Supporting Information

Trausch et al. 10.1073/pnas.1312918111

SI Methods and Materials

Preparation of RNA. All RNAs used in this study were prepared by in vitro transcription by T7 RNA polymerase, using standard methods (1). Briefly, DNA transcription templates were made using synthetic DNA oligonucleotides (Integrated DNA Technologies) by PCR, using 1 μM outer 5′- and 3′-primers and 50 μM each dNTP. RNAs were synthesized in a transcription reaction containing 30 mM Tris·Cl at pH 8.0, 20 mM DTT, 0.1% (vol/vol) Triton X-100, 3.4 mM spermidine, 32 mM $MgCl_2$, 4 mM each ribonucleotide triphosphate, 0.25 mg/mL T7 RNA polymerase, and 50 nM DNA template, incubated at 37 °C for 2 h. The resultant RNA was ethanol precipitated and resuspended in loading buffer [2:1 (vol/vol) ratio of 8 M urea and 0.5 mM Na_2EDTA in ddH_2O and 100% formamide] and purified by denaturing PAGE [12% (wt/vol) 29:1 acrylamide:bisacrylamide]. The appropriate-length band, as visualized by ultraviolet shadowing, was excised and the RNA eluted into 0.5× T.E. buffer (10 mM Tris·Cl at pH 8.0, 1 mM EDTA) at 4 °C, using the crush-and-soak method. RNA was concentrated and buffer exchanged into T.E., using a 10,000 molecular weight cutoff centrifugal concentrator (Amicon). RNA concentrations were determined using calculated extinction coefficients and stored at −80 °C until use.

Crystallization of the "env87" SAM-RNA Complex. Sequences used for screening were taken from a phylogenetic alignment presented by Weinberg and colleagues (2). The *env87*(ΔU92) aptamer–S-adenosyl-L-methionine (SAM) complex was prepared for crystallization by making a solution of 300 μM RNA and 1 mM SAM. One microliter of this solution was mixed with 1 μL of a mother liquor containing 50 mM Na-cacodylate at pH 7.0, 250 mM ammonium acetate, 20 mM $MgCl_2$, and 16% (vol/vol) 2-methyl-2,4-pentanediol, and the complex was crystallized using the hanging drop crystallization method at 30 °C. Long hexagonal prisms formed in 2–3 d. For cryoprotection, 10 μL mother liquor was added to the drop. Crystals were scooped using a nylon loop and flash frozen in liquid nitrogen.

Diffraction data were collected on beamline 8.2.2 at the Advanced Light Source at a wavelength of 0.9999. Data were indexed, integrated, and scaled, using iMosflm (3). Molecular replacement was implemented using PHASER (4). The conserved binding pocket and surrounding residues (30% sequence composition) from the SAM-I riboswitch (2GIS) were used as the search model. This resulted in a single solution with a high translation function Z-score (~9). The molecular replacement generated a map that showed strong electron density for peripheral helices (Fig. S2*C*).

The model of the *env87*(ΔU92) SAM-I/SAM-IV complex was built using COOT with the RCRANE plug-in to assist in the building process (5, 6). The model was refined against the diffraction data extending to 3.2 Å resolution, using PHENIX (7). Iterative rounds of model building and simulated annealing/B-factor refinement were implemented. After the building of the RNA, metal ions were added to the model, and finally SAM, to produce the full model (Fig. S2*A*). Statistics on the data collection and refinement are shown in Table S4.

Isothermal Titration Calorimetry. RNA for ITC was transcribed and purified as described earlier, and calorimetric measurements were performed as described previously for other riboswitches (8). RNA samples were prepared for ITC by dialysis overnight at 4 °C in a solution containing 20 mM K-Hepes at pH 8.0, 135 mM KCl, 15 mM NaCl, and 10 mM $MgCl_2$. The RNA was then diluted with dialysis solution to the concentration used in the experiment (exact concentrations and c-values are shown in Table S3). Two hundred seventy microliters RNA was loaded into the cell. *S*-adenosylmethionine (Sigma Aldrich A2408) was brought up from a solid stock in the dialysis solution, and 40 μL was loaded into the syringe (exact concentrations are shown in Table S3). All titrations were done with 21 injections; the first injection contained a volume of 0.2 μL, the following 19 contained 2 μL, and the final injection contained 1.6 μL. The first injection was omitted from data processing. All injection durations were two times the volume in seconds and had a filter period setting of 5 s. All experiments were done on the MicroCal iTC_{200} system at a temperature of 37 °C, a reference power of 11, with initial delay of 60 s and the feedback mode set to high. Data were fit to a single-site binding model, using Origin ITC software (MicroCal Software Inc.). Examples of thermograms are shown in Fig. S4. All reported values represent the average of three independent experiments. Error values are reported as SD.

