

SUPPORTING INFORMATION

Single-Molecule Imaging of an *In Vitro* Evolved RNA Aptamer Reveals Homogeneous Ligand Binding Kinetics

Mark P. Elenko[‡], Jack W. Szostak[†], and Antoine M. van Oijen^{*‡}

Program in Biological and Biomedical Sciences[‡], Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, and Howard Hughes Medical Institute and Center for Computational and Integrative Biology[†], Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School[‡], 240 Longwood Avenue, Boston, Massachusetts 02115

Materials and Methods

Primary Reagents: DNA sequences were synthesized by the Massachusetts General Hospital DNA Synthesis Core Facility and the W.M. Keck Facility at Yale University. DNA templates were amplified by PCR, and transcribed *in vitro* with T7 RNA Polymerase. The Class V aptamer previously found by *in vitro* selection and optimized¹ (68 nt long) was modified by the addition of a 3' A₂₀ tail to produce the RNA sequence:

5' -GG GGG CAT TTT GGT AGG TCG GTC GCT GCT TCG GCA GTG AGG GGT AGG
CAT TGC TGG CCT AGG GTC CCC AAA AAA AAA AAA AAA AAA AA-3'

RNA was purified by PAGE followed by ethanol precipitation. A commercial dT₂₀ 5'-biotinylated oligonucleotide (Roche) was used to couple folded RNA to the functionalized surface.

The buffer used throughout was based on the original aptamer selection^{1,2} conditions plus BSA (NEB) to block nonspecific binding: 200 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 0.1 mg/ml BSA, 0.001% Tween-20, 10 mM potassium phosphate pH 6.2. A vacuum desiccator was used to degas the buffer before use. The imaging buffer also included the oxygen scavenging system components: 50 nM protocatechuate 3,4-dioxygenase (Sigma), 2.5 mM 3,4-dihydroxybenzoic acid (protocatechuic acid) (Sigma), and 1 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) (Sigma).

Labeled GTP was produced by reacting a five-fold excess of guanosine-5'-[γ -thio]triphosphate (Sigma) with tetramethylrhodamine-5-iodoacetamide dihydroiodide (Molecular Probes/Invitrogen) in PBS pH 7.2 for 2 hours at room temperature and gel purifying to yield the GTP-TMR. Gels were imaged using a fluorescence scanner (GE Storm) and reaction product bands were excised. Formation of the GTP-TMR product was verified by electrospray ionization mass spectrometry (ES-MS) (Bruker Esquire 3000 Plus and Bruker Esquire 6000). Analogue studies of the Class V aptamer, and the other GTP aptamers found in the same selections, indicated that modifications, including deletion, of the GTP β and γ phosphates did not significantly change the binding affinity³. Surprisingly, the K_d of the GTP-TMR conjugate was approximately 30-fold tighter than the K_d of GTP, which we speculate may reflect the stacking of the TMR moiety on nearby nucleotides and/or electrostatic interactions of the negatively charged RNA with the positively charged dye.

Fluorescent carboxylated microspheres (0.89 μ m Flash Red, Bangs Labs) were biotinylated by coupling to biotinamidocaproyl labeled BSA (Sigma) using EDAC (Invitrogen) and sulfo-NHS (Pierce), based on the supplier's protocol.

Slide Functionalization and Flow Cell Construction: Detailed methods have been published previously⁴. Glass coverslips (VWR) were cleaned, silanized with 3-aminopropyl triethoxysilane (Sigma), functionalized with 5 kD PEG (Nektar), 1-2% of which was biotinylated, and kept under vacuum. Functionalized coverslips were incubated with 1 mg/ml streptavidin (Sigma) in PBS pH 7.5, rinsed, and then incorporated into flow cells. A \sim 3 mm x \sim 40 mm channel was cut into 0.12 mm double-sided tape (Grace BioLabs) to produce a cell volume of \sim 15 μ L. The tape was used to join the coverslip and a quartz slide with pre-drilled holes into which polyethylene tubing (0.76 mm inner diameter, 1.22 mm outer diameter; Becton Dickinson) was placed. The flow cell and tubing joints were sealed with epoxy. The assembled flow cell was incubated with buffer to establish non-specific surface blocking.

