A rare bacterial RNA motif is implicated in the regulation of the purF gene whose encoded enzyme synthesizes phosphoribosylamine

SARAH N. MALKOWSKI,1 RUBEN M. ATILHO,2 ETIENNE B. GREENLEE,3 CHRISTINA E. WEINBERG,3,5 and RONALD R. BREAKER2,3,4

1Department of Chemistry, 2Department of Molecular Biophysics and Biochemistry, 3Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8103, USA
4Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520-8103, USA

ABSTRACT

The Fibro-purF motif is a putative structured noncoding RNA domain that was discovered previously in species of Fibrobacter by using comparative sequence analysis methods. An updated bioinformatics search yielded a total of only 30 unique-sequence representatives, exclusively found upstream of the purF gene that codes for the enzyme amidophosphoribosyltransferase. This enzyme synthesizes the compound 5-phospho-D-ribosylamine (PRA), which is the first committed step in purine biosynthesis. The consensus model for Fibro-purF motif RNAs includes a predicted three-stem junction that carries numerous conserved nucleotide positions within the regions joining the stems. This architecture appears to be of sufficient size and complexity for the formation of the ligand-binding aptamer portion of a riboswitch. In this study, we conducted biochemical analyses of a representative Fibro-purF motif RNA to confirm that the RNA generally folds according to the predicted consensus model. However, due to the instability of PRA, binding of this ligand candidate by the RNA could not be directly assessed. Genetic analyses were used to demonstrate that Fibro-purF motif RNAs regulate gene expression in accordance with predicted PRA concentrations. These findings indicate that Fibro-purF motif RNAs are genetic regulation elements that likely suppress PRA biosynthesis when sufficient levels of this purine precursor are present.

Keywords: aptamer; gene control; noncoding RNA; PRA; purine; ribose

INTRODUCTION

Riboswitches are structured noncoding RNAs that regulate gene expression by binding to small ligands such as metabolites, signaling molecules, or elemental ions (Roth and Breaker 2009; Serganov and Nudler 2013; Sherwood and Henkin 2016). Nearly 50 riboswitch classes have been discovered and validated to date (McCown et al. 2017), and many additional “orphan” riboswitch candidates exist whose ligand specificities have yet to be established (Weinberg et al. 2017a; Greenlee et al. 2018). Well represented among the known riboswitch classes are those that sense purines or their various derivatives to regulate purine transport, salvage, and de novo biosynthetic genes among many other processes. The ligands for these various metabolite-sensing RNAs include guanine (Mandal et al. 2003), adenine (Mandal and Breaker 2004), 2′-deoxyguanosine (Kim et al. 2007; Weinberg et al. 2017b), and ADP (Sherlock et al. 2019). Likewise, purine biosynthetic precursors such as PRPP (phosphoribosylpyrophosphate) (Sherlock et al. 2018a) and ZTP (5-aminoimidazole-4-carboxamide riboside 5′-triphosphate) (Kim et al. 2015) have corresponding riboswitch classes, as do purine nucleotide derivatives such as the signaling molecules c-di-GMP (Sudarsan et al. 2008; Lee et al. 2010), c-di-AMP (Nelson et al. 2013), c-AMP-GMP (Kellenberger et al. 2015; Nelson et al. 2015), and ppGpp (Sherlock et al. 2018b). Recently, a riboswitch class that senses the...
oxidized purine derivatives xanthine and uric acid has been reported (Yu and Breaker, 2020). Finally, there are many riboswitch classes that sense purine-containing enzyme cofactors (McCown et al. 2017), among which is a riboswitch for NAD+ that at a minimum selectively recognizes the ADP moiety of this ubiquitous biomolecule (Malkowski et al. 2019; Huang et al. 2020).

Given the striking abundance and diversity of riboswitches for purines and purine-related compounds, we were intrigued by the discovery of a novel, but rare RNA motif located upstream of the purine biosynthetic gene purF in some species of Fibrobacter (Weinberg et al. 2017a). Fibro-purF motif RNAs exhibit an overall secondary structure that is vaguely like the three-stem junction architecture formed by previously described riboswitch aptamers for guanine (Mandal et al. 2003), adenine (Mandal and Breaker 2004), and 2′-deoxyguanosine (Kim et al. 2007; Weinberg et al. 2017b). However, RNAs conforming to this new motif are distinct in both their nucleotide sequences and gene associations.

Here we report that representatives of the Fibro-purF motif appear to be involved in regulating gene expression most likely in response to changing concentrations of phosphoribosylamine (PRA), which is the product of the first committed step in the pathway for de novo purine biosynthesis. Genetic analyses demonstrate that expression of a reporter gene fused downstream from a Fibro-purF motif RNA is repressed in response to high concentrations of PRA. Unfortunately, the poor chemical stability of PRA precluded our attempts to assess direct binding of this compound by Fibro-purF RNAs. Thus, we cannot yet be certain that the conserved motif functions as the aptamer component of a metabolite-binding riboswitch.

