Supporting Information

Trevino et al. 10.1073/pnas.1107113108

SI Text

Aptamer Dissociation Constant Determination by NMR. The binding equilibrium of the aptamer and ligand can be described by the simple model: $A + nL = AL_n$, where A is the unbound aptamer, L is the free ligand, and n is the number of binding sites. The microscopic dissociation constant, $[K_d']^h = [A][L]^n/[AL_n]$, where h is the Hill coefficient, which approaches the value of n for fully cooperative binding. Because the number of GTP aptamer binding sites remains unknown, one binding site was assumed. Additionally, the DNA ATP aptamer is known to have two binding sites (1). Thus, for the calculation of free ligand concentration in the ATP aptamer titrations (see below), we assumed two binding sites to conserve proper stoichiometry. However, because our data indicate that the extent of ligand cooperativity is small, a Hill coefficient of 1 was assumed (indicating no negative or positive cooperativity). In this case, $[K_d']^h$ is equal to the apparent dissociation constant, K_d, which can be determined by plotting complex AL_n formation as a function of free ligand concentration.

Dissociation Constant Determination by Water-Ligand Observed Gradient Spectroscopy (waterLOGSY). The waterLOGSY signal intensities of titrated ligand (ATP or GTP), starting at a concentration of 80–160 μM, were recorded both in the absence or presence of a constant amount of aptamer (56–300 μM). The hydrated free ligand generates a weak, positive NOE, whereas bound ligand within ligand-aptamer complex generates a strong, negative NOE. Because the exchange rate between bound and free state ligand is faster than the NMR timescale, the signal resonances of bound and free state ligand are merged together. Therefore, the directly observed waterLOGSY signal is a combined weighted average resulting from both the free and bound ligand NOE signals at equilibrium.

To eliminate the contribution from free ligand and obtain the waterLOGSY signal intensity due exclusively to bound ligand—which is proportional to the formation of complex AL_n—the waterLOGSY signal intensities recorded in the absence of aptamer were subtracted from those recorded in the presence of aptamer.

Nonlinear fitting of the data was executed with Prism 5.0 (GraphPad Software Inc.) according to the Hill equation, Eq. **S1**:

$$I_{\text{wlogsy}} = I_{\max}([L_{\text{free}}]^h/[K_d']^h + [L_{\text{free}}]^h), \qquad \textbf{[S1]}$$

where I_{wlogsy} is the intensity of NOE signal attributed to bound ligand (arbitrary units), $I_{\max}$ is the saturation level of the signal (corresponding to 100% AL_n complex formation), h is the Hill coefficient, K_d' is the microscopic dissociation constant, and $[L]_{\text{free}}$ is the concentration of ligand in the free state and can be calculated by using Eq. **S2**:

$$[L]_{\text{free}} = [L]_{\text{total}} - n[A]_{\text{total}} I_{\text{wlogsy}}/I_{\max}, \qquad \textbf{[S2]}$$

where $[L]_{\text{total}}$ is the total concentration of ligand, $[A]$ is the total concentration of aptamer, and the ratio of $I_{\text{wlogsy}}/I_{\max}$ represents the fractional saturation level.

Starting with an approximated $I_{\max}$ value, $[L]_{\text{free}}$ can be estimated from Eq. **S2**. $[L]_{\text{free}}$ and I_{wlogsy} are used as inputs to nonlinearly fit the Hill equation, Eq. **S1** thereby resulting in a more accurate $I_{\max}$ value as well as an estimated K_d value. This process can be iterated several times until the K_d and $I_{\max}$ values are in fair agreement as indicated by the R^2 value.

