Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems

Madeline E. Sherlock[a], Narasimhan Sudarsan[b], and Ronald R. Breaker[a,b,c,1]

[a]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8103; [b]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520-8103; and [c]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520-8103

Contributed by Ronald R. Breaker, April 20, 2018 (sent for review November 27, 2017; reviewed by Daniel A. Lafontaine and Jörg Vogel)

Riboswitches are noncoding portions of certain mRNAs that bind metabolite, coenzyme, signaling molecule, or inorganic ion ligands and regulate gene expression. Most known riboswitches sense derivatives of RNA monomers. This bias in ligand chemical composition is consistent with the hypothesis that widespread riboswitch classes first emerged during the RNA World, which is proposed to have existed before proteins were present. Here we report the discovery and biochemical validation of a natural riboswitch class that selectively binds guanosine tetraphosphate (ppGpp), a widespread signaling molecule and bacterial "alarmone" derived from the ribonucleotide GTP. Riboswitches for ppGpp are predicted to regulate genes involved in branched-chain amino acid biosynthesis and transport, as well as other gene classes that previously had not been implicated to be part of its signaling network. This newfound riboswitch–alarmone partnership supports the hypothesis that prominent RNA World signaling pathways have been retained by modern cells to control key biological processes.

branched chain amino acid | guanosine tetraphosphate | logic gate | stringent response | *ykkC* motif RNA

Among the 38 distinct riboswitch (1–3) classes validated previously (4), five respond to RNA-derived signaling molecules that might have emerged in an RNA World, but have been retained to trigger regulatory processes in modern cells. We have proposed (5, 6) that riboswitches for the RNA-derived molecules c-di-GMP (7, 8), c-di-AMP (9), ZTP (10), and c-AMP-GMP (11, 12) represent modern versions of primordial signaling processes built entirely from RNA.

A widely distributed RNA signaling molecule that is essential for many bacteria is ppGpp (Fig. 1*A*), a guanosine tetraphosphate alarmone originally called magic spot (13). This compound, along with its natural precursor pppGpp, is widely implicated in regulating RNA polymerase and other protein factors to control metabolic and physiological responses to various stresses (14–16), including the long-studied stringent response. The stringent response is a major starvation-induced adaptation that can be triggered by amino acid depletion (17–19). In Proteobacteria, this adaptation is accomplished at least in part via the direct inhibition of RNA polymerase by ppGpp binding (16).

In Firmicutes, transcription regulation by ppGpp is known to act indirectly. GTP levels drop as a result of (p)ppGpp synthesis to specifically down-regulate transcription of genes that initiate with a G nucleotide, which affects the production of certain mRNAs and also rRNAs. GTP depletion still permits transcription of genes that initiate with an A nucleotide, such as those for branched-chain amino acid (BCAA) biosynthesis (20, 21). Specific gene activation also occurs by the dissociation of the CodY transcriptional repressor protein when GTP levels drop as a result of (p)ppGpp synthesis (22). This protein-based system thus allows ppGpp to indirectly activate CodY-controlled genes. Herein we describe another mechanism for ppGpp-triggered regulation of gene expression in Firmicutes mediated by numerous examples of selective ppGpp-sensing riboswitches.

Results and Discussion

Identification of ppGpp Riboswitch Candidates. The widespread distribution of ppGpp in biology, its role in regulating fundamental biological processes, and its construction from an RNA mononucleotide are consistent with an ancient RNA World origin (5). Given these characteristics, we previously speculated that ppGpp would be an ideal ligand candidate for a riboswitch class that had yet to be discovered (5, 6, 10, 12). Thus, we were well positioned to test ppGpp as a candidate ligand when we recently (23) identified a riboswitch subtype (Fig. 1*B*) of unknown function that commonly associates with genes for BCAA metabolism (Fig. 1*C*). The ppGpp riboswitch candidates were recognized as "subtype 2a" variants of what we originally called a *ykkC* motif RNAs (24).

The predominant members (subtype 1) of the *ykkC* motif collection were recently determined to function as guanidine-sensing RNAs, which are collectively called guanidine-I riboswitches (23). In this same study, it was discovered that certain *ykkC* motif RNAs (subtype 2) carry distinctive mutations at nucleotide positions found directly in the binding pocket (25, 26). Members of the *ykkC* subtype 2 collection do not bind guanidine, confirming that the binding site differences indeed alter riboswitch specificity.

Although ppGpp had previously been tested for binding to *ykkC* RNAs (27), it was determined in retrospect that the specific RNA sequences tested belong to the guanidine-I riboswitch class. After further separation of the subtypes into four additional classes, called subtype 2a through 2d (based on distinctive gene associations and aptamer sequence variations), ppGpp became a likely candidate ligand because of the association of many subtype 2a representatives with genes for BCAA biosynthesis. Furthermore, BCAA biosynthesis is activated during the stringent response in Firmicutes, which are the predominant organisms that contain subtype 2a *ykkC* motif variants.

Significance

Bacteria and other organisms make extensive use of signaling molecules that are derived from ribonucleotides or their derivatives. Previously, five riboswitch classes had been discovered that sense the four RNA-derived signaling molecules: c-di-GMP, c-di-AMP, c-AMP-GMP, and ZTP. We now report the discovery and biochemical validation of bacterial riboswitches for the widespread alarmone guanosine tetraphosphate (ppGpp), which signals metabolic and physiological adaptations to starvation. These findings expand the number of natural partnerships between riboswitches and ribonucleotide-like signaling molecules, and provide RNA-based sensors for detecting ppGpp production in cells.