RESULTS AND DISCUSSION

Bioinformatic analyses suggest that Fibro-purF motif RNAs are distinct regulatory elements associated with purine metabolism

Initially, only 13 distinct-sequence representatives of the Fibro-purF motif were reported (Weinberg et al. 2017a), in which nearly all were present in a cow rumen metagenome sequence database. We conducted additional bioinformatic searches using an expanded collection of bacterial genomic sequence databases (See Materials and Methods) to identify a total of 30 nonredundant representatives (Fig. 1). However, all examples are exclusively present in the phylum Fibrobacter and metagenomic data sets.

An updated consensus sequence and structural model was created based on the alignment of all 30 representatives (Fig. 2A). The additional examples revealed that some of the nucleotides originally annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations. Even with the increase in bioinformatic data, the narrow phylogenetic distribution and limited number of representatives make it difficult to assess the importance of the nearly 30 positions that are annotated as highly conserved (<97%) appear to tolerate mutations.

FIGURE 1. Sequence alignment of 30 distinct representatives of the Fibro-purF motif. Shaded nucleotides identify base complementarity corresponding to predicted hairpins P1, P2, and P3. Consensus row letters in red, black, and gray identify nucleotides present in >97%, >90%, or >75% of the representatives, respectively.
the known PRPP riboswitch class are usually found associated with purF genes only when they are part of the entire purine biosynthetic operon. In contrast, Fibro-purF motif representatives are associated exclusively with the purF gene, and only when it is not part of an operon with other genes related to purine biosynthesis. This gene association pattern suggests that a compound other than PRPP might be the ligand. Similarly, it seems unlikely that an end-product of the purine biosynthetic pathway would be used to regulate purF gene expression, but not the other genes in this process as suggested by the Fibro-purF motif distribution. These considerations caused us to favor the tentative hypothesis that the RNA senses the direct end-product of the PurF enzyme, PRA.

Given the various technical challenges noted above, we considered employing genetic experiments that could demonstrate gene control function by Fibro-purF motif RNAs and help reveal the ligand that triggers changes in expression. Gene expression driven by Fibro-purF motif RNAs should be repressed if the ligand is PRA. In contrast, expression would likely be activated if the ligand is the PurF enzyme substrate PRPP. A short distance following each Fibro-purF motif representative is the ribosome binding site (RBS) and AUG translation start codon for the purF open reading frame (ORF) (Fig. 1). Thus, we predict Fibro-purF motif RNAs could use a mechanism for gene regulation involving the control of translation initiation. However, again the limited number of examples makes it difficult to confidently predict a likely anti-RBS structure or the directionality of the putative genetic switch. Genetic assays fusing a Fibro-purF motif RNA to a reporter gene were subsequently used to further investigate the function of this rare noncoding RNA motif.

**Fibro-purF motif RNAs demonstrate genetic “OFF” switch character**

Given the distinctive structural and genetic context characteristics noted above, we first sought to determine if Fibro-purF motif RNAs regulate gene expression. A representative Fibro-purF motif from *Fibrobacter succinogenes*, along with its native promoter, was fused upstream of a β-galactosidase (lacZ) gene and transformed into *Escherichia coli* (Fig. 3A). Reporter constructs containing the wild-type (WT) Fibro-purF motif RNA were examined for evidence of differential gene expression in rich and minimal media growth conditions. Cells containing WT Fibro-purF RNA exhibited essentially no reporter gene expression when grown in rich medium (Fig. 3B). Cells can import purines from rich media, and thus PRA should be abundant because cells are not using it for de novo purine biosynthesis. In minimal medium, when de novo purine biosynthesis is active, this same reporter strain exhibits elevated gene expression compared to cells grown in rich medium. PRA...
should be depleted in these cells due to its use in making purines. Thus, these reporter gene assays are consistent with the hypothesis that Fibro-purF motif RNAs function as genetic “OFF” switches, and that the regulatory signal is abundant when cells are grown in rich medium.

It is important to note that several constructs containing mutations of highly conserved nucleotides (Fig. 3A: mutations G18A, G38A and C39A, or G43A) were evaluated, and they exhibited gene expression levels similar to WT in rich medium (data not shown). This data calls into question the hypothesis that the Fibro-purF motif RNAs function as an aptamer. However, as discussed above, the limited number of distinct representatives and the small evolutionary distance between these examples makes it challenging to gauge the importance of the apparent highly conserved nucleotides presented in the consensus model (Fig. 2A).

