Aptamers Selected for Higher-Affinity Binding Are Not More Specific for the Target Ligand

James M. Carothers,[†] Stephanie C. Oestreich,[‡] and Jack W. Szostak*

Contribution from the Department of Molecular Biology, and Center for Computational and Integrative Biology 7215, Simches Research Center, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114

Received February 15, 2006; E-mail: szostak@molbio.mgh.harvard.edu

Abstract: Previous study of eleven different in vitro-selected RNA aptamers that bind guanosine triphosphate (GTP) with K_ds ranging from 8 μM to 9 nM showed that more information is required to specify the structures of the higher-affinity aptamers. We are interested in understanding how the more complex aptamers achieve higher affinities for the ligand. In vitro selection produces structural solutions to a functional problem that are are as simple as possible in terms of the information content needed to define them. It has long been assumed that the simplest way to improve the affinity of an aptamer is to increase the shape and functional group complementarity of the RNA binding pocket for the ligand. This argument underlies the hypothesis that selection for higher-affinity aptamers automatically leads to structures that bind more specifically to the target molecule. Here, we examined the binding specificities of the eleven GTP aptamers by carrying out competition binding studies with sixteen different chemical analogues of GTP. The aptamers have distinct patterns of specificity, implying that each RNA is a structurally unique solution to the problem of GTP binding. However, these experiments failed to provide evidence that higher-affinity aptamers bind more specifically to GTP. We suggest that the simplest way to improve aptamer K_ds may be to increase the stability of the RNA tertiary structure with additional intramolecular RNA−RNA interactions; increasingly specific ligand binding may emerge only in response to direct selection for specificity.

Introduction

RNA aptamers are functional oligonucleotides that fold into distinct three-dimensional shapes and bind target ligands.[1−3] In the laboratory, RNA aptamers can be generated by searching large combinatorial pools of random sequences for rare functional molecules using in vitro selection.[4] In vitro selection can yield a diverse array of distinct binding structures that vary significantly in their affinities for the ligand.[2,5,6] With respect to their frequency in sequence space (the ensemble of all possible sequences of a particular length), the products of in vitro selection from random sequence pools tend to be the simplest, most abundant structural solutions to a given problem.[7] For this reason, in vitro-selected aptamers are likely to utilize the least complex structural mechanisms possible to attain the binding affinities that they exhibit. Therefore, by defining the

mechanistic and structural factors that underlie differences in affinity for a set of RNA aptamer structures from the same selection, it should be possible to uncover the simplest ways to improve binding.

It is perhaps easiest to imagine that tighter binding to a target molecule, measured by a lower apparent dissociation constant (K_d), emerges through improved shape and functional group complementarity between the ligand and the aptamer binding pocket.[3,8,9] If the binding pocket is rigid, a high degree of interaction and geometric complementarity with the ligand could also be expected to lead to high-specificity binding,[3,9] where specificity is defined as preferential binding to a target ligand compared to other molecules.[9,10] In fact, it has long been assumed that selection for higher-affinity binding automatically leads to aptamers that bind more specifically to a target molecule.[9] However, this view discounts the potential relevance of contributions to the free energy of binding from intramolecular contacts within the RNA itself, as opposed to intermolecular contacts between the RNA and the ligand.[10] To date, this issue has not been explicitly addressed through the analysis of specificity as a function of affinity for a set of aptamers selected against the same target.

The set of 11 different GTP-binding aptamers that we previously isolated[5,6] provides an ideal testing ground for the

[†] Current address: California Institute for Quantitative Biomedical Research and Center for Synthetic Biology, University of California, Berkeley, 717 Potter Street MC # 3224, Berkeley, CA 94720-3224.

[‡] Current address: Novartis Institutes for BioMedical Research, Inc., 400 Technology, Square, Cambridge, MA 02139.

(1) Ellington, A. E.; Szostak, J. W. *Nature* **1990**, *346*, 818−822.
(2) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1990**, *68*, 611−647.
(3) Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820−825.
(4) Carothers, J. M.; Szostak, J. W. In *The Aptamer Handbook: Functional Oligonucleotides and their Applications*; Klussmann, S., Ed.; Wiley-VCH: Berlin, 2006; pp 3−28.
(5) Davis, J. D.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11616−11621.
(6) Carothers, J. M.; Oestreich, S. C.; Davis, J. H.; Szostak, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 5130−5135.
(7) Lorsch, J. R.; Szostak, J. W. *Acc. Chem. Res.* **1996**, *29*, 103−110.

(8) Lancet, D.; Sadovsky, E.; Seidemann, E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3715−3719.
(9) Eaton, B. E.; Gold, L.; Zichi, D. A. *Chem. Biol.* **1995**, *2*, 633−638.
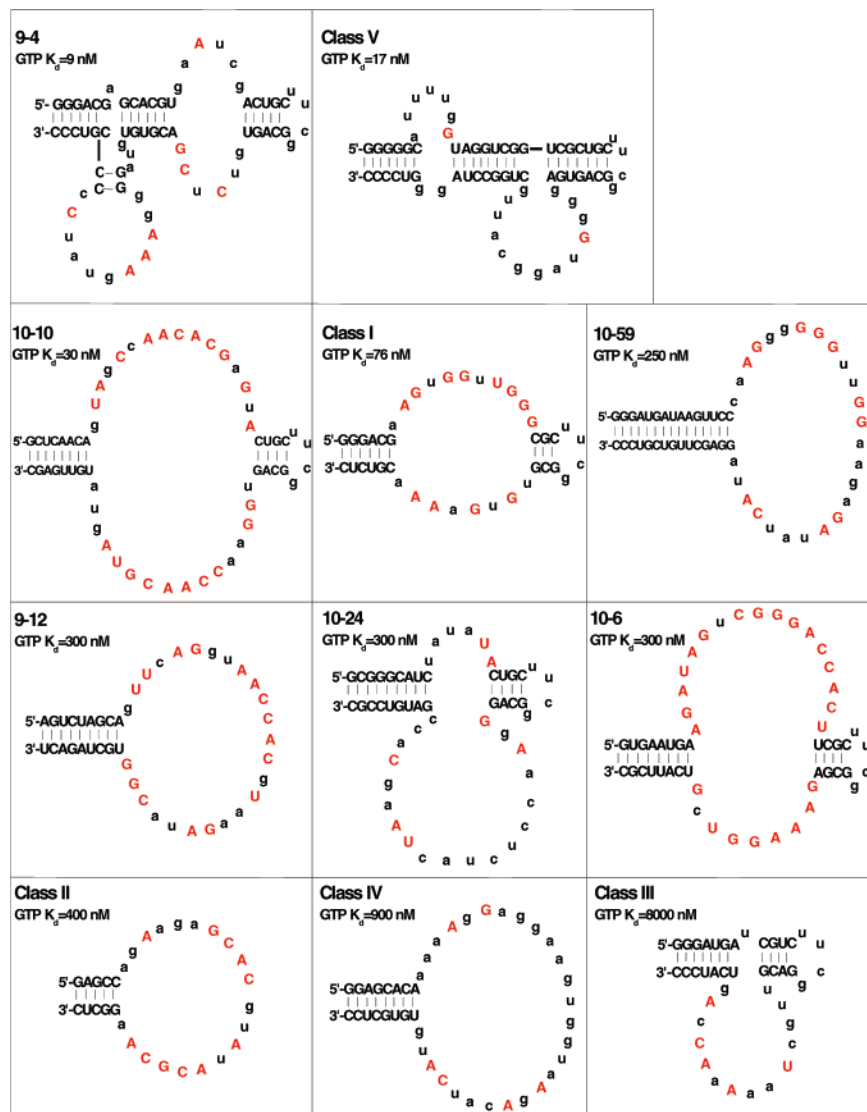(10) Szwajkajzer, D.; Carey, J. *Bioploymers* **1997**, *44*, 181−198.