Aptamer Affinity, Rates, and Specific Binding

Previous work found that the Class V aptamer has an affinity for GTP of $17 \text{ nM}^{1,3}$ with slow on and off rate constants of $< 1/100 \text{ min}^{-1} \text{ nM}^{-1}$ and $1/10 \text{ min}^{-1}$, respectively (unpublished data, Szostak lab).

Specific binding of the GTP-TMR by several GTP aptamers was verified by bulk fluorimetry in solution (Varian/Cary Eclipse Fluorescence Spectrophotometer). The guanosine has a quenching effect on the TMR when they are coupled, enabling a bulk fluorimetry assay based on the dequenching of the TMR upon binding of the GTP by the RNA. This assay can be used to determine the on rate by fitting the dequenching curve with a single exponential (Figure S1); the off rate can be found by measuring the re-quenching upon addition of excess (1000-fold) unlabeled GTP (Figure S2). This yields an on rate of $0.082 \text{ min}^{-1} \text{ nM}^{-1}$ and an off rate of 0.044 min^{-1} , producing a calculated K_d for GTP-TMR of 0.54 nM . Specific binding in the flow cells was verified by control experiments using several aptamers, nonbinding mutant aptamers (Figure S3), nonbinding labeled alternate ligand (UTP-TMR), and various other system components. In the absence of RNA the GTP-TMR produces an average of only 11 transient nonspecific spots per frame (Figure S4). Binding is specific for GTP-TMR as compared to unconjugated TMR (Figure S5). Long run testing without the RNA was performed to check for any long-term nonspecific binding in the flow cell.

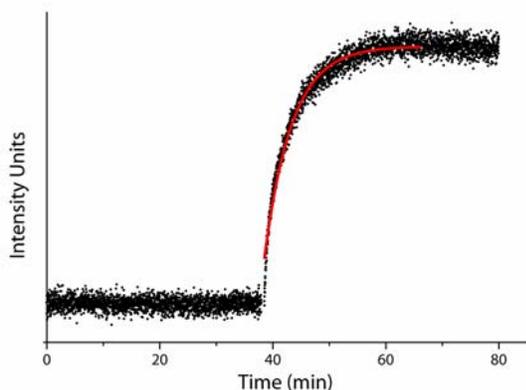


Figure S1. Class V bulk fluorimetry. Folded Class V RNA is added to GTP-TMR causing unquenching (black points). Fitting with a single exponential (red line) yields an on rate.

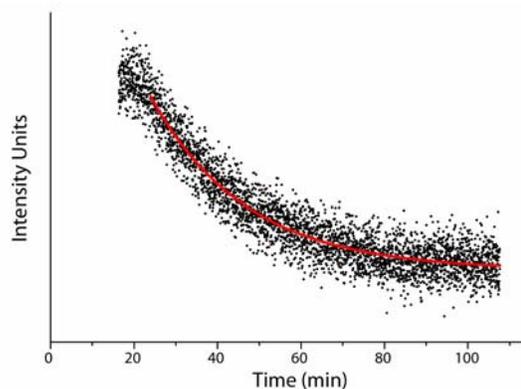


Figure S2. Class V bulk fluorimetry. A 1000-fold excess of GTP is added to folded Class V RNA and GTP-TMR. As GTP-TMR unbinds from the RNA where it is replaced by GTP, it is re-quenched (black points). Fitting with a single exponential (red line) yields an off rate.

