Nucleic Acids Research, 1996, Vol. 24, No. 6 **1029–1036**

RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity

Albert Geiger, Petra Burgstaller, Herbert von der Eltz[1], Albert Roeder[1] and Michael Famulok*

Institut für Biochemie, Genzentrum der Ludwig-Maximilians-Universität München, Würmtalstraße 221, 81375 München, Germany and [1]Boehringer Mannheim GmbH, Research and Development, Werk Penzberg, Nonnenwald 2, 82372 Penzberg, Germany

Received December 18, 1995; Revised and Accepted February 6, 1996

ABSTRACT

A completely randomized RNA pool as well as a degenerate pool comprised of an RNA sequence which binds citrulline with a dissociation constant of 60 μM were used to select for tight binding arginine specific RNA aptamers. A modified *in vitro* selection scheme, based on affinity chromatography was applied to allow the enrichment of high affinity solution binders. The selection scheme included a negative selection with the non-cognate ligand citrulline, and a heat denaturation step prior to affinity elution with an excess of the cognate ligand arginine. After 20 cycles the majority of the pools bound specifically to the arginine matrix even after denaturation/renaturation in the presence of 20 mM of a non-cognate amino acid. When denatured and eluted in the presence of 20 mM arginine, the selected RNAs quantitatively washed off the column. These RNA aptamers were cloned and sequenced. Equilibrium dialysis performed with the most abundant clone among the selected sequences revealed K_d values of 330 nM for the RNA/arginine affinity, which is nearly a 200-fold improvement over the tightest binding arginine binding RNAs known to date. Arginine recognition by this RNA is highly enantioselectice: L-arginine is bound 12 000-fold better than D-arginine. Chemical modification analysis revealed that the secondary structure of the aptamer might contain a pseudoknot motif. Our tight binding arginine aptamers join a number of natural and *in vitro* selected RNAs which recognize arginine. The RNAs described here compare in their binding affinity with the tightest binding RNA aptamers for low molecular weight molecules isolated in other *in vitro* selection experiments.

INTRODUCTION

Interactions between amino acids and RNA play substantial roles in a number of biological systems (1). For example, arginine inhibits the self-splicing reaction of the group I intron of *Tetrahymena* by substituting for two H-donor sites of the guanosine cofactor which contact the G^{264}–C^{311} base pair in the ribozyme's guanosine binding site (2,3). Recently, the editing reactions of aminoacyl tRNA synthetases have been viewed as an example of RNA dependent amino acid recognition (4). These editing reactions involve RNA dependent steps which eliminate errors of amino acid activation and aminoacylation (5). A third example is the interaction of the HIV-1 TAT protein with a stem–loop structure of 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 (6). Short oligopeptides resembling the basic region as well as free arginine bind specifically to TAR, although weaker than within the context of the whole protein (7). The TAT–TAR interaction provided the first example to show that in protein–RNA recognition RNA structures are involved which interact with individual amino acid side chains in the protein. It seems likely that other, yet undiscovered RNA–protein interactions exist in which single amino acid side chains within a protein or peptide form specific contacts to structural elements provided by RNA to largely determine specificity, functionality and strength of binding (1).

The isolation and characterization of RNA sequences which specifically recognize individual amino acids might facilitate a better understanding of biologically relevant protein–RNA or RNA–amino acid interactions. A powerful tool to obtain amino acid binding RNAs is *in vitro* selection (8–11). RNA aptamers which specifically recognize amino acids, such as immobilized tryptophan (12), arginine (13,14), citrulline (14) and valine (15) have been extracted from pools of up to 10^{15} different RNA sequences. The reported affinities ranged from 60 μM (14) to 12 mM (15) with a high level of discrimination against other amino acids being obtained in each case. Among these amino acid binding RNAs, the arginine specific aptamers might be especially relevant to protein–RNA recognition because arginine side chains carrying a positive charge at neutral pH seem to be particularly suited to form specific contacts with a negatively charged nucleic acid (16). For example, the HIV-1 Rev protein contains a basic region in which 10 out of 17 amino acids between positions 34 and 50 are arginines. A corresponding 17mer peptide binds to the Rev responsive element (RRE) RNA IIB hairpin in the same way as

* To whom correspondence should be addressed

within the context of the full-length protein (17) with four arginines being important for specificity (18). In the BIV-TAR/Tat complex an arginine and an isoleucine residue were found to be critical for binding and specificity (19). Furthermore, the Rex-protein of HTLV-I might interact with its natural RNA binding element XBE through arginine residues (20).

To explore the range of affinities which can be achieved in RNA–arginine recognition we set out to obtain RNA sequence motifs which tightly bind to arginine. We started with a completely randomized pool and a partially randomized pool of RNAs comprised of a citrulline binding sequence which was previously isolated by *in vitro* selection (14). Previously, arginine binding sequences which differed from the citrulline motif in three base positions were identified from the latter pool (14). In the present study, a modified *in vitro* selection scheme based on affinity chromatography was applied to extract tight binding arginine specific RNA aptamers from the two pools. A number of sequences were identified which bound tightly to arginine and did not share significant sequence homologies to each other or any of the other previously identified arginine aptamers (13,14). One of these sequences was shown to bind L-arginine with high enantioselectivity and with affinities in the order of magnitude of the tightest binding RNA aptamers for low molecular weight molecules obtained so far (21–24). Our study shows that the problem of arginine recognition can be achieved by many different RNA sequences over a broad range of binding affinities.

