Published in final edited form as:

Anal Biochem. 2011 July 15; 414(2): 246–253. doi:10.1016/j.ab.2011.03.018.

A primer-free method that selects high affinity single stranded DNA aptamers using Thermostable RNA Ligase

Yi-Tak Lai and Jeffrey J. DeStefano*

Department of Cell Biology and Molecular Genetics, University of Maryland College Park, College Park, MD 20742

Abstract

This report describes a method for selecting single stranded DNA (ssDNA) molecules that bind with high affinity (aptamers) to specific target proteins. This SELEX (Systematic Evolution of Ligands by EXponential enrichment) method is similar to other “primer-free” approaches where the random sequence ssDNA starting pool has no fixed sequences at the 5′ and 3′ termini. Therefore there are no predetermined sequences that could bias selection. Like other SELEX methods, repeated cycles (typically 5–15) of selection then amplification and re-selection are used. The method differs from other primer-free approaches in that the key step for regenerating new material for subsequent rounds is ligation of the selected ssDNA to a defined sequence oligonucleotide using Thermostable RNA Ligase. Under specific conditions, this ligase ligated 30 nucleotide random sequence ssDNA (5′-N₃₀-3′) to a specified 20 nucleotide ssDNA with ~50% efficiency. Efficiency was improved to ~90% by addition of a single T residue to the 3′ end (5′-N₂₉T-3′). High efficiency in this step is critical, especially early in the procedure as any selected material that is not ligated is lost. In this report, human immunodeficiency virus reverse transcriptase was used as the target protein but the method could be applied to essentially any protein.

Keywords

SELEX; HIV reverse transcriptase; Primer-free SELEX; Aptamer

Introduction

Systematic evolution of ligands by exponential enrichment (SELEX) is a method for selecting small nucleic acid (RNA or DNA) molecules that bind with high affinity to specific targets. It was originally developed by three separate groups in the early 1990s [1; 2; 3]. The selected nucleic acids are called “aptamers”. Though aptamer targets are most often proteins, they also includes other small molecules, cultured cells, and viruses among others [4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16]. Aptamers have many potential uses including therapeutic and diagnostic applications, as well as replacement of antibodies in ELISA and other technologies [17; 18; 19; 20; 21; 22; 23; 24; 25]. Essentially, any technique that involves highly specific recognition of protein molecules could benefit from

© 2011 Elsevier Inc. All rights reserved.

*Corresponding Author: Department of Cell Biology and Molecular Genetics, University of Maryland College Park, Biosciences Research Bldg., College Park, MD 20742, (P) 301-405-5449, (F) 301-314-9489, jdestefa@umd.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

aptamer-based approaches. Although several compounds developed from SELEX-based methods are currently in clinical trials, so far only a single drug for macular degeneration, Macugen (EyeTech Pharmaceuticals Inc. New York, NY USA), has been approved by the Food and Drug Administration (FDA).

Most aptamers are identified through standard SELEX procedures that include a starting pool of nucleic acids with a region of random nucleotides of approximately 20–50 bases long. The random nucleotides are flanked on the 3' and 5' ends by fixed sequences of approximately 20 nucleotides. This results in oligonucleotides of ~60–90 nucleotides that are used for the selection process. After several rounds of selection tight binding aptamers that contain the flanking sequences are selected. Typically, these sequences are modified to determine the minimal bases that are required for tight binding. A major advantage of this approach is that the flanking sequences provide a simple means to amplify oligonucleotides that are selected in each round of selection. A clear disadvantage is the potential for the fixed sequences to bias the selection by interacting with nucleotides in the random region of the oligonucleotide. Many, if not most selected aptamers require sequences from the fixed flanking regions to maintain tight binding.

Recently, new methods have been published for “primer-free” SELEX, in which selection occurs in the absence of any flanking, fixed nucleotides [16; 26; 27], and “minimal primer” SELEX, in which the fixed sequences are reduced to fewer nucleotides [28; 29], or other approaches in which primer ends are block or switched during the selection process [30]. By eliminating fixed flanking regions (completely in the case of primer-free SELEX) during the selection step these methods abrogate the biases noted above. The trade-off is the methods are generally more times consuming than fixed sequence approaches, as several enzymatic and hybridization steps are involved to introduce flanking sequences after the selection step. These additional steps can introduce biases related to enzymatic reactions, and the additional steps can lead to loss of some material. The latter problem is especially important in the first round of selection where all or most sequences are unique.

In this report we introduce a new primer-free method for selecting DNA aptamers. The method takes advantage of a commercially available Thermostable RNA Ligase that functions at 60°C. Under the right conditions as defined in this report, the ligase is capable of ligating DNAs with high efficiency (between ~50% and >90% depending on the sequence used), more efficiently than T4 RNA Ligase 1, which is commonly used for this application. Performance of ligations at high temperature also minimizes the potential effects of second structures on ligation. This method, although not solving all the problems inherent in primer-free SELEX, provides another tool for selection of high affinity aptamers using an approach that does not bias the outcome with fixed sequences.

Materials and Methods

Materials

Human immunodeficiency virus reverse transcriptase (HIV-RT) was from Worthington Biochemical Corporation. *Taq* polymerase, restriction enzymes, Klenow DNA polymerase, T4 polynucleotide kinase (PNK) and T4 RNA Ligase 1 were from New England Biolabs. Thermostable RNA Ligase and CircLigase™ were from Epicentre Biotechnologies. Calf intestinal alkaline phosphatase (CIAP) and dNTPs were from Roche. Rapid DNA Ligation kits were from Promega. Nitrocellulose filters were from Whatman. *Pfu*Turbo® DNA polymerase and StrataClone™ Blunt PCR cloning kits were from Agilent Technologies. Miniprep DNA preparation kits were from Qiagen. Radiolabeled compounds were obtained from Perkin Elmer. Sephadex G-25 spin columns were from Harvard Apparatus. All oligonucleotides were from Integrated DNA Technologies. The 30 nucleotide random

sequence oligomer (5'-N₃₀-3') used as a starting pool for selection was synthesized using "hand-mixed" ratios of the four nucleotides in order to assure an equal probability of incorporation of each nucleotide at each position. This material was used directly in the SELEX protocol. All other chemicals were from Sigma, Fisher Scientific, or VWR Scientific.

