nature portfolio

Peer Review File

Title: A small RNA that cooperatively senses two stacked metabolites in one pocket for gene control

<b>REVIEWER COMMENTS</B>

Reviewer #1 (Remarks to the Author):

Wedekind and colleagues present a highly novel crystal structure of a type-I preQ-1 riboswitch that reveals the binding of two ligand that directly interact with one another. In support of their structural analysis, the authors present a detailed calorimetric analysis of cooperative binding in several representative type-I preQ-1 riboswitches to make the case that this unusual mode of ligand binding is highly likely to be a central characteristic of this grouping. Finally, a cell based reporter assay demonstrates that ablation of either ligand binding site causes a significant shift in the extracellular preQ-1 concentration required to elicit the regulatory response. Together, the authors present a solid structure-function analysis that reveals a new mode of ligand recognition in RNA.

Overall, this manuscript is well-written and provides clear and mostly compelling data (see below, with respect to the cell-based assay). The crystallographic and calorimetric data is rigorously analyzed and the authors present a convincing set of conclusions. The figures are also clear and highlight key points in the text well. This work is highly worthy of publication, once the authors address a set of minor points, as follows.

Lines 76-77. I somewhat disagree with the statement “Although several riboswitches can recognize two effectors, these sites are positioned in separate domains”. This essentially comes down to how the authors define “domain” in RNA. I tend to think of a domain as an independently folding element, such as the classic P4-P6 domain of the T. tetrahymena group I intron. In that light some of the two ligand binding riboswitches are single folding domains, such as the THF riboswitch. This RNA clearly folds as a single domain with the main folding center around the 3WJ site that enables the formation of the pseudoknot site. I really don’t think that this detracts from the author’s main point, and they should consider modifying this claim (also made on line 137).

Line 183. Given the nature of the cell-based reporter assay, I would argue that this is a two significant figure experiment rather than three (i.e., 90 ± 3 nM). Also, I am confused by the assertion that the two phases correspond to a preQ1 binding event based upon the ITC data. What do the authors specifically envision each state to be? I assume they think that the second transition is a single ligand event? If so, why would this event elicit the greatest repression (as opposed to the first event, which contributes moderately). I admit that I am a bit confused by the observed biphasic transition given the degree of cooperativity in binding.

Line 188 – 189. The statement that the riboswitch sensing “acts a ‘dimmer’ switch rather than a ‘digital’ switch” is predicated on the assumption that the concentration of ligand added to the medium is the same as in the cell, which is not known by the authors. For some metabolites, high affinity influx pumps can significantly concentrate the compound and the shape of the curve can reflect the behavior of that component of metabolism. The authors should use caution in interpreting their cell-based data. In the discussion the authors also reiterate this conclusion (lines 206 – 207). I would like to see a more

concrete discussion as to why the authors think that a positively cooperative system leads to a less steep response curve, since this observation is counter to what was hypothesized.

Figure 3. In interpreting the first phase of the transition, the error bars seem quite large and overlap between the two baselines of that transition. I am surprised given the size of the errors in this transition that the stated error is as low as cited.

Reviewer #2 (Remarks to the Author):

The manuscript of Schroeder et al describes the structure of a type I preQ1 riboswitch from Canobacterium antarticus as well as extensive biophysical characterization to corroborate the structural findings. The main finding is that the aptamer domain of this riboswitch binds two stacked preQ1 molecules, something that had not been observed before. Strong binding data support a cooperative binding model and in vivo data corroborate the importance of the two binding model. Overall, this is a very strong manuscript. The data are of excellent quality and the structure provides new information that helps understand preQ1 riboswitches better. The observation of two binding ligands is novel for riboswitches and helps support several important ideas regarding evolution of function in the RNA world. There are a few minor points that need to be addressed:
1. The authors conclude that binding of two preQ1 molecules is a hallmark of all type I preQ1 riboswitches. While the evidence presented is good, I think a couple of mutagenesis experiments to support it would enhance the manuscript. Similar experiments to the ones done with the Can riboswitch to show the involvement of C31 and U17 but using a different organism would help address this point.
2. A multiple sequence alignment of many preQ1 riboswitches should be added to the Supplemental Materials to highlight the common features.
3. Additional discussion on the differences amongst the three types is needed, highlighting what was learned from the current work.
4. Extended Data Figure 4 is very confusing. Panels mix with each other. The figure needs to be redesigned for additional clarity.

Reviewer #3 (Remarks to the Author):

In this manuscript, Schroeder and co-workers report a detailed structural, biochemical, and biological study describing a Class I preQ1 riboswitch from Carnobacterium antarcticus (and providing information more broadly about this class of riboswitch). Remarkably, their crystallographic studies reveal that the riboswitch binds two copies of the preQ1 metabolite, exhibiting positive cooperativity, and stacking directly next to each other in the binding site. Biological studies using a reporter system reveal a “dimmer switch”-type response to ligand over a wide concentration range. This work is notable as it is the first example of an RNA binding to 2 copies of a single metabolite, with implications for riboswitch

evolution, RNA catalysis and potentially synthetic biology as well. The manuscript is highly rigorous and well written and methodology is sound. References are appropriately cited. This work will be of broad interest and will definitely be appropriate for the readers of Nature Communications. I have several minor comments:

1. The title is somewhat misleading. There are other examples of riboswitches that bind 2 different metabolites, but this is notably the first example of a switch that binds 2 molecules of the same metabolite (in the same binding pocket!). More precise language would better emphasize the important advance reported here – I suggest "A small RNA senses its effector in a tandem stacked mode for cooperative gene control" or something similar.

2. It is somewhat non-obvious that the system exhibits positive cooperativity but the response is increased in concentration range (a "dimmer switch"). This is counter to the typical example where positive cooperativity generally results in an increased Hill coefficient (and a reduced range of effect, as noted in the discussion). Can the authors comment further? Note, the Hill coefficient model was developed for enzymes, which are of course a fundamentally different system on several levels. Since this is the first example of a riboswitch that senses 2 copies of a single molecule it has fundamental importance and would be worth discussing for readers outside this specific field interested in this unique biochemical phenomenon.

3. In the case of this example, the SDS is partially embedded within the aptamer domain. However, in Class III preQ1 switches, the SDS is found further outside. Do the authors think such dual binding events as observed here are uniquely effective when the SDS is within the aptamer? This might be worth adding a sentence or two in the discussion.

To be clear, these very minor comments do not dampen enthusiasm for what I consider to be a very strong and rigorous study. I strongly support the publication of this manuscript once the issues above have been addressed.

We thank each of the reviewers for providing detailed and insightful comments to improve the quality of our manuscript. We are especially gratified by the uniformly positive evaluations of the work. Reviewer #1 stated, *"… this manuscript is well-written and provides clear and mostly compelling data… highly worthy of publication …*". Reviewer #2 wrote, "*… this is a very strong manuscript. The data are of excellent quality and the structure provides new information*". Reviewer #3 wrote, "[the] *manuscript is highly rigorous and well written and methodology is sound ... This work will be of broad interest...*". Moreover, each of the reviewers stated that the work would be suitable for publication following "*minor*" revisions. We have positively addressed all comments as requested with one exception (noted below). As a result, we believe that the manuscript has improved significantly and that the work is now ready for publication.

To track our changes, we created the ensuing point-by-point response. Our responses are in red with specific changes highlighted in yellow (reviewer #1), green (reviewer #2) or cyan (reviewer #3). A marked-up manuscript with color coding for each reviewer is appended to this document. We have also provided our complete written responses following each reviewer's questions (below).

Our noteworthy revisions are summarized as follows:

i. The title was changed to be clearer, as requested by rev. #2

ii. The abstract was shortened and references were removed to be consistent with the journal style. The Extended Data were moved to Supplementary Information to be consistent with the journal style.

iii. A more circumspect approach was taken when interpreting the GFP*uv* assay based on the critiques of rev. #1 and rev #3. Specifically, the interpretation of the in-cell assays as evidence for a dimmer switch response was removed in light of the limitations of this assay (noted in the revision) and positive cooperativity observed in our ITC analysis.

iv. We produced a new multisequence alignment as Supplementary Figure 1 based on comments from rev. #2. We described the sequence hallmarks that differentiate the various $preQ_1$-I riboswitch types in light of the consensus model, the new multisequence alignment and known co-crystal structures.

v. We added a new discussion about the prospect that other riboswitches use dual stacked effector recognition, as suggested by rev. #3.

vi. Based on our own reading of the manuscript, we found and corrected errors in Fig 1c regarding the highlighted aSDS. We also added a missing H-bond in Fig. 1e (and removed underlying H-bonds for clarity). We corrected an error in the base pairing of Fig. 3b, which now shows the correct SDS-aSDS interaction. We added individual data points to the bar plots in Fig. 3d to match the journal style. We clarified the average K_D value reported for WT *Can* in the text and the representative metrics shown in 2d. These parameters were updated in Supplementary table 2. Other self-identified minor typos are colored red but remain unhighlighted.

vii. We included all of our raw data from ITC and GFPuv assays in a single spreadsheet, as per journal policy.

viii. We made minor corrections to the Supporting Table 1 of refinement statistics that we discovered during the final stages of PDB deposition.

ix. We made changes to the Methods to match journal policy and to describe the treatment of errors for EC_{50} fitting, based on comments from rev. #1.

The detailed point-by-point responses to reviewers are as follows:

Reviewer #1 (Remarks to the Author):

Lines 76-77. I somewhat disagree with the statement "Although several riboswitches can recognize two effectors, these sites are positioned in separate domains". This essentially comes down to how the authors define "domain" in RNA. I tend to think of a domain as an independently folding element, such as the classic P4-P6 domain of the T. tetrahymena group I intron. In that light some of the two ligand binding riboswitches are single folding domains, such as the THF riboswitch. This RNA clearly folds as a single domain with the main folding center around the 3WJ site that enables the formation of the pseudoknot site. I really don't think that this detracts from the author's main point, and they should consider modifying this claim (also made on line 137).

We agree that the THF riboswitch uses one folding unit. In terms of effector binding, we also agree that two spatially separated sites (i.e., a 3-way junction and a pseudoknot region) are used to recognize individual effectors. However, our rationale was that each site fits the definition of a 'domain' as described by [Pley *et al*. (1994) Nature **372**, 68-74]. In this manuscript, McKay and co-workers state, "*a hammerhead RNA-DNA ribozyme-inhibitor complex at 2.6 Å resolution reveals that the base-paired stems are A-form helices and that the core has two structural domains. The first domain is formed by the sequence 5′-CUGA following stem I and is a sharp turn identical to the uridine turn of transfer RNA, whereas the second is a non-Watson-Crick three-base-pair duplex with a divalent-ion binding site*". There is a detailed description of *domain I* and *domain II* in this paper. As such, this represents a somewhat different perspective about a domain. Obviously, unlike the THF riboswitch, the *Can* riboswitch's binding sites are not spatially separated. We simply wanted to convey the latter point, but we see how our nomenclature could be confusing to the audience. The reviewer's concept of a domain is also more rigorous and is widely accepted by the structural biology community.

To avoid ambiguities, we changed our text to state (new lines 75-78) *"Although several riboswitches can recognize two effectors, these binding pockets are spatially separated* [30-35]*. In this respect, the Can preQ$_1$-I$_I$ riboswitch is exceptional because the metabolites stack tandemly, forming an unprecedented ligand-ligand interface within a single pocket."* These changes avoid the definition of a "domain" while clarifying our meaning.

Similarly, we changed line 137: *"Although other riboswitches bind two effectors, these examples involve distinct binding pockets that spatially separate the ligands*[4, 6, 7, 43]*"*

On Line 139 we added: "*recognition of two interacting ligands in a single aptamer pocket is unprecedented in RNA biology."*

Line 216 of the Discussion. We clarified, "*We described the structure and cooperative binding of a small riboswitch that senses two stacked effectors in a single binding pocket*."

Line 183. Given the nature of the cell-based reporter assay, I would argue that this is a two significant figure experiment rather than three (i.e., 90 ± 3 nM).

We agree that two significant figures are more appropriate. We have changed all EC_{50} measurements reported in the text (e.g., line 190 &193) to reflect this point. In **Supplementary Table 4**, we changed the number of significant figures to two as well.

Also, I am confused by the assertion that the two phases correspond to a preQ1 binding event based upon the ITC data. What do the authors specifically envision each state to be? I assume they think that the second transition is a single ligand event? If so, why would this event elicit the greatest repression (as opposed to the first event, which contributes moderately). I admit that I am a bit confused by the observed biphasic transition given the degree of cooperativity in binding.

Given the concern of reviewer #1 and #3 on this topic and comments below, we have taken a more circumspect interpretation of the two-phase dose response curve. Accordingly, we removed suggestions that the two-phase binding curve represents two discrete binding events of the $preQ_1$ metabolite, which appears to be an overinterpretation of the data. We thank the reviewer for point this out. Accordingly, we have made the following changes to the text:

Lines 81-83. We deleted, "~~*Unexpectedly, the Can preQ$_1$-I$_I$ riboswitch showed an extended effector-sensing range that is more akin to a dimmer switch than a digital switch*[2, 36]~~."

Lines 80-81. We added the following text instead: *"Mutants at each effector site reduce binding affinity and raise the concentration of preQ$_1$ required for gene repression in a bacterial reporter assay."*

Line 192-193. To clarify the binding by the $preQ_1$-II riboswitch, "*for the Lactobacillus rhamnosus (Lrh) preQ$_1$-II riboswitch*[37]*, which binds a single ligand with an EC$_{50}$ of 15 nM*[13]".

Lines 195-197. We changed the text to be more cautious, *"Notably, the Can riboswitch sensing range is broader than the Lrh riboswitch in this assay, suggesting that it detects preQ$_1$ over a wider range of effector concentrations. At present, the basis for this apparent sensing difference is uncertain (see below)."*

Line 200-204. We removed emphasis on the growth curve, *"In accord with ITC data, C17U and C31U mutants each showed poorer EC$_{50}$ values that were ~60-fold higher and ~210-fold higher than WT (**Figs. 3c-e & Supplementary Table 4**). While each mutant retains dual binding in vitro, the elevated EC$_{50}$ values imply that preQ$_1$ levels must be significantly higher inside cells to elicit an efficient gene-regulatory response, underscoring the importance of each effector binding site for gene regulation."*

The observation that one phase of the biphasic WT curve seems to impart more gene regulatory activity than the other is a keen one. Although we agree that we cannot directly relate our *in vitro* ITC data to our cell assay, our structure suggests that the β effector could provide a platform on which the ceiling can stack, thus ordering the P2 region and stabilizing the gene-off state. This prediction is supported by the observation that the C31U mutant regulates gene expression more poorly than C17U.

Lines 205-212: We added a passage that relates our in-cell mutant data to our structure. We wrote, *"Although our data cannot differentiate a preferred order of preQ$_1$ binding, impairment of the β site had a more pronounced effect on gene regulation (**Figs. 3d,e**). While C17U elicited a 6-fold repression, the C31U variant repressed GFPuv expression by only 2-fold (**Figs. 3d,e**). This functional disparity — also reflected by poorer C31U K_{D1} and K_{D2} values (**Supplementary***

*Table 3) — could be due to the requirement of the β effector to serve as a scaffold that supports the binding pocket ceiling via stacking (**Fig. 1f**). In this manner, the β site orders P2 in the gene off state while binding at the α site either orders the β site pocket or stabilizes effector binding at the β site."*

We addressed the interpretation of the biphasic curve in more detail in the next point.

Line 188 – 189. The statement that the riboswitch sensing "acts a 'dimmer' switch rather than a 'digital' switch" is predicated on the assumption that the concentration of ligand added to the medium is the same as in the cell, which is not known by the authors. For some metabolites, high affinity influx pumps can significantly concentrate the compound and the shape of the curve can reflect the behavior of that component of metabolism. The authors should use caution in interpreting their cell-based data. In the discussion the authors also reiterate this conclusion (lines 206 – 207). I would like to see a more concrete discussion as to why the authors think that a positively cooperative system leads to a less steep response curve, since this observation is counter to what was hypothesized.

We see the reviewer's point and believe it is worth mentioning in the main text. As such, we revised the text to more circumspectly consider the GFPuv assay results and the positive cooperativity measured by ITC. We also added a statement that competition between $preQ_1$ and other metabolites could affect the shape of the dose-response curve. We noted that other riboswitches interact with other metabolites in a cellular context, as reported for the *glmS* riboswitch/ribozyme [Watson, P. Y. & Fedor, M. J. *Nat. Struct. Mol. Biol.* **18**, 359-363 (2011)].

Lines 229-238 state:

*"Cooperative riboswitches are posited to show a steep "digital" dose-response[2, 42, 43], yet the Can riboswitch exhibits a broad, biphasic dose-response in our in-cell GFP*uv *assay despite the positive cooperativity we observe in our ITC analysis (**Supplementary Table 3**). Although it is tempting to associate each transition in our GFP*uv *assay with an individual $preQ_1$ binding event, the intracellular concentration of $preQ_1$ is not known in such assays and depends on multiple factors, such as the efficacy of 7-deazapurine transporters[44,45]. Additionally, we cannot rule out possible competition between $preQ_1$ and other metabolites in the cellular milieu[46], as observed for the* glmS *riboswitch[47]. These, or other factors, likely influence the shape of the Can riboswitch dose-response curve (**Fig. 3c**); nevertheless, the $preQ_1$-I riboswitch is expected to maintain positive cooperativity inside the cell[46]."*

Figure 3. In interpreting the first phase of the transition, the error bars seem quite large and overlap between the two baselines of that transition. I am surprised given the size of the errors in this transition that the stated error is as low as cited.

This is an excellent observation and we see the reviewer's point. **Fig. 3c** depicts the average of three datasets with the standard error of the mean shown for each point. In response to the reviewer's question, we contacted GraphPad for insight into how best to address this question and whether we used an appropriate approach. The representative recommended that we employ the "compare datasets" function to assess whether each replicate in a dataset should be analyzed in separate columns or together in sub-columns using GraphPad. Indeed, the different modes of analysis alter how the software calculates the errors. In the case of the WT *Can* datasets, the software recommended analyzing each dataset separately. For the other constructs, the software recommended analyzing all three replicates together in sub-columns.