Figure 1. GTP aptamer secondary structures. The secondary structure of each GTP aptamer used in this analysis is shown with invariant positions in the recognition loops outlined in red.

Table 1. Aptamer Properties

aptamer	K_d (nM)[a]	IC (bits)	size (nt)	$\Delta G_{2°(folding)}$ (−kcal mol^{-1})[b]	# analogues with log $K_{d\ (analog)}/K_{d\ (GTP)}$ < 2
9−4	9 ± 1	65.0 ± 1.0	69	26	10
Cl V	17 ± 4	55.0 ± 1.0	68	31	9
10−10	30 ± 6	67.0 ± 1.0	60	19	9
Cl I	76 ± 3	45.0 ± 0.5	41	11	11
10−59	250 ± 20	43.0 ± 0.5	56	22	11
9−12	300 ± 50	53.0 ± 0.5	43	10	10
10−24	300 ± 50	44.0 ± 1.5	55	17	12
10−6	300 ± 100	67.0 ± 1.0	54	13	9
Cl II	400 ± 200	40.0 ± 1.0	30	6	12
Cl IV	900 ± 200	33.0 ± 1.0	43	10	9
Cl III	8000 ± 1000	42.0 ± 1.0	41	14	8

[a] Apparent K_d and information content (IC) are from ref 6. Other values were measured or calculated as described in the text. [b] $\Delta G_{2°(folding)}$: calculated free energy of secondary structure formation (−kcal mol^{-1}).

affinity-specificity hypothesis (Figure 1, Table 1). The optimized versions of these aptamers vary greatly in sequence and affinity; the K_d for GTP of the weakest binder is 8 μM, whereas the strongest binder in the set has a K_d for GTP of 9 nM. We used in vitro selection to determine the number of conserved nucleotides and base pairs needed for binding in each aptamer

by identifying functional sequence variants of the original isolates.[6] From these data, the information content[11−13] necessary to specify the aptamer structures in the conditions of the selection was calculated. We found that tighter-binding aptamers are larger, have more intricate secondary structures, and are more informationally complex than weaker-binding aptamers. On average, approximately 10 bits of additional information content, equivalent to specifying 5 more conserved positions or base pairs within an RNA, are required to define structures capable of 10-fold better binding. Although this analysis proved that more nucleotide or base-pair positions must be specified in tighter-binding aptamers, there has not yet been a mechanistic or structural explanation for how the more informationally complex aptamers accomplish higher-affinity binding.[6,14]

Here, we present a detailed biochemical characterization of the binding specificities of the set of 11 GTP aptamers. We determined the aptamer affinities for a set of 16 different

(11) Schneider, T. D.; Stormo, G. D.; Gold, L.; Ehrenfeucht, A. *J. Mol. Biol.* **1986**, *188*, 415−431.
(12) Adami, C.; Cerf, N. J. *Physica D (Amsterdam)* **2000**, *37*, 62−69.
(13) Adami, C. *Phys. Life Rev.* **2004**, *1*, 3−22.
(14) Lau, M. W. L.; Cadieux, K. E. C.; Unrau, P. J. *J. Am. Chem. Soc.* **2004**, *126,* 15686−15693.

nucleotide and nucleoside analogues of GTP in competition binding assays and compared these to the aptamer affinities for GTP. First, we find that each aptamer exhibits a unique pattern of ligand-binding specificity compared to the others in the set. Taken together with the fact that there are no sequence motifs common to more than one of the GTP aptamers, this study provides further evidence that each aptamer is a distinct structural solution to the problem of binding GTP with moderate to high affinity.[6] Furthermore, contrary to expectations,[9] these data do not demonstrate that binding specificity is closely related to binding affinity; aptamers with higher affinities for GTP do not bind more specifically to GTP than aptamers with lower affinities for GTP. We infer that the more complex, higher-affinity aptamers do not make more intermolecular aptamer−ligand functional group interactions or have higher degrees of steric complementarity with the ligand. Instead, we propose that much of the observed[6] informational cost of improving binding reflects the need to increase the number and quality of the contacts within the RNA itself, rather than between the RNA and the ligand.

Experimental Section

Binding Assays. The apparent K_ds of the aptamers for the GTP analogues were measured in competition binding experiments using a spin filtration method[15] as previously described[5] in a buffer of 200 mM KCl, 10 mM potassium phosphate, 5 mM $MgCl_2$, 0.1 mM EDTA, pH 6.2. RNA was prepared at a concentration of 4 times the K_d for GTP (final concentration = $2 \times K_d$) in 100 μL of H_2O, heated to 80 °C for 5 min, and then equilibrated overnight at room temperature. Next, 100 μL of $2\times$ binding buffer containing 0.22 nM α-^{32}P-GTP and varying concentrations of unlabeled analogue were added. After equilibrating the samples overnight at room temperature, they were transferred to the top chamber of a Microcon YM-30 spin filter. The tubes were centrifuged at 13 000 g for 5 s to saturate the membrane. The filters were placed into new collection tubes and again spun at 13 000 g for 20 s, allowing 25 μL of solution to flow through the membrane. This process is similar to equilibrium dialysis in that the solution that remains above the molecular-weight cutoff membrane contains free GTP ($[L]_f$), free analogue ($[I]_f$), free RNA ($[R]_f$), RNA-bound GTP ($[RL]$), and RNA-bound analogue ($[RI]$); the filtrate contains only free GTP ($[L]_f$) and free analogue ($[I]_f$). A 20-μL sample was removed from each side of the filter, and the radioactivity from the α ^{32}P-GTP was measured by scintillation counting for 1 min. The radioactive counts were baseline-corrected to account for background retention in the upper chamber of the spin filter.[16] The fraction of labeled GTP bound to the aptamer, $[RL]/[L]_t$ (where the total concentration of GTP, $[L]_t = [RL] + [L]_f$), was normalized such that when the concentration of analogue competitor was zero, $[RL]/[L]_t$ equaled 100%. The K_ds were derived from a nonlinear least-squares fit of the normalized fraction of labeled GTP bound to the aptamer versus the concentration of analogue [I] using R: A Language and Environment for Statistical Computing version 2.0.0 for Linux (R Foundation for Statistical Computing, Vienna, Austria) (see Supporting Methods).

