Molecular Recognition of Amino Acids by RNA-Aptamers: An L-Citrulline Binding RNA Motif and Its Evolution into an L-Arginine Binder

Michael Famulok

Contribution from Fachbereich Chemie, Institut für Biochemie der Ludwig-Maximilians-Universität München, Am Klopferspitz 18a, 82152 Martinsried, Germany

Received October 22, 1993®

Abstract: A pool of $\sim 10^{15}$ different RNA molecules was enriched by in vitro selection for specific binding to L-citrulline [L-(+)-2-amino-5-ureidovaleric acid] in solution. Cloning and sequencing of the binding pool revealed two highly conserved regions present in all of the selected RNAs. These consensus sequences are 10 and 6 bases in length and are part of a defined secondary structure. All of the sequenced RNAs had the potential to fold into the same binding motif. On the basis of this motif, a 44 nucleotide long RNA was constructed. The dissociation constant of the 44-mer RNA and L-citrulline was determined by equilibrium gel filtration and by analytical column chromatography and was found to be 62–68 μM. Specificity of the L-citrulline-aptamer was tested by affinity elution with D-citrulline, L-arginine [L-(+)-2-amino-5-guanidinovaleric acid], L-albizziin [L-(+)-2-amino-3-ureidopropionic acid], L-glutamine, and urea. The binding to L-citrulline was weakly enantioselective: D-citrulline had a K_d of 180 μM which corresponds to a $\Delta\Delta G$ of 2.5 kJ mol^{-1} (0.6 kcal mol^{-1}). The other citrulline analogues had no or very weak affinity to the RNA-aptamer. One of the L-citrulline-binding sequences (clone 16) was resynthesized with a 30% mutation rate per base position. This doped pool was subjected to further cycles of selection and amplification on agaroses that were derivatized with L-citrulline, L-arginine, L-lysine, L-albizziin, and L-glutamine. After three cycles, the pool bound to L-citrulline and after four cycles, aptamers for L-arginine showed up. No aptamers to L-albizziin, L-glutamine, or L-lysine could be enriched. Of 22 sequences derived from the arginine pool, 11 had the potential to fold into a motif that was only slightly different from the L-citrulline-binding motif and was highly specific for its cognate ligand. However, the arginine-binding pool also contained a large number of different and much less abundant motifs with no similarity to the "parent" motif. The most abundant RNA had values of K_d of 56–76 μM for L-arginine and a K_d of 410 μM for D-arginine, which corresponds to a $\Delta\Delta G$ of 4.6 kJ mol^{-1} (1.1 kcal mol^{-1}). This study provides the first example of the reselection of an RNA-aptamer to change its binding specificity.

Introduction

Noncovalent interactions between amino acids,[1] peptides,[2] proteins,[3] or antibiotics[4] with RNA play substantial roles in many biological systems. Therefore, a wide interest exists in gaining accurate information about those factors which contribute to binding specificity, functionality, conformational transitions, and other critical features in such systems. Extensive effort has been devoted to the physical study of the interactions between dsDNA and small organic molecules or proteins.[5] Model systems for high-resolution physical studies are much easier obtained for DNA than for RNA. This is primarily due to the ease of synthesis of DNA relative to RNA, the well-defined and essentially invariant DNA secondary structure, and greater stability.[6] In contrast, RNA provides numerous different structural elements that can be used for specific contacts between the nucleic acid and a bound ligand.[3,7]

A powerful approach to obtain model systems for the study of RNA/ligand interactions is the recently developed technique of in vitro selection. In vitro selection allows the isolation of RNA molecules that bind ligands with high specificity from a large random pool of RNA oligonucleotides.[8] Notable examples of this technique include the selection of RNAs that specifically bind to organic dye molecules,[9] the isolation of RNAs that bind to L-arginine,[10] and the creation of a 40-mer RNA motif that binds very tightly to ATP.[11] Furthermore, in vitro selection has recently been used for the isolation of ribozymes with novel catalytic functions.[12] These examples show that this technique can lead to RNA molecules with features that allow a systematic exploration of factors involved in the molecular recognition, functionality, and catalysis by RNA. In this study, I use in vitro selection to isolate RNA motifs that specifically recognize the

® Abstract published in *Advance ACS Abstracts*, February 1, 1994.

(1) (a) Tao, J.; Frankel, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2723. (b) Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R. *Science* **1992**, *257*, 76. (c) Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*. 3680. (d) Yarus, M. *Science* **1988**, *240*, 1751. (e) Hicke, B. J.; Christian, E. L.; Yarus, M. *EMBO J.* **1989**, *8*, 3843. (f) Yarus, M.; Illangesekare, M.; Christian, E. *Nucleic Acids Res.* **1991**, *19*, 1297. (g) Yarus, M. *Biochemistry* **1989**, *28*, 980. (h) Yarus, M. *New Biol.* **1991**, *3*, 183. (i) Famulok, M.; Szostak, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 3990.

(2) Weeks, K. M.; Ampe, C.; Schultz, S. C.; Steitz, T. A.; Crothers, D. M. *Science* **1990**, *249*, 1281.

(3) (a) Ellington, A. D. *Curr. Biol.* **1993**, *3*, 365. (b) Green, M. R. *Ibid.* **1991**, *1*, 245.

(4) (a) Von Ahsen, U.; Schroeder, R. *Nucleic Acids Res.* **1991**, *19*, 2261. (b) Von Ahsen, U.; Davies, J.; Schroeder, R. *Nature* **1991**, *353*, 368. (c) Von Ahsen, U.; Davies, J.; Schroeder, R. *J. Mol. Biol.* **1992**, *226*, 935. (d) Davies, J.; Von Ahsen, U.; Schroeder, R. In *The RNA World*; Gesteland, R. F., Atkins, J. F., Eds.; CSHL Press: Cold Spring Harbor, NY, 1993; p 185. (e) Zapp, M. L.; Stern, S.; Green, M. R. *Cell* **1993**, *74*, 969.

(5) Dickerson, R. E.; Kopka, M. L.; Pjura, P. E. In *DNA Ligand Interactions*; Guschlbauer, W., Saenger, W., Eds.; Plenum Press: New York, 1987; pp 45–62.

(6) Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; IRL Press; Oxford, U.K., 1990; p 295.

(7) Steitz, T. A. Similarities and Differences between RNA and DNA Recognition by Proteins. In *The RNA World*; Gesteland, R. F., Atkins, J. F., Eds.; CSH Laboratory Press: Cold Spring Harbor, NY, 1993; p 219.

(8) Reviews: (a) Famulok, M.; Szostak, J. W. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 979. (b) Szostak, J. W. *Trends Biochem. Sci.* **1992**, *17*, 89. (c) Famulok, M.; Szostak, J. W. In *Nucleic Acids and Molecular Biology*; Eckstein, F., Lilley, D. M. J., Eds.; Springer Verlag: Berlin, 1993; Vol. 7, p 271.

(9) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818.

(10) Connell, G. J.; Illangesekare, M.; Yarus, M. *Biochemistry* **1993**, *32*, 5497.

(11) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550.

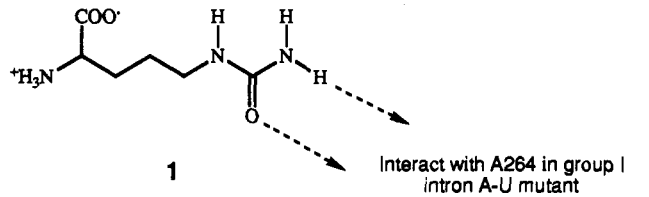
(12) Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411.

