4. Results

4.1 Testing the Effectiveness of Enriching Functional RNA Sequences to Optimize Selection Conditions

Special selection strategies are required for isolating high-information content aptamers and disfavoring minimally complex solutions. Experiments were therefore conducted to establish optimal selection conditions for recovering the desired high-affinity aptamers.

4.1.1 Inhibition of Aptamer Column Binding Activity by Pool RNA

In order to examine the extent of the inhibition of aptamer function caused by annealing of pool RNA to aptamer RNA, the following experiments were performed:

10 nM ^{32}P-labeled RNA of aptamer 9-4 (the GTP aptamer with the highest affinity) was denatured and incubated with GTP immobilized on thiopropyl sepharose beads for 24 h at room temperature alone and together with 10 μM unlabeled pool RNA respectively. When the column-binding behavior of the two samples was tested, incubation of both components together resulted in an almost 95% inhibition compared to the column-binding behavior of aptamer RNA alone (44.70% of 9-4 were eluted after 10 column volumes of washing (see green curve in Fig. 8) compared to 2.60% after incubation together with pool RNA; see blue curve in Fig. 8).

In order to counter this effect, 0.2 nM labeled 9-4 RNA in 0.2 μM pool RNA were denatured and annealed together in a large volume (10 ml) of 1x binding buffer. After the volume was reduced to 200 μl by spinfiltration, the resulting mixture of 10 nM labeled 9-4 RNA in 10 μM pool RNA was incubated with GTP immobilized on thiopropyl sepharose. Even though approximately 40% of the RNA was lost during the volume-reduction process, column binding of 9-4 could be restored to approximately half the level observed without pool RNA (22.36%, see pink curve in Fig. 8). In addition,

undesired unspecific binding of aptamer RNA with the column material was reduced to a background level with this incubation method (see resin fraction, pink curve, Fig. 8):

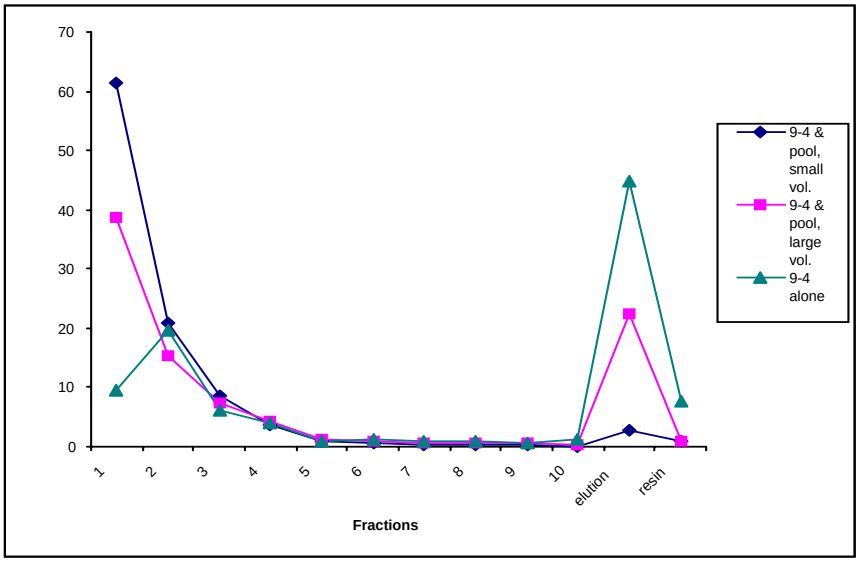


Fig. 8: Inhibition of Aptamer Activity by Pool RNA. Elution profile of aptamer 9-4 RNA when incubated without pool RNA (green curve), together with pool RNA and incubation in a large volume (pink curve) and 9-4 RNA together with pool RNA after incubation in a small volume (blue curve). Fractions 1-10: 1-column volume wash fractions; elution: elution fraction; resin: RNA remaining bound to the column material after elution.

4.1.2 Pre-Column and Pre-Incubation of the Column Material with tRNA

Using a pre-column containing uncharged thiopropyl sepharose and incubation of both charged and uncharged column material with tRNA could lower unspecific binding of aptamer RNA to the column material by more than two-fold (data not shown).

4.1.3 Amount of Washing

Increasing the amount of washing with 1x binding buffer after incubation of the RNA with GTP immobilized on the column material from 10 to 200 column volumes resulted in a 4-fold improvement of the ratio of specific to unspecific binding (data not shown).

4.1.4 Examining Aptamer Association and Dissociation

In order to test the optimal incubation and elution times to recover high-affinity aptamers such as 9-4, its association and dissociation were measured by spinfiltration and column binding.

Unlabeled 9-4 RNA was denatured as described above, and an equal volume of 2x binding buffer was added containing ^{32}P-labeled GTP. After incubation for up to 4 h at RT, the amount of GTP bound to 9-4 RNA was determined by spinfiltration. After 4 h of incubation, still only slightly more than 70% of GTP was bound to the RNA (see Fig. 9).

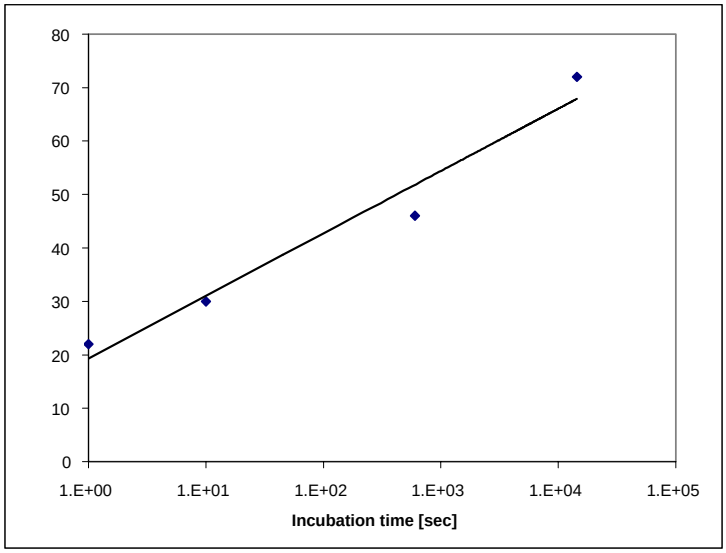


Fig. 9: Testing the Association of 9-4 RNA with the Spinfiltration Method. The incubation time before eluting is plotted against the percentage of labeled GTP bound to 9-4 RNA.

^{32}P-labeled pool RNA was denatured and incubated with GTP immobilized on thiopropyl sepharose and its incubation time before elution was compared to equally treated 9-4 RNA (see Fig. 10). A more than 442-fold enrichment of aptamer RNA compared to pool RNA could be observed.

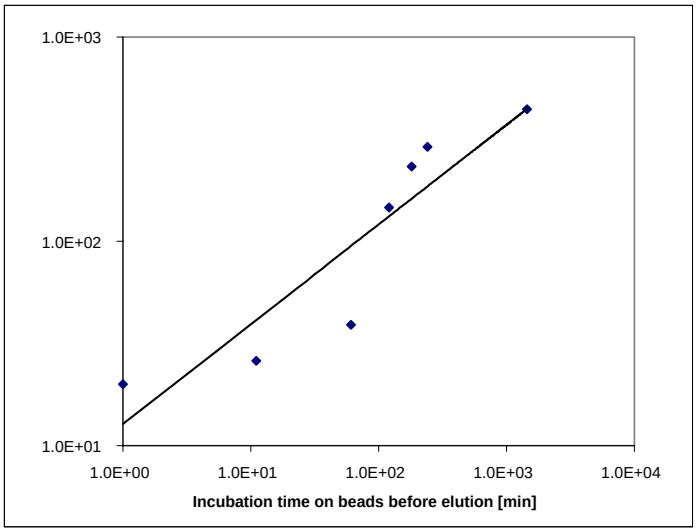


Fig. 10: Testing the Incubation Time before Elution of 9-4 RNA with the Column Binding Method. The incubation time before elution is plotted against the enrichment of aptamer vs. pool RNA.

To determine the amount of time that is necessary for 9-4 RNA to elute from GTP, ^{32}P-labeled 9-4 RNA was denatured and incubated with GTP immobilized on thiopropyl sepharose as described above for up to 24 h and its elution profile was compared to that of labeled pool RNA treated in the same way (see Fig. 11). A 577-fold enrichment of aptamer RNA compared to pool RNA could be observed.

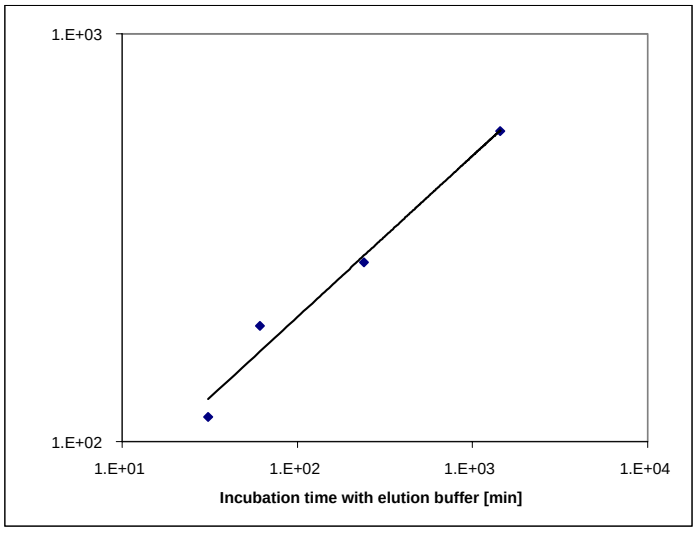


Fig. 11: Testing the Elution Time of 9-4 RNA with the Column Binding Method. The incubation time with elution buffer is plotted against the enrichment of aptamer vs. pool RNA.