Calculating Free Energy of Secondary Structure Formation. The free energy of aptamer secondary structure formation for the 11 optimized sequences was calculated using the program RNAeval from the ViennaRNA Package version 1.6 for Intel Linux.[17] The optimized sequences were folded at 37 °C using the default settings with the secondary structure models shown in Figure 1.

(15) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425−1429.
(16) Motulsky, H.; Christopoulos, A. *Fitting Models to Biological Data using Linear and Nonlinear Regression*; Graphpad: San Diego, CA, 2003.
(17) Hofacker, I. L.; Fontana, W.; Stadler, P. F.; Bonhoeffer, L. S.; Tacker, M.; Schuster, P. *Monatsh. Chem.* **1994**, *125*, 167−188.
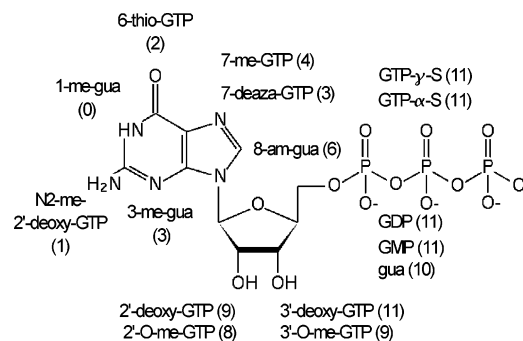
Figure 2. GTP functional groups and their importance for aptamer binding. The role of each functional group in binding was tested by comparison of analogue and GTP binding affinities. The number of aptamers (out of 11 total) that suffered no more than a 100-fold loss in affinity to the analogue (log $K_{d(Analog)}/K_{d(GTP)}$ < 2) is shown in parentheses next to each functional group. Functional groups on the nucleobase play a more significant role in aptamer binding than functional groups on the sugar and triphosphate.

To calculate the free energy of secondary structure formation predicted from sequence length, a set of 1000 random sequences was computer generated for each aptamer. The free energy of secondary structure formation for each sequence in the set was then calculated without specifying the secondary structure using the program RNAfold from the ViennaRNA package version 1.6. Sequences were folded at 37 °C using the default settings. In the same manner, 1000 sets of 11 aptamer-length sequences were generated and folded using RNAfold to estimate the likelihood that all of the aptamers in the set would, by chance, have free energies of secondary structure formation more stable than the median values for same-length random sequences.

Correlation Coefficients and Regressions. The Spearman rank correlation coefficient (r_s) was applied as a nonparametric test of association, because it does not require specific assumptions about data distribution and is insensitive to deviations from the normal distribution.[18] Confidence intervals of 95% were used for all rankings, and significance was determined under the null hypothesis of no correlation with a one-tailed test, $\alpha = 0.05$. P values were calculated from the test statistic as described.[19] To quantify the number of aptamer−analogue pairs with log $K_{d(analog)}/K_{d(GTP)}$ values at a given threshold, significance was determined with the K_ds and standard deviations obtained from nonlinear least-squares fits of the analogue binding data under the null hypothesis of no difference with a two-tailed test, $\alpha = 0.05$ (Note: none of the major conclusions would be different if smaller α values had been used). Kendall's robust line-fit method is a nonparametric approach to curve fitting that does not require independent and normally distributed data.[19] The slope is given by the median slope of all pairwise combinations of points. The intercept is the median intercept of all points computed with the median slope.

Results

Measuring Aptamer Specificity for GTP. To determine the extent to which the aptamers (Figure 1) bind specifically to GTP, we measured their affinities for two different sets of chemical analogues of GTP (Figure 2). Aptamer affinities for each analogue were obtained from competition binding assays with radiolabeled ^{32}P-GTP and unlabeled analogue. Details of the assays are described in the Experimental Section and the Supporting Information.

The 10 analogues in the first set (Supporting Figure 1A) have bulky substituents and were chosen to provide a way of

(18) Sokol, R. R.; Rohlf, J. *Biometry: The Principals and Practice of Statistics in Biological Research*, 3rd ed.; W. H. Freeman & Co.: New York, 1995.
(19) Rosner, B. A. *Fundamentals of Biostatistics*, 4th ed.; Duxbury Press: Belmont, CA, 1995.

Table 2. Aptamer Binding to GTP Analogs Bearing Bulky Substituents[a,b,c]

aptamer	GTPγ-S	GTP-α-S	2'-O-me-GTP[d,e]	3'-O-me-GTP[e]	1-me-gua[d,e]	N2-me-2'-deoxy-gua[e]	3-me-gua[e]	6-thio-GTP[d,e]	7-me-GTP[d,e]	8-amino-gua[d,e]
9−4	<1	<2	<1	<2	≫**3**	≫**3**	≫**3**	≫**3**	≫**3**	<1
CIV	<1	<2	<2	<2	≫**3**	≫**3**	≫**3**	≫**3**	≫**3**	>2
10−10	<1	<2	>**2**	<2	≫**3**	≫**3**	≫**3**	<1	<1	>**2**
Cl I	<1	<2	<2	<2	>**2**	≫**3**	<2	≫**3**	>**3**	<1
10−59	<1	<2	<1	<2	>**2**	≫**3**	<1	>**2**	≫**3**	<1
9−12	<1	<1	<2	<1	≫**3**	≫**3**	≫**3**	<2	≫**3**	>**3**
10−24	<1	<1	<2	<1	>**2**	<2	>**3**	>**3**	>**2**	<1
10−6	<1	<1	≫**3**	<1	≫**3**	≫**3**	≫**3**	≫**3**	<1	≫**3**
Cl II	<2	<2	<2	<2	≫**3**	≫**3**	≫**3**	>**2**	<2	<2
Cl IV	<1	<2	≫**3**	≫**3**	≫**3**	≫**3**	<2	≫**3**	≫**3**	<1
Cl III	<1	<1	<1	≫**2**	≫**2**	≫**2**	≫**2**	≫**2**	<2	≫**2**