MATERIALS AND METHODS

Materials

L-Arginine, D-arginine, L-citrulline, agmatine and 1-cyclohexyl-3-(2-(morpholino-ethyl)-carbodiimide metho-p-toluenesulphonate (CMCT) were purchased from Sigma. L-Homoarginine and dimethyl sulfate (DMS) were from Aldrich, kethoxal (KE) from ICN. L-[2,3,4,5-^{3}H]arginine–HCl, [α-^{32}P]ATP and [γ-^{32}P]ATP (3 Ci/μmol) from Amersham. dNTPs, NTPs, glycogen and all enzymes and cloning plasmids used were from Boehringer Mannheim unless otherwise noted. Sequencing was performed with a digoxigenin labelled primer (DIG *Taq* DNA sequencing; Boehringer Mannheim). PCR primers were synthesized on an Millipore Expedite oligonucleotide synthesizer using standard phosphoramidite chemistry, and purified by HPLC, unless otherwise noted. Epoxy activated agarose was purchased from Pharmacia.

Agaroses

Arginine–agaroses were obtained by derivatization of epoxy activated Sepharose 6B according to the manufacturer's protocol. The agarose used in the selection was prepared with 3.0 μM arginine per millilitre of swollen agarose (3.0 mM), the agarose used in the binding assays were derivatized to 4.0 mM arginine. The coupling efficiency was determined by addition of L-[2,3,4,5-^{3}H]arginine–HCl to a coupling reaction containing 100 μl of activated agarose, followed by quantification of immobilized radioactivity by Cerenkov counting. Preselection agarose was derivatized with glycine in the same way.

RNA pool

The synthesis of pool 1 is described previously (14). For pool 2 we synthesized a 111mer DNA with a 30% doped insert of 74 nucleotides between a 18mer and a 19mer primer binding site on an Applied Biosystems DNA synthesizer. The sequences of the primer binding sites and the primers are shown in Figure 1A.

Pool synthesis was performed as described previously (14), with the modification that during synthesis of the partially randomized insert each phosphoramidite solution contained 10% of the other three amidites. Pool complexity was estimated from the percentage of the total synthesized, gel purified DNA (182 μg; 5 nmol) which was PCR-amplifiable to the full-length 111mer as described in the legend of Figure 1B. The whole 182 μg synthetic DNA were PCR-amplified in six cycles, purified and *in vitro* transcribed as described previously (14).

Selections

Ligand-derivatized agarose (0.5 ml) was equilibrated with several column volumes of binding buffer (250 mM NaCl, 50 mM Tris–HCl pH 7.6, 5 mM $MgCl_2$). Aqueous RNA solution (10–15 μg; 0.4 nmol) was diluted with water to a final volume of 80 μl, heated to 90°C for 10 min, cooled to 23°C for 15 min and 20 μl 5× binding buffer was added. After incubation for 5 min at 23°C this RNA was loaded onto the column, rinsed with 5 column volumes (5.0 ml) of binding buffer. Then the column was washed with 6 column volumes of a solution of 20 mM citrulline in binding buffer. The agarose was heated to 90°C for 10 min, cooled to 23°C for 15 min and washed with 5 column volumes of 20 mM citrulline in binding buffer. Elution of arginine specific aptamers was done with a 20 mM solution of arginine in binding buffer for 15 column volumes, followed by a heat denaturation step (90°C, 10 min, 23°C, 10 min) before elution with 5 column volumes of 20 mM arginine in binding buffer (Fig. 1C). Only the fraction of RNA which eluted after the heat denaturation/renaturation step in the presence of arginine was collected, phenol extracted to remove the amino acid, precipitated and amplified as described previously (14).

K_d determinations

Measurements of dissociation constants were performed by the methods of equilibrium dialysis (22,25,26) using dialysis chambers with 500 μl volumes. The membrane (cellulose triacetate filters; Sartorius) had a molecular weight cutoff of 5000 Da. Typically, 1.0 μM RNA concentrations were used. Measurements were taken at three different concentrations of L-[2,3,4,5-^{3}H] arginine–HCl (250, 25 and 2.5 μM). Sample volumes were 200 μl on each side of the membrane. Samples were equilibrated at 23°C overnight, after which the volume on both sides of the membranes was determined to ensure equal volumes on each side. Aliquots were withdrawn and subjected to scintillation counting. From the specific activity of arginine, the concentrations of bound ligand $[ES]_{eq}$, free amino acid $[S]_{eq}$ and free RNA $[E]_{eq}$, at equilibrium, were calculated. The K_d was calculated from the equation:

$$K_d = ([E]_{eq} \times [S]_{eq}) \times [ES]_{eq}^{-1}$$

This analysis assumes the formation of a 1:1 complex between the RNA and the amino acid. The K_d for the binding of the clone ag.06 to D-arginine was determined by analytical affinity chromatography (27). When applied to the 4.0 mM D-arginine agarose, this RNA elutes in the void volume of the column (170 μl agarose, see Fig. 5B). We calculate the K_d to be ≥4 mM.

