A chemical probe based on the PreQ₁ metabolite enables transcriptome-wide mapping of binding sites

Sumirtha Balaratnam¹,⁷, Curran Rhodes¹,⁷, Desta Doro Bume¹, Colleen Connelly¹, Christopher C. Lai¹, James A. Kelley¹, Kamyar Yazdani¹, Philip J. Homan²,³, Danny Incarnato⁴, Tomoyuki Numata⁵,⁶ & John S. Schneekloth Jr ⁴✉

The role of metabolite-responsive riboswitches in regulating gene expression in bacteria is well known and makes them useful systems for the study of RNA-small molecule interactions. Here, we study the PreQ₁ riboswitch system, assessing sixteen diverse PreQ₁-derived probes for their ability to selectively modify the class-I PreQ₁ riboswitch aptamer covalently. For the most active probe (11), a diazirine-based photocrosslinking analog of PreQ₁, X-ray crystallography and gel-based competition assays demonstrated the mode of binding of the ligand to the aptamer, and functional assays demonstrated that the probe retains activity against the full riboswitch. Transcriptome-wide mapping using Chem-CLIP revealed a highly selective interaction between the bacterial aptamer and the probe. In addition, a small number of RNA targets in endogenous human transcripts were found to bind specifically to 11, providing evidence for candidate PreQ₁ aptamers in human RNA. This work demonstrates a stark influence of linker chemistry and structure on the ability of molecules to crosslink RNA, reveals that the PreQ₁ aptamer/ligand pair are broadly useful for chemical biology applications, and provides insights into how PreQ₁, which is similar in structure to guanine, interacts with human RNAs.

¹ Chemical Biology Laboratory, National Cancer Institute, Frederick, MD 21702, USA. ² Center for Cancer Research Collaborative Bioinformatics Resource, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. ³ Advanced Biomedical Computational Science, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA. ⁴ Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Groningen, The Netherlands. ⁵ Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 744 Motooka, Nishi-Ku Fukuoka-Shi Fukuoka 812-8582, Japan. ⁶ Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba-shi, Ibaraki 305-8566, Japan. ⁷ These authors contributed equally: Sumirtha Balaratnam, Curran Rhodes. ✉email: schneeklothjs@mail.nih.gov

https://doi.org/10.1038/s41467-021-25973-x
Riboswitches are naturally occurring RNA sequences that influence bacterial gene expression by binding directly to small molecules. As they can bind specifically and selectively to small molecules, these important functional elements provide invaluable systems to study RNA-small molecule interactions. One such system is the PreQ1 riboswitch, which binds to the small, modified nucleotide 7-aminomethyl-7-deazaguanine (known as PreQ1). PreQ1 riboswitches have been divided into three classes based on differences in function and size. Upon recognition of PreQ1, the RNA changes conformation and alters gene expression at either transcriptional level or translational level. This system serves as a metabolic feedback sensor for PreQ1 levels, allowing bacteria to control the expression of genes involved in one carbon metabolism in response to metabolite levels. While many bacteria contain diverse PreQ1-responsive riboswitches in mRNAs, human mRNAs appear to be devoid of analogous elements. To date, the study of PreQ1 in human transcriptomes has mostly been limited to covalent modification of tRNAs by queuosine (of which PreQ1 is a metabolic precursor in bacteria). PreQ1 has also been utilized in a series of intriguing studies aimed at enzymatic tagging of complex RNAs. Owing to the high selectivity and affinity of the PreQ1/aptamer interaction, as well as ease of chemical modification, it could find broad use in chemical biology and synthetic biology applications. Additionally, PreQ1 is structurally very similar to guanine and thus a PreQ1-derived probe could provide insights into how this and other similar nucleobases recognize human RNA structures. Thus, a broad assessment of PreQ1 binding in human transcriptomes would not only clarify the selectivity of the PreQ1 aptamer interaction but also enable an unbiased examination of metabolite binding to human RNAs.

Reactive molecules that covalently modify RNAs have played a substantial role in understanding RNA biology by probing structure, controlling gene expression, imaging RNAs, tagging RNAs with functional handles, and demonstrating target engagement for RNA-binding small molecules. However, examples of structurally characterized RNA/ligand complexes are relatively rare. Still, highly specific RNA/ligand pairs can find broad use in diverse applications. As interest in RNA as a target for small molecule drugs has increased, so has the need for information about the design of RNA-targeting probes. Understanding the chemical features that impact RNA crosslinking efficiency would influence probe design, as well as inform interpretation of results from experiments that use these probes. To date, there has been limited study on the factors that govern covalent crosslinking efficiency to RNA. Furthermore, the impact of tagging small ligands with reactive handles could be substantial. Aptamers can discriminate between highly similar ligands, even distinguishing fluoride from chloride, potentially representing a challenge in probe design. Drawing inspiration from the approaches above, we set out to design covalent probes targeting a PreQ1 riboswitch, and to apply them for a broad assessment of binding in complex systems.

The class-I PreQ1 riboswitch aptamers (PreQ1-I) represent an ideal system for this study due to their relatively small size and its conserved ligand binding sites. For example, the PreQ1 riboswitch from Bacillus subtilis (Bs-preQ1-RS) consists of just 34 nucleotides. Upon binding to PreQ1, it folds into a H-type pseudoknot; this conformational change results in the transcriptional downregulation of queCDEF operon. Similarly, Staphylococcus saprophyticus (Ss-preQ1-RS) carries a related class-I PreQ1 riboswitch that also functions by transcriptional termination. For the PreQ1 riboswitch from Thermoaerobacter tengcongensis (Tt-preQ1-RS) binding to PreQ1 also results in the formation of an H-type pseudoknot structure. However, in this case the riboswitch acts at the translational level by regulating the accessibility of the Shine-Dalgarno sequence and ribosome binding.

While the mechanism and sequences of these PreQ1-I class riboswitches differ they all have tight and selective interactions with the PreQ1 ligand. For example, the Tt-preQ1-RS has an equilibrium dissociation constant of 2 nM for PreQ1, while the Bs-preQ1-RS constant has been reported as 20 nM from in-line probing. In addition to this tight binding, these aptamers have been studied extensively using both X-ray crystallography and NMR, providing detailed information about the mode of ligand binding. Other studies on this system include how it responds to other natural metabolites, engineering to remodel ligand specificity, and investigations for inhibitor design. PreQ1 riboswitches from other organisms have been studied in detail as well, revealing multiple distinct classes of independently evolved riboswitches that function by different mechanisms. In fact, PreQ1 riboswitches are among the most frequently occurring riboswitches in Nature.

Here, we report the design and evaluation of a series of covalent probes for the PreQ1-I class riboswitch. In total, a series of 16 probes are reported, including both electrophilic and photochemically activated crosslinkers. Since the PreQ1-I class riboswitches have a small aptamer domain and have conservation of PreQ1 binding, we utilize multiple PreQ1-I class aptamers (Bs, Ss, Tt) from different species in this study to ensure that the best probes could be broadly recognized by these riboswitches. Gel-based and mass spectrometry experiments are conducted to quantify crosslinking efficiency and characterize the covalent adduct. We report a remarkable effect of linker length and crosslinker structure on crosslinking efficiency. None of the electrophilic probes showed evidence of covalent modification, however photocrosslinking probes had varying efficiency. One of these probes was selected for further in-depth characterization. The probe was found to photocrosslink only to the bases of the aptamer, not sugar or phosphate moieties. Additionally, the photocrosslinking probe retained activity in a transcriptional termination assay. An X-ray co-crystal structure of the probe in complex with the aptamer demonstrated that the mode of binding was unperturbed by the introduction of the crosslinker sidechain. We also demonstrate the potential utility of this approach for labeling, target engagement and enrichment studies. Transcriptome-wide mapping studies (Chem-CLIP) in the presence of the bacterial aptamer reveal an unexpectedly high specificity for the bacterial aptamer over all other human RNAs. In total human RNA (without aptamer added), the probe enriched several transcripts significantly. Validation of these genes with competitive assays revealed several specific interactions between PreQ1 and human RNAs.

