A Novel, Modification-Dependent ATP-Binding Aptamer Selected from an RNA Library Incorporating a Cationic Functionality[†]

Narendra K. Vaish,[‡,§,#] Rosa Larralde,[§] Andrew W. Fraley,[‡,⊥] Jack W. Szostak,[*,§] and Larry W. McLaughlin[*,‡]

Department of Chemistry, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts 02467, and Howard Hughes Medical Institute, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114

Received December 12, 2002; Revised Manuscript Received May 9, 2003

ABSTRACT: An analogue of uridine triphosphate containing a cationic functional group was incorporated into a degenerate RNA library by enzymatic polymerization. In vitro selection experiments using this library yielded a novel receptor that binds ATP under physiological pH and salt conditions in a manner completely dependent on the presence of the cationic functionality. The consensus sequence and a secondary structure model for the ATP binding site were obtained by the analysis of functional sequences selected from a partially randomized pool based on the minimal parental sequence. Mutational studies of this receptor indicated that several of the modified uridines are critical for ATP binding. Analysis of the binding of ATP analogues revealed that the modified RNA receptor makes numerous contacts with ATP, including interactions with the triphosphate group. In contrast, the aptamer repeatedly isolated from natural RNA libraries does not interact with the triphosphate group of ATP. The incorporation of a cationic amine into nucleic acids clearly allows novel interactions to occur during the molecular recognition of ligands, which carries interesting implications for the RNA world hypothesis. In addition, new materials generated from such functionalized nucleic acids could be useful tools in research and diagnostics.

The RNA world hypothesis (*1*) postulates that early organisms were dependent upon RNA, not protein, molecules for biochemical transformations (*2*−*4*), and that over the course of evolution the present day DNA−RNA−protein world developed from this ancestral RNA-based state. This hypothesis was strengthened with the discovery of catalytic RNA molecules ("ribozymes") (*5*, *6*). The development of in vitro selection methodologies (*7*−*9*) subsequently led to the discovery of hundreds of novel RNA and DNA motifs that can act as catalysts and high affinity receptors (*10*, *11*). Nevertheless, the functional capacity of RNA is probably limited by the narrow range of functionality and geometric constraints available from its limited set of nucleotide building blocks (*12*, *13*). The dearth of functional groups in RNA compared with proteins is probably why proteins replaced the majority of functional RNAs, and has been suggested to explain why in vitro evolved nucleic acids perform relatively poorly as catalysts compared with natural enzymes and certain catalytic antibodies (abzymes) (*14*), particularly when comparing a specific reaction type (*15*−*18*). On the other hand, it is difficult to disentangle the effects of functional group diversity from the limitations of particular selection technologies. If the paucity of functional groups is a real limitation for nucleic acids, then endowing nucleic acids with additional functionalities should improve their structural and catalytic properties (*13*). From an evolutionary perspective, some of the more than 90 posttranscriptional RNA modifications isolated from biotic sources (*19*) are of potential catalytic importance, and may be molecular fossils derived from the RNA world.

Naturally occurring RNAs containing monomers with a positive charge are rare, but archaeosine (the amidine derivative of 7-deazaguanosine) (*20*) is present in the tRNA of all archaea. Under physiological pH, this aromatic amidine would be expected to protonate and exist largely with a positive charge. Archaeosine is found in place of G15, normally a semi-invariant residue in bacterial and eukaryal tRNAs. This residue is involved in maintaining the tertiary interactions between the D-loop and TΨC loop. In the elevated temperature environment of the archaea, the additional complementary ion-pair interactions involving the protonated archaeosine and one or more negatively charged phosphodiesters might prove critical to the structure and thus the function of archaeal tRNAs.

One way of increasing the functional group diversity available to nucleic acids is the rational or directed approach of using altered building blocks (*21*−23) or including other organic molecules as cofactors for ribozymes (*24*, *25*). An expanded genetic alphabet could also enhance the structural and functional diversity of nucleic acids (*26*). In vitro selection is a combinatorial technique that generates receptors, ligands, and catalysts from nucleic acid libraries

[†] This work was supported by grants from the NSF (MCB-0066776) to L.W.M. and from the NIH (GM-53936) to J.W.S.

* Correspondence to Jack W. Szostak (Phone: 617-726-5981. Fax: 617-726-6893. E-mail: szostak@molbio.mgh.harvard.edu) and Larry W. McLaughlin (Phone: 617-552-3622. Fax: 617-552-2705. E-mail: larry.mclaughlin@bc.edu).

[‡] Boston College.

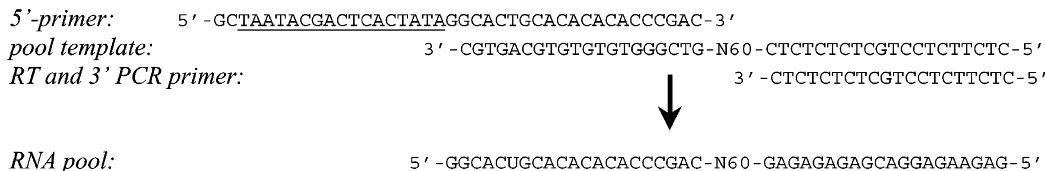
[§] Howard Hughes Medical Institute, and Massachusetts General Hospital.

[#] Present address: Sirna Therapeutics, Inc., 2950 Wilderness Place, Boulder, CO 80303, USA. E-mail: vaishn@sirna.com.

[⊥] Present address: ULP−Faculté de Pharmacie, Laboratoire de Chimie Génétique, 74 route du Rhine − BP 24, 67401 Illkirch cedex, France. E-mail: fraley@bioorga.u-strasbg.fr.

10.1021/bi027354i CCC: $25.00 © 2003 American Chemical Society
Published on Web 07/04/2003

a)



b)

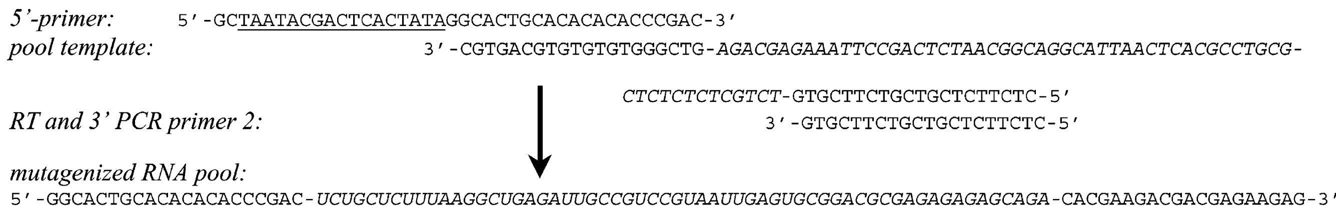


FIGURE 1: Pool construction. (a) random pool, and (b) mutagenized pool based on selected 10N23 aptamer. The 60 italicized bases (mutagenized at 21%) include 47 nucleotides derived from the 3′ end of the 10N23 random region and 13 adjacent nucleotides from the original 3′ constant region. The T7 RNA polymerase promoter sequence in the 5′ primer is underlined; transcription initiates immediately following the underlined sequence. The last T in the mutagenized-pool random region was introduced to extend the outer base-paired stem of the aptamer.

containing as many as $10^{14}-10^{16}$ different molecules and is a powerful method to explore the sequence space of biopolymers in search of functional domains. Some effort has been devoted to developing in vitro selection processes that use modified monomer triphosphates as substrates for polymerases (*27*−*36*), and several in vitro selection experiments have been undertaken to select for catalysts or aptamers whose functions depend on the presence of modified nucleotides (*37*−*48*), with notable success (*39*, *40*, *43*−*47*). On the other hand, modified building blocks have not proven to always be critical in obtaining superior receptors or catalysts (*40*, *42*, *48*, *49*). For example, comparable or better Diels−Alderase ribozymes were obtained from pools assembled with natural nucleotides (*49*) vs corresponding modified pools (*41*). Similarly, no significant catalytic advantage for modified DNA is evident when phosphodiesterase activity was the selection criterion for sequences containing imidazole-modified deoxyuridine in the presence of divalent zinc, relative to sequences constructed of natural deoxynucleotides in the presence of either divalent magnesium (*37*, *50*) or divalent zinc (*51*).

