



ARTICLE

<https://doi.org/10.1038/s41467-019-09493-3>

OPEN

Synthetic ligands for PreQ₁ riboswitches provide structural and mechanistic insights into targeting RNA tertiary structure

Colleen M. Connelly¹, Tomoyuki Numata^{2,3}, Robert E. Boer¹, Michelle H. Moon¹, Ranu S. Sinniah¹, Joseph J. Barchi¹, Adrian R. Ferré-D'Amaré ² & John S. Schneekloth Jr. ¹

Riboswitches are naturally occurring RNA aptamers that regulate gene expression by binding to specific small molecules. Riboswitches control the expression of essential bacterial genes and are important models for RNA-small molecule recognition. Here, we report the discovery of a class of synthetic small molecules that bind to PreQ₁ riboswitch aptamers. These molecules bind specifically and reversibly to the aptamers with high affinity and induce a conformational change. Furthermore, the ligands modulate riboswitch activity through transcriptional termination despite no obvious chemical similarity to the cognate ligand. X-ray crystallographic studies reveal that the ligands share a binding site with the cognate ligand but make different contacts. Finally, alteration of the chemical structure of the ligand causes changes in the mode of RNA binding and affects regulatory function. Thus, target- and structure-based approaches can be used to identify and understand the mechanism of synthetic ligands that bind to and regulate complex, folded RNAs.

¹Chemical Biology Laboratory, National Cancer Institute, Frederick, MD 21701, USA. ²Biochemistry and Biophysics Center, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA. ³Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8566, Japan. These authors contributed equally: Colleen M. Connelly, Tomoyuki Numata. Correspondence and requests for materials should be addressed to J.S.S.Jr. (email: schneeklothjs@mail.nih.gov)

An explosion of interest in RNA biology in recent years has revealed a multitude of new regulatory functions for noncoding RNA^{1,2}. Many noncoding RNAs regulate gene expression or protein function and are often dysregulated in cases of infectious disease or cancer^{3,4}. The observation that regulatory, noncoding RNAs can directly influence disease states⁵ has led to the suggestion that such RNAs could be suitable as targets for small molecules^{6–8}. Thus, the goal of developing RNA-binding small molecules as both therapeutics and chemical probes is an area of increasing interest. However, the discovery of molecules that bind to RNA with good affinity and specificity is a major challenge owing to RNA's highly flexible, dynamic structure and largely solvent-exposed binding pockets, in addition to the fact that many RNAs are actively unfolded in the cell⁹. Further, understanding the function of noncoding RNA has proven highly challenging in many cases¹⁰. Although it is now routine to utilize structure to rationalize the function of protein-binding small molecules, high-resolution structures of RNA are often difficult to determine. The challenge in characterizing the specific contacts small molecules make with RNAs is often a substantial impediment to understanding and developing small molecule probes of RNA.

One class of regulatory, noncoding RNAs where major advances in structure determination have been seen is riboswitches. Riboswitches are structured RNA elements that occur in the untranslated regions of mRNA, most often in bacteria¹¹. Riboswitches act as regulators of gene expression through recognition of a small-molecule ligand that induces a conformational change in the RNA, thereby regulating downstream gene expression¹². Riboswitches often recognize ligands related to the function of the associated gene product¹³ and, in some examples, the ligand can be a product of the biosynthetic pathway regulated by the riboswitch. Molecules recognized can range from nucleobases, cofactors, and amino acids to metal ions¹³. These structured RNA elements have two domains, an aptamer domain that recognizes the cognate ligand through specific interactions, and an expression platform that changes conformation upon binding to modulate gene expression¹². Riboswitch-mediated gene regulation can occur at either the transcriptional or translational level. For transcriptional riboswitches, cognate ligand binding usually promotes an RNA conformation that contains a terminator helix that triggers dissociation of RNA polymerase and leads to transcription termination upstream of the coding sequence. In the absence of the ligand, the RNA folds into an alternate antiterminator, therefore, allowing transcription of the downstream gene. In translational riboswitches, ligand binding causes a conformational change that produces a helix that occludes the Shine-Dalgarno sequence and prevents ribosomal binding. In the absence of the ligand, an alternate structure forms, increasing the availability of the ribosomal binding region and facilitating translation initiation. The ability of riboswitches to control essential bacterial genes has stimulated their evaluation as antibacterial targets. Furthermore, riboswitches also serve as important model systems for highly specific RNA-small molecule interactions^{14–22}.

The PreQ₁ riboswitch governs the expression of genes responsible for the biosynthesis of queuosine (Q)²³. The riboswitch regulates downstream gene expression in response to its cognate ligand PreQ₁ (7-aminomethyl-7-deazaguanine). PreQ₁ is a modified, guanine-derived nucleobase that is incorporated into tRNAs at the wobble position, where it is further altered to produce Q²⁴. The role of Q is linked to the preservation of translational fidelity and aids reading of degenerate codons. Further, removal of enzymes involved in Q biosynthesis results in multiple phenotypic defects in a variety of organisms^{25,26}. This pathway is thought to be essential for bacterial virulence in some

cases²⁷. Many organisms contain PreQ₁ riboswitches, and to date, three distinct PreQ₁ riboswitch classes have been discovered that differ in both sequence and mechanism²³. Importantly, class I PreQ₁ (PreQ₁-I) riboswitches, such as that found in *Bacillus subtilis* (*Bs*), are among the smallest known riboswitches, with the *Bs* riboswitch having a minimal aptamer domain of just 34 nucleotides (nt). In *Bs*, the PreQ₁ riboswitch regulates the *queCDEF* operon that is involved in Q biosynthesis. Upon ligand binding, the *Bs*PreQ₁ riboswitch folds into an H-type pseudoknot²⁸ causing a conformational change that induces a terminator hairpin and downregulates transcription. Similarly, the PreQ₁-I riboswitch from *Thermoanaerobacter tengcongensis* (*Tt*) also adopts an H-type pseudoknot structure²⁹ in the ligand-bound state. In contrast, the *Tt* riboswitch regulates expression of a PreQ₁ transporter³⁰ and does so via translational inhibition by a conformational change that sequesters the Shine-Dalgarno sequence. The small and well-characterized aptamer domains of the PreQ₁ riboswitches from *Bs* and *Tt* make them ideal model systems for assessing the ability to target riboswitches with synthetic small molecules to regulate downstream gene expression³¹.

In this study, we report the discovery of a new class of synthetic small molecules that bind directly to PreQ₁ riboswitches. We performed a small molecule microarray (SMM) screen on the aptamer domain of the PreQ₁ riboswitch found in *Bs*. Hit compounds that showed selective binding to the riboswitch over other RNAs were validated through a series of biophysical experiments. One hit compound from this screen exhibits a dissociation constant of ~ 500 nM to the *Bs*PreQ₁ aptamer by multiple orthogonal fluorescence titrations. Further, in-line probing suggests that the synthetic ligand induces a riboswitch conformation that is different from that of the cognate ligand-bound form. The compound is also shown to bind the structurally similar, but functionally distinct *Tt*PreQ₁ riboswitch. Co-crystal structures with this aptamer show that the compound binds in the PreQ₁-binding site but makes key interactions with conserved nucleotides that are different from those of the cognate ligand. Importantly, in vitro transcription termination assays demonstrate that the small molecule is capable of regulating riboswitch function upon binding. Structure-guided alteration of the chemical structure of the ligand impacts both the mode of binding and activity of the ligands. The co-crystal structures will further be used for developing additional compounds that target the PreQ₁ riboswitch and the Q biosynthetic pathway.

Results

SMM screening of a PreQ₁ riboswitch aptamer. To identify drug-like small molecules that bind to the PreQ₁ riboswitch, we used a SMM-screening strategy^{32–40}. In this approach, small molecules are spatially arrayed and covalently linked to a glass surface. Next, a fluorescently tagged RNA of interest is incubated with the arrays. Slides are washed, imaged, and the fluorescence intensity is measured for each location on the array. For each compound (printed in duplicate), a composite Z-score is generated, reflecting the increase in fluorescence upon addition of labeled RNA. In parallel, other RNAs may be counter screened to evaluate selectivity and prioritize hit compounds that arise from the screen.

Toward this end, we designed a 5'-Cy5-labeled RNA consisting of the 34-nt aptamer domain of the PreQ₁ riboswitch from *Bs* (5'-Cy5-*Bs*PreQ₁-RS, Fig. 1A)³⁰. Once purity was confirmed by gel electrophoresis, the RNA was annealed and incubated on SMM slides to perform the screen. In parallel, we also screened analogous 5'-Cy5-labeled constructs consisting of aptamer domains from S-adenosylmethionine (SAM-II)⁴¹ and thiamine pyrophosphate (TPP)⁴² riboswitches (Supplementary Table 1).

