# Isolation of high-affinity GTP aptamers from partially structured RNA libraries

Jonathan H. Davis* and Jack W. Szostak†

Howard Hughes Medical Institute, and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114

Aptamers, RNA sequences that bind to target ligands, are typically isolated by *in vitro* selection from RNA libraries containing completely random sequences. To see whether higher-affinity aptamers can be isolated from partially structured RNA libraries, we selected for aptamers that bind GTP, starting from a mixture of fully random and partially structured libraries. Because stem-loops are common motifs in previously characterized aptamers, we designed the partially structured library to contain a centrally located stable stem-loop. We used an off-rate selection protocol designed to maximize the enrichment of high-affinity aptamers. The selection produced a surprisingly large number of distinct sequence motifs and secondary structures, including seven different aptamers with $K_d$s ranging from 500 to 25 nanomolar. The engineered stem-loop was present in the three highest affinity aptamers, and in 12 of 13 independent isolates with a single consensus sequence, suggesting that its inclusion increased the abundance of high-affinity aptamers in the starting pool.

*I*n vitro selection, or SELEX (1, 2), was developed as a method for isolating novel nucleic acid sequences with desired activities. Selections for RNAs that bind ligands specifically (*aptamers*) have uncovered RNA sequences that bind to a wide range of target molecules, including proteins, antibiotics, sugars, bile acids, vitamins, cofactors, nucleic acids, and amino acids (3–5). Recently, there has been increased interest in the application of aptamers as biosensors for the detection and measurement of biological or environmental ligands (6–8), and as *in vivo* probes of biological function (9). Aptamer selections are also of interest at a more fundamental level, because they provide a means for exploring structural features that influence binding affinity and specificity.

Essentially, all *de novo* aptamer selections reported to date begin with a fully random RNA sequence library, in which the random region varies from 30 to 70 nt in length. It seems likely that higher-affinity aptamers will be significantly less common in a library than lower-affinity aptamers. If so, very high-affinity aptamers might have a low probability of being present in the starting library. We hypothesized that the chances of finding complex and therefore rare aptamers could be improved by designing the starting library to include common structural components of RNA aptamers. Many aptamers have at least one internal stem-loop that appears to act as a structural anchor for recognition loops (for examples, see refs. 10–13). The base-paired stems of these stem-loops can generally be of any sequence (10, 11). We therefore created a library that contained a stable stem-loop centered in the random region (Fig. 1*A*). To see whether this partially structured library would be a better source of high-affinity aptamers, we carried out a selection for GTP aptamers from a 1:1 mix of this partially structured library and a fully random library. We used a stringent off-rate selection protocol, similar to that described by Geiger *et al.* (12), to select for high-affinity aptamers to GTP. Subsequent analysis of these aptamers revealed that most of those with the highest affinity for GTP did in fact originate from the partially structured library.

## Materials and Methods

**Design and Synthesis.** The library was made by combining two separate DNA templates, each of which consisted of two fixed primer-binding regions flanking a 64-base variable region. In the random library, all 64 bases in the variable region were random, whereas in the partially structured library, two 26-base random stretches were separated by a 12-base fixed sequence (5′-CTGCCGAAGCAG-3′), to form a 4-base pair stem and a stable UUCG tetraloop in the transcribed RNA (ref. 14; Fig. 1). The primers were B8 (3′-AAGTAAGTCAACCGCGGAGGATAT-CACTCAGCATAATGTA-5′), which includes the T7 promoter, and RQ (5′-GTGACGCGACTAGTTACGGA-3′), complementary to the 3′ end of the library RNA. One-micromole syntheses were performed on an Applied Biosystems synthesizer. Phosphoramidites for the random positions were mixed in an A:C:G:T ratio of 3:3:2:2 to achieve approximately equal incorporation of each base (15).