Author contributions: M.E.S., N.S., and R.R.B. designed research; M.E.S. and N.S. performed research; M.E.S., N.S., and R.R.B. analyzed data; and M.E.S. and R.R.B. wrote the paper.

Reviewers: D.A.L., Université de Sherbrooke; and J.V., Helmholtz Institute for RNA-Based Infection Research.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: ronald.breaker@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1720406115/-/DCSupplemental.

Published online May 21, 2018.

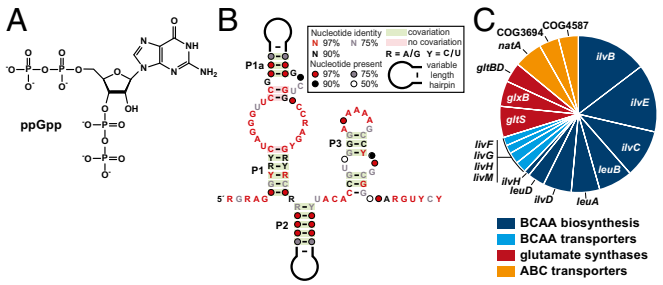

Fig. 1. ppGpp riboswitches and their regulatory network. (*A*) The chemical structure of ppGpp (guanosine-3′,5′-bisdiphosphate). (*B*) Consensus sequence and secondary structure model for ppGpp riboswitches derived from 115 unique representatives. (*C*) Genes predicted to be regulated by ppGpp riboswitches.

Subtype 2a *ykkC* Motif RNAs Selectively Bind ppGpp. To test for ppGpp riboswitch function, we first arbitrarily chose to examine a BCAA-associated candidate RNA construct from the *ilvE* gene of *Thermosediminibacter oceani* (Fig. 2*A*) by using in-line probing. This method exploits the differential instability of RNA phosphoester bonds in unstructured versus structured regions, to provide evidence that ligand docking alters the folding state of the ligand-binding aptamer (28, 29). We observed a pattern of ppGpp-dependent RNA structural modulation with this construct (Fig. 2*B*) to yield a binding curve consistent with 1:1 complex formation and an apparent dissociation constant (K_D) of ∼10 nM (Fig. 2*C*). Nucleotides that exhibit structural modulation are located in regions that are known to form the binding pocket for guanidine-I riboswitches (25, 26). Remarkably, these two riboswitch classes utilize the same general architecture but have altered key nucleotides to create binding pockets capable of recognizing completely different ligands.

Individual mutation of six highly conserved nucleotides (M1 through M6) within the *T. oceani* RNA (*SI Appendix*, Fig. S1*A*), predicted to be directly involved in ligand binding, each cause at least a three order-of-magnitude loss in ppGpp binding affinity (*SI Appendix*, Fig. S1*B*). We speculate that these nucleotide positions might form the binding pocket for ppGpp because their equivalent positions are directly involved in contacting the ligand in the highly similar members of the guanidine-I riboswitch class. Reversing the nucleotide identity of a conserved base pair (M7) similarly reduces ligand affinity by about one order-of-magnitude. These findings demonstrate that conserved nucleotides are critical for optimal binding by the aptamer.

Riboswitch binding specificity was assessed via in-line probing by using a collection of ppGpp analogs comprised of guanosine nucleotide derivatives. In *Bacillus subtilis*, the most abundant natural version of the signaling molecule is the pentaphosphate species (pppGpp), which is directly synthesized when GTP is used as the biosynthetic precursor. The riboswitch displays equal affinities for ppGpp (Fig. 2*C*) and pppGpp (*SI Appendix*, Fig. S2*B*), indicating they are both likely to trigger gene expression in cells. At higher concentrations, similar compounds, such as pGp, GTP, and GDP, modulate the RNA structure in the same regions as pppGpp and ppGpp (*SI Appendix*, Fig. S3*A*). Both GDP and GTP require >1,000-fold higher concentration than ppGpp to half-maximally saturate the riboswitch aptamer (*SI Appendix*, Fig. S3*B*), whereas other nucleotides have no effect regardless of the concentrations tested (*SI Appendix*, Fig. S3*C*).

We hypothesized that the ppGpp aptamer RNA might recognize the guanine base of its ligand through Watson–Crick pairing and sought to identify a conserved cytosine residue that might perform this function. The conserved cytosine between the P2 and P3 stems was identified as a likely candidate due to its role in the guanidine-I riboswitch structure (25, 26). Specifically, this C forms a long-range base pair with a conserved G in the internal loop of the P3 stem, which appears to be absent in ppGpp riboswitches because the corresponding nucleotide is no longer a conserved G nucleotide. It seemed plausible that the C nucleotide (C71 in the *T. oceani* 112 *ilvE* RNA), which is in a region of reduced scission upon ligand binding in-line probing reactions (Fig. 2*A*), might be available to base pair to the nucleobase of ppGpp. Indeed, introducing a C to U mutation at this position (M8) (Fig. 3*A*) switches the specificity of the aptamer from a G-containing ligand to an A-containing ligand, as demonstrated by ability of the mutant RNA to preferentially recognize pAp over pGp or ppGpp (Fig. 3*B*).

