen

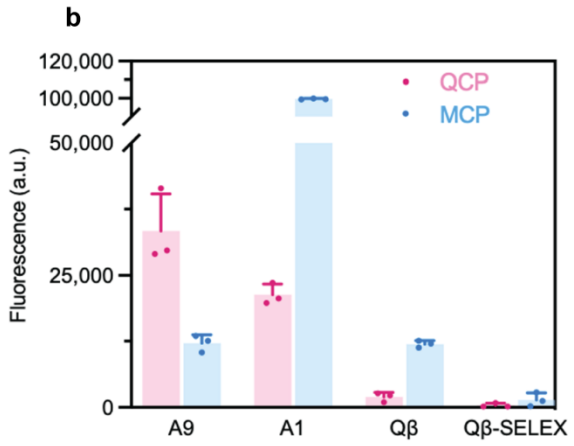
bond with the exposed C1, R59 is stabilized by interactions with both C2 and G12 phosphate groups. K63 interacts with C2 phosphate.



Supplemental Figure 16c, 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.

2. *"Also, authors should explain how this approach (directed evolution) gives different results compared to a conventional aptamer selection protocol."*

Our direct evolution approach is conducted intracellularly, linking aptamer activity directly to the selection process. Only the active aptamers, which successfully recruit the target protein and enhances reporter expression in cells, are enriched. This method inherently excludes aptamers that fail to induce the desired cellular function, ensuring that the isolated aptamers are both active and specific in cells. In contrast, conventional aptamer selection, such as SELEX, typically occurs in vitro, where aptamer activity cannot be simultaneously assessed under physiological condition and in cells. As a result, aptamers selected through SELEX may not always function as intended in cells. For example, the QCP-binding aptamer previously discovered by Hirao *et al*. [*Mol Divers* 4, 75–89 (1998)] using SELEX, Qβ-SELEX, showed significant lower activity and low specificity than A9 in both *E. coli* cells (Suppl. Fig. 9b) and mammalian cells (Fig. 5b).



Supplemental Figure 9b, A9 and A1 showed strong binding activity to QCP, while A9 exhibited reduced affinity for MCP. The Qβ RNA and Qβ-SELEX not only showed negligible on-target binding activity to QCP in vivo, but also exhibited enhanced off-target binding 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.

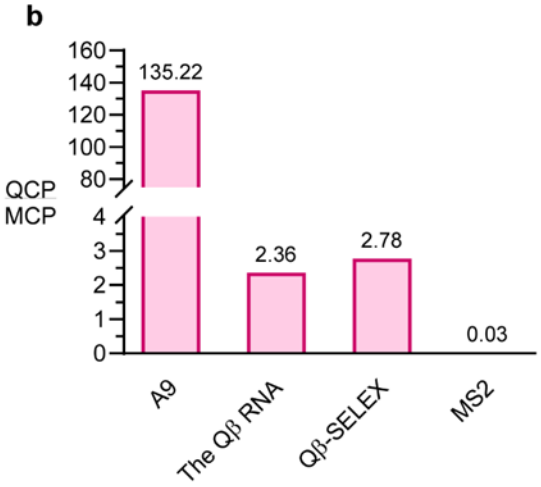


Figure 5b, A9 is highly specific for cognate target QCP over MCP, compared to the Qβ RNA and the MS2 aptamer, and an *in vitro*-selected aptamer Qβ-SELEX. Binding activities measured using luciferase reporter, ($n = 4$ technical replicates), and measurements are plotted in ratios of QCP over MCP, as well as MCP over QCP.

3. *"How the aptamers selected are different from the aptamers that would have been selected (or have been selected) using SELEX or other more standard selection approaches?"*

We compared the performance of the QCP aptamers obtained via SELEX, known as Qβ-SELEX, to A9 that was discovered using our intracellular selection platform CRISPR-Hybrid. A9 not only outperformed Qβ-SELEX by a wide margin in transcriptional activation assays in bacterial and mammalian cells, but also showed significantly higher specificity over a naturally occurring QCP-binding Qβ RNA in those cells (Suppl. Fig. 9b and Fig. 5b).

4. *"At the end the specificity/orthogonality that was searched by the authors is not really that high, so I wonder if a more conventional approach would have given better results."*

We respectfully disagree with R3 that the specificity/orthogonality of the aptamer identified by our approach is not high. We confirmed A9's high specificity and orthogonality in mammalian cells using a luciferase activation assay, where minimal cross-binding was observed between A9 and non-cognate proteins (Fig. 5a)

A9's orthogonality was further validated in a multiplexing gene editing experiment. When pairing A9 with QCP-associated transcriptional activators and pairing MS2 with MCP-associated transcriptional repressors, A9 and MS2 were able to simultaneously and specifically recruit transcriptional activators and repressors, respectively, without interference, demonstrating robust orthogonality in a complex cellular context (Fig. 5c).

Response to Reviewers

Reviewer #3

1. *"The authors this time put more efforts in characterizing their aptamer (only A9). I remain quite skeptical about their overall conclusions. This is particularly evident in Figure 4 results where the authors conclude that the affinity and specificity of the selected aptamer is due to its specific sequence. However, the mutated aptamers (especially A18 and A20) do not really perform that bad in comparison in terms of affinity.*

We appreciate the reviewer's positive comments, and respectfully disagree with their conclusion that the affinity and specificity of the selected aptamer A9 are not due to its specific sequence. While we agree that many mutated aptamers, such as A18 and A20, showed K_D values that are only slightly higher than A9 in in vitro SPR experiments, these aptamers are significantly less specific intracellularly. The difference between their in vitro and intracellular activities highlighted the advantages of A9 as a highly functional aptamer in the intracellular environment. To address this comment from R3, we have added the specificity performance of A20 in cells to Fig. 4b, which showed that A20 is notably less specific that A9.

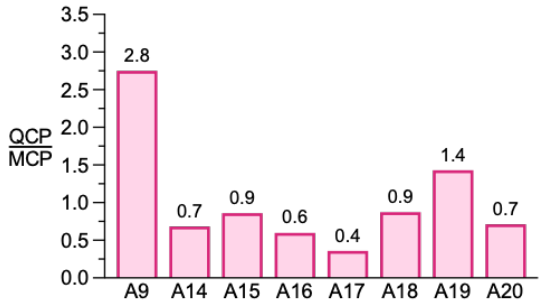


Figure 4b, Selectivity of A9 variants with rationally designed mutations.

2. "*Also, no results are shown for A15 and A19, why?"*
To address this question from R3, we have added the SPR characterizations of A15 and A19. Please see the revised Fig. 4c. A15 and A19 were found to have significantly lower affinities compared to A9, suggesting that the bulge mutations carried by these mutants caused important disruption to the structure of the aptamer.

	K_D (nM)	K_{on} ($M^{-1}s^{-1}$)	K_{off} (s^{-1})
A9	10.2	1.05×10^5	1.08×10^{-3}
A14	17.8	3.27×10^4	5.81×10^{-4}
A15	86.2	4.55×10^4	3.93×10^{-3}
A16	28.9	4.32×10^5	1.25×10^{-3}
A17	90.8	1.73×10^4	1.57×10^{-3}
A18	16	4.01×10^4	6.43×10^{-4}
A19	80.4	4.78×10^4	3.85×10^{-3}
A20	18	6.01×10^4	1.08×10^{-3}
A21	119	9.80×10^5	1.16×10^{-3}

Figure 4c, Binding properties of A9 variants and QCP by SPR.