Evidence that the *nadA* motif is a bacterial riboswitch for the ubiquitous enzyme cofactor NAD^+

SARAH N. MALKOWSKI,[1] TARA C.J. SPENCER,[2,6] and RONALD R. BREAKER[3,4,5]

[1]Department of Chemistry, Yale University, New Haven, Connecticut 06520-8103, USA

[2]Department of Biology, Howard University, Washington, D.C. 20059, USA

[3]Department of Molecular, Cellular and Developmental Biology, [4]Department of Molecular Biophysics and Biochemistry, [5]Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520-8103, USA

ABSTRACT

The *nadA* motif is a riboswitch candidate present in various Acidobacteria species that was previously identified by bioinformatic analysis of bacterial DNA data sets. More than 100 unique representatives have been identified exclusively upstream of *nadA* genes, which code for an enzyme in the biosynthetic pathway of the ubiquitous coenzyme NAD^+. The architecture of *nadA* motif RNAs suggests they use structurally similar tandem ligand-binding aptamer domains to control translation initiation. Biochemical analyses reveal that the first domain selectively binds ligands carrying an adenosine 5′-diphosphate (5′ ADP) moiety, including NAD^+ and its reduced form, NADH. Genetic analyses indicate that a tandem *nadA* motif RNA suppresses gene expression when NAD^+ is abundant, and that both aptamer domains are required for maximal gene regulation. However, we have not observed selective binding of the nicotinamide moiety of NAD^+ or binding by the second putative aptamer in vitro, despite sequence and structural similarities between the tandem domains.

Keywords: adenosine-5′-diphosphate; aptamer; gene control; nicotinamide adenine dinucleotide; noncoding RNA

INTRODUCTION

Many bacteria exploit metabolite- or inorganic ion-binding riboswitches (Breaker 2012; Peselis and Serganov 2014; Sherwood and Henkin 2016; Bastet et al. 2018; Lots and Suess 2018) to regulate gene expression in response to changes in important metabolites and other chemical signals. Approximately two-thirds of the more than 40 riboswitch classes experimentally validated to date (McCown et al. 2017) selectively respond to compounds that are derived from RNA, including numerous nucleotide-like enzyme cofactors. Specifically, known riboswitch classes sense and respond to several common enzyme cofactors such as adenosylcobalamin (AdoCbl) (Nahvi et al. 2002) and its derivatives (Nahvi et al. 2004; Johnson et al. 2012), thiamin pyrophosphate (TPP) (Mironov et al. 2002; Winkler et al. 2002a), flavin mononucleotide (FMN) (Mironov et al. 2002; Winkler et al. 2002b), *S*-adenosylmethionine (SAM) (Wang and Breaker 2008; Poiata et al. 2009; Weinberg et al. 2010; Mirihana Arachchilage et al. 2018), tetrahydrofolate (THF) (Ames et al. 2010; Chen et al. 2019), and molybdenum and tungsten cofactors (MoCo, WCo) (Regulski et al. 2008a). The striking abundance of riboswitches that sense these and numerous other RNA-like ligands supports the hypothesis that some modern riboswitch classes are descendants from molecular switches used by ancient RNA World life-forms (Breaker 2009, 2012).

Conspicuously absent from this list of nucleotide-derived cofactors that are sensed by riboswitches are two ubiquitous biomolecules: nicotinamide adenine dinucleotide (NAD^+) (Fig. 1A) or any of its naturally modified derivatives and coenzyme A (CoA) or any of its many thioester derivatives (Breaker 2011). Thus, we were particularly intrigued by the discovery of a novel but rare riboswitch candidate called the *nadA* motif (Weinberg et al. 2017). Although this small motif exhibited only marginal evidence for riboswitch function based on its initial reported characteristics, we chose to subject this motif to further bioinformatic, genetic, and biochemical analyses. Intriguingly, all *nadA* motif RNAs carry additional conserved features, including

[6]**Present address:** Division of Biological Sciences, University of California San Diego, La Jolla, California 92093, USA

Corresponding author: ronald.breaker@yale.edu

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.072538.119. Freely available online through the *RNA* Open Access option.

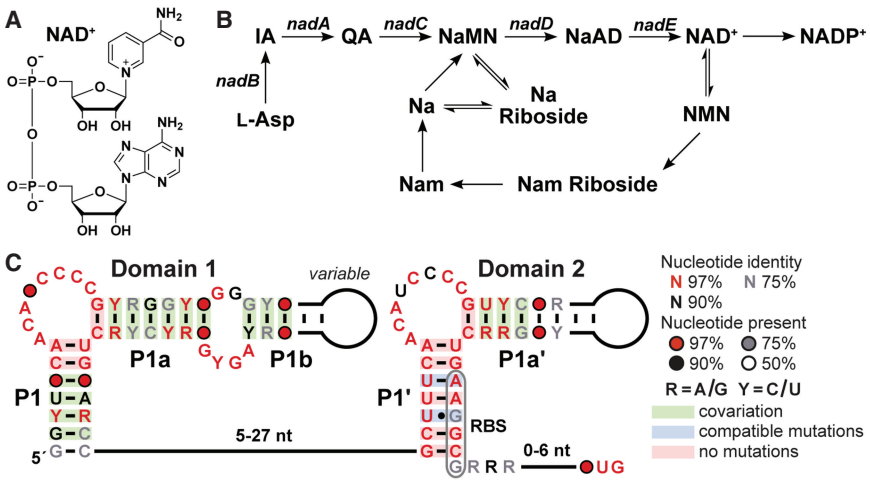

FIGURE 1. Tandem *nadA* motif domains are found upstream of genes in the biosynthetic pathway for NAD^+. (*A*) The chemical structure of NAD^+. (*B*) Biosynthetic pathway of NAD^+ and related molecules (Stancek et al. 2005; Zhou et al. 2011). Note that all *nadA* motif representatives are found exclusively upstream of *nadA* genes. (L-Asp) L-aspartate, (IA) iminoaspartate, (QA) quinolinic acid, (NaMN) nicotinic acid mononucleotide, (NaAD) nicotinic acid adenine dinucleotide, (NMN) nicotinamide mononucleotide, ($NADP^+$) nicotinamide adenosine dinucleotide phosphate, (Na) nicotinic acid, (Nam) nicotinamide. Riboside derivatives carry a ribose moiety linked by an N-glycosidic linkage to the nicotinic or nicotinamide ring nitrogen atom. (*C*) Consensus sequence and secondary structure model for 101 unique examples of tandem *nadA* motif RNAs. The putative ribosome binding site (RBS) is followed by a predicted start codon for the adjacent *nadA* open reading frame (ORF).

what appears to be a tandem aptamer architecture. Our findings strongly suggest that this tandem system functions as a riboswitch to regulate the expression of genes involved in NAD^+ biosynthesis.

Despite this data supporting our conclusion that *nadA* motif RNAs function as NAD^+ riboswitches, there remain some unusual findings with this riboswitch class. For example, in vitro assays reveal that NAD^+ is bound with an apparent 1:1 stoichiometry, and with K_D values that approach 100 µM. However, this binding is entirely due to recognition of the 5′ ADP moiety of the cofactor by the first aptamer domain. In addition, we have not yet observed ligand-induced structural modulation of the second putative aptamer, either in its native tandem arrangement or when examined independently. These biochemical characteristics observed in vitro suggest that there are unusual structural and functional features of this riboswitch class that remain to be established.

RESULTS AND DISCUSSION

Bioinformatic analyses of *nadA* motif RNAs reveal a tandem architecture

A consensus sequence and structural model for *nadA* motif RNAs as first reported (Weinberg et al. 2017) included a single hairpin structure with an internal bulge carrying seven highly conserved nucleotide positions. The original consensus model was based on 240 representatives with an average length of 66 nt. Despite the relatively simple structure and short length, this weak riboswitch candidate remained intriguing to us because of its common association with the *nadA* gene. This gene encodes an enzyme converting iminoaspartate to quinolinic acid, which is produced as an early step in the biosynthesis of NAD^+ (Fig. 1B; Gazzaniga et al. 2009).

While computationally searching for additional representatives, we noticed that each original *nadA* motif representative almost always resides adjacent to a second highly similar motif. Indeed, it is likely that all *nadA* motif RNAs are naturally formed using two similar conserved stem-bulge structures. Instances in which only a single stem-bulge structure is observed are all derived from metagenomic sequence data, in which the nucleotide positions that would correspond to the second putative aptamer are not available in the data set. These observations suggest that the tandem stem-bulge arrangement is important for the biochemical function of *nadA* RNAs.