SHAPE Chemical Footprinting. RNA for chemical probing was constructed using 5′- and 3′-structure cassettes (9, 10). RNA was purified as described earlier. Chemical modification of the RNA was done at 100 mM K-Hepes at pH 8.0, 100 mM NaCl, 1 or 10 mM $MgCl_2$, and 0.1 μM RNA, all in 10 μL reaction volume. Ligand was added to select reactions to 1 mM. One microliter of the 10 μL reaction was DMSO (control) or 130 mM *N*-methylisotoic anhydride (NMIA), dissolved in DMSO. All probing was done at the indicated temperatures for a period consistent with five half-lives of NMIA at that temperature. Samples were reverse transcribed by adding 3 μL ^{32}P 5′-end-labeled DNA oligomer and incubated at 65 °C for 5 min, 35 °C for 10 min, and 54 °C for 1 min. Seven microliters of 167 mM Tris·Cl at pH 8.3, 250 mM KCl, 1.67 mM each dNTP, 16.7 mM DTT, and 0.33 units of SuperScript III reverse transcriptase (Invitrogen) were added to the reaction. Reactions containing 1 mM $MgCl_2$ during the probing reaction were supplemented to a concentration of 5 mM during reverse transcription. Sequencing reactions were doped with ddNTPs at a final concentration of 0.5 mM. Reactions were held at 54 °C for 12 additional minutes. Reverse transcription was terminated with the addition of 1 μL 4 M NaOH and incubation at 95 °C for 5 min. Twenty-nine microliters of acid stop mix (160 mM unbuffered Tris, 85% formamide, 42 mM EDTA) was added, and samples continued to incubate for an additional 5 min at 95 °C. Two microliters were loaded on a 10% denaturing polyacrylamide gel and ran at 55 W for 3.5 h. Gels were imaged using a Typhoon PhosphorImager (Molecular Dynamics).

***lacZ* Reporter Assay.** Riboswitch activity within *Escherichia coli* was monitored by use of the Miller assay (11). Briefly, the level of *lacZ* expression is monitored by its ability to enzymatically convert a colorless reactant to a colored product. The switch was cloned upstream of a *lacZ* reporter gene under control of a proD promoter (12). The DNA of the insert was constructed by PCR, using overlapping oligonucleotides. The sequences for all constructs are shown in Table S3. Inserts were cloned into a pBR327 plasmid, using NsiI and HindIII restriction sites. Sequence-verified plasmids were transformed into either Keio parental (BW25113) or Keio JW3909-1 (Δ*metJ725*) (13). All plates for Keio parental cells were made using carbenicillin (75 μg/mL), and all liquid cultures were supplemented with ampicillin (100 μg/mL). All plates for Keio JW3909-1 (Δ*metJ725*) cells were made using carbenicillin (75 μg/mL) and kanamycin (30 μg/mL), and all liquid cultures were supplemented with ampicillin (100 μg/mL) and kanamycin (30 μg/mL).

