In Vitro Selection of Dopamine RNA Ligands†

Cecilia Mannironi,‡ Alessia Di Nardo,‡,|| Paolo Fruscoloni,‡ and G. P. Tocchini-Valentini*,‡,§

Institute of Cell Biology, CNR, Viale Marx 43, 00137 Rome, Italy, and Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637

Received January 13, 1997; Revised Manuscript Received May 20, 1997⊗

ABSTRACT: RNA aptamers that specifically bind dopamine have been isolated by *in vitro* selection from a pool of 3.4×10^{14} different RNA molecules. One aptamer (dopa2), which dominated the selected pool, has been characterized and binds to the dopamine affinity column with a dissociation constant of 2.8 μM. The specificity of binding has been determined by studying binding properties of a number of dopamine-related molecules, showing that the interaction with the RNA might be mediated by the hydroxyl group at position 3 and the proximal aliphatic chain in the dopamine molecule. The binding domain was initially localized by boundary experiments. Further definition of the dopamine binding site was obtained by secondary selection on a pool of sequences derived from a partial randomization of the dopa2 molecule. Sequence comparison of a large panel of selected variants revealed a structural consensus motif among the active aptamers. The dopamine binding pocket is built up by a tertiary interaction between two stem and loop motifs, creating a stable framework in which five invariant nucleotides are precisely arrayed. Minimal active sequence and key nucleotides have been confirmed by the design of small functional aptamers and mutational analysis. Enzymatic probing suggests that the RNA might undergo a conformational change upon ligand binding that stabilizes the proposed tertiary structure.

Repeated cycles of enrichment and enzymatic amplification allow the isolation of rare molecules with specific properties from large populations ($10^{14}-10^{15}$) of random DNA or RNA sequences. These techniques, commonly called *in vitro* selection or SELEX (Tuerk & Gold, 1990; Green et al., 1991; Gold et al., 1995; Uphof et al., 1996), have been used to identify, from such combinatorial libraries, RNA and single-stranded DNA (aptamers) that specifically recognize proteins and a wide variety of low molecular weight ligands. Even though small molecules provide fewer opportunities for specific interactions, *in vitro* selection experiments have been successfully aimed at low molecular weight ligands, providing a simple system for the study of the versatility of the four chemical building block (Gold et al., 1995). The small molecules recognized by RNA motifs range from natural molecules, such as amino acids (Famulok & Szostak, 1992; Connel et al., 1993; Famulok, 1994; Majerfeld & Yarus, 1994; Geiger et al., 1996), cofactors (Lorsch & Szostak, 1994; Burgstaller & Famulok, 1994), nucleotides (Sassanfar and Szostak, 1993; Connel et al., 1994), the alkaloid theophylline (Jenison et al., 1994), and aminoglycoside antibiotics (Wang et al., 1996; Famulok & Hüttenhofer, 1996), to abiotic organyc dyes (Ellington & Szostak, 1990) and transition state analogs (Morris et al., 1994). Recently an RNA motif that binds Zn ions has been isolated (Ciesiolka et al., 1995; Ciesiolka & Yarus, 1996). RNA motifs for small molecules were found to have different structures, from the relatively simple internal loops, to more complex pseudoknots. The solution structures for different RNA−ligand complexes have been solved at atomic-level resolution, defining in these cases the principles associated with RNA folding and molecular recognition (Yang et al., 1996; Dieckmann et al., 1996; Jiang et al., 1996; Fan et al., 1996). In addition, such RNA or DNA aptamers, because of their extremely high affinities and specificities, have potential as diagnostic and pharmacological tools (Jellinek et al., 1993; Binkley et al., 1995; Nieuwlandt et al., 1995; Gold, 1995).

We have been interested in isolating RNA molecules able to bind tightly and specifically the small neurotransmitter dopamine. Dopamine is a pivotal molecule in the central nervous system, playing a major role in the regulation of movement, and is implicated in the pathophysiology of Parkinson's and Huntington's diseases, psychosis, and drug addiction (Nestler, 1994). No interaction between dopamine and nucleic acids *in vivo* has been reported. In this paper we describe the isolation from a random RNA library of high-affinity dopamine RNA aptamers, the high specificity of which was assessed by the study of the binding properties of related compounds. Covariation analysis, site-directed mutagenesis and nuclease mapping indicate that the dopamine binding site is likely to be in a cleft between two loops, which interact through Watson−Crick or wobble base pairing. We hypothesize that the ligand specificity depends on the identity of five highly conserved bases.

MATERIALS AND METHODS

Materials. Sepharose CL-4B, dopamine, norepinephrine, catechol, tyramine, phenethylamine, ethylamine, and L-dopa were purchased from Sigma, dopamine-agarose was from

† This Research was supported by Progetto Finalizzato "Ingegneria Genetica", by Progetto Finalizzato "Biotecnologie e Biostrumentazione" and by Progetto Strategico "Nucleotidi Antisenso" (CNR). A.D.N. was supported by a fellowship from the Fondazione Adriano Buzzati-Traverso.

* Author to whom correspondence should be addressed.

‡ CNR.

|| Present address: EMBL, Postfach 102209, 69012 Heidelberg, Germany.

§ University of Chicago.

⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1997.

[1] Abbreviations: K_d, equilibrium dissociation constant; PCR, polymerase chain reaction; nt, nucleotide; ss, single strand; ds, double strand.

S0006-2960(97)00063-9 CCC: $14.00

ICN, and 3-*O*-methyldopamine was from Aldrich. All radiolabeled nucleotides were from Amersham. T7 RNA polymerase and Taq DNA polymerase were kindly provided by Kelly Williams; all other enzymes were purchased from commercial sources. Sequencing was performed either using a T7 DNA polymerase sequencing kit for conventional dideoxy sequencing, or using a dideoxy terminator cycle sequencing kit for sequence analysis on a 373A DNA Sequencer (Applied Biosystems). Oligonucleotides were synthesized on a 392 DNA/RNA synthesizer using standard phosphoramidite chemistry and purified by OPC cartridges (Applied Biosystems), unless otherwise noted.

RNA Pools. An initial pool (pool 1) of 4×10^{15} individual DNA molecules, 114 nucleotides (nt) long, containing 80 bases of random sequence flanked by defined regions at the 5′ and 3′ ends, was chemically synthesized and purified on denaturing polyacrylamide gel. The 5′ PCR primer, 5′TAATACGACTCACTATAGGGAATTCCGCGTGTGC, contained the sequence of the T7 RNA polymerase promoter and the *Eco*RI restriction site. The 3′ PCR primer, 5′GAGGATCCCGAACGGAC, contained the *Bam*HI restriction site. It was found that 8.4% of the random oligos could be fully extended by Taq DNA polymerase, giving an effective complexity of the DNA pool of 3.4×10^{14} different sequences.

Transcription templates were created by PCR amplification of degenerate oligos; a unique PCR cycle was done on the PCR products to rectify heteroduplex accumulated during the last PCR cycles (Green et al., 1991). The RNA pool was synthesized by T7 RNA polymerase transcription and purified on denaturing polyacrylamide gels.

The mutagenized pool 2, utilized in selections 2.1 and 2.2, was constructed by randomizing the original sequence of the dopa2 aptamer so as to yield 30% mutation (Bartel & Szostak, 1993) in the region spanning from the first nucleotide to nucleotide 93. A new constant region containing the sequence of the T7 RNA polymerase promoter (5′TAATACGACTCACTATAGGGAAGCTTGTACAGGG) was created at the 5′ end as an upstream PCR primer. The 3′ PCR primer was the same used for the amplification of pool 1. From the chemical synthesis yield and Taq DNA polymerase primer extension experiment we estimated that the complexity of pool 2 was 3×10^{13} different molecules.