After the natural tandem arrangement of this system was recognized, we developed distinct consensus sequence and structural models for the first and second domains to produce a revised depiction of *nadA* motif RNAs (Fig. 1C). Both domains form base-paired regions, called P1 and P1a, that flank an internal bulge. A prime symbol is used on the stem names for domain 2 to distinguish them from the similar stems of domain 1. Both domains carry a near identical "ACANCCCC" sequence in the P1 or P1a internal bulge. However, domain 1 is distinct in that it also carries an internal loop (with the sequence "AGYG") between stem P1a and an additional stem called P1b. Furthermore, domain 2 carries a purine-rich region in the right shoulder of stem P1′ that we predict serves as the RBS. We expect that ribosome access to the RBS is required for efficient initiation of translation from the neighboring start codon for the *nadA* open ORF.

Given the genomic location and the tandem arrangement of *nadA* motif RNAs, we initially hypothesized that each representative might function as a riboswitch to cooperatively bind two identical ligands related to NAD^+ biosynthesis. The same type of arrangement involving two highly similar aptamers and a single expression

platform is observed for many glycine riboswitches, in which evidence of cooperative ligand binding has been observed (Mandal et al. 2004; Butler et al. 2011). Cooperative binding would yield a steeper dose–response curve, thereby allowing cells to more responsively control the expression of a key gene for the production of this essential enzyme cofactor. However, it also seemed possible that juxtaposition of two highly similar aptamers could yield one of several other possible biochemical effects. For example, tandem aptamers that sense different compounds can yield Boolean logic gates (Stoddard and Batey 2006; Sudarsan et al. 2006; Lee et al. 2010), as has been most recently observed for tandem aptamers for guanine and phosphoribosyl pyrophosphate (PRPP) (Sherlock et al. 2018). Guanine binding suppresses gene expression related to purine biosynthesis via the tandem aptamer system, whereas PRPP binding overrides this suppression to promote the biosynthesis of ATP.

Regardless of the precise interplay between the two putative *nadA* aptamers, the architecture of each representative also strongly implicates a mechanism whereby ligand binding to domain 2 causes inhibition of translation initiation by exploiting a single expression platform. Specifically, the formation of P1a′ includes nucleotides of the RBS, and therefore the accumulation of NAD^+ (or a related compound) is predicted to promote the formation of the "OFF" state of the riboswitch. We speculate that the most likely ligand for such a riboswitch is NAD^+, because this essential coenzyme is present in bacterial cells at a concentration (~2.6 mM) that is much higher than any other similar compound in the biosynthetic pathway or among the collection of natural derivatives of this coenzyme, such as NADPH (~120 μM) or NADH (~83 μM) (Bennett et al. 2009).

Domain 1 of *nadA* motif RNAs binds NAD^+, NADH, and other compounds with a 5′ ADP moiety

Ligand-binding functions of *nadA* RNAs were assessed by conducting in-line probing assays (Soukup and Breaker 1999; Regulski and Breaker 2008b) using various-length constructs from several different bacterial species. For example, we examined a 61-nt RNA construct encompassing domain 1 of the *nadA* motif from *Acidobacterium ailaaui* (termed 61 *nadA*) (Fig. 2A). The banding pattern changes observed for the products of spontaneous RNA cleavage in the absence or presence of various concentrations of NADH (Fig. 2B) reveal that this enzyme cofactor causes substantial structural alterations in the RNA. Specifically, the nucleotides involved in these structural changes are largely confined to the bulge and internal loop regions that carry many highly conserved nucleotides. Furthermore, the dose–response curve generated from this data indicates that ligand binding most likely occurs with a 1-to-1 interaction and an apparent dissociation constant (K_D) of ~300 μM (Fig. 2C). These results suggest that the conserved nucleotides in the bulge and loop regions help form a binding pocket of an aptamer that becomes stabilized when the ligand is docked.

Very similar results are observed for domain 1 of a 75-nt *nadA* RNA construct from *Edaphobacter aggregans* (termed 75 *nadA*) (Fig. 2D). Notably, on the addition of NAD^+, we observed modulation of the regions of this RNA construct (Fig. 2E) that generally correspond to those undergoing modulation with the 61 *nadA* construct. A K_D of ~90 μM is estimated for NAD^+ binding by the 75 *nadA* RNA construct (Fig. 2F). Very similar results are observed for 75 *nadA* RNA binding to NADH (Supplemental Fig. S1). Also, mutations to highly conserved nucleotides of the 75 *nadA* RNA (Fig. 2D) cause a complete loss of NAD^+ binding at the highest concentration tested (1 mM, data not shown), indicating that these positions are critical for the ligand-binding function of the RNA.

The 75 *nadA* RNA construct was used to conduct a broader structure–activity relationship (SAR) analysis by examining the ability of various analogs or fragments of NAD^+ and NADH to be bound by a domain 1 aptamer (Fig. 3A). Initially, in-line probing assays were conducted at a concentration of 1 mM of each candidate ligand (Supplemental Fig. S2). Binding assays were then expanded to establish K_D values for those compounds that triggered structural modulation. From these values (Fig. 3A, B; Supplemental Fig. S3), we can conclude that domain 1 of the *E. aggregans nadA* RNA exhibits selective binding only of the 5′ ADP moiety of the NAD^+ cofactor. Similar results are observed for domain 1 constructs from other organisms (data not shown).

By examining the ligand function of a series of 5′ ADP analogs (Fig. 3A; Supplemental Fig. S3), an interaction map was developed to identify key functional groups on the nucleotide moiety important for selective recognition by this aptamer (Fig. 3C, left). Removing the 2′ hydroxyl group or replacing with a bulkier phosphate is not tolerated (Supplemental Fig. S3), suggesting the presence of an important hydrogen bond contact between the hydroxyl group and the aptamer. When changing the number of phosphates on the 5′ end of ADP, optimal binding is observed with di- or triphosphates. This observation suggests that the aptamer binding pocket might use hydrogen bonding or metal ion coordination with two phosphates of the ligand.

The results of modifications to the adenine base further highlight key interactions. Introducing bulky functional groups at various sites on the purine ring decrease or entirely eliminate binding, suggesting the RNA structure uses a binding pocket shape that rejects most analogs that carry additional bulk on the purine ring. One exception is observed for the exocyclic amine at position 6. Binding appears to be unaffected by the presence of a methyl group at this position (Fig. 3; N^6-CH_3-ATP).