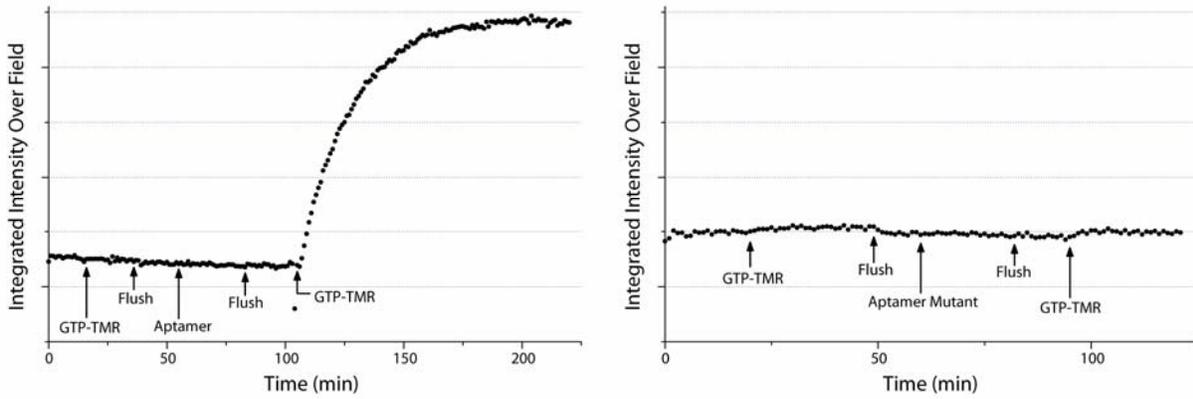


Figure S3. Example of specific binding control for the flow cell setup. Reagents are flowed into the flow cell at the points indicated by arrows. The traces show the integrated intensity over the whole field for each frame with a frame rate of 1 min^{-1} . (Left) Run with the 9-4 aptamer^{1,2} (another *in vitro* selected GTP aptamer) showing no increase in signal owing to GTP-TMR alone, and an increase in signal due to binding of GTP-TMR when the aptamer is present. (Right) Run under the same conditions with the same concentrations using a 9-4 mutant with a single mutation (position 54, A \rightarrow C) at a position critical for binding¹. The y axes of the two graphs are the same.

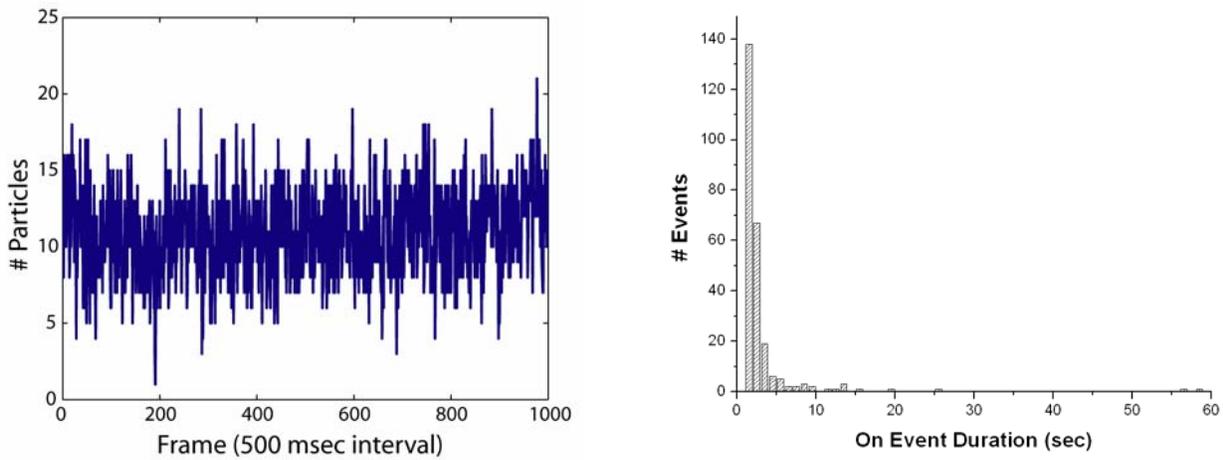


Figure S4. Nonspecific binding in the absence of aptamer. (Left) A 500 msec interval experiment with GTP-TMR in the absence of aptamer produces an average of 11 spots per frame, (Right) these are short transient surface interactions.

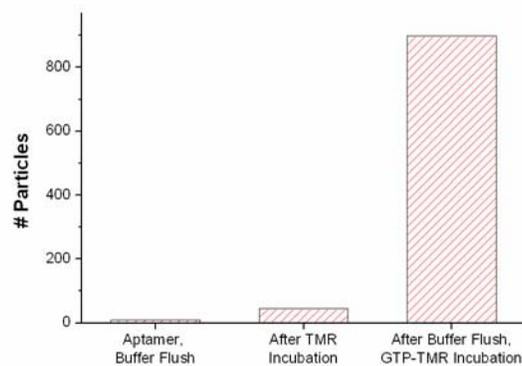


Figure S5. Specificity for ligand and not dye. A flow cell experiment performed with a GTP aptamer (10^{-10}) shows little binding upon incubation with unconjugated TMR (which then washed away) compared to the high level of ligand-specific binding upon incubation with GTP-TMR.