Methods

5' end-labeling of oligonucleotides—Reactions for labeling of the starting material (5'-N₃₀-3') were done in a 100 µl volume containing 1500 pm of 5'-N₃₀-3', 70 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 5 mM dithiothreitol (DTT), 10 µls of γ-³²P ATP (3000 Ci/mmol, 10 µCi/µl) and 2 µls (20 units) of T4 polynucleotide kinase. Because of the large excess of oligonucleotide over label, only a small fraction of the primers have phosphate added to the 5' end. This limits circularization of this oligonucleotide and favors ligation to oligonucleotide 1 (see below) in the ligation reaction with Thermostable RNA Ligase. The reaction mixture was incubated for 30 minutes at 37°C and then the PNK was heat inactivated for 10 minutes at 70°C according to manufacturer's recommendation. The material was then run through a Sephadex G-25 spin column.

Labeling of oligo 1 (5'-/5'Phos/GTCGACGAGGTCTAGAATA/3'dideoxyC/- 3', MnII site underlined) was performed as above, except that 200 pmoles of oligo 1 (without a phosphate at the 5' end) were incubated in a 50 µl volume as above for 15 minutes. After this non-radioactive ATP was added to a final concentration of 200 µM and the reaction was continued for 20 minutes then processed as described above.

Selection of material with HIV-RT using nitrocellulose filters for rounds 1 and 2 and subsequent processing—In the initial round, 3 separate reactions with approximately 500 pm each of a 5' ³²P end labeled 30 nucleotide oligonucleotide of random sequence (5'-N₃₀-3') and 10 pm of HIV-RT were prepared in 50 µl of buffer containing 50 mM Tris-HCl (pH=8), 80 mM KCl, 6 mM MgCl₂, and 1 mM DTT. The reactions were incubated at room temperature for 1 hour, then filtered over a nitrocellulose filter disk (25 mm, 0.2 µm pore size) with suction. Disks were preincubated in the above buffer prior to use, then washed twice with 500 µl of buffer containing 25 mM Tris-HCl (pH=8) and 10 mM KCl. Filter disks were then cut into small pieces and placed in 500 µl of 50 mM Tris-HCl (pH=7) and 1 mM EDTA (pH=8). The material was then incubated at 90°C for 5 minutes then 70°C for 10 minutes. After allowing the material to cool, it was extracted with 500 µl of 25:24:1 phenol: chloroform: isoamyl alcohol, followed by 500 µl of chloroform. The aqueous phase was precipitated by adding 50 µl of 3 M sodium acetate (pH=7), 40 µg glycogen, and 1 ml of ethanol. Precipitated material was collected by centrifugation after incubation at -20°C for several hours. The pellet was washed with 70% ethanol, then dried in a DNA SpeedVac Concentrator. The precipitated material was resuspended in 75 µl of water and passed over Sephadex G-25 spin columns equilibrated in water. Collected material was then dried. Nucleic acid from all 3 reactions (~20 pm total) was then combined and re-selected as above but in a 20 µl final volume using 2 pm of HIV-RT. This step was necessary to reduce the total amount of nucleic acid for the ligation reactions. Nucleic acid (~0.3 pm total) was extracted as described above, except glycogen was omitted from the precipitation step and 30 pm of 5' ³²P end labeled oligo 1 (see above) was added. The collected precipitate was once again run over a Sephadex G-25 spin column as described above and dried. The dried material was resuspended in 21 µl of water and ligations were performed in 3 separate ligation reactions using Thermostable RNA Ligase. All reactions were in 20 µl final volume containing: 50 units (0.5 µls) Thermostable RNA Ligase, ~10 pm 5' ³²P end labeled oligo 1, ~0.1 pm 5'-N₃₀-3', 50 mM MOPS (pH=7.5), 5 mM MgCl₂, 10 mM KCl, 1 mM DTT, 50 µM ATP, 1 mM hexamine cobalt chloride ([Co(NH₃)₆]Cl₃), and