[a] The order of magnitude increase in aptamer−analogue K_d compared to the aptamer−GTP K_d is presented for each aptamer−analogue pair. [b] Binding curves used to calculate the aptamer−analogue K_ds are presented in Supporting Figure 2. Supporting Table 1 gives the $K_{d(analog)}/K_{d(GTP)}$ values for this set of analogs. [c] Me: methyl; Gua: guanosine. [d] Analogues for which log $K_{d(analog)}/K_{d(GTP)}$ > 2 are highlighted with bold text. [e] Analogues for which no binding was observed are marked ≫3 or ≫2, where applicable (see text).

screening for close approaches between the aptamer RNA and the GTP ligand. The set consists of six different methylated analogues, 8-amino-guanosine, and three analogues in which a larger sulfur atom replaces an oxygen atom: 6-thio GTP, GTP-α-S, and GTP-γ-S. Such bulky analogues could have altered binding affinity compared to GTP for one, or more, of the following reasons. There may be drastic reductions in binding affinity (increases in K_d) if the aptamer cannot sterically accommodate the bulky substituent without radically restructuring the binding pocket. Relative to GTP, 3-methyl-guanosine and 7-methyl-guanosine have potential hydrogen bond acceptors that are blocked; the hydrogen-bonding ability of the sulfur-substituted analogues is also reduced.[20] Other analogues, 2'-O-methyl-GTP, 3'-O-methyl-GTP, 1-methyl-guanosine, and *N*-2-methyl-2'-deoxy-GTP, have potential hydrogen bond donors that are eliminated. If the potential hydrogen bond donor or acceptor is utilized as such in the aptamer-ligand complex, its absence should have a negative effect on affinity. Finally, with the exception of the sulfur-containing analogues, the addition of the bulky substituents changes the hydrophobicity of the ligand, where the effect on the binding affinity could be either positive or negative.[21]

The six molecules in the second set (Supporting Figure 1B) of analogues are characterized by the removal or alteration of individual functional groups. These modifications include the loss of phosphate groups, sugar hydroxyls, or the exocyclic amine, or change of N7 in the nucleobase to CH. Here, a difference in binding affinity for one of these analogues compared to GTP captures the net energetic effect of changes in solvation, hydrogen bonding, and van der Waals and electrostatic interactions.

The apparent dissociation constant (K_d) for an aptamer−ligand complex describes the ratio of the concentration of the complex to the concentrations of free aptamer and free ligand at equilibrium. Explicitly, K_d = [RNA][Ligand]/[RNA:Ligand], where the equilibrium constant is directly related to the change in standard free energy, $\Delta G = -RT \ln 1/K_d$.[22] Specificity can be measured as the difference in the free energy of binding to one ligand (L) versus another (L')[9,10]:

$$\Delta\Delta G = \Delta G_L - \Delta G_{L'} = -RT \ln \frac{K_{d(L')}}{K_{d(L)}} \qquad (1)$$

Taking

$$\log_{10} \frac{K_{d(analog)}}{K_{d(GTP)}} \qquad (2)$$

quantifies the magnitude by which an aptamer binds specifically to GTP compared to an analogue. For example, the 10−10 aptamer K_d for GTP is 30 ± 6 nM (Table 1), whereas its K_d for 2'-deoxy-GTP (6500 ± 200 nM) is more than 2 orders of magnitude higher (log $K_{d(2'-deoxy-GTP)}/K_{d(GTP)}$ = 2.3) (Table 2, Supporting Figure 1C). In other words, the 10−10 aptamer favors binding to GTP more than 200-fold compared to 2'-deoxy-GTP. Equivalently, the 10−10 specificity for GTP compared to 2'-deoxy-GTP could be expressed as $\Delta\Delta G$ = 3.2 kcal mol^{-1} at 25 °C. In contrast, the 9−4 aptamer K_d for 2'-deoxy-GTP (10 ± 3 nM) (Supporting Figure 1A) is the same as the 9−4 aptamer K_d for GTP (9 ± 1 nM) (Table 1). Hence, for the 9−4 aptamer, log K_d $G_{(2'-deoxy-GTP)}/K_{d\ (GTP)}$ and the corresponding $\Delta\Delta G$ are close to zero, indicating that the 9−4 aptamer does not bind specifically to GTP compared to 2'-deoxy-GTP.

The log $K_{d\ (analog)}/K_{d\ (GTP)}$ values for all of the aptamers are plotted in Figure 3 (also see Tables 2 and 3 and Supporting Table 1). For 34% of the aptamer−analogue pairs, aptamer binding was not very specific for GTP relative to the analogue; log $K_{d(Analog)}/K_{d(GTP)}$ was <1 and $\Delta\Delta G$ < 1.4 kcal mol^{-1}. For 38% of the aptamer−analogue pairs, aptamer binding was moderately specific for GTP relative to the analogue; log $K_{d(Analog)}/K_{d(GTP)}$ was >1 but <2 (1.4 kcal mol^{-1} < $\Delta\Delta G$ < 2.7 kcal mol^{-1}). In the remaining 28% of the aptamer−analogue pairs, binding was very specific for GTP, with log $K_{d(Analog)}/K_{d(GTP)}$ > 2 ($\Delta\Delta G$ > 2.7 kcal mol^{-1}). For many of these cases, the ^{32}P-GTP was not detectably displaced in the competition binding assays, i.e., no binding was detected, corresponding to log $K_{d(Analog)}/K_{d(GTP)}$ ≫ 3 except in the case of the low affinity Class III aptamer, for which no binding corresponds to log $K_{d(Analog)}/K_{d(GTP)}$ ≫ 2.

Overall, the aptamers have very little specificity for the ligand triphosphate region and only slightly more discrimination for