0002-7863/94/1516-1698$04.50/0

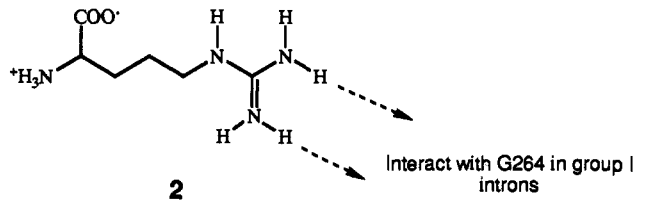
```
                T7-promotor
M38.27 TCTAATACGACTCACTATAGGAGCTCAGCCTTCACTGC-------->
                          SacI
M111.1                  GGAGCTCAGCCTTCACTGC---(N74)---GGCACCACGGTCGGATCC-3'
                                                      BamHI
M20.106                                  <--------CCGTGGTGCCAGCCTAGGTG-5'

Primers:
5'-end
M38.27        TCTAATACGACTCACTATAGGAGCTCAGCCTTCACTGC
M27.39  CCGAATTCTAATACGACTCACTATAGG  (primer to add EcoRI site to M111.1)
     EcoRI
3'-end
M20.106  GTGGATCCGACCGTGGTGCC
```

Figure 1. Design of the random 111-mer oligonucleotide. The underlined positions indicate restriction sites, the boxed region the T7-promoter sequence. The 20-mer primer has a two-base overhang at the 5'-end, resulting in a 113-mer RNA after PCR and in vitro transcription. For cloning, the EcoRI restriction site was introduced by PCR using primer M27.39 after the selections.

Scheme 1. Structures of L-Citrulline (**1**) and L-Arginine (**2**)[a]



1



2

[a] The positions depicted interact with bases at position 264 in *Tetrahymena* group I intron RNA and a mutant *Tetrahymena* group I intron RNA through hydrogen bonding.[13]

amino acids L-citrulline and L-arginine. Furthermore, I report the first example of the evolution of an RNA-aptamer to change its binding specificity.

L-Citrulline (cit, **1**) and L-arginine (arg, **2**) differ in two ways: (i) the urea group in **1** is neutral whereas the guanidino group in **2** is basic; (ii) the C=O in the urea group of **1** is a hydrogen bond acceptor while the C=N—H in the guanidino group of **2** is a hydrogen bond donor. The goal of this study was to first isolate a binding motif for L-citrulline and then to evolve this motif for specific recognition of L-arginine. This selection strategy was chosen in order to obtain Cit/Arg-related sequences in which the changes in the genotype can be directly related to changes in the phenotype. In addition, the reselection of a mutated RNA motif that recognizes an amino acid with a neutral side chain was expected to reduce the selection of aptamers that interact with arginine mainly via electrostatic interactions through the guanidinium group to a minimum.

Several examples in biological systems exist in which interactions between amino acids, particularly **1** and **2**, and RNA play an important role. For example, the self-splicing reaction of the group I intron of *Tetrahymena*, which requires guanosine as a cofactor, is inhibited by arginine.[1d] One critical feature of this inhibition is the contact of G^{264} and the guanidino group of arginine, which can substitute for the two H-donor sites of guanine. A mutant ribozyme, in which the G^{264}·C^{311} base pair is changed into an A·U requires 2-aminopurine instead of guanosine as the cofactor. This mutant is effectively inhibited by L-citrulline (**1**), while L-arginine (**2**) has no effect[13] (Scheme 1). A second example is the interaction of the human immunodeficiency virus (HIV) TAT protein, a transcriptional activator, with a stem–loop structure of transactivation response (TAR) RNA, located at the 5'-end of HIV-1 mRNA. Critical for the recognition of TAT and TAR is a single arginine within a basic region of TAT.[14] L-Arginine binds specifically to TAR, but with a higher K_d (4 mM) than within the context of the protein or a short peptide.[15]

(13) Michel, F.; Hanna, M.; Green, R.; Bartel, D. P.; Szostak, J. W. *Nature* **1989**, *342*, 391.

Experimental Section

1. Materials. dNTPs, NTPs, glycogen, all restriction enzymes, and DNase I were purchased from Boehringer Mannheim, α^{32}P-GTP, L-[2,3,4,5-^{3}H]citrulline, and L-[2,3,4,5-^{3}H]arginine·HCl from Amersham, Taq-polymerase from Stratagene, Superscript reverse transcriptase (RNaseH$^-$) from Gibco-BRL, and RNAsin from Promega. T7-RNA polymerase was purified as described.[16] Sephadex G25 F was from Pharmacia. Oligonucleotides were synthesized on an Applied Biosystems oligonucleotide synthesizer using standard phosphoramidite chemistry, unless otherwise stated.

2. Agaroses. All agaroses were obtained by derivatization of epoxy-activated Sepharose 6B from Pharmacia with the respective amino acids, according to the manufacturers protocol. The agarose has a hydrophobic 11-atom spacer arm. Agaroses were prepared with 3.0, 6.0, 15.0, and 33.0 μM ligand per milliliter of agarose (3.0–33.0 mM). For preselection, a glycine-derivatized epoxyagarose (15.0 mM) was used. L-Citrulline was also coupled to *N*-hydroxysuccinimide-activated agarose (CH-Sepharose 4B) from Pharmacia.

3. RNA-Pools. For the starting selection on L-citrulline, a 111-mer DNA with a random insert of 74 nucleotides was synthesized on a Milligen/Biosearch DNA synthesizer Model 8750. The sequences of the primer binding sites and the primers are shown in Figure 1. To reduce depurination reactions, the detritylation time was reduced by 20 s in each cycle. The synthesized 111-mer was deprotected in 30% ammonia and precipitated with n-butanol.[17] The DNA was redissolved in 400 μL of water and purified on an 8% denaturing (8.5% urea) polyacrylamide gel. The band was visualized by UV-shadowing, cut out, and eluted with 3 mL of 300 mM NaCl. Pure ssDNA (175 μg, 4.7 nmol) was obtained, and ~36% of this DNA was PCR-amplifiable to the full-length 111-mer. All 175 μg of DNA was PCR-amplified in a total volume of 200 mL in 20 tubes, each containing 10 mL of PCR-mix (10 mM Tris-HCl, pH = 8.3; 50 mM KCl; 0.001% gelatine; 1.5 mM $MgCl_2$; 0.3% TWEEN 100; 0.2 mM dNTPs; primers M38.27 (0.5 μM) and M20.106 (1.1 μM)) in six cycles (temperature cycle conditions: 94 °C, 4 min; 42 °C, 7 min; 72 °C, 7 min). The PCR-DNA was extracted with phenol/$CHCl_3$, precipitated with ethanol, dissolved in 1.3 mL of TE (10 mM Tris-HCl, pH = 7.6; 1 mM EDTA, pH = 8.0), and purified by gel filtration on Sephadex G-50 (Pharmacia). For in vitro transcription, 400 μg of PCR-amplified DNA (180 μL) was used in a 10-mL transcription reaction (40 mM Tris-HCl, pH = 7.9; 6.8 mM spermidine; 22 mM $MgCl_2$; 0.01% Triton X-100; 5 mM NTPs; 10 mM DTT; 5000 U T7-RNA polymerase; 50 μCi [α-^{32}P]-GTP; 40 U RNAsin) and transcribed at 37 °C for 16 h.

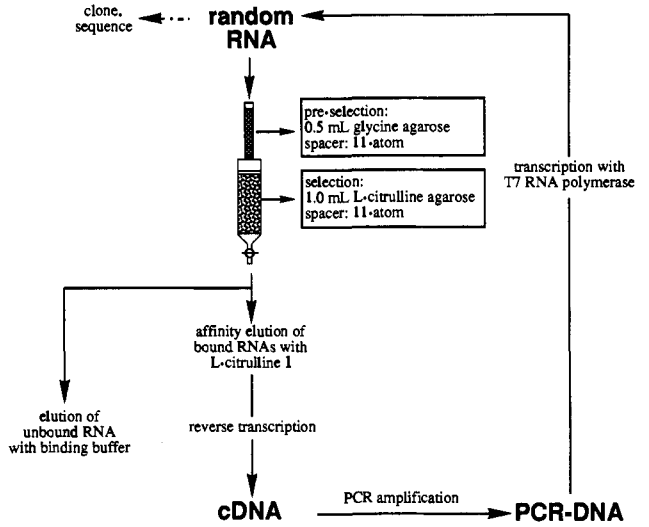
(14) Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Science* **1991**, *252*, 1167.