The outcome of an in vitro selection experiment depends on many parameters, including the selection conditions, the number of cycles performed, replication biases for or against the incorporation of modified nucleotides, and the nature of the fitness landscape (the map of fitness onto sequence space). This multiplicity of factors makes a valid comparison of natural and modified RNAs experimentally difficult, but at a minimum, selection experiments to be compared should be carried out under similar conditions. In the present study, we have incorporated a cationic functional group into RNA and selected for receptors for ATP. We have then compared the sequences obtained with this added functionality with those obtained from natural sequences subjected to selection under identical conditions. Here we describe the structure and binding properties of a modification-dependent aptamer that emerged from this selection experiment.

MATERIALS AND METHODS

Materials. The sequences of the primers and pool DNA templates are shown in Figure 1. Oligodeoxynucleotides were synthesized using standard phosphoramidite chemistry on an Applied Biosystems 381A DNA synthesizer and purified as described (*52*). To counteract the coupling efficiency bias for the various nucleoside phosphoramidites a 3:3:2:2 mixture of A:U:G:C was used to synthesize randomized positions (*53*).

[α-^{32}P]CTP (3000 Ci/mmol, 250 μCi), [α-^{32}P]ATP (6000 Ci/mmol and [γ-^{32}P]ATP (6000 Ci/mmol, 250 μCi) were purchased from ICN (Costa Mesa, CA). T7 RNA polymerase (200 U/μL) was purchased from Ambion (Austin, TX) or prepared from overproducing plasmid pAR1219 according to the method of Zawadzki and Gross (*54*). Superscript II RNase H−reverse transcriptase (200 U/μL) was purchased from Life Technologies (Grand Island, NY). *Taq* DNA polymerase, RNase T1, T4 PNK, and calf intestinal alkaline phosphatase were purchased from Boeringer (Mannheim, Germany). The common nucleotide triphosphates were purchased from Pharmacia (Piscataway, NJ) and the 5-(3-aminopropyl)uridine triphosphate (UNH_2) was prepared as reported (*30*). The TA cloning kit was purchased from Invitrogen (Carlsbad, CA). Polyacrylamide gel electrophoresis was performed in an apparatus from Life Technologies (Grand Island, NY). Gels were imaged in a Phosphorimager 425 from Molecular Dynamics (Sunnyvale, CA). PCR reaction was performed in PTC-200 Peltier thermocycler (MJ Research, Waltham, MA). Epoxy agarose, ATPγS, inosine-5′-triphosphate (ITP), 2-chloro ATP, 5′-deoxyadenosine, 3′-deoxyATP, and 8-bromo ATP were purchased from Sigma. 6-(*N*-methyl)-ATP, 1-methyl ATP, 3′-*O*-methyl ATP, 2′-*O*-methyl ATP and 2-amino ATP were from Trilink BioTechnologies (San Diego, CA) and 7-deaza ATP from IBA GmbH (Göttingen, Germany). 2′-Deoxy ATP was purchased from USB (Cleveland, Ohio). Microcon filters were from Amicon.

DNA Pools. Synthetic ssDNA containing a 60-residue random region flanked by two defined primer binding sites

of 20 residues (1.77 nmol of pool A ssDNA) was used as template in a 20 mL PCR reaction for the synthesis of a dsDNA pool (Figure 1). Approximately 15% of the synthetic DNA template, or $\sim 2 \times 10^{14}$ unique molecules, could be PCR amplified. The PCR reaction mixture contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.25 mM each of the four deoxynucleoside triphosphates, 4 μM primers 1 and 2 and 0.025 U/μL *Taq* polymerase. Three PCR cycles were performed with a thermal cycle of 92 °C for 2 min, 55 °C for 1 min, and 72 °C for 2 min. Subsequently, 1.1 equiv of EDTA was added to the PCR reaction to chelate the Mg^{2+} ions and the PCR reaction was concentrated 10-fold by extraction with *n*-butanol followed by phenol/$CHCl_3$ extraction. The dsDNA pool was ethanol precipitated and redissolved in 1 mL of TE buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 50 mM NaCl). The doped dsDNA library was prepared similarly using oligonucleotide pool B and primers 1 and 3.

Preparation of the ATP-Agarose Affinity Column. The ATP-agarose affinity column was prepared by stirring epoxy-agarose (2 g) with ATPγS (5 mM) in sodium phosphate buffer (pH 8.0, 100 mM) in a total volume of 10 mL for 16 h at 37 °C. Upon the basis of the remaining ATP in the supernatant solution, the epoxy-agarose contained 3 mM ATP. The unreacted epoxy-agarose was quenched by reacting it with mercaptoethanol (10 mM, pH 8.0) for 16 h. The ATP-agarose was washed with 5 column volumes of 1 M NaCl and stored in 20 mM Tris-HCl (pH 7.12) and 1 M NaCl as a 50% slurry at 4 °C. A preselection agarose gel was prepared from epoxy-agarose by treatment with mercaptoethanol (10 mM, pH 8.0) for 16 h at 37 °C and stored in 20 mM Tris-HCl (pH 7.12) and 1 M NaCl.

Primary Selection. The dsDNA pool library was transcribed (*30*) to generate RNA libraries containing either natural nucleotides only (UTP pool), or the 5-(3-amino)-propyl uridine triphosphate in place of UTP (UNH_2 pool). RNA was internally labeled with α^{32}P-CTP to monitor the course of the selection. Both RNA libraries were purified by 10% denaturing polyacrylamide gel electrophoresis (PAGE) and isolated in water by the crush and soak method. The RNA pools were preselected using the preselection agarose column to minimize the enrichment of RNA molecules interacting with column material (*55*). Preselection agarose and ATP-agarose (0.3 mL each) were equilibrated with 20 column volumes of selection buffer (100 mM NaCl; 1 mM $MgCl_2$; 20 mM Tris-HCl, pH 7.12). RNA (3400 pmol) was denatured in water (95 °C, 5 min) and immediately mixed with 5× selection buffer to bring the final concentration to 1× selection buffer at 5 μM RNA and the mixture was allowed to cool to ambient temperature (10 min). The equilibrated RNA was applied to the preselection column, eluted with 2 volumes of selection buffer, and applied directly to the selection column (ATP-agarose). After the ATP-agarose column was washed with 10 column volumes of selection buffer, bound RNAs were affinity eluted with 5 column-volumes of selection buffer containing 5 mM ATP and 5 mM $MgCl_2$. The eluted RNAs were ethanol precipitated in the presence of 100 μg of glycogen and desalted on a NAP-5 desalting column (Pharmacia). After reverse transcription and PCR amplification, DNA templates were transcribed as noted above and the resulting RNA pool was used in the next round of selection. Successive selection cycles were performed as for round 1 except that only ~1000 pmol of RNA was added to the column. A total of 10 rounds of iterative selection were performed. The dsDNA from round 10 was cloned into the vector pCR 2.1-TOPO using a TOPO TA cloning kit according to manufacturer's instructions. DNA for sequencing was amplified from single white colonies with M13 forward and reverse primers. For the analysis of RNA from individual clones, DNA from individual colonies was amplified using the PCR primers previously used for library amplification, and transcribed as above.

Defining the Minimal Aptamer Domain. To determine the 5′-end of the functional region of this aptamer, RNA transcribed with UNH_2 was digested with $NaHCO_3$ and then the neutralized RNA was fractionated on ATP-agarose as described above. RNA samples from the column flow-through and the ATP-eluted fractions were reverse transcribed using 5′-^{32}P labeled RT-primer 2. Dideoxy-cytidine (ddC) and dideoxy-guanosine (ddG) sequencing ladders were generated for use as markers by using a 3.3:1 ratio of ddNTP and dNTP in a reverse transcription reaction on intact RNA. Reverse transcription to generate the ddC ladder was carried out at 5 mM Tris (pH 8.3), 25 mM KCl, 5 mM $MgCl_2$, 20 mM DTT, 100 μM each of dATP, dTTP and dGTP, ddCTP (82.5 μM) and dCTP (50 μM), and Superscript II RNase H^- reverse transcriptase (8 U/μL). The ddG ladder was created similarly. Samples were analyzed by 6% denaturing PAGE (Figure 5A).