Such mutational studies could become more focused on key nucleotides if the phylogenetic data are expanded in the future.

Although our leading hypothesis is that Fibro-purF motif RNAs operate as riboswitches, it is possible that this RNA class functions as another type of genetic control element, such as a protein-binding RNA. However, the parent species for this the Fibro-purF motif construct is in a different phylum (Fibrobacteres) from E. coli (Proteobacteria). Thus, by examining the function of the Fibro-purF motif representative from F. succinogenes in the surrogate organism E. coli, we expect that any possible protein-based gene control factor necessary in the native host would not be present, or at least would be unlikely to bind the RNA. Similarly, a protein factor that might be required to control gene expression via the native F. succinogenes promoter is expected to be absent in E. coli, or poorly match to the sequence of the foreign promoter region. Also, any possible sRNA factor that might be used to regulate purF gene expression is unlikely to be present and functionally equivalent in E. coli. Thus, the more likely explanation for these specific findings is that Fibro-purF motif RNAs are riboswitches that sense a metabolite to regulate the adjacent ORF.

PRA levels affect reporter gene expression mediated by Fibro-purF motif RNAs

To seek additional evidence regarding the natural signaling molecule relevant to the Fibro-purF motif, E. coli strains carrying knockouts (KOs) of specific genes in the purine biosynthetic pathway were evaluated. A KO strain that maintains a reduced concentration of the ligand is expected to exhibit increased reporter gene expression if the regulatory system functions as a genetic “OFF” switch. In contrast, a KO strain that can still produce the ligand should exhibit reduced reporter gene expression.

Strains with disruptions to their de novo purine biosynthetic pathway require an exogenous purine source, such as adenine, to grow on minimal media. Therefore, we established agar-diffusion assays wherein the purine...
biosynthesis gene KO strains carrying the WT Fibro-\(\text{purF}\) reporter construct (Fig. 3A) were grown on minimal media with a filter disk spotted with adenine. Two KO strains were examined that disrupt purine biosynthesis either before the production of PRA (\(\Delta\text{purF}\)) or the step occurring immediately after its production (\(\Delta\text{purD}\)) (Fig. 2A). These KO strains were chosen because they are expected to most directly alter the in vivo concentration of PRA.

Agar diffusion assays with the \(E.\ coli\ \Delta\text{purF}\) strain yielded cell growth only near the adenine-infused filter disk, wherein a halo of high reporter gene expression was observed (Fig. 3C, top). The increased intensity at the interface between robust cell growth and no growth is expected if cells deplete any residual PRA present (see further discussion below) via activation of de novo purine biosynthesis at the limit of adenine diffusion on the plate. Similar results with analogous agar-diffusion assays have been observed for known riboswitch classes, such as those for ZTP (Kim et al. 2015), HMP-PP (Atilho et al. 2019; Stav et al. 2019), and NAD\(^+\) (Malkowski et al. 2019). In contrast, the analogous \(\Delta\text{purD}\) strain resulted in a similar cell growth pattern, but these cells exhibit almost complete suppression of reporter gene expression (Fig. 3C, bottom). The PurD protein, phosphoribosylamine-glycine ligase, converts PRA to \(N^1\)-(5-phospho-D-ribosyl)glycinamide, and therefore PRA should accumulate to high levels in the \(\Delta\text{purD}\) strain. These results are consistent with the hypothesis that PRA is the compound that causes suppression of reporter gene expression.

In \(\Delta\text{purF}\) cells, we do not anticipate that PRA is completely absent. PRA can be formed by the spontaneous coupling of ammonia and PRPP (Nierlich and Magasanik 1965). Also, substantial amounts presumably can be generated by two alternative enzymatic routes, one beginning with threonine using enzymes from the isoleucine and tryptophan biosynthetic pathways (Bazurto et al. 2016), and another by diverting a histidine biosynthetic intermediate (Koenigsknecht and Downs 2010; Koenigsknecht et al. 2012). One or more of these alternative routes to PRA might explain the fact that \(\Delta\text{purF}\) cells nearest to the adenine-infused disk exhibit somewhat lower reporter gene expression than those in the halo of reporter gene expression (Fig. 3C, top).

To provide additional evidence that the reporter gene expression characteristics observed are the result of the changes in PRA levels rather than more general cellular distress, expression was monitored in three additional, arbitrarily chosen, \(E.\ coli\) strains that are auxotrophic for the amino acids lysine, methionine, or histidine. With the \(\Delta\text{lysA}\) (lysine auxotroph) and \(\Delta\text{metE}\) (methionine auxotroph) strains, constant expression of the Fibro-\(\text{purF}\) reporter construct is observed in cells whose growth is supported by the addition of the missing amino acids on the filter disk (Fig. 3D). These cells presumably are producing PRA to support de novo purine biosynthesis, which is necessary for their growth on minimal media. However, the continuous use of PRA for purine biosynthesis should deplete its levels, thereby permitting robust reporter gene expression. Importantly, these strains with gene KOs unrelated to PRA metabolism do not exhibit a reporter gene expression halo like that observed for the \(\Delta\text{purF}\) strain (Fig. 3C).

