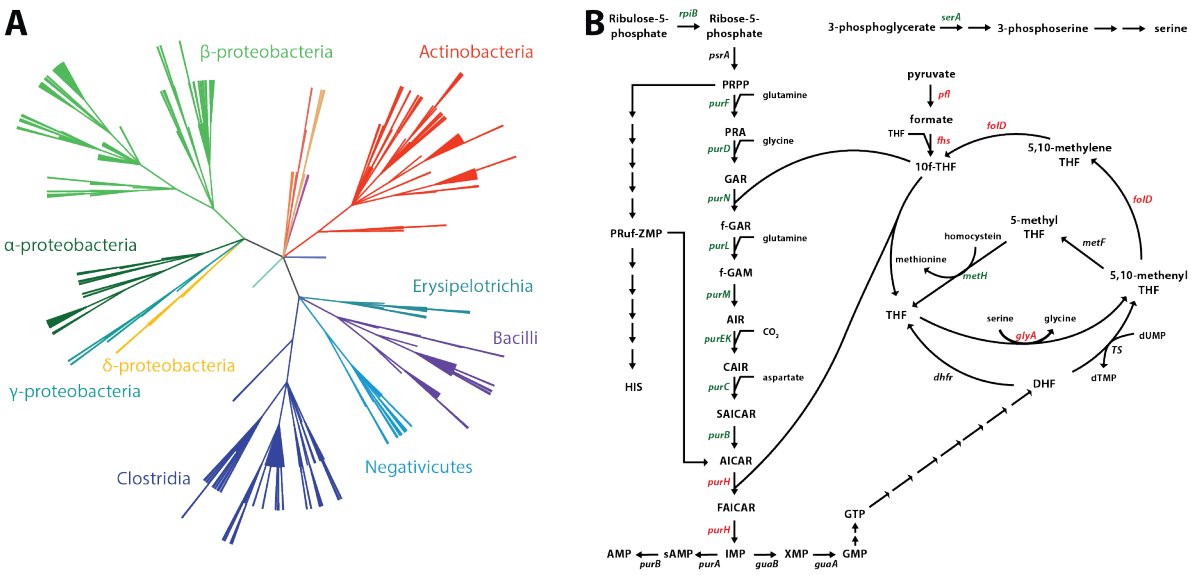

Figure S1. The *pfl* Element Is a Widespread and Highly Conserved RNA Motif in Bacteria that Regulates Genes Associated with Purine Synthesis and One-carbon Metabolism, Related to Figure 1

(A) Phylogenetic tree of fully sequenced organisms that have *pfl* motif RNAs identified by bioinformatics (See **Extended Experimental Procedures** and **Supplemental Data**). In addition to about 600 examples of *pfl* motif RNAs found in fully sequenced organisms, there are about 6,000 additional examples found in environmental sequences. The latter number includes duplicates identified from near-identical genomic DNAs, and so there are about 1,400 unique representatives among environmental sequences.

(B) Genes most frequently associated with *pfl* motif RNAs are highlighted in red and genes that are occasionally associated are highlighted in green. Genes in gray have not been observed to be associated with *pfl* motif RNAs.

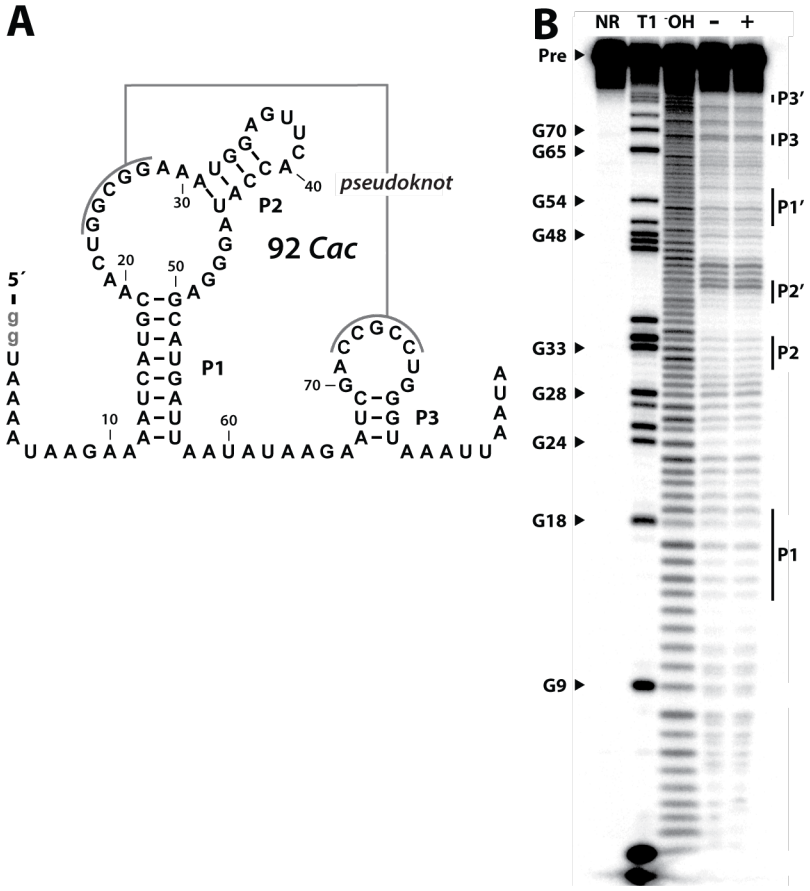

Figure S2. 92 *Cac* RNA Misfolds Under In-line Probing Reaction Conditions, Related to Figure 2

(A) Sequence and secondary structure model for the 92 *Cac* RNA from *C. acetobutylicum*. Nucleotides in gray were added to facilitate transcription.

