In Vitro Selection of RNA Ligands to Substance P†

Dan Nieuwlandt,* Matt Wecker, and Larry Gold

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309

Received November 22, 1994; Revised Manuscript Received February 21, 1995®

ABSTRACT: RNA ligands to the tachykinin substance P have been selected from a large pool of random sequence RNA molecules. Substance P is an undecapeptide that plays a variety of roles as a neurotransmitter and neuromodulator in the central and peripheral nervous system of mammals. A systematic evolution of ligands by exponential enrichment (SELEX) procedure was used to isolate RNAs that bind substance P immobilized on a solid support. RNAs that also bind substance P in solution were identified, and the tightest binder was subjected to mutagenesis in a second SELEX procedure to evolve ligands with a higher affinity for the peptide. A comparative analysis of 36 ligands isolated from the second SELEX experiment revealed two main sequence classes with highly conserved secondary structures within each class. Dissociation constants for the interaction of these ligands with substance P in solution were determined by equilibrium dialysis. The amino acid residues involved in the interaction with the highest affinity ligand (190 nM K_d) were mapped by determining which of a set of overlapping fragments of substance P can compete with the intact peptide for binding. A binding competition experiment also demonstrated the ability of the same ligand to discriminate between substance P and the reverse orientation of the same amino acid sequence. The results from this study demonstrate that SELEX can yield high affinity RNA ligands to small nonconstrained peptides.

A recently developed methodology for the Systematic Evolution of Ligands by EXponential enrichment (SELEX) has permitted the isolation of nucleic acids that bind target molecules with high affinity and specificity. SELEX (Tuerk & Gold, 1990; Irvine et al., 1991) involves the affinity selection of target binders present among a large pool of random sequence nucleic acid molecules followed by amplification of selected species by the polymerase chain reaction to yield an enriched pool of nucleic acids for the next cycle of selection and amplification. These cycles are repeated until nucleic acids with the highest affinity for the target have evolved from the population. Reported applications include the selection of nucleic acid ligands to nucleic acid binding proteins (Tuerk & Gold, 1990; Bartel et al., 1991; Schneider et al., 1992; Tuerk et al., 1992), non-nucleic acid binding proteins (Gold et al., 1993; Bock et al., 1992; Jellinek et al., 1993), organic dyes (Ellington & Szostak, 1990, 1992), ATP (Sassanfar & Szostak, 1993), cyanocobalamin (Lorsch & Szostak, 1994), and theophylline (Jenison et al., 1994).

We have been interested in isolating ligands that bind tightly and specifically to small bioactive peptides. In addition to providing a relatively simple system for the study of RNA−protein interactions, such ligands would have potential as diagnostic and pharmacological tools. The feasibility of isolating ligands with high affinity for small linear peptides was uncertain, however, since they have intrinsically flexible structures and usually exist in solution as an equilibrium of multiple conformers (Rizo & Gierasch, 1992). Binding affinities may be limited by the conformational entropy lost upon binding a flexible peptide. The pharmacologically important peptide substance P (SP) was chosen as the target in this study to investigate the possibility of isolating RNA ligands to a nonconstrained peptide.

SP is an 11 amino acid peptide that belongs to the tachykinin family of neuropeptides. Known mammalian tachykinins (neurokinins) include neurokinin A, neurokinin B, neuropeptide K, and neuropeptide γ. All tachykinins share the carboxy-terminal sequence Phe-X-Gly-Leu-Met-NH_2. The mammalian tachykinins are produced by neurons in the central and peripheral nervous system where they are predominantly localized in the nerve terminals (Escher & Regoli, 1989). Neurotransmitter and neuromodulator functions of SP include peripheral vasodilation, smooth muscle contraction, pain transmission (nociception), stimulation of exocrine secretions, and immunomodulation [for a review, see Escher and Regoli (1989)]. Recent studies also suggest that SP has a role in angiogenesis and may be a contributor to angiogenic diseases such as rheumatoid arthritis, atherosclerosis, diabetic retinopathy, and cancer (Fan et al., 1993). There is also evidence that SP has memory-modulating and reinforcing effects, and Huston et al. (1993) have suggested a possible link between these roles of SP and the impairment in associative functioning accompanying Alzheimer's disease.

In this report, we describe the isolation of RNA ligands with high affinity for SP from a large pool of random sequence RNA molecules. Preliminary investigations into the interaction of one of these ligands with SP demonstrated that multiple amino acid residues are involved in binding and that a peptide with the same amino acid sequence in the opposite orientation (rSP) is a poor competitor for binding.

EXPERIMENTAL PROCEDURES

Materials. Synthetic single-stranded DNAs (ssDNA) were obtained from Operon (Alameda, CA). Cys-SP and Cys-

† This work was supported by grants to Larry Gold from the National Institutes of Health (GM-28685 and GM-19963) and a Research Fellowship Award to Dan Nieuwlandt from the National Institutes of Health (1 F32 GM16005-01).

* Present address and to whom correspondence should be sent: Nexstar, Inc., 2860 Wilderness Place, Boulder, CO 80301. Telephone: (303) 546-7694. Fax: (303) 444-0672.

® Abstract published in *Advance ACS Abstracts,* April 1, 1995.

0006-2960/95/0434-5651$09.00/0

OLIGONUCLEOTIDES

Starting ssDNA template (experiment A):

5'- GCCGGATCCGGGCCTCATGTCGAA [-60N-] TTGAGCGTTTATTCTGAGCTCCC -3'

5' PCR Primer:

Hind III

5'- CCGAAGCTTAATACGACTCACTATAGGGAGCTCAGAATAAACGCTCAA -3'

T7 Promoter

3' PCR Primer:

Bam HI

5'- GCCGGATCCGGGCCTCATGTCGAA -3'

Starting RNA (experiment A):

5'- GGGAGCUCAGAAUAAACGCUCAA [-60N-] UUCGACAUGAGGCCCGGAUCCGGC -3'

Starting RNA (experiment B, ligand A13):

5'- GGGAGCUCAGAAUAAACGCUCAAGGGCAACGCGGGCACCCCGACAGGUGCAAAAA CGCACCGACGCCCGGCCGAAGAAGGGGAUUCGACAUGAGGCCCGGAUCCGGC -3'

PEPTIDES

SP:

N-Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH_2

Cys-SP:

Ac-Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-Cys-C

Cys-rSP:

N-Cys-Met-Leu-Gly-Phe-Phe-Gln-Gln-Pro-Lys-Pro-Arg-C

FIGURE 1: Starting ssDNA template and RNAs, PCR primers, and peptides used in this study. For Cys-SP, "Ac" indicates that the peptide was synthesized with an acetylated N-terminus.

rSP were synthesized and purified by Macromolecular Resources (Fort Collins, CO). Thiopropyl-activated Sepharose 6B, SP, and all SP fragments with the exception of SP 1−6 were purchased from Sigma. SP 1−6 was purchased from Peninsula Laboratories, Inc. (Belmont, CA). [2-L-Prolyl-3,4-^{3}H(N)]-SP and all radionucleotides were obtained from NEN Research Products (Dupont; Wilmington, DE). Enzymes were purchased from commercial sources.

Immobilizing SP on Thiopropyl-Activated Sepharose. Cys-SP (Figure 1) was covalently coupled (disulfide bond) to thiopropyl-activated Sepharose 6B through an interaction of the Cys-thiol group of peptides with hydroxypropyl-2-pyridyl disulfide ligands of the matrix. The coupling reaction consisted of 1.4 mg of Cys-SP and 1 g of preswollen thiopropyl-Sepharose in a final volume of 7 mL in the following buffer: 500 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 8.4. The suspension was gently mixed for 1 h at 4 °C and then transferred to a chromatography column and washed with 30 mL of 100 mM sodium acetate, pH 6.0. Remaining hydroxypropyl-2-pyridyl disulfide ligands were reacted by suspending the matrix in 7 mL of 100 mM sodium acetate plus 5 mM β-mercaptoethanol, followed by gentle mixing at room temperature for 45 min. The matrix was washed in SP binding buffer (10 mM HEPES, pH 7.0, 150 mM NaCl, 5 mM KCl, 5 mM $CaCl_2$) and the concentration of bound peptides was quantitated by ninhydrin assays. Storage was in SP binding buffer with 0.1% sodium azide at 4 °C. Thiopropyl Sepharose 6B utilized for counterselection and the dilution of SP-Sepharose was prepared as described above except for the omission of SP.

Random Sequence RNA Pool. Template DNA for the initial random sequence RNA population was generated from a synthetic random sequence ssDNA pool. The ssDNAs contained 60 nt of contiguous random sequence flanked by defined 5′ and 3′ ends that permit primer hybridization (Figure 1). Double-stranded DNA (dsDNA) molecules, synthesized initially by Klenow enzyme, and subsequently (following cycles of selection) by *Taq* DNA polymerase, have a T7 RNA polymerase promoter at the 5′ end. Klenow reaction products were purified on an 8% polyacrylamide gel and *in vitro* transcription of 500 pmol (~3×10^{14} unique sequences) of the resulting dsDNA template yielded the initial pool of uniformly [α-^{32}P]GTP-labeled 107-nt random sequence RNAs.

Table 1: Summary of SELEX Experiment A Selection Cycle Results[a]

selection cycle	pmol of RNA[a]	counterselection column?[b]	percent eluted with DTT[c]
1[d]	500	N	0.47
1	500	N	0.35
2	200	N	1.92
3	50	N	0.77
4	50	N	0.61
5	50	Y	2.57
6	50	Y	1.52
7	50	Y	2.27
8	50	Y	13.8
9	50	Y	19.8
10	50	Y	17.5
11	50	Y	19.4
12	50	Y	12.0

[a] Pmol of RNA applied to a 100-μL column volume of SP-Sepharose (80 μM SP). [b] Whether a thiopropyl-Sepharose column counterselection step was incorporated into the selection scheme prior to affinity selection on SP-Sepharose is indicated by either N (no) or Y (yes). [c] Percentage of the applied RNA that was eluted with 10 column volumes of SP buffer containing 100 mM DTT following an extensive wash with 20 column volumes of SP buffer. [d] Results from a control experiment in which the initial random sequence RNA pool was applied to a thiopropyl-Sepharose column in the absence of substance P.

SELEX Experiment A. Uniformly ^{32}P-labeled RNAs were suspended in 25 μl of SP binding buffer, heated at 70 °C for 5 min, and then cooled to room temperature. The quantity of RNA used for each selection cycle is indicated in Table 1. The RNA suspension was applied to a 100-μL SP-Sepharose column (80 μM SP) at room temperature, followed by 10 200-μL SP binding buffer wash volumes. Peptide-bound RNAs were then recovered with five 200-μL volumes of binding buffer containing 100 mM dithiothreitol (DTT). DTT reduces the linker disulfide bond resulting in the release of peptide from the matrix. The DTT eluate was extracted once with phenol and the RNAs were recovered by ethanol precipitation with 20 μg of yeast tRNA as carrier. Reverse transcription, PCR amplification, and T7 RNA polymerase transcription were performed essentially as described elsewhere (Tuerk & Gold, 1990). Transcription of PCR products yielded the RNA pool for the next cycle of selection and amplification.

SELEX Experiment B. Random mutagenesis (by the PCR mutagenesis procedure described below) of the highest affinity ligand from experiment A (A13) provided the template DNA used to initiate experiment B. Uniformly ^{32}P-labeled RNAs (200 pmol for each selection cycle) were suspended in 400 μL of 188 mM NaCl, denatured by heating at 90 °C for 90 s, and then quick-cooled on ice. After the addition of 100 μL of 5× SP binding buffer (minus NaCl), the RNA was combined with a 100-μL column volume of SP-Sepharose suspended in 400 μL of SP binding buffer. SP concentrations (as a function of column volume) used for each selection cycle are listed in Table 2. The 1 mL suspension was mixed on a rocking platform at room temperature for 30 min. For cycles in which a counterselection column (no peptide) preceded the SP-Sepharose column in the selection scheme (indicated in Table 2), an identical binding procedure was followed with the coun-

Table 2: Summary of SELEX Experiment B Selection Cycle Results[a]

selection cycle	[SP][b] (mM)	counterselection column?[c]	percent eluted with DTT[d]
ligand A13[e]	80	N	9.97
1	80	N	1.01
2	80	N	0.92
3	20	N	1.87
4	20	N	5.52
5	20	N	3.85
6	20	N	4.46
7	20	N	29.2
8	5.0	Y	16.3
9	2.5	Y	10.5
10	1.3	Y	8.0
11	1.3	Y	12.6
12	1.3	Y	14.4

[a] PCR mutagenesis of the ligand pool (of ligand A13 for selection cycle 1) preceded selection cycles 1−6; standard PCR amplification of the ligand pool preceded selection cycles 7−12. [b] Concentration of substance P contained on the SP-Sepharose affinity matrix as a function of column volume (100 μL). Binding reactions occurred in well-mixed 1-mL suspensions. [c] Whether a thiopropyl-Sepharose column counterselection step was incorporated into the selection scheme prior to affinity selection on SP-Sepharose is indicated by either N (no) or Y (yes). [d] Percentage of the applied RNA that was eluted with 10 column volumes of SP binding buffer containing 100 mM DTT following an extensive wash with greater than 20 column volumes of SP binding buffer. [e] Results from a control experiment with nonmutagenized ligand A13.

terselection matrix. Bound RNAs were pelleted with the counterselection matrix by centrifugation (~1000*g*, 5 s), and a 100-μL column volume of SP-Sepharose was suspended in the supernatant. Following the 30 min incubation, the SP-Sepharose was pelleted as above and the supernatant was removed. The matrix was resuspended in 400 μL of binding buffer and transferred to a syringe column (shortened 1-mL syringe with a small quantity of glass wool at the bottom). The flow-through volume was collected, and the column was washed with an additional 10 200-μL volumes of binding buffer. Peptide-bound RNAs were eluted with five 200-μL volumes of binding buffer containing 100 mM DTT. RNAs were recovered as in experiment A and amplified. PCR amplification following selection cycles 1−5 were by the PCR mutagenesis procedure described below. Standard PCR amplification followed selection cycles 6−12.

PCR Mutagenesis. Essential features of a modified PCR procedure for the introduction of random point mutations into DNA have been described elsewhere (Leung et al., 1989; Cadwell & Joyce, 1992; Bartel & Szostak, 1993). In experiment B of this study, the reaction mixture for selection cycle one consisted of 5 pmol of ligand A13 dsDNA (for selection cycles 2−6, all of the cDNA recovered from the prior selection cycle served as template), 1 mM dCTP, 1 mM dTTP, 0.2 mM dGTP, 0.2 mM dATP, 7 mM $MgCl_2$, 0.55 mM $MnCl_2$, 100 pmol of each primer, and 3 units of *Taq* polymerase (Promega) in 100 μL of *Taq* buffer (supplied with enzyme). The mixture was subjected to three PCR cycles: each cycle was 93 °C, 45 s; 50 °C, 45 s; and 72 °C, 2 min. A 13-μL volume of the reaction was transferred to a new tube and brought up to 100 μL with fresh reaction buffer (same composition minus template) followed by an additional three cycles of PCR mutagenesis. This dilution-PCR mutagenesis procedure was repeated nine more times for a total of 11 three-cycle reactions (33 doublings). Products from the eleventh three-cycle reaction were amplified further by diluting 13 μL of the reaction in 87 μL of a standard PCR reaction mixture (0.4 mM dNTPs, 3.75 mM $MgCl_2$, 250 pmol of each primer, and 5 units of *Taq* polymerase in 1× *Taq* buffer) followed by eight PCR cycles. Products from the final reaction were used as template for T7 RNA polymerase transcription. To analyze the mutation results, PCR products from the final amplification of PCR-mutagenized ligand A13 (for selection cycle 1) were cloned and sequenced.

DNA Sequencing. PCR products were cloned into the *Hin*dIII and *Bam*HI restriction sites of pUC18 and sequenced by the dideoxynucleotide termination method using modified T7 DNA polymerase (Sequenase 2.0; United States Biochemical) and universal forward and reverse primers.

Equilibrium Dialysis. Equilibrium dialysis experiments were performed with a Spectra/Por equilibrium dialyzer, Spectra/Por microcell dialysis chambers (200-μL chamber volumes), and 12 000−14 000 MWCO dialysis membranes (Spectrum; Houston, TX). These membranes allow passage of SP but not 107-nt RNAs. Sample volumes were 200 μL on each side of the membrane (i.e., 200 μL per chamber). All experiments were at room temperature in 1× SP binding buffer. RNAs (unlabeled) were denatured prior to re-equilibration in binding buffer as described above for experiment B. For binding curves (Figure 4), RNA concentrations between 50 nM and 20 μM were utilized. Single-point K_d estimates for individual experiment B ligands were performed with 200 nM or 1 μM RNA. The concentration of [^{3}H]SP applied to one side of the membrane (P_o; this symbol is used throughout this report to denote the peptide concentration as applied to one side of the membrane) was 20 nM for binding curves and single-point K_d estimates. For binding competition experiments, equal concentrations (Figures 6 and 7) of the competing peptide were applied to each side of the membrane at the start of dialysis. The dialysis cells were rotated at 10 rpm to shorten the time required to reach equilibrium. Equilibrium dialysis initiated with [^{3}H]-SP on one side of the membrane and in the absence of RNA was performed to determine the length of time required to attain equilibrium (~2.5 h for Spectra/Por-2 membranes); dialysis times for binding measurements were at least one hour longer than this determined time. Following equilibration, samples were withdrawn from the dialysis chambers, added to scintillation fluid, and counted. Equilibrium dissociation constants (K_d) were defined by the following equation (Rosen et al., 1980):

$$K_d = [R_f][P_f]/[P_b]$$

where $[R_f]$ is the concentration of unbound RNA (i.e., total RNA minus $[P_b]$), $[P_f]$ is the concentration of unbound peptide (concentration on the side of the membrane that does not contain RNA), and $[P_b]$ is the concentration of peptide bound to RNA ([P] on side of membrane that contains RNA minus $[P_f]$). Donnan effects were neglected in K_d measurements because it was assumed that this problem would be overcome by the high NaCl concentration (Karush & Sonnenberg, 1949). The fraction of SP bound was calculated by dividing $[P_b]$ by $[P_t]$, where $[P_t]$ is the total peptide concentration in the dialysis chamber containing RNA.

RESULTS

RNA ligands with affinity for SP were isolated in SELEX experiment A by selecting for RNAs present in a random

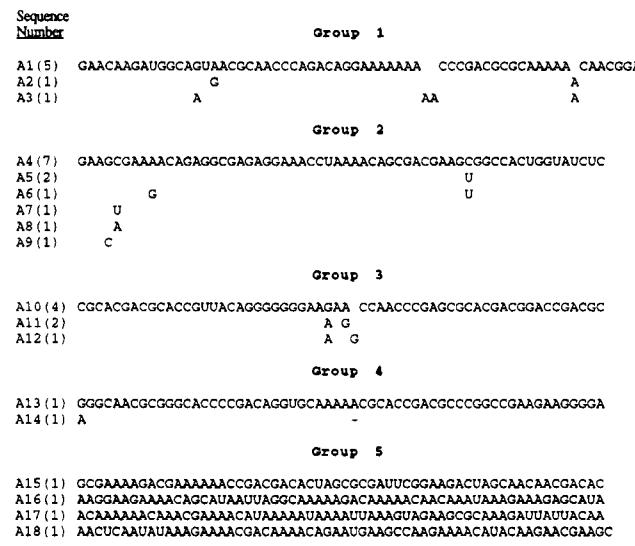

FIGURE 2: Sequences from the 60N regions only of RNA ligands selected in SELEX experiment A. Sequence numbers are preceeded by the letter "A" to designate their selection in experiment A. The number of identical sequences among the 33 clones analyzed is indicated in parentheses next to the sequence number. Groups 1−4 each represent a single "parental" sequence with varients resulting from point mutations presumably introduced by the polymerases used in the SELEX protocol. Sequences without close relatives among the sequenced population were placed in group 5. For sequences listed below the first sequences in groups 1−4, only differences from the first sequence are shown. Nucleotides listed below gaps in the first sequence indicate insertions. The hyphen in the ligand A14 sequence denotes the deletion of this nucleotide.

sequence pool that bind SP immobilized on a solid support. This was followed by mutagenesis of the highest affinity ligand and continued affinity selection in SELEX experiment B. The initial random sequence RNA population for experiment A consisted of approximately 3×10^{14} different molecules, each with 60 nt of contiguous random sequence. RNA 5′ and 3′ defined ends and their complimentary primer sequences are shown in Figure 1. The dissociation constant of the unselected random sequence RNA pool was roughly estimated at 1.2 mM as indicated by an equilibrium dialysis experiment with 20 μM RNA and 20 nM (P_o) of [^{3}H]SP where the fraction of SP bound was 0.017. The constraints of the dialysis system prohibited the use of significantly higher concentrations of RNA in the analysis. For the first selection cycle, 500 pmol of RNA was used; this quantity was reduced as the copy number of individual species increased (200 pmol in cycle 2 and 50 pmol for the remaining cycles). RNA pools subjected to selection cycles 5−12 were first counterselected on thiopropyl-Sepharose 6B (see Experimental Procedures) to remove RNAs with affinity for Sepharose or the linker arm. A significant increase in binding to SP was observed after the seventh selection cycle, with only a moderate improvement in subsequent cycles (Table 1). AMV reverse transcriptase sequencing of RNA pools showed a significant decrease in sequence randomness following selection cycles 7−12 (data not shown). No additional change in the sequence pattern was observed for RNA pools generated from selection cycles 10−12, suggesting that no further enrichment was occurring under these conditions.

PCR products from the twelfth cycle of selection and amplification were cloned and sequenced. Of the 33 clones sequenced, 18 are unique (Figure 2). However, 10 of the unique sequences almost certainly resulted from point mutations within a selected sequence during the amplification or cloning procedures (see sequence groups 1−4, Figure 2). Representatives of each sequence group were analyzed by equilibrium dialysis for their ability to bind SP in solution. Only group 3 and group 4 ligands demonstrated binding to free SP at RNA and peptide concentrations of 4 and 2 μM (P_o), respectively. Ligand A13, with an estimated K_d of 14 μM under these conditions, exhibited the highest affinity for SP in solution. The K_d for ligand A13 was subsequently more accurately estimated at 5.8 μM with a five-point binding curve (Figure 4), an improvement in affinity of about 200-fold over the initial unselected random sequence RNA pool. Sequence groups which did not exhibit binding to free SP under these conditions were presumably selected for their affinity to an SP conformation that is more prevalent when the peptide is coupled to the matrix or to an epitope that includes portions of both the peptide and the linker arm. Alternatively, ligands with affinity for the linker arm alone may have escaped the counterselection process resulting in their subsequent elution with DTT.

Higher affinity ligands to SP were produced and isolated in experiment B. Using ligand A13 dsDNA as the starting template source, an additional 12 cycles of selection and amplification were performed with PCR mutagenesis preceding the first six selection cycles. PCR products from the initial mutagenic PCR amplification of ligand A13 (RNA produced from this template pool was used for the first selection cycle) were cloned and sequenced to investigate the mutagenesis procedure. Ninety-six point mutations were identified within the 60N regions of 25 clones sequenced, representing a mutation rate of 0.064 per nucleotide position. At this rate, an average of 3.8 point mutations was expected per RNA per each of the first six selection cycles. Transition and transversion frequencies were equal (48 of each), and the mutations appeared to be randomly distributed throughout the sequence space. However, as observed by Bartel and Szostak (1993), there was a bias in the types of mutations induced. With the nonmutated 60N region of clone A13 having a nucleotide representation of 22 G, 20 C, 17 A, and 1 U, the following point mutations were identified: A·T to T·A (30), A·T to G·C (22), C·G to T·A (14), G·C to A·T (10), C·G to A·T (8), G·C to T·A (4), G·C to C·G (2), A·T to C·G (2), T·A to C·G (2), T·A to A·T (2), T·A to G·C (0), and C·G to G·C (0).

Selective pressure for the tightest binders was also increased in experiment B by reducing the peptide concentration following cycles in which binding had significantly increased (Table 2). The matrix-coupled SP concentrations were effectively lower than those listed in Table 2 (given as a function of column volume; 100 μL), since binding occurred in a well-mixed 1-mL suspension. For selection cycles 8−12, a counterselection thiopropyl Sepharose matrix was used prior to SP affinity selection. An RNA pool generated from cycle 12 PCR products exhibited a dissociation constant of 0.80 μM for SP in an equilibrium dialysis binding experiment (Figure 4), a 7-fold improvement in binding over ligand A13. PCR products from the twelfth selection cycle were cloned and sequenced. Of the 33 clones sequenced, all were unique. However, all but three of the sequences can clearly be placed into two major sequence classes (Figure 3).