Both methods revealed that at a least a 24 h-incubation and elution were required for a maximum of the aptamer RNA to bind to its target molecule and to competitively elute with free GTP.

4.2 Characterization of GTP Aptamers

4.2.1 GTP Selection

In order to isolate high-affinity GTP aptamers, a GTP selection starting with $5x10^{14}$ random sequences was performed in ten rounds. Analysis of these sequences revealed no dominant sequence, class, or motif, but a wide variety of unrelated sequences, yielding seven aptamers with dissociation constants ranging from 25 to 500 nM (Davis and Szostak, 2002)[1].

Additional sequences derived from rounds seven to ten of this selection were characterized in order to identify further distinct GTP aptamers with a wider range of affinities and to follow the presence of known aptamers throughout the selection. 249 additional sequences were obtained based on their multiple independent occurrence in the pool (71 from round 7; 58 from round 8; 120 from rounds 9 and 10). For more than half of these (151 of 249 sequences), GTP solution binding experiments were conducted, leading to the identification of four additional aptamers. Together with the seven previously obtained aptamers, these eleven aptamers represented more than half of all active selected sequences and most likely the majority of the simplest solutions to the problem of moderate- to high-affinity GTP binding.

4.2.2 Aptamer Optimization

In order to determine the optimal sequence and secondary structure of the selected aptamers and to calculate the informational and functional complexity of each optimized

[1] The GTP selection performed by Jonathan Davis (Davis and Szostak, 2002) served as the basis for this project. Aptamer optimization and characterization was performed in collaboration with James Carothers. My main contribution was the complete optimization of the Class II and Class III aptamers and significant involvement in the optimization process of the remaining aptamers.

aptamer, the following steps were undertaken: First, the K_ds and column binding behaviors of the aptamers were examined. End mapping experiments were then used to define the minimal sequences with conserved domains necessary for binding to GTP.

4.2.2.1 Doped Re-Selections

The starting pool of a selection usually represents only a small fraction of all sequences that are theoretically possible to synthesize, in this case $5x10^{14}$ out of 10^{38} possible sequences. The sequences obtained after *in vitro* selection with such a pool represent a small set of most likely sub-optimal solutions that can however be regarded as a valid basis for further investigation by *in vitro* selection. Since most RNA pools for *in vitro* selections only search a small portion of sequence space, re-selections are usually carried out as a second step to further refine the search and identify the best possible sequence for any given aptamer. Therefore re-selections were performed with mutagenized sequence libraries based on the minimal functional region of each aptamer. From these libraries, approximately 10^{12} independent sequences were transcribed, which allowed the almost complete search of the sequence neighborhood surrounding each original isolate to a distance of about five mutations and sparser sampling of the surrounding region.

All doped pools were enriched for functional sequences in three to seven rounds of re-selection. Re-selections were conducted until either more than 20% of the RNA bound and eluted or when the same amount of binding was observed in three consecutive rounds. The selection protocol was similar to that of the original selection (Davis and Szostak, 2002), but lacked the increase in stringency. 30-80 clones from each re-selection were then sequenced and aligned (see examples in Figs. 13 and 14).

4.2.2.2 Aptamer Sequence Alignments and Comparisons

A number of Watson-Crick and GU wobble-base pair covariations were observed in the alignments of sequences resulting from the doped re-selections. These permitted the prediction of base-paired regions and the generation of secondary structure models

(Fig. 15), since covariation confirms the existence of stems. All models were further characterized by examining mutations by solution and column binding. The observation of disrupted base pairs in the sequence alignments suggested that the stems of some aptamers were longer than necessary to survive the low stringency of the re-selection, since shorter stems would most likely have provided sufficient stability. In order to rule out limitations that stem lengths might impose on binding affinity, the outer stem of any construct was extended by three GC base pairs. Any further additions to the stems did not improve aptamer activity. Aptamers Class II and Class III will serve as examples for elucidating the aptamer optimization process.

4.2.2.2.1 Optimization of the Class II Aptamer

The secondary structure of the Class II aptamer was determined by alignment of 48 independent sequences derived from the doped re-selection (Fig. 12). This aptamer consists of a stem-loop structure with possibly two GC-base pairs inserted into the loop (Fig. 15). The alignment in Fig. 12 shows that eleven positions in the binding loops are completely conserved (marked in orange in the parent sequence at the top of Fig. 12), which suggests that these bases must be invariant for the aptamer to be functional. Seven additional positions are restricted to just two bases (marked in green in the parent sequence at the top of Fig. 12). Alignment of the Class II sequences revealed the high sequence conservation of the loop. The repeated covariation (Fig. 12, marked in red) provides strong evidence for the functional significance of the stem.

Interestingly, a mutation from G to A in position 6 corresponds to the same mutation in position 25 (see Fig. 12, marked in blue) and these mutations occurred together at a significantly higher rate than expected purely by chance. When testing the equilibrium dissociation constants of sequences with G-A mutations by spinfiltration, a more than 10-fold improvement was observed with As in positions 6 and 25 compared to Gs. G in either position combined with A in the other position led to a 15-fold decrease in K_d. When examining the remainder of the loop sequence to obtain the optimal sequence, mutations from U to A and C respectively in position 17 were tested due to the high

mutation rate observed in this position, but exhibited a 2.5-fold decrease in K_d compared to U. In addition, when examining the insertion of a U between positions 7 and 8, binding activity of the aptamer was abolished (data not shown).

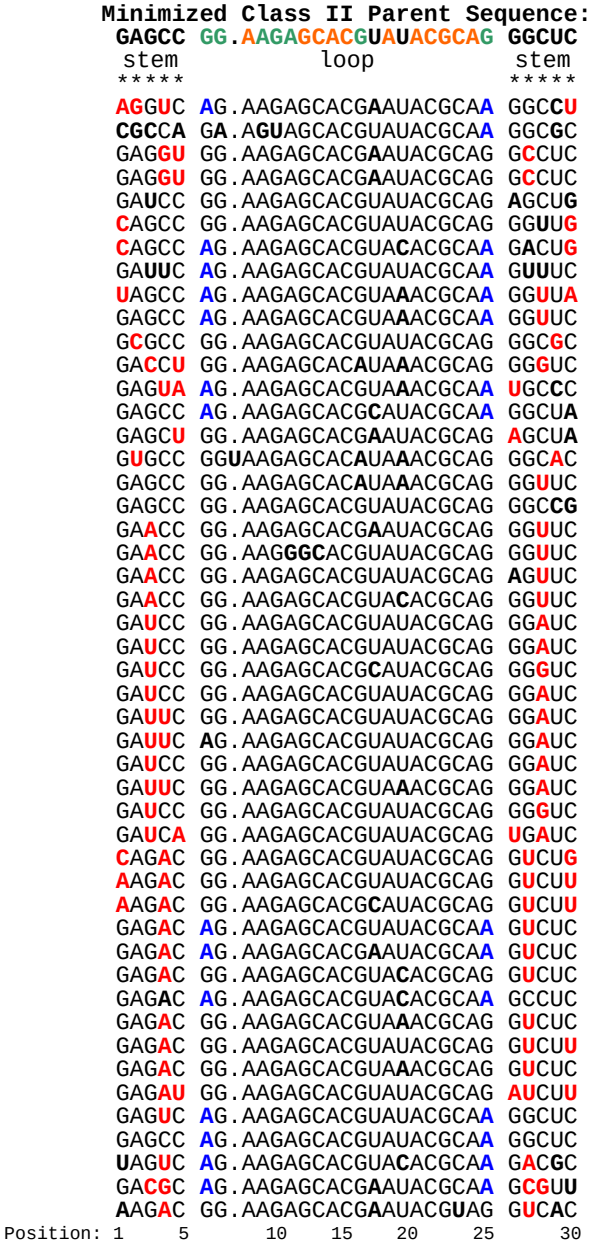


Fig. 12: Class II Re-Selection Sequence Alignment. The minimal-construct sequence doped for the re-selection is shown at the top in bold. Mutations that differ from the parent sequence are marked in bold font, covariations in red. Bases that form Watson-Crick or GU wobble-base pairs in the stems are marked by asterisks (*). The position of stems and loops is also indicated ("stem" / "loop"). Co-mutations are marked in blue. Invariant positions are marked in orange, positions that are restricted to two bases are shown in green in the parent sequence at the top.