Selection Procedures. Selections were performed using dopamine-agarose affinity columns containing 1.7 mM dopamine, linked through its amino groups to cyanogen bromide-activated agarose. In cycles 1, 4, and 6−9 the RNA pools were first counterselected on CL-Sepharose 4B to remove RNAs with affinity for Sepharose. ^{32}P-labeled RNA, water dissolved, was denaturated for 5 min at 65 °C and cooled to room temperature for 10 min in column buffer before being loaded onto the affinity column pre-equilibrated with 50 vol of column buffer. Column buffer contained 50 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$, and 0.5 M NaCl for selections 1 and 2.2 or 0.15 M NaCl for selection 2.1. Because of the high sensitivity of dopamine to oxygen, 0.02% of ascorbic acid was always included in the column buffers. In the first round of selection 1, 30 nmol of the initial RNA pool, corresponding to 48 copies of individual sequences, were loaded onto a 1 mL dopamine-agarose column. In all subsequent rounds of selection 1−3 nmol of RNA were applied to 0.2 mL columns. RNA was allowed to bind to the column ligand for 18 min, and unbound RNA molecules were then extensively washed with column buffer, until eluted radioactivity had reached a plateau level (10−90 column volumes were necessary). After the wash step, 12 vols of column buffer containing 0.1 M dopamine (in selection 1) or 1 mM dopamine (in selections 2.1 and 2.2.) were applied to the column and incubated for 20 min. The affinity-eluted RNAs were phenol and ether extracted, ethanol precipitated in the presence of 20 μg of glycogen, and quantitated by Cerenkov counting. Water-dissolved RNA was reverse transcribed and amplified by PCR.

Cloning and Sequencing. PCR-amplified DNA from the round 9 pool of the selection 1 was cloned by restriction enzyme cleavage at the *Eco*RI and *Bam*HI sites in the 5′ and 3′ primer regions, respectively, followed by ligation with an *Eco*RI−*Bam*HI-digested pUC 13 vector. PCR-amplified DNA from round 6 of the secondary selections was blunt-ended by Klenow 3′− 5′ exonuclease activity and cloned into a *Sma*I-digested pUC13 vector. Individual clones were sequenced by the dideoxy methods.

The Pileup DNA multiple sequence alignment program (Feng & Doolittle, 1987) was used for comparative sequence analysis. Prediction of RNA secondary structure by free energy minimization was done by the MulFold program (Jaeger et al., 1989).

Partial Alkaline Hydrolysis. The 5′- and 3′-^{32}P-labeled dopa2 RNA was partially digested under mild alkaline conditions (Pan & Uhlenbeck, 1992). Selection of the digestion products was performed under *in vitro* selection conditions eluting bound RNAs with 0.1 M dopamine. Aliquots of the collected fractions were loaded onto a 8% polyacrylamide gel containing 8 M urea for autoradiography.

Dissociation Constant Measurements. Dissociation constants (K_d) were measured by either isocratic elution from a 1 mL of dopamine-agarose column (Arnold et al., 1986) or by equilibrium filtration (Jenison et al., 1994).

Ligand Specificity. Relative elution of the dopamine-agarose-bound dopa2/c.1 RNA by dopamine-related molecules was determined to study the specificity of ligand binding. All ligand specificity experiments were done applying a batch procedure of affinity chromatography, using 50 μL of dopamine-agarose matrix in 1.5 mL Eppendorf tubes. From 10 to 15 pmol of ^{32}P-labeled RNA was applied to the affinity matrix in 0.5 M NaCl, 50 mM Tris-HCl, 5 mM $MgCl_2$, and 0.02% ascorbic acid. Unbound RNA was washed with 0.6 mL of column buffer and dopamine-agarose-bound RNA eluted with the same volume of 1 mM solution of each molecule. Eluted fractions were quantitated by Cerenkov counting in a scintillation counter.

Enzymatic Probing. RNA transcripts were 5′-^{32}P-labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase. RNAs were denatured at 65 °C for 5 min and cooled to room temperature for 10 min in 0.15 M NaCl, 50 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$, and 0.02% ascorbic acid. For the study of nuclease cleavage in the presence of the ligand, 1 mM dopamine was dissolved in the reaction buffer. T1 cleavage reaction was performed at 37 °C for 30 s at a final concentration of 0.1 unit/μL, V1 at room temperature for 5 min at 0.024 unit/μL, S1 at 37 °C for 1 min at 10 units/μL, 1 mM $ZnCl_2$ and 1 mM spermine were included in the S1 reaction buffer. An equal volume of a gel loading buffer containing 9 M urea/50 mM EDTA, pH 8, was added and the reaction products were immediately separated on 14% polyacrylamide gels containing 8 M urea.

Affinity Elution Experiments on Modified Aptamers. RNAs were obtained by *in vitro* transcription from DNA templates

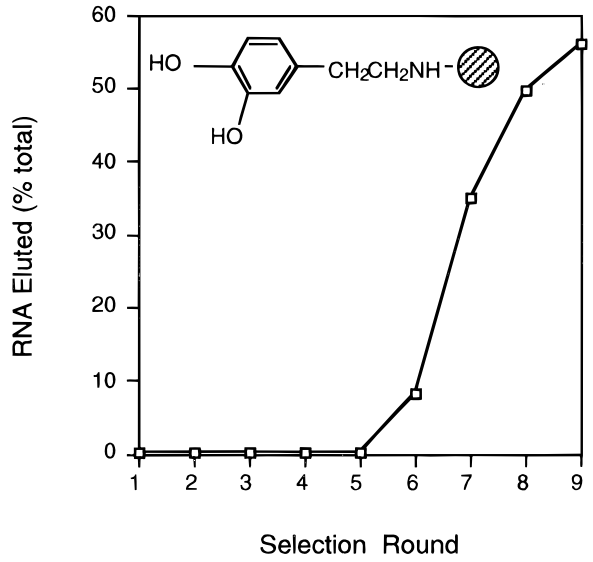


FIGURE 1: Summary of the selection for dopamine aptamers from the random sequence RNA pool (selection 1). The curve represents the percent of the total RNA loaded on the column that was specifically eluted by 0.1 M dopamine. The structure of dopamine, linked to a Sepharose bead, is shown (inset).

containing T7 RNA polymerase promoter. DNA templates were generated by PCR amplification of synthetic oligonucleotides or selected clones. Point mutations on the dopa2/c.1 and dopa1.30/c.30 molecules were introduced by PCR primers, and the presence of mutations was confirmed by ribonuclease analysis in the *in vitro* transcribed RNAs.

From 10 to 15 pmol of ^{32}P-labeled RNA was applied in high salt buffer (0.5 M NaCl, 50 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$, and 0.02% ascorbic acid) to 0.2 mL dopamine-agarose columns. Affinity chromatography was performed as in the selection procedure.

RESULTS

Selection of Dopamine-Binding RNAs. RNA aptamers specific for dopamine were selected from a pool of RNA molecules with an 80 nt random region and 5′ and 3′ constant regions for reverse transcription and PCR. The complexity of the initial pool was approximately 3.4×10^{14} independent sequences. Aptamers were isolated by affinity chromatography on a dopamine-agarose columns. RNA was loaded in high salt column buffer (0.5 M NaCl, 50 mM Tris-HCl, pH 7.4, and 5 mM $MgCl_2$) to prevent nonspecific interactions between the RNA and the column matrix and counterselected against a sepharose matrix. Prior to eluting the bound RNA with 0.1 M dopamine, the column was extensively washed with column buffer. A significant increase in binding to dopamine was observed after six rounds of selection, and after nine rounds 56% of the applied RNA was affinity eluted from the column (Figure 1).