**Selection Protocol.** The affinity matrix was prepared by incubating thiopropyl Sepharose 6-B (Pharmacia) with 4 mM GTP-γ-S (Sigma) for 1 h at 4°C in 10 mM Tris·HCl (pH 7.5) and 10 mM EDTA. The immobilized GTP-γ-S concentration was measured by washing an aliquot of beads with buffer, eluting the bound GTP-γ-S by incubating for 10 min in 6 mM DTT, then estimating the GTP-γ-S concentration by comparison with standards after separation from other UV-absorbing species by TLC on polyethyleneimine-cellulose with 1.5 M LiCl and 1.5 M formic acid. The GTP-γ-S concentration on the beads was typically 300–600 μM.
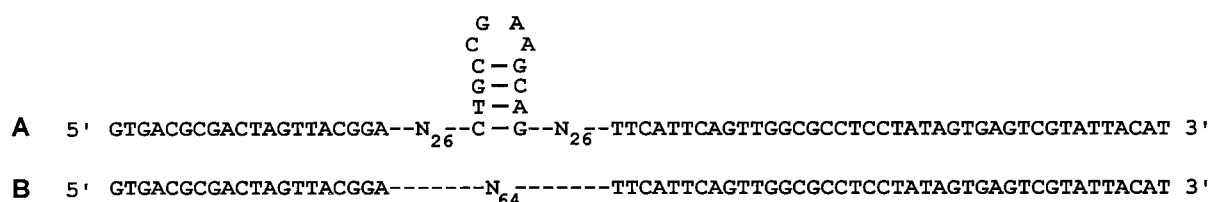
Selection buffer was chosen to be conducive for NMR spectrometry (200 mM KCl/10 mM potassium phosphate/5 mM MgCl₂/0.1 mM EDTA, pH 6.2), and elution buffer was the same with the addition of 5 mM GTP and 7 mM MgCl₂. The starting library contained $2.5 \times 10^{14}$ unique DNA template molecules from each of the two libraries described above. *In vitro* T7 transcription (16) was used to generate approximately 50 RNA equivalents of the starting library. After each affinity enrichment step, the eluted RNA was added to 200–400 μl of reverse transcription–PCR (Ready-To-Go RT-PCR beads, Pharmacia), and 2–5 rounds of PCR were performed to make enough DNA for transcription. T7 transcription then generated 100–1,000 RNA copies of each template molecule (0.2–2 mg RNA). Preliminary to the selection step, PAGE-purified RNA was denatured in 200 μl H₂O at 80°C for 1 min, then allowed to cool to room temperature. An equal volume of 2× selection buffer was added, and the RNA was then loaded onto a 0.5-ml affinity resin column. To purge the population of matrix binders, for the first three rounds of selection, a precolumn containing 0.5 ml of the same matrix without bound GTP-γ-S was placed above the selection matrix, and the RNA dripped through the precolumn directly onto the selection column. After 3 column volumes of washing, the precolumn was removed.

Table 1 provides detailed conditions for each selection step. To ensure that tightly bound aptamers with slow off-rates were

```
                                    G   A
                                  C       A
                                  C  —  G
                                  G  —  C
                                  T  —  A
A   5' GTGACGCGACTAGTTACGGA--N26--C — G--N26--TTCATTCAGTTGGCGCCTCCTATAGTGAGTCGTATTACAT 3'


B   5' GTGACGCGACTAGTTACGGA-------N64-------TTCATTCAGTTGGCGCCTCCTATAGTGAGTCGTATTACAT 3'
```

**Fig. 1.** Design of the initial libraries. Engineered (*A*) and random (*B*) libraries, both 124 bases long with 64 bases between the constant primer binding regions, were used to generate the starting RNA. The random library contains 64 random bases, whereas the engineered library contains 52 random bases with a 4 bp stem and stable UUCG tetraloop in the middle.

efficiently recovered, elution was achieved by four successive incubations of 30 min each with elution buffer containing GTP, and the four resulting fractions were combined for amplification. In later rounds, weak complexes were selected against by a pre-elution step in which one column volume of elution buffer was incubated on the column for 0.5–10 min (the time was increased over the course of the selection), then replaced with fresh elution buffer for further pre-elution, or for the first long elution step (12). Pre-elution fractions were not included in reverse transcription and amplification.

**Sequencing and Affinity Determination.** Enriched libraries were cloned and sequenced by using a TOPO TA cloning kit (Invitrogen) and an Applied Biosystems 377 sequencer. Individual RNA clones were screened for binding activity either by Biacore or by ultrafiltration $K_d$ assays (11). For Biacore screening, biotinylated GTP and ATP were produced by reacting GTP- or ATP-$\gamma$-S with biotin-HPDP (Pierce). A Biacore streptavidin chip (SA) was loaded with the biotinylated GTP on one cell, and with the biotinylated ATP on the other to serve as a control. Aptamer RNA was presented to the chip and allowed to bind, then the off-rate was measured first during elution in buffer alone, and then, to test for rebinding, during elution with buffer containing 5 mM GTP. Ultrafiltration assays were performed by incubating freshly annealed RNA at varying concentrations with 0.5–1.0 nM [$\alpha$-$^{32}$P]GTP for at least 2 h, then loading 200 $\mu$l into a Microcon (Millipore) spin filter with a 10-kDa cutoff. The filters were centrifuged at 13,000 $\times$ $g$ for 30 s to saturate the membranes, then fitted with fresh collection tubes and centrifuged again for 90 s. Twenty-five microliters each from the top (retentate) and bottom (filtrate) were counted by Cerenkov scintillation. The bottom counts represent the free ligand, [$L$], and the top counts represent total (bound plus free) ligand, [$L$] + [$A{\cdot}L$]. Seven values for the fraction of bound ligand, [$L$]/([$L$] + [$A{\cdot}L$]) were measured and plotted against [$A$], and the $K_d$ was determined from the value where half the ligand is bound. Because [$A$] $\gg$ [$L$], the concentration of free aptamer is not significantly affected by ligand binding. $K_d$ values were un-

changed with incubations of 4–24 h, suggesting that equilibrium had been reached.