Dissociation Constant Determination by ^{1}H NMR Line Width. Watergate 1D ^{1}H watergate NMR spectra were recorded for ligand titration series in the presence of a constant amount of aptamer (corresponding to aptamer concentrations of 60–80 μM). For each ligand concentration, the line width of the NMR signal attributed to the nucleotide H8 or H2 was measured at half peak maximum. The observation of a consistently sharp internal reference line width signal indicated satisfactory field homogeneity for each titrated sample. In this method, broadening of the H2 or H8 ligand signal is indicative of an increased transverse relaxation rate (R_2) adopted by the ligand while complexed with the aptamer. Attenuation of line broadening occurs as more ligand remains in the unbound state (2). Because of relatively weak binding (micromolar), the exchange rate between the bound and free states is much faster than the NMR timescale, which causes the ligand signal resonances due to both bound and free states to merge. Fast exchange does not significantly contribute to extra line width broadening and may be neglected.

Thus, the observed line width signal is related to $R_{2\text{obs}} = F_f \times R_{2f} + F_b \times R_{2b}$, where F_f and F_b are the fraction of ligand in the free and bound states, respectively, and R_{2f} and R_{2b} are the transverse relaxation rates of ligand in the free and bound states, respectively. This equation can be arranged to $F_b = (R_{2\text{obs}} - R_{2f})/(R_{2b} - R_{2f})$, to reflect the fraction of bound ligand. R_{2f} was determined by a titration of ligand in the absence of aptamer.

The dissociation constant (K_d) can be determined by plotting changes in the observed line width of NMR signal as a function of the free concentration of ligand. according to the Hill equation:

$$\theta = [L_f]^h/([K_d']^h + [L_f]^h),$$

where θ is the fraction of aptamer that is bound to ligand, or:

$$\theta = F_b[L_{\text{total}}]/n[A_{\text{total}}],$$

where $[A_{\text{total}}]$ is the total concentration of aptamer, n is the number of binding sites, and $[L]_f$ is the concentration of ligand in the free state. Because ligand concentration is in excess to aptamer concentration, the total concentration of ligand, $[L_{\text{total}}] \approx [L_f]$. Additionally, a Hill coefficient of 1 is assumed such that K_d' is approximately equal to K_d.

Thus, we obtain the equation:

$$R_{2\text{obs}} = R_{2f} + n[A](R_{2b} - R_{2f})/(K_d + [L]_f). \qquad \textbf{[S3]}$$

This approximation allows for the calculation of K_d and R_b by iteratively refitting the data (Eq. **S3**) until K_d and R_{2b} values are in fair agreement as indicated by the R^2 value. Nonlinear fitting of the data was executed with Prism 5.0 (GraphPad Software Inc.).

1. Lin CH, Patel DJ (1997) Structural basis of DNA folding and recognition in an AMP DNA aptamer complex: Distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP. *Chem Biol* 4:817–832.