**Results**

**Synthesis and evaluation of electrophilic probes.** To design covalent probes targeting the PreQ1-I class riboswitch aptamers, a structure-informed strategy was taken. From multiple available structures of ligand-bound riboswitches it was apparent that the pendant amine from PreQ1 was involved with binding but was also mostly solvent exposed. Furthermore, proximal to the binding site were multiple different potential nucleophilic groups. For example, the 2′-OH of G11 or base side chains of G4
and G5 from the *Thermoanaerobacter tengcongensis* (Tt-preQ₁-RS)(PDB: 3Q50) aptamer are all near the binding site. We thus designed and synthesized a series of probes that contain commonly used electrophilic warheads for covalent modifications that could potentially react with these groups. As described in Table 1, these compounds included Michael acceptors (1–3), acylating reagents (4–9), and the chlorambucil derivative that has been used in Chem-CLIP experiments (10) and has been reported to successfully alkylate RNA in a non-specific manner.

The reactivity of each of these probes towards PreQ₁-RS aptamers was evaluated using mass spectrometry (e.g., MALDI-TOF) and denaturing PAGE experiments. Initial in vitro analysis of the reaction mixtures between the Tt-preQ₁-RS aptamers and the electrophilic probes showed some evidence of higher molecular weight species corresponding to potential adducts in low yields. However, these species did not correlate to a specific product. Furthermore, when these reaction mixtures were subjected to denaturing PAGE analysis, no evidence was seen for specific or modified RNA (Supplementary information Fig. S1). Efforts to optimize the reaction conditions by altering reaction temperature or buffer conditions did not affect the outcome. Based on these findings, we speculate that these electrophiles may simply not be reactive enough to yield high percentages of covalently modified RNAs, or that the electrophiles could not achieve the proper geometry for covalent modification. Similarly, others have reported that modification of the amine at the C7 position of PreQ₁ with various aminoaalkyl chains results in decreased activity in an in vivo *E. coli* reporter system. To overcome this challenge, we turned our attention towards the synthesis and evaluation of more reactive, diazirine-containing photoaffinity crosslinkers.

**Design and optimization of photocrosslinking covalent probes.** Similar to electrophilic probes, photoaffinity probes are commonly employed for labeling studies with proteins. Diazirine functionalized nucleosides have been used previously to map nucleic acid-ligand interactions or conjugate dyes to aptamers. However, typically only one probe was reported, and the effect of linker structure or crosslinking group was not broadly evaluated. In order to assess how photoaffinity linker structure influences crosslinking efficiency, a series of photoaffinity probes were designed and synthesized (Fig. 1 and Table 1). Additional details for the synthesis and characterization of these compounds are provided in the supplementary information (Section 3: synthesis and characterization). Included in this series are probes varied in crosslinker structure (stabilized and unstabilized diazirines), as well as linker length. The Bsp-preQ₁-RS was treated with 50-fold excess of probe in the presence of riboswitch buffer containing 2 mM MgCl₂, irradiated with 365 nm light sources for 15 minutes, and PAGE and MALDI MS analysis was used to assess the formation of crosslinks. Gratifyingly, the expected adducts were observed in all cases, however only trace crosslinking was observed with 14 in MALDI. Crosslinking efficiencies were above 30% for unstabilized probes 11, 12, and 13, and approximately 20% with stabilized probes 15 and 16 (Supplementary information Fig. S2).

**Effect of linker length and composition on photocrosslinking efficiency.** The behavior of compounds 14–16 highlight the importance of the linker structure and length in preserving the native interaction between PreQ₁ and the riboswitch as well as enabling crosslinking. Crosslinking was not observed for 14 in denaturing PAGE analysis, potentially due to steric clashes that could occur between the phenyl ring in the linker and the RNA that prevent the compound from inserting fully into the binding pocket. The results for compounds 15 and 16 support this conclusion as some crosslinking can be restored when the phenyl ring is spaced further away from the PreQ₁ core by PEG chains. Overall, we showed that both compounds 11 and 13 can modify the aptamer with high efficiency (>30%) under these conditions. We chose to proceed with compound 11 because it retained the positive charge in the natural PreQ₁ ligand and contained an alkylene handle for further modification.

We next evaluated the effects of crosslinking 11 to the PreQ₁ aptamer as a function of both time and concentration. In both denaturing PAGE and MALDI-TOF assays, clear dose and time-dependent crosslinking was observed (Figs. 2a, b and S3). Importantly, even at higher doses or longer times of irradiation, a single crosslinked species was the predominant product, indicating a highly specific crosslinking event.

**Competitive crosslinking experiments demonstrate selectivity in vitro.** With an optimized probe in hand, we next explored whether the photoaffinity labeling with compound 11 achieves proximity induced site-specificity or if the observed adducts were due to indiscriminate reactivity of diazirine with the RNA. Consequently, we performed labeling experiments in the presence of increasing concentrations of unmodified PreQ₁. PreQ₁ showed a dose-dependent inhibition of crosslinking of 11, with complete inhibition of crosslinking observed at higher concentrations (Fig. 2c). To further establish the selectivity of 11, we evaluated the labeling efficiency in the presence of up to tenfold excess of tRNA with respect to the aptamer. No significant decrease in the labeling efficiency was observed under these conditions, indicating that 11 retains selectivity for the aptamer in the presence of excess tRNA (Fig. 2d). We also performed labeling experiments with other structured RNAs such as microRNA 21 (miR-21) and SAM II riboswitch with compound 11. No detectable high molecular weight (slow migrating) crosslink product was observed in denaturing PAGE gels under these conditions (Fig. 2e).

**Mass spectrometry and crystallography establish mode of interaction with 11.** In an effort to probe the site-specificity of the crosslink further, the photocrosslinked RNA was subjected to a nucleoside digestion using NEB Nucleoside Digestion Mix. The resulting reaction mixture was analyzed by Orbitrap-Liquid chromatography–mass spectrometry (LC/MS) and we identified guanosine-compound 11 adduct as a primary modified species (Fig. 3a, b). Moreover, further fragmentation and MS/MS analysis of guanosine adduct using a higher-energy collisional dissociation (HCD) and collision-induced dissociation (CID) techniques showed either a loss of the sugar or the PreQ₁ portion of the probe (Fig. 3c–e). Based on these fragmentation patterns, we conclude that the photoaffinity probe exclusively reacts with guanosine on the nucleobase rather than on the sugar. These findings are consistent with proximity-induced alkylation of guanines located near the binding site of the ligand as observed in crystal structures.