Three individual colonies for each construct were picked and grown overnight at 37 °C in 5 mL CSB rich defined media (14). Five microliters of the saturating culture was used to inoculate 5 mL fresh CSB media. Cultures were grown to midlog at 37 °C for ∼6 h ($O.D._{600}$, 0.4–0.6) in the Keio parental cell line and ∼9 h ($O.D._{600}$, 0.4–0.6) in the Keio JW3909-1 (Δ*metJ725*) cell line. The cultures were then placed on ice for 45 min. The OD_{600} was then measured in a Greiner 96-well polystyrene plate (catalog number 655079) in a Tecan Infinite M200 Pro plate reader, using 300 μL culture. Then, 20 μL culture was added to 80 μL permeabilization solution (200 mM Na_2HPO_4, 20 mM KCl, 2 mM $MgSO_4$, 0.8 mg/mL hexadecyltrimethylammonium bromide, 0.4 mg/mL sodium deoxycholate, and 5.4 μg/mL β-mercaptoethanol). All solutions were then incubated at 30 °C for 30 min. To start the reaction, 600 μL substrate solution (60 mM Na_2HPO_4, 20 mM NaH_2PO_4, 10 mM KCl, 20 μg/mL hexadecyltrimethylammonium bromide, 10 μg/mL sodium deoxycholate, 2.7 μg/mL β-mercaptoethanol, and 1 mg/mL *ortho*-nitrophenyl-β-galactoside) was added to the tubes. The reaction was allowed to proceed for 30 min. The reaction was stopped by the addition of 700 μL of 1 M Na_2CO_3. Samples were then spun down in a microcentrifuge at 15,000 rpm for 5 min. The OD_{420} of 300 μL of sample taken from the top of the tube was measured using the plate and plate reader described earlier. The Miller Units were calculated using the following equation:

$$1\ \text{Miller unit} = \frac{OD_{420}}{OD_{600} \times 30\ \text{min} \times 20\ \mu L} \times 1{,}000.$$

In an effort to background subtract, data shown are Miller units minus the Miller units of pBR327 control vector. Each biological replicate was preformed in triplicate, and the error shown is the SD of these trials.

Bioinformatics Analysis. Seed alignments of RF00162, RF00634, and RF01725 families were downloaded from the Rfam database (6) and manually aligned. The unalignable regions were removed, along with gap-rich and hypervariable columns (gap fraction larger than 0.8, and the 10% highest-entropy positions). The final alignment was used to build the phylogenetic tree with FastTree2 (15), using the generalized time-reversible model of nucleotide evolution. The alignment is supplied in a separate FASTA nucleic acid file.

1. Edwards AL, Garst AD, Batey RT (2009) Determining structures of RNA aptamers and riboswitches by X-ray crystallography. *Methods Mol Biol* 535:135–163.
2. Weinberg Z, et al. (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol* 11(3):R31.
3. Leslie A, Powell H (2007) *Evolving Methods for Macromolecular Crystallography*, eds Read RJ, Sussman JL (Springer, Dordrecht).
4. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
5. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
6. Keating KS, Pyle AM (2010) Semiautomated model building for RNA crystallography using a directed rotameric approach. *Proc Natl Acad Sci USA* 107 (18):8177–8182.
7. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
8. Gilbert SD, Batey RT (2009) Monitoring RNA-ligand interactions using isothermal titration calorimetry. *Methods Mol Biol* 540:97–114.
9. Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc* 127(12):4223–4231.
10. Wilkinson KA, Merino EJ, Weeks KM (2006) Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): Quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc* 1(3):1610–1616.
11. Miller JH (1992) *A Short Course in Bacterial Genetics — A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).
12. Davis JH, Rubin AJ, Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* 39(3):1131–1141.
13. Baba T, et al. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* 2, 2006 0008.
14. Ceres P, Garst AD, Marcano-Velazquez JG, Batey RT (2013) Modularity of select riboswitch expression platforms enables facile engineering of novel genetic regulatory devices. *ACS Synth Biol* 2(8):463–472.
15. Price MN, Dehal PS, Arkin AP (2010) FastTree2: Approximately maximum likelihood trees for large alignments. *PLoS One* 5(3):e9490.