To determine the 3′-end of the functional region, the aptamer was transcribed using UNH_2. The terminal 5′-triphosphate was removed by the action of calf intestinal alkaline phosphatase and the RNA was 5′-end labeled using γ^{32}P-ATP and T4 polynucleotide kinase. The 5′-labeled RNA was partially digested with 50 mM $NaHCO_3$ (pH 9) containing 2 mM EDTA at 90 °C for 8 min and neutralized with HCl to generate an alkaline ladder. The base hydrolyzed RNA was fractionated on ATP-agarose under selection conditions as described above. RNA from column flow-through and ATP-eluted fractions was precipitated and desalted using a NAP-5 desalting column. An alkaline ladder and T1 digest of 5′-labeled RNA transcribed with UTP were also generated as approximate size markers because an effective alkaline hydrolysis ladder could not be generated from RNA transcribed with UNH_2 (Figure 5B).

Secondary Selection. The essential residues (from boundary mapping) of clone 10N23 were mutagenized at a rate of 21% by the chemical synthesis of a new degenerate oligonucleotide (Figure 1), and the resulting library was subjected to nine rounds of selection under conditions similar to those described above. The stringency of the selection procedure (column washing) was increased from 10 column volumes in rounds 1 and 2 (selection rounds 11 and 12) to 20 column volumes in rounds 3 and 4, 50 column volumes in round 5, 80 column volumes in round 6, and 100 column volumes in rounds 7−9 (selection rounds 17−19). The dsDNA after round nine (selection round 19) was cloned and 30 clones were sequenced.

K_d Determination. Solution binding experiments were performed by ultrafiltration of equilibrated mixtures of aptamer and target molecules using Microcon 30 spin-filters. RNA concentration was measured by UV absorption of thermally denatured samples, as monitored by melting curve

determination (Cary UV spectrometer). RNAs were subjected to a folding protocol consisting of a brief incubation at 80 °C in water, followed by mixing with 2× buffer and two cycles of heating and cooling (3 min ramp to 94 °C, 2 min at 94 °C, 10 min ramp to 20 °C, 2 min at 20 °C) (Perkin-Elmer thermal cycler). For K_d determinations, serial dilutions of folded RNA were equilibrated with ≤1 nM (γ^{32}P)-ATP or (α^{32}P)-ATP for at least 1 h. Equilibrated solutions (150 or 200 μL) were transferred to spin-filters and centrifuged for 11 s at 16 000 rpm to yield about 10 μL of filtrate, which was discarded. After an additional spin of 16−17 s, 15 or 20 μL aliquots of both the filter chamber ("top") and the receiving tube ("bottom") were collected and counted by Cerenkov radiation. Data were analyzed using Deltagraph 4.0 (DeltaPoint). K_d values were determined using the equation: $y = b + c(x/(x + k))$; where y is the fraction of labeled ATP bound (("top" − "bottom")/"top"), x is RNA concentration, k is the K_d, b represents nonspecific binding to the aptamer and the filter (ideally zero), and c is the maximum fraction of counts that can be bound (b, c, and k are parameters fitted through nonlinear regression). The K_d values obtained with either α- or γ-labeled ATP were not substantially different. IC_{50}'s were determined in a similar fashion, with RNA concentration kept constant (1.25 μM) and varying concentrations of cold competitor added to prefolded aptamer preequilibrated with trace (α^{32}P)-ATP. After reequilibration and a spin-filtration of the sample as above, data analysis was conducted with the same equation, with x representing the concentration of competitor (determined through UV absorbance), k being the IC_{50}, b representing total initial binding (specific binding, a function of both the RNA concentration and the quality of the labeled ATP, plus nonspecific binding), c representing the maximum decrease in binding, and $b - c$ reflecting nonspecific binding.

RESULTS

One obvious difference between the building blocks used for proteins and those used for nucleic acids is the presence of the positively charged aliphatic amine of lysine. This type of residue if present in RNA could result in new types of charge−charge interactions, especially involving the phosphate residues. To test this idea, we performed two in vitro selection experiments in parallel, one with UTP tethering a primary aliphatic amino group (present for all uridines) in an RNA combinatorial library, and a second using only natural NTPs. We chose to select for ATP binding motifs because at least three laboratories have selected for RNA aptamers for ATP, or biological molecules carrying the adenosine moiety, and each set of experiments reported the same solution (*56−58*). This consensus sequence and secondary structure in all cases bound only to the adenosine portion of the molecule and did not interact with the negatively charged phosphate groups. It is possible that it is difficult to evolve interactions between the RNA aptamer and the phosphate residues of ATP, especially under relatively low-salt and/or low Mg^{2+} conditions, because of the negatively charged phosphate backbone of the RNA. At the outset, we speculated that by including a cationic functional group in selection experiments, modified RNA might have an advantage over its natural counterpart because of the possibility of salt-bridge interactions or additional hydrogen bonds. Moreover, a protonated amine might interact with the charged phosphate backbone of RNA aptamers and facilitate folding of the RNA into new types of active complexes by stabilizing inter- or intramolecular duplexes (*59*).

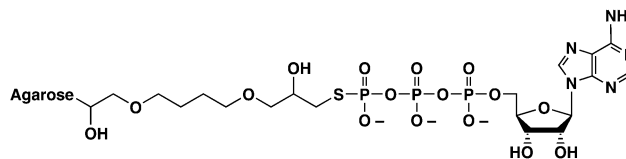


FIGURE 2: ATP-agarose used in selection procedure.

5-(3-Aminopropyl)-uridine, designated as UNH_2, was synthesized as described previously (*30*). The side chain primary amine, with an expected pK_a of ~9.5, is positively charged at physiological pH. We previously reported that UNH_2 can be incorporated into RNA by T7 RNA polymerase transcription, and could thus be used to generate RNA libraries with increased chemical diversity. We demonstrated that UNH_2 satisfies all the conditions for in vitro selections, namely: (i) the nucleotide triphosphate must act as a substrate for RNA polymerase; (ii) reverse transcriptase must be able to accurately copy the resulting modified RNA into cDNA; (iii) the modified nucleotide must replace the native nucleotide entirely to preserve information.

Two in vitro selections were performed in parallel to select for RNA receptors that bind ATP, one using a library synthesized from A, C, G, and U (to explore natural sequence space) and a second using a library synthesized from A, C, G, and UNH_2 (to explore unnatural sequence space) (Figure 1). The selection procedure of Sassanfar and Szostak (*58*) was followed with two modifications. Low ionic strength conditions (1 mM $MgCl_2$ and 100 mM NaCl, pH 7.1) were chosen to resemble physiological conditions. Another difference was the presentation of ATP on the selection column. Sassanfar and Szostak immobilized the ATP through its C8 position using a diaminohexyl linker to cyanogen bromide-activated agarose. In our selections, ATPγS was immobilized on an epoxy agarose column via the γ-thiotriphosphate (Figure 2), thus allowing a less hindered approach of an aptamer to the adenosine moiety.

