SUPPLEMENTARY INFORMATION

Protein unties the pseudoknot: S1-mediated unfolding of RNA higher order structure

Paul E. Lund[1,†], Surajit Chatterjee[1, †], May Daher[1,3] and Nils G. Walter[1,2,*]

[1]Single Molecule Analysis Group, Department of Chemistry, University of Michigan, Ann Arbor, Michigan, USA. [2]Center for RNA Biomedicine, University of Michigan, Ann Arbor, Michigan, USA. [3]Current address: Department of Chemistry, University of Detroit Mercy, Detroit, MI 4822, USA.

*To whom correspondence should be addressed. Tel: +1-734-615-2060; Email: nwalter@umich.edu

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Contents

SUPPLEMENTARY METHODS

Mutagenesis, expression, and purification of *E. coli* ribosomal protein S1

The parent pCA24N plasmid containing the rpsA gene, encoding the *E. coli* ribosomal protein S1, was prepared with an additional N-terminal His-tag from the ASKA(-) clone JW0894 (National BioResource Project – *E. coli* at National Institute of Genetics) (1).The 7 amino-acid linker sequence, TDPALRA, present in the parent plasmid between the 6×His-tag and the second native amino acid encoded by rpsA was changed to the recognition sequence (ENLYFQ^G) for TEV protease (2) using site-directed mutagenesis. Additionally, a second mutagenesis primer was designed to insert an additional stop codon at the C-terminus, removing additional amino acids not encoded in the native rpsA gene, but instead resulting from the original generalized cloning strategy. Mutagenesis primer sequences were designed with the aid of the QuikChange Primer Design online tool from Agilent (https://www.genomics.agilent.com/primerDesignProgram.jsp). Candidate primer sequences were further optimized using the OligoAnalyzer online tool v3.1 from Integrated DNA Technologies (IDT; http://www.idtdna.com/calc/analyzer) with the default settings (50 mM Na^+, 0 mM Mg^{2+}, 25 °C) to minimize the stability of self-dimers (typically predicted ΔG > -10 kcal/mol, or <10 – 20% of the ΔG of hybridization to a fully complementary sequence), and the propensity to form internal hairpins (typically T_m < 55 °C), ultimately yielding the mutagenesis primers 5′-gag caa aag att cag tGC CCT GAA AAT ACA GAT TCT Cat ggt gat ggt gat gg-3′ and 5′-aaa gca gct aaa ggc gag TAA cta tgc ggc cgc taa ggg-3′ (complementary flanking sequences are shown in lower case).

Mutagenesis conditions for site-directed mutagenesis were based on protocols provided by the lab of Prof. Roger Sunahara, adapted from Sawano and Miyawaki (3). Briefly, primers were purchased from Life Technologies or IDT and 5′ phosphorylated for 30 min at 37 °C at 6 μM final primer concentration in 1X NEBuffer 2.1 supplemented with 1 mM ATP, using T4 Polynucleotide Kinase (New England Biolabs, M0201), which was then heat-inactivated at 65 °C for 20 min. The mutagenesis reaction was performed in a final volume of 50 μL containing 1 nM of the parent plasmid DNA as the template, 0.25 μM of mutagenesis primer, 1 mM of each dNTP, 2.5 U of Pfu Ultra DNA polymerase (Agilent), 20 U of Taq DNA Ligase (New England Biolabs, M0208), 0.5X Pfu Ultra HF Buffer (Agilent), and 0.5X Taq DNA Ligase Buffer (New England Biolabs). The reaction conditions were: initial denaturation at 95 °C for 30 s,

followed by 20 cycles of denaturation (95 °C for 30 s), annealing (55 °C for 60 s), and extension (68 °C for 12.5 min), and then chilled to 4 °C. When the reaction was complete, 20 U of DpnI (New England Biolabs, R0176) were added to the reaction and incubated at 37 °C for 1 h, and then chilled to 4 °C. 25 µL of the JM109 strain of competent *E. coli* cells (Promega) were transformed with 1 µL of the reaction mixture following the transformation protocol provided in the QuikChange Multi Site-Directed Mutagenesis Kit manual (Agilent) using appropriately scaled volumes, and plated onto LB-agar plates containing 170 µg/mL chloramphenicol. Clones carrying the plasmids with the desired mutation(s) were identified by Sanger sequencing with the following sequencing primers: 5′-CAG GAA ACA GCT ATG ACC-3′, 5′-ATT CGT GCG TTC CTG CCA-3′, 5′-GTC TGA CAT CTC CTG GAA CG-3′, and 5′-CGA GCG TTC TGA ACA AAT CC-3′. The final plasmid (pCA24N_6xHis_TEV_rpsA) is available through Addgene (www.addgene.org).

pCA24N_6xHis-TEV_rpsA was expressed in the BLR(DE3) strain of *E. coli* using conditions loosely based on those described by Lancaster *et al.* (4). 1 L of LB-Miller broth containing 68 µg/mL chloramphenicol was inoculated 1:500 from a saturated overnight culture and grown with shaking at 37 °C, and induced with 1 mM IPTG at an OD_{600} ~0.6. The culture was harvested 2 h post-induction by centrifugation at 5,000 rpm for 15 min at 4 °C in a Beckman JLA 8.100 rotor. All subsequent steps were performed at 4 °C or on ice. The cell pellet was resuspended in 30 mL of buffer B (15 mM Tris-HCl [pH 7.05 at 25 °C], 30 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM β-mercaptoethanol, 0.1 mM PMSF), then pelleted by centrifugation at 6,800 × g for 10 min, and then stored at -80 °C for later use. The pellet was resuspended in 30 mL buffer B and lysed in two passes through an M-110L Microfluidizer processor (Microfluidics). The lysate was cleared by centrifugation at 10,400 × *g* for 45 min, and then combined with 5 mL of Ni-NTA Agarose resin (Qiagen, 30210) that has been pre-equilibrated in buffer B. The mixture was tumbled for ~2.5 h, then transferred to a disposable Econo-Pac column (Bio-Rad Laboratories, 9704652) and drained. The resin was washed with 25 mL of buffer C (15 mM Tris-HCl [pH 7.05 at 25 °C], 30 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM β-mercaptoethanol, 10 mM imidazole [pH 8.0]) with 500 mM NaCl to reduce the amount of co-purifying RNA, and then washed again with 25 mL of buffer C to remove excess Na^+. Bound protein was eluted from the resin in 4 fractions using 15 mL total of buffer D (15 mM Tris-HCl [pH 7.05 at 25 °C], 30 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM β-mercaptoethanol, 250 mM imidazole [pH 8.0]). Fractions containing

significant amounts of 6×His-TEV-S1 were pooled and the concentration of 6×His-TEV-S1 was estimated from the A_{280} of the solution using a Nanodrop2000 spectrophotometer and an estimated ε_{280} = 48,930 M^{-1} cm^{-1} (ExPASy ProtParam, Swiss Institute of Bioinformatics). Approximately 41 mg of 6×His-TEV-S1 in 7.5 mL of buffer D was thoroughly mixed with ~0.5 mg of TEV protease prepared in-house and transferred to 10,000 MWCO dialysis tubing and dialyzed overnight into buffer E (15 mM Tris-HCl [pH 7.05 at 25 °C], 5 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM β-mercaptoethanol). The dialyzed solution was combined with 5 mL of Ni-NTA Agarose resin (Qiagen) that has been pre-equilibrated in buffer E, and tumbled for ~3 h. The flow-through from this second nickel affinity column was directly loaded onto a 5 mL Q Sepharose Fast Flow anion exchange column (GE Healthcare, 17-0510-01), pre-equilibrated with buffer E. The column was washed with 15 mL of buffer E, and then eluted with increasing amounts of buffer F (15 mM Tris-HCl [pH 7.05 at 25 °C], 600 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM β-mercaptoethanol) in buffer E. A step-wise gradient was from 0 – 80% buffer F over 100 mL with a 10 mL step-size was used, and protein S1 eluted between ~30 – 50% buffer F. S1-containing fractions were pooled and concentrated using an Amicon Ultra-15 10,000 MWCO centrifugal filter unit (EMD Millipore) to a final volume of ~5 mL and transferred to 10,000 MWCO dialysis tubing, and dialyzed into storage buffer A(10) (25 mM Tris-HCl [pH 7.05 at 22 °C], 100 mM NH_4Cl, 10 mM $MgCl_2$, 10% [v/v] glycerol, 6 mM β-mercaptoethanol). The S1 concentration was measured from the solution A_{280} after dialysis using a Nanodrop2000 spectrophotometer and an estimated ε_{280} = 47,440 M^{-1} cm^{-1} (ExPASy ProtParam, Swiss Institute of Bioinformatics), then aliquoted, snap frozen with liquid nitrogen, and stored at -80 °C. This final protein solution had a measured 260/280 absorbance ratio of 0.74, suggesting the absence of any co-purifying nucleic acid. Protein aliquots were removed for use and thawed on ice; aliquots were kept for up to one week stored at -20 °C and then discarded.