Intriguingly, assays conducted with the \(\Delta\text{hisB}\) strain reveal that cells exhibit a more complex gene expression pattern, presumably due to the relationship between histidine and purine biosynthesis (Figs. 2B, 4A). Our interpretation of the different reporter expression characteristics (Fig. 4B) is described as follows. At high histidine concentration near the filter disk, cells are expected to use PRA to make purines, and reporter gene expression reaches an intermediate level similar to the \(\Delta\text{lysA}\) and \(\Delta\text{metE}\) strains. However, histidine concentrations decrease when moving away from the filter disk, thereby eventually requiring cells to attempt to produce histidine. Normally, cells generate
histidine by using a biosynthetic pathway initiated by coupling PRPP to the purine ring of ATP (Fig. 4A; Koenigsknecht et al. 2012). This pathway should diminish the level of PRPP that otherwise could be used to sustain the production of PRA required for purine biosynthesis. Thus, histidine starvation should first cause PRA depletion and an intensifying of reporter gene expression, but then the reporter levels should diminish, not by reduced expression from each reporter construct, but by simply a reduction in the total number of cells. This interpretation of the data is consistent with our hypothesis that Fibro-purF motif RNAs down-regulate purine biosynthesis in response to PRA.

A representative Fibro-purF motif RNA largely conforms to the predicted secondary structure

For most riboswitch validation studies, we use biochemical assays to demonstrate selective ligand binding by the aptamer domain. However, in some instances, there are technical limitations that preclude such demonstrations. For example, the proposed riboswitches for molybdenum cofactor and tungsten cofactor were demonstrated using only genetic experiments due to the chemical instability of these redox-active coenzymes (Regulski et al. 2008). Similarly, PRA cannot readily be tested for ligand binding because it is not commercially or otherwise available likely due to its short half-life (Nierlich and Magasanik 1965), which is measured in seconds under biologically relevant conditions (Chen et al. 1987; Schindel et al. 1988). In an attempt to assess ligand binding, we contracted the synthesis of the carbocyclic analog (C-PRA) in which the furanose ring oxygen of PRA was replaced with a methylene group (Gebauer et al. 2012). In-line probing assays (Soukup and Breaker 1999; Regulski and Breaker 2008) were conducted to evaluate both the secondary structure model and the ability of C-PRA to bring about a structural change indicative of ligand binding. In-line probing exploits the relative differences in phosphodiester linkage stability based on the local structure of the RNA chain (Soukup and Breaker 1999) and can reveal details regarding the structure and binding characteristics of RNA aptamers (Soukup et al. 2001).

In-line probing assays were conducted using a 5′ 32P-labeled representative Fibro-purF RNA sequence from Fibrobacter sp. UWB11, which included 82 nt spanning the conserved motif through the start codon of the adjoining purF gene. This RNA construct, termed 82 Fibro-purF (Fig. 5A), exhibits a pattern of spontaneous RNA breakdown (Fig. 5B) that is largely consistent with the proposed secondary structure model. For example, the nucleotides proposed to form the base-pairs of stems P1 and P2 undergo little spontaneous breakdown. However, there is some ambiguity in the region originally proposed to form the P3 stem. This stem appears to be poorly formed, and the banding pattern from the in-line probing data suggests that the nucleotides at positions 41 through 44 might be involved in a different interaction.

Whereas genetic analyses show strong support for the hypothesis that PRA is the natural ligand for Fibro-purF motif RNAs, evidence for direct binding in biochemical

![In-line probing analysis of a Fibro-purF motif RNA is consistent with the predominant secondary structure features.](image-url)
analyses has not been observed. Specifically, we do not observe evidence for robust binding by the C-PRA analog at concentrations as high as 1 mM (Fig. 5B). Similar in-line probing results are observed with three constructs derived from other species (data not shown), and these findings suggest that the constructs are properly folding. Furthermore, the banding pattern from in-line probing assays remains essentially unchanged regardless of the presence (1 mM) or absence of other ligand candidates (Fig. 5B) such as ribose, ribose 5-phosphate, or over 60 molecules related to purine metabolism as reported previously (Weinberg et al. 2017a). Thus, direct evidence for ligand binding remains elusive.