(20) Christian, E. L.; Yarus, M. *J. Mol. Biol.* **1992**, *228*, 743−758.
(21) Koizumi, M.; Breaker, R. R. *Biochemistry* **2000**, *39*, 8983−8992.
(22) Eisenberg, D.; Crothers, D. M. *Physical Chemistry with Applications to the Life Sciences*; Benjamin: Menlo Park, CA, 1979.
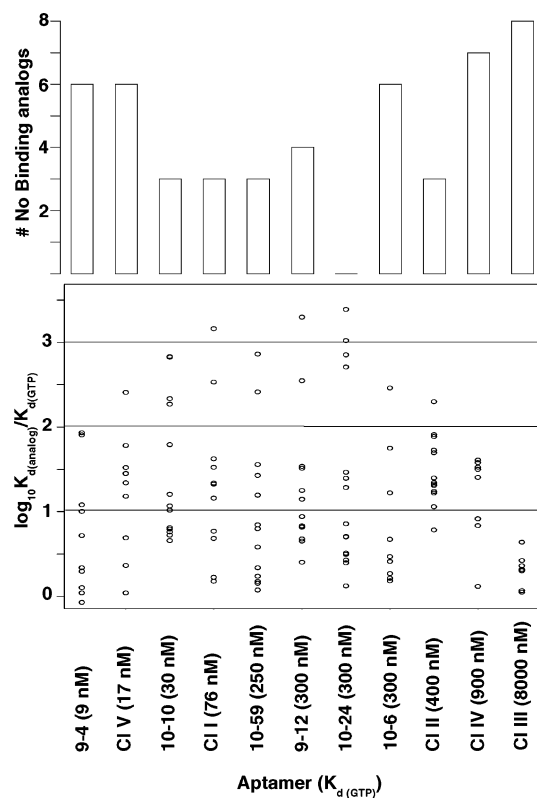
Figure 3. Aptamer specificities for GTP analogues. $\log K_{d(Analog)}/K_{d(GTP)}$, which is the magnitude by which binding is specific for GTP compared to an analogue, is plotted for each aptamer/analogue pair. The aptamers are arranged from left to right in order of affinity for GTP (shown in parentheses). For each aptamer, the number of analogues to which no binding was observed is represented in the bar graph at the top.

Table 3. Aptamer Binding to GTP Analogs with Removed or Altered Functional Groups[a,b]

aptamer	GDP	GMP	guanosine[c,d]	2′-deoxy-GTP[c]	3′-deoxy-GTP	7-deaza-GTP[c,d]
9−4	<1	<2	<1	<1	<2	≫**3**
Cl V	<1	<2	<2	<2	<1	≫**3**
10−10	<1	<1	<2	>**2**	<2	>**2**
Cl I	<1	<1	<1	<2	<1	≫**3**
10−59	<1	<1	<1	<1	<2	≫**3**
9−12	<1	<1	<2	<2	<1	>**2**
10−24	<1	<2	<1	<1	<2	<1
10−6	<1	<1	<2	>**2**	<1	<2
Cl II	<2	<2	<2	<1	<2	<2
Cl IV	<1	<1	<2	<2	<2	≫**3**
Cl III	<1	<1	≫**2**	<1	<1	≫**2**

[a] This table presents the order of magnitude increase in aptamer−analogue K_d compared to the aptamer−GTP K_d for each aptamer−analogue pair. [b] Binding curves used to calculate the aptamer−analogue K_ds are presented in Supporting Figure 2. Supporting Table 2 gives the $K_{d(analog)}/K_{d(GTP)}$ values for this set of analogs. [c] Analogues for which $\log K_{d(analog)}/K_{d(GTP)} > 2$ are highlighted with bold text. [d] Analogues for which no binding was observed are marked ≫3 or ≫2, where applicable (see text).

modifications to the ribose sugar. All of the aptamers are clearly most specific for the GTP nucleobase (Figure 2).

Only the Class III aptamer showed greatly reduced binding ($\log K_{d(Analog)}/K_{d(GTP)} > 2$) to analogues where the triphosphate had been modified or deleted. The Class III aptamer does not bind to guanosine even though the Class III aptamer K_ds for GDP, GMP, GTP-γ-S and GTP-α-S are nearly the same as the K_d for GTP (Tables 1 and 2). Note that due to limited analogue solubility, we could not accurately measure the Class III aptamer affinity for any analogue where the K_d was >2 orders of magnitude larger than the Class III aptamer K_d (8 μM) for GTP.

Nonetheless, we can conclude that the Class III aptamer K_d for guanosine is at least 2 orders of magnitude worse than the Class III K_d for GTP.

Across the entire set, the aptamers display slightly more specificity for the ligand ribose sugar than the triphosphate region. Only two of the aptamers had >100-fold increases in K_d when the 2′ OH was removed (Table 2), and none were affected that strongly by loss of the 3′ OH. Aptamers 10−10, 10−6, and Class IV bound poorly to 2′-O-methyl-GTP and neither the Class IV nor the Class III aptamers exhibited binding to 3′-O-methyl-GTP.

Modifications to the nucleobase moiety (Figure 2) caused by far the greatest effects on binding. In fact, just 25% (19 of 77) of the aptamer−analogue combinations involving analogues with nucleobase changes or deletions exhibited binding within 100-fold that of GTP ($\log K_{d(Analog)}/K_{d(GTP)} < 2$). When free in solution, guanosine nucleotides and nucleosides typically adopt syn conformations about their glycosidic bonds, in contrast to most other nucleotides and nucleosides that adopt anti conformations.[23] Inosine triphosphate (ITP) differs from GTP only in the loss of the 2-amino group, one effect of which is that ITP exists in the anti conformation.[23] The importance of the glycosidic bond conformation and the sensitivity of the selected aptamers to modifications on the ligand nucleobase are illustrated by the fact that none of the aptamers bind detectably to ITP or to ATP, CTP, TTP, or UTP (data not shown).

Distinct Patterns of Binding Specificity. We used the results from the competition binding assays to diagram the pattern of binding specificity for each aptamer (Tables 2 and 3; Figure 4). Ligand positions where modification resulted in substantially reduced binding are denoted by red shading in Figure 4. Interestingly, each aptamer displays a pattern of binding specificity that is different from the other aptamers in the set. Here, specific binding to GTP compared to a given analogue was demarcated by a $\log K_{d(Analog)}/K_{d(GTP)}$ value >2.

Changing the cutoff used to signify specific binding; for example, $\log K_{d(Analog)}/K_{d(GTP)} > 1$ alters the patterns illustrated in Figure 4. However, even with different cutoffs to signify specific binding, the aptamers appear sensitive to distinct subsets of ligand modifications (also see Supporting Figure 3). We interpret this to mean that each aptamer has a distinct ligand-binding surface.

Affinity and Specificity. We sought evidence that aptamers with high affinity for GTP are also more specific for GTP binding than aptamers with low affinity for GTP. We expected that aptamers with high affinity (lower K_d) for GTP would bind better to a smaller number of analogues than aptamers with low affinity (higher K_d) for GTP. For each aptamer, we plotted the number of analogues to which the RNA exhibited high-affinity binding relative to GTP against the aptamer K_d for GTP (Figure 5). Initially, we examined analogues with $\log K_{d(Analog)}/K_{d(GTP)}$ values <2 (colored red in Figure 5). All of the aptamers displayed binding in this range to 8−12 of the analogues. Surprisingly, we do not see any evidence that the number of analogues to which each aptamer binds in this range is correlated with binding affinity for GTP ($r_s = 0.07$, $P = 0.42$).