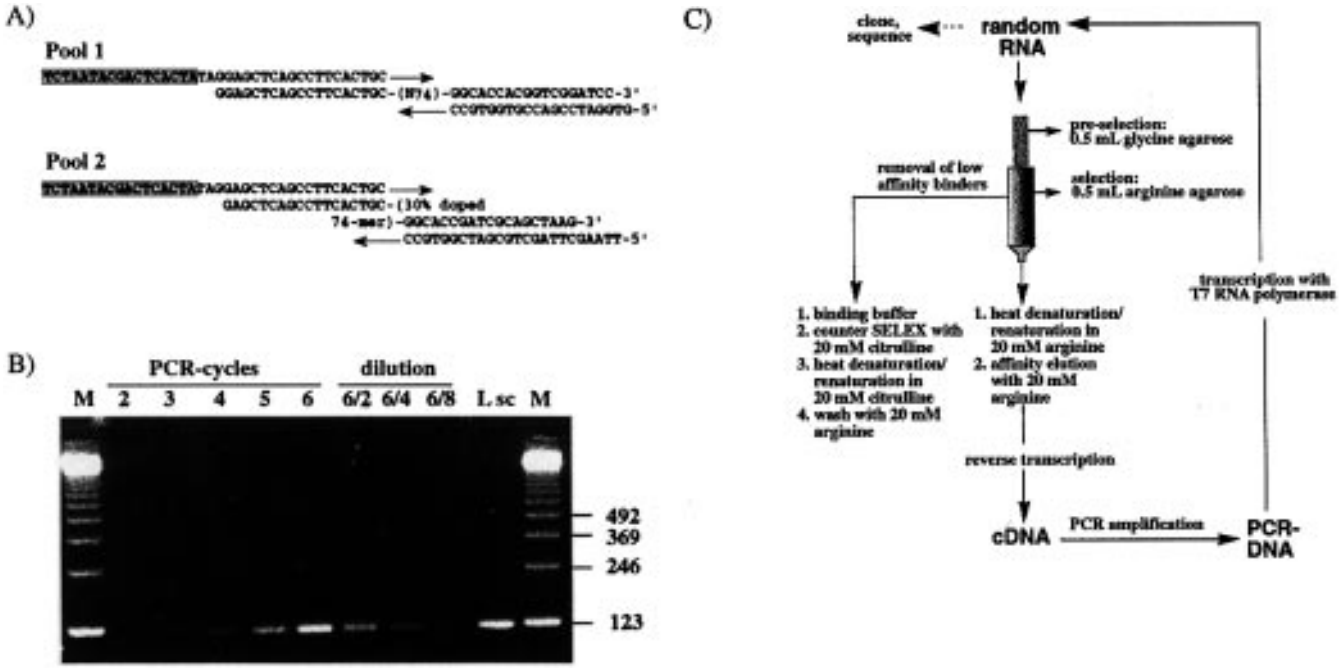

Figure 1. (**A**) Design of the completely randomized pool 1 and the degenerate pool 2. The boxed region shows the T7-promoter sequence. The 3′ primer of pool 1 has a two-base overhang at the 5′ end, resulting in a 113mer RNA after PCR and *in vitro* transcription. The 3′ primer of pool 2 has a five base overhang at the 5′ end and the 5′ primer has one additional G at the 3′ end of the T7 promoter, resulting in a 117mer RNA after PCR and *in vitro* transcription. (**B**) Demonstration of exact doubling of the synthesized pool DNA during PCR cycle 6. The PCR-DNA from cycle 6 was diluted 1:1 (6/2), 1:2 (6/4) and 1:4 (6/8) and the intensities (resulting from ethidium bromide staining) compared with the PCR-DNA from cycle 5, 4 and 3. The intensity of the band at cycle 6 is the same as the intensity of the 123 band in the 123 ladder (M) which corresponds to 50 ng of DNA. Lane L shows 50 ng of DNA from a large scale PCR of pool 2. With 2.2 ng of synthetic ssDNA in cycle 0, 6 cycles of PCR would have yielded 138 ng DNA. Therefore, only 36% of the synthetic input DNA were extendable to the full-length dsDNA PCR product. This value was used to evaluate the pool complexity, which was calculated from the total amount of synthetic, gel purified ssDNA (5.0 nmol) to be $610^{23} \times 5 \cdot 10^{-9} \times 0.36 = 1.1 \times 10^{15}$ molecules. (**C**) Schematic for the *in vitro* selection cycle. RNA is loaded onto the preselection column and eluted for 2.5 column volumes onto the selection column. The preselection column is removed. Non-binding and low affinity binding RNAs are removed by a buffer wash, an affinity elution with 20 mM citrulline, a heat denaturation/renaturation in 20 mM citrulline, and a wash with 20 mM arginine. RNAs which still bind to the column are heat denatured and renatured in 20 mM arginine in binding buffer, eluted with 20 mM arginine/buffer, phenol extracted and reverse transcribed. The cDNA is PCR-amplified, the PCR-DNA*in vitro* transcribed to yield an enriched RNA pool which can be used for the next selection cycle.

Chemical modification

RNA ag.06 (1–5 pmol) was denatured for 3 min at 95°C and renatured in the absence or presence of 50 µM of amino acid in 250 mM NaCl, 50 mM sodium cacodylate pH 7.4 and 5 mM $MgCl_2$ for 10 min at room temperature. For probing with CMCT 50 mM potassium borate pH 8.0 instead of sodium cacodylate was used. Chemical modification was performed by addition of 1 µl DMS (1:5 dilution in 96% ethanol), 1 µl kethoxal (1:5 dilution in H_2O, stock solution: 37 mg/ml) or 12.5 µl CMCT (32 mg/ml in H_2O) to a final reaction volume of 50 µl, followed by incubation at 25°C for 20 min. After precipitation in the presence of 10 µg tRNA, samples were dissolved in H_2O (DMS and CMCT modified samples) or 25 mM potassium borate pH 7.0 (kethoxal modified samples). Detection of modified positions by primer extension and polyacrylamide gel electrophoresis was performed as described previously (28,29) using 5′-^{32}P end-labeled primer M20.106.