As a result, we now report the standard error between the three separately determined EC_{50} values for WT *Can*. (Line 190) The new EC_{50} parameters for the WT *Can* construct using two significant figures are:

$EC_{50,1}$ = 96 ± 14 nM (previously 86 ± 3 nM)

$EC_{50,2}$ = 7100 ± 360 nM (previously 6800 ± 200 nM)

Although the values did not change appreciably, the errors did increase and are more in line with what is expected from the curves shown in **Fig. 3c**. The graph shown in **Fig. 3c** remains the same even though these new errors are used.

We also changed the $EC_{50,1}$ and $EC_{50,2}$ values in **Supplementary Table 4** and recalculated the Fold EC_{50} change metrics reported in Fig. 3e based on this comment

Lines 473-475: We added brief description of the GraphPad consideration in the Methods, *"The replicates in each construct were compared using the "compare datasets" function before analysis."*

Reviewer #2 (Remarks to the Author):

1. The authors conclude that binding of two preQ1 molecules is a hallmark of all type I preQ1 riboswitches. While the evidence presented is good, I think a couple of mutagenesis experiments to support it would enhance the manuscript. Similar experiments to the ones done with the Can riboswitch to show the involvement of C31 and U17 but using a different organism would help address this point.

We appreciate the reviewer's suggestion for more experiments; however, we believe that sufficient evidence already exists based upon our current analysis and the addition of a new multisequence alignment (recommended by reviewer #2 in the next point). Our rationale is that the cost-to-benefit ratio for such new experiments will be large given the time and resources involved. Specifically, these experiments are not standard ITC experiments because they require large amounts of RNA and $preQ_1$ for the VP-ITC due to the poor K_D values of mutants, as well as the need to capture the full parabolic character of cooperative binding for analysis by our Python program.

As the reviewer noted, "*the evidence presented is good*" and we believe such experiments would be merely incremental. The reviewer's suggestion of a multisequence alignment combined with the existing consensus model strongly bolsters our conclusion that all type I $preQ_1$ riboswitches appear to use dual, stacked effector recognition. Indeed, nucleotides that engage in $preQ_1$ binding at the *Can* riboswitch α and β sites are 97% conserved based on all known type I sequences (>1,500 representatives) [McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. *Chem Biol* **21**, 880-889 (2014).].

As the reviewer observed, our existing experimental data firmly support the requirement for nucleotides C17 and C31 at the respective α and β sites for dual $preQ_1$ binding and function. With the benefit of our new *Can* $preQ_1$-I_I riboswitch co-crystal structure, it is clear that all known type I riboswitch sequences possess these two bases — and other nucleotides — required for α and β site recognition. Please see new **Supplementary Fig. 1a**.

Moreover, our manuscript also demonstrates that the WT *Can, Ngo and Hin* preQ$_1$-I$_I$ riboswitches —spanning multiple phyla — each bind two preQ$_1$ molecules based on ITC. All three species show the characteristic parabolic response to preQ$_1$ when the experiment is performed at 37 °C (**Supplementary Figs. 5a,g,h**), which is best described by a two-interdependent-sites binding model (**Supplementary Fig. 5b**) wherein the macroscopic cooperativity constants, γ, support positive cooperativity (**Supplementary Table 2**). Collectively the data suggest that the mode of binding is the same in all three sequences (**Supplementary Fig 6a**), and that this analysis extends to the entire type I subclass (**Supplementary Fig. 1a**).

2. A multiple sequence alignment of many preQ1 riboswitches should be added to the Supplemental Materials to highlight the common features.

We agree with the reviewer and we have added a multiple sequence alignment in new **Supplementary Fig. 1**. We carefully selected phylogenetically diverse sequences from each type of class I preQ$_1$ riboswitch to provide the greatest diversity — albeit type III is found almost exclusively in gamma proteobacteria. These sequence alignments and the consensus models — which were derived from all known representatives analyzed by McCown *et al*. [McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. *Chem Biol* **21**, 880-889 (2014).] — collectively illustrate that the nucleobases involved in type I recognition are absolutely conserved but are absent in type II and type III preQ$_1$-I riboswitches.

To accentuate differences in the signature residues engaged in preQ$_1$ binding by the type I and type II preQ$_1$-I riboswitches, our new figure includes sequences from riboswitches that were crystallized previously (bolded genus and species). In the columns above each base, we denoted positions that contact each preQ$_1$ effector (bolded) and positions that form the P1 helix (underlined). The structural mapping upon the riboswitch sequences nicely explains the type I and type II covariation models [McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. *Chem Biol* **21**, 880-889 (2014).]. Importantly, the type I covariation model considers >1,500 sequences and our crystallographic data account for why specific bases are conserved (i.e., because they engage in α and β-site preQ$_1$ recognition). For the sake of brevity, we chose to sample a small number of diverse sequences, although the covariation models — also included I the figure — were derived from all sequence representatives from each riboswitch type [adapted from McCown *et al*. & Breaker (2014*) Chem & Biol* **21**, 880].

For type III riboswitches, there is no crystallographic data to indicate which nucleobases contact preQ$_1$ beyond the conserved cytidine in loop L2 — which presumably contacts the WC face of preQ$_1$ [McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. *Chem Biol* **21**, 880-889 (2014).]. However, the alignment and consensus model clearly indicate that nucleobases required for beta site preQ$_1$ recognition are absent. Moreover, we previously published ITC data from a representative type III riboswitch and found 1:1 binding stoichiometry [Liberman, J. A., Bogue, J. T., Jenkins, J. L., Salim, M. & Wedekind, J. E. ITC analysis of ligand binding to preQ1 riboswitches. *Meth. Enzymol.* **549**, 435-50 (2014).].

The new supplementary figure legend states:
"Supplementary Figure 1 | Covariation model and multisequence alignments of preQ$_1$ class I riboswitches. (**a**) Type I Covariation models generated from the full group of known sequence representatives (adapted from ref. 2); red, black and gray positions indicate 97%, 90% and 75% sequence conservation. Multisequence alignments were generated using a handful of representatives derived from phylogenetically diverse bacteria (reported by McCown *et al.*[2]). Positions in bold within the alignment recognize preQ$_1$ based on the *C. antarcticus* co-crystal structure of this investigation. PreQ$_1$ binding nucleobases at the α and β sites are each denoted

in the covariation model and the sequence alignment as α or β. Here and elsewhere, bolded organisms have been structurally characterized (this work). In addition to the greatest number of representative sequences (indicated in italics), preQ$_1$-I$_I$ riboswitches exhibit the greatest taxonomic diversity[2]. (**b**) same as (**a**), but with type II sequences. Characterized sequences are *T. tencongensis*[3, 4] and *B. subtilis*[5]. Asterisks denote conserved preQ$_1$ recognition positions. (**c**) same as (**a**) and (**b**) but with type III sequences. Due to a lack of structural characterization, the canonical specificity base is the only predicted preQ$_1$ recognition position[2]. Alignments were created in JALVIEW[6]."

3. Additional discussion on the differences amongst the three types is needed, highlighting what was learned from the current work.

We agree with the reviewer, and have added the following paragraph to the first paragraph of our discussion:

The new text describing this analysis is on Lines 216-228:

*"We described the structure and cooperative binding of a small riboswitch that senses two stacked effectors in a single binding pocket. Examination of all known preQ$_1$-I sequences encompassing multiple phyla revealed that nucleobases that compose the α and β binding sites are conserved only within preQ$_1$-I$_I$ sequences (**Supplementary Fig. 1**). In contrast, only nucleobases associated with α site recognition are conserved within preQ$_1$-I$_{II}$ sequences, consistent with known Tte and Bsu riboswitches structures (**Supplementary Figs. 1b, 4b,c & 6**) and previous bioinformatic analysis[28]. Although experimental analysis of the preQ$_1$-I$_{III}$ riboswitch is sparse, it appears that nucleobases associated with α site recognition are conserved in preQ$_1$-I$_{II}$ representatives, but not those associated with β site recognition (**Supplementary Fig. 1c**). This is consistent with previous ITC experiments, which demonstrated that this riboswitch binds with a 1:1 stiochiometry[41] — like preQ$_1$-I$_{II}$ representatives. Accordingly, the unprecedented mode of dual effector recognition appears to be a hallmark of the most common and taxonomically diverse preQ$_1$ riboswitch group[30, 35], the preQ$_1$-I$_I$ riboswitch, which has been overlooked until now."*

4. Extended Data Figure 4 is very confusing. Panels mix with each other. The figure needs to be redesigned for additional clarity.

We understand the reviewer's concern. As requested, we reformatted the figure to clearly differentiate among the panels. In the revision, we rearranged the ITC panels, increased the spacing between them and made the schematic diagram larger, thus increasing the readability. The revised figure is presented as **Supplementary Fig. 5**.

Reviewer #3 (Remarks to the Author):

1. The title is somewhat misleading. There are other examples of riboswitches that bind 2 different metabolites, but this is notably the first example of a switch that binds 2 molecules of the same metabolite (in the same binding pocket!). More precise language would better emphasize the important advance reported here – I suggest "A small RNA senses its effector in a tandem stacked mode for cooperative gene control" or something similar.

We agree with the reviewer that the title (which must be fifteen-words or less) should convey better the novel mode of effector recognition. We believe we have captured this point with the new title:

A small RNA that cooperatively senses two stacked metabolites in one pocket for gene control

2. It is somewhat non-obvious that the system exhibits positive cooperativity but the response is increased in concentration range (a "dimmer switch"). This is counter to the typical example where positive cooperativity generally results in an increased Hill coefficient (and a reduced range of effect, as noted in the discussion). Can the authors comment further? Note, the Hill coefficient model was developed for enzymes, which are of course a fundamentally different system on several levels. Since this is the first example of a riboswitch that senses 2 copies of a single molecule it has fundamental importance and would be worth discussing for readers outside this specific field interested in this unique biochemical phenomenon.

In response to this comment and those from reviewer 1, we have altered our interpretation of the in-cell experiments. Upon reflection, we agree with rev. #1 that it is not possible to relate the observed positive cooperativity from ITC to the curve shapes derived from the GFP*uv* reporter assay for the reasons s/he stated. There are several reasons for this:

1. We do not know the intracellular $preQ_1$ concentration.

2. We do not know the efficiency of associated 7-deazapurine influx transporters.

3. The experiment takes place in a complex cellular environment. In this scenario, $preQ_1$ likely competes with other metabolites for binding, altering the apparent concentration of $preQ_1$ needed to elicit a specific regulatory response.

Our response to reviewer #1 in this regard appears on Lines 231-241. We wrote:

"Cooperative riboswitches are posited to show a steep "digital" dose-response[2, 42, 43], yet the Can *riboswitch exhibits a broad, biphasic dose-response in our in-cell GFP*uv *assay despite the positive cooperativity we observe in our ITC analysis (**Supplementary Table 3**). Although it is tempting to associate each transition in our GFP*uv *assay with an individual $preQ_1$ binding event, the intracellular concentration of $preQ_1$ is not known in such assays and depends on multiple factors, such as the efficacy of 7-deazapurine transporters[44,45]. Additionally, we cannot rule out possible competition between $preQ_1$ and other metabolites in the cellular milieu[46], as observed for the* glmS *riboswitch[47]. These, or other factors, likely influence the shape of the Can riboswitch dose-response curve (**Fig. 3c**); nevertheless, the $preQ_1$-I_I riboswitch is expected to maintain positive cooperativity inside the cell[46]."*

However, as rev. #3 suggested, we also considered the benefits of positive cooperativity for gene regulation. This discussion provides background and context for the broader community.

On lines 239-251 of the Discussion we wrote,

*"Our data allow us to conclude that dual-effector recognition is critical for efficient gene regulation by $preQ_1$-I_I riboswitches — as indicated by the deleterious effects caused by specific α and β site mutants. However, we can only speculate on the reason why cooperativity evolved in $preQ_1$-I_I riboswitches but not in other types or classes of the $preQ_1$ riboswitch family. Our data suggest that the level of regulation attained is similar between the $preQ_1$-I_I Can riboswitch and the $preQ_1$-II Lrh riboswitch, despite differences in $preQ_1$ binding stiochiometry[37, 48] (**Fig. 3d**). This result suggests that these two disparate riboswitch folds evolved equally effective chemical networks to sense a common effector for gene regulation. Yet, cooperativity is expected to provide notable benefits in regulation efficiency. One such advantage is that gene expression is*

permitted when metabolite levels are low (Fig. 3b, middle panel), while assuring the ability to quickly attenuate expression before excess effector accrues in the cell[46]. This is reasonable considering that many preQ$_1$-I$_I$ riboswitches control the translation of transporters that salvage Q-precursor metabolites from the extracellular environment[23, 28, 49].”

3. In the case of this example, the SDS is partially embedded within the aptamer domain. However, in Class III preQ1 switches, the SDS is found further outside. Do the authors think such dual binding events as observed here are uniquely effective when the SDS is within the aptamer? This might be worth adding a sentence or two in the discussion.

The reviewer raises a very interesting point. We have added the following brief paragraph on Lines 252-264 that considers this possibility. We also took the last sentence from the previous version of the manuscript and added it here to close the paragraph.

“*Although the Can and Lrh riboswitches differ in terms of binding stoichiometry and overall fold, each positions its expression platform near the binding pocket. This organization raises the question of whether dual, stacked metabolite binding could be effective to regulate folds in which the aptamer is located distally from the expression platform. PreQ$_1$-III riboswitches exemplify this organization, wherein the expression platform can be as far as 40 Å away from the aptamer[38]. Communication between the single-effector pocket and an orthogonal SDS-antiSDS helix is mediated by an A minor base that makes a T-shaped contact with the edge of preQ$_1$[38] (**Supplementary Fig. 4e**). The preQ$_1$-II riboswitch uses a similar pocket[37] and the A-minor base was shown to be essential for gene-regulatory function[40]. It is conceivable that dual, stacked effector recognition could be used by the preQ$_1$-III riboswitch fold, if the effectors promoted coaxial helical stacking, and one or both were detected by an A-minor motif. Accordingly, we predict that additional riboswitches that bind dual, stacked effectors exist in nature.”*

We also modified the figure legend of **Supplementary Figure 4e** to support the A-minor statements added in the main text.

“A70 and A84 are inclined A-minor bases that originate from an orthogonal A-form helix that abuts the effector edge[8, 9]. In the preQ$_1$-II riboswitch, these bases are important for gene regulation and dynamics[10-12].”

To be clear, these very minor comments do not dampen enthusiasm for what I consider to be a very strong and rigorous study. I strongly support the publication of this manuscript once the issues above have been addressed.

We thank the reviewer for this clarification and appreciate the thoughtful review. We thank the other reviewers as well.

A small RNA that cooperatively senses two stacked metabolites in one pocket for gene control

Griffin M. Schroeder[1,2], Chapin E. Cavender[1,2], Maya E. Blau[3], Jermaine L. Jenkins[1,2], David H. Mathews[1,2] and Joseph E. Wedekind[1,2]

[1.] *Department of Biochemistry & Biophysics, University of Rochester School of Medicine & Dentistry, Rochester, NY 14642, USA.*

[2.] *Center for RNA Biology, University of Rochester School of Medicine & Dentistry, Rochester, NY 14642, USA.*

[3.] *University of Rochester, 120 Trustee Road, Rochester, NY 14627, USA.*

To whom correspondence should be addressed. Tel: +1 585 273 4516;

Email: joseph.wedekind@rochester.edu

ORCID's: 0000-0001-6354-752X (GMS), 0000-0002-5899-7953 (CEC), 0000-0002-8948-4982(MEB), 0000-0003-2548-3275 (JLJ), 0000-0002-2907-6557 (DHM) and 0000-0002-4269-4229 (JEW)

Abstract

Riboswitches are structured non-coding RNAs often located upstream of essential genes in bacterial messenger RNAs. Such RNAs regulate expression of downstream genes by recognizing a specific cellular effector. Although nearly 50 riboswitch classes are known, only a handful recognize multiple effectors. Here, we report the 2.60-Å resolution co-crystal structure of a class I type I $preQ_1$-sensing riboswitch that reveals two effectors stacked atop one another in a single binding pocket. These effectors bind with positive cooperativity *in vitro* and both molecules are necessary for gene regulation in bacterial cells. Stacked effector recognition appears to be a hallmark of the largest subgroup of $preQ_1$ riboswitches, including those from pathogens such as *Neisseria gonorrhoeae*. We postulate that binding to stacked effectors arose in the RNA World to closely position two substrates for RNA-mediated catalysis. These findings expand known effector recognition capabilities of riboswitches and have implications for antimicrobial development.

Introduction

Riboswitches are found primarily in the 5′ leader sequences of bacterial mRNAs where they regulate the expression of genes by recognizing a cognate effector[1-3]. These RNA-control elements usually comprise two domains: an aptamer that recognizes a metabolite with high specificity and an expression platform that contains gene-regulatory sequences[1]. Upon ligand binding, the expression platform undergoes conformational changes that alter the accessibility of key regulatory regions, such as the Shine-Dalgarno sequence (SDS), which must be unobstructed to initiate translation[3]. Direct observation of ligand-mediated transitions in riboswitches has enriched our understanding of RNA allostery and folding[4-7]. Riboswitches are also promising antimicrobial targets due to their presence in numerous human pathogens[2] and the finding that riboswitch dysregulation can compromise bacterial virulence[8].