To ensure that choosing $\log K_{d(Analog)}/K_{d(GTP)} < 2$ as a cutoff did not obscure more subtle differences in specificity, we also

(23) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
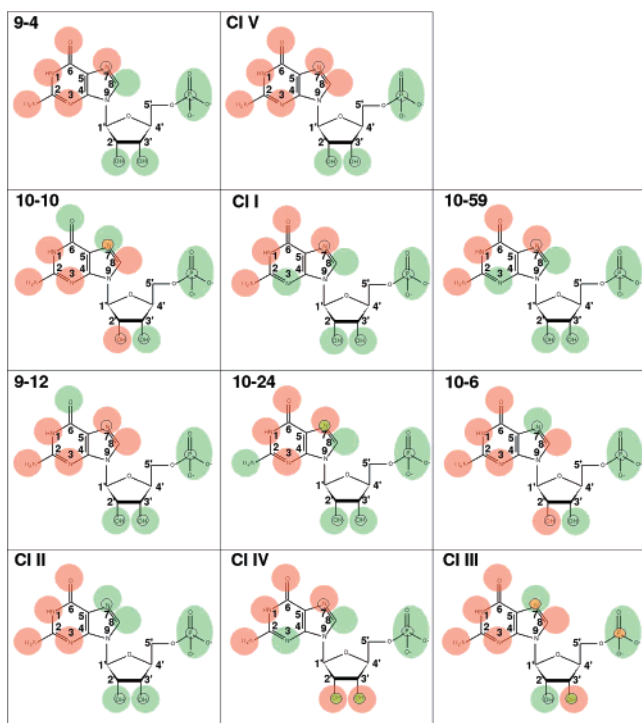
Figure 4. Each aptamer has a unique pattern of binding specificity. For each aptamer, ligand positions where chemical modification resulted in large reductions in aptamer binding affinity relative to GTP (i.e., log $K_{d(Analog)}/K_{d(GTP)} > 2$) are colored red. Ligand positions where chemical modifications did not reduce binding affinity more than 100-fold relative to GTP are colored green. The large shaded circles represent results for analogues with bulky modifications, whereas the smaller inset circles represent results obtained with modified or deleted functional groups. Only the GMP moiety of the ligand is shown here because none of the modifications to the ligand β or γ phosphates that we tested resulted in substantially reduced binding affinity relative to GTP. Also see Supporting Figure 2.

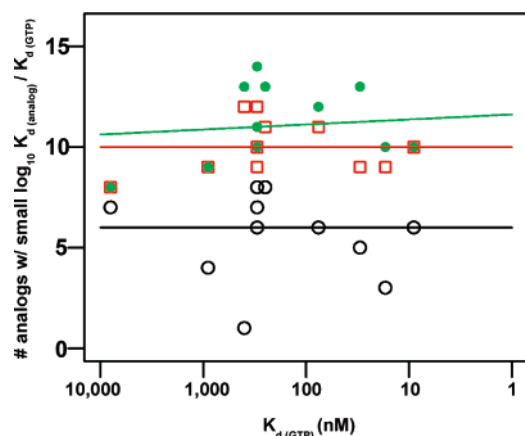**Figure 5.** Higher affinity aptamers are not more specific. For each aptamer, the number of analogues for which log $K_{d(Analog)}/K_{d(GTP)} <$ N is plotted against $K_{d\ (GTP)}$. Three separate values of N were employed to visualize different degrees of specificity: log $K_{d(Analog)}/K_{d(GTP)} < 1$ (open black circles), < 2 (open red squares), and < 3 (solid green circles). The lines were generated by applying Kendall's robust line-fit method[19] to the data indicated by the respective colors. Note that several of the plot points are the same for different values of N, resulting in overlap between red squares and solid green circles.

computed the correlation statistic using log $K_{d(Analog)}/K_{d(GTP)} < 1$ and < 3 as the cutoffs (Figure 5). There was no significant change in the correlation statistic whether the log $K_{d(Analog)}/K_{d(GTP)}$ cutoff for specific binding was <1, <2, or <3. Thus, we do not find that aptamers that bind with more affinity for
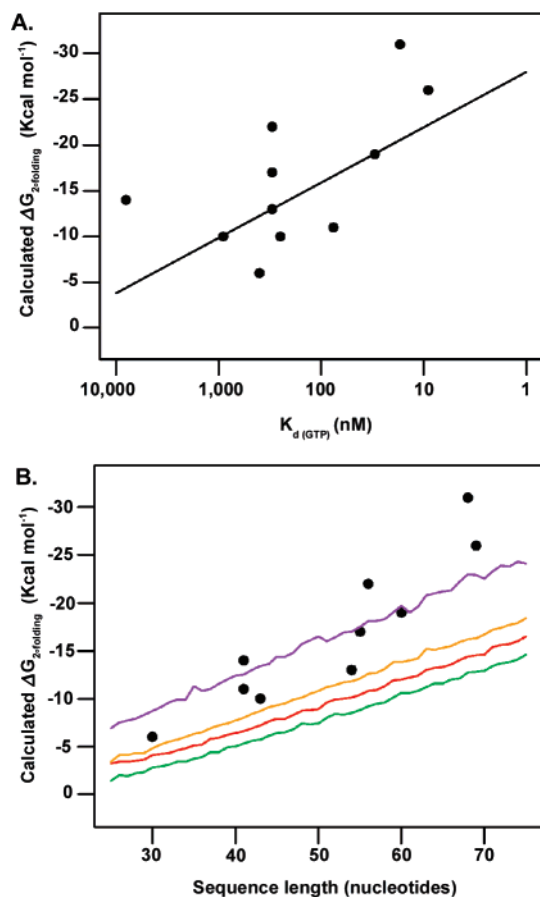
Figure 6. Aptamer secondary structures are not unusually stable. (A) Free energy of secondary structure formation ($\Delta G_{2^{\circ}(folding)}$) calculated for each aptamer using the program RNAeval (see Experimental Section) is plotted against $K_{d(GTP)}$ in nanomolar. The line was generated by applying Kendall's robust line-fit method. (B) Calculated $\Delta G_{2^{\circ}(folding)}$ for each aptamer is plotted against aptamer sequence length (●). The red line indicates the median $\Delta G_{2^{\circ}(folding)}$ values of 1000 random sequences plotted as a function of sequence length; the 33rd, 67th, and 95th percentiles are signified by the green, orange, and purple lines, respectively.

GTP are more specific for GTP binding compared to chemically related analogues.

Free Energy of Secondary-Structure Formation. Within the set of selected aptamers, the higher-affinity molecules tend to have longer sequences, are more informationally complex, and have more intricate secondary structures[6] (Figure 1; Table 1). It is known that selection for more stable RNA tertiary structures can produce RNAs with more stable secondary structures.[24] We wondered whether selection for tighter-binding yields aptamers with unusually stable secondary structures.