A total of 10 rounds of iterative selection and amplifications were performed. Approximately 5% of the modified RNA was bound and eluted after round 10. By comparison, only 0.5% of the natural RNA was bound and eluted after round 10 (Figure 3). Further rounds of selection did not result in an increase of eluted material. Therefore, eluted RNA from round 10 of both selections was PCR amplified and sequenced. Fourteen sequences were obtained from the UNH_2 pool, and 16 sequences from the UTP pool (Figure 4). Comparison of the sequences from the two pools revealed that UTP and UNH_2 incorporation accounted for ~25% of the residues, suggesting that the UNH_2 was not strongly discriminated against during repetitive cycles of pool amplification. Sequences from the UTP pool fell into four sequence families, A, B, C, and D, and a heterogeneous group E of unique, unrelated sequences. Sequences from families A and B carried the motif reported by Sassanfar and Szostak (the Sassanfar aptamer) (*58*), and family D carries a mutant version of this motif. Since selections were carried out in a low ionic strength buffer, sequences of class C, D, and E from the UTP pool were assessed for binding both under selection conditions and under those reported by

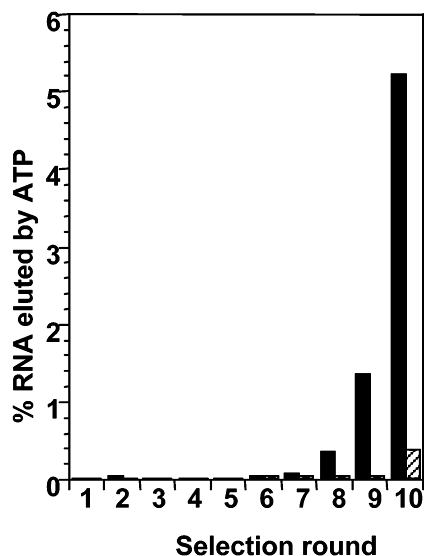


FIGURE 3: Progress and results of the selection. Bars represent the percentage of ^{32}P counts from each pool that bound to the ATP-column and eluted with free ATP. Filled bars, UNH_2 pool; diagonally hatched bars, UTP pool.

Sassanfar and Szostak (10 mM $MgCl_2$, 300 mM NaCl, 10 mM Tris-HCl, pH 7.6). None of these sequences bound more than 2% either in low or high ionic strength buffers.

Sequences from the UNH_2 pool fell into two families. Family 1 contained 13 sequences that differed from each other by only a few point mutations and thus appeared to arise from single parent sequence. All these sequences carried a mutant version of the Sassanfar motif that is similar to the mutant motif in family D above. Family 2 contained only one sequence, 10N23; it did not resemble the other sequence solutions. To assess the contribution of the UNH_2 nucleotides to the binding of receptors to ATP, all of the sequences from the UNH_2 pool were transcribed with either UNH_2 or UTP, and the resulting RNAs tested for binding to ATP-agarose. All sequences from family 1 bound to the ATP-agarose similarly (2−4% binding) after 10 column volume washes, irrespective of the modification. Although the Sassanfar motif contains one highly conserved (95%) uridine in its recognition loop (*56*), modification at this position does not appear to influence ATP binding. However, sequence 10N23 showed 4% binding to ATP-agarose when transcribed with UNH_2 but less than 0.1% when transcribed with UTP. Thus, the column-binding behavior of clone 10N23 was highly dependent on the presence of the amino modification.

To define the minimal region of the 10N23 sequence that was required for ATP binding, we mapped the 5′ and 3′ boundaries of the binding domain (Figure 5). 5′-End-mapping by reverse-transcription of functional RNA fragments (see Methods) showed that the entire 5′ constant region and as many as 12 nucleotides from the 5′ side of the random region could be deleted without loss of ATP-binding ability. In contrast, 3′ end mapping suggested that 8−10 nts of the 3′ constant region were required for ATP-binding; such requirements are often observed in selection experiments. Together, these experiments showed that a contiguous stretch of ∼60 nucleotides was required to contain the entire functional ATP-binding structure (Figure 6). Mfold analysis (*60*, *61*) of this region of the 10N23 sequence suggested a possible secondary structure (Figure 8).

To further define the nucleotides of the 10N23 RNA that were critical for ATP binding, a highly degenerate pool of sequences was prepared by partially randomizing residues 34−94 of clone 10N23 (Figures 1 and 6), a region that includes 12 nucleotides from the original 3′ constant region. Each nucleotide position was doped at 21% (i.e., at each doped position the parental base was present at a frequency of 0.79, and each of the other three bases was present at a frequency of 0.07). A synthetic oligonucleotide library consisting of 60 mutagenized bases flanked by the original 5′ constant region and a new 3′ constant region was used as a template for PCR amplification and T7 transcription (Figure 1). Selections were carried out similarly to the primary selection process; however, the selection column was washed with increasing amounts of selection buffer to progressively increase the stringency of the selection. In the final round, the selection column was washed with 100 column volumes of selection buffer and approximately 5% of the applied RNA eluted specifically with 5 mM ATP. No further rounds of selection were done to maintain as much variability as possible in the selected pool. Thirty clones were sequenced from the final round; the aligned sequences are shown in Figure 7. Fifteen of the clones (19N1−19N17) were tested for binding to ATP-agarose and all behaved similarly to the parental clone 10N23 (data not shown). A secondary structure diagram of the aptamer, along with a summary of the results from the reselection experiment, is shown in Figure 8. The presence of the internal base-paired stem is strongly supported by compensatory base-pair covariations at three positions. The presence of the hypothesized outer stem was not supported by simple covariation, and the sequences are quite variable, although all clones retained the potential for significant base-pairing in this general region. To further test the significance of this potential base-pairing, we prepared a shortened aptamer lacking the outer stem entirely (10N23-core) (Figure 6). This construct, corresponding to the sequence from G_{46} to C_{78}, exhibited a K_d of 13 μM, an affinity within 10-fold that of optimized versions of the full sequence (∼2 μM, see below), indicating that the outer-stem stabilizes the aptamer, but is not essential for function. A careful analysis of the sequence of 10N23 suggested that alternate outer stems could form by slipping the pairing sequences by two or four residues; thus, it could potentially exist in several conformations. To avoid having a mixture of competing structures, we redesigned the outer stem of constructs used for further characterization such that they could only exist in a single conformation (10N23−S and 10N23−6) (Figure 6). A 13-nucleotide region in loop I was highly conserved. Several potential short base-paired stems can be formed within this conserved region, or between it and the closing stem−loop; however, the significance of these potential stems is unclear due to sequence invariance and lack of covariation.

To investigate which amino-uridines were essential for the binding of ATP, each UNH_2 within the core region was replaced by cytidine, using construct 10N23−S as a starting point (K_d = 3.5 μM) (Figures 6 and 8). Synthetic T to C mutations were PCR amplified followed by transcription to generate mutant RNA. Although not an ideal substitution, this strategy allowed the substituted aptamers to be synthesized by in vitro transcription; individual substitutions with U or some other U analogue would have required an

Sequences from round 10 of UNH_2 pool:

```
Family 1:
10N5(2): GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGCGGCCT C.TTGTTGGATTGT
10N4(4): GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGTGGCCC C.TTGTTGGATTGT
10N14  : GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGGGTGGCCC C.TTGTTGGATTGT
10N15  : GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGCGGCCT CCTTGTTGGATTGT
10N22  : GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGTGGCTT C.TTGTTGGATTGT
10N3   : ..TGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGTGGCTT C.TTGTTGGATTGT
10N6   : GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGTGGCCC C..TGTTGGATTGT
10N8   : GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGTGGCCT ...TGTTGGATTGT
10N9   : GCTGA GGGCCGCGGACGTAATTGCGCGTGGAAGACCCTGCGTGGCTC ...TGTTGGATTGT

Family 2:
10N23  : GCTACCAGCTATTCTGCTCTTTAAGGCTGAGATTGCCGTCCGTAATTGAGTGCGGACGCGagagagagcag
```

