Cobalamin riboswitches exhibit a broad range of ability to discriminate between methylcobalamin and adenosylcobalamin

Received for publication, March 24, 2017, and in revised form, May 4, 2017 Published, Papers in Press, May 8, 2017, DOI 10.1074/jbc.M117.787176

Jacob T. Polaski1, Samantha M. Webster1, James E. Johnson, Jr.2, and Robert T. Batey3

From the Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309

Edited by Joel Gottesfeld

Riboswitches are a widely distributed class of regulatory RNAs in bacteria that modulate gene expression via small-molecule-induced conformational changes. Generally, these RNA elements are grouped into classes based upon conserved primary and secondary structure and their cognate effector molecule. Although this approach has been very successful in identifying new riboswitch families and defining their distributions, small sequence differences between structurally related RNAs can alter their ligand selectivity and regulatory behavior. Herein, we use a structure-based mutagenic approach to demonstrate that cobalamin riboswitches have a broad spectrum of preference for the two biological forms of cobalamin in vitro using isothermal titration calorimetry. This selectivity is primarily mediated by the interaction between a peripheral element of the RNA that forms a T-loop module and a subset of nucleotides in the cobalamin-binding pocket. Cell-based fluorescence reporter assays in Escherichia coli revealed that mutations that switch effector preference in vitro lead to differential regulatory responses in a biological context. These data demonstrate that a more comprehensive analysis of representative sequences of both previously and newly discovered classes of riboswitches might reveal subgroups of RNAs that respond to different effectors. Furthermore, this study demonstrates a second distinct means by which tertiary structural interactions in cobalamin riboswitches dictate ligand selectivity.

Regulation of mRNA expression by structural switches embedded in their 5'-leader is a widespread and diverse form of RNA-based regulation in bacteria (1). One type of molecular switch, the riboswitch, directly binds small molecule effectors to inform the expression machinery via the interplay of two functional domains: the ligand-binding receptor domain (aptamer domain) and a downstream regulatory switching domain (expression platform) (2). Currently, almost 40 distinct classes of riboswitches have been identified, each defined by covariation analysis-based secondary structural models and associated patterns of nucleotide conservation, along with the specific metabolite, signaling molecule, or ion to which the RNA responds (3, 4). Structural analyses of receptor domains of most riboswitch classes in complex with their cognate effector molecule have revealed that each is composed of a highly structured RNA element containing a ligand-binding pocket that typically discriminates between its cognate ligand and chemical analogues, often at levels rivaling proteins (5–7).

However, recently it has been observed that within some classes of riboswitches are rare sequence variants whose primary and secondary structure closely conforms to the consensus but have altered ligand selectivity. For example, a bioinformatics search for new guanine and adenine riboswitches uncovered a rare set of sequences found in a single organism, Mesoplasma florum, that bind 2'-deoxyguanosine (2'-dG)4 (8). Similarly, sequence differences within the binding pocket of a subset of GEMM-I (genes for the environment, membranes, and motility) riboswitches were shown to result in a switch from cyclic di-GMP to cyclic AMP-GMP binding (9). Other riboswitch classes contain rare variants whose cognate effector is not known but is not the assigned molecule (10). Thus, a deeper computational and experimental examination of diversity within riboswitch classes may reveal unexpected differences in metabolite responses among related RNAs.

One of the most broadly distributed classes of riboswitches across bacteria was originally characterized as specifically responding to 5'-deoxyadenosylcobalamin (AdoCbl) (11, 12). Cobalamin is a complex protein cofactor consisting of a corrin ring that chelates CoIII and a flexible linker with a terminal dimethylbenzimidazole moiety that coordinates with the α-axial position of the cobalt (Fig. 1A). The biologically active forms of cobalamin are 5'-deoxyadenosylcobalamin and methylcobalamin (MeCbl), which bear the active group in the β-axial position; aerobic photolysis of the CoIII–carbon bond yields an inactive form that can be found in the cell, hydroxocobalamin (HyCbl) (13). Sequence alignment of cobalamin-binding ribo-

4The abbreviations used are: 2'-dG, 2'-deoxyguanosine; AdoCbl, 5'-deoxyadenosylcobalamin; MeCbl, methylcobalamin; HyCbl, hydroxocobalamin; L, loop; IL, internal loop; IT, isothermal titration calorimetry; PE, peripheral element; SL, stem loop.
switches yielded a consensus secondary structure with several regions of highly conserved nucleotides, including a four-way junction between helices P3–P6, a T-loop-mediated tertiary interaction, and a variable peripheral element containing a highly conserved internal loop (IL11/10) (14).

Crystal structures of two cobalamin riboswitches in complex with AdoCbl revealed the core of the ligand-binding pocket is defined by strands J3/4 and J6/3 of the four-way junction that is organized by a universally conserved T-loop-mediated tertiary interaction between L4 and IL6/7 (15, 16). Interaction of the 5'-deoxyadenosyl moiety of AdoCbl with the junction is enabled by direct contacts between nucleotides in J6/3 and IL11/10 that pull J6/3 away from J3/4 to provide space for the sterically bulky 5'-deoxyadenosyl moiety (Fig. 1C). The majority of cobalamin riboswitches—greater than 90%, which we will refer to as the Cbl-I class (Rfam accession RF00174)—have the P8–P12 extension and associated IL11/10 element (17), which we refer to as the Cbl-IIa class (Rfam accession RF01689). Structural and biochemical analysis of a variant from the Cbl-IIa class revealed an inability to bind AdoCbl but rather bind MeCbl and HyCbl, both of which have a small β-axial moiety (15). In this RNA, an internal bulge between P1 and P3 (J1/3, which comprises peripheral element 2 (Fig. 1B, PE2)) interacts with J6/3 in a fashion that positions it adjacent to J3/4 in the core, allowing the riboswitch to only accommodate cobalamins with small β-axial moieties (Fig. 1D).