To further validate the binding of 11 to the PreQ₁ aptamer, we performed X-ray crystallography with the PreQ₁ aptamer from *T. tengcongensis* (Tt-preQ₁-RS), using a previously reported method. Efforts to determine the X-ray crystal structure of 11 photocrosslinked to the Tt-preQ₁-RS aptamer were unsuccessful. Co-crystals were obtained using wild type (WT) and an abasic mutant at positions 13 and 14 (ab_13-14) without photocrosslinking, and the complex structures were solved by the molecular replacement method at 2.80 and 1.57 Å resolution respectively (Fig. 4a, b and Fig. S4, and Table S1). Positions A13 and A14 of the Tt-preQ₁-RS are important for stabilization of the expression platform. However, when we co-crystallized the wild-type Tt-
Table 1 Structures and crosslinking efficiencies for electrophilic and photoaffinity probes used for proximity induced covalent modification of the Tt-preQ₁-RS aptamer (compounds 1–10) or Bs-preQ₁-RS aptamer (compounds 11–16). Error bars indicate standard deviation, \( n = 3 \). (Molar ratio between the aptamer and probe is 1:50).

| Compound ID | Compound structure | % Crosslink efficiency (PAGE) |
|-------------|--------------------|------------------------------|
| 1           | ![Compound 1 structure](image1) | Not observed                  |
| 2           | ![Compound 2 structure](image2) | Not observed                  |
| 3           | ![Compound 3 structure](image3) | Not observed                  |
| 4           | ![Compound 4 structure](image4) | Not observed                  |
| 5           | ![Compound 5 structure](image5) | Not observed                  |
| 6           | ![Compound 6 structure](image6) | Not observed                  |
| 7           | ![Compound 7 structure](image7) | Not observed                  |
| 8           | ![Compound 8 structure](image8) | Not observed                  |
| 9           | ![Compound 9 structure](image9) | Not observed                  |
| 10          | ![Compound 10 structure](image10) | Not observed                  |
| 11          | ![Compound 11 structure](image11) | 31.3 ± 2.9                    |
| 12          | ![Compound 12 structure](image12) | 49.2 ± 3.2                    |
| 13          | ![Compound 13 structure](image13) | 39.9 ± 5.3                    |
| 14          | ![Compound 14 structure](image14) | Trace                         |
preQ1-RS with the compound 11, only small co-crystals were obtained that diffracted to 2.8 Å resolution. In contrast, when using abasic mutant at positions 13 and 14, larger co-crystals were obtained that diffracted to 1.6 Å resolution. Importantly, the binding mode of the compound is conserved between the WT and abasic mutant crystals and the RNA structures are quite similar with the exception of the nucleotides from positions 12 to 14. No specific interaction is observed between the alkyne handle and the RNA, suggesting that positions 12 to 14 are not involved in the direct interaction for binding of the compound 11. Based on these findings, it was concluded that a high-resolution co-crystal structure of the abasic mutant is more informative in terms of evaluating the chemical determinants of this interaction.

From inspection of the structure, it is clear that the PreQ1 portion of 11 binds to the Tt-preQ1-RS aptamer in an analogous mode to the native ligand, making near identical contacts. As shown in Fig. 4b, the methylamine group hydrogen-bonds with O6 of G5. This hydrogen bonding pattern is similar to the interaction observed in our previously reported structure (PDB ID: 6E1W), as well as that of the Bs riboswitch structure (PDB ID: 3K1V). However, the hydrogen bond between N7 of G5 and the methylamine group was not observed in the present crystal structure. This is likely due to the impact of the extended linker on N7, which is flexible and may interfere with this interaction. The sidechain of the probe projects out of the binding pocket toward the solvent and is proximal to two guanosines (G4 and G5) that it could potentially modify upon irradiation. Thus, it is clear that modification of the exocyclic amine of PreQ1 does not dramatically alter its mode of binding to RNA, though it does decrease affinity. To further identify which guanosine residue in the active site is primarily modified by 11, we designed three different mutants for Tt-preQ1-RS aptamers (G4A, G5A, and G11A) by replacing each guanine near the binding site with an adenine. When we performed photocrosslinking experiments for the wild type Tt-preQ1-RS aptamer and the three mutants, we observed approximately 40% reduction in crosslinking efficiency to the G4A mutant (Tt-G4A). In contrast, the other two mutants (Tt-G5A and Tt-G11A) showed similar crosslinking efficiency to the wild type Tt-preQ1-RS aptamer (Tt-WT) (Fig. 4c and Fig. S5). Thus, the identity of the nucleotide at position 4 is most important for photocrosslinking.

Probe 11 retains transcriptional termination activity. Single-round transcription termination assays were performed to assess the function of 11. We used a previously reported protocol to assess the activity of both the cognate ligand and the photocrosslinking probe. The unmodified ligand (PreQ1) achieved a maximal termination efficiency of ~85% with an EC50 of 7.4 nM, while probe 11 achieved a maximal termination efficiency of ~65% with an EC50 of 7.1 µM (Fig. 4d and Fig. S6), demonstrating that probe 11 functions in a similar way to the cognate ligand, albeit requiring a higher concentration. The decreased activity in the termination assay is consistent with decreased affinity of 11 when compared to the cognate ligand (Fig. S7). In Microscale thermophoresis (MST) experiments with both the Bs-preQ1-RS and Tt-preQ1-RS aptamers, we observed that probe 11 binds approximately 50-fold weaker than the unmodified ligand. (Fig. S7). Importantly, modification of the N7 of PreQ1 with alkyl tethers has been shown to reduce cell activity by others as well. Since the termination assays were performed without irradiation, these experiments confirm that the presence of the photocrosslinker
Fig. 2 Biochemical optimization of photocrosslinking and competition experiments to evaluate the specificity of compound 11 for the Ss-preQ₁-RS aptamer. a Dose-dependent crosslinking efficiency of compound 11. b Time-dependent crosslinking efficiency of compound 11. Competition experiments in the presence of c PreQ₁ and d tRNA. e PAGE analysis to assess crossreactivity of compound 11 with other structured RNAs shows no quantifiable modification of either miR-21 RNA or SAM-II RNA. Values for crosslinking efficiency are reported as the mean ± standard deviation of three replicate experiments. Source data are provided as a Source Data file.
sidechain does not ablate the inhibitory activity of the ligand or negatively impact selectivity.

Probes 11 retains selectivity in cell lysates and total human RNA. Having demonstrated the selectivity of 11 towards the PreQ₁ aptamer in vitro and confirmed that the photocrosslinking probe binds to the RNA like the unmodified ligand, we next explored the potential applications of ligand-based strategy for labeling applications, including fluorophore- and biotin-labeling. To do this, we synthesized biotin-conjugated 11 (Bio-11) using copper catalyzed click chemistry with biotin-PEG₆-azide and performed photocrosslinking experiments with the Ss-preQ₁-RS under our optimized reaction conditions.
conditions. MALDI-TOF analysis of the reaction mixture showed the expected molecular weight adducts (Supplementary information Fig. S8). Concurrently, the photocrosslinked Ss-preQ1-RS aptamer was also subjected to copper catalyzed click reaction with biotin-PEG$_2$-azide and Cy5-picolyl azide and the expected click product was observed in both experiments by mass spectrometry and denaturing PAGE analysis (Supplementary information Figs. S9 A and B and S10).

Bio-11 was prepared and added to MCF-7 cell lysates that had been supplemented with Ss-preQ$_1$-RS aptamer. After irradiation, the lysates were analyzed by denaturing PAGE. As shown in Fig. 5a, when the gel was stained with SYBR gold stain, a crosslinked higher molecular weight band was observed. In order to confirm the presence of the biotinylated, crosslinked product, RNA samples from the gel were then transferred to a positively charged nylon transfer membrane and the blot was incubated with streptavidin-HRP overnight. As shown in Fig. 5a, biotin modified RNA bands were only observed when the lysates had been photoirradiated in the presence of probe Bio-11.

Next, we demonstrated that 11 could be used for both the photocrosslinking and biotin labeling in a complex environment. We isolated total cellular RNA from MCF-7 cells, added 1 µM of Bs-preQ$_1$-RS (a 70 nt construct containing both the aptamer and expression platforms), and treated with compound 11 in riboswitch buffer. After incubation, samples were irradiated at 365 nm for 15 minutes. Subsequently, we performed copper catalyzed click reaction with TAMRA azide and ran the samples on denaturing PAGE. In the absence of the MCF-7 total RNA, a strong TAMRA fluorescent band was observed at the expected molecular weight indicating that the photocrosslinking and click reactions were successful (Fig. 5b, lane 2). Additionally, only weak fluorescent signals were observed when the Bs-preQ$_1$-RS aptamer was not added into the mixture of 11 and MCF-7 total RNA (Fig. 5b, lane 3). In contrast, as shown in lane 4 of Fig. 5b, compound 11 was able to selectively modify the PreQ$_1$ riboswitch even in the presence of other cellular RNAs.