**Sequence Mapping.** We performed 3′ end mapping as described (17). Partially hydrolyzed 5′-$^{32}$P-labeled RNA was passed over a GTP column, washed, and eluted as in the selection (without pre-elution). The flow-through and GTP-eluted RNA were analyzed by PAGE to determine the 3′ end of the region required for GTP binding. We performed 5′-end mapping with unlabeled partially hydrolyzed RNA, which was subjected to affinity enrichment for GTP as above, then reverse transcribed with 5′-end $^{32}$P-labeled primers. The resultant cDNA represented the lengths of the RNA in each fraction and was analyzed in a similar manner to the 3′ mapping.

### Results

**Library and Selection Design.** The partially structured library (consisting of $2.5 \times 10^{14}$ independent DNA sequences) was designed with a central 12-nt sequence encoding a 4-bp stem closed by a stable UUCG loop, flanked by 26 random bases on each side, whereas the fully random library (also consisting of $2.5 \times 10^{14}$ independent DNA sequences) contained 64 contiguous random bases. RNA transcribed from these two initial libraries was mixed together in a 1:1 ratio before the start of the selection, giving each library an approximately equal opportunity to provide winning sequences.

In the first phase of the selection, both low- and high-affinity GTP aptamers were enriched by incubating the RNA pool with the GTP-derivatized affinity resin, followed by a relatively low-stringency washing procedure. Care was taken to recover high-affinity aptamers present in the pool by prolonged elution in the presence of a high concentration of competing GTP. In the first three rounds, <1% of the input sequences were bound and then specifically eluted, but in round 4 the specifically eluted RNA increased to 8%.

In the second phase of the selection, from round 5 on, we attempted to eliminate low- and moderate-affinity aptamers from the pool by increasingly stringent washing (Table 1). To enrich for aptamers with low off-rates, the affinity column was

**Table 1. Selection protocol**

| Round | No. col. vols. of washes | No. pre-elutions | Time of each pre-elution | % surviving 6 col. vols. | % eluted and amplified |
|-------|--------------------------|------------------|--------------------------|--------------------------|------------------------|
| 1  | 6  | 0 | —      | 0.1  | 0.1  |
| 2  | 6  | 0 | —      | 0.04 | 0.04 |
| 3  | 6  | 0 | —      | 0.4  | 0.4  |
| 4  | 6  | 0 | —      | 7.5  | 7.5  |
| 5  | 20 | 0 | —      | 20   | 10   |
| 6  | 10 | 2 | 0.5 min | 45  | 5    |
| 7  | 10 | 3 | 4 min  | 50   | 1    |
| 8  | 10 | 3 | 4 min  | 60   | 4.7  |
| 9  | 10 | 3 | 4 min  | 71   | 13   |
| 10 | 10 | 4 | 10 min | 76   | 13   |