Sequences from round 10 of UTP pool:

```
Family A:
10U10  : AAAAGGGAAGAAACTGCGCGCTCAGCTGGTTTAGCTGGGGGCGGCTTTTGAGAAGGCGTT
10U2   : AAAAGGGAAGAAACTGCGCGCTCCGCTGGTTTAGCTGGGGGCGGCTTTTGAGAAGGTGTT

Family B:
10U6(2): GTGATTTGTCCACGGGGAAGAAACTGCGCTGTATCCCAGCGGCCGTAGAGAGATCAATCC
10U22  : GTGATCTGTCCACGGGGAAGAAACTGCGCTATATCCCAGCGGCCGTAGAGAGATCAGTCC

Family C:
10U8(5): CGAATTACCGGTCAAAGATCCGATGAAATGCGTTAATTCAACTGCTTTGTGGAGGCCAC
10U24  : CGAATTACCGGTCAAAGACCCGATGAAATGCGTTAATTCAACTGCTTTGTGGAGGCCAC

Family D:
10U3   : CGGGCTCGGAAATCGCGTGGAAGATCCTGCGGCTAAGCCGGACGCGCCGGCCGGAAGG
10U20  : CGGGCTCGGAAATCGCGTGGAAGATCCTGCGGCTAAGCCGGACGCGCCGGCCGGAGGG

Family E:
10U1   : CATGGNAAGAGACTGCACCACTTATTGACTTATTTGGTGGATGGCGGCGATGTGCTGCTTC
10U21  : CGGTTGAGCGATAGCGAGGGTGCTGCAGGCCGTAACTGTGCAGTGTGCCCAACTTCTGCT
10U7   : CTGCCCCCGAAGTTTAGCTCAACTAGTTCAGCTTTATTATGTCCCTGTTCTTCTGTCAATA
```

FIGURE 4: Sequences of the selected ATP binding RNAs from the modified (UNH_2) and natural (UTP) pools. Numbers in parentheses next to clone names indicate the number of times each sequence was observed. Dots in class 1 sequences indicate gaps introduced to maintain alignment. Lower case residues at the end of sequence 10N23 indicate the part of the 3′ PCR primer (primer 2) binding site that was included in the active structure. Underlined sequences denote the Sassanfar motif and its variants.

extensive synthetic effort. Individual replacement of U_{49}, U_{64}, U_{67}, U_{68}, or U_{72} by C resulted in K_d values greater than 100 μM. However, replacing U_{54} by C resulted only in a modest loss of binding ($K_d = 6$ μM), and deletion of U_{55} did not affect binding at all ($K_d = 2.3$ μM). We did not mutate U_{60} since this residue mutated to C at the frequency expected for a neutral change after the secondary selection (Figures 7 and 8); furthermore, a clone carrying this mutation showed normal binding to ATP-agarose. Thus, U_{49}, U_{64}, U_{67}, U_{68}, and U_{72} seem to play critical roles in the recognition of ATP. Loss of binding upon mutation of U_{64} to C is notable because the U is thought to form a wobble base pair within the internal stem. To ascertain that the U_{72} to C mutation was not due to the disruption of a Watson−Crick base pair, we replaced the A_{65}−U_{72} base pair with G−C. This change also resulted in a very high K_d (>100 μM). These results together suggest that the critical interactions of U_{64}−G_{73} and A_{65}−U_{72} either do not involve base-pairing, or involve additional interactions. In at least one study, an exogenous cation (propyl- or ethylamine) was added to the natural sequence to rescue binding activity (*39*). We did not perform this experiment since it is difficult to differentiate between specific and nonspecific cationic stabilization in the aptamer-ligand complex.

Aptamer 10N23−6, in which the inner and outer stems should form in a unique nondegenerate manner, is our current optimal version of this aptamer, exhibiting a K_d of ~2 μM at 1 mM Mg^{2+}. Spin-filtration binding experiments using 2 μM RNA and trace levels of labeled ATP were used to determine the optimal Mg^{2+} concentration for binding, by comparing the fraction of ATP bound under concentrations of $MgCl_2$ varying from 0 to 10 mM. These experiments revealed that binding is compromised at low Mg^{2+} concentrations, improving significantly as Mg^{2+} is increased to 3 mM Mg^{2+}, above which no further changes were observed. Further characterization of the aptamer was therefore conducted at 6 mM $MgCl_2$, so that addition of ATP or other possible Mg^{2+}-chelating agents would not perturb the free Mg^{2+} concentration to an extent that might affect binding. Figure 9 shows the ATP binding curve of 10N23−6

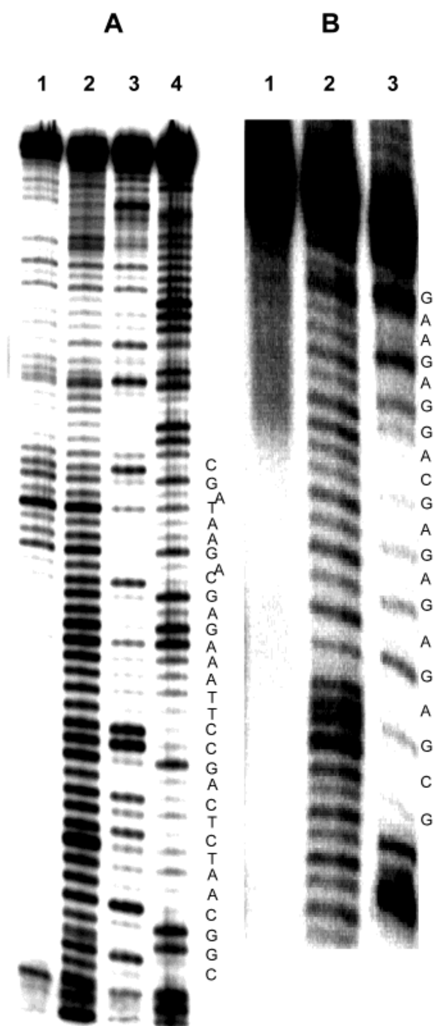


FIGURE 5: Minimal sequence required for binding of 10N23 RNA to ATP-agarose. (A) Determination of the 5′ end. RNA was partially digested with $NaHCO_3$ and neutralized before applying to ATP-agarose. Lane 1: cDNA ladder generated from the reverse transcription of the ATP-eluted fraction using 5′ ^{32}P labeled primer 2. Lane 2: cDNA ladder generated by reverse transcription of the flow-through. Lane 3: ddC ladder. Lane 4: ddG ladder. (B) Determination of the 3′ end. Lane 1: ATP eluted fraction from the ATP-agarose. The RNA was transcribed using UNH_2, 5′ end labeled, digested partially with $NaHCO_3$ and neutralized before applying to ATP-agarose. Lane 2: Partial alkaline hydrolysis ladder generated from 5′ end labeled RNA transcribed using UTP. Lane 3: Partial RNase T1 digest generated from UTP transcribed RNA.

determined by varying the RNA concentration, from which a K_d of 1.0−1.1 μM was derived. A dissociation constant calculated in this manner represents the K_d of the whole RNA sample, regardless of what fraction of it is actually active. When labeled 10N23−6 RNA was carefully folded (see Methods for K_d determination) in "high Mg^{2+}" selection buffer (6 mM $MgCl_2$; 100 mM NaCl; 20 mM Tris-HCl, pH 7.12), approximately 45% of the material bound to an ATP-agarose affinity matrix and specifically eluted with ATP. Thus, at least 45% (and possibly more) of the RNA folds into the active conformation under conditions used to measure the K_d, so the affinity of the functioning aptamer for ATP lies between 450 nM and 1.1 μM.

The specificity of the ligand−receptor interaction was studied by determining the ability of various ATP-analogues to displace labeled ATP from 10N23−6 RNA in solution. The resulting data are presented as IC_{50}'s, normalized to displacement by ATP (Figure 10). These competition binding experiments indicate that 10N23−6 RNA makes functional interactions with several distinct parts of the ATP molecule. No competition was observable with GTP, CTP, or UTP. ADP bound similarly to ATP, AMP about 2-fold more weakly, and adenosine (and 5′-deoxy adenosine) bound about 4-fold more weakly, suggesting a significant interaction of the aptamer with the α- and β-phosphates. Both the relative and absolute binding to ATP and adenosine is unaffected by high-salt (600 mM NaCl, 6 mM Mg^{2+}). A possible explanation for this observation is that all interactions with charged groups (such as the phosphates) may occur in a buried environment sheltered from the effects of external salt. Removing the ribose 2′- or 3′-hydroxyl groups resulted in a modest loss of binding. Interestingly, 3′-*O*-Me ATP bound as well as ATP, suggesting that *O*-3′ may be acting as a hydrogen bond acceptor; in contrast, 2′-*O*-Me ATP bound similarly to 2′-dATP, suggesting that the 2′-OH may be acting as a hydrogen bond donor. The extremely weak binding of ITP, 6-N-Me ATP, 7-deaza ATP, and 1-Me ATP are suggestive of several critical interactions with the nucleobase. By comparison, the Sassanfar motif bound to ATP, ADP, AMP, and adenosine with essentially equal affinities under the same conditions (K_d values 2−3 μM, data not shown) indicating that the UTP-derived motif does not recognize the triphosphate group. The negatively charged triphosphate would likely repel the RNA phosphate backbone and not contribute to binding.