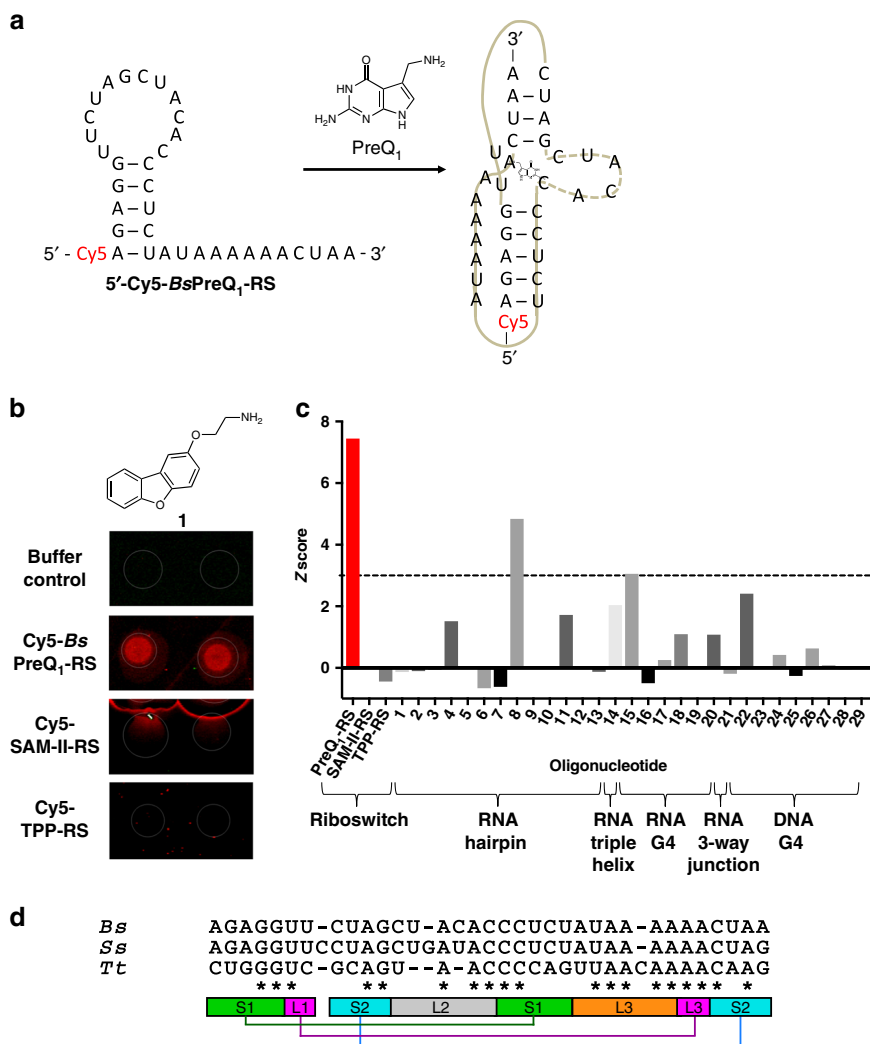


Fig. 1 Small molecule microarray screening. **a** Sequence of the Cy5-labeled aptamer domain of the BsPreQ₁ riboswitch (5'-Cy5-BsPreQ₁-RS) used for SMM screening. The conformational change induced by PreQ₁ binding is also shown. **b** SMM images for the hit compound **1** (compound is printed in duplicate). Spot diameter (white circle) is 150 μm. **c** Selectivity of the hit compound across various oligonucleotides screened using the SMM platform (Supplementary Table 2) as measured by composite Z-score. Horizontal line corresponds to a “hit” cutoff of 3. **d** Sequence alignment of *Bs*, *Ss*, and *Tt*PreQ₁ riboswitch aptamers. Conserved nucleotides are marked by asterisks. Secondary structure of the class I PreQ₁ riboswitch aptamers is below the alignment and base-pairing interactions are indicated by colored bars

Each of these three aptamer domains are known to recognize different biologically active ligands and have well defined, complex three-dimensional structures, making them suitable controls for selectivity. To further measure selectivity, the Z-score for each compound is compared across many different SMM screens (in this case 29 different oligonucleotides, Supplementary Table 2). We identified 243 hits from a collection of 26,227 compounds screened in the SMM, for an initial hit rate of 0.93%. After ruling out compounds that bound promiscuously to other riboswitches or other RNAs and DNAs previously screened against the library, we generated a list of 86 candidate compounds for further study (Supplementary Table 3). The TPP and SAM-II riboswitches had slightly fewer numbers of “selective” hits, with rates of 51 hits (0.19%) and 61 hits (0.23%), respectively. Twenty hits identified as selective for the PreQ₁ aptamer were purchased for further analysis (Supplementary Table 4 and Supplementary Fig. 1). As a representative example, direct binding on SMM slides (Fig. 1B) and selectivity data (Fig. 1C) are shown for compound **1**.

Ligand observed NMR validates binding. Each of the purchased compounds was evaluated using Water-Ligand Observed via Gradient Spectroscopy (WaterLOGSY) NMR⁴³. In this experiment, each compound was subjected to a standard ¹H NMR pulse sequence (to assess solubility in aqueous buffer), WaterLOGSY without RNA (to assess aggregation), and WaterLOGSY in the presence of unlabeled BsPreQ₁-RS RNA (to assess binding to the riboswitch aptamer) (Fig. 2A). For compounds that bind directly to the RNA, peaks phase positively only in the presence of RNA. In contrast, aggregating compounds phase positively even in the absence of RNA. Peaks for compounds that are soluble but do not bind to the RNA are phased negatively in both cases (as can be seen with *N*-methyl-L-valine, included here as a convenient internal, non-binding control). From these experiments, we identified five compounds with suitable solubility and RNA binding in the solution phase. We also evaluated binding of each compound against 31 different structured RNAs and DNAs from previous SMM screens, where compound **1** was found to have the best profile of selective binding (Fig. 1C).

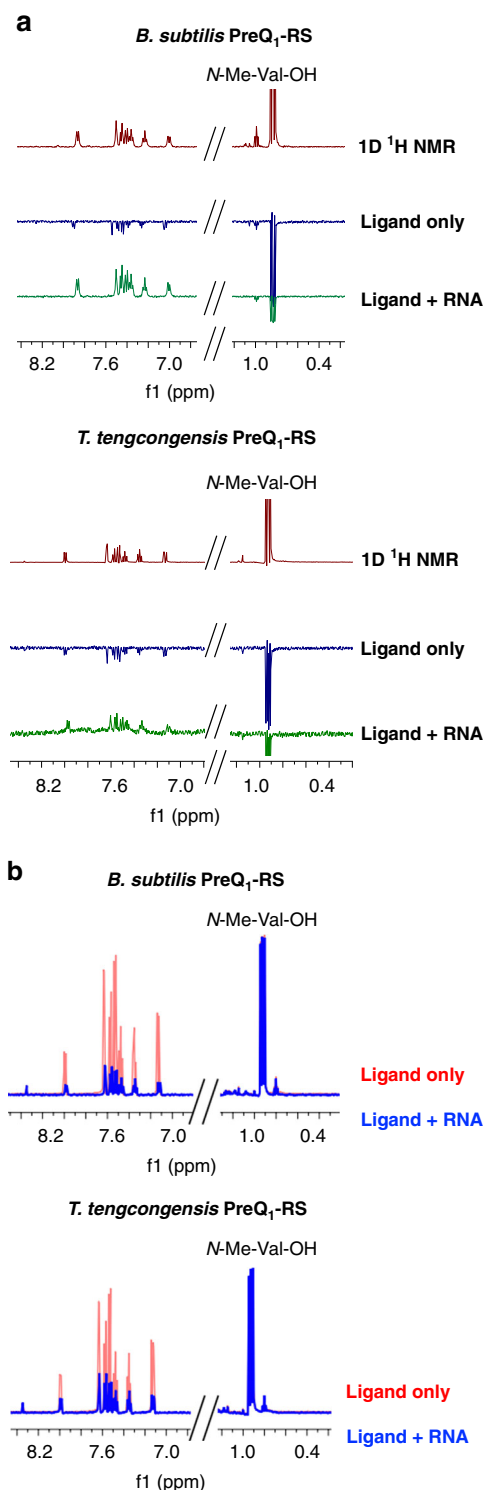


Fig. 2 NMR validation of hit compound **1**. **a** ¹H NMR of **1** and *N*-methyl-L-valine (non-binding control) (Top, red spectrum), WaterLOGSY NMR of **1** and *N*-methyl-L-valine in the absence (middle, blue spectrum) and presence (bottom, green spectrum) of either the *Bs* or *Tt*PreQ₁ riboswitch aptamer. **b** CPMG of **1** and *N*-methyl-L-valine in the absence (red) and presence (blue) of either the *Bs* or *Tt*PreQ₁ riboswitch aptamer

To further validate these results, we also evaluated **1** by WaterLOGSY NMR with a second PreQ₁ riboswitch aptamer domain from *Tt* (unlabeled *Tt*PreQ₁-RS, Supplementary Table 1)²⁹. Although divergent in sequence and genetic control mechanism⁴⁴, this riboswitch also recognizes the same cognate

ligand and has a mostly conserved binding site. Gratifyingly, **1** also bound directly to the *Tt*PreQ₁ aptamer domain by WaterLOGSY (Fig. 2A). In addition to WaterLOGSY experiments, we evaluated the behavior of **1** in Carr-Purcell-Meiboom-Gill (CPMG) ¹H NMR experiments, which can be used to observe alterations in T₂ relaxation times of small molecules that bind large molecules in solution (with binding shortening relaxation time)^{45,46}. Here, **1** was subjected to a CPMG pulse sequence in the absence and presence of both the *Bs* and *Tt* aptamers used in the WaterLOGSY experiments. In both cases, substantial peak attenuation of **1** with no attenuation of the control was observed in the presence of RNA, indicating shortening of the T₂ relaxation time of protons in **1** (Fig. 2B). Thus, CPMG experiments further validate the binding of **1** to both aptamers. Because the two aptamers have conserved binding sites but diverge moderately in structure, this suggested that **1** could potentially share a binding site with the cognate ligand.

Fluorescence titrations demonstrate submicromolar binding.

To estimate the affinity of **1** to the PreQ₁ riboswitch aptamer domain, we used orthogonal fluorescence titrations. First, we measured changes in fluorescence of the Cy5-labeled *Bs* aptamer domain construct used for SMM screening (5'-Cy5-*Bs*PreQ₁-RS) as a function of compound concentration. In parallel, we evaluated effects on an AlexaFluor 647-labeled *Tt* aptamer domain construct (5'-AF647-*Tt*PreQ₁-RS, Supplementary Table 1) to both measure affinity and further rule out effects associated with the fluorophore (Fig. 3A). In the case of the *Bs* aptamer, we measured an apparent dissociation constant (K_D) of 534 ± 123 nM for **1**. Similarly, **1** had a K_D of 457 ± 202 nM for the *Tt* aptamer domain. For comparison, we also measured the affinity of PreQ₁ itself in this assay. The affinity of PreQ₁ for the *Bs* aptamer was 4.1 ± 0.6 nM, and for the *Tt* aptamer it was 2.8 ± 0.4 nM (Supplementary Fig. 2). This is in good agreement with literature values measured by other methods^{29,30,44}. Owing to the presence of a conjugated π -electron system, we evaluated **1** for inherent fluorescence and found that **1** was fluorescent, with $\lambda_{ex} = 300$ nm and $\lambda_{em} = 340$ nm. Thus, by holding the concentration of **1** constant and titrating in increasing quantities of unlabeled riboswitch aptamers, we could measure K_D values by observing changes in fluorescence as a function of RNA concentration (Fig. 3B). Here, we measured a K_D of 490 ± 368 nM for the *Bs* aptamer domain and 99 ± 38 nM for the *Tt* aptamer domain, respectively. Thus, there is good agreement for K_D measurements between multiple different techniques. In addition to the selectivity measured by comparing different SMM screens (Fig. 1C), we also used fluorescence titration to evaluate the binding of **1** to tRNA. Upon titrating tRNA into a solution of **1**, only non-specific binding was observed, and fluorescence quenching did not fit to a 1:1 binding model, indicating only weak binding to tRNA at high micromolar levels (Supplementary Fig. 3).