Consistent with earlier observations with the Ss-preQ$_1$-RS aptamer, photocrosslinking efficiency is significantly reduced in the presence of free PreQ$_1$ competitor (Fig. 5b, lanes 5 & 6). This observed competition between 11 and PreQ$_1$ demonstrates the importance of proximity induced site-specific modification within the binding pocket. Additionally, no fluorescent band was observed when the reaction mixture was not irradiated with UV, confirming that irradiation of 11 is required for detection of the aptamer with this method (Fig. 5b, lane 7). In comparison, when compound 17, a small molecule containing alkyne-diazirine fragment that lacks the PreQ$_1$ targeting portion, was used to treat total RNA under similar conditions, no detectable labeling of the riboswitch was observed, again confirming that the recognition event between the PreQ$_1$ scaffold of 11 and the aptamer is critical for diazirines to be useful as photocrosslinking partners (Supplementary information Fig. S11).

Chem-CLIP and c-Chem-CLIP sequencing experiments to quantitatively assess selectivity of 11 in human transcriptomes. To further investigate crossreactivity of 11 with other human cellular RNAs, a Chem-CLIP experiment was also performed without exogenous Bs-preQ$_1$-RS. In the absence of Bs-preQ$_1$-RS, 11 selectively enriched 16 transcripts including TERC, HIST1H3F and HIST2H2BF (cutoff: log$_2$fold-change > 0.95, adjusted p-value = 2.93e$^{-21}$ (Fig. 6a)). As evident in Fig. 6b and c, no other RNAs were significantly enriched when comparing 11 to 17 in this analysis. This high level of selectivity makes this compound-aptamer an ideal pair for chemical biology applications that require specific labeling in complex environments.

Chem-CLIP and c-Chem-CLIP sequencing experiments confirm target engagement in total human RNA. Chem-CLIP experiments were also performed to further probe the selectivity of 11 for the full-length Bs-preQ$_1$-RS aptamer in the presence of MCF-7 derived total cellular RNA. After UV photocrosslinking in the presence of 11, total RNA was click labeled with biotin-PEG$_2$-azide and enriched with streptavidin-conjugated magnetic beads. Enriched RNA was then reverse transcribed, and samples were submitted for next-generation sequencing. Compound 17, which lacks the PreQ$_1$ targeting warhead, was again used as a negative control. In samples where Bs-preQ$_1$-RS was spiked into total RNA, differential gene expression analysis between 11 and 17 treated samples revealed that the aptamer sequence was significantly enriched by compound 11 (log$_2$fold-change > 2.5, adjusted p-value = 2.93e$^{-21}$ (Fig. 6a)). As evident in Fig. 6b and c, no other RNAs were significantly enriched when comparing 11 to 17 in this analysis. This high level of selectivity makes this compound-aptamer an ideal pair for chemical biology applications that require specific labeling in complex environments.

Competitive reverse transcription PCR (RT-qPCR) experiments (c-Chem-CLIP) were also performed with TERC, HIST1H2F and HIST2H2BF to validate individual enrichment events and assess biochemical competition with unlabeled probes. RT-qPCR, TERC was enriched approximately 1.5-fold in 11 treated samples compared to the negative control (17), in good agreement with sequencing data. Furthermore, this enrichment was competitive with free PreQ$_1$ ligand, which lacks the diazirine-alkyne crosslinking warhead. Similar results were observed with HIST1H3F and HIST2H2BF. Enrichment was also observed for HIST2H2BF, however, competition with PreQ$_1$ was only evident at the highest concentration of competitor (100 µM).
Together, the results of these sequencing and qPCR experiments indicate that these RNAs selectively bind to this ligand. Further investigation is underway to map these cross-links and fully characterize these interactions. Extension of this c-Chem-CLIP crosslinking pull-down strategy to live cells in an effort to confirm the relevance of these potential interactions with 11 is currently underway.

Discussion

Herein, we employ a structure-informed strategy to design and characterize covalent probes of the PreQ1 riboswitch aptamer. Initial attempts to design electrophilic probes were unsuccessful, however photocrosslinking probes reacted with efficiency approaching 50% modification. We applied multiple different biophysical and biochemical assays to assess the reactivity and site-specificity of the photoaffinity probe towards the target RNA in crosslinking experiments. Digestion of the crosslinked aptamer and analysis by mass spectrometry revealed that the photocrosslinker only reacts with guanine bases, with no crosslinks observed to the sugars. Given that there are only 3 conserved guanines in the class-I preQ1 aptamer (Bs, Tt and Ss), two of which are proximal to the binding site, this observation confirms a highly specific crosslinking event. The carbene-based reactivity typically associated with diazirines would be expected to predominate in stabilized diazirine probes such as compounds 14–17 and to alkylate multiple chemically diverse sites due to high reactivity\(^41,42\). However, here we observe unstabilized diazirines as having higher crosslinking efficiency and specific reactivity, suggesting that crosslinking may occur through an alkylation event via a rearranged diazo-intermediate (as observed elsewhere\(^48,49\)). Based on the high reactivity of this intermediate it is likely that 11 can form adducts with multiple different atoms in multiple different nucleobases in the binding pocket, and more than one product is possible. However, we observed that mutation of G4 to adenosine resulted in a roughly 40% reduction in crosslinking, suggesting that this residue is the majority site of modification.

Importantly, binding studies, X-ray crystallography and transcription termination assays revealed that labeling PreQ1 with the crosslinking sidechain does not ablate binding or function and preserves the majority of the contacts that the compound makes with the aptamer. Despite a 50-fold reduction in affinity, the specificity of this interaction enables selective photocrosslinking of 11 to a PreQ1 riboswitch in cell lysates, which further illustrates the robustness of this approach to selectively target the RNA of interest under biologically relevant conditions. Thus, this PreQ1/aptamer pair may find substantial utility in chemical biology applications for targeting RNA due to the ease of manipulation of the PreQ1 scaffold and robustness of the proximity induced crosslinking. This straightforward approach relies exclusively on...
chemical recognition to achieve site-specific modification of an RNA. By fully characterizing this specific recognition event, we were additionally able to probe the transcriptome for potential PreQ1 binding sites on human RNAs. Through Chem-CLIP and competitive qPCR experiments in total cellular RNA, we were able to show that \textit{11} crosslinked to the aptamer with exceptional selectivity. In the absence of aptamer, \textit{11} selectively enriched 16 RNAs when compared to the negative control probe \textit{17}. The high level of structural similarity between PreQ1 and guanine is particularly intriguing in this regard, as it is possible that guanine can also specifically interact with structures in these RNAs. In this way, evaluation of bacterial riboswitch ligands like PreQ1 can shed light on potential existing RNA aptamers for human metabolites.

![Chemical recognition to achieve site-specific modification of an RNA.](image)

**Fig. 5** Diazirine probe \textit{11} selectively labels the PreQ1 aptamer in cell lysates. \textbf{a} In denaturing PAGE experiments Bio-\textit{11} selectively crosslinks the Ss-preQ1-RS aptamer in the presence of cell lysates. The crosslinked product is observed by SYBR gold staining and streptavidin-HRP detection. \textbf{b} Competitive photocrosslinking experiments with Bs-preQ1-RS full-length aptamer (70 nts) spiked into MCF-7 total RNA and imaged by TAMRA labeling (top) and SYBR gold stain (bottom). Source data are provided as a Source Data file.