Experimental Setup

See Supplementary Figure S6. Imaging was done on an inverted microscope (Olympus IX71) equipped with a 60X 1.45 NA oil immersion objective (Olympus) driven by a piezoelectric drive (Physik Instrumente). In addition to laser-quality excitation, dichroic, and emission filters (Chroma HQ535/50x, z532rdc, 51007m dual band and HQ545lp) mounted in a filter cube in the microscope filter turret, an additional emission filter (Chroma HQ 635/10x) was placed in the lamp condenser assembly. The light source for single-molecule imaging was a 532 nm solid-state laser (CrystaLaser) as attenuated by optics and neutral density filters (Thor Labs). The photobleaching rate for the laser power used in these experiments was checked by bleaching GTP-TMR on clean glass slides, and was found to be $\sim 0.045 \text{ sec}^{-1}$; a rate constant of ~ 22 seconds for continuous illumination. This is stretched out to ~ 220 minutes by performing 200 ms image acquisitions at 2 minute intervals (see below); this is ~ 9 -fold longer than the mean on event durations found for the Class V aptamer. The light source for autofocusing was the microscope lamp. The filtered lamp light is at such a low power, and a wavelength shifted far enough away from the TMR absorption peak to ensure that there is no TMR photobleaching during autofocusing. Shutters (Ludl) were placed in front of the laser and lamp. Images were captured using an EM-CCD (Hamamatsu) running at $-65 \text{ }^\circ\text{C}$ with an acquisition time of 200 ms. The system was shielded from external light sources.

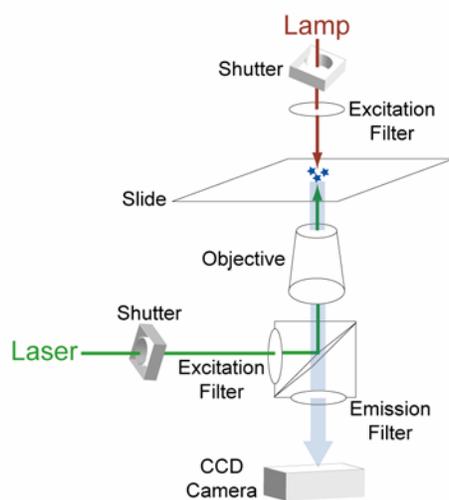


Figure S6. Imaging setup.

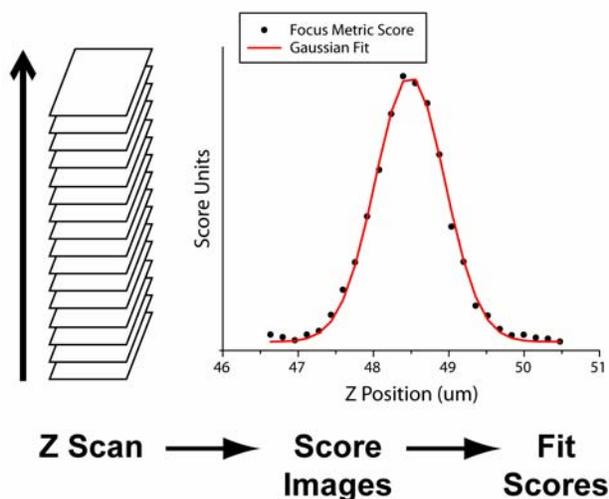


Figure S7. Autofocusing.

The setup was controlled using custom software written in LabVIEW (National Instruments). This software recorded images using the laser illumination at specified intervals (2 minutes for these experiments) by controlling the camera and laser shutters. This interval allows the lifetime of a fluorophore captured at the surface to be stretched out over a long period and also dictates the effective time resolution of the run. Between data image acquisitions the software performed autofocusing (Figure S7) by opening the lamp shutter and scanning through the z-axis using the piezoelectric objective drive. This produced images of the beads alone which were evaluated using a focus metric. The center of a gaussian curve fitted to the focus score data produced a current focal plane value which was used to position the piezoelectric drive. This approach to autofocusing by imaging beads was originally pioneered by the lab of Thomas Pollard⁵.

Solutions were fed into the flow cell by placing the input tubing in the current solution and creating flow using a syringe pump (Harvard Apparatus) connected to the flow cell output tubing. The pump was used only as a mechanism to smoothly add/remove reagents, at other times the pump was turned off.

Experimental Protocol

After being placed on the microscope the flow cell was flushed with buffer. The RNA and the DNA capture oligonucleotide were combined at a 1:1 ratio in water, heated to 80 °C for 1 minute, and then left to cool to room temperature in order to fold the RNA as per the original selection^{1,2} and to anneal the RNA and DNA. The RNA-DNA complex in buffer (80 pM for these experiments) was added to the flow cell, incubated 20 minutes, and then the flow cell was washed with buffer. The fluorescent microspheres, at a concentration found to yield ~3-6 beads/field, were added, left to incubate for 20 minutes, and then washed with buffer. The run was started via the LabVIEW software. After initial baseline acquisition, and at ~3.5 hour subsequent intervals, GTP-TMR (300 pM) in imaging buffer was added. The resulting movie, showing the initial appearance of numerous spots followed by spot disappearances and reappearances, was saved for analysis (See Supplementary Movie online; movie

shows every third frame of run and has been compressed and converted to reduce file size; the seven large constant bright spots are the fluorescent microspheres).

RNA Photodamage

Loss of aptamer activity as evidenced by the cessation of binding events (Figure S8) was found to be a serious hindrance to collecting aptamer binding data over long periods. RNA immobilized at the surface remains binding-competent for long periods. The loss of activity is correlated with the presence of a photo-reactive dye under laser illumination, suggesting that the effect is due to photochemically induced damage to the RNA. Given that the mean duration of final on events is not significantly different from the mean durations of the preceding events, and that exposure to a nonbinding labeled ligand also appears to induce the loss of activity, it is likely that such damage to the RNA is not simply caused by the dye attached to the bound ligand. We hypothesize that photochemically generated reactive oxygen species are damaging the RNA. This is consistent with the finding that the problem is ameliorated by use of the oxygen scavenging system as described above. This approach increases the average length of the trajectories, raising the maximum number of events found in a trace by as much as 4-fold.

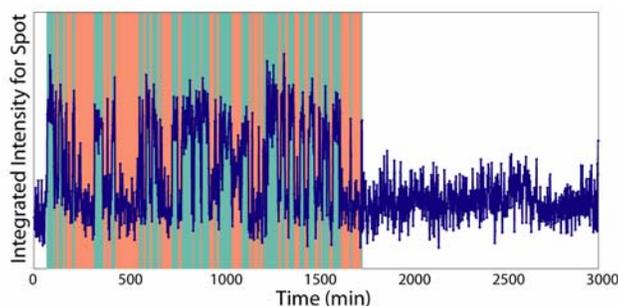


Figure S8. Loss of aptamer binding activity. An intensity trace from a single aptamer shows a sudden cessation of binding events indicating irreversible damage to the RNA.

Data Processing and Analysis

Initial Processing: Most data processing and analysis was done using MATLAB (MathWorks), with additional work done using ImageJ (<http://rsb.info.nih.gov/ij/>) and Origin (OriginLab). Some MATLAB particle finding routines were from the online code repository of the lab of Daniel Blair

(<http://physics.georgetown.edu/matlab/>). In brief, scripts were used to automatically find candidate spots using thresholding, peak finding, and centroid location. The same methods were used to find the fluorescent bead centers in each image. The bead centers were used to calculate the lateral drift for each frame and then to align all the frames to eliminate that drift. A window of 3.3 hours (out of 50.4 hours) in the middle of the movie was chosen for spot selection, a time point at which roughly half of the peak number of spots remained (see RNA Photodamage section above). A subset of candidate spots which did not overlap other spots was generated. The integrated intensity of 3x3 pixel windows around the spots was calculated for all movie frames to generate intensity trajectories over time. The intensity trajectories were then evaluated for candidate on and off events by hand and another custom program was used to mark all events and calculate event durations. These durations were then used for all subsequent analysis.

Duration Distribution Analysis: The event distributions (Fig 2B) are well fit by single exponentials, indicating single-rate on and off processes for which the rate is equal to both the mean and the standard deviation. Each single-molecule trace is a finite sampling of these distributions and the mean single-molecule on and off event durations should in the limit match their respective rate constants. The distribution of such single-molecule duration means should have a standard deviation $\sigma_{\text{means}} = \sigma_{\text{events}} / \sqrt{N}$, where N is the number of events, so given the single exponential event distribution for which σ_{events} is equal to the mean, we expect that $\sigma_{\text{means}} = \text{mean} / \sqrt{N}$. This is true of simulations for various particle counts and rates (simple simulations randomly drew event durations from single exponential distributions based on the given on and off rates, simulation data below is from runs using 103 simulated particles with rates matching those found for Class V), and is also true for the measured Class V single-molecule mean on event durations the distribution of which is thus completely explained by the sampling limit (Table S9). The relationship also holds well for the measured Class V single-molecule mean off event durations, except for $\sim 3.4\%$ of the events which form a right-hand side tail that extends beyond the expected range for the single exponential fit. These long outlier off events are

present in 59% of the molecules in the analysis set so that any given molecule has zero or one of these events. Thus this is not indicative of a subpopulation of molecules. While it is theoretically conceivable that these events correspond to a very rarely accessed state with a lower rate of binding, the fact that it is observed so infrequently over a long timescale makes for a much less likely explanation than the many alternative possible technical causes. These include not just the inevitable undersampling of short on events owing to the time resolution, but also the possibilities of unseen “dark” on events produced by binding unlabeled (hydrolyzed) GTP or photodamaged GTP-TMR, and below threshold “dim” on events caused by fluorescent signal variability. When the single-molecule mean durations are calculated without the outlier off events then the single-molecule outliers in “kinetic space” are no longer found (Figure S10).

	<u>N =</u>					
Sample Simulations:	<u># Events</u>	<u>Mean</u>	<u>σ_{means}</u>	<u>Mean / sqrt(N)</u>	<u>Difference</u>	<u>% Difference</u>
On Event Durations	5	22.80	10.18	10.20	0.02	0.2%
Off Event Durations	5	29.00	13.40	12.97	0.43	3.2%
On Event Durations	10	25.90	8.67	8.19	0.48	5.5%
Off Event Durations	10	30.45	9.84	9.63	0.21	2.1%
On Event Durations	18	25.07	5.67	5.91	0.24	4.2%
Off Event Durations	18	29.58	6.45	6.97	0.52	8.1%
On Event Durations	50	24.23	3.77	3.43	0.34	9.1%
Off Event Durations	50	30.93	4.54	4.37	0.17	3.7%
On Event Durations	500	24.90	1.01	1.11	0.10	10.3%
Off Event Durations	500	30.29	1.25	1.35	0.10	8.4%
	<u>N = Mean</u>					
Class V Measured:	<u># Events</u>	<u>Mean</u>	<u>σ_{means}</u>	<u>Mean / sqrt(N)</u>	<u>Difference</u>	<u>% Difference</u>
On Event Durations	18.33	24.95	6.26	5.83	0.43	6.9%
Off Event Durations	17.33	30.39	8.88	7.30	1.57	17.7%
Off Event Durations (Ignoring 3.4% Longest Events)	17.33	28.40	7.01	6.82	0.18	2.6%

Table S9. Distribution analysis.

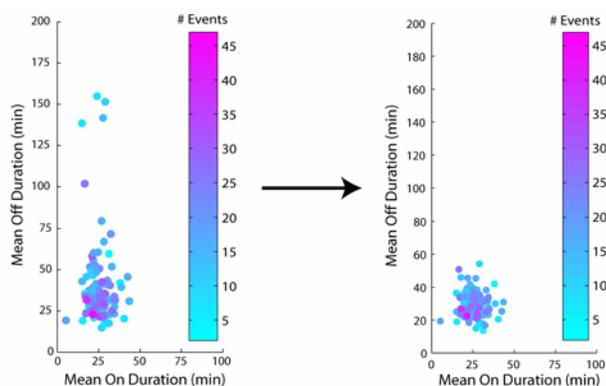


Figure S10. Scatter plot of the mean on and off durations for each single-aptamer trajectory as in Figure 3B (left), and done ignoring the longest 3.4% of off events (right). The color indicates the number of on events captured for each molecule.

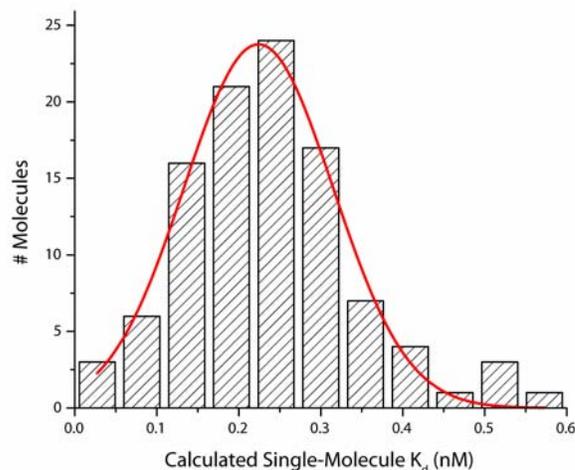


Figure S11. Single-molecule K_d distribution for Class V aptamer with fitted Gaussian curve in red

The single-molecule mean rates can be combined to yield single-molecule K_{ds} (Figure S11) from which gaussian fitting finds a K_d value of 0.22 nM, a close match to the bulk solution fluorimetry estimate of 0.54 nM. Applying a correction to account for expected photobleaching (discussed above), yields a K_d value of 0.25 nM. If the single-molecule mean duration gaussian fitting results (Figure 3A) are used, the resulting K_d is 0.25 nM; correcting this for expected photobleaching produces a K_d of 0.28 nM.

Short Timescale Experiments

The kinetics observed using the 2 minute interval are those that are expected from bulk solution measurements. Experiments with shorter intervals were performed to explore the possibility that some aptamers might display kinetics at shorter timescales. The shortest interval possible using the current system is 500 msec. Autofocusing is turned off to allow for fast imaging and because the bleaching time is shorter than the drift time. On event durations can be well fit with a single exponential (Figure S12) producing a rate which is comparable to that found for photobleaching (photobleaching is discussed

above, the comparison is made in units of exposure frames). Thus at short timescales there is a single rate process which appears to simply be photobleaching.

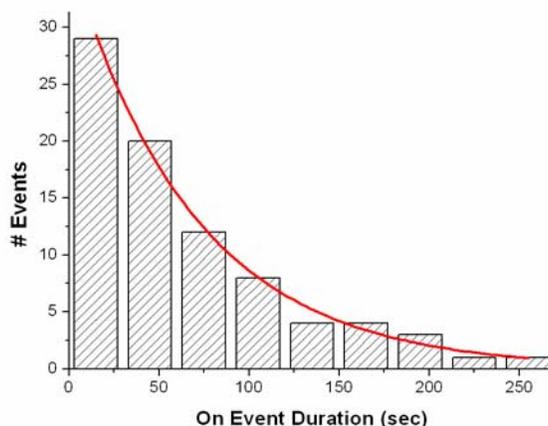


Figure S12. Short timescale kinetics. Experimental setup was as before but GTP-TMR was added and allowed to bind for 1 hour before imaging was performed using a 500 msec interval and no autofocusing. On event durations are binned and fit with a single exponential with a decay constant of 69.5 sec ($R^2 = 0.99$) or 139 frames.

Temporal Dynamics Analysis

The autocorrelation function captures the correlation of the length of any event n with the length of subsequent event $n+i$ for each possible lag of i events, yielding a sensitive measure of any molecular memory. The autocorrelation function $r(m)$ for m event lags, where $n = (\# \text{ Events Maximum}) - m$, can be calculated as⁶:

$$r(m) = \frac{n \sum_i t_i t_{i+m} - \left(\sum_i t_i \right)^2}{n \sum_i t_i^2 - \left(\sum_i t_i \right)^2}$$

2D histograms correlating the duration of event t with event $t+1$ or event $t+10$ were also generated, normalized by the number of events, and the $t+10$ results were subtracted from the $t+1$ results to show any subtle features indicating molecular memory. Memory would manifest itself as a diagonal vs. off-diagonal pattern, and is not seen (Figure S13).

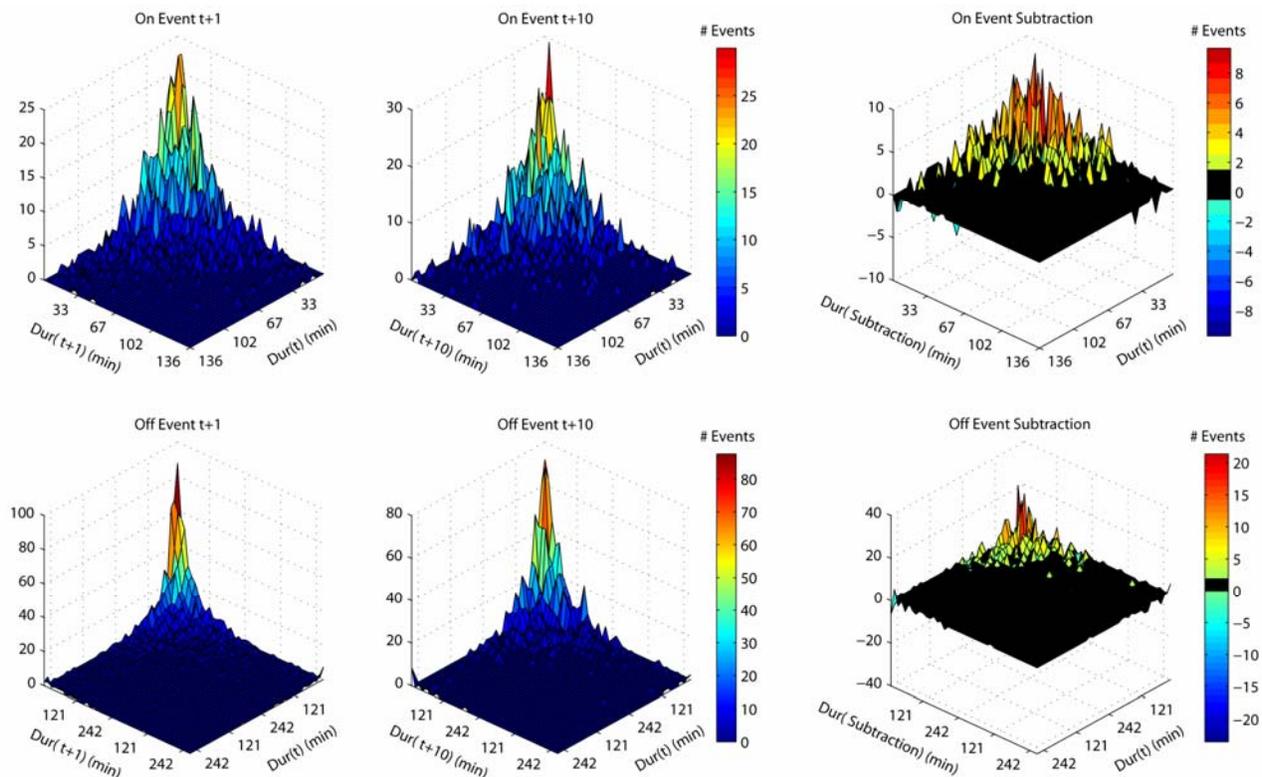


Figure S13. 2D event correlation histograms for on event durations (top row) and off event durations (bottom row) for particles with more than ten events. The first column compares event t to event $t+1$, the second column compares event t to event $t+10$, and the third column is the result of subtracting one histogram from the other. Bins were normalized by the number of events in order to perform the subtraction.

References

- (1) Carothers, J. M.; Oestreich, S. C.; Davis, J. H.; Szostak, J. W. *J Am Chem Soc* **2004**, *126*, 5130-5137.
- (2) Davis, J. H.; Szostak, J. W. *Proc Natl Acad Sci U S A* **2002**, *99*, 11616-11621.
- (3) Carothers, J. M.; Oestreich, S. C.; Szostak, J. W. *J Am Chem Soc* **2006**, *128*, 7929-7937.
- (4) van Oijen, A. M.; Blainey, P. C.; Crampton, D. J.; Richardson, C. C.; Ellenberger, T.; Xie, X. S. *Science* **2003**, *301*, 1235-1238.
- (5) Kuhn, J. R.; Pollard, T. D. *Biophys J* **2005**, *88*, 1387-1402.
- (6) Lu, H. P.; Xun, L.; Xie, X. S. *Science* **1998**, *282*, 1877-1882.