Preparation of DNA templates for in vitro transcription

DNA oligonucleotides (5'-CCC TTG TTT TGT TAA CTG GGG TTA CTG CGA CCC AGG ACC TAT AGT GAG TCG TAT TAA ATT-3'; 5'-AAT TTA ATA CGA CTC ACT ATA GG-3') designed to give a partially double-stranded template for transcription of the wild-type *Tte* pseudoknot were gel purified prior to use. DNA oligonucleotides were purchased from IDT and 200 – 400 µg of oligonucleotide was electrophoresed on an 18 cm 20% Urea-PAGE gel. The oligonucleotide band was visualized by UV shadowing briefly

using a 312-nm lamp. The top half of the band was cut out from the gel, and extruded through a needle-less 3 mL syringe. The oligonucleotide was eluted from the gel pieces overnight at 4 °C in 3 mL of elution buffer (500 mM NH_4OAc, 0.1% [w/v] SDS and 0.1 mM EDTA). The eluted solution was extracted twice with an equal volume of chloroform saturated with TE (pH 6.5) and then the oligonucleotide was precipitated from the aqueous phase by adding 2 volumes of absolute ethanol, stored overnight at -20 °C, and collected by centrifugation at 12,800 × g for 45 min at 4 °C. The pelleted material was washed once with 1 mL of cold 70% (v/v) ethanol, dried under vacuum, and finally resuspended in milliQ water.. Extinction coefficients estimated using OligoCalc (5) were used to calculate the concentration of purified oligonucleotide from the A_{260}. Successful removal of truncated sequences was confirmed by electrophoresing a 50 ng sample on a 20% Urea-PAGE gel, followed by staining with a 1:10,000 dilution of SYBR Gold nucleic acid stain (Life Technologies) in 1X TBE, and visualizing on a UV transilluminator. The partially double-stranded duplex template was assembled using conditions similar to those described in the MEGAshortscript T7 transcription kit manual (Life Technologies): 10 µM each of the (+) strand and respective (-) strand were combined in a final volume of 50 µL with 10 mM Tris-HCl (pH 8.0 at 22 °C) and 100 mM NaCl, heated for 10 min a 90 °C copper bead bath, and then allowed to cool to room temperature over 15 min.

Site-directed mutagenesis of the pUC19_TTE1564 plasmid (Addgene ID 61000), which carries the wild-type pseudoknot sequence (6), was performed to generate the P2-deletion mutant. The mutagenesis reaction was performed according to the procedure described above, using the mutagenesis primer 5'-AAC AAA ATG CTC ACC TGG GTT CGC CCA GTT AAC AAA ACA AGG-3' and the following reaction conditions: initial denaturation at 95 °C for 30 s, followed by 20 cycles of denaturation (95 °C for 30 s), annealing (55 °C for 60 s), and extension (68 °C for 8 min). The final plasmid (pUC19_TTE1564_UUCG-loop) is available through Addgene.

The template for in vitro transcription of the P2-deletion mutant was prepared by PCR from the pUC19_TTE1564_UUCG-loop plasmid using PAGE-purified DNA primers 5'-TTT CCC AGT CAC GAC GTT-3' and 5'-GGG CAC AAA ATT ACC TC-3'. The PCR reaction was performed using 10 ng of the plasmid DNA, 0.5 µM of each DNA primer, 1 U Phusion High-Fidelity DNA polymerase (New England Biolabs, M0530), 1X Phusion HF Buffer, 200 µM of each dNTP in a 50 µL final volume. The PCR reaction

conditions were: initial denaturation at 98 °C for 10 s, followed by 30 cycles of denaturation (98 °C for 10 s), annealing (57 °C for 15 s), and extension (72 °C for 7 s). When the reaction was complete, 10 µL of 3 M NaOAc (pH 5.2) was added to each reaction and the PCR product was purified using the QIAquick PCR purification kit (Qiagen), and the concentration measured using a Nanodrop2000 spectrophotometer.

RNA preparation for EMSA and melting curve studies

RNA pseudoknots were generated by *in vitro* transcription using N-terminally His-tagged T7 RNA polymerase prepared in-house using a method adapted from that described by He *et al.* (7) with the following modifications: T7 was expressed in a different strain of BL21 *E. coli*, 250 mM NaCl was added to all buffers to reduce co-purification of other proteins, and batch-binding to the nickel resin was used in place of column loading. Reaction conditions for in vitro transcription were modeled after those described previously (8,9). The wild-type *Tte* riboswitch pseudoknot was transcribed from the partially double-stranded oligonucleotide template described above. Briefly, 150 µL transcription reactions containing 300 nM template, 120 mM HEPES-KOH (pH 7.6 at 22 °C), 0.01% (v/v) Triton X-100, 30 mM $MgCl_2$, 7.5 mM of each NTP, 40 mM DTT, 2 mM spermidine trihydrochloride, 0.2 mg/mL T7 RNA polymerase, and 0.01 U/µL Inorganic pyrophosphatase (MP Biomedicals) were incubated in a circulating water bath at 37 °C for 16-18 h, and then mixed with an equal volume of 2X gel loading buffer (95% (v/v) formamide, 18 mM EDTA, 0.025% [w/v] each of SDS, bromophenol blue, and xylene cyanol) to stop the reaction. The reaction with loading buffer was heated in a 90 °C copper bead bath for 3 min and then snap cooled on ice. The transcript was gel purified as described above on a 20% Urea-PAGE gel, and identified by UV-shadowing with a 312-nm lamp. The bottom half of the band was cut from the gel and extruded through a needle-less 3 mL syringe. The RNA was eluted from the gel, precipitated and resuspended in milliQ water as described above. Extinction coefficients estimated using OligoCalc were used to calculate the concentration of purified RNA from the A_{260}. The complete sequence of the wild-type pseudoknot used in this study, with the exception of smFRET experiments (see below), is 5'-GGU CCU GGG UCG CAG UAA CCC CAG UUA ACA AAA CAA GGG-3'.

In vitro transcription reactions to prepare the P2-deletion mutant construct was performed in a similar manner, except with a reaction volume of 60 µL, 150 nM PCR product as template, and an incubation time of 4.5 hr. The sequence of the P2-deletion mutant is 5'-GGG CAG UGA GCA ACA AAA UGC UCA CCU GGG uuc gCC CAG UUA ACA AAA CAA GGG AGG UAA UUU UGU GCC C-3', where lower case indicates the site of mutation.

3' Fluorophore labeling of RNA

RNAs prepared by transcription as described above were labeled with a Cy3 fluorophore at their 3′ end following a method described previously by Willkomm and Hartmann (10) with several modifications. Briefly, RNAs were first oxidized by incubating 5 µM RNA in 100 mM NaOAc (pH 5.2) with freshly prepared 2.5 mM sodium (meta) periodate (Fluka, 71859) on ice for 70 min, protected from light. Subsequently, the oxidized RNA was precipitated with the addition of 0.1 V of 3 M NaOAc (pH 5.2) and 2.5 V of cold absolute ethanol, followed by incubated on dry ice until frozen. The solution was inverted until just thawed and then centrifuged at 20,800 × g for 45 min at 4 °C to pellet the RNA. The supernatant was removed by pipetting, and the pellets were then washed with ~0.3 V of cold 70% (v/v) ethanol and centrifuged again for 20 min. The wash was removed by pipetting and the pellets were dried under vacuum.