20% (w:v) PEG-8000. Reactions were incubated 12–18 hours at 60°C in a thermocycler with a heated lid. After ligation reactions, 20 µl of 2X loading buffer (90% formamide, 20 mM EDTA (pH=8), 0.025% (w:v) each of bromophenol blue and xylene cyanol) was added and samples were run on a 12% polyacrylamide/7M urea denaturing gel (in 1X Tris/borate/EDTA buffer) with a molecular weight marker as described [31]. Ligated material (50 nucleotides in length) was located using a phosphorimager (Fujifilm FLA7000), excised, and then eluted by the crush-and-soak method [31] in 500 µl 50 mM Tris-HCl (pH=7), 1 mM EDTA (pH= 8). Forty µg of glycogen was added to the eluate and the nucleic acid was recovered by ethanol precipitation as described above. The 5' ends of the recovered nucleic acid were phosphorylated in a standard PNK reactions as described above, except the total volume was 25 µl, 20 mM KCl was included, radiolabeled ATP was replaced with 0.4 mM cold ATP and 10 units of PNK were used. After inactivating the PNK (see above), 10 µg of oligo 2 (5'-CTGGTATTCTAGACCTCGTC-3', MnlI site underlined) was added. This oligo was complementary to a region on oligo 1 (see Fig. 1). Oligo 2 was hybridized to the ligated product by heating to 65°C, and slow cooling to room temperature. Then an extension reaction was performed with Klenow DNA polymerase (1 unit total) by adding dNTPs (final concentration 100 µM) and bringing the volume up to 50 µl, followed by incubation at 37°C for 10 minutes. The reaction volume was then increased to 100 µls with water and the material was extracted and precipitated. The recovered double stranded DNA was blunt-end ligated to a duplex formed using oligo 3 (5'-ATAGCATGAATTCCCGAAGACGC-3', BbsI site underlined) and oligo 4 (5'-GCGTCTTCGGGAATTCATGC-3', BbsI site underlined) using a rapid DNA ligation kit and T4 DNA ligase. Two pmoles of the oligo 3-oligo 4 duplex was used in the round 1 reactions, and 1 pmole was used in subsequent rounds. The manufacturer's recommended protocol was used in the rapid ligation reactions which were incubated for 30 minutes at room temperature. The entire 10 µl ligation reaction (round 1) or 5 µl of the reaction (all subsequent rounds) was used in a polymerase chain reaction (PCR). Reactions were performed in the manufacturer's supplied buffer in 100 µl total volume, and included fifty pmoles each of oligo 2 and oligo 5 (5'-GCATGAATTCCCGAAGACGC-3'), 200 µM each of dATP, dCTP, and dGTP, 25 µM dTTP, 2 µls of α -³²P dTTP (800 Ci/mmol, 10 µCi/µl), and 2.5 units of *Taq* polymerase. A total of 15–18 cycles was performed using 94°C (30 seconds), 50°C (1 min), and 72°C (1 min) followed by an additional cycle of 5 min at 72°C. Native gel loading buffer (20 µl: 40% sucrose, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) was added and reactions were run on a 12% native polyacrylamide (29:1 acrylamide:bisacrylamide) gel and products of the correct size (73 base pairs) were located and recovered as described above using a phosphorimager. Approximately 1/40th (for each reaction) of the recovered material was used in a second set of 4 PCRs. The PCRs were performed as described above for 6, 8, 10, or 12 cycles. A range of cycles was performed to insure that a sufficient quantity of the correct product was recovered. If product recovery was too low, another PCR was performed using an optimal cycle number based on the results with the variable cycle reactions. Because of the random 30 base region of the product, too many cycles results in the generation of multiple single strands that do not form complete duplexes. Products of the correct size from the 4 PCRs were recovered from 12% native gels as described above. The eluted material was precipitated and digested in 30 µl total volume with 7.5 units each of BbsI and MnlI in the manufacturer's buffer for BbsI for 2 hours at 37°C. The sample was heated to 90°C for 3 min, then cooled rapidly on ice. Four units of CIAP were added and incubation was continued at 37°C for an additional 30 min. Twenty-five µl of 2X loading buffer was added and the sample was run on a 12% denaturing gel as described above. A marker for the 30 nucleotide single stranded dephosphorylated product was included on the gel. The correct product was located with a phosphorimager and recovered using 40 µg of glycogen to aid precipitation as described above. The recovered material (~10–20 pmoles) was used for a second round of selection with nitrocellulose filters using the same protocol with the changes noted.

Selection procedure for all rounds after round 2 performed by gel-shift—

Following the second round, subsequent rounds were conducted using native 6% polyacrylamide gels to recover the selected material by gel-shift. The recovered 30 nucleotide product from the previous round was incubated with HIV-RT (see below) in 10 μ l final volume using the conditions stated above for 30 minutes. Two μ l of native gel loading buffer was then added and the material was run on a 6% native polyacrylamide gel at ~100 V until the blue dye marker was ~2–3 inches below the wells. Gel-shifted nucleic acid was located with a phosphorimager and eluted as described above. Ten pmoles of radiolabeled oligo 1 was added to the eluate and precipitation in the absence of glycogen was carried out as described above. The recovered material was processed as described above to regenerate the 30 nucleotide random nucleic acid for the next round of selection. After round 6, the KCl concentration in the selection step was increased from 80 to 150 mM. Various amounts of HIV-RT were used in the selection process. Our goal was to recover approximately 1/25th of the total nucleic acid material from each round as shifted material on the gels. In round 3, the first round where the gel-shift was used, 5 pmoles of HIV-RT was added to the selection reaction. In subsequent rounds the level was decreased as the amount of recovered material from the previous round increased. By round 6, 0.5 pmoles of HIV-RT was being used and this amount was used in all subsequent rounds.

Sequences analysis of products recovered from selection rounds 6, 11, and 12—

The sequence of the 30 nucleotide random region from products selected in rounds 6, 11, and 12 was determined by performing a PCR with *Pfu*Turbo® DNA polymerase, and then cloning the 73 base pair PCR product using StrataClone™ Blunt PCR cloning kit from Agilent Technologies and the manufacturer's suggested protocol. Mini-preps were prepared by the QIAprep Miniprep Kit from Qiagen. Products were sequenced by Genewiz Inc.

Determination of equilibrium dissociation constants (K_d)—Selected material (~2 nM) from various rounds of selection (see Results) or other designed sequences (see Fig. 5) were either labeled internally, or 5' end labeled with ³²P and mixed with various amounts of HIV-RT (as stated in Fig. legends) in 10 μ l of buffer containing 50 mM Tris-HCl (pH=8), 1 mM DTT, 80 mM KCl, and 6 mM MgCl₂, for 10 min at room temperature. Two μ l of 6X native loading buffer was added and the samples were run on a 6% native polyacrylamide gel as described above for the gel-shift experiments. The amount of shifted vs. non-shifted material was quantified using a phosphorimager. Values for K_d were determined by plotting the concentration of shifted product (nM) vs. the concentration of HIV-RT and fitting the data by nonlinear least square fit to the quadratic equation: $[ED] = 0.5([E]_t + [D]_t + K_d) - 0.5(([E]_t + [D]_t + K_d)^2 - 4[E]_t[D]_t)^{1/2}$, where $[E]_t$ is the total enzyme concentration and $[D]_t$ is the total primer-template concentration [32]. In cases where the affinity of HIV-RT for the nucleic acid was low ($K_d \geq 1 \mu$ M), a more precise K_d value was not determined.