Ligand induces conformational change to riboswitch aptamers.

A key aspect of riboswitch biochemistry is that biological effects are driven by ligand-induced conformational changes in the RNA upon binding. Thus, we aimed to evaluate the effects of **1** on the conformation of the PreQ₁ aptamers by in-line probing experiments³⁰. In-line probing experiments are routinely used to monitor alterations of riboswitch conformation and can report on ligand-induced effects on RNA structure. First, we performed in-line probing using a fluorescently tagged 36-mer *Bs* aptamer domain (5'-AF647-*Bs*PreQ₁-RS, Supplementary Table 1) with PreQ₁. Consistent with previous reports using a radiolabeled 36-mer construct³⁰, we observed an increase in cleavage at C12 and a corresponding decrease in cleavage at U32 (based on 34-

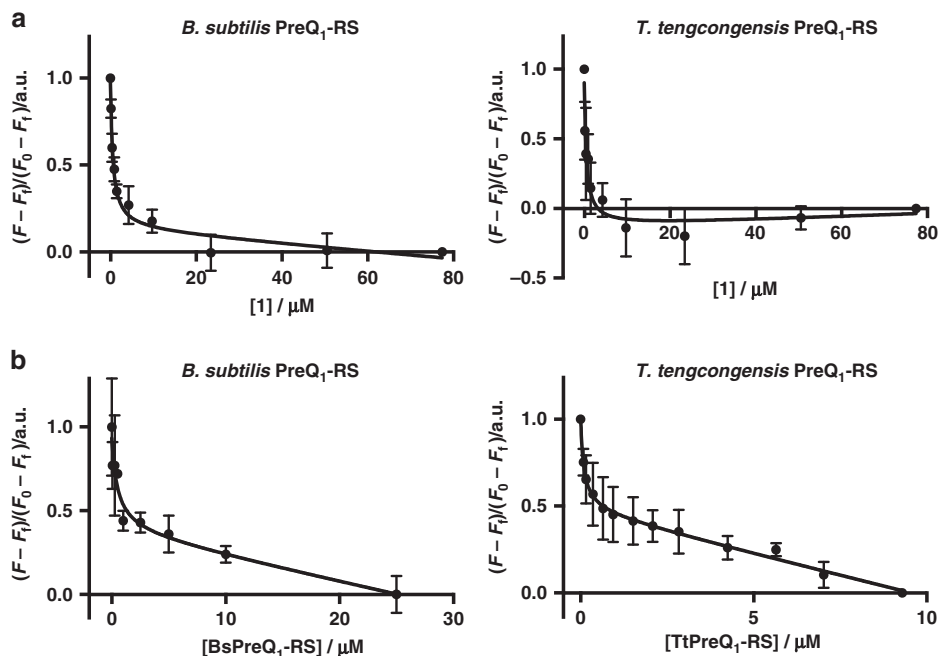


Fig. 3 Affinity of **1** for PreQ₁-RS aptamers. **a** Fluorescence intensity assay of 5'-Cy5-labeled *Bs*PreQ₁-RS or 5'-AlexaFluor 647-labeled *Tt*PreQ₁-RS RNA in the presence of increasing concentration of **1**. **b** Inherent fluorescence titration of **1** with increasing concentration of unlabeled *Bs*PreQ₁-RS or *Tt*PreQ₁-RS RNA. Error bars indicate the standard deviation determined from three independent measurements. Source data are provided as a Source Data file

mer numbering, Fig. 1D). Thus, in-line probing using a fluorescently tagged construct accurately matches literature reports on ligand-induced conformational changes. Next, we performed an analogous experiment in the presence of synthetic ligand **1** (Fig. 4A). In contrast to the cognate ligand, **1** caused an increase in cleavage at C8, U9, U13, C15, U22, and U24, whereas it did not induce prominent increase in cleavage at C12. Although these changes were consistent from experiment to experiment, smaller changes in cleavage at other residues (for example nucleotides 17–19) were less consistent from experiment to experiment and could not be attributed to a specific binding interaction. Thus, although the compound binds directly to the *Bs* aptamer domain and alters the conformation of the RNA, the effects of **1** in this assay are distinct from those of PreQ₁. We also performed in-line probing with the *Tt* riboswitch (5'-AF647-*Tt*PreQ₁-RS, Supplementary Table 1). In this case, with PreQ₁ the most notable increase in cleavage occurred at U12 (Fig. 4B). Upon addition of compound **1**, a similarly small increase in cleavage occurred at U12 and A13. Thus, **1** may induce a conformational change in the *Tt* aptamer similar to PreQ₁. The *Bs* and *Tt* riboswitches have been previously described to have similar conformational ensembles but differ in their ability to recognize ligands by conformational selection and induced fit mechanisms, respectively⁴⁴. This distinction may be reflected in the effects observed in in-line probing experiments.

Ligand modulates riboswitch transcriptional termination.

Having demonstrated that **1** binds and causes conformational changes in both the *Bs* and *Tt* aptamers, we next evaluated **1** in a functional assay. As described above, the *Bs* riboswitch regulates the transcription of downstream genes upon ligand binding. Thus, we examined the capacity of **1** to modulate transcriptional termination. DNA templates were designed containing the aptamer domain and terminator hairpin of the *Bs*PreQ₁ riboswitch, followed by an elongation sequence (*Bs*PreQ₁ TTA, Supplementary Table 5). In vitro transcription of these templates produces

an RNA sequence containing the riboswitch. This transcription can either be halted at the terminator hairpin in the presence of the cognate ligand or produce a read-through product that is transcribed to the end of the template. Despite repeated attempts, efforts to observe transcriptional termination of the *Bs* riboswitch by PreQ₁ were unsuccessful in this assay. *Staphylococcus saprophyticus* (*Ss*) also contains a related PreQ₁ riboswitch sequence with a conserved binding site that functions by transcriptional termination (Fig. 1D). Dose-dependent transcriptional termination was observed with a template containing the *Ss* riboswitch (*Ss* PreQ₁ TTA, Supplementary Table 5). Near-complete termination was observed at the highest PreQ₁ concentrations tested with an EC₅₀ for PreQ₁ of 36 ± 5 nM (Fig. 5). Similarly, addition of increasing concentrations of **1** also resulted in near-complete termination of transcription with an EC₅₀ of 359 ± 23 μ M. Thus, **1** functions by a mechanism similar to PreQ₁ itself. We also confirmed that **1** binds to the *Ss* aptamer by WaterLOGSY experiments (Supplementary Fig. 4). In addition, the K_D of **1** for the *Ss* aptamer was determined to be by 42 ± 6 nM by fluorescence titration (Supplementary Fig. 5). As a further control, the chemically unrelated ligand 5-aminoimidazole-4-carboxamide ribonucleotide, which binds to other riboswitches⁴⁷, was tested in transcriptional assays and had no effect on the *Ss* reporter construct (Supplementary Fig. 7).

X-Ray crystal structure of the ligand/aptamer complex.

To establish a molecular basis for its selective binding, we solved crystal structures of **1** bound to the *Tt*PreQ₁ riboswitch aptamer domain. Initial efforts using the wild-type *Tt*PreQ₁ riboswitch aptamer domain failed, yielding structures indistinguishable from the previously reported ligand-free aptamer domain structure of the RNA, in which the nucleobase of A14 occupies the PreQ₁-binding site²⁹. To alleviate competition by the intramolecular interaction with binding of the exogenous ligand, we designed aptamer domains in which the nucleobase at position 14, as well as one or two adjacent disordered loop nucleobases were

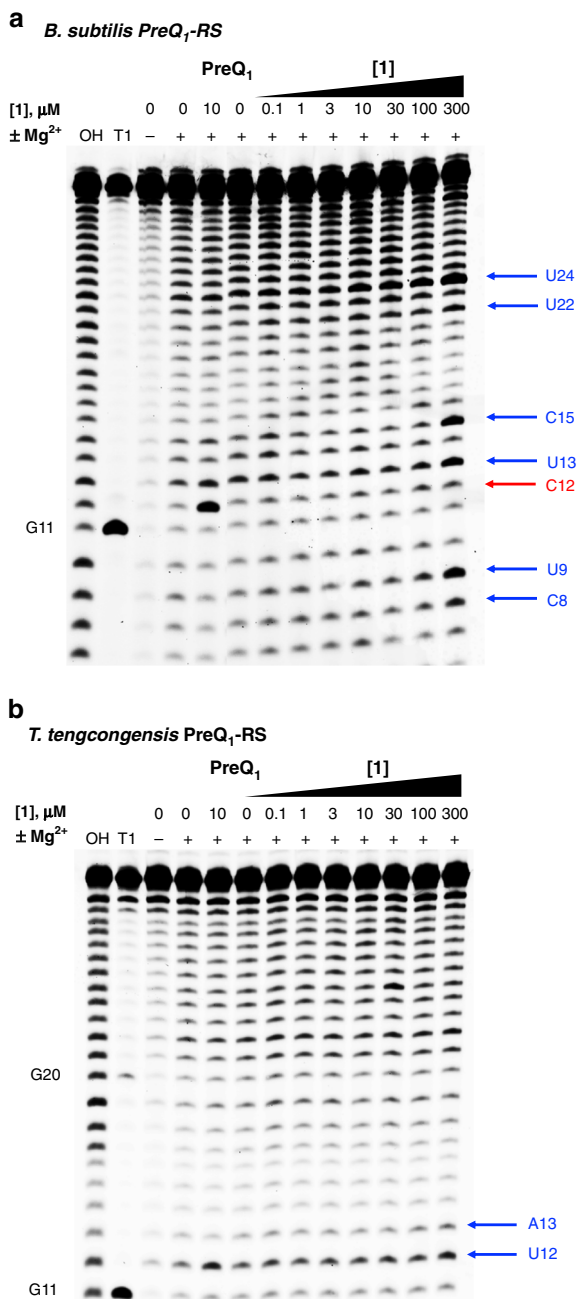


Fig. 4 Ligand-induced conformational changes. In-line probing of **a** 5'-AlexaFluor 647-labeled *Bs*PreQ₁-RS RNA and **b** 5'-AlexaFluor 647-labeled *Tt*PreQ₁-RS RNA after treatment with **1** at increasing concentrations or a DMSO control in the absence (-) or presence (+) of 1 mM MgCl₂. Treatment with PreQ₁ at a concentration of 10 μM is used as a positive control. OH and T1 are a partial alkaline hydrolysis ladder and ribonuclease T1 digestion, respectively. Arrows designate nucleotide positions where the cleavage efficiency was significantly altered by compound treatment (blue) or PreQ₁ treatment (red)

removed. (ab13_14 and ab13_14_15, Supplementary Table 1). Co-crystal structures of these abasic *Tt*PreQ₁ riboswitch aptamer domains complexed with **1** (ab13_14-1, and ab13_14_15-1) were determined by the molecular replacement (MR) method and both refined at 1.8 Å resolution. These two structures are near-identical (r.m.s.d. is 0.14 Å for 638 non-hydrogen atom pairs). The following discussion focuses on the ab13_14-1 complex. We also determined the crystal structure of ab13_14 in complex with

PreQ₁ (ab13_14-PreQ₁) at 1.69 Å resolution, for comparison (Methods and Supplementary Table 7).

