B₁₂ cofactors directly stabilize an mRNA regulatory switch

James E. Johnson Jr*, Francis E. Reyes†*, Jacob T. Polaski† & Robert T. Batey†

Structures of riboswitch receptor domains bound to their effector have shown how messenger RNAs recognize diverse small molecules, but mechanistic details linking the structures to the regulation of gene expression remain elusive1,2. To address this, here we solve crystal structures of two different classes of cobalamin (vitamin B₁₂)-binding riboswitches that include the structural switch of the downstream regulatory domain. These classes share a common cobalamin-binding core, but use distinct peripheral extensions to recognize different B₁₂ derivatives. In each case, recognition is accomplished through shape complementarity between the RNA and cobalamin, with relatively few hydrogen bonding interactions that typically govern RNA–small molecule recognition. We show that a composite cobalamin–RNA scaffold stabilizes an unusual long-range intramolecular kissing-loop interaction that controls mRNA expression. This is the first, to our knowledge, riboswitch crystal structure detailing how the receptor and regulatory domains communicate in a ligand-dependent fashion to regulate mRNA expression.

An mRNA leader identified as controlling expression of a cobalamin transport gene (btuB) in Escherichia coli14 was the first validated riboswitch shown to interact directly with cellular metabolites in the absence of proteins5,6. Cobalamin riboswitches have since been found to regulate B₁₂ biosynthesis in bacteria widely7,8, and they are one of the most broadly distributed riboswitches in biology9,10. The cobalamin riboswitch family comprises two classes5,8 distinguished by peripheral extensions surrounding a common core. The secondary structure of both classes contains a central four-way junction (P3–P6 helices) forming the core receptor domain responsible for cobalamin binding (Fig. 1b and Supplementary Fig. 1). The other shared element is a kissing-loop (KL) interaction between loop L5 of the receptor and L13 of the regulatory domain that instructs the expression machinery. For cobalamin riboswitches that regulate translation, L13 typically contains the ribosome-binding site (RBS). Within the first cobalamin riboswitch class, the KL is generally linked to a downstream secondary structural switch11 akin to that used by most riboswitches9. The classes are primarily differentiated by the presence of a large peripheral

---

Figure 1 | Structures of cobalamins and cobalamin riboswitches. a, Cobalamins contain a corrin ring with a cobalt atom coordinated by an α-axial dimethylbenzimidazole (DMB) and a variable β-axial group. R denotes adenosylocobalamin (AdoCbl) (i), methylcobalamin (MeCbl) (ii), aquocobalamin (AqCbl) (iii) or cyanocobalamin (CNCbl) (iv). b, Secondary structure of the cobalamin riboswitch family. The conserved core is shown in blue, the KL interaction in green, and peripheral extensions distinguishing the two classes are shown in black (AdoCbl) and red (AqCbl). c, Chemical probing of the env8AqCbl riboswitch in the presence of cobalamins. The A and G sequencing lanes are shown to the left, followed by no probing reagent or ligand controls. NMIA, N-methylisatoic anhydride. d, Repression of GFPuv expression by the env8AqCbl riboswitch in E. coli. Red circles and orange triangles denote wild-type env8AqCbl riboswitch in the presence of AqCbl and AdoCbl, respectively; blue diamonds denote a leader sequence that spans the crystallized riboswitch sequence; and black squares denote an L5 mutant that cannot form the KL with L13. Error bars represent ± s.d. The inset shows plates of the wild-type and L5 mutant grown in the absence (−) or presence (+) of AqCbl.

---

©2012 Macmillan Publishers Limited. All rights reserved
extension of P6 in one class (representing approximately 95% of the sequences of the cobalamin family\(^{10}\)) that is absent or severely truncated in the other class\(^{8,9}\) (Fig. 1b). Another peripheral extension between P1 and P3 further defines the classes. A unique feature of numerous cobalamin riboswitches, particularly those without the P6-extension, is that regulation seems to be achieved through tertiary structure formation (the KL interaction).