In-line probing was also used to estimate the K_D of pAp binding by the M8 RNA construct (Fig. 3*C*). The binding curve generated from these data (Fig. 3*D*) reveals a K_D of ∼50 μM for one-to-one binding of pAp by M8 RNA. This K_D is similar to that measured for pGp binding to the WT version of this same RNA (*SI Appendix*, Fig. S3*B*). These data suggest that this C nucleotide at position 71 of the 112 *ilvE* aptamer is the primary determinant of nucleobase recognition for the riboswitch ligand. Currently, we have no evidence for a naturally occurring aptamer for ppApp (or a similar adenosine-containing molecule) because this position of the aptamer is conserved as a C nucleotide across all natural examples of the sequences resembling ppGpp riboswitches examined to date.

Transcription Termination Regulated by a ppGpp Riboswitch. To determine whether ppGpp binding by a member of this riboswitch class directly affects gene expression, single turnover in vitro transcription termination assays (30) were performed as previously described (31) with a representative from the 5′UTR of the *ilvE* gene of *Desulfitobacterium hafniense* (Fig. 4*A*). This construct was chosen because the expression platform exhibited an architecture that suggested the *D. hafniense* riboswitch uses a transcription termination mechanism, and this construct was

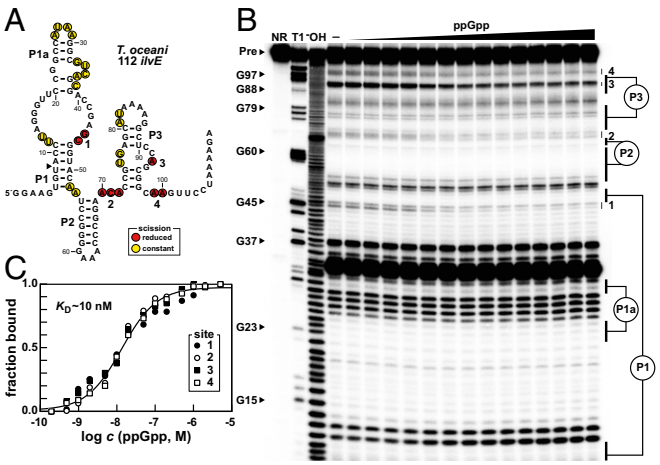

Fig. 2. Ligand binding by a ppGpp riboswitch aptamer. (*A*) Sequence and secondary structure of the 112 nucleotide RNA derived from the *ilvE* gene of *T. oceani*. Data collected in *B* were used to determine regions of constant or reduced scission upon the addition of ppGpp to in-line probing reactions. (*B*) PAGE analysis of the products of in-line probing reactions with 5′-^{32}P–labeled 112 *ilvE* RNA in the presence of increasing ppGpp concentrations. NR, T1, and ⁻OH represent RNA undergoing no reaction, partial digest with RNase T1 (cleaves after guanosine nucleotides), or partial digest under alkaline conditions (cleaves after every nucleotide), respectively. Selected bands corresponding to RNA strand scission after G residues are annotated. In-line probing reactions contained either no ligand (–) or ppGpp ranging from 200 pM to 5 μM. Regions 1–4 of the RNA exhibit structural stabilization in a ppGpp-dependent manner. (*C*) Plot of the fraction of RNA bound to ppGpp, as determined by the normalized fraction of RNA scission at regions 1–4 as a function of the logarithm of ppGpp concentration (*c*). The dissociation constant (K_D) was estimated by a sigmoidal fit of the average of the data points from all regions of RNA structural modulation using a procedure described in *Materials and Methods*. Data shown are representative of at least three separate in-line probing experiments.