Unbiased $|F_o| - |F_c|$ residual electron-density maps unambiguously located **1** at the interhelical interface of stems S1 and S2, surrounded by the L2 and L3 loops, in both co-crystal structures (Fig. 6A, ab13_14-1; Supplementary Fig. 8, ab13_14_15-1). Thus, the synthetic ligand occupies a position similar to that of the cognate ligand PreQ₁. The dibenzofuran of **1** is sandwiched between G11 and the G5•C16 pair, with the furan oxygen engaging in a hydrogen bond (3.2 Å) with the N6 atom of A29 (Fig. 6B and Supplementary Fig. 5). The same oxygen is in van der Waals contact with the N1 atom of A29 (3.7 Å). The nucleotides that interact with **1** are phylogenetically conserved among PreQ₁ riboswitches (Fig. 1D), being responsible for PreQ₁ recognition (Supplementary Fig. 9)²⁸. In contrast to the solvent-inaccessible heterocycle of **1**, its amine-bearing sidechain (largely coplanar with the dibenzofuran heterocycle) is solvent exposed, with the amine hydrogen bonding to water, and therefore not directly recognized by the RNA (Fig. 6B). However, to avoid steric clash with the L2 loop, **1** must bind with its sidechain emerging from the PreQ₁-binding pocket away from L2 loop of the riboswitch. Of note, lack of direct recognition of the amine of **1** provides a plausible explanation for how the binding event occurred when **1** was chemically conjugated to the SMM slide. Although the sugar of residue 13 of the ab13_14-1 structure lacked electron density and is presumed disordered, in the ab13_14_15-1 structure, the ribose ring of residue 13 is ordered and in van der Waals contact with ring C of the dibenzofuran of **1** (Supplementary Figure 8).

Structure-guided changes alter mode of binding and activity.

Next, we asked whether alteration of the chemical structure of **1** would have effects on riboswitch recognition or activity. We hypothesized that alteration of the basicity of the pendant amine of **1** might enable it to engage in hydrogen-bonding contacts with the RNA. Thus, we prepared the dimethylamino derivative **2**, containing a more basic tertiary amine (Fig. 7A). We evaluated the affinity of **2** for both the *Bs* and *Tt* aptamers by monitoring changes in fluorescence of the 5'-Cy5-*Bs*PreQ₁-RS and 5'-AF647-*Tt*PreQ₁-RS RNAs (Supplementary Table 1) upon titration with **2** and determined affinities of 0.4 ± 0.1 μM and 0.6 ± 0.2 μM, respectively (Supplementary Fig. 10). In addition to the pendant amine, we asked whether alteration of the core heterocyclic scaffold would have effects on RNA binding. We noted that in the structure of **1** bound to the riboswitch, the exocyclic amine of A29 makes a hydrogen bond with the oxygen atom of the dibenzofuran. In this structure, A29 is not coplanar with **1** and sits at an angle of 40°, where N1 and N6 bisect the plane of the dibenzofuran core. Nucleotides are amphiphilic and are capable of serving as both hydrogen bond donors and acceptors. We therefore synthesized a carbazole derivative of **2**, compound **3** (Fig. 7A), replacing the dibenzofuran oxygen with a nitrogen atom (that may potentially donate a hydrogen bond). Affinities of **3** for the *Bs* and *Tt* riboswitch aptamer domains were measured to be 0.1 ± 0.08 μM and 0.1 ± 0.04 μM, respectively, using fluorescence titrations (Supplementary Fig. 10). Thus, all three compounds have similar affinities for the riboswitch aptamers. Finally, **2** and **3** were evaluated in transcriptional termination assays (Fig. 7B, Supplementary Fig. 11). Here, **2** behaved similarly to **1** in terms of maximal effect. However, **3** was markedly inferior in the functional assay.

To further understand the interactions of **2** and **3** with the riboswitches, we also solved crystal structures of each ligand in complex with the *Tt* riboswitch. The co-crystal structures of abasic mutant riboswitch aptamer domains bound to **2** and **3**

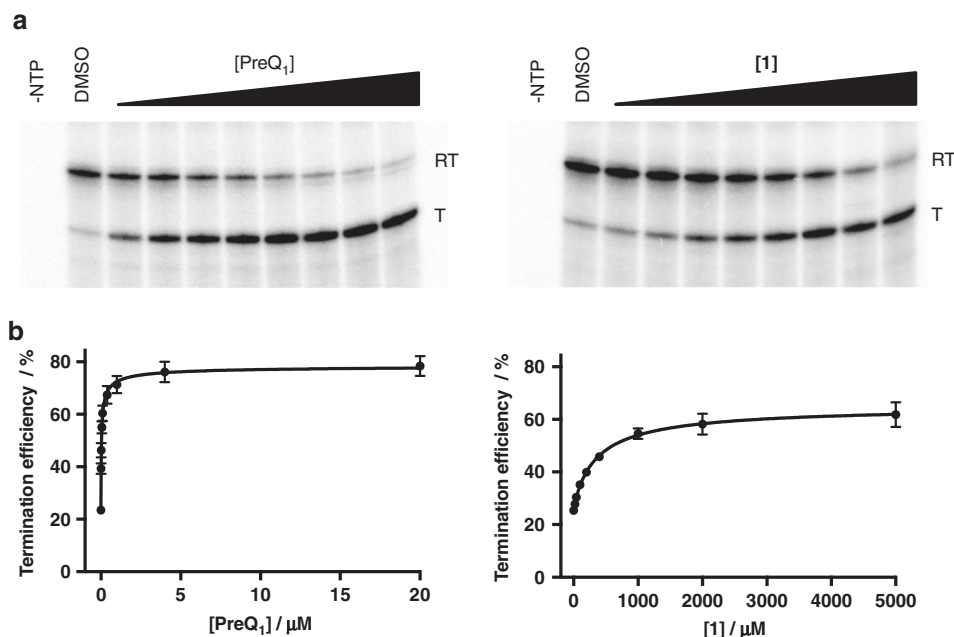


Fig. 5 Transcription termination assay of **1**. **a** Representative gel images of the ³²P-labeled RNA products of in vitro transcription of the *Ss* PreQ₁-RS template in the presence of increasing concentrations of PreQ₁ (left) and **1** (right). Bands corresponding to the read-through transcription product (RT) and terminated transcription product (T) are indicated. Full gel images are provided in the Supplementary Information. **b** Quantification of transcription termination efficiency of increasing concentrations of PreQ₁ (left) or **1** (right). Error bars indicate standard deviation, $n = 3$. Source data are provided as a Source Data file

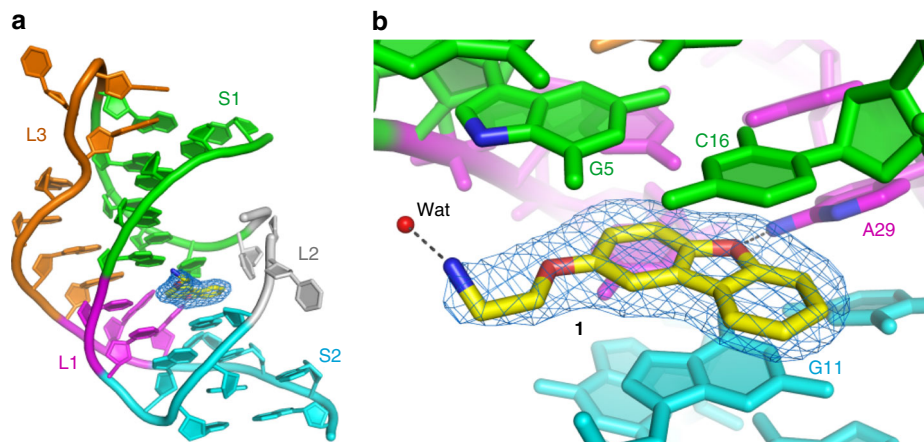


Fig. 6 Co-crystal structure of the aptamer complexed with **1**. **a** Overall structure of the complex. S1, S2, L1, L2, and L3 are colored green, cyan, magenta, gray, and orange, respectively. The nucleotides between L3 and S2, which interact with L1, are also in magenta. Unbiased $|F_o| - |F_c|$ electron density map for the compound **1** (blue mesh) contoured at 3.0 σ . **b** Detail of the ligand-binding site. Nucleotides that interact with **1** are labeled. Dotted lines denote hydrogen bonds between **1** and a bound water molecule and with A29

(ab13_14-2 and ab13_14_15-3, respectively) were solved by the MR method and refined at 1.94 Å and 2.56 Å resolution, respectively (Methods and Supplementary Table 7). The overall structure of ab13_14-2 is similar to that of the riboswitch bound to **1**, with **2** in a similar binding pose, directly in the PreQ₁ binding site (Fig. 8A). In this structure, the sidechain of **2** bends upwards at the methylene preceding the amine, bringing the pyramidal tertiary amine of **2** within hydrogen-bonding distance (2.9 Å) of the N7 atom of the phylogenetically conserved G5 of the riboswitch (Fig. 1D). In contrast, the primary amine of **1** lies 3.6 Å from the same RNA atom. Next, we solved the crystal structure of **3** in complex with the *Tt* riboswitch aptamer domain. Similar to **1** and **2**, compound **3** occupies the PreQ₁-binding site

(Fig. 8B). However, in this structure, **3** has shifted by ~1 Å in the direction of RNA loop L2 (Fig. 8C) to support a hydrogen bond between the new donor atom in the heterocycle and the RNA (the distances between the carbazole nitrogen atom of **3** and the N1 and N6 atoms of A29 are 3.2 Å and 3.5 Å, respectively). This confirms that alteration of the chemical structure of the initial hit can lead to changes in binding modes and somewhat altered affinity. Possibly owing to the ~1 Å shift in its binding pose (Fig. 8C), the conformation of the sidechain of the bound **3** resembles that of the bound **1** rather than that of **2**. The sidechain of the bound **3** is largely coplanar with the carbazole, exposed to solvent, and does not lie within hydrogen-bonding distance of G5 of the RNA.