2. Ludwig C, Guenther UL (2009) Ligand based NMR methods for drug discovery. *Front Biosci* 14:4565–4574.

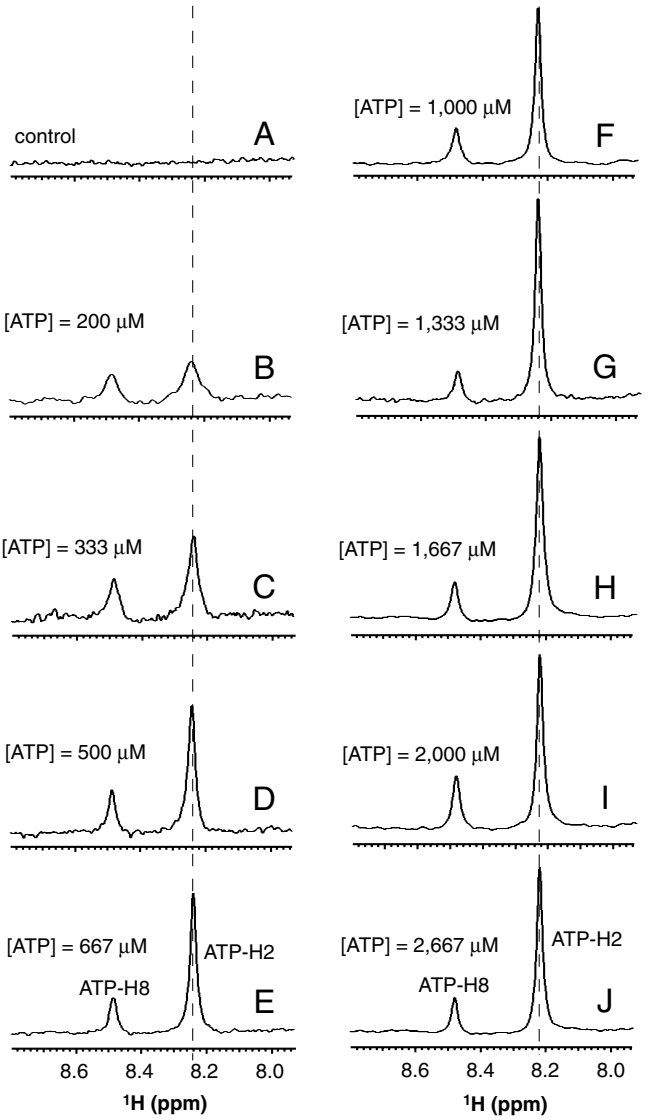

Fig. S1. NMR experimental data for the titration of the mosaic nucleic acid (MNA) ATP aptamer 74 with ATP. One-dimensional ^{1}H NMR waterLOGSY spectra of a series of ATP ligand concentrations (indicated) generated in the presence (*A–J*) of approximately 80 μM MNA ATP aptamer 74. The expanded spectral region that contains ATP aromatic nucleobase proton signals is shown. The signal intensities are on a uniform arbitrary scale throughout the titration.

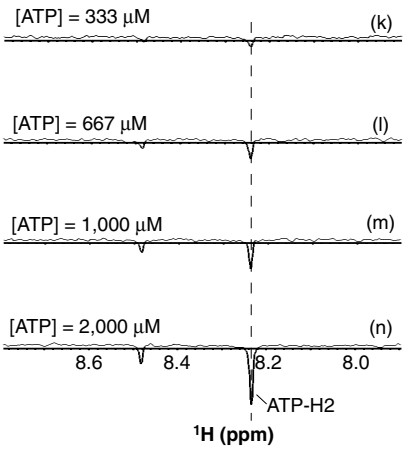

Fig. S2. NMR experimental data for the titration of the mosaic nucleic acid (MNA) ATP aptamer 74 with ATP. One-dimensional ^{1}H NMR waterLOGSY spectra of a series of ATP ligand concentrations (indicated) generated in the absence (*K–N*) of approximately 80 μM MNA ATP aptamer 74. The expanded spectral region that contains ATP aromatic nucleobase proton signals is shown. The signal intensities are on a uniform arbitrary scale throughout the titration.