(15) (a) Weeks, K. M.; Ampe, C.; Schultz, S. C.; Steitz, T. A.; Crothers, D. M. *Ibid.* **1990**, *249*, 1281.

(16) Zawadzki, V.; Gross, H. J. *Nucleic Acids Res.* **1991**, *19*, 1948. I thank H. Gross for his advice and for details regarding the purification procedure.

(17) Sawadogo, M.; Van Dyke, M. W. *Nucleic Acids Res.* **1991**, *19*, 674.

Scheme 2. Schematic for the In Vitro Selection Cycle[a]



[a] RNA is loaded onto the preselection column and eluted for 2.5 column volumes onto the selection column. The preselection column is removed. Nonbinding RNAs are eluted off the selection cholumn for five column volumes. Bound RNA is eluted with a solution of the ligand in binding buffer and converted to cDNA. The cDNA is PCR-amplified. The PCR-DNA is transcribed to RNA, at which point the cycle can be repeated.

700 U DNase I (RNase-free) was added and incubation was continued for 1 h. After addition of 2 mL of 500 mM EDTA (pH = 8.0), the mix was extracted twice with phenol/$CHCl_3$, adjusted to 300 mM NaCl, precipitated by addition of 3.0 volumes of ethanol, washed with 70% ethanol, dried in vacuo, and redissolved in 2.0 mL of water. The RNA was purified by gel filtration on Sephadex G-50 (eluent: water) and used for the selections.

4. Selections. To minimize the enrichment of RNA molecules that require parts of the agarose in addition to the ligand, or that bind to the α-aminocarboxy-derivatized spacer, the pool was preselected on glycine-derivatized epoxyagarose. Preselection agarose (0.5 mL) and 1.5 mL of ligand-derivatized agarose (3.0–6.0 mM) were equilibrated with several column volumes of binding buffer (250 mM NaCl; 50 mM Tris-HCl, pH = 7.6; 5 mM $MgCl_2$). A 490-μL sample of the RNA-pool (400 μg) in binding buffer was loaded onto the preselection column. The preselection column was rinsed with 760 μL of binding buffer (2.5 column volumes), and the eluting RNA was dripped directly onto the selection column. This column was rinsed with 250 μL and then with four more column volumes (6.0 mL) of binding buffer. Elution of aptamers was done with a 6.0 mM solution of ligand in binding buffer for three column volumes. The eluted RNA was precipitated with ethanol in the presence of 100 μg of glycogen, washed with 70% ethanol, redissolved in 50 μL of water and reverse transcribed at 42 °C for 45 min (50 mM Tris-HCl, pH = 8.3; 75 mM KCl; 3.0 mM $MgCl_2$; 10 mM DTT; 0.3% TWEEN 20; 0.4 mM dNTPs; 20 U RNAsin; 200 U reverse transcriptase). The resulting cDNA was PCR-amplified, purified as described above, and transcribed. The purified RNA (25%, ca. 1.0 nmol) was used as the input for the next selection cycle.

5. K_d Determinations. Measurements of dissociation constants were performed by the methods of isocratic elution[18] and by equilibrium gel filtration (Figure 2).[19] Isocratic elutions were done in the absence and in the presence of ligand in solution using 1.0 mL of 15.0 mM (L-citrulline) or 33.0 mM (L-arginine) ligand-derivatized agarose columns (diameter = 5.0 mm). The column was equilibrated with binding buffer. The ^{32}P-labeled aptamer-RNA was applied to the column and eluted with binding buffer in 0.75-mL fractions until no radioactivity was detectable on the column. The elution volume was determined from the peak maximum of the eluted RNA. The experiment was repeated with the same column that was now equilibrated with binding buffer in the presence of 0.5–8.0 mM amino acid or amino acid analogues, respectively. The

(18) Arnold, F. H.; Schofield, S. A.; Blanch, H. W. *J. Chromatogr.* **1986**, *355*, 1.

(19) (a) Fersht, A. In *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman and Co.: New York, 1985; p 186. (b) Hummel, J. P.; Dreyer, W. J. *Biochem. Biophys. Acta* **1962**, *63*, 530. (c) Connors, K. A. *Binding Constants*; J. Wiley: New York, 1987; p 314.

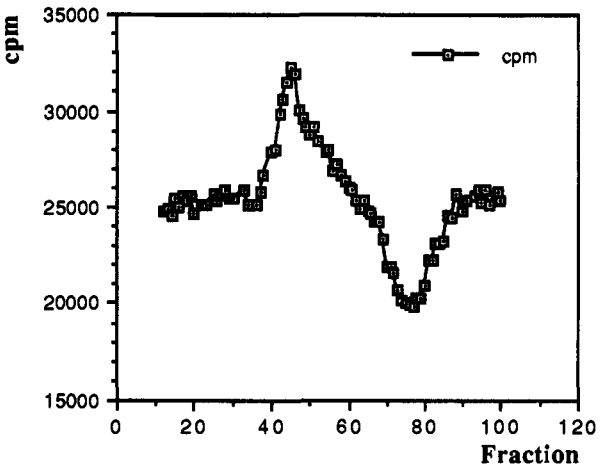


Figure 2. Dissociation constant measurements by equilibrium gel filtration. A 15-cm column (diameter 0.5 cm) was filled with 2.0 mL of Sephadex-G25 F and equilibrated with binding buffer containing 5.0 μM L-[2,3,4,5-^3H]citrulline. 44.Cit11 (5.0 nmol) was loaded and eluted. The radioactivity of individual fractions (26–28 μL) was quantified in a scintillation counter. The area under the peak was integrated to quantify the amount of bound **1** and the K_d.

K_d was calculated using the equation[10] $K_d = [L]\{(V_L - V_0) \times (V_e - V_L)^{-1}\}$, where [L] is the concentration of ligand in the buffer, V_0 is the void volume of the column; V_L is the elution volume of the RNA in the presence of ligand in solution, and V_e is the elution volume of RNA in the absence of ligand in solution (for amino acid analogues, the term K_d might more correctly be replaced by K_c, the competition constant).

For equilibrium gel filtration a 15-cm column (diameter = 0.5 cm) was filled with 2.0 mL of Sephadex G 25 F and equilibrated with binding buffer containing 5 μM radiolabeled amino acids. A 5-nmol sample of RNA was dissolved in the same ligand solution (50 μL) and loaded onto the column. Fractions of 26–28 μL were collected, and the radioactivity was measured by scintillation counting. The amount of bound amino acid was determined by integration of the area under the first peak and was used to calculate the K_d (Figure 2).

Results

The L-Citrulline Selection. From a pool of 10^{15} different 113-mer RNAs, molecules that bind to L-citrulline (**1**) were selected. The amino acid was coupled to an epoxy-activated agarose through its amino group with an 11-atom hydrophobic spacer arm between the agarose and the ligand. RNAs with no affinity to the column were washed away with buffer; molecules that bound to the column were eluted with a 6.0 mM solution of **1** in binding buffer. Eluted RNAs were reverse transcribed, amplified by the polymerase chain reaction (PCR) and by in vitro transcription, and the enriched RNA-pool was used for a new selection cycle. The percentage of bound RNA to the L-citrulline column in selection cycles 1–5 was between 0.04 and 2.5% (cycle 5); in cycle 6, 11% of the RNA bound to the column; in cycle 7, 50% binding was detected. This RNA did not bind to agaroses derivatized with L-albizziin, L-arginine, L-histidine, or glycine (15 mM). The RNA from cycle 7 was reverse transcribed, PCR-amplified, and transcribed. This RNA-pool was then tested for specificity to L-citrulline by affinity elution with different L-citrulline analogues. Solutions of 6.0 mM L-asparagine, L-arginine, and L-albizziin in binding buffer did not elute the RNA off a 6.0 mM citrulline–agarose column. Fifteen percent of the RNA was eluted by 6.0 mM urea in three column volumes, whereas a 1.0 mM L-citrulline solution eluted the RNA almost quantitatively within two column volumes.