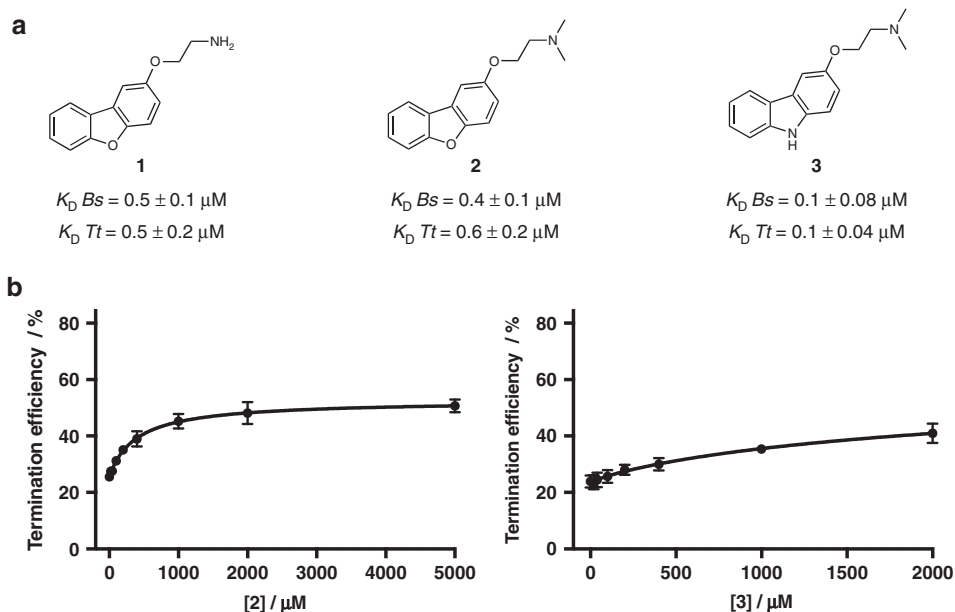


Fig. 7 Evaluation of modified analogs of **2** and **3**. **a** Apparent K_D values for **1-3** as determined by fluorescence intensity assays with 5'-Cy5-BsPreQ₁-RS and 5'-AlexaFluor 647-TtPreQ₁-RS RNA. Values are the mean \pm s.d. (error bars) of three replicate measurements. **b** Quantification of termination efficiency of increasing concentrations of **2** (left) and **3** (right) in transcription termination assays of a DNA template containing the *Ss* PreQ₁ riboswitch. Error bars indicate the standard deviation determined from three independent measurements. Source data are provided as a Source Data file. Gel images are provided in the Supplementary Information

Discussion

In this work, we report the discovery of synthetic ligands that bind to different PreQ₁ riboswitch aptamer domains. Compounds discussed here are characterized by multiple orthogonal biophysical techniques, and exhibit submicromolar dissociation constants to the RNA. Furthermore, structural probing experiments indicate clear ligand-induced conformational changes. As a consequence of binding, the compounds exhibit activity in transcriptional termination assays, inducing premature termination analogous to those produced by the cognate ligand itself. Thus, the compounds function by a mechanism similar to PreQ₁, despite having no obvious chemical similarity. Structural characterization of the ligand-aptamer complexes by X-ray crystallography provides insights into both the mode of recognition and mechanism of action of these compounds.

Structural comparison between ab13_14-1 and the PreQ₁-bound wild-type TtPreQ₁ riboswitch (referred to as WT-PreQ₁) reveals conformational differences in four nucleotides (positions 12–15) of the L2 loop and two nucleotides at the 3'-end (A32 and G33) of these complexes. In the WT-PreQ₁ structure, C15 recognizes PreQ₁ through Watson-Crick base pairing and serves as a starting point of the consecutive base stack, C15-A14-A13-A32-G33. The last two nucleotides (A32 and G33) of the base stack constitute the Shine-Dalgarno sequence and form stem S2 by base-pairing with A10 and C9, respectively. Therefore, the continuous base stack is proposed to be crucial to inhibit gene expression through sequestering the Shine-Dalgarno sequence from the recognition by the ribosome. Because **1** is bulkier than PreQ₁, compound-binding interferes with the correct positioning of C15. Consequently, C15 in the ab13_14-1 structure shifts position and its base is placed perpendicular to that of the WT-PreQ₁ structure to escape the steric clash with **1** (Fig. 8D and G). Moreover, A32 and G33 are located downward (far away from ligand binding site) compared to those of the WT-PreQ₁ structure. Therefore, the ab13_14-1 structure is similar to the free state of the riboswitch, despite the difference that the ligand-binding site was occupied by A14 in the latter structure (Fig. 8E and G).

The striking contrast between the ab13_14-1 and WT-PreQ₁ structures is predominantly attributable to the binding of the bulkier compound to the riboswitch. Consistent with this, the crystal structure of ab13_14-PreQ₁ showed a similar conformation to WT-PreQ₁ rather than ab13_14-1, where C15 in ab13_14-PreQ₁ is well superimposed to that in WT-PreQ₁ and A32 and G33 in ab13_14-PreQ₁ shift toward the ligand-binding site compared with the ab13_14-1 structure (Fig. 8D–G and Supplementary Fig. 5). In addition, U12 of ab13_14-PreQ₁ is placed in a similar location to A14 of WT-PreQ₁ and hydrogen bonds with G11 in order to compensate for the absence of the base at position 14. Therefore, PreQ₁ binding to the ligand binding site results in a conformation similar to WT-PreQ₁ despite the absence of the bases at positions 13 and 14. From these findings, **1** binding to the wild-type TtPreQ₁ riboswitch may induce a sub-optimal formation of the continuous base stack that transmits the ligand binding information to the expression platform of the riboswitch. This suboptimal conformation may lead to the decreased maximum termination efficiency in the transcription termination assay for **1** in comparison with the cognate ligand.

Like the translational class I PreQ₁ riboswitch, transcriptional class I PreQ₁ riboswitches also have a continuous base stack from the ligand-binding site to the expression platform. Given that the transcriptional and translational PreQ₁ riboswitches control ON/OFF switching using similar structural motifs, the **1**-bound TtPreQ₁ riboswitch structure can explain why this compound has an EC₅₀ of $359 \pm 23 \mu M$ in transcription termination assays (*Ss* aptamer). This value is considerably higher than that for PreQ₁ ($36 \pm 5 \text{ nM}$) despite affinities of $0.5 \pm 0.1 \mu M$ and $4 \pm 0.6 \text{ nM}$, respectively (*Bs* aptamer). These findings suggest that the region of **1** situated near C15 would be a good candidate for synthetic efforts to improve the potency for downregulating gene expression. Our co-crystal structures are the starting point for developing additional compounds that target the bacteria-specific Q biosynthetic pathway.

Although **1**, **2**, and **3** have similar binding affinities for the Tt and Bs aptamers, they do not have identical activity in