All cobalamin riboswitches are proposed to bind adenosylcobalamin (AdoCbl), despite experimental validation of only a few member sequences\(^{5,6,10}\). Unexpectedly, we found some riboswitches recognize methylcobalamin (MeCbl) and aquacobalamin (AQcbl) with more than 500-fold higher affinity than AdoCbl (Supplementary Fig. 2 and Supplementary Table 1). Chemical probing\(^{12}\) shows \(B_2\)-dependent reactivity changes in a riboswitch lacking the P6-extension only in the presence of derivatives with small \(\beta\)-axial moieties (Fig. 1a), unlike the \(E.\ coli\ btuR\) riboswitch that selectively binds adenosylcobalamin\(^{5,6}\) (Fig. 1c; Supplementary Fig. 3). Considerable reactivity changes are localized to three regions: the central junction (J6/3), L5 and L13, the latter two suggestive of KL formation. Cobalamin selectivity and the role of KL formation in gene regulation were validated by cell-based assays. Control of reporter gene expression by a riboswitch lacking the P6-extension was tested in a \(\Delta btuR\) (cobalamin adenosyl transferase-deficient) strain of \(E.\ coli\) incapable of converting AQcbl to AdoCbl (ref. 13).

Expression of the green fluorescent protein variant GFPuv is repressed ~8.5-fold by AQcbl, but less than 2-fold by AdoCbl (Fig. 1d). A mutation of L5 in the mRNA (L5-GAAA) that retains cobalamin binding but eliminates KL formation abolishes repression, directly establishing the essential role of tertiary structure formation in regulating gene expression. Notably, similar sequences group on a branch of a cobalamin riboswitch phylogenetic tree highly populated by marine cyanobacterial\(^{8}\) and environmental (\(env\)) metagenomes of samples from the ocean surface\(^{14}\). The prevalence of these riboswitches in these bacteria suggests adaptation of the RNA to environments in which the free cobalamin pool is predominantly AQcbl owing to rapid photolysis of AdoCbl (ref. 15). Therefore, we refer to riboswitches containing the P6-extension as AdoCbl, and those lacking this extension as AQcbl, reflecting their probable biological effectors.

To determine the mechanistic basis for cobalamin-dependent regulation we solved structures of the \(env8\) AQcbl riboswitch in complex with AQcbl (Fig. 2a) and the \(Thermoanaerobacter tengcongensis\) (\(Tte\)) AdoCbl riboswitch bound to AdoCbl (Fig. 2b and Methods). AdoCbl riboswitches are the largest of the known riboswitches—at more than 200 nucleotides they are the size of the \(Azoarcus\) group I self-splicing intron\(^{16}\). The \(env8\) AQcbl riboswitch is the first structure containing both the receptor and regulatory domain. Insertion of the sequence spanning the crystal structure upstream of a GFPuv reporter confers cobalamin-dependent regulatory activity, albeit at lower efficiency than the wild-type riboswitch (Fig. 1d, crystal). Thus, the \(env8\) AQcbl structure corresponds to a completely functional riboswitch encompassing all of the sequence necessary and sufficient to impart biological activity.

The global architecture of both RNAs is defined by organization of their common secondary structure into two coaxial stacks, P1–P3–P6 and P4–P5–P13 (Fig. 2, blue and green, respectively), consistent with comparative analysis of known RNA structures\(^{17}\) and mutual information sequence analysis\(^{18}\). These stacks are joined by a T-loop–T-loop motif (L4–L6), a common module of RNA tertiary architecture\(^{18,19}\). In \(Tte\) AdoCbl, the T-loop (L4) interacts with an internal loop between P6 and P7 that partially mimics the structure of the T-loop (Supplementary Fig. 4), rather than another T-loop as observed in the AQcbl and flavin mononucleotide riboswitches\(^{20}\). Class-specific peripheral extensions flank the core (Fig. 2, magenta and cyan) that contact J6/3, a key element for cobalamin recognition.