Although C-PRA is chemically similar to the cognate ligand, a putative PRA receptor might discriminate against C-PRA for several reasons. For example, the electronic character of the ring will be different than that for PRA, which could preclude the formation of a hydrogen bond. Also, the structure of the original ribose ring could be altered sufficiently by the replacement of the oxygen with a tetrahedral methylene group such that the binding site is a poor fit for the shape of the ring. Previous research has demonstrated that protein enzymes can utilize C-PRA, albeit with lower efficiency than PRA (Caperelli and Liu 1992). Therefore, the differences between PRA and C-PRA affect the function of protein enzymes, and similarly could disrupt the function of an RNA aptamer.

Concluding remarks

The bioinformatic and genetic data presented herein are largely consistent with our hypothesis that Fibro-purF motif RNAs function as genetic control elements for the purine biosynthetic intermediate PRA. When purines are abundant, PRA levels are expected to be elevated, leading to suppression of gene expression by binding to the aptamer domain represented by the Fibro-purF motif. Unfortunately, the fact that PRA is highly unstable makes it impractical to use conventional biochemical assays to confirm direct binding activity. Moreover, without stable analogs that can dock to the putative aptamer, further biophysical studies to create atomic-resolution structural models also could be challenging. Perhaps a more comprehensive screening of analogs by using additional constructs might provide a Fibro-purF motif RNA docked to a ligand that can be used for in vitro studies.

This analysis of the Fibro-purF riboswitch candidate highlights challenges associated with studying rare noncoding RNA motifs. Originally reported with just 13 examples (Weinberg et al. 2017a), and now expanded to 30, Fibro-purF motif RNAs would constitute one of the rarest riboswitch classes. The consistent association between this motif and the purF gene strongly implicates PRA as the ligand. Indeed, genetic knockouts in the purine biosynthetic pathway bracketing PRA reveal that gene expression is repressed when PRA levels are expected to be high (Fig. 3C). These findings, as well as the distinctive reporter gene expression results observed with the ΔhisB strain (Fig. 3D), implicate PRA as the ligand responsible for triggering the genetic switch. However, we cannot rule out the possibility that some other ligand or some other mechanism is responsible for gene control.

It has previously been proposed that a large number (perhaps many thousands) of additional rare riboswitch classes await discovery (Ames and Breaker 2010; McCown et al. 2017). Uncovering a rare riboswitch candidate that alters gene expression in response to changing PRA concentrations highlights the tremendous opportunities for bacterial noncoding RNA research. Some of these rare classes will undoubtedly be found as variants of known riboswitch classes (Weinberg et al. 2017b), and so hypotheses regarding their gene control function will be less doubtful. However, it is expected that many distinct, but exceedingly rare riboswitch classes also exist, and some of these will prove to be difficult to experimentally validate. The Fibro-purF RNA riboswitch candidate conforms to this latter subset, and therefore requires additional evidence to determine its possible riboswitch function.

Many bacterial species use protein factors, such as PurR (Cho et al. 2011), or riboswitches specific for guanine (Mandal and Breaker 2004), or PRPP (Sherlock et al. 2018a) to regulate purine transport and biosynthesis. The Fibrobacter species that use Fibro-purF motif domains to regulate this first committed step for purine biosynthesis therefore are unusual. Indeed, the rarity of this motif might be due in part to the instability of PRA, which makes an unusual choice for a regulatory ligand. Despite the narrow distribution of the Fibro-purF motif, RNA structures with structural similarities might have existed in the RNA World where they allowed ancient organisms to sense and respond to this essential purine biosynthetic precursor.

MATERIALS AND METHODS

Compounds, DNA oligonucleotides, and bacterial strains

Compounds were purchased from Sigma-Aldrich, except for C-PRA [IUPAC: 1, 2-Cyclopentanediol, 3-amino-5-[(phosphonoxy)methyl], ammonium salt (1:2), (1R,2S,5R), which was synthesized (Gebauer et al. 2012), purified and validated by Alfa Chemistry. Radiolabeled [γ-32P]-ATP was purchased from PerkinElmer. All enzymes were purchased from New England BioLabs unless otherwise noted. DNA oligonucleotides were purchased from Sigma-Aldrich. The Coli Genetic Stock Center (CGSC) at Yale University supplied E. coli strain BW25113 and the corresponding mutant strains: ΔpurF (CGSC #9854), ΔpurD (CGSC #10855), ΔlysA (CGSC #10193), ΔhisB (CGSC #9650), and ΔmetE (CGSC #10758).
Bioinformatic analyses

Additional representatives of Fibro-
\textit{purF} motif RNAs were identified using Infernal 1.1 (Nawrocki and Eddy 2013) from RefSeq (O’Leary et al. 2016) version 96 and several metagenomic databases as described previously (Weinberg et al. 2017a). A total of 30 unique representatives were used to generate an updated consensus and secondary structure model (Fig. 2A), which was created in part by using the program R2R (Weinberg and Breaker 2011).