Analysis of Dopamine Aptamers. RNA from the final cycle was reverse transcribed, amplified, and cloned into a pUC13 vector for dye terminator automatic sequencing. Of the 44 sequenced clones, the most abundant one (clone 2) occurred 20 times; the others were differently represented as indicated in Figure 2. A search for primary sequence homology among the 14 unique clones revealed no apparent consensus sequence or obvious structural similarities among the aptamers. RNA was made from the more represented clones and binding activity was determined on a dopamine-agarose matrix. The K_ds of clones 2 and 41, measured by isocratic elution, are 2.8 and 3.5 μM, respectively, 6.4 μM is the K_d value measured for the round 8 selected pool. The most abundant clone 2 (hereafter, dopa2) was chosen for further analysis.

Determination of Minimal Sequence Requirements. In order to define minimal structures of the dopa2 aptamer for dopamine binding, we determined the 5′ and 3′ boundaries of the binding domain (Figure 3, panels A and B). Partial alkaline hydrolysis of end-labeled dopa2 RNA and subsequent analysis of those RNA fragments still capable of binding to the affinity matrix showed that the 5′ and 3′ boundaries are localized at nucleotides 6 and 85, respectively. The MulFold dopa2 RNA structure is shown in panel C of Figure 3; the RNA secondary structure was confirmed by S1 nuclease digestion, except for the region from nucleotide 72 to nucleotide 80, that appeared to be a single-strand loop region (data not shown). Any RNA molecules shorter than the defined boundary positions did not bind to the dopamine column. This result suggests that the 5′ PCR primer sequence (in small letters in Figure 3, panel C) is part of the active molecule, in agreement with the fact that *in vitro* evolution has selected nucleotides able to form a duplex with the 5′ constant region. The 3′ border of the minimal binder lies within stem 3 of the dopa2 RNA, excluding all the region of stem and loop 4 (SL4). The analysis of binding activity of *in vitro* transcribed RNAs from synthetic oligonucleotides (data not shown) showed that an RNA spanning from position 6 to position 85 binds to the column and is specifically eluted (35% of the input RNA is affinity eluted by 0.1 M dopamine). As shown in panel A of Figure 3 (small closed arrows), RNA molecules 3′-labeled, with 5′ ends comprised between positions 41 and 46, bound to the column and were affinity eluted. An RNA encoding positions 42−113 was made, but no binding activity was

SEL.1

```
2  gggaattccgcgtgtgcGCCGCGGAAGACGTTGGAAGGATAGATACCTACAACGGGGAATATAGAGGCCAGCACATAGTGAGGCCCTCCTCCCAAGgtccgttcgggatcctc (20)++
23 gggaattccgcgtgtgcCAGCCACGTAGAATCGTCTCTGCGCCGGGGAGAGTGTCATGGTTAACGGGTAACAGCTCGTGGAGACGACCTCGGTTTCGgtccgttcgggatcctc(4) +
41 gggaattccgcgtgtgcTGGCGGGGAGAACTTACATGGATTAGAGAGTGGTGCTTCATAAGAAGGACAACTGCGTTCCAGCCGCACCCCCCCCCCAgtccgttcgggatcctc (3) ++
7  gggaattccgcgtgtgcACAGCCAGGGTAAGAGTCTAGAGTAGTTCATTCCTAGTATCTGGAGGCTGGTACTATAAGGGTAGTACTGCCCGCGTACTgtccgttcgggatcctc(3) +
20 gggaattccgcgtgtgcCTTTCCATGTATTAGAGAGGAGAGCGCAAACACGAAGGCTCGTCGTTACAGCCGAAGAGCTCCAGTACCATGGTTCATGAgtccgttcgggatcctc(3) +
22 gggaattccgcgtgtgcATGGTTAATGGAATCCTGGGCTCCGGAGATCACCCATCGGAGGTAATTTCTTTAGAATGCTAGACGGACACTACCAGCGTgtccgttcgggatcctc(2) +
5  gggaattccgcgtgtgcTAAGTACTTGGCGGGATAACGTCCAAGAGACGAGAGTCACGCACCTTAAACAGCTGTACTACGTGCCCCCGTAGGTgtccgttcgggatcctc    (2) +
25 gggaattccgcgtgtgcTTAACCTAATGGAATACAATTCTAACAGGAATAGGAGGGTAGCTCCAAGTACTAACGACCGTTCCGGTAGCCGGCTGGCTgtccgttcgggatcctc(1) +/-
24 gggaattccgcgtgtgcCGAATTGCCGTTGCACCGCATGTCAAATCATGCGGCGGTTACATGTCGACGGCTCGGAAAAAATCGGACAGCGCgtccgttcgggatcctc       (1)
36 gggaattccgcgtgtgcAGGGTTATCCAGAGCTGGGCATCGGGTACCCGCCCCAATGCTAATTGACCACTTGGGCGAGATCATATCTAGTCACAGTCgtccgttcgggatcctc(1)
15 gggaattccgcgtgtgcCACCAATCACATCTACCCGCGGTGTGGAATTGGCGTAAACGGGCAAGGCTATTTCAGGGTTATCAGAGAAATCCTTGCCAgtccgttcgggatcctc(1)
18 gggaattccgcgtgtgcAGGAACAGCCACGGTGTTCAAACTTGAACAACAGGATCGGTACAAGTGCTAAGACTCACGGTTCTTAGCTATCTAGCCCGgtccgttcgggatcctc(1)
19 gggaattccgcgtgtgcGAGGAGCAGGGTTTAAAGAGCTTAGGACCGCGGATATTAGGGTGAGTACGAGAAAACTCCTATTCCGACAGACTACCCGCgtccgttcgggatcctc(1)
38 gggaattccgcgtgtgcCACTACCGGCGTGGTAAGAAGAGGGCGCGAGCGATCGTGTACGGCTTAAAGAGTGGGGTCGCTCCATCATGGTGAAAGCgtccgttcgggatcctc (1)
```

FIGURE 2: Sequences of unique clones from the final round of the dopamine selection 1. Nucleotides in small letters represent the 5′ and 3′ fixed primer binding sequences. Clone identification numbers are given to the left of each sequence, and the number of identical clones is shown in parentheses on the right. Binding efficiency of the more represented aptamers was assayed: (++) 50−80%; (+) 20−50%; (+/−) <20% retention on the dopamine column. Clone size differences may be artifacts of cloning.