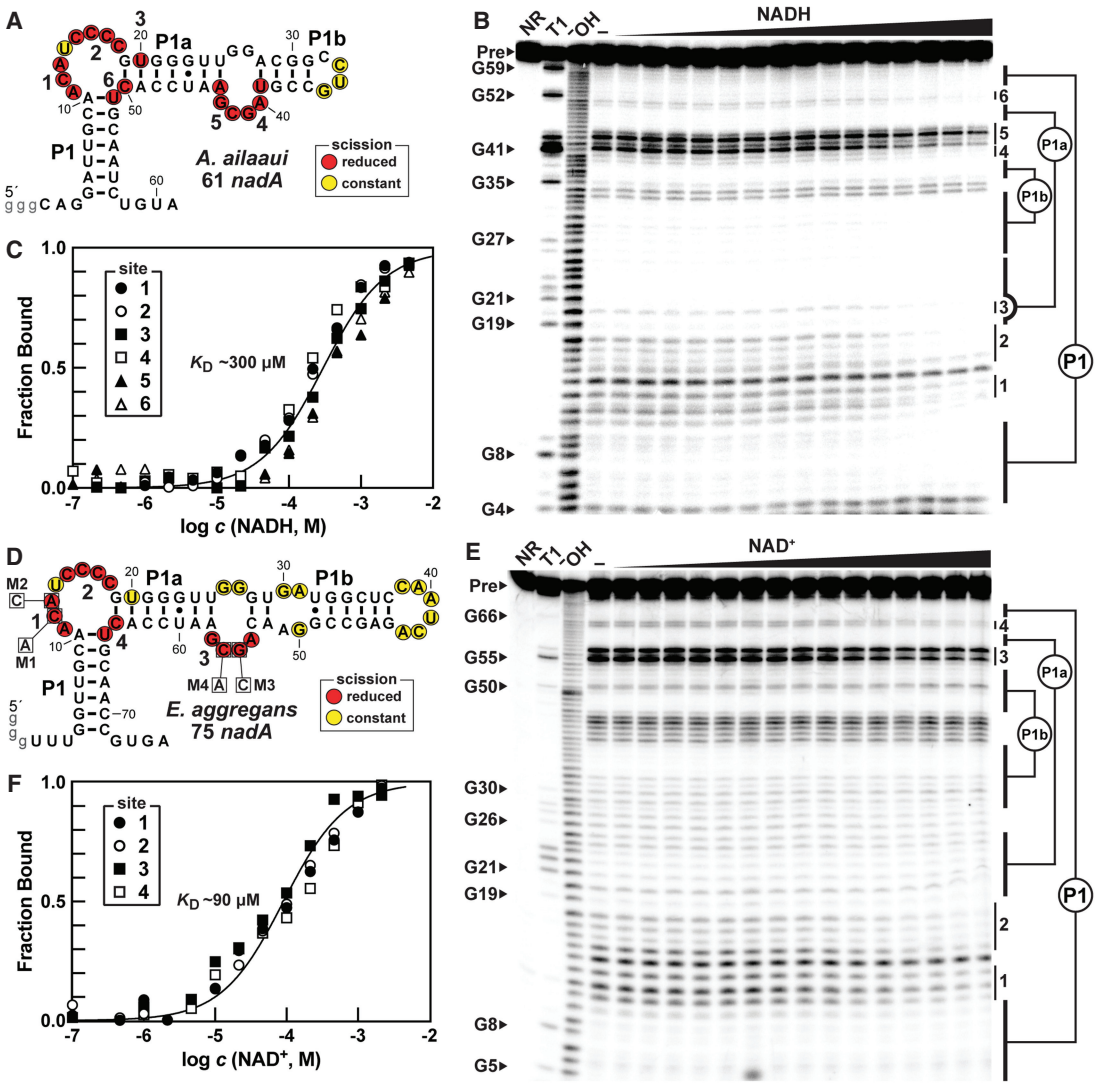

FIGURE 2. Domain 1 of *nadA* motif RNAs bind NAD^+ and NADH in vitro. (*A*) Sequence and secondary structure model of the 61 *nadA* RNA construct from *A. ailaaui*. Lowercase nucleotides on the 5′ terminus designate G residues added to facilitate production by T7 RNA polymerase. Sites of spontaneous RNA cleavage revealed by an in-line probing assay depicted in *B* are circled. Red circles identify nucleotide linkages that undergo suppression of spontaneous cleavage in the presence of added ligand, whereas yellow circles identify linkages that cleave at the same rate regardless of ligand addition. Numbered regions were used to quantify ligand binding as presented in *C*. (*B*) Denaturing (8 M urea) polyacrylamide gel electrophoresis (PAGE) analysis of in-line probing reactions conducted with 5′ ^{32}P-labeled 61 *nadA* RNA in the absence (–) or presence of NADH, ranging from 100 to 4.64 mM. NR, T1, and $^-$OH indicate no reaction, partial digestion with T1 ribonuclease (cleaves after every G), and partial digestion under alkaline conditions (cleaves after every nucleotide), respectively. Precursor RNA (Pre) and certain G residues are annotated, and regions of modulation (vertical bars) are numbered as designated in *A*. (*C*) Plot of the fraction of RNA bound to ligand versus the logarithm of the NADH concentration as determined by quantification of the PAGE data in *B*. See the Materials and Methods section for details. (*D*) Sequence and secondary structure model of the 75 *nadA* RNA construct from *E. aggregans*, containing only domain 1. Mutations M1–M4 alter highly conserved (>97%) nucleotides (boxed and annotated with the mutant nucleotide identity) to assess their requirement for ligand binding activity. (*E*) PAGE analysis of in-line probing reactions of 5′ ^{32}P-labeled 75 *nadA* RNA in the absence (–) or presence of NAD^+ ranging from 100 to 4.64 mM. Annotations are as describe for *B*. (*F*) Plot of the fraction of RNA bound to ligand versus the logarithm of the NAD^+ concentration as determined by quantification of the PAGE data in *E*.

However, the absence of the amine group, as with inosine 5′-triphosphate (ITP) (Supplemental Fig. S3), causes a loss of binding affinity. These data suggest the RNA forms a single hydrogen bond with the exocyclic amine of the adenine nucleobase, even when the nitrogen is presented as a secondary amine.

The binding assay results described above indicate that domain 1 portions of tandem *nadA* motif RNAs selectively bind ADP and various derivatives. However, they do not recognize the nicotinamide ring nor do they bind the ribose portion of the NMN moiety. The affinities measured for ADP binding are more than 10-fold better than the

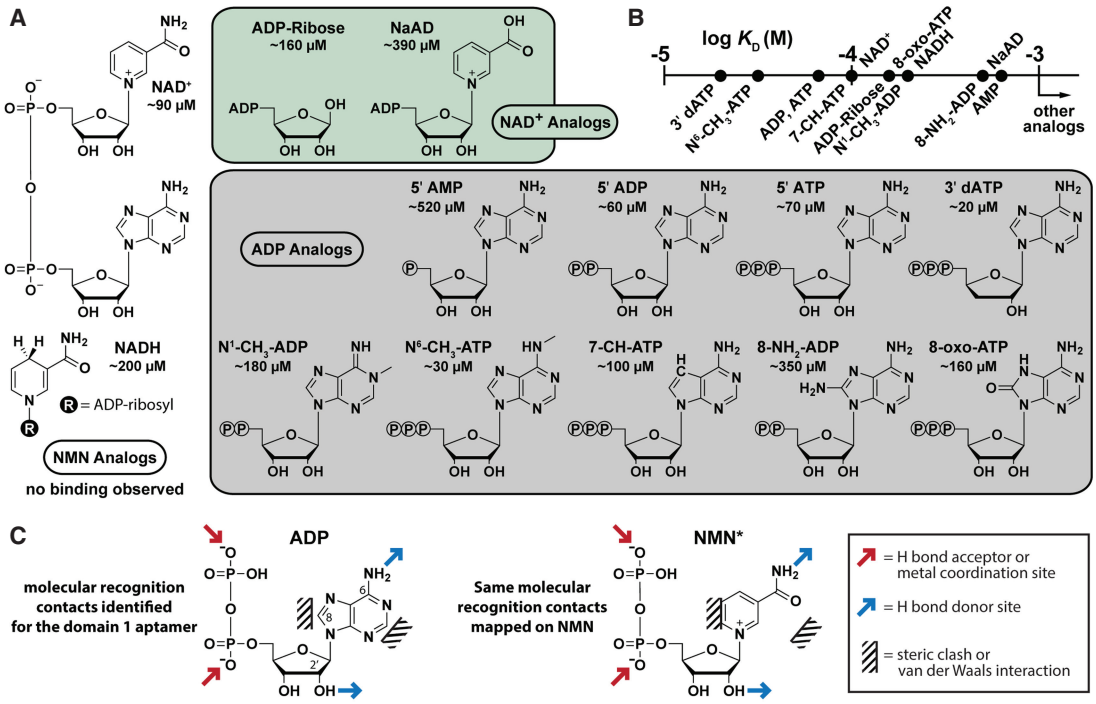

FIGURE 3. SAR data for ligand binding by domain 1 of a *nadA* RNA construct. (*A*) Chemical structures of various ligand candidates examined for binding of the 75 *nadA* RNA construct from *E. aggregans* (Fig. 2D) by using in-line probing assays. All compounds were initially screened for binding at 1 mM, and a dissociation constant was obtained for compounds that induced structural modulation. (*B*) The plot of the K_D values for compounds recognized by domain 1, as determined by quantification of PAGE analyses of in-line probing reactions (data not shown). (*C*) (*Left*) The chemical structure of 5′ ADP and the proposed sites of molecular recognition by the RNA aptamer formed by domain 1. Arrows represent proposed sites for hydrogen bond acceptor (red) and donor (blue) function. (*Right*) The chemical structure of phosphorylated NMN (NMN*) depicted with the recognition sites used by domain 1 to bind ADP.

cellular concentration of NAD^+ (2.6 mM), ADP (0.56 mM), and ATP (9.6 mM), as measured for *Escherichia coli* (Bennett et al. 2009). Likewise, the best K_D value measured is for dADP (~20 μM) (Fig. 3A,B). However, some riboswitches are known to operate as kinetically driven gene control elements (Wickiser et al. 2005a,b; Gilbert et al. 2006; Haller et al. 2011; Lemay et al. 2011; Frieda and Block, 2012), and therefore the functional relevance of comparisons between K_D values and in vivo ligand concentrations must be assessed with caution.