Future studies will require a more molecular-level investigation of aptamers containing cationic sites. Appropriate biophysical or X-ray analysis would provide insight into the number of cationic sites and their optimal locations necessary in an RNA aptamer to properly shield the triphosphate moiety of the ligand from the destabilizing charge−charge repulsion interactions of the phosphodiesters within the receptor sequence.

DISCUSSION

The most striking aspect of our selection for ATP-binding aptamers from a UNH_2 substituted RNA pool is the emergence of a new RNA motif that is distinct from the ATP-binding aptamer that has previously been isolated from normal RNA libraries (the Sassanfar aptamer). Moreover, the new aptamer is completely dependent upon the UNH_2 substitution for its ATP-binding function. This change, the replacement of U with 5-aminopropyl U, has allowed new functional structures to form. This observation parallels previous observations that aptamers and deoxy-ribozymes selected from DNA pools usually (but not always) have very different structures than aptamers and ribozymes selected from RNA pools. Once again, a small chemical difference is translated into the emergence of completely distinct macromolecular structures (*40*, *56*−*58*, *62*).

We have shown that at least five UNH_2 residues cannot be mutated to C without complete loss of ATP binding, but chemical synthesis of aptamers with individual UNH_2 to U substitutions will be required to determine which individual aminopropyl groups are required for activity, and further structural work will be required to show how the aminopropyl groups contribute to the structure or function of the

Clone 10N23:

```
 1        10        20        30        40        50        60        70        80        90        100
5'-GGCACTGCACACACACCCGACGCUACCAGCUAUUCUGCUCUUUAAGGCUGAGAUUGCCGUCCGUAAUUGAGUGCGGACGCGAGAGAGAGCAGGAGAAGAG-3'
```

Minimal region defined by end-mapping: 5'-CUGCUCUUUAAGGCUGAGAUUGCCGUCCGUAAUUGAGUGCGGACGC-GAGAGAGAGCAGG-3'

mutagenized RNA pool:

```
5'-GGCACTGCACACACACCCGAC-     UCUGCUCUUUAAGGCUGAGAUUGCCGUCCGUAAUUGAGUGCGGACGC-GAGAGAGAGCAGA--
                                                                                --CACGAAGACGACGAGAAGAG-3'
```

Construct 10N23-core: 5'-GGCUGAGAUUGCCGUCCGUAAUUGAGUGCGGAC-3'

Construct 10N23-6: 5'-GGCAGACGUAAGGCUGAGAUUGCCGUCCGUAAUUGAGUGCGGACGCGAGACGUCUGCC-3'

Construct 10N23-S: 5'-GGCAGCUAAGGCUGAGAUUGCCGUCCGUAAUUGAGUGCGGACGCGAGAGCUGCC-3'

FIGURE 6: Sequences of the parental clone 10N23, the minimal active region defined by end-mapping, the mutagenized pool used for secondary selection, and three derivatives used in binding studies. The first and last underlined portions of 10N23−6 and −S are designed to pair in a single register; the two inner underlined regions represent the inner base-paired stem. Construct 10N23−6 was used for measurement of the K_d for ATP and affinity for derivatives and analogues; UNH_2 to C mutations were examined within the framework of construct 10N23−S.

```
                  34   40              50              60             70              80            90
10N23             TCTGCTCT.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.GCGAGAGAGAGCAGA

19N1              TC.GCTCT.TTGAGGCTGAGATTGCC.GTCAGTAATTGAGTGCTGAC.GCGAGAGGGTGTAGA
19N2              TCTGCTCT.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGACCGCCAGAGACGGGTGA
19N3              TCTGCTCT.TCTAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.GCCAGAGACGGG.GA
19N4              TCTGCTCT.TCCAGGCTGAGATTGCC.GTTCGTAATTGAGTGCGAAC.GCCAGAGAGAGGGTGAC
19N5              TCTGTTCT.TCAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.ATGAGAGAGGGTGGA
19N6              TCTGATAT.TTATGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.CCGAGAGGGTGTAGA
19N7              .CTGCTCT.TCAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.ACCAGTGAGGGTAGA
19N9              TCTGCTCT.TC.AGGCTGAGAT.GCCCGTCCTTAATTGAGTGGGGAC.ACCAGAGAGGGTGAA
19N10             TGT.CCTC.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.AGGATGGAGGGTGGA
19N11             TCTG.CTC.TTAAGGCTGAGAT.GCCCGTCCGTAATTGAGTGCGGAC.GTCAGAGTTGGAAGGA
19N12             TCTGCTCT.ATATGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.GAGAGAGAGAGTAGA
19N14             TCTGCTCT.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAT.ACCAGAGAGGGTAAA
19N15             CCTGCTCT.TCAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.GCCAGGGTGGGTAGA
19N16             TCTGACTT.CTAGGGCTGAGATTGCC.GCCCGTAATTGAGTGCGGAC.CAGAGCGAGTGTAGA
19N17             CCTGCTCT.T.CCGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.CCGGGAGAGGGTAGA
19N18             TCTG.TTT.TTGAGGCTGAGATTGCC.GTTCGTAATTGAGTGCGAAC.ACAGAGTTGGAGCAGA
19N20             TCTGGT.T.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC ACCAGAGGGAGGGGA
19N21             TCTGCTCT.TCAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAT.ACCAGAGAGAGGTGT
19N24             TCTGCTCT.TTAAGGCTGAGATTGCC.GTCTGTAATTGAGTGCGGAC.GCGGGAGAGGGTAAA
19N26             TCTGCGCT.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.ACCAGAGGGGGTAGA
19N27             TCTGCTCT ATAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAT GCGAGAGAGAGTTGA
19N29             TCTGATTT.TCA.GGCTGAGATTGCC.GCCCGTAATTGAGTGCGGTC.ACAGGAGAGGGTGAA
19N31             TCTCCGCG.TTATGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC..CGAGAGAGACGGGA
19N34             TTTGCTCT.TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.CCGAGTCGGGAGCAAA
19N37             TCTC.TCT.TCAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAT.ACGTGAGCGGGTGAA
19N38             TCTGCTCT.TTAAGGCTGAGATTGCC.GTCAGTAATTGAGTGCTGAC.GCCAGAGAGAGGGTAA
19N39             TCTGCTCT.TCTAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGGC.ACCAGAGGGGGTAGA
19N40             TCTGTTCT..TAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC.ACTAGAGAGAGGTGG
19N41             TCTGCTC....AAGGCTGAGAT.GCCAGTCCGTAATTGAGTGCGGAC.GCCAGAGAGAGGTGT
19N42             TCTGTTG..TTAAGGCTGAGATTGCC.GTCCGTAATTGAGTGCGGAC...TAGAGCGGGTAAA
```

FIGURE 7: Aligned sequences from secondary selection. The first 15 of these sequences were examined and all showed binding to ATP-agarose that was similar to that of the parental 10N23 clone. Mutations are shown in bold; a large central region of high conservation is evident. Within this region, covariations and mutations that conserve wobble-pairing are underlined; these mutations support the existence and functional requirement for base-pairing of the internal stem of the aptamer.

aptamer. Our new UNH_2-substituted aptamer engages in weak but favorable interactions with the α- and β-phosphates of ATP, in contrast to the Sassanfar aptamer which binds equivalently to ATP, ADP, AMP, and adenosine. At this point, we have no evidence that the observed phosphate interactions are mediated by the amino-groups of one or more of the modified UNH_2 residues of the aptamer.