**Fig. 6** Selective enrichment of Bs-preQ1-RS aptamer by \textit{11} in total human RNA with Bs-preQ1-RS aptamer spiked in. \textbf{a} Comparison of the normalized counts of the Bs-preQ1-RS aptamer across \textit{11} and \textit{17} treated samples shows significantly higher counts in all four replicates of \textit{11} treated samples. Statistical significance was calculated by two-tailed \textit{t}-test analysis. \textbf{b} Correlation plot showing differential gene enrichment for compounds \textit{11} and \textit{17}. The gene entry corresponding to the Bs-preQ1-RS is labeled in red. \textbf{c} Volcano plot from differential expression analysis between \textit{11} and \textit{17}. The gene entry corresponding to the Bs-preQ1-RS is labeled in red. \(p = \text{adjusted } p\)-value. All analyses performed on four independent replicate samples treated with \textit{11} or \textit{17}. 

---

**ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-25973-x**
Humans lack the biosynthetic machinery to produce PreQ₁ and related metabolites such as queuine and queuosine, and they must be acquired from diet or commensal bacteria such as *B. subtilis*. However, they are important metabolites that specifically modify certain human tRNAs to promote protein folding. Moreover, queuosine modification occurs at the wobble position of the tRNA with GUN anticodons and this modification regulates the translation efficiency. The specific role that these metabolites play in humans remains under investigation, though there is at least some evidence that they may also have anti-cancer activity in addition to their regulatory roles in commensal bacteria. Here, we demonstrate that 11 also has specific interactions with important regulatory human RNAs, including TERC, HIST1H3F, and HIST2H2BF, all of which have been shown previously to form stable RNA structures. The structure of vertebrate TERC has been studied extensively and this RNA has been shown to form multiple stable structures including a pseudoknot domain that is required for binding to hTERT. Telomerase activity is low in the majority of healthy somatic cells, however, it is significantly increased in various cancers, and TERC has been implicated as a potential therapeutic target in many of these cancers. While the biological significance of the proposed interaction between 11 and TERC is currently unknown, small molecule probes that can selectively recognize this RNA would be valuable to researchers studying telomerase biology.

**Methods**

**General information.** Unless otherwise stated, all reagents and solvents were purchased from commercial vendors and used without further purification. All ¹H, ¹³C, and ¹⁹F NMR spectra were collected on a 500, 126, and 471 MHz NMR.
spectrometers in (CD$_3$)$_2$SO, respectively. The $^1$H, $^{13}$C, and $^{19}$F NMR chemical shifts are given in parts per million (δ) with respect to an internal tetramethylsilane (TMS, δ = 0 ppm relative to TMS). NMR data are processed, annotated, and reported in the following format: chemical shift (multiplicity (s = singlet, br = broad singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), Integration) using MestReNova. The large-scale synthesis of 7-azadeguanylic acid (PreQ$_3$) was carried out according to previously reported protocol.$^{40}$

Oligonucleotides. Deprotected and HPLC purified riboswitch aptamers or full-length riboswitch were purchased from Horizon Discovery Inc. or IDT with the following sequences and used without any further purifications.

11-mer: $^{5}$-CGGGGAGGCAUUCCACCCUUUAACAAAACAAG-3

11-mer: $^{5}$-CCCCAGUUAACAAAACAAG-3

UV photocrosslinking experiments

For the single screening of compounds 11–16 binding to Bs-preQ$_3$-RS aptamer, the Bs-preQ$_3$-RS (5 µM, 50 µL) in riboswitch buffer (50 mM Tris (pH 7.5), 100 mM KCl, 2 mM MgCl$_2$) was heated for 75 °C for 5 min and allowed to cool to RT over the course of 1 h. For compound treatment, 10 µL of total volume was prepared by adding 9.5 µL of the compound mixture to 0.5 µL of DMSO containing compounds. After pre-incubation, the reaction mixture was UV photocrosslinked for 15 min as described above. Then sample mixtures (contains a mixture of photocrosslinked and unmodified RNA) were assessed for extent of crosslinking by negative ion MALDI-TOF-MS analysis and digestion. The RNA was subjected to enzymatic digestion to generate individual nucleosides and was analyzed by LC/MS/MS system at 30,000 resolutions. High-resolution deconvolution of oligonucleotide anion charge-states was carried out using the Xtract program.$^{52}$ Crosslinked RNA was subjected to enzymatic digestion to generate individual nucleosides and was analyzed by LC/MS/MS to determine where crosslinks were occurring and to identify possible structures. For the enzymatic digestion study, photocrosslinked RNA was transferred to a PCR tube and treated with NEB Nucleosome Digestion Mix® Reaction Buffer (10 µL, 10×) and Nucleosome Digestion Mix (5 µL) and incubated at 37 °C for 1 h. Then, the reaction mixture was analyzed by MS and MS/MS using both CID and HCD. A narrow-bore C18 column system, eluted with CH$_3$OH/H$_2$O gradients tailored to the components of interest, was used for separations. High-resolution CID and HCD MS/MS of the guanosine adduct MH$^+$ indicated that adduction occurred exclusively on the guanine base.

Click reaction on photocrosslinked RNA for labeling with fluorophores and biotin-azides. Biotin or fluorophore-azide probes were crosslinked to aptamer RNA by collecting a 50 mM Tris (pH 7.5) solution of each 10 µM of Probe (5 µM) at room temperature for 30 min. Then samples were UV crosslinked as described above. Then sample mixtures (contains a mixture of photocrosslinked and unmodified RNA) were assessed for extent of crosslinking by negative ion MALDI-TOF-MS analysis and digestion. The RNA was subjected to enzymatic digestion to generate individual nucleosides and was analyzed by LC/MS/MS to determine where crosslinks were occurring and to identify possible structures. For the enzymatic digestion study, photocrosslinked RNA was transferred to a PCR tube and treated with NEB Nucleosome Digestion Mix® Reaction Buffer (10 µL, 10×) and Nucleosome Digestion Mix (5 µL) and incubated at 37 °C for 1 h. Then, the reaction mixture was analyzed by MS and MS/MS using both CID and HCD. A narrow-bore C18 column system, eluted with CH$_3$OH/H$_2$O gradients tailored to the components of interest, was used for separations. High-resolution CID and HCD MS/MS of the guanosine adduct MH$^+$ indicated that adduction occurred exclusively on the guanine base.

Denaturing PAGE analysis (general procedure). For PAGE analysis, RNA samples were mixed with equal volume of denaturing loading dye and heated at 95 °C for 5 min. Equal volumes of the samples were then subjected to 17% PAGE in the presence of 7 M urea. The gel was stained in 1X SYBR® Gold Nucleic Acid Gel Staining Solution (Molecular Probes) for 15 min at room temperature. The gel images were obtained by scanning the gel on an AMERSHAM TYPHOON Biomolecular Imaging. Image software was used for processing. Crosslinking efficiency was calculated by quantifying the higher molecular weight band intensity (crosslink adduct band). Values reported are the average of three replicates.

Oligonucleotides

5′-GUGGGAGGCAUUCCACCCUUUAACAAAACAAG-3′

5′-CCCCAGUUAACAAAACAAG-3′

Oligonucleotides

5′-AGGAUGGCUAGUCUGAUCCCUUUAACAAAACAAG-3′

5′-CG5-3′

5′-AGAGGCUACGCUAACCCUUUAACAAAACAAG-3′

5′-BS5-3′

5′-miR21-5p: 5′-GGGGUGACUGUUGAUAUUGCAGGCCAACC-3′
buffer), and 3 µL of NTPs mix (200 µM ATP, 200 µM CTP, 200 µM UTP, 100 µg/ml heparin, and 250 mM KCl in 1 x transcription buffer), and incubated at 37 °C for 2 h. Finally, the membrane was incubated with the appropriate primer and 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 x TBE). The reaction mixture was separated by 8% denaturing PAGE and visualized by phosphorimage. Gels and images are provided as Supplementary Figures. The band intensity was analyzed by ImageQuant software (GE Healthcare). Termination efficiency was calculated by dividing the intensity for the terminated RNA band by those for the total (terminated and antiterminated) RNAs.