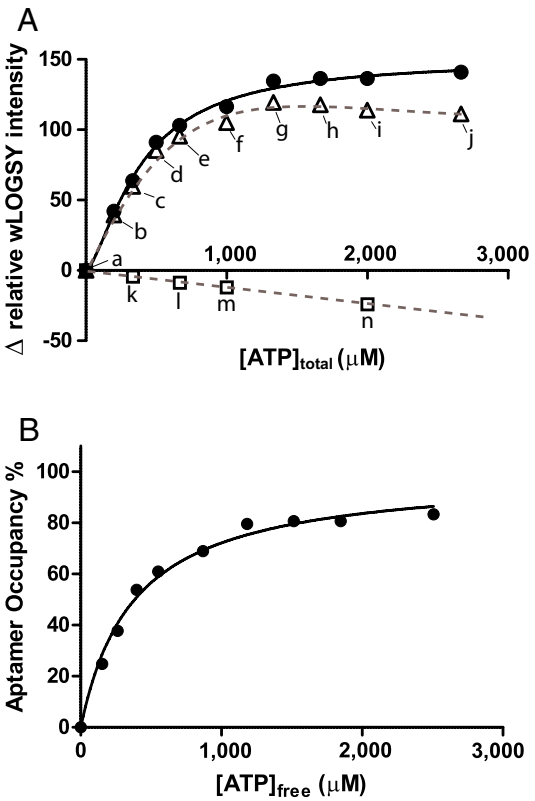

Fig. S3. Determination of the MNA ATP aptamer 74 dissociation constant for ATP. (*A*) The calibrated change in relative waterLOGSY intensity for the ATP H2 proton as a function of increasing total ligand concentration (●) calculated by subtracting the observed signal acquired in the absence of aptamer (□) from observed signals acquired in the presence of aptamer (△). Lowercase letters correspond to experimental data in Fig. S1 and Fig. S2. (*B*) The change in fractional occupancy as a function of free ligand concentration, calculated by nonlinear fitting of the experimental data of the change in calibrated relative waterLOGSY intensity against ATP concentrations (see *Materials and Methods*).

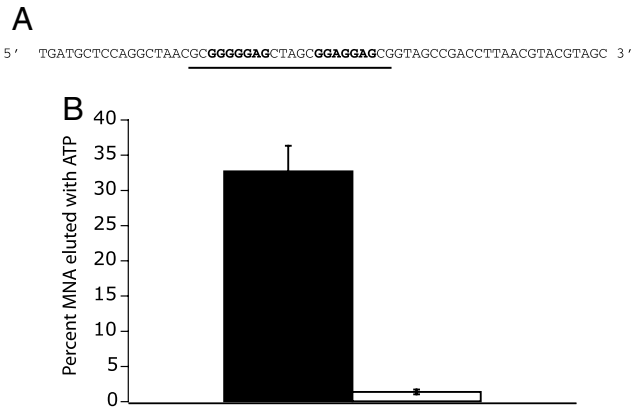

Fig. S4. Mosaic nucleic acid (MNA) ATP aptamer 74 column binding assays. (*A*) DNA sequence of the MNA ATP aptamer 74 flanked by the PBS regions GGGAGAGGAGAGAAACG and GGATCGTTACGACTAGCATCGATG (5′ to 3′). The previously identified DNA ATP aptamer motifs are shown in bold. Bases deleted in the MNA ATP aptamer 74 deletion mutant are underlined. (*B*) The full-length MNA ATP aptamer 74 sequence (black, $n = 6$), or the MNA ATP aptamer 74 deletion sequence (white, $n = 3$) was incubated with ATP agarose (2 mM) for 5–15 min, washed with 12 column volumes of binding buffer followed by three column volumes of buffer containing free ATP ligand (2 mM). Percent of MNA eluted from the column by free ATP is shown. After all wash regimes, background binding by the MNA to the derivatized agarose ranged from 15 to 40%.