(B) PAGE analysis of in-line probing products generated in the absence or presence of 1 mM ZMP. If properly folded, nucleotides that form stems are expected to be more resistant to RNA cleavage. However, substantial cleavage product bands are generated by cleavage at nucleotides predicted to reside in P1 and P2 stems, suggesting that these highly conserved secondary structures are not folded properly under our reaction conditions.

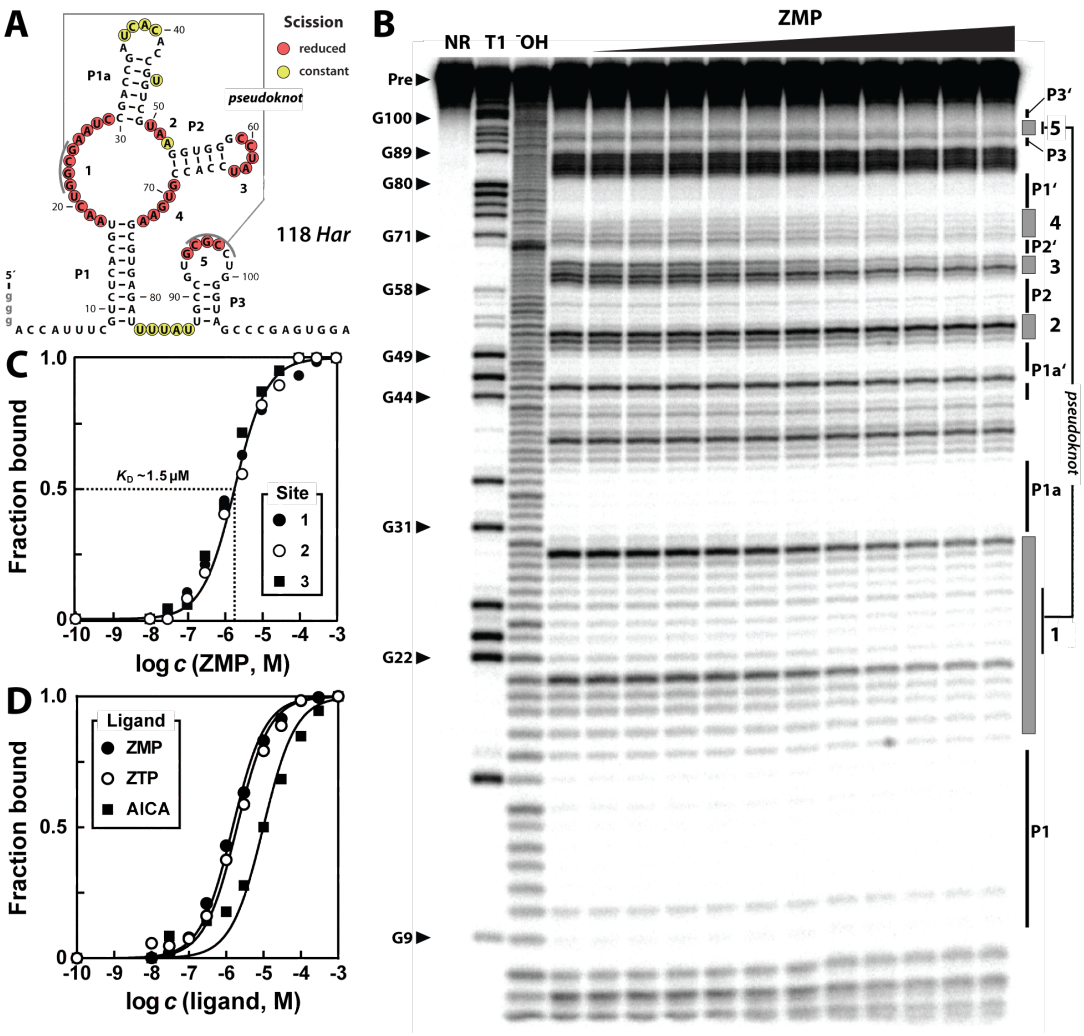

Figure S3. ZMP Binding by 118 *Har* from *H. arsenicoxydans*, Related to Figure 2

(A) Sequence and secondary structure model for 118 *Har* RNA. Annotations are as described for **Figure 2**. Data are derived from the in-line probing results presented in B.

(B) PAGE analysis of in-line probing assay product generated in the presence of various concentrations of ZMP. Annotations are as described for **Figure 2**.

(C) Plot of the normalized fraction of RNA cleavage versus the concentration ZMP using the data presented in B. The solid line represents a theoretical binding curve expected for a one-to-one interaction.

(D) Comparison of the apparent K_D values for selected ZMP analogs with 118 *Har* RNA as determined by in-line probing (data not shown). Additional annotations are as described for C.