We calculated the free energy of secondary-structure formation using the program RNAeval[17] (Table 1). There is a significant correlation ($P < 0.02$) between the apparent K_d for GTP and the free energy of secondary-structure formation ($\Delta G_{2^{\circ}(folding)}$) (Figure 6A). We compared the calculated $\Delta G_{2^{\circ}(folding)}$ value for each aptamer with 1000 computer-generated random sequences of the same length (Figure 6B). All eleven aptamers have $\Delta G_{2^{\circ}(folding)}$ values more stable than the median same-length random sequence values. The likelihood that eleven aptamer-length sequences would have calculated $\Delta G_{2^{\circ}(folding)}$ values larger than the median by chance is very small ($P < 0.001$), suggesting

(24) Juneau, K.; Podell, E.; Harrington, D. J.; Cech, T. R. *Structure (Camb)* **2001**, *9*, 221–31.

that selection for secondary structural stability may have taken place. However, only the 9−4, Class V, 10−59, and Class III aptamers have calculated free energies of secondary structure formation that are significantly more stable than the same-length random sequences (i.e., $\Delta G_{2°\text{(folding)}}$ values more stable than 95% of the same-length random sequence values).

It is possible that the calculated $\Delta G_{2°\text{(folding)}}$ values do not accurately capture the magnitude of the actual secondary structure free energies because they do not consider the potential influence of the ligand or tertiary interactions within the RNA. Still, to a first approximation, these results indicate that selection for higher-affinity binding leads to RNAs with more stable secondary structures, but not necessarily to aptamers with secondary structures that are substantially more stable than predicted on the basis of size.

Discussion

We investigated the relationship between RNA aptamer binding affinity and binding specificity for a target ligand. The test set of evolved aptamers consisted of molecules with K_ds for GTP that range from 8 μM to 9 nM. We assessed the affinity of each of the aptamers for a series of sixteen different chemical analogues of GTP along with ATP, CTP, UTP, TTP, and ITP in competition binding assays. Aptamer specificity for GTP was quantified as log $K_{d\text{(Analog)}}/K_{d\text{(GTP)}}$. It has been argued that selection for higher-affinity binding to a ligand automatically leads to aptamers with more specificity for the ligand.[9] Surprisingly, we did not find any evidence that the higher-affinity aptamers are more specific for GTP than the lower-affinity aptamers (Table 1; Figures 3 and 5).

We examined a wide range of degrees of specificity to maximize our chances of observing an affinity−specificity correlation. We initially set the cutoff for significant differences in specificity as log $K_{d\text{(Analog)}}/K_{d\text{(GTP)}} > 2$ because previous studies have suggested that differences of that magnitude could be attributed to readily observable structural features of the aptamer−ligand complex such as a steric clash or loss of a hydrogen bond.[15,21,25−27] Solution NMR studies of the Class I GTP aptamer[28] demonstrate that analogues for which log $K_{d\text{(Analog)}}/K_{d\text{(GTP)}} > 2$ have modifications that introduce discernible steric clashes and/or the loss of hydrogen-bonding interactions with the aptamer. Likewise, biochemical and structural[15,27] studies of an RNA aptamer that binds theophylline showed that ligand modifications that result in approximately 100-fold or greater increases in K_d relative to theophylline correspond to particular intermolecular aptamer−ligand contacts. As noted above, we could not measure the Class III aptamer affinity for any analogue whose K_d was > 100-fold higher than its K_d for GTP (8 μM) because of limited analogue solubility. Therefore, both empirical observations of aptamer structures and practical considerations indicate that the most conservative treatment of the data was to use log $K_{d\text{(Analog)}}/K_{d\text{(GTP)}} > 2$ as the cutoff for meaningful differences in binding specificity. However, using log $K_{d\text{(Analog)}}/K_{d\text{(GTP)}}$ cutoffs of >1 or >3 did not lead to either quantitatively or qualitatively different outcomes (Figure 5).

The GTP analogues that we tested for aptamer binding had bulky substituents and functional group substitutions and deletions in the nucleobase, sugar, and triphosphate regions. It is possible that differences in binding specificity are limited to sites for which we could not screen the effect of chemical group modifications, such as the sugar 1′ and 4′ positions or the nucleobase 4 and 5 positions. There were also many cases where the analogues bound so poorly that the binding affinity could only be described qualitatively as "no binding detected". Neither of these caveats alters the overall picture: there is no obvious trend toward substantially more binding specificity for GTP as the aptamer affinity for GTP increases.

It is notable that all of the aptamers in our set have unique patterns of ligand-binding specificity (Figure 4; Supporting Figure 3). NMR structural studies of the Class I GTP aptamer[28] illustrate that the RNA−ligand contacts are mediated by platforms of aptamer purines that sandwich the planar portion of the ligand and a base−base interaction between the GTP and an aptamer guanosine. This general mode of recognition was seen previously in RNA aptamers that bind to ATP,[25,26,29] FMN,[30,31] and theophylline.[3,15,27] Further high-resolution structural studies will be required to determine whether this motif is present in any of the other GTP aptamers. Although some of the GTP aptamers may employ common modes of ligand binding, the fact that they have unique patterns of binding specificity implies that there are differences in the shapes and compositions of their binding pockets. These results, combined with the observation that the GTP aptamer recognition bulges do not share sequence identity[6] (Figure 1), reinforce the notion that large sets of structurally diverse aptamers can be recovered from sparse searches of sequence space.[5,6]

The GTP aptamers were isolated from pools of molecules containing long stretches of random sequence.[5] The aptamers most likely to be recovered are the simplest structures, defined as those lowest in informational complexity, with the binding affinity required to survive the process of selection.[7,32] Therefore, the physical mechanisms used to attain the required K_ds should be the simplest to implement structurally. Earlier efforts to understand the mechanisms by which aptamers achieve low K_ds focused exclusively on the energetics of the local intermolecular aptamer−ligand interactions.[3,9] Those analyses assumed that maximizing the intermolecular geometric complementarity and the quality and quantity of aptamer−ligand functional group interactions within the binding pocket was the easiest way to improve the aptamer K_d for a ligand (e.g., aligning hydrogen-bond donors within the RNA with hydrogen-bond acceptors in the ligand). All else being equal, increased aptamer−ligand steric and functional group complementarity would cause higher-affinity aptamers to bind more specifically to the target ligand than lower-affinity aptamers.[9] Conversely, because we found that higher-affinity aptamers do not bind more specifically to the target ligand (Table 1, Figure 5), we conclude that their binding pockets are unlikely to have better shape and functional group complementarity with the ligand than the binding pockets of lower-affinity aptamers.