MICROBIOLOGY

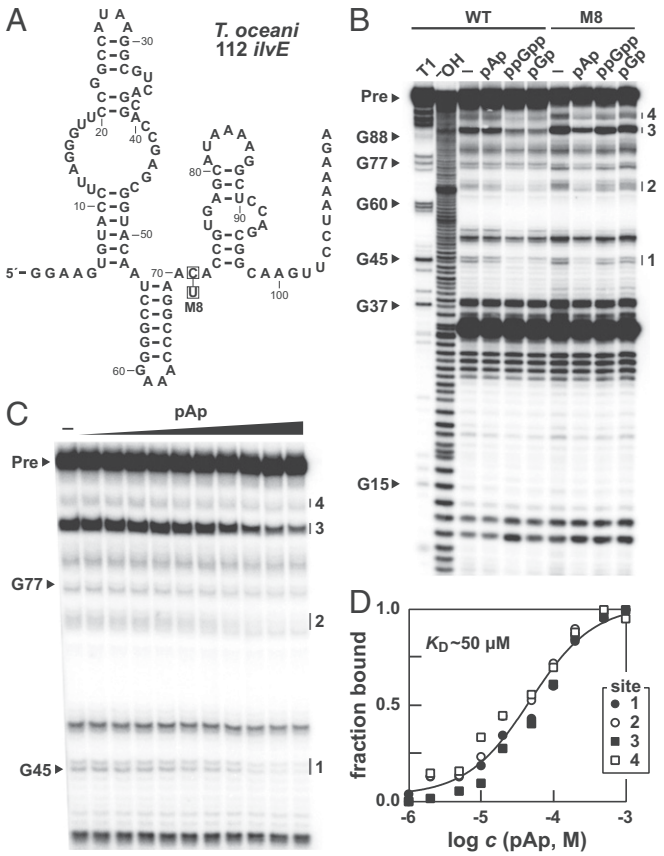

Fig. 3. Mutation of a conserved nucleotide changes ligand specificity. (*A*) Sequence and secondary structure of the WT 112 *ilvE* RNA from *T. oceani*. The boxed nucleotide identifies the M8 mutation. (*B*) PAGE analysis of in-line probing assays with WT and M8 112 *ilvE* RNA in the absence (–) or presence of 1 mM pAp, ppGpp, or pGp. Additional annotations are as described in the legend to Fig. 2*B*. (*C*) PAGE analysis of in-line probing assays of M8 112 *ilvE* RNA with various concentrations of pAp ranging from 0 (–) to 1 mM. (*D*) Plot of the fraction of RNA bound to pAp. Additional annotations are as described in the legend to Fig. 2*C*.

confirmed to operate with RNA polymerase from *Escherichia coli*. Specifically, we predicted that the P3 helix of the aptamer acts as an antiterminator stem, and so transcription of the full-length mRNA should increase when ppGpp is bound. As expected, the WT riboswitch permits RNA polymerase to proceed through the stop site of the adjoining intrinsic terminator stem with greater efficiency only when sufficient ppGpp is added to the reaction (Fig. 4*B*). An RNA (M9) with a single mutation at a nucleotide position that disrupts the binding pocket terminates with the same frequency regardless of the addition of ppGpp. This finding demonstrates that transcription read-through requires ppGpp binding by the riboswitch, rather than the binding of this signaling molecule to the RNA polymerase protein complex.

The concentration of ppGpp needed to achieve half-maximal termination (T_{50}) is 6 μM for the *D. hafniense ilvE* RNA (Fig. 4 *C* and *D*). This is nearly 1,000-fold higher than the K_D value measured for direct binding of ppGpp to the *T. oceani* 112 *ilvE* RNA construct (Fig. 2*C*). This difference might be partly due to the distinct sequences of the two ppGpp riboswitches, or to the different conditions in which the two types of assays are conducted. However, it is well established (31–34) that some riboswitches operate under kinetic control (31) (likely experienced during transcription termination assays), rather than reaching thermodynamic equilibrium (as attained during in-line probing assays). Riboswitches operating under kinetic control have been shown to require higher concentrations of ligand to trigger gene control compared with K_D values of the same constructs measured under thermodynamic equilibrium. In other words, K_D values measured in a test tube do not reflect the concentration of ligand naturally required to trigger large changes in gene expression because each riboswitch aptamer does not have time to reach thermodynamic equilibrium before RNA polymerase transcribes beyond the intrinsic terminator stem.

The amount of full-length transcript is quite high even when no ppGpp is present and the terminator is expected to form (Fig. 4*B*). This is again likely due to kinetic aspects, such as the speed of RNA polymerase, which can influence the amount of ligand needed to regulate riboswitch-mediated termination, and the absence of the transcription factor NusA, which can increase the dynamic range of transcription termination yields (31, 35). Also, the modest dynamic range of the *D. hafniense ilvE* riboswitch in vitro might be due in part to the fact that the run of U residues traditionally following a terminator stem is interrupted with a C nucleotide. A construct wherein this C residue is altered to a U