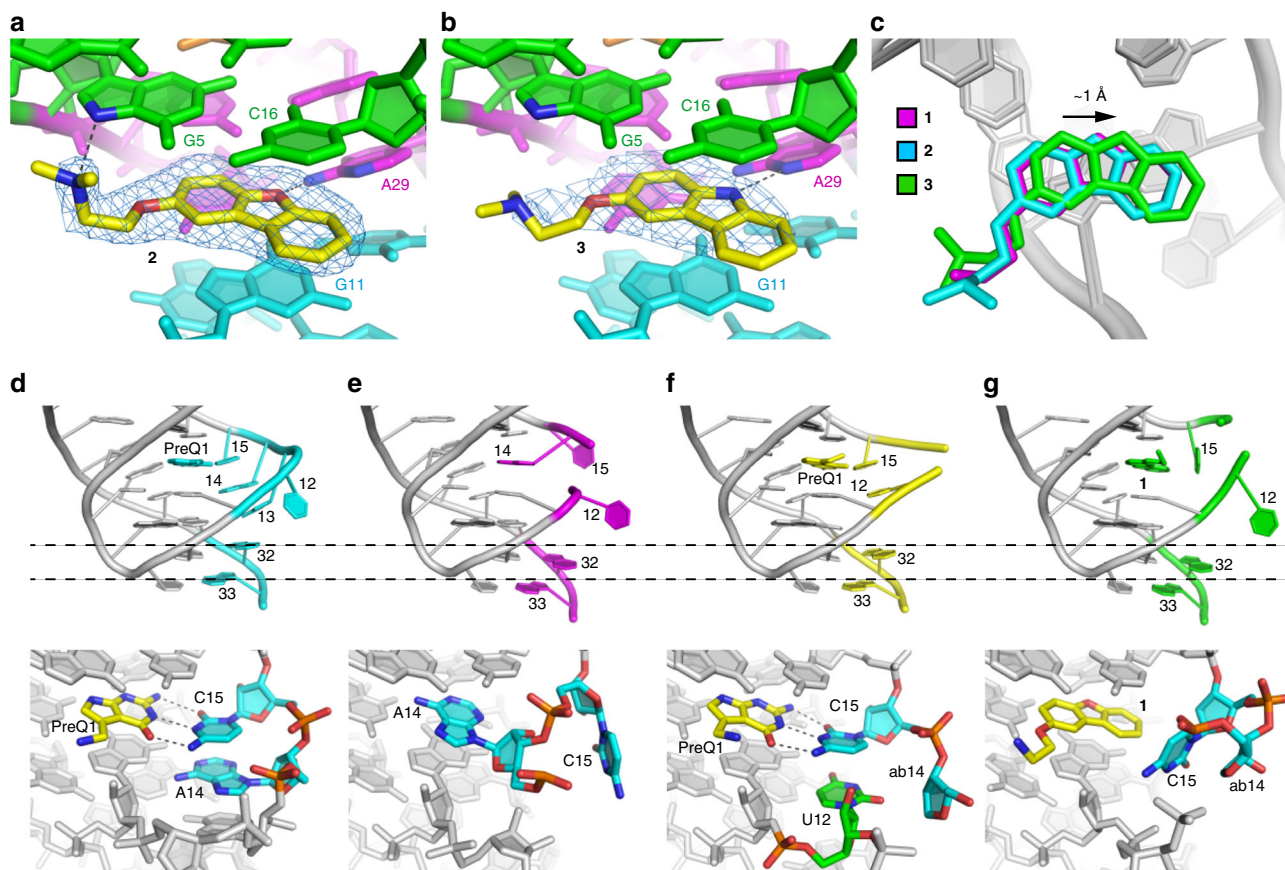


Fig. 8 Structural analysis of the bound *TtPreQ1* aptamers. **a** Ligand-binding site of the ab13_14-2 co-crystal structure, superimposed on the $|F_o| - |F_c|$ electron-density map calculated before addition of the ligand to the crystallographic model (blue mesh, 3.0σ contour). Hydrogen bonds are indicated as dotted lines. **b** Ligand-binding site of the ab13_14_15-3 co-crystal structure. The $|F_o| - |F_c|$ electron density map for the compound is colored blue and contoured at 3.0σ . **c** Comparison of binding modes of **1** (magenta), **2** (cyan), and **3** (green). Ligand-binding site and continuous base stack of the aptamer domains of the **d** WT-PreQ1, **e** WT free, **f** ab13_14-PreQ1, and **g** ab13_14-1 forms. (Upper) Cartoon representations with ligands and key nucleotides labeled and colored cyan, magenta, yellow, and green in the WT-PreQ1, WT free, ab13_14-PreQ1, and ab13_14-1 forms, respectively. To compare the locations of A32 and G33 among these structures, dashed lines are indicated. (Lower) Detail of the ligand-binding site. PreQ1 and **1** are in yellow. Key nucleotides responsible for forming the base stack are colored cyan. U12 in the ab13_14-PreQ1 structure, which occupies a similar location to A14 in the WT-PreQ1 structure, is in green. Hydrogen bonds between PreQ1 and C15 are indicated as dotted lines

transcription termination with the *Ss* riboswitch. This difference in activity may be due to subtle differences in binding modes to the RNA, reflected in the lack of a hydrogen bond to G5 and the altered pose of ligand **3**. Importantly, the synthetic ligand has considerably decreased solubility at concentrations needed for transcription termination, which may also play a role in decreased function. Although it has been previously demonstrated that binding affinity is not the sole parameter that governs the activity of RNA-binding compounds⁴⁸, these observations are often not accompanied by mechanistic or structural rationalization. Our work demonstrates that subtle alterations in the mode of recognition of ligands may also have a role in the ability of a compound to induce conformational or functional effects. This work highlights a challenge in developing ligands for functional RNAs, namely that affinity is not the only property that governs the activity of a compound. Finally, the work described herein illustrates the role structure has in understanding the behavior of small synthetic compounds that bind to and modulate the function of complex RNAs.

Methods

General RNA methods. To avoid RNase contamination, all buffers were prepared with diethyl pyrocarbonate-treated water and all surfaces and equipment were decontaminated with RNaseZap (Ambion) prior to RNA handling. Deprotected

and high-performance liquid chromatography purified oligonucleotides were purchased from Dharmacon (ThermoFisher) or IDT DNA. Sequences of the RNA oligonucleotides used for biochemical and biophysical experiments can be found in Supplementary Table 1.

SMM screening. SMM slides were prepared according to established procedures^{36,38}. In brief, γ -aminopropyl silane microscope slides (Corning) were functionalized with an Fmoc-protected amino polyethylene glycol spacer (Fmoc-8-amino-3,6-dioxaoctanoic acid) in *N,N*-dimethylformamide (DMF). Following piperidine deprotection, 1,6-diisocyanatohexane was coupled to the surface to provide isocyanate-functionalized microarray slides for immobilization of small molecule library members. A total of 26,227 small molecules (10 mM in DMSO) containing at least one primary or secondary alcohol or amine were purchased from commercial vendors including ChemBridge and ChemDiv. The libraries were printed on seven array slides containing ~3840 distinct molecules printed in duplicate, in addition to dyes and controls used for quality-control validation. The arrays were exposed to pyridine vapor to facilitate covalent attachment to the isocyanate-functionalized slides. Slides were then incubated with a 1:20 polyethylene glycol:DMF (v/v) solution to quench unreacted isocyanates. The 5'-Cy5-BsPreQ1-RS RNA was dissolved in 50 mM Tris, 100 mM KCl, 1 mM MgCl₂, pH 7.5, diluted to 5 μ M, and was annealed by heating to 75 °C for 5 min, followed by slowly cooling to room temperature for 30 min. The annealed RNA was then further diluted to 1 μ M in 50 mM Tris, 100 mM KCl, 1 mM MgCl₂, pH 7.5 for screening. Printed microarray slides were incubated with the RNA at a concentration of 1 μ M for 2 h using a LifterSlip (Electron Microscopy Sciences). Following incubation, slides were washed three times with 4 mL of 50 mM Tris, 100 mM KCl, 1 mM MgCl₂, pH 7.5 buffer for 2 min in a four-well Nunc plate, and the slides were dried by centrifugation for 2 min at 4000 \times g. Slides were imaged for fluorescence (650 nm

excitation, 670 nm emission) on an Innopsys Innoscan 1100 AL Microarray Scanner with a resolution of 5 μm . The scanned image was aligned with the corresponding GenePix Array List (GAL) file to identify individual features. Hits were determined based on signal-to-noise ratio (SNR), defined as (mean foreground—mean background)/standard deviation of background, and Z-score, defined as: $Z = (\text{Mean SNR}_{635\text{compound}} - \text{Mean SNR}_{635\text{library}}) / (\text{SD SNR}_{635\text{library}})$ with the following criteria: (a) coefficient of variance of duplicate spots < 100 , (b) Average Z-score for a compound > 3 , (c) $[(Z\text{-score}_{\text{RNA incubated}} - (Z\text{-score}_{\text{Control Array}})) / Z\text{-score}_{\text{Control Array}}] > 3$, (d) no activity with any other nucleic acid structures screened in parallel. Other Cy5-labeled riboswitches screened in parallel included TPP and SAM-II, which were screened using the same method described above for the PreQ₁ riboswitch. Hits were further validated by visual inspection of array images and compounds for further study were purchased from original suppliers.

General ligand observed NMR methods. All NMR spectra were recorded at 293 K on a Bruker AVANCE III 500 MHz spectrometer equipped with a TCI cryoprobe. NMR buffer was composed of 50 mM Tris-*d*₁₁, pH 7.5, 100 mM KCl, 1 mM MgCl₂ in 100% UltraPure distilled H₂O (Invitrogen). Subsequently prepared NMR samples contained 10% D₂O and 5% DMSO-*d*₆ to improve compound solubility. All compounds were first dissolved to a 10 mM stock concentration in 100% DMSO-*d*₆ prior to NMR sample preparation. Purchased RNA was buffer exchanged (3 kDa MWCO spin column, EMD Millipore) into the NMR buffer prior to use. For WaterLOGSY and CPMG experiments, a reference 1D-¹H spectrum was collected for each sample using the “zgesgp” excitation sculpting water suppression pulse sequence from Bruker, with 128 scans. All data was processed and visualized with MestReNova software (Version 8.1.2-11880).

WaterLOGSY NMR. *BsPreQ1*-RS and *TiPreQ1*-RS RNA were buffer exchanged into 50 mM Tris-*d*₁₁, pH 7.5, 100 mM KCl, 1 mM MgCl₂ in UltraPure distilled H₂O using centrifugal filtration (3 kDa MWCO, EMD Millipore) and were annealed by heating to 75 °C for 5 min, followed by slowly cooling to room temperature for 30 min. NMR samples containing each SMM compound at a concentration of 375 μM and a negative control *N*-methyl-L-valine (Chem-Impex-International) at 375 μM were prepared in 50 mM Tris-*d*₁₁, pH 7.5, 100 mM KCl, 1 mM MgCl₂ in UltraPure distilled H₂O containing 10% D₂O and a final concentration of 5% DMSO-*d*₆. For each compound, samples were prepared with and without each RNA at a concentration of 15 μM . Samples were incubated on ice for 10 min, degassed at 20 °C for 15 min, and transferred to a Shigemitsu NMR tube. A reference 1D-¹H and 1D WaterLOGSY spectra with and without RNA were recorded. NMR spectra were recorded at 293 K on a Bruker AVANCE III 500 MHz spectrometer equipped with TCI cryoprobe⁴³.

CPMG NMR. *BsPreQ1*-RS and *TiPreQ1*-RS RNA were buffer exchanged into 50 mM Tris-*d*₁₁, pH 7.5, 100 mM KCl, 1 mM MgCl₂ in UltraPure distilled H₂O using centrifugal filtration (3 kDa MWCO, EMD Millipore) and were annealed by heating to 75 °C for 5 min, followed by slowly cooling to room temperature for 30 min. Samples of **1** (300 μM) with or without each RNA (3 μM) were prepared in 50 mM Tris-*d*₁₁, pH 7.5, 100 mM KCl, 1 mM MgCl₂ in UltraPure distilled H₂O containing *N*-methyl-L-valine (300 μM), 10% D₂O, and 5% DMSO-*d*₆. Samples were incubated on ice for 10 min, degassed at 20 °C for 15 min, and transferred to a Shigemitsu NMR tube. NMR spectra were acquired at 298 K on a Bruker AVANCE III 500 MHz spectrometer fitted with a TCI Cryoprobe with Z-gradient. CPMG data sets with Presaturation Water Suppression (cpmgrp1d) were acquired using a 200 ms T₂ relaxation filter (d₂0=0.001 s, L4=200) with a 10-second relaxation delay. O1 was set to 2349.67 Hz, RG to 203, and data sets were acquired with 64 scans. The spectra of the sample with and without RNA were aligned and scaled so that the peak heights of the internal standard *N*-methyl-L-valine were equivalent. Significantly attenuated ligand peaks ($\geq 20\%$) in the presence of RNA were considered indicative of small molecule binding⁴⁵.

Fluorescence intensity assay. Fluorescence titrations were performed using either a 5'-Cy5-*BsPreQ1*-RS or AF647-*TiPreQ1*-RS. Each RNA was diluted into 50 mM Tris, 100 mM KCl, 1 mM MgCl₂, pH 7.5 to a concentration of 10 μM and was annealed by heating to 75 °C for 5 min, followed by slowly cooling to room temperature for 30 min. Small molecule solutions were prepared as serial dilutions in 10% DMSO in buffer. In a cuvette, RNA was diluted to a final concentration of 50 nM in 50 mM Tris, 100 mM KCl, 1 mM MgCl₂, pH 7.5 buffer containing 10% DMSO. Small molecules were titrated to final concentrations ranging from 0 to 77.4 μM maintaining a 10% final DMSO concentration. After delivery of the small molecule, the sample was allowed to equilibrate for 8 min at room temperature. The fluorescence emission spectrum was then measured at room temperature using a Photon Technology International, Inc. QuantaMaster 600™ Spectrofluorometer equipped with Felix GX 4.2.2 software. Fluorescence was recorded at an excitation wavelength of 649 nm and an emission of 655–800 nm. Total area under the peak from 655–700 nm was quantified and was then normalized to the values obtained for RNA incubated with a DMSO control. Normalized fluorescence for three independent replicates were averaged and plotted against

small molecule concentration. K_D values were determined using a single site-binding model.