4.2.2.2.2 Optimization of the Class III Aptamer

The secondary structure of the Class III aptamer was determined by alignment of 38 independent sequences derived from the re-selection (Fig. 13). This aptamer consists of a stem, a bulge and a pre-engineered stem-loop (Fig. 15). The high sequence conservation of the loop underlines its importance for the proper function of the aptamer. The repeated covariation (Fig. 13, marked in red), including a selected mutation from CUGC - CAGU to CGUC - GACG (positions 9-12 / 17-20) in order to extend the stem, provides strong evidence for the functional significance of the stem. Eight positions in the binding loops are completely conserved (marked in orange in the parent sequence at the top of Fig. 13), which suggests that these bases must be invariant for the aptamer to be functional. Ten additional positions are restricted to just two bases (marked in green in the parent sequence at the top of Fig. 13).

When testing point mutations in aptamer loop regions, aptamer binding to GTP was completely abolished. The following mutations were examined by spinfiltration (data not shown):

- insertion of a U between positions 7 and 8 of Class II
- G→U mutation in position 23 of Class III
- C→U mutation in position 24 of Class III
- A→U mutation in position 29 of Class III

In contrast, positively selected variations that occurred more frequently than would have been predicted by chance were observed in the loop regions, and when tested generally improved the binding affinity of the aptamers. Examples are:

- G→A mutation in position 6 of Class II (10-fold improvement)
- G→A mutation in position 25 of Class II (10-fold improvement)
- A→U mutation in position 8 of Class III (5-fold improvement)
- CUGC→CGUC mutations in positions 9-12 of Class III (2-fold improvement)
- CAGU→GACG mutations in positions 17-20 of Class III (2-fold improvement)
- C→G mutation in position 34 of Class III (2-fold improvement)

Minimized Class III Parent Sequence:

```
CCCAUGA A CUGC UUCG CAGU UUGCUAAAAACCAC UCGUGGG
  stem     stem loop stem      loop       stem
*******   ****      ****               *******
UCCAUGA U CGUC CUCG CACG UUGCUAAAAACCAG UCGUGGA
UCCAUGA U CGUC CUCG CACG UUGCUAAAAACCAG UCGUGGA
CCCAUGA U CGUC UGCG UACG UUGCUAGAAACCAG UCGUGGG
CCCAAGA U CAGC GUCG UCUG UUGCUAAAAACCAG UCUUGGG
CUCACGA U CAUC UUCG AAUG UUGCUAAAAACCAG UCUGUUU
CCCUUGA U CAGC UUUG .CUG UUGCUUAAAACCAG UCUUGGG
CCCAUUU U GUGC UUCG CAGU UUUCUAAAUACAAG UCGUAGU
UCCAUGA U CGUC CUCG CACG UUGCUAAAAACCAG UCGUGGA
CCCGUGA U CGCC UUCG CACG UUGCUAAAAACCAG UCGCGGG
CGCAUGA U CGAC UUCG AUCG UUGCUAAAAACCAG UCGUGUG
CUUAGGA U CAGC UUUG CCUG UUGCUAAAAACCAG UCCUGGC
GCCAUGA U CGUC UUCG AACG UUGCUAAAAACCAG UUGUGGC
CCCAUUA A CAGC UUCG CCCG UUGCUAGAAACCAG UGUAGAG
ACCAUGA A CUCC UCCG CCCG UUGCUAAAAACCAA UUGUGGG
CCAUUGC A CGGA UCCG CCCG UUGCUAAAAACCAG UGUGGAG
CACAUGA U CGUC CUUG CACG UUGCUAAAAACCAG UCUGUGU
CACAUGA U CAGC UUAA UCUG UUGCUAAAAACCAG UCUGUGU
CUCAUGA A CGGU UUGG CCCG UUGCUAAAAACCAG UGUGGGU
CCCCUGA A CUGA UUUG UCCG UUGCUAGAAACCAG UGUGGGU
CCCAUGA U CAUC UUCG CAUG UUGCUAAAAACCAG UCCCGGG
CCCAUGA U CAUC UUCG CAUG UUGCUAAAAACCAG UCCCGGG
CCCAUGA U CAGC UUCG .CUG UUGCUAAAAACCAG UCAUGGG
CCCAUGA U CAGC UUCG .CUG UUGCUAAAAACCAG UCAUGGG
CCCUUGA U CAGC UUCG .CUG UUGCUAAAAACCAG UCGUGGG
CCCAUGA U CAGC UUCG CCUG UUGCUAAAAACCAG UCGUGGG
CCCAAGA U CAGC UUCG .CUG UUGCUCAAAACCAG UCGUGGG
CCCCUGA U CAUC UUUG AAUG UUGCUAAAAACCAG UCGAGGG
CCAAGGA U CAGC UUCG CCUG UUGCUAAAAACCAG UCCUGGU
CUCAUGA U CAGC CUCG CCUG UUGCUAAAAACCAG UCGUGAG
CCCAUGA U CAGC CAUC GCUG UUGCUAAAAACCAG UCAUGAG
CUCAUGA U CGUC UUCG CACG UUGCUAAAAACCAG UCGUGAG
CCCAUGA U CGUC AUCG CACG UUGCUAAAAACCAG UCUGUGG
CCCAUGA U CGCC UUCG UACG UUGCUAAAAACCAG UCGUGGG
UCCACGA U CGUC UCCG CACG UUGCUAAAAACCAG UCGUGGG
CACAUGA U CGUC UUCG CACG UUGCUAAAAACCAG UCGUGUG
CUCAUGA U CGUC AUCG UACG UUGCUAAAAACCAG UCGUGAG
CCAUUGA U CGUC AUCG CACG UUGUUAAAAACCAG UCAUCGG
CCC...A U CGAC UACG UUCG UUGCUAAAAACCAG U...GGG
Position: 1    5      10     15    20     25   30     35    40
```

Fig. 13: Class III Re-Selection Sequence Alignment. The minimal-construct sequence doped for the re-selection is shown at the top in bold. Mutations that differ from the parent sequence are marked in bold font, covariations in red. Bases that form Watson-Crick or GU wobble-base pairs in the stems are marked by asterisks (*). The position of stems and loops is also indicated (“stem” / “loop”). Invariant positions are marked in orange, positions that are restricted to two bases are shown in green in the parent sequence at the top.

4.2.2.2.3 Results of Aptamer Optimization

In the following, aptamers Class II and Class III again serve as examples for illustrating the results of sequence optimization of all GTP aptamers. Fig. 14 contains three DNA sequences for each type of aptamer. The top sequence ("Original Sequence") represents the sequence of the clone that was identified from the original GTP aptamer selection (Davis and Szostak, 2002). The middle sequence ("Minimized Construct") is the minimal sequence required for binding to GTP after endmapping and was employed as the template for the re-selection. The last sequence ("Optimized Construct") represents the optimized version of each aptamer type. The corresponding apparent dissociation constants are also noted. Sequence changes compared to the parent sequence are marked in red.

Class II

	K_d (µM)
Original Sequence:	
GGGACGGCTGCTAGCGCGGTAGAAAACCGAGCCGGAAGAGCACGTATACGCAGGGCTCAACTACA	4.6
Minimized Construct:	
GGGACGCTGAACGAGCCGGAAGAGCACGTATACGCAGGGCTCAACTAG	2.5
Optimized Construct:	
GGGAGCCAGAAGAGCACGTATACGCAAGGCTC	0.4

Class III

	K_d (µM)
Original Sequence:	
GGGACACCCTAAAAGGATACCCATGAACTGCTTCGCAGTTTGCTAAAAACCACTCGTGGGTACCT	112.0
Minimized Construct:	
GGGACACCCTAAAAGGATACCCATGAACTGCTTCGCAGTTTGCTAAAAACCACTCGTGGGTACCT	112.0
Optimized Construct:	
GGGATGATCGTCTTCGCACGTTGCTAAAAACCAGTCATCCC	8.0

Fig. 14: Results of Sequence Optimization. Original sequence, minimized construct (after endmapping), optimized construct (after re-selection and testing of constructs) and apparent K_ds for Class II and Class III. Sequence changes relative to the original isolate are marked in red. Nucleotides added at the 5' end to increase transcriptional yield are shown in blue.

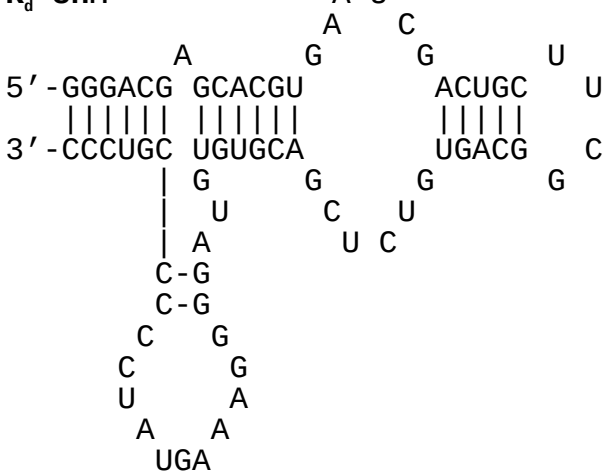
The optimization procedure showed different degrees of efficiency for different aptamers: relative to the original isolate, aptamers 10-10, 10-59, and 9-12 showed no increase in binding affinity, while 9-4, Class I, 10-24, 10-6, and Class IV improved slightly (2-3 fold). However, optimization of three other aptamers (Class II, III and V) resulted in larger improvements of activity of 12, 14 and 235-fold, respectively (data not shown). This was expected since the initial isolates were significantly sub-optimal,

because only a minute portion of sequence space had been sampled by the original selection. The starting library contained $5x10^{14}$ different molecules, representing only a small fraction of the total number of possible sequences with a length of 64 bases ($4^{64} = 10^{38}$ possible sequences). Therefore, except for those aptamers with structures simple enough to have been isolated independently multiple times, the initial pool was not expected to contain sequences with maximum activity.