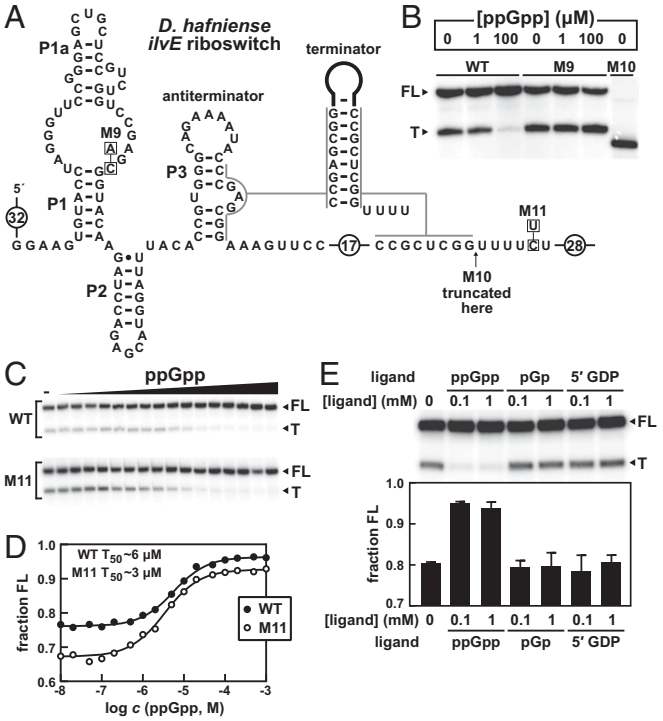

Fig. 4. ppGpp regulates riboswitch-dependent transcript termination. (*A*) Sequence and secondary structure of the ppGpp riboswitch derived from the *ilvE* gene of *D. hafniense*. An alternative "terminator" RNA structure, which only forms in the absence of ppGpp, is depicted wherein a hairpin forms followed by a U-rich tract. (*B*) PAGE analysis of single-round transcription termination assays with WT, M9, and M10 *D. hafniense ilvE* riboswitches conducted in the absence or presence of ppGpp, as indicated. Construct sequences are as depicted in *A*. FL and T denote full-length and terminated transcripts, respectively. Construct M10 provides a size marker for a transcript that terminates several nucleotides before the product generated when the intrinsic terminator forms. (*C*) PAGE analysis of single-round transcription termination assays of the WT and M11 *D. hafniense ilvE* RNAs with either no ligand (–) or ppGpp ranging from 10 nM to 1 mM. (*D*) Plot of the fraction of FL WT or M11 *D. hafniense ilvE* riboswitch transcripts contributing to the total number of transcripts (FL plus T) as a function of the logarithm of ppGpp molar concentration. The concentration of ppGpp required to cause half-maximal change in termination (T_{50}) was determined by a sigmoidal curve fit. (*E*, *Upper*) PAGE analysis of single-round transcription termination assays of the WT *D. hafniense ilvE* RNA with either no ligand (–) or each ligand as indicated at a concentration of 0.1 or 1 mM. Data are representative of multiple trials. (*Lower*) Plot of the fraction of FL WT *D. hafniense ilvE* riboswitch transcripts contributing to the total number of transcripts (FL plus T) in the absence of ligand (–) or with each ligand as indicated. Values reported are the average and SD of three separate trials, including the PAGE image depicted above.

(M11) exhibits a modestly improved dynamic range (Fig. 4 *C* and *D*). Regardless of the factors affecting the dynamic range of the riboswitch when measured in vitro, similar dynamic ranges for other riboswitch classes examined by in vitro transcription assays have been observed previously (36–38).

To further assess ppGpp as the natural ligand for this riboswitch class, we tested other guanine nucleotide compounds using in vitro transcription termination assays. Although pGp and GDP cause modulation of the RNA under in-line probing conditions, albeit with decreased affinity compared with ppGpp, neither is able to cause a shift from terminated to full-length product even at a concentration of 1 mM (Fig. 4*E*). GTP was not tested because it is already present at 150 μM in the transcription reaction. These results demonstrate that these RNAs selectively regulate gene expression in response to ppGpp or pppGpp, but not any other guanine nucleotide derivative.

Gene Associations of ppGpp Riboswitches. The ligand binding and in vitro regulatory function of ppGpp riboswitch constructs described above are entirely consistent with the predominant gene associations observed for members of this riboswitch class. The link between BCAA metabolism and ppGpp signaling is well established (13–18), and our observations that ppGpp and pppGpp are bound by members of this riboswitch class are consistent with this natural gene association.

However, ppGpp riboswitches are also commonly associated with members of two other gene classes, which are connections that were not well established previously. The first includes genes encoding various glutamate synthase domains, including *gltS*, *gltB*, *gltD*, and *glxB*. The second includes operons containing three ATP-binding cassette (ABC) transporters, one of which is predicted to be *natA* while the other two (COG3694 and COG4587) have unknown functions. Representative RNAs associated with glutamate synthase (*SI Appendix*, Fig. S4) and *natA* (*SI Appendix*, Fig. S5) gene classes bind ppGpp with dissociation constants similar to that observed for the *T. oceani* 112 *ilvE* riboswitch aptamer (Fig. 2).

Organisms that use ppGpp riboswitches to control glutamate synthase genes typically also use ppGpp riboswitches to regulate BCAA biosynthesis-related gene expression, suggesting that these two metabolic processes are linked. Glutamate levels are known to be elevated during the stringent response (39), and this riboswitch discovery provides a direct mechanism for up-regulation of glutamate biosynthesis during this stress response. Moreover, the known connection between these biochemical responses reinforces our hypothesis that the natural ligand for some *ykkC* subtype 2 RNAs indeed is ppGpp or its immediate biosynthetic precursor pppGpp.

There was no prior established link between ppGpp signaling and the ABC transporters associated with ppGpp riboswitches. Organisms with ppGpp riboswitches upstream of these transporter genes are mutually exclusive to organisms with ppGpp riboswitches upstream of BCAA and glutamate synthesis genes. A possible explanation for this gene-association pattern comes from the fact that one of these transporters shares homology with NatA, which has a predicted cation transport function and is known to be expressed under membrane stress conditions (40). Various cell envelope stresses are also known to activate the *relP* ppGpp synthase gene in *B. subtilis* (41–43). Therefore, it seems logical for these transporters, if they are indeed NatA homologs, to be expressed when ppGpp levels are high, although in this case ppGpp production is not directly related to the stringent response.

Some ppGpp Riboswitches Reside in Tandem with T Box RNAs. In ~40% of the instances when the ppGpp riboswitch is found in the 5′UTR of BCAA biosynthesis genes, a leucine T box RNA is found either directly upstream or downstream of the riboswitch in the same UTR. T box RNAs are regulatory domains that control gene expression in response to binding certain uncharged tRNAs, thereby indirectly detecting a deficiency of an amino acid (44). These two RNA elements presumably function independently because each has its own expression platform. When found in tandem, they function as a two-input Boolean AND logic gate, which would be similar to riboswitch logic gate architectures reported previously (45). Specifically, both uncharged $tRNA^{Leu}$ and ppGpp would be necessary for transcription of the downstream protein coding regions.