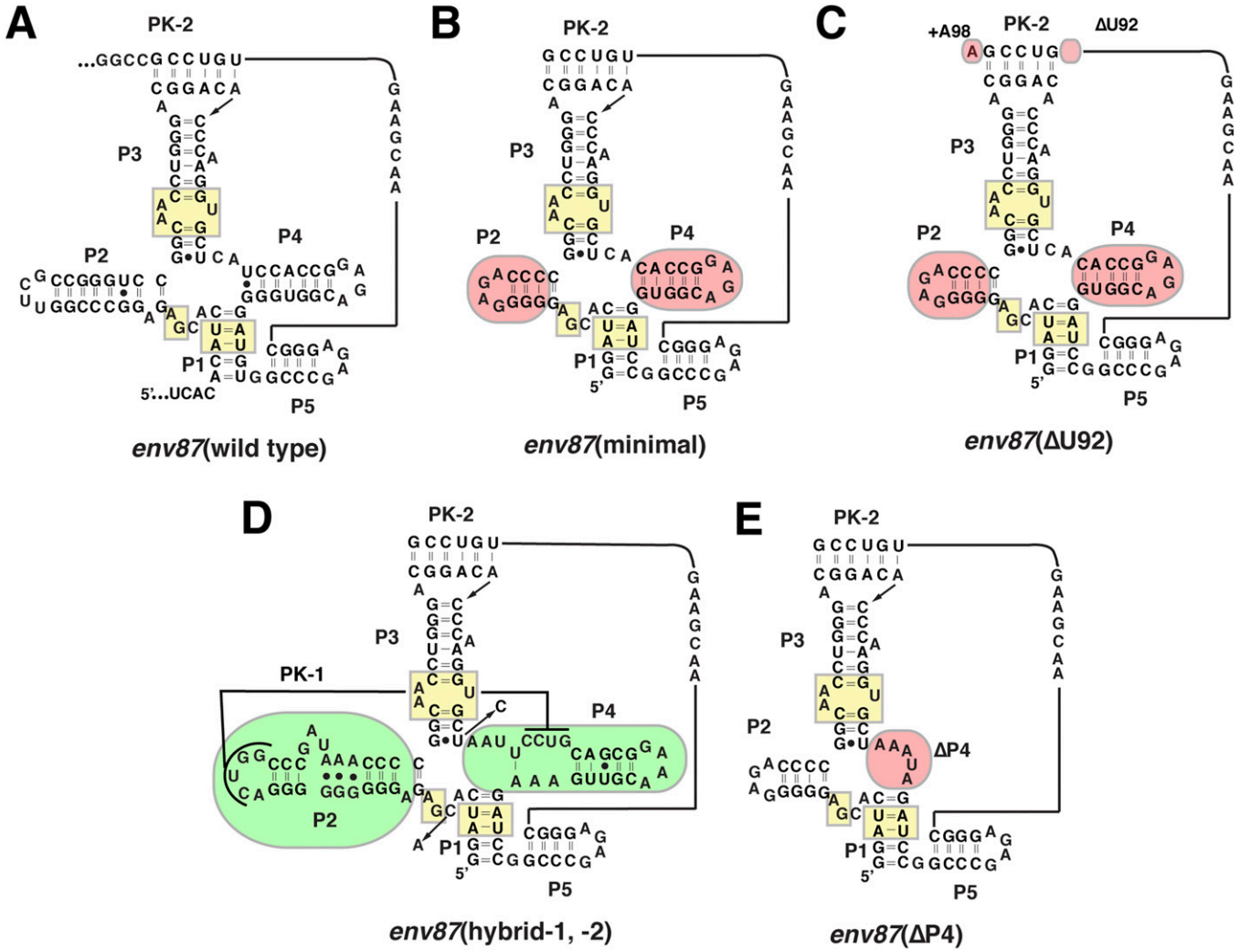

Fig. S1. Sequence and secondary structures of RNAs used in this study. In each secondary structure, the yellow box denotes nucleotides that directly contact SAM, and red and green ovals denote regions that have been altered relative to the wild-type sequence. (*A*) Predicted secondary structure of the wild-type *env87* aptamer sequence. (*B*) Secondary structure of the minimized env87 aptamer with truncations of the nonconserved P2 and P4 helices. (*C*) Secondary structure of the RNA used for determination of the crystal structure that incorporates the ΔU92 deletion and the A98 insertion. (*D*) Secondary structure of a hybrid aptamer containing the full pseudoknot 1 (PK-1) of *Thermoanearobacter tengcongensis* (*Tte*) SAM-I (green) appended into the *env87*SAM-I/IV aptamer. The "hybrid 2" sequence has two additional point substitutions (denoted by arrows) in the core that reflect the *Tte*SAM-I sequence. (*E*) Secondary structure of the env87(ΔP4) aptamer with the five-nucleotide linker (red) between P3 and P1.