The oxidized RNA was then coupled with a hydrazide derivative of the fluorophore Cy3 (GE Healthcare, PA13120). A typical 100 µL coupling reaction contained ~0.2 – 1.0 nmol of RNA, 50 nmol of Cy3 hydrazide (dye) dissolved in 10 µL of DMSO, and 100 mM NaOAc (pH 5.2). Solutions were degassed prior to the addition of dye, and the headspace above fully assembled reactions was flushed with nitrogen before capping the reaction tube. Reactions were protected from light and incubated at room temperature for 4 h with agitation. In all subsequent steps, solutions were protected from light. After the end of the incubation, the Cy3-labeled RNA was precipitated with the addition of 0.1 V of 3 M NaOAc (pH 5.2) and 2.5 V of cold absolute ethanol, followed by incubated on dry ice until frozen. The solution was inverted until just thawed and then centrifuged at 20,800 × g for 45 min at 4 °C to pellet the RNA. The supernatant was removed by pipetting, and the pellets were then washed with 2 V of cold 70% ethanol and

centrifuged again for 20 min. The wash was removed by pipetting and the pellets washed again with 0.5 V of cold 70% ethanol and centrifuged again for 15 min. This final wash was removed by pipetting and the RNA pellets were dried under vacuum and resuspended in 30 µL cold milliQ water, and then desalted using Illustra MicroSpin G-50 columns (GE Healthcare) that had been pre-equilibrated in milliQ water. The final concentration of RNA in the recovered solution was determined spectrophotometrically using a Nanodrop2000 spectrophotometer, using the extinction coefficient ε_{260} = 485437 M^{-1} cm^{-1} for the RNA and ε_{550} = 150,000 M^{-1} cm^{-1} for Cy3. The contribution of dye to the absorbance at 260 nm was accounted for as follows: $A_{260,\,RNA} = A_{260} - 0.08 \times A_{550}$.

Analysis of melting curves

Custom scripts were written in Matlab (The MathWorks) to partially automate the processing and analysis of data from melting curve experiments. As advocated by Owczarzy (11), an approximated second derivative of the absorbance *vs* temperature was plotted to aid the user in the appropriate selection of upper and lower baseline regions (12,13). The fraction folded (α) as well as the fraction unfolded (θ) as a function of temperature (T) were calculated from the baseline-corrected absorbance values, according to the equations $\alpha = (\mathrm{A_U} - \mathrm{A})/(\mathrm{A_U} - \mathrm{A_L}), \text{and } \theta = (\mathrm{A} - \mathrm{A_L})/(\mathrm{A_U} - \mathrm{A_L})$, where A, A_U, and A_L are the absorbance, absorbance of the upper baseline, and absorbance of the lower baseline, respectively. The resulting plot of α versus T was smoothed using a Savitzky-Golay FIR smoothing filter with a polynomial order of 1. The window for smoothing was varied from 5 – 17 points depending on the quality of the data. The local maxima in the first derivative plot of $d\alpha/d(1/T)$ corresponds to the melting temperature (T_m) for the special case of intramolecular unfolding present here. This method of determining T_m is less sensitive to the choice of baseline than other methods, such as the maximum of the first derivative of absorbance versus temperature (dA/dT). Because many of the melting curves displayed clear 2-step melting behavior, the T_m for each apparent transition was determined by fitting the plot of $d\alpha/d(1/T)$ versus T with the sum of one or more Gaussians.

Rate constant analysis from smFRET data

The smFRET traces were idealized with a two-state model using hidden-Markov modeling (HMM) with a segmental *k*-means algorithm as implemented in the program QuB (14,15) for all experimental conditions, with the exception of experimental conditions where a third FRET state was visually evident in the

smFRET histograms (i.e., RNA + S1, RNA + Gua + S1, and RNA + DAP + S1), in which case a three-state model was used.

Dwell times in the undocked and docked states were extracted from all the idealized traces, and the cumulative dwell time distributions were fit with an exponential function of the form $y = A(1 - e^{-kt})$ or $y = A_1(1 - e^{-k_1 t}) + A_2(1 - e^{-k_2 t})$ to obtain the rate constants k_{dock} and k_{undock}, respectively. In cases where 2-exponential fitting was merited, the reported rate constant is the amplitude-weighted average of the two fit coefficients k_1 and k_2. The error associated with the measured rate constants was estimated by bootstrapping using a custom Matlab script: bootstrap_TDP_TODP_v02_1.m. Briefly, this script loads all traces from a given experimental condition and attempts to perform 1- and 2-exponential fitting on this original dataset thus determining the measured rate constant(s) for the condition, and then performs bootstrap fitting on 1000 sets of replicate data that are chosen by sampling *M* traces from traces in original dataset with replacement, where *M* is equal to the number of traces present in the original data set. The standard deviation of the resulting 1000 replicate fit-coefficients is reported as the error of measured rate constant.

Single molecule cluster analysis (SiMCAn) of smFRET traces

To identify and visualize the various behavioral patterns in the smFRET trace data, single molecule cluster analysis (SiMCAn) was performed in MATLAB as described previously (16), with minor modifications. Because of the requirements of the SiMCAn pipeline, three-column trace data files ([time] [donor fluorescence] [acceptor fluorescence]) were converted to two-column data files ([donor fluorescence] [acceptor fluorescence]) and then re-idealized in vbFRET (17) with the following parameters: Fitting attempts = 10, Min FRET states = 1, Max FRET states = 5. A maximum of 5 FRET states was used to allow for any additional dynamic complexity present in the data to be captured. Overall, there was good agreement between vbFRET and QuB idealizations with respect to the number of FRET states used to fit data from a given condition. Peak centers from fitting of the smFRET histograms were used as initial guesses for FRET states. All trace data from a single condition were idealized together, yielding a single PATH.dat file for each experimental condition.

First-round clustering to identify common patterns of trace behavior was then performed as follows: the SiMCAn FRET_processing.m script was then used to reassign the idealized FRET state values for each trace into 10 evenly spaced bins (centered at 0.05, 0.15, 0.25, …, 0.95), yielding a *.datR file for each condition. Re-binned data files from all conditions were then loaded in SiMCAn Process_HMM_data.m, FRET state values were set from files, and the analysis was run with Threshold Length = 30 and with Remove Static traces selected. The number of dynamic clusters (i.e., 3) was chosen as the minimum number of clusters needed to maximize the magnitude of the difference between inter- and intra-cluster distances, while simultaneously minimizing the ratio of intra- to inter-cluster distance. The Display Processed Data function was then used to generate additional output plots for all identified static and dynamic clusters, for example the *.mldatx containing the distance and rank for each trace in the cluster. Minor changes were made to the script to include additional output files reporting the fraction of molecules from a given experimental condition present in each cluster, however the underlying algorithms are unchanged from the original implementation.

Second-round clustering to identify experimental conditions with common behavioral profiles was then performed. Given the relatively small number of traces present in the static FRET clusters with binned FRET efficiencies centered at 0.35, 0.45, 0.55 and 0.65, these clusters were first collapsed into a single static cluster S-0.35 to 0.65. In contrast to the original implementation of SiMCAn, in which clusters of trace behavior (first-round clusters, row labels) were grouped into clades based only on abundance profiles across conditions, here the clustergram (heat map) was generated by considering both experimental condition and trace behavior cluster. The result is that clusters of trace behavior were organized by similarity of their abundance profiles (i.e., fraction of molecules in a given cluster for each condition) across conditions (row dendrogram), and experimental conditions were also grouped by similarity of the patterns of trace behavior observed (column dendrogram). In addition, here the fraction of molecules in a given cluster for a given condition was used directly, whereas in the previous implementation of SiMCAn (16), the fractions were first normalized to a mean of zero with unit variance. These Matlab scripts, input data, and resulting analysis files associated with this work are available in DeepBlue from the University of Michigan library.

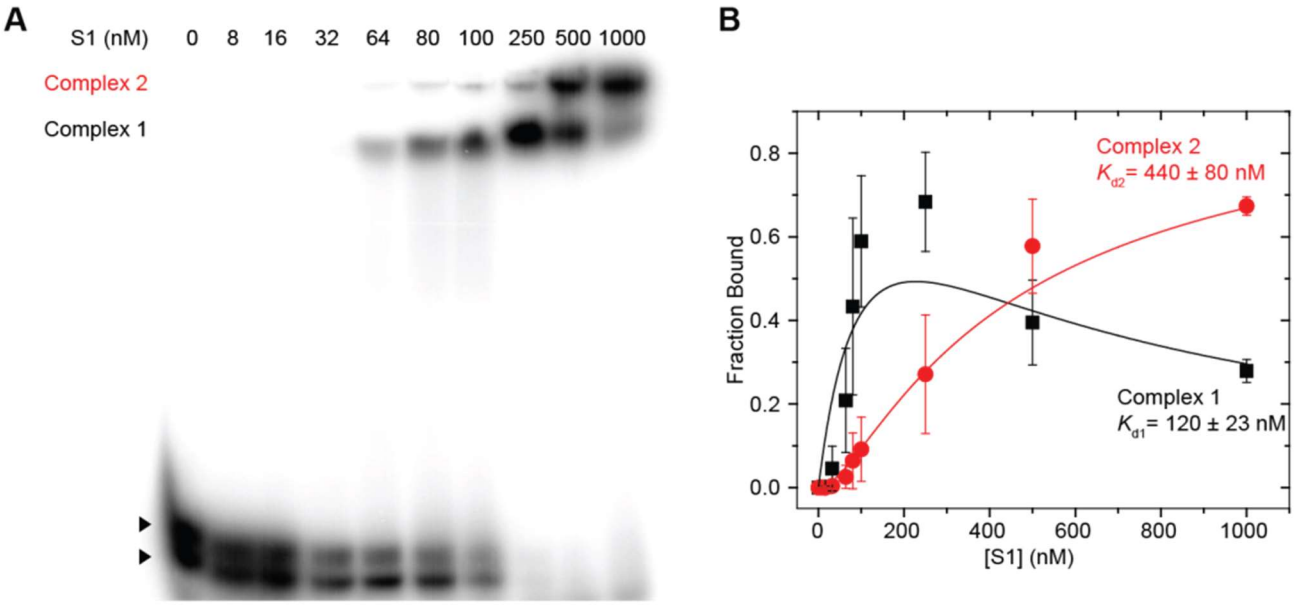

Figure S1. Assessment of S1 binding to the *Tte* pseudoknot

(A) Titration of trace amounts of 5'-^{32}P-labeled *Tte* pseudoknot with S1. (B) Quantification of fraction of RNA bound to S1. Dissociation constants were determined by using two-site binding model and fitting the results with Eq 1, 2 (see Materials and Methods) gives the values of K_{d1} = 120 ± 23 nM and K_{d2} = 440 ± 80 nM. Data represent the mean ± S.D. of n = 3 independent experiments. The complex 1 isotherm shows a decrease at the highest S1 concentrations (>250 nM). We posit that this is not due to dissociation of complex 1, but because of its conversion to complex 2, possibly mediated by direct protein-protein interactions.

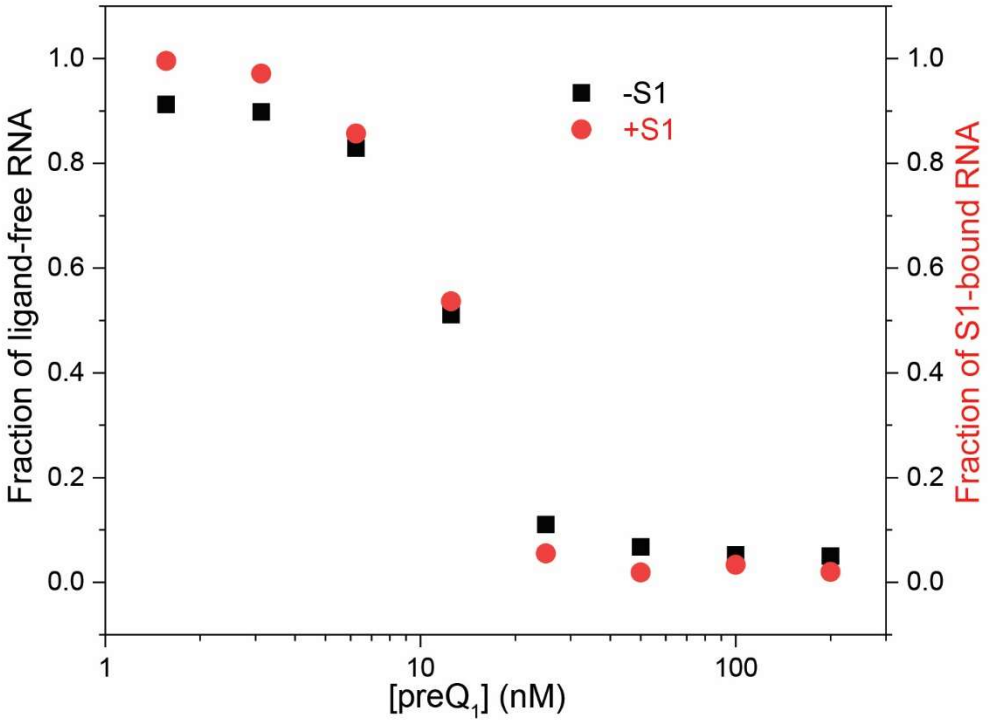

Figure S2. Quantifications of bands for ligand-free RNA (from Figure 2B) and S1-bound RNA (from Figure 2C).

The fraction of S1-bound RNA at a certain concentration of $preQ_1$ is similar to the fraction of ligand-free RNA present in the solution at that concentration of $preQ_1$ in the absence of S1, suggesting that S1 is only interacting with the ligand-free RNA.

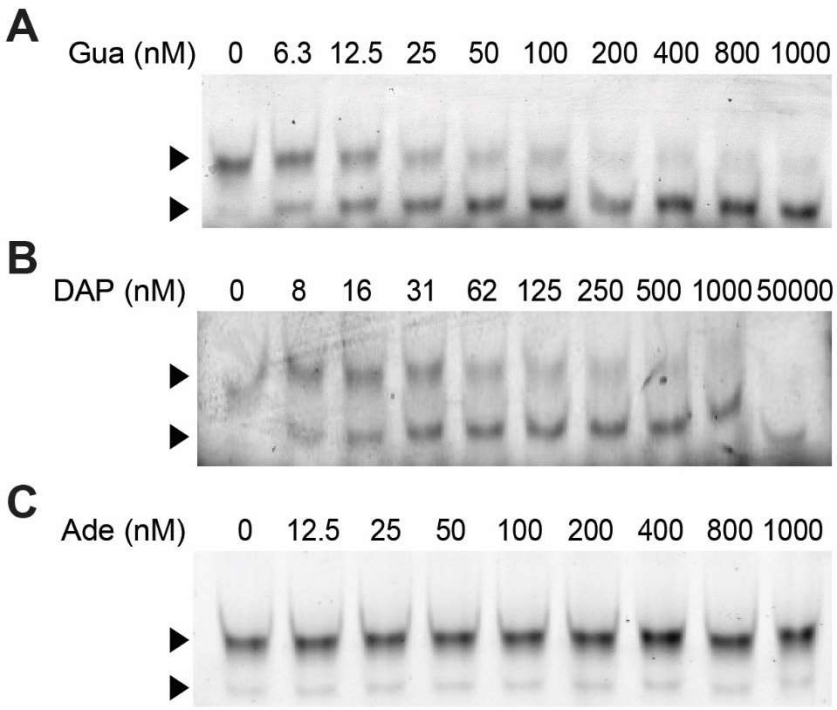

Figure S3. Assessment of weak ligands binding to the *Tte* pseudoknot

EMSA titration of the 3'-Cy3-labeled *Tte* pseudoknot with (A) guanine (Gua), (B) 2,6-diaminopurine (DAP), and (C) adenine (Ade). The RNA concentration was 10 nM in all lanes.

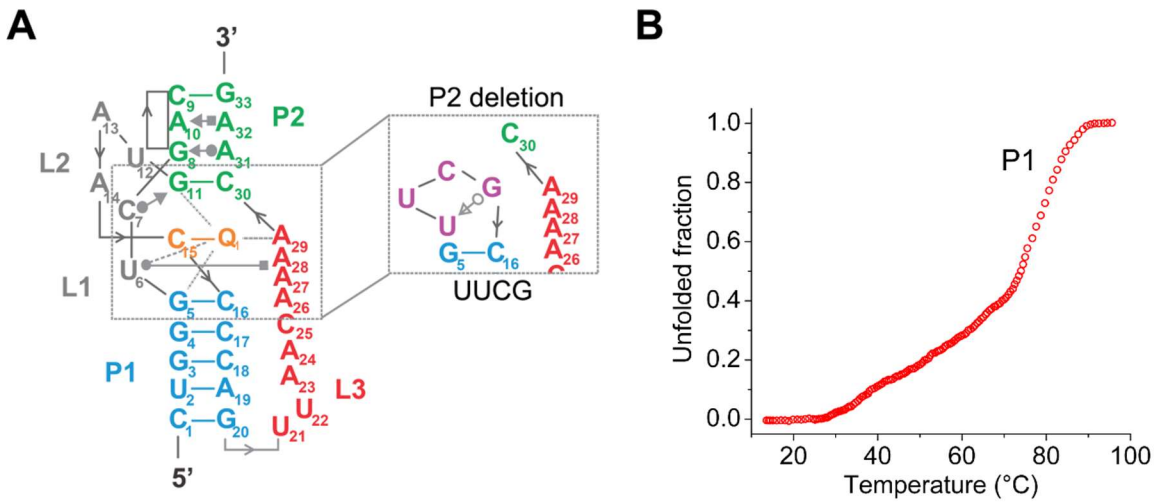

Figure S4. Characterization of pseudoknot variant through UV melting curves

(A) A variant of the *Tte* preQ$_1$ riboswitch was designed to completely abolish the pseudoknot structure through deletion of P2. (B) Melting curve of the pseudoknot variant shown in (A), revealing a single transition at >70°C for the melting of P1, which allows for the first transition of the wild-type to be confidently assigned to melting of P2.

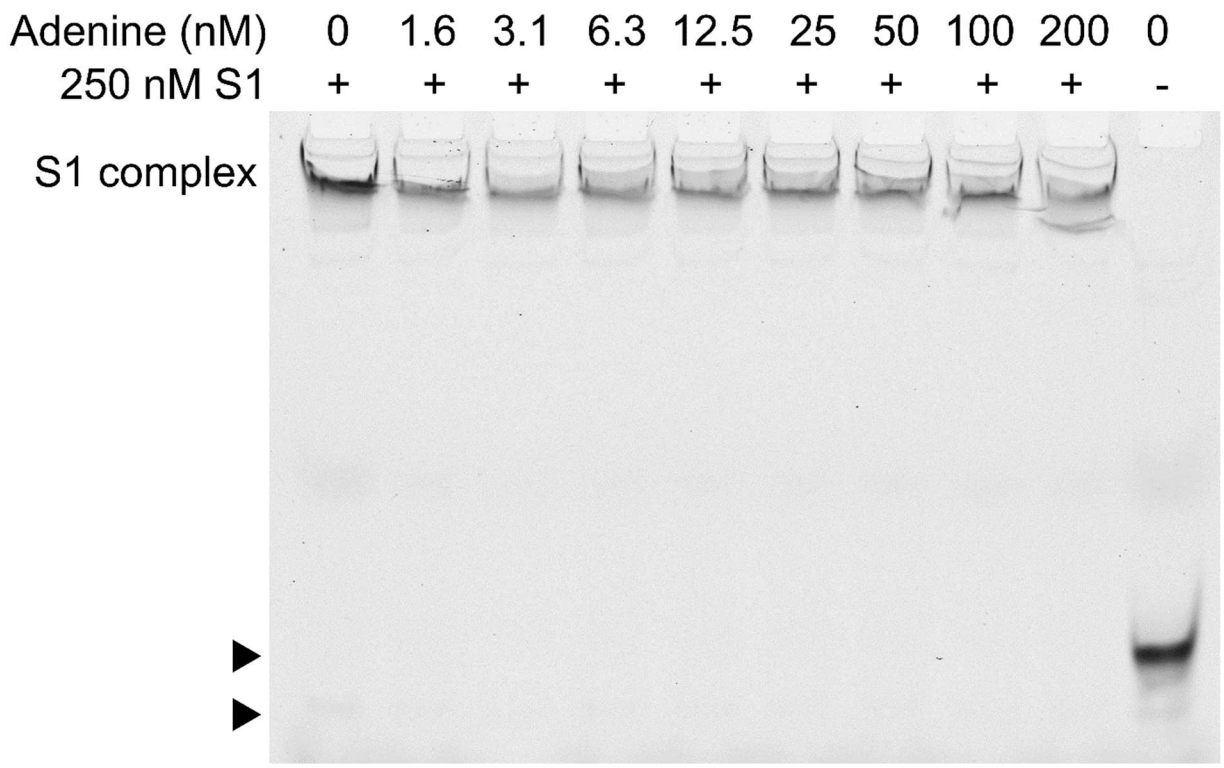

Figure S5. S1 binding to the *Tte* pseudoknot is not affected by adenine

Changes in S1 binding affinity are specific to ligands such as $preQ_1$ that alter RNA stability, whereas titration with adenine, which is not a near-cognate ligand, has no effect. Arrowheads indicate the position of the slower and faster migrating pre-folded and fully-folded forms, respectively, of the free RNA pseudoknot. RNA is present at 10 nM in all lanes.

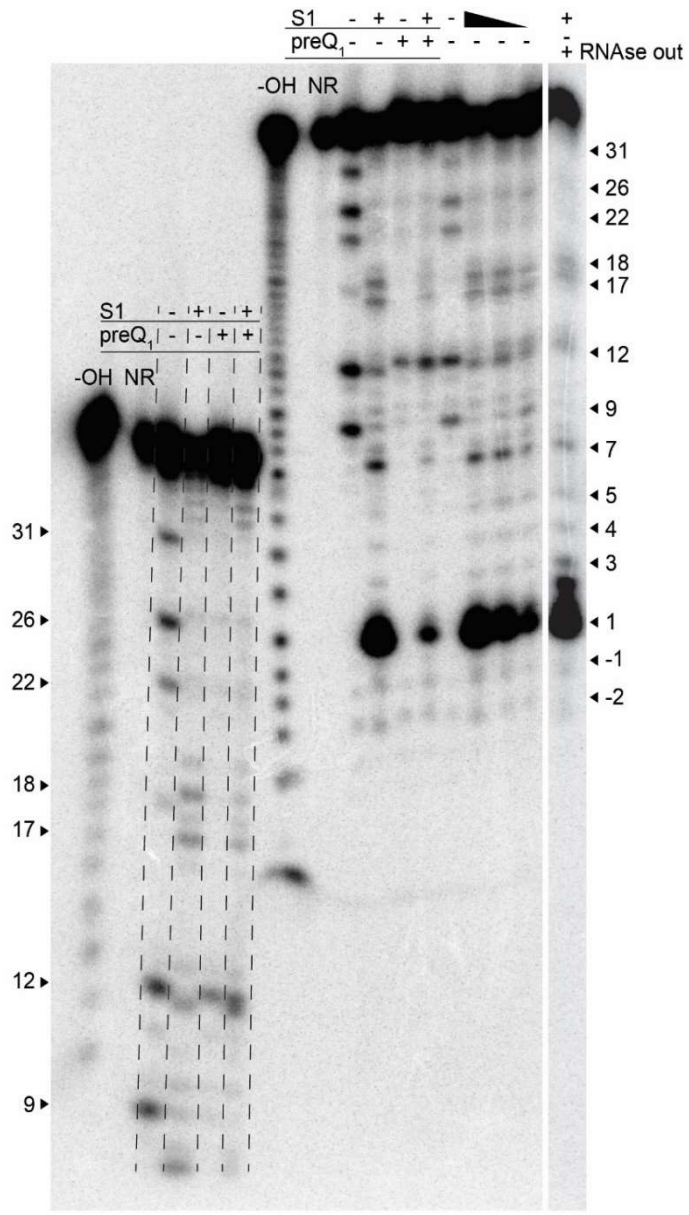

Figure S6. In-line probing of the RNA pseudoknot in the presence of S1 and preQ$_1$

In-line probing analysis of 5'-^{32}P-labeled *Tte* pseudoknot (as in Figure 4B), where the gel was doubly loaded to better observe the changes that occur in P2, particularly in nucleotides 8-10 and 31-33. The first six lanes were loaded and run for ~3 h, at which point the remaining ten lanes were loaded and the gel run for an additional ~1 h. The last four lanes before the break show that incubation of the radiolabeled *Tte* pseudoknot with a decreasing amount of S1 results in lesser destabilization of the pseudoknot. Concentrations of S1 were 2, 1, 0.5 µM. NR, no reaction; -OH, alkaline hydrolysis ladder. The last lane is in-line probing analysis of the 5'- 32 P-labeled *Tte* pseudoknot in the presence of S1 and RNase inhibitor (RNAseOUT, ThermoFisher Scientific) to test that the cleavage pattern of the pseudoknot are not arising due to presence of RNAse.

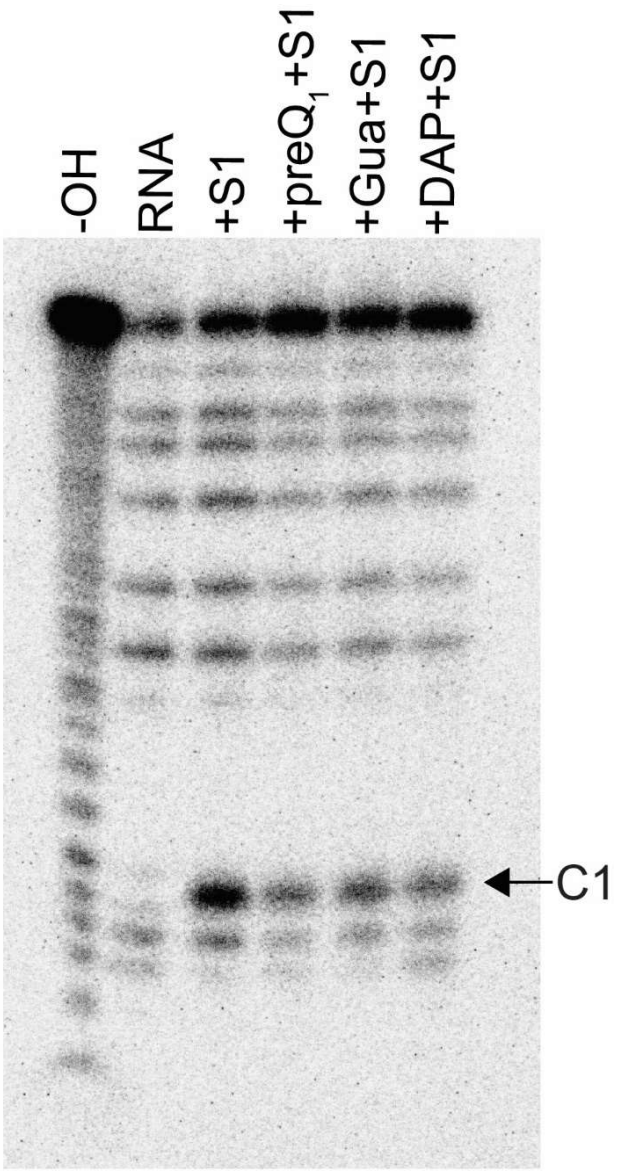

Figure S7. Cleavage of 5'-^{32}P-labeled *Tte* pseudoknot by RNase A in the absence and presence of S1 and ligands preQ$_1$, guanine (Gua), and 2,6-diaminopurine (DAP).

Ligand concentrations were 200 nM preQ$_1$, 1 μM Gua and 50 μM DAP.

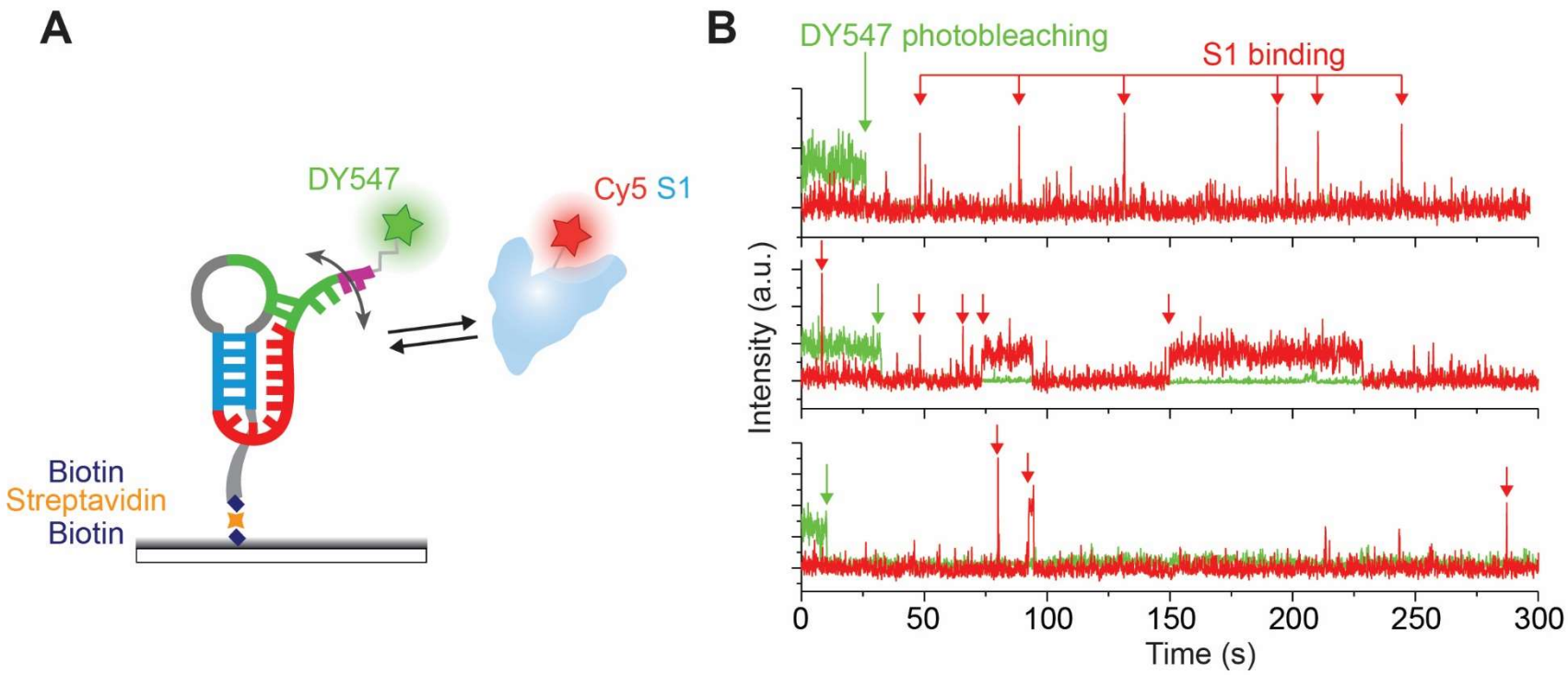

Figure S8. Reversible binding of Cy5 labeled S1 to surface-immobilized DY547 labeled *Tte* pseudoknot

(A) Protein S1 was sparsely labeled with Cy5 to observe the reversible binding to 3'-DY547 labeled *Tte* pseudoknot attached on the microscope slide. (B) Representative time traces indicate repeated, reversible S1 binding. Green arrows indicate the photobleaching of DY547, whereas S1 binding events are indicated with red arrows. Of note, some of the longer binding events may be curtailed by photobleaching.

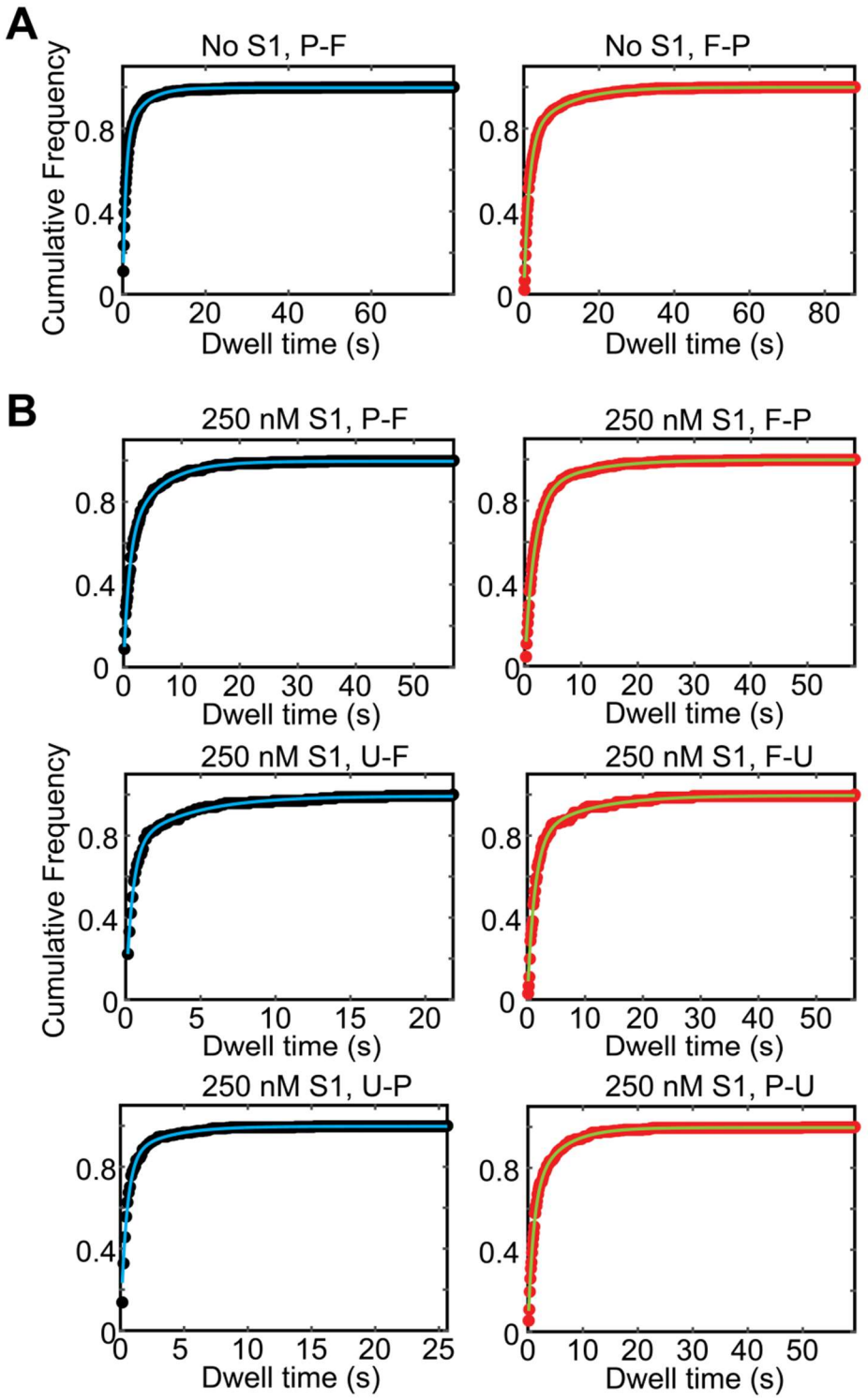

Figure S9. Representative cumulative dwell time distributions in the pre-folded (P), folded (F) and unfolded (U) states of the *Tte* pseudoknot in the absence (A) and presence (B) of 250 nM S1

These cumulative dwell-time distributions were fit with double-exponential functions to estimate the rate constants.

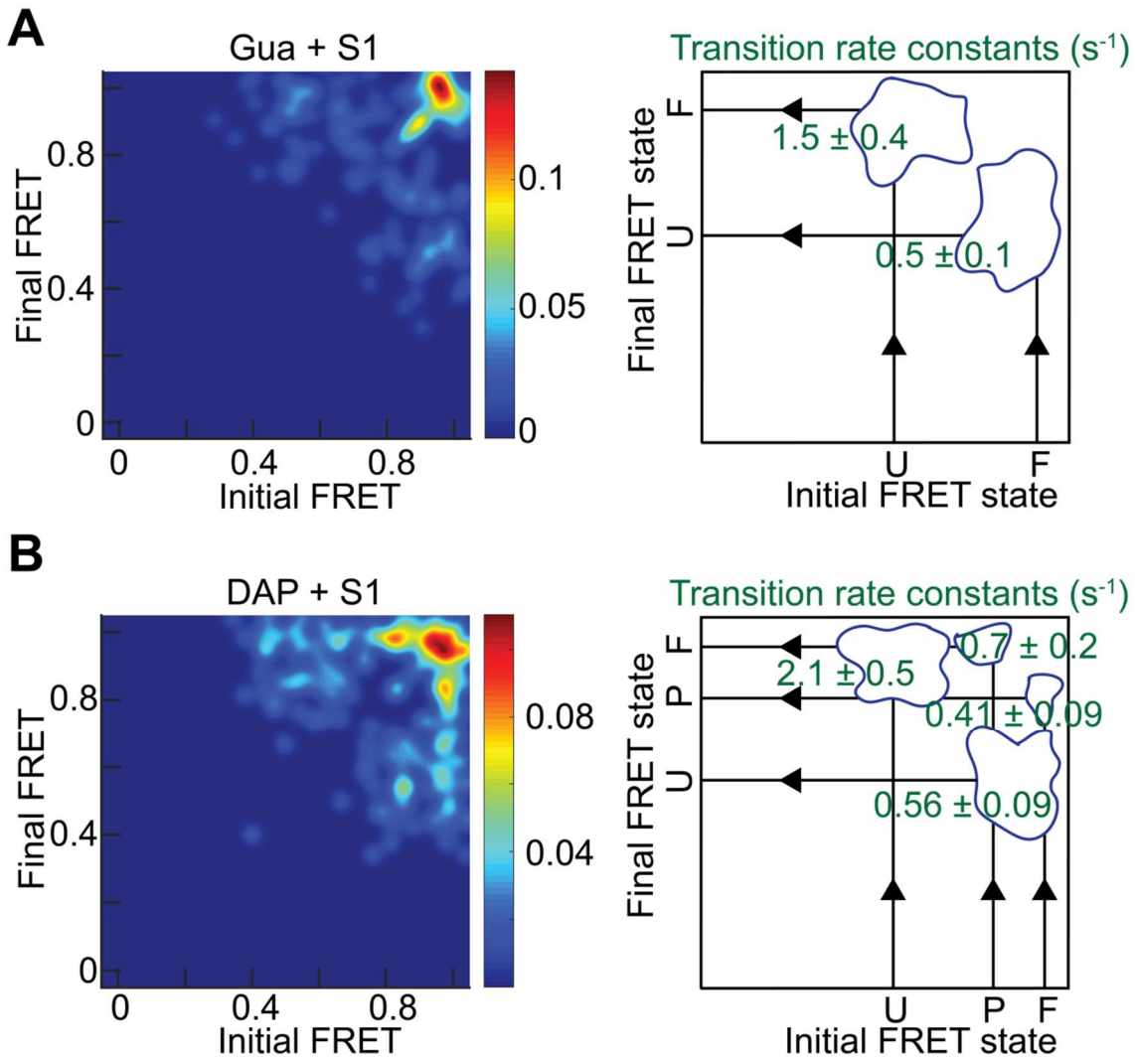

Figure S10. Single molecule FRET Transition Occupancy Density Plots for the *Tte* pseudoknot in the presence of S1 and weakly stabilizing ligands

Transition occupancy density plots (TODPs) illustrating the most common FRET transitions (left panel) and corresponding mean rate constants (right panel) for Gua (A) and DAP (B) in the presence of S1. Outlined regions show the population of molecules that exhibit a particular FRET transition considered in the calculation of each rate constant. U, unfolded; P, pre-folded; F, folded.

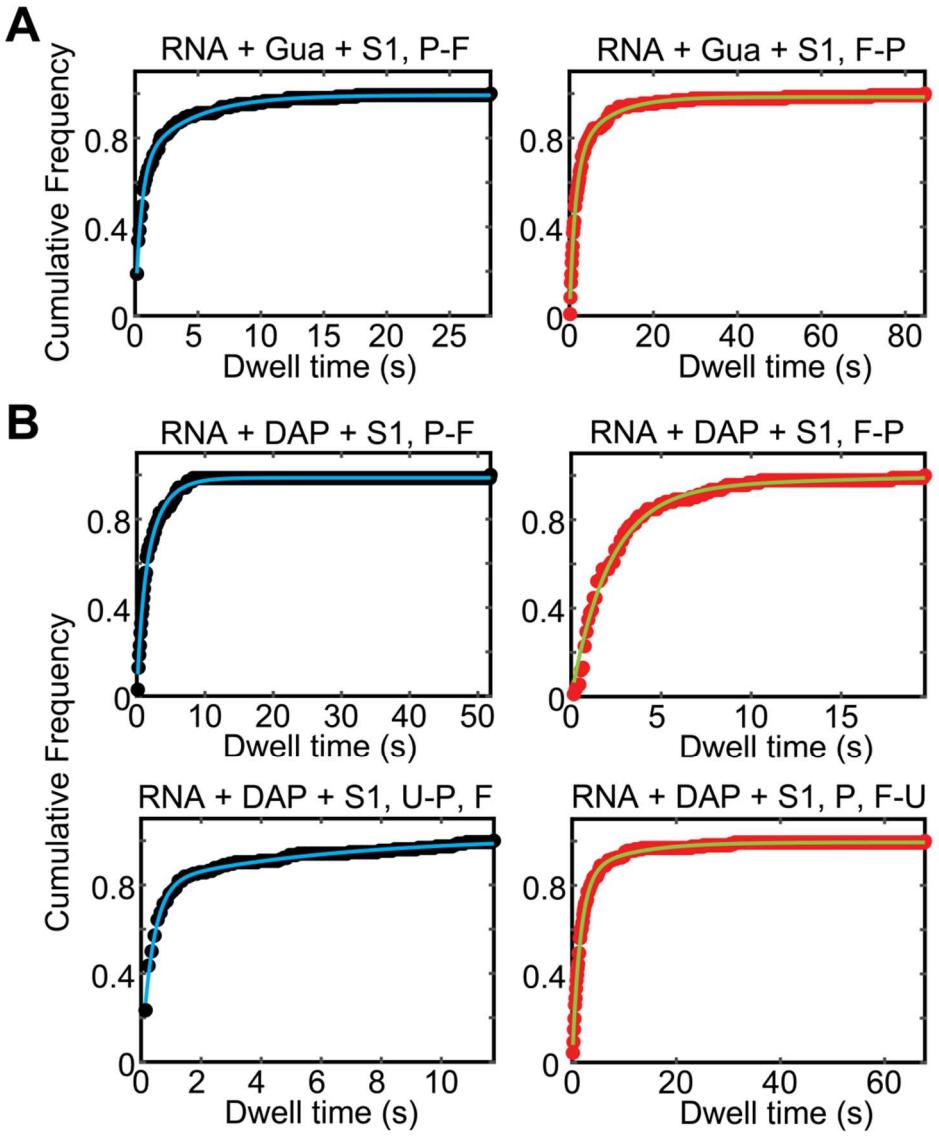

Figure S11. Representative cumulative dwell time distributions in the pre-folded (P), folded (F) and unfolded (U) states of Guanine-(Gua) (A) and DAP-bound (B) pseudoknot in the presence of protein S1

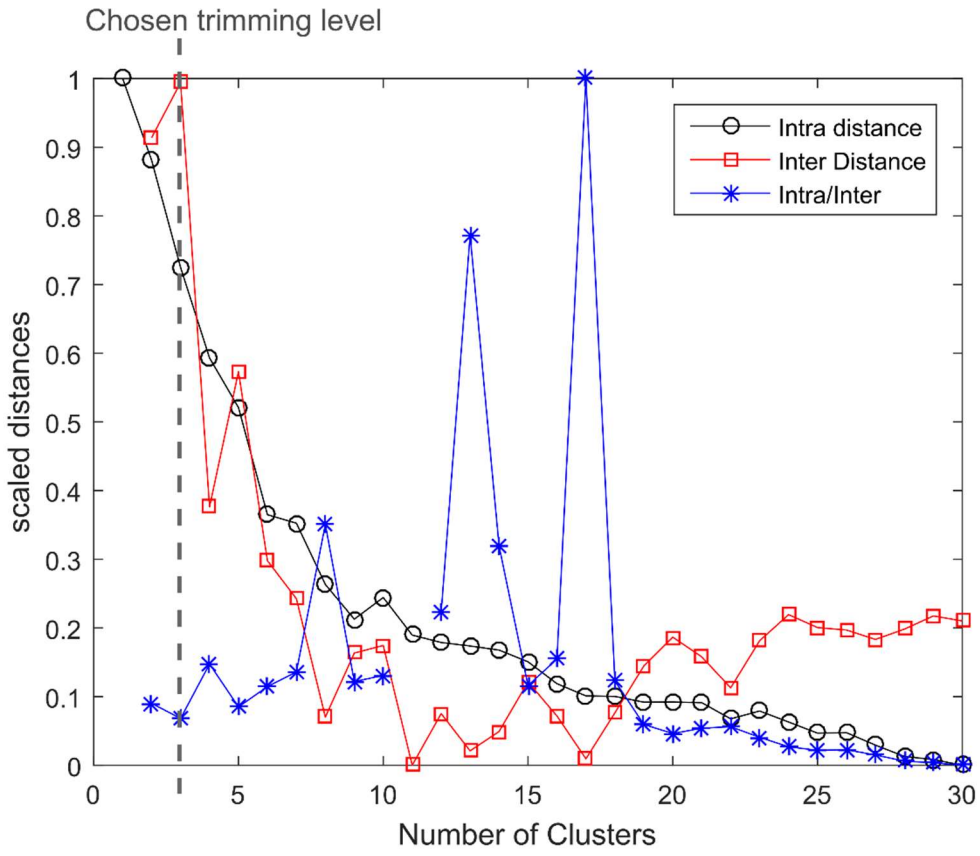

Figure S12. Distance plot generated in SiMCAn for choosing number of dynamic clusters

The level at which the hierarchical tree is trimmed was chosen by finding the number of clusters that maximizes the inter-cluster distances (i.e., the behavioral pattern of each cluster is most different from other clusters) as well as minimizes the intra-cluster distance (i.e., minimizes differences between molecular behaviors from molecules within the same cluster) and minimizes the ratio of intra/inter-cluster distances (i.e., minimizes the magnitude of the differences between molecules within a cluster relative to the magnitude of differences between clusters). Note that the inter- and intra-cluster distances have been scaled between 0 and 1 for the purposes of plotting.

SUPPLEMENTARY REFERENCES

1. Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H. and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.*, **12**, 291-299.
2. Kapust, R.B., Tozser, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D. and Waugh, D.S. (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.*, **14**, 993-1000.
3. Sawano, A. and Miyawaki, A. (2000) Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. *Nucleic Acids Res.*, **28**, e78.
4. Lancaster, L. and Noller, H.F. (2005) Involvement of 16S rRNA nucleotides G1338 and A1339 in discrimination of initiator tRNA. *Mol. Cell*, **20**, 623-632.
5. Kibbe, W.A. (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.*, **35**, W43-46.
6. Rinaldi, A.J., Lund, P.E., Blanco, M.R. and Walter, N.G. (2016) The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs shows ligand-dependent accessibility bursts. *Nat. Commun.*, **7**, 8976.
7. He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., McAllister, W.T. and Durbin, R.K. (1997) Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expr. Purif.*, **9**, 142-151.
8. Gurevich, V.V., Pokrovskaya, I.D., Obukhova, T.A. and Zozulya, S.A. (1991) Preparative *in vitro* mRNA synthesis using SP6 and T7 RNA polymerases. *Anal. Biochem.*, **195**, 207-213.

9. Cunningham, P.R. and Ofengand, J. (1990) Use of inorganic pyrophosphatase to improve the yield of *in vitro* transcription reactions catalyzed by T7 RNA polymerase. *Biotechniques*, **9**, 713-714.

10. Willkomm, D.K. and Hartmann, R.K. (2008), *Handbook of RNA Biochemistry*. Wiley-VCH Verlag GmbH, pp. 86-94.

11. Owczarzy, R. (2005) Melting temperatures of nucleic acids: discrepancies in analysis. *Biophys. Chem.*, **117**, 207-215.

12. Mergny, J.L. and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515-537.

13. Puglisi, J.D. and Tinoco Jr, I. (1989) Absorbance melting curves of RNA. *Methods Enzymol.*, **180**, 304-325.

14. Qin, F. and Li, L. (2004) Model-based fitting of single-channel dwell-time distributions. *Biophysical Journal*, **87**, 1657-1671.

15. Blanco, M. and Walter, N.G. (2010) Analysis of complex single-molecule FRET time trajectories. *Methods in Enzymology*, **472**, 153-178.

16. Blanco, M.R., Martin, J.S., Kahlscheuer, M.L., Krishnan, R., Abelson, J., Laederach, A. and Walter, N.G. (2015) Single Molecule Cluster Analysis dissects splicing pathway conformational dynamics. *Nat. Methods*, **12**, 1077-1084.

17. Bronson, J.E., Fei, J.Y., Hofman, J.M., Gonzalez, R.L. and Wiggins, C.H. (2009) Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data. *Biophys. J.*, **97**, 3196-3205.