This tandem arrangement likely exploits two required inputs to ensure that the expression of associated genes is reduced when nutrients are plentiful. Additionally, because RelA is allosterically regulated by any uncharged tRNA, the ppGpp riboswitch will turn on gene expression when any amino acid is limiting, whereas each T box RNA indirectly detects a limitation only of its cognate amino acid. At the promoter level, BCAA biosynthesis genes are regulated by the CodY transcriptional repressor, which uses GTP and isoleucine as regulatory inputs (22, 46). It is unclear whether all three of these mechanisms are utilized simultaneously because, while organisms with ppGpp riboswitches typically have the *codY* gene, the canonical CodY binding site consensus sequence (46) was not found upstream of ppGpp riboswitch sequences.

Bioinformatic and Biochemical Data Reveal Additional Riboswitch Classes. All other RNAs examined that were grouped with *ykkC* subtype 2 fail to bind either ppGpp or guanidine, and therefore represent distinct riboswitch classes. Subtypes 2a through 2d were identified by sorting RNA representatives based on their distinct gene associations, and then by comparative sequence analysis (23), as was conducted to identify *ykkC* subtype 2a RNAs and to verify that they function as ppGpp riboswitches (Fig. 1). We recently determined that some of these RNAs (subtype 2b) bind phosphoribosyl pyrophosphate, a precursor of nucleotide biosynthesis (47). However, we speculate that at least two additional candidate riboswitch classes (subtypes 2c and 2d) whose ligands remain undiscovered are represented among the remaining *ykkC* RNA variants.

Conclusions

A riboswitch class for ppGpp has long been anticipated (5, 6, 10, 12), but representatives of this particular class were challenging to identify due to their striking similarity to riboswitches for guanidine (23). At one point early in evolutionary history, ppGpp riboswitches might have been extremely widespread, but it appears that their function has been replaced by the CodY repressor protein factor in many organisms.

Notably, the discovery of ppGpp riboswitches in tandem with T box regulators serves as an additional demonstration of how complex molecular devices can be formed purely from RNA. The Boolean AND gate architecture described herein involves a tRNA sensed by a T box RNA and a riboswitch RNA that senses the RNA-derived alarmone ppGpp. This all-RNA system is similar to that for the two-input NOR logic gate formed by tandem riboswitches for *S*-adenosylmethionine and adenosylcobalamin (45). Similarly, it has previously been demonstrated that a riboswitch located in tandem with a group I self-splicing ribozyme functions as a two-input AND gate that requires both the signaling molecule c-di-GMP and the nucleotide GTP to properly splice and express a messenger RNA from *Clostridium difficile* (8, 48). Such complex all-RNA systems suggest that riboswitch-mediated responses to nucleotide-like signaling molecules such as ppGpp could be of ancient RNA World origin.

Materials and Methods

Chemicals and Reagents. Chemicals were purchased from Sigma-Aldrich with the exceptions of guanosine 5′,3′-bisdiphosphate (ppGpp), guanosine pentaphosphate (pppGpp) (Jena Bioscience), and guanosine 5′,3′-bisphosphate (pGp) (TriLink Biotech). [γ-^{32}P]-ATP was purchased from Perkin-Elmer and used within 2 wk of delivery. All bulk chemicals were purchased from J. T. Baker and enzymes were purchased from New England Biolabs, unless otherwise noted. All solutions were prepared using deionized water (dH_2O) and were either autoclaved or filter-sterilized (using 0.22-μm filters; Millipore) before use. DNA oligonucleotides used in this study (*SI Appendix*, Table S1) were purchased from Sigma-Aldrich and Integrated DNA

Technologies. Strains of *B. subtilis* were obtained from the *Bacillus* Genetic Stock Center at The Ohio State University.

Bioinformatics Analyses. Additional examples of *ykkC* motif RNAs were identified using Infernal 1.1 (49) to search RefSeq v76 plus additional environmental microbial databases (50). Iterative searches for new sequences were performed based on the previously published alignment of *ykkC* subtype 2 RNAs (23). These sequences were sorted by nucleotide identity at a specific position as previously indicated (23) to exclude guanidine-I riboswitches, then further manually sorted by downstream gene association. This analysis identified 105 examples of subtype 2a RNAs (ppGpp riboswitches), of which 78 are upstream of BCAA metabolism genes, 15 are upstream of various glutamate synthase genes, and 12 are upstream of *natA* and other ABC transporter genes. It should be noted that this is an underestimate of the true number of representatives of this riboswitch class in this dataset, as examples with incomplete gene contexts were not evaluated. The consensus sequence and secondary structure model of these 105 ppGpp riboswitch RNAs was constructed using R2R software (51).

RNA Oligonucleotide Preparation. Preparation of RNAs for use in in-line probing experiments was performed as previously described (23, 52) with the following modifications. All double-stranded DNA (dsDNA) templates used for RNA transcription were assembled via extension of overlapping synthetic DNAs (*SI Appendix*, Table S1) using SuperScript II reverse transcriptase (Thermo Fisher Scientific). Forward primers contain the T7 RNA polymerase (T7 RNAP) promoter followed by guanosine residues (when not present in the WT sequence) to enable transcription. Thereafter, RNA constructs were prepared from these dsDNAs by in vitro transcription, purified, and subsequently 5′ ^{32}P-labeled as previously described (23, 52).