Interactions between the RNA and cobalamin are mediated primarily through van der Waals shape complementarity with few direct hydrogen bonds. In both structures, cobalamin is sandwiched between the minor grooves of the P3–P6 coaxial stack and the helix created by base pairing of L5 and L13, forming the regulatory KL (Fig. 2). The J3/4 and J6/3 strands are central to the receptor side of the binding pocket; in \(env8\) AQcbl, stacking of four purines from these strands creates a relatively flat surface (G19, A20, A67 and A68; Fig. 3a). The \(\beta\)-axial face of cobalamin projects directly towards this surface, so the plane of the corrin ring is almost perpendicular to the bases of the purine stack. Despite numerous propionamide and acetamide groups surrounding the corrin ring, only one acetamide contacts the minor groove edge of G19 in \(env8\) AQcbl and its equivalent in \(Tte\) AdoCbl (G49) (Fig. 3b, c). A SELEX-generated aptamer uses a similar strategy to bind cyano-cobalamin, although the molecular details of recognition are quite different\(^{21}\) (Supplementary Fig. 5).

Selectivity between cobalamin derivatives is achieved through conformational differences in J6/3 mediated by the peripheral extensions. In \(env8\) AQcbl, proximity of A20 and A68 to the corrin ring sterically occludes the 5’-deoxyadenosyl moiety of AdoCbl, establishing its selectivity for cobalamins with small \(\beta\)-axial moieties (Fig. 3b). The conformation of J6/3 that blocks AdoCbl binding is enforced through base pairing with the J1/3 peripheral extension (G10-U69 and C11-G70). In \(Tte\) AdoCbl, placement of J6/3 further from J3/4 allows the Hoogsteen face of A162, the equivalent of A68 in \(env8\) AQcbl, to base pair with the 5’-deoxyadenosyl moiety (Fig. 3c). Positioning of J6/3 is reinforced by two highly conserved adenosines (A130 and A131) in the internal loop between P10 and P11 of the P6-extension (Fig. 3c and Supplementary Fig. 6), in support of the finding that this extension is
required for AdoCbl specificity (Supplementary Table 1). These adenosines are strongly protected from chemical modification only in the presence of AdoCbl (J11/10, Supplementary Fig. 3), suggesting that the peripheral domain docks with the core after initial cobalamin binding. Because other cobalamins can bind the AdoCbl riboswitch with ~80-fold lower affinity (Supplementary Table 1), we propose that the P6-extension may allow these riboswitches to regulate gene expression differentially on the basis of whether the intracellular cobalamin pool is dominated by AdoCbl or its photolysed product AqCbl. Notably, a recently discovered repressor protein in *Myxococcus xanthus* uses B12 to photoregulate carotenoid biosynthesis22. Use of peripheral extensions by a riboswitch to modulate its properties has not been previously observed, but is common with other large RNAs23,24.

The receptor–cobalamin complex presents a composite surface composed of RNA and ligand that binds the regulatory domain. Superimposition of the full env8AqCbl riboswitch with a structure of the receptor domain alone indicates that the receptor–ligand complex provides a relatively rigid surface for docking of P13 (Supplementary Fig. 7). RNA–RNA interactions between the two domains are through base-pairing of nucleotides in L5 and L13 to form the KL. In both structures, the α-axial face of B12 interacts with the KL through structure-specific rather than sequence-specific contacts. Dimethylbenzimidazole and the aminopropyl linker interact with the ribose-phosphate backbone of L5 and L13 primarily through van der Waals contacts; the only direct hydrogen bond is between the 2'-hydroxyl group of the cobalamin ribosyl moiety and the 2'-hydroxyl group of C91/C194 (Fig. 4a). Further contacts are made to the RNA by propionamide and acetamide groups with the only direct base contact being to the sugar edge of U44 (*env8* AqCbl) or its equivalent, G74 (*Tte* AdoCbl), in L5. The KL–cobalamin interface is nearly identical between the two structures despite low sequence similarity in L5 and L13, indicating structure-specific recognition of the KL by B12.