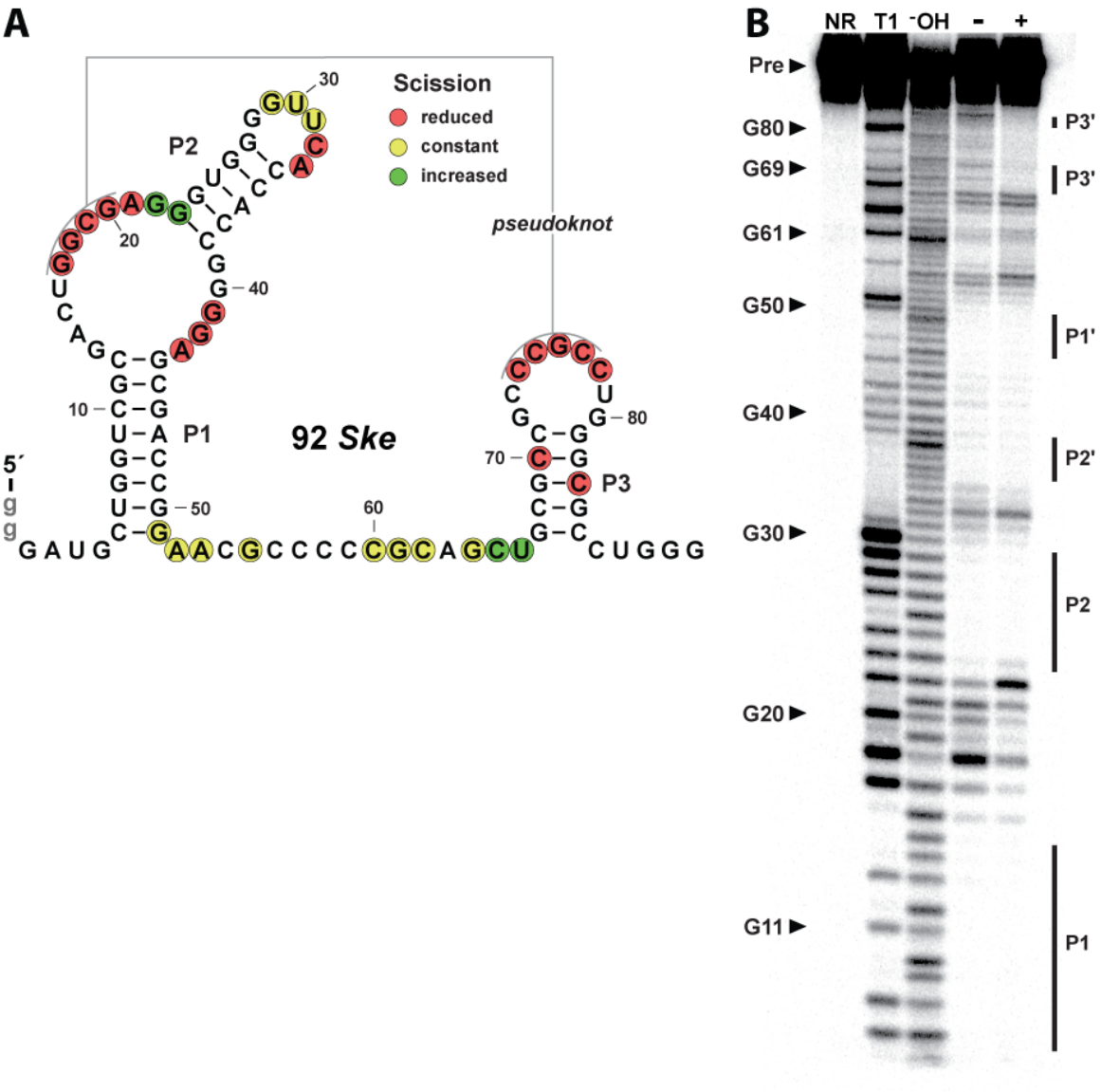

Figure S4. ZMP Binding by 92 *Ske* from *S. keddieii*, Related to Figure 2

(A) Sequence and secondary structure model for 92 *Ske* RNA. Annotations are as described in **Figure 2**. Data are derived from the in-line probing results presented in B.

(B) PAGE analysis of in-line probing assay products generated in the absence or presence of 1 mM ZMP.