Results and Discussion

The efficiency of the selection process and its affect on the recovery of tight binding aptamers

In the initial selection round with an oligonucleotide of 30 bases (5'-N₃₀-3', see Fig. 1), nearly all nucleic acid sequences are unique. With 500 pmoles (5×10^{-10} moles) of starting material, only ~0.026% of the total number of possible sequences is represented ($((5 \times 10^{-10}) \times (6.023 \times 10^{23}))/4^{30} \times 100$). On the surface this would suggest that many, if not most sequences that may bind very tightly to a protein, are unlikely to be present. However, proteins often recognize shorter nucleotide runs or specific structures that can form from a diverse array of sequences. Even if a particular protein bound tightly to a run of 20 specific nucleotides, such a run would be represented ~270 times in the 500 pmoles pool ($((5 \times$

$10^{-10}) \times (6.023 \times 10^{23})/4^{20}$). Because of the way proteins recognize nucleic acids, it is highly likely that most of the possible recognition motifs/sequences that could form with a nucleic acid that is 30 nucleotides or shorter would be present in a 500 pmoles starting pool. More complex structures, for example duplexes resembling primer-templates with recessed 3' termini, or nucleic acids that can simultaneously interact with more than one domain of the protein, are less likely to occur in the starting material, mostly because of the short 30 nucleotide length.

Assuming that the starting pool has reasonable but not complete diversity, it is important that those oligos in the pool that can bind with high affinity to the protein (HIV-RT in this case) are recovered. Since nearly all the sequences in the pool are unique, any sequence not recovered from round 1 is essentially lost. This is not a major issue in subsequent rounds since PCR amplification after round 1 generates many copies of each recovered sequence.

In the protocol used here the first and second rounds of selection were performed using nitrocellulose filters, while subsequent rounds were performed using gel-shift (Fig. 1). Filter selection was faster, allowing rapid processing of larger amounts of material and high recovery. For these reasons it was advantageous to use this approach, at least for rounds 1 and 2. Later rounds used gel-shift because of preference to this approach due to lower background and higher selectivity. However, many proteins are unable to shift nucleic acids on gels. In such cases it is feasible to continue with filter selection, and we are currently doing this for some other proteins being tested in this protocol (data not shown).

Even with the high recovery efficiency of the nitrocellulose approach, it is impossible to achieve 100% recovery of tight binding sequences. To minimize the potential loss of high affinity sequences, for round 1 three separate selections with 500 pmoles each were conducted then combined. The combined material was reselected in order to decrease the total amount of material for the first ligation step. This was necessary to achieve the highest efficiency in the ligation reaction which required a high ratio of donor (5' phosphorylated oligo 1) to acceptor (selected nucleic acid). Since the amount of recovered material was still relatively high after the two sequential selections, three separate ligations with Thermostable RNA Ligase were performed (see steps between products A and B in Fig. 1). An experiment was performed to test the efficiency of the ligation process. A drawback of using T4 RNA Ligase 1 in a DNA selection is its relatively low efficiency at ligating DNA. The efficiency can be improved using hexaminecobalt chloride and polyethylene glycol [33]. Even using these optimal conditions, T4 RNA Ligase 1 only ligated ~30% (see Fig. 2 legend) of the 5'-N₃₀-3' DNA to oligo 1 to produce the specific 50 nucleotide ligation product as shown in the experiment in Fig. 2A. Thermostable RNA Ligase was more efficient ligating about 50% of the material. Thermostable RNA Ligase also produced less nonspecific ligation products than T4 RNA Ligase 1. A second thermostable ligase, CircLigase™, was about as efficient as T4 RNA Ligase 1, though as expected it did circularize the 5'-N₃₀-3' DNA with higher efficiency (data not shown). We should note that in order to make detection and quantification of ligated products easier, the ligation tests shown in Fig. 2A (and Fig. 5 below) were conducted using 5' end-labeled (and thus 5' phosphorylated) 5'-N₃₀-3' and unlabeled oligo 1, while in the ligation reactions used for selection the 5' end of 5'-N₃₀-3' was not phosphorylated. Phosphorylating the 5' end of the acceptor increases the likelihood of circularization, which would decrease the level of acceptor available for ligation to oligo 1. This could in turn lead to an underrepresentation of the actual level of ligated product in the selection reactions. Although some circularization did occur (see Fig. 2A), ligation to oligo 1 was more pronounced, probably because of the 100-fold excess of the oligo 1 donor to the 5'-N₃₀-3' acceptor. Also, this approach did not seem to misrepresent ligation efficiency in the normal selection reactions. This was confirmed by conducting a ligation reaction with non-phosphorylated 5'-N₃₀-3' and oligo 1, then end-labeling the resulting

products in an exchange-reaction with PNK (Fig. 2B). The proportion of correctly ligated (50 nucleotide product) vs. non-ligated material was similar (~50% efficiency) in this experiment in comparison to those with 5' end-labeled 5'-N₃₀-3'. Although circularization did not significantly interfere with quantification of ligation using 5'-N₃₀-3', it may interfere with oligos of a specific sequence or with longer oligos that could circularize more easily. Recently published primer-free or minimal primer (2 fixed nucleotides at the 3' and 5' ends flanking a random region) approaches that used a more complex ligation step with a "bridging" nucleic acid and T4 DNA ligase report >90% efficiency for the ligation step of the minimal primer approach, while the efficiency in the primer-free approach was not specified [27]. A second method termed "primer-free genomic SELEX" reported somewhat lower efficiencies [16]. Given the complexities of these various methods and the one used in this manuscript, it is difficult to compare their overall efficiency by comparing results from a single step or even multiple steps in the protocols. It would be interesting to compare the results of these approaches to see if they select similar aptamers.

Although the efficiency of the ligation step results in a significant loss of diversity in round 1, some of this loss is made up by conducting multiple selections and ligations in the first round. As noted above, this issue is less important in subsequent rounds. There is also the possible loss of diversity due to the Thermostable RNA Ligase having a bias for ligating specific sequences or nucleic acids with less structure. T4 RNA ligase has been shown to ligate specific DNA sequences preferentially [34]. Although this is also the case with Thermostable RNA Ligase (see below), the higher ligation efficiency suggests that the bias may be less pronounced. Inhibition of ligation because of high structure is also likely to be less pronounced because the reactions are performed at 60°C.