(25) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550−553.
(26) Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. *RNA* **1996**, *2*, 628−640.
(27) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J. P.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644−649.
(28) Carothers, J. M.; Davis, J. H.; Chou, J. J.; Szostak, J. W. *RNA* **2006**, *12*, 567−579.
(29) Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. *Nature* **1996**, *382*, 183−186.
(30) Burgstaller, P.; Famulok, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1084−1087.
(31) Fan, P.; Suri, A.; Fiala, R.; Live, D.; Patel, D. *J. Mol. Biol.* **1996**, *258*, 480−500.
(32) Salehi-Ashtiani, K.; Szostak, J. W. *Nature* **2001**, *414*, 82−84.

How might aptamers achieve higher-affinity binding without an increased number of intermolecular interactions and better shape complementarity with the ligand? Obviously, the local energetics of the aptamer−ligand interaction are an important aspect of the binding affinity.[3,8,9] However, the aptamer−ligand K_d is determined by the ΔG between the bound and unbound states of the entire system, not just the local aptamer−ligand interactions.[10] In fact, structural studies have revealed that ligand binding and RNA aptamer folding are often tightly coupled,[3,26−28,33] underscoring the potential relevance of tertiary intramolecular RNA−RNA interactions in improving the aptamer−ligand K_d. We hypothesize that differences in the tertiary stabilities of the RNA folds could account for much of the observed variation in aptamer−GTP K_ds. If so, the same physical mechanisms that improve the aptamer K_d for the target ligand would also improve the aptamer K_d for chemically related ligands. In turn, this would explain why we do not see evidence that higher-affinity aptamers bind more specifically to GTP.

There is a positive correlation between aptamer−GTP K_d and the calculated ΔG of secondary-structure formation (Figure 6A). This is consistent with a role for the stability of the RNA secondary structure fold in differentiating the higher-affinity from the lower-affinity aptamers. Yet, only four of the eleven aptamers (Figure 6B) have calculated ΔGs of secondary structure formation that are significantly more stable than predicted from their sequence lengths. One possibility is that although the higher-affinity aptamers could have more stable tertiary folds, there was no direct selection pressure for more stable secondary structures.[34] Because RNA tertiary folding is highly cooperative and facilitated by nonlocal interactions,[10] it is difficult to separate the ligand binding energies from RNA folding energies except where only simple base pairing is involved.[35] Here, the analysis was restricted to consideration of the calculated free energies of secondary-structure formation. Clearly, high-resolution structure models will be necessary to fully understand the role of the RNA−RNA interactions, separate from the RNA−ligand interactions, in determining tertiary structural stability and aptamer−ligand affinities.

We know from sequence comparison data that more nucleotides in the higher-affinity GTP aptamers are subject to functional constraints.[6] The functional constraints are a consequence of the need to make intermolecular contacts with the ligand in the context of a folded RNA secondary and tertiary structure. For this set of aptamers, the informational cost to improve binding affinity 10-fold is approximately 10 bits, enough to specify 5 conserved positions or 5 base pairs in an RNA structure in the conditions of the original selection.[5] The findings presented here suggest that increasing the RNA shape and functional group complementarity for the ligand is not the simplest way to improve binding affinity. Rather, the informational cost of improving aptamer affinity may be a function of the increased complexity required to form more stable tertiary RNA structures.

One prediction arising from the suggestion that higher-affinity binding may be achieved through the formation of more stable tertiary structures is that selection for stability could inadvertently result in molecules that are more specific for the target ligand. Selection for stability could produce a molecule with more rigidity in the elements surrounding a binding pocket, eliminating the structural flexibility needed to accommodate chemically related analogues of a given ligand.[36] Consistent with this expectation, a nonbiological ATP binding protein subjected to selection for increased stability in the presence of a chemical denaturant showed more specific binding to ATP when tested with a panel of analogues.[36]

We have previously suggested[6] that RNA aptamers that bind planar, polar ligands, such as GTP, require less sequence information to specify than structures that bind flexible, nonpolar targets with the same affinity in similar environments. We further argued that a particular degree of improvement in binding affinity could have the same informational cost, regardless of the chemical nature of the ligand, if increasing the number of interactions within the RNA, rather than between the RNA and the ligand, is generally the simplest way to improve affinity. Determining whether the observations and conclusions made here about GTP can be extended to other nonplanar, or nonpolar ligands will be an important test of this hypothesis.

Increasing aptamer affinity by stabilizing the tertiary fold of the RNA, or by improving the quality of the interactions with the ligand, does not seem mutually exclusive. Why, then, is more specific binding, predicted from the latter process, not observed in our set of aptamers? We propose several potentially relevant factors. Across the whole set of aptamers, (Figures 2 and 4) binding is greatly affected by modifications to the ligand nucleobase. It may be difficult for additional RNA−ligand nucleobase interactions to come into play. At the same time, modifications to the ligand ribose or triphosphate frequently have only small negative effects on binding. Because the aptamers were selected from a random sequence pool, we only see the most likely solutions. Improving aptamer K_d by increasing the tertiary stability of the RNA must require structures that are less informationally complex than those needed to improve aptamer K_d by interacting more strongly with the sugar or triphosphate moieties of the ligand.[37] Although higher-affinity aptamers that bind more specifically to the target ligand may exist in sequence space, their greater informational complexities mean that they can only be isolated by selecting for both affinity and specificity.

Supporting Information Available: Methods for K_d determination (Supporting Methods). Chemical analogues of GTP employed in this study (Supporting Figure 1). Aptamer−analogue competition binding curves (Supporting Figure 2).

(33) Sussman, D.; Nix, J. C.; Wilson, C. *Nat. Struct. Biol.* **2000**, *7*, 53−57.
(34) Guo, F.; Ceck, T. R. *Nat. Struct. Biol.* **2002**, *9*, 855−861.
(35) Hertel, K. J.; Peracchi, A.; Uhlenbeck, O. C.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8497−8502.
(36) Chaput, J. C.; Szostak, J. W. *Chem. Biol.* **2004**, *11*, 865−874.
(37) Sazani, P. L.; Larralde, R.; Szostak, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 8370−8371.

Aptamer patterns of binding specificity (Supporting Figure 3). $K_{d^-(\text{Analog})}/K_{d(\text{GTP})}$ values for analogues bearing bulky substituents (Supporting Table 1). $K_{d(\text{Analog})}/K_{d(\text{GTP})}$ values for analogues with modified or deleted functional groups (Supporting Table 2). This information is available free of charge via the Internet at http://pubs.acs.org.

JA060952Q