The secondary structure of the UNH_2 substituted aptamer is unusual, in that the outer base-paired stem makes only a modest contribution to the binding affinity. A small core region of 32 nucleotides, lacking this outer stem, shows only ∼10-fold weaker binding than the larger aptamer with the outer stem. Furthermore, the sequences adjacent to the outer stem are highly variable, and seem to play a role as spacers.

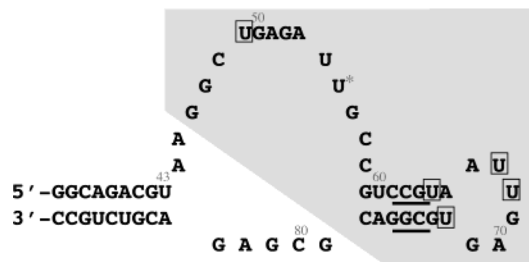


FIGURE 8: Proposed secondary structure model of construct 10N23−6. Sequence 10N23−6 was generated from clone 10N23 by changing residues in the outer stem to allow base-pairing in only one register. Numbering of residues in sequence 10N23−6 is as in full-length clone 10N23. The highly conserved region identified in the secondary selection (nts 46−78), corresponding to the core region defined by construct 10N23-core, is within the shaded box. The underlined section of the inner stem is the location of the base-pair covariation mutations observed in the secondary selection. U_{55} (labeled with an asterisk) can be deleted with little if any loss of binding affinity. Mutation of the five boxed UNH_2 residues to C decreases binding to undetectable levels.

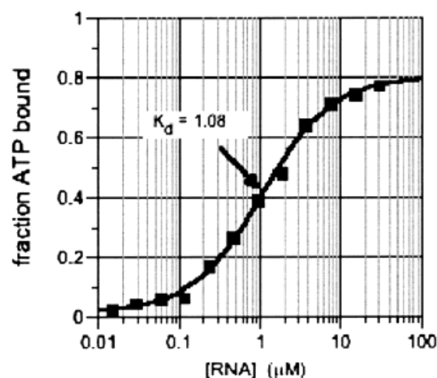


FIGURE 9: Dissociation constant of 10N23−6 and ATP complex. K_d determination by titration of RNA concentration by spin filtration method. The fitted equation is $y = 0.014 + 0.788(x/(x + 1.079))$ $R^2 = 0.996$.

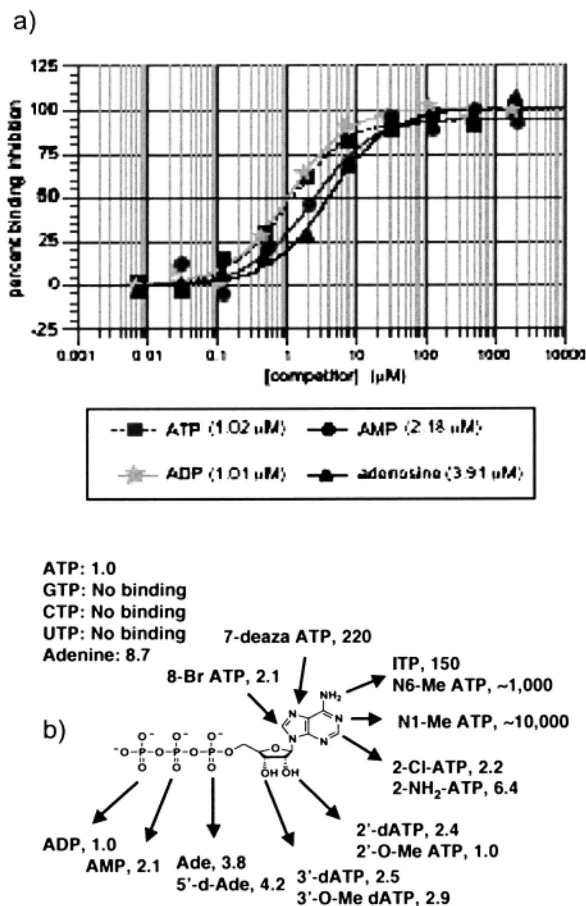


FIGURE 10: Ligand specificity. Positions involved in interaction with RNA were determined by the ability of ATP analogues to displace bound ^{32}P-ATP from 1.25 μM 10N23−6 RNA. (a) Comparison of binding to ATP, ADP, AMP, and adenosine. (b) Values indicate IC_{50}'s of various ATP analogues, normalized to the IC_{50} of ATP.

It is possible that the outer stem provides a modest structural constraint on the remainder of the aptamer, helping to stabilize the active conformation. In most stem-bulge-stem-loop aptamers, the outer stem is more important, and disruption of the stem results in complete loss of binding. Analysis of additional sequence variants, and ultimately high-resolution structural analysis, will be required to fully understand the architecture of this unusual RNA motif and the role that the aminopropyl groups play in the overall structure.

Selections for ATP-binding structures have now been carried out from a number of nucleic acids (RNA, DNA, aminopropyl-U RNA, and amino-propynyl-dU DNA, fluoresceinated U, and boronated derivatives) (*40*, *56−58*, *62−64*) and from a random sequence polypeptide library (*65*). It is of interest to compare the molecules that have emerged from these distinct libraries. Both RNA selections yielded functional ATP-receptors with affinities in the low micromolar range (*56−58*). One modification-dependent aptamer was recovered from the UNH_2−RNA pool, and this aptamer appears to have a structure that is quite distinct from that of the Sassanfar aptamer which has been repeatedly isolated from normal RNA pools (*56−58*). The UNH_2-modified aptamer does recognize the phosphates of ATP, but these interactions contribute rather little to the overall binding energy. Both the normal DNA and the modified DNA pools yielded very similar aptamers; in all cases these appear to form pseudo-symmetric structures that bind two molecules of ATP (*40*, *62*). No modification-dependent aptamers were recovered from this selection. Overall, the introduction of an amino functionality appears to have made surprisingly little difference to the outcome of these selection experiments. Both the modified RNA and DNA pools still allowed the emergence of the same receptor structures that were seen from the unmodified RNA and DNA pools (*40*, *56−58*, *62*). We suspect that the full potential of these modifications to influence the results of a selection might not be seen until selective pressures are applied that cannot be overcome by normal RNA and DNA. For example, selective pressure for a major, essential interaction with the triphosphate might lead to the recovery of aptamers from amino-modified pools but not from normal pools. Thus, further work will be required to define conditions under which the restricted chemical functionality of normal RNA and DNA actually becomes limiting.

ACKNOWLEDGMENT

We thank members of the J.W.S. and L.W.M. laboratories for helpful discussions.