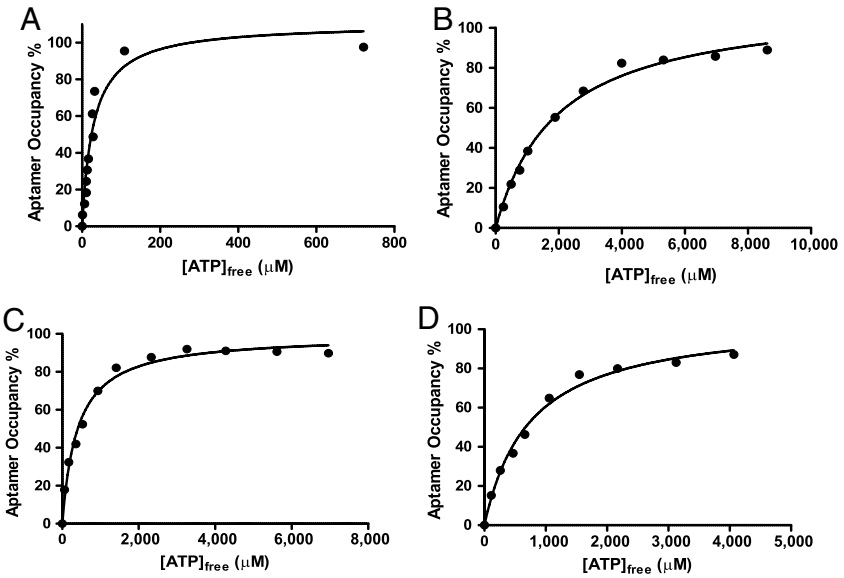

Fig. S5. ATP aptamer binding curves for ATP. Binding curves for (*A*) DNA, (*B*) RNA, (*C*) RNA/DNA 74-1, and (*D*) 74-2 versions of the minimal ATP aptamer sequence. Data collected by the waterLOGSY methodology (see *Materials and Methods*).

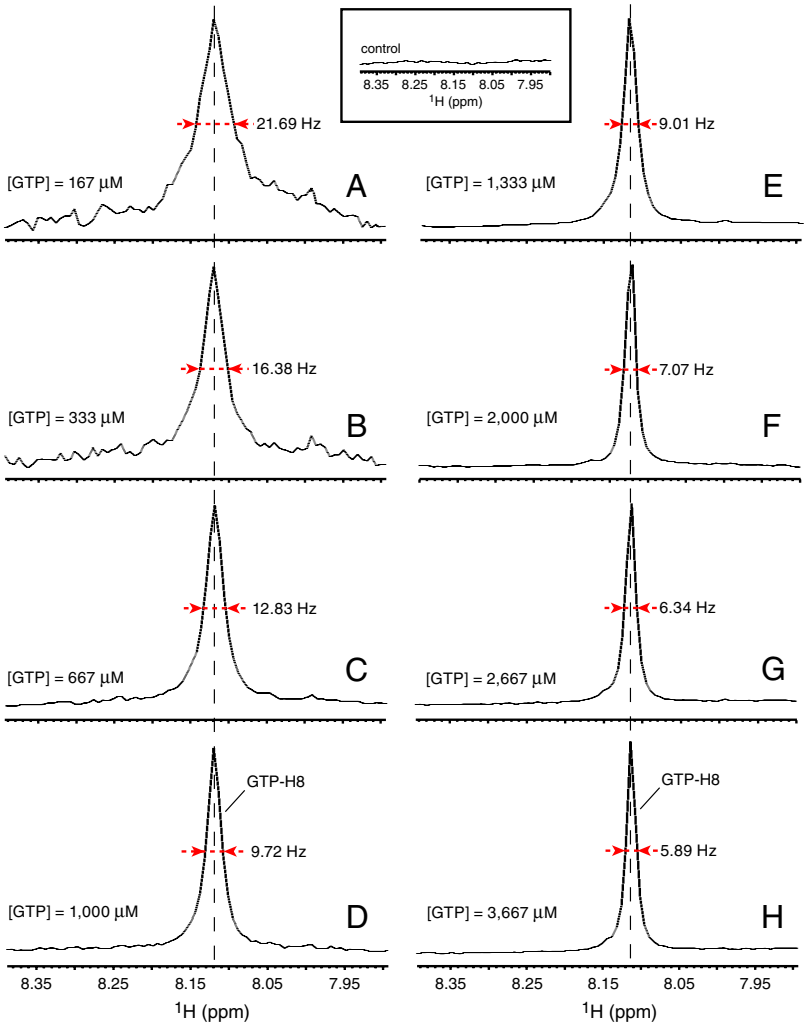

Fig. S6. NMR experimental data for the titration of the mosaic nucleic acid (MNA) GTP aptamer 812 with GTP. One-dimensional ^{1}H spectra of a series of GTP (*A*–*H*) ligand concentrations (indicated) in the presence of approximately 80 μM MNA GTP aptamer 812. The expanded spectral region that contains aromatic nucleobase proton signals is shown. The peak line widths are as labeled. The height of displayed signals is normalized to the same level.