## A  Selection



## B



## C
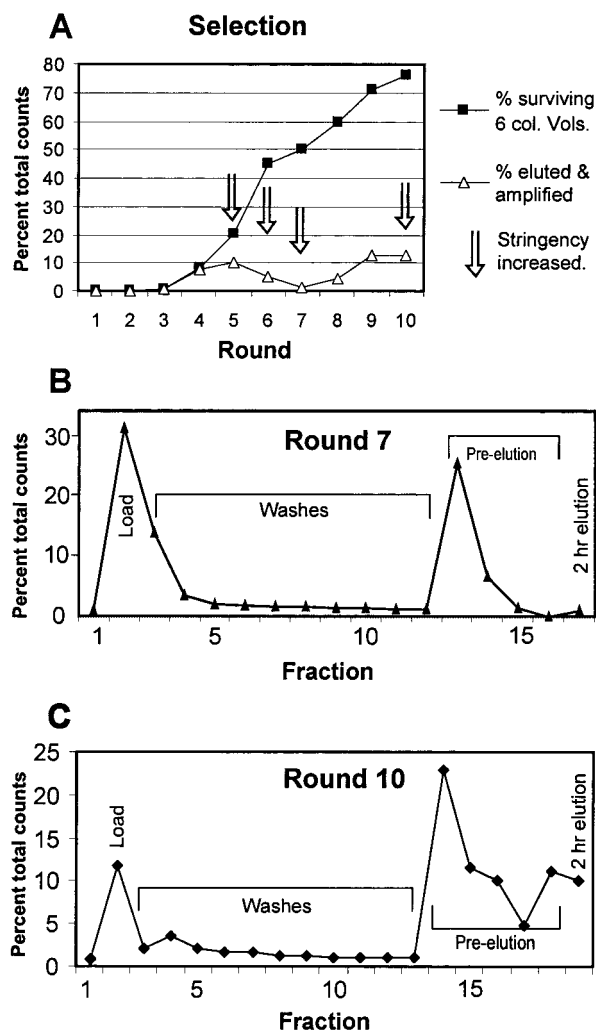


**Fig. 2.** (*A*) Progress of the selection over 10 rounds. Filled squares represent the fraction still bound to the column after 6 column volumes of washing. Open triangles represent the fraction specifically eluted after surviving the full washing and the pre-elution, if used. Arrows indicate the rounds where stringency was increased, either by more extensive washing or more pre-elution (see Table 1 for details). (*B*) Profile of round 7, expressed as percentage of total counts in each fraction. (*C*) Profile of round 10, illustrating the shift toward more stable complexes. In round 7, most of the bound RNA eluted in the first pre-elution fraction, indicating a fast $k_{off}$, whereas by round 10 less than half of the bound aptamers eluted in that fraction.

washed with selection buffer containing GTP, and RNAs remaining bound to the column were then eluted by prolonged incubation with GTP. Both the number of GTP-wash steps, and the incubation time of those steps, were increased from rounds 5 through 10. By round 10, 10% of the input RNA survived four 10-min GTP-washes. To provide a consistent measure of the binding properties of the bulk RNA over all of the selection cycles, the percentage of RNA binding to the column after six column volumes of buffer wash was determined for each round. This percentage increased gradually after round 4, reaching 75% by round 10 (Fig. 2*A*). Fig. 2*C* shows the column binding, washing, and elution profile of round 10.

**Aptamer Sequences.** Molecules derived from rounds 8, 9, and 10 of the selection were cloned and sequenced. Analysis of these sequences revealed no one dominant sequence, class, or motif, but rather a wide variety of unrelated sequences. In the final

## A



## B
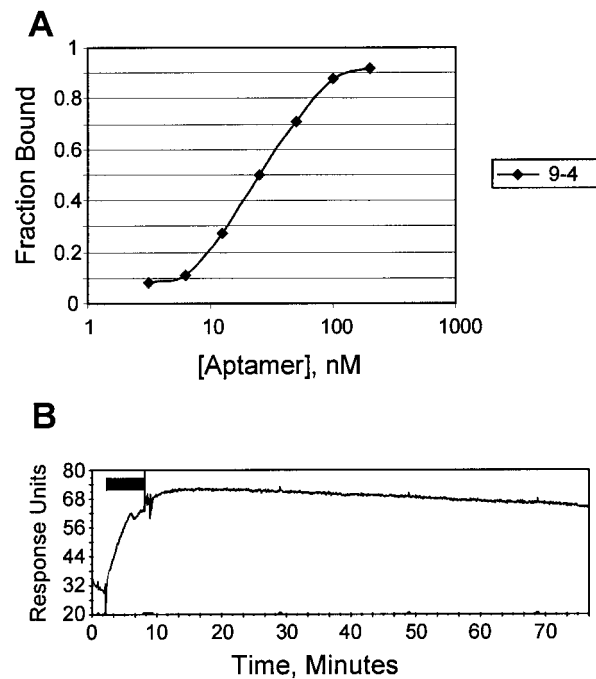


**Fig. 3.** (*A*) Binding curve of clone 9-4, as measured by ultrafiltration assays. The *y* axis is the fraction bound, whereas the *x* axis is the concentration of aptamer in the assay. The concentration at which the curve intersects the horizontal line representing one half bound (0.5) is the $K_d$, in this case 25 nM. (*B*) Biacore trace of clone 9-4, the aptamer with the lowest off-rate. RNA was allowed to bind to the chip for 6 min (horizontal bar), then the chip was washed with selection buffer at 25 $\mu$l/min. The signal from the bound aptamer peaked at 40 response units (RU) above baseline, and after 1 h it had dropped by 6 RU, or 15%. This result corresponds to a $k_{off}$ of 0.0027 min$^{-1}$, or 0.16 h$^{-1}$. The off-rate was similar when washed with buffer containing 5 mM GTP (not shown).

round, the largest family of closely related sequences represented only 9% of all round 10 sequences cloned (9 of 100). These aptamers, exemplified by clone 9-4, are clearly descendants of a single common ancestral sequence. Other closely related sequence families had between two and nine members, from the total pool of 157 sequences from rounds 8–10. A single group of otherwise unrelated sequences (Class I) shared highly conserved sequence motifs flanking the designed stem-loop and were found in each of the last three rounds; these Class I sequences comprised about 10% of all sequences (16 of 157) from rounds 8–10.