RESULTS AND DISCUSSION

The stringent selection scheme led to tight binding sequences

Two RNA-pools of ~10^{15} different sequences (Fig. 1) were subjected to cycles of *in vitro* selection for arginine binding. One pool (pool 2; 117mer) was comprised of a 74mer oligonucleotide sequence which was previously shown to bind citrulline with a dissociation constant of 65 µM (14). Previously, in a standard *in vitro* selection experiment, a mutant RNA aptamer specific for arginine was isolated from this pool which differed from the citrulline binding RNA by three mutations (14). We have now re-selected the same pool as well as a completely randomized pool (pool 1; 113mer) for tight binding to arginine, by increasing the stringency of the selection so that weak binders are removed and only the tight binding arginine specific aptamers are amplified. This was achieved by a modified selection scheme which included a counter selection (23) with 20 mM of the non-cognate amino acid citrulline. To increase the stringency of competition between free citrulline and immobilized arginine, the column bound RNA was heat denatured and renatured in the presence of 20 mM citrulline. Elution was continued with citrulline solution for several column volumes to remove non-bound RNA. We then equilibrated the column with 20 mM arginine in binding buffer for several column volumes, and carried out a heat denaturation/renaturation step in 20 mM arginine. Our rationale for this treatment was that molecules with a slow dissociation rate (k_{off}) which affinity-elute slowly can be assumed to be tight binders. When heat denatured, these RNAs separate from the immobilized ligand and renature in the presence of an excess of ligand in solution. This method should avoid the

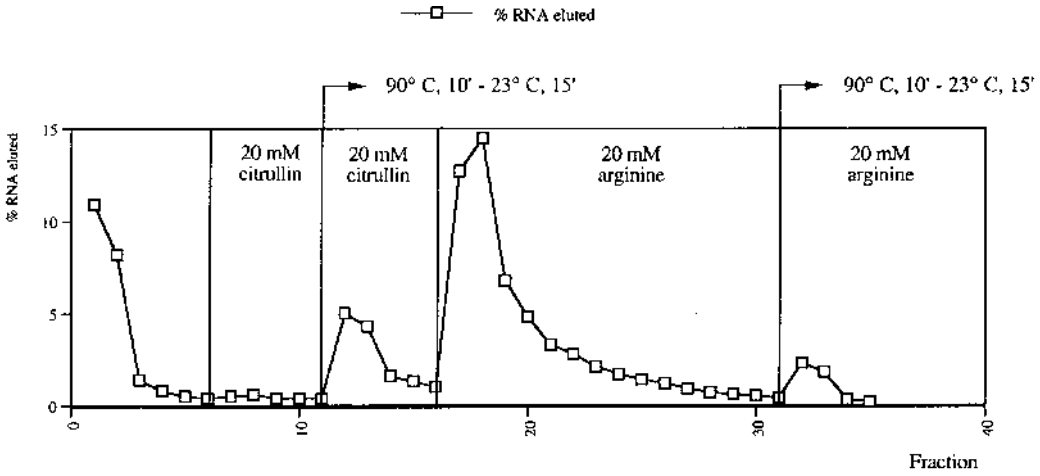

Figure 2. Elution profile of the enriched RNA pool 1 from cycle 20. With pool 2 the same elution pattern was obtained. The fraction that eluted with 20 mM arginine in binding buffer after heat denaturation/renaturation contained 5% of the total RNA loaded to the column.

undesired enrichment of matrix binders or molecules which require functional groups provided by the matrix in addition to the ligand itself for binding, examples of which have been reported in previous selection experiments in which the bound nucleic acid was eluted by denaturation only (12).

In this way, we were able to increase the amount of RNAs which elute with arginine after denaturation/renaturation from undetectable amounts in the first 8 cycles to 5% after 20 cycles. The elution profile obtained at cycle 20 is shown in Figure 2. Elution profiles were essentially identical for pools 1 and 2.

We sequenced the isolated RNA aptamers from both pools after cycles 10 and 20, respectively. Thirty-five percent of the 36 clones sequenced after cycle 10 showed the arginine motif previously described (14), whereas 65% of the sequences were new. After selection cycle 20, however, none of the obtained sequences resembled the previously described motif, indicating that binders with K_ds >60 μM, the K_d of the previously isolated arginine motif (14), have been quantitatively removed during the high stringency selection. The sequences isolated after cycle 20 are listed in Figure 3A.

The variability of the isolated sequences from cycle 20 is still remarkably high: in 31 sequenced clones one sequence appeared eight times, one four times, four sequences appeared twice, and eleven sequences only once. The sequences do not share any obvious primary sequence homologies or relationships although a majority of the clones (~70%) contains tracts of pyrimidines between 6 and 11 bases long which, however, are variable in their sequence and C/U ratios. To determine whether the different arginine binding RNAs contained common secondary structure motifs rather than common sequence motifs, secondary structure predictions were generated using the Zuker program (30). No predominant secondary structural feature appears in the selected RNAs, nor are the pyrimidine tracts utilized to form any particular structural motif. As expected, a higher variability of sequences was found in the completely randomized pool: 60% of the unrelated sequences are from this pool, whereas 40% of the unrelated sequences result from the doped pool. Only one clone, ag.06 which is also the most abundant clone to be isolated was present in both pools. The ratio of 60/40 of unrelated sequences reflects the diversity of the two pools: while the sequence space covered by the completely randomized pool 1 was of unrestricted divergency, the sequence space covered by pool 2 was restricted by its partial randomization. The finding that we obtained a rather divergent set of functional molecules after 20 cycles of stringent selection indicates that many different RNA sequences and possibly tertiary structure solutions capable of tight and specific binding to arginine appear to exist. This underlines the potential importance of arginine side chains as 'handles' for the specific recognition of RNA by RNA-binding proteins. This result is also consistent with structural data obtained by NMR for the HIV-Tar/arginine (16,31), the BIV-Tar/BIV-tat (32,33), and the aptamer/arginine (14) complexes (unpublished data) which show that critical contacts to arginines are achieved by entirely different RNA tertiary structures (34).