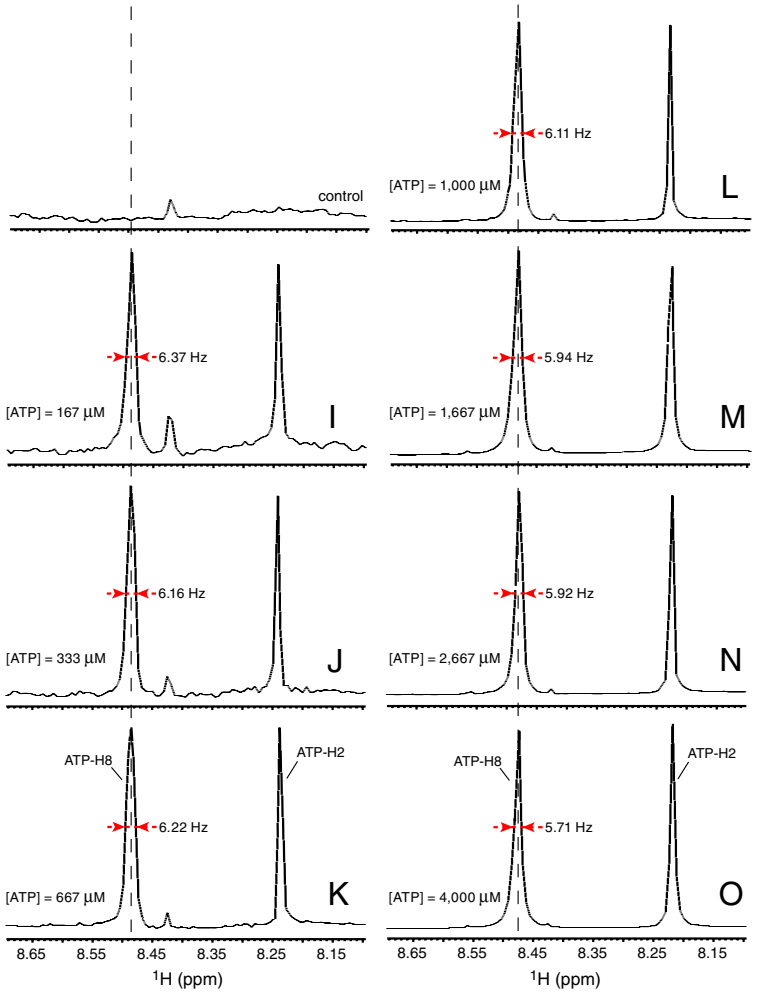

Fig. S7. NMR experimental data for the titration of the mosaic nucleic acid (MNA) GTP aptamer 812 with ATP. One-dimensional ^{1}H spectra of a series of ATP (*I–O*) ligand concentrations (indicated) in the presence of approximately 80 μM MNA GTP aptamer 812. The expanded spectral region that contains aromatic nucleobase proton signals is shown. The peak line widths are as labeled. The height of displayed signals is normalized to the same level.