Microscale thermophoresis (MST)-binding measurements for PreQ1 and 11. Fluorescently labeled Bs-preQ1-RS aptamer and Tr-preQ2-RS aptamer were prepared in 1 x Aptamer binding buffer (50 mM Tris (pH 7.5), 100 mM KCl, and 2 mM MgCl2). Both pre-Q1/riboswitch aptamers were annealed by heating to 75 °C for 5 min and then cooled to room temperature over 1 h. A twofold dilution series of PreQ1 or probe 11 was prepared in 100 nM Cy5-labeled aptamer and 5% DMSO, with the final concentrations of PreQ1 or 11 ranging from 500 nM to 0.0153 nM. Samples were incubated at room temperature for at least 15 min. The compounds were added to premium coated capillaries (NanoTemper Technologies) and the sample mixtures were subjected to MST analysis. MST experiments were conducted in triplicate on a Monolight NT.115 system (NanoTemper Technologies). The results were analyzed by Tjump analysis, and the values obtained were normalized and plotted against the PreQ1 or 11 concentrations. The apparent dissociation constant was then determined using a single-site model to fit the curve.

X-ray crystallography experiments and data collection. For X-ray crystallography, we used the preQ1 riboswitch aptamer domain from Thermoanaerobacter tengcongensis (Tt). The wild-type aptamer and abasic mutant (ab13_14), in which the nucleobases at positions 13 and 14 were removed, were employed for crystallization. The crystals of Tt-preQ1-RS aptamer domain in complex with 11 were prepared at 20 °C and cryoprotected by the same method as reported previously. X-ray diffraction data were collected at the beamlines BL32XU and BL45XU of the SPring-8 (Hyogo, Japan) with the aid of an automatic data-collection system ZOZO and BL-1A of the Photon Factory (Ibaraki, Japan). Diffraction data were integrated and scaled with the programs KAMO and XDS. Data processing statistics are summarized in Table S1.

Structure determination and refinement. The co-crystal structures were solved by the molecular replacement method with the program PHASER, using the previously determined Tt-preQ1-RS aptamer domain structure (PDB ID: 6EIW) as the search model. The solutions were subjected to simulated annealing to uncouple the working and free R values in the refinement process, as well as energy minimization, restrained isotropic B-factor, and TLS refinement with PHENIX. The resulting electron density maps revealed the locations of the compound. The atomic models were then refined by iterative cycles of refinement with PHENIX and manual rebuilding with COOT. The current co-crystal structures of ab13_14 and wild-type Tt-preQ1-RS aptamer domain in complex with 11 were refined to an Rfree of 0.190 (0.208) and 0.193 (0.235) at 1.57 and 2.80 Å resolution, respectively.

Cell culture and RNA isolation. MCF-7 cells were grown in 757 flasks in Dulbecco’s modified Eagle’s medium (DMEM) with high glucose supplemented with 10% fetal bovine serum and 1% antibiotics (streptomycin and penicillin) at 37 °C in 5% CO2 in humified incubator according to ATCC’s recommendations. On the day of the experiment (~80% confluency) the cells were harvested by trypsinization, washed and resuspended in 1x PBS. Total RNA was isolated from MCF-7 cells using the RNeasy mini kit (Qiagen, Cat #: 74104) with column DNase digestion according to the manufacturer protocol. The concentration of eluted RNA was measured using a nanodrop.

Photocrosslinking experiments with biotinylated probe 11 (Bio-11) in the presence of cell lysates. Cell lysates were extracted from the MCF-7 cells using RIPA buffer (Santa Cruz). Then 1 µL of mixed Bs-preQ1-RS aptamer (34 nucleotide length, see above) was added into the 3 µg of cell lysate and photocrosslinked to Bio-11 (250 µM) as described above. Next, equal volumes of the samples were subjected to 17% denaturing PAGE in the presence of 7 M urea. Meanwhile, pre-cut nylon membrane, whatman filter paper and sponge were soaked in 0.5X TBE buffer. Once the gel was run, the transfer sandwich was prepared, and the sandwich was placed into the BioRad Trans-Blot Cell and filled with pre-chilled 0.5x TBE. The Trans-Blot Cell was run at 80 V for 60 min at 4 °C with continuous stirring. Then the membrane was UV crosslinked at 1200 mJ for 3 min in a UVcrosslinker. Then the membrane was transferred into blocking buffer (ThermoFisher TM Pierce starting blocking buffer at room temperature for 1 h). Then the membrane was washed with 1 x Tris buffered saline. The blotting was conducted using the ECL kit (BioRad). The blot was then washed with 1 x Tris buffered saline. The blotting Luminol Reagent (Cell signaling Technology #7003) in the ChemiDoc imager.

Photocrosslinking experiments with 11 in the presence of total cellular RNA. Total RNA isolated from the MCF-7 cells was split into different tubes (5 µg RNA condition) and was diluted up to 50 µL total volume with riboswitch buffer (50 mM Tris (pH 7.5), 100 mM KCl and 2 mM MgCl2). One micromolar of the full-length Bs-preQ1-RS (70 nts) was added to the appropriate tubes during the dilution with riboswitch buffer. Before the addition, the RNA was folded as described above, by heating at 75 °C for 5 min and allowing to cool to RT over the course of 1 h. Next, 20 mM compounds 11 and 17 were prepared in DMSO and the compounds were diluted to a final concentration of 250 µM with the RNA in riboswitch buffer (5% final DMSO). Compounds 11 or 17 were mixed with RNA samples and incubated in the dark for 30 min to establish an equilibrium. After incubation, the solutions were spotted on a P84 glass filter plate for UV crosslinking. The control sample (Fig. 6, lane 7), that would not be irradiated, was left in the dark in the original tube. In the case of competition experiments, samples were prepared with the free PreQ1 ligand (500 µM and 1000 µM concentrations) titrated into the solution containing compound 11 and the total RNA to allow for competition before UV crosslinking. Samples in the plate were irradiated for 15 min at 365 nm using a SPECTROLINKER XL-1000 UV Crosslinker. After UV irradiation the sample mixtures were conjugated to TAMRA azide using the click labeling protocols described above. Finally, RNA was ethanol precipitated after TAMRA conjugation and samples were analyzed by PAGE.

RNA sequencing experiments. Total RNA (5 µg) isolated from MCF-7 cells was diluted up to 50 µL total volume with riboswitch buffer (50 mM Tris (pH 7.5), 100 mM KCl and 2 mM MgCl2). One micromolar of the full-length Bs-preQ1-RS (70 nts) was added into the tubes during the dilution with riboswitch buffer. Before the addition, Bs-preQ1-RS was folded as described above, by heating at 75 °C for 5 min and allowing to cool to RT over the course of 1 h. 200 nM stock solutions of compounds 11 and 17 were prepared in DMSO and the compounds were diluted to a final concentration of 250 µM with the RNA in riboswitch buffer (5% final concentration of DMSO). The compound and RNA mixtures were incubated at room temperature for 1 h and samples were UV crosslinked as described above. The modified RNA was labeled with biotin-PEG3-azide using the click protocol described above. After labeling the RNA was subjected to ethanol precipitation to remove excess biotin-PEG3-azide. The RNA pellets were dissolved in 50 µL of incubation buffer (50 mM Tris Ph 7.5, 100 mM KCl, 0.01% Tween-20) and 50 µL of streptavidin linked magnetic beads (NEB) was added to each tube. After 30 min, the beads were pulled down on a magnetic tower and washed three times with 500 µL incubation buffer. Finally, 50 µL of elution buffer (95% formamide/H2O, 10 mM EDTA) was added to each tube and the tubes were placed on a 95 °C heat block for 4 min. The supernatant containing the eluted RNA was removed on the magnetic tower and the solution was diluted with RNase free water and ethanol precipitated overnight. The concentration of the reconstituted RNA was measured using a Qubit RNA HS Assay Kit (Invitrogen). Two-hundred nanograms of each sample was subsequently used to prepare sequencing libraries with the QiaSeq Stranded Total RNA Library Kit (Qiagen). Completed libraries were submitted to Novogene for sequencing on a HiSeq 2500 (2 x 150 reads). Four replicates of each treatment condition were prepared and sequenced. We followed the same protocol as described above but without addition of the Bs-preQ1-RS aptamer to a crosslinking to other RNAs in MCF7 total RNA. All sequencing experiments were performed with four independent replicates for each condition (n = 4). In experiments without aptamer spike-in one replicate of 17 failed QC and was not included in this analysis.