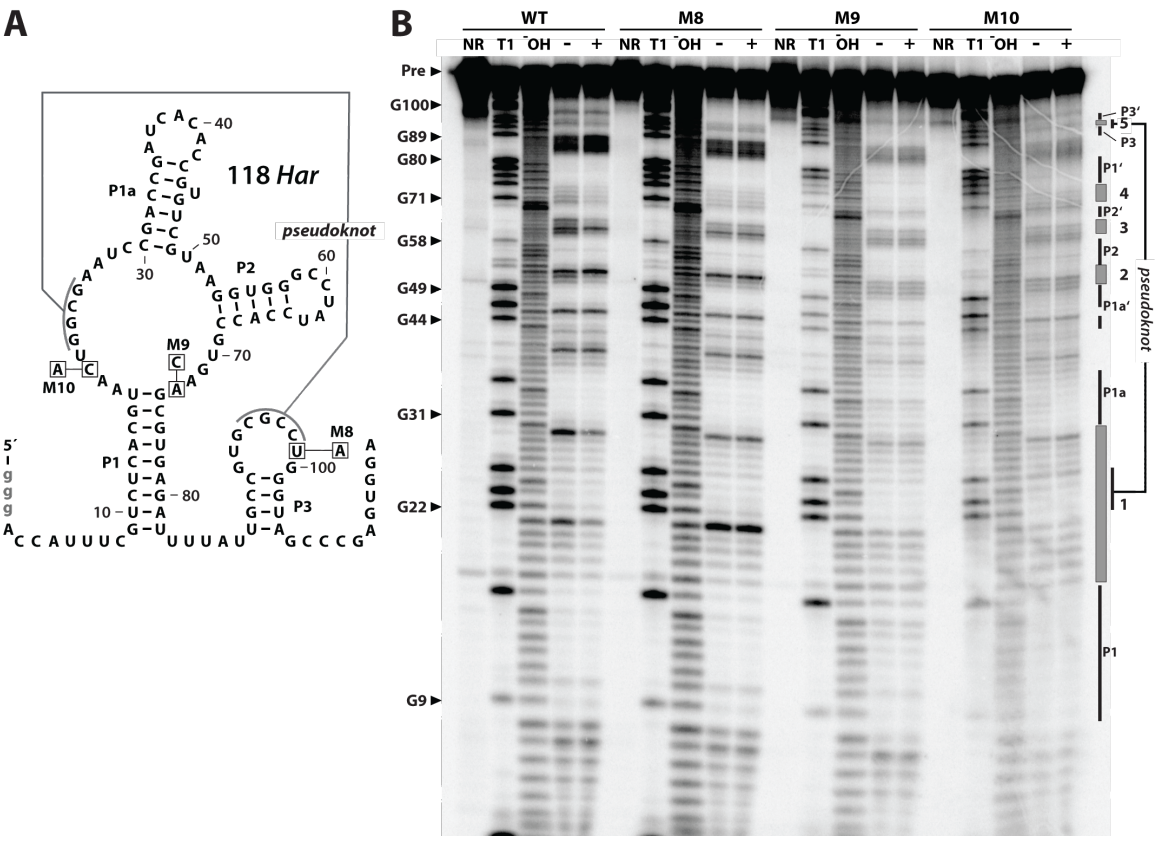

Figure S5. Mutational Analysis of 118 *Har* RNA Mutants by In-line Probing, Related to Figure 3

(A) Sequence and secondary structure of the 118 *Har* RNA construct and various mutations used to evaluate the importance of certain sequences and substructures.

(B) PAGE image of in-line probing reactions for each of the construct after incubation with 100 μM ZMP. Annotations are as described for **Figure 2**.

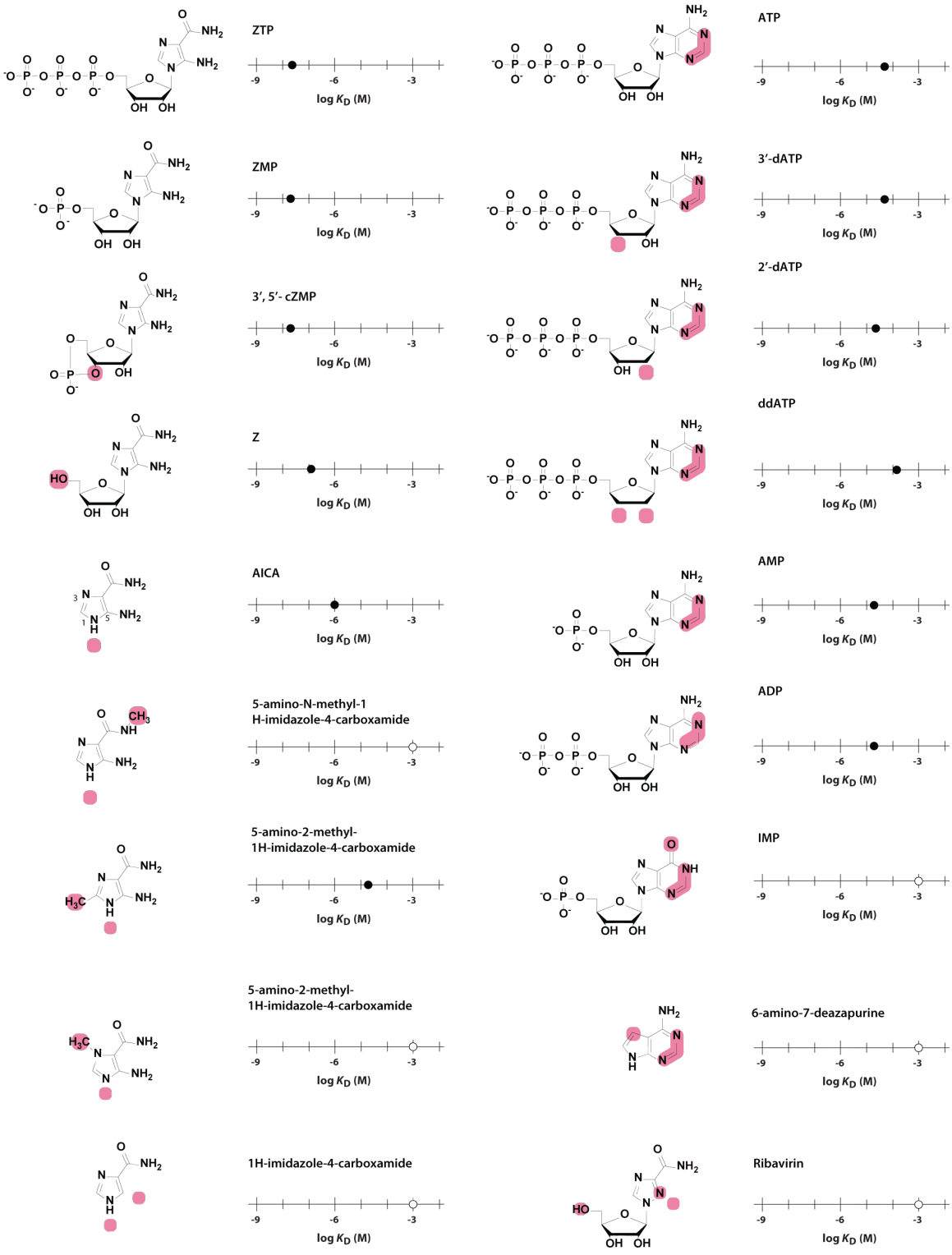

Figure S6. Molecular Recognition Characteristics of a ZMP/ZTP Aptamer, Related to Figure 4

Chemical structures and apparent K_D values for various ZMP analogs as established by in-line probing with 104 *Cba* RNA. Differences in chemical structure relative to ZMP are highlighted. Open circles indicate no observed binding at the concentrations designated.