A second group of variants, which we refer to as Cbl-IIb (Rfam accession RF01482), represent a diverse set of RNAs associated with ethanolamine utilization genes in a number of bacteria, including Enterococcus faecalis and Listeria monocytogenes (18) that also lack the P8–P12 PE1 subdomain, but significantly differ from Cbl-IIa representatives in the nucleotide composition of J6/3 of the binding core. Additionally, in contrast to members of the Cbl-IIa class, Cbl-IIb riboswitches from E. faecalis and L. monocytogenes have been shown to interact highly selectively with AdoCbl (18) but do not have a known sequence element that interacts with J6/3. Therefore, the structural basis for their AdoCbl selectivity remains unknown. Interestingly, close examination of both Cbl-IIa and Cbl-IIb repre-
sentatives reveals significant variation in the sequence and secondary structure of region J1/3, suggesting that these groups of riboswitches may have differing abilities to selectively bind the biological forms of cobalamin. To better define selectivity across cobalamin riboswitches, the ability of a subset of RNAs to discriminate between AdoCbl and MeCbl was examined using isothermal titration calorimetry (ITC). This analysis revealed a broad spectrum of ability to discriminate between the two biologically active forms of cobalamin, including members that cannot discriminate between the two. A mutagenic examination of two Cbl-IIa RNAs, one that selectively binds MeCbl and one that strongly prefers AdoCbl, reveals that the interaction between J6/3 in the binding core and the peripheral junction J1/3 is one mechanism by which members of this class achieve selective cobalamin binding. These results are mirrored in a biological context, where genetic reporter assays conducted in *Escherichia coli* demonstrate that mutations made to these regions alter the form of cobalamin that elicits a regulatory response. 

### Results

**Cobalamin riboswitches share a conserved core but differ widely in the identity of key sequence elements**

Unlike Cbl-I riboswitches where nucleotides comprising the IL11/10-J6/3 interaction are highly conserved, there is considerable variation within Cbl-IIa/b riboswitches, both in the four-way junction and the associated PE2. The two most highly conserved sequence elements around the contact site for the $\beta$-axial moiety of cobalamin are J3/4, which has the consensus sequence GRAA in Cbl-IIa and GGAA in Cbl-Iib and on the other side of the four-way junction J6/3 with an RYG consensus sequence in Cbl-IIa and a UCU consensus (18) in Cbl-Iib. More strikingly, across Cbl-IIa/b RNAs PE2 differs substantially, with variants containing large internal bulges (*env4Cbl-IIa, env8Cbl-IIa, env4Cbl-IIa, env47Cbl-IIa, and env50Cbl-IIa*), three-way helical junctions (*Synechococcus elongatus hupE Cbl-IIa and Acaryochloris marina hupE Cbl-IIa*), and the near absence of nonhelical features (*E. faecalis eutG Cbl-Iib*). This structural diversity suggests the J1/3-J6/3 interaction observed in the crystal structure of the *env8Cbl-IIa* riboswitch may not be universal to this class, which might impact the ligand selectivity of individual variants for different chemical forms of cobalamin. Note that the “*env*” designations for environmental sequences are those given in their original discovery (17), and their Rfam accession numbers are given in supplemental Table S1.

### Cbl-IIa cobalamin riboswitches exhibit a range of selectivity for AdoCbl versus MeCbl

To ascertain ligand selectivity among Cbl-IIa/b riboswitches, a set of RNAs was chosen that span sequence and secondary structure diversity found across this class but skewed toward members of Cbl-IIa, where the greatest differences in PE2 is observed. As benchmarks, the *env4Cbl-IIa* and *env8Cbl-IIa* riboswitches, which were previously shown to be highly selective for MeCbl, and the *E. faecalis Cbl-Iib* (*EfaCbl-Iib*) riboswitch that controls the *eutG* operon that strongly prefers AdoCbl (18, 19) were used in this analysis (Table 1). Two notable differences between the sequences of the *env4/env8Cbl-IIa* and *EfaCbl-Iib* riboswitches may account for their difference in ligand selectivity: 1) the *EfaCbl-Iib* riboswitch does not possess a significant J1/3 bulge, suggesting the absence of a J1/3-J6/3 interaction, and 2) the sequence composition of its J6/3 region differs significantly by containing only pyrimidine nucleotides (supplemental Fig. S1).