Sequencing data analysis. For the analysis of these pulldown experiments, a reference was built by using the longest isoform for every human RefSeq transcript, plus the sequence of the Bs-preQ1-RS full-length aptamer (see above). Reads were aligned to this reference using Bowtie2 and the “--very-sensitive-local” preset. Resulting BAM files were then processed to TSV count files, containing the number of reads mapping to each transcript. The DESeq2 package adjusted-p-value < 0.05 was used to identify transcripts significantly enriched by compound 11 in the pulldown with respect to negative control. For the sequencing experiments without the aptamer spiked-in, the volcano plot generated by the DESeq2 package was used to identify potential hits that were significantly enriched (FC > 1.5-fold, adjusted-p-value < 0.05 when comparing 11 treated samples to 17 treated samples).

Quantitative reverse transcription PCR (RT-qPCR). Samples for these experiments were prepared using the same protocol as described above for the RNA-sequencing experiments (without spiking Bs-preQ1-RS aptamer). However, in this case extra samples were prepared with the free PreQ1 ligand (25 µM and 100 µM concentrations) titrated into the solution containing compound 11 and the total RNA to allow for competition before UV crosslinking. RNA concentration was determined using a Qubit RNA HS Assay Kit (Invitrogen). For the sample treatment condition, 200 ng of RNA was used for cDNA synthesis with qScript cDNA SuperMix (Quintana Biosciences) according to the manufacturer’s protocol. The 20 µL
reactions were incubated in a thermocycler (Bio-rad) for 5 min at 22 °C, 30 min at 42 °C, 5 min at 85 °C, then held at 4 °C. Then 20 ng of cdNA was subject to qPCR using a Perfecta SYBR Green Super Mix (Quanta Bioscience) on an Eppendorf Mastercycler RealPlex2 in the presence of the appropriate set of primers. The reactions were incubated at 95 °C for 10 min, followed by 40 cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 20 s. Threshold cycles (Ct) is defined as the first cycle that showed a detectable increase in fluorescence due to the formation of PCR products and was used to determine the template amount in each sample. The relative fold change in expression was measured using the Livak method and were normalized to the relative level of RNAs of the control experiment (compound 17). For example, to calculate the ΔΔCt between each gene of interest and the average of the control samples: ΔCt(ΔCt) = Ctgapdh) – Ctgapdh (treatment or knockdown) – ΔCt(ΔCt) control; fold change = 2−ΔΔCt. The primers used in these experiments are shown in the Supplementary Information Table S2.

**Data availability**

Atomic coordinates and the structure factors of the co-crystal structures of ab13.14 and the wild-type TrpE preQ1, 85 aptamer domain in complex with 11 have been deposited in the Protein Data Bank, under the accession codes 7E9E and 7E9L, respectively. Previously published crystals 3Q50, 6E1W and 3B1V were accessed from the PDB and used for reference. Raw fastq files along with processed alignment files and counts files from these sequencing experiments have been deposited on the NCBI's gene expression omnibus (GEO) database.

**Received: 12 March 2021; Accepted: 26 August 2021; Published online: 06 October 2021**

**References**

1. Breaker, R. R. Riboswitches and the RNA world. Cold Spring Harb. Perspect. Biol. 4, 000356 (2012).
2. Serganov, A. & Nudler, E. A decade of riboswitches. Cell 152, 17–24 (2013).
3. Roth, A. & Breaker, R. R. The structural and functional diversity of metabolite-binding riboswitches. Annu. Rev. Biochem 78, 305–34 (2009).
4. Howe, J. A. et al. Selective small-molecule inhibition of an RNA structural element. Nature 526, 672–7 (2015).
5. Warner, K. D. et al. Validating fragment-based drug discovery for biological RNAs: lead fragments bind and remodel the TPP riboswitch specifically. Nat. Chem. 21, 591–597 (2014).
6. Blount, K. F., Wang, X. J., Lim, J., Sudarsan, N. & Breaker, R. R. Antibacterial GSE168353. The source data underlying Fig. 2A–E, 4C–D, 5A–B, 7B–D and Supplementary Figs. S1, S2, S5, S6A–B, S7A–D and S10 are provided as a Source Data file. All other relevant data are available from the authors upon request. Source data are provided with this paper.

14. Ayele, T. M., Loya, T., Valdez-Simon, A. N., Bassell, G. J. & Heemstra, J. M. Imaging and Tracking mRNA in Live Mammalian Cells Via Fluorescent Photobinding Labeling. (Cold Spring Harbor Laboratory, 2020).
15. Zhou, C. Y., Alexander, S. C. & Devaraj, N. K. Fluorescent turn-on probes for wash-free mRNA imaging. Chem. Sci. 8, 7169–7173 (2017).
16. Wang, J., Schulz, P. G. & Johnson, K. A. Mechanistic studies of a small-molecule modulator of SMN2 splicing. Proc. Natl Acad. Sci. USA 115, 20160–201612 (2018).
17. Mortison, J. D. et al. Tetracyclines modify translation by targeting key mRNA substructures. Cell Chem. Biol. 25, 1506–1518.e13 (2018).
18. Suresh, B. M. et al. A general fragment-based approach to identify and optimize bioactive ligands targeting RNA. Proc. Natl Acad. Sci. USA 117, 33197–33203 (2020).
19. Mukherjee, H. et al. PEARL-seq: a photoaffinity platform for the analysis of small molecule-RNA interactions. ACS Chem. Biol. 15, 2374–2381 (2020).
20. Hargrove, A. E. Small molecule-RNA targeting: starting with the fundamentals. Chem. Commun. (Camb.) 56, 14744–14756 (2020).
21. Connelly, C. M., Moom, M. H. & Schneekloth, J. S. The emerging role of RNA as a therapeutic target for small molecules. Cell Chem. Biol. 23, 1077–1090 (2016).
22. Meyer, S. M. et al. Small molecule recognition of disease-relevant RNA structures. Chem. Sci. Rev. 49, 7167–7199 (2020).
23. Baker, J. L. et al. Widespread genetic switches and toxicity resistance proteins for fluorokide. Science 335, 233–235 (2012).
24. Suath, B., Ishi, K. & Tabor, T. Microsecond folding of preQ1 riboswitch and its biological significance revealed by two-dimensional fluorescence lifetime correlation spectroscopy. J. Am. Chem. Soc. 143, 7988–7998 (2021).
25. Klein, D. J., Edwards, T. E. & Ferré-D’Amaré, A. R. Co-crytal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. Nat. Struct. Mol. Biol. 16, 343–344 (2009).
26. Connelly, C. M. et al. Synthetic ligands for PreQ1 Nat. Commun. 10, 15019 (2019).
27. Jenkins, J. L., Kruclinska, J., McCarty, R. M., Bandarian, V. & Wedekind, J. Comparison of a preQ1 riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. J. Biol. Chem. 286, 24626–37 (2011).
28. Schroeder, G. M. et al. Analysis of a preQ1 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).
29. Santner, T., Rieder, U., Kreutz, C. & Micura, R. Pseudoknot reorganization of the preQ1 class I riboswitch. J. Am. Chem. Soc. 139, 11928–31 (2012).
30. Eichhorn, C. D., Kang, M. & Feigon, J. Structure and function of preQ1. Biochim. Biophys. Acta 1839, 939–950 (2014).
31. Dutta, D. & Wedekind, J. E. Nucleobase mutants of a bacterial preQ1 J. Biol. Chem. 295, 2555–2567 (2020).
32. McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E. & Breaker, R. R. Riboswitch diversity and distribution. RNA 23, 995–1011 (2017).
33. Guan, L. & Disney, M. D. C. D. Pseudoknot preorganization of small-molecule RNA complex formation enables cellular profiling of small-molecule-RNA interactions. Angew. Chem. Int. Ed. 52, 10010–13 (2013).
34. Behm-Ansamt, I., Helm, M. & Motorin, Y. Use of specific chemical reagents for detection of modified nucleotides in RNA. J. Nucleic Acids 2011, 1–17 (2011).
35. Neuner, E., Fremer, M., Lusser, A. & Micura, R. Superior cellular activities of azido- over amino-functionalized ligands for engineered preQ1 riboswitches in E.coli. RNA Biol. 15, 1376–1383 (2018).
36. Li, Z. et al. Design and synthesis of minimalist terminal alkylane-containing diazirine photo-crosslinkers and their incorporation into kinase inhibitors for cell- and tissue-based proteome profiling. Angew. Chem. Int. Ed. 52, 8551–8556 (2013).
37. Gao, J., Mfuh, A., Amako, Y. & Woo, C. M. Small molecule interactome mapping by photoaffinity labeling reveals binding site hotspots for the NSAI Ds. J. Am. Chem. Soc. 140, 4259–4268 (2018).
38. Brunnner, J., Sein, H. & Richards, F. M. 3-Fluorouracil-3-phenylazidirine. A new carbene generating group for photoafflunit labeling. J. Biol. Chem. 285, 3313–8 (1980).
39. Dubinsky, L., Krom, B. P. & Meijler, M. Diazirine based photoaffinity labeling. Bioorg. Med. Chem. 20, 554–70 (2012).
40. Artsimovich, I. & Henkin, T. M. In vitro approaches to analysis of transcription termination. Methods 47, 37–43 (2004).
41. Williams, A. S. & Marzluff, W. F. The sequence of the stem and flanking sequences at the Y’end of histone mRNA are critical determinants for the binding of the stem-loop binding protein. Nucleic Acids Res. 23, 654–662 (1995).
42. Dávila López, M. & Samue1sson, T. Early evolution of histone mRNA 3’ end processing. RNA 14, 1–10 (2008).
43. Maida, Y. et al. An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature 461, 230–235 (2009).
47. Su, Z. et al. Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in C9FTD/ALS. *Neuron* **84**, 239 (2014).