For example, a riboswitch for the bacterial second messenger c-di-AMP can bind ATP with an affinity near the concentration of cellular ATP concentrations (Watson and Fedor 2012; Nelson et al. 2013). Despite these binding characteristics, ATP does not trigger c-di-AMP riboswitches during in vitro transcription reactions, despite the high concentration of ATP present in such mixtures (Nelson et al. 2013). This is most likely due to the need for the ligand to bind rapidly, before RNA polymerase reaches a point in the template beyond the intrinsic terminator stem (Wilson and von Hippel 1995; Yarnell and Roberts 1999), when it becomes committed to synthesizing the complete mRNA.

Furthermore, because *nadA* motif RNAs regulate gene expression related to NAD^+ biosynthesis, it seems advantageous to selectively recognize this cofactor directly, rather than more broadly respond to ADP, ATP, or the general pool of compounds that carry the ADP (or dADP) moiety. A superficial comparison of the molecular recognition contacts made by the domain 1 ADP aptamer (Fig. 3C, left) and the structure of NMN (or its pyrophosphate analog) (Fig. 3C, right) reveal that there is a remarkable similarity in the structures of the two major chemical moieties of NAD^+ that is recognized by the domain 1 aptamer. However, we predict that the steric block to modifications at the C8 position of ADP would likely weaken the binding of the NMN moiety of NAD^+ by the domain 1 aptamer.

Given the observations above, it is tempting to speculate that the unusual characteristics of both aptamers might still be explained by considering both the imperfect tandem architecture of *nadA* motif RNAs and the imperfect symmetry of NAD^+. The consensus models for domains 1 and 2 are highly similar, but not identical. Most notably, domain 2 lacks the internal bulge that forms the exceptionally well-conserved junction between P1a and P1b stems of domain 1 (Fig. 1C). Perhaps domain 1 has evolved to selectively bind the 5′-ADP (or 5′-AMP) portion of the cofactor, whereas domain 2 has evolved to recognize only the NMN portion. Together, the two aptamers could collaborate to form a near-symmetrical architecture

that selectively binds and responds to NAD^+. Aspects of this hypothesis were explored in greater detail as described below.

Ligand binding by domain 2 of *nadA* motif RNAs has not been observed

Despite the strong similarity in both sequence and secondary structure between the two domains of *nadA* motif RNAs, we have been unable to induce structural modulation in domain 2 by using 5′ ADP, NAD^+, or any of the various analogs or fragments presented in this report (data not shown). Various additional RNA constructs (Supplemental Fig. S4) carrying either domain 2 alone (10 constructs from different organisms) or both domains 1 and 2 (12 constructs from different organisms) have been subjected to in-line probing analyses. Intriguingly, ligand binding by domain 1 is consistently observed with all constructs carrying this aptamer, but we have not detected ligand-mediated structural modulation of domain 2 with any of the single or tandem arrangements examined.

In other natural tandem aptamer (cooperative riboswitch) systems, ligand binding by one aptamer improves the binding affinity of its neighboring aptamer (Mandal et al. 2004; Sherlock et al. 2017). Thus, if aptamer function by domain 2 is completely dependent on ligand binding by the adjacent domain 1 aptamer, then it is not surprising that no evidence can be found for ligand binding by RNA constructs formed using only domain 2 of a naturally tandem *nadA* motif representative. Even when *nadA* motif constructs carry both domains, it is possible that the second aptamer might be rendered inoperable because they are thermodynamically trapped in an inactive state. This is the case for some riboswitches in which the expression platform precludes ligand binding because this mutually exclusive structure is more stable than the ligand-binding aptamer conformation (Wickiser et al. 2005a; Atilho et al. 2019a). In some instances, in-line probing results indicate that domain 2 indeed is poorly formed. However, we have not observed structural modulation, for any of the tandem constructs examined in the current study, that are indicative of binding to domain 2. It seems unlikely that all constructs tested have folding problems that preclude the ligand-binding function of domain 2.

There remain other reasonable possible explanations for the absence of binding by domain 2. For example, perhaps there is a co-ligand missing from our in vitro binding assays. Some other riboswitch classes appear to rely on the presence of specific ion partners to bind their target ligand, such as a K^+ ion involved in ligand binding by lysine riboswitches (Serganov et al. 2008) and Mg^{2+} ions required by fluoride riboswitches (Ren et al. 2012). Although both ions are abundant in our in-line probing reactions, it is possible that some other co-ligand is absent.

Nicotinic acid biosynthesis is important for gene repression by a *nadA* motif element

Given the unusual characteristics of *nadA* motif constructs used for in vitro ligand binding assays, we sought additional evidence of gene regulation function in cells. Genetic fusions between a specific natural *nadA* motif representative and a β-galactosidase (*lacZ*) gene from *E. coli* were created and examined under various conditions in vivo. The *nadA* representative from the bacterium *Koribacter versatilis* (Fig. 4A) was chosen for in vivo experiments because this organism was one of only three fully sequenced genomes containing the *nadA* RNA motif. This choice also allowed us to ensure both the accuracy of the *nadA* motif sequence and the inclusion of the complete natural genetic context when preparing the reporter-fusion construct. This construct should overcome any possible construct or reaction condition issues experienced during the in vitro binding assays described above.

Reporter constructs carrying the WT and various mutant *K. versatilis nadA* motif sequences were examined for gene expression differences in *E. coli* cells grown in rich (LB) media (Fig 4B, filled bars). The lowest level of reporter gene expression is exhibited by cells carrying the WT riboswitch reporter construct. When nucleotides are mutated to disrupt base-pairing of the P1′ stem (M1) and better expose the RBS, gene expression increases by approximately threefold. Although this extent of gene expression change is modest, it is similar to the level of gene expression change observed with some other riboswitches (Lee et al. 2009; Caron et al. 2012). Also, additional results discussed in greater detail below indicate that the natural dynamic range is much greater than the value suggested by comparing only WT and M1 constructs.

Mutant constructs carrying a single mutation to a highly conserved nucleotide in the internal bulge of domain 1 (M2), domain 2 (M3), or both (M4) exhibit intermediate gene expression levels compared to the WT and M1 constructs. This data is consistent with our hypothesis that nucleotides within the highly conserved bulge joining P1 and P1a of domain 1, and the analogous nucleotides of domain 2, are important for riboswitch function. Furthermore, the results are consistent with a genetic "OFF" function for the riboswitch upon ligand binding. Mutations that disrupt the anti-RBS structure (P1′), or that disrupt the putative ligand binding pockets (M2 and M3) all cause an increase in gene expression. Importantly, the combination of mutations M2 and M3 to create construct M4 causes higher gene expression than observed for either of the independent M2 and M3 mutations. This is consistent with the possibility that ligand binding by two aptamers is involved in suppressing gene expression.

Given the modest dynamic range between WT and M4 in WT *E. coli* cells, we sought ways to manipulate the