Genetic reporter constructs

Reporter constructs were designed to include the Fibro-
\textit{purF} motif RNA from \textit{Fibrobacter succinogenes} (NC_017448.1), its natural promoter, and the first five codons of the downstream \textit{purF} ORF. The construct was inserted into the plasmid vector pRS414 as a translational fusion to the \textit{E. coli lacZ} gene. The resulting WT and mutant reporter constructs were transformed into \textit{E. coli} BW25113 and into the knockout strains as indicated.

Genetic reporter assays

Reporter gene assays were performed largely as described previously (Malkowski et al. 2019). Briefly, \textit{E. coli} strains carrying Fibro-
\textit{purF} motif reporter constructs were grown overnight at 37°C in Lysogeny Broth (LB) with appropriate antibiotic(s). Cells were plated on M9 minimal medium agar plates containing the appropriate antibiotic(s) and X-gal (100 µg mL$^{-1}$). Filter disks were prepared from 0.35 mm thick pure cellulose chromatography paper (Fisher Scientific) and placed on inoculated agar plates before 5 µL of the compound noted was added to the appropriate antibiotic(s) and X-gal (100 µg mL$^{-1}$). Cells were plated on M9 minimal medium agar plates containing the appropriate antibiotic(s) and X-gal (100 µg mL$^{-1}$). Filter disks were prepared from 0.35 mm thick pure cellulose chromatography paper (Fisher Scientific) and placed on inoculated agar plates before 5 µL of the compound noted was added to the disk. Plates were incubated at 37°C for 24 to 48 h before analysis.

RNA oligonucleotide preparation and in-line probing

RNA oligonucleotides were prepared by in vitro transcription from synthetic DNA templates, purified, enzymatically 5'$\sp{32}$P-labeled, and repurified as previously described (Malkowski et al. 2019). In-line probing assays were performed as previously described (Soukup and Breaker 1999; Regulski and Breaker 2008; Malkowski et al. 2019). DNA oligonucleotides used for the construction of a double-stranded DNA template for transcription of the \textit{Fibrobacter} sp. UW811 (NZ_FSTR01000002.1) 82 Fibro-
\textit{purF} RNA were 5’-TAATACGACTCACTATAGGTCGTCCGTTT CCTCGCTTGCACCCGTCCGGGTGG and 5’-CATAATGCCTCT TGTAATGATGTTTCCCTACATTATGTTCAAGCCACC GGAGCGGGTG. The underlined text corresponds to a T7 RNA polymerase promoter sequence. These were used to generate a double-stranded template by primer extension using SuperScript II reverse transcriptase (Thermo Fisher Scientific) as described previously (Malkowski et al. 2019). Other RNA constructs were prepared in a similar fashion with the appropriate synthetic oligonucleotides.

ACKNOWLEDGMENTS

We thank Neil White, Narasimhan Sudarsan, Kevin Perkins, and other members of the Breaker laboratory for helpful discussions. We also thank Kenneth Brewer for assistance with the bioinformatic searches for additional Fibro-
\textit{purF} motif representatives. S.N.M. and R.M.A. were supported by the National Science Foundation Graduate Research Fellowship Program (DGE1122492). C.E.W. was supported by the Deutsche Forschungsgemeinschaft (LU1889/1-1). This work was also supported by a National Institutes of Health (NIH) grant to R.R.B. (GM022778) and by the Howard Hughes Medical Institute.

Author contributions: All authors participated in the design of experiments. C.E.W. and E.B.G. designed the WT Fibro-
\textit{purF} reporter construct and performed initial scouting experiments. S.N. M. and R.M.A created and tested additional reporter constructs. S.N.M. designed and generated RNA constructs for biochemical analysis, performed in-line probing assays with C-PRA and related monomers, and executed all genetic assays depicted. S.N.M. and R.R.B. evaluated and interpreted all the results presented in this report. S.N.M. and R.R.B. wrote the manuscript with input from all authors.

Received July 16, 2020; accepted August 14, 2020.

REFERENCES

Ames TD, Breaker RR. 2010. Bacterial riboswitch discovery and analysis. In The chemical biology of nucleic acids (ed. Mayer G). Wiley, Chichester, UK.

Atiho RM, Mirihana Arachchilage G, Greenlee EB, Knecht KM, Breaker RR. 2019. A bacterial riboswitch class for the thiamin precursor HMP-PP employs a terminator-embedded aptamer. \textit{Elife} 8: e45210. doi:10.7554/eLife.45210

Bazurto JV, Farley KR, Downs DM. 2016. An unexpected route to an essential cofactor: \textit{Escherichia coli} relies on threonine for thiamine biosynthesis. \textit{MBio} 7: e01840-15. doi:10.1128/mbio.01840-15

Caperelli CA, Liu D. 1992. Carbocyclic substrates for \textit{de novo} purine biosynthesis. \textit{J Biol Chem} 267: 9783–9787.