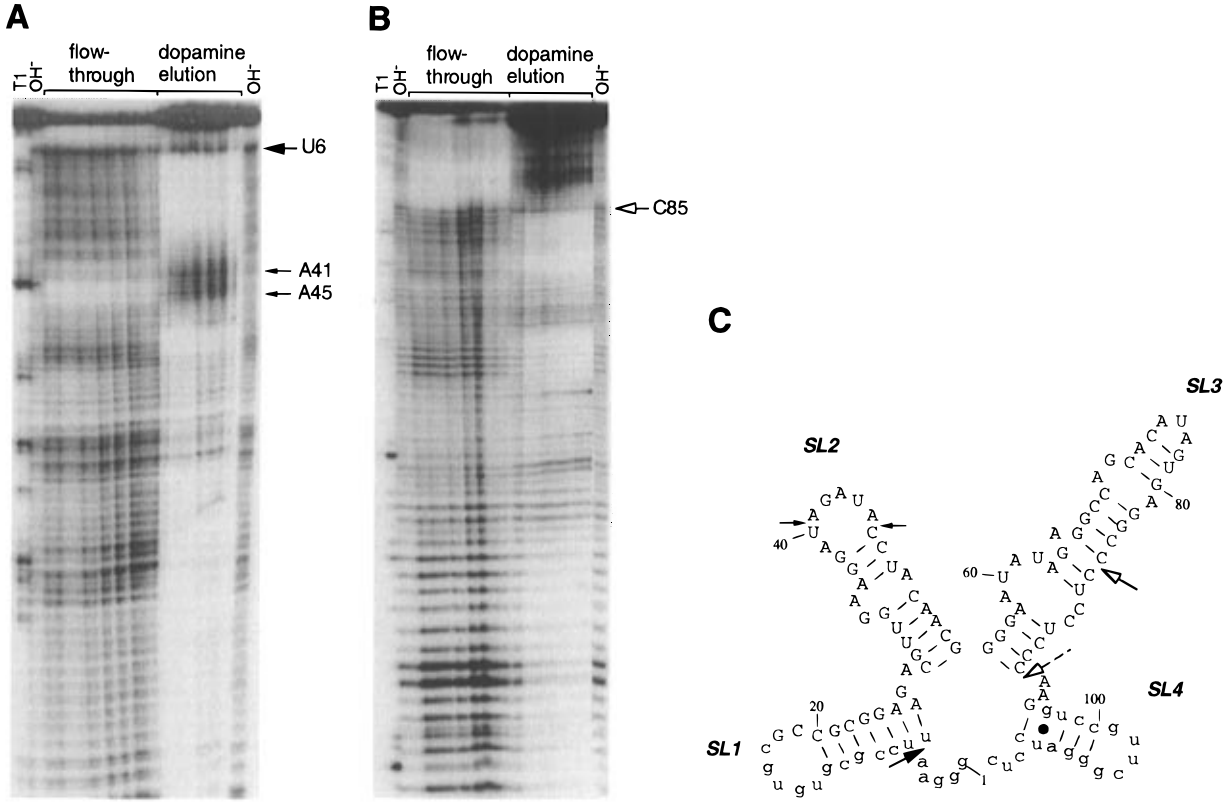


FIGURE 3: Minimal sequence requirement for the binding of the selected dopa2 RNA to dopamine-agarose. 3′-(A) or 5′-(B) ^{32}P end-labeled dopa2 RNA was partially digested under mild alkaline conditions. The digestion products were loaded onto a dopamine-agarose column and washed with 10 column volumns of selection buffer, and bound molecules were affinity eluted with 0.1 M dopamine. Aliquots of the fractions collected from the column were loaded onto a 8% polyacrylamide gel for autoradiography, along with alkaline hydrolysate (OH^-), and partial RNase T1 digest (T1). The positions of the boundaries from 3′-^{32}P-labeled (closed arrows) and 5′-^{32}P-labeled (open arrows) RNA are shown. (C) Predicted secondary structure of the dopa2 RNA. Stem and loop motifs (SL) are numbered, 5′ and 3′ fixed primer binding sequences are reported in small letters, 5′ and 3′ boundaries are indicated (closed and open arrows, respectively). Dashed open arrow indicates the 3′ end of dopa2/c.1 construct.

found. We think that RNAs with 5′ ends between positions 41 and 46 are able to form active complexes with the complete aptamers in the hydrolysis mixture. The best folding and binding activity (65% of the input RNA is affinity eluted) is obtained with an aptamer containing a full-length 5′ end and a 3′ end at position 93 of the dopa2 molecule (dopa2/c.1; 3′ end indicated in panel C by a dashed open arrow). The percentage of a single sequence RNA preparation applied to the affinity column that is specifically eluted might depend on the fraction of molecules able to fold in the active conformation for the binding to the dopamine-agarose. In fact, as previously discussed (Uhlenbeck, 1995) a sample of RNA molecules with identical sequences generally contains a mixture of conformations only a fraction of which retains the desired biochemical activity. As minimal aptamer for further studies we chose the dopa2/c.1 RNA, whose affinity to the dopamine in solution, as estimated by equilibrium filtration, is 1.6 ± 0.17 μM.

Specificity of Dopamine Binding. To study the specificity of the ligand−aptamer interaction the ability of various dopamine analogs to elute dopamine-agarose-bound dopa2/c.1 RNA was assessed (Figure 4). Substitutions at either the carbon atoms of the aliphatic chain (i.e. norepinephrine and L-dopa) decreased elution effectiveness as well as the complete deletion of the aliphatic chain (i.e. catechol). Methylation of the hydroxyl group at position 3 of the benzene ring (i.e. 3-*O*-methyldopamine) dramatically reduced the elution in 10 column volumes, and removal of the 3-hydroxyl (i.e. tyramine) or the 3- and 4-hydroxyls (i.e. phenethylamine) abolished it. Similarly, a compound lacking

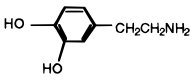
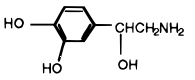
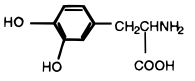
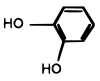
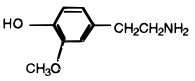
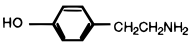
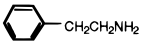
COMPOUND	STRUCTURE	RELATIVE ELUTION (% of dopamine elution)
DOPAMINE		100
NOREPINEPHRINE		58 ± 13
L-DOPA		30 ± 13
CATECHOL		23 ± 8
3-O-METHYLDOPAMINE		14 ± 8
TYRAMINE		2 ± 1
PHENETHYLAMINE		2 ± 1
ETHYLAMINE		6 ± 4

FIGURE 4: Ligand specificity. The ability of dopamine related molecules to elute bound dopa2/c.1 RNA from dopamine-agarose was studied by batch affinity chromatography. The relative ability of different eluants used at 1 mM concentration is expressed as a percentage of dopamine effectiveness. Negative control value (elution with no ligand in the buffer) has been subtracted from each eluant data. The relative elution values for each dopamine analog is an average of three independent experiments. Mean ± standard deviations are reported.