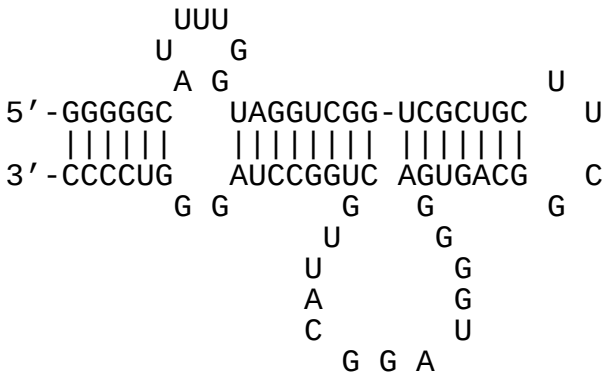
4.2.2.2.4 Optimized GTP Aptamers

Sequence comparison yielded insights into which parts of the aptamer require complete conservation, while others are permitted to vary without affecting aptamer function. In addition, covariation of bases confirms the existence of a stem. Models of optimized structures could thus be devised and tested in order to further optimize the aptamers. After their best functioning sequence, secondary structure and apparent K_d had been determined, 11 GTP aptamers were identified whose affinity spread over almost 3 orders of magnitude (9 nM - 8μM). These values represent apparent K_ds, which result in actual K_ds when being divided by the fraction of the molecules that is correctly folded. Aptamers derived from sequences that are similar, but had evolved independently in the selection as proven by diverse flanking sequences, were named with the "Class" nomenclature, whereas the names of those for which only a single original sequence could be identified reflected the round of the original selection in which these aptamers had first been observed (aptamer "9-4" was for instance first observed in round 9; see Fig. 15).

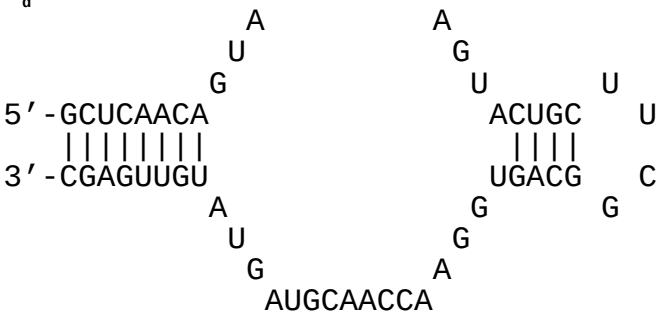
9-4
K_d=8nM



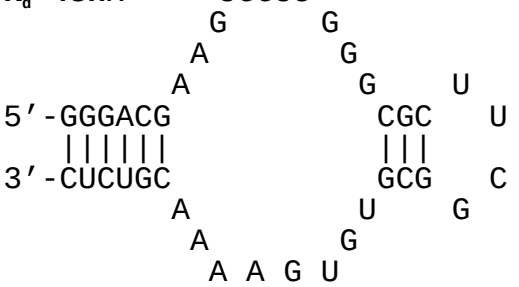
Class V
K_d=13nM



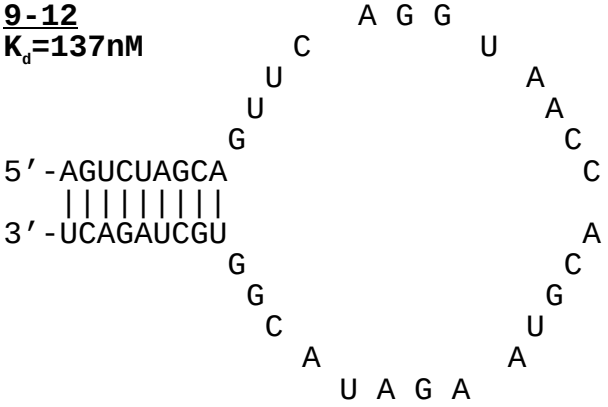
10-10
K_d=18nM



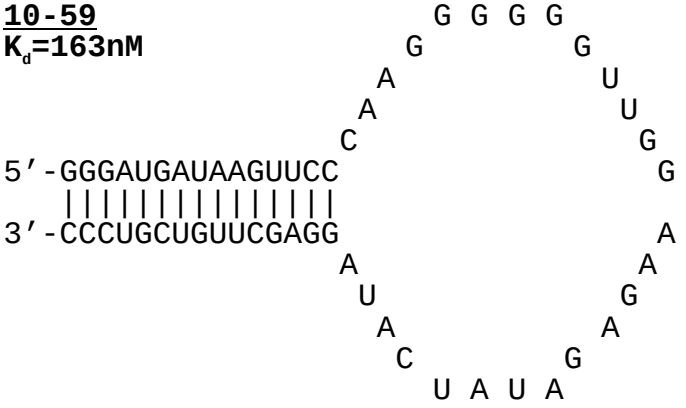
Class I
K_d=43nM



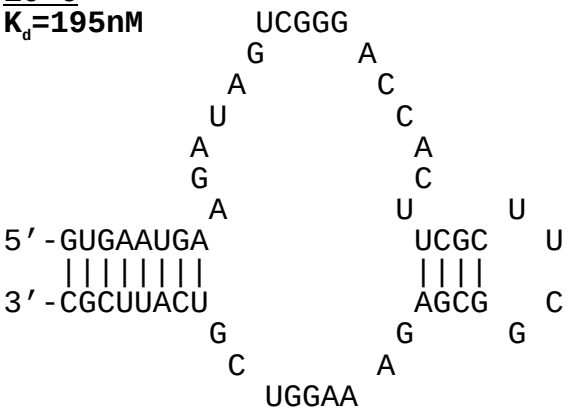
9-12
K_d=137nM



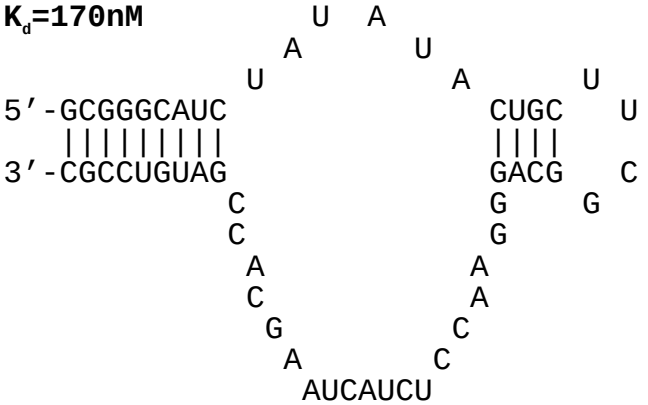
10-59
K_d=163nM



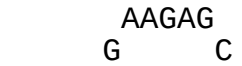
10-6
K_d=195nM



10-24
K_d=170nM



Class II
K_d=248nM



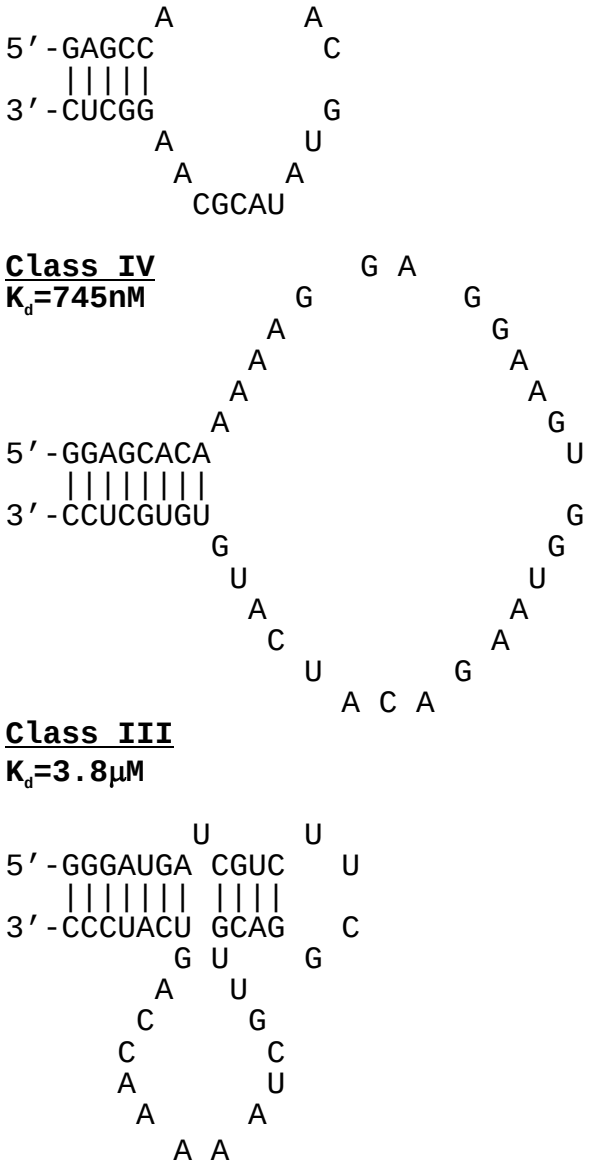


Fig. 15: Proposed Secondary Structures and Real K_d Values for GTP Aptamers.