To quantify the strength of binding we measured the dissociation constant (K_d) of one of the arginine binding sequences (clone ag.06) in solution by equilibrium dialysis. We chose the most abundant clone which appeared eight times in our enriched pool with two of the isolated clones showing a point mutation at different positions (Fig. 3). This analysis revealed a K_d of 330 nM for clone ag.06 which corresponds to a nearly 200-fold improvement in binding affinity compared to the tightest binding arginine aptamers described so far (14). Specific recognition of arginine by RNA spans over a broad range of binding affinities from 4 mM as found for the arginine/Tar interaction (35) to 330 nM as demonstrated here.

The RNA aptamers described here also compare well in their binding affinity with the best aptamers obtained in other *in vitro* selection experiments. For example, a complex between an RNA aptamer and theophylline interacts with a K_d of 320 nM (23). An aptamer specific for cyanocobalamine binds its ligand with a K_d of 88 nM (22), and the best aptamers for the aminoglycoside antibiotics tobramycin (36), kanamycin, lividomycin (37) and neomycin (38) interact with their cognate ligands with values of K_d between 20 and 220 nM.

Secondary structure elucidation

The computer generated folding of clone ag.06 is shown in Figure 3B. The folding proposes the formation of three relatively long consecutive stems; stem 1 consists of 7 base pairs (bp) which

Figure 3. (**A**) Sequences of the arginine binding RNAs from cycle 20. Only the inserts corresponding to the randomized regions are shown. Sequences resulting from pool 2 are marked by an asterisk (*, magenta), sequences which appeared in both pools are marked with a double-cross (#, magenta). The sequences are listed with respect to their abundance. Point mutations (PM) are indicated in red. (**B**) Computer generated folding of the tight binding sequence ag.06. Primer sequences are shown in blue, small letters; the sequence corresponding to the randomized region is shown in green. Two point mutations (red) were identified in clones ag.04, and ag.16 and appear in stem 3.

are separated by a bulged C from the 10 bp stem 2, and stem 3 containing 8 bp. The formation of stem 3 is supported by the point mutations found in clones ag.04, and ag.16, respectively, which result in the substitution of two G·U wobble pairs for G·C pairs (Fig. 3B).

To examine the secondary structure of clone ag.06, a chemical modification analysis with DMS, kethoxal and CMCT was carried out in the presence and absence of L-arginine (Fig. 4).

In the presence of the ligand the chemical probing data confirm the formation of the Watson–Crick paired regions. In its absence, only the base pairs formed by A25–U65 and A26–U64 which close stem 2 are accessible to chemical modification. Stem 3 forms regardless whether the amino acid is present since protections from chemical modification at Watson–Crick positions of stem 3 are identical in presence and absence of L-arginine. Only the positions G49 and U53 which are likely to be paired, because of the U to C mutation in clone ag.04, are weakly modified. Among the unpaired base positions in the multibranched loop region only A27, A31, A41, G61, A62 and G63 are accessible to modification in the absence of arginine. In the presence of the amino acid, all adenosine residues (A27, A31, A41 and A62) in the multibranched region are protected. G61 shows identical degrees of modification in the presence and absence of the ligand. G63 is weakly protected in the presence of 50 µM arginine. Table 1 summarizes the differences in the modification pattern observed for the free RNA and the complex.

An interesting result is the lack of modification at positions C28, C29 and G30 both in the presence and absence of arginine. Likewise, we only observed very weak instead of the expected strong modification in the UGG loop which closes the 3 bp stem formed by U32–G40, G33–C39 and C34–G38. This suggests the formation of a pseudoknot structure (39,40) due to Watson–Crick pairing at positions C28–G37, C29–G36 and possibly G30–U35. A pseudoknot structure was also the critical recognition element in the cyanocobalamine aptamer isolated by Lorsch and Szostak (22). This pseudoknot bound to cyanocobalamine with values of K_d <100 nM.

The relatively high percentage of Watson–Crick pairing found in clone ag.06 and other clones (data not shown) might be the result of our applied selection scheme involving heat denaturation/renaturation; those sequences which fold back into a defined secondary structure within the relatively short renaturation period of 15 min were preferentially enriched compared to other potentially tight