Kissing-loops are inherently stable structures widely used to promote tertiary architecture formation and RNA–RNA interactions25,26. A unique feature of the KL in both structures is the presence of non-Watson–Crick pairs and bulged nucleotides in the helix formed by pairing of L5 and L13 (Fig. 4a, labelled MM and B). Mismatched pairs are strongly destabilizing to the KL27, suggesting a mechanism for creating a cobalamin-dependent switch. To show bound cobalamin promotes KL formation, its structure was probed as a function of Mg2+, which promotes long-range tertiary interactions in RNA28. Chemical probing of the unbound AqCbl riboswitch shows that at least 15 mM Mg2+ is required before reactivity protections consistent with the bound state. The receptor–ligand complex of *env8* AqCbl (Fig. 3) defines the shape complementarity between RNA and ligand, emphasizing the extensive structural changes in the receptor domain that accommodate B12, consistent with the data presented in Supplementary Fig. 3.
with KL formation are observed, whereas bound ligand allows its formation under physiological Mg$^{2+}$ concentrations (0.5–1 mM) (Fig. 4b and Supplementary Fig. 8). This same trend is observed for the _E. coli btuB_ AqCbl riboswitch (Supplementary Fig. 9). Stabilizing the KL with mutations in L13 that enable perfect Watson–Crick pairing with L5 (G95U, ΔG97) results in cobalamin-independent KL formation, equivalent to a constitutively repressed mRNA due to sequestration of the RBS (Fig. 4c). Together, these data show that cobalamin riboswitches use ligand-dependent formation of a tertiary RNA module as the basis for gene regulation. This is a structurally distinct but functionally equivalent mechanism to most other RNA switches that use mutually exclusive secondary structures to achieve regulatory activity.

**METHODS SUMMARY**

RNAs were transcribed by T7 RNA polymerase, purified by denaturing methods and refolded before crystallization with either AqCbl or AdoCbl (ref. 29). Diffraction data was collected at the Advanced Light Source (beamlines 5.0.2, 8.2.1 and 8.2.2), and solved using a combination of multiwavelength anomalous dispersion and molecular replacement approaches. Data collection and refinement statistics are presented in Supplementary Tables 2 and 3. Biochemical analysis of an AdoCbl riboswitch was performed using a sequence that regulates the _E. coli btuB_ gene; the env4AqCbl and env8AqCbl sequences were derived from the Pacific Ocean metagenome (85% identity shared between them). All experiments using AdoCbl or MeCbl were performed using a red safe light to minimize photo-lysis of this compound.

Full Methods and any associated references are available in the online version of the paper.

Received 2 April; accepted 25 September 2012.