Following recovery of the ligated 50 nucleotide product (Fig. 1, product B), a set of reactions was performed to produce a blunt-ended 53 base pair duplex, with a 5' phosphate at the end containing the random region (Fig. 1, product C). The 3' end of the upper strand in product C is derived from Oligo 1, however, the 3' dideoxycytidine residue is replaced by a deoxycytidine and is followed by a 5'-CAG-3' sequence. This depiction was based on the assumption that the 3'-5' exonuclease activity of the Klenow polymerase would excise the dideoxy residue and add four template-directed nucleotides. Oligo 2, which was used to prime the step with Klenow, was not complementary to the first three 5' nucleotides of oligo 1. This is due to a previous version of this oligo with complete complementarity produced primer dimers in the subsequent PCRs. The 5' phosphate at the random end of product C was added by PNK before the Klenow step. It directed ligation by T4 DNA ligase of the oligo 3-oligo 4 duplex to this end. The oligo 3-oligo 4 duplex was staggered in order to force blunt end ligation in the correct orientation (Fig. 1, product D). The blunt-end ligation step using the rapid ligation kit and a vast excess of the oligo 3-oligo 4 duplex typically showed a ligation efficiency of >60% (data not shown).

After producing product D, an initial PCR step was performed for 15–18 cycles (see Methods). Material from this step was isolated on a native gel. About 1/40th of the total recovered material was used in a second PCR step that was performed in 4 separate reactions using a range of cycles (see Methods). Amplification steps must always be performed carefully in SELEX protocols, as too many cycles produce products that cannot form a complete duplex, due to the random region while too few cycles produce a low yield. This is easily followed on a native gel, where the desired product appears and builds up over a range of cycles then is lost as too many cycles are performed. Although the approach used here was effective, in many cases a single PCR step with multiple reactions conducted for 15–18 cycles (depending on the level of the starting material) would produce sufficient yield for the next round.

Cleavage of the PCR products to reproduce the 5'-N₃₀-3' product (Fig. 1, product E) was highly efficient (Fig. 3). Note that the system can be designed to select single stranded products in a wide range of sizes, provided that cleavage of the PCR product yields a product of unique size for the selected material. This may require using primers in the PCR step, or linkers that are different sizes than the ones employed here.

The analysis of the SELEX procedure by monitoring binding to HIV-RT, and the nature of the recovered aptamers

Thirty nucleotide material (Fig. 1, product E) from select rounds was analyzed in a gel-shift assay to estimate binding affinity for HIV-RT (Fig. 4). Results were compared to the starting material and K_d values were estimated. Note that this approach does not yield a true K_d value, as gel-shift is not an equilibrium binding method. However, the derived values are useful in comparing the binding affinity of material from the various rounds. The starting material bound weakly to HIV-RT with K_d value of $\sim 1 \mu\text{M}$. Attempts to measure the value more precisely were not performed as large amounts of enzyme would be required, and this resulted in a smearing effect in the gel-shift. The K_d values for 30 nucleotide material selected in rounds 2, 6, 10 and 11 were $\sim 1 \mu\text{M}$, $170 \pm 56 \text{ nM}$, $191 \pm 42 \text{ nM}$, and $142 \pm 31 \text{ nM}$ (rounds 6, 10, and 11 are an average of 3 exp. $\pm$ standard deviation), respectively. Interestingly, the affinity did not increase appreciably after round 6, despite five more rounds of selection at higher salt concentrations (see Methods). A second independent SELEX procedure was also performed. The K_d value after 9 rounds was similar to the value for the round 11 material in the first SELEX and the sequences recovered were nearly identical (see below).

Aptamer sequences recovered from round 11 of the first SELEX and round 9 of the second were nearly identical and had little secondary structure

Ten clones were sequences from the round 11 material. Only two different sequences (denoted R11-1 and R11-2 in Table 1) were recovered and both were closely related, differing only by the first nucleotide. These same sequences were also the predominant oligos found in material sequenced from round 6, which explains the similar affinity of material from these rounds for HIV-RT. Gel-shift assays with each sequence revealed that they had similar affinity for HIV-RT. Interestingly, using the DNA mfold program [35], both sequences had a very low level of predicted secondary structure with the most stable predicted structures having ΔG values of -0.61 kcal/mole for both sequences. A possible explanation for this is that Thermostable RNA Ligase preferentially ligates material with low structure (see below). This would seem to make sense as availability of the 3' end of the nucleic acid would be pivotal in the ligation step. However, this factor should be mitigated to some extent by the 60°C reaction temperature. Also notable was the regular arrangement of 4 separate G dinucleotides in the sequences, which suggests the possible formation of a higher ordered structure. Several sequences closely related to R11-1 and R11-2 were also recovered from round 9 of the second SELEX experiment (Table 1). Experiments are currently being conducted to determine why the selected aptamers bind tightly to HIV-RT.

The ligation efficiency of Thermostable RNA Ligase improves with low structure and a T residue at the 3' terminus of the acceptor

The significant increase in binding affinity of the round 11 selected material to HIV-RT indicated that the RT binding step likely played a major role in the selection process for this protocol. However, this does not rule out the possibility that Thermostable RNA Ligase may have also influenced selection. To test this further, the ligation efficiencies of the starting material and material selected from rounds 2, 6 and 11 were tested (data not shown). Starting material (see above) and material from round 2 each showed $\sim 50\%$ ligation

efficiency in standard 8 hour ligation reactions, while efficiency was greater than 90% for round 6 and 11 material.

The more efficient ligation of the selected material suggests that Thermostable RNA Ligase may have played some role in the selection process. However, it is also possible that material that binds tightly to HIV-RT just happens to ligate efficiently because of other properties. For example, low structure may favor binding to RT by allowing the aptamer to be flexible, while it also promotes higher ligation efficiency. Also, it was notable that all the aptamers recovered from round 11 (and round 6 (data not shown)) terminated with a T residue at the 3' end. Several oligonucleotides were designed to try and determine why the selected material ligated more efficiently. These included homopolymeric 30-mers composed of A (5'-dA₃₀-3') or T (5'-T₃₀-3') which presumably contain little or no secondary structure (Fig. 5A). A set of 30-mer oligos with sequences related to the round 11 selected material but with the 3' terminus either protruding or bound in a stem structure (3'-overhang and 3'-recessed, respectively); and finally, a highly structured sequence 30 nt G-C rich sequence (G-C flip-back). Ligation reactions for these oligos and the R11-1 aptamer and 5'-N₃₀-3' starting material were performed for 30 min, 2 hr and 8 hr and plotted as shown in Fig. 5B. Both R11-1 and 5'-T₃₀-3' ligated with very high efficiency, reaching ~70% ligated product by 30 min and over 90% by 2 hours. The ligation advantage was not due solely to a lack of structure, as 5'-dA₃₀-3' ligated with much lower efficiency that was comparable to 5'-N₃₀-3'. Interestingly, the 3'-overhang and 3'-recessed oligos ligated better than starting material or 5'-dA₃₀-3' and were comparable, despite the latter essentially "protecting" the 3' end in a secondary structure. Finally, the highly structured G-C flip-back ligated very poorly, although this may have been at least partially due to the 3' G residue (see below).

To test the effect of a T residue at the 3' end of the acceptor oligo, three derivatives of R11-1 in which the 3' terminal nucleotide was replaced with A, G, or C were also tested in ligation reactions along with R11-1 and a version of the random starting material with a T at the 3' end (5'-N₂₉T-3') (Fig. 5C). While oligos terminating in T, A, or C all showed greater than 90% ligation efficiency in 8 hour reactions, the T terminated oligo ligated far more efficiently at early time points followed by A then C in ligation efficiency. The version of R11-1 with a 3' G residue ligated much more poorly than the others. The importance of a T residue at the 3' end is emphasized most strongly by ligation experiments with 5'-N₂₉T-3' (Fig. 5C, filled squares). This oligo differed from the starting material (5'-N₃₀-3') by just the addition of a single 3' T residue. Ligation with 5'-N₂₉T-3' essentially mimicked that with R11-1(A) and reached a level of ~90% after 8 hours while 5'-N₃₀-3' was only between 50–60% (Fig. 5B, filled circles) at this time point. This finding suggests that it may be advantageous to start the selection process with random oligos that include a fixed T residue at the 3' end. This would decrease the loss of diversity in the early rounds of the protocol, though it would also create some bias in the results.

The above results indicate that Thermostable RNA Ligase showed a clear preference for a T residue at the 3' end of the acceptor, at least for ligation to the oligo 1 donor that was used here. Low structure was also advantageous, as modestly structured oligos closely related to R11-1 (3'-recessed and 3'-overhang) ligated with lower efficiency. Although strong binding to HIV-RT clearly played an important role in the selection process, it is not clear if, and to what extent Thermostable RNA Ligase sequence/structural preferences influenced the process. Selections using this method with other proteins are ongoing and should help in answering this question. The fact that only sequences with a 3' terminal T were recovered suggests that the ligase played at least some role, as deletion or replacement of the T does not affect binding to RT (data not shown). If this method preferentially selects low structure aptamers then it may be particularly useful for finding aptamers that could potentially be

used as inhibitors in cells, as highly structured nucleic acids have a tendency to activate innate cellular responses [36].

Conclusions and significance

Just a few protocols for primer-free SELEX are currently described in the literature (see Introduction). Primer-free SELEX is advantageous because it allows selection of nucleic acid aptamers from random pools without interference from fixed flanking regions. Though all these methods are considerably more complex than SELEX approaches using flanking fixed sequences, the method proposed in this report takes advantage of a new Thermostable RNA ligase. In comparison to other primer-free methods, this simplifies the procedure for regenerating nucleic acids for subsequent rounds of selection. In practice each primer-free method may have specific advantages or biases and it would be interesting to compare the aptamers selected from the various methods using the same target protein.

Acknowledgments

This work was supported by National Institute of General Medicine grant number GM051140.

References

1. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. *Nature*. 1990; 346:818–822. [PubMed: 1697402]
2. Joyce GF. Amplification, mutation and selection of catalytic RNA. *Gene*. 1989; 82:83–87. [PubMed: 2684778]
3. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. 1990; 249:505–510. [PubMed: 2200121]
4. Blank M, Weinschenk T, Priemer M, Schluesener H. Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels, selective targeting of endothelial regulatory protein pigpen. *J Biol Chem*. 2001; 276:16464–16468. [PubMed: 11279054]
5. Cerchia L, Duconge F, Pestourie C, Boulay J, Aissouni Y, Gombert K, Tavitian B, de Franciscis V, Libri D. Neutralizing aptamers from whole-cell SELEX inhibit the RET receptor tyrosine kinase. *PLoS Biol*. 2005; 3:e123. [PubMed: 15769183]
6. Daniels DA, Chen H, Hicke BJ, Swiderek KM, Gold L. A tenascin-C aptamer identified by tumor cell SELEX: systematic evolution of ligands by exponential enrichment. *Proc Natl Acad Sci U S A*. 2003; 100:15416–15421. [PubMed: 14676325]
7. Hicke BJ, Marion C, Chang YF, Gould T, Lynott CK, Parma D, Schmidt PG, Warren S. Tenascin-C aptamers are generated using tumor cells and purified protein. *J Biol Chem*. 2001; 276:48644–48654. [PubMed: 11590140]
8. Hicke BJ, Stephens AW, Gould T, Chang YF, Lynott CK, Heil J, Borkowski S, Hilger CS, Cook G, Warren S, Schmidt PG. Tumor targeting by an aptamer. *J Nucl Med*. 2006; 47:668–678. [PubMed: 16595502]
9. Missailidis S, Perkins A. Update: aptamers as novel radiopharmaceuticals: their applications and future prospects in diagnosis and therapy. *Cancer Biother Radiopharm*. 2007; 22:453–468. [PubMed: 17803440]
10. Pan W, Craven RC, Qiu Q, Wilson CB, Wills JW, Golovine S, Wang JF. Isolation of virus-neutralizing RNAs from a large pool of random sequences. *Proc Natl Acad Sci U S A*. 1995; 92:11509–11513. [PubMed: 8524793]
11. Perkins AC, Missailidis S. Radiolabelled aptamers for tumour imaging and therapy. *Q J Nucl Med Mol Imaging*. 2007; 51:292–296. [PubMed: 17464274]
12. Pieve CD, Perkins AC, Missailidis S. Anti-MUC1 aptamers: radiolabelling with (99m)Tc and biodistribution in MCF-7 tumour-bearing mice. *Nucl Med Biol*. 2009; 36:703–710. [PubMed: 19647177]

13. Tang Z, Shangguan D, Wang K, Shi H, Sefah K, Mallikratchy P, Chen HW, Li Y, Tan W. Selection of aptamers for molecular recognition and characterization of cancer cells. *Anal Chem*. 2007; 79:4900–4907. [PubMed: 17530817]
14. Wang C, Zhang M, Yang G, Zhang D, Ding H, Wang H, Fan M, Shen B, Shao N. Single-stranded DNA aptamers that bind differentiated but not parental cells: subtractive systematic evolution of ligands by exponential enrichment. *J Biotechnol*. 2003; 102:15–22. [PubMed: 12668310]
15. Wang J, Jiang H, Liu F. In vitro selection of novel RNA ligands that bind human cytomegalovirus and block viral infection. *RNA*. 2000; 6:571–583. [PubMed: 10786848]
16. Wen JD, Gray DM. Selection of genomic sequences that bind tightly to Ff gene 5 protein: primer-free genomic SELEX. *Nucleic Acids Res*. 2004; 32:e182. [PubMed: 15601993]
17. Brody EN, Gold L. Aptamers as therapeutic and diagnostic agents. *J Biotechnol*. 2000; 74:5–13. [PubMed: 10943568]
18. Deisingh AK. Aptamer-based biosensors: biomedical applications. *Handb Exp Pharmacol*. 2006:341–357. [PubMed: 16594624]
19. Gold L. The SELEX process: a surprising source of therapeutic and diagnostic compounds. *Harvey Lect*. 1995; 91:47–57. [PubMed: 9127985]
20. James W. Aptamers in the virologists' toolkit. *J Gen Virol*. 2007; 88:351–364. [PubMed: 17251551]
21. Joshi PJ, Fisher TS, Prasad VR. Anti-HIV inhibitors based on nucleic acids: emergence of aptamers as potent antivirals. *Curr Drug Targets Infect Disord*. 2003; 3:383–400. [PubMed: 14754437]
22. Mok W, Li Y. Recent progress in nucleic acid aptamer-based biosensors and bioassays. *Sensors*. 2008; 8:7050–7084.
23. Nimjee SM, Rusconi CP, Sullenger BA. Aptamers: an emerging class of therapeutics. *Annu Rev Med*. 2005; 56:555–583. [PubMed: 15660527]
24. Porschewski P, Grattinger MA, Klenzke K, Erpenbach A, Blind MR, Schafer F. Using aptamers as capture reagents in bead-based assay systems for diagnostics and hit identification. *J Biomol Screen*. 2006; 11:773–781. [PubMed: 16943392]
25. Zhang Z, Blank M, Schluesener HJ. Nucleic acid aptamers in human viral disease. *Arch Immunol Ther Exp (Warsz)*. 2004; 52:307–315. [PubMed: 15507871]
26. Pan W, Clawson GA. Primer-free aptamer selection using a random DNA library. *Methods Mol Biol*. 2010; 629:369–385. [PubMed: 20387162]
27. Pan W, Xin P, Clawson GA. Minimal primer and primer-free SELEX protocols for selection of aptamers from random DNA libraries. *Biotechniques*. 2008; 44:351–360. [PubMed: 18361789]
28. Jarosch F, Buchner K, Klussmann S. In vitro selection using a dual RNA library that allows primerless selection. *Nucleic Acids Res*. 2006; 34:e86. [PubMed: 16855281]
29. Vater A, Jarosch F, Buchner K, Klussmann S. Short bioactive Spiegelmers to migraine-associated calcitonin gene-related peptide rapidly identified by a novel approach: tailored-SELEX. *Nucleic Acids Res*. 2003; 31:e130. [PubMed: 14576330]
30. Shtatland T, Gill SC, Javornik BE, Johansson HE, Singer BS, Uhlenbeck OC, Zichi DA, Gold L. Interactions of Escherichia coli RNA with bacteriophage MS2 coat protein: genomic SELEX. *Nucleic Acids Res*. 2000; 28:e93. [PubMed: 11058143]
31. Sambrook, J.; Russell, DW. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2001.
32. Hsieh JC, Zinnen S, Modrich P. Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J Biol Chem*. 1993; 268:24607–24613. [PubMed: 7693703]
33. Tessier DC, Brousseau R, Vernet T. Ligation of single-stranded oligodeoxyribonucleotides by T4 RNA ligase. *Analytical Biochemistry*. 1986; 158:171–178. [PubMed: 3799962]
34. Harada K, Orgel LE. In vitro selection of optimal DNA substrates for T4 RNA ligase. *Proc Natl Acad Sci U S A*. 1993; 90:1576–1579. [PubMed: 8434020]
35. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003; 31:3406–3415. [PubMed: 12824337]

36. Hornung V, Latz E. Intracellular DNA recognition. *Nat Rev Immunol.* 2010; 10:123–130. [PubMed: 20098460]

Figure 1.

Protocol for primer-free SELEX. A schematic diagram of the protocol for selecting 30 nucleotide aptamers is illustrated. Various intermediates are shown (Products A-E) and labeled for reference in the text. Primers used in the protocols are shown at the bottom. Details of individual steps are provided in the Methods section.

Figure 2.

(A and B). Ligation reactions with T4 RNA Ligase 1 and Thermostable RNA Ligase. (A) In these reactions a small fraction of the 5'-N₃₀-3' starting material was labeled with ³²P at the 5' end (see Methods) in order to follow the ligation process. The starting material was ligated to oligo 1 (with cold phosphate at the 5' end) in Figure 1 (oligo 1:starting material, 10 pm:0.1 pm) using either T4 RNA Ligase 1 (10 units) or Thermostable RNA Ligase (0.5 units) for ~ 8 hours as described under Materials and Methods. Reactions were run on 12 % denaturing polyacrylamide gels and dried gels were visualized using a phosphorimager. These reactions differ slightly from standard SELEX reactions where oligo 1 was radioactively labeled at the 5' end while the 5'-N₃₀-3' nucleic acid was dephosphorylated in all rounds except round 1, where a small fraction was end-labeled as described above. Positions of the starting material and 50 nucleotide correctly ligated product (Product B in Fig. 1) are indicated along with locations of aberrant ligation products which may include circularized products or products with more than one oligo 1 or 5'-N₃₀-3'. The location of circularized 5'-N₃₀-3' monomers, which occur due to the 5' ³²P labeling (see text), is also indicated. A DNA molecular size marker indicating positions (in nucleotides) of specific size standard is also shown (lane M). Ligation efficiency for the formation of the 50 nucleotide product was 31 ± 8 % (ave. of 2 exp. ± standard deviation) and 51 ± 7 % (ave. of

4 exp.) for T4 RNA Ligase 1 and Thermostable RNA Ligase, respectively. (B) Ligation experiment performed with unlabeled 5' dephosphorylated 5'-N₃₀-3' and unlabeled 5' phosphorylated oligo 1. After the reaction products were radioactively labeled with T4 PNK using an exchange-reaction [31] and visualized as described above. Ligation efficiency was ~ 50 %.

Figure 3.

Cleavage of PCR product with BbsI and MnlI to regenerate 30 nucleotide selected material. A cleavage reaction run on a 12% denaturing polyacrylamide gel is shown for cleavage of the 73 nucleotide double stranded PCR product (equivalent to Product D in Fig. 1 minus the 5' overhang) with BbsI and MnlI. Correct cleavage would produce 23, 24, 25 and 30 nucleotide products on a denaturing gel with the last being the selected material. Some undigested and partially digested material is also evident in this gel and was typically observed in experiments.

Figure 4.

Gel-shift analysis of starting material and material from rounds 6, 10, and 11 of primer-free SELEX. An example of a gel-shift experiment run on a native 6% polyacrylamide gel is shown. Increasing amounts of HIV-RT (as indicated above wells, note higher amounts used for starting material) were mixed with 30 nucleotide 5'-N₃₀-3' derived from the starting material or rounds 6, 10, and 11 as described under Methods. Starting material was 5' end-labeled with ³²P, while rounds 6, 10, and 11 were internally labeled during the SELEX procedure (see Methods). Positions of shifted and unshifted material are indicated. Refer to the Results and Discussion section for quantification of experiments.

Figure 5. (A, B, and C). Ligation of various 30 nucleotide oligos to oligo 1 with Thermostable RNA Ligase. Ligations were carried out for 0.5, 2, or 8 hours using conditions described in Fig. 2A. Products were visualized as described in Fig. 2 and quantified using a phosphoimager. (A) Predicted structures (using mfold [35]) of the 30 nucleotide oligos ligated to oligo 1 (see Fig. 1) in panel B. Only those oligos with a predicted secondary structure are shown. The predicted ΔG value in kcal/mol using the conditions of 80 mM NaCl, 6 mM MgCl₂ and 37°C is also listed for each structure. (B) Plot of the percent of 30 nucleotide oligo ligated to oligo 1 vs. time. (C) Ligation plots for R11-1 (terminal 3' T) or R11-1 in which the 3' terminal T was changed to A, C, or G, and 5' N₂₉T-3'. All experiments were repeated at least once and similar results were obtained.

Table 1Sequences recovered from SELEX experiments and K_d for binding HIV-RT

Name ¹	Sequence ²	# of times recovered in round	K_d for binding HIV-RT (nM) ³
Starting material	5'-N ₃₀ -3'	NA	~1000
SELEX Experiment 1			
R11-1	53-AGGAAGGCTTTAGGTCTGAGATCTCGGAAT-33	9/10	82 +/- 7
R11-2	53- <u>G</u> GGAAGGCTTTAGGTCTGAGATCTCGGAAT-33	1/10	102 +/- 15
SELEX Experiment 2			
E2R10-1	53-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGGAT-33	1/10	ND
E2R10-2	53-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGTAT-33	2/10	ND
E2R10-3	53-AGGAAGGCTTTAGGTCTGAGATCTCGGAAT-33	3/10	ND
E2R10-4	53-AGGAAGGCTT <u>C</u> AGGTCTGAGAT <u>TT</u> CGGAAT-33	1/10	ND
E2R10-5	53-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGT <u>GT</u> -33	1/10	ND
E2R10-6	53-AGGAAGGCTTTAGGTCTGAGATCTCGGTAT-33	1/10	ND
E2R10-7	53-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGTAT-33	1/10	ND

¹ For experiment 1 names indicate the round of selection from which the sequence was recovered followed by a designated number. For experiment 2, the name is preceded by an "E2" designation.

² Underlined nucleotides in the recovered sequences indicate differences from the R11-1 sequence recovered in the first SELEX experiment.

³ Equilibrium dissociation constants (K_d) were determined using a gel shift assay as described under Material and Methods. Results for R11-1 and R11-2 are an average of 3 experiments +/- standard deviation. The result for "Starting material" was an estimate based on repeated experiments.

ND- Not determined; NA-Not applicable.