An RNA that may provide a clue as to which of these differences is predictive of ligand selectivity is a riboswitch that has features of both *env4/env8Cbl-IIa* and *EfaCbl-Iib*. Like *EfaCbl-Iib*, the *T. elongatus cbiX Cbl-IIa* (*TelCbl-IIa*) riboswitch is pre-

### Table 1

| Riboswitch     | $K_{D,\text{Ado}}$ $^a$ (nM) | $K_{D,\text{Me}}$ $^a$ (nM) | $K_{\text{rel}1}$ $^b$ (nM) | $K_{D,\text{Me}}$ $^a$ (nM) | $K_{\text{rel}2}$ $^b$ (nM) |
|----------------|-------------------------------|-------------------------------|-----------------------------|-------------------------------|-----------------------------|
| *env8Cbl-IIa*  | >1e7$^c$                      | 6 ± 3                         | >2e5$^c$                    | 480 ± 40                      | >2000$^c$                   |
| *env4Cbl-IIa*  | >1e7$^c$                      | 44 ± 23                       | >2e5$^c$                    | 3000 ± 1000                   | >300$^c$                    |
| *T. elongatus cbiX Cbl-IIa* | 2700 ± 400 | 17 ± 3                         | 200                          | 300 ± 70                      | 7                           |
| *A. marina hupE Cbl-IIa* | 6900 ± 100 | 510 ± 300                      | 10                           |                               |                             |
| *env62Cbl-IIa* | 110 ± 20                      | 56 ± 20                       | 2                            |                               |                             |
| *S. elongatus hupE Cbl-IIa* | 640 ± 50   | 3800 ± 3000                    | 0.2                          |                               |                             |
| *env127Cbl-IIa* | 200 ± 30                      | 1000 ± 300                    | 0.2                          |                               |                             |
| *env47Cbl-IIa* | 440 ± 30                       | 4200 ± 1000                    | 0.1                          |                               |                             |
| *env77Cbl-IIa* | 890 ± 200                     | 7700 ± 2000                    | 0.1                          |                               |                             |
| *env50Cbl-IIa* | 6.2 ± 5.0                     | 140 ± 10                      | 0.04                         |                               |                             |
| *E. faecalis eutG Cbl-Iib* | 100 ± 40     | 36000 ± 20000                   | 0.003                         |                               |                             |

$^a$ Binding affinities are shown as the average ± S.D. from three independent titrations. For *env8Cbl-IIa, env4Cbl-IIa*, and *T. elongatus cbiX Cbl-IIa* riboswitches, the values for $K_{D,\text{Ado}}$ and $K_{D,\text{Me}}$ were calculated using a multiple ligand-binding model.  

$^b$ $K_{\text{rel}1} = (K_{D,\text{Ado}})/(K_{D,\text{Me}})$.  

$^c$ The values are based upon an estimate of the minimum value for $K_D$ based upon the ITC measurements.
Cofactor recognition by cobalamin riboswitches

Figure 2. Mutations to the binding core and peripheral elements of env50Cbl-IIa switch its selectivity from AdoCbl to MeCbl. Individual nucleotides and base pairs of the wild-type env50Cbl-IIa riboswitch and various mutant constructs are depicted as open circles. The locations of mutations designed to convert env50Cbl-IIa into env8Cbl-IIa are shown as filled colored circles, with the $K_{\text{rel}}$ for each construct shown above the corresponding structures. Constructs are grouped according to mutated regions of the RNA, where the colored circles next to each $K_{\text{rel}}$ correspond to the mutations mapped onto the secondary structures. For clarity, the outlined yellow circle at the 3'-end of each RNA denotes a constant sequence that does not interfere with folding.

Predicted to have a P1/P3 region that is largely helical with few bulged nucleotides, such that there is a near absence of a J1/3 bulge, but a J6/3 sequence that is identical to env4/env8Cbl-IIa riboswitches. Unlike EfaCbl-IIb, the TelCbl-IIa riboswitch shows a 200-fold preference for MeCbl over AdoCbl, suggesting that the sequence composition of J6/3 plays a significant role in dictating ligand selectivity (Table 1).

Surprisingly, many of the other RNAs examined that also have very similar J6/3 sequences do not display high selectivity for either AdoCbl or MeCbl. Instead, most sequences have a modest (2–10-fold) ability to discriminate between the two forms of cobalamin. For example, the env47Cbl-IIa and env77Cbl-IIa riboswitches are predicted to have secondary structures nearly identical to those of env4/env8Cbl-IIa, except for an extended stem loop 13 (SL13) and a J3/4 strand with one additional nucleotide, yet bind both MeCbl and AdoCbl with an ~10-fold preference for AdoCbl (Table 1). Similarly, the env127Cbl-IIa riboswitch has a sequence and secondary structure nearly identical to those of env77Cbl-IIa and binds AdoCbl with 5-fold higher affinity than MeCbl (Table 1). The env62Cbl-IIa riboswitch, which has a nearly identical sequence composition in its four-way junction compared with env4/env8Cbl-IIa but no J1/3 bulge, weakly discriminates between MeCbl and AdoCbl with a 2-fold preference for MeCbl. Two RNAs with J3/4 elements that differ significantly compared with other tested sequences, A. marina hupE Cbl-IIa and S. elongatus hupE Cbl-IIa, are 10-fold selective for MeCbl and 6-fold selective for AdoCbl, respectively (Table 1). Together, these results suggest that J6/3 alone does not dictate selectivity; sequences outside the four-way junction influence ligand-binding preferences. Furthermore, these data suggest that a substantial fraction of Cbl-IIa riboswitches may productively bind either form of cobalamin with nanomolar affinity.