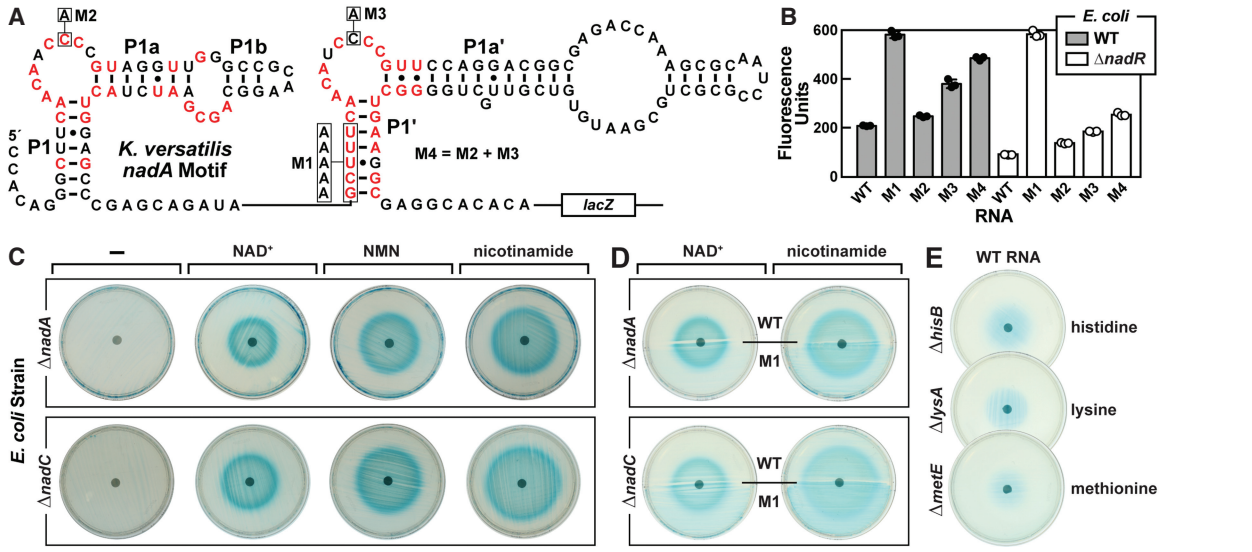

FIGURE 4. Gene regulation by a tandem *nadA* motif RNA. (*A*) Sequence and secondary structure of the wild-type (WT) *nadA* RNA aptamer from *K. versatilis*. Red nucleotides are >97% conserved as depicted in the *nadA* motif consensus model (Fig. 1C). Mutant constructs (M1–M4) carry nucleotide changes at the indicated positions. (*B*) The plot of the relative β-galactosidase expression of an *E. coli* strain carrying a WT or mutant *nadA* riboswitch-*lacZ* reporter fusion construct in WT cells or a genomic disruption to the *nadR* gene. Cells were grown in LB for 6 h before quantification. (*C*) M9 minimal media agar plates containing 100 μg mL^{-1} X-gal were streaked with *E. coli* strains carrying a WT riboswitch reporter fusion construct and with genomic disruptions to *nadA* or *nadC* genes. Plates contain a filter disk supplemented with 5 μL dH_2O (–) or with 5 μL of a 10 mM solution containing a nicotinamide derivative. (*D*) M9 minimal media agar plates containing 100 μg mL^{-1} X-gal were streaked with *E. coli* strains with genomic disruptions to *nadA* or *nadC* genes and carrying a WT or M1 riboswitch reporter fusion construct. Plates contain a filter disk supplemented with the nicotinamide derivative noted. (*E*) Agar diffusion assays using the WT riboswitch reporter fusion construct hosted by various additional strains of *E. coli* that are auxotrophic for compounds unrelated to NAD^+. Filter disks were supplemented with 5 μL of a 10 mM solution containing the amino acid noted. Additional details are as described for *C*.

cellular concentration of NAD^+ to determine if the riboswitch could more robustly modulate gene expression. Previous studies have revealed that NAD^+ concentration in various bacterial species, including *E. coli*, is regulated by a protein repressor called NadR (Grose et al. 2005). Deletion of the *nadR* gene is expected to increase the concentration of NAD^+, which should further suppress reporter gene expression if the *nadA* motif functions as an "OFF" switch when this enzyme cofactor is bound by the RNA. We used an *E. coli* strain carrying a knockout of the *nadR* gene to evaluate the same *nadA* motif reporter gene constructs described above (Fig. 4A). Indeed, gene expression levels are lowered for the WT reporter construct and for the mutant constructs M2–M4 (Fig. 4B, open bars). However, the same trend in gene expression is retained, again suggesting that ligand binding to domains 1 and 2 provide the largest reductions in gene expression.

We also manipulated the concentrations of NAD^+ in cells by using *E. coli* strains that carry knockouts of *nadA* or *nadC* genes, which produce proteins essential for the production of this enzyme cofactor. For growth, these strains require supplementation with compounds carrying the nicotinamide moiety when grown on otherwise minimal medium (such as M9). Agar diffusion assays either with *E. coli* strain Δ*nadA* or Δ*nadC* carrying the WT *nadA* motif reporter construct yield growth and gene expression patterns that again are consistent with our hypothesis that *nadA* motif RNAs function as genetic "OFF" riboswitches for NAD^+. In each instance, the addition of compounds that carry a nicotinamide moiety permit cell growth near the filter disk carrying the supplemented compound, and reporter gene expression is greatest at the boundary between robust cell growth and the lack of growth when cells are starved for NAD^+ (Fig. 4C). This gene expression "halo" pattern has been observed for other riboswitch classes subjected to similar agar diffusion assays (Kim et al. 2015; Atilho et al. 2019a).

To make direct comparisons between cells carrying the WT or M1 reporter constructs on agar media, samples of each type (either in the Δ*nadA* or Δ*nadC* genetic background) were plated on hemispheres of the same Petri dish. Agar diffusion assays (Fig. 4D) reveal that the plated M1 cells exhibit a higher background level of gene expression, as expected given the mutations in this construct weaken the anti-RBS stem (P1′) and increase reporter gene expression (Fig. 4B). In addition, the M1 construct exhibits a maximum level of gene expression that visually appears to be similar to that exhibited by the WT construct when both are in cells starved for NAD^+ (blue halo intensity). This observation indicates that M1 only incompletely activated gene expression, and that the mutations in the M1 *nadA* RNA construct do not otherwise prevent partial

suppression of gene expression when the ligand is present in abundance.

To rule out the possibility that cells undergoing general metabolite starvation cause an up-regulation of the WT *nadA* reporter construct, we evaluated reporter gene expression in three additional *E. coli* strains that are auxotrophic for histidine, lysine, or methionine. None of these strains exhibit the gene expression halo on the addition of their essential metabolite (Fig. 4E), indicating that the riboswitch is selective for NAD^+, rather than responding to a lack of other metabolites or general cellular distress.

Concluding remarks

In the current study, we provide evidence for the existence of a naturally occurring NAD^+-sensing riboswitch. These RNAs use a unique architecture featuring two nearly identical domains in tandem (Fig. 1C). The majority of the previously validated riboswitch classes sense metabolites that are derived from RNA nucleotides or their precursors (McCown et al. 2017). This pattern supports the hypothesis that some riboswitch classes might have their evolutionary origins in the RNA World (Breaker 2009, 2012) and thus are also being triggered by ancient biological compounds. Consistent with this view is the fact that the common nucleotide-derived enzyme cofactors are sensed by the largest number of riboswitch classes compared to any other ligand category. These cofactors have been proposed to have been first formed by RNA World organisms (Woese 1967; White 1976; Benner et al. 1989) and are essential for biochemical functions in species from all domains of life. Thus, the existence of riboswitches for another nucleotide-derived coenzyme, NAD^+, fits well with the remarkable bias observed for riboswitch ligands from modern cells.

Riboswitches that sense the enzyme cofactors TPP, AdoCbl, and SAM constitute the top three most abundant riboswitch classes in bacteria (McCown et al. 2017). Moreover, riboswitches for the cofactors FMN, THF, and MoCo/WCo are also relatively abundant. Given the abundances of these prominent riboswitches for coenzymes from the RNA World, the previous absence of riboswitches for two ubiquitous nucleotide-based coenzymes NAD^+ and CoA had become increasingly conspicuous (Breaker 2011). NAD^+ is essential for cellular function, and this compound might have interacted with RNA molecules, including riboswitches, long before the emergence of proteins in evolution (Breaker 2009, 2012). Thus, it is not surprising that a candidate riboswitch class was encountered that likely serves as a sensor for NAD^+. If the *nadA* system indeed constitutes an NAD^+ riboswitch, this provides further support for the existence of an RNA world. However, there are several unusual characteristics of the *nadA* motif system that remain to be investigated, as briefly highlighted below.

The observation that examples of the original *nadA* motif RNAs always reside in tandem allowed us to create distinct consensus models for the two putative aptamer domains. The relative similarity between these two domains suggested to us that they might bind two separate, but identical ligands, as occurs for other homologous tandem aptamer systems. For example, the tandem aptamers in glycine riboswitches cooperatively bind two separate glycine molecules (Mandal et al. 2004; Butler et al. 2011). Likewise, tandem arrangements of independently functioning riboswitches with highly similar aptamers also can bind two of the same ligands, such as observed for TPP (Welz and Breaker 2007). Given these precedents, the simplest explanation for the tandem *nadA* system associated with a single expression platform is that they would cooperatively bind two identical ligands to repress translation initiation.