Sequencing of 21 aptamers revealed 19 different sequences with two sequences (clones Cit-3 and Cit-16) occurring twice (Figure 3). All sequences contained two consensus sites, CG-GUUAGGUC and GGAGUG. Twelve sequences had the 10-mer consensus site 5′ of the 6-mer site, nine sequences had it 3′

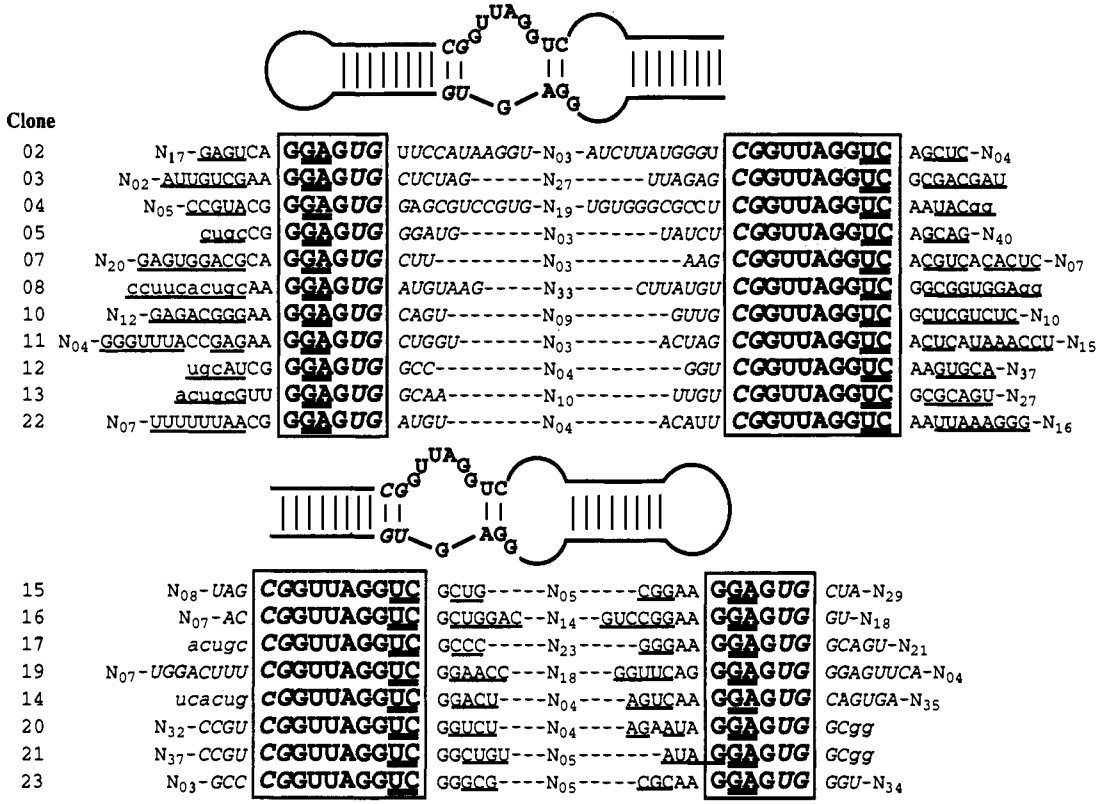
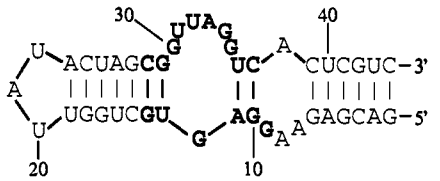


Figure 3. Sequences of L-citrulline-binding RNAs. RNA from cycle 7 was reverse transcribed, and the cDNA was amplified for eight PCR cycles with primers M20.106 and M38.27, diluted 1:1000, and PCR-amplified with primers M20.106 and M27.39 (for primer sequences see Figure 1). The resulting DNA was digested with EcoRI and BamHI, gel purified, and cloned into the vector pGem3Z (Promega.) Sequencing was done by the dideoxy method. The sequences are aligned relative to the two consensus sequences, shown in bold. The framed sequences are invariant among the different clones. Bases in small letters are part of the primers.



44.Cit11

Figure 4. Sequence and secondary structure of the 44-mer aptamer 44.Cit11. The 10- and 6- mer consensus sequences are shown in bold.

of the 6-mer, but all sequences had the potential to fold into the same secondary structure (Figure 3). Clones Cit-3, Cit-11, Cit-14, and Cit-16 all bound to the column and could be eluted with L-citrulline in the same way as the RNA from cycle 7. On the basis of the similarities of the individual sequences, a 44-mer RNA was constructed (44.Cit11, 5′-GAC GAG AAG GAG UGC UGG UUC UAC UAG CGG UUA GGU CAC UCG UC-3′). This RNA was used for further characterization of the aptamer. The secondary structure of 44.Cit11 is shown in Figure 4.

The bulge 5′-GUUAGG-3′ lies between two flanking stems and is 100% conserved in all clones. The short stem at the 3′-end of the bulge (double underlined in Figure 3) consists of only an A^{11}·U^{36} and a G^{10}·C^{37} base pair. This small stem is also 100% conserved and is required for binding: abolishment of one base pair results in complete loss of binding. At the 5′-end of the bulge, the stem (shown in italic in Figure 3) is flanked by two guanines, G^{12} and G^{30}. These Gs might have the potential to base-pair in a noncanonical way. Such noncanonical base pairs have been shown to exist in several functional RNA molecules and to be essential for functionality.[20] Changing the G^{30} into a C, so that a Watson–Crick base pair can be formed with G^{12}, resulted in a complete loss of binding. It should be noted, however, that the guanines at positions 34 and 35 could also pair with G^{12}. Chemical probing reactions[21] to map the binding site of the amino acid and to particularly determine the nature of the Gs at positions 12, 30, 34, and 35 are currently being performed in this laboratory. The stem ends with a base pair formed by a U at position 13 and a G at position 29. These bases are 100% conserved. When U^{13} was changed into a C, resulting in a change of the U^{13}·G^{29} into a C^{13}·G^{29} base pair, the binding was 5–10-fold weaker. Base pair G^{14}·C^{28} is also 100% conserved. The rest of this stem is variable in sequence and in length. Comparison of all clones revealed significant variation at the unpaired loop and the stem composed of bases 1–9 and 38–44 (underlined in Figure 3). Nine clones have the sequence AAG in addition to one or two purines in the unpaired region. Other clones have completely different sequences in the three-base unpaired region. There are also clones that contain only two (clone Cit-20) or no (clone Cit-21) unpaired bases in this region. The underlined stem varies highly both in length and in sequence.

Analytical affinity chromatography revealed a K_d of 68 μM for 44.Cit11; by equilibrium gel filtration a K_d of 62 μM was measured. To determine the influence of the magnesium concentration on the binding constant, the equilibrium gel filtration was performed at 20 mM $MgCl_2$ instead of 5.0 mM $MgCl_2$. At 20 mM $MgCl_2$, the K_d increased to 113 μM, which corresponds to about 2-fold weaker binding than at 5.0 mM $MgCl_2$. By analytical affinity chromatography, the ability of D-citrulline to elute 44.Cit11 off the L-citrulline column was tested using a 1.0 mM solution of D-citrulline in binding buffer (5.0 mM $MgCl_2$). With D-citrulline a K_d of 180 μM was found. This

(20) (a) Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W. *Cell* **1991**, *67*, 529. (b) Hou, Y.-M.; Westhof, E.; Giegé, R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6776. (c) Giver, L.; Bartel, D. P.; Zapp, M. L.; Pawul, A.; Green, M. R.; Ellington, A. D. *Nucleic Acids Res.* **1993**, *21*, 5509.