Unexpectedly, one Cbl-IIa variant was found to be moderately selective for AdoCbl. The env50Cbl-IIa riboswitch shares a highly similar core structure with env4/env8Cbl-IIa, env47Cbl-IIa, and env62Cbl-IIa but binds AdoCbl with over 20-fold higher affinity than MeCbl (Table 1). Structural differences between these sequences are subtle, with env50Cbl-IIa having a J1/3 region that is one nucleotide smaller than env4/

env8Cbl-IIa and env47Cbl-IIa, an internal loop in P6 (IL6/7), and an extended SL13, any of which might contribute to the observed differences in ligand preference. Unfortunately, for Cbl-IIa representatives, inspection of the sequences in J1/3 and J6/3 do not yield any clear insights into predicting the selectivity of different riboswitch variants.

Two purines in J1/3 at site of direct contact with the β-axial moiety of cobalamin influence selectivity

To gain insights into the determinants of selectivity, a series of changes were made to the AdoCbl-selective env50Cbl-IIa riboswitch that replace key sequence identities with those of the MeCbl-selective env8Cbl-IIa RNA. To accomplish a selectivity switch, sequences in nonhelical regions of the four-way region closest to the β-axial moiety of cobalamin in the ligand-binding pocket were mutated first, followed by more distal regions (Fig. 2 and supplemental Fig. S2). Specifically, nucleotides in J3/4 and J6/3 were converted to their identity in env8Cbl-IIa, then J1/3, and finally the interaction between L4 and the IL6/7 element. This strategy was successfully used previously to determine the nucleotides most involved in a selectivity switch between guanine-selective and 2’-deoxyguanosine-binding RNAs within the purine riboswitch family (20).

Mutations in the four-way junction of env50Cbl-IIa that convert this region into that of env8Cbl-IIa yield an RNA with little ligand selectivity. Within the four-way junction there are only two nucleotides that differ between the env50Cbl-IIa and env8Cbl-IIa variants: G20 and G77, which are both adenosines in env8Cbl-IIa. The G20A or G77A mutations decreased AdoCbl affinity by ~4-fold, while only modestly increasing affinity for MeCbl, such that these RNAs have a small preference for AdoCbl (2–4-fold) (Fig. 2, variants (a) and (b), and supplemental Table S2). In combination, these two changes convert env50Cbl-IIa into an RNA that has moderately high affinity for both AdoCbl and MeCbl but is unable to discriminate between them (variant (c); Fig. 2 and supplemental Table S2). This further supports the above findings that nucleotide identities in the four-way junction do not fully confer selectivity but rather act in concert with other elements in the RNA.
To determine the influence of J1/3 alone on ligand selectivity, this element of env50Cbl-IIa was replaced with that of env8Cbl-IIa along with the two flanking Watson-Crick base pairs in P1 and P3 (variants (d) and (e), respectively). Although these mutations had little impact on the affinity of these RNAs for AdoCbl, these variants have significantly increased affinity for MeCbl (>10-fold), resulting in riboswitches that have either a slight preference for AdoCbl or cannot discriminate between the two ligands (Fig. 2, variants (d) and (e), and supplemental Table S2). Together, these data suggest that J1/3 has a moderate influence on ligand selectivity but, like J6/3, is not solely responsible for highly selective binding of either AdoCbl or MeCbl.

An interaction between J6/3 of the binding core and peripheral element J1/3 is the main determinant of ligand selectivity

The above data clearly establish that sequence elements within the four-way junction and J1/3 alone do not establish the selectivity of Cbl-IIa riboswitches. Combining the mutations in J3/4 and J6/3 from variant (c) with those in J1/3 from either variant (d) or (e) to create variants (f) and (g), respectively, causes selectivity to be substantially shifted toward MeCbl. In the case of variant (g), the riboswitch has a 6-fold preference for MeCbl, whereas variant (f) is 20-fold selective, demonstrating that, together, regions J1/3 and J6/3 of env8Cbl-IIa are key for conferring MeCbl selectivity (Fig. 2, variants (f) and (g), and supplemental Table S2). It is important to note that the combination of mutations in the four-way junction and PE2 represent a gain of function with relation to MeCbl binding.

In the crystal structure of the env8Cbl-IIa riboswitch in complex with HyCbl, J6/3 directly interacts with J1/3 through a set of base-base and base-backbone interactions (Fig. 3A and the N2 exocyclic amine of G10 (Fig. 3A). The module is further reinforced by an adenosine at position 12 that intercalates between the C11-G70 and G10-U69 base pairs.

Additional changes to the aptamer domain made by transplanting the peripheral L4-L6 interaction of env8Cbl-IIa to create variants (h) and (i) further increase selectivity for MeCbl to 10- and 70-fold, respectively (Fig. 2, variants (h) and (i), and supplemental Table S2). Notably, variants (h) and (i) have roughly equal affinities for MeCbl (7.8 and 5.1 nM, respectively), but variant (h) displays stronger selectivity for MeCbl than variant (i) because it has weaker affinity for AdoCbl. The difference between these two constructs is that variant (i) carries a noncanonical A-C base pair in P3 and an A-U base pair in P1 that flanks J1/3, rather than two G-C base pairs as observed in env50Cbl-IIa. Together, these data reveal that selectivity for MeCbl by the env8Cbl-IIa riboswitch is a combination of nucleotide composition in the four-way junction and their interaction with J1/3, whereas more peripheral sequence features can further bias selectivity.