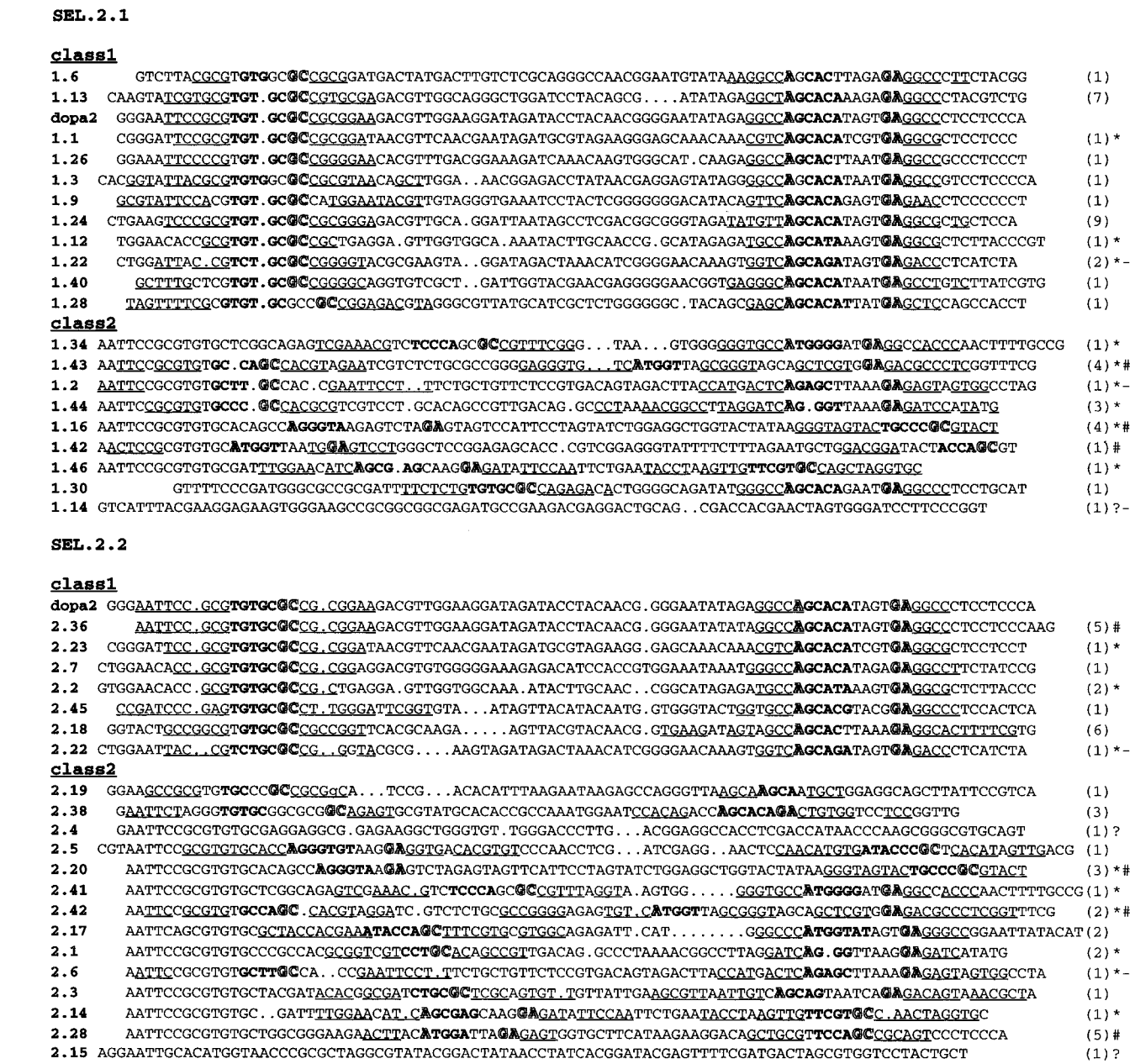


FIGURE 5: Consensus motifs within the sequences of the unique clones from the selections 2.1 and 2.2. Alignments of clone sequences isolated in the secondary selections 2.1 and 2.2 are shown. Only the 93 nt long selected regions are reported. The sequence of the parental dopa2 aptamer is included in the alignment, and the aptamers have been assigned to 2 classes, on the basis of sequence similarity with dopa2 RNA. Conserved structural motifs and invariable nucleotides are indicated in the figure (stem regions are underlined, potentially long-range base-pairing nucleotides are in bold, and invariant nucleotides are outlined). On the left side the clone identification numbers are indicated, and in parentheses on the right the number of clones with identical sequence are shown. Symbols reported on the right indicate (*) clones common to the selections 2.1 and 2.2; (#) clones common to secondary and primary selections; (−) clones with no binding activity; (?) and clones that did not fit the structural model.

the catechol ring (i.e. ethylamine) was unable to elute bound RNA. These results indicate that the hydroxyl at position 3 strongly contributes to the aptamer binding. In addition, the aliphatic chain might interact with the RNA, as indicated by the fact that the catechol ring by itself is 23% active compared to dopamine.

Secondary Selections. To further define the nucleotides of the dopa2 RNA that were critical for the dopamine-binding activity, we prepared a highly degenerate pool of sequences, partially randomizing the dopa2 sequence, including the original 5′-constant region. A synthetic oligonucleotide of a total length of 129 nt, consisting of 93 mutagenized bases flanked by a new primer binding site at the 5′ end, was used as template for PCR amplification and T7 RNA transcription. The bound RNAs were eluted with 1 mM dopamine. In order to study possible changes in the aptamer sequences as a function of a decrease of the buffer ionic strength, secondary selections were done at two different salt concentrations (0.15 M NaCl, 50 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$ for selection 2.1; 0.5 M NaCl, as in the original selection, for selection 2.2). After six rounds of selection and amplification the binding was restored to comparable levels for both salt concentrations (over 40% of the applied RNAs were affinity eluted from the columns by 1 mM dopamine). No further selection rounds were done in order to maintain as much variability as possible in the selected pools. The round 6 RNAs from both selections 2.1 and 2.2 were reverse transcribed, PCR amplified and cloned. 43 clones from selection 2.1 and 42 clones from selection 2.2 were sequenced. Figure 5 shows sequence alignment of

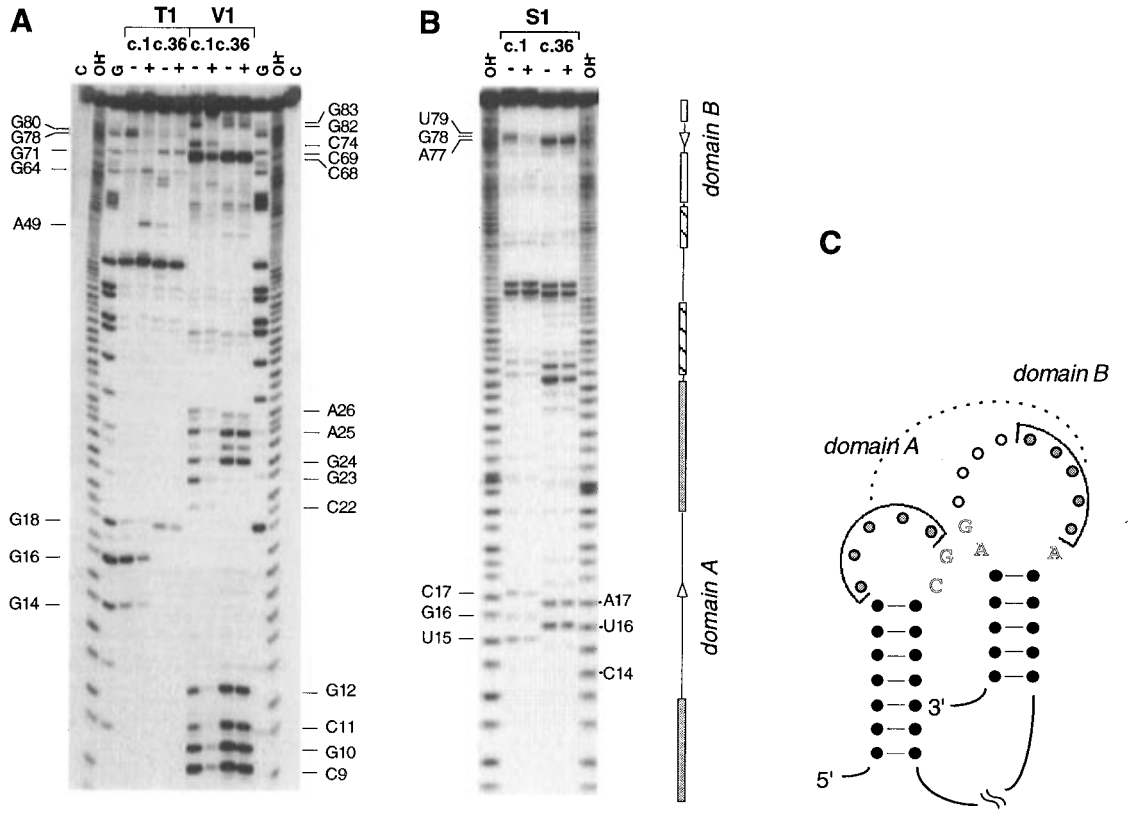


FIGURE 6: Enzymatic probing of the dopa2/c.1 and of the inactive mutant dopa2/c.36 RNAs. ^{32}P-labeled RNAs were partially digested with RNase T1, V1 (A), and S1 (B), in the presence (+) or in the absence (−) of 1 mM dopamine during nuclease digestion. Aliquots of enzymatic digestions were loaded onto a 14% polyacrylamide, 8 M urea gel along with undigested controls (C), alkaline hydrolysate (OH^-) and partial denaturing RNase T1 digest (G), of the dopa2/c.1 RNA (on the left side of the gels), or of dopa2/c.36 RNA (on the right side). Residues that are protected by nuclease digestion are indicated. Positions discussed in the text are also shown. The vertical bars to the right of the autoradiograms indicate the predicted positions of stems and loops. Boxes and lines represent stem and loop regions, respectively; arrows indicate the antiparallel orientation of complementary sequences in the domain A and B loops. Panel C shows a schematic representation of the dopamine cage. Loop interaction is indicated by a dashed line.

unique sequences from selections 2.1 and 2.2. No bias towards a specific primary or secondary structure as a function of salt concentration was found. In fact a number of sequences selected in high salt buffer were identical to sequences selected at lower salt concentration (indicated by an asterisk; as an example, clone 1.1 is 99% identical to clone 2.23). Sequences from both secondary selections, analyzed by a multiple sequence alignment program, were split into two classes, depending on the similarity with the wild-type dopa2 sequence: class 1 comprises molecules with a similarity to the parental sequence, class 2 sequences are heterogeneous and apparently unrelated to dopa2.