(21) Stern, S.; Moazed, D.; Noller, H. F. *Methods Enzymol.* **1988**, *164*, 481.

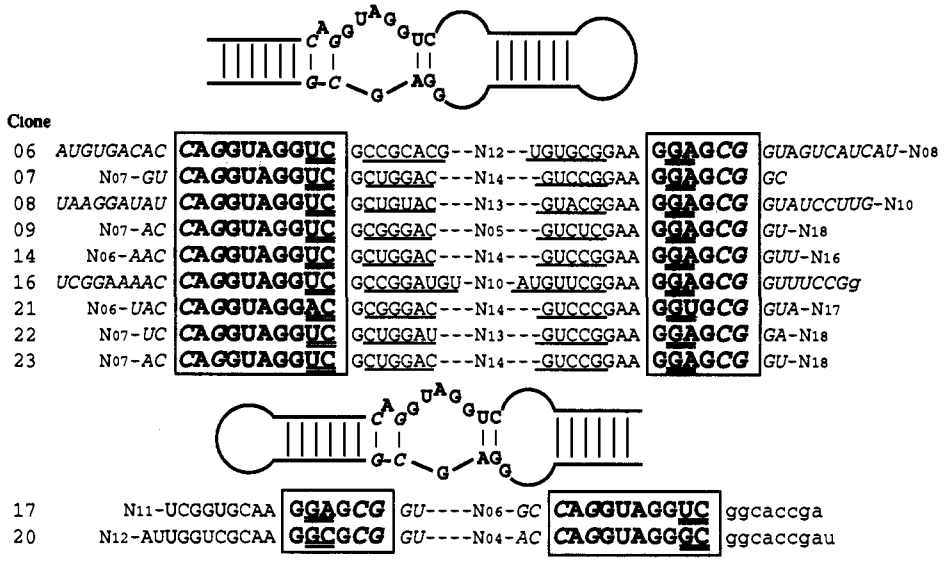
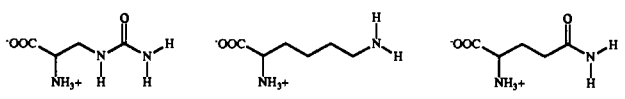


Figure 5. Sequences of L-arginine-binding RNAs. RNA from cycle 5 of the reselection of the doped Cit-16-pool was reverse transcribed and PCR-amplified with the primers M38.27 and Cit24.1(5′-TTAAGCTTAGCTGCGATCGGTGCC-3′). The PCR-DNA was digested with SacI and HindIII (underlined site in primer Cit24.1), gel purified, and cloned into pGem3z. The alignment of the sequences is the same as in Figure 3.

difference in the binding affinities corresponds to a $\Delta\Delta G$ of 2.5 kJ mol^{-1} (0.6 kcal mol^{-1}) and shows that citrulline recognition by 44.Cit11 is weakly enantioselective. Both this RNA and the "full-length" sequences (clones Cit-11 and Cit-16) had a more than 10-fold weaker affinity to the ligand bound to the agarose as compared to binding in solution. The K_d for binding to L-citrulline–agarose was between 0.5 and 0.8 mM. However, this K_d should be considered an upper estimate since the fraction of bound amino acid that is accessible to the RNA is unknown.

Binding of 44.Cit11 to **1** is specific. As with the isolated full-length sequences, urea has some potential to elute this RNA off the column, but albizziin (**3**) has not, although it contains the urea portion. This is consistent with the proposal that steric hinderance by the amino acid moiety prevents **3** from occupying the urea-binding site in the aptamer.

Reselection of a Mutated Pool Derived from a Citrulline-Binding Sequence. The DNA of one of the citrulline-binding sequences, clone Cit-16, was chemically resynthesized, such that each position was mutagenized at a level of 30%.[22] This was achieved by doping each phosphoramidite solution with 10% of each of the other bases. The 3′-primer was changed in order to avoid the unintended amplification of possible "wild-type" contaminants. As before, this oligonucleotide synthesis gave a complexity of roughly 10^{15} different sequences. The DNA was PCR-amplified and converted into RNA by in vitro transcription. This RNA was then used for further cycles of selection. To test if the doped pool contained mutants which had become able to specifically bind to amino acids with some structural similarity to **1**, selections on agaroses derivatized with different "citrulline analogues" were performed. These selections included L-arginine (**2**), L-albizziin (**3**), L-lysine (**4**), and L-glutamine (**5**). As a control, it was also selected for



L-citrulline binding. Furthermore, the selection of **3–5** included a negative selection on glycine agarose to remove RNAs that bind to the agarose or to the α-aminocarboxy-derivatized spacer.

(22) The phylogeny of the selected L-citrulline aptamers shows that only 16 bases are invariant whereas the flanking stems vary both in length and in sequence. Thus, mutations in this invariant region should be most crucial for the determination of binding specificity. The applied degeneracy of 30% per position means that the molecules resulting from the synthesis of the degenerate pool contain an average of five mutations in this region.

Table 1. Reselection of 30% Overall Mutated Cit-16 RNA for Binding to Different Amino Acid Agaroses[a]

cycle	L-Cit (1)	L-Arg (2)	L-Alb (3)	L-Lys (4)	L-Gln (5)
1	0.01	0.01	<0.01	0.01	<0.01
2	0.12	0.10	0.07	0.09	0.11
3	12.1	0.21	0.05	0.33	0.07
4	41.7	33.7	0.09	0.13	0.11
5			0.8	0.30	0.09
6			1.0	1.1	0.6
7			0.7	1.2	0.4

[a] The values show the relative percentage of input RNA that bound to the columns and eluted with the ligand in solution after washing away the nonbinding RNAs with five column volumes of binding buffer. For the reselection, a new 3′-primer binding site was designed to avoid the amplification of possible Cit-16 "wild-type" contaminants.

To avoid the selection of RNAs with affinities to both **1** and **2**, aptamers to **2** were subjected to a negative selection on L-citrulline–agarose. The results of the selection are shown in Table 1. In the third cycle, the RNA-pool bound to L-citrulline (**1**), whereas four cycles were required to enrich the pool for L-arginine (**2**) binding. No aptamers to **3–5** could be isolated in seven selection cycles.

The L-Arginine-Aptamers. The RNA selected for arginine binding from cycle 4 was reverse transcribed. The resulting cDNA was amplified by PCR and in vitro transcribed into RNA. This RNA did not bind to agaroses derivatized with **3–5** or **1**. A 6.0 mM solution of **2** in binding buffer eluted the aptamers quantitatively off a 15.0 mM L-arginine column within two column volumes, with the majority of the binders eluting in the first column volume of affinity elution. Solutions of 20 mM **3–5** or **1** were unable to elute the arginine aptamers off the 15.0 mM column.