**Aptamer Affinities and Off-Rates.** Based on the fact that a significant fraction of the input RNA in round 10 survived four 10-min washes in the presence of competing GTP, we expected that many of the selected aptamers in this pool would have high affinities for GTP, and very low off-rates. We measured the solution GTP affinity for the 9-4 sequence family (Fig. 3*A*) and for all other sequence families, for the Class I sequences, and for an assortment of unique sequences; $K_d$ values ranged from 25 nM to >100 $\mu$M (Table 2 shows the data for the best binders). Off-rates were measured for a subset of these sequences on a Biacore instrument, by using a streptavidin chip saturated with biotinylated GTP; this measurement revealed a similarly wide range of off-rates, from >1 s$^{-1}$ to as slow as 0.0027 min$^{-1}$ (Fig. 3*B*).

The most prevalent single sequence (9-4) had both the highest affinity ($K_d$ = 25 nM) and the lowest off-rate (0.0027 min$^{-1}$), consistent with the strong off-rate selection applied from rounds 5 through 10 of the selection. The various Class I sequences also

## Table 2. Characteristics of tightest binding aptamers

| Aptamer | No. sequences | | $K_d$, nM | $k_{off}$, min$^{-1}$ | Pool of origin |
| --- | --- | --- | --- | --- | --- |
| | Unique | Total | | | |
| 9-4 | 1 | 10 | 25 | 0.0027 | Designed |
| 10-10 | 1 | 1 | 28 | 0.01 | Designed |
| Class I | 13 | 16 | 60–300 | 0.06 | 12 designed/1 random |
| 10-59 | 1 | 6 | 200 | 0.027 | Random |
| 9-12 | 1 | 9 | 250 | 0.03 | Random |
| 10-24 | 1 | 1 | 250 | — | Designed |
| 10-6 | 1 | 1 | 500 | — | Designed |

had high affinity for GTP ($K_d$ of different class members range from 60 to 300 nM), with an off rate of 0.06 min$^{-1}$ for a Class I aptamer with a $K_d$ of 300 nM. The Biacore elution traces were not affected by the addition of GTP to the wash buffer, demonstrating that the observed off-rate represents the actual lifetime of the aptamer–GTP complex. Five other aptamers exhibited $K_d$s for GTP that ranged between 500 nM and 25 nM. Of the seven aptamers identified with $K_d$s lower than 500 nM, five were derived from the designed library, and two from the random library. All three aptamers with affinities stronger than 100 nM were derived from the partially structured library. Even more striking, 12 of 13 independent Class I aptamers were derived from the partially structured library. The implications of these observations are discussed below.

**Aptamer Secondary Structures.** We have begun to characterize the secondary structures of the selected aptamers, to provide an additional basis for comparison between the aptamers selected from the designed vs. random libraries.

The secondary structure of the Class I aptamer was determined by comparison of the sequences of the 13 independent class members (Fig. 4). This aptamer consists of an internal asymmetric loop flanked on both sides by base-paired stems (one of which was the designed stem-loop) (Fig. 5C). Alignment of the Class I sequences revealed highly conserved 11- and 8-nt sequences immediately flanking the designed stem-loop. In all cases, these recognition loops were anchored by an outer stem of at least six base pairs, although many were interrupted by mismatches or bulges. The sequence of this stem was different in every independent isolate; the repeated identification of complementarity independent of sequence provides strong evidence for the functional significance of this stem. No other sequence conservation or covariation was observed in comparing the Class I sequences. We made constructs containing shortened stems, and a molecule with only four base pairs in the outer stem and two in the hairpin still bound GTP in solution, although about

100-fold more weakly than the original sequence. The alignment in Fig. 4 shows that 14 positions in the binding loops are completely conserved, and two more are restricted to just two bases.