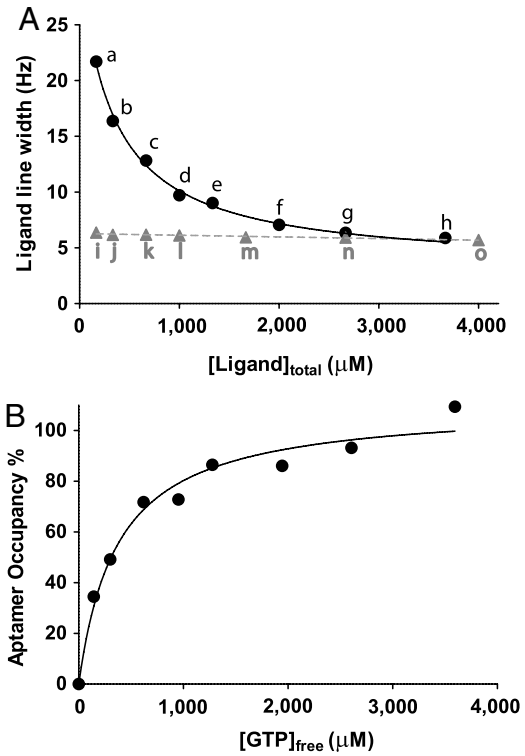

Fig. S8. Mosaic nucleic acid (MNA) GTP aptamer 812 dissociation constant determination for GTP. (*A*) Observed NMR line width upon titration with GTP (●) or ATP (△). Lowercase letters correspond to experimental data in Fig. S6 and Fig. S7. (*C*) The fraction of bound ligand as a function of increasing free GTP concentration was determined by nonlinear fitting to the observed data (see *Materials and Methods*).

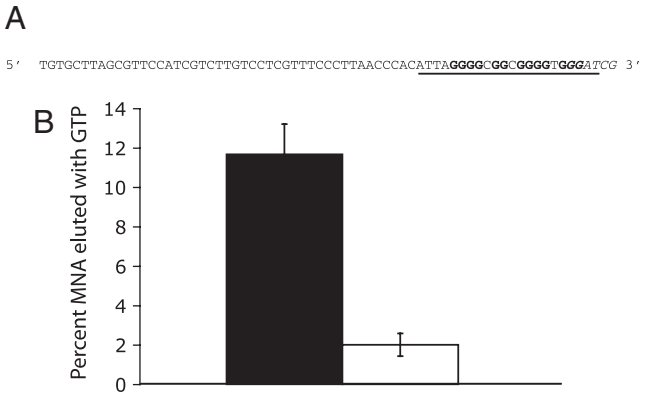

Fig. S9. Mosaic nucleic acid (MNA) GTP aptamer 812 column binding assays. (*A*) DNA sequence of the MNA GTP aptamer 812 including six bases complement to the reverse PBS (italic). The G-rich 812-1 sequence that is deleted in the MNA GTP aptamer 812 deletion is underlined. (*B*) The full-length MNA GTP aptamer 812 sequence (black, $n = 6$), or the MNA GTP aptamer 812 deletion sequence (white, $n = 3$) was incubated with GTP agarose (5 mM) for 5–15 min, washed with 12 column volumes of binding buffer followed by three column volumes of buffer containing free GTP ligand (5 mM). Percent of MNA eluted from the column by free GTP is shown. After all wash regimes, background binding by the MNA to the derivatized agarose ranged from 15 to 40%.

Table S1. Determination of the approximate fraction of ribose and deoxyribose in mosaic nucleic acid (MNA)

Sequence	Ratio (d/r)							RNA in full-length, %
Nucleotide fragment size			mono-	di-	tri-	tetra-	penta	
Starting pool	1:1	% total fragments	73	13.5	9	3	0.7	84
		%RNA	73	6.8	3.3	0.8	0.1	
Starting pool	9:1	% total fragments	24	12	24	18	18	46
		%RNA	24	6	8	4.5	3.6	
Aptamer 74	9:1	% total fragments	21	13	19	22	25	44.1
		%RNA	21	6.5	6.3	5.3	5	
Aptamer 812	9:1	% total fragments	28	11	28	21	12	50.5
		%RNA	28	5.5	9.3	5.3	2.4	

MNA was digested by potassium hydroxide to cleave MNA at ribonucleotide linkages under conditions that fully digest an RNA transcript to mononucleotides but do not digest a DNA primer. The resultant product is a complex pool of nucleotide fragments that each contain one ribonucleotide. Because each fragment is characterized by a difference in charge, nucleotide monomers, dimers, trimers, and so forth can be resolved by HPLC (260 nm absorbance), where RNA and DNA absorb similarly. Thus, absorbance due to ribonucleotides for the monomer fraction is 100% whereas absorbances due to ribonucleotides for dimer and trimer fractions are 50% and 33%, respectively. Based on the relative absorbance of all observed fractions, it is possible to estimate the amount of RNA in the original MNA transcript. In all cases, peaks that eluted later than pentamers had combined absorbances (<5%) and were discounted from calculations.