Cloning and sequencing of 22 aptamers revealed 22 different sequences. Eleven clones had the potential to fold into a common secondary structure (Figure 5). Nine of these clones contained the 10-mer consensus sequence CAGGUAGGUC and the 6-mer consensus sequence GGAGCG. The 10-mer consensus is a double mutant of the related L-citrulline 10-mer sequence, whereas the 6-mer differs in one position from the related L-citrulline consensus. Two clones, Arg-20 and Arg-21, showed covariations in the two-base stem, double underlined in Figure 5. Of the 22 sequences, 11 did not have the potential to fold into the most abundant secondary structure. These sequences did not show any obvious conserved sequence similarities to the "parent" Cit-16 motif. Two of these clones (Arg-2 and Arg-5), both 58 bases in length, showed 60% sequence identity to each other and shared

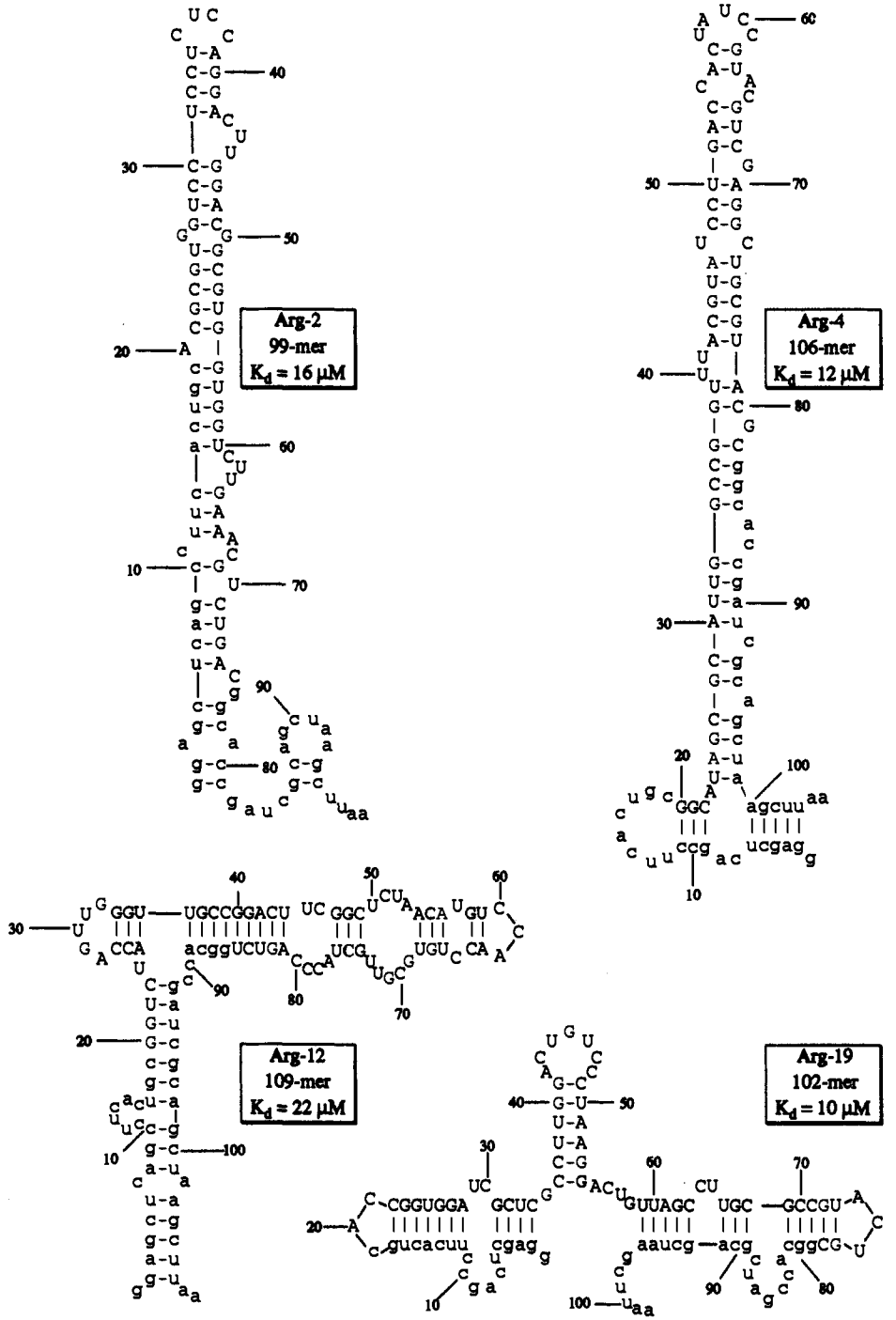


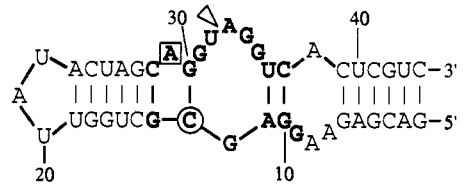
Figure 6. Computer generated foldings of the sequences Arg-2, Arg-4, Arg-12, and Arg-19. The binding constants are solution K_ds and were determined by isocractic elution. Bases in small letters represent primer binding sites.

a 13-mer, a 9-mer, a 7-mer, and a 6-mer consensus sequence. Arg-2, Arg-4, Arg-12, and Arg-19 were individually tested for binding and had binding constants between 10 and 22 μM (tested by isocratic elution). These sequences could not be eluted with L-citrulline (**1**) or guanidine (**7**) and were very specific for L-arginine (**2**) binding. Computer analysis[23] of the sequences tested for binding revealed folding patterns which were both highly structured and low energy. The folding patterns are shown in Figure 6.

A 44-mer RNA was synthesized (44.Arg11, GAC GAG AAG GAG CGC UGG UUC UAC UAG CAG GUA GGU CAC UCG UC) that differs from 44.Cit11 only in the three positions within the 10-mer and the 6-mer consensus sequences indicated in the secondary structure in Figure 7 (circle = substitution; square = insertion; triangle = deletion).

44.Arg11 was used for further characterization of the arginine-binding aptamer. When compared to the secondary structure of the related 44.Cit11 RNA, the three mutations result in the following differences: The bulge 5′-$G^{31}U^{32}AGG^{35}$-3′ lacks one of the two 5′-Us found in the same region in 44.Cit11 (indicated by the triangle). The arginine aptamer contains a C^{13}·G^{30} base pair instead of the U^{13}·G^{29} base pair found in the citrulline aptamer (indicated by the circle). Furthermore, a bulged A appears in position A^{29} in 44.Arg11 (indicated by the square) while no bulged base is found in 44.Cit11. When this bulged A^{29} was deleted, L-arginine binding was completely abolished. Changing the potential G^{12}·G^{31} base pair into an isosteric A^{12}·A^{31} base pair resulted in a complete loss of binding. To test if the highly variable

(23) Zucker, M.; Stiegler, P. *Nucleic Acids Res.* **1981**, *9*, 133.



44.Arg11

Figure 7. Sequence and secondary structure of the 44-mer aptamer 44.Arg11. The 10- and 6-mer consensus sequences are shown in bold. The box, circle, and triangular represent changes in the aptamer relative to 44Cit.11 (boxed A at position 29, insertion; circled C at position 13, substitution, triangle between positions 32 and 33, deletion of one U).

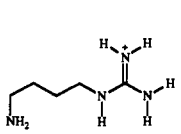
Table 2. Summary of the Binding Behavior of 44.Cit11 and 44.Arg11 Mutants, Including Selected Clones with Base Positions That Vary among Different Clones[a]

citrulline binders	arginine binders	base(s) → change	binding
44.Cit11		$G^{30} \rightarrow C^{30}$	(–)
		$U^{13} \rightarrow C^{13}$	(±)
10 clones		$A^{38} \rightarrow G^{38}$	(+)
clones 4, 12, 22		$A^{38} \rightarrow AA$	(+)
clones 21, 23		$A^{38} \rightarrow GG$	(+)
clone 21		$A^7A^8G^9$ deleted	(+)[b]
clone 20		A^7A^8 deleted	(+)
	44.Arg11	A^{29} deleted	(–)
		$G^{12}.G^{31} \rightarrow A^{12}A^{31}$	(–)
		G^1-G^9 and A^{38}-$C^{44} \rightarrow$ UUCG	(–)
	clone 21	$A^{11}.U^{36} \rightarrow U^{11}.A^{36}$	(+)
	clone 20	$A^{11}.U^{36} \rightarrow C^{11}.G^{36}$	(+)

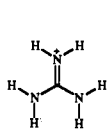
[a] Base numbers refer to the numbering in Figures 4 and 7. [b] Other clones, such as 2, 4, 5, 7, 12, 13, 19, and 22, show variations in these three bases.

stem–loop region 5′-G^1ACGAGAAG-3′ and 5′-A^{38}CUCGUC (underlined in Figure 5) that flanks the two-base-pair stem (double underlined in Figure 5) is required for binding, this region was substituted by a closing loop 5′-UCCG,[24] leaving away the closing loop U^{20}AU. This substitution resulted in a significant loss of binding activity. Table 2 summarizes the binding behavior of mutants of 44.Cit11 and 44.Arg11, respectively. Base positions that vary among different clones are also included. For greater clarity, the numbering of the bases refers to the numbering in 44.Cit11 and 44.Arg11, respectively.