48. Chee, G. L., Yalowich, J. C., Bodner, A., Wu, X. & Hasinoff, B. B. A diazirine-based photoaffinity etoposide probe for labeling topoisomerase II. *Bioorg. Med. Chem.* **18**, 830–830 (2010).

49. West, A. et al. Labeling preferences of diazirines with protein biomolecules. *J. Am. Chem. Soc.* **143**, 6691–6700 (2020).

50. Genome, M. B. et al. Anchoring A. Q. The queuosine microenvironment: charting a course from microbe to man. *Nutrients* **7**, 2897–929 (2015).

51. Vinayak, M. & Pathak, C. Queuosine modification of tRNA: its divergent role in cellular machinery. *Biosci. Rep.* **30**, 135–48 (2009).

52. Tuorto, F. et al. Queuosine-modified tRNAs confer nutritional control of protein translation. *EMBO J.* **37**, e70777 (2018).

53. Xu, D. et al. PreQ0 base, an unusual metabolite with anti-cancer activity from Streptomyces qinglanensis. *Anticancer Agents Med. Chem.* **15**, 285–90 (2015).

54. Zhang, J. et al. tRNA Queuosine modification enzyme modulates the growth and microbiome recruitment to breast tumors. *Cancers (Basel)* **12**, 628 (2020).

55. Artandi, S. E. & Depinho, R. A. Telomeres and telomerase in cancer. *Carcinogenesis* **31**, 9–18 (2010).

56. Baena-Del Valle, J. A. et al. MYC drives overexpression of telomerase RNA (hTERT/TERC) in prostate cancer. *J. Pathol.* **244**, 11–24 (2018).

57. Cherkaoui Jaouad, I. et al. Novel mutation and structural RNA analysis of the noncoding RNase MRP gene in cartilage-hair hypoplasia. *Mol. Syndromol.* **6**, 77–82 (2015).

58. Zhang, Q., Kim, N.-K. & Feigon, J. Architecture of human telomerase RNA. *Proc. Natl Acad. Sci. USA* **108**, 20325–20332 (2011).

59. Jafri, M. A., Ansari, S. A., Alqahtani, M. H. & Shay, J. W. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Nat. Rev. Cancer* **16**, 221 (2016).

60. Klepper, F., Polborn, K. & Carell, T. Robust synthesis and crystal-structure analysis of 7-cyano-7-deazaguanine (PreQ0 Base) and 7-(aminomethyl)-7-deazaguanine (PreQ2 Base). *Helv. Chim. Acta* **88**, 2610–2616 (2005).

61. Waltzthofer, T. et al. xTract: software for characterizing conformational changes of protein complexes by quantitative cross-linking mass spectrometry. *Nat. Methods* **12**, 1185–1190 (2015).

62. Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A. Hitchhiker's guide to click-chemistry with nucleic acids. *Chem. Rev.* **121**, 7122–7154 (2021).

63. Hirata, K. et al. ZOO: an automatic data-collection system for high-throughput structure analysis in protein microcrystallography. *Acta Crystallogr. D. Sci. Biol. Crystallogr.* **75**, 138–150 (2019).

64. Yamashita, K., Hirata, K. & Yamamoto, M. KAMO: towards automated data processing for microcrystals. *Acta Crystallogr. D. Sci. Biol. Crystallogr.* **74**, 441–449 (2018).

65. Kabsch, W. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–32 (2010).

66. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

67. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–21 (2010).

68. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–32 (2004).

69. Langbead, R. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–9 (2012).

70. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

71. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–8 (2001).

Acknowledgements
The synchrotron-radiation experiments at SPring-8 (Hyogo, Japan) were performed on beamlines BL32XU and BL45XU with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) under proposal numbers 2019B2580 and 2020A2568. We also thank the beam-line staff at BL-1A and BL-17A of the Photon Factory (Ibaraki, Japan) for technical assistance during data collection. This research was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. Project numbers Z01 BC011585 07 (PI, J. S. Schneekloth, Jr). This work was supported in part by the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Numbers 16KK0166 and 20K21281) and a grant from The Uehara Memorial Foundation to T.N. D.I. was supported by the Dutch Research Council (Netherlands Organisation for Scientific Research, NWO) as part of the research programme NWO Open Competitive ENW-XX (project number OCENW.XS.044), and by the Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen.

Author contributions
S.B. performed photocrosslinking experiments, qPCR, and helped write and edit the paper. C.R. performed photocrosslinking and Chem-CLIP experiments, and helped write and edit the paper. D.D.B. performed chemical synthesis and characterization and mass spectral analyses. C.C.L. and J.A.R. performed mass spectral analyses. K.Y., P.J.H., and D.I. performed bioinformatic analyses. T.N. performed X-ray crystallographic analyses and transcription termination assays. J.S.S. conceived the project and helped write and edit the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-25973-x.

Correspondence and requests for materials should be addressed to John S. Schneekloth Jr.

Peer review information Nature Communications thanks Joseph Wedekind and other anonymous, reviewers for their contributions to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

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 2021