A complementary set of mutations made to a MeCbl selective RNA switches its ligand preference to AdoCbl

If the interaction between the core four-way junction and J1/3 primarily dictates selectivity among Cbl-IIa riboswitches, then env8Cbl-IIa should be able to be converted into an AdoCbl-selective RNA via alterations to these two elements (Fig. 4 and supplemental Fig. S3). Similar to the strategy above, mutations were made to the binding core of env8Cbl-IIa either alone or in combination with mutations to peripheral element J1/3. The first set of mutations were made to J3/4 and J6/3, where A20 and A68 were substituted with guanosines from corresponding positions in env50Cbl-IIa, in addition to substituting the A-U pair in P1 and the noncanonical A-C pair in P3 with two G-C base pairs to create variant (b’’) (supplemental Fig. S3). This variant is weakly selective for AdoCbl (~2-fold) and shows a significant decrease in affinity for MeCbl (250- and 3-fold relative to K_{D1} and K_{D2}, respectively, with a corresponding increase in affinity for AdoCbl (~1000-fold; variant (b’’)) (Fig. 4 and supplemental Table S3). Further mutation of variant (b’’) by transplanting peripheral element J1/3 of env50Cbl-IIa.
into the RNA to create variant (d') significantly enhances its selectivity for AdoCbl (75-fold) and provides additional support for a model in which the interaction between J1/3 and J6/3 principally modulates ligand preference (Fig. 4, variant (d'), and supplemental Table S3).

**Mutations that alter ligand selectivity in vitro are recapitulated in a cellular context**

To determine whether regions of the RNA shown to be important for ligand selectivity by calorimetry also have the same role in a biological context, mutant env8Cbl-IIa riboswitches (variants (b') and (d')) were tested for their ability to regulate gene expression in the presence of either AdoCbl or MeCbl. The wild-type env8Cbl-IIa riboswitch, an RNA that carries mutations in L5 previously shown to ablate regulatory activity (23), and variants (b') and (d') were placed under control of a constitutive promoter and upstream of a reporter gene encoding the fluorescent protein mNeon (Fig. 5A and supplemental Table S4) (24). This reporter scheme was previously used to determine the activity of wild-type env8Cbl-IIa along with a suite of mutants designed to elucidate the regulatory mechanism of Cbl-IIa/b riboswitches that regulate translation initiation (15, 23). These reporters were transformed into either *E. coli* strain BW25113, in which conversion of cobalamins to AdoCbl is coupled to import, and thus this compound dominates the intracellular cobalamin pool or the same *E. coli* strain with a deletion of the cobalamin adenosyltransferase gene (*E. coli* knock-out variant) is unable to form the regulatory loop-loop interaction between L5 and SL13. Mutations A20G and A68G (shown in red) in the binding core along with conversion of the A-U pair in P1 and the A-C pair in P3 to G-C pairs (boxed in red) were made to create variant (b'). Variant (d') carries the same changes as variant (b') along with mutations G12A, G10U, and A8G (shown in blue). *B*. values of fold repression from cell-based genetic reporter assays using the riboswitches depicted in *A*. Each circle represents the average fold repression of three technical replicates from a single biological replicate. Red circles represent average fold repression values from experiments performed using the *E. coli* BW21153: ΔbttR cell strain with MeCbl, and the blue circles represent average fold repression values from experiments performed using the parental *E. coli* BW21153 cell strain with AdoCbl. Horizontal black lines represent the average fold repression of triplicate measurements from three biological replicates, and fold repression was calculated by dividing average fluorescence values of the no ligand condition by average fluorescence values from *E. coli* grown in the presence of cobalamin.

Cofactor recognition by cobalamin riboswitches

![Figure 4](https://example.com/figure4.png)

**Figure 4. Mutations to the binding core and peripheral elements of env8Cbl-IIa switch its selectivity from MeCbl to AdoCbl.** Individual nucleotides and base pairs of the wild-type env8Cbl-IIa riboswitch and various mutant constructs are depicted as open circles. The locations of mutations designed to convert env8Cbl-IIa into env50Cbl-IIa are shown as filled colored circles, with the *K* <sub>rel</sub> for each construct shown above the corresponding structure. Constructs are grouped according to mutated regions of the RNA, where the colored circles next to each *K* <sub>rel</sub> correspond to the mutations mapped onto the secondary structures.

![Figure 5](https://example.com/figure5.png)

**Figure 5. Genetic reporter assays in *E. coli* support that interactions between J1/3 and J6/3 mediate ligand selectivity in a cellular context.** *A*. Sequence and secondary structure of the wild-type env8Cbl-IIa riboswitch is shown in black. The riboswitch carrying mutations in L5 (boxed in green, knock-out variant) is unable to form the regulatory loop-loop interaction between L5 and SL13. Mutations A20G and A68G (shown in red) in the binding core along with conversion of the A-U pair in P1 and the A-C pair in P3 to G-C pairs (boxed in red) were made to create variant (b'). Variant (d') carries the same changes as variant (b') along with mutations G12A, G10U, and A8G (shown in blue). *B*. values of fold repression from cell-based genetic reporter assays using the riboswitches depicted in *A*. Each circle represents the average fold repression of three technical replicates from a single biological replicate. Red circles represent average fold repression values from experiments performed using the *E. coli* BW21153: ΔbttR cell strain with MeCbl, and the blue circles represent average fold repression values from experiments performed using the parental *E. coli* BW21153 cell strain with AdoCbl. Horizontal black lines represent the average fold repression of triplicate measurements from three biological replicates, and fold repression was calculated by dividing average fluorescence values of the no ligand condition by average fluorescence values from *E. coli* grown in the presence of cobalamin.