4.3 Aptamer Affinity and Stability

To examine the stability of the aptamers and of the complex that they form with GTP, and to test the hypothesis that aptamers with better binding activity make more and stronger contacts with their ligand, the following aptamer parameters were investigated:

- equilibrium dissociation constants
- fraction of correctly folded aptamer
- free energy of secondary structure formation
- number and intensity of contacts that the aptamer is making with GTP
- information content

4.3.1 RNA Secondary Structure Prediction

When using the program mfold (Zuker, 2003) for RNA secondary structure prediction, structures were returned that showed high similarity to those obtained by end mapping and subsequent sequence alignment and testing of optimized sequences after the doped re-selection. In general, mfold however produced structures with additional base pairs in loop regions that are most likely unpaired due to the lack of covariation observed by sequence comparison. In addition, inserting very few base pairs would not have a sufficient stabilizing effect in the middle of a loop to justify their formation.

While yielding satisfying results when asked to fold optimized structures, the program is however not returning the correctly folded aptamer structures when given a non-optimized aptamer sequence from the original selection as a starting point. This result is most likely due to the tendency of the algorithm to produce a secondary structure with the highest possible amount of base pairing. To improve the program, its emphasis should therefore be shifted from an attempt to form the largest number of base pairs even in limited numbers inside a loop towards assigning a higher value and accuracy in predicting base stacking and the contribution of loops to the free energy of secondary structure formation. Mfold is thus only of limited use for sequence optimization purposes, and therefore end mapping and sequence comparison have to be performed to determine the optimized minimal structure of RNA aptamers.

In addition, while providing important insights into energetic contributions to RNA secondary structure, the program does not allow any conclusions about RNA tertiary structure.

4.3.2 Equilibrium Dissociation Constants and Aptamer Fraction Correctly Folded

After a significant correlation between aptamer size and information content (P<0.025) and aptamer K_d and information content (P<0.050) had been shown previously (Carothers et al., 2003), it was tested whether a correlation could be established between K_d and the fraction of each aptamer that is correctly folded.

Apparent dissociation constants were determined by measuring the fraction of trace ligand bound as a function of the RNA concentration. These values represent the quotient of the K_d measured with correctly folded RNA and the fraction of correctly folded RNA (see Methods section and Table 1).

	9-4	Class V	10-10	Class I	9-12	10-59	10-24	10-6	Class II	Class IV	Class III
apparent Kd [nM]	9	17	30	76	300	240	300	300	400	900	8000
real Kd [nM]	7.86	13.26	17.56	43.19	136.81	163.24	195.26	170.38	247.95	744.79	3800
IC [bits]	86.31	56.58	81.3	47.9	57.44	59.37	52.56	70.66	38.63	44.36	41.69
% folded	89.9	73.59	63.96	58.26	53.85	64.05	66.06	57.87	60.77	37.57	44.1
total size [bp]	69	68	60	41	43	50	55	54	30	43	41

Table 1: Aptamer Parameters: Apparent K_d [nM]: Quotient of real K_d and fraction of correctly folded RNA; Real K_d [nM]; IC: Information content [bits]; % folded: Fraction of correctly folded aptamer RNA [%]; Total size [base pairs].

Aptamers with high affinity (9-4 and Class V) showed a large percentage of folding into their correct secondary and tertiary structure already prior to binding to GTP, and the contrary was true for lower-affinity aptamers (Class III and Class IV). However, aptamers 10-10 and Class I, two well-performing aptamers, showed only intermediate folding fractions similar to aptamers with up to 10-times decreased affinity, such as 9-12, 10-59, 10-24, 10-6 and Class II (Fig. 16). These results suggest that the fraction correctly folded does contribute to the improved K_d and higher information content, but additional factors are also part of improving aptamer performance.

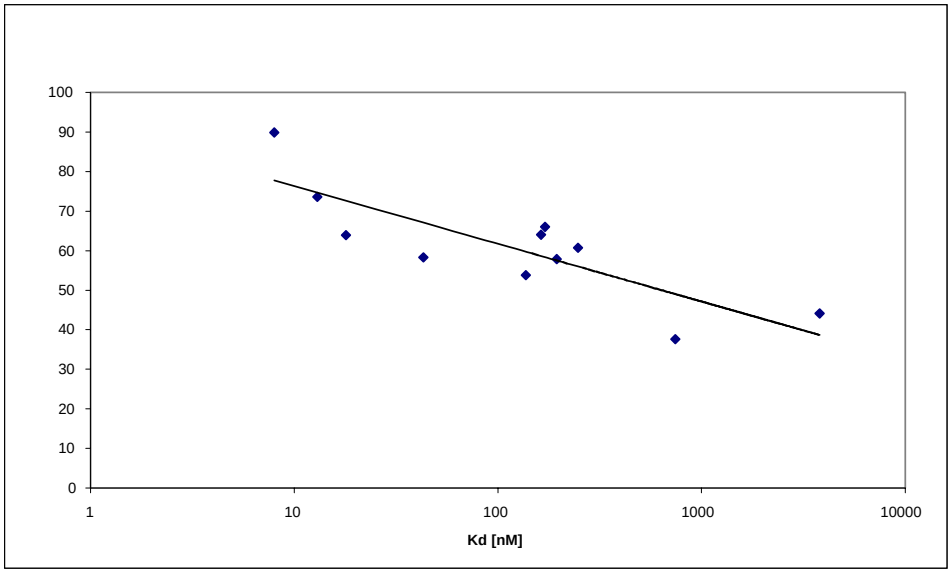


Fig. 16: Relationship of K_d [nM] and Aptamer Fraction Correctly Folded.

4.3.3 Relationship of Aptamer Quality and Free Energy of Secondary Structure Formation

Since the relationship between K_d and aptamer fraction correctly folded (Fig. 16) suggested that both aptamer stability and stability of the complex formed with GTP play a role for aptamer quality, we also examined the relationship between free energy of secondary structure formation and aptamer performance.

Free energies of secondary structure formation of evolved RNA sequences from the GTP selection rounds 7-10 (Davis and Szostak, 2002) were calculated using the program RNAfold from the ViennaRNA Package version 1.4, a dynamic programming algorithm (Zuker and Stiegler, 1981; Zuker and Sankoff, 1984; Jaeger et al., 1989 and 1990; Fontana et al., 1993; Hofacker and Fontana, 1995) and compared to the results obtained when applying this algorithm to randomized RNA sequences of the same base composition and length.

The experiments described above revealed that large sets of randomized RNA free energies of secondary structure formation are normally distributed (Fig. 17, red bars)

and free-energy minimization calculations of evolved RNA sequences are more negative than randomized RNA sequences of the same length (Fig. 17, yellow bars).

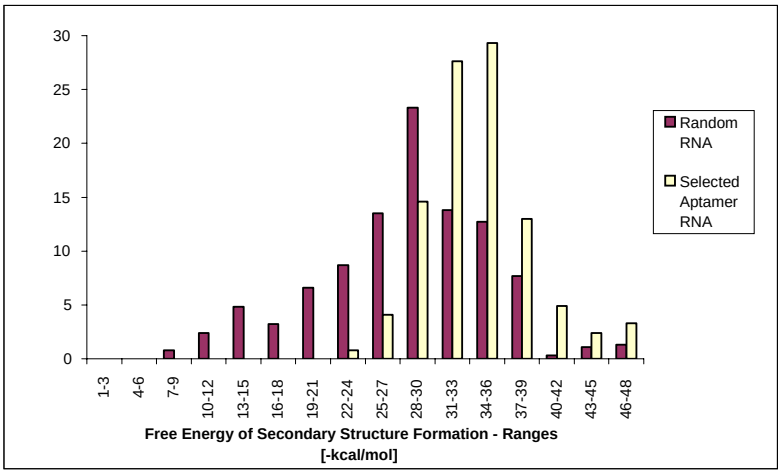


Fig. 17: Distribution of Aptamer Free Energy of Secondary Structure Formation. Free energy of secondary structure formation [-kcal/mol] for all aptamer sequences from rounds 7-10 of the original GTP selection (yellow) compared to a random RNA pool containing sequences of equal length (red) calculated by the program RNAfold from the ViennaRNA Package version 1.4.

In addition, when monitoring the free energy of secondary structure formation of the aptamers as the selection progressed, the tendency towards more negative free energy of secondary structure formation was confirmed (Fig. 18).