Table S2. Sequences of the ATP and GTP mosaic nucleic acid (MNA) aptamers obtained from in vitro selections in this study

Selection	Name	Sequence	Copy no.
ATP		GTGGCAGCGGTACGC**GGGGGAG**TCTGCTTCGGCAGGC**GGAGGAG**CGCACAGGAGCCGCTCG	11
		CCCCTAGAGGAACTAAGTATCAGGCC**GGAGGAG**CACATGC**GGGGGAG**ACCAATAGCTACTCTAG	10
	74	TGATGCTCCAGGCTAACGCG**GGGGAG**CTAGC**GGAGGAG**CGGTAGCCGACCTTAACGTACGTAGC	5
		ATCTCCATGCCTGAGCGGCGGAGCTCCTGCTTC**GGCAGGAC**GC**GGGGGAG**CTCTACCATGTAGT	4
		TGGA**GGAGGAG**GATATTGAGGAAGAGTACCTCC**GGGGGAG**CGTAACATCTAACTCCCCTCGGTGC	4
		CTACGTCCCAAAGCGATCTGCCTGCCGC**GGAGGAG**GCGCGCTTAGCGCGCT**GGGGGAG**CCGAGC	3
GTP		TACCGGCCCTTGACCACTGCTCGTCCATACCTTTCCCTACATCTATTCCA**GGGGTGGCGGGGTGGGAT**	19
		TGCGATGCTCAGTCACCCTGAGAACTACACGTAACCTCTCCGTACCACCA**GGGG**A**GGC**G**GGGG**T**GGG**AT	6
	812	TGTGCTTAGCGTTCATCGTCTTGTCCTCGTTTCCCTTAACCCACATTA**GGGGTGGCGGGGTGGG**AT	5
		TTGTAGCTGCAACCTTTAGGATGGGGCCTACGTACCTCCTCCACCCTAAT**GGGGA**G**GC**G**GGGGT**G**GG**AT	5
		GCCTTGCCCCTAGGCTCAACTAGGCCCTCGATACCACTACCTGTAACTT**GGGGCGGTGGGGA**GG**GG**AT	4
		CTCACCCATAGGCCCCAGAAGCCCCTCTGCTTCGGCAGGCCGTATACCCA**GGGGCGGCGGGGTGGG**AT	4
		ACCTCCCGGACGAACCTGGCACCCGTTACATGCCCTCTCCCCC**GGGGTGGGGGGGGGG**AT	4
		TGTCCTCGGACGCTTACTTTAGGATTCCTCCACCCTTAAACACAATAGA**GGGGCGGCGGGGTGGG**AT	3
		TTCAGTCTCCGATACTCCCAGGCTGCCTGCTTCGGCAGCTATCCCTCAC**GGGG**A**GGTGGGGC**G**GG**AT	3
		CCTCCGCGGACTGTATACCCGATGCCTACTCCC**GGGCGGGGTGGTGGGG**TGTCCTTCCCGATGC	3
		ACACATCAGAGTCACACAAGACGTGTCCCCACTGGCCCTTCATCTCCCA**GGGGTGGCGGGG**A**GG**G**GG**AT	3

The region shown corresponds to the initial random region of the library. Only sequences for which three or more isolates were obtained are shown. Common G-rich motifs for the ATP aptamer sequences (GGGGAG and GGAGGAG) and for the GTP selection (GGGG, GG, GGGG, GGG) are shown in bold. Partial reverse primer binding sites are underlined.