Supplementary Material

Single-molecule force spectroscopy of the *add* adenine riboswitch relates folding to regulatory mechanism

Krishna Neupane, Hao Yu, Daniel A.N. Foster, Feng Wang, and Michael T. Woodside

Supplementary Materials and Methods

RNA construct design and assembly

A 77-bp DNA fragment coding for the *add* adenine riboswitch aptamer and 3 flanking nucleotides on each side (see Table S1) was inserted between the BamH I and Spe I sites of a pMLuc-1 plasmid (Novagen). A 2,226-bp, linear, double stranded (ds) DNA transcription template was amplified by PCR from this plasmid, with a T7 promoter placed in the upstream primer. The completed template consisted of 1,289 bp of linker sequence upstream of the aptamer, followed by the aptamer itself, and then an additional 842 bp of linker downstream of the aptamer. These 2,208 bp of DNA were transcribed with MEGAscript high yield transcription kit (Ambion). Two single stranded (ss) DNA handles designed to be complementary with the upstream and downstream linker sequences flanking the *add* A-riboswitch aptamer were created by asymmetric PCR (1): primary PCR was first performed with both forward and reverse primers for each handle on its own, then the purified ds PCR products were used as templates in asymmetric PCR reaction with only one primer present. The 3´ end of the DNA handle complementary to the upstream (5´ end) linker section of the transcript was labelled with dig-ddUTP using terminal transferase (Roche). The DNA handle complementary to the downstream (3´) linker of the transcript was functionalised with biotin on the 5´ end of the PCR primer.

The RNA transcript and the two ssDNA handles complementary to the linker sequences on either side of the aptamer were annealed by gradually increasing the temperature to 85°C, then stepping down to 62 and 52°C for 90 minutes each in a thermal cycler. The RNA-DNA hybrid thereby created consisted of a 1,289-nt RNA-DNA duplex handle labelled with digoxigenin at one end, 77 nt of RNA containing the aptamer, and a 842-bp RNA-DNA duplex handle labelled with biotin at the other end. The aptamer is separated from the duplex handles by 3 nt of unstructured ssRNA, to uncouple interactions between the handles and the aptamer.

Constructs containing the expression platform alone and the full-length riboswitch were created in the same way, using either a 58-bp DNA insert coding for the *add* adenine riboswitch expression platform, or a 117-bp insert coding for the full-length riboswitch, in both cases with flanking nucleotides on each side (Table S1).

The final hybrid constructs were diluted to a concentration of 160 pM and mixed with equal volumes of 600-nm and 820-nm diameter polystyrene beads (Bangs) at concentrations of 250 pM. The smaller beads were coated with avidin DN (Vector Labs), and the larger beads were coated with polyclonal anti-digoxigenin (Roche). The bead-construct mixture was incubated for 1 hour at room temperature, then diluted 100-fold into RNAse-free measuring buffer: 50 mM MOPS pH 7.5, 130 mM KCl, 4mM $MgCl_2$, and 200 U/mL RNAse inhibitor (SUPERase•In, Ambion). An oxygen scavenging system consisting of 40 U/mL glucose oxidase (Sigma), 185 U/mL catalase (EMD), and 250 mM D-glucose (Sigma) was also included in the buffer. The diluted dumbbells were placed in a flow chamber prepared on a microscope slide, and inserted into the optical trap.

Optical trap and data acquisition

A custom-built dual-optical trapping instrument (optical tweezers) was used to perform the force spectroscopy measurements. A 1,064-nm diode-pumped solid-state infrared laser (Spectra-Physics) was used to create two traps with orthogonal polarisations. The bead position in each trap was monitored by measuring the light of a 633 nm He-Ne laser (Melles-Griot), split into two

beams by polarisation and scattered off the two beads, using position sensitive detectors (Pacific Silicon Sensor). The trap stiffnesses were controlled independently by adjusting the power applied to acousto-optic deflectors (AODs) in the paths of the trapping beams. The stiffness of each trap was calibrated from an average of three different methods as described previously (2). The position of each trap was controlled in the pulling axis using an electro-optic deflector (Conoptics), and in the orthogonal axis using an AOD (AA Optoélectronique). The optical trapping instrument was located in a low-vibration environment where the temperature of the room was controlled to 20.0 ± 0.1°C during the measurements. All measurements were performed using customised Labview software (National Instruments). Data were acquired at 20 kHz and filtered at 10 kHz with an 8-pole Bessel filter (Krohn-Hite).