The sequence homology among class 1 clones is limited to two defined regions that correspond to the dopa2 sary structure motifs stem and loop 1 and stem and loop 3 (SL1 and SL3 in Figure 3). Hereafter we will refer to them as domain A and B, respectively. Underlined sequences can be folded into stem structures. A number of compensatory base changes within the potential double-stranded regions are consistent with the proposed secondary structures. Several nucleotides in the two loop domains are highly conserved (bold and outlined nucleotides): nucleotides in bold in the two domains were found to be complementary and in an antiparallel orientation. The fact that complementary nucleotides have been conserved in the two domains, as in the dopa2 molecule, is consistent with a tertiary interaction between the two regions through canonical or wobble base pairings. In addition, five nucleotides of domains A and B (outlined in the figure) are invariant in all class 1 sequences and in the same relative position: a GC dinucleotide at 3′ of the domain A loop, an A at the 5′, and a GA at the 3′ of the domain B loop.

A careful analysis of class 2 sequences showed that what is common to almost all class 2 aptamers and between class 1 and class 2 sequences is a higher-order structural motif. In fact, two loop domains containing a stretch of complementary nucleotides (from 3 to 6 variable nucleotides, indicated in bold in Figure 5) and 5 conserved residues (outlined in the figure) were found along the primary sequences of almost all active aptamers. RNA binding activity had been checked for each of the unique secondary selection clones: two unique sequences, that showed the consensus motifs, did not bind to the dopamine-agarose matrix. Only three of the secondary selection clones did not show the consensus motifs, one of them did not bind to dopamine agarose matrix.

Interestingly, some of the secondary selection clones are similar and in some cases almost identical to clones derived from selection 1 (for example, clones 1.43 and 2.42 are 97% identical to clone 23). A re-examination of the selection 1 aptamers showed that the two structural domains were even present in 12 out of the 14 unique sequences. In conclusion, the base pairing interaction between the two loops might be crucial for the presentation of the invariable nucleotides. The relative position of the consensus motifs can be inverted, as seen for example for clone 41 of selection 1 and clones 1.46 and 2.5 of secondary selections, where domain B is upstream domain A in the primary sequence. Another variable is the

relative distance separating consensus domains (see for example dopa2-like clones compared to clones 1.30 or 2.17).

Enzymatic Probing. The proposed secondary/tertiary structures of the dopamine aptamers were confirmed by nuclease mapping (Figure 6, panels A and B). Single-strand (ss, T1 and S1 nucleases)- and double-strand (ds, V1 nuclease)-specific enzymes were used to identify those positions of dopa2/c.1 RNA (see Figure 3, panel C) and of the derived inactive mutant dopa2/c.36, that are protected from enzyme digestion in the free RNA or in the presence of 1 mM dopamine. The c.1 and c.36 RNAs have identical sequences, except for three nucleotides in the domain A loop (the c.1 residues G14, G16, and C17 were changed to C14, U16, and A17 in the c.36 RNA). These base substitutions abolish binding to dopamine (no binding to the affinity matrix), probably by preventing the proposed base-pairing interaction between domains A and B.

Significant differences between c.1 e c.36 in the cleavage pattern of both ss- and ds-specific nucleases were found only in the regions of A and B domains. G14, G16, and G18 are protected from T1 G-specific nuclease in the presence of the ligand in the active aptamer. No protection was found for the c.36 G18 in the presence of dopamine. There are a few nonspecific cleavage products in the T1 hydrolysis pattern, like the bands corresponding to base A49 in both c.1 and c.36 RNAs, which might be due to a 2′-*O*-mediated hydrolysis at a pyrimidine-adenosine dinucleotide (Kierzek, 1992). The hydrolysis at G64 is enhanced in the c.1 RNA in the presence of dopamine, presumably because of a conformational constraint induced by ligand binding. Protection from T1 recognition, in the presence of the ligand, of domain B positions G71, G78, and G80 was seen for the active c.1 RNA but not for the inactive c.36. G71 is less accessible to the T1 nuclease in the c.1 than in the c.36 RNA (more evident after c.1 ligand binding), while G78 and G80 in the free c.1 RNA are highly recognized compared to the c.36 RNA. This particular behavior might be due to the double helical structure of the sequence 5′GCACA and to a consequent loop constrain in the downstream loop nucleotides.

The ds-specific V1 nuclease recognized both 5′ and 3′ helices at the domain A of either c.1 and c.36 free RNAs. After dopamine binding a strong protection from V1 recognition was seen only in the c.1 RNA. Whether the V1 protection depends on a distortion of the helix after ligand binding or on a quenching effect on the phosphate backbone by the positive charged dopamine enclosed in the binding pocket remains to be assessed. Furthermore, a V1 protection after dopamine binding, specific for c.1 but not for c.36, is present at the base-paired nucleotides C68 and C69 and at the corresponding G82 and G83. Interestingly, the presence of a band at nucleotide C74 in c.1, a band that is absent in c.36, is consistent with the double-helical conformation of the 5′GCACA domain B region.

Finally, the pattern of the ss-specific endonuclease S1 indicates a protection of c.1 domain B loop region upon ligand binding, at nucleotides A77, G78, and U79. No protection was found in the corresponding region of c.36 RNA. Protection upon ligand binding was found also in the domain A region (U15, G16, and C17), though at a lesser extent than in the domain B.

A schematic diagram of the proposed dopamine binding pocket is given in panel C of Figure 6.

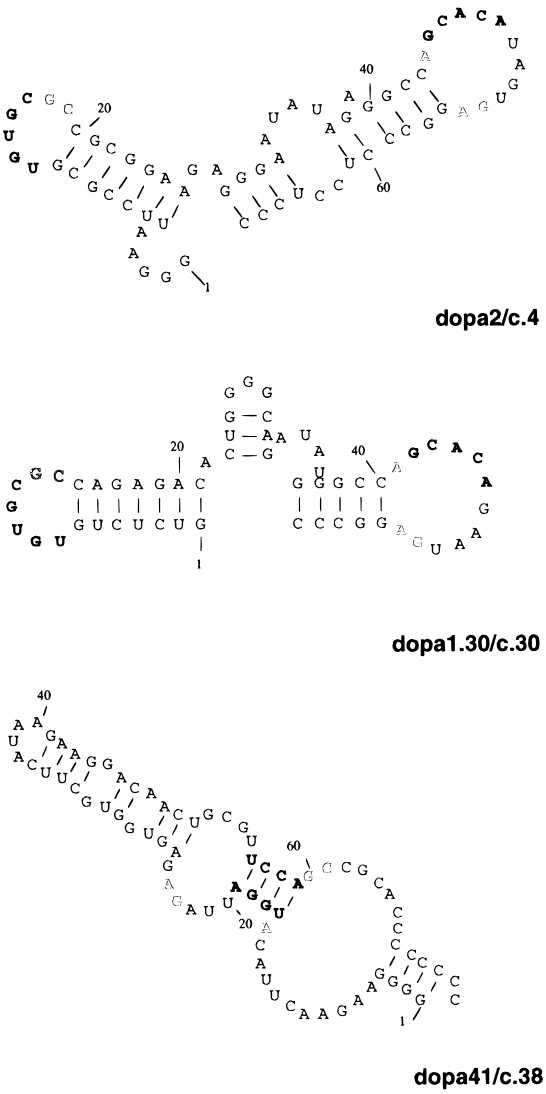


FIGURE 7: Minimal dopamine aptamers. The minimized secondary structures of dopa2/c.4, dopa1.30/c.30, and dopa41/c.38 RNA are reported. Highly conserved motifs are indicated in the structures (complementary and invariant nucleotides are indicated in bold and outlined, respectively).