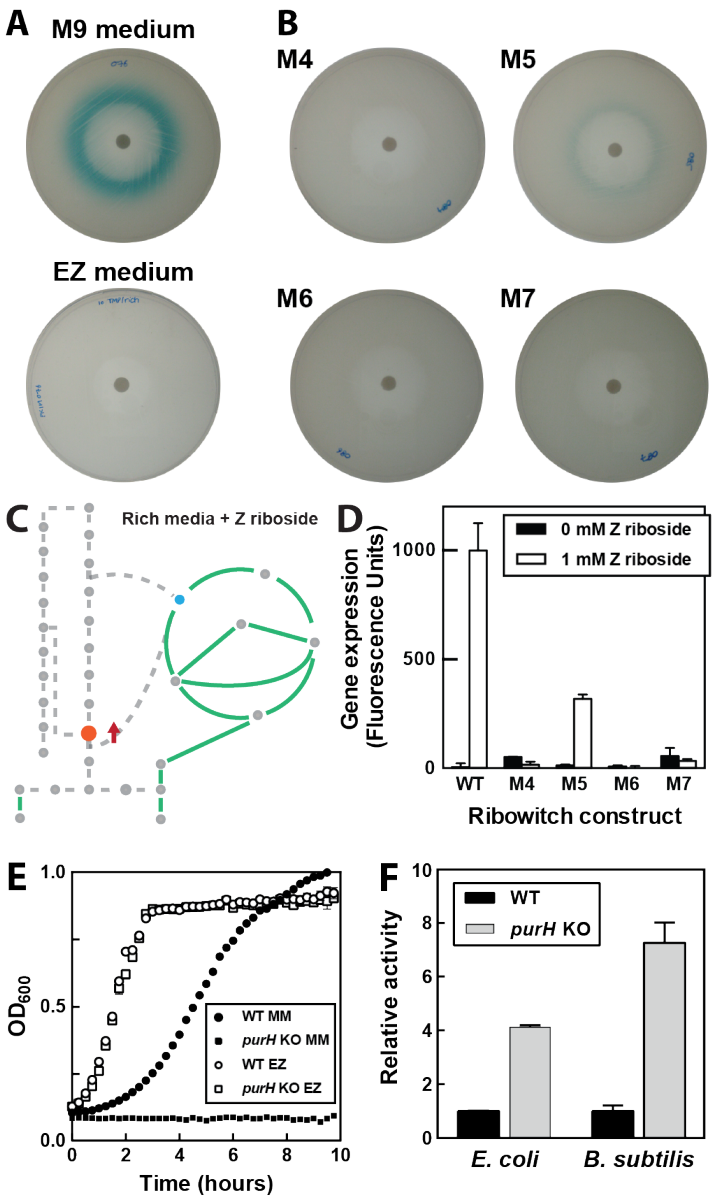

Figure S7. Assessment of Gene Control by the *P. carotovorum* ZTP Riboswitch, Related to Figure 6

(A) Agar diffusion assays of *E. coli* transformed with WT ZTP riboswitch reporter construct. The reporter strain was grown on agar plates with M9 minimal medium or EZ rich medium supplemented with X-gal. The disk contains the antibiotic TMP (1 μg mL^{-1}).

(B) Agar diffusion assays of *E. coli* transformed with mutant versions of ZTP riboswitch reporter construct. The mutant reporter strains were grown on agar plates with M9 minimal medium supplemented with X-gal. The disk contains the antibiotic TMP (1 μg mL^{-1}). Mutants M4 through M8 are as depicted in **Figure 6**.

(C) Schematic of the expected metabolic flux upon addition of Z riboside in rich medium. (D) Plot of the expression levels of WT and various mutant reporter constructs (**Figure 6A**) in cells treated with Z riboside.

(E) Comparison of the growth of *E. coli* WT and *purH* KO strains in minimal (M9) and rich (EZ) medium.

(F) Plots of reporter gene expression levels after growth in rich (EZ) medium for 20 hr at 37ºC.