Chen Y, Murray M, Schendel F, Otvos J, Wehrli S, Stubbe J. 1987. Chemical characterization of phosphoribosylamine, a substrate for newly discovered trifunctional protein containing glycineamide ribonucleotide synthetase activity. \textit{Adv Enzyme Regul} 26: 319–333. doi:10.1016/0065-2571(87)90021-5

Cho B-K, Federowicz SA, Embree M, Park Y-S, Kim D, Palsson BØ. 2011. The PurR regulon in \textit{Escherichia coli} K-12 MG1655. \textit{Nucleic Acids Res} 39: 6456–6464. doi:10.1093/nar/gkr307

Gebauer P, Heindl D, Horn C. 2012. Method and substances for preparation of N-substituted pyridinium compounds. U.S. patent no. WO 2012/101095 A1.

Gots JS, Love SH. 1954. Purine metabolism in bacteria. II. Factors influencing biosynthesis of 4-amino-5-imidazolecarboxamide by \textit{Escherichia coli}. \textit{J Biol Chem} 210: 395–405.

Greenlee EB, Stav S, Atiho RM, Brewer KI, Harris KA, Malkowski SN, Mirihana Arachchilage G, Perkins KR, Sherlock ME, Breaker RR. 2018. Challenges of ligand identification for the second wave of orphan riboswitch candidates. \textit{RNA Biol} 15: 377–390. doi:10.1080/15476286.2017.1403002

Huang L, Wang J, Lilley DMJ. 2020. Structure and ligand binding of the ADP-binding domain of the NAD$^+$ riboswitch. \textit{Nucleic Acids Res} 26: 878–887. doi:10.1261/ma.074898.120

Kellenberger CA, Wilson SC, Hickey SF, Gonzalez TL, Su Y, Hallberg ZF, Brewer TF, Iavarone AT, Carlson HK, Hsieh YF, et al. 

www.rnajournal.org 1845
2015. GEMM-I riboswitches from Geobacter sense the bacterial second messenger cyclic AMP-GMP. Proc Natl Acad Sci 112: 5383–5388. doi:10.1073/pnas.1419328112

Kim JN, Roth A, Breaker RR. 2007. Guanine riboswitch variants from Mesoplasma florum selectively recognize 2′-deoxyguanosine. Proc Natl Acad Sci 104: 16092–16097. doi:10.1073/pnas.0705884104

Kim PB, Nelson JW, Breaker RR. 2015. An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism. Mol Cell 57: 317–328. doi:10.1016/j.molcel.2015.01.001

Koenigsknecht MJ, Downs DM. 2010. Thiamine biosynthesis can be used to dissect metabolic integration. Trends Microbiol 18: 240–247. doi:10.1016/j.tim.2010.03.003

Koenigsknecht MJ, Lambracht JA, Fenlon LA, Downs DM. 2012. Perturbations in histidine biosynthesis uncover robustness in the metabolic network of Salmonella enterica. PLoS One 7: e48207. doi:10.1371/journal.pone.0048207

Lee ER, Baker JL, Weinberg Z, Sudarsan N, Breaker RR. 2010. An allosteric self-splicing ribozyme triggered by a bacterial second messenger. Science 329: 845–848. doi:10.1126/science.1190713

Malkowski SN, Spencer TCJ, Breaker RR. 2019. Evidence that the nadA motif is a bacterial riboswitch for the ubiquitous enzyme cofactor NAD⁺. RNA 25: 1616–1627. doi:10.1261/rna.072538.119

Mandal M, Breaker RR. 2004. Adenine riboswitches and gene activation by disruption of a transcription terminator. Nat Struct Mol Biol 11: 29–35. doi:10.1038/nsmb710

Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. 2003. Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. Cell 113: 577–586. doi:10.1016/S0092-8674(03)00391-X

McCown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR. 2017. Riboswitch diversity and distribution. RNA 23: 995–1011. doi:10.1261/ma.061234.117

Mirihana Arachchiilage G, Weinstein Z, Breaker RR. 2018. SAM-VI RNAs selectively bind S-adenosylmethionine and exhibit similarities to SAM-III riboswitches. RNA Biol 15: 371–378. doi:10.1080/15476286.2017.1399232

Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics 29: 2933–2935. doi:10.1093/bioinformatics/bt3059

Nelson JW, Sudarsan N, Furukawa K, Weinberg Z, Wang JX, Breaker RR. 2013. Riboswitches in eubacteria sense the second messenger c-di-AMP. Nat Chem Biol 9: 834–839. doi:10.1038/nchembio.1363

Nelson JW, Sudarsan N, Phillips GE, Stav S, Lünse CE, McCown PJ, Breaker RR. 2015. Control of bacterial exoelectrogenesis by c-AMP-GMP. Proc Natl Acad Sci 112: 5389–5394. doi:10.1073/pnas.1419264112

Nierlich DP, Magasanik B. 1965. Phosphoribosylglycinamidase synthetase of Aerobacter aerogenes. Purification and properties, and nonenzymatic formation of its substrate 5′-phosphoribosylamine. J Biol Chem 240: 366–374.