The remaining aptamers could not be analyzed by alignment because of the lack of independent comparison sequences. We used 5′ and 3′ end-mapping to determine the minimal regions required for full function, for the other six aptamers with $K_d$s ≤500 nM. With the exception of aptamer 9-4, the 5′ and 3′ ends of the minimal sequences were complementary, and this analysis therefore served to identify an essential base-paired stem. Minimum-length constructs were then transcribed and all were found to exhibit binding activity in solution.
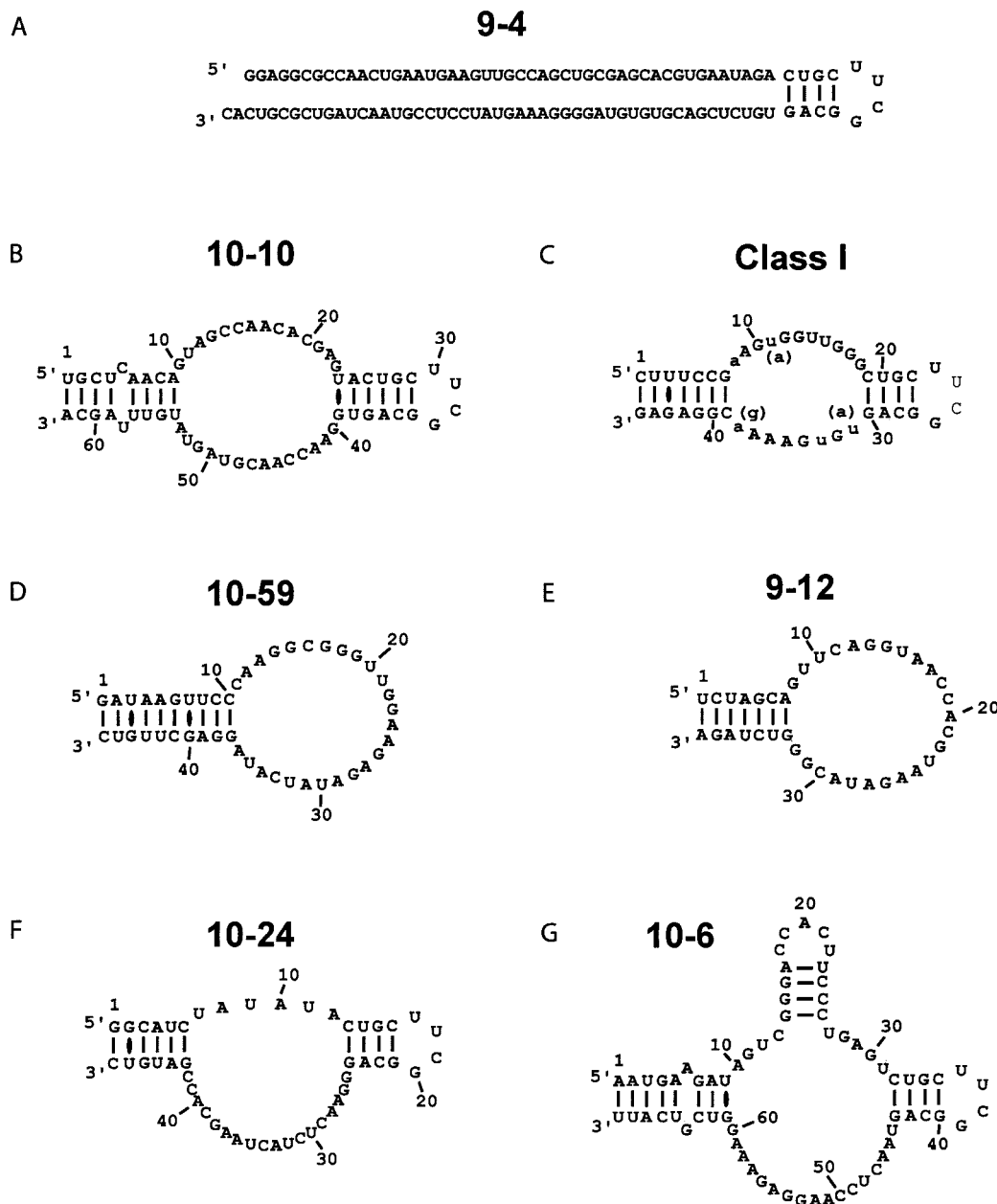
End mapping of 9-4 suggested that most of the full-length sequence was necessary for binding. Because there are numerous short regions of sequence complementarity within this sequence, we cannot yet propose a specific secondary structure for this aptamer. On the basis of end-mapping and the complementarity of the terminal regions, we propose stem-bulge-stem-loop secondary structures for 10-10 and 10-24 that are similar to that of Class I (Fig. 5). Of course, similarity of secondary structure does not imply similarity of tertiary structure. The proposed secondary structure for 10-6 is also similar, except that one of the recognition loops itself appears to contain a stable stem-loop. Remarkably, both aptamers derived from the random pool (10-59 and 9-12) appear to have simple stem-loop secondary structures, although it is certainly possible that there are short stretches of base-pairing within the large recognition loops.