Measurements and analysis: Force extension curves

Force extension curves (FECs) were recorded by moving the traps apart (for unfolding) and/or together (for refolding) at constant velocities between 10–3,000 nm/s, producing loading rates ranging from ~5–400 pN/s. Data were analysed using custom software in Igor Pro (Wavemetrics). In order to remove slow drift over the course of the measurements (up to several hours), individual FECs measured on the same molecule under the same pulling conditions were aligned vertically using the low-force (F ~ 2–3 pN) part of the data, where the FECs have very low slope, and horizontally using the high-force (F ~ 20 pN) part of the data, where the FECs have very high slope. Typical drift correction was less than a few nm. All analysis was performed on aligned FECs.

Contour length changes during unfolding were determined by partitioning the aligned FEC data into separate states having different contour lengths, then fitting each state to an extensible worm-like chain (WLC) model that consisted of two WLCs in series: one for the duplex handles, and one for the ssRNA that was unfolded in each state. The extensible WLC interpolation formula (3),

$$F(x) = \frac{k_B T}{L_p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c} + \frac{F}{K}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{K}\right], \qquad \text{(S1)}$$

depends on the persistence length L_p, the contour length L_c, and the elastic modulus K. The parameters L_p, L_c, and K describing the duplex handles were first determined by fitting the FEC for the fully-folded state. Then the FECs for the partially- and fully-unfolded states were fit by treating L_p, L_c, and K as fixed variables for both the duplex handles (values determined from the previous fit) and the unfolded ssRNA (values determined from the literature (4): L_p = 1 nm, K ~ 1,500 pN). The contour length of unfolded RNA was therefore the only fitting parameter for the partially- and fully-unfolded states.

To convert ΔL_c into the number of nucleotides unfolded, n_{nt}, we included the effects of changing the number of folded helices: $n_{nt} = (\Delta L_c + \Delta n_h d_h)/L_c^{nt}$, where Δn_h is the number of helices unfolded, d_h = 2.2 nm is the diameter of an A-form dsRNA helix, and L_c^{nt} = 0.59 nm is the contour length per ssRNA nucleotide (5). For example, when one hairpin was unfolded, 2.2 nm (the width of one helix) was added to the fitted contour length change, and the total divided by 0.59 nm/nt to obtain the number of nucleotides involved in the unfolding.

Measurements and analysis: Identification of P2 and P3 folding

The folding transitions for hairpins P2 and P3 were primarily identified by the number of nucleotides involved. The constant force measurements were sufficiently precise so as to be able to distinguish between P2 and P3: the standard error on the mean number of nucleotides involved in the transitions was only 0.4 nt. Hence the 2-nt length difference between P2 and P3 represents 5 standard deviations, and misidentification is extremely unlikely. Nevertheless, to confirm the identification, we also measured the folding at constant force using 10 μM of an anti-sense DNA oligomer complementary to nucleotides 25-38 (the 5′ arm and loop of hairpin P2; sequence listed in Table S1) to prevent formation of P2. In this case, we should observe P3 folding but none of the other states, since they all depend on the folding of P2. As seen in Fig. S1C, the oligomer indeed prevented all folding transitions except the one occurring at the highest forces, which was identified from both the force range and the extension change (19.8 ± 0.5 nt) as being P3 folding.