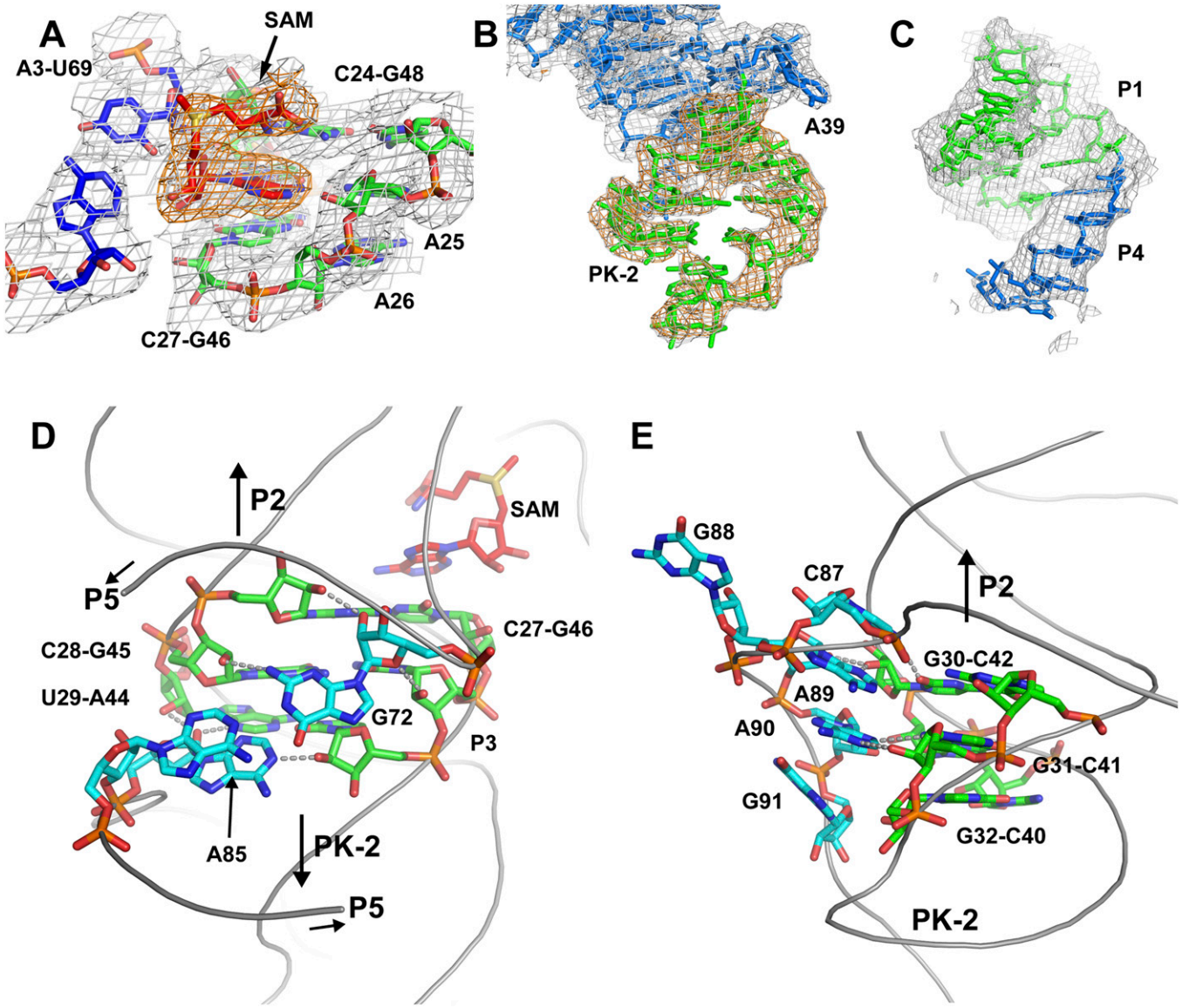

Fig. S2. (*A*) The binding pocket of SAM overlaid with the 2Fo-Fc simulated annealing omit map at 1.0 σ (gray) and the Fo-Fc simulated annealing omit map at 3.0 σ (orange). The color of the residues matches Fig. 2. (*B*) The PK-2 of SAM-I/IV overlade with the 2Fo-Fc simulated annealing omit map at 1.0 σ (gray) and the Fo-Fc simulated annealing omit map at 3.0 σ (orange). The residues colored green (33–38, 92–97) were removed to generate the omit map; the rest of the aptamer is colored in blue. (*C*) Representative 2Fo-Fc map at 1 σ of electron density of the P4 domain that was not used in the search model for molecular replacement. Residues in green represent the part of the SAM-I core (taken from 2GIS) used as the search model from molecular replacement. Residues in blue (3′ side of P4) represent some of the bases that were built in, using this initial electron density. Architecture of the P5/PK-2 subdomain. (*D*) Interactions of the universally conserved first G–A pair of P5 and first adenine residue of J5/PK-2 with P3. Extensive hydrogen binding between G72 and A86 with the minor groove of P3 (gray dashed lines) enables the perpendicular packing of P5. (*E*) Interactions between nucleotides in J5/PK-2 and the minor groove of P3. These interactions are dominated by type I and type II A-minor triplet interactions mediated by A89 and A90.

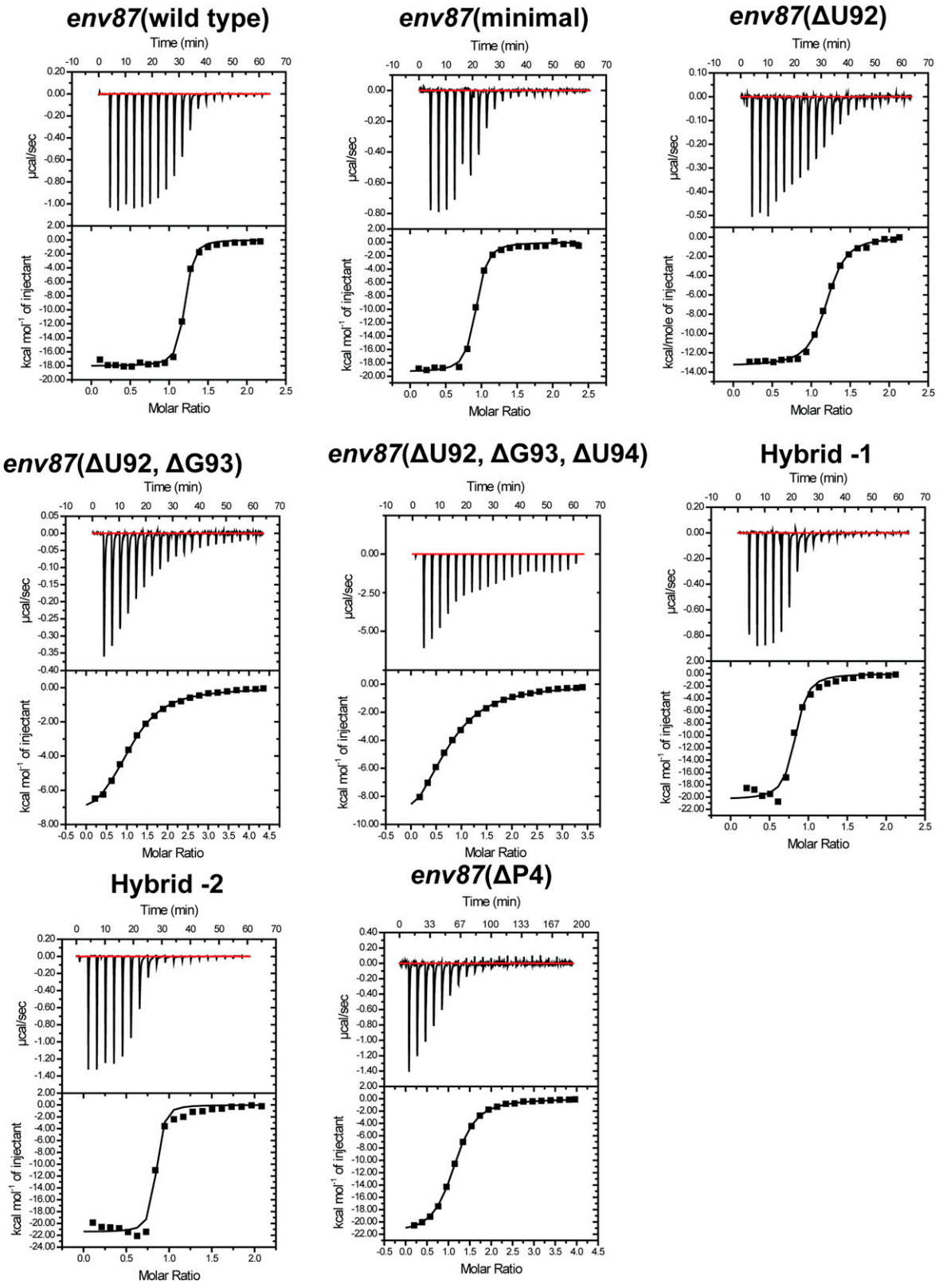

Fig. S3. Example thermogram and binding isotherm graphs for each unique RNA/condition reported by isothermal titration calorimetry. The red line represents the baseline used for the integration. Data from hybrid 1 and hybrid 2 contain residuals that may suggest an alterative binding model. Attempts to fit the data to a two, non-identical-site binding model resulted in a convergence back to a single site.

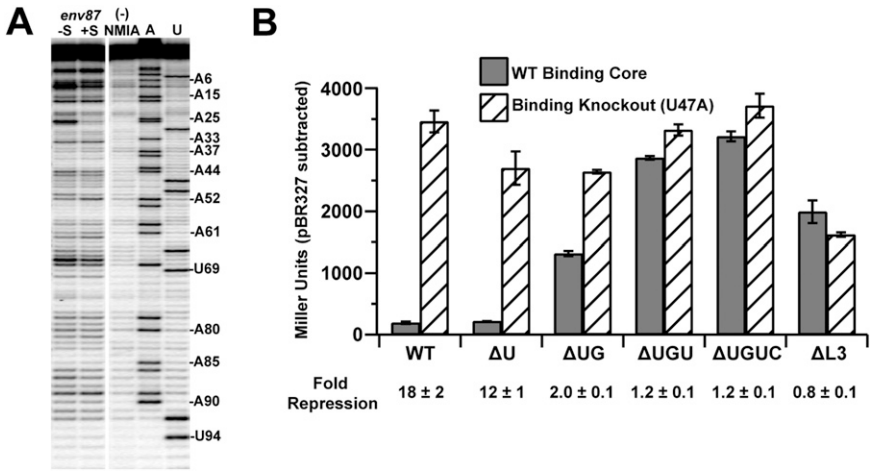

Fig. S4. (*A*) Raw data of selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) structure probing. Sequencing gel of chemical probing of the *env87*(minimal) RNA. −S and the +S correspond to the absence and presence of 100 μM SAM, respectively. The −NMIA lane is a control for RNA degradation, as well as intrinsic reverse transcription stops. A and U are sequencing reactions. (*B*) In vivo *lacZ* reporter assay in Δ*metJ* Keio cell line. Gray boxes represent the wild-type binding core of SAM-I/IV, and hashed boxes represent the U47A binding knockout. ΔU represents ΔU92. ΔUG represents ΔU92,ΔG93. ΔUGU represents ΔU92,ΔG93 and ΔU94. ΔUGUC represents ΔU92, ΔG93, ΔU94, and ΔC95. ΔL3 represents capping of L3 with a UUCG tetraloop. Fold repression is shown below the graph. All errors are the SD of three individual biological replicates.

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)
Table S4 (DOCX)
Dataset S1 (FNA)