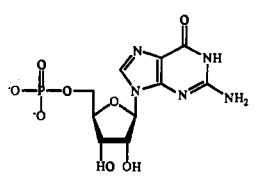
The K_d of 44.Arg11 was determined by analytical affinity chromatography using a 33.0 mM L-arginine–agarose (56 μM) and by equilibrium gel filtration (76 μM). The binding buffer contained 5.0 mM $MgCl_2$. The K_ds of a series of arginine analogues were measured by isocratic elution and were calculated using the elution volume of the aptamer in the presence of the arginine analogue in solution (V_L) and the elution volume of 44.Arg11 in the absence of affinity ligand (V_e) (Table 3.)



Agmatine 6



Guanidine 7



GMP 8

L-arginine is bound with a K_d of 56 μM, D-arginine with 412 μM. The differences in the binding affinities between D- and L-arginine correspond to a $\Delta\Delta G$ of 4.6 kJ mol^{-1} (1.1 kcal mol^{-1}).

(24) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D. R.; Gayle, M.; Guild, N.; Stormo, G.; d'Aubenton-Carafa, Y.; Uhlenbeck, O. C.; Tinoco, I.; Brody, E. N.; Gold, L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1364.

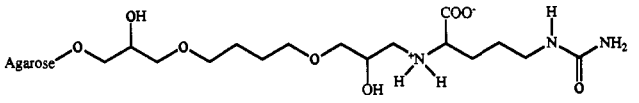
Table 3. Dissociation Constants for Binding of RNA 44.Arg11 to L-Arginine (**2**) and the Analogues D-Arginine, Agmatine (**6**), Guanidine (**7**), Guanosine (**8**), and L-Citrulline (**1**)[a]

ligand	concentration of affinity eluent [L] (mM)	elution volume (mL)	K_d (mM)
L-arginine (**2**)	0.0	49.5	–
L-arginine (**2**)	0.5	5.6	0.06
D-arginine	2.0	9.1	0.41
agmatine (**6**)	0.5	9.3	0.11
guanidine (**7**)	4.0	29.3	5.65
GMP (**8**)	8.0	40.7	36.3
L-citrulline (**1**)	5.0	29.5	7.19

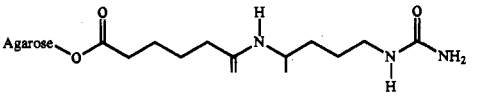
[a] The K_ds refer to binding in solution. They were determined by affinity elution of radiolabeled 44.Arg11 RNA off a column that contained 1.1 mL of a highly derivatized L-arginine–agarose (33.0 mM). Different concentrations of the affinity eluent in binding buffer were used. The void volume of the column was 0.75 mL. K_ds were calculated using the equation $K_d = [L]\{(V_L - V_0)(V_e - V_L)^{-1}\}$. [L] is the concentration of the ligand in the elution buffer.

Discussion

The 10- and 6-mer consensus sequences found in the L-citrulline-aptamers show different orientations relative to each other: there is roughly a 1:1 ratio of individual sequences in which the 10-mer consensus sequence is located 5′- of the 6-mer, and of sequences in which the 10-mer appears 3′- of the 6-mer. Since in both cases the consensus sequences are flanked by bases that can pair with each other, this result particularly supports the proposed folding of the aptamers. A sequence of 44 bases in length (44.Cit11) is sufficient for specific binding. The isolated sequences probably represent the smallest possible motif for L-citrulline binding with an affinity that corresponds to the applied selection stringency. This is supported by the observed "sequence homogeneity": all isolates contain the 10- and 6-mer consensus sequences but are otherwise entirely different. Shorter sequences are more abundant in the initial pool than sequences that fulfill the same or better binding criteria, but require longer consensus sequences. Another critical feature that led to the selection of the motif described here is the way in which the amino acid is coupled to the matrix. This can be concluded from an unsuccessful attempt to select L-citrulline-binding aptamers with an agarose in which **1** was attached through an amide bond. No aptamers could be enriched in seven selection cycles with this agarose. The main difference between the two agaroses is that the amino acid has the potential to remain in its zwitterionic state when coupled to epoxy-activated agarose. This is not possible with the *N*-hydroxysuccinimide-activated agarose. Sassanfar and Szostak have recently isolated



1 coupled to epoxy-activated agarose



1 coupled to N-hydroxy succinimide-activated agarose

a 40-mer RNA that binds to ATP.[11] All of their selected aptamers contained an 11-mer consensus sequence embedded in a common secondary structure. The observed "sequence homogeneity" was explained in the same way as described above.

One of the main purposes of this study was to test if the binding motif for L-citrulline recognition could be evolved to bind related amino acids. The doped Cit-16-pool contained only variants which

could bind to L-arginine (**2**). The failure to select aptamers to L-albizziin (**3**) from this pool suggests that despite their structural similarity the molecular recognition of **1** and **3** by RNA requires completely different and unrelated RNA binding motifs. There might be many cases where closely related RNA structures do not exist for closely related binding specificities so that it may often be difficult to evolve from one function to a related function, making it necessary to start anew in such cases. This leads to the question of whether this is also true for proteins, i.e. whether they are more flexible in an evolutionary sense than RNAs. We are currently repeating the selection of aptamers for **3–5** and other amino acids using the same pool that was initially used for the L-citrulline selection.

It is arguable whether this approach deserves to be termed "evolution". The selection described here is certainly not *Darwinian* evolution since mutations are introduced exclusively at the level of the initial synthesis of the doped pool. This pool might have contained a few sequence variants with a specific affinity for **3**, **4**, or **5** that was too low to allow these RNAs to survive the applied selection stringency during several selection cycles. Such moderate binders would have had a chance to improve their affinity gradually during each selection cycle in a *Darwinian* in vitro evolution system similar to the one recently reported by Joyce and co-workers.[25] In this system, mutations are introduced at the PCR-amplification step at each selection cycle. The reselection of the doped Cit-16-pool for binding to **3–5**, in combination with Darwinian evolution in vitro, may lead to tight binding sequences. We are currently modifying the PCR protocol described by Joyce et al. (which has been developed and optimized for the in vitro evolution of RNAs about 400 bases in length)[26] to allow the amplification of 40–80-mers with a higher mutation rate.

The common arginine-binding motif which resulted from the selection of the doped pool is a triple mutant of the parent L-citrulline structure. However, two clones, Arg-20 and Arg-21, are 5-error mutants with respect to the parent aptamer. The two additional mutations in these clones reflect a covariation in the two-base-pair stem within the consensus sequences. The comparison of the selected sequences shows that mutations in the amino acid binding site occurred in molecules with many other unrelated changes in the irrelevant parts of the sequence. Studies on group I intron RNA have demonstrated that the inhibition specificity of the self-splicing reaction can be changed from **2** to **1** when the G^{264}·C^{311} base pair is changed to an A·U base pair.[13] These amino acids were proposed to interact with G^{264} and A^{264}, respectively, as shown in Scheme 1. The two H-donor/H-acceptor sites, indicated in the aforementioned scheme, do not contact 44.Arg11 and 44.Cit11 in the same way as the group I intron does. Otherwise, the reselection of the doped Cit-16-pool for arginine binding ought to have resulted in the covariation of one base pair for *all* of the selected sequences and not just for Arg-20 and Arg-21.