Minimal Aptamers. To test whether the two conserved domains were necessary and sufficient for dopamine binding, a minimal RNA carrying only the two stem and loop motifs derived from the dopa2 molecule (SL1 and SL3) was synthesized (dopa2/c.4, Figure 7). The dopa2/c.4 RNA binds to dopamine agarose and 35% of the RNA loaded on the affinity column is specifically eluted by 1 mM dopamine in high salt buffer (0.5 M NaCl, 50 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$).

In some of the selected clones (clones 1.30, 1.34 and 2.4, Figure 5) domains A and B are close to each other and minimal active aptamers were made by removing sequences upstream and downstream of the defined secondary structure motifs. Figure 7 shows the MulFold secondary structure of the active construct dopa1.30/c.30 RNA (30 % of the input RNA is affinity eluted).

The secondary structure of construct 41/c.38, obtained from clone 41 after deletion of sequences to the 5′ and 3′ side of the conserved domains, is shown. Interestingly, the minimized c.38 secondary structure reproduces in the second dimension the proposed tertiary interaction between domains A and B, in which Watson−Crick base pairing between four complementary nucleotides (nucleotides in bold) juxtapose

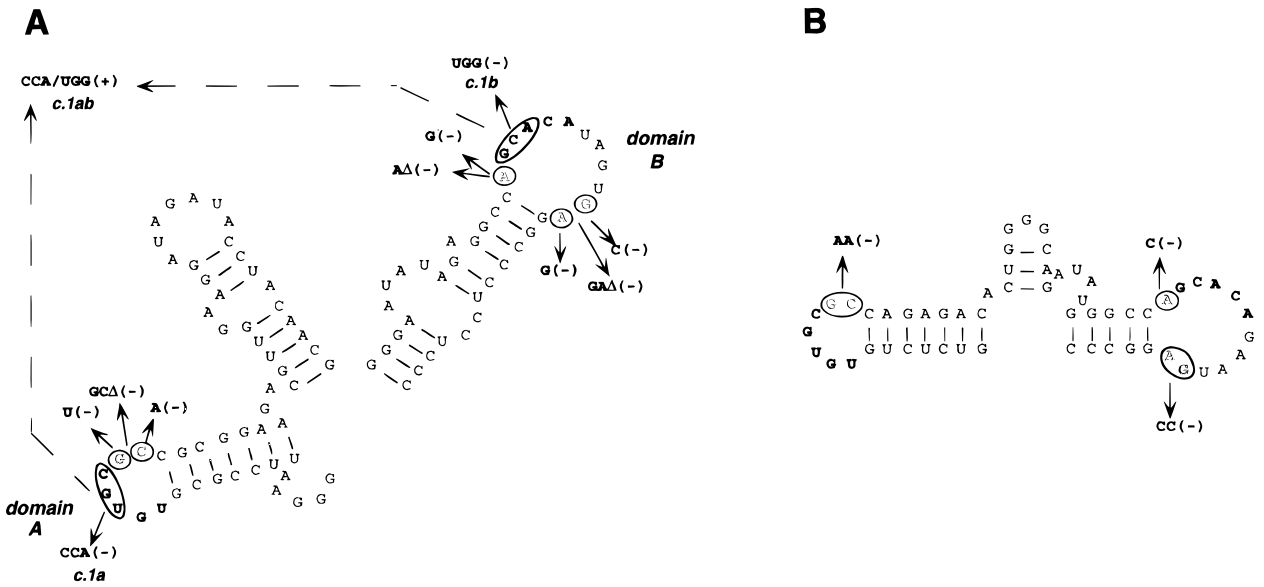


FIGURE 8: Mutational analysis of dopa2/c.1 and dopa1.30/c.30 RNAs. Predicted secondary structures of dopa2/c.1 (A) and dopa1.30/c.30 (B) are shown. Relevant nucleotides are indicated in bold and outlined. Base change and △-deletion (circled nucleotides) are indicated by arrows. The arrows illustrate base substitution. The results of mutation analysis are indicated (+, binding; −, no binding).

the five conserved residues (outlined nucleotides). 57% of c.38 RNA applied to the affinity matrix is eluted by 1 mM dopamine.

Aptamer Mutational Analysis. To verify that the proposed base-pairing interaction between the two loops is necessary for creating a dopamine binding pocket, in which recognition is mediated by the five invariant nucleotides, a mutational analysis on dopa2/c.1 and dopa1.30/c.30 RNA was carried out (Figure 8, panels A and B). Base substitutions or deletions were chosen so as not to interfere with the secondary structure of the molecules.

Mutants of dopa2/c.1 in which the base pairing between three of the five complementary nucleotides was disrupted were made: in mutant c.1a 5′UGC in domain A was changed to 5′CCA, in mutant c.1b 5′GCA in domain B was changed to 5′UGG. Neither of the two mutants bound to the affinity matrix. The pairwise compensatory mutation in the c.1ab double mutant restored binding to wild type levels (65% of the input RNA was affinity eluted from dopamine agarose by 1 mM dopamine).

Any point mutation or deletion of the five conserved nucleotides in both dopa2/c.1 and dopa1.30/c.30 abolished activity completely (no binding to the matrix).

DISCUSSION

In this paper we describe the isolation of RNA aptamers that specifically bind the monoaminic neurotransmitter dopamine, among the smallest organic molecules at which an *in vitro* selection experiment has been successfully aimed.

The K_d of the RNA−ligand binding to dopamine in solution, estimated by equilibrium filtration, is 1.6 μM. A broad range of affinities has been obtained by *in vitro* selection protocols with small molecule targets, from nanomolar values, as found for cyanocobalamine (Lorsch & Szostak, 1994) and theophylline (Jenison et al., 1994), to the higher millimolar values as for valine (Majerfeld & Yarus, 1994). The aptamer affinity might depend on the selection strategy utilized or on the stability of the ligand binding site. The high specificity of dopamine recognition is indicated by our study of binding activity of a number of related molecules: we hypothesize that the dopamine−aptamer interaction is mediated by an hydrogen bond involving the hydroxyl group at the position 3 and an acceptor in the RNA binding pocket. Both the benzene ring and the proximal region of the aliphatic chain might be embedded into the RNA binding pocket.

In contrast to the majority of selected aptamers that bind small molecules, a relatively large RNA sequence is needed to bind dopamine (the minimal aptamer is 57 nt long). The ligand binding site resides in a cleft between two helical regions likely to be stabilized by such base interactions as hydrogen bonds and base stacking. This structural framework might be required for the contact between the ligand and five invariant nucleotides, which we hypothesize are involved in the recognition of dopamine. Significantly, all point mutations of each of the conserved nucleotides abolish binding activity. The structure of the ligand binding site could be deduced only after partial randomization of the sequence that dominated the first selection final pool and reselecting for dopamine-binding RNAs. In fact no consensus sequence had been found among the 14 unique clones selected in the primary selection, implying that there were many "sequence solutions" for the binding of a relatively simple molecule such dopamine. From a careful analysis of sequence alignment of clones selected in the two ionic strength secondary selections, we deduced the minimal region in the parental sequence necessary for the binding to dopamine and conserved in the evolutionary selection.The highly conserved structural motif is shared by almost all aptamers from both primary and secondary selections. Our model for the structure of the dopamine binding pocket is based upon sequence conservation and analysis of compensatory mutations, confirmed by mutational analysis and enzymatic probing on minimal aptamers.