For the FEC measurements, the standard error of 1 nt was sufficiently large to leave some uncertainty in the identification of P2 and P3. As additional surety that the transitions corresponding to hairpins P2 and P3 were correctly identified, we used two different anti-sense DNA oligomers (sequences listed in Table S1) to block, separately, the formation of P2 or P3: one complementary to nucleotides 25-38 (the 5′ arm and loop of hairpin P2), the other to nucleotides 59-72 (the loop and 3′ arm of hairpin P3). The first oligomer would be expected to prevent the folding of P2 (and all other states), but leave P3 folding unaffected, whereas the second should prevent P3 folding (and all other states) but leave P2 folding unaffected.

FECs measured in the presence of 10 μM of the P2-blocking oligomer (Figure S1A, red) show that indeed the only unfolding transition that occurs is the one at the highest forces which had been previously identified by the extension change as P3. The contour length change for this remaining transition corresponds to 19.5 ± 0.5 nt from 550 FECs. Since 19 nt are expected for P3, these results confirm the identification of the P3 folding transition. In contrast, FECs measured in the presence of 10 μM of the P3-blocking oligomer (Figure S1A, blue) show a single unfolding event but with a longer contour length: 22 ± 1 nt determined from 165 FECs, matching the 21 nucleotides expected for P2 unfolding. Just as importantly, the unfolding force distribution for P2 (i.e. P3 blocked: Fig. S1B, blue) was significantly different from the distribution for P3 (i.e. P2 blocked: Fig. S1B, red), with a mean force 2 pN lower matching the force difference recorded in the constant data (Table S3).

As a final check on the identification of the transitions in Fig. 1, we compared the distribution of unfolding forces for the transition attributed to P3 in Fig. 1 (Fig. S1B, black) to the distributions expected for P2 unfolding (Fig. S1B, blue) or P3 unfolding (Fig. S1B, red). It clearly matches the distribution expected for P3 unfolding, and is inconsistent with the distribution for unfolding P2.

Measurements and analysis: Energy landscape parameters from constant force measurements

Since folding of the aptamer was apparently sequential, each transition could be analysed as a two-state system. For each transition, the force at equipoise (where the molecule spends equal time in folded and unfolded states), $F_{1/2}$, was determined by fitting the probability distribution for the unfolded state, $P_u(F)$, to the Boltzmann relation (6): $P_u(F) = \{1+\exp[(F_{½}-F)\,\Delta x/k_BT]\}^{-1}$. The position of the transition state along the reaction coordinate, $\Delta x^{‡}$, was determined from the force-dependence of the kinetics (6): $k(F) = k_0\exp(-F\Delta x^{‡}/k_BT)$. This method was used for the

transition state distances from both the folded and unfolded state, based on the respective folded and unfolded state kinetics. We assumed that the position of the barrier is force-independent (borne out fairly well over the relatively narrow range of forces measured in each transition by the linear force dependence of the logarithm of the rates). As a consistency check, $F_{1/2}$ was also determined from the force at which folding and unfolding rates were equal; this value was found to agree (within experimental uncertainty) with the value obtained from the probability distribution.

The free energy difference between adjacent states within the landscape was calculated from $\Delta G = (F_{½} - F) \cdot \Delta x$. The barrier height was calculated at $F_{½}$ from the two-state kinetics, $\Delta G^{‡}_{½} = -k_B T \ln(k_{½}/k_0)$, then corrected for the final force by.$(F_{½} - F)\Delta x^{‡}$. The energy landscapes therefore include the free energy for stretching the unfolded ssRNA. The constant force results used to recreate the energy landscapes are listed in Table S3. All energies and positions in the landscape were calculated with respect to the state P2P3.

Measurements and analysis: Identification of misfolded state

The extension change from U to the misfolded state "M" (the most-occupied off-pathway state) was determined from constant-force extension histograms (Figure S2B). Assuming the creation of a single helix during the misfolding, the measured extension change corresponds to 31 ± 1 nt; if other structures such as pseudoknots formed, then the number of nucleotides involved in the misfolding might differ from this value, depending on the geometry of the structure.

The two potential structures found in a search for sub-optimal structures are illustrated in Figure 4B: a potential pseudoknot involving nucleotides 17-45 and a hairpin involving nucleotides 37-68. In principle only the hairpin can fully explain the set of states observed at constant force: the pseudoknot does not interfere with P3 folding, hence the misfolded state would not occur solely from the unfolded state. Nevertheless, to distinguish between these possibilities and confirm our hypothesis that the misfolding involves the hairpin, we performed constant force measurements with two different anti-sense oligomers: one complementary to the 5′ strand of P1 (nucleotides 13-22), which would inhibit the formation of the pseudoknot, and one complementary to the junction loop J2/3 between P2 and P3 (nucleotides 46-53). Misfolded states were still observed in the presence of 10 μM of the first oligomer (Figure 4C), definitively ruling out the pseudoknot as a possibility, whereas they were not seen in the presence of 10 μM of the second oligomer.

Measurements and analysis: Predicted thermodynamics of folding

The free energies of secondary structures were calculated using *mfold* (7), with an estimated 25% reduction to account for the different buffer ionic strength and temperature for our measurement conditions (8). For example, the prediction for the stability of P1 and the junction loops is −11.5 kcal/mol at 37 °C and 1 M monovalent salt, which would be equivalent to ~ −8.6 kcal/mol in our experiments.

The relative thermodynamic stabilities of the translationally-active ("on") and repressed ("off") states of the riboswitch were estimated from the predicted secondary structure energy of the full-length riboswitch in the "on" and "off" states, adding in the tertiary structure stabilities measured from force spectroscopy. The secondary structure alone is predicted to be more stable by 1.4 kcal/mol in the "off" state (−17.9 kcal/mol) than in the "on" state (−16.5 kcal/mol). In the absence of adenine, tertiary structure in the "on" state (mostly the loop-loop interaction) brings

another −1.5 kcal/mol of stability. However, since the loop-loop interaction can still form when P1 is prevented from folding (as seen by using the P1 blocking oligomer, Figure 4C), the difference between "on" and "off" states remains unchanged. Since adenine binding provides an additional stability of −8 kcal/mol, with adenine bound the "on" state is more stable by 6.6 kcal/mol, whereas without adenine the "off" state is more stable by 1.4 kcal/mol. Errors in these numbers are estimated at ~ 1 kcal/mol.

Measurements and analysis: Full-length riboswitch FECs

The "on" and "off" states of the riboswitch were identified from the total contour length change upon complete unfolding of the full-length riboswitch. In the case of the "on" state, when only the aptamer is folded this should be ΔL_c = 39.7 nm (71 nt and 1 helix). In the case of the "off" state, this should be 53.4 nm (98 nt and 2 helices, assuming that the loop-loop interaction between P2 and P3 is still formed). The two initial states were also characterised by different unfolding forces: the "on" state unfolded at high force (as for the adenine-bound aptamer), whereas the "off" state unfolded at lower force and moreover displayed additional partially-unfolded intermediates.

Supplementary References

1. Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1986) Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature*, **324**, 163-166.
2. Svoboda, K. and Block, S.M. (1994) Biological applications of optical forces. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 247-285.
3. Wang, M.D., Yin, H., Landick, R., Gelles, J. and Block, S.M. (1997) Stretching DNA with optical tweezers. *Biophys. J.*, **72**, 1335-1346.
4. Seol, Y., Skinner, G.M. and Visscher, K. (2004) Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.*, **93**, 118102-118105.
5. Saenger, W. (1984) *Principles of Nucleic Acid Structure*. Springer, New York.
6. Woodside, M.T., Behnke-Parks, W.M., Larizadeh, K., Travers, K., Herschlag, D. and Block, S.M. (2006) Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 6190-6195.
7. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406-3415.
8. Liphardt, J., Onoa, B., Smith, S.B., Tinoco, I.J. and Bustamante, C. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*, **292**, 733-737.

Table S1 Sequences of RNA molecules measured: *add* riboswitch aptamer, expression platform, full-length riboswitch, and oligomer sequences complementary to portions of the aptamer.

Aptamer sequence	5'-*CGC*GGCTTCATATAATCCTAATGATATGGTTTGGGAGTTTCTAC CAAGAGCCTTAAACTCTTGATTATGAAGTC*TGT*-3' The single-stranded linkers to the duplex handles are listed in italics.
Expression platform	5'-*CCA*TTATGAAGTCTGTCGCTTTATCCGAAATTTTATAAAGAGAA GACTCATGAA*TCCC*-3'
Full-length sequence	5'-*CGC*GGCTTCATATAATCCTAATGATATGGTTTGGGAGTTTCTAC CAAGAGCCTTAAACTCTTGATTATGAAGTCTGTCGCTTTATCCGAA ATTTTATAAAGAGAAGACTCATAA*ATC*-3'
Oligomers complementary to aptamer	
complementary to junction J2/3 between P2 and P3 (46-53)	5'-GTAGAAAC-3'
complementary to 5′ strand of P1 (13-22)	5'-ATATGAAGCC-3'
complementary to 5′ strand and loop of P2 (25-38)	5'-CCATATTATTAGGA-3'
complementary to loop and 3′ strand of P3 (59-72)	5'-CAAGAATTTAAGGC-3'

Table S2. Aptamer unfolding contour length changes. The expected contour length change from unfolding different structural components of the aptamer are listed, along with the values obtained from analysis of force-extension curves in the presence and absence of adenine. Uncertainties represent the standard error.

Transition	**F to U**	**F to P1_U**	**F to P2P3**	**F to P2**	**F to P3**
ΔL_c expected	39.7 nm (71 nt)	10.6 nm (18 nt)	20.5 nm (31 nt)	29.5 nm (50 nt)	30.7 nm (52 nt)
ΔL_c measured (no adenine)	39.6 ± 0.5 nm		19 ± 1 nm		30.3 ± 0.5 nm
ΔL_c measured (with adenine)	40.1 ± 0.8 nm		20.6 ± 0.8 nm		30.4 ± 0.4 nm

Table S3. Constant force results. "P3" represents folding of hairpin P3, "P2" represents folding of hairpin P2, "P1_U" represents the P1-unfolded state, "NAF" represents complete folding in the absence of adenine, and "AF" represents complete folding in the presence of adenine.

Folding transition	Δx (nm)	Δx (nt)	$F_{1/2}$ (pN)	$\ln(k_{1/2})$ (s^{-1})	$\Delta G^{\ddagger}_{1/2}$ (kcal/mol)	$\Delta x^{\ddagger}_{f}$ (nm)	$\Delta x^{\ddagger}_{u}$ (nm)
P3	5.2 ± 0.2	19.0 ± 0.5	12.9 ± 0.2	4.2 ± 0.1	4.3 ± 0.1	2.7 ± 0.2	4.1 ± 0.3
P2	6.3 ± 0.2	21.3± 0.5	10.8 ± 0.2	4.1± 0.1	4.4 ± 0.1	2.5 ± 0.3	4.6 ± 0.4
P1_U	5.1 ± 0.2	8.6 ± 0.6	10.1± 0.2	4.3 ± 0.2	4.3 ± 0.1	2.1 ± 0.2	3.9 ± 0.2
NAF	7.7 ± 0.4	22 ± 1	10 ± 2	2 ± 1	6 ± 1	N/A	N/A
AF	7.7 ± 0.4	22 ± 1	N/A	N/A	N/A	N/A	N/A

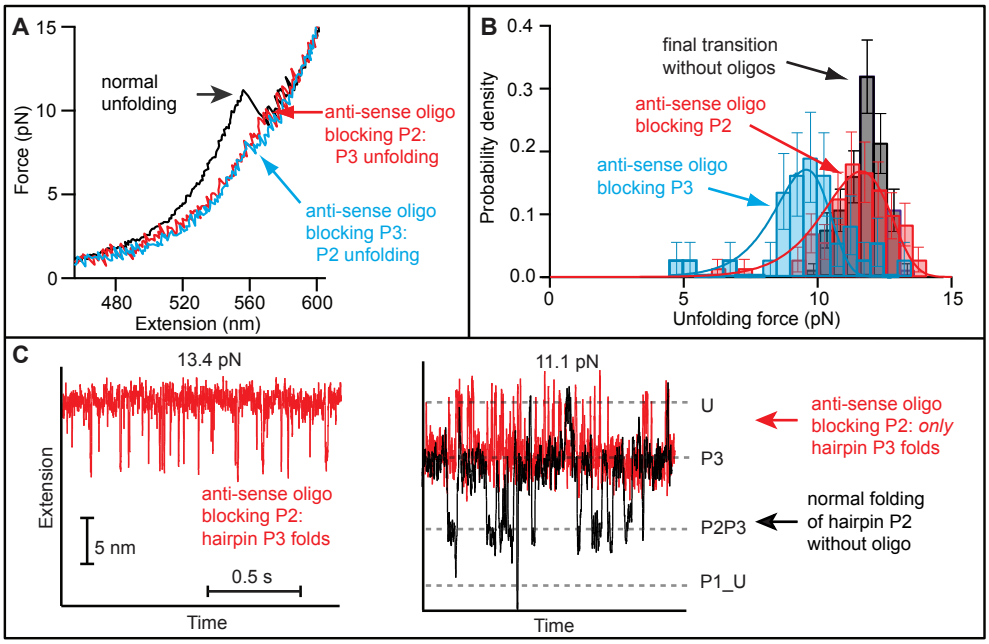

Figure S1. Confirmation of hairpin folding identification using anti-sense oligos. (A) FECs under normal conditions (black), with a DNA oligomer preventing formation of hairpin P2 (red), and with a DNA oligomer preventing formation of hairpin P3 (blue). The high-force unfolding transition under normal conditions matches the unfolding when P2 is blocked, when that the only part of the aptmaer that can fold is P3. (B) The unfolding force distribution for P2 (when P3 is blocked, blue) differs from the distribution for P3 (when P2 is blocked, red). The distribution of unfolding forces for the last transition seen without oligomers (black) matches the distribution for P3 unfolding. (C) Under constant force, again only the high force transition is seen when A DNA oligomer blocks the formation of P2 (red), confirming that this transition involves P3 folding.

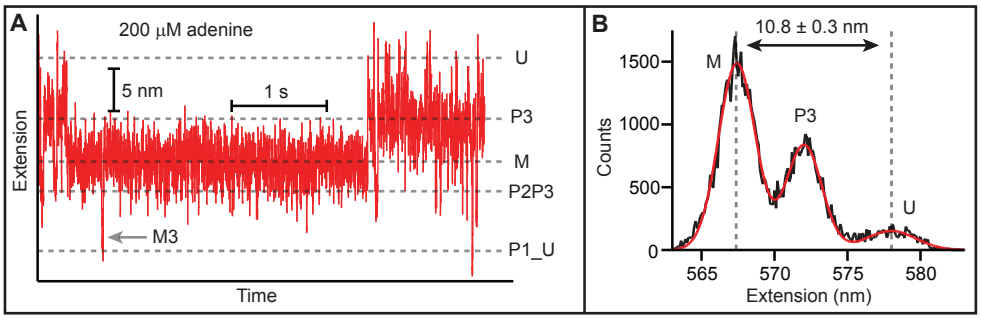

Figure S2. Misfolded states. (A) The same misfolded states were observed with adenine present as without adenine. (B) The extension change from the unfolded state (U) to the misfolded state (M) was determined from extension histograms.