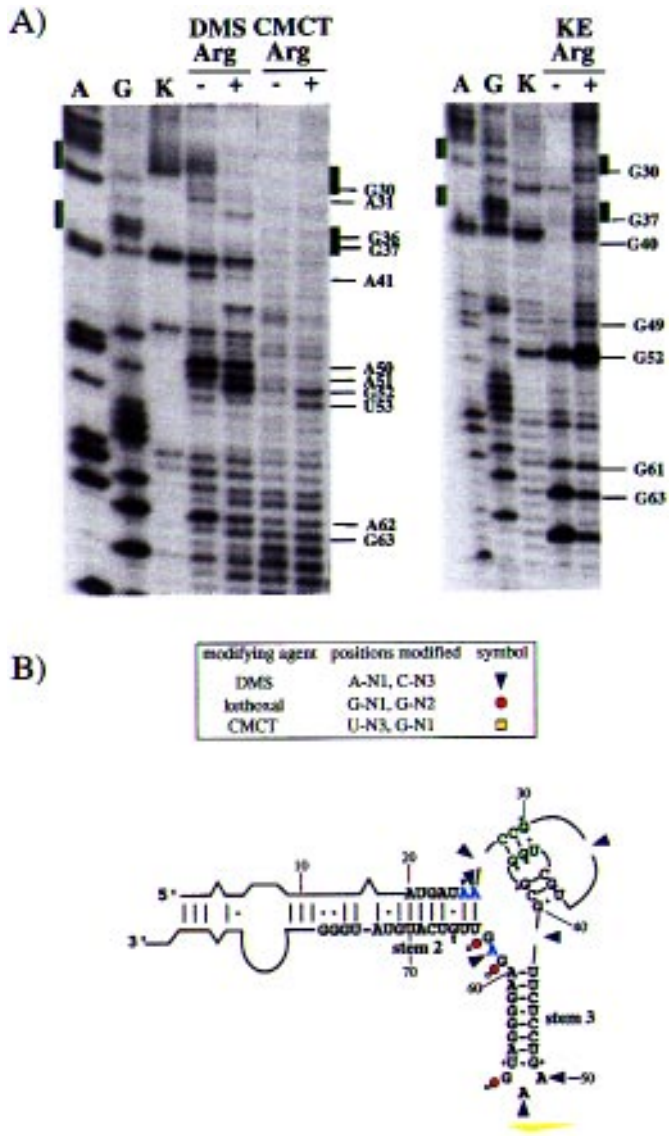

Figure 4. Refined secondary structure of clone ag.06. (**A**) Chemical probing and footprinting of the secondary structure of ag.06 RNA with DMS, kethoxal and CMCT. A and G indicate the dideoxy sequencing lanes, K the primer extension of unmodified RNA. Chemical protection analysis was performed in the free RNA (–) and in the presence (+) of L-arginine (Arg). The regions corresponding to residues C28C29G30 and U35G36G37, respectively, which are not accessible to chemical probing and presumably base pair to form a pseudoknot structure, are indicated by a green bar. (**B**) Refined secondary structure of ag.06 RNA as proposed from the chemical modification analysis. The constant regions which belong to the primer binding sites are represented schematically. The reactivity of nucleotides in the free RNA is summarized by the coloured labels (blue triangle, DMS-modified; red circle, kethoxal modified; yellow square, CMCT modified). The size of the label reflects the degree of modification. Base positions which are protected from modification in the complex are shown in light blue. Residues C28C29G30 and U35G36G37 are shown in green.

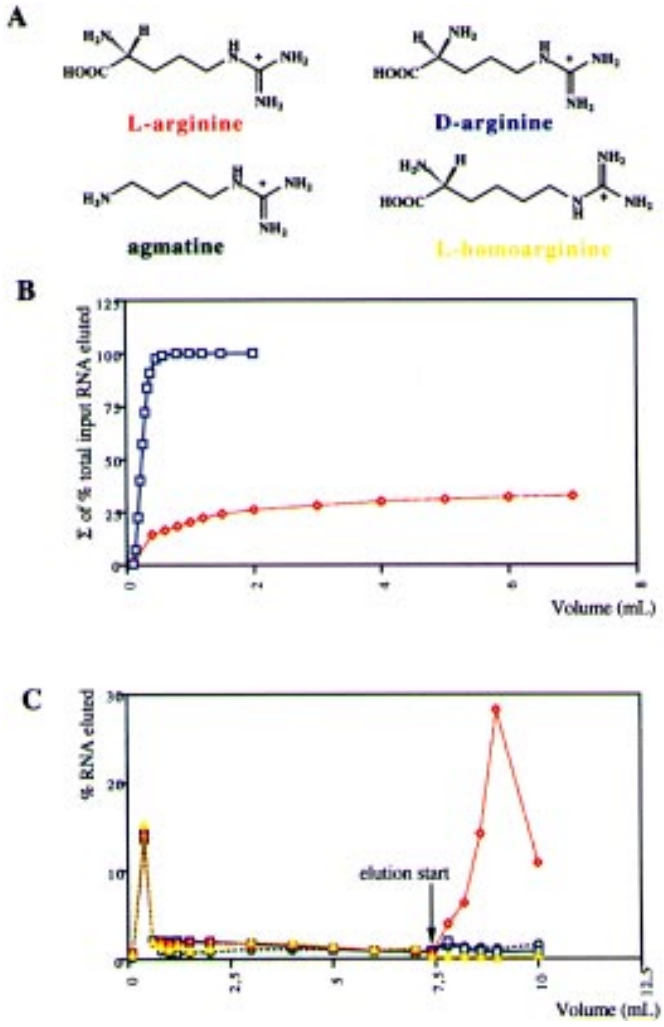

Figure 5. Specificity of L-arginine recognition. (**A**) Chemical structures of ligands used for the specificity determination. (**B**) Elution of ag.06 from a 4 mM L-arginine agarose (red) and from a 4 mM D-arginine agarose (blue). Cumulative output c.p.m. are expressed as the percentage of the input c.p.m. (**C**) Affinity elution of ag.06 RNA from an L-arginine column by L-arginine and competition by arginine analogues. Competitive affinity elution was performed with 12.5 mM solutions of L-arginine (red), D-arginine (blue), agmatine (green) and L-homoarginine (yellow). The column volume was 0.17 ml of swollen agarose.

binding molecules which have slower folding kinetics. This is supported by the finding that the folding proposed for all RNAs suggest a significantly higher content of G·C base pairs than A·U or G·U base pairs. For example, clone ag.01 contains 22 G·C pairs, 8 A·U pairs and 7 G·U pairs (data not shown), which means that 75% of all bases in this clone are base paired.