The most logical ligand candidate for *nadA* RNAs is NAD^+, based on the fact that it is the primary product of the biosynthetic process associated with this candidate riboswitch class (Weinberg et al. 2017), and based on the fact that NAD^+ is present in the highest concentration compared to all other derivatives or biosynthetic intermediates of this coenzyme (Bennett et al. 2009). The expression platform for *nadA* RNAs is predicted to repress gene expression in the presence of ligand, functioning as an "OFF" switch. This logic is similar to previously validated riboswitches that recognize coenzymes (Winkler et al. 2002a,b; Nahvi et al. 2004). All *nadA* RNA representatives are found upstream of *nadA*, a gene in the biosynthetic pathway of NAD^+. Indeed, our genetic data support the hypothesis that *nadA* motif RNAs function as riboswitches for NAD^+ (or a closely related compound). However, other more complicated explanations for these findings cannot be ruled out at this time.

Surprisingly, ligand binding studies in vitro reveal that domain 1 selectively binds ADP or its close derivatives, whereas domain 2 is unresponsive to these and other ligands tested. These findings leave several important unanswered questions. For example, why is only the ADP moiety of NAD^+ sensed by the first aptamer? It is possible the second domain is a relic of a once active aptamer and simply serves as a structure to support the first aptamer and to host the expression platform? There is precedence for the use of "ghost" aptamers by glycine riboswitches, wherein the second aptamer-like structure actually does not bind ligand (Ruff et al. 2016). However, it seems unlikely that ghost aptamer function is involved in the *nadA* system. Domain 2 of each *nadA* RNA carries most of the same nucleotides that are conserved in domain 1, which seems unnecessary if domain 2 does not bind a similar ligand.

Although domain 1 RNA constructs bind NAD^+ and NADH with biologically relevant affinities (Figs. 2 and 3), the riboswitch system would need to prevent being inappropriately triggered by compounds such as ADP and

ATP, which are present in cells at concentrations nearly equivalent to NAD^+. It seems unlikely to us that ADP (or ATP) is the natural ligand. NAD^+ and NADH play essential roles in the oxidative phosphorylation process that cells use to most efficiently convert ADP to ATP, which provides a reason why cells might use a riboswitch for ADP to regulate the biosynthesis of NAD^+. However, the tandem *nadA* motif never associates with genes unrelated to NAD^+ biosynthesis, whereas a riboswitch for ADP or ATP might be expected to be used to regulate other genes as well.

One way to preclude a genetic response to compounds carrying an ADP moiety, and thereby only respond to the desired ligand, such as NAD^+, is to exploit differences in the rate constant for ligand association. This biochemical strategy is used by a c-di-AMP riboswitch to prevent misregulation by high concentrations of ATP (Nelson et al. 2013). To use this strategy, the aptamer system must use chemical interactions to favor the swifter binding of the desired molecule compared to the natural ligand analogs. Thus, if NAD^+ is the true ligand for *nadA* motif RNAs, then domain 2 seems most likely to provide the missing pocket for selective recognition of this enzyme cofactor.

Our findings could be explained if the two domains function as a roughly symmetrical aptamer system wherein domain 1 binds the ADP moiety and domain 2 binds the NMN moiety, such that the tandem system actually recognizes only one ligand, NAD^+. The positioning of the pyridine and amide moieties of the nicotinamide base is similar to that of the purine ring and primary amine moieties of ADP (Fig. 3C). Furthermore, both the phosphorylated NMN moiety and ADP carry identical ribose and pyrophosphate groups. Thus, domain 2 of *nadA* motif RNAs might recognize phosphorylated NMN with similar interactions as used by domain 1 to recognize 5′ ADP. However, this proposal does not explain why structural modulation has not been observed in domain 2 upon the addition of candidate ligands to RNA constructs carrying both domains, or domain 2 alone.

Neither of the two domains of *nadA* motif RNAs is found independent of the other in nature, wherein the consensus sequences match perfectly. This suggests that ligand recognition and/or riboswitch function in cells requires both domains in the order observed. However, we carefully considered the possibility that the domain 1 aptamer, which strongly binds molecules such as ADP, dADP, and ATP, might have homologs in other bacterial species (Fig. 3A, B). Intriguingly, we have identified nine representatives of a remarkably close variant class in several species of *Nocardia* that differ from *nadA* domain 1 aptamers in only a few key locations (Supplemental Fig. S5). This novel variant is called the *ubiB* motif because it associates with the gene of the same name (also called *aarF*), which codes for a monooxygenase enzyme that catalyzes the first step in ubiquinone biosynthesis (Poon et al. 2000). Moreover, there appears to be a strong hairpin that includes the RBS sequence for the adjacent coding region, which suggests that *ubiB* motif RNAs regulate translation initiation. Also, given the known ligand binding properties of *nadA* domain 1, it is intriguing to note that the ubiquinone cofactor is a critical component of the electron transport chain, which ultimately converts ADP into ATP (Anraku 1988).

For the *nadA* motif, we currently favor the hypothesis that the second aptamer might require a co-ligand that is needed to support the binding of the nicotinamide moiety of NAD^+. This co-ligand might not be directly related to the abundance of NAD^+ but is an ordinary and abundant constituent of bacterial cells that assists binding. This arrangement would be similar to the role of Mg^{2+} in many riboswitch classes, or perhaps more similar to the role of K^+ in lysine riboswitches (Serganov et al. 2008). To address the question of ligand binding by domain 2, a search could be conducted to determine if cell extracts can promote ligand binding. Also, the determination of the structure of the complex between ADP and a domain 1 RNA construct would provide insight into the roles of the conserved nucleotides in this aptamer. This would hopefully reveal the nature of the putative ligand binding site in domain 2, based on its conserved nucleotides.

Even if *nadA* motif RNAs are proven by biochemical or biophysical methods to selectively recognize NAD^+ or a natural compound related to this enzyme cofactor, this still would not entirely address the relative scarcity of riboswitches for this essential compound. Unlike riboswitches for some of the other ubiquitous cofactors, perhaps there are competing protein factors or other riboswitch classes that nature has used to distribute the molecular sensing and gene regulation duties. We have uncovered another apparently unrelated riboswitch candidate in certain bacteria that uniformly associate with genes for NMN transport (Higgs G, Breaker RR, unpubl.). Although this motif is also rare, it might help explain the narrow distribution and relative rarity of *nadA* motif RNAs. If representatives of NAD^+ riboswitches can be further experimentally validated, then the absence of riboswitch classes for CoA, and also for cofactors such as pyridoxal phosphate and biotin will become even more conspicuous.

MATERIALS AND METHODS

Chemicals, DNA oligonucleotides, and bacterial strains

Chemicals were purchased from Sigma-Aldrich with the following exceptions: NADPH (Santa Cruz Biotechnology); 7-deaza-ATP, 2-amino-ATP, N^6-methyl-ATP, and 8-oxo-ATP (TriLink BioTechnologies); 2-hydroxy-ATP, 1-CH_3-ATP, 3′-dATP (Jena Bioscience); 8-NH_2-ADP and β-nicotinamide-1,N^6-ethenoadenine dinucleotide (Biolog). Also, the radiolabeled nucleotide [γ-^{32}P]-ATP was purchased from PerkinElmer. All enzymes were purchased from New England BioLabs, unless otherwise noted.

DNA oligonucleotides were purchased from Sigma-Aldrich. A list of oligonucleotides used in this study can be found in Supplemental Table S1. *Escherichia coli* strain BW25113 and the corresponding mutant strains *ΔnadA* (CGSC 8797), *ΔnadC* (CGSC 8387), and *ΔnadR* (CGSC 11585) were obtained from the Coli Genetic Stock Center (CGSC) at Yale University.

Bioinformatic analyses

Additional homologous *nadA* motif RNAs were identified using Infernal 1.1 (Nawrocki and Eddy 2013). Bioinformatic searches were conducted using RefSeq (O'Leary et al. 2016) version 80 and various metagenomic databases as described previously (Weinberg et al. 2017). A total of 101 unique representatives, all exhibiting tandem architectures, were used to generate an updated consensus model (Fig. 1C). The consensus sequence and structural models were constructed with the aid of the program R2R (Weinberg and Breaker 2011).