PreQ$_1$-I (class I) riboswitches are the founding group of bacterial gene regulators that control the cellular concentration of queuosine (Q)[9] (**Fig. 1a**) — a hypermodified 7-deazapurine nucleobase required for translational fidelity in mammals and bacteria[10-12]. Although Q is not essential in bacteria, Q deficiency is associated with slow mid-log growth[13], compromised stationary-phase viability[11] and loss of virulence[14]. Previous preQ$_1$-I riboswitch structures revealed an H-type pseudoknot fold, which recognizes a single preQ$_1$ ligand that completes coaxial stacking between flanking helices, thus stabilizing the expression platform[15-18]. The small size and well-defined fold of this class have spurred investigations of its folding and dynamics[4, 15, 19, 20 21], effector specificity[22, 23], the ligand-free to bound-state transition[4, 18, 19, 22, 24] and targeting with drug-like molecules[25, 26]. Multiple bacterial species exhibit 1:1 riboswitch-to-preQ$_1$ stoichiometry[15-18, 22, 25, 27], which is the prevailing ligand-binding mode of most riboswitches[2].

Importantly, the latter preQ$_1$-I riboswitch analyses have considered relatively few sequences. Recent work further classified preQ$_1$-I riboswitches into three subgroups called types I-III[28]. Inspection of the associated consensus models reveals that types I and II adopt similar secondary structures (**Fig. 1b &** **Supplementary Figure 1a,b**). Although preQ$_1$-I$_{II}$ (type

II) sequences prefer adenosine before the cytidine specificity base, preQ$_1$-I$_I$ sequences prefer uracil followed by CUA in the 3′-expression platform[23, 28]. This observation and the results we describe in this study suggest that all previously studied sequences are preQ$_1$-I$_{II}$ riboswitches. Importantly, preQ$_1$-I$_I$ riboswitches, found in gram-positive and -negative bacteria, are more represented than all other preQ$_1$ riboswitch subgroups and classes combined[28].

To elucidate the gene-regulatory properties of preQ$_1$-I$_I$ riboswitches, we determined the co-crystal structure of a preQ$_1$-I$_I$ riboswitch from *Carnobacterium antarcticus*[29] (*Can*). The H-type pseudoknot structure unexpectedly reveals two bound preQ$_1$ effectors in a single aptamer (**Fig. 1c**, **d**). Although several riboswitches can recognize two effectors, these binding pockets are spatially separated[31-36]. In this respect, the *Can* preQ$_1$-I$_I$ riboswitch is exceptional because the metabolites stack tandemly, forming an unprecedented ligand-ligand interface within a single pocket. Using isothermal titration calorimetry (ITC) with in-house software that models two interdependent binding sites, we demonstrated that two preQ$_1$ effectors bind with positive cooperativity. Mutants at each effector site reduce binding affinity and raise the concentration of preQ$_1$ required for gene repression in a bacterial reporter assay. ~~Unexpectedly, the *Can* preQ$_1$-I$_I$ riboswitch showed an extended effector-sensing range that is more akin to a dimmer switch than a digital switch~~. We also found that additional preQ$_1$-I$_I$ sequences from *Haemophilus influenzae* (*Hin*) and *Neisseria gonorrhoeae* (*Ngo*) sense two preQ$_1$ effectors with positive cooperativity, suggesting that tandem, stacked effector binding is a hallmark of all preQ$_1$-I$_I$ riboswitches. Use of a single binding pocket to recognize two effectors has implications for the development of new antimicrobials that utilize a chemical scaffold that avoids cross-reactivity with naturally occurring metabolites.

Results

Features of the Can Riboswitch Fold

To identify a suitable preQ$_1$-I$_I$ riboswitch for structural and functional analysis, we searched previously curated type I sequences[28] (**Fig. 1b** and **Supplementary Fig. 1a**) for a strong SDS (5′-AGGAG-3′) to use in a bacterial reporter assay[13]. We found several candidates, such as that from *Paenibacillus terrae*, but NCBI BLAST searches led to the discovery of an unreported sequence from *Can*[29]. This riboswitch crystallized readily from low salt solutions and the preQ$_1$-bound co-crystal structure was determined by molecular replacement. The structure was refined to 2.60 Å-resolution yielding R_{work}/R_{free} values of 0.23/0.27 with acceptable quality-control metrics (**Supplementary Table 1**). Three crystallographically independent molecules were built, which showed varied quality in electron-density maps. Chains A and B are well defined, but the chain C P1-L3 junction shows a break (**Supplementary Fig. 2a**). Importantly, both effectors and the core aptamer are well-resolved in each chain (**Supplementary Fig. 2b**), providing a firm foundation to guide functional experiments.

The overall fold of the *Can* preQ$_1$-I$_I$ riboswitch is an H-type pseudoknot (**Fig. 1c,d**). P1 is a canonical A-form helix whose minor groove is recognized by six A-amino-kissing interactions donated by the A-rich patch in L3 (**Supplementary Fig. 3**). This stabilizing segment culminates with a A28•G5-C18•A29 base-triple variation that forms the binding pocket floor and is reminiscent of preQ$_1$-I$_{II}$ riboswitch structures from *Thermoanaerobacter tengcongensis* (*Tte*)[16, 18, 22] and *Bacillus subtilis* (*Bsu*)[15, 17] (**Fig. 1e**). The pocket ceiling comprises an C8•A12•U32 base triple derived entirely from P2 (**Fig. 1f**). This configuration contrasts with preQ$_1$-I$_{II}$ riboswitches, in which the ceiling is formed by bases from both P2 and the L2 loop[16, 17, 23]. The preference for C8 and U32 in preQ$_1$-I$_I$ riboswitches appears to be incompatible with the base quadruple ceiling observed in preQ$_1$-I$_{II}$ riboswitches that require an adenine immediately before the cytosine specificity base (**Fig. 1b & Supplementary Figs. 1a,b**).

P2 also contains the expression platform, wherein the Watson-Crick (WC) face of A33 of the SDS pairs non-canonically with G9 and the 2´-hydroxyl interacts with the WC face of A11 (**Fig. 1f**). These SDS-anti(a)SDS interactions presumably attenuate translation. Although we

hypothesize that SDS G34 makes a WC pair with C10, the former is involved in a crystal contact (**Supplementary Fig. 2c**). Notably, the *Tte* preQ$_1$-I$_{II}$ riboswitch forms the expected aSDS-SDS C-G intramolecular WC pair and exhibits non-canonical pairing in its expression platform[16, 18, 22], as observed here for the *Can* preQ$_1$-I$_I$ riboswitch.

Stacked Metabolites in a Small Aptamer

A distinguishing feature of our structure is two preQ$_1$ molecules, which we term α and β, stacked in a single aptamer pocket (**Fig. 2a**). Recognition at the α site is conserved among preQ$_1$-I$_I$ and preQ$_1$-I$_{II}$ riboswitches, wherein specificity is conferred by a cytidine that recognizes preQ$_1$ by a *cis* WC interaction. Other conserved α-site contributions include the WC face of A30, the major-groove edge of U6 and the major-groove edge of G5, which interacts with the preQ$_1$ methylamine (**Fig. 2b** & **Supplementary Figs. 4a-c**). In contrast, the β site has not been observed previously. Bases C31 and U7, which are highly conserved among type I sequences, confer specificity for preQ$_1$ by contributing three hydrogen bonds that recognize the metabolite edge (**Fig. 2c** & **Supplementary Fig. 1a**). The β-site preQ$_1$ interacts with the α-site effector through aromatic stacking and donation of a hydrogen bond from the methylamine to both the keto oxygen of the α-site effector and O4 of U16 (**Figs. 2a,c**). The mode of β-site effector recognition differs from all known preQ$_1$ riboswitches, including preQ$_1$-II[37] and preQ$_1$-III[38], which utilize *trans* WC-pairing to read the preQ$_1$ face (**Supplementary Figs. 4d,e**). Although other riboswitches bind two effectors, these examples involve distinct binding pockets that spatially separate the ligands[34-36, 39]. To our knowledge, recognition of two interacting ligands in a single aptamer pocket is unprecedented in RNA biology.

Stacked Recognition is Cooperative

Interacting ligands should cooperatively influence each other's binding. Analysis of the *Can* riboswitch by ITC at 25 °C showed that the wildtype (WT) sequence binds $preQ_1$ with an average macroscopic K_D of 32.0 ± 2.0 nM and a ligand-to-receptor ratio (N) of 1.8 (**Fig. 2d** & **Supplementary Table 2**). Fitting to a single-phase isotherm supports binding with positive cooperativity, in accord with our structure. Enthalpy drives binding and offsets the predicted entropic cost of ordering two ligands, producing a favorable ΔG° (**Supplementary Table 2**). Analysis at 37 °C to accentuate cooperative binding produced a parabolic thermogram best described by a binding model wherein two interdependent ligands occupy non-equivalent sites (**Supplementary Figs. 5a,b**). We implemented this model to assess the macroscopic binding constant of each interaction, which yielded K_{D1} of 891 nM and K_{D2} of 461 nM for the first and second binding events. The improved affinity observed for the second $preQ_1$ relative to the first indicates positive cooperativity, exemplified by the macroscopic cooperativity constant, γ, of 7.7 (**Supplementary Table 3**).

We next generated *Can* riboswitch mutants to probe recognition at the α and β sites. Position 17 is a major determinant of α-site specificity (**Fig. 2a,b**) and the C17U mutation severely weakened binding as indicated by macroscopic K_{D1} and K_{D2} values of 3.13 μM and 1.30 μM (**Supplementary Fig. 5c** & **Supplementary Table 3**). This result is consistent with the position of C17 in our structure and an equivalent nucleobase in the *Bsu* $preQ_1$-I_I aptamer[23]. C17U showed a parabolic isotherm suggesting retention of two binding events. Likewise, position 31 shows a prominent role in β-site specificity (**Fig. 2a,c**). The C31U mutation produced a comparable parabolic isotherm, corresponding to K_{D1} and K_{D2} values of 6.64 μM and 10.26 μM (**Supplementary Fig. 5d & Supplementary Table 3**). As expected from the structure, C31U severely affects β-site recognition. Both C17U and C31U retain positive cooperativity with γ values of 9.6 and 2.6 (**Supplementary Table 3**).

Dual-Binding Signatures in Other $preQ_1$-I_I Riboswitches

C17 and C31 are highly conserved in the type I consensus model (**Fig. 1b & Supplementary Fig. 1a**) and the importance of each is confirmed by our structure and ITC experiments. We next asked if dual binding is evident in other type I riboswitches. We used ITC at 25 °C to evaluate sequences from *Hin* and *Ngo*, which belong to the *Proteobacteria* phyla rather than the *Firmicutes* (**Supplementary Fig. 1a**). The *Hin* riboswitch binds preQ$_1$ with a K_D of 52.9 ± 0.2 nM whereas the *Ngo* riboswitch binds with a K_D of 50.5 ± 1.3 nM; like WT *Can*, each binds with an N of ~2 (**Supplementary Table 2** & **Supplementary Figs. 5e,f**). Analysis at 37 °C accentuates the cooperative character of isotherms (**Supplementary Figs. 5g, h**), resulting in γ values of 26.7 and 32.9 that indicate substantial positive cooperativity for each (**Supplementary Table 3**). We note a high degree of sequence identity exists in the binding pocket of the *Can, Hin* and *Ngo* riboswitches (**Supplementary Fig. 1a**). Significantly, each possesses key nucleobases required for α and β site preQ$_1$ recognition including U6, U7, U16, C17, A30 and C31 (**Supplementary Fig. 6**). Given ITC evidence of cooperativity for all three riboswitches, it appears that each riboswitch uses a similar mode of dual, stacked preQ$_1$ recognition.

Gene Regulation Requires Two Effectors

Using a GFP*uv* reporter gene[13, 40] controlled by the *Can* riboswitch in live cells, we asked whether both preQ$_1$ molecules were required for effective gene regulation (**Fig. 3a**). We hypothesized that when both sites are occupied the SDS would be less accessible, leading to greater repression of GFP*uv* translation (**Fig. 3b**); likewise, intermediate levels of translation would occur if one site is occupied. Dose-response analysis of the WT riboswitch produced a biphasic curve with EC$_{50}$ values of 96 ± 14 nM (EC$_{50,\,1}$) and 7100 ± 360 nM (EC$_{50,\,2}$) (**Fig. 3c & Supplementary Table 4**). Collectively, both binding events confer 15.4-fold repression, comparable to the 14.9-fold repression observed for the *Lactobacillus rhamnosus* (*Lrh*) preQ$_1$-II riboswitch[37], which binds a single ligand with an EC$_{50}$ of 15 nM[13] (**Figs. 3c, *inset*, 3d &**

194 **Supplementary Table 4**). Notably, the *Can* riboswitch sensing range is broader than the *Lrh*
195 riboswitch in this assay, suggesting that it detects $preQ_1$ over a wider range of effector
196 concentrations. At present, the basis for this apparent sensing difference is uncertain (see
197 below). To ensure that the changes in GFP*uv* expression were riboswitch driven, we evaluated
198 a positive control containing an SDS without an upstream riboswitch and a negative control
199 lacking the SDS[13]. As expected, neither control responded to changes in $preQ_1$ concentration
200 (**Fig. 3c,e**). In accord with ITC data, C17U and C31U mutants each showed poorer EC_{50} values
201 that were ~60-fold higher and ~210-fold higher than WT (**Figs. 3c-e & Supplementary Table**
202 **4**). While each mutant retains dual binding *in vitro*, the elevated EC_{50} values imply that $preQ_1$
203 levels must be significantly higher inside cells to elicit an efficient gene-regulatory response,
204 underscoring the importance of each effector binding site for gene regulation.
205 Although our data cannot differentiate a preferred order of $preQ_1$ binding, impairment of
206 the β site had a more pronounced effect on gene regulation (**Figs. 3d,e**). While C17U elicited a
207 6-fold repression, the C31U variant repressed GFP*uv* expression by only 2-fold (**Figs. 3d,e**).
208 This functional disparity — also reflected by poorer C31U K_{D1} and K_{D2} values (**Supplementary**
209 **Table 3**) — could be due to the requirement of the β effector to serve as a scaffold that supports
210 the binding pocket ceiling via stacking (**Fig. 1f**). In this manner, the β site orders P2 in the gene
211 off state while binding at the α site either orders the β site pocket or stabilizes effector binding at
212 the β site.
213
214 **Discussion**
215 We described the structure and cooperative binding of a small riboswitch that senses
216 two stacked effectors in a single binding pocket. Examination of all known $preQ_1$-I sequences
217 encompassing multiple phyla revealed that nucleobases that compose the α and β binding sites
218 are conserved only within $preQ_1$-I_1 sequences (**Supplementary Fig. 1a**). In contrast, only

nucleobases associated with α site recognition are conserved within preQ$_1$-I$_{II}$ sequences, consistent with known *Tte* and *Bsu* riboswitches structures (**Supplementary Figs. 1b, 4b,c & 6**) and previous bioinformatic analysis[28]. Although experimental analysis of the preQ$_1$-I$_{III}$ riboswitch is sparse, it appears that nucleobases associated with α site recognition are conserved in preQ$_1$-I$_{II}$ representatives, but not those associated with β site recognition (**Supplementary Fig. 1c**). This is consistent with previous ITC experiments, which demonstrated that this riboswitch binds with a 1:1 stiochiometry[41] — like preQ$_1$-I$_{II}$ representatives. Accordingly, the unprecedented mode of dual effector recognition appears to be a hallmark of the most common and taxonomically diverse preQ$_1$ riboswitch group[30, 35], the preQ$_1$-I$_I$ riboswitch, which has been overlooked until now.

Cooperative riboswitches are posited to show a steep "digital" dose-response[2, 42, 43], yet the *Can* riboswitch exhibits a broad, biphasic dose-response in our in-cell GFP*uv* assay despite the positive cooperativity we observe in our ITC analysis (**Supplementary Table 3**). Although it is tempting to associate each transition in our GFP*uv* assay with an individual preQ$_1$ binding event, the intracellular concentration of preQ$_1$ is not known in such assays and depends on multiple factors, such as the efficacy of 7-deazapurine transporters[44,45]. Additionally, we cannot rule out possible competition between preQ$_1$ and other metabolites in the cellular milieu[46], as observed for the *glmS* riboswitch[47]. These, or other factors, likely influence the shape of the *Can* riboswitch dose-response curve (**Fig. 3c**); nevertheless, the preQ$_1$-I$_I$ riboswitch is expected to maintain positive cooperativity inside the cell[46].

Our data allow us to conclude that dual-effector recognition is critical for efficient gene regulation by preQ$_1$-I$_I$ riboswitches — as indicated by the deleterious effects caused by specific α and β site mutants. However, we can only speculate on the reason why cooperativity evolved in preQ$_1$-I$_I$ riboswitches but not in other types or classes of the preQ$_1$ riboswitch family. Our data suggest that the level of regulation attained is similar between the preQ$_1$-I$_I$ *Can* riboswitch and

the preQ$_1$-II *Lrh* riboswitch, despite differences in preQ$_1$ binding stiochiometry[37, 48] (**Fig. 3d**). This result suggests that these two disparate riboswitch folds evolved equally effective chemical networks to sense a common effector for gene regulation. Yet, cooperativity is expected to provide notable benefits in regulation efficiency. One such advantage is that gene expression is permitted when metabolite levels are low (**Fig. 3b**, *middle panel*), while assuring the ability to quickly attenuate expression before excess effector accrues in the cell[46]. This is reasonable considering that many preQ$_1$-I riboswitches control the translation of transporters that salvage Q-precursor metabolites from the extracellular environment[23, 28, 49].

Although the *Can* and *Lrh* riboswitches differ in terms of binding stoichiometry and overall fold, each positions its expression platform near the binding pocket. This organization raises the question of whether dual, stacked metabolite binding could be effective to regulate folds in which the aptamer is located distally from the expression platform. PreQ$_1$-III riboswitches exemplify this organization, wherein the expression platform can be as far as 40 Å away from the aptamer[38]. Communication between the single-effector pocket and an orthogonal SDS-antiSDS helix is mediated by an A minor base that makes a T-shaped contact with the edge of preQ$_1$[38] (**Supplementary Fig. 4e**). The preQ$_1$-II riboswitch uses a similar pocket[37] and the A-minor base was shown to be essential for gene-regulatory function[40]. It is conceivable that dual, stacked effector recognition could be used by the preQ$_1$-III riboswitch fold, if the effectors promoted coaxial helical stacking, and one or both were detected by an A-minor motif. Accordingly, we predict that additional riboswitches that bind dual, stacked effectors exist in nature.