Binding specificity of ag.06

As expected, binding of clone ag.06 to the non-cognate amino acid L-citrulline was not detectable, consistent with the negative selection against L-citrulline. To test the specificity of clone ag.06 for L-arginine, we examined the binding affinity of this clone to several arginine analogues, D-arginine, agmatine and L-homoarginine (Fig. 5A). Figure 5B shows the result of a binding assay performed with 4 mM D-arginine agarose and 4 mM L-arginine agarose, respectively. After a 7 ml buffer wash, corresponding to 35 column volumes, 65% of the total input clone ag.06 RNA still bound to the L-arginine column. From the D-arginine column this RNA elutes in the void volume which means that the K_d of the ag.06 RNA for D-arginine is at best 4.0 mM. With the solution K_d of 330 nM for L-arginine recognition, ag.06 discriminates at least

12 000-fold between the D- and L-enantiomer of arginine. This is the highest factor of discrimination between similar ligands or enantiomers by an RNA molecule known to date. Other RNA aptamers with similar discrimination factors are known (41). For example, an aptamer isolated for theophylline recognition discriminates theophylline from caffeine with a 10 000-fold difference in the binding affinity (23). The enantioselectivity of the ag.06 RNA corresponds to a difference in the relative binding energies (ΔΔG) of 30 kJ·mol^{-1} (7.2 kcal·mol^{-1}). Several other RNA aptamers have been isolated that can discriminate between L- and D-amino acids (12–15) although with relatively poor discrimination factors.

Figure 5C shows elution profiles of clone ag.06 bound to L-arginine agarose. Affinity elutions with 12.5 mM solutions of different arginine analogues were performed after an extensive buffer wash. Only the L-arginine affinity elution was capable of eluting most of the RNA off the column within 12 column volumes, whereas D-arginine, agmatine or L-homoarginine were inactive. The results of the affinity elutions are quantified in Table 2.

Table 1. Summary of the modified positions in clone ag.06 in the presence and absence of L-arginine

Base position	Modified in free RNA	Modified in complex
A25	+	–
A26	++	–
A27	++	–
C28	–	–
C29	–	–
G30	±	±
A31	++	–
G33	–	–
G36	±	±
G37	±	±
G38	±	±
G40	–	–
A41	++	–
G49	+	+
A50	++	++
A51	++	++
G52	++	++
U53	+	+
G61	++	++
A62	++	–
G63	++	++

++, Indicates a high degree modification; +, indicates a medium degree of modification, ±, indicates a low degree of modification; –, indicates protection from modification.

The elution data performed with the arginine analogues are consistent with the clone ag.06 RNA being highly specific for L-arginine recognition. In a recent publication it has been argued that selection for high-affinity binding automatically results in highly specific binding (42). The results presented here confirm these suggestions; although we did not select for enantioselective binding, the tight binding RNA aptamer ag.06 binds L-arginine with higher enantio- and ligand selectivity than other RNA aptamers with lower binding affinity to their cognate ligands (12–15).

Table 2. Affinity elution of L-arginine-bound ag.06 RNA with different arginine analogues

Ligand	Percent of total input RNA eluted from L-arginine column
L-arginine	64.1
D-arginine	6.3
agmatine	7.3
L-homoarginine	1.6

The concentration of L-arginine on the column was 4.0 µmol/ml of swollen agarose (4.0 mM) as determined by derivatization in the presence of ^{3}H-labelled arginine. Affinity elution was performed with 12.5 mM solutions of ligands in binding buffer over 12 column volumes (2.0 ml) at ambient temperature within 15 min of incubation.

CONCLUSION

An RNA molecule ('aptamer') which binds tightly and highly specifically to L-arginine in solution has been isolated by applying a modified *in vitro* selection scheme. A series of arginine binding sequences was selected, which do not share primary sequence relationships to each other. This demonstrates that the tight and specific molecular recognition of arginine can be achieved by many different RNA sequences. One of the selected RNA aptamers was shown to exhibit the highest factor of discrimination between two low molecular weight enantiomers, L- and D-arginine, reported to date. The relative binding energies for L- and D-arginine binding differ by 30 kJ.mol^{-1} (7.2 kcal.mol^{-1}) and correspond to a 12 000-fold difference in the K_d values. Binding to other arginine analogues, such as agmatine and L-homoarginine cannot be detected. The secondary structure of the tight binding L-arginine aptamer was refined by chemical modification analysis for the complex and the free RNA. This analysis revealed the formation of a compact pseudoknot structure. In the free RNA, all unpaired adenosine residues within the core region which forms the pseudoknot are accessible to modification with DMS, whereas these positions are protected in the complex. This suggests that these residues might be involved in a direct contact to the bound amino acid, or in the formation of tertiary interactions induced by the ligand. The specificity of the isolated RNA aptamer for L-arginine might indicate that every substituent carried by the amino acid is directly contacting the RNA.

The high specificity obtained for the molecular recognition of L-arginine by the ag.06 RNA aptamer raises the question whether such highly specialized motifs can be expected to play a role in complexes between arginine-rich peptides or proteins and RNA. As recently pointed out by Sundquist (34), the interactions in the full length HIV-1 Tat–protein–TAR complex cannot be mimicked by arginine alone, but high resolution NMR analyses (16,31) have allowed to relate structural differences observed in the free TAR–RNA and the Tat–TAR complex to individual arginine residues. Analogously, the identification and characterization of DNA (43) or RNA motifs such as the one described here will help

to give new insights into the principles of RNA–protein and RNA–ligand interactions.

ACKNOWLEDGEMENTS

We thank A. Hüttenhofer for discussions and critical reading of the manuscript. M.F. expresses his gratitude to E.-L. Winnacker and W. Wolf for their support. This work was supported by grants from Boehringer Mannheim to A.G. and M.F., the Deutsche Forschungsgemeinschaft (grant no: Fa275/1-2) and the European Union (grant no: Biot-CT93-0345) to M.F.