The resulting intracellular concentra-
Cofactor recognition by cobalamin riboswitches

AdoCbl with >10-fold higher affinity than variant (b'), shows enhanced regulatory activity (~6-fold repression) when expressed in cells grown in the presence of AdoCbl (Fig. 5B). Like variant (b'), variant (d') does not regulate gene expression in the presence of MeCbl, consistent with an ~4-fold decrease in affinity for MeCbl relative to variant (b'). Together, these experiments fully demonstrate that nucleotide identities in J6/3 of the four-way junction and J1/3 dictate cobalamin selectivity within at least a subset of Cbl-IIa riboswitches.

Discussion

Generally it is assumed that all members of a bioinformatically defined class of riboswitches recognize the same effector ligand—a premise that is only experimentally validated with a few members and is becoming challenged (10). In this work, it is shown that the Cbl-IIa class of cobalamin riboswitches, as defined by the lack of the P8–P12 peripheral extension and the RGY consensus sequence in J6/3, do not all recognize the same chemical form of cobalamin or discriminate between them with similar selectivity. Instead, a broad spectrum of selectivity is observed across this class with representatives displaying high selectivity for either AdoCbl or MeCbl, as well as members showing little ability to discriminate between them. A detailed mutagenic survey of a highly AdoCbl selective RNA reveals that the regions of the RNA most directly involved in determining selectivity are J6/3 in the ligand binding site and peripheral element J1/3, which are observed to directly interact in the env8Cbl-IIa crystal structure.

Although most riboswitches recognize their cognate ligand with very high selectivity over near-cognate compounds found in the cell (28, 29), some riboswitches are faced with the problem of monitoring a pool of several chemically related metabolites as found with cobalamin. A simple solution for recognizing chemical diversity is to evolve a receptor that ignores variable ligands as found with cobalamin. A simple solution for recognizing chemical diversity is to evolve a receptor that ignores variable ligands as found with cobalamin. A simple solution for recognizing chemical diversity is to evolve a receptor that ignores variable ligands as found with cobalamin. A simple solution for recognizing chemical diversity is to evolve a receptor that ignores variable ligands as found with cobalamin. A simple solution for recognizing chemical diversity is to evolve a receptor that ignores variable ligands as found with cobalamin.

Conversely, the cobalamin riboswitch directly confronts the reduced state of the pterin moiety of folates (30, 31). The riboswitch does not bind the fully oxidized pterin moiety of folic acid, whereas reduced forms of dihydrofolate and tetrahydrofolate are bound with high affinity. However, in bacteria the reduced folate pool carries a diverse set of one-carbon groups on its N5/N10 atoms of the pterin ring and can be polyglutamylated such that the overall pool is chemically heterogeneous (32). To cope with this, the RNA only recognizes the faces of the pterin moiety that are not involved in carrying one carbon compounds along with completely ignoring the glutamyl moiety.

Figure 6. Mutations that promote unfavorable sterics in the binding core of env8Cbl-IIa allow the RNA to accommodate AdoCbl. A, packing of A20 (red spheres) in J3/4 and A68 (blue spheres) in J6/3 against HyCbl (yellow sticks) in the wild-type env8Cbl-IIa riboswitch. B, A20G and A68G mutations modeled into the env8Cbl-IIa structure (Fig. 4; variant (b')) illustrate potential steric interference between the exocyclic amine of G68 and cobalamin. Measurements between atoms were calculated in PyMOL and are represented in Å.

For cobalamin riboswitches, ligand selectivity is driven by a complex interplay between nucleotides within the binding core of the RNA and interactions between the core and peripheral regions. Modulation of ligand selectivity resulting from cross-talk between peripheral domains and the binding pocket has been previously observed in the M. florum 2'-dG binding riboswitch. Here, it was shown that a single-point mutation of a nucleotide that directly contacts the ligand in the B. subtilis xpt/pbuX guanine-binding riboswitch causes a moderate switch in the selectivity of this RNA from guanine to 2'-dG, and further mutation of the binding pocket improved selectivity at the cost of affinity for both guanine and 2'-dG (20). However, these alterations in combination with mutations to peripheral elements were needed to fully recapitulate the binding properties of the wild-type M. florum riboswitch. Regions of the RNA that are peripheral to the binding core not only play a role in dictating ligand selectivity of riboswitches. Bioinformatic analysis of several hundred purine riboswitches identified a region of the RNA that does not directly participate in ligand recognition but was shown to be important for ligand binding kinetics, as well as the regulatory response (34). Thus, like their protein counterparts, resides in the second shell of the binding pocket.
and beyond have a profound influence on the ligand binding characteristics of riboswitches. In this regard, a recent bioinformatic analysis of variant riboswitches that restricts its analysis to the first shell nucleotides (10) may not fully reveal potential diversity in the various riboswitch classes.