Extant riboswitches can also provide clues about the organization of extinct ribozymes[1]. Riboswitches that utilize distal binding domains to accommodate a single ligand suggest how the folds of early ribozymes were organized to position substrates[1]. Our findings extend this concept to single-domain ribozymes. In particular, the *Can* aptamer shows how a ribozyme could position two substrates in one pocket to promote covalent bond formation. Intriguingly, the

α-site primary amine is solvent accessible (**Supplementary Fig. 7**), providing a key functional group absent from the RNA chemical repertoire[50]; in contrast, the β-preQ$_1$ WC face is solvent accessible. Notably, O6-methyl preQ$_1$ shows site-specific preQ$_1$-I$_{II}$ riboswitch methylation[51], providing a precedent for ligand-mediated chemical transformation of RNA. These observations collectively suggest how a ribozyme could position two substrates within a single compact fold to facilitate chemistries required for prebiotic metabolism.

PreQ$_1$-I$_I$ riboswitches are prominent in human pathogens[23, 28] including *Ngo*, an urgent public-health threat[52]. The mode of effector recognition by the preQ$_1$-I$_I$ riboswitch provides new opportunities to target such regulatory RNAs. For example, a single small molecule that simultaneously occupies both α and β binding sites could reduce cross-reactivity with targets that recognize preQ$_1$-like molecules (e.g., guanine), yielding greater potency and reduced toxicity. Our results suggest that such riboswitches merit further exploration for their potential as antimicrobial targets.

MATERIALS AND METHODS

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

RNA Purification

RNA strands were synthesized by Dharmacon (Lafayette, CO) as described by the manufacturer except that deprotection heating was 30 min at 65 °C. RNA was purified by 15% denaturing PAGE and DEAE chromatography[53]. DEAE buffer was replaced with 0.02 M Na-HEPES pH 6.8, 0.10 M ammonium acetate, and 0.002 M EDTA; care was taken to minimize UV

exposure[54]. After ethanol precipitation of pooled DEAE fractions, RNA was dissolved in Nanopure™ UV/UF (ThermoFisher) water and desalted on a PD-10 column (GE Healthcare). Quality was assessed by analytical PAGE stained with SYBR Gold (Thermo-Fisher) and visualized on a GelDoc (BioRad XR+). Yield was measured spectrophotometrically. Lyophilized RNA was stored at −20 °C.

Structure Determination

Lyophilized RNA was dissolved in 20 µL of 0.01 M sodium cacodylate pH 7.0 and concentrated to 800 µM by centrifugation. Separate volumes of the concentrated riboswitch and an equal volume of folding buffer (0.004 M $MgCl_2$, 0.01 mM sodium cacodylate pH 7.0 and 0.0016 M $preQ_1$) were heated at 65 °C for 3 min. The folding mix was added dropwise to the RNA and heated 3 min at 65 °C, followed by slow cooling to 24 °C.

Crystals were grown from VDX plates (Hampton Research) by hanging-drop vapor-diffusion. A 1 µL volume of RNA was combined with 1 µL of precipitant drawn from 1 mL in the well. Crystals grew from solutions of 30% (v/v) 2-methyl-2,4-pentanediol, 0.08 M KCl, 0.012 M NaCl, 0.04 M sodium cacodylate pH 5.5, and 0.002 M hexammine cobalt (III) chloride. Crystals grew in 3 weeks at 20 °C as hexagonal rods of size 0.125 mm × 0.040 mm × 0.040 mm. Crystals were cryo-protected by 2 min transfers into well solution supplemented with 40% to 60% (v/v) 2-methyl-2,4-pentanediol. Single rods were captured in nylon loops using 16 mm copper pins (Hampton Research) with the *c**-axis oriented parallel to the φ axis. Crystals were plunged into $N_2(l)$ for shipping to the Stanford Synchrotron Radiation Lightsource (SSRL).

X-ray data were collected remotely on beamline 12-2 using Blu-Ice software and the Stanford Auto-Mounter[55-57] at a λ of 0.9800 Å with a Δφ of 0.15°, an exposure time of 0.7 s per image with 450 total images, and a sample-to-detector distance of 425 mm at 100 K. All data were recorded on a PILATUS 6M detector (Dectris Inc). Data-collection strategies were generated using Web-Ice[55]. Diffraction data were reduced with autoxds[58] using XDS,

POINTLESS, AIMLESS and TRUNCATE[59-61].The structure was determined by molecular replacement in PHENIX[61] starting from the *B. subtilis* preQ$_1$-I riboswitch (Protein Data Bank entry 3FU2). The top solution for three molecules in the asymmetric unit produced a TFZ of 9.2 and a log-likelihood gain of 289. The structure was built in COOT with additional refinement in PHENIX[61]. Intensity and refinement statistics are in **Supplementary Table 1**. Cartoons, schematic diagrams and surfaces of coordinates were generated in PyMOL (Schrödinger LLC). In **Supplementary Fig. 7**, preQ$_1$ atoms were colored by solvent accessible surface area using the color area (solvent) function in PyMOL (Schrödinger LLC). Reported solvent accessible surface area were calculated in PISA[62] (PDBe PISA v1.52) for chain A.

Isothermal Titration Calorimetry

Each sample was folded by dissolving lyophilized RNA in 250 µL 0.01 M sodium cacodylate pH 7.0. RNA was heated to 65 °C for 3 min and mixed with an equal volume of preheated folding buffer at 65 °C comprising 0.01 M sodium cacodylate pH 7.0 and 0.004 M $MgCl_2$. The combined solution was heated for an additional 3 min, then slow cooled to 24 °C followed by overnight dialysis against 2 L of ITC buffer (0.050 M Na-HEPES pH 7.0, 0.10 M NaCl and 0.004 M $MgCl_2$) using a 3,500 MWCO Slide-A-Lyzer Dialysis Cassette G2 (Thermo-Scientific). PreQ$_1$ from a 0.020 M stock in water was diluted to 0.0010 M in ITC buffer.

ITC was conducted using two different instruments. Experiments with WT *Can*, *Ngo* and *Hin* riboswitches at 25 °C were conducted on a PEAQ-ITC (Malvern) with RNA in the cell and preQ$_1$ in the syringe over 19 injections. Experiments were carried out with an injection volume of 4 µL (0.5 µL technical injection) and a spacing of 150 s. These thermograms were analyzed with MicroCal PEAQ-ITC Analysis software (Malvern Panalytical, Inc) using a ‘single-sites’ binding model, which corresponds to the independent sites model below.

To obtain additional data points for cooperativity analysis, WT experiments were also conducted at 37 °C on a VP-ITC (MicroCal). Experiments were carried out with an injection

volume of 10 µL (6 µL technical injection) and a spacing of 240 s with RNA in the cell and $preQ_1$ in the syringe over 29 injections. Mutant riboswitches were analyzed similarly but at 25 °C due to poor binding. These thermograms were analyzed using a 'two-interdependent non-equivalent sites' model (**Supplementary Fig. 5b**) and described below.

In each case, at least two measurements were performed for each RNA sample on the appropriate instrument. Representative thermograms and curve fits are provided in **Supplementary Fig. 5**. Thermodynamic parameters for experiments performed on the PEAQ ITC are in **Supplementary Table 2** and experiments on the VP-ITC in **Supplementary Table 3**. Macroscopic $\Delta G°$ values for mutant riboswitches represent the sum of microscopic $\Delta G°$ values, which were obtained by calculating K_{rel} at each binding event versus the WT riboswitch at 25 °C. The concentrations of RNA and $preQ_1$ used in ITC experiments are reported in the source data file.

Least-squares regression analysis of ITC experiments (Two Interdependent Non-Equivalent Sites Model)

ITC experiments performed on the VP instrument produced parabolic thermograms indicative of cooperativity but these could not be satisfactorily fit with conventional ITC software as noted[39]. Structural evidence indicates that the $preQ_1$ ligands interact in their respective binding pockets, suggesting that a cooperative binding model in which the two effector-binding sites are non-equivalent and interdependent was appropriate. We implemented this model (**Supplementary Fig. 5b**) in a custom Python program based on the binding polynomial theory[63].

Rather than fitting an apparent stoichiometry, we fixed the number of binding sites to exactly two and fit a nuisance parameter that represents the effective concentration of active riboswitch RNA in the ITC cell relative to the recorded concentration[63, 64]. Although the binding model describes a binding enthalpy and a microscopic dissociation constant for each of four

distinct binding equilibria (**Supplementary Fig. 5b**), there are only three independent microscopic dissociation constants:

$$K_{D,A1} = \frac{[R][L]}{[RL_A]} \quad K_{D,B1} = \frac{[R][L]}{[RL_\beta]} \quad K_{D,A2} = \frac{[RL_B][L]}{[RL_{AB}]} \quad K_{D,B2} = \frac{[RL_A][L]}{[RL_{AB}]} \tag{1}$$

$$K_{D,A1}K_{D,B2} = K_{D,B1}K_{D,A2} = \frac{[R][L]^2}{[RL_{AB}]} \tag{2}$$

Likewise, there are only three independent binding enthalpies because enthalpy is a state function; completing a thermodynamic cycle must result in no enthalpy change.

$$\Delta H^\circ_{A1} + \Delta H^\circ_{B2} - \Delta H^\circ_{A2} - \Delta H^\circ_{B1} = 0 \tag{3}$$

The binding polynomial results in a cubic equation in the concentration of free ligand $[L]$

$$[L]^3 + \left(2R_T - L_T + K_{D,A2} + K_{D,B2}\right)[L]^2$$

$$+ \left((R_T - L_T)\left(K_{D,A2} + K_{D,B2}\right) + K_{D,A1}K_{D,B2}\right)[L] - L_T K_{D,A1}K_{D,B2} = 0 \tag{4}$$

where $R_T = [R] + [RL_A] + [RL_B] + [RL_{AB}]$ is the total concentration of RNA in the ITC cell and $L_T = [L] + [RL_A] + [RL_B] + 2[RL_{AB}]$ is the total concentration of preQ$_1$ in the ITC cell. We solved this cubic equation analytically by choosing the root that satisfies $[L] = 0$ when $L_T = 0$.

Following a recent approach[65], we explicitly accounted for the dilution of all chemical species present due to displacement of the liquid in the ITC cell by the injection volume. The differential changes in the concentrations of bound species due to a differential injected volume dV are

$$d[RL_A] = \frac{1}{V_0}(-[RL_A]dV + d\Phi_{A1} - d\Phi_{B2})$$

$$d[RL_B] = \frac{1}{V_0}(-[RL_B]dV + d\Phi_{B1} - d\Phi_{A2}) \tag{5}$$

$$d[RL_{AB}] = \frac{1}{V_0}(-[RL_{AB}]dV + d\Phi_{A2} + d\Phi_{B2})$$

where V_0 is the volume of the ITC cell and Φ_i is the flux through the binding equilibrium i. The enthalpy can be expressed as a function of the total injected volume V

$$H(V) = \frac{1}{V_0}\left(\Delta H^{\circ}{}_{A1}\frac{d\Phi_{A1}}{dV} + \Delta H^{\circ}{}_{B1}\frac{d\Phi_{B1}}{dV} + \Delta H^{\circ}{}_{A2}\frac{d\Phi_{A2}}{dV} + \Delta H^{\circ}{}_{B2}\frac{d\Phi_{B2}}{dV}\right) \quad (6)$$

The enthalpy change associated with a particular injection that brings the stoichiometric ratio of ligand to receptor S from S_{i-1} to S_i is given by the average value of the enthalpy over this interval[65].

$$\Delta H_i = \frac{1}{S_i - S_{i-1}}\int_{S_{i-1}}^{S_i} dS\, H(S) \quad (7)$$

Inserting Eqs. 1–3 and Eqs. 5–6 into Eq. 7 and using integration by parts gives the injection enthalpy change in terms of the fit parameters, the ITC cell volume, the initial concentrations of RNA in the ITC cell R_0 and of ligand in the syringe L_0, and the concentration of free ligand obtained as the solution to Eq. 4.

$$\Delta H_i = \frac{\Omega(S_i) - \Omega(S_{i-1})}{L_0(S_i - S_{i-1})}$$

$$\Omega(S) = \frac{\left(\frac{L_0}{R_0} + S\right)(L_T - [L])\left(\Delta H_{A1}K_{D,B2} + \Delta H_{B1}K_{D,A2} + (\Delta H_{A1} + \Delta H_{B2})[L]\right)}{K_{D,A2} + K_{D,B2} + 2[L]} \quad (8)$$

We used a trust region reflective algorithm[66] implemented in the optimize.least_squares() method of SciPy[67] to minimize the following cost function:

$$F(\vec{\theta}, \lambda) = \sum_{i=1}^{N}\left(\Delta H_i(\vec{\theta}) - \Delta H_{i,obs}\right)^2 + \lambda\sum_{j=1}^{M}\left(\frac{\theta_j - \theta_{j,0}}{w_j}\right)^2 \quad (9)$$

where N is the number of observed injections and M is the number of fit parameters. The first term is a least-squares term describing the goodness-of-fit between the estimated and observed injection enthalpy changes. The second term is an L2 regularization term—whose relative strength is controlled by the hyperparameter λ—that prevents overfitting by penalizing deviations of the fit parameters θ from a target value θ_0. In a Bayesian framework, this penalty is interpreted as a Gaussian prior on the fit parameters with mean θ_0 and standard deviation w[68]. For the three independent microscopic dissociation constants, regularization was applied to the natural

logarithm of the dissociation constant. The regularization targets were set to the values of the fit parameters from a binding model assuming two independent and equivalent binding sites, i.e. the model used by most commercial ITC software. We derived analytical derivatives of the cost function given by Eq. 9 with respect to the fit parameters to take advantage of computationally efficient gradient-based optimization methods.

For each RNA sequence, we performed a global fit to obtain a single set of fit parameters informed by multiple experiments in which the initial concentrations of riboswitch receptor and ligand vary in order to interrogate different regions of the resulting thermogram. One offset parameter, a constant added to the estimated injection enthalpy changes, was fit for each experiment. The hyperparameter λ controlling the relative strength of the regularization term was optimized for each RNA sequence individually by cross validation across experiments. Each experiment was fit individually for a sequence of λ with logarithmic spacing—$\log_{10} \lambda$ was varied from −6 to +6 in steps of 0.125. The resulting fit parameters were used to estimate the value of the cost function for the other experiments involving the same RNA sequence. The value of λ with the smallest average value of the cost function for experiments not used to train the parameters was chosen for the global fit.

Two sets of values were used for the regularization weights w. For *Can* WT, *Can* C31U and *Hin* WT, the regularization weights were $16\ k_B T$ for the binding enthalpies and $\log 10$ for the logarithms of the dissociation constants, where k_B is the Boltzmann constant and T is the absolute temperature. However, these weights produced poor quality fits for *Can* C17U and *Ngo* WT, as revealed by fit parameters with large bootstrapped uncertainties. As such, the regularization weights for these sequences were $16\ k_B T$ for both the binding enthalpies and the logarithms of the dissociation constants. For all RNA sequences, the regularization weights were 1 kcal mol^{-1} for the offsets and 0.05 for the nuisance parameter describing the effective RNA concentration.

To derive estimates and 95% confidence intervals for the fit parameters, we used a bootstrapping method to resample the fitting target in the nonlinear regression[69]. In each bootstrap iteration, we added the residual from the initial fit multiplied by a random number sampled from a standard normal distribution to the observed injection enthalpy changes. The resulting distributions of fit parameters are non-normal, and so we report the estimate of each fit parameter as the median of the bootstrap parameter distribution. We also report a 95% confidence interval as the (2.5, 97.5) percentiles of this bootstrap distribution.

In addition to the fit parameters, we also calculate the following derived parameters: cooperativity C, macroscopic dissociation constants $K_{D,1}$ and K_{D2}, and a macroscopic cooperativity γ.

$$C = \frac{K_{D,A1}}{K_{D,A2}}$$

$$K_{D,1} = \left(\frac{1}{K_{D,A1}} + \frac{1}{K_{D,B1}}\right)^{-1}$$

$$K_{D,2} = K_{D,A2} + K_{D,B2}$$

$$\gamma = \frac{4K_{D,1}}{K_{D,2}} \quad (10)$$

We obtained estimates and 95% confidence intervals for derived parameters by calculating the derived parameters for each bootstrap iteration and then reporting the median and (2.5, 97.5) percentile of the bootstrap parameter distribution.

In-cell GFPuv reporter assay

The WT *Can* riboswitch was placed into the pBR327-*Lrh*(WT)-GFPuv plasmid upstream of the GFPuv reporter gene (**Fig. 3a**). Riboswitch mutants were prepared by site-directed mutagenesis (GenScript Inc.) on the WT sequence, which were verified by DNA sequencing.

462 For experiments involving the *Lrh* riboswitch, the parent pBR327-*Lrh*(WT)-GFPuv plasmid was
463 used[13].

464 The assay was performed as described[13, 40] with some exceptions. *E. coli* strain JW2765
465 Δ*queF* cells (Coli Genetic Stock Center, Yale University) — incapable of $preQ_1$ biosynthesis —
466 were transformed with the desired plasmid and grown on CSB agar plates containing both
467 ampicillin (100 μg mL^{-1}) and kanamycin (50 μg mL^{-1}). Single colonies were isolated to inoculate
468 overnight liquid cultures of 3 mL CSB-amp-kan media. These were used to inoculate 2 mL of
469 fresh CSB-amp-kan media with varying concentrations of $preQ_1$: 0, 1 nM, 10 nM, 50 nM, 75 nM,
470 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 2.5 μM, 5 μM, 7.5 μM 10 μM, 50 μM, 100 μM, 500 μM,
471 1 mM, and 3 mM; the highest concentration corresponds to the solubility limit of $preQ_1$ in CSB[13].
472 Three or more biological replicates were measured for each concentration. All measurements
473 and analysis were performed as described[13] using Prism (GraphPad Software, Inc). The
474 replicates in each construct were compared using the "compare datasets" function before
475 analysis. The WT *Can* curve showed a biphasic model whereas others were best described by
476 a log(inhibitor) dose versus response (3 parameter).

477 An unpaired student's t-test with a Welch's correction was used to analyze fold
478 repression data. The *p* value for WT *Can* vs WT *Lrh* was 0.6429 (t = 0.503, degrees of freedom
479 (df) = 3.76, 95% confidence interval = -4.515 to 6.453). The *p* value for WT *Can* vs C17U *Can*
480 was 0.0125 (t = 5.94, df = 2.72, 95% confidence interval = -15.47 to -4.266). The *p* value for WT
481 *Can* vs C31U *Can* was 0.0112 (t = 9.23, df = 2.02, 95% confidence interval = -20.61 to -7.583).
482 The *p* value for WT *Can* vs the negative control was 0.0106 (t = 9.59, df = 2.01, 95% confidence
483 interval = -21.16 to -8.706). The *p* value for WT *Can* vs the positive control was 0.0103 (t = 9.71,
484 df = 2.01, 95% confidence interval = -21.3400 to -8.2800).

485 Notably, fluorescence emission in the absence of $preQ_1$ is comparable between all
486 riboswitch constructs and the positive control; moreover, the WT *Can* and *Lrh* sequences
487 repress GFPuv fluorescence emission to a level comparable to the negative control —

demonstrating the rigor of the assay (**Supplementary Fig. 8**). An unpaired student's t-test with a Welch's correction was also used to analyze fluorescence emission data. The *p* value for WT *Can* was 0.0037 (t = 14.96, df = 2.09, 95% confidence interval = -173452 to -98511). The *p* value for C17U *Can* was 0.0175 (t = 6.64, df = 2.18, 95% confidence interval = -212164 to -53149). The *p* value for C31U *Can* was 0.1263 (t = 2.05, df = 3.22, 95% confidence interval = -107817 to -21313). The *p* value for WT *Lrh* was 0.0037 (t = 14.96, degrees of freedom (df) = 2.09, 95% confidence interval = -173452 to -98511). The *p* value for the positive control was 0.7441 (t = 0.35, df = 3.45, 95% confidence interval = -61691 to -48527). The *p* value for the negative control was 0.0631 (t = 2.60, df = 3.80, 95% confidence interval = -11377 to 488.8).

DATA AVAILABILITY

Structure factor amplitudes and coordinates for the *Can* preQ$_1$-I$_I$ riboswitch were deposited in the Protein Data Bank. Publicly available PDB entries used in this study are: 6VUI, 3FU2, 4RZD and 4JF2. Source data files are available in the Supplementary Data, which includes injection data for ITC, and fluorescence emission and cell growth readings for in-cell assays.

CODE AVAILABILITY

The ITC fitting software and parameter fits are available on GitHub at https://github.com/chapincavender/itc_two_site_fit, distributed under the MIT license.

REFERENCES

1. Nahvi, A. et al. Genetic control by a metabolite binding mRNA. *Chem Biol* **9**, 1043-1049 (2002).

2. McCown, P. J. et al. Riboswitch diversity and distribution. *RNA* **7**, 995-1011 (2017).

3. Breaker, R. R. Riboswitches and Translation Control. *Cold Spring Harb. Perspect. Biol.* **10**, 1-14 (2018).

4. Suddala, K. C. et al. Single transcriptional and translational preQ1 riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure. *Nucleic Acids Res.* **41**, 10462-75 (2013).

5. Hua, B. et al. Real-time monitoring of single ZTP riboswitches reveals a complex and kinetically controlled decision landscape. *Nat. Commun* **11**, 4531-4531 (2020).

6. Stagno, J. R. et al. Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* **541**, 242-246 (2017).

7. Frieda, K. L. & Block, S. M. Direct Observation of cotranscriptional folding in an adenine riboswitch. *Science* **338**, 397-400 (2012).

8. Babina, A. M., Lea, N. E., & Meyer, M. M. *In vivo* behavior of the tandem glycine riboswitch in *Bacillus subtilis*. *mBio* **8**, 5 (2017).

9. McCarty, R. M. & Bandarian, V. Biosynthesis of pyrrolopyrimidines. *Bioorg. Chem.* **43**, 15-25 (2012).

10. Harada, F. & Nishimura, S. Possible anticodon sequences of tRNAHis, tRNAAsn, and tRNAAsp from *Escherichia coli*. Universal presence of nucleoside O in the first position of the anticodons of these transfer ribonucleic acid. *Biochemistry* **11,** 301-308 (1972).

11. Noguchi, S., Nishimura, Y., Hirota, Y. & Nishimura, S. Isolation and characterization of an Escherichia coli mutant lacking tRNA-guanine transglycosylase. Function and biosynthesis of queuosine in tRNA. *J Biol. Chem.* **257**, 6544-50 (1982).

12. Tuorto, F. et al. Queuosine-modified tRNAs confer nutritional control of protein translation. *Embo J* **37,** e99777 (2018).

13. Dutta, D., Belashov, I. A. & Wedekind, J. E., Coupling green fluorescent protein expression with chemical modification to probe functionally relevant riboswitch conformations in live bacteria. *Biochemistry* **57**, 4620-4628 (2018).

14. Hurt, J. K., Olgen, S. & Garcia, G. A., Site-specific modification of *Shigella flexneri* virF mRNA by tRNA-guanine transglycosylase *in vitro*. *Nucleic Acids Res.* **35,** 4905-4913 (2007).

15. Kang, M.; Peterson, R. & Feigon, J. Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol. Cell* **33**, 784-90 (2009).

16. Spitale, R. C., Torelli, A. T., Krucinska, J., Bandarian, V. & Wedekind, J. E. The structural basis for recognition of the $preQ_0$ metabolite by an unusually small riboswitch aptamer domain. *J Biol. Chem.* **284**, 11012-11016 (2009).

17. Klein, D. J., Edwards, T. E. & Ferre-D'Amare, A. R., Cocrystal structure of a class I $preQ_1$ riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat. Struc. Mol Biol.* **16,** 343-344 (2009).

18. Schroeder, G. M. et al. Analysis of a $preQ_1$-I riboswitch in effector-free and bound states reveals a metabolite-programmed nucleobase-stacking spine that controls gene regulation. *Nucleic Acids Res.* **48**, 8146-8164 (2020).

19. Rinaldi, A. J., Lund, P. E., Blanco, M. R. & Walter, N. G.. The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs shows ligand-dependent accessibility bursts. *Nat Commun* **7,** 8976 (2016).

20. Banas, P., Sklenovsky, P., Wedekind, J. E., Sponer, J. & Otyepka, M., Molecular mechanism of $preQ_1$ riboswitch action: a molecular dynamics study. *J. Phys. Chem. B* **116**, 12721-12734 (2012).

21. Rieder, U., Lang, K., Kreutz, C., Polacek, N. & Micura, R.. Evidence for pseudoknot formation of class I $preQ_1$ riboswitch aptamers. *Chembiochem* **10**, 1141-4 (2009).

22. Jenkins, J. L., Krucinska, J., McCarty, R. M., Bandarian, V. & Wedekind, J. E., Comparison of a $preQ_1$ riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J. Biol. Chem* **286**, 24626-37 (2011).

23. Roth, A. et al. A riboswitch selective for the queuosine precursor $preQ_1$ contains an unusually small aptamer domain. *Nat Struct Mol Biol* **14**, 308-17 (2007).

24. Suddala, K. C., Wang, J., Hou, Q. & Walter, N. G. Mg^{2+} shifts ligand-mediated folding of a riboswitch from induced-fit to conformational selection. *J. Am. Chem. Soc.* **137**, 14075-14083 (2015).

25. Connelly, C. M. et al. Synthetic ligands for $PreQ_1$ riboswitches provide structural and mechanistic insights into targeting RNA tertiary structure. *Nat. Commun.* 10, 1501 (2019).

26. Neuner, E., Frener, M., Lusser, A. & Micura, R. Superior cellular activities of azido- over amino-functionalized ligands for engineered $preQ_1$ riboswitches in *E. coli*. *RNA Biol* **15**, 1376-1383 (2018).

27. Eichhorn, C. D., Kang, M. & Feigon, J., Structure and function of $preQ_1$ riboswitches. *Biochim. Biophys. Acta* **1839**, 939-950 (2014).

28. McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. Structural, functional, and taxonomic diversity of three $preQ_1$ riboswitch classes. *Chem Biol* **21**, 880-889 (2014).

29. Zhu, S et al. Complete genome sequence of hemolysin-containing *Carnobacterium* sp. strain CP1 isolated from the antarctic. *Genome Announc.* **4**, e00690-16 (2016).

30. Leontis, N. B. & Westhof, E., Geometric nomenclature and classification of RNA base pairs. *RNA* **7**, 499-512 (2001).

31. Ren, A. & Patel, D. J. c-di-AMP binds the ydaO riboswitch in two pseudo-symmetry-related pockets. *Nat. Chem. Biol.* **10**, 780-786 (2014).

32. Jones, C. P. & Ferre-D'Amare, A. R. Crystal structure of a c-di-AMP riboswitch reveals an internally pseudo-dimeric RNA. *EMBO J.* **33**, 2692-703 (2014).

33. Reiss, C. W. & Strobel, S. A., Structural basis for ligand binding to the guanidine-II riboswitch. *RNA* **23**, 1338-1343 (2017).

34. Trausch, J. J., Ceres, P., Reyes, F. E. & Batey, Robert T., The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer. *Structure* **19**, 1413-1423 (2011).

35. Huang, L., Wang, J. & Lilley, D. M. J. The structure of the guanidine-II riboswitch. *Cell Chem. Biol.* **24**, 695-702.e2 (2017).

36. Gao, A. & Serganov, A. Structural insights into recognition of c-di-AMP by the ydaO riboswitch. *Nat. Chem. Biol.* **10**, 787-792 (2014).

37. Liberman, J. A., Salim, M., Krucinska, J. & Wedekind, J. E. Structure of a class II $preQ_1$ riboswitch reveals ligand recognition by a new fold. *Nat. Chem. Biol.* **9**, 353-355 (2013).

38. Liberman, J. A. et al. Structural analysis of a class III $preQ_1$ riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3485-94 (2015).

39. Huang, L., Serganov, A. & Patel, D. J. Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. *Mol. Cell.* **40**, 774-786 (2010).

40. Dutta, D. & Wedekind, J. E., Nucleobase mutants of a bacterial $preQ_1$-II riboswitch that uncouple metabolite sensing from gene regulation. *J. Biol. Chem.* **295**, 2555-2567 (2020).

41. Liberman, J. A., Bogue, J. T., Jenkins, J. L., Salim, M. & Wedekind, J. E. ITC analysis of ligand binding to $preQ_1$ riboswitches. *Meth. Enzymol.* **549**, 435-50 (2014).

42. Sudarsan, N. et al. Tandem riboswitch architectures exhibit complex gene control functions. *Science* **314**, 300-304 (2006).

43. Jones, C. P. & Ferre-D'Amare, A. R. Long-range interactions in riboswitch control of gene expression. *Annu. Rev. Biophys.* **46**, 455-481 (2017).

44. Zallot, R., Yuan, Y. & de Crecy-Lagard, V. The *Escherichia coli* COG1738 member YhhQ is involved in 7-Cyanodeazaguanine $preQ_0$ transport. *Biomolecules* **7,** 1-13 (2017).

45. Yuan, Y. et al. Discovery of novel bacterial queuine salvage enzymes and pathways in human pathogens. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 19126-19135 (2019).

46. Breaker, R. R., Riboswitches and the RNA world. *Cold Spring Harb. Perspect. Biol.* **4,** a003566 (2012).

47. Watson, P. Y. & Fedor, M. J. The *glmS* riboswitch integrates signals from activating and inhibitory metabolites *in vivo*. *Nat. Struct. Mol. Biol.* **18**, 359-363 (2011).

48. Kang, M., Eichhorn, C. D. & Feigon, J. Structural determinants for ligand capture by a class II preQ$_1$ riboswitch. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E663-E671 (2014).

49. Rodionov, D. A. et al. A novel class of modular transporters for vitamins in prokaryotes. *J Bacteriol.* **191**, 42-51 (2009).

50. Wilson, T. J. & Lilley, D. M. J., The potential versatility of RNA catalysis. *Wiley Interdiscip. Rev. RNA.* **12**, e1651 (2021).

51. Flemmich, L., Heel, S., Moreno, S., Breuker, K. & Micura, R. A natural riboswitch scaffold with self-methylation activity. *Nat. Commun.* **12**, 3877 (2021).

52. CDC Biggest Threats and Data: 2019 Antibiotic Resistance Threats Report. https://www.cdc.gov/drugresistance/biggest-threats.html.

53. Lippa, G. M. et al. Crystallographic analysis of small ribozymes and riboswitches. *Methods Mol Biol* **848**, 159-84 (2012).

54. Kladwang, W., Hum, J. & Das, R. Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. *Sci Rep* **2**, 517 (2012).

55. Gonzalez, A. et al. Web-Ice: integrated data collection and analysis for macromolecular crystallography. *J.App. Crystall.* **41**, 176-184 (2008).

56. Soltis, S. M. et al. New paradigm for macromolecular crystallography experiments at SSRL: automated crystal screening and remote data collection. *Acta Crystallogr D Biol Crystallogr* **64,** 1210-21 (2008).

57. McPhillips, T. M. et al. Blu-Ice and the distributed control system: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J Synchrotron. Radiat.* **9**, 401-406 (2002).

58. Gonzalez, A. &. Tsai, Y. Autoxds. (2010) http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script.

59. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**,125-132 (2010).

60. Kabsch, W., Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133-414 (2010).

61. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.D* **66**, 213-221 (2010).

62. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774-797 (2007).

63. Freire, E., Schön, A. & Velazquez-Campoy, A. Chapter 5 Isothermal Titration Calorimetry: general formalism using binding polynomials. *Methods Enzymol.* **455**, 127-155 (2009).

64. Duvvuri, H., Wheeler, L. C. & Harms, M. J.. pytc: open-source Python software for global analyses of isothermal titration calorimetry data. *Biochemistry* **57**, 2578-2583 (2018).

65. Dumas, P. Rigorous equations for isothermal titration calorimetry: theoretical and practical consequences. Preprint at https://www.biorxiv.org/content/10.1101/512780v3 (2019).

66. Branch, M. A., Coleman, T. F. & Li, Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM J. on Sci. Comput.* **21**, 1-23 (1999).

67. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261-272 (2020).

68. Williams, P. M. Bayesian regularization and pruning using a laplace prior. *Neural Comput.* **7**, 117-143 (1995).

69. Wu, C. F. J. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Stat.* **14**, 1261-1295 (1986).

SUPPLEMENTARY INFORMATION

Supplementary Tables, Figures and Data are available at Nature Communications online.

ACKOWLEDGMENTS

We thank members of the Wedekind and Mathews labs for technical assistance and helpful discussions.

FUNDING

Support for this research was provided by the National Institutes of Health National Institute of General Medical Sciences (NIH NIGMS) grants R01 GM132185 to D.H.M and R01 GM063162 to J.E.W. G.M.S. was supported by training grant T32 GM118283, and C.E.C was supported by training grant T32 AI049815 to J.E.W. and an Elon Huntington Hooker graduate fellowship. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by NIH

NIGMS (P30 GM133894). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIH or NIGMS.

AUTHOR CONTRIBUTIONS

GMS, CEC, DHM and JEW designed experiments. GMS, MEL performed experiments. GMS, CEC, JLJ, DHM and JEW analyzed the data. GMS, CEC and JEW wrote the paper. All authors reviewed the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing interests.

FIGURES

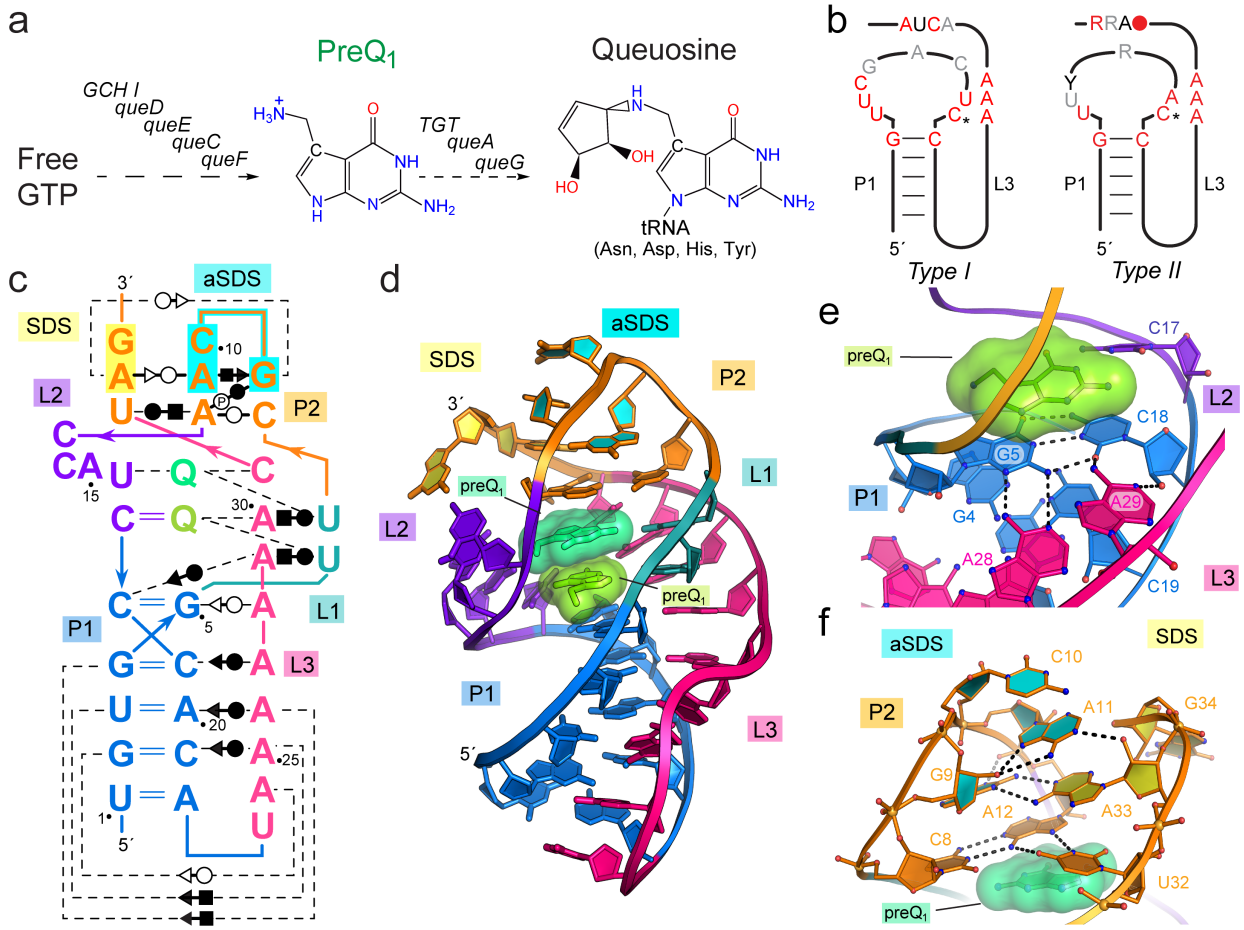


Figure 1: Queuosine biosynthesis, the preQ$_1$ riboswitch consensus model and co-crystal structure of the *Carnobacterium antarcticus* (*Can*) riboswitch. (**a**) The queuosine (Q) biosynthetic pathway proceeds through the 7-deazapurine metabolite preQ$_1$[9]. (**b**) PreQ$_1$-I riboswitch subtypes shown as secondary structures based on covariation[28]. Red, black and gray positions indicate 97%, 90% and 75% sequence conservation. Asterisk indicates a specificity base predicted to recognize preQ$_1$. (**c**) Secondary structure of the *Can* riboswitch. Colors correspond to specific pseudoknot base pairing (P) and loop (L) sequences. PreQ$_1$ is depicted as "Q". Non-canonical pairing is indicated by Leontis–Westhof symbols[30]. The Shine-Dalgarno sequence (SDS) and anti-(a)SDS are highlighted in yellow and cyan. (**d**) Ribbon diagram of the global *Can* riboswitch fold. (**e**) Binding pocket floor overview wherein floor bases comprise the A28•G5-C18 base triple. Dashed lines depict hydrogen bonds here and elsewhere. (**f**) Overview of the pocket ceiling, which comprises the U32•A10•C8 base triple. The view highlights P2 bases in the aSDS and SDS.

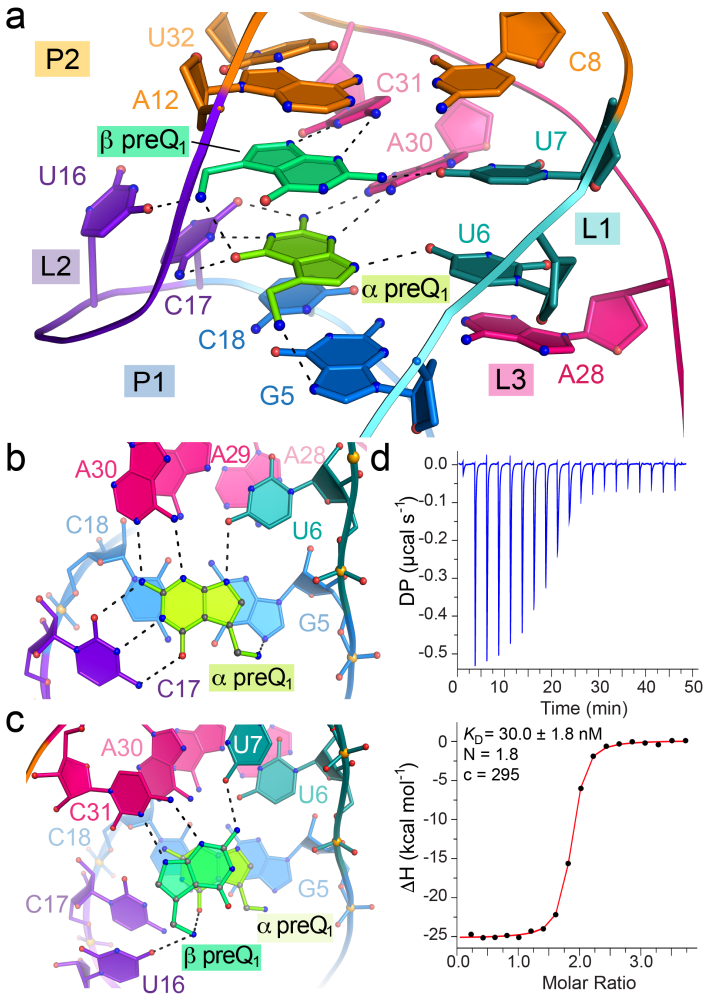


Figure 2: The *Can* preQ$_1$-I$_I$ riboswitch pocket with two preQ$_1$ ligands and confirmation of ligand-to-RNA stoichiometry. (**a**) Overview of fully occupied binding pocket. Interactions in the (**b**) α site and (**c**) β site. (**d**) Representative ITC experiment with titration of preQ$_1$ into WT *Can* RNA. The binding constant K_D, ligand-to-RNA stoichiometry N, and c value are shown.

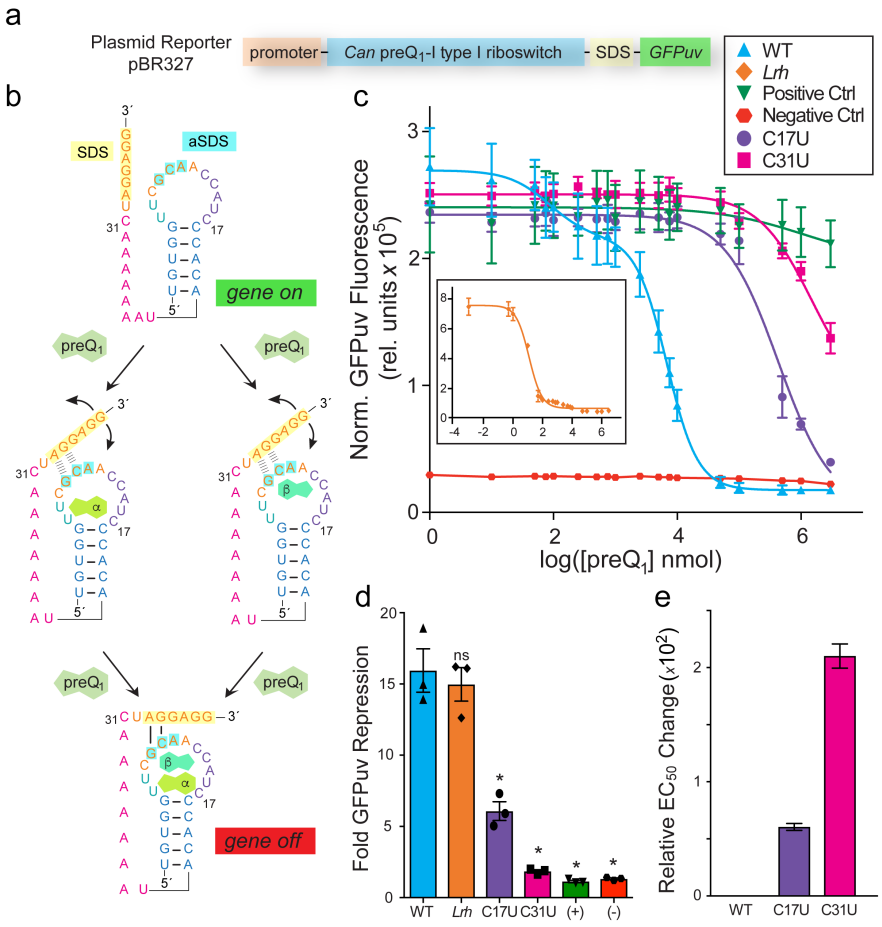


Figure 3: Riboswitch reporter assay and dose response in live bacteria. (**a**) Schematic of the plasmid reporter. (**b**) Two-site binding model wherein $preQ_1$ can bind either site first. (**c**) Average GFP*uv* emission dependence on $preQ_1$; (*inset*) one-site binding by the *Lrh* $preQ_1$ riboswitch[40]. (**d**) Bar graph showing fold repression of GFP*uv* emission for the *Can, Lrh* and mutant riboswitches with individual points shown. (**e**) Bar graph showing fold change in average EC_{50} relative to *Can* riboswitch $EC_{50,2}$. Significance was determined by a Student's *t*-test with Welch's Correction (n = 3). S.E.M. is shown in **c** and **d**; propagated errors are shown in **e**.

Supplementary Information

A small RNA that cooperatively senses two stacked metabolites in one pocket for gene control

Griffin M. Schroeder[1,2], Chapin E. Cavender[1,2], Maya E. Blau[3], Jermaine L. Jenkins[1,2], David H. Mathews[1,2] and Joseph E. Wedekind[1,2]

[1.] *Department of Biochemistry & Biophysics, University of Rochester School of Medicine & Dentistry, Rochester, NY 14642, USA.*

[2.] *Center for RNA Biology, University of Rochester School of Medicine & Dentistry, Rochester, NY 14642, USA.*

[3.] *University of Rochester, 120 Trustee Road, Rochester, NY 14627, USA.*

To whom correspondence should be addressed. Tel: +1 585 273-4516; Email: joseph.wedekind@rochester.edu

Supplementary Table 1: **Data collection and refinement statistics (molecular replacement)**

	$PreQ_1$ Bound *Can* (PDB Entry 7rex)
Data collection	
Space group	*P* 3_2 2 1
Cell dimensions	
a = *b*, *c* (Å)	57.8, 153.6
α = β, γ (°)	90, 120
Resolution (Å)	35.8 – 2.60 (2.72 – 2.60)*
R_{merge} (%)	9.6 (88.0)
$R_{P.I.M.}$ (%)	7.1 (66.7)
I / σ(I)	8.6 (2.0)
Completeness (%)	99.6 (99.1)
Redundancy	4.7 (4.8)
CC1/2	0.99 (0.86)
Refinement	
Resolution (Å)	35.8 – 2.60 (2.72 – 2.60)*
No. reflections	17418
R_{work} / R_{free}	0.232/0.272
No. atoms	
RNA	2042
$preQ_1$/ion	78/4
water	1
B-factors (Å^2)	
RNA	87
$preQ_1$	52
r.m.s. deviations from ideal	
Bond lengths (Å)	0.005
Bond angles (°)	0.849
clash score per 1000 atoms	2.5

*Parenthetical values indicate data in the highest resolution shell.

Supplementary Table 2: Average Thermodynamic Parameters for the Wildtype Type I PreQ$_1$-I Riboswitches

Sequence	$K_{D, app}$ (nM)[a]	N	ΔH_1 (kcal mol^{-1})	$-T\Delta S$ (kcal mol^{-1})	ΔG (kcal mol^{-1})
C. antarcticus	32.0 ± 2.0	1.8 ± 0.01	−25.5 ± 0.2	15.3 ± 0.3	−10.3 ± 0.1
H. influenzae[b]	52.9 ± 0.2	2.2 ± 0.03	−25.6 ± 1.1	15.6 ± 1.0	−10.0 ± 0.1
N. gonorrohoeae[b]	50.5 ± 1.3	2.2 ± 0.06	−21.8 ± 1.0	11.8 ± 1.0	−10.0 ± 0.1

[a] *Measured at 25 °C.*

[b] *Classified as type I preQ$_1$-I riboswitch based on Roth et al*[1]

Supplementary Table 3: Average Thermodynamic Parameters for WT and Mutant Type I PreQ$_1$ Riboswitches

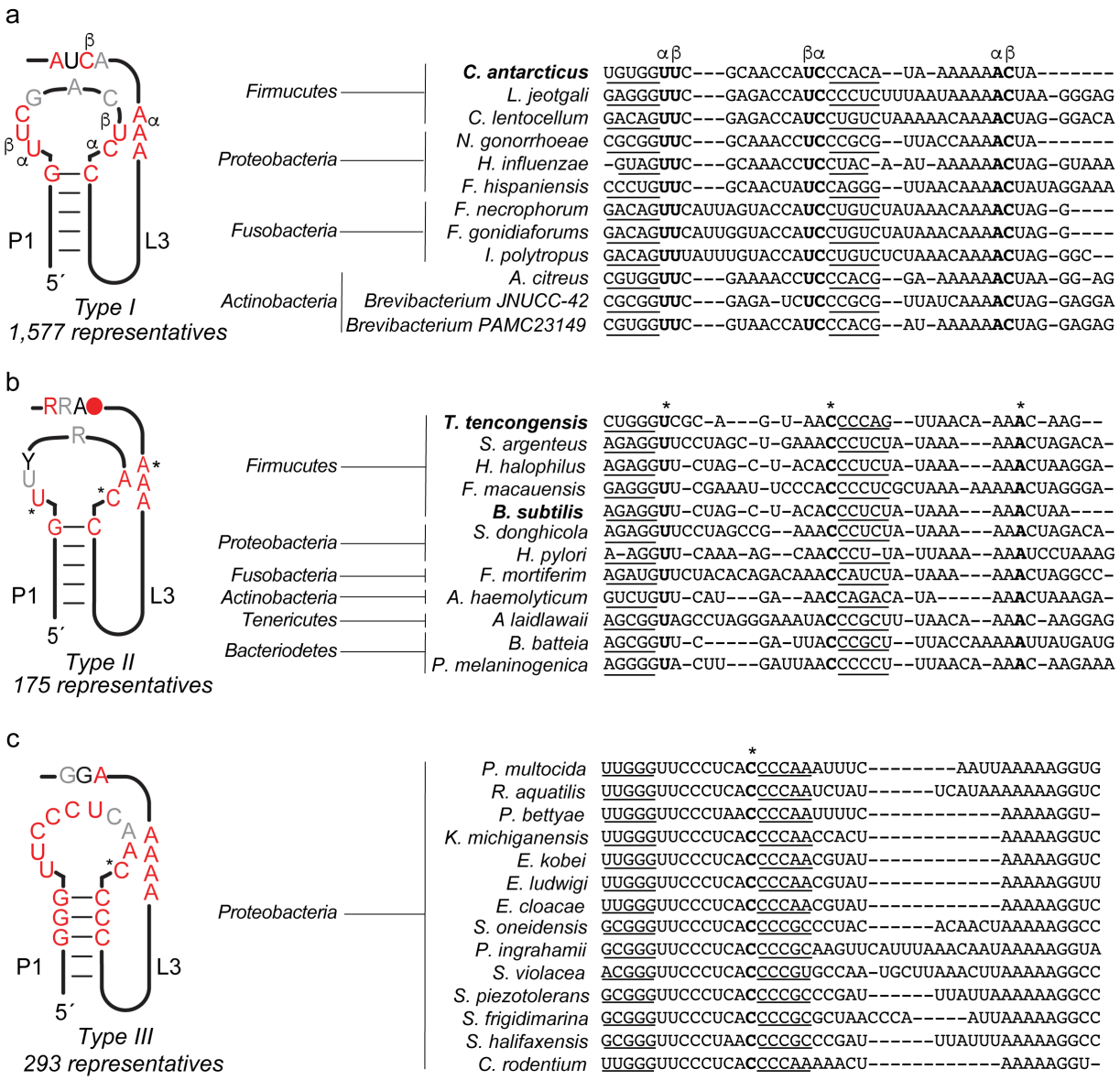
Parameter	*Can*[a] WT	*Can*[b] C17U	*Can*[b] C31U	*Hin*[c] WT	*Ngo*[d] WT
ΔH_{A1} (kcal mol^{-1})	−36.9 (−37.8, −36.3)[e]	−1.6 (−1.8, 0.2)	8.6 (7.6, 9.4)	−40.0 (−40.4, −39.6)	−35.4 (−36.3, −34.6)
ΔH_{B1} (kcal mol^{-1})	−29.9 (−31.4, −28.6)	0.4 (−1.9, 0.8)	−17.7 (−18.7, −15.4)	−26.0 (−26.3, −25.8)	−31.0 (−31.0, −30.9)
ΔH_{A2} (kcal mol^{-1})	−47.4 (−48.8, −45.7)	−4.3 (−4.9, −2.3)	−23.5 (−25.6, −22.4)	−55.1 (−55.5, −54.7)	−37.6 (−39.0, −36.2)
ΔH_{B2} (kcal mol^{-1})	−40.3 (−40.9, −39.5)	−2.3 (−4.4, −2.2)	−49.7 (−51.3, −48.0)	−41.2 (−41.5, −40.9)	−33.2 (−34.7, −31.7)
$K_{D,A1}$ (nM)	1480.0 (1190.0, 1840.0)	5690.0 (5200.0, 6190.0)	9660.0 (8640.0, 11050.0)	3710.0 (3380.0, 4020.0)	1402.0 (1190.0, 1670.0)
$K_{D,B1}$ (nM)	2260.0 (1300.0, 3650.0)	6950.0 (6210.0, 8180.0)	21300.0 (17700.0, 25300.0)	9320.0 (8660.0, 9960.0)	15500.0 (4920.0, 30900.0)
$K_{D,A2}$ (nM)	182.0 (126.0, 273.0)	587.0 (492.0, 803.0)	3200.0 (2430.0, 4180.0)	113.0 (107.0, 120.0)	13.0 (7.0, 39.0)
$K_{D,B2}$ (nM)	278.0 (249.0, 305.0)	716.0 (605.0, 981.0)	7020.0 (6560.0, 7400.0)	284.0 (266.0, 307.0)	142.0 (120.0, 165.0)
C[f]	8.1 (4.5, 14.3)	9.7 (7.3, 11.2)	3.0 (2.6, 3.6)	32.8 (29.7, 35.8)	109.0 (31.0, 249.0)
K_{D1}[g] (nM)	891.0 (630.0, 1208.0)	3130.0 (3030.0, 3220.0)	6640.0 (6420.0, 6830.0)	2650 (2470.0, 2820.0)	1280.0 (974.0, 1580.0)
K_{D2} *(nM)*	461.0 (380.0, 565.0)	1300.0 (1140.0, 1750.0)	10260.0 (9460.0, 10980.0)	398.0 (379.0, 420.0)	156.0 (127.0, 200.0)
γ[h]	7.7 (4.5, 12.6)	9.6 (7.2, 11.0)	2.6 (2.5, 2.7)	26.7 (23.6, 29.6)	32.9 (19.6, 49.1)

[a] *Carnobacterium antarcticus* (*Can*)
[b] ITC was recorded at 25 °C; all other measurements were recorded at 37 °C.
[c] *Haemophilus influenzae* (*Hin*)
[d] *Neisseria gonorrohoeae* (*Ngo*)
[e] Values reported are the median and confidence intervals (2.5, 97.5) from bootstrapping (See **Methods**)
[f] The ratio of microscopic binding constants yields the cooperativity constant C, which shows positive cooperativity when greater than unity.
[g] For simplicity, the microscopic binding constants can be used to generate the macroscopic binding constants, K_{D1} and K_{D2}, corresponding to the first and second ligand binding steps (**Supplemental Fig. 5b**).
[h] The ratio of macroscopic binding constants (K_{D1}, K_{D2}) multiplied by a statistical factor of 4 yields the macroscopic cooperativity constant γ as described in the Methods.

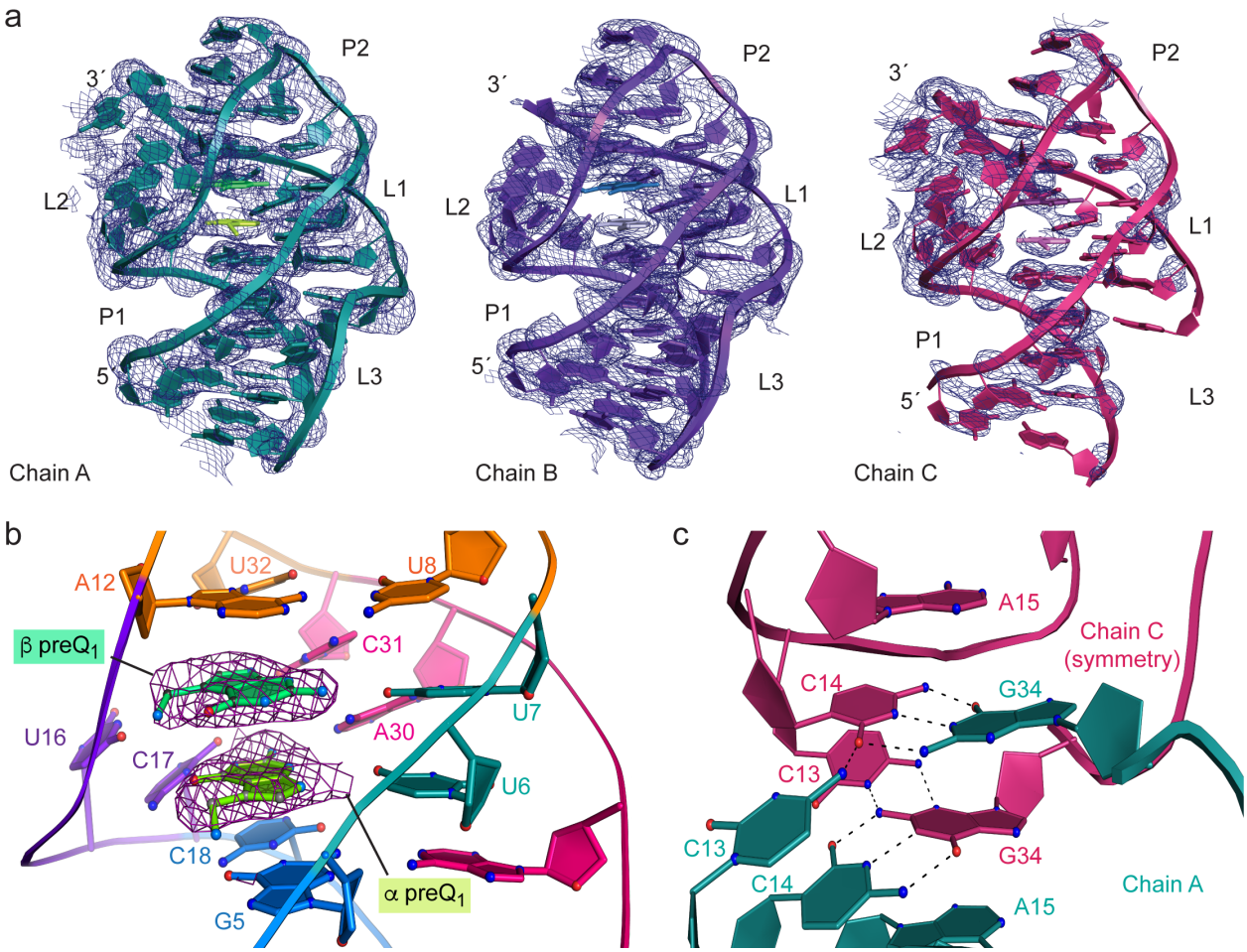
Supplementary Table 4: EC_{50} and fold change in $preQ_1$-induced reporter-gene repression

Riboswitch Sequence	$EC_{50,1}$ (nM)	$EC_{50,2}$ (nM)	EC_{50} Fold Change	Fold Repression
C. antarticus WT[a]	96 ± 14	7100 ± 360	N/A	15.4 ± 1.5
L. rhamnosus WT[b]	15 ± 0.1	N/A	N/A	14.9 ± 1.2
C. antarticus C17U[b]	4.3×10^5 ± 1	N/A	60 ± 3	5.9 ± 0.7
C. antarticus C31U[b]	1.5×10^6 ± 1	N/A	210 ± 10	1.9 ± 0.1

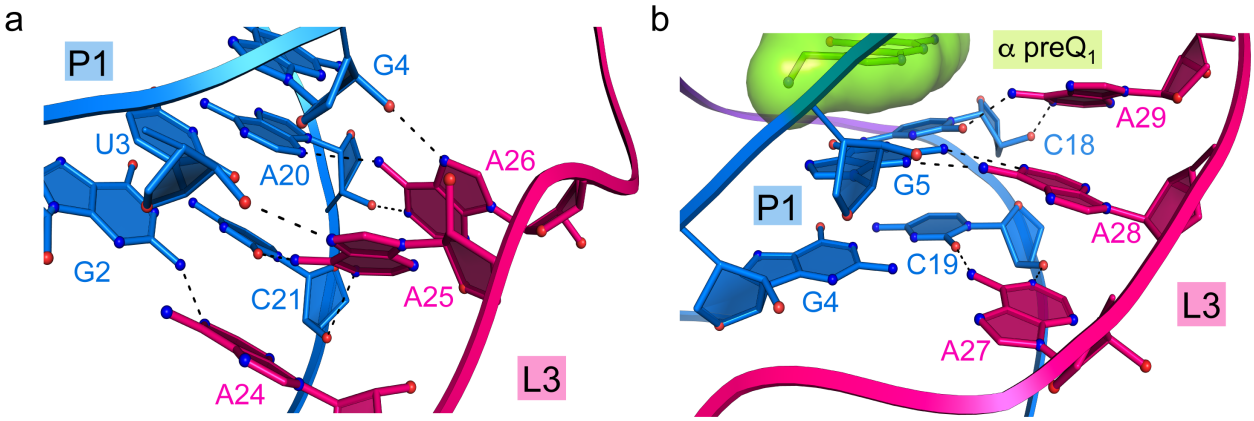
[a] Fit with a biphasic model (see **Methods**).
[b] Fit with log(inhibitor) vs response (three parameters) (see **Methods**).



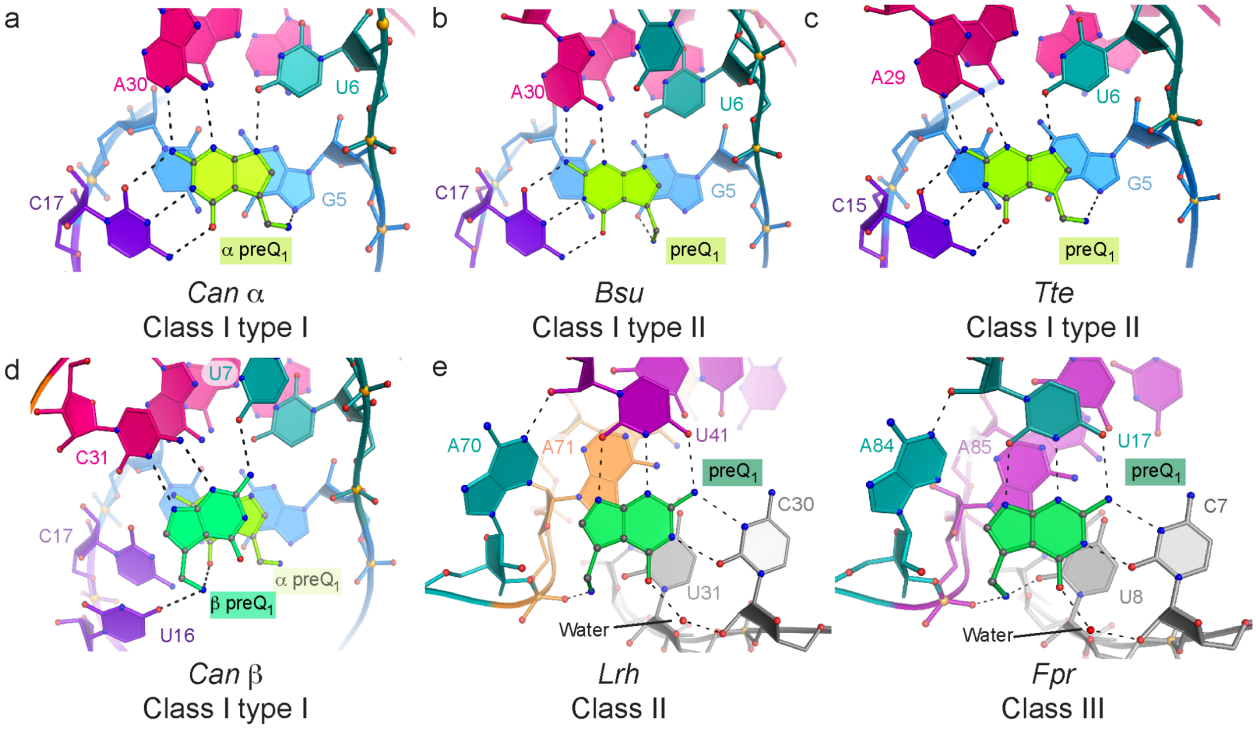
Supplementary Figure 1 | Covariation model and multisequence alignments of preQ$_1$ class I riboswitches. (**a**) Type I Covariation models generated from the full group of known sequence representatives (adapted from ref. 2); red, black and gray positions indicate 97%, 90% and 75% sequence conservation. Multisequence alignments were generated using a handful of representatives derived from phylogenetically diverse bacteria (reported by McCown *et al.*[2]). Positions in bold within the alignment recognize preQ$_1$ based on the *C. antarcticus* co-crystal structure of this investigation. PreQ$_1$ binding nucleobases at the α and β sites are each denoted in the covariation model and the sequence alignment as α or β. Here and elsewhere, bolded organisms have been structurally characterized (this work). In addition to the greatest number of representative sequences (indicated in italics), preQ$_1$-I$_i$ riboswitches exhibit the greatest taxonomic diversity[2]. (**b**) same as (**a**), but with type II sequences. Characterized sequences are *T. tencongensis*[3, 4] and *B. subtilis*[5]. Asterisks denote conserved preQ$_1$ recognition positions. (**c**) same as (**a**) and (**b**) but with type III sequences. Due to a lack of structural characterization, the canonical specificity base is the only predicted preQ$_1$ recognition position[2]. Alignments were created in JALVIEW[6].



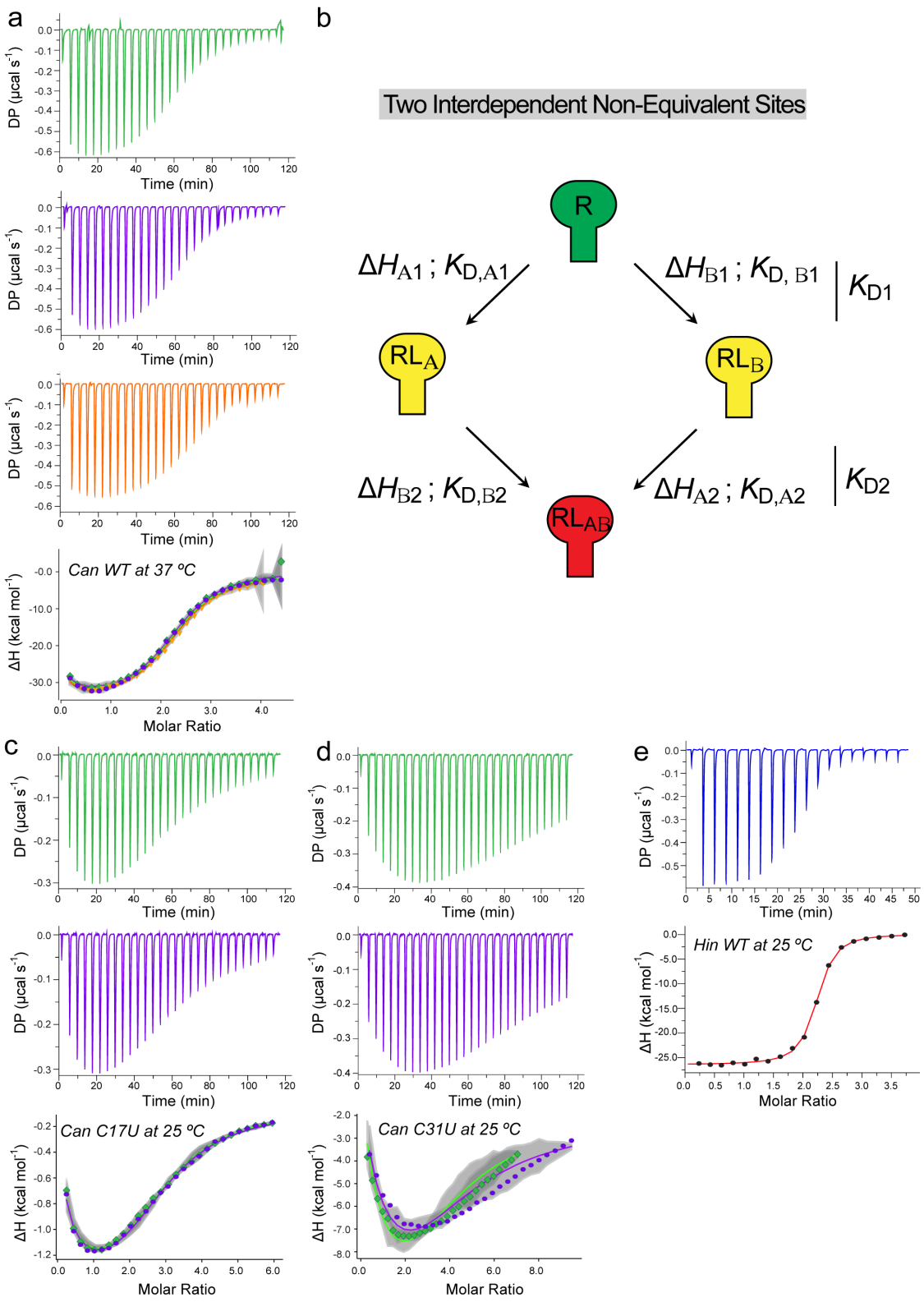
Supplementary Figure 2 | Structural quality and details of the *C. antarcticus* preQ$_1$-I type I riboswitch aptamer and expression platform. (**a**) Reduced bias 2mF_o-DF_c electron-density maps contoured at 1.2 σ around each riboswitch chain in the asymmetric unit. Chains A and B show electron density bathing the entire model, while chain C reveals a break at the junction between P1 and the L3 loop. Individual nucleotides are shown as a cartoon diagram for simplicity. Here and elsewhere, preQ$_1$ is depicted as a ball-and-stick model (green shades in chain A). (**b**) Feature-enhanced, composite-omit map showing the quality of preQ$_1$ models fit to unbiased electron density[7]. (**c**) The expression platforms of chains A and C form a crystal contact that likely takes the place of an intramolecular WC pair between C10 and G34.

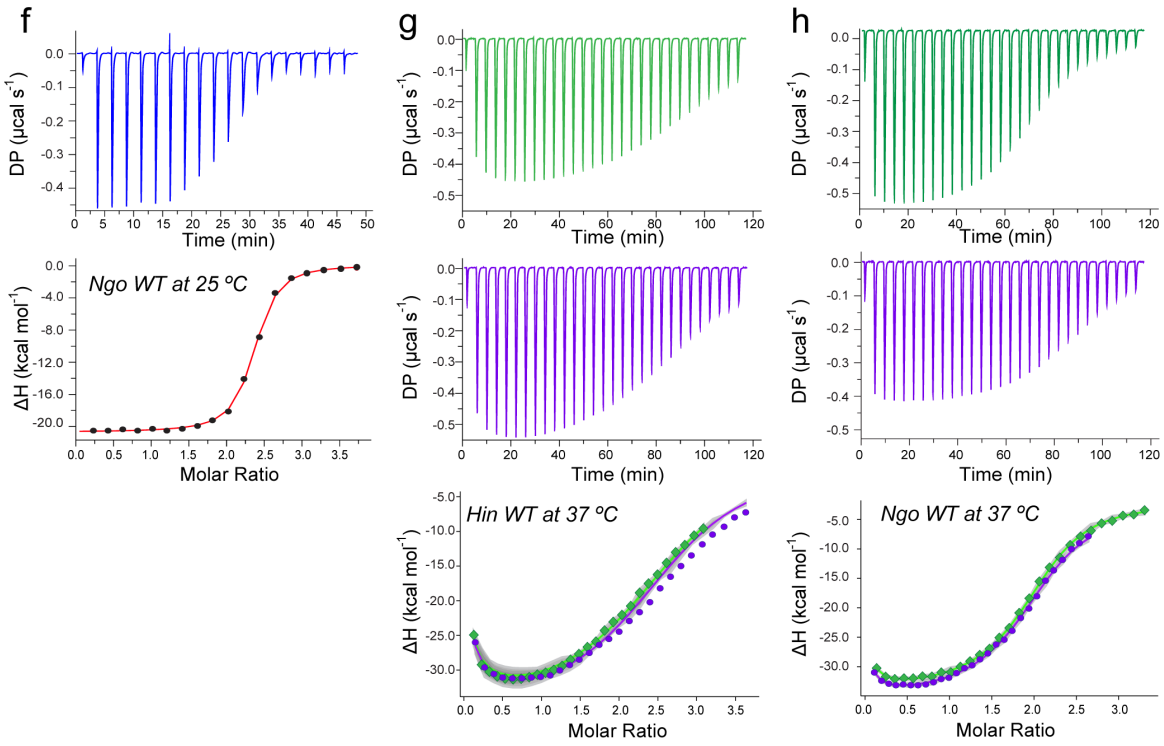


Supplementary Figure 3 | **A-amino kissing interactions between Loop 3 and the minor groove of P1.** (**a**) View of interactions near the base of P1 where the stem transitions to loop L3. Adenines 24, 25 and 26 engage in sugar edge interactions to P1 via their WC faces and Hoogsteen edges. (**b**) View of the floor of the α-preQ$_1$ binding site. Adenines 27, 28 and 29 use their WC faces to pair with the sugar edges of P1 nucleotides, supporting the floor of the ligand binding pocket.

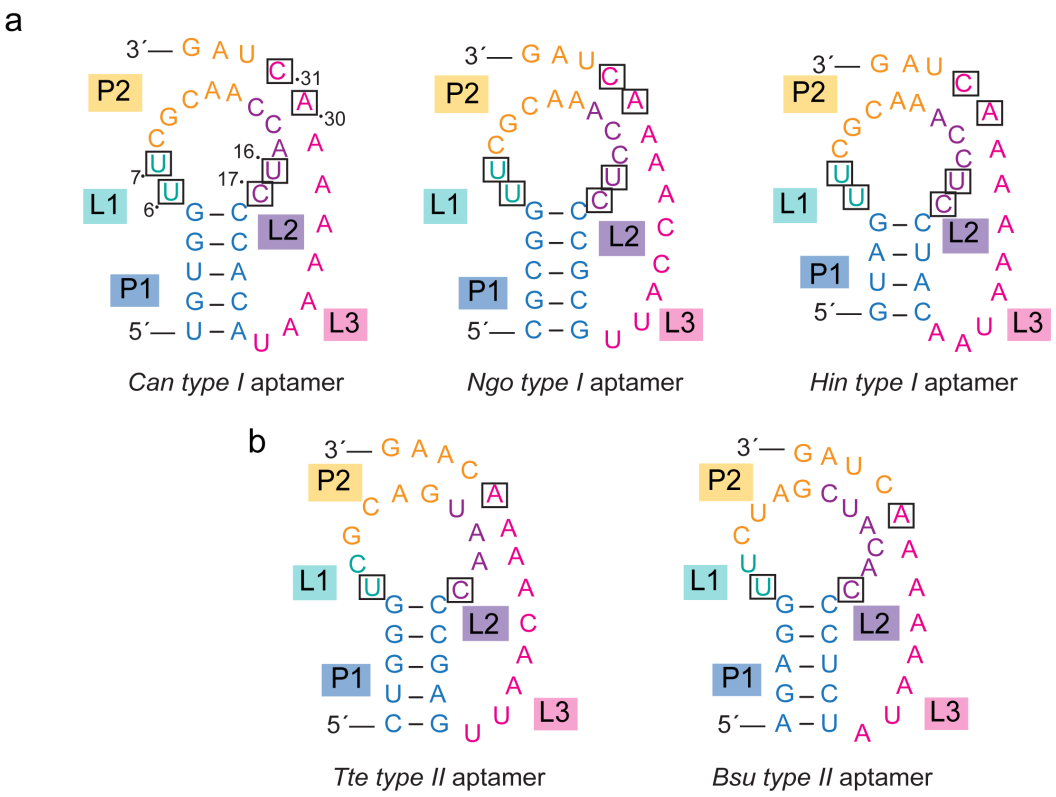


Supplementary Figure 4 | Binding pockets of the *Can* preQ$_1$-I$_I$ riboswitch of this investigation compared to other preQ$_1$ riboswitches. (**a**) The α preQ$_1$ binding site of the *Can* riboswitch. Like known class I structures in (**b**) and (**c**), preQ$_1$ sensing at the α site occurs by canonical *cis* Watson-Crick (WC) pairing. (**b**) Binding pocket of the *B. subtilis* (*Bsu*) class I type II preQ$_1$ riboswitch[5]. Effector readout at this site is equivalent to the *Can* riboswitch α site. (**c**) The binding pocket of the *T. tengcongensis* (*Tte*) class I type II preQ$_1$ riboswitch[4] is equivalent to **b**. (**d**) The β site of the *Can* riboswitch of this investigation displays a new mode of preQ$_1$ recognition of the minor-groove edge equivalent by hydrogen bonds from U7 and C31, whereas the WC face of the ligand is open. (**e**) PreQ$_1$ binding pockets of the class II riboswitch from *L. rhamnosus* (*Lrh*)[8] (*left*) and the class III riboswitch from *F. prausnitzii* (*Fpr*)[9] (*right*). Despite adopting different global folds, both riboswitches use a similar constellation of ten bases for ligand recognition that involves *trans* WC interactions with specificity base C30 or C7. A70 and A84 are inclined A-minor bases that originate from an orthogonal A-form helix that abuts the effector edge[8, 9]. In the preQ$_1$-II riboswitch, these bases are important for gene regulation and dynamics[10-12].

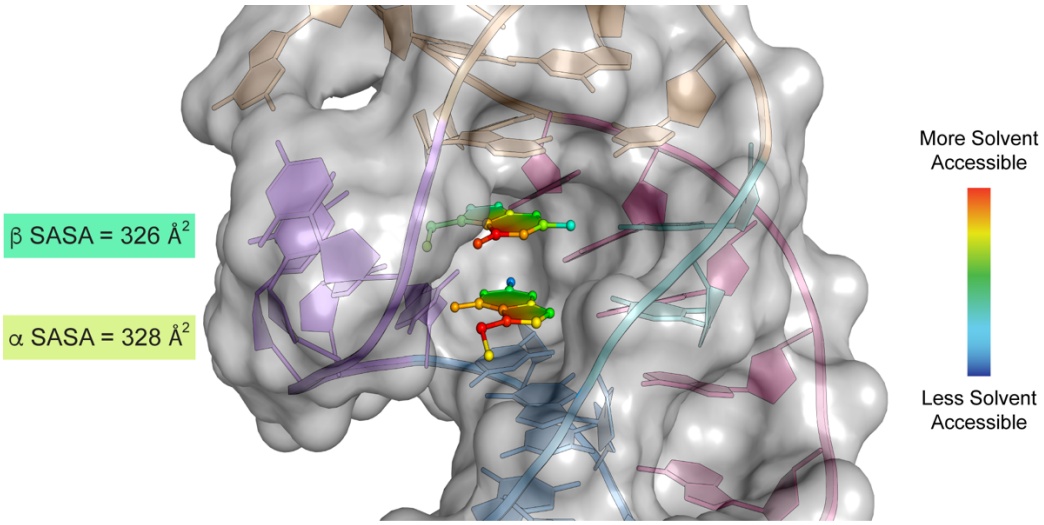

a
DP (µcal s-1)
Time (min)
Can WT at 37 ºC
ΔH (kcal mol-1)
Molar Ratio
b
Two Interdependent Non-Equivalent Sites
R
ΔHA1 ; KD,A1
ΔHB1 ; KD, B1
KD1
RLA
RLB
ΔHB2 ; KD,B2
ΔHA2 ; KD,A2
KD2
RLAB
c
Can C17U at 25 ºC
d
Can C31U at 25 ºC
e
Hin WT at 25 ºC



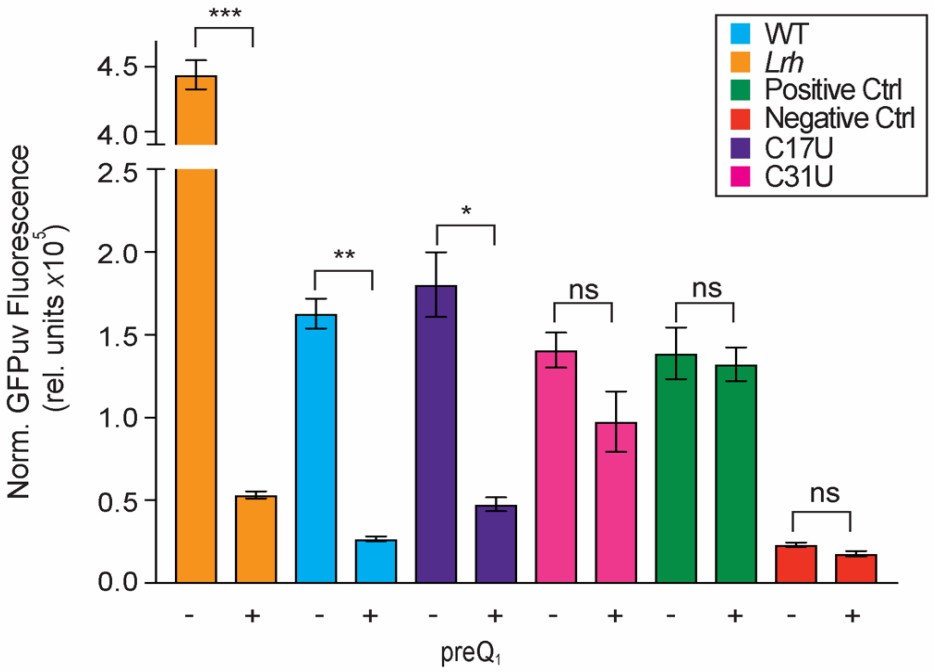
Supplementary Figure 5 | Representative ITC thermograms and corresponding fits: (**a**) Replicate thermograms and global fit of WT *C. antarcticus* (*Can*) riboswitch at 37 °C. Here and elsewhere, the gray areas represent 95% confidence intervals. (**b**) Schematic diagram of the two-interdependent-sites binding model used to fit cooperative isotherms of this investigation (See **Methods**). (**c**-**d**) Replicate thermograms and global fits of *Can* riboswitch mutants C17U (**c**) and C31U (**d**) at 25 °C. (**e**-**f**) Representative thermograms of WT *H. influenzae* (*Hin*) (**e**) and *N. gonorrhoeae* (*Ngo*) (**f**) performed at 25 °C and fit with an independent sites (single set of sites) binding model (Malvern Panalytical, Inc). Thermodynamic parameters are listed in **Supplementary Table 2**. (**g**-**h**) Replicate thermograms and global fits of WT *Hin* (**g**) and *Ngo* (**h**) riboswitches at 37 °C. Cooperative thermograms (**a**, **c**, **d**, **g** & **h**) were analyzed using a non-linear least-squared minimization fitting model developed in our lab (see **Methods**). Thermodynamic parameters for **a**, **c**, **d**, **g** & **h** are in **Supplementary Table 3**. Additional experimental parameters for **e** & **f** are listed in **Supplementary Table 2.**



Supplementary Figure 6 | Secondary Structure Models of Known Class I Riboswitches. (**a**) Type I riboswitch sequences from *C. antarticus* (*Can*), *N. gonorrohoeae* (*Ngo*) and *H. influenzae* (*Hin*). The latter two sequences were identified previously[1]. The secondary structure diagram of the *Can* riboswitch was derived from the co-crystal structure of this investigation. Here and elsewhere, positions that contact preQ$_1$ are boxed. Color-codes correspond to specific pseudoknot base pairing (P) and loop (L) sequences as defined[13]. (**b**) Structurally characterized type II aptamers that bind one preQ$_1$ equivalent are from *T. tengcongensis*[3, 4, 14, 15] (*Tte*) and *B. subtilis* (*Bsu*)[1, 5].



Supplementary Figure 7 | Solvent accessibility of bound effectors. Bound *C. antarcticus* preQ$_1$-I$_I$ riboswitch with the atomic surface shown (semi-transparent gray). The underlying RNA is colored as in **Fig. 1a** with backbone shown as a cartoon for simplicity. Atoms in each preQ$_1$ effector are colored by solvent-accessible surface area (SASA).



Supplementary Figure 8 | Fluorescence emission for GFP*uv* constructs in live bacteria. Normalized fluorescence emission for each riboswitch construct shows fluorescence emission in the absence of $preQ_1$ for all constructs except the negative control (red). Changes in fluorescence from GFP*uv* in the absence and presence of saturating $preQ_1$ (3 μM) resulted in significantly decreased emission when the reporter gene was under control of WT *L. rhamnosus* (*Lrh*, p = 0.0006) or *C. antarcticus* (*Can*, p = 0.0037) riboswitches; partial reduction of GFPuv emission was observed for *Can* mutants C17U (p = 0.017) and C31U ($p > 0.05$). The positive and negative controls were unaffected by $preQ_1$. The mean and standard error of the mean (S.E.M.) are reported (n = 3). Significance was determined by a Student's *t*-test with Welch's correction (*p = 0.05, **p= 0.005, ***p = 0.001).

References

1. Roth, A. et al. A riboswitch selective for the queuosine precursor preQ$_1$ contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* **14**, 308-17 (2007).

2. McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. Structural, functional, and taxonomic diversity of three preQ$_1$ riboswitch classes. *Chem. Biol.* **21**, 880-889 (2014).

3. Jenkins, J. L., Krucinska, J., McCarty, R. M., Bandarian, V. & Wedekind, J. E., Comparison of a preQ$_1$ riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J. Biol. Chem* **286**, 24626-37 (2011).

4. Schroeder, G. M. et al. Analysis of a preQ$_1$-I riboswitch in effector-free and bound states reveals a metabolite-programmed nucleobase-stacking spine that controls gene regulation. *Nucleic Acids Res.* **48**, 8146-8164 (2020).

5. Klein, D. J., Edwards, T. E. & Ferre-D'Amare, A. R., Cocrystal structure of a class I preQ$_1$ riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat. Struc. Mol. Biol.* **16,** 343-344 (2009).

6. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J., Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 1189-1191 (2009).

7. Afonine, P. V. et al. FEM: feature-enhanced map. *Acta Crystallogr. D* **71**, 646-666 (2015).

8. Liberman, J. A., Salim, M., Krucinska, J. & Wedekind, J. E. Structure of a class II preQ$_1$ riboswitch reveals ligand recognition by a new fold. *Nat. Chem. Biol.* **9**, 353-355 (2013).

9. Liberman, J. A. et al. Structural analysis of a class III preQ$_1$ riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3485-94 (2015).

10. Dutta, D. & Wedekind, J. E., Nucleobase mutants of a bacterial preQ$_1$-II riboswitch that uncouple metabolite sensing from gene regulation. *J. Biol. Chem.* **295**, 2555-2567 (2020).

11. Soulière, M. F. et al. Tuning a riboswitch response through structural extension of a pseudoknot. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E3256–E3264 (2013).

12. Kang, M., Eichhorn, C. D. & Feigon, J. Structural determinants for ligand capture by a class II preQ$_1$ riboswitch. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E663-E671 (2014).

13. Peselis, A. & Serganov, A. Structure and function of pseudoknots involved in gene expression control. *Wiley Interdiscip. Rev. RNA* **5**, 803-822 (2014).

14. Spitale, R. C., Torelli, A. T., Krucinska, J., Bandarian, V. & Wedekind, J. E. The structural basis for recognition of the preQ$_0$ metabolite by an unusually small riboswitch aptamer domain. *J Biol. Chem.* **284**, 11012-11016 (2009).

15. Kang, M.; Peterson, R. & Feigon, J. Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol. Cell* **33**, 784-90 (2009).

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

In this revised manuscript, the authors have provided a detailed set of point-by-point responses with associated changes to the manuscript, where appropriate. In light of the comments raised by the reviews, the authors have reconsidered their analysis and interpretation of some of their data, such as the cell-based reporter assays, and provided additional information such as the phylogenetic alignment to provide further information about this novel subclass of preQ1 riboswitches. None of these changes significantly altered the original findings or conclusions, but rather serve to clarify key observations and conclusions. These revisions, in the opinion of this reviewer, have fully addressed all of the concerns raised and the current form of the manuscript is suitable for publication.

Reviewer #2 (Remarks to the Author):

All my concerns were addressed appropriately. The very colorful, as in full of colors, response addresses well all previous points raised.

Reviewer #3 (Remarks to the Author):

Thank you to the authors for responding to our comments. The authors have addressed our concerns and I think revised manuscript is suitable for publication.