**Published online 14 October 2012.**

1. Garst, A. D., Edwards, A. L. & Battey, R. T. Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.* 3, a003533 (2011).
2. Breaker, R. R. Riboswitches and the RNA world. *Cold Spring Harb. Perspect. Biol.* 4, a003566 (2012).
3. Lundrigan, M. D., Koster, W. & Kadner, R. J. Transcribed sequences of the _Escherichia coli btuB_ gene control its expression and regulation by vitamin B$_{12}$. *Proc. Natl Acad. Sci. USA* 88, 1479–1483 (1991).
4. Franklund, C. V. & Kadner, R. J. Multiple transcribed elements control expression of the _Escherichia coli btuB_ gene. _J. Bacteriol._ 179, 4039–4042 (1997).
5. Nahvi, A., Barrick, J. E. & Breaker, R. R. Coenzyme B$_{12}$ riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.* 32, 143–150 (2004).
6. Nahvi, A. et al. Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9, 1043 (2002).
7. Rodionov, D. A., Viteschak, A. G., Mironov, A. A. & Gelfand, M. S. Comparative genomics of the vitamin B$_{12}$ metabolism and transport in bacteria by a conserved RNA structural element. *RNA* 9, 1084–1097 (2003).
8. Viteschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Regulation of the vitamin B$_{12}$ metabolism and transport in bacteria by a conserved RNA structural element. *RNA* 9, 1084–1097 (2003).
9. Barrick, J. E. & Breaker, R. R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* 8, R239 (2007).
10. Gardner, P. P. et al. Rfam: updates to the RNA families database. *Nucleic Acids Res.* 37, D136–D140 (2009).
11. Fox, K. et al. Multiple posttranscriptional regulatory mechanisms partner to control ethanolamine utilization in _Enterococcus faecalis_. *Proc. Natl Acad. Sci. USA* 106, 4435–4440 (2009).
12. Wilkinson, K. A., Merino, E. J. & Weikel, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nature Protocols* 1, 1610–1616 (2006).
13. Lundrigan, M. D. & Kadner, R. J. Altered cobalamin metabolism in _Escherichia coli btuR_ mutants affects btuB gene regulation. _J. Bacteriol._ 171, 154–161 (1989).
14. Weinberg, Z. et al. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metabolomes. *Genome Biol.* 11, R31 (2010).
15. Taylor, R. T., Smucker, L., Hanna, M. L. & Gill, J. Aerobic photolysis of alkylcobalamins: quantum yields and light-action spectra. *Arch. Biochem. Biophys.* 196, 521–533 (1975).
16. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. Crystal structure of a self-splicing group I intron with both exons. *Nature* 430, 45–50 (2004).
17. Jaeger, L., Verzemnieks, E. J. & Geary, C. The UA handle: a versatile submotif in large RNA molecules. *Nucleic Acids Res.* 37, 215–230 (2009).
18. Nagaswamy, U. & Fox, G. E. Frequent occurrence of the T-loop RNA folding motif in ribosomal RNAs. *RNA* 8, 1112–1119 (2002).
19. Krasnikov, A. S. & Mondragon, A. On the occurrence of the T-loop RNA folding motif in large RNA molecules. *RNA* 9, 640–643 (2003).
20. Sernov, A., Huang, L. & Patel, D. J. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* 458, 233–237 (2009).
21. Sussman, D., Nix, J. C. & Wilson, C. The structural basis for molecular recognition by the vitamin B₁₂ RNA aptamer. *Nature Struct. Biol.* 7, 53–57 (2000).
22. Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S. & Elias-Arnanz, M. Light-dependent gene regulation by a coenzyme B₁₂-based photoreceptor. *Proc. Natl Acad. Sci. USA* 108, 7565–7570 (2011).
23. Vicens, Q. & Cech, T. R. Atomic level architecture of group I introns revealed. *Trends Biochem. Sci.* 31, 41–51 (2006).
24. Melnikov, S. et al. One core, two shells: bacterial and eukaryotic ribosomes. *Nature Struct. Mol. Biol.* 19, 560–567 (2012).
25. Xin, Y., Laing, C., Leontis, N. B. & Schlick, T. Annotation of tertiary interactions in RNA structures reveals variations and correlations. *RNA* 14, 2465–2477 (2008).
26. Butcher, S. E. & Pyle, A. M. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res.* 44, 1302–1311 (2011).
27. Gregorian, R. S. Jr & Crothers, D. M. Determinants of RNA hairpin loop-loop complex stability. *J. Mol. Biol.* 248, 968–984 (1995).
28. Pyle, A. M. Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem.* 7, 679–690 (2002).
29. Reyes, F. E., Garst, A. D. & Batey, R. T. Strategies in RNA crystallography. *Methods Enzymol.* 469, 119–139 (2009).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by grants from the National Institutes of Health (GM073850 and 1S10RR026516) to R.T.B. and by a Colorado Diversity Initiative Fellowship and NIH Ruth L. Kirschstein fellowship (F32GM095121) to J.E.J. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract no. DE-AC02-05CH11231.

**Author Contributions** F.E.R. discovered the specificity of the AqCbl class and performed all aspects of the crystallography with assistance from J.T.P.; J.E.J. performed all biochemical experiments and fully characterized the specificities of the cobalamin family; J.T.P. obtained all in vivo data; and all authors contributed to the analysis of the data and the writing of this paper.