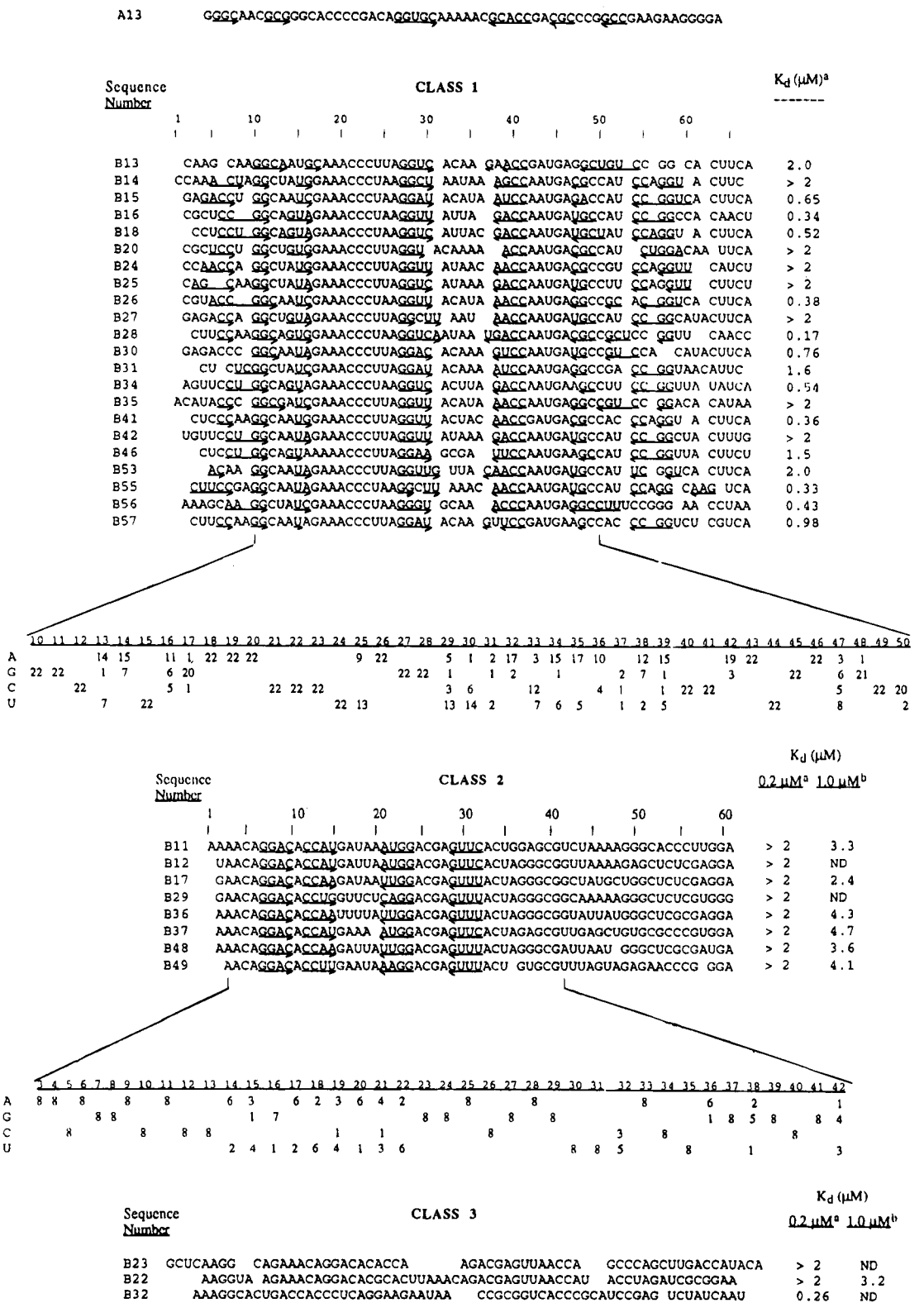

FIGURE 3: Alignments of sequences from the 60N (selected) regions of experiment B ligands. Ligand A13, the parental ligand used as the starting material for experiment B, is shown above the alignment for comparison. The ligands have been assigned to three classes on the basis of sequence and secondary structure similarities. Sequence numbers are preceeded by the letter B to designate their selection in experiment B. Nucleotide positions are numbered consecutively from a starting position dictated by aligned consensus sequences. Gaps in the sequences represent the absence of nucleotides at those numbered positions. K_d measurements from single-point equilibrium dialysis binding experiments are shown to the right of the individual ligands. Binding reactions consisted of 20 nM [^{3}H]SP (P_o) and either 0.2 μM RNA (a) or 1.0 μM RNA (b). ND = not determined. Nucleotide frequencies at the indicated positions are shown below the class 1 and 2 alignments.

All 33 clones were screened for their ability to bind SP in solution by single-point equilibrium dialysis measurements (Figure 3). Ligands assigned to class 1 generally have the highest affinity for SP in solution, with ligand B28 having the lowest K_d (measured at 170 nM in this screen). A five-point binding curve was subsequently performed by equilibrium dialysis to obtain the more reliable ligand B28 K_d measurement of 190 nM (Figure 4). Experiment B, there-

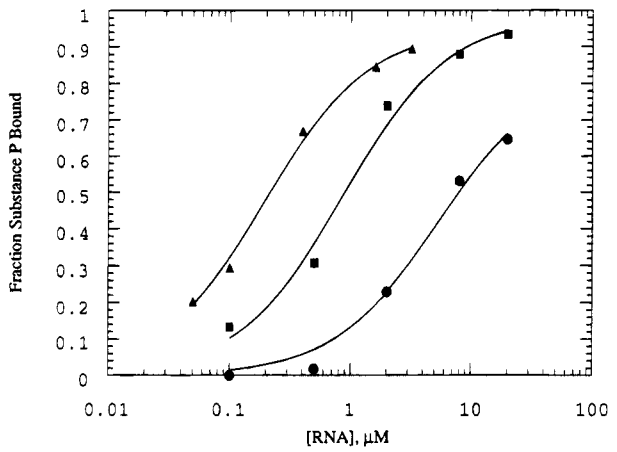

FIGURE 4: Binding curves for ligand A13 (●), experiment B selection cycle 12 pooled RNAs (■), and ligand B28 (▲). The fraction of [^{3}H]substance P bound in equilibrium dialysis experiments is plotted as a function of total RNA concentration. In each experiment, the concentration of substance P applied to one side of the dialysis membrane (P_o) was 20 nM. All binding reactions were at room temperature in 1× substance P binding buffer.

fore, yielded ligands with binding affinities up to 30-fold better than their ancestral ligand A13 and approximately 6000-fold better than the initial unselected random sequence RNA pool. With the exception of class 3 ligand B32, all class 2 and class 3 ligands exhibited a K_d above 2 μM (Figure 3). A comparison of nucleotide positions 4−22 of the high affinity class 3 ligand B32 with positions 10−28 of class 1 ligands reveals a significant sequence similarity ($GGCN_2$-$UN_4ACCCUNAGG$), indicating the probable importance of this region in SP binding. The lower affinity class 3 ligands (B22 and B23) share significant stretches of sequence homology with the relatively low affinity class 2 ligands: ACAGGACAC and GACGAGUU at positions 4−12 and 24−31 in the class 1 alignment, respectively (Figure 3).

A comparative analysis of sequences within class 1 and 2, with base-pairing decisions influenced by observed covariation [phylogenetic comparative approach; see Fox and Woese (1975) and Noller and Woese (1981)], led to the prediction of possible secondary structures (Figure 5). Despite the high degree of sequence variability introduced during experiment B, a consensus structure predicted for the class 1 ligands resembles that predicted for A13 (Figure 5). Conserved nucleotides within the large asymmetric loop present in both the predicted ligand A13 structure and the class 1 consensus structure suggest a role for these nucleotides in binding. The majority of the nucleotides within this loop are completely conserved among class 1 ligands. In addition, the sequence within the high affinity class 3 ligand (B32) that is shared with the class 1 ligands (ACCCUNAGG) is present within this loop.

Experiment B class 2 ligands share a high degree of primary and secondary structure similarity; all can assume a stem−loop structure with two internal asymmetric loops (Figure 5). Most of the internal loop nucleotides are conserved, suggesting their involvement in the SP interaction. The nucleotides conserved among the lower affinity class 3 ligands (B22 and B23) and the class 2 ligands (listed above) are partially present in the two internal loops and form the stem that separates them.

To determine which amino acids are required for the interaction with ligand B28, overlapping fragments of the peptide were tested for their ability to compete with intact SP for binding in equilibrium dialysis experiments. Each competition experiment contained 1.6 μM B28 RNA and 1.6 μM SP (P_o; including 20 nM [^{3}H]SP), in addition to 32 μM (P_o) of the competing peptide fragment (20:1 ratio of competing peptide to SP). In the absence of competing peptide, 63% of the SP was bound under these conditions. The percent inhibition observed with 32 μM competing unlabeled SP (as a positive control) was 86%. The results (Figure 6) indicate that the four C-terminal residues (Phe^8-Gly^9-Leu^{10}-Met^{11}-NH_2) are not necessary for binding. The presence of Phe^7 appears to be required for optimal binding but is not entirely necessary. The involvement of Arg^1 in the interaction is indicated by the inability of fragment 2−11 to compete under these conditions. However, the four N-terminal residues (Arg^1-Pro^2-Lys^3-Pro^4) alone did not exhibit significant binding, suggesting a requirement for one or both of the Gln residues. Taken together, the results suggest that, minimally, Arg^1 and Gln^5-Gln^6-Phe^7 are involved in the interaction with ligand B28.

A high specificity of ligand B28 for SP was suggested by the ability of the ligand to discriminate between SP and the reverse orientation of the same peptide (rSP, Figure 1). An equilibrium dialysis competition experiment was performed as above, except the competing concentrations (P_o) of rSP were either 1.6, 6.4, 25, or 100 μM. Competition experiments with unlabeled SP at the same concentrations were performed for comparison. In the absence of added competitor, 69% binding of SP was observed under these conditions. The data indicates that rSP is a poor competitor for binding to ligand B28 (Figure 7).

DISCUSSION

In vitro evolution methods have yielded nucleic acid ligands to non-nucleic acid binding proteins with dissociation constants in the 10^{-10} to 10^{-8} range (Gold et al., 1993; Bock et al., 1992; Jellinek et al., 1993). Ligands to cyanocobalamin (Lorsch & Szostak, 1994) and theophylline (Jenison et al., 1994) represent the highest affinity ligands reported to small constrained molecules with dissociation constants as low as 88 and 100 nM, respectively. In this report, we describe the isolation of RNA ligands to the nonconstrained peptide substance P with dissociation constants as low as 190 nM, representing an improvement in binding over the initial unselected random sequence RNA pool of about 6000-fold. Ligands that bind this tightly were not necessarily expected since a conformational entropy loss should accompany binding of the recognized conformation of the flexible peptide, resulting in a relatively low complex stability. A parallel situation has been seen in studies of tRNA anticodon/anticodon (Eisinger & Gross, 1974; Moras et al., 1986) and tRNA anticodon/codon (Yoon et al., 1975) interactions. The relatively high affinity of the anticodon of one tRNA molecule for a complementary anticodon of another tRNA has, at least in part, been explained by the small entropy of binding; in an intact tRNA, the anticodon loop has a fairly rigid conformation (Watson−Crick helical structure) that undergoes only minor changes upon duplex formation (Moras et al., 1986). On the contrary, trinucleotides in solution visit numerous conformations, many of which do not allow simultaneous hydrogen bonding with the three bases of a complimentary trinucleotide. Accordingly, complexes formed between a tRNA anticodon and its

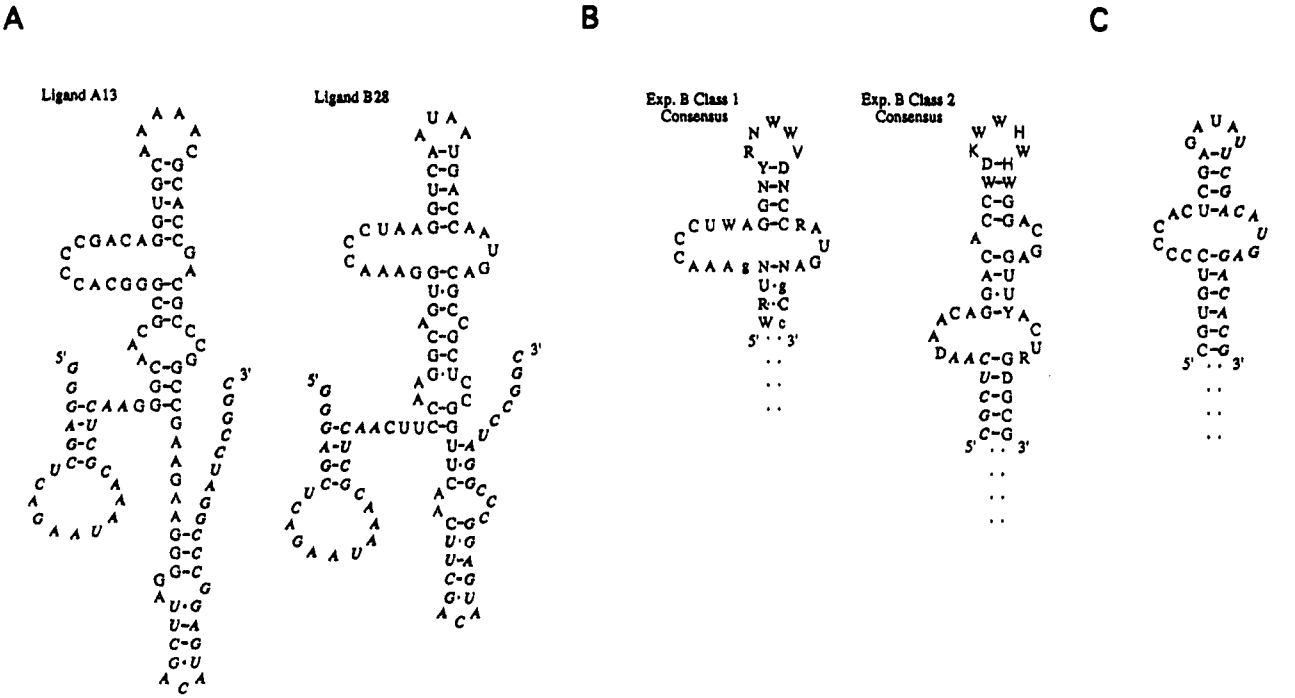

FIGURE 5: (A) Predicted secondary structures for ligands A13 and B28. The 23-nucleotide 5′ and 25-nucleotide 3′ fixed sequences complementary to the PCR primers are shown in bold italicized type. (B) Consensus sequences and predicted secondary structures for the highly conserved regions of experiment B class 1 and 2 ligands. Universally conserved nucleotides are shown as normal capital letters. Lowercase letters are used to indicate positions were a specific nucleotide is not universally conserved but occurs at a frequency of >90%. The following symbols are used to indicate other nucleotide patterns at individual locations: N = any base; R = A or G; Y = C or U; W = A or U; V = A or C; K = G or U; D = A, G, or U; H = A, C, or U; ... = base-pairing is sometimes, but not always, possible between these two nucleotide positions. In the class 2 consensus structure, the terminal seven nucleotides of the 5′ fixed sequence are shown in bold italicized type. (C) An RNA ligand selected for binding to arginine by Connell et al. (1993) is shown for comparison. Only the region of the RNA required for binding is shown. The sequence in bold italicized type represents the portion of the 3′ fixed primer binding sequence that is involved in the structure.

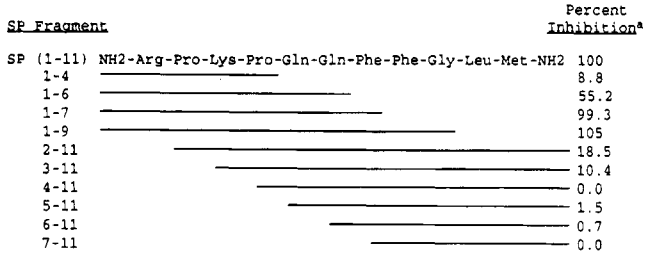

FIGURE 6: Competition between substance P fragments and intact substance P for binding to ligand B28. Equilibrium dialysis binding reactions consisted of 1.6 μM ligand B28 RNA, 1.6 μM substance P (P_o; including 20 nM [^{3}H]substance P), and 32 μM (20-fold excess) of the competing peptide fragment. The fraction of [^{3}H]-substance P bound in the absence of added competitor was 0.63. An 86.1% decrease (inhibition) in ^{3}H-substance P bound was observed in the presence of 32 μM (P_o) competing unlabeled substance P. The percent decrease (inhibition) in [^{3}H]substance P bound in the presence of individual substance P fragments is expressed relative to intact substance P = 100% inhibition.

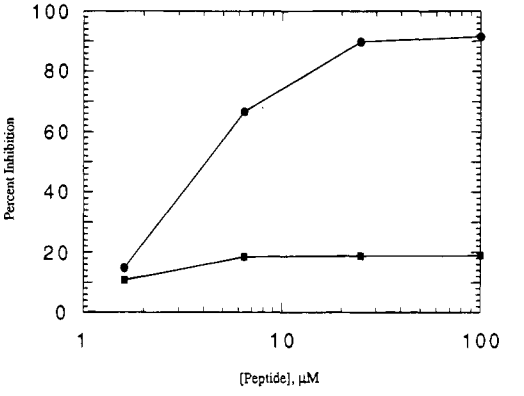

FIGURE 7: Plot of percent inhibition of [^{3}H]substance P bound to ligand B28 in the presence of varying concentrations of unlabeled substance P (●) or a peptide (rSP) which contains the same amino acid sequence as substance P but in the reverse orientation (■). Each equilibrium dialysis binding reaction consisted of 1.6 μM ligand B28 RNA and 1.6 μM substance P (P_o; including 20 nM [^{3}H]substance P) in addition to the competing peptide at a concentration of 0, 1.6, 6.4, 25, or 100 μM. The percent decrease (inhibition) in the fraction of [^{3}H]substance P bound in the presence of competing peptide is expressed relative to the fraction of [^{3}H]-substance P bound in the absence of added competitor (0.69).

trinucleotide codon have binding affinities 2−3 orders of magnitude weaker than seen with the tRNA/tRNA interactions. The affinities of complementary trinucleotides (Eisinger, 1971), where neither polymer is constrained, are about 6 orders of magnitude weaker.

An affinity matrix partitioning method, in which peptide-bound RNAs were recovered by cleaving the matrix linker, was used in this study to isolate RNAs with affinity for SP. An alternative method would involve the elution of target bound ligands by competition with free target molecules, as this should assure the recovery of ligands specific for the free target. However, attempts at applying this method to SELEX with SP have not been successful thus far, possibly due to peptide aggregation in the matrix environment. As an example, ligand A13 bound to SP-Sepharose was not effectively eluted with eight column volumes of 1 mM SP, even though this ligand has demonstrated a dissociation constant of 5.8 μM for free SP in equilibrium dialysis experiments. Following application of the free SP, the bound RNAs could not be eluted with DTT either (without prior

application of free SP, bound RNAs are rapidly eluted with DTT). When rSP (Figure 1) is substituted for SP in the above experiment, RNAs are not eluted with 1 mM rSP but are rapidly eluted with DTT (indicating the RNAs remained bound to column-coupled SP). An explanation for this phenomenon is that ligands bind to free SP aggregates trapped in the matrix. Two points should be noted however: the inclusion of [^{3}H]SP in the applied 1 mM SP solution demonstrated that SP rapidly elutes from the matrix and, in the equilibrium dialysis experiments, SP rapidly (2−2.5 h) equilibrated across the 12 000−14 000 MW membrane at all concentrations used in this study (up to 100 μM). SP has been shown to aggregate in basic, acidic, and saline aqueous solutions (Poujade et al., 1983; Billeter et al., 1982). In this study, the salinity of the buffer would most likely be responsible if aggregation occurs because SP exists as a monomer at concentrations at least as high as 5 mM in aqueous solution at neutral pH (Chassaing et al., 1986). Nonetheless, the ligand elution procedure described in this study would be useful in other affinity selection procedures utilizing immobilized targets where elution with the free target is not a viable option.

The PCR mutagenesis procedure introduced in SELEX experiment B resulted in the selection of RNAs with little sequence resemblance to the "parental" ligand A13, yet, at least for class 1 ligands, the selected RNAs can assume a similar secondary structure (Figure 5). The large number of selected mutations is not completely unexpected since the random sequence RNA pool used to initiate experiment A contained about 10^{14} unique sequences, only a minute fraction of the 4^{60} possible sequences 60 nt in length. Under these conditions, the evolution of a more favorable primary sequence solution to higher ordered structures would be essentially unavoidable.

A phylogenetic comparison of ligand A13 and the ligands recovered from experiment B led to the secondary structures predicted in Figure 5. Conserved sequences within the asymmetric internal loops of all inspected ligands suggest their importance in the SP interaction. The discovery that the SP N-terminal arginine residue is involved in the interaction with ligand B28 prompted an inspection of RNA ligands previously selected for binding to arginine by Connell et al. (1993). One of three structural motifs in their report (motif 3) is similar to the predicted class 1 consensus structure. Their predicted motif (Figure 5) consists of a stem−loop structure with an internal asymmetric loop. As in the class one consensus structure, a four base-pair stem separates the internal loop from a five-nucleotide terminal loop. Similarities are also seen in the internal loop sequences. However, the similarities seen in the 3′ half of the internal loop may be entirely coincidental since this region of the "motif 3" ligand is part of a 3′ fixed primer binding sequence; only its structural location has been selected.

Competitive binding experiments (Figure 7) indicated that ligand B28 can discriminate between SP and the reverse orientation of the same amino acid sequence (rSP, Figure 1). Knowing that multiple amino acid residues are involved in the interaction (Figure 6), it can be assumed that SP possesses a unique structure, with unique relative positions of amino acid side chains, as it is recognized by the ligand. Although this recognized structure might include the N-terminal arginine α-amine of free SP, the N-terminus of the column-coupled peptide (Cys-SP, Figure 1), to which ligands were selected, was acetylated. While circular dichroism and NMR studies (Mehlis et al., 1980; Chassaing et al., 1986) indicate that SP lacks an ordered structure in water (or undergoes a rapid equilibrium between different conformations), Chassaing et al. (1986) have proposed a preferred conformation of SP in methanol consisting of a flexible Arg^1-Pro^2-Lys^3 N-terminus, an α-helical structure involving residues Pro^4-Gln^5-Gln^6-Phe^7-Phe^8, and an interaction of the C-terminal carboxamide with the primary amides from both glutamines. An α-helical structure may not exist in the core residues of rSP, for instance, because these residues would not have the benefit of the helix-nucleating properties of proline (Presta & Rose, 1988).

This study demonstrates an application of SELEX technology for the selection of ligands with high affinity for a small nonconstrained peptide. A detailed analysis of the structure of these ligands and of their interaction with SP should increase our understanding of amino acid-specific RNA structural motifs. In regards to the pharmacological utility of these ligands, further study is required to determine whether they antagonize the biological activity of SP.

ACKNOWLEDGMENT

We thank Drs. Britta Singer and David Pribnow for helpful discussions. The excellent technical assistance provided during portions of this study by Terri Gibson and Tali Carasali is gratefully acknowledged. We also thank Nexstar, Inc., and the W. M. Keck Foundation for their generous support of RNA science on the Boulder campus.

REFERENCES

Bartel, D. P., & Szostak, J. W. (1993) *Science 261*, 1411−1418.
Bartel, D. P., Zapp, M. L., Green, M. R., & Szostak, J. W. (1991) *Cell 67*, 529−536.
Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321−346.
Bock, L., Griffin, L., Latham, J., Vermaas, E., & Toole, J. (1992) *Nature 355*, 564−566.
Cadwell, R. C., & Joyce, G. F. (1992) *PCR Methods Appl. 2*, 28−33.
Chassaing, G., Convert, O., & Lavielle, S. (1986) *Eur. J. Biochem. 154*, 77−85.
Connell, G. J., Illangesekare, M., & Yarus, M. (1993) *Biochemistry 32*, 5497−5502.
Eisinger, J. (1971) *Biochem. Biophys. Res. Commun. 43*, 854−861.
Eisinger, J., & Gross, N. (1974) *J. Mol. Biol. 88*, 165−174.
Ellington, A. D., & Szostak, J. W. (1990) *Nature 346*, 818−822.
Ellington, A. D., & Szostak, J. W. (1992) *Nature 355*, 850−852.
Escher, E., & Regoli, D (1989) in *Peptide Hormones as Prohormones: Processing, Biological Activity, Pharmacology* (Martinez, J., Ed.) pp 26−52, Ellis Horwood Limited, West Sussex, England.
Fan, T.-P. D., Hu, D.-E., Guard, S., Gresham, G. A., & Watling, K. J. (1993) *Br. J. Pharmacol. 110*, 43−49.
Fox, G., & Woese, C. (1975) *Nature 256*, 505−507.
Gold, L., Tuerk, C., Allen, P., Binkley, J., Brown, D., Green, L., MacDougal, S., Schneider, D., Tasset, D., & Eddy, S. (1993) in *The RNA World* (Gespeland, A., & Atkins, J., Eds.) pp 497−509, Cold Spring Harbor Laboratory Press, Plainview, NY.
Huston, J. P., Hasenohrl, R. U., Boix, F., Gerhardt, P., & Schwarting, R. K. W. (1993) *Psychopharmacology 112*, 147−162.
Irvine, D., Tuerk, C., & Gold, L. (1991) *J. Mol. Biol. 222*, 739−761.

Jellinek, D., Lynott, C. K., Rifkin, D. B., & Janjic, 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.
Karush, F., & Sonnenberg, M. (1949) *J. Am. Chem. Soc. 71*, 1369−1376.
Leung, D. W., Chen, E., & Goeddel, D. V. (1989) *Technique 1*, 11−15.
Lorsch, J. R., & Szostak, J. W. (1994) *Biochemistry 33*, 973−982.
Mehlis, B., Rueger, M., Becker, M., Bienert, M., Niedrich, H., & Oehme, P. (1980) *Int. J. Pept. Protein Res. 15*, 20−28.
Moras, D., Dock, A.-C., Dumas, P., Westhof, E., Romby, P., Ebel, J-P., & Giege, R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 932−936.
Noller, H. F., & Woese, C. R. (1981) *Science 212*, 403−410.
Poujade, C., Lavielle, S., Chassaing, G., & Marquet, A. (1983) *Biochem. Biophys. Res. Commun. 114*, 1109−1116.
Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632−1641.
Rosen, D., Okamura, M. Y., & Feher, G. (1980) *Biochemistry 19*, 5687−5692.
Rizo, J., & Gierasch, L. M. (1992) *Annu. Rev. Biochem. 61*, 387−418.
Sassanfar, M., & Szostak, J. (1993) *Nature 364*, 550−553.
Schneider, D., Tuerk, C., & Gold, L. (1992) *J. Mol. Biol. 228*, 862−869.
Tuerk, C., & Gold, L. (1990) *Science 249*, 505−510.
Tuerk, C., MacDougal, S., & Gold, L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6988−6992.
Yoon, K., Turner, D. H., & Tinoco, I., Jr. (1975) *J. Mol. Biol. 99*, 507−518.

BI9427072