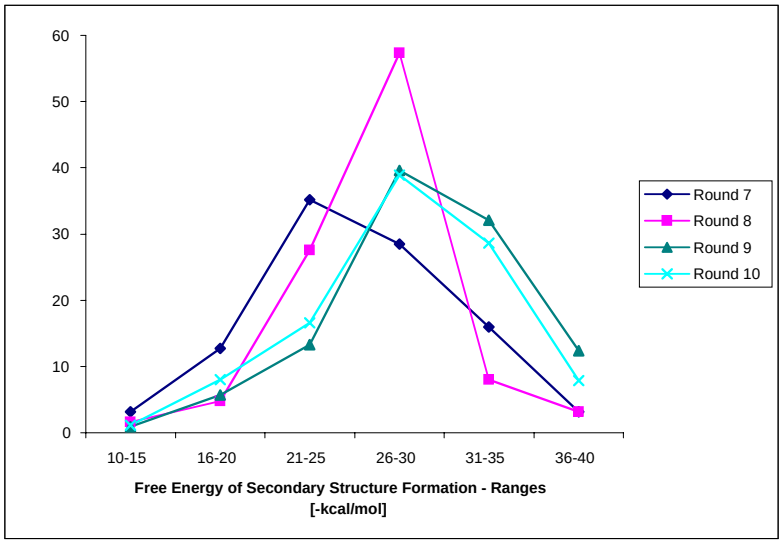


Fig. 18: Distribution of Aptamer Free Energy of Secondary Structure Formation by Round of Selection. Free energy of secondary structure formation [-kcal/mol] for all aptamer sequences from rounds 7-10 of the original GTP selection were calculated by the program RNAfold from the ViennaRNA Package version 1.4. The same amount of RNA bound and eluted in rounds 9 and 10, hence the similarity of the respective graphs.

When examining the free energy of secondary structure formation of individual aptamers using the program mfold version 3.1 (Zuker, 2003; Mathews et al., 1999; see also Table 2), for some of them (9-4, Class V, 10-10, 10-59, and 10-24) a significant contribution to the quality of the aptamer seems to stem from the free energy of secondary structure formation (Fig. 19). Low free energy of secondary structure formation on the other hand was also observed for some aptamers (Class I, 9-12, Class II, Class IV), which are not necessarily weak binders, but simply do not seem to require high free energy of secondary structure formation in order to function.

	9-4	Class V	10-10	Class I	9-12	10-59	10-6	10-24	Class II	Class IV	Class III
FE [kcal/mol]	26.4	31.0	18.6	10.9	9.9	21.9	13.4	16.7	5.5	10.4	14.1

Table 2: Free Energy of Secondary Structure Formation (FE) [kcal/mol].

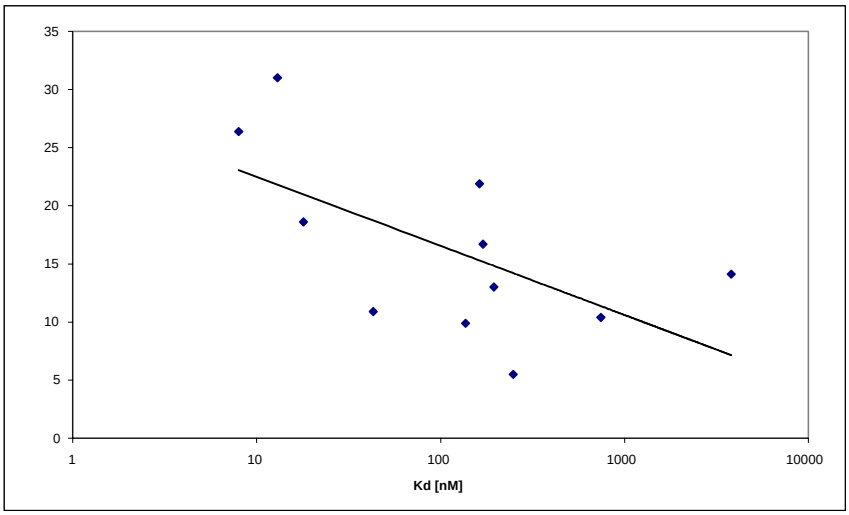


Fig. 19: Relationship of K_d [nM] and Free Energy of Secondary Structure Formation (FE SSF).

Most of the free energy of secondary structure formation is contributed by the stems. Larger and high-affinity aptamers, such as 9-4 and Class V, show a collapse of

their loops in favor of additional stems that provide more structure and therefore stability. The loop region is however the part through which most aptamers interact with their target, but there is no direct way to accurately determine the free energy contributed by the loops, even though the entropic effect of loop size is taken into account to some extent by the folding algorithm employed here (Mathews et al., 1999). When determining the RNA concentration by conducting thermal melting experiments, on average 20% higher A_{260} was observed (Fig. 3). This hyperchromicity might in part be due to additional base stacking in the loops. The stabilizing effect of the stems and their contribution to the free energy of secondary structure formation is therefore an important, but by no means the only parameter contributing to aptamer performance.

4.3.4 Investigation of Aptamer Binding Mechanisms with GTP Analogs

To test the hypothesis that aptamers with better binding activity make more and stronger contacts with their target, competition binding experiments were conducted with trace levels of radioactive GTP in the binding buffer, the RNA concentration at twice the apparent K_d and titrations of 23 different unlabeled GTP analogs as competitors. A diverse set of analogs was selected to cover most positions in the GTP molecule: For voluminous substitutions, methylated analogs were chosen to cover positions 1, 2, 3 and 7 in the nucleobase as well as positions 2' and 3' in the sugar. In addition, substitution analogs were selected to examine the effect of altering the hydrogen-bond (7-Deaza-GTP, 7-Deaza-Guanosine) and electron-donor-and-acceptor structure (6-Thio-GTP, 8-Amino-Guanosine). Finally, GTP analogs were tested that lacked one or several of the phosphates to examine to which extent the aptamers are interacting with the phosphate region of GTP (GTP-γ-S, GDP, GMP, Guanosine and Guanosine-5'-O-(1-thiotriphosphate) (Fig. 20).

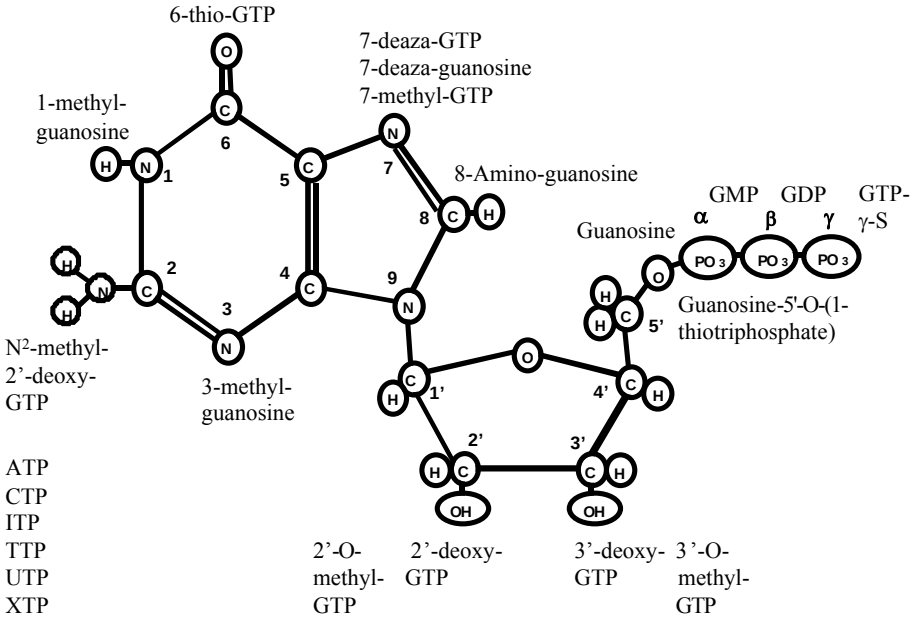


Fig. 20: GTP Analogs indicating the respective position that was modified in the molecule (GTP structure: courtesy of Rosa Larralde).

4.3.4.1 Overall Interaction Pattern

To examine aptamer interaction with GTP analogs, first IC50 values were determined with each aptamer and each GTP analog and normalized to the IC50 obtained with GTP as an unlabeled competitor. Investigation of the mechanism of aptamer binding in this way yielded no uniform pattern of interaction, but multiple structural solutions and rather idiosyncratic behavior of each aptamer (Table 3).