**Author Information** Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession numbers 4FRG (env8AqCbl(D13,P13)), 4FRN (env8AqCbl) and 4QMA (7eAdoCbl). Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.T.B. (robert.batey@colorado.edu) or F.E.R. (francis.reyes@colorado.edu).
RESEARCH LETTER

METHODS

RNA preparation. RNA constructs (full sequences are shown in Supplementary Fig. 1) were prepared using DNA templates generated from PCR amplification using established protocols. DNA templates for the E. coli btuB cobalamin riboswitch were amplified from E. coli genomic DNA, whereas the env4, env8, and T. tengcongensis cobalamin riboswitches were amplified using a series of overlapping oligonucleotides. PCR products were used as a template for transcription reactions using T7 RNA polymerase and the RNA purified using denaturing PAGE (8% or 12%, 29:1 acrylamide:bisacrylamide). The RNA was buffer exchanged and concentrated into 0.5× TE buffer and frozen at −80°C until use. Isothermal titration calorimetry. RNA was dialysed overnight into a buffer containing 5 mM sodium-MES, pH6.0, 100 mM KC1, and 5 mM MgCl2. RNA was diluted up to a final concentration of 10 μM and titrated with AdoCbl, AqCbl or MeCbl dissolved in the dialysis buffer at concentrations tenfold in excess of the RNA. Titrations were all performed at 25°C using a MicroCal iTC200 microcalorimeter. Data analysis and fitting was performed with the Origin software suite as previously described.

Chemical probing with NMIA. Chemical probing of RNAs were performed using slight modifications to established protocols. Purified RNA constructs were refolded by incubation at 70°C for 3 min, room temperature for 5 min, and on ice for a minimum of 5 min. For ligand comparison experiments, 10 μl solutions were prepared to a final concentration of 0.1 μM RNA, 100 mM potassium-phosphate buffer pH8.0, 100 mM NaCl, 6 mM MgCl2, and 6.5 mM NMIA. Final ligand concentrations for the E. coli btuB experiments were 0.5 μM AdoCbl, 2.5 mM AqCbl, CNCbl and MeCbl. For the env4-AqCbl experiments, final ligand concentrations were 100 μM for all compounds. For magnesium titrations, the conditions were the same as the above mentioned conditions with the final concentrations of ligands as 500 μM AdoCbl or 100 μM MeCbl. Samples were reverse transcribed as previously described. Products were separated using 12% denaturing polyacrylamide gel and visualized using a Typhoon Phosphorimager (Molecular Dynamics).

In vivo reporter assay. For all in vivo assays, E. coli ΔbtuR (Keio collection) cells were grown in a rich, chemically defined medium that was supplemented with (100 μg ml−1) ampicillin and a varying amount of a cobalamin. For titration experiments, 5 μl of overnight culture was added to 5 ml of media and incubated for 3 h at 37°C. Fluorescence and A600nm measurements were performed on 300 μl of cells from each replicate in clear-bottom 96-well plates. GFPuv fluorescence was read at an excitation wavelength of 395 nm and a 510 nm emission wavelength using an Xfluor SafireII fluorimeter (Tecan). All data shown represent average fluorescence values of three biological replicates that were normalized to the A600 nm in each well. All fluorescence measurements were background corrected by taking identical fluorescence and A600nm measurements for E. coli ΔbtuB cells transformed with pHJ322 vector that did not contain the reporter gene. Background fluorescence was subtracted from total fluorescence and the fold-repression was calculated by dividing the average normalized/background corrected fluorescence values for the unrepressed construct (−Cbl) by the average normalized/background corrected fluorescence value for each repressed construct (+Cbl). Titration data was fit to a two state equation to determine half-maximum effective concentration (EC50) values, and error bars represent the standard deviation for each triplicate fluorescence measurement.