REFERENCES

1. Gilbert, W. (1986) *Nature 319*, 618.
2. Crick, F. H. C. (1968) *J. Mol. Biol. 38*, 367−379.
3. Orgel, L. E. (1968) *J. Mol. Biol. 38*, 381−393.
4. Woese, C. (1967) The evolution of the genetic code. In *The Genetic Code*, pp 179−195, Harper & Row, New York.
5. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) *Cell 31*, 147−157.
6. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) *Cell 35*, 849−857.
7. Ellington, A. D., and Szostak, J. W. (1990) *Nature 346*, 818−822.
8. Robertson, D. L., and Joyce, G. F. (1990) *Nature 344*, 467−468.
9. Tuerk, C., and Gold, L. (1990) *Science 249*, 505−510.
10. Osborne, S. E., and Ellington, A. D. (1997) *Chem. Rev. 97*, 349−370.
11. Breaker, R. R. (1997) *Chem. Rev. 97*, 371−390.
12. Benner, S. A., Allemann, R. K., Ellington, A. D., Ge, L., Glasfeld, A., Leanz, G. F., Krauch, T., MacPherson, L. J., Moroney, S., Piccirilli, J. A., and et al. (1987) *Cold Spring Harb. Symp. Quant. Biol. 52*, 53−63.
13. Tarasow, T. M., and Eaton, B. E. (1998) *Biopolymers 48*, 29−37.
14. Narlikar, G. J., and Herschlag, D. (1997) *Annu. Rev. Biochem. 66*, 19−59.
15. Prudent, J. R., Uno, T., and Schultz, P. G. (1994) *Science 264*, 1924−1927.
16. Uno, T., Ku, J., Prudent, J. R., Huang, A., and Schultz, P. G. (1996) *J. Am. Chem. Soc. 118*, 3811−3817.
17. Morris, K. N., Tarasow, T. M., Julin, C. M., Simons, S. L., Hilvert, D., and Gold, L. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 13028−13032.
18. Hilvert, D., Hill, K. W., Nared, K. D., and Auditor, M.-T. M. (1989) *J. Am. Chem. Soc. 111*, 9261−9262.
19. Limbach, P. A., Crain, P. F., and McCloskey, J. A. (1994) *Nucleic Acids Res. 22*, 2183-2196.
20. Gregson, J. M., Crain, P. F., Edmonds, C. G., Gupta, R., Hashizume, T., Phillipson, D. W., and McCloskey, J. A. (1993) *J. Biol. Chem. 268*, 10076−10086.
21. Bashkin, J. K., McBeath, R. J., Modak, A. S., Sample, K. R., and Wise, W. B. (1991) *J. Org. Chem. 56*, 3168−3176.
22. Reynolds, M. A., Beck, T. A., Say, P. B., Schwartz, D. A., Dwyer, B. P., Daily, W. J., Vaghefi, M. M., Metzler, M. D., Klem, R. E., and Arnold, L. J. (1996) *Nucleic Acids Res. 24*, 760−765.
23. Tarasow, T. M., Tarasow, S. L., Tu, C., Kellogg, E., and Eaton, B. E. (1999) *J. Am. Chem. Soc. 121*, 3614−3617.
24. Roth, A., and Breaker, R. R. (1998) *Proc. Nat. Acad. Sci. U.S.A. 95*, 6027−6031.
25. Peracchi, A., Beigelman, L., Usman, N., and Herschlag, D. (1996) *Proc. Nat. Acad. Sci. U.S.A. 93*, 11522−11527.
26. Benner, S. A., Burgstaller, P., Battersby, T. R., and Jurczyk, S. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 163−181, Cold Spring Harbor, New York.
27. Lee, S. E., Sidorov, A., Gourlain, T., Mignet, N., Thorpe, S. J., Brazier, J. A., Dickman, M. J., Hornby, D. P., Grasby, J. A., and Williams, D. M. (2001) *Nucleic Acids Res. 29*, 1565−1573.
28. Brakmann, S., and Löbermann, S. (2001) *Angew. Chem., Int. Ed. 40*, 1427−1429.
29. Thum, O., Jaeger, S., and Famulok, M., (2001) *Angew. Chem., Int. Ed. 40*, 3990−3993.
30. Vaish, N. K., Fraley, A. W., Szostak, J. W., and McLaughlin, L. W. (2000) *Nucleic Acids Res. 28*, 3316−3322.
31. Matulic-Adamic, J., Daniher, A. T., Karpeisky, A., Haeberli, P., Sweedler, D., and Beigelman, L. (2000) *Bioorg. Med. Chem. Lett. 10*, 1299−1302.
32. Perrin, D. M., Garestier, T., and Helene, C. (1999) *Nucleosides Nucleotides 18*, 377−391.
33. Sakthivel, K., and Barbas, C. F. (1998) *Angew. Chem., Int. Ed. Eng. 37*, 2872−2875.
34. Ito, Y., Teramoto, N., Kawazoe, N., Inada, K., and Imanishi, Y. (1998) *J. Bioact. Compat. Polym. 13*, 114−123.
35. Dewey, T. M., Mundt, A. A., Crouch, G. J., Zyzniewski, M. C., and Eaton, B. E. (1995) *J. Am. Chem. Soc. 117*, 8474−8475.
36. Lin, Y., Qiu, Q., Gill, S. C., and Jayasena, S. D. (1994) *Nucleic Acids Res. 22*, 5229−5234.
37. Santoro, S. W., Joyce, G. F., Sakthivel, K., Gramatikova, S., and Barbas, C. F., 3rd. (2000) *J. Am. Chem. Soc. 122*, 2433−2439.
38. Teramoto, N., Imanishi, Y., and Ito, Y. (2000) *J. Bioact. Compat. Polym. 15*, 297−308.
39. Teramoto, N., Imanishi, Y., and Ito, Y. (2000) *Bioconjug. Chem. 11*, 744−748.
40. Battersby, T. R., Ang, D. N., Burgstaller, P., Jurczyk, S. C., Bowser, M. T., Buchanan, D. D., Kennedy, R. T., and Benner, S. A. (1999) *J. Am. Chem. Soc. 121*, 9781−789.
41. Tarasow, T. M., Tarasow, S. L., and Eaton, B. E. (1997) *Nature 389*, 54−57.
42. Wiegand, T. W., Janssen, R. C., and Eaton, B. E. (1997) *Chem. Biol. 4*, 675−683.
43. Latham, J. A., Johnson, R., and Toole, J. J. (1994) *Nucleic Acids Res. 22*, 2817−2822.
44. Proske, D., Gilch, S., Wopfner, F., Schatzl, H. M., Winnacker, E. L., and Famulok, F. (2002) *Chembiochem 3*, 717−725.
45. Teramoto, N., Ichinari, H., Kawazoe, N., Imanishi, Y., and Ito, Y. (2001) *Biotechnol. Bioeng. 75*, 463−8.
46. Lermer, L., Roupioz, Y., Ting, R., and Perrin, D, M. (2002) *J. Am. Chem. Soc. 124*, 9960−9961.
47. Ito, Y., Suzuki, A., Kawazoe, N., and Imanishi, Y. (2001) *Bioconj. Chem. 1*2, 850−544.
48. Zinnen, S. P., Domenico, K., Wilson, M., Dickinson, B. A., Beaudry, A., Mokler, V., Daniher, A. T., Burgin, A., and Beigelman, L. (2002) *RNA 8*, 214−228.
49. Seelig, B., and Jaschke, A. (1999) *Chem. Biol. 6*, 167−176.
50. Santoro, S. W., and Joyce, G. F. (1997) *Proc. Nat. Acad. Sci. U.S.A. 94*, 4262−4266.
51. Li, J., Wenchao, Z., Know, A. H., and Lu, Y. (2000) *Nucleic Acids Res. 28*, 481−488.
52. Ellington, A. D., and Pollard, J. D. (1998) *Curr. Protoc. Mol. Biol.* 2.11.1−2.11.25.
53. Bartel, D. P., and Szostak, J. W. (1993) *Science 261*, 1411−8.
54. Zawadzki, V., and Gross, H. J. (1991) *Nucleic Acids Res.* 19, 1948.
55. Famulok, M. (1994) *J. Am. Chem. Soc. 116*, 1698−1706.
56. Burke, D. H., and Gold, L. (1997) *Nucleic Acids Res. 25*, 2020−2024.
57. Burgstaller, P., and Famulok, M. (1994) *Angew. Chem., Int. Ed. Eng. 33*, 1084−1087.
58. Sassanfar, M., and Szostak, J. W. (1993) *Nature 364*, 550−553.
59. Antony, T., Thomas, T., Shirahata, A., and Thomas, T. J. (1999) *Biochemistry 38*, 10775−10784.
60. Jaeger, J. A., Turner, D. H., and Zucker, M. (1990) *Methods Enzymol. 183*, 281−306.
61. Zucker, M., Jaeger, J. A., and Turner, D. H. (1991) *Nucleic Acids Res. 19*, 2707−2714.
62. Huizenga, D. E., and Szostak, J. W. (1995) *Biochemistry 34*, 656−665.
63. Lato, S. M., Ozerova, N. D., He, K., Sergueeva, Z., Shaw, B. R., and Burke, D. H. (2002) *Nucleic Acids Res. 30*, 1401−1407.
64. Jhaveri, S., Rajendran, M., and Ellington, A. D. (2000) *Nat. Biotechnol. 18*, 1293−1297.
65. Keefe, A. D., and Szostak, J. W. (2001) *Nature 410*, 715−718.

BI027354I