Table S1. Oligonucleotides Used In This Study

Name	DNA sequence (5' to 3')	Description
1	TAATACGACTCACTATAGGGAAATGGTCAAGTGA CTGGTGGAAACATCAAAAGATGCGTAGGATAACC TACAGGGAGCTTGATTATATAT	In-line probing, WT 104 *Cba* fwd; Two-step PCR
2	CATACGTGCCCAGGCGGTCGGCAATTTCTAAATA TATAATCAAGCTCCC	In-line probing, WT, M1, M2 104 *Cba* rev; Two-step PCR
3	TAATACGACTCACTATAGGGAAATGGTCAAGTGA CAGGTGGAAACATCAAAAGATGCGTAGGATAAC CTACAGGGAGCTTGATTATATAT	In-line probing, M1 104 *Cba* fwd; Two-step PCR
4	TAATACGACTCACTATAGGGAAATGGTGTAGTGA CTGGTGGAAACATCAAAAGATGCGTAGGATAACC TACAGGGAGCTTGATTATATAT	In-line probing, M2 104 *Cba* fwd; Two-step PCR
5	TAATACGACTCACTATAGGGAAATGGTGTAGTGA CTGGTGGAAACATCAAAAGATGCGTAGGATAACC TACAGGGAGCTACATTATATAT	In-line probing, M3 104 *Cba* fwd; Two-step PCR
6	CATACGTGCCCAGGCGGTCGGCAATTTCTAAATA TATAATGTAGCTCCC	In-line probing, M3 104 *Cba* rev; Two-step PCR
7	TAATACGACTCACTATAGGGATGCTGGTCGCGAC TGGCGAGGGTGGGGTTCACCACCGGGGAGCGAC CGGAACGCCCCCGCAGCTGCGCC	In-line probing, WT 92 *Ske* fwd; Two-step PCR
8	CCCAGGCGCCCAGGCGGGCGGCGCAGCTGCGG	In-line probing, WT 92 *Ske* rev; Two-step PCR
9	TAATACGACTCACTATAGGGACCATTTCGTCTCA CGTAACTGGCGAATCCGACCGATCACACCGTGTC GTAAGGTGGGCCTATCCACCG	In-line probing, 118 *Har* WT, M8, M9 fwd; Two-step PCR
10	TCCACTCGGGCTACCCAGGCGCACGGCAATAAAA TCTCACGCTTCACGGTGGATAGGCCCACCTTACG	In-line probing, 118 *Har* WT, M10 rev; Two-step PCR
11	TCCACTCGGGCTACCCTGGCGCACGGCAATAAAA TCTCACGCTTCACGGTGGATAGGCCCACCTTACG	In-line probing, 118 *Har* M8 rev; Two-step PCR
12	TCCACTCGGGCTACCCAGGCGCACGGCAATAAAA TCTCACGCGTCACGGTGGATAGGCCCACCTTACG	In-line probing, 118 *Har* M9 rev; Two-step PCR
13	TAATACGACTCACTATAGGGACCATTTCGTCTCA CGTAAATGGCGAATCCGACCGATCACACCGTGTC GTAAGGTGGGCCTATCCACCG	In-line probing, 118 *Har* M10 fwd; Two-step PCR
14	TAATACGACTCACTATAGGTAAAATAAGAAAATC ATGCAACTGGCGGAAATGGAGTTCACCATAGGGA GCATGAT	In-line probing, 92 *Cac* fwd; Two-step PCR
15	TATTAATTTACCCAGGCGGTCGATTCTTATATTAA TCATGCTCCCTATGGTGAACT	In-line probing, 92 *Cac* rev; Two-step PCR
16	ATAAGTGTAAAATTATTCTTATTGACAAACTAA	Transcription termination assay, *Cbe pfl* motif and expression platform fwd; PCR from *C. beijerinckii* spores
17	ATCCTTCCCATTGTTTAAAC	Transcription termination assay, *Cbe pfl* motif and expression platform rev; PCR from *C. beijerinckii* spores
18	TAATACGACTCACTATAGGGTATTAGTCATATGA CTGACGGAAGTGGAGTTACCACATGAAGTATGAC TAGGCATAT	In-line probing, 102 *Cbe* fwd; Two-step PCR
19	TTTTTTTGCCCAGACGGTCGGCTTTTTGTGGCATA TAAGATAATATGCCTAGTCATACTTCATGTGG	In-line probing, 102 *Cbe* rev; Two-step PCR
20	TCATAAATGCGCTCATGTTACTGACGCTTAAAGA TAGATCACTATCGTGA	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* WT, M7 fwd; Two-step PCR

21	AAAAGTAAGTTATGCCCAGACGCGCGGCTCGTAATGACCCATGCTTCACGATAGTGATCTATC	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* WT, M4, M6 rev; Two-step PCR
22	AAAGAATTCGCATGACCCCATCGTTGACAACCGCCCCGCTCACCCTTTATTTATAAATGTACTCATAAATGCGCTCATG	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* PCR WT fwd; PCR using 20-21 PCR product
23	AAAGGATCCAAGGAAATGTACGTGATCAGGCATGACTTATCTCCAAAAGTAAGTTATGCCCA	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* PCR WT rev; PCR using 20-21 PCR product
24	TCATAAATGCGCTGTTGTTACTGACGCTTAAAGATAGATCACTATCGTGA	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* M4, M5 fwd; Two-step PCR
25	AAAAGTAAGTTATGCCCAGACGCGCGGCTCGTAATGACCGTTGCTTCACGATAGTGATCTATC	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* M5 rev; Two-step PCR
26	AAAGAATTCGCATGACCCCATCGTTGACAACCGCCCCGCTCACCCTTTATTTATAAATGTACTCATAAATGCGCTGTTG	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* PCR M4, M5 fwd; PCR using either 24-21 (for M4) or 24-25 (for M5) PCR product as the template
27	TCATAAATGCGCTCATGTTAATGACGCTTAAAGATAGATCACTATCGTGA	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* M6 fwd; Two-step PCR
28	AAAAGTAAGTTATGCCCTGACGCGCGGCTCGTAATGACCCATGCTTCACGATAGTGATCTATC	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* M7 rev; Two-step PCR
29	AAAGGATCCAAGGAAATGTACGTGATCAGGCATGACTTATCTCCAAAAGTAAGTTATGCCCT	Cloning of *P. carotovorum* ZTP riboswitch and first 8 codons of *rhtB* PCR M7 rev; PCR using 20-28 PCR product as template
30	TACTAAGACCAACAAACATAACGAC	Cloning of *Bacillus* SG-1 ZTP riboswitch and first 8 codons of *pfl* PCR fwd; PCR from *Bacillus* SG-1 genomic DNA
31	AAAGGATCCTTATTCCCTCCTGGTTAATGTCAGT	Cloning of *Bacillus* SG-1 ZTP riboswitch and first 8 codons of *pfl* PCR BamHI rev; PCR from *Bacillus* SG-1 genomic DNA
32	AAAGAATTCAAAAATAATGTTGTCCTTTTAAATAAGATCTGATAAAATGTGAACTAATACTAAGACCAACAAACATAACGAC	Cloning of *Bacillus* SG-1 ZTP riboswitch and first 8 codons of *pfl* PCR EcoRI fwd (introduces *Bacillus subtilis lysC* promoter); PCR using 31-31 PCR product as the template

EXTENDED EXPERIMENTAL PROCEDURES

Chemical Reagents

ZMP and AICA were purchased from Sigma-Aldrich. ZTP and 5′-amino-5′-deoxyadenosine were purchased from BioLog Life Science Institute. Z riboside, 3′, 5′-cZMP, ribavirin, mizoribine, and various AICA analogs were purchased from Toronto Research Chemicals. [γ-^{32}P] ATP and [α-^{32}P] GTP were purchased from PerkinElmer and used within two weeks of receipt. Bulk chemicals were purchased from J.T. Baker, unless otherwise noted. All solutions were prepared using deionized water (dH_2O) and either autoclaved or filter sterilized (0.22-μm filters, Millipore) before use.

Preparation of RNA Oligonucleotides

RNAs were prepared by in vitro transcription using DNA templates generated by primer extension of overlapping synthetic DNA oligonucleotides containing the promoter sequence for T7 RNA polymerase (Sigma-Aldrich or Integrated DNA Technologies). All primers used to generate wild-type and mutant DNAs are found in **Table S1**. In vitro transcription reactions were carried out using bacteriophage T7 RNA polymerase (T7 RNAP) in 80 mM N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES, pH 7.5 at 23ºC), 40 mM dithiothreitol (DTT), 24 mM $MgCl_2$, 2 mM spermidine, and 2 mM of each nucleoside 5′-triphosphate (NTP). RNA was purified using denaturing (8 M urea) 8% polyacrylamide gel electrophoresis (PAGE). The appropriately sized band, visualized by UV shadowing, was excised and the RNA was eluted from the crushed gel slice using 10 mM Tris-HCl (pH 7.5 at 23ºC), 200 mM NaCl, and 1 mM EDTA (pH 8.0). The RNA was subsequently precipitated with ethanol and pelleted by centrifugation.

To generate 5′ ^{32}P-labeled RNAs, the 5′-terminal triphosphate was removed using rAPid alkaline phosphatase (Roche Applied Science). The RNAs were then radiolabeled with [γ-^{32}P] ATP (Perkin Elmer) using T4 polynucleotide kinase (New England Biolabs). 5′ ^{32}P-labeled RNAs were purified and isolated as described above.

Single-round in vitro Transcription Termination Assays

Approximately 2 pmol of DNA template was added to transcription initiation buffer (20 mM Tris-HCl [pH 8.0 at 25 °C], 20 mM NaCl, 14 mM $MgCl_2$, 0.1 mM EDTA, 10 μg mL^{-1} BSA, 130

μM ApA, 1% glycerol, 0.04 U μL^{-1} *E. coli* RNA polymerase holoenzyme [Epicenter], 1.5 μM GTP, 2.5 μM UTP, and 2.5 μM ATP). Approximately 8 μCi [α-^{32}P]-GTP was added and transcription was allowed to proceed for 10 minutes at 37°C, leading to formation of a stalled complex seven nucleotides into the predicted *pfl* aptamer region. 8 μL aliquots of the resulting mixture were delivered to separate tubes containing 1 μL of 10x elongation buffer (1x initiation buffer plus 150 μM ATP, GTP, and CTP, 50 μM UTP, and 1 mg/mL heparin) and 1 μL of ligand at 10x the desired final concentration. Transcription elongation was allowed to proceed for 20 minutes at 37°C, at which time it was halted by the addition of 2x gel loading buffer (8 M urea, 20% (w/v) sucrose, 0.1% (w/v) SDS, 0.05% (w/v) bromophenol blue sodium salt, 0.05% (w/v) xylene cyanol, 90 mM Tris–HCl, 90 mM borate, and 1 mM EDTA [pH 8.0 at 25°C]). Transcription products were separated and visualized by PAGE as described above for in-line probing assays. The extent of elongation was determined as described previously (Kim et al., 2007).

Bacterial Strains and Growth Conditions

E. coli BW25113 strain and its isogenic derivative BW25113 *ΔpurH::kan* were obtained from the Coli Genetic Stock Center (Yale University). Reporter vector pRS414 was a gift from W.W. Simons (UCLA). *Bacillus subtils 1A1* (*trpC2*), *1A293* (*guaB3 metC7 purH1 trpC2*), and the *B. subtilis* integration vector pDG1661 were obtained from the Bacillus Genetic Stock Center (The Ohio State University). Cells were grown in EZ defined rich medium (Teknova) or M9 defined minimal medium supplemented with 0.2% glucose and 20 mM uracil unless otherwise noted. When required, growth medium was supplemented with antibiotics at the following concentrations: carbenicillin, 100 μg mL^{-1}; kanamycin, 50 μg mL^{-1}; chloramphenicol, 5 μg mL^{-1}. Growth curves for *E. coli* WT and *ΔpurH::kan* strains were generated using a Bioscreen C instrument (Oy Growth Curves Ab Ltd, Finland).

