

# Supporting Information

Grigg et al. 10.1073/pnas.1222214110

## SI Materials and Methods

**Cloning and Transcription.** Template DNAs for generation of T box RNA or tRNA were assembled from nested PCR reactions and inserted between the EcoRI and SmaI sites in a modified pUC19 vector, preceded by the T7 RNA polymerase promoter and followed by the hepatitis D virus ribozyme (1). T box constructs were designed from the leader of *Bacillus subtilis glyQ* (BSU25270) and *Geobacillus kaustophilus glyQ* (GK3430) and tRNA constructs were designed from coding regions for glycine tRNA (tRNA<sup>Gly</sup>) (GKT020) and isoleucine tRNA (tRNA<sup>Ile</sup>) (Tfu\_R0001) engineered to include a 3' CCA tail. Plasmids were purified using Qiagen Mega Prep kits (for crystallization templates) or Invitrogen Maxi Prep kits (templates for all biochemical assays) (1). Purified plasmid DNA was linearized with BamHI or HindIII, located 3' to the hepatitis D virus ribozyme, and used for in vitro transcription with T7 RNA polymerase in 1 mL (biochemical assays) or 10 mL (crystallization) reactions, following previously established protocols (1). RNA was separated by denaturing polyacrylamide gel electrophoresis (PAGE) and visualized by UV-shading (1). Target RNA bands were excised, and RNA was passively eluted by dialysis into 5 mM Hepes, pH 7.0. RNAs were then buffer exchanged by repeated passes using Millipore centrifugation columns (3K or 10K cutoff). After the final exchange, RNA samples were diluted to 5  $\mu$ M in 5 mM Hepes, pH 7.0, 50 mM NaCl. RNA was refolded by heat denaturation at 92 °C for 3 min, followed by cooling to 65 °C over 5 min and addition of 10 mM MgCl<sub>2</sub>. At this point, stem I<sub>57</sub> for crystallization was incubated for one additional minute before rapid cooling on ice. The larger RNA constructs were slow cooled to 30 °C over ~10 min before cooling on ice. Samples were then concentrated using Millipore centrifugation columns (3K or 10K cutoff) and used directly or flash-frozen in liquid nitrogen and stored at -80 °C until needed. All experiments were performed in buffer A (10 mM Hepes, pH 7.0, 50 mM NaCl, 10 mM MgCl<sub>2</sub>) unless stated otherwise.

**Size Exclusion Chromatography.** RNA samples for size exclusion chromatography were refolded in the presence of tRNA. Folding was carried out as described earlier, but prefolded tRNA was added at 60 °C during refolding. The refolded samples were separated using a Superdex 200 10/300 column (GE Healthcare) equilibrated in buffer A on a DuoFlow FPLC system (BioRad). A total of 250  $\mu$ L of the T box RNA-tRNA samples (~5  $\mu$ M) was applied to the column at a flow rate of 0.5 mL/min. Fractions were collected and analyzed by 8% denaturing PAGE in 0.5 $\times$  Tris-borate-EDTA buffer to verify the components of the peak. A standard curve was generated to approximate the eluted complex mass by using elution times for individual T box and tRNA constructs of known mass.

**Dynamic Light Scattering.** T box RNA (5  $\mu$ M) and tRNA (5  $\mu$ M) samples were prepared in buffer A as described for size exclusion chromatography. Measurements were performed in a Protein Solutions DynaPro-MSTC instrument. All measurements were performed at 20 °C in a 20- $\mu$ L cuvette. A calibration curve was determined by plotting hydrodynamic radius ( $R_H$ ) against RNA molecular weight standards (various stem I truncations and tRNA samples individually). The data were fit by the equation  $R_H = (a \cdot \text{molecular weight})^b$  to determine the constants  $a$  and  $b$ . These values were used directly in Dynals software to calculate the molecular weight of the complex based on the average measured  $R_H$  from 10 readings.

**Electrophoretic Mobility Shift Assay.** T box RNA samples were diluted to 0.5  $\mu$ M in buffer A and incubated for 1 h at room temperature with varied concentrations of tRNA, as specified in *Results*. Samples were then stored on ice until further use. Samples were mixed with 10% glycerol (vol/vol), and 10  $\mu$ L of the reaction was separated by 8% native PAGE (0.5 $\times$  Tris-borate buffer, 10 mM MgCl<sub>2</sub>). Separation was performed at 4 °C at 10 W (~350 V) with run times of ~4 h. Gels were stained with ethidium bromide and band intensity was quantified by using ImageJ (2). All titrations were performed in experimental triplicate.

**Crystallization and Data Collection.** RNA was concentrated to ~5 mg/mL in buffer A, and initial crystallization leads were determined from the Natrix HT screen (Hampton Research) in sitting drop vapor diffusion plates with 1  $\mu$ L RNA solution and 1  $\mu$ L crystallization solution. Crystals were optimized by hanging drop vapor diffusion at 18 °C in 2- to 4- $\mu$ L drops with equal parts of RNA and reservoir solution, containing 80 mM NaCl, 12 mM KCl, 20 mM MgCl<sub>2</sub>, 40 mM sodium cacodylate, pH 6.0, 12 mM spermine, and 26% to 32% (vol/vol) 2-methyl-1,3 propanediol. A microcrystalline precipitate first formed, then resolved within 2 wk to form thin rectangular plates that continued to grow for approximately 3 wk. Crystals were briefly washed in reservoir solution supplemented with 35% (vol/vol) 2-methyl-1,3 propanediol before flash-cooling to 100 K directly on the beam-line cryosystem. Osmium hexamine soaked crystals were transferred to a crystal washing solution supplemented with 1 mM osmium hexamine for 4 to 6 h before flash cooling. Data for native and osmium-soaked crystals were collected on the Cornell High Energy Synchrotron Source beam line A1 at 0.9767 Å wavelength.

**Data Processing and Structure Solution.** Data were processed and scaled by using HKL2000 (3) in the space group  $C2$ . The thin osmium-soaked crystals decayed during collection, so data from three crystals were scaled together to achieve high redundancy. Initial osmium sites were located by single-wavelength anomalous dispersion by using ShelxCD (4, 5) in the HKL2Map interface (6). The initial sites were used in AutoSol (7) from the Phenix suite (8) to refine the sites and perform density modification. Data were used to 3.6-Å resolution with an initial figure of merit of 0.46. Helices could be identified in the density and the AG bulge and distal loop could be roughly traced manually by using ideal RNA structures in Coot (9), although poor map quality limited conclusive loop building and base placement. At this point, ER-RASER (10, 11) from Rosetta (12) was used to minimize and rebuild the structure by using initial experimental maps. Following further density modification and phase extension using Resolve, the remainder of the structure could be unambiguously modeled in the density and was refined using Phenix.Refine (8). The final solution contains two molecules in the asymmetric unit that superimpose with an rmsd of 1.3 Å over all atoms. All analysis was performed by using molecule A.

**Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension.** All selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) experiments were performed as previously described (13, 14), with minor modifications. Stem I<sub>101</sub>, stem I<sub>86</sub>, and tRNA were modified to include a 5' hairpin (5'-GGCCUUCGGGC-CAA) and two 3' hairpins with a reverse transcriptase primer-binding site embedded (5'-UCGAUCCGGUUCGCCGGAUC-CAAUCGGGCUUCGGUCCGGUUC). Folded RNA samples (1  $\mu$ M) were exchanged into SHAPE buffer: 100 mM Hepes, pH

8.0, 50 mM NaCl, 10 mM MgCl<sub>2</sub>. Samples were mixed with fivefold excess of binding partner or the same volume of SHAPE buffer and incubated at room temperature for ~1 h. Sample (8 μL) was then combined with 2 μL of 100 mM 1-methyl-7-nitroisatoic anhydride and incubated at 37 °C for ~2 min. The remaining SHAPE procedure was performed as previously described (13). Reverse transcription was performed using a hexachlorofluorescein-labeled cDNA primer (5'-GAACCGGACCGAAGCCCG). Samples were analyzed by fragment analysis using an Applied Biosciences 3730xl and analyzed using ShapeFinder software (15). Integrated data from three separate readings were scaled by using the 2% to 8% rule (16), averaged, and normalized to 0 to 1 by using GraphPad Prism software. The relative reactivity is calculated by subtracting the reactivity of the SHAPE construct alone from the reactivity of the SHAPE construct in the presence of its binding partner. Numbering is shifted by one nucleotide to account for reverse transcriptase blockage.

**UV Cross-Linking.** Cross-linking experiments were performed by using stem I<sub>101</sub>, stem I<sub>86</sub>, and tRNA SHAPE constructs. The SHAPE construct (stem I<sub>SHP101</sub>, stem I<sub>SHP86</sub>, or tRNA<sub>SHP</sub>; 2 μM) was incubated in the presence or absence of its binding partner (stem I or tRNA; 4 μM) at room temperature for ~1 h. Samples were then placed 8 cm from a handheld UV lamp (254 nm, UVGL-58 Mineralight Lamp). Cross-linking was initially detected as a band shift in denaturing PAGE with SYBR Gold staining. Five minutes of exposure was optimal to observe a prominent slow-migrating band, while minimizing background. Samples were split into two 20-μL fractions and separated on opposite sides of a single denaturing PAGE gel. Half the gel was stained with SYBR Gold to identify cross-linked band mobility. Samples were excised from the unstained half of the gel, crushed, and RNA was eluted by overnight exchange into 500 μL of 0.5× Tris-EDTA buffer. Solutions were adjusted to 0.3 M NaCl and precipitated by the addition of 2 μL GlycoBlue (15 mg/mL; Invitrogen) and 1 mL of cold (-20 °C) 100% ethanol followed by incubation at -80 °C for 30 min. RNA was pelleted by centrifugation at 4 °C and 16,000 × g for 30 min. Supernatants were removed and pellets were air dried for 20 min before resuspension in 10 μL of 0.5× Tris-EDTA buffer. RNA was then reverse-transcribed using the same HEX-labeled primer and procedure described for SHAPE analysis earlier. Reactions were separated on a 10% sequencing gel and visualized by using a Typhoon 9400 (GE Healthcare).

**Small-Angle X-Ray Scattering Data Collection and Processing.** Refolded solutions of stem I<sub>86</sub>, tRNA, or the complex were prepared as described earlier in buffer A with 15 mM MgCl<sub>2</sub>, and the final concentrations used for small-angle X-ray scattering (SAXS) data collection and analysis are 0.44, 0.40, and 0.93 mg/mL, respectively. All SAXS data were collected at the Cornell High Energy Synchrotron Source G1 beam line. The X-ray energy is 10.69 keV, and the sample-detector distance is 1 m. For each sample, 60 images with 1-s exposure time were recorded by a PILATUS 100K detector. An oscillating capillary sample holder was used to reduce radiation damage to the RNA molecules. The 1.5-mm capillary

requires only a small amount of sample (~30 μL). The amplitude and speed of oscillation are monitored by a video camera and can be adjusted remotely during experiments to keep the sample in the beam. Buffer backgrounds were measured before and after each sample to ensure the capillary remained clean. A silver-behenate standard was used to calibrate the scattering angles. A small amount of beam transmitted through a molybdenum beam stop provided intensity normalization and the position of the beam center. All mathematical manipulations of the data (azimuthal integration, normalization, averaging, and buffer subtraction) as well as error propagation were carried out by using MatLab programs (Mathworks).

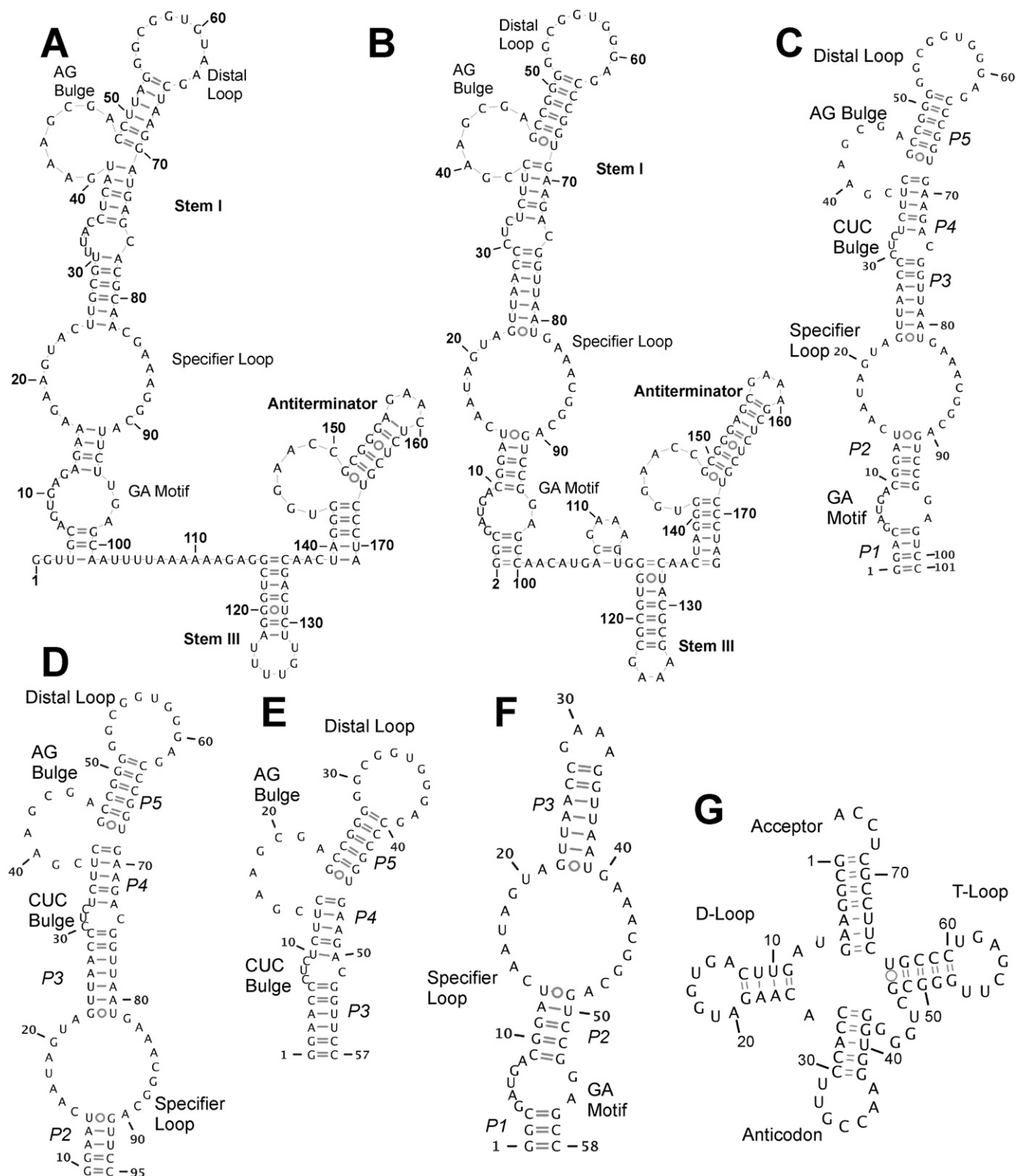
**Shape Reconstruction.** Because of the good agreement between the experimental and theoretical scattering profiles of the tRNA, stem I<sub>86</sub>, and the complex,  $D_{max}$  was estimated from the dimensions of the ribosome-bound *Escherichia coli* tRNA<sup>Phe</sup> structure [Protein Data Bank (PDB) ID 2J00 (17)] and the T box stem I<sub>86</sub> and complex models. The pair distance distribution function  $P(r)$  for each molecule was calculated by using the GNOM program. The GNOM output files were used as input to DAMMIF to perform ab initio shape reconstruction. Twenty reconstructed bead models for each molecule were superimposed and averaged by using DAMAVER in the automatic mode. The average normalized spatial discrepancies (NSDs) are listed in Fig. S7F.

Three different search volumes (Fig. S8 A–C) were tested in MONSA: a sphere (diameter, 120 Å); an oblate spheroid (major axis, 60 Å; minor axis, 30 Å); and a probability map computed by DAMAVER from the DAMMIF complex reconstruction. (More specifically, we modified the damstart.pdb file to allow two phases for each dummy atom by replacing “1 2 201” or “0 2 201” at the end of each atom line with “0 2 3012.”) The oblate shape is suggested by the  $P(r)$  of the complex. After the PDB file of the search volume and the raw SAXS data of the complex and its two components are input, MONSA generates 10 two-phase bead models. The models of phase 1 (tRNA), phase 2 (stem I<sub>86</sub>), and the combined phase (the complex) were averaged by using DAMAVER, and Fig. S7F shows that the models calculated through searching the DAMSTART probability map yields the lowest NSDs, especially for the combined phase representing the stem I<sub>86</sub>-tRNA complex. This demonstrates that model dephasing and refinement are possible by using MONSA when the shape of the complex is known. Because these models shows high convergence numerically (Fig. S7F) and visually (Fig. S8 D–F), we decided to align the averaged model (damfilt.pdb files) of each phase to one reference model (the last model in Fig. S8 D–F is the reference model DAMAVER selected for the complex model alignment) by using the SUPCOMB program (18). Fig. S5F shows that the phase information is well preserved after phase averaging and alignment. The next convergent models were obtained through annealing from the oblate spheroid (Fig. S8 G–I). The crystal or model structures were docked into the low-resolution SAXS-reconstructed envelopes by using SUPCOMB (enantiomorphs enabled).

1. Ke A, Doudna JA (2004) Crystallization of RNA and RNA-protein complexes. *Methods* 34(3):408–414.
2. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.
3. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, ed Carter, Jr. CW (Academic, San Diego), Vol 276, pp 307–326.
4. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* 58(pt 10 pt 2):1772–1779.
5. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64(pt 1):112–122.
6. Pape T, Schneider TR (2004) HKL2MAP: A graphical user interface for phasing with SHELX programs. *J Appl Cryst* 37:843–844.
7. Terwilliger TC, et al. (2009) Decision-making in structure solution using Bayesian estimates of map quality: The PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* 65(pt 6):582–601.
8. 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.
9. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(pt 4):486–501.
10. Chou F, Sripakdeevong P, Dibrov SM, Hermann T, Das R (2013) Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat Methods* 10(1):74–76.
11. Sripakdeevong P, Kladwang W, Das R (2011) An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc Natl Acad Sci USA* 108(51):20573–20578.
12. Das R, Baker D (2008) Macromolecular modeling with rosetta. *Annu Rev Biochem* 77:363–382.
13. Lu C, et al. (2011) Variable sequences outside the SAM-binding core critically influence the conformational dynamics of the SAM-III/SMK box riboswitch. *J Mol Biol* 409(5):786–799.
14. 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.

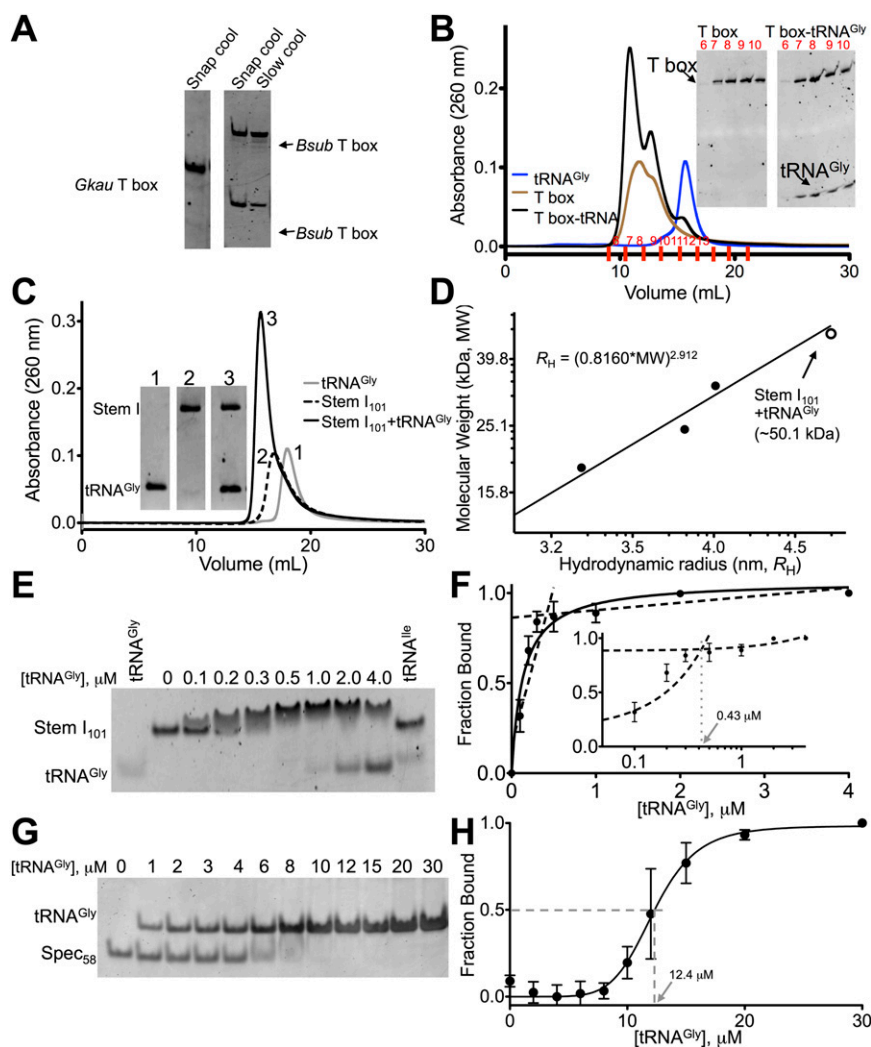
15. Vasa SM, Guex N, Wilkinson KA, Weeks KM, Giddings MC (2008) ShapeFinder: A software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA* 14(10):1979–1990.
16. Duncan CD, Weeks KM (2008) SHAPE analysis of long-range interactions reveals extensive and thermodynamically preferred misfolding in a fragile group I intron RNA. *Biochemistry* 47(33):8504–8513.

17. Selmer M, et al. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313(5795):1935–1942.
18. Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Cryst* 34:33–41.

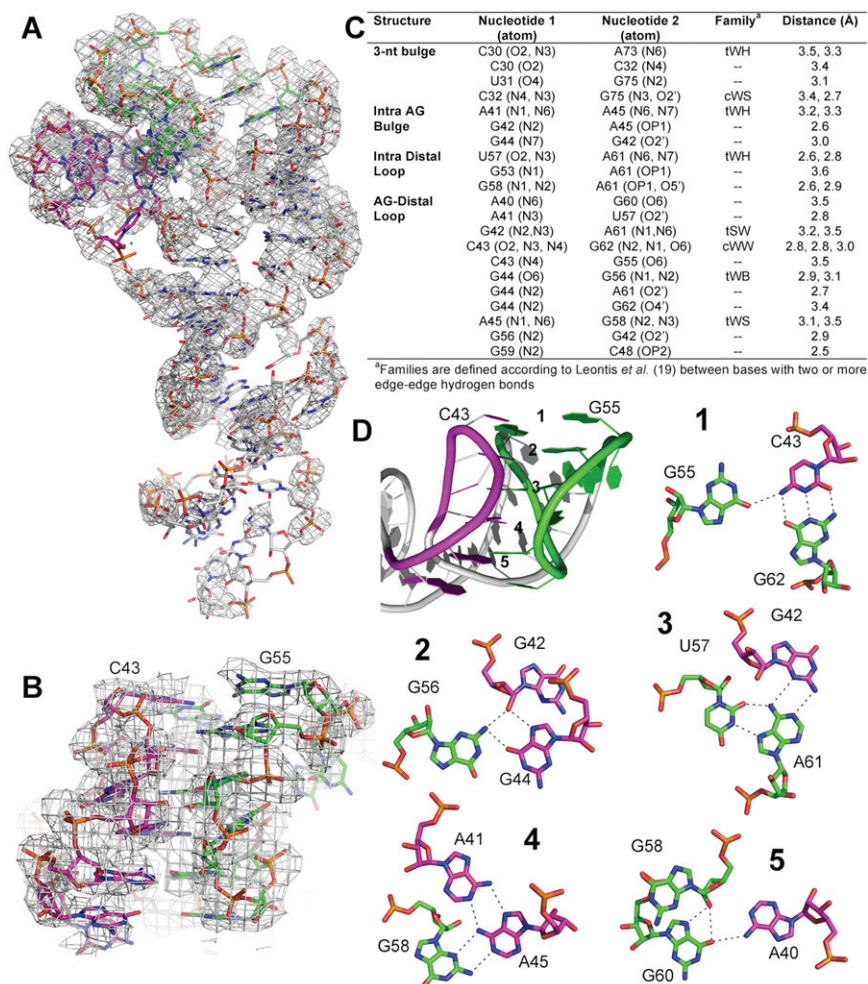


**Fig. S1.** Predicted secondary structure for T box RNAs and tRNA used in this study. (A) *Bacillus subtilis* (*Bsub*) full-length, (B) *Geobacillus kaustophilus* (*Gkau*) full-length, (C) *Gkau* stem I<sub>101</sub>, (D) *Gkau* stem I<sub>86</sub>, (E) *Gkau* stem I<sub>57</sub>, (F) *Gkau* spec<sub>58</sub>, and (G) *Gkau* tRNA<sup>Gly</sup>.





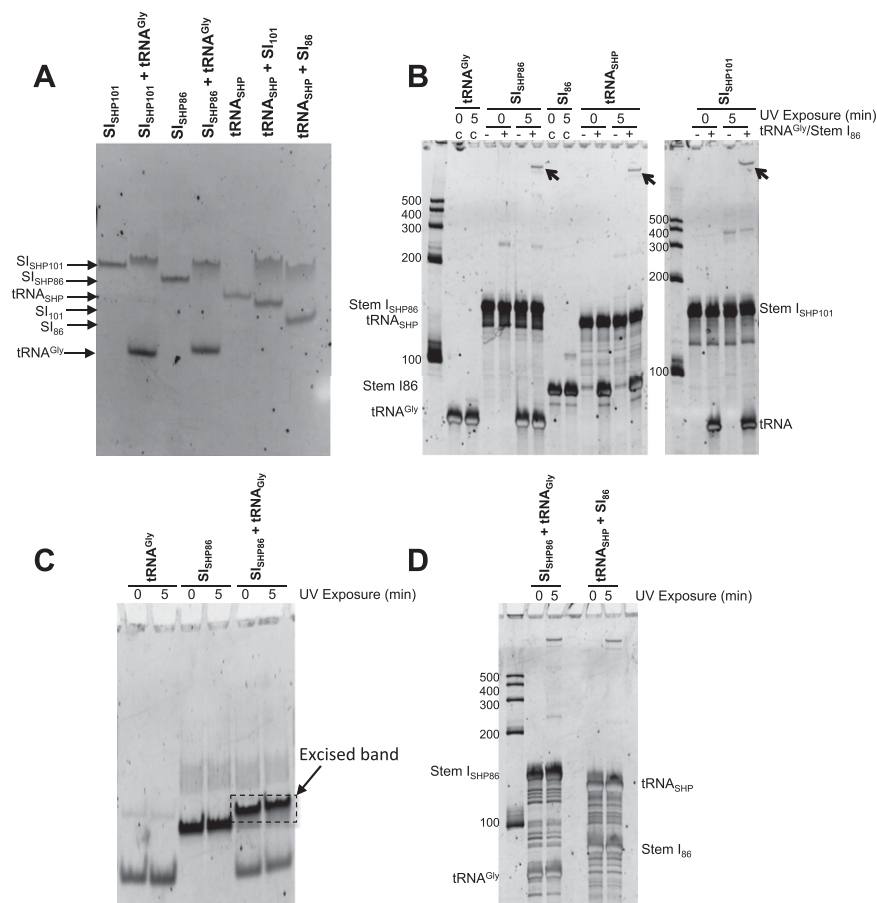
**Fig. S2.** Refolding and tRNA binding of in vitro transcribed T box RNAs. (A) Native PAGE of full-length *Gkau* and *Bsub* T box RNAs. (B) Superdex 200 30/100 elution profiles for *Bsub* T box RNA without tRNA<sup>Gly</sup> (brown) and with tRNA<sup>Gly</sup> (black) and tRNA<sup>Gly</sup> alone (blue). (Inset) Denaturing PAGE images of the corresponding fractions. (C) Superdex 200 30/100 elution profile for stem I<sub>101</sub> with tRNA<sup>Gly</sup>. (Inset) Denaturing PAGE images of labeled peaks. Peak 3 corresponds to a molecular weight of ~57 kDa, as described in *SI Materials and Methods*. (D) Dynamic light scattering standard curve showing experimental data for stem I<sub>101</sub>-tRNA (open circle), corresponding to a mass of ~50 kDa. (E) EMSA for tRNA<sup>Gly</sup> binding by stem I<sub>101</sub> and (F) quantified band intensities on a linear (main) and logarithmic (Inset) x-axis. Data are roughly fit by a linear titration to saturation (dashed line, 0.43 μM). The fit of a single site-binding model with  $K_d < 0.25$  μM is shown as a black curve. The data indicate 1:1 binding and establish an upper bound for the  $K_d$  of 0.25 μM. (G) Native PAGE for spec<sub>58</sub> titration with tRNA<sup>Gly</sup> and (H) quantified spec<sub>58</sub> band intensities. Data reach half maximal saturation at ~12 μM, but are best modeled by multiple sites or nonspecific binding.



**Fig. 53.** The stem I<sub>57</sub> crystal structure. (A) Initial osmium hexamine experimental electron density maps after density modification (3.6 Å resolution, contoured at 1 $\sigma$ ) are shown in gray with the final stem I<sub>57</sub> model shown in sticks. The AG bulge and distal loop are highlighted in magenta and green, respectively. (B) Close-up of the distal stem I loop-loop interaction. (C) Nonhelical hydrogen bond distances in stem I<sub>57</sub>. (D) Distal loop-loop hydrogen bonds shown as dashed lines. Nucleotides are colored as in A and B.



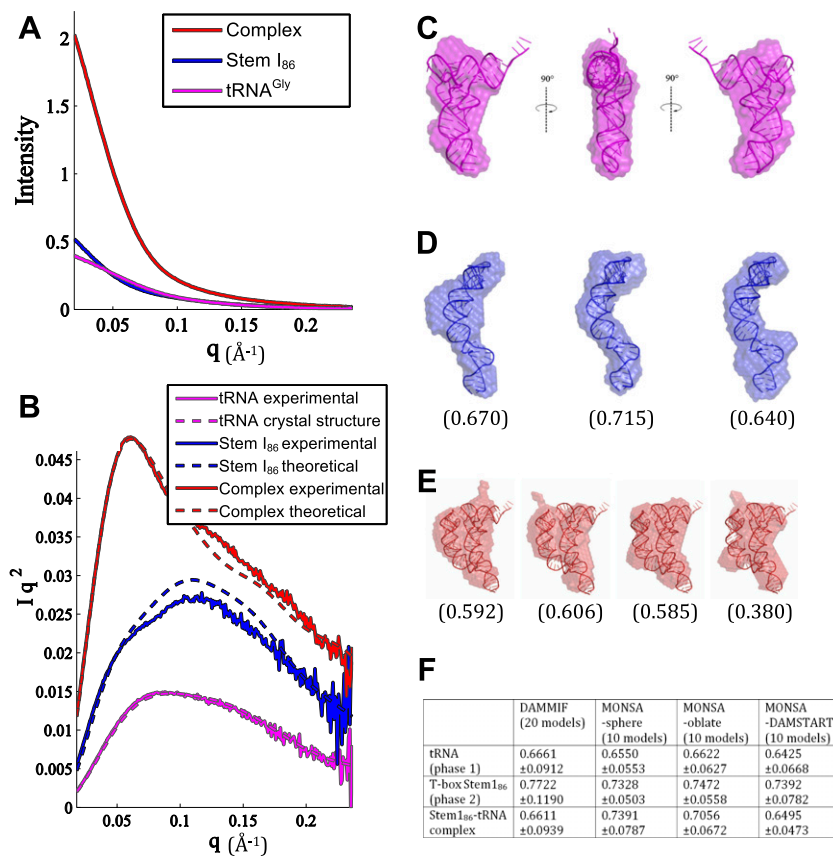




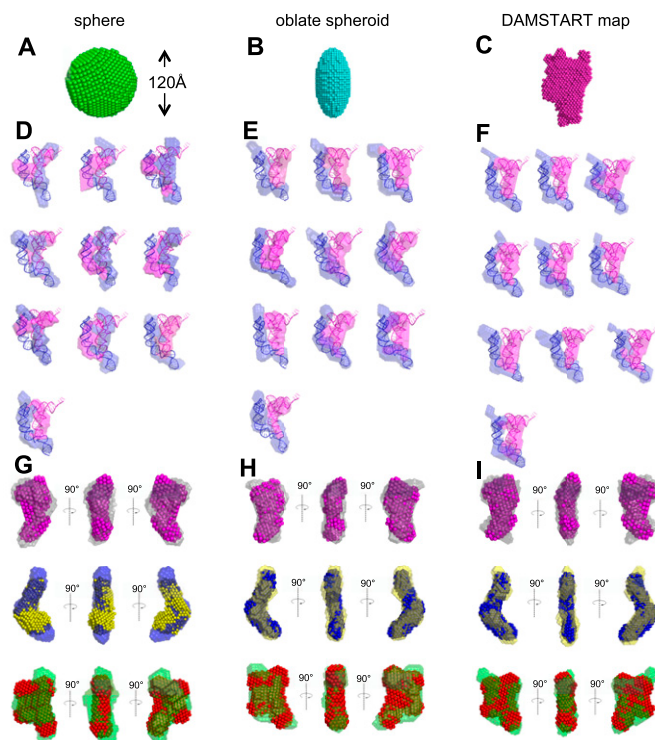
**Fig. 55.** PAGE of the cross-linked stem I-tRNA complex. (A) EMSA for complex formation between modified SHAPE constructs of stem I<sub>101</sub> (SI<sub>SHP101</sub>), stem I<sub>86</sub> (SI<sub>SHP86</sub>), and tRNA<sup>Gly</sup> (tRNA<sub>SHP</sub>) and their unmodified binding partner (tRNA<sup>Gly</sup>, stem I<sub>101</sub>, or stem I<sub>86</sub>). (B) Denaturing PAGE of cross-linked stem I and tRNA SHAPE constructs. Samples are shown without (0 min) or with UV exposure (5 min) in the absence (-) or presence (+) of their binding partner or the binding partner alone (C). Arrows indicate the prominent UV cross-linked sample formed only in the presence of stem I and tRNA that were excised for reverse transcription (Fig. 4E). (C) Representative native PAGE for samples after cross-linking, as shown in B. Bands corresponding to the 1:1 complex (dashed box) were visualized by UV shadowing and excised to be analyzed by denaturing PAGE. (D) Denaturing PAGE for the excised complex bands, as indicated in C. The slow-migrating species in denaturing PAGE comigrates with the non-cross-linked 1:1 complex by native PAGE. This suggests the retarded mobility product represents the cross-linked 1:1 complex rather than a higher-order oligomer. Ladders indicate linear RNA size standards.







**Fig. S7.** SAXS data and DAMMIF reconstruction. (A) Molar concentration normalized scattering profiles of tRNA<sup>Gly</sup>, stem I<sub>86</sub>, and the complex. (B) Comparison between the experimental and theoretical SAXS Kratky curves. Models for stem I<sub>86</sub> and the complex are described in *Results* and Fig. 5. *E. coli* tRNA<sup>Phe</sup> coordinates were extracted from the ribosome-bound crystal structure (PDB ID 2J00; P-site) (17). (C) Best docking of a tRNA structure into the averaged model of 20 SAXS reconstructions by using the program SUPCOMB. (D) and (E) are different groups of stem I<sub>86</sub> and stem I<sub>86</sub>-tRNA complex models clustered by the program DAMCLUS. Only the average of each group is shown. Average NSD values are shown in the parentheses. (F) Average NSD values of DAMMIF and MONSA reconstruction.



**Fig. S8.** Two-phase MONSA reconstruction. (A–C) Three different search volumes used for reconstruction in the program MONSA, including a sphere (diameter, 120 Å; A), an oblate spheroid (major axis, 60 Å; minor axis, 30 Å; B), and a probability map computed by DAMAVER from the complex alone reconstruction (C). (D–F) Ten two-phase complex bead models reconstructed by the program MONSA using the sphere (D), the oblate spheroid (E), and the DAMSTART model (F) as the search volume. Fig. S7F shows that the most and least convergent models are generated through searching the DAMSTART probability map and the sphere, respectively. (G–I) The overlay of the averaged models reconstructed by the programs DAMMIF and MONSA using the sphere (G), the oblate spheroid (H), and the DAMSTART model (I) as the search volume.

**Table S1. X-ray data collection and processing statistics**

Statistic	Osmium hexamine	Native
Data collection		
Wavelength, Å	0.9767	0.9767
Resolution range, Å	35.0–3.6 (3.66–3.60)	35.0–2.65 (2.70–2.65)
Space group	C2	C2
Unit cell dimensions		
<i>a</i> , Å	119.5	116.5
<i>b</i> , Å	41.5	41.5
<i>c</i> , Å	92.0	90.4
β, °	109.1	106.8
Unique reflections	5,440	11,578
Completeness, %	96.6 (77.8)	96.3 (77.1)
Average <i>I</i> /σ	19.1 (3.8)	15.5 (2.9)
Multiplicity	9.4 (6.0)	3.6 (3.1)
<i>R</i> <sub>merge</sub>	0.187 (0.412)	0.068 (0.484)
Wilson B, Å <sup>2</sup>	—	48.9
Refinement		
<i>R</i> <sub>work</sub> , <i>R</i> <sub>free</sub>	—	20.2 (24.1)
<i>B</i> -factors, Å <sup>2</sup>		
All atoms	—	71.1
RNA	—	71.4
Solvent	—	43.0
rmsd bond length, Å	—	0.0012
rmsd angles, °	—	0.419
PDB ID code	—	4JRC

Values in parentheses represent the highest-resolution shell. PDB, Protein Data Bank.