## Discussion

To provide a direct comparison between partially structured and fully random RNA libraries, we made a mixture of two such pools and then selected for high-affinity GTP aptamers from the mixed pools. It was not obvious *a priori* that the designed pool would be a superior source of high-affinity aptamers compared with the random pool: even though the designed library has the
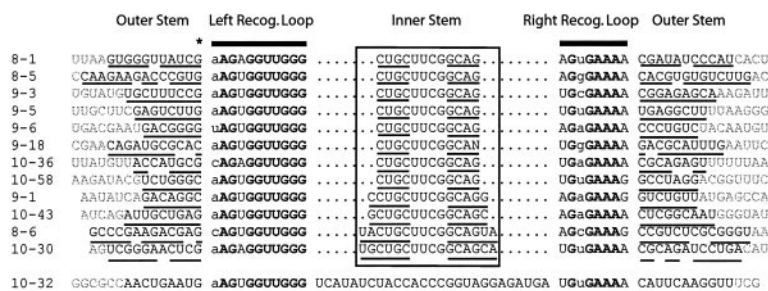
**Fig. 4.** Alignment of unique Class I sequences. The top 12 derive from the engineered library, whereas 10-32 is the lone class representative from the random library. The box contains the fixed 12-base central hairpin, with one or two extra base pairs in several of the sequences (as expected by chance). Bases that form Watson–Crick or G-U wobble pairs in the stems are underlined. The recognition loops are under the two heavy bars; bold text denotes 100% conserved bases, regular capital letters are semivariant bases (i.e., two species are allowed at that position), and lowercase are fully variable. The last base in the top strand of the outer stem, marked by an asterisk (*), is always a G or C, so it can also be considered semivariant.

## 9-4

```
5' GGAGGCGCCAACUGAAUGAAGUUGCCAGCUGCGAGCACGUGAAUAGA C U G C  U  U
                                                  | | | |      
3' CACUGCGCUGAUCAAUGCCUCCUAUGAAAGGGGAUGUGUGCAGCUCUGU G A C G  G  C
```

B

## 10-10

```
                  A G C C A A C A      20
          G U                     C G
     10                              A
   5' U G C U  C A A C A          U A C U G C  U
                                              U
      | | | |  | | | |           | | | | | |    U
   3' A C G A  U U G U           G U G A C G  G  C
     1      U                  A        A   40
          60  U G               G
                  A U G C A A C C A
                  50
```

C

## Class I

```
            10
         A G u G G U U G G      20
       a            (a)   G
   5' C U U U C C G        C U G C  U
                                     U
      | | | | | | |        | | | |    U
   3' G A G A G G C (g)  (a) G A C G  G  C
     40        a A A A G u G u   30
```

D

## 10-59

```
                A G G C G G G      20
        C A                   U
     10                         U
   5' G A U A A G U U C C      G G
                               A
      | | | | | | | | |        G
   3' C U G U U C G A G G      G
     1            U A          A
              40    A C U A U A G
                        |
                        30
```

E

## 9-12

```
              10
         U U C A G G U A
                        A
     5' U C U A G C A    C
        1          G      C  - 20
        | | | | | |       A
     3' A G A U C U G      C
                   G       A
                 30 G C A U A G A A U
```

F

## 10-24

```
                  10
            U A U A U
     5' G G C A U C          A C U G C  U
        1                            | | | |    U
        | | | | | |          G A C G  G  C
     3' C U G U A G C                        20
              C                   A
              C A              A
            40   C G A A U C A U C U C
                      |
                      30
```

G

## 10-6

```
                      20
                   A C
                 C   C
                 C   U
                 A — U
                 G — C
                 G — C
                 G — C
            10        U G A  30
   5' A A U G A  G A U  C         U C U G C  U
      1        A    A            | | | |    U
      | | | | | |               G A C G  G  C
   3' U U A C U  C U G  60         U   A  40
               G                    A
                A                   A
                 G      50    U
                  G G A A C  C
```

**Fig. 5.** Sequences of the highest-affinity RNA aptamers, in order of their measured binding constants (see Table 2), and their proposed secondary structures. The lengths were determined by end mapping, and the secondary structures based on the assumption that complementary termini form a duplex, and that the engineered hairpin is formed whenever present. *G* (10-6) has a potential 4-base stem within the binding loop, and *A* (9-4) has a number of possible structures, so no particular structure is presented. *C*, a Class I aptamer, is the best characterized structure, because 13 independent variants emerged from the selection. The variable positions in the binding loop (see Fig. 4) are denoted by lowercase letters.

advantage that a stable stem-loop is always present, this result must be balanced against the disadvantage that the recognition loops must form adjacent to this stem-loop. In contrast, the random library has the advantage that the recognition loops can occur in many different registers within the random region, but the disadvantage that in most of these cases the required stem-loop will be missing. The balance between these opposing factors (number of registers vs. chance of having a stem-loop) depends to a great extent on several factors that are difficult to assess quantitatively, including the destabilizing effect of large loops in the internal stem-loop, and the destabilizing effects of wobble and mismatch base-pairs in either stem. Both of these destabilizing effects can be compensated for to some extent by longer, more stable stems.