The protection from ss-specific nucleases in the loop regions of both domains A and B could be due to either the double-helical conformation of specific residues, stabilized by ligand binding (as G14 and G16 in the domain A), or the interaction of base functional groups with the ligand, presumably stacked into the binding pocket (as for G18 in the domain A). Tertiary interaction might be possible even

in the absence of the ligand, as shown by the presence of the ds-specific V1 nuclease site at position C74. Further studies on chemical binding interference will clarify the role of each of the conserved nucleotide.

No retention to the affinity matrix was found in the absence of [Mg^{2+}] in the column buffer (data not shown), indicating that the correct aptamer folding depends on the presence of [Mg^{2+}] ions. In addition, 1 mM spermine was necessary to stabilize the dopa2/c.1 RNA in the S1 cleavage buffer (see Materials and Methods). As previously discussed (Quigley et al., 1978) both [Mg^{2+}] and spermine stabilize secondary and tertiary structural features of tRNA in solution. We hypothesize that the described dopamine binding pocket can only be generated when stabilizing interactions, such as hydrogen bonds and base stacking in the presence of [Mg^{2+}] ions, are stronger than destabilizing interactions, such as electrostatic repulsion of the multiple charges on the loop phosphate backbone.

We do not know whether the isolated aptamers are the most tightly binding RNAs; our repeatedly isolated structures are probably the most frequent structures in our experimental protocol (depending for example on the choice of an 80 nt initial randomized region). It is possible that the functional groups (the five conserved nucleotides) that mediate dopamine recognition can be incorporated into a different stable framework; for instance, a minimal active aptamer derived from clone 36 from the selection 1 (from G15 to A61 of the clone 36 nucleotide sequence, Figure 2) can be folded into a pseudoknot-like structure wherein the 5′-loop (L1; Pleij et al., 1985; Pleij, 1990) is lengthened while the joining region (L2) is shortened to one nucleotide, and the five conserved bases reside in the loop regions. In addition the selection of an active aptamer with unknown structure (clone 2.15) suggests that completely different dopamine binding sites could be formed by RNA molecules.

In conclusion we propose that the described dopamine cage may be able to act as a stable scaffold for the embedding of a small molecule. As previously discussed (Uphoff et al., 1996), selection experiments can provide insights into RNA structure and can also be used to refine existing structures; we suggest that a structure like the one described for the dopamine aptamers could form by long-range interactions in natural RNA molecules.

Given that specificity of dopamine binding probably depends on five conserved nucleotides, one might be able to modify ligand specificity by changing the identity of the functional residues. Efforts in this direction are in progress.

ACKNOWLEDGMENT

We thank K. Williams for helpful discussions and E. Lund for critical reading of the manuscript. We also thank J. Amatruda for the help with the manuscript, G. Di Franco for technical assistance, and A. Sebastiano for secretarial assistance.

REFERENCES

Arnold, F. H., Schofield, S. A., & Blanch, H. W. (1986) *J. Chromatogr. 355*, 1−12.
Bartel, D. P., & Szostak, J. W.(1993) *Science 261*, 1411−1418.
Binkley, J., Allen, P., Brown, D. M., Green, L., Tuerk, C., & Gold, L. (1995) *Nucleic Acids Res. 23*, 3198−3205.
Burgstaller, P., & Famulok, M. (1994) *Angew. Chem., Int. Ed. Engl. 33*, 1084−1087.
Ciesiolka, J., & Yarus, M. (1996) *RNA 2*, 785−793.
Ciesiolka, J., Gorski, J., & Yarus, M. (1995) *RNA 1*, 538−550.
Connel, G. J., & Yarus, M. (1994) *Science 264*, 1137−1141.
Connel, G. J., Illangesekare, M., & Yarus, M. (1993) *Biochemistry 32*, 5497−5502.
Dieckmann, T., Suzuki, E., Nakamura, G. K., & Feigon, J. (1996) *RNA 2*, 628−640.
Ellington, A. D., & Szostak, J. W. (1990) *Nature 346*, 818−822.
Famulok, M. (1994) *J. Am. Chem. Soc. 116*, 1698−1706.
Famulok, M., & Szostak, J. W. (1992) *J. Am. Chem. Soc. 114*, 3990−3991.
Famulok, M., & Hüttenhofer, A. (1996) *Biochemistry 35*, 4265−4270.
Fan, P., Suri, A. K., Fiala, R., Live, D., & Patel, D. J. (1996) *J. Mol. Biol. 258*, 480−500.
Feng, D. F., & Doolittle, R. F. (1987) *J. Mol. Evol. 35*, 351−360.
Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A., & Famulok, M. (1996) *Nucleic Acids Res. 24*, 1029−1036.
Gold, L. (1995) *J. Biol. Chem. 270*, 13581−13584.
Gold, L., Polisky, B., Uhlenbeck, O., & Yarus, M. (1995) *Annu. Rev. Biochem. 64*, 763−797.
Green, R., Ellington, A. D., Bartel, D. P., & Szostak, J. W. (1991) *Methods Enzymol. 2*, 75−86.
Jaeger, J. A., Turner, D. H., & Zuker, M. (1989) *Methods Enzymol. 183*, 281−306.
Jellinek, D., Lynott, C. K., Rifkin, D. B., & Janjić, N. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11227−11231.
Jenison, R. D., Gill, S. C., Pardi, A., & Polisky, B. (1994) *Science 263*, 1425−1429.
Jiang, F., Kumar, R. A., Jones, R. A., & Patel, D. J. (1996) *Nature 382*, 183−186.
Kierzek, R. (1992) *Nucleic Acids Res. 20*, 5079−5084.
Lorsch, J. R., & Szostak, J. W. (1994) *Biochemistry 33*, 973−982.
Majerfeld, I., & Yarus, M. (1994) *Nat. Struct. Biol. 1*, 287−292.
Morris, K. N., Tarasow, T. M., Julin, C. M., Simons, S. L., Hilvert, D., & Gold, L. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 13028−13032.
Nestler, E. J. (1994) *Cell 79*, 923−926.
Nieuwlandt, D., Wecker, M., & Gold, L. (1995) *Biochemistry 34*, 5651−5659.
Pan, T., & Uhlenbeck, O. C. (1992) *Biochemistry 31*, 3887−3895.
Pleij, C. W. A. (1990) *Trends Biochem. Sci. 15*, 143−147.
Pleij, C. W. A., Rietveld, K., & Bosch, L. (1985) *Nucleic Acids Res. 13*, 1717−1731.
Quingley, G. J., Teeter, M. M., & Rich, A. (1978) *Biochemistry 75*, 64−68.
Sassanfar, M., & Szostak, J. W. (1993) *Nature 364*, 550−553.
Tuerk, C., & Gold, L. (1990) *Science 249*, 505−510.
Uhlenbeck, O. C. (1995) *RNA 1*, 4−6.
Uphoff, K. W., Bell, S. D., & Ellington, A. D. (1996) *Curr. Opin. Struct. Biol. 6*, 281−288.
Wang, Y., Killian, J., Hamasaki, K., & Rando, R. R. (1996) *Biochemistry 35*, 12338−12346.
Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E., & Famulok, M. (1996) *Science 272*, 1343−1347.

BI9700633