RNA In-Line Probing Analyses. In-line probing assays were performed as previously described (28, 29). The values for the amount of RNA aptamer bound to ligand or "fraction bound" were estimated based on the ligand-mediate changes in the band intensities from in-line probing assays also as reported previously (23, 52). Briefly, quantified changes in band intensities at designated sites reflect the fraction of RNAs that are bound to ligand, such that half-maximal increases or decreases in band intensities represent the K_D for the RNA–ligand interaction.

Single-Round Transcription Termination Assays. Template dsDNA constructs for single-round in vitro transcription were designed to include the aptamer and expression platform of the *D. hafniense* riboswitch starting at the predicted natural transcription start site and ending 28 nucleotides beyond the terminator stem. The promoter used to enable transcription was derived from the *B. subtilis lysC* gene, which is compatible with *E. coli* RNA polymerase. The described WT and mutant templates were purchased as synthetic dsDNAs and subsequently amplified by PCR. Approximately 2 pmol of the resulting, purified DNA template was added to a transcription initiation mixture [20 mM Tris·HCl (pH 7.5 at 23 °C), 75 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 10 μg mL^{-1} BSA, 130 μM ApA dinucleotide, 1% glycerol, 0.04 U μL^{-1} *E. coli* RNA polymerase holoenzyme, 2.5 μM GTP, 2.5 μM ATP, and 1 μM UTP]. Approximately 1 μCi [α-^{32}P]-UTP was added per 8-μL reaction and transcription was allowed to proceed at 37 °C for 10 min, leading to formation of a stalled polymerase complex at the first cytidine nucleotide of each transcript. For each 8-μL transcription reaction, 1 μL of 10× elongation buffer [20 mM Tris·HCl (pH 7.5 at 23 °C), 75 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 2 mg mL^{-1} heparin, 1.5 mM ATP, 1.5 mM GTP, 1.5 mM CTP, and 250 μM UTP] as well as 1 μL of a 10× solution of the ligand of interest were added sequentially. Because ppGpp is a known inhibitor of *E. coli* RNA polymerase initiation (16), the holoenzyme was first allowed to assemble into a stalled, stable ternary complex before ppGpp was added concurrently with the elongation mixture. Transcription reactions were then incubated at 37 °C for an additional 60 min.

The transcription products were subsequently analyzed via denaturing 10% PAGE and visualized using a phosphorimager (GE Healthcare Life Sciences). Fraction full-length (FL) values were calculated by varying the ligand concentration in separate reactions and quantifying the changes in band intensity of both FL and terminated (T) transcription products using the formula (FL)/(FL + T). The T_{50} values were determined by plotting the fraction FL as a function of the logarithm of ligand concentration and using a sigmoidal four parameter logistic fit in GraphPad Prism 7.

ACKNOWLEDGMENTS. We thank Shira Stav, Adam Roth, Kimberly Harris, and other members of the R.R.B. laboratory for helpful discussions. This work was supported by NIH Grants GM022778 and DE022340 (to R.R.B.). M.E.S. was supported by NIH Cellular and Molecular Biology Training Grant T32GM007223. R.R.B. is also supported by the Howard Hughes Medical Institute.