REFERENCES

1 Draper,D.E. (1995) *Annu. Rev. Biochem.*, **64**, 593–620.
2 Yarus,M. (1988) *Science*, **240**, 1751–1758.
3 Michel,F., Hanna,M., Green,R., Bartel,D.P. and Szostak,J.W. (1989) *Nature*, **342**, 391–395.
4 Schimmel,P. and Schmidt,E. (1995) *Trends Biochem. Sci.*, **20**, 1–2.
5 Jakubowski,H. and Goldman,E. (1992) *Microbiol. Rev.*, **56**, 412–429.
6 Calnan,B.J., Tidor,B., Biancalana,S., Hudson D. and Frankel,A.D. (1991) *Science*, **252**, 1167–1171.
7 Weeks,K.M., Ampe,C., Schultz,S.C., Steitz,T.A. and Crothers,D.M. (1990) *Science*, **249**, 1281–1285.
8 Famulok,M. and Szostak,J.W. (1992) *Angew. Chem. Int. Ed. Engl.*, **31**, 979–988.
9 Klug,S.J. and Famulok,M. (1994) *Mol. Biol. Rep.*, **20**, 97–107.
10 Gold,L., Polisky,B., Uhlenbeck,O.C. and Yarus,M. (1995) *Annu. Rev. Biochem.*, **64**, 763–797.
11 Joyce,G.F. (1994) *Curr. Opin. Struct. Biol.*, **4**, 331–336.
12 Famulok,M. and Szostak,J.W. (1992) *J. Am. Chem. Soc.*, **114**, 3990–3991.
13 Connell,G.J., Illangesekare,M. and Yarus,M. (1993) *Biochemistry* , **32**, 5497–5502.
14 Famulok,M. (1994) *J. Am. Chem. Soc.*, **116**, 1698–1706.
15 Majerfeld,I. and Yarus,M. (1994) *Nature Struct. Biol.*, **1**, 287–292.
16 Aboul-ela,F., Karn,J. and Varani,G. (1995) *J. Mol. Biol.*, **253**, 313–332.
17 Kjems,J., Calnan,B.J., Frankel,A.D. and Sharp,P. (1992) *EMBO J.*, **11**, 1119–1129.
18 Tan,R.Y., Chen,L., Buettner,J.A., Hudson,D. and Fraenkel,A.D. (1993) *Cell*, **73**, 1031–1040.
19 Frankel,A.D. (1994) In Nagai,K. and Mattaj,I.W. (eds) *RNA–protein Interactions*. IRL Press, NY, pp. 221–247.
20 Baskerville,S., Zapp,M. and Ellington,A.D. (1995) *J. Virol.*, **69**, 7559–7569.
21 Sassanfar,M. and Szostak,J.W. (1993) *Nature*, **364**, 550–553.
22 Lorsch,J.R. and Szostak,J.W. (1994) *Biochemistry*, **33**, 973–982.
23 Jenison,R.D., Gill,S.C., Pardi,A. and Polisky,B. (1994) *Science*, **263**, 1425–1429.
24 Burgstaller,P. and Famulok,M. (1994) *Angew. Chem., Int. Ed. Engl.*, **33**, 1084–1087.
25 Fersht,A. (1985) *Enzyme Structure and Mechanism*, 2nd Edition. W. H. Freeman and Co., NY, pp. 186.
26 Connors,K.A. (1987) *Binding Constants*. Wiley, NY, pp. 314.
27 Arnold,F.H., Schofield,S.A. and Blanch,H.W. (1986) *J. Chomatogr.*, **355**, 1–12.
28 Burgstaller,P., Kochoyan,M. and Famulok,M. (1995) *Nucleic Acids Res.*, **23**, 4769–4776.
29 Stern,S., Moazed,D. and Noller,H.F. (1988) *Methods Enzymol.*, **164**, 481–489.
30 Zuker,M. (1989) *Science*, **244**, 48–52.
31 Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) *Science*, **257**, 76–80.
32 Puglisi,J.D., Chen,L., Blanchard,S. and Frankel,A.D. (1995) *Science*, **270**, 1200–1203.
33 Ye,X., Kumar,R.A. and Patel,D.J. (1995) *Chem. Biol.*, **2**, 827–840.
34 Sundquist,W.I. (1996) *Nature Struct. Biol.*, **3**, 8–11.
35 Tao,J. and Frankel,A. D. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 2723–2726.
36 Wang,Y. and Rando,R.R. (1995) *Chem. Biol.*, **2**, 281–290.
37 Lato,S.M., Boles,A.R. and Ellington,A.D. (1995) *Chem. Biol.*, **2**, 291–303.
38 Wallis,M.G., Von Ahsen,U., Schroeder,R. and Famulok,M. (1995) *Chem. Biol.*, **2**, 543–552.
39 Mans,R.M.W. and Pleij,C.W.A.(1993) In Eckstein,F. and Lilley,D.M.J. (eds) *Nucleic Acids and Molecular Biology*. Vol. **7**, Springer Verlag, Berlin, pp. 250–270.
40 Puglisi,J.D., Wyatt,J.R. and Tinoco,I., Jr (1990) *J. Mol. Biol.*, **214**, 437–453.
41 Ellington,A.D. (1994) *Curr. Biol.*, **4**, 427–429.
42 Eaton,B.E., Gold,L. and Zichi,D.A. (1995) *Chem. Biol.*, **2**, 633–638.
43 Harada,K. and Frankel,A.D. (1995) *EMBO J.* **14**, 5798–5811.