Analogs	9-4	Class V	10-10	Class I	9-12	10-59	10-24	10-6	Class II	Class IV	Class III
GTP	1	1	1	1	1	1	1	1	1	1	1
GTP-γ-S	1.42	1.47	2.07	0.61	2.68	1.4	1.65	0.92	3.45	2.01	1.08
GDP	1.36	1	1.65	0.84	2.57	1.03	1.72	0.89	7.94	2	1.38
GMP	1.88	8.96	2.02	1.25	2.22	2.09	7.4	1.4	3.94	1.96	1.32
Guanosine	0.88	19.8	3.75	1.09	14.09	0.68	1.4	0.42	9.65	5.86	NB
GTP-1-S	9.18	9.33	2.61	2.84	2.02	7	2.88	2.4	15.62	3.98	1.5
2'-Deoxy-GTP	1.18	7.39	30.66	8.07	5.63	3.14	1.5	32.27	1.54	4.51	1.11
2'-O-Me-GTP	3.1	10.03	81.9	16.17	10.22	3.47	8.81	NB	3.19	NB	1.09
3'-Deoxy-GTP	5.21	1	4.9	2.54	2.92	4.59	4.29	0.86	2.13	4.77	1.14
3'-O-Me-GTP	15.4	8.58	19.5	8.09	3.58	8.96	5.71	1.82	3.08	NB	NB
1-Me-Guanosine	NB	NB	NB	34.05	NB	60	100	NB	NB	NB	NB
N2-Me-2'-Deoxy-GTP	NB	NB	NB	NB	NB	NB	10.53	NB	NB	NB	NB
3-Me-Guanosine	NB	NB	NB	7.99	NB	0.63	131	NB	NB	3.38	NB
6-Thio-GTP	NB	NB	2.65	NB	7.03	120	327	NB	12.75	NB	NB
7-Me-GTP	NB	NB	1.95	118	NB	NB	111	0.8	4.25	NB	2.5
7-Deaza-Guanosine	NB	NB	18.75	NB	NB	NB	4.39	2.6	19.13	NB	NB
7-Deaza-GTP	NB	NB	28.51	NB	95.07	NB	2.88	12.14	8.5	NB	NB
8-Amino-Guanosine	1.07	64.17	120	0.63	410	0.42	1.44	NB	31.89	0.46	NB

Table 3: Normalized IC50 Values. IC50 values were determined for each aptamer with each GTP analog and normalized to the IC50 obtained with unlabeled GTP as a competitor. NB: no binding observed; GTP-1-S: Guanosine-5'-O-(1-thiotriphosphate); Me: methyl.

When grouping the strength of the interactions on a log-scale (<5-fold worse than interaction with GTP, 5-50-fold, 50-500-fold, and >500-fold = no binding observed (NB), Table 4), some aptamers like Class IV and Class III seemed to very closely interact with GTP by exhibiting no detectable interaction with 8 and 9 of the GTP analogs respectively. Other aptamers like 10-10, Class I, Class II and especially 10-24 seem to permit weaker interactions with GTP, showing 0-4 non-binding interactions. This is also reflected by the fact that, with the exception of 10-24, in general only up to 2 weak interactions (50-500-fold worse than with GTP) could be observed, whereas 5 out of 11 aptamers show no weak interactions. On the other hand, 5 out of 11 aptamers make contacts with 6-9 positions in the GTP molecule that are sufficiently tight and important for aptamer function that any modification of these positions decreases aptamer affinity by more than three orders of magnitude compared to interaction with wild-type GTP.

Affinity	9-4	Class V	10-10	Class I	9-12	10-59	10-24	10-6	Class II	Class IV	Class III
0-5xGTP	7	3	3	7	6	9	9	9	7	8	8
5-50xGTP	3	6	4	5	4	2	4	2	7	1	0
50-500xGTP	0	1	2	1	2	2	4	0	0	0	0
>500xGTP	7	7	3	4	5	4	0	6	3	8	9 (>100xGTP)

Table 4: Number of Contacts of GTP Aptamers with GTP Analogs Grouped by Intensity. Due to the limited solubility of GTP and its analogs, for the weakest-binding aptamer, Class III, only values for up to 100-fold decreased interaction compared to GTP could be obtained.

Only a slight increase in specificity could be observed in aptamers 9-4 and Class V (7 non-binding positions compared to 3-6 for most weaker-binding aptamers), and neither the number nor the strength of interactions of aptamers with GTP analogs seem to correlate with aptamer performance.

While aptamers Class IV and Class III show an even higher amount of strong interactions (8 and 9 non-binding interactions with GTP analogs), these data have to be interpreted with caution. Due to the limited solubility of GTP and its analogs, for Class III only data for up to 100-fold decreased interaction with GTP analogs compared to GTP could be obtained. Even though data for Class III and Class IV show high similarity, this

result stems most likely from the limitations of the method used here for the accurate examination of weaker-binding aptamers.

When converting IC50 values into real K_ds and calculating the percentage of correctly folded aptamer for all interactions, the latter value varies among different aptamers, but is homogeneous when examining the binding of an aptamer to GTP and several of its analogs (Table 5).

Real Kd	9-4	Class V	10-10	Class I	9-12	10-59	10-24	10-6	Class II	Class IV	Class III
GTP	7.86 nM	13.26 nM	17.56 nM	43.19 nM	136.81 nM	163.24 nM	170.38 nM	195.26 nM	247.95 nM	744.79 nM	3.8 μM
GTP- γ-S	11.44 nM	19.67 nM	36.97 nM	26.21 nM	370.45 nM	228.94 nM	274.79 nM	179.96 nM	1.19 μM	1.50 μM	4.13 μM
GDP	11.80 nM	13.70 nM	29.06 nM	35.97 nM	346.41 nM	167.61 nM	297.98 nM	84.27 nM	2.78 μM	1.47 μM	5.25 μM
GMP	13.65 nM	119.54 nM	36.19 nM	53.32 nM	305.27 nM	339.44 nM	1.24 μM	273.00 nM	1.34 μM	1.46 μM	5.01 μM
Guanosine	6.95 nM	265.46 nM	65.58 nM	46.13 nM	1.93 μM	112.8 nM	244.92 nM	725.19 nM	3.38 μM	4.34 μM	NB
GTP-1-S	68.07 nM	113.88 nM	45.58 nM	122.87 nM	271.74 nM	1.19 μM	497.65 nM	463.84 nM	5.35 μM	3.02 μM	5.70 μM
2'-Deoxy-GTP	8.75 nM	95.91 nM	537.54 nM	346.80 nM	772.42 nM	514.13 nM	256.36 nM	6.29 μM	547.50 nM	3.35 μM	4.16 μM
2'-O-Me-GTP	25.65 nM	123.45 nM	1.44 μM	692.58 nM	1.42 μM	565.62 nM	1.51 μM	NB	1.11 μM	NB	4.15 μM
3'-Deoxy-GTP	45.32 nM	12.95 nM	86.19 nM	105.10 nM	378.11 nM	746.59 nM	721.65 nM	167.00 nM	737.79 nM	3.52 μM	4.33 μM
3'-O-Me-GTP	114.70 nM	112.60 nM	335.20 nM	347.04 nM	482.48 nM	1.46 μM	951.90 nM	361.74 nM	1.06 μM	NB	NB
1-Me-Guanosine	NB	NB	NB	1.42 μM	NB	9.77 μM	17.04 μM	NB	NB	NB	NB
N2-Me-2'-Deoxy-GTP	NB	NB	NB	NB	NB	NB	1.79 μM	NB	NB	NB	NB
3-Me-Guanosine	NB	NB	NB	339.22 nM	NB	102.04 nM	22.10 μM	NB	NB	2.50 μM	NB
6-Thio-GTP	NB	NB	46.38 nM	NB	957.11 nM	19.74 μM	58.72 μM	NB	4.44 μM	NB	NB
7-Me-GTP	NB	NB	34.55 nM	5.07 μM	NB	NB	19.14 μM	157.91 nM	1.51 μM	NB	9.52 μM
7-Deaza-Guanosine	NB	NB	321.61 nM	NB	NB	NB	725.76 nM	507.64 nM	6.78 μM	NB	NB
7-Deaza-GTP	NB	NB	500.26 nM	NB	12.98 μM	NB	505.56 nM	2.35 μM	2.96 μM	NB	NB
8-Amino-Guanosine	7.60 nM	872.17 nM	2.15 μM	25.95 nM	55 μM	68.27 nM	247.09 nM	NB	11.12 μM	340.41 nM	NB
% Folded Correctly	9-4	Class V	10-10	Class I	9-12	10-59	10-24	10-6	Class II	Class IV	Class III
GTP	89.90	73.59	63.96	58.26	53.85	64.05	57.87	66.06	60.77	37.57	44.10
GTP- γ-S	89.00	73.50	64.28	58.66	53.94	64.62	57.18	66.26	60.04	37.11	44.41
GDP	90.67	74.66	64.07	59.16	53.75	63.91	58.70	66.13	59.33	37.13	44.51
GMP	90.30	73.52	63.98	58.19	53.56	64.40	57.16	66.34	60.56	37.27	44.23
Guanosine	87.47	73.59	63.66	58.56	53.17	65.00	58.81	66.46	60.74	37.20	NB
GTP-1-S	89.95	73.18	63.65	58.14	53.63	64.11	58.79	66.50	60.61	37.48	44.49
2'-Deoxy-GTP	88.49	74.59	63.10	60.48	53.68	64.51	57.63	66.20	61.09	38.00	44.00
2'-O-Me-GTP	86.00	72.43	63.27	58.14	53.60	64.07	57.73	NB	60.24	NB	44.35
3'-Deoxy-GTP	89.57	73.52	63.86	58.51	53.13	64.31	57.67	66.42	60.00	37.29	44.40
3'-O-Me-GTP	89.17	72.37	63.82	58.66	53.61	64.16	57.01	66.20	60.39	NB	NB
1-Me-Guanosine	NB	NB	NB	57.32	NB	64.30	57.79	NB	NB	NB	NB
N2-Me-2'-Deoxy-GTP	NB	NB	NB	NB	NB	NB	57.64	NB	NB	NB	NB
3-Me-Guanosine	NB	NB	NB	58.34	NB	64.48	57.13	NB	NB	37.81	NB
6-Thio-GTP	NB	NB	63.77	NB	53.12	64.53	56.91	NB	60.13	NB	NB
7-Me-GTP	NB	NB	63.86	58.33	NB	NB	58.16	66.69	60.88	NB	44.76
7-Deaza-Guanosine	NB	NB	63.32	NB	NB	NB	56.99	66.49	60.70	NB	NB
7-Deaza-GTP	NB	NB	63.31	NB	53.63	NB	57.91	66.04	60.32	NB	NB
8-Amino-Guanosine	90.09	74.34	64.51	58.11	53.56	64.34	57.61	NB	60.48	34.04	NB

Table 5: Real K_ds and Fraction of Correctly Folded Aptamers. NB: no binding observed; GTP-1-S: Guanosine-5'-O-(1-thiotriphosphate); Me: methyl.