1. Roth A, Breaker RR (2009) The structural and functional diversity of metabolite-binding riboswitches. *Annu Rev Biochem* 78:305–334.
2. Serganov A, Nudler E (2013) A decade of riboswitches. *Cell* 152:17–24.
3. Sherwood AV, Henkin TM (2016) Riboswitch-mediated gene regulation: Novel RNA architectures dictate gene expression responses. *Annu Rev Microbiol* 70:361–374.
4. McCown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR (2017) Riboswitch diversity and distribution. *RNA* 23:995–1011.
5. Nelson JW, Breaker RR (2017) The lost language of the RNA World. *Sci Signal* 10:1–10.
6. Breaker RR (2009) Riboswitches: From ancient gene control systems to modern drug targets. *Future Microbiol* 4:771–773.
7. Sudarsan N, et al. (2008) Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* 321:411–413.
8. 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.
9. Nelson JW, et al. (2013) Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat Chem Biol* 9:834–839.
10. 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.
11. Kellenberger CA, et al. (2015) GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP. *Proc Natl Acad Sci USA* 112:5383–5388.
12. Nelson JW, et al. (2015) Control of bacterial exoelectrogenesis by c-AMP-GMP. *Proc Natl Acad Sci USA* 112:5389–5394.
13. Cashel M (1969) The control of ribonucleic acid synthesis in *Escherichia coli*. IV. Relevance of unusual phosphorylated compounds from amino acid-starved stringent strains. *J Biol Chem* 244:3133–3141.
14. Potrykus K, Cashel M (2008) (p)ppGpp: Still magical? *Annu Rev Microbiol* 62:35–51.
15. Dalebroux ZD, Swanson MS (2012) ppGpp: Magic beyond RNA polymerase. *Nat Rev Microbiol* 10:203–212.
16. Hauryliuk V, Atkinson GC, Murakami KS, Tenson T, Gerdes K (2015) Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Microbiol* 13:298–309.
17. Cashel M, Rudd K (1987) The stringent response. *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed Neidhardt FC (American Society for Microbiology, Washington, DC), pp 1410–1438.
18. Traxler MF, et al. (2008) The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol Microbiol* 68:1128–1148.
19. Shivers RP, Sonenshein AL (2004) Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. *Mol Microbiol* 53:599–611.
20. Krásný L, Tiserová H, Jonák J, Rejman D, Sanderová H (2008) The identity of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in *Bacillus subtilis*. *Mol Microbiol* 69:42–54.
21. Wolz C, Geiger T, Goerke C (2010) The synthesis and function of the alarmone (p) ppGpp in firmicutes. *Int J Med Microbiol* 300:142–147.
22. Sonenshein AL (2005) CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. *Curr Opin Microbiol* 8:203–207.
23. Nelson JW, Atilho RM, Sherlock ME, Stockbridge RB, Breaker RR (2017) Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. *Mol Cell* 65:220–230.
24. Barrick JE, et al. (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci USA* 101:6421–6426.
25. Reiss CW, Xiong Y, Strobel SA (2017) Structural basis for ligand binding to the guanidine-I riboswitch. *Structure* 25:195–202.
26. Battaglia RA, Price IR, Ke A (2017) Structural basis for guanidine sensing by the *ykkC* family of riboswitches. *RNA* 23:578–585.
27. Meyer MM, et al. (2011) Challenges of ligand identification for riboswitch candidates. *RNA Biol* 8:5–10.
28. Soukup GA, Breaker RR (1999) Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* 5:1308–1325.
29. Regulski EE, Breaker RR (2008) In-line probing analysis of riboswitches. *Methods Mol Biol* 419:53–67.
30. Landick R, Wang D, Chan CL (1996) Quantitative analysis of transcriptional pausing by *Escherichia coli* RNA polymerase: His leader pause site as paradigm. *Methods Enzymol* 274:334–353.
31. Wickiser JK, Winkler WC, Breaker RR, Crothers DM (2005) The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* 18:49–60.
32. Wickiser JK, Cheah MT, Breaker RR, Crothers DM (2005) The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* 44:13404–13414.
33. Gilbert SD, Stoddard CD, Wise SJ, Batey RT (2006) Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *J Mol Biol* 359:754–768.
34. Haller A, Soulière MF, Micura R (2011) The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Acc Chem Res* 44:1339–1348.
35. Lemay J-F, et al. (2011) Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms. *PLoS Genet* 7:e1001278.

36. Winkler WC, Cohen-Chalamish S, Breaker RR (2002) An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci USA* 99:15908–15913.
37. McDaniel BAM, Grundy FJ, Artsimovitch I, Henkin TM (2003) Transcription termination control of the S box system: Direct measurement of *S*-adenosylmethionine by the leader RNA. *Proc Natl Acad Sci USA* 100:3083–3088.
38. Winkler WC, Nahvi A, Sudarsan N, Barrick JE, Breaker RR (2003) An mRNA structure that controls gene expression by binding *S*-adenosylmethionine. *Nat Struct Biol* 10:701–707.
39. Imaizumi A, Kojima H, Matsui K (2006) The effect of intracellular ppGpp levels on glutamate and lysine overproduction in *Escherichia coli*. *J Biotechnol* 125:328–337.
40. Cheng J, Guffanti AA, Krulwich TA (1997) A two-gene ABC-type transport system that extrudes Na^+ in *Bacillus subtilis* is induced by ethanol or protonophore. *Mol Microbiol* 23:1107–1120.
41. Nanamiya H, et al. (2008) Identification and functional analysis of novel (p)ppGpp synthetase genes in *Bacillus subtilis*. *Mol Microbiol* 67:291–304.
42. Geiger T, Wolz C (2014) Intersection of the stringent response and the CodY regulon in low GC Gram-positive bacteria. *Int J Med Microbiol* 304:150–155.
43. Geiger T, Kästle B, Gratani FL, Goerke C, Wolz C (2014) Two small (p)ppGpp synthases in *Staphylococcus aureus* mediate tolerance against cell envelope stress conditions. *J Bacteriol* 196:894–902.
44. Green NJ, Grundy FJ, Henkin TM (2010) The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett* 584:318–324.
45. Sudarsan N, et al. (2006) Tandem riboswitch architectures exhibit complex gene control functions. *Science* 314:300–304.
46. den Hengst CD, et al. (2005) The *Lactococcus lactis* CodY regulon: Identification of a conserved cis-regulatory element. *J Biol Chem* 280:34332–34342.
47. Sherlock ME, Sudarsan N, Stav S, Breaker RR (2018) Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria. *eLife* 7:e33908.
48. Chen AG, Sudarsan N, Breaker RR (2011) Mechanism for gene control by a natural allosteric group I ribozyme. *RNA* 17:1967–1972.
49. Nawrocki EP, Eddy SR (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29:2933–2935.
50. Weinberg Z, et al. (2015) New classes of self-cleaving ribozymes revealed by comparative genomics analysis. *Nat Chem Biol* 11:606–610.
51. Weinberg Z, Breaker RR (2011) R2R—Software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics* 12:3.
52. Sherlock ME, Malkowski SN, Breaker RR (2017) Biochemical validation of a second guanidine riboswitch class. *Biochemistry* 56:352–358.