The results of the selection experiment described in this paper suggest that the presence of the designed stable stem-loop made the partially structured library a superior source of high-affinity aptamers. The arguments supporting this conclusion are summarized here. The most direct, albeit qualitative, argument is that the three highest-affinity aptamers, and five of the top seven, derived from the partially structured pool. A somewhat more quantitative argument can be made based on the Class I aptamers, which are characterized by the presence of highly conserved recognition loop sequences flanking a central stem-loop. Among the 16 sequences representing Class I aptamers were 12 independent sequences derived from the partially structured library and three others that were duplicates of these.

Only one Class I sequence came from the fully random library. Therefore, Class I aptamers appear to be roughly 12 times as common in the initial partially structured pool as in the fully random pool. Three other high-affinity aptamers include recognition loops flanked by two stems (Fig. 5), and all of these include the engineered stem. Thus, the partially structured library contained at least 15 sequences encoding high-affinity aptamers with a stem-loop-stem-loop secondary structure, compared with one from the random pool. The probability that this ratio would occur by chance if the two pools were not different is less than 0.01%. One reason the engineered library has such a significant effect (in addition to the simple presence of the stem-loop) may be that the stable tetraloop nucleates folding, thereby preventing the formation of many alternative structures. On the other hand, the fully random pool yielded two high-affinity aptamers with simple stem-loop secondary structures, whereas none were observed from the partially structured pool. Thus, it appears that our attempt to devise a library design that yields larger numbers of high-affinity aptamers has strongly increased the abundance of certain aptamer structures, while possibly decreasing the abundance of other structures.

Why did we observe multiple independent sequences belonging to the Class I type of aptamer, whereas all other high-affinity aptamers were represented by only one independent sequence? The simplest explanation for this observation is that the unique aptamers have a very high information content, i.e., many of the nucleotides in the recognition loops must be uniquely specified, so that the probability of occurrence of multiple independent members of these classes in a pool of random sequences is very low. If this interpretation is correct, it may be that even higher-affinity aptamers exist, but their structures may be so complex that they can be recovered only from much larger libraries, or from libraries that incorporate additional designed elements.

1. Ellington, A. D. & Szostak, J. W. (1990) *Nature (London)* **346,** 818–822.
2. Tuerk, C. & Gold, L. (1990) *Science* **249,** 505–510.
3. Wilson, D. S. & Szostak, J. W. (1999) *Annu. Rev. Biochem.* **68,** 611–647.
4. Hesselberth, J., Robertson, M. P., Jhaveri, S. & Ellington, A. D. (2000) *J. Biotechnol.* **74,** 15–25.
5. Famulok, M. (1999) *Curr. Opin. Struct. Biol.* **9,** 324–329.
6. Famulok, M. & Mayer, G. (1999) *Curr. Top. Microbiol. Immunol.* **243,** 123–136.
7. Jhaveri, S., Rajendran, M. & Ellington, A. D. (2000) *Nat. Biotechnol.* **18,** 1293–1297.
8. Osborne, S. E., Matsumura, I. & Ellington, A. D. (1997) **1,** 5–9.
9. Famulok, M., Mayer, G. & Blind, M. (2000) **33,** 591–599.
10. Sassanfar, M. & Szostak, J. W. (1993) *Nature (London)* **364,** 550–553.
11. Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. (1994) *Science* **263,** 1425–1429.
12. Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A. & Famulok, M. (1996) *Nucleic Acids Res.* **24,** 1029–1036.
13. Fan, P., Suri, A. K., Fiala, R., Live, D. & Patel, D. J. (1996) *J. Mol. Biol.* **258,** 480–500.
14. Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., d'Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Jr., *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85,** 1364–1368.
15. Bartel, D. P. & Szostak, J. W. (1993) *Science* **261,** 1411–1418.
16. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15,** 8783–8798.
17. Pan, T. & Uhlenbeck, O. C. (1992) *Biochemistry* **31,** 3887–3895.
18. Pace, N. R., Thomas, B. C. & Woese, C. R. (1999) in *The RNA World*, eds. Gesteland, R. F., Cech, T. R. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY).