In vitro selection of RNAs that bind to the human immunodeficiency virus type-1 gag polyprotein

Michael A. Lochrie*, Sheela Waugh, Dickson G. Pratt, Jr, Jared Clever[1], Tristram G. Parslow[1] and Barry Polisky

NeXstar Pharmaceuticals, Inc., 2860 Wilderness Place, Boulder, CO 80301, USA and [1]Department of Pathology and Department of Microbiology and Immunology, Box 506, 513 Parnassus Avenue, University of California, San Francisco, CA 94143-0506, USA

ABSTRACT

RNA ligands that bind to the human immunodeficiency virus type-1 (HIV-1) gag polyprotein with 10^{-9} M affinity were isolated from a complex pool of RNAs using an *in vitro* selection method. The ligands bind to two different regions within gag, either to the matrix protein or to the nucleocapsid protein. Binding of a matrix ligand to gag did not interfere with the binding of a nucleocapsid ligand, and binding of a nucleocapsid ligand to gag did not interfere with the binding of a matrix ligand. However, binding of a nucleocapsid ligand to gag did interfere with binding of an RNA containing the HIV-1 RNA packaging element (ψ), even though the sequence of the nucleocapsid ligand is not similar to ψ. The minimal sequences required for the ligands to bind to matrix or nucleocapsid were determined. Minimal nucleocapsid ligands are predicted to form a stem–loop structure that has a self-complementary sequence at one end. Minimal matrix ligands are predicted to form a different stem–loop structure that has a CAARU loop sequence. The properties of these RNA ligands may provide tools for studying RNA interactions with matrix and nucleocapsid, and a novel method for inhibiting HIV replication.

INTRODUCTION

The HIV-1 gag protein is a polyprotein composed of matrix, capsid, nucleocapsid, p1, p2 and p6 proteins. There are estimated to be about 2000 gag proteins per virus. Gag is processed into individual proteins by the HIV-1 protease after the virus is released from infected cells.

Gag has multiple functions in the life cycle of HIV-1 that are reflective of the functions of its individual protein components (1). Nucleocapsid is a zinc finger protein that has nucleic acid binding, melting and annealing activities which are required at several points in the viral life cycle (2–4). For example, nucleocapsid binding to a structured RNA element called ψ at the 5′ end of the HIV-1 genomic RNA plays a critical role in RNA packaging into nascent virions (5). Capsid surrounds the nucleocapsid–RNA complex. Capsid multimerizes to form the viral core

and binds the cellular protein cyclophilin A (6). Matrix surrounds the capsid and is found just under the viral membrane. Matrix also multimerizes to form the viral core, is required for incorporation of envelope proteins into assembling viruses (7), and plays a role in the entry of viral DNA into the nucleus prior to integration (8). The p6 protein plays a role in budding of viruses from cells (9) and binds to the HIV-1 vpr protein leading to its incorporation into viruses. The p1 and p2 peptides have been proposed to regulate the proteolysis of gag (11). A gag–pol polyprotein precursor is also encoded by the HIV genome. The pol gene encodes enzymes such as protease, integrase, reverse transcriptase and RNAse H. Multimerization of gag–pol results in the incorporation of these enzymes into virions.

Systematic evolution of ligands by exponential enrichment (SELEX) is an *in vitro* selection and evolution method that has been used to isolate nucleic acid polymers that bind to various molecules (12,13). The SELEX process and similar methods have been used previously to isolate RNA ligands that bind to the HIV-1 rev (14–17), tat (18), reverse transcriptase (19,20) and integrase (21) proteins. Another screening method has been used to isolate nucleic acids that bind to the HIV-1 envelope protein (22). Here we report the use of the SELEX method to isolate RNA molecules that bind to the HIV-1 gag polyprotein. Such molecules could be used to inhibit gag functions or study gag/RNA interactions.

MATERIALS AND METHODS

Reagents

The HIV-1 p55 (gag), p7 (nucleocapsid) and p15 (p7 nucleocapsid–p6) proteins from HIV-1 strain LAI were expressed in *Escherichia coli* as fusions to glutathione *S*-transferase (GST) and purified as described (5). They are referred to here as GST–gag, GST–p7 and GST–p15, respectively. The GST protein was also purified using the same procedures. The HIV-1$_{SF-2}$ p24 capsid protein was obtained from the NIH AIDS Research and Reference Reagent Program and differs from HIV-1$_{LAI}$ p24 at one amino acid position (23). The HIV-1$_{BH-10}$ p17 matrix protein was obtained from the MRC AIDS Reagent Project and is identical in sequence to HIV-1$_{LAI}$ p17 (23). All DNA oligonucleotides were synthesized by Operon, Inc. (Alameda, CA). Restriction enzymes, T4 RNA ligase and T4 polynucleotide kinase were purchased from New

---

*To whom correspondence should be addressed. Tel: +1 303 546 7661; Fax: +1 303 444 0672; Email: mlochrie@nexstar.com

England Biolabs (Beverly, MA). T7 RNA polymerase was purchased from Enzyco, Inc. (Denver, CO). RNase T_1 and bacterial alkaline phosphatase were purchased from Boehringer Mannheim (Indianapolis, IN). All radioisotopes were purchased from DuPont NEN Research Products (Boston, MA). The pUC9 plasmid was obtained from Gibco BRL (Gaithersburg, MD).

In vitro selection of RNA ligands

RNA ligands that bind to the HIV-1 gag protein were isolated by the SELEX method essentially as described (24–26), using the GST–gag protein as a target, but with the following modifications. SELEX experiment A used a 2′-hydroxyl RNA library with 50 randomized positions (50N) that has the following sequence: 5′-GGGAGACAAGAAUAAACGCUCAA-50N-UUCGACAG-GAGGCUCACAACAGGC-3′. The 5′ and 3′ 'fixed' sequences were required for the reverse transcription and PCR steps of SELEX. The RNA library was bound to GST–gag in a buffer that consisted of 50 mM Tris, pH 7.5, 200 mM KOAc, 5 mM MgCl$_2$ and 1 mM dithiothreitol. Binding of the RNA pool to GST–gag was done at 37°C for 5 min. Nitrocellulose filtration (HAWP, Millipore, Inc., Bedford, MA) was used to separate free from protein-bound RNA. After filtration, the protein-bound RNA was recovered and reverse transcribed with avian myeloblastosis virus reverse transcriptase (Life Sciences, Inc., St Petersburg, FL) and a 3′ primer (5′-GCCTGTTGTGAGCCTCCTGTCGAA-3′). *Taq* polymerase was used to amplify the cDNA by PCR using a 5′ primer (5′-TAATACGACTCACTATAGGGAGACAAGAATA-AACGCTCAA-3′, which contains the sequence of the T7 RNA polymerase promoter and of the 5′ fixed region) and the 3′ primer described above. The PCR product was then transcribed with T7 RNA polymerase to generate an enriched RNA library for use in the next round of SELEX. The concentration of GST–gag was reduced from 180 nM to 10 pM during the course of SELEX A and the molar ratio of the RNA library to GST–gag ranged from 3 to 160.

SELEX experiment B was similar to SELEX A except that an RNA library that has the following sequence was used: 5′-GGGA-AAAGCGAATCATACACAAGA-50N-GCTCCGCCAGAGAC-CAACCGAGAA-3′. This RNA library was bound to GST–gag at 37°C for 15 min in a buffer that consisted of 50 mM Tris, pH 7.4, 140 mM KCl, 5 mM NaCl, 5 mM MgCl$_2$ and 1 mM dithiothreitol. Murine leukemia virus reverse transcriptase (Superscript; Gibco, Inc., Gaithersburg, MD) and a 3′ primer (5′-TTCTCGGTTGGTCT-CTGGCGGAGC-3′) was used for the reverse transcription step. *Taq* polymerase, a 5′ primer (5′-TAATACGACTCACTATAGGGAA-AAGCGAATCATACACAAGA-3′ which contains the sequence of the T7 RNA polymerase promoter and of the 5′ fixed region) and the 3′ primer were used for the PCR steps. During SELEX B the concentration of GST–gag was reduced from 500 nM to 40 pM and the molar ratio of RNA library to GST–gag ranged from 5 to 100. In addition, a 10–200-fold molar excess of heparin (molecular weight 5000; Calbiochem, Inc., La Jolla, CA) over RNA was used in each round of SELEX B, except for the first round.

Subcloning of ligands

At the conclusion of the SELEX procedures, the RNA pools from SELEX A were amplified by PCR using the following primers: 5′-CCG<u>AAGCTT</u>AATACGACTCACTATAGGGAGACAAG-AATAAACGCTCAA-3′ and 5′-GCC<u>GGATCC</u>GCCTGTTGTG-

AGCCTCCTGTCGAA-3′ and the RNA pools from SELEX B were amplified by PCR using the following primers: 5′-CG-C<u>GGATCC</u>TAATACGACTCACTATAGGGAAAAGCGAATC-ATACACAAGA-3′ and 5′-GGC<u>GAATTC</u>TTCTCGGTTGGTC-TCTGGCGGAGC-3′. The PCR products were cut with *Hin*dIII and *Bam*HI (SELEX A; underlined) or *Eco*RI and *Bam*HI (SELEX B; underlined), and then subcloned into pUC9. All ligations were transformed into *E.coli* strain DH5α.

Sequence analysis of cloned ligands

Plasmids were prepared from tranformants (27) and ligand sequences were determined using the Sequenase, version 2.0, enzyme (US Biochemicals, Inc., Cleveland, OH) according to the manufacturer's instructions. A DNA primer, corresponding to the sequence of the T7 RNA polymerase promoter (5′-TAATACGA-CTCACTATA-3′), was used as the sequencing primer. Sequences were analyzed and aligned using the GCG Sequence Analysis Software Package, version 7.2 (Genetics Computer Group, Madison, WI).

Binding affinity of RNA ligands for HIV-1 proteins

The binding affinity of RNA ligands for various proteins was measured by a nitrocellulose filtration method essentially as described (24,26,28). Radiolabeled RNA (~2 fmol, ~10 000 c.p.m.) was bound to protein that typically ranged in concentration from 10^{-10} to 10^{-6} M. Protein and RNA were incubated at 37°C for 10 min in a 25 µl reaction composed of SELEX B buffer. Since the amount of RNA added to each reaction varies, the actual concentration of RNA in the binding reactions varied from ~50 to 200 pM, i.e., at least 5-fold below any apparent K_d. Therefore, for RNAs with apparent K_ds near 1 nM, the K_d may be an underestimation as a result of the RNA competing with itself to some extent for binding. After binding RNA to protein the reactions were vacuum-filtered through nitrocellulose filters. A reaction with no protein was also incubated and filtered in order to determine the amount of RNA that binds to the filter in the absence of protein. The amount of labeled RNA retained on the filter in a protein-dependent manner was determined. The percentage of RNA bound to the protein was plotted against the concentration of the protein using Kaleidograph computer software (Synergy, Inc., Reading, PA). The apparent K_d is defined as the protein concentration at which 50% of maximal RNA binding occurs. Binding of RNA to GST–gag, GST–p15 or GST–p7 proteins was done in the presence of a 10 000-fold molar excess of yeast tRNA over the RNA ligand concentration.

The ψ-456 RNA was made by transcription of *Xho*I-digested pBSP (29). ψ-456 is a 456 base RNA that consists of 441 bases from the HIV-1$_{HXB2R}$ RNA (containing nucleotide positions 712–1152 of HIV-1$_{HXB2R}$ RNA; 30) plus 15 bases of vector-derived sequence at the 5′ end. The concentration of radiolabeled ψ-456 was 75 pM (Fig. 1) or 159 pM (Fig. 2).

Competition binding reactions

Two radiolabeled RNA ligands (SW8.4 or SW8.27) were transcribed with T7 RNA polymerase as described above. A constant concentration of radiolabeled ligand (0.3 nM SW8.27 or 0.4 nM SW8.4) was mixed with various amounts of unlabeled competitor RNAs (SW8.4, SW8.27, ψ-456, or the round 0 RNA library used for SELEX A) ranging in concentration from 10^{-12}
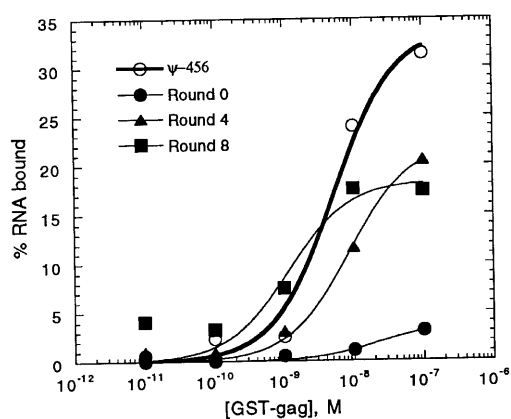
Figure 1. Binding of selected RNA pools to GST–gag protein. RNA pools from rounds 0, 4 and 8 of *in vitro* selection 'B' or the ψ-456 RNA were bound to varying concentrations of GST–gag protein as described in Materials and Methods.

Figure 2. Binding of cloned RNA ligands to GST–gag protein. RNA prepared from cloned RNA ligands SW8.4 and SW8.27 was bound to varying concentrations of GST–gag protein and compared to the binding of the random RNA or ψ-456 RNA as described in Materials and Methods.

to 10^{-6} M. The mixed RNAs were then added to 1 nM GST–gag protein, a concentration approximately equal to the apparent K_d of the protein for the radiolabeled RNA ligand. The reaction was incubated in binding buffer and then filtered as described above.

Ligand boundary determinations

The boundaries of RNA ligands were determined essentially as described by Jellinek *et al*. (28). Ligands were labeled at the 5′ end by dephosphorylating the RNA with bacterial alkaline phosphatase and then labeling the RNA by using [γ-^{32}P]ATP and T4 polynucleotide kinase. Ligands were labeled at the 3′ end using cytidine 3′,5′[^{32}P]bis(phosphate) and T4 RNA ligase. The labeled RNAs were partially hydrolyzed in separate tubes for 5 min in 50 mM NaCO$_3$, pH 9.0 at 95°C. The hydrolyzed RNA was bound to the appropriate HIV-1 protein at a concentration that was equal to or within a factor of 5 of the apparent K_d. The binding reactions were incubated and filtered as described above. Bound RNA was eluted from the filter in 66% phenol, 2 M urea. The eluted RNAs were precipitated and analyzed on an 8% polyacrylamide, 8 M urea gel electrophoresed in 1× TBE buffer (27). To align the boundary with the ligand sequence, the partially alkali-hydrolyzed ligand and a RNAse T$_1$ partial digest of the ligand were electrophoresed on the same gel. The HIV-1 p17 matrix protein was used to determine the boundaries of matrix ligands and the GST–p15 protein was used to determine the boundaries of nucleocapsid ligands. In some binding reactions, especially with high affinity ligands, a 10 000-fold molar excess of heparin or yeast tRNA over the ligand was added to suppress nonspecific binding to short RNA fragments.

Truncated and mutated RNA ligands

Truncated RNA ligands were generated by two methods. In one method, an oligonucleotide consisting of the antisense sequence of the T7 RNA polymerase promoter plus the antisense sequence of the desired truncated ligand was synthesized. This was hybridized to an oligonucleotide with the sequence corresponding to the sense strand of the T7 promoter. Such partial duplexes can be used as a template for transcription by T7 RNA polymerase (30,31).
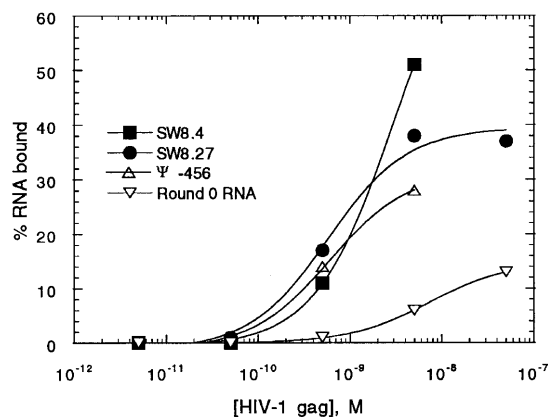
In another method, two oligonucleotides were synthesized. The sequence of one oligonucleotide contained the sense sequence of the T7 RNA polymerase promoter and part of the 5′ end of the truncated ligand. The sequence of the second oligonucleotide encoded the antisense sequence of the 3′ end of the truncated ligand. The oligonucleotides were designed so that they overlapped. These oligonucleotides were subjected to three rounds of PCR in order to generate a double-stranded T7 RNA polymerase transcription template. Mutated ligands were generated in a similar manner by incorporating the desired base changes into the oligonucleotides used to generate the transcription templates.

RESULTS

In vitro selection of RNA ligands

Two independent experiments (SELEX A and B) were done to select RNAs that bind to the HIV-1 gag protein *in vitro* with high affinity. The number of RNAs present in the initial (round 0) library of RNA molecules was estimated to be 8×10^{14} for SELEX A and 7×10^{14} for SELEX B. The two selection experiments used different binding buffers, reverse transcriptases and RNA libraries. In addition, heparin was used in SELEX B as a competitor for nonspecific or low affinity interactions.

A GST–gag fusion protein was used for selection of gag-binding RNAs. There was minimal concern about isolating GST-binding RNAs since the affinity of the round 0 RNA for GST was low (>10 μM), while the affinity of round 0 RNA for GST–gag was ~0.1–1 μM. Analysis of the affinity of RNA pools for GST–gag after successive rounds of *in vitro* selection revealed a steady improvement in affinity (Fig. 1). By round 10 (SELEX A) or round 8 (SELEX B) the affinity of the RNA pools for GST–gag was beginning to plateau at ~1 nM (Fig. 1). However, the proportion of nitrocellulose filter-binding RNAs in the population had also steadily increased to ~10–30% of the total in later rounds. The bulk sequence of the pool of RNA molecules shifted from random to nonrandom at round 5 of both SELEX experiments, indicating the complexity of the pool of RNAs had been reduced (data not shown). Based on the affinity of the RNA populations for GST–gag, the nonrandomness of their sequence, and the estimated proportion of filter-binding RNAs, the selected RNAs

Table 1. Properties of RNA ligands that bind HIV-1 gag

Ligand[a]	Sequence[b,c]	Binds to[d]	N[e]	K_d[f]
ML6.1	ACAGCGAUUUGUGUGCUCGUCGGCUGUUUGCAGAUGUAUGGAACCAUGCA	nucleocapsid	1	
ML6.2	GGGCUUGGGGUGGGGUCUUGCAGACGGAUAUGGUAAGCUGCUGCUGGCU	GST-gag	1	
ML6.3	AUUGGGAACGUGGCGCAUCGGGAGUGUAUGGUGGCUUGUGGGCGGGUUA	GST-gag	1	
ML6.6	AGCCAGGUCUUCAAUGGUAGUUGUUGAUUGGUGGUGCUUUGCAGAUGUAU	nucleocapsid	1	11 nM
ML6.8	UCGGUUAUGGAGAGGCGGUAGAGAGCGAGUAGCAAGUAACGGUGCAAAAG	matrix	1	3 nM
ML6.9	CCGCACUUCGAUUGCAUUGUCGAGGAGAGCGGGUGCAGAUGUACGGGUUUG	nucleocapsid	1	7 nM
ML6.12	AGACUUGGUACGCGUGGUUAGCUAGGAGGGGGGGAGAAGAUAAGGCGAUAG	GST-gag	1	
ML6.16	GAGUUUUGCUUCGACAUGGGGUUAGCUGUGGACGCAGGGUGACCUGGGACC	GST-gag	1	
ML6.17	UGGAGUGGGUAGGCGUGGGAGGUAUGGCAUCUUGUGUGUGUGCCAAGAUGGU	?	1	4 nM
ML8.7	UGUGAUGGGUGUGAUUUUGGUUAGCUGUGGAGGGCCAUGUGGGCUGGACG	matrix	7	12 nM
ML8.11	GGUGUGUGUGCGCCGCGAGGUGUCUGAGGAGAGAGGUGGCAGUGGAGGGU	GST	1	13 nM
ML8.14	UGUGAUGGGUGUGAUUUUGGUUAGCUGUGGAGGGUUAUGUGGGUCGGACG	matrix	2	12 nM
ML8.20	CGUUCGGUGGUGGACAGGGUAAUGUGGGAGGGACCGGGUGAUUGUGUAUGU	nucleocapsid	2	10 nM
ML8.22	UCACGACGGAGGGCGGAUGAAGGGGGGAAGGUCGUGAGUCCAUGCCGUGU	nucleocapsid	1	10 nM
SW8.2	UGAACCUCUGAGUUCUUCAUAGAACUGAUAUAUGUUCAAUAAAGGGUAGAUG	matrix	1	
SW8.4	CGUCUUCUACAGGGAACCGUGGUGCAUCUGUGAAGUUGUAGAUUCCUAGU	nucleocapsid	12	1 nM
SW8.6	GGGUUGAUUACCAAAUCGUAACCUGUACCCUGCCUACACUAGGACAACGG	matrix	1	
SW8.24	CGAGCUUUUAGUAAGACUCAUGCCGAGAAAUCGGUCGUGAUGCUGUCGAG	nucleocapsid	2	5 nM
SW8.25	GCCCGGGAAUUGCAUGUUGUGCGUGCCGGGAGUCCAAGUCAGCAUCCUCA	nucleocapsid	1	5 nM
SW8.27	GGGUUGAUUACCAAAUCGUAACCUGUACCCUGCCUACCCUAGGACAACGA	matrix	3	3 nM
SW8.30	AUCCCACUCGGUCGUGACCUGACGUGAAAACGGAUAGGAUCGCACGUCAA	matrix	1	
SW8.31	GUGUCCCUUGUCACCUGGGACUGGGCCGUUUGAACUGACAUUCUAUACGA	nucleocapsid	1	41 nM
SW8.37	GGCCAACGUCCGAGUUGGUUUCCCAUAGCUGAGCACAGGACGGCCUUCUGC	nucleocapsid	1	50 nM
SW10.12	UCUCGUAUCAGCACACCGGGUACAAAGAGGAUGCAAAUACGCCUGUGACC	matrix	1	30 nM
SW10.22	CGUCUUCUACAGGGAACCGUGGUGCAUCUGUGAAGUUGUUGUAGAUUCCUAGA	nucleocapsid	1	1 nM
SW10.28	UUGUUAGCCUGGUGAAGCAUGUACUGCGAAGUGGAGGGAAUGUGGAAGGG	nucleocapsid	1	8 nM
SW10.39	CGAGCCUUUAGUAAGACUCAUGCCGAGAAAUCGGUCGUGAUGCUGUCGAG	nucleocapsid	1	2 nM

[a]SW, RNA ligands isolated from SELEX 'A'; ML, RNA ligands isolated from SELEX 'B'. The first number before the period refers to the round of the SELEX procedure from which the ligand was isolated. The number after the period is a clone designation number.
[b]The sequence of the selected region of each ligand is shown. The entire ligand also includes the 5′ and 3′ fixed regions as designated in Materials and Methods. Note that ligand SW10.22 has four bases different from SW8.4, ligand ML8.7 differs from ML8.14 by four bases, ligand SW8.6 differs from SW8.27 by two bases, and ligand SW8.24 differs from SW10.39 by one base.
[c]The 5′ fixed sequence of SW8.30 is missing the A at position 9.
[d]GST–gag, ligand binds to the GST–p55 (HIV-1 gag) protein used in the SELEX procedure, but that binding to individual components of the HIV-1 gag polyprotein has not been determined. Nucleocapsid, binds to the GST–p7 (HIV-1 nucleocapsid) protein; matrix, binds to the HIV-1 p17 matrix protein; GST, binds to the glutathione *S*-transferase protein; ?, ligand binds to the GST–p55 (HIV-1 gag) protein, but does not bind to GST, GST–p7 (HIV-1 nucleocapsid), GST–p15 (HIV-1 nucleocapsid p7–p6), p17 HIV-1 matrix or p24 HIV-1 capsid proteins.
[e]N, number of clones of each ligand isolated.
[f]K_d, apparent affinity of ligand for GST–gag as determined by the method described in Materials and Methods.

from rounds 8 and 10 of SELEX A and rounds 6 and 8 of SELEX B were subcloned so that the sequences and protein binding properties of individual RNAs could be determined.

Sequence of RNA ligands

The sequences of 27 different RNA ligands (representing 49 subclones) that bind to GST–gag are shown in Table 1. Several features of the sequences are notable. First, the sequences from SELEX A are not similar to those from SELEX B. SW8.4 and ML8.20 are the most closely related, with 60% identity within the randomized region. Second, a comparison of the entire randomized region of each ligand to the sequence of HIV-1$_{LAI}$ genomic RNA revealed that none of the sequences was >50% identical to HIV RNA or the ψ region. Third, one sequence from each SELEX procedure was more frequent than any of the others (Table 1). Ligand SW8.4 and a similar ligand (SW10.22) represented 48% (13/27) of the RNAs isolated from SELEX A. Ligand ML8.7 and a similar ligand (ML8.14) represented 41% (9/22) of the RNAs isolated from SELEX B. Most of the remaining sequences were unique isolates. Minor variants of some RNAs were also isolated (Table 1, legend). These are presumed to have arisen from errors during the PCR step of SELEX.

Affinity of RNA ligands

RNA ligands were tested for binding to GST–gag (at 1 or 10 nM) or to nitrocellulose filters. Only those ligands that bound to GST–gag and not to filters are listed in Table 1. Some of these ligands were bound to GST–gag over a wider range of concentrations to obtain an apparent binding affinity (K_d). An example of such a binding study is shown in Figure 2 for the two most frequently isolated ligands, SW8.4 and SW8.27. The affinity of 19 ligands for GST–gag is summarized in Table 1. Most of the ligands bound to GST–gag with a K_d of ~1–10 nM.

As a positive control for binding to gag, a 456 base RNA (ψ-456) was used (see Materials and Methods). ψ-456 encompasses a region in the HIV-1 RNA that contains the HIV-1 RNA packaging signal (ψ), which is bound specifically by the HIV-1 gag and nucleocapsid proteins. ψ-456 typically bound to GST–gag with a K_d of ~1 nM (for example see Figs 1 and 2). Therefore, the *in vitro* selected RNA ligands described here are able to bind to gag with an affinity that is approximately the same as the native HIV-1 ψ element.

Binding of RNA ligands to different sites on gag

Twenty three GST–gag ligands were screened for binding to GST, HIV-1 capsid, HIV-1 nucleocapsid (p7 and p15) and HIV-1 matrix
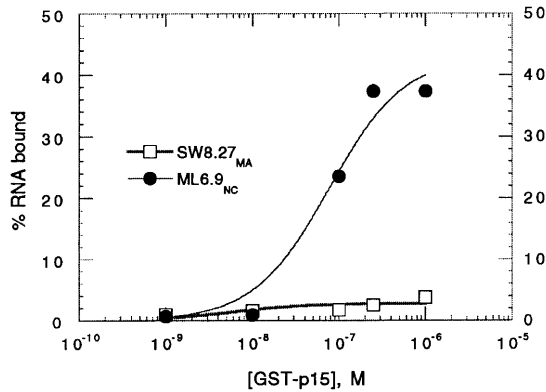
Figure 3. Binding of cloned RNA ligands to GST–HIV-1 p15 nucleocapsid protein. RNA prepared from cloned RNA ligands SW8.28$_{MA}$ and ML6.9$_{NC}$ was bound to varying concentrations of GST–p15 protein as described in Materials and Methods. NC, nucleocapsid ligand; MA, matrix ligand.
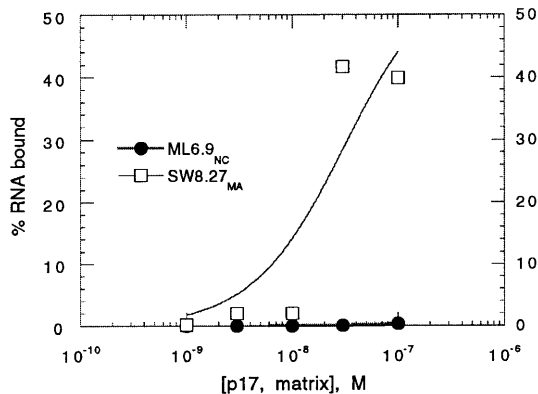

Figure 4. Binding of cloned RNA ligands to HIV-1 matrix protein. RNA prepared from cloned RNA liigands SW8.27$_{MA}$ and ML6.9$_{NC}$ was bound to varying concentrations of HIV-1 p17 matrix protein as described in Materials and Methods. NC, nucleocapsid ligand; MA, matrix ligand.

proteins. The binding specificity of the ligands fell mainly into two classes: those that bound to matrix and those that bound to nucleocapsid (Table 1). Binding was mutually exclusive. Matrix ligands did not bind with high affinity to nucleocapsid (Fig. 3) and nucleocapsid ligands did not bind with high affinity to matrix (Fig. 4). Nucleocapsid ligands were isolated more frequently from SELEX A and matrix ligands were isolated more frequently from SELEX B (see Table 1). The bias towards matrix ligands in SELEX B may have occurred because the heparin that was used in SELEX B may have effectively competed with RNA for binding to nucleocapsid and forced some ligands to bind to matrix.

The nucleocapsid ligands were able to bind to gag with the same affinity either in the presence or in the absence of a 10 000-fold molar excess of yeast tRNA. In contrast, the apparent affinity of round 0 RNA was reduced by a factor of at least 100-fold in the presence of a 10 000-fold molar excess of yeast tRNA (data not shown). Therefore, the nucleocapsid ligands not only bind to gag with high affinity, but also with specificity.

Two ligands were identified that bound to GST–gag, but not to nucleocapsid or matrix. ML8.11 bound to GST. ML6.17 bound to GST–gag with high affinity ($K_d = 4$ nM), but did not bind to the
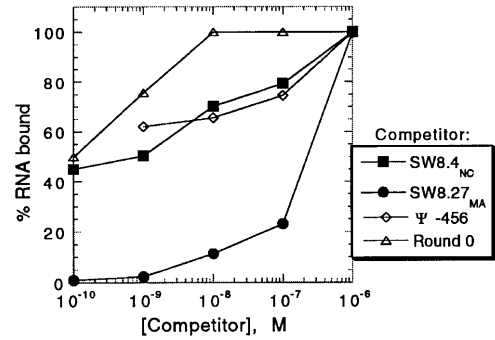

Figure 5. Competition between gag ligands and RNAs for binding to HIV-1 gag. The HIV-1 matrix RNA ligand SW8.27$_{MA}$ was labeled with ^{32}P and mixed with varying amounts of the indicated RNA competitors. The RNA mixture was incubated with GST–gag protein and filtered as described in Materials and Methods.
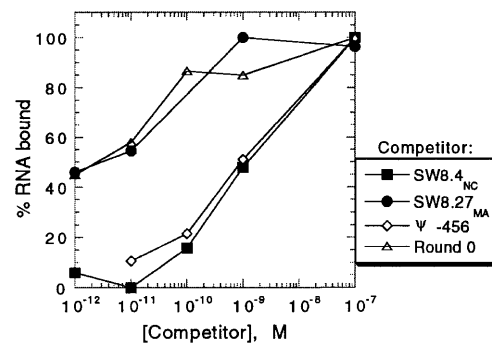

Figure 6. Competition between gag ligands and RNAs for binding to HIV-1 gag. The HIV-1 nucleocapsid RNA ligand SW8.4$_{NC}$ was labeled with ^{32}P and mixed with varying amounts of the indicated RNA competitors. The RNA mixture was incubated with GST–gag protein and filtered as described in Materials and Methods.

p17 matrix, p24 capsid, GST–p7 (nucleocapsid), GST–p15 (p7 nucleocapsid–p6) or GST proteins. It is possible that ML6.17 binds to more than one protein component within GST–gag.

Competition between RNA ligands for binding to gag

Ligands were isolated that bind to two separate domains of the HIV-1 gag protein (matrix and nucleocapsid). Competition binding reactions were done between matrix ligands, nucleocapsid ligands and ψ-456 to determine whether ligands that bound to the same or separate sites would compete for binding to gag.

Ligands that bound to the same region on gag interfered with each other for binding to gag. That is, a matrix ligand interfered with the binding of itself (Fig. 5) or of another matrix ligand (data not shown) to the gag protein. A nucleocapsid ligand interfered with the binding of itself (Fig. 6) or of another nucleocapsid ligand (data not shown) to the gag protein. In addition, a nucleocapsid ligand interfered with the binding of ψ-456 to gag (Fig. 6). Therefore there may be only one binding site for RNA on matrix or nucleocapsid, but there are at least two RNA binding sites on the gag polyprotein.

In contrast, a matrix ligand did not interfere with the binding of either a nucleocapsid ligand or ψ-456 to gag (Fig. 5). Also a nucleocapsid ligand did not interfere with the binding of a matrix

Table 2. Boundaries of RNA ligands that bind HIV-1 gag

Ligand[a]	Boundary[b]
ML6.1$_{NC}$	GGGAAAAGCGAA<u>UCAUACACAAGAACAGCGAUUUGUGUGCUCGUCGGCU</u>GUUugcagaugua UGGAACCAUGCA GCUCCGCCAGAGACCAACCGAGAA
ML6.6$_{NC}$	GGGAAAAGCGAAUC<u>AUACACAAGAAGCCAGGUCUUCAAUGGUAGUUGUUGAUUGGUGGUGCUU</u>ugcagauguaU GCUCCGCCAGAGACCAACCGAGAA
ML6.8$_{MA}$	GGGAAAAGCGAAUCAUACACAAGAUCGGUUAUGGAG<u>AGGCGGUAGAGACGCAGUAGCAAGUAACGGUGCAAAAG</u>GCUCCGCCAGAGACCAACCGAGAA
ML6.9$_{NC}$	GGGAAAAGCGAAUCAUACACAAGACCGCA<u>CUUCGAUUGCAUUGUCGAGGAGAGCGG</u>gugcagauguacGGGGUUUGGCUCCGCCAGAGACCAACCGAGAA
ML8.7$_{MA}$	GGGAAAAGCGAAUCAUACACAAGAUGUGAUGGGUGUGAUUUU<u>GGUUAGCUGUGGAGGGCCAUGUGGGCUGGA</u>CG GCUCCGCCAGAGACCAACCGAGAA
ML8.20$_{NC}$	GGGAAAAGCGAAUCAUACACAAGA<u>CAAGACGUUCGGUGGUGGACAGGGUAAUGUGGAGGGA</u>CCGGGUGAUuguguaugu gCUCCGCCAGAGACCAACCGAGAA
ML8.22$_{NC}$	GGGAAAAGCGAAUCAUACACAAGA<u>UCACGACGGAGGGCGGAUGAAGGGGGGAAGGUCGUGAGUC</u>caugcccgugu gCUCCGCCAGAGACCAACCGAGAA
SW8.4$_{NC}$	GGGAGACAAGAAUAAACGCUCAACGUCUUC<u>UACAGGGAACCGU</u>gugucauc<u>UG</u>UGAAGUUGUAGAUUCCUAGUUUCGACAGGAGGCUCACAACAGGC
SW8.24$_{NC}$	GGGAGACAAGAAUAAACGCUCAACGA<u>GCUUUUAGUAAGACUCAUGCCGAGAAAUCGGUCGUGAUG</u>cugucgaguucgacagGAGGCUCACAACAGGC
SW8.25$_{NC}$	GGGAGACAAGAAUAAA<u>CGCUCAAG</u>cccgggAAUUGCAUGUUGUGCGUGCCGGGAGUCCAAGUCAGCAUCCUCAUUCGACAGGAGGCUCACAACAGGC
SW10.12$_{MA}$	GGGAGACAAGAAUAAACGCUCAAUCUCGU<u>AUCAGCACACCGGGUACAAAGAGGAUG</u>CAAAUACGCCUGUGACCUUCGACAGGAGGCUCACAACAGGC
SW10.28$_{NC}$	GGGAGACAAGAAUAAAACGCUCAAUUGuua<u>accugguga</u>aaGCAUGUACUGCGAAGUGGAGGGAAUGUGGAAGGGUUCGACAGGAGGCUCACAACAGGC

[a]Ligands are named as in Table 1.

[b]Underlined regions designate the minimal nucleic acid ligand sequence that is predicted to bind to the HIV-1 gag protein, based on the position of the 5′ and 3′ boundaries, which are at the left and right ends of the minimal protein-binding sequence, respectively. Lower case lettering indicates a self-complementary sequence found in nucleocapsid ligands.

Minimum sequences required for ligands to bind gag

In order to identify smaller ligands that might be suitable for structural studies and to attempt to elucidate sequence motifs in common among the ligands, the minimum sequences that are required for ligands to bind to the HIV-1 gag protein with high affinity were determined. This information was obtained through boundary determination and ligand truncation studies. Boundary experiments were performed on nine nucleocapsid ligands and three matrix ligands. The results of these experiments are summarized in Table 2. The length of the predicted minimal ligands ranges from 14 to 59 bases. Seven of the ligands required part of the fixed region sequence to bind to gag, while five required only the initially randomized region to bind to gag.

Minimal nucleocapsid ligand sequences

Mimimal RNA sequences that bind nucleocapsid are predicted to form stable stem–loop structures. The stems often contain purine-rich bulges. However, the lengths of the stems and loops are variable, the positions of the purine-rich bulges are variable, and no consensus sequence was found among the sequences of the minimal nucleocapsid ligands. Therefore, the structure of nucleocapsid ligands may be more important than their sequence.

Nevertheless, a striking feature of the minimal nucleocapsid-binding ligand sequences is that the boundaries of all nine ligands examined end at a self-complementary sequence that ranges in size from 8 to 16 bases (Table 2). In all cases, except one (SW10.28), the self-complementary sequence is located at the 3′ boundary. (The self-complementary sequence in ligand SW10.28 is located at the 5′ boundary.) The lengths of these self-complementary sequences are similar to the binding site size determined for nucleocapsid, which can vary from ~7 to 15 bases, depending on buffer conditions and the protein/nucleotide ratio (21,52). The predicted structure of the minimal ML 8.20 ligand (ML8.20$_{NC}$t14-75) is shown in Figure 7B and is representative of the predicted structures of other minimal nucleocapsid-binding ligands.
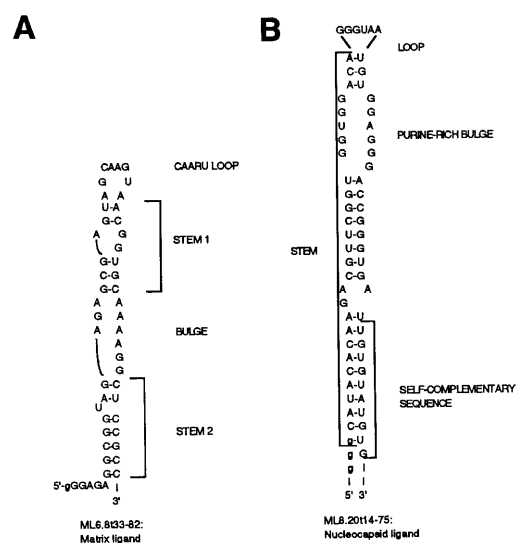

Figure 7. Predicted HIV-1 gag ligand structures. Minimal gag-binding sequences shown in Table 2 were folded using the MulFOLD (version 2.0) program (53). (**A**) Predicted structure of matrix ligand ML6.8t33-82. (**B**) Predicted structure of nucleocapsid ligand ML8.20t14-75.

Five truncated nucleocapsid ligands (ML6.1$_{NC}$t13-63, ML6.6-$_{NC}$t15-74, ML8.20$_{NC}$t14-75, ML8.22$_{NC}$t23-79, SW8.4$_{NC}$t30-53) were generated based on the boundary data shown in Table 2 in order to validate the results of the individual 5′ and 3′ boundary experiments and the predicted minimal ligand structures. The results of binding these truncated ligands to nucleocapsid are shown in Table 3. All five truncated nucleocapsid ligands bound to nucleocapsid. The smallest truncate tested that bound to nucleocapsid was SW8.4$_{NC}$t30-53, which is 29 bases long.

The binding behavior of truncated and mutated ligands highlights the importance of the self-complementary sequence for binding nucleocapsid (Table 3). Truncated ligand ML8.20$_{NC}$t21-66 lacks the self-complementary sequence, but maintains base pairing in the predicted stem and, in fact, adds three additional GC base pairs to the end of the stem. However ML8.20$_{NC}$t21-66 does not bind nucleocapsid. Mutant ligand SW8.4$_{NC}$G47C (Table 3) is a full length version of SW8.4 that has a single point mutation at position 47, which changes a guanine to a cytosine within the

Table 3. Binding properties of truncated and mutant HIV-1 gag ligands

Ligand[a]	Sequence of ligand[b]	Binding[c]
ML6.1$_{NC}$t13-63	gggUCAUACACAAGAACAGUGAUUUGUGUGCUCGUCGGCUGUU<u>UGCAGAUGUA</u>U	+
ML6.6$_{NC}$t54-78	g<u>GGUGGUGCUU</u><u>UGCAGAUGUA</u>UGCUC	-
ML6.6$_{NC}$t15-74	gggAUACACAAGAAGCCAGGUCUUCAAUGGUAGUUGUUGAUUGGUGGGUGCUU<u>UGCAGAUGUA</u>U	+
ML6.8$_{MA}$t33-82	gGGAGAGGCGGUAGAGAGCGAGUAGCAAGUAACGGUGCAAAAGGCUCCGCC	+
ML6.9$_{NC}$t44-68	gggUCGAGGAGAGCG<u>GGUGCAGAUGUAC</u>	-
ML8.20$_{NC}$t21-66	gggAAGACGUUCGGUGGUGGACAGGGUAAUGUGGAGGGACCGGGUGAUUccc	-
ML8.20$_{NC}$t14-75	gggCAUACACAAGACGUUCGGUGGUGGACAGGGUAAUGUGGAGGGACCGGGUGAU<u>UGUGUAUGUG</u>	+
ML8.22$_{NC}$t23-79	gGAUCACGACGGAGGGCGGAUGAAGGGGGGAAGGUCGUGAGUC<u>CAUGCCGUGUG</u>CUCC	+
SW8.4$_{NC}$t35-68	GGGAACCGU<u>GGUGCAUC</u>UGUGAAGUUGUAGAUUC	-
SW8.4$_{NC}$t30-53	Gggag<u>CUACAGGGAACCGU</u><u>GGUGCAUC</u>UG	+
ML6.8$_{MA}$(Δ56-57)	GGGAAAAGCGAAUCAUACACAAGAUCGGUUAUGGAGGAGGCGGUAGAGAGCGAGUAAAGUAACGGUGCAAAAGGCUCCGCCAGAGACCAACCGAGAA	-
SW8.4$_{NC}$G47C	GGGAGACAAGAAUAAACGCUCAACGCUCUUCUACAGGGAACCGU<u>GGUCCAUC</u>UGUGAAGUUGUAGAUUCCUAGUUUCGACAGGAGGCUCACAACAGGC	-
SW8.4$_{NC}$-GGCGCGCC	GGGAGACAAGAAUAAACGCUCAACGCUCUUCUACAGGGAACCGU<u>GGCGCGCC</u>UGUGAAGUUGUAGAUUCCUAGUUUCGACAGGAGGCUCACAACAGGC	-

[a]Ligands are named as in Table 1. The lower case 't' indicates the RNA is a truncated ligand. Numbers following the 't' indicate the nucleotide positions from the full length ligand that are included in the truncated ligand. Mutant ligands are described in the text.

[b]Lower case letters designate nucleotides added to facilitate transcription of the truncated ligand. Self-complementary sequences are underlined.

[c]+, ligand binds to the appropriate HIV-1 protein at least 10-fold better than round 0 RNA; –, ligand does not bind to the appropriate HIV-1 protein any better than round 0 RNA.

self-complementary sequence. This mutation also eliminated binding to nucleocapsid. The fact that matrix ligands and ligands that bind to other proteins (13) do not contain self-complementary sequences at the ends of their boundaries adds more support for the importance of these sequences in the nucleocapsid ligands. However not all self-complementary sequences may allow ligands to bind to nucleocapsid. Mutant ligand SW8.4$_{NC}$-GGCGCGCC contains the self-complementary sequence GGCGCGCC in place of the self-complementary sequence GGUGCAUC, and is not able to bind nucleocapsid (Table 3).

In spite of the apparent importance of self-complementary sequences, these sequences alone may not ensure binding to nucleocapsid. Ligands ML6.6$_{NC}$t54-78, ML6.9$_{NC}$t44-68 and SW8.4$_{NC}$t35-68 (Table 3) contain the self-complementary sequences found at the 3′ ligand boundary, but lack the 5′ side of the predicted stem. None of these truncated ligands binds to nucleocapsid. In summary, a specific self-complementary sequence in the context of a stem–loop structure may be required for efficient binding of the RNAs described here to nucleocapsid.

Different structures and sequences within ψ may be required for distinct dimer formation, gag binding and RNA packaging steps. Guanine quartets (33) and kissing loops (34–39) have been proposed to be important for some of these steps. We considered whether the nucleocapsid-binding RNAs described here could form such structures.

The nucleocapsid ligands do not appear to form guanine quartets for two reasons. First, there were no consistent sequence patterns found in all of the nucleocapsid ligands that would suggest guanine quartets could form. Second, the binding of nucleocapsid ligand SW8.4 to GST–gag was not sensitive to potassium as would be expected if guanine quartets were involved. SW8.4 bound to gag as well in 140 mM LiCl or 140 mM NaCl as in the 140 mM KCl used in the SELEX procedure.

Kissing loops are also not likely to be involved in the binding of SELEX-derived nucleocapsid ligands to gag. While there are sequences within some nucleocapsid ligands that can form kissing loop-like structures (e.g., positions 29–65 of SW8.4), these sequences are not within the minimal nucleocapsid-binding ligand. Also, truncated ligands that lack putative kissing loops can bind to nucleocapsid.

Minimal matrix ligand sequences

A common primary sequence and structural motif could be discerned within the boundaries of the three HIV-1 matrix ligands for which boundaries were determined and in matrix ligand SW8.27, which was not amenable to boundary determination. The sequence of each minimal matrix-binding ligand is predicted to form a stable stem–loop structure. The stem has several nucleotides that are conserved among the matrix ligands and consists of two substems that are separated by a gap or a bulge. A CAARU sequence is found in the sequence of the loop. The predicted structure of ligand ML6.8$_{MA}$t33-82, which is similar to the predicted structure of the other matrix ligands, is shown in Figure 7A. Note that the predicted structures of matrix ligands isolated from SELEX A are circularly permuted relative to those from SELEX B.

Besides sequence similarities, three other experimental results support the structural model shown in Figure 7A. First, the truncated matrix ligand ML6.8$_{MA}$t33-82 (Fig. 7A), which corresponds to the predicted minimal boundaries of ligand ML6.8 (Table 2), binds to the HIV-1 matrix protein with an apparent K_d ~2-fold lower than the full length ML6.8 ligand. Second, deletion of a GC dinucleotide overlapping the CAARU loop motif (GCAAGU) in ligand ML6.8 reduced binding to matrix ~100-fold, to the same as that of round 0 RNA (Table 3). Third, sequence variants are predicted to maintain base pairing in the stem. Ligand ML8.7 differs from ML8.14 at four positions, but each binds with the same affinity to matrix (Table 1). These four differences are predicted to maintain base pairing in stem 1 and as such support the structural model shown in Figure 7A.

DISCUSSION

This report describes RNA ligands that were selected *in vitro* to bind to the HIV-1 gag polyprotein with high affinity. The ligands bind to two different proteins within gag, matrix and nucleocapsid. This result has several implications. First, it demonstrates that ligands which bind to a protein can recognize smaller domains within that protein. This implies that ligands which bind to smaller domains of a protein may also bind to the entire protein.

Second, it demonstrates that there can be more than one nucleic acid binding site on a protein and that one can isolate ligands that bind to those sites within one SELEX experiment. Third, if a protein has more than one RNA binding site, it may be possible to manipulate the frequency at which ligands that bind to the two sites can be isolated.

Matrix and nucleocapsid ligands may be able to bind to gag simultaneously. While the competition binding experiments reported here do not provide rigorous evidence that this can occur they demonstrate that the two classes of RNA ligands do not inhibit each other from binding to gag. The ability of matrix and nucleocapsid ligands to bind simultaneously to gag might be expected since matrix and nucleocapsid are at opposite ends of an elongated protein reported to be ~85 Å long and 35 Å wide (40).

The matrix and nucleocapsid ligands have different sequence and structural motifs. Each of the nucleocapsid ligands contain a self-complementary sequence at one end and are predicted to form stable stem–loop structures. However they do not have substantial primary sequence homology to one another or to matrix ligands. The actual structures formed by the nucleocapsid ligands are not known. They could exist as duplexes hybridized through the self-complementary sequences, solitary stem–loops, duplexes hybridized along the length of the stem–loop, or variations of these structures. In the presence of nucleocapsid protein, some structures may be more stable than others.

The individual roles of the stem–loop and the self-complementary sequences found in the nucleocapsid ligands is not clear, but it seems likely that both are important for binding to nucleocapsid. The presence of the self-complementary sequences in all nine minimal nucleocapsid-binding ligands examined and the identification of a single base change within the sequence of one ligand that eliminates high-affinity binding argues strongly that they are significant for binding. The predicted structure of SELEX-derived nucleocapsid ligands may represent structures that occur in the ψ region that result in nucleocapsid binding or dimerization of ψ-containing RNAs, or they may represent structures that occur during reverse transcription since nucleocapsid is also required for efficient reverse transcription.

In order for HIV RNA to dimerize, an intermolecular interaction is thought to occur between identical stem–loops within the ψ region. The loop contains a six base self-complementary sequence (GCGCGC). These sequences, referred to as 'kissing' loops, can spontaneously and stably hybridize in the absence of nucleocapsid. Once kissing loops have formed, it has been proposed that nucleocapsid melts the stems which formed the loops and reanneals them to form a duplex. This duplex would contain several purine rich bubbles. For SELEX-derived ligands an initial interaction may occur between the self-complementary sequences of two ligand molecules. Then nucleocapsid may accelerate the destabilization of that structure and the formation of the more stable, fully duplex structure. Duplexes formed from nucleocapsid ligands would also contain purine rich bubbles.

The matrix ligands have a primary sequence motif in common among them that is predicted to form a stem–loop structure with a CAARU loop. Matrix ligands do not contain the self-complementary sequences found in nucleocapsid ligands. The ligands reported here could be used to characterize the molecular interactions between matrix and RNA, since at this time the nature of those interactions is not known. In the structures of the HIV-1 (41–43) and SIV matrix (44) proteins there is a patch of basic amino acids near the amino terminus that is located on one

face of the protein. This region serves as the nuclear localization signal of matrix (8). Three other RNA-binding HIV-1 proteins (tat, rev, nucleocapsid) are known to utilize basic amino acids to bind RNA. It is possible that matrix also utilizes this theme. Matrix has been reported to bind double-stranded RNA (45). The physiological relevance of this activity is not known, but matrix is not required for HIV-1 RNA packaging (46). Since the matrix ligands reported here lack any extensive sequence homology to the HIV-1 RNA sequence we have not gained any new insights about the significance of RNA binding by matrix.

Several reports have been published describing methods for directly inhibiting gag protein function. These include the use of sense RNA (47,48), benzamides (49), peptides (50), antibodies (51) and dominant negative proteins (52). The RNAs described here may provide an alternative strategy for directly inhibiting gag protein function. RNAs isolated by the SELEX method could be introduced into or expressed within cells. Gag is an attractive target for inhibition by RNA ligands for several reasons. For example, gag is an RNA binding protein. Therefore, the *in vitro* selection process can result in RNAs that can inhibit critical RNA binding functions of gag. Also since gag and gag–pol are multifunctional, the binding of RNA ligands to them may have a variety of inhibitory effects. Nucleocapsid ligands might prevent RNA packaging. The matrix ligands may provide a novel inhibition strategy. Also, since matrix and nucleocapsid ligands do not interfere with each other for binding to gag it may be possible to use them simultaneously to achieve a synergistic inhibitory effect.

NOTE ADDED IN PROOF

After submission of this manuscript Berglund *et al.* published a paper describing *in vitro* selection of RNAs that bind to the HIV-1$_{NL4-3}$ nucleocapsid protein (*Nucleic Acids Res.* **25**, 1042–1049). A nucleocapsid-binding sequence they identified can be found in several of our nucleocapsid ligands.

ACKNOWLEDGEMENTS

REFERENCES

1 Gorelick,R.J. and Henderson,L.E. (1994) In Myers,G., Korber,B., Wain-Hobson,S., Jeang,K.-T., Henderson,L.E. and Pavlakis,G.N. (eds), *Human Retroviruses and AIDS 1994: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos Natl. Lab, Los Alamos, NM, pp. III-2–III-10.
2 Darlix,J.-L., Lapadat-Tapolsky,M., Rocquigny,H. and Roques,B.P. (1995) *J. Mol. Biol.*, **254**, 523–537.
3 Lapadat-Tapolsky,M., Pernelle,C., Borie,C. and Darlix,J.-L. (1995) *Nucleic Acids Res.*, **23**, 2434–2441.
4 You, J.C. and McHenry,C.S. (1994) *J. Biol. Chem.*, **269**, 31491–31495.
5 Clever,J., Sassetti,C. and Parslow,T.G. (1995) *J. Virol.*, **69**, 2101–2109.
6 Luban,J., Bossolt,K.L., Franke,E.K., Kalpamna,G.V. and Goff,S. (1994) *Cell*, **73**, 1067–1078.

7   Freed,E.O. and Martin,M.A. (1995) *J. Virol.*, **69**, 1984–1989.
8   Buckrinsky,M.I., Haggerty,S., Dempsey,M.P., Sharova,N., Adzhubei,A., Spitz,L., Lewis,P., Goldfarb,D., Emerman,M. and Stephenson,M. (1993) *Nature (London)*, **365**, 666–669.
9   Gottlinger,H.G., Dorfman,T., Sodroski,J.G. and Haseltine,W.A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3195–3199.
10 Huang,L.-M. and Jeang,K.-T. (1995) In Myers,G., Korber,B., Hahn,B., Jeang,K.-T., Mellors,J.W., McCutchan,F.E., Henderson,L.E. and Pavlakis,G.N. (eds), *Human Retroviruses and AIDS 1995: A Ccompilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos Natl. Lab, Los Alamos, NM, pp.III-3–III-9.
11 Pettit,S.C., Moody,M.D., Wehbie,R.S., Kaplan,A.H., Nantermet,P.V., Klein,C.A. and Swanstrom,R. (1994) *J. Virol.*, **68**, 8017–8027.
12 Gold,L., Tuerk,C., Allen,P., Binkley,J., Brown,D., Green,L., MacDougal,S., Schneider,D., Tasset,D. and Eddy,S.R. (1993) RNA: the shape of things to come. In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. 497–510.
13 Gold,L., Polisky,B., Uhlenbeck,O. and Yarus,M. (1995) *Annu. Rev. Biochem.*, **64**, 763–797.
14 Bartel,D.P., Zapp,M.L., Green,M.R. and Szostak,J.W (1991) *Cell*, **67**, 529–536.
15 Giver,L., Bartel,D., Zapp,M., Pawul,A., Green,M. and Ellington,A.D. (1993) *Nucleic Acids Res.* **23**, 5509–5516.
16 Jensen,K.B., Green,L., MacDougal-Waugh,S. and Tuerk,C. (1994) *J. Mol. Biol.*, **235**, 237–247.
17 Tuerk,C., MacDougal-Waugh,S., Hertz,G.H. and Gold,L. (1994) In Ferre,R., Mullis,K., Gibbs,R. and Ross,A. (eds) *The Polymerase Chain Reaction*. Birkhauser, Springer-Verlag, New York, pp. 233–243.
18 Tuerk,C. and MacDougal-Waugh,S. (1993) *Gene*, **137**, 33–39.
19 Schneider,D.J., Feigon,J., Hostomsky,Z. and Gold,L. (1995) *Biochemistry*, **34**, 9599–9610.
20 Tuerk,C., MacDougal,S. and Gold,L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6988–6992.
21 Allen,P., Worland,S. and Gold,L. (1995) *Virology*, **209**, 327–336.
22 Wyatt,J.R., Vickers,T.A., Roberson,J.L., Buckheit,R.W.,Jr, Klimkait,T., DeBaets,E., Davis,P.W., Rayner,B., Imbach,J.L. and Ecker,D. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 1356–1360.
23 Myers,G., Korber,B., Wain-Hobson,S., Smith,R.F. and Pavlakis,G.N. (1993) *Human Retroviruses and AIDS 1993: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos Natl. Lab, Los Alamos, NM.
24 Fitzwater,T. and Polisky,B. (1996) *Methods Enzymol.* **267**, 275–301.
25 Irvine,D., Tuerk,C. and Gold,L. (1991) *J. Mol. Biol.*, **222**, 739–761.
26 Tuerk,C. and Gold,L. (1990) *Science*, **249**, 505–510.
27 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual (second edition)*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
28 Jellinek,D., Green,L.S., Bell,C. and Janjic,N. (1994) *Biochemistry*, **33**, 10450–10456.
29 Luban,J. and Goff,S. (1991) *J. Virol.*, **65**, 3203–3212.
30 Milligan,J.F. and Uhlenbeck,O.C. (1989) *Methods Enzymol.*, **180**, 51–62.
31 Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) *Nucleic Acids Res.*, **15**, 8783–8798.
32 Khan,R. and Giedroc,D.P. (1994) *J. Biol. Chem.*, **269**, 22538–22546.
33 Sundquist,W.I. and Heaphy,S. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3393–3397.
34 Haddrick,M., Lear,A.L., Cann,A.J. and Heaphy,S. (1996) *J. Mol. Biol.*, **259**, 58–68.
35 Laughrea,M. and Jétte,L. (1996) *Biochemistry*, **35**, 1589–1598.
36 Muriaux,D., Girard,P.M., Bonnet-Mathoniere,B. and Paoletti,J. (1995) *J. Biol. Chem.*, **270**, 8209–8216.
37 Paillart,J.C., Marquet,R., Skripkin,E., Ehresmann,B. and Ehresmann,C. (1994) *J. Biol. Chem.*, **269**, 27486–27493.
38 Paillart,J.-C., Skripkin,E., Ehresmann,B., Ehresmann,C. and Marquet,R. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 5572–5577.
39 Skripkin,E., Paillart,J.-C., Marquet,R., Ehresmann,B. and Ehresmann,C. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 4945–4949.
40 Nermut,M.V., Hockley,D.J., Jowett,J.B.M., Jones,I.M., Garreau,M. and Thomas,D. (1994) *Virology*, **198**, 288–296.
41 Hill,C.P., Worthylake,D., Bancroft,D.P., Christensen,A.M. and Sundquist,W.I. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 3099–3104.
42 Massiah,M.A., Starich,M.R., Paschall,C., Summers,M.F., Christensen,A.M. and Sundquist,W.I. (1994) *J. Mol. Biol.*, **244**, 198–223.
43 Matthews,S., Barlow,P., Boyd,J., Barton,G., Russell,R., Mills,H., Cunnigham,M., Meyers,N., Burns,N., Clark,N., *et al.* (1994) *Nature (London)*, **370**, 666–668.
44 Rao,Z., Belyaev,A.S., Fry,.E., Roy,P., Jones,I.M. and Stuart,D. (1996) *Nature (London)*, **378**, 743–747.
45 Leis,J., Scheible,P. and Smith,R. (1980) *J. Virol.*, **35**, 722–731.
46 Lee,P.P. and Linial,M.L. (1994) *J. Virol.*, **68**, 6644–6654.
47 Berkout,R. and van Wamel,J.L. (1995) *Antiviral Res.*, **26**, 101–115.
48 Cohli,H., Fan,B., Joshi,R.L., Ramezani,A., Li,X. and Joshi,S. (1994) *Antisense Res. Dev.*, **4**, 19–26.
49 Rice,W.G., Supko,J.G., Malspeis,L., Buckheit,R.W., Clanton,D., Bu,M., Graham,L., Schaeffer,C.A., Turpin,J.A., Domagal,J., *et al.* (1995) *Science*, **270**, 1194–1197.
50 Niedrig,M., Gelderblom,H.R., Pauli,G., Marz,J., Bickhard,H., Wolf,H. and Modrow,S. (1994) *J. Gen. Virol.*, **75**, 1469–1474.
51 Tanchou,V., Delaunay,T., De Rocquigny,H., Bodeus,M., Darlix,J.-L., Roques,B. and Benarous,R. (1994) *AIDS Res. Human Retroviruses*, **10**, 983–993.
52 Trono,D., Feinberg,M. and Baltimore,D. (1989) *Cell*, **59**, 113–120.
53 Zuker,M. (1989) *Science*, **244**, 48–52.