Deconvolution of a Complex Target Using DNA Aptamers*

Stephen Fitter1 and Robert James2
From the Department of Gastroenterology and Hepatology, Flinders Medical Centre, Adelaide, South Australia 5042

In vitro selection of single-stranded nucleic acid aptamers from large random sequence libraries is now a straightforward process particularly when screening with a single target molecule. These libraries contain considerable shape diversity as evident by the successful isolation of aptamers that bind with high affinity and specificity to chemically diverse targets. We propose that aptamer libraries contain sufficient shape diversity to allow deconvolution of a complex mixture of targets. Using unfractionated human plasma as our experimental model, we aim to develop methods to obtain aptamers against as many proteins as possible. To begin, it is critical that we understand how aptamer populations change with increasing rounds of selection. Certain aptamer families were apparent after only three selection rounds. Two additional cycles saw a decline in the relative abundance of these families and the emergence of yet another family that accounted for more than 60% of sequences in the pool. To overcome this population convergence, an aptamer-based target depletion method was developed, and the library screen was repeated. The previous dominant family effectively disappeared from the selected populations but was replaced by other aptamer families. Insights gained from these initial experiments are now being applied in the creation of second generation plasma protein screens and also to the analysis of other complex biological targets.

It is becoming increasingly apparent that our understanding of the complexity of many biological processes is limited by studies that focus on just a few molecules. This realization is largely due to the recent development of technologies that allow simultaneous measurement of thousands of different molecules. Information obtained from this type of assay not only puts a scale (number of genes or proteins) to the system under study but allows more accurate classification of biological states. Clearly, more detailed insights into a biological system or process can be gained if the number and identity of components is known (1, 2). Array-based platforms for profiling gene expression have already provided important new insights into many areas of biology (3–6). In particular, the assignment of genotypic and phenotypic signatures to various biological states is not only useful diagnostically but has also increased our understanding of disease progression (7–9).

Although analysis of nucleic acids has and will continue to be important, extension of the same technology to proteins is eagerly anticipated.

* This work was supported in part by a grant from Flinders Technologies Pty. Ltd. with additional support provided from the South Australia Government and an Australian Industry Biotechