

**Supporting Information For:**

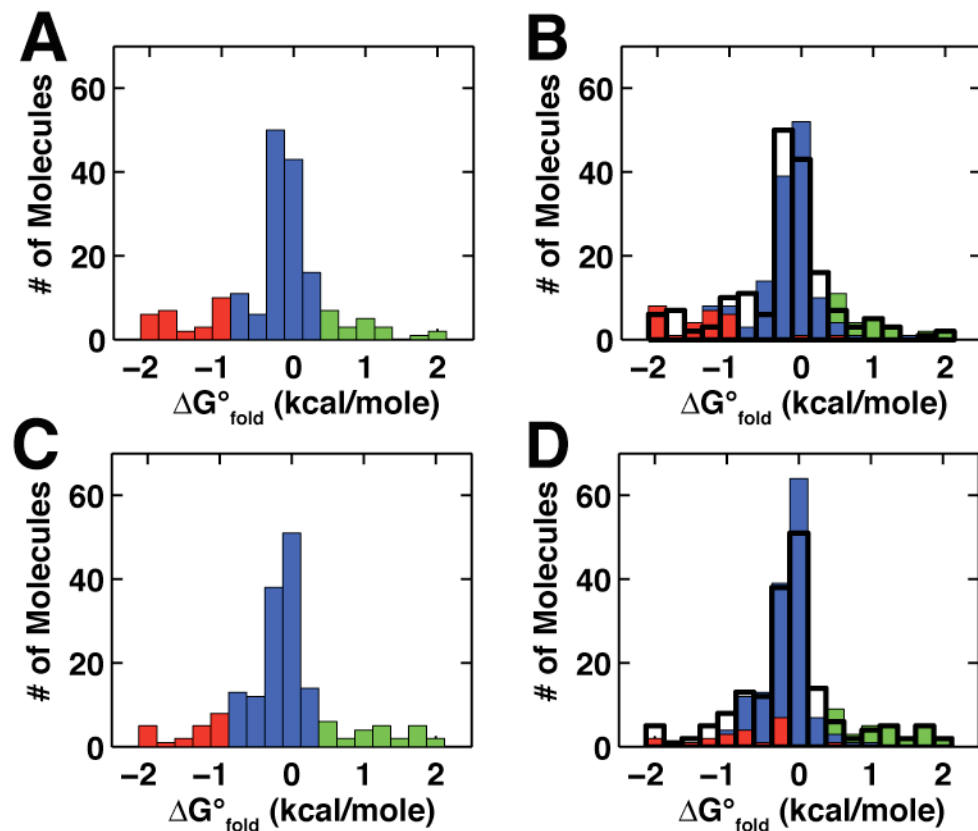
**Removal of Covalent Heterogeneity Reveals Simple Folding Behavior for P4-P6 RNA**

Max Greenfeld<sup>1,2</sup>, Sergey V. Solomatin<sup>2</sup>, and Daniel Herschlag<sup>1,2,3</sup>

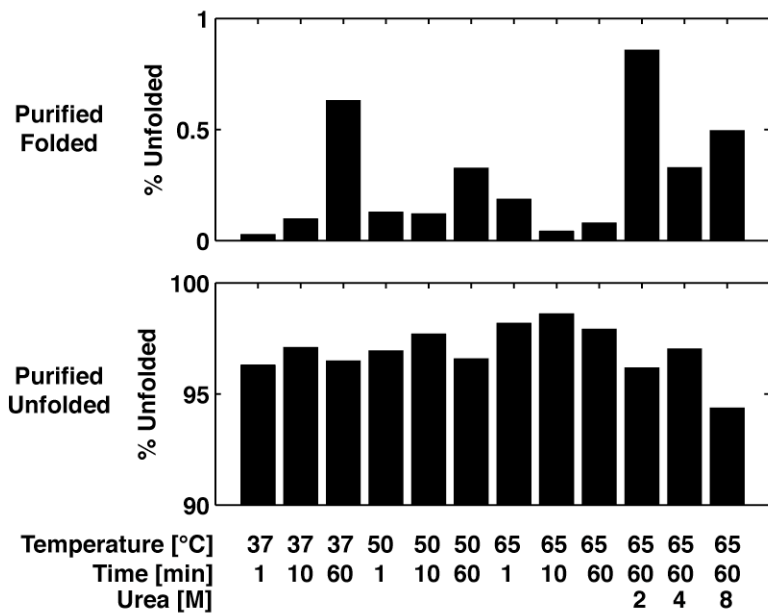
<sup>1</sup>Department of Chemical Engineering and <sup>2</sup>Department of Biochemistry, Stanford University,  
Stanford, CA 94305

<sup>3</sup>Corresponding author:  
Dept. of Biochemistry, B400  
Stanford University  
Stanford, CA, 94305-5307  
650-723-9442 (telephone)  
650-723-6783 (fax)  
herschla@stanford.edu

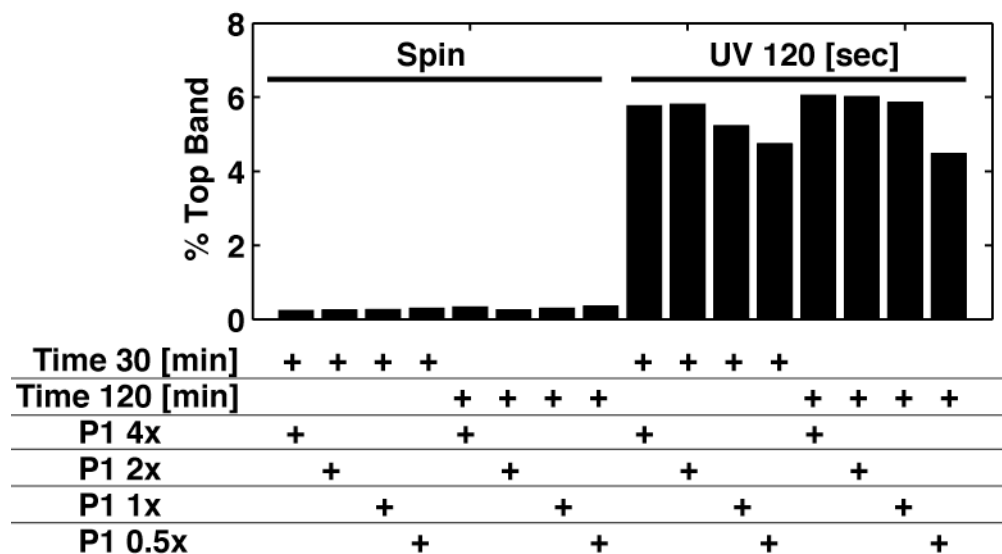
SUPPLEMENTAL FIGURES



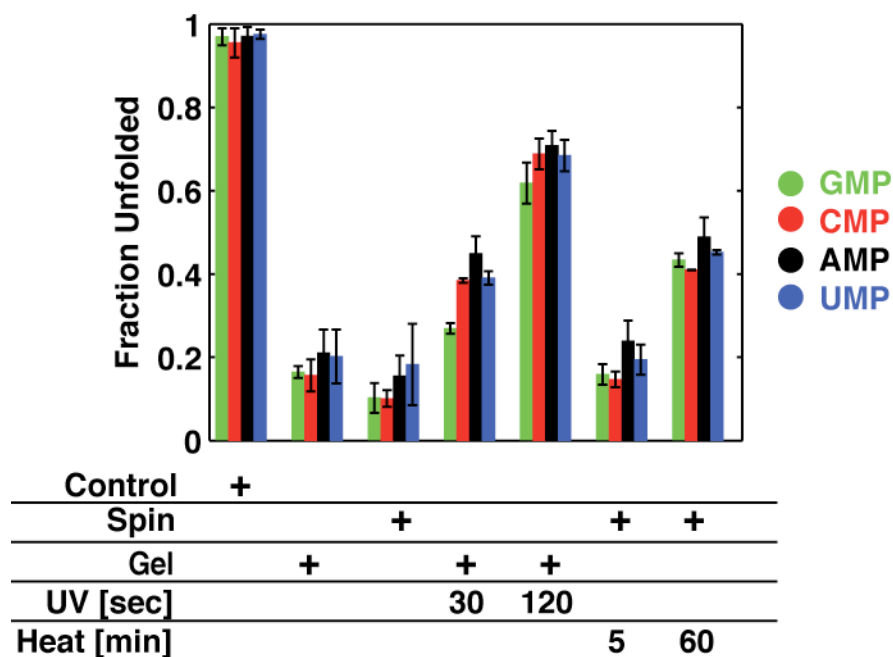
**Figure S1.** smP4-P6 states are persistent. (A and C) Histogram of free energies of folding of smP4-P6 (N = 175) in 15 mM  $\text{Ba}^{2+}$  observed initially for 1 minute. (B) Same molecules as in (A) imaged after the laser was turned off for 10 minutes. (D) Same molecules as in (C) after unfolding in 10 mM EDTA and 100 mM  $\text{Na}^{+}$  for 10 minutes and then refolding in 15 mM  $\text{Ba}^{2+}$  for 2 minutes prior to re-imaging. In (B) and (D) colors correspond to the initial positions of the molecules in (A) and (C), respectively, and the black outlines demarcate the initial distributions.



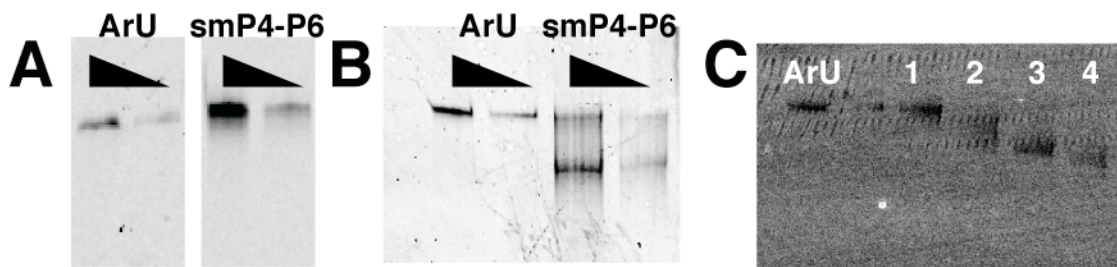
**Figure S2.** Folded and unfolded P4-P6 RNA purified by native gel (1 mM Mg<sup>2+</sup>) do not undergo significant exchange. Purified folded and purified unfolded P4-P6 were subjected to heat and urea denaturation with the goal of redistributing the purified material to the initially observed distributions, and each sample was rerun on a native gel with 1 mM Mg<sup>2+</sup>. Minimal redistribution of 0-3% between the populations was observed. The absence of significant exchange is consistent with covalent differences between the molecules.



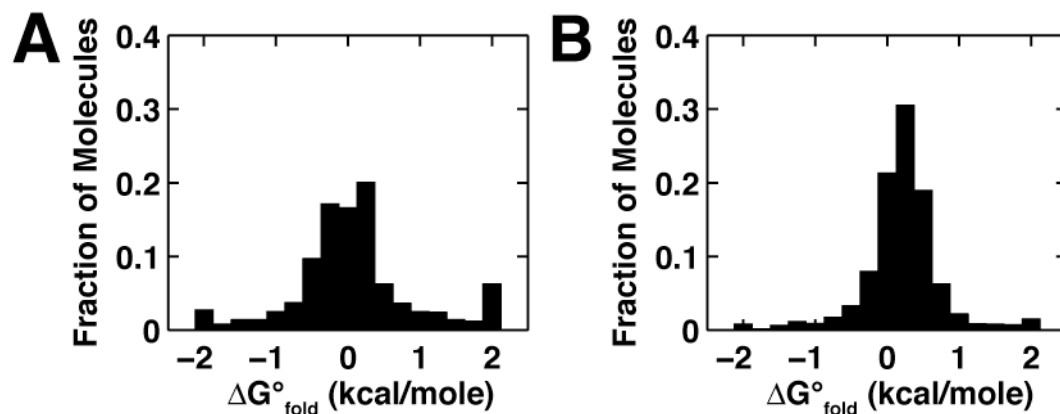
**Figure S3.** Control demonstrating complete digestion of RNA purified by spin column and denaturing PAGE plus 120 seconds of UV-shadowing purified [ $\alpha$ - $^{32}$ P]-UTP labeled P4-P6 is obtained under standard P1 nuclease reaction conditions involving a 30 minute incubation with 1x P1 (see Experimental Procedures). The amount of undigestible RNA did not vary substantially with variation in P1 concentration and incubation time, indicating complete digestion.



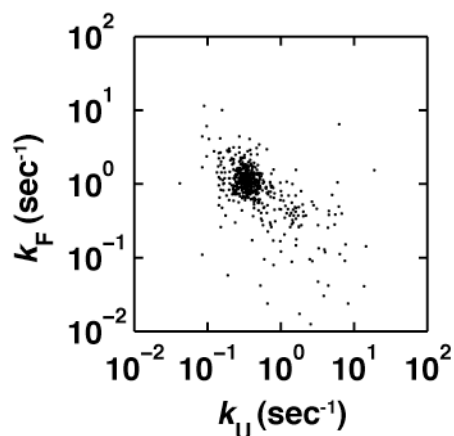
**Figure S4.** Heat and UV-treatment affect the amount of P4-P6 running as the unfolded band on a 1 mM Mg $^{2+}$  native gel. Quantification of the experiments described in Fig. 5C for P4-P6 transcribed with each [ $\alpha$ - $^{32}$ P]-labeled XTP (X = G, C, A, or U).



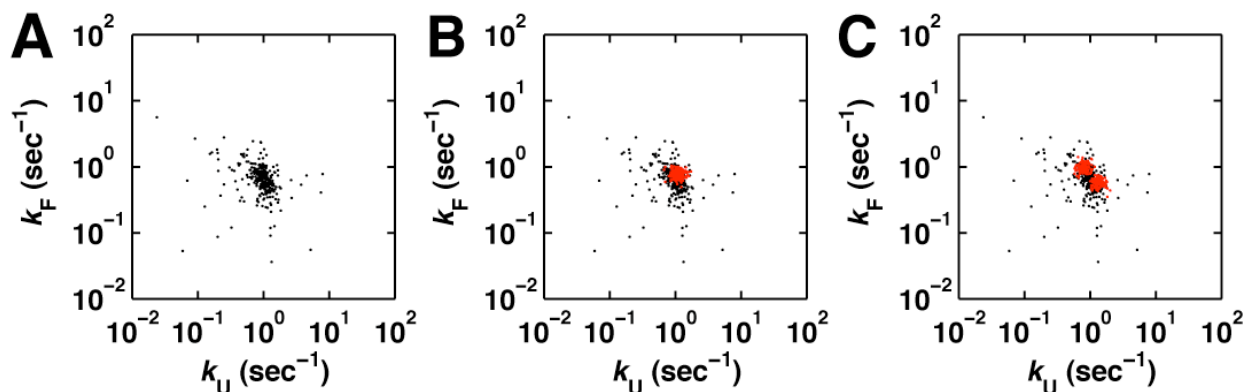
**Figure S5.** Purification of smP4-P6 by native gel. (A) smP4-P6 ligated from pieces that were not UV-shadowed and were annealed for 1 minute at 95 °C ran as one band on a 8% denaturing polyacrylamide gel. (B) This same RNA ran as two distinct bands on a 1 mM Mg<sup>2+</sup> native gel. The RNA was purified by native gel for smFRET experiments. ArU is an unfolded control (see main text). (C) Native gel of purified fractions of smP4-P6. Four fractions of smP4-P6 were collected. Intensities of fractions in the gel shown do not reflect the actual abundance of the species in the unpurified material. The major fractions (labeled 1 and 3), corresponding to the two bands seen in (B), were analyzed using smFRET. For gels with two lanes containing the same sample the amount loaded was 1x and 3x.



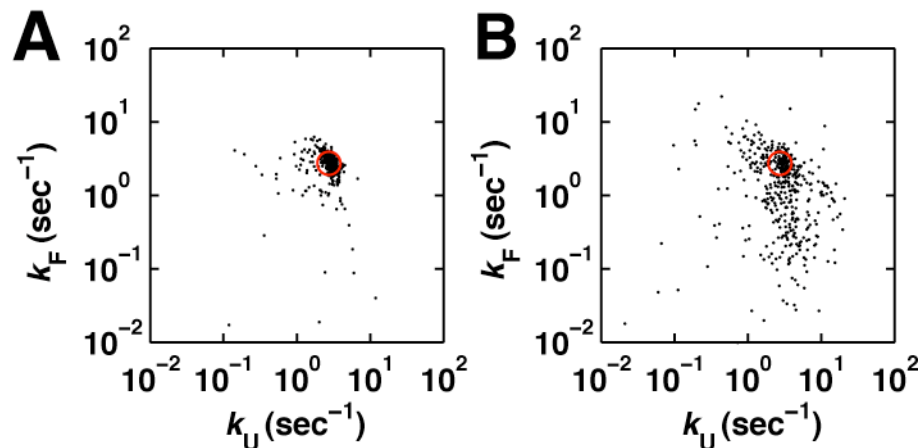
**Figure S6.** Comparison of folding free energies for individual smP4-P6 molecules prior to native gel purification (A) and subsequent to native gel purification (B) for RNA prepared without UV-shadowing. The experiment was carried out in 2.5 mM Mg<sup>2+</sup>. Average values of  $\Delta G^{\circ}_{\text{fold}}$  of  $0.10 \pm 0.94$  kcal/mol (N = 987) and  $0.20 \pm 0.59$  kcal/mol (N = 1953) were obtained for (A) and (B) respectively. Histogram shown in (B) is the same as in Fig. 7A, but the ordinate scale is changed and simulations are not included.



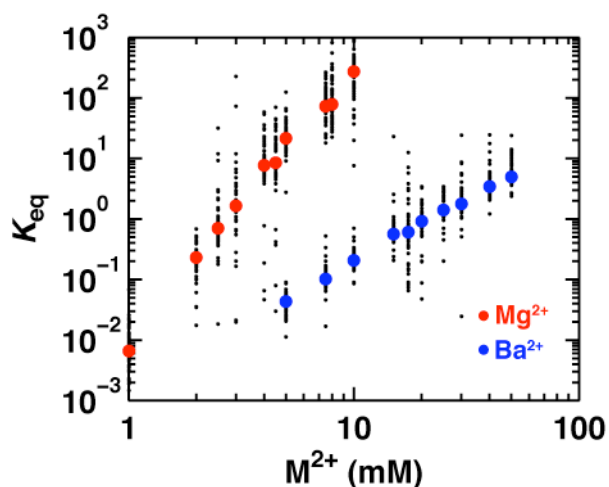
**Figure S7.** Kinetics of the initial smP4-P6 preparation (prepared with UV-shadowing, 5 minute heat treatments, and not native gel purified) in 2.5 mM  $Mg^{2+}$ . All the data taken ( $N = 566$ ) are included (compared with  $N = 300$  in Fig. 2C).



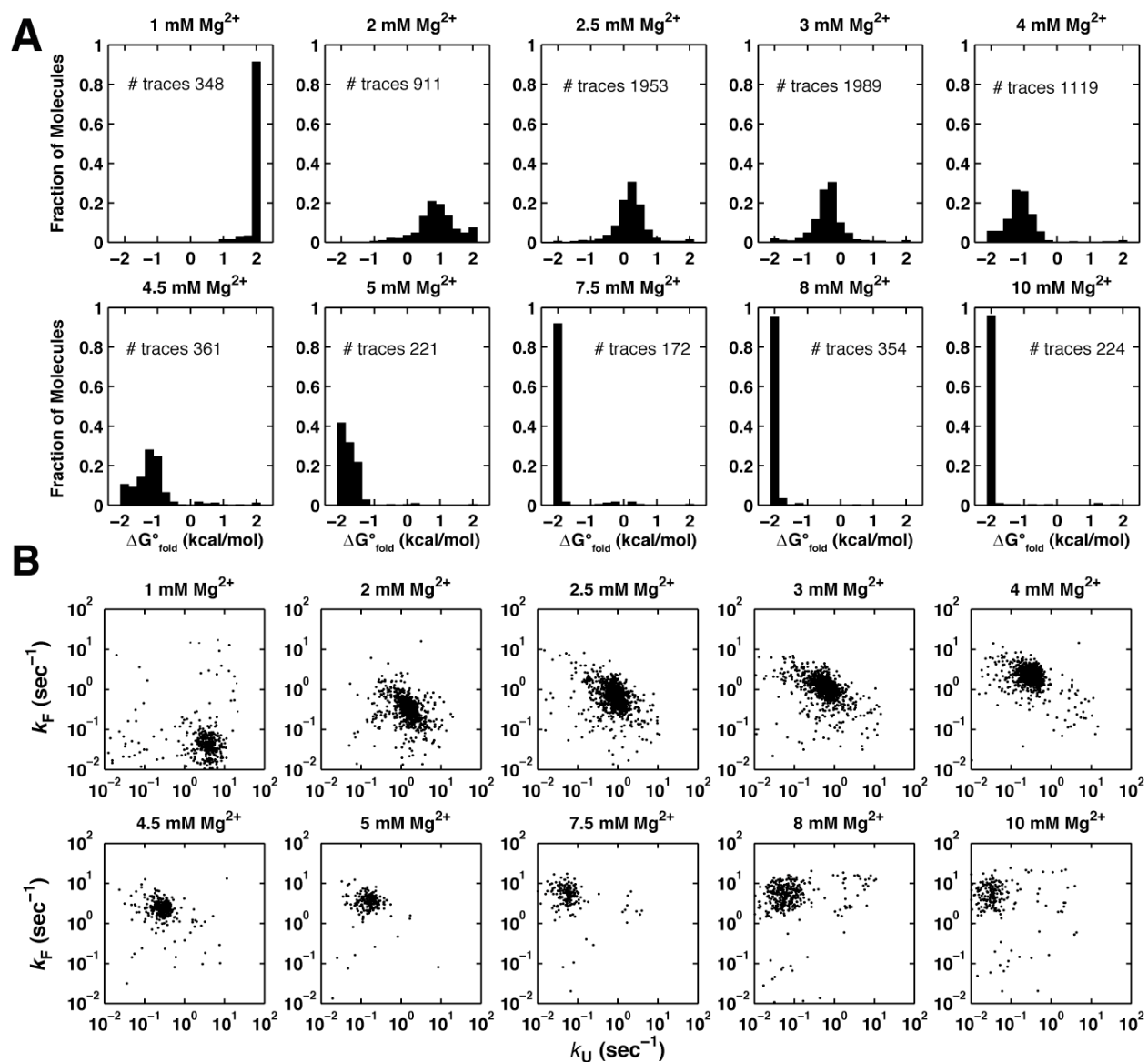
**Figure S8.** Comparison of observed and simulated kinetics at 2.5 mM  $Mg^{2+}$  for individual smP4-P6 molecules prepared without UV-shadowing and purified by native gel. (A) The kinetics of 300 out of a total of 1953 smP4-P6 traces; these same data are repeated in (B) and (C) (black points). (B) Simulated kinetics (red) from traces ( $N = 300$ ) obtained using mean measured unfolding and folding rate constants ( $k_F = 0.75 \text{ sec}^{-1}$  and  $k_U = 1.06 \text{ sec}^{-1}$ ) and noise equal to that observed in experiments (see Experimental Procedures). In this comparison,  $\sim 45\%$  of smP4-P6 molecules lie in the region covered by the simulated traces. (C) Simulated kinetics (red) from traces ( $N = 300$ ) obtained assuming two non-exchanging populations of molecules (Population 1:  $k_F = 0.95 \text{ sec}^{-1}$  and  $k_U = 0.80 \text{ sec}^{-1}$  and Population 2:  $k_F = 0.57 \text{ sec}^{-1}$  and  $k_U = 1.35 \text{ sec}^{-1}$ ). In this comparison  $\sim 80\%$  of smP4-P6 molecules lie in the region covered by the simulated traces. This comparison demonstrates smP4-P6 molecules are largely described by rate constants within a two-fold range.



**Figure S9.** Kinetics of purified populations of smP4-P6 in 20 mM  $\text{Ba}^{2+}$  ( $N = 441$ ). Kinetics of purified smP4-P6 cluster tightly (A), whereas those of damaged smP4-P6 (B) show more heterogeneity. The red circle is drawn to guide the eye to the regions where the folding behavior of molecules are identical in the two populations of molecules (this circle is not a simulation). The large regions of overlapping molecules suggests that native gel purification in 1 mM  $\text{Mg}^{2+}$  separates some smP4-P6 molecules based on properties different than those that determine folding in  $\text{Ba}^{2+}$ . Covalent modifications in the MC/MCR tertiary motif could account for this effect.

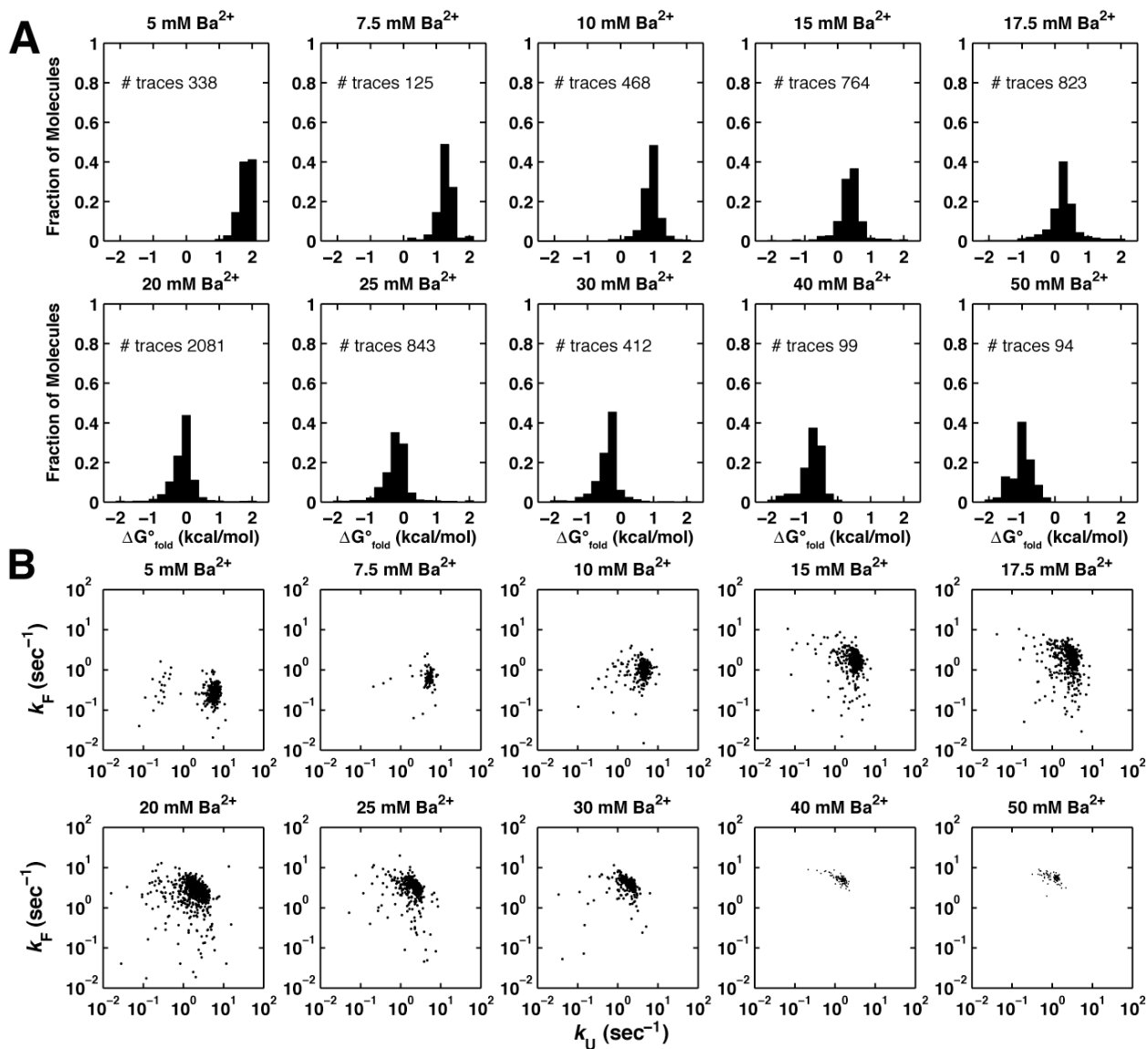


**Figure S10.** Titration comparing folding equilibrium constants of smP4-P6 in  $\text{Mg}^{2+}$  (red) and  $\text{Ba}^{2+}$  (blue) (all buffers contained 100 mM  $\text{Na}^+$ ). Large circles represent the median of all molecules analyzed at the indicated ion concentration; smaller black circles represent 50 randomly chosen molecules presented to convey the spread in the data.

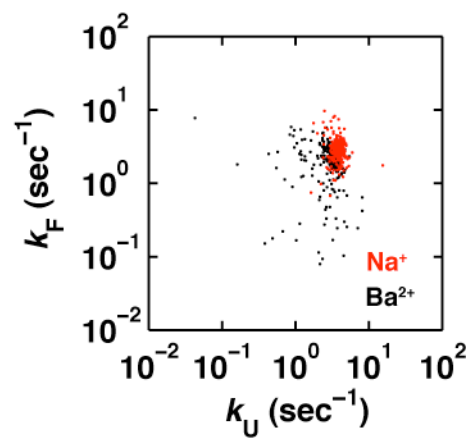


**Figure S11.** Summary of all the thermodynamic (A) and kinetic (B) data for purified smP4-P6 in  $\text{Mg}^{2+}$ .





**Figure S12.** Summary of all the thermodynamic (A) and kinetic (B) data for purified smP4-P6 in  $\text{Ba}^{2+}$ .



**Figure S13.** Kinetics of smP4-P6 (N=300) in 17.5 mM  $\text{Ba}^{2+}$ (black) and in 800 mM  $\text{Na}^+$  (red). Ion concentrations were chosen to be near the folding midpoint in  $\text{Ba}^{2+}$  and  $\text{Na}^+$ .

## EXPERIMENTAL PROCEDURES

### Reagents: RNA Preparation

His-tagged T7 RNA polymerase and T4 DNA ligase were overexpressed and purified in-house. T4 polynucleotide kinase (T4 PNK), shrimp alkaline phosphatase (SAP) and bovine serum albumin (BSA) were purchased from New England BioLabs (MA). P1 Nuclease from *Penicillium citrinum* was purchased from Sigma-Aldrich (MO). Synthetic DNA oligonucleotides were purchased from Integrated DNA Technologies (IA) and synthetic RNA oligonucleotides were purchased from Dharmacon (CO). NTPs (Sigma-Aldrich, MO) (95-99% purity), dNTPs (Fermentas, MD), and guanosine 5'-monophosphate (Sigma-Aldrich, MO) (>99% purity) were used for *in vitro* transcription reactions or PCR. [ $\alpha$ - $^{32}$ P]-labeled NTPs (6000 Ci/mmol) were from MP Biomedical (OH). All other reagents were of standard molecular biology grade or higher. Ammonium persulfate (APS), N-cyclohexyl-2-aminoethanesulfonic acid (CHES), dithiothreitol (DTT), imidazole, magnesium chloride, manganese(II) chloride, 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS), hydrogen peroxide, 2-mercaptoethanol, sodium hydroxide, spermidine, tetramethylethylenediamine (Temed), tris(hydroxymethyl)aminomethane (Tris) and tris-saturated phenol were from Sigma-Aldrich (MO). Boric acid, chloroform, hydrochloric acid, OmniPure 40% acrylamide solutions 19:1 acrylamide to bisacrylamide, sodium acetate, sodium chloride, sulfuric acid, and urea were from EMD Chemicals (NJ). Ethanol was from Gold Shield (CA). Ethylenediaminetetraacetic acid (EDTA) and potassium hydroxide were from J.T. Baker (NJ).

### Reagents: Single Molecule

Chemical reagents used in single molecule imaging were of the highest purity commercially available (>99% with trace metal grade used when available) from Sigma-Aldrich (MO), including 3-(N-morpholino)propanesulfonic acid (MOPS), sodium chloride, magnesium chloride, barium chloride, (+)-6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox) and D-glucose. An oxygen scavenging system of Glucose oxidase Type VII from *Aspergillus niger* (Sigma-Aldrich, MO) and Catalase (Roche) was employed. Streptavidin and biotinylated bovine serum albumin (bBSA) (Sigma-Aldrich, MO) were used for surface immobilization.

**Table S1.** Sequences of DNA and RNA used.

P4-P6 WT	GG AAT TGC GGG AAA GGG GTC AAC AGC CGT TCA GTA CCA AGT CTC AGG GGA AAC TTT GAG ATG GCC TTG CAA AGG GTA TGG TAA TAA GCT GAC GGA CAT GGT CCT AAC CAC GCA GCC AAG TCC TAA GTC AAC AGA TCT TCT GTT GAT ATG GAT GCA GTT CA
P4-P6 ArU (Unfolded Control)	GG AAT TGC GGG AAA GGG GTC AAC AGC CGT TCA GTA CCA AGT CTC AGG GGA AAC TTT GAG ATG GCC TTG CAA AGG GTA TGG <b>TTT TTT TCT</b> GAC GGA CAT GGT CCT AAC CAC GCA GCC AAG TCC TAA GTC AAC AGA TCT TCT GTT GAT ATG GAT GCA GTT CA
GG104-148+DNAzI	GGA ATT GCG GGA AAG GGG TCA ACA GCC GTT CAG TAC CAA GTC TCA GGT GGG TGC GAG
163-234+DNAzI	GGC CTT GCA AAG GGT ATG GTA ATA AGC TGA CGG ACA TGG TCC TAA CCA CGC AGC CAA GTC CTA AGT CAA CAG TGG GTG CGA G
250-261+T2 DNAzI	GGA TGC AGT TCA ACC AAA ATC AAC CTA AAA CTT ACA CAT GGG TGC GAG
Splint 148/149 - 162/163	AC CAT ACC CTT TGC AAG GCC ATC TCA AAG TTT CCC CTG AGA CTT GGT ACT
Splint 234/235 - 249/250	TGA ACT GCA TCC ATA TCA ACA GAA GAT CTG TTG ACT TAG GAC
DNAzI-148	CGC ACC CAG GCT AGC TAC AAC GAC TGA GAC T
DNAzI-234	CGC ACC CAG GCT AGC TAC AAC GAT GTT GAC TT
DNAzI-T2	CGC ACC CAG GCT AGC TAC AAC GAG TGT AAG T
T2b	biotin - TGT GTA AGT TTT AGG TTG ATT TTG GT
149-162 U155	GGA AAC (5-N-U)UU GAG AU
235-249 U241	AUC UUC (5-N-U)GU UGA UAU
T7 Promoter	CCA AGT AAT ACG ACT CAC TAT A

## Native Gel Assays and P1 Nuclease Digestions

P4-P6 was transcribed from a double stranded PCR-amplified template consisting of a T7 polymerase promoter and the P4-P6 sequence (Table S1). Final transcription conditions were 25 mM MgCl<sub>2</sub>, 40 mM Tris-HCl pH 8.3, 30 mM DTT, 2 mM spermidine, 0.0001% Triton X-100, 2.5 mM of each nucleotide triphosphate (NTP), 100 nM of DNA template, T7 RNA polymerase and 20 pmoles of either [ $\alpha$ -<sup>32</sup>P]-labeled ATP, UTP CTP or GTP (6000 Ci/mmmole). Transcription reactions were allowed to proceed at 37 °C for three hours before purification and were then purified in three ways.

One-third of the material underwent six rounds of 10X concentration followed by dilution in 500  $\mu$ L of water on a Microcon™ YM-30 (Millipore, MA). One-third was purified on an 8% PAGE gel, with 7 M urea and 1X TBE (100 mM Tris, 83 mM boric acid and 1 mM EDTA running buffer (all gels described are polymerized with 1/100 freshly made 10% APS and 1/1000 Temed). The RNA was visualized using autoradiography to expose film and excised using sterile scalpels. Water (200  $\mu$ L) was added to the excised samples before it underwent three freeze thaws and overnight elution. Finally one-third was purified by denaturing PAGE and extracted under the same conditions as described above except with the addition of a UV-shadowing step prior to excising the bands. In this step a 6 Watt, 252 nm UV light (UVP, CA) was held approximately 5 cm over the sample for 30 or 120 seconds. All samples were diluted to an activity of 50,000 CPM/ $\mu$ L.

In all analytical gels ~10,000 CPM of sample was loaded per lane. For native and denaturing gels, samples were diluted in a final volume of 10  $\mu$ L of 25 mM Na-MOPS, pH 7.0, and 5 mM EDTA buffer prior to treatment. Heat treatments at 95 °C were carried out using a standard laboratory thermo cycler. For the 5 min and 60 min incubation at 95 °C only the time spent at 95 °C was varied. Nuclease digests were carried out in 100 mM Na-acetate, pH 5.3, and 10 mM ZnCl<sub>2</sub> with 1  $\mu$ g (equivalent to 1X in Fig. S3) of P1 nuclease for two hours at 50 °C in a 12  $\mu$ L reaction volume.

Denaturing gels to test length homogeneity were 10% polyacrylamide, with 7 M urea and 1X TBE. Native gels were 10% polyacrylamide, 1X TB (100 mM Tris, 83 mM boric acid) and 1 mM MgCl<sub>2</sub>. Native gels were run at 20 W and maintained at ~4 °C with the use of liquid-cooled gel boxes. Denaturing PAGE gels of P1 nuclease-digested samples were 20% polyacrylamide, with 7 M urea and 1X TBE.

Autoradiography was carried out by exposing storage phosphor screens (GE Healthcare, NJ) of dried gels. Two time points were taken to ensure measurements were within the dynamic range of the screens. Storage phosphor screens were imaged on a Typhoon™ 9400 (GE Healthcare, NJ). Densitometry was completed using TotalLab Quant (TotalLab Limited, Newcastle, UK).

### smP4-P6 Assembly

smP4-P6 was assembled using splinted ligations (1-3) from five RNA pieces (Table S1). Two pieces were synthetic RNA oligonucleotides (149-162 U155 and 235-249 U241), with 5-amino-allyl-uridine site-specifically incorporated at positions U155 and U241. These samples were

labeled with cy3 and cy5 using post-labeling reactive dye packets (GE Healthcare, NJ) using protocols described elsewhere (2). Three pieces of smP4-P6 (GG104-148, 163-234 and 250-261) were *in vitro* transcribed from PCR-amplified templates (GG104-148+DNAzI, 163-234+DNAzI and 250-261+T2 DNAzI, Table S1) that contained the T7 promoter, P4-P6 sequence, and 10 nucleotide tail necessary for the DNAzyme reaction (4). Final transcription conditions were 17-25 mM MgCl<sub>2</sub>, 40 mM Tris-HCl, pH 8.3, 30 mM DTT, 2 mM spermidine, 0.0001% Triton X-100, 0.8 mM of each NTP, 100 nM of DNA template, T7 RNA polymerase and 3 mM GMP for the 163-234+DNAzI and 250-261+T2+DNAzI pieces. Transcription reactions were carried out for two hours at 37 °C. The next steps varied between the UV-shadowed and non-UV-shadowed smP4-P6 preparations.

1. In the initial preps of smP4-P6 transcription reactions were terminated with ethanol precipitation and were purified by denaturing PAGE, and UV-shadowing was used to visualize the RNA. Purified samples were then DNAzyme-treated by incubating 10 μM RNA for two hours at 37 °C in 10 mM MnCl<sub>2</sub>, 150 mM NaCl, 100 mM EPPS, pH 8.3, and 1 μM complementary DNAzyme (Table S1). The DNAzyme reaction was terminated by ethanol precipitation of the RNA and the RNA was purified by denaturing PAGE with UV-shadowing to visualize the RNA.
2. For non-UV treated samples the complementary DNAzyme (Table S1) was added to the transcription reaction to a final concentration of 5 μM and the DNAzyme reaction was allowed to proceed for an additional two hours before the reaction was stopped and concentrated by ethanol precipitation. Samples were desalted on PD-10 columns (GE Healthcare, NJ) and concentrated again using Microcon™ YM-10 (Millipore, MA) columns.

The 2'-3' cyclic phosphate left after DNAzyme treatment was removed by incubating transcribed RNA in 90 mM imidazole, pH 6.5, 9 mM 2-mercaptoethanol, 22 μg/mL BSA, 9 mM MgCl<sub>2</sub>, and 1.4 U/μL of T4 PNK for 10 hours at 37 °C. The reaction was stopped by the addition of 1 M CHES buffer, pH 11, to a final concentration of 50 mM followed by phenol chloroform extraction. Ethanol precipitation was used for final desalting and concentration.

Ligation reactions were carried out in two parts. The first step involved combining the five RNA pieces (GG104-148, 149-162-U155, 163-234, 235-249-U241 and 250-261+T2) and two DNA splints (Splint 148/149 - 162/163 and Splint 234/235 - 249/250) at a putative concentration of ~0.4 μM in 1 mM Tris, pH 8.0, and 0.1 mM EDTA. Initial preparations of smP4-P6 with UV-shadowing were then annealed at 95 °C for 5 min. For subsequent preparations of smP4-P6 the time spent at 95 °C was decreased to 1 min. This annealed mixture was then diluted four fold into the final ligation reaction conditions of 66 mM Tris, pH 7.6, 8 mM MgCl<sub>2</sub>, 10 mM DTT, 1.3 mM ATP, 0.004% Triton-X 100, RNasin 8 U/mL and T4 DNA ligase.

The ligation reaction was allowed to proceed at 37 °C for 18 hours. The reaction was stopped and the RNA was concentrated by ethanol precipitation. This reaction mixture was then separated on an 8% denaturing PAGE gel. Ligated bands were visualized on a Typhoon™ 9400 (GE Healthcare, NJ) using cy3 and cy5 fluorescence. Bands were excised using sterile scalpels. The resulting gel pieces underwent three freeze-thaw cycles and were eluted into water. smP4-P6

prepared in this fashion ran as one band on a denaturing PAGE gel (Fig. S5A). An aliquot of the eluted sample (~10%) was set aside for analysis without further treatment. Depending on the ligation preparation, between 20% and 40% of this RNA ran as unfolded material on a 1 mM Mg<sup>2+</sup> native gel. The remainder of the denaturing PAGE purified sample was speedvaced to a volume sufficient to be loaded onto a 1 mM Mg<sup>2+</sup> native gel. This RNA ran mainly as two bands (Fig. S5 B), but regions between the two main bands and slightly below the primary fastest running band were also excised. Excised bands were eluted into water. All excised bands when rerun on a 1 mM Mg<sup>2+</sup> (Fig. S5 C) native gel maintained positions, which corresponded to the initial purification native gel.

### **Preparations of smP4-P6 for Imaging**

Purified smP4-P6 was annealed to the DNA oligonucleotide T2b (Table S1) by incubating eluted smP4-P6 with 2 nM T2b, 100 mM Na-MOPS, pH 7.0, and 1 mM EDTA at 50 °C for 15 minutes. This stock is used over multiple days of experiments. Replicate measurements taken over multiple days generated with the same stock do not show significant variation.

Molecules were imaged at the surface of a quartz slide (G. Finkerbeiner). Slides were drilled to create access ports to the imaging flow chamber. Slides were used multiple times and were cleaned before storage in water with piranha solutions (sulfuric acid and hydrogen peroxide) and then etched for 30 minutes using potassium hydroxide and sonication. A ~5  $\mu$ L flow chamber, accessible by the drilled ports, was created between a No. 1 glass cover slip (VWR, PA) and the slide using double stick tape (3M, MN) and 5 minute epoxy (Devcon, MA).

Flow chambers were first washed with 100  $\mu$ L buffers (all washes were carried out with 50 mM MOPS, pH 7.0, and 100 mM NaCl). This washing was followed by the addition of 10  $\mu$ L of 1 mg/mL of bBSA for a 10 minute incubation. The chamber was then washed with 100  $\mu$ L of buffer, followed by the addition of 10  $\mu$ L of 0.1 mg/mL of streptavidin for a 10 minute incubation. The chamber was washed again with 100  $\mu$ L of buffer before depositing smP4-P6 molecules. Typically a five minute incubation with 5  $\mu$ L of a 10X to 100X dilution of the T2b annealed smP4-P6 stock was sufficient to deposit smP4-P6 at a surface density such that 50-100 molecules are present in a typical microscope field of view. After deposition, the chamber was washed again with 100  $\mu$ L of buffer before adding 50  $\mu$ L of smFRET buffers. Buffers used in smFRET measurements contained 1-10 mM MgCl<sub>2</sub>, 5-50 mM BaCl<sub>2</sub>, or 700 mM additional NaCl in a background of 50 mM MOPS (pH 7.0), 100 mM NaCl, 2 mg/ml glucose, 1.8 mM Trolox, 100 U/ml glucose oxidase, 1000 U/ml catalase.

### **TIRF Microscope and Single Molecule Data Analysis**

Slides were mounted in an inverted Nikon Eclipse TE300 (JP) microscope. The microscope was customized so that molecules could be illuminated using prism-based total-internal-reflection (5,6). A dichroic mirror was used to combine a 532 nm green laser (Laser Quantum, UK) and a 635 nm red laser (Hitachi, JP) that were focused through the prism on to the quartz slide. Images were collected through a 60X water-immersion objective with a 1.2 numerical aperture (Nikon, JP). Images were passed through a 550-nm long pass filter to eliminate scattered light and separated into images corresponding to donor (cy3/green) and acceptor (cy5/red) emissions using

band-pass filters. Each image was focused onto half of an electron multiplying chard-coupled device (EMCCD) of a Cascade 128+ (Photometrics, AZ) camera, which has a resolution 128×128 pixels. Movies were collected with the EMCCD set to have a conversion gain of 3, multiplication gain of 3,200 and acquisition time of 40 ms. Green laser intensity was adjusted such that the average signal-to-noise ratio was ~6.

Movies were analyzed first by running an algorithm that examines the first 30 to 50 frames of the acceptor image looking for pixels that exceeded the background typically by a threshold of  $5\sigma$ , where  $\sigma$  is the standard deviation of the background fluorescence. The corresponding image in the donor channel was found by applying linear offsets to points in the acceptor channel, which were determined by imaging fluorescent beads that are visible in both channels. For acceptor channel pixels that pass the threshold and the corresponding donor channel, the local background (7×7 pixel) was subtracted and the background-subtracted intensities were saved as a putative single molecule trace.

Discontinuous experiments were completed in a similar fashion to those describe by Solomatin *et al.* (5). Traces were extracted from movies generated during the initial observation and the latter observation. Traces were identified as appearing in both observations if they were within 1 pixel in both movies. In some experiments linear offsets of 1-2 pixels were applied to all putative traces in the second movie before the final alignment, this translation likely arises from drift in the microscope stage over the course of the experiment.

Standard criteria for the identification of a single FRET labeled molecule are as follows: (1) single step photobleaching; (2) anticorrelated donor and acceptor channels; (3) total intensity consistent with a single molecule; and (4) stable total intensity. All putative traces were manually screened using these criteria to select a subset as arising from smP4-P6 molecules. Identified traces were then analyzed to extract thermodynamic and kinetic parameters.

To analyze the thermodynamics of individual molecules FRET was calculated using the recorded donor and acceptor intensity according to equation 1

$$FRET = \frac{I_{red} - I_{cross-talk}}{I_{green} + I_{red} - I_{cross-talk}} \quad (1)$$

where  $I_{red}$  is the intensity in the acceptor channel,  $I_{green}$  is the intensity in the donor channel and  $I_{cross-talk}$  is the intensity in the acceptor channel due to imperfect separation of the donor channel by the dichroic mirrors. The resulting FRET traces were thresholded at a value of 0.45. The total time spent in the high and low FRET states was used to calculate equilibrium constants (which in turn are used for calculating  $\Delta G^{\circ}_{fold}$ ) and fractional time spent folded. Good agreement was observed between equilibrium constants calculated with thresholding and equilibrium constants calculated using the extracted kinetic parameters (see below) of traces.

Kinetic analysis of traces was carried out by fitting the donor and acceptor channels to a two state Hidden Markov Model (HMM). This analysis was accomplished through maximum likelihood estimation using the Baum-Welch algorithm (7,8). A two state kinetic model was fit for transitions between folded and unfolded states ( $k_F$  and  $k_U$ ), providing two fitted parameters. Unfolded and folded states in the donor and acceptor channels were each fit with two Gaussian distributions (each with a distinct mean intensity and standard deviation), this resulted in eight



additional fitting parameters. Fitting and data analysis was accomplished using software written in-house and was implemented in Matlab™ 2008b.

Variation in measured thermodynamic and kinetic parameters can arise from intrinsic differences among the molecules or from noise inherent in the measurement. To gauge these differences we turned to simulation where simulated traces are generated with the mean folding and unfolding rates of all traces measured. To estimate the noise inherent in the measurements due to variation in trace length and from signal-to-noise, a simulated trace was generated for each measured trace that has a similar signal-to-noise and identical trace length to a molecule in the measured distribution. To accomplish this, the simulated trace was generated using as an input parameter the mean folding and unfolding rates for all traces, the fitted donor and acceptor channel parameters for a measured trace and measured trace length. Simulated traces were then subjected to the same thermodynamic and kinetic analysis of measured traces.

### SUPPLEMENTAL REFERENCES

1. Moore, M. J., and Sharp, P. A. (1992) *Science* **256**, 992-997
2. Solomatin, S., and Herschlag, D. (2009) *Methods Enzymol.* **469**, 47-68
3. Akiyama, B. M., and Stone, M. D. (2009) *Methods Enzymol.* **469**, 27-46
4. Silverman, S. K., and Baum, D. A. (2009) *Methods Enzymol.* **469**, 95-117
5. Solomatin, S. V., Greenfeld, M., Chu, S., and Herschlag, D. (2010) *Nature* **463**, 681-684
6. Sattin, B. D., Zhao, W., Travers, K., Chu, S., and Herschlag, D. (2008) *J. Am. Chem. Soc.* **130**, 6085-6087
7. Welch, L. R. (2003) *IEEE Information Theory Society Newsletter* **53**
8. Cappé, O., Moulines, E., and Rydén, T. (2005) *Inference in Hidden Markov Models*, Springer Science & Business Media, New York