Structure solution of the env8AqCbl–AqCbl complex. The original env8AqCbl riboswitch (GenBank accession AACY021350931.1/557–442) was modified with the following mutations: G12A, A14G, A31U, G42C and C62G, and the replacement of the linker region with that of a sequence from the original alignment (GenBank accession AACY023653040/384–265). Each RNA was synthesized as refolded as described earlier. Crystals grew within 3 months at 20°C under 10% (v/v) 2-methyl-1,3-propanediol, 40 mM sodium cacodylate, pH 7.0, 12 mM spermine tetrahydrochloride, 80 mM KC1 and 20 mM BaCl2, by hanging drop vapour diffusion. A complete data set was collected on a crystal flash frozen in liquid nitrogen at the cobalt anomalous edge and reduced with autoPROC. The structure was solved via molecular replacement using env8AqCbl(A11/13,P13)–AqCbl. The solution was confirmed by anomalous difference fourier maps showing the location of aquocobalamin. BUSTER refinement produced clear density for P5 and P13 (Supplementary Fig. 10c, d). Model building was assisted with CRANDER within COOT. The final model was prepared using restrained refinement within BUSTER. Data collection, phasing and refinement statistics for the env8-AqCbl and env8AqCbl(A11/13,P13)–AqCbl can be found in Supplementary Table 2.

Structure solution of the TtAdoCbl–AdoCbl complex. The sequence of the TtAdoCbl riboswitch (GenBank accession AE008691.1/395133–395733) was modified to substitute the human U1A-binding protein RNA motif and a stable GAAA closing tetraloop for the non-conserved P2 and P10 helices, respectively. The RNA was synthesized and refolded as described earlier, to which 2 μl of a 4 mM solution of AdoCbl was added. Crystals grew over the course of 3 days at 30°C by the hanging drop method in 10% isopropanol, 300 mM MgCl2 and 100 mM sodium-hepes, pH7.5, in the absence of U1A and cryoprotected in 30% glycerol. Heavy atom derivatives were prepared by including 2–10 mM compound in the mother liquor during crystallization.

All data sets were collected at the cobalt anomalous edge for native data sets and the anomalous edges for TaBr and Ir derivatives. Molecular replacement in Phaser was using env8AqCbl RNA yielded a solution sufficient to locate heavy atoms in all data sets using anomalous difference data. Cross-crystal dispersive differences for cobalt were used to pair native data sets with heavy atom derivatives to create three sets of experimental phases (Supplementary Table 3). The phases were combined via DMMULTI resulting in traceable density maps. Iterative cycles of model building, refinement in BUSTER and PHENIX, multiple crystal averaging and heavy atom phasing, allowed most of the RNA to be built; the U1A loop at the apex of P2, the junction between P8 and P10, and J1/J3 remain unresolved in the electron density. Supporting electron density maps are shown in Supplementary Fig. 12. Data collection, phasing and refinement statistics can be found in Supplementary Table 3.

30. Edwards, A. L., Garst, A. D. & Batey, R. T. Determining structures of RNA aptamers and riboswitches by X-ray crystallography. Methods Mol. Biol. 535, 135–163 (2009).
31. Milligan, J. F. & Uhlenbeck, O. C. Synthesis of small RNAs using T7 RNA polymerase. Methods Enzymol. 180, 51–62 (1989).
32. Gilbert, S. D. & Batey, R. T. Monitoring RNA–ligand interactions using isothermal titration calorimetry. Methods Mol. Biol. 540, 97–114 (2009).
33. Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 6006.0008 (2006).
34. Vorhein, C. et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D 67, 293–302 (2011).
35. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D 65, 688–697 (2009).
36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of COOT. Acta Crystallogr. D 67, 213–221 (2001).
37. Blanc, E. et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D 66, 2210–2221 (2004).
38. Keating, K. S. & Pyle, A. M. Semiautomated model building for RNA crystallography using a directed rotamer approach. Proc. Natl Acad. Sci. USA 107, 8177–8182 (2010).
39. Ferré-D’Amare, A. R. Use of the spliceosomal protein U1A to facilitate crystalization and structure determination of complex RNAs. Methods 52, 159–167 (2010).
40. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).