Design of Reporter Gene Constructs

Sequences of DNA primers used for cloning are included in **Table S1**. For the *E. coli* plasmid reporter construct, a DNA fragment containing the *E. coli lysC* promoter and the region encompassing the ZTP riboswitch from *Pectobacterium carotovorum subsp. carotovorum PC1* (nucleotides -106 to +24 relative to the *rhtB* translation start site*)* was prepared by PCR using

synthetic DNA primers. This fragment was cloned into the translational reporter vector pRS414, wherein the 8th codon of the ORF was fused in-frame to the 9th codon of the *lacZ* reporter gene. The resulting plasmid was transformed into *E. coli* BW25113 or BW25113 *ΔpurH::kan* strains. DNA fragments for riboswitch mutants were also prepared using synthetic primers.

The reporter construct for *B. subtilis* was prepared as described previously (Baker et al., 2012). Briefly, nucleotides –211 to +1 relative to the *folD* translation start site was amplified by PCR from *Bacillus sp. SG-1* genomic DNA. Subsequently, the *B. subtilis lysC* promoter was introduced via PCR amplification. This fragment was cloned into the vector pDG1661 as a transcriptional fusion with a *lacZ* reporter gene. The resulting plasmid was integrated into the *amyE* locus of *B. subtilis* 1A1 or 1A293.

Liquid-based β-galactosidase Assays

Liquid-based β-galactosidase assays were performed as previously described (Nelson et al., 2013; Vidal-Aroca et al., 2006). For β-galactosidase assays in liquid medium, a single colony of *E. coli* reporter strain was picked and grown overnight in either EZ or M9 medium with carbenicillin. The next day, cells were diluted 1/200th and subjected to various growth conditions (i.e., trimethoprim addition). The resulting cultures were incubated for ~20 hours at 37ºC. 80 µL of the resulting cultures were transferred to Costar black 96-well clear-bottom plates and the absorbance at 595 nm was measured using a Tecan Synergy 2 plate reader. Subsequently, 80 µL of Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$) was added to each well and 4-methylumbelliferyl-β-D-galactopyranoside (40 µl of a 1mg ml^{-1} solution) was added to each well. Plates were incubated at room temperature for 15 min, and the reaction was halted using 40 µL of 1 M Na_2CO_3. Excitation and emission was measured at 360/460 nm using a Tecan Synergy 2 plate reader. Fluorescence units were calculated as previously described (Vidal-Aroca et al., 2006).

RNA Homology Searches

Sequences from the bacterial and archaeal subsets of version 56 of the RefSeq database (Pruitt et al., 2007) and various environmental sequence databases (IMG/M (Markowitz et al., 2008), CAMERA (Sun et al., 2011), MG-RAST (Meyer et al., 2008), GenBank, and assembled Human Microbiome Project sequences (Human Microbiome Project, 2012)) were used to conduct

homology search for *pfl* motif RNAs, as previously described (Baker et al., 2012). We identified about 2000 distinct sequences of *pfl* motif in this manner. This number excludes redundant sequences (included in **Supplemental Information**), which arise mainly in metagenome and environmental sequence datasets. The RNA consensus diagram was drawn using R2R (Weinberg and Breaker, 2011). Protein-coding genes were predicted as described previously (Roth et al., 2014). Conserved protein domains were predicted using the Conserved Domain Database version 2.25 as described previously (Baker et al., 2012).

SUPPLEMENTAL REFERENCES

Baker, J.L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R.B., and Breaker, R.R. (2012). Widespread genetic switches and toxicity resistance proteins for fluoride. Science *335*, 233-235.

Human Microbiome Project, C. (2012). A framework for human microbiome research. Nature *486*, 215-221.

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.

Markowitz, V.M., Ivanova, N.N., Szeto, E., Palaniappan, K., Chu, K., Dalevi, D., Chen, I.M., Grechkin, Y., Dubchak, I., Anderson, I.*, et al.* (2008). IMG/M: a data management and analysis system for metagenomes. Nucleic Acids Res *36*, D534-538.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A.*, et al.* (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics *9*, 386.

Nelson, J.W., Sudarsan, N., Furukawa, K., Weinberg, Z., Wang, J.X., and Breaker, R.R. (2013). Riboswitches in eubacteria sense the second messenger c-di-AMP. Nat Chem Biol.

Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res *35*, D61-65.

Roth, A., Weinberg, Z., Chen, A.G., Kim, P.B., Ames, T.D., and Breaker, R.R. (2014). A widespread self-cleaving ribozyme class is revealed by bioinformatics. Nat Chem Biol *10*, 56-60.

Sun, S., Chen, J., Li, W., Altintas, I., Lin, A., Peltier, S., Stocks, K., Allen, E.E., Ellisman, M., Grethe, J.*, et al.* (2011). Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. Nucleic Acids Res *39*, D546-551.

Vidal-Aroca, F., Giannattasio, M., Brunelli, E., Vezzoli, A., Plevani, P., Muzi-Falconi, M., and Bertoni, G. (2006). One-step high-throughput assay for quantitative detection of beta-galactosidase activity in intact gram-negative bacteria, yeast, and mammalian cells. BioTechniques *40*, 433-434, 436, 438 passim.

Weinberg, Z., and Breaker, R.R. (2011). R2R--software to speed the depiction of aesthetic consensus RNA secondary structures. BMC Bioinformatics *12*, 3.