O’Leary NA, Wright MW, Brister JR, Ciufio S, Haddad D, McVeigh R, Rajput B, Robbertse SM, Smith-Wheat B, Ako-Adjie D, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44: D733–D745. doi:10.1093/nar/gkv1189

Regulski EE, Breaker RR. 2008. In-line probing analysis of riboswitches. Methods Mol Biol 419: 53–67. doi:10.1007/978-1-59745-033-1_4

Regulski EE, Moy RH, Weinberg Z, Barrick JE, Yao Z, Ruzzo WL, Breaker RR. 2008. A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. Mol Microbiol 68: 918–932. doi:10.1111/j.1365-2958.2008.06208.x

Roth A, Breaker RR. 2009. The structural and functional diversity of metabolite-binding riboswitches. Annu Rev Biochem 78: 305–334. doi:10.1146/annurev.biochem.78.070507.135565

Schindel FJ, Cheng YS, Otvos JD, Wehlé S, Stubbbs L. 1988. Characterization and chemical properties of phosphoribosylamine, an unstable intermediate in the de novo purine biosynthetic pathway. Biochemistry 27: 2614–2623. doi:10.1021/bi00470a052

Serganov A, Udler E. 2013. A decade of riboswitches. Cell 152: 17–24. doi:10.1016/j.cell.2012.12.024

Sherlock ME, Sudarsan N, Stav S, Breaker RR. 2018a. Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria. Elife 7: e33908. doi:10.7554/eLife.33908

Sherlock ME, Sudarsan N, Breaker RR. 2018b. Riboswitches for the alrmame ppGpp expand the collection of RNA-based signaling systems. Proc Natl Acad Sci 115: 6052–6057. doi:10.1073/pnas.1720461115

Sherlock ME, Sadeeshkumar H, Breaker RR. 2019. Variant bacterial riboswitches associated with nucleotide hydrolase genes sense nucleoside diphosphates. Biochemistry 58: 401–410. doi:10.1021/acs.biochem.8b00617

Sherwood AV, Henkin TM. 2016. Riboswitch-mediated gene regulation: novel RNA architectures dictate gene expression responses. Annu Rev Microbiol 70: 361–374. doi:10.1146/annurev-micro-091014-104306

Soukup GA, Breaker RR. 1999. Relationship between intermolecular linkage geometry and the stability of RNA. RNA 5: 1308–1325. doi:10.1261/sin15583829999891

Soukup GA, DeRose EC, Koizumi M, Breaker RR. 2001. Generating new ligand-binding RNAs by affinity maturation and disintegration of allosteric ribozymes. RNA 7: 524–536. doi:10.1261/rna.1355838201002175

Stav S, Atiño RM, Arachchiilage GM, Nguyen G, Higgs G, Breaker RR. 2019. Genome-wide discovery of structured noncoding RNAs in bacteria. BMC Microbiol 19: 66. doi:10.1186/s12866-019-1433-7

Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, Breaker RR. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science 321: 411–413. doi:10.1126/science.1159519

Weinberg Z, Breaker RR. 2011. R2R-software to speed the depiction of aesthetic consensus RNA secondary structures. BMC Bioinformatics 12: 3. doi:10.1186/1471-2105-12-3

Weinberg Z, Lünse CE, Corbino KA, Ames TD, Nelson JW, Roth A, Perkins KR, Sherlock ME, Breaker RR. 2017a. Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. Nucleic Acids Res 45: 10811–10820. doi:10.1093/nar/gkw499

Weinberg Z, Nelson JW, Lünse CE, Sherlock ME, Breaker RR. 2017b. Bioinformatic analysis of riboswitch structures uncovers variant classes with altered ligand specificity. Proc Natl Acad Sci 114: E2077–E2085. doi:10.1073/pnas.1619581114

Yu D, Breaker RR. 2020. A bacterial riboswitch class senses xanthine and uric acid to regulate genes associated with purine oxidation. RNA doi:10.1261/ma.075218.120

Zalkin H. 1983. Structure, function, and regulation of amidophosphoribosyltransferase from prokaryotes. Adv Enzym Regul 21: 225–237. doi:10.1016/0065-2571(83)90016-X