The results of this study raise several biological and evolutionary questions about cobalamin riboswitches. First, why do members of the Cbl-I class exhibit highly selective binding for AdoCbl (12), whereas Cbl-IIa representatives feature a broad spectrum of affinities for different forms of the cofactor? One potential explanation may be that this relaxed binding selectivity in Cbl-IIa riboswitches reflects the conditions of the natural environments that organisms harboring these RNAs inhabit. For example, cyanobacteria in aquatic habitats that possess Cbl-IIa riboswitches experience changes in light exposure throughout the day, meaning the predominant intracellular cobalamin species could fluctuate between AdoCbl and its photolysis product HyCbl. In this scenario, having a riboswitch that responds to both forms of cobalamin would allow for gene expression to be dynamically regulated, depending on the degree of light exposure. A similar strategy is employed in Myxococcus xanthus, which uses AdoCbl as a proxy for light exposure to regulate the expression of carotenoid biosynthetic genes (35). Here, when light exposure is low, an interaction between the transcriptional repressor CarH and AdoCbl promotes CarH tetramer formation to enhance its affinity for an operator sequence to decrease carotenoid production. Alternatively, exposure to light decreases the concentration of intracellular AdoCbl, which increases carotenoid levels by inhibiting CarH tetramer formation to weaken its interaction with the operator. Alternatively, this flexibility in cobalamin binding may ensure proper regulation regardless of the status of the cobalamin pool that changes throughout the daily cycle.

However, if organisms like cyanobacteria can use a single riboswitch to respond to multiple forms of cobalamin, why are some Cbl-IIa variants highly selective for HyCbl/MeCbl and unable to recognize AdoCbl? Although these representatives are rare, their existence suggests that some bacteria may use multiple cobalamin riboswitches to regulate gene expression. In support of this, the bacterium Desulfitobacterium hafniense has been shown to harbor 18 cobalamin riboswitches in its genome from both Cbl-I and Cbl-IIa/b classes (36). Although in vitro binding experiments using three of these sequences showed a broad range of affinity for AdoCbl (0.027–90 μM), quantitative RT-PCR demonstrated that all three riboswitches repress gene expression when cells were grown in the presence of AdoCbl. Despite this discrepancy, which may have arisen from high intracellular AdoCbl concentrations based on experimental conditions, it suggests that some organisms may employ multiple riboswitches from the same family with varying selectivity profiles to more finely tune gene expression.

**Experimental procedures**

**RNA synthesis and preparation**

For all in vitro binding experiments, DNA templates for transcription reactions were amplified by PCR and transcribed by T7 RNA polymerase (37). Transcription reactions were purified using the appropriate percentage polyacrylamide gel (8M urea, 29:1 acrylamide: bisacrylamide) based on RNA length. Transcripts were visualized by UV shadowing, excised from the gel and extracted by soaking at 4 °C in 0.5× TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). Buffer exchange and concentration were performed using centrifugal concentrators (Amicon) with the appropriate molecular weight cutoff. Final RNA concentration was calculated as the summation of the individual bases and absorbance of 260-nm light.

**Isothermal titration calorimetry**

RNAs were dialyzed overnight at 4 °C into 1× ITC buffer (5 mM Na-MES, pH 6.0, 100 mM KCl, 5 mM MgCl2). RNA and either AdoCbl or MeCbl were diluted to the desired final concentration in 1× ITC buffer, and titrations were performed at 25 °C using a MicroCal ITC200 microcalorimeter. For titrations done with AdoCbl, ligand stocks were prepared under red light in a dark room and the instrument was covered with aluminum foil to limit exposure to light. Data analysis and curve fitting were performed using the Origin software suite as previously described (38).

**Cell-based reporter assays**

Plasmids containing cobalamin riboswitches controlling the expression of a fluorescent reporter gene (mNeon) were transformed into either *E. coli* strain BW25113: ΔbttuR or parental BW25113 and plated onto Petri dishes containing 2× YT medium containing 1.2% agar and supplemented with 100 μg/ml carbenicillin. Three individual colonies were picked for each riboswitch and grown to saturation at 37 °C overnight in 5 ml of a rich, chemically defined growth medium (CSB medium) supplemented with 100 μg/ml ampicillin (39). For gene expression assays, 5 μl of each saturated overnight culture was added to three separate tubes containing 5 ml of fresh CSB medium with or without MeCbl or AdoCbl and supplemented with 100 μg/ml ampicillin. For assays performed using AdoCbl, ligand stocks were prepared and added to bacterial cultures under red light, and culture tubes were wrapped with foil to limit light exposure. Cultures were grown to mid-log phase at 37 °C, at which time 200 μl of culture from each tube was added to wells in a Costar® 96-well half area microplate. Expression of mNeon was measured at an excitation wavelength of 490 nm and a 517-nm emission wavelength using a Tecan Infinite 200® PRO plate reader. Each data point shown represents cobalamin-dependent repression of mNeon expression from three technical replicates of a single biological replicate. Fold repression was determined using fluorescence values that were calculated by normalizing the relative expression levels of mNeon in each well to the cell density (A600), which were then background corrected using cell density normalized fluorescence from wells containing a pBR327 empty vector. Fold repression was calculated by dividing the average normalized background corrected fluorescence for the unpressed condition (absence of cobalamin) by the average normalized background corrected fluorescence for the repressed condition (presence of cobalamin).
Cofactor recognition by cobalamin riboswitches

Author contributions—J. T. P., J. E. J., and R. T. B. designed the study, and J. T. P. and R. T. B. wrote the manuscript. J. E. J. collected the initial calorimetry data relating to Table 1, and J. T. P. and S. M. W. collected the final set of data. S. M. W. and J. T. P. collected and analyzed the data related to Figs. 2 and 4, and J. T. P. collected data related to Fig. 5. All authors contributed to analyzing the results and approved the final manuscript.