Binding of L-arginine (**2**) requires the presence of the two-base-pair stem, shown double underlined in Figure 5. The covariations in Arg-20 and Arg-21 allow the conclusion that the sequence of this stem does not affect binding. Of the 11 sequences that can fold into the proposed common secondary structure, 9 share the same sequence in the stem as shown in Figure 5. This is not surprising because the parent Cit-16 has the same sequence in this stem. In Arg-20, the base pair A^{11}·U^{36} is changed into a C^{11}·C^{36}, and in Arg-21, it is changed to a U^{11}·A^{36} pair. This covariation in particular indicates that the U in the highly conserved bulge 5′-GUAGG is unlikely to form a base triple with the A·U base pair in the stem. Puglisi et al.[1c] have demonstrated by NMR spectroscopy that the formation of a base triple is critical for the binding of L-arginine (**2**) to an RNA stem–loop structure of transactivation response RNA (TAR). In 44.Arg11, all other bulged bases are purines and the only conserved A·U base pair is the one in the double-underlined two-base stem. Therefore, the tertiary structure of 44.Arg11 most likely does not contain any base triples of the U–A·U or C^+–G·C-type.

To determine which functional groups are involved in binding to the aptamer, K_ds of individual arginine analogues were measured. The affinity of 44.Arg11 for guanidine (**7**) is almost 100-fold weaker than for **2**. In contrast, agmatine (**6**) binds almost as well as **2** and about 4-fold better than D-arginine. From these results the following conclusions can be drawn: the carboxylic acid portion of the amino acid does not significantly contribute to binding energy, suggesting that it does not interact electrostatically with the aptamer. If any hydrogen bonds to the RNA are formed at all, they must be very weak. The guanidino group in **2** contacts the RNA through hydrogen bonds rather than electrostatic interactions and thus determines to a large extent the specificity of the aptamer for **2**. This can be concluded from the fact that **7** alone has a relatively weak affinity to the RNA, but in combination with the δ-amino group in **2** or **6**, binding is tight and specific. Replacement of the guanidino group in **2** for a urea group results in a more than 100-fold weaker binding. The presence of the amino group in **6** and **2** results in an increase in the binding energy of at least 11.3 kJ mol^{-1} (2.7 kcal mol^{-1}) with respect to that of **7**. Such a dramatic increase in binding energy must be due to electrostatic interactions, probably between the amino group and the phosphate backbone of the aptamer. This proposal is further supported by the finding that at 20 mM $MgCl_2$ the 44.Cit11 binds less strongly to L-citrulline than at 5.0 mM $MgCl_2$. Electrostatic interactions between ammonium ions and the phosphate backbone of nucleic acids are weakened at increasing Mg^{2+} concentrations due to charge neutralization by Mg^{2+}. The same effect has been observed with the antibiotic viomycin, which inhibits the self-splicing of group I intron RNA and probably interacts electrostatically with the RNA through guanidinium and/or ammonium groups; at 3–5 mM $MgCl_2$, viomycin inhibited splicing, whereas at 10 mM $MgCl_2$, the inhibition was dramatically decreased.[27] The opposite effect was found with the ATP-binding aptamer described by Sassanfar and Szostak, which binds with a K_d of 5 μM at 5.0 mM Mg^{2+} and with a K_d of 2 μM at 20 mM Mg^{2+}.[11] This effect was explained by the increased neutralization of the negative charges at the triphosphate group of ATP at higher Mg^{2+} concentrations and hence decreased repulsion by the phosphate backbone of the aptamer.

The relatively good enantioselectivity is consistent with a well-defined binding pocket in which the L-amino acid forms close contacts to the RNA. The carboxylic acid of the D-enantiomer probably disrupts some of these contacts due to unfavorable steric interactions, resulting in an almost 7-fold decrease in the binding constant. This enantioselectivity in binding of L- and D-arginine is in the same order of magnitude as the enantioselectivity shown by group I ribozymes. In these RNAs, L-arginine is bound roughly 2-fold better than D-arginine,[1c] but examples are reported in which the preference for L- over D-arginine is 10-fold.[28]

Connell et al.[10] recently obtained three short RNA motifs by in vitro selection which were reported to bind L-arginine (**2**) in solution with affinities around 1.0 mM. One of these sequences had a slightly higher affinity for D- over L-arginine [K_d(D-arginine) ≈ 0.6 mM; K_d(L-arginine ≈ 1.0 mM]. The affinities of the selected sequences for guanidinium (**7**), methyl guanidinium, or agmatine (**6**) were about the same as for **2**, as shown by affinity elution with the respective ligands. As the self-splicing group I intron, these motifs bind guanosine monophosphate (GMP), but they do not show any similarity to the G-binding site in group I intron RNA.

(25) (a) Beaudry, A. A.; Joyce, G. F. *Science* **1992**, 257, 635. (b) Lehman, N.; Joyce, G. F. *Nature* **1993**, *361*, 182.
(26) (a) Cadwell, R. C.; Joyce G. F. *PCR Methods Appl.* **1992**, *2*, 28. (b) See also footnote 12 in ref 12.
(27) Wank, H.; Von Ahsen, U.; Davis, J.; Schroeder, R. *J. Mol. Biol.*, in press.
(28) Yarus, M.; Majerfeld, I. *J. Mol. Biol.* **1992**, *225*, 945.

44.Arg11 has no affinity to GMP and does not show any sequence or structural similarities to the motifs described by Connell et al.

Why did the selection experiment described here not lead to the isolation of previously discovered naturally occuring citrulline- or arginine-binding RNAs, such as group I intron RNA or transactivation-response RNA (TAR)? The most simple explanation would be that I selected for much higher binding affinities than the affinities observed in the natural RNAs. The K_ds for arginine in group I introns or TAR are in the millimolar range. Furthermore, the aptamers described here presumably require the free α-amino group of the amino acid for Coulomb interactions to the phosphate backbone. TAR recognizes arginine within the context of the TAT protein in which the α-amino group is substituted for an amide group.

Conclusion

I have shown that from a pool of 10^{15} different RNA sequences, randomized in 74 positions, individual RNA molecules ("aptamers") can be selected that bind specifically to the amino acid L-citrulline (**1**) in solution. The selection led to the enrichment of one common structural motif that consists of two consensus sequences, 10 and 6 bases in length. Furthermore, this study has demonstrated that, from a pool derived from a highly mutated L-citrulline-binding sequence, a new motif with binding specificity for a related amino acid, L-arginine, can be evolved. This new motif is a triple mutant of the parent motif which has entirely lost its affinity for L-citrulline.

By the in vitro selection method presented here, RNAs with high affinity to structurally similar amino acids can be obtained whose sequences are different but related. Minor alterations in the selected sequence motifs can result in very large changes in substrate specificities. This approach should help to increase our knowledge about the specificity of amino acid side chain recognition by RNAs and hence about the origin of the genetic code.[29] It seems quite plausible that, in the early stages of the evolution of the genetic code, RNA molecules were involved that could specifically discriminate between similar amino acids.[30] Such RNAs should have evolved to bind to structurally related amino acids in cumulative selection processes rather than to have originated independently. The in vitro selection method used here provides a model for the study of these and related problems.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft and by the European Community (Grant No. Biot-CT93-0345). I would like to express by gratitude to E.-L. Winnacker for his encouragement, to R. Schroeder, G. F. Joyce, F. Michel, and P. Burgstaller for fruitful discussions, and to E. Westhof, H. F. Noller, J. W. Szostak, and M. G. Wallis for critical reading of the manuscript.

(29) Szathmáry, E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9916 and literature cited therein.

(30) Yarus, M. In *The RNA World*; Gesteland, R. F.; Atkins, J. F., Eds.; CSHI Press: Cold Spring Harbor, NY, 1993; p 205.