Inherent ligand fluorescence titration. Fluorescence titrations were performed using either unlabeled *BsPreQ1*-RS or *TiPreQ1*-RS. Each RNA was resuspended to a concentration of ~500 μM in 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM MgCl₂, and annealed by heating to 75 °C for 5 min, followed by slowly cooling to room temperature for 30 min. Serial dilutions of each RNA were prepared in 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM MgCl₂. Inherent fluorescence titrations were performed with compound **1** at a final concentration of 500 nM in 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM MgCl₂ with 5% DMSO. For *BsPreQ1*-RS, in a black 384-well plate, **1** was diluted to a final concentration of 500 nM in buffer with a 5% final DMSO concentration. RNA was added to final concentrations ranging from 0–50 μM in triplicate, the plate was centrifuged (1000 rpm, 2 min), and samples were allowed to incubate for 30 min at room temperature with shaking. The fluorescence intensity was then measured on a Synergy Mx microplate reader (BioTek) at an excitation wavelength of 300 nm and an emission wavelength of 340 nm (gain = 120). The fluorescence intensity was then normalized to the values obtained for **1** in the absence of RNA and was plotted against RNA concentration. The dissociation constants were determined using a single site binding model.

For *TiPreQ1*-RS, in a cuvette, **1** was diluted to a final concentration of 500 nM in 50 mM Tris, pH 7.5, 100 mM KCl, 1 mM MgCl₂ buffer with a 5% final DMSO concentration. RNA was titrated to final concentrations ranging from 0 to 7 μM and the sample was allowed to equilibrate for 8 min at room temperature after each addition. The fluorescence emission spectrum was then measured at room temperature using a Photon Technology International, Inc. QuantaMaster 600™ Spectrofluorometer equipped with Felix GX 4.2.2 software. Fluorescence was recorded at an excitation wavelength of 300 nm and an emission of 320–460 nm. Total area under the peak from 320–380 nm was quantified and was then normalized to the values obtained for **1** in the absence of RNA. Normalized fluorescence for three independent replicates were averaged and plotted against RNA concentration. K_D values were determined using a single site binding model.

Analysis of Mg²⁺-induced in-line cleavage. This procedure was adapted from a related procedure that used radiolabeled oligonucleotides³⁰. 5'-AF647-*BsPreQ1*-RS or 5'-AF647-*TiPreQ1*-RS RNA was diluted into 50 mM Tris, 100 mM KCl, pH 7.5 to a concentration of 10 μM and were annealed by heating to 75 °C for 5 min, followed by cooling on ice for 10 min. The annealed RNA was incubated at a final concentration of 1 μM in 50 mM Tris, pH 8.3, 10 mM KCl buffer with either a DMSO control (5% final concentration) or compound **1** at a concentration of 0.1, 1, 3, 10, 30, 100, or 300 μM . For 5'-AF647-*BsPreQ1*-RS, MgCl₂ was added at a final concentration of 2 mM, and the reactions were incubated at room temperature in darkness for 40 hrs. For 5'-AF647-*TiPreQ1*-RS, MgCl₂ was added at a final concentration of 2 mM, and the reactions were incubated at room temperature in darkness for 72 h to reach the desired level of cleavage. Alkaline hydrolysis was performed in 50 mM NaHCO₃, pH 9.0 at 95 °C for 5 min. Ribonuclease T1 digestion was carried out with 0.1 U of ribonuclease T1 (Ambion) in 20 mM Tris, pH 7.5, 50 mM NaCl, 0.1 mM MgCl₂ at room temperature for 20 min. The RNase T1 reaction was stopped by adding 0.2 volumes of 5 mM EDTA. Equal volumes of loading buffer containing 7 M urea, 1 × Tris-borate-EDTA (TBE), and 0.01% direct red dye was added to each reaction, and the samples were heated to 95 °C for 5 min. 10 μL of each sample was analyzed by electrophoresis on a denaturing polyacrylamide sequencing gel (20% polyacrylamide, 19:1 crosslinking, 7 M urea) at 60 W, 45 °C for 3.5 hrs. The gel was visualized by fluorescence of the 5'-Cy5 label (630 nm excitation, 670 emission) with Typhoon FLA 9500 Phosphorimager (GE Healthcare Life Sciences) and was analyzed with ImageQuant software.

X-ray crystallography RNA preparation. Sequences of the RNAs employed in this study are based on *TiPreQ1* riboswitch aptamer domain and listed in Supplementary Table 1. RNAs were purchased from Dharmacon and deprotected according to the manufacturer's instructions. After lyophilization, RNAs were dissolved in water and stored at –25 °C.

Crystallization and diffraction data collection. RNA solutions containing 11.8 mM sodium cacodylate (pH 7.0) were heated at 65 °C for 2 min. MgCl₂ and synthetic compound were then added to the final concentrations of 10 mM and 0.5 mM, respectively. The solutions were incubated at 65 °C for 3 min, and then cooled down gradually to the room temperature. *TiPreQ1* riboswitch was crystallized by the hanging-drop vapor diffusion method at 21 °C, under conditions containing 5–15 mM Mg acetate, 50 mM MES (pH 5.6), and 2.3–2.7 M ammonium sulfate (Condition I) and containing 0.1–0.3 M potassium sodium tartrate, 100 mM sodium citrate (pH 5.6), and 2.0–2.8 M ammonium sulfate (Condition II). Hanging drops were prepared by mixing 1 μL of the RNA solution (0.38 mM RNA in 10 mM sodium cacodylate, pH 7.0, 10 mM MgCl₂, and 0.5 mM synthetic compound) with 1 μL of the reservoir solution and were equilibrated against 400 μL of reservoir solution. The RNA crystals grew within 2 weeks to maximum dimensions of 300 × 100 × 100 μm^3 . For data collection, the RNA crystals were transferred to the cryoprotectant solutions containing 10 mM Mg acetate, 27.5 mM MES (pH 5.6), and 2.15 M lithium sulfate for the crystals obtained with the Condition I, and 60 mM

potassium sodium tartrate, 25 mM sodium citrate (pH 5.6), and 2.15 M lithium sulfate for the crystals under the Condition II. The crystals were mounted in a nylon loop and flash frozen by plunging into liquid nitrogen. X-ray diffraction data were collected at the beamlines 5.0.1 and 5.0.2 of the Advanced Light Source (ALS), Lawrence Berkeley National Laboratory. Diffraction data were integrated and scaled with the program DIALS⁴⁹. Data processing statistics are summarized in Supplementary Table 7.

Structure determination and refinement. The structures were solved by the MR method using the previously determined *T. tengcongensis* PreQ₁ aptamer structure (PDB ID: 3Q51 [<https://doi.org/10.2210/pdb3Q51/pdb>]) as a search model with the program PHASER⁵⁰. The solutions were subjected to simulated annealing, energy minimization, restrained isotropic B-factor, and TLS refinement with PHENIX⁵¹, and the resulting electron-density maps revealed the locations of the small molecules. Iterative cycles of refinement and manual rebuilding⁵² produced the current co-crystal structures of ab13_14-1, ab13_14_15-1, ab13_14-2, ab13_14_15-3, and ab13_14-PreQ₁ with R_{free} of 20.5, 20.5, 20.3, 24.8, and 21.0% at 1.80, 1.80, 1.94, 2.56, and 1.69 Å resolution, respectively. Refinement statistics are summarized in Supplementary Table 7. Molecular graphics were produced with PyMol (<http://www.pymol.org/>). Stereo images are provided in Supplementary Figure 12 and 13.

Single-round transcription termination assay. The transcription termination assays were carried out according to established protocols^{53,54} with several modifications. The DNA plasmid containing λ_{PR} promoter and 26-nt C-less sequence followed by the *Staphylococcus saprophyticus* PreQ₁ riboswitch and its downstream sequence cloned into pIDTSMART-AMP was purchased from Integrated DNA Technologies. The DNA template was amplified by PCR from the plasmid using forward and reverse primers (Supplementary Table 6), and then was gel-extracted after agarose gel electrophoresis for purification. Halted transcription complexes were prepared in a solution containing 1 μM GTP, 5 μM ATP, 5 μM UTP, 100 μM ApU, [α -³²P] GTP, 75 nM DNA template, 0.0167 U/ μL *Escherichia coli* RNA polymerase holoenzyme (New England BioLabs) in 1 \times transcription buffer (20 mM Tris-HCl, pH 8.0, 2 mM NaCl, 1 mM MgCl₂, 4% glycerol, 0.1 mM DTT, and 0.1 mM EDTA), and incubated at 37 °C for 15 min. A DNA oligonucleotide complementary to the 26-nt C-less sequence was added to the reactions at 1.1 μM final concentration in 1 \times transcription buffer and incubated at room temperature for 5 min, in order to prevent undesired non-specific interactions between the 26-nt C-less sequence and riboswitch. Elongation was restarted by combining 9 μL of halted transcription complex, 3 μL of the compound (0–25 mM compound and 25% DMSO in 1 \times transcription buffer), and 3 μL of NTPs mix (200 μM ATP, 200 μM CTP, 200 μM GTP, 200 μM UTP, 100 $\mu\text{g}/\text{mL}$ heparin, and 250 mM KCl in 1 \times transcription buffer), and incubated at 37 °C for 20 min. To remove the DNA template, 0.5 U of RQ1 RNase-Free DNase (Promega) was added to the reactions and incubated at 37 °C for 10 min. The reactions were stopped by adding equal volume of loading dye (8 M urea, 20% sucrose, 0.05% bromophenol blue, and 0.05% xylene cyanol in 2 \times TBE). The reaction mixture was separated by 8% denaturing PAGE and visualized by phosphorimager. Full gel images are provided as Supplementary Figures. The band intensity was analyzed by ImageQuant software (GE Healthcare). Termination efficiency was calculated with dividing the intensity for the terminated RNA band by those for the total (terminated and anti-terminated) RNAs. The sequences of the DNA oligonucleotides used in the transcription termination assay are summarized in Supplementary Table 5.

Synthetic procedures and characterization. Chemical synthesis and compound characterization are provided in the Supplementary Methods.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Atomic coordinates and the structure factors of the co-crystal structures of ab13_14-1, ab13_14_15-1, ab13_14-2, ab13_14_15-3, and ab13_14-PreQ₁ have been deposited in the Protein Data Bank, under the accession codes 6EIS, 6EIT, 6EIU, 6EIV and 6EIW, respectively. The source data underlying Figs. 3A–B, 5B, 7B and Supplementary Figs 2, 3, 5, 7B, and 10 are provided as a Source Data file. All other relevant data are available from the authors upon request.

Received: 7 September 2018 Accepted: 8 March 2019

Published online: 02 April 2019

References

- Cech, T. R. & Steitz, J. A. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **157**, 77–94 (2014).
- Ulitsky, I. & Bartel, D. P. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**, 26–46 (2013).
- Sahu, A., Singhal, U. & Chinnaiyan, A. M. Long noncoding RNAs in cancer: from function to translation. *Trends Cancer* **1**, 93–109 (2015).
- Cheatham, S. W., Gruhl, F., Mattick, J. S. & Dinger, M. E. Long noncoding RNAs and the genetics of cancer. *Br. J. Cancer* **108**, 2419–2425 (2013).
- Cooper, T. A., Wan, L. & Dreyfuss, G. RNA and disease. *Cell* **136**, 777–793 (2009).
- Thomas, J. R. & Hergenrother, P. J. Targeting RNA with small molecules. *Chem. Rev.* **108**, 1171–1224 (2008).
- Guan, L. R. & Disney, M. D. Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.* **7**, 73–86 (2012).
- Connelly, C. M., Moon, M. H. & Schneekloth, J. S. Jr. The emerging role of RNA as a therapeutic target for small molecules. *Cell Chem. Biol.* **23**, 1077–1090 (2016).
- Guo, J. U. & Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **353**, aaf5371 (2016).
- Leone, S. & Santoro, R. Challenges in the analysis of long noncoding RNA functionality. *FEBS Lett.* **590**, 2342–2353 (2016).
- Breaker, R. R. Riboswitches and the RNA world. *Cold Spring Harb. Perspect. Biol.* **4**, a003566 (2012).
- Serganov, A. & Nudler, E. A decade of riboswitches. *Cell* **152**, 17–24 (2013).
- Roth, A. & Breaker, R. R. The structural and functional diversity of metabolite-binding riboswitches. *Annu. Rev. Biochem.* **78**, 305–334 (2009).
- Blount, K. F. & Breaker, R. R. Riboswitches as antibacterial drug targets. *Nat. Biotechnol.* **24**, 1558–1564 (2006).
- Howe, J. A. et al. Selective small-molecule inhibition of an RNA structural element. *Nature* **526**, 672–677 (2015).
- Furukawa, K. et al. Identification of ligand analogues that control c-di-GMP riboswitches. *ACS Chem. Biol.* **7**, 1436–1443 (2012).
- Blount, K. F. et al. Novel riboswitch-binding flavin analog that protects mice against *Clostridium difficile* infection without inhibiting cecal flora. *Antimicrob. Agents Chemother.* **59**, 5736–5746 (2015).
- Warner, K. D. et al. Validating fragment-based drug discovery for biological RNAs: lead fragments bind and remodel the TPP riboswitch specifically. *Chem. Biol.* **21**, 591–595 (2014).
- Blount, K. F., Wang, J. X., Lim, J., Sudarsan, N. & Breaker, R. R. Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.* **3**, 44–49 (2007).
- Vicens, Q. et al. Structure-activity relationship of flavin analogues that target the flavin mononucleotide riboswitch. *ACS Chem. Biol.* **13**, 2908–2919 (2018).
- Daldrop, P. et al. Novel ligands for a purine riboswitch discovered by RNA-ligand docking. *Chem. Biol.* **18**, 324–335 (2011).
- Gilbert, S. D., Reyes, F. E., Edwards, A. L. & Batey, R. T. Adaptive ligand binding by the purine riboswitch in the recognition of guanine and adenine analogs. *Structure* **17**, 857–868 (2009).
- McCown, P. J., Liang, J. J., Weinberg, Z. & Breaker, R. R. Structural, functional, and taxonomic diversity of three preQ1 riboswitch classes. *Chem. Biol.* **21**, 880–889 (2014).
- Iwata-Reuyl, D. An embarrassment of riches: the enzymology of RNA modification. *Curr. Opin. Chem. Biol.* **12**, 126–133 (2018).
- Meier, F., Suter, B., Grosjean, H., Keith, G. & Kubli, E. Queuosine modification of the wobble base in tRNAHis influences 'in vivo' decoding properties. *EMBO J.* **4**, 823–827 (1985).
- Rakovich, T. et al. Queuosine deficiency in eukaryotes compromises tyrosine production through increased tetrahydrobiopterin oxidation. *J. Biol. Chem.* **286**, 19354–19363 (2011).
- Durand, J. M., Dagberg, B., Uhlin, B. E. & Bjork, G. R. Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of *Shigella flexneri*: the expression of the virF gene. *Mol. Microbiol.* **35**, 924–935 (2000).
- Klein, D. J., Edwards, T. E. & Ferre-D'Amare, A. R. Cocystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat. Struct. Mol. Biol.* **16**, 343–344 (2009).
- Jenkins, J. L., Krucinska, J., McCarty, R. M., Bandarian, V. & Wedekind, J. E. Comparison of a preQ1 riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J. Biol. Chem.* **286**, 24626–24637 (2011).
- Roth, A. et al. A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* **14**, 308–317 (2007).
- Wu, M. C. et al. Rational re-engineering of a transcriptional silencing PreQ1 riboswitch. *J. Am. Chem. Soc.* **137**, 9015–9021 (2015).
- Hergenrother, P. J., Depew, K. M. & Schreiber, S. L. Small-molecule microarrays: covalent attachment and screening of alcohol-containing small molecules on glass slides. *J. Am. Chem. Soc.* **122**, 7849–7850 (2000).
- Kuruvilla, F. G., Shamji, A. F., Sternson, S. M., Hergenrother, P. J. & Schreiber, S. L. Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays. *Nature* **416**, 653–657 (2002).
- Duffner, J. L., Clemons, P. A. & Koehler, A. N. A pipeline for ligand discovery using small-molecule microarrays. *Curr. Opin. Chem. Biol.* **11**, 74–82 (2007).

35. Bradner, J. E. et al. A robust small-molecule microarray platform for screening cell lysates. *Chem. Biol.* **13**, 493–504 (2006).
36. Bradner, J. E., McPherson, O. M. & Koehler, A. N. A method for the covalent capture and screening of diverse small molecules in a microarray format. *Nat. Protoc.* **1**, 2344–2352 (2006).
37. Connelly, C. M., Boer, R. E., Moon, M. H., Gareiss, P. & Schneekloth, J. S. Jr. Discovery of inhibitors of MicroRNA-21 processing using small molecule microarrays. *ACS Chem. Biol.* **12**, 435–443 (2017).
38. Connelly, C. M., Abulwerdi, F. A. & Schneekloth, J. S. Jr. Discovery of RNA binding small molecules using small molecule microarrays. *Methods Mol. Biol.* **1518**, 157–175 (2017).
39. Abulwerdi, F. A. & Schneekloth, J. S. Jr. Microarray-based technologies for the discovery of selective, RNA-binding molecules. *Methods* **103**, 188–195 (2016).
40. Sztuba-Solinska, J. et al. Identification of biologically active, HIV TAR RNA-binding small molecules using small molecule microarrays. *J. Am. Chem. Soc.* **136**, 8402–8410 (2014).
41. Gilbert, S. D., Rambo, R. P., Van Tyne, D. & Batey, R. T. Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat. Struct. Mol. Biol.* **15**, 177–182 (2008).
42. Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956 (2002).
43. Dalvit, C. et al. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J. Biomol. NMR* **18**, 65–68 (2000).
44. 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**, 10462–10475 (2013).
45. Hoffmann, M. M., Sobstyl, H. S. & Badali, V. A. T2 relaxation measurement with solvent suppression and implications to solvent suppression in general. *Magn. Reson. Chem.* **47**, 593–600 (2009).
46. Headey, S. J., Pearce, M. C., Scanlon, M. J. & Bottomley, S. P. Blind Man's bluff - elaboration of fragment hits in the absence of structure for the development of antitrypsin deficiency inhibitors. *Aust. J. Chem.* **66**, 1525–1529 (2013).
47. Kim, P. B., Nelson, J. W. & Breaker, R. R. An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism. *Mol. Cell* **57**, 317–328 (2015).
48. Rizvi, N. F. et al. Discovery of selective RNA-binding small molecules by affinity-selection mass spectrometry. *ACS Chem. Biol.* **13**, 820–831 (2018).
49. Winter, G. et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. D. Struct. Biol.* **74**, 85–97 (2018).
50. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
51. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
52. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
53. Artsimovitch, I. & Henkin, T. M. In vitro approaches to analysis of transcription termination. *Methods* **47**, 37–43 (2009).
54. Jones, C. P. & Ferre-D'Amare, A. R. Recognition of the bacterial alarmone ZMP through long-distance association of two RNA subdomains. *Nat. Struct. Mol. Biol.* **22**, 679–685 (2015).

Acknowledgements

We thank the staff of beamlines 5.0.1 and 5.0.2 of the ALS for diffraction data collection, and S. Bachas, C. Jones, and R. Trachman III for discussions. We thank J. Kelley for assistance with HRMS work. This work was supported in part by the Center for Cancer Research of the National Cancer Institute, NIH (1 ZIA BC011585 04), the Intramural Program of the National Heart, Lung and Blood Institute, NIH, the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number 16KK0166), and a grant from the Astellas Foundation for Research on Metabolic Disorders.

Author contributions

C.C. designed experiments, performed SMM screens, fluorescence titrations, in-line probing, NMR experiments, and transcription termination assays. T.N. designed and performed transcription termination assays and crystallography. R.B. performed synthetic chemistry and characterization. M.M. performed SMM screens and affinity titrations. R.S. and J.B. performed NMR experiments. A.F. designed experiments and helped write the manuscript. J.S. conceived the project, designed experiments, and helped write the manuscript. All authors contributed in editing the manuscript.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-09493-3>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Journal Peer Review Information: *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2019