Acknowledgments—We acknowledge the input of Dr. Francis E. Reyes during the early stages of this project and Otto Kletzien for assistance with some of the cell-based assays.

References
1. Roth, A., and Breaker, R. R. (2009) The structural and functional diversity of metabolite-binding riboswitches. *Annu. Rev. Biochem.* 78, 305–334
2. Garst, A. D., Edwards, A. L., and Batey, R. T. (2011) Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.* 3, a003533
3. Sherwood, A. V., and Henkin, T. M. (2016) Riboswitch-mediated gene regulation: novel RNA architectures dictate gene expression responses. *Annu. Rev. Microbiol.* 70, 361–374
4. Seregnanov, A., and Patel, D. J. (2012) Metabolite recognition principles and molecular mechanisms underlying riboswitch function. *Annu. Rev. Bioophys.* 41, 343–370
5. Batey, R. T. (2011) Structure and mechanism of purine-binding riboswitches. *Q. Rev. Bioophys.* 45, 345–381
6. Seregnanov, A. (2009) The long and the short of riboswitches. *Curr. Opin. Struct. Biol.* 19, 251–259
7. Kim, J. N., Roth, A., and Breaker, R. R. (2007) Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2′-deoxyguanosine. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16092–16097
8. Nahni, A. M., Gaffney, B. L., Jones, R. A., and Strobel, S. A. (2011) Differential analogue binding by two classes of c-di-GMP riboswitches. *J. Am. Chem. Soc.* 133, 15578–15592
9. Weinberg, Z., Nelson, J. W., Lüschn, C. E., Sherlock, M. E., and Breaker, R. R. (2017) Bioinformatic analysis of riboswitch structures uncovers variant classes with altered ligand specificity. *Proc. Natl. Acad. Sci. U.S.A.* 114, E2077–E2085
10. Barrick, J., and Breaker, R. R. (2007) The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* 8, R239
11. Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002) Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9, 1043
12. Kozlowski, P. M., Garabato, B. D., Łodowski, P., and Jaworska, M. (2016) Photolytic properties of cobalamin: a theoretical perspective. *Dalton Trans.* 45, 4457–4470
13. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A., and Gelfand, M. S. (2003) Regulation of the vitamin B12 metabolism and transport in bacteria: evaluation of potential for antibacterial drug discovery. *Bioessays* 25, 637–648
14. Gilbert, S. D., Reyes, F. E., Edwards, A. L., and Batey, R. T. (2009) Adaptive ligand binding by the purine riboswitch reveals a role in regulatory switching. *Science* 325, 463–472
15. Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S., and Elias-Arranz, M. (2011) Light-dependent gene regulation by a coenzyme B12-dependent photoreceptor. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7565–7570
16. Choudhary, P. K., Duret, A., Rohrbach-Brandt, E., Hölger, C., Sigel, R. K., and Maillard, J. (2013) Diversity of cobalamin riboswitches in the cori–noid-producing organohalide respirer *Desulfobacterium hafniense* *H11032*. *J. Bacteriol.* 195, 5186–5195
17. Edwards, A. L., Garst, A. D., and Batey, R. T. (2009) Monitoring RNA-ligand interactions using isothermal titration calorimetry. *Methods Mol. Biol.* 540, 97–114
18. Ceres, P., Garst, A. D., Marciano-Velázquez, J. G., Knight, R., and Batey, R. T. (2013) Nucleotides adjacent to the ligand-binding pocket are linked to activity tuning in the purine riboswitch. *J. Mol. Biol.* 425, 1596–1611
19. Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S., and Elias-Arranz, M. (2011) Light-dependent gene regulation by a coenzyme B12-dependent photoreceptor. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7565–7570
20. Choudhary, P. K., Duret, A., Rohrbach-Brandt, E., Hölger, C., Sigel, R. K., and Maillard, J. (2013) Diversity of cobalamin riboswitches in the cori–noid-producing organohalide respirer *Desulfobacterium hafniense* *H11032*. *J. Bacteriol.* 195, 5186–5195
21. Edwards, A. L., Garst, A. D., and Batey, R. T. (2009) Determining structures of RNA aptamers and riboswitches by X-ray crystallography. *Methods Mol. Biol.* 535, 135–163
22. Gilbert, S. D., and Batey, R. T. (2009) Monitoring RNA-ligand interactions using isothermal titration calorimetry. *Methods Mol. Biol.* 540, 97–114
23. Ceres, P., Garst, A. D., Marciano-Velázquez, J. G., and Batey, R. T. (2013) Modularity of select riboswitch expression platforms enables facile engineering of novel genetic regulatory devices. *ACS Synth. Biol.* 2, 463–472