4.3.4.2 Interaction with the Phosphate Region of GTP

Most aptamers bind equally well or with only slightly decreased affinity to GTP-triphosphate analogs, indicating weak or no interaction with the phosphate region. Binding to GTP-γ-S was also examined, since this analog had been used when immobilizing GTP to thiopropyl-sepharose for *in vitro* selection. Results of these experiments showed only a slight decrease in aptamer binding activity (up to 3-fold compared with GTP), and in the case of aptamers Class I and 10-6 even improved binding, thus ensuring comparable behavior of GTP and GTP-γ-S when used for *in vitro* selection purposes.

4.3.4.3 Interaction with the Sugar Region of GTP

Interaction with the sugar region is also weak, but is impaired more severely by introduction of voluminous groups such as methyl groups into the GTP molecule compared to modifications where interacting groups were removed: Aptamers interacted on average three-fold less with 2'-O-methyl-GTP and 3'-O-methyl-GTP respectively than with 2'-deoxy-GTP and 3'-deoxy-GTP respectively. Class III and Class IV are the only aptamers that do not bind detectably to 3'-O-methyl-GTP.

4.3.4.4 Interaction with the Nucleobase Region of GTP

Aptamers interacted mostly with the nucleobase region of GTP, which is reflected by the large number of interactions that are abolished by introduction of modified groups in this region.

Almost half of the aptamers (9-4, Class I, 10-59, and Class IV) interacted better or equally well with 8-amino-guanosine compared to GTP, while modifications in most other positions in the nucleobase severely impeded binding. This result might stem from a steric effect caused by stabilization of GTP in the *syn* conformation due to the attached amino group. Aptamers whose interaction is severely impeded by introduction of the amino group (Class V, 10-10, 9-12, 10-24, 10-6, Class II, and Class III) on the other hand could prefer binding to GTP in the *anti* conformation.

The fact that no binding was observed with triphosphate-analogs with multiple alterations in the nucleobase region such as ATP, CTP, ITP, TTP, UTP, and XTP also adds to the importance of the nucleobase region for aptamer binding and to the evidence that interactions with different positions are independent from each other, since no uniform binding pattern could be observed.

4.3.4.5 Differences in Binding Energy

In order to further evaluate aptamer interaction with GTP analogs, differences in binding energy (ΔΔG values) between aptamer interaction with GTP compared to GTP analogs were calculated from the ratio of experimental K_d values according to the following formula:

$$\Delta\Delta G = -RT \ln(K_{d[GTP]} / K_{d[Analog]})$$

Negative f□□□□□□□□□□□□□□□□□□□□, observed for instance for the interaction of some aptamers (9-4, Class I, 10-59 and Class IV) with 8-amino-guanosine, are favored interactions, whereas positive values represent decreased binding activity up to values (about 12 kJ/mol) that were equivalent to half of the free energy of secondary structure formation for the Class II aptamer (Table 6).

Analogs	9-4	Class V	10-10	Class I	9-12	10-59	10-6	10-24	Class II	Class IV	Class III
GTP	0	0.00	0	0	0	0	0	0	0	0	0
GTP-γ-S	921	967.67	1827	-1226	2444	830	-200	1173	3849	1718	204
GDP	997	80.10	1236	-449	2280	65	-2062	1372	5931	1668	793
GMP	1354	5395.83	1775	517	1969	1796	822	4871	4140	1652	678
Guanosine	-302	7353.57	3233	162	6495	-907	3220	891	6411	4325	NB
GTP-1-S	5297	5276.81	2341	2566	1684	4875	2123	2630	7537	3435	995
2'-Deoxy-GTP	263	4855.39	8396	5112	4248	2815	8521	1003	1944	3690	222
2'-O-Me-GTP	2902	5474.81	10814	6809	5742	3049	NB	5354	3678	NB	216
3'-Deoxy-GTP	4299	-58.05	3904	2182	2495	3731	-384	3542	2676	3811	320
3'-O-Me-GTP	6578	5249.07	7237	5113	3093	5376	1513	4222	3565	NB	NB
1-Me-Guanosine	NB	NB	NB	NB	NB	10041	NB	11301	NB	NB	NB
N2-Me-2'-Deoxy-GTP	NB	NB	NB	NB	NB	NB	NB	5771	NB	NB	NB
3-Me-Guanosine	NB	NB	NB	5058	NB	-1153	NB	11939	NB	2972	NB
6-Thio-GTP	NB	NB	2383	NB	4774	11767	NB	14337	7080	NB	NB
7-Me-GTP	NB	NB	1661	NB	NB	NB	-521	11586	4433	NB	2254
7-Deaza-Guanosine	NB	NB	7135	NB	NB	NB	2345	3556	8119	NB	NB
7-Deaza-GTP	NB	NB	8219	NB	11171	NB	6105	2669	6085	NB	NB
8-Amino-Guanosine	-83	10272.51	11797	-1250	NB	-2139	NB	912	9333	-1921	NB

Table 6: ΔΔG Values from Competition Experiments with GTP [J/mol]. NB: no binding observed; GTP-1-S: Guanosine-5'-O-(1-thiotriphosphate); Me: methyl.

4.3.4.5.1 Comparison of Binding Energy Distributions

ΔΔG values for interactions of each aptamer with each competitor were determined, grouped in ranges of 1000 J/mol and plotted against the frequency of the values in these ranges (Fig. 21).

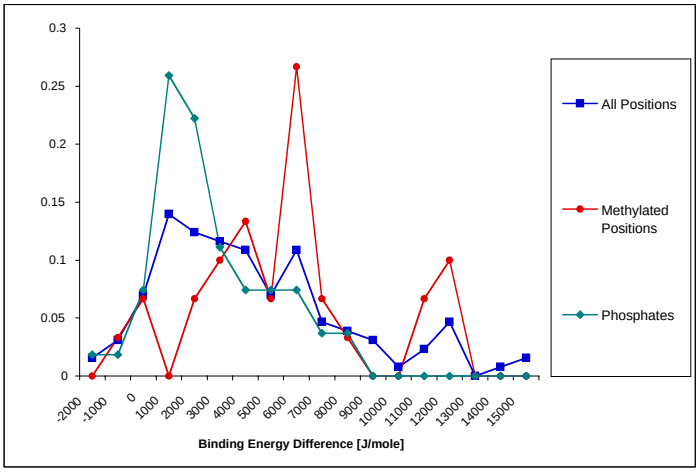


Fig. 21: Differences in Binding Energy (ΔΔG). ΔΔG values were grouped in ranges of 1000 J/mol and plotted against their frequency of occurrence. Blue: Interaction with all positions in the GTP molecule; Green: interaction with GTP analogs modified in the phosphate region of GTP (GTP-γ-S, GDP, GMP, Guanosine, Guanosine-5'-O-(1-thiotriphosphate)); Red: interaction with methylated GTP analogs (2'-O-Methyl-GTP, 3'-O-Methyl-GTP, 1-Methyl-Guanosine, N^2-Methyl-2'-Deoxy-GTP, 3-Methyl-Guanosine, and 7-Methyl-GTP).

When comparing the interaction of GTP aptamers with GTP analogs modified in the phosphate region (GTP-γ-S, GDP, GMP, Guanosine, Guanosine-5'-O-(1-thiotriphosphate); see data marked in green in Fig. 22) to those with analogs with methylation modifications (2'-O-Methyl-GTP, 3'-O-Methyl-GTP, 1-Methyl-Guanosine, N^2-Methyl-2'-Deoxy-GTP, 3-Methyl-Guanosine, and 7-Methyl-GTP; see data marked in red in Fig. 22) a shift towards more positive ΔΔG values with broader distribution was observed. These results confirm that the interaction of aptamers with the sugar and phosphate region of GTP is both less significant and more uniform across aptamers compared to interactions with the nucleobase region. Most interactions thus occur in the nucleobase region, but seem to be independent since no uniform binding pattern could be observed across all aptamers.