The *nadA* domain 1 variant motif *ubiB* also was identified using Infernal 1.1 from RefSeq version 80 and environmental microbial databases as described elsewhere (Atilho et al. 2019a). The seed alignment contained only the first aptamer of *nadA* motif RNAs. In addition, the *E*-value threshold was raised to 1000 to search for distal variants as described elsewhere (McCown et al. 2017). One *ubiB* motif representative was identified from this search. Additional examples were subsequently identified by using Infernal 1.1 and the initial hit as the seed sequence. Iterative searches were conducted as the alignment of *ubiB* motif RNAs expanded.

RNA oligonucleotide preparation

RNA oligonucleotide preparation was performed as previously described (Nelson et al. 2017; Sherlock et al. 2017). Briefly, DNA oligonucleotides containing a T7 RNA polymerase promoter sequence upstream of the desired sequence were used to prepare templates for in vitro RNA transcription. Synthetic DNA oligonucleotides (<120 nt) were combined in equimolar amounts with their complementary strand to form a double-stranded DNA template. Longer DNA oligonucleotides (>120 nt) containing overlapping regions were extended using SuperScript II reverse transcriptase (Thermo Fisher Scientific). RNA constructs were then prepared by in vitro transcription overnight at 37°C in a 50 µL reaction. The resulting RNA transcripts were purified by denaturing (8 M urea) 10% PAGE. Purified RNA was dephosphorylated with rAPid alkaline phosphatase (Roche Applied Science) according to the manufacturer's protocol, enzymatically 5′ ^{32}P-labeled with [γ-^{32}P]-ATP, and purified as described above.

In-line probing analyses

In-line probing assays were performed as previously described (Soukup and Breaker 1999; Regulski and Breaker 2008b). Briefly, trace amounts of 5′ ^{32}P-labeled RNAs were incubated at 23°C for 42–72 h in the presence or absence of different concentrations of desired ligand candidates in the presence of 20 mM $MgCl_2$, 100 mM KCl, and 50 mM Tris-HCl (pH 8.3 at 23°C). The products of spontaneous RNA cleavage were separated by denaturing 10% PAGE and visualized with a Typhoon PhosphorImager (GE Healthcare).

Dissociation constants were calculated by quantifying the changes in band intensities using ImageQuaNT software (GE Healthcare) to identify internucleotide linkages that exhibit structural modulation. Values for band intensities were normalized to a band that did not alter upon addition of ligand and plotted as a function of the logarithm of the ligand concentration using GraphPad Prism 7. Each set of points was fit to a sigmoidal four-parameter logistic fit with the Hill coefficient constrained to 1. When bands did not reach a plateau in cleavage intensity for a given in-line probing experiment, the theoretical minimal or maximal values were computed. These values were used to scale the intensities at each ligand concentration used, and thus yielded the fraction of RNA bound to ligand values for each in-line probing reaction. These values were plotted as a function of the logarithm of the ligand concentration, and the points were again fit to a sigmoidal four-parameter logistic fit with the following constraints: top value set to 1 (maximum fraction bound), bottom set to 0, and Hill coefficient set to 1. The apparent K_D value was obtained from this data.

Riboswitch reporter constructs

The riboswitch reporter fusion construct was designed to include the *B. subtilis lysC* promoter, which has no known regulation and therefore is constitutively active. The RNA transcript includes a *nadA* motif representative from *K. versatilis* (NC_008009.1) and the first nine codons of the downstream *nadA* ORF. This construct was synthesized by GeneArt and inserted into the vector pRS414 as a translational fusion with the ninth codon of the *E. coli lacZ* reporter gene. Templates for mutant riboswitch constructs were constructed using synthetic DNA oligonucleotides and the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent). The resulting WT and mutant reporter constructs were transformed into *E. coli* BW25113 cells to conduct gene expression assays.

Riboswitch reporter assays

Riboswitch reporter assays were performed in a manner similar to that described previously (Atilho et al. 2019b). Briefly, bacterial cultures were grown overnight at 37°C in Lysogeny Broth (LB) with the appropriate antibiotic(s). For WT cells carrying the plasmid vector, carbenicillin (100 µg mL^{-1}) was used. For genetic knockout cells, carbenicillin and kanamycin (100 µg mL^{-1}) were used. To establish liquid-based β-galactosidase assays for visual interpretation, cells were diluted to an OD_{600} of 0.05 or 0.1 in LB supplemented with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, 100 µg mL^{-1}) and the appropriate antibiotic(s) before incubation at 37°C with shaking. Similarly, for reporter expression quantification by spectrophotometer, LB media containing antibiotic(s) were inoculated with cells to an OD_{600} of 0.05 or 0.1 and quantified at the desired time point (4–7 h) using MUG (4-methylumbelliferyl β-D-galactopyranoside) as previously described (Atilho et al. 2019b). Fluorescence units were calculated as described previously (Vidal-Aroca et al. 2006).

For agar diffusion assays, *E. coli* strains carrying riboswitch reporter constructs were grown overnight at 37°C in 3 mL of LB.

Cells were washed three times with between 1 and 2 mL of 1X PBS (Dulbeccos's phosphate buffered saline, without calcium or magnesium, was purchased from Thermo Fisher Scientific) and resuspended in 5 mL of M9 medium. Cells were then plated on M9 agar plates containing the appropriate antibiotics and X-gal (100 µg mL^{-1}). M9 minimal medium broth (1X M9 salts [42 mM Na_2HPO_4, 24 mM KH_2PO_4, 9 mM NaCl, 19 mM NH_4Cl], 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 2% glucose, 0.5 µg/mL thiamin) was purchased from Teknova. Filter disks were prepared from 0.35-mm-thick pure cellulose chromatography paper (Fisher Scientific) and placed on prepared agar plates before 5 µL of the compound was added to the disk. Plates were incubated for 48 h at 37°C before analysis.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Neil White, Ruben Atilho, Kevin Perkins, Narasimhan Sudarsan, and other members of the Breaker laboratory for helpful discussions. We especially thank Etienne Greenlee for his contributions to the discovery of the *ubiB* motif. S.N.M. was supported by the National Science Foundation Graduate Research Fellowship Program (DGE1122492). T.C.J.S. was funded by the Howard Hughes Medical Institute EXROP (Exceptional Research Opportunities Program) and was a participant in the Yale BioMed SURF (Summer Undergraduate Research Fellowship). This work was also supported by National Institutes of Health (NIH) grants to R.R.B. (GM022778). In addition, research in the Breaker laboratory is supported by the Howard Hughes Medical Institute.

Received June 28, 2019; accepted July 29, 2019.

REFERENCES

Ames TD, Rodionov DA, Weinberg Z, Breaker RR. 2010. A eubacterial riboswitch class that senses the coenzyme tetrahydrofolate. *Chem Biol* **17:** 681–685. doi:10.1016/j.chembiol.2010.05.020

Anraku Y. 1988. Bacterial electron transport chains. *Annu Rev Biochem* **57:** 101–132. doi:10.1146/annurev.bi.57.070188.000533

Atilho RM, Mirihana Arachicilage G, Greenlee EB, Knecht KM, Breaker RR. 2019a. A bacterial riboswitch class for the thiamin precursor HMP-PP employs a terminator-embedded aptamer. *eLife* **8:** e45210. doi:10.7554/eLife.45210

Atilho RM, Perkins KR, Breaker RR. 2019b. Rare variants of the FMN riboswitch class in *Clostridium difficile* and other bacteria exhibit altered ligand specificity. *RNA* **25:** 23–34. doi:10.1261/rna.067975.118

Bastet L, Turcotte P, Wade JT, Lafontaine DA. 2018. Maestro of regulation: riboswitches orchestrate gene expression at the levels of translation, transcription and mRNA decay. *RNA Biol* **15:** 679–682. doi:10.1080/15476286.2018.1451721

Benner SA, Ellington AD, Tauer A. 1989. Modern metabolism as a palimpsest of the RNA world. *Proc Natl Acad Sci* **86:** 7054–7058.

Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* **5:** 593–599. doi:10.1038/nchembio.186

Breaker RR. 2009. Riboswitches: from ancient gene-control systems to modern drug targets. *Future Microbiol* **4:** 771–773. doi:10.2217/fmb.09.46

Breaker RR. 2011. Prospects for riboswitch discovery and analysis. *Mol Cell* **43:** 867–879. doi:10.1016/j.molcel.2011.08.024

Breaker RR. 2012. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* **4:** a003566. doi:10.1101/cshperspect.a003566

Butler EB, Xiong Y, Wang J, Strobel SA. 2011. Structural basis of cooperative ligand binding by the glycine riboswitch. *Chem Biol* **18:** 293–298. doi:10.1016/j.chembiol.2011.01.013

Caron M-P, Bastet L, Lussier A, Simoneau-Roy M, Massé E, Lafontaine DA. 2012. Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc Natl Acad Sci* **109:** E3444–E3453. doi:10.1073/pnas.1214024109

Chen X, Mirihana Arachchilage G, Breaker RR. 2019. Biochemical validation of a second class of tetrahydrofolate riboswitches in bacteria. *RNA* **25:** 1091–1097. doi:10.1261/rna.071829.119

Frieda KL, Block SM. 2012. Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* **338:** 397–400. doi:10.1126/science.1225722

Gazzaniga F, Stebbins R, Chang SZ, McPeek MA, Brenner C. 2009. Microbial NAD metabolism: lessons from comparative genomics. *Microbiol Mol Biol Rev* **73:** 529–541. doi:10.1128/MMBR.00042-08

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. doi:10.1016/j.jmb.2006.04.003

Grose JH, Bergthorsson U, Roth JR. 2005. Regulation of NAD synthesis by the trifunctional NadR protein of *Salmonella enterica*. *J Bacteriol* **187:** 2774–2782. doi:10.1128/JB.187.8.2774-2782.2005

Haller A, Souliére MF, Micura R. 2011. The dynamic nature of RNA as a key to understanding riboswitch mechanisms. *Acc Chem Res* **44:** 1339–1348. doi:10.1021/ar200035g

Johnson JE Jr, Reyes FE, Polaski JT, Batey RT. 2012. B_{12} cofactors directly stabilize an mRNA regulatory switch. *Nature* **492:** 133–137. doi:10.1038/nature11607

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

Lee ER, Blount KF, Breaker RR. 2009. Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. *RNA Biol* **6:** 187–194. doi:10.4161/rna.6.2.7727

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

Lemay J-F, Desnoyers G, Blouin S, Heppell B, Bastet L, St-Pierre P, Massé E, Lafontaine DA. 2011. Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms. *PLoS Genet* **7:** e1001278. doi:10.1371/journal.pgen.1001278

Lots TS, Suess B. 2018. Small-molecule-binding riboswitches. *Microbiol Spectr* **6:** RWR-0025-2018.

Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL, Breaker RR. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* **306:** 275–279. doi:10.1126/science.1100829

McCown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR. 2017. Riboswitch diversity and distribution. *RNA* **23:** 995–1011.

Mirihana Arachchilage G, Sherlock ME, Weinberg 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

Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, Perumov DA, Nudler E. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111:** 747–756.

Nahvi A, Sudarsan N, Ebert MS, Zho X, Brown KL, Breaker RR. 2002. Genetic control by a metabolite binding mRNA. *Chem Biol* **9:** 1043–1049. doi:10.1016/S1074-5521(02)00224-7

Nahvi A, Barrick JE, Breaker RR. 2004. Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res* **32:** 143–150.

Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29:** 2933–2935. doi:10.1093/bioinformatics/btt509

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, 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. doi:10.1016/j.molcel.2016.11.019

O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei 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

Peselis A, Serganov A. 2014. Themes and variations in riboswitch structure and function. *Wiley Interdiscip Rev RNA* **5:** 803–822. doi:10.1002/wrna.1247

Poiata E, Meyer MM, Ames TD, Breaker RR. 2009. A variant riboswitch aptamer class for *S*-adenosylmethionine common in marine bacteria. *RNA* **15:** 2046–2056. doi:10.1261/rna.1824209

Poon WW, Davis DE, Ha HT, Jonassen T, Rather PN, Clarke CF. 2000. Identification of *Escherichia coli ubiB*, a gene required for the first monooxygenase step in ubiquinone biosynthesis. *J Bacteriol* **182:** 5139–5146. doi:10.1128/JB.182.18.5139-5146.2000

Regulski EE, Moy RH, Weinberg Z, Barrick JE, Yao Z, Ruzzo WL, Breaker RR. 2008a. 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

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

Ren A, Rajashankar KR, Patel DJ. 2012. Fluoride ion encapsulation by Mg^{2+} ions and phosphates in a fluoride riboswitch. *Nature* **486:** 85–89. doi:10.1038/nature11152

Ruff KM, Muhammad A, McCown PJ, Breaker RR, Strobel SA. 2016. Singlet glycine riboswitches bind ligand as well as tandem riboswitches. *RNA* **22:** 1728–1738. doi:10.1261/rna.057935.116

Serganov A, Huang L, Patel DJ. 2008. Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature* **455:** 1263–1267. doi:10.1038/nature07326

Sherlock ME, Malkowski SN, Breaker RR. 2017. Biochemical validation of a second guanidine riboswitch class in bacteria. *Biochemistry* **56:** 352–358. doi:10.1021/acs.biochem.6b01270

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. doi:10.7554/eLife.33908

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 internucleotide linkage geometry and the stability of RNA. *RNA* **5:** 1308–1325. doi:10.1017/S1355838299990891

Stancek M, Schnell R, Rydén-Aulin M. 2005. Analysis of *Escherichia coli* nicotinate mononucleotide adenylyltransferase mutants *in vivo* and *in vitro*. *BMC Microbiol* **6:** 16.

Stoddard CD, Batey RT. 2006. Mix-and-match riboswitches. *ACS Chem Biol* **1:** 751–754. doi:10.1021/cb600458w

Sudarsan N, Hammond MC, Block KF, Welz R, Barrick JE, Roth A, Breaker RR. 2006. Tandem riboswitch architectures exhibit complex gene control functions. *Science* **314:** 300–304. doi:10.1126/science.1130716

Vidal-Aroca F, Giannattasio M, Brunelli E, Vezzoli A, Plevani P, Muzi-Falconi M, Bertoni G. 2006. One-step high-throughput assay for quantitative detection of β-galactosidase activity in intact Gram-negative bacteria, yeast, and mammalian cells. *BioTechniques* **40:** 433–440. doi:10.2144/000112145

Wang JX, Breaker RR. 2008. Riboswitches that sense *S*-adenosylmethionine and *S*-adenosylhomocysteine. *Biochem Cell Biol* **86:** 157–168. doi:10.1139/O08-008

Watson PY, Fedor MJ. 2012. The *ydaO* motif is an ATP-sensing riboswitch in *Bacillus subtilis*. *Nat Chem Biol* **8:** 963–965.

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, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, Breaker RR. 2010. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* **11:** R31. doi:10.1186/gb-2010-11-3-r31

Weinberg Z, Lünse CE, Corbino KA, Ames TD, Nelson JW, Roth A, Perkins KR, Sherlock ME, Breaker RR. 2017. Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res* **45:** 10811–10823. doi:10.1093/nar/gkx699

Welz R, Breaker RR. 2007. Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA* **13:** 573–582. doi:10.1261/rna.407707

White HB. 1976. Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* **7:** 101–104.

Wickiser JK, Cheah MT, Breaker RR, Crothers DM. 2005a. The kinetics of ligand binding by an adenosine-sensing riboswitch. *Biochemistry* **44:** 13404–13414. doi:10.1021/bi051008u

Wickiser JK, Winkler WC, Breaker RR, Crothers DM. 2005b. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* **18:** 49–60. doi:10.1016/j.molcel.2005.02.032

Wilson KS, von Hippel PH. 1995. Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc Natl Acad Sci* **92:** 8793–8797. doi:10.1073/pnas.92.19.8793

Winkler W, Nahvi A, Breaker RR. 2002a. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419:** 952–956. doi:10.1038/nature01145

Winkler WC, Cohen-Chalamish S, Breaker RR. 2002b. An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci* **99:** 15908–15913. doi:10.1073/pnas.212628899

Woese C. 1967. *The genetic code: the molecular basis for genetic expression*. Harper and Row, New York.

Yarnell WS, Roberts JW. 1999. Mechanism of intrinsic transcription termination and antitermination. *Science* **284:** 611–615. doi:10.1126/science.284.5414.611

Zhou Y, Wang L, Yang F, Lin X, Zhang S, Zhao ZK. 2011. Determining the extremes of the cellular NAD(H) level by using an *Escherichia coli* NAD$^+$-auxotrophic mutant. *Appl Environ Microbiol* **77:** 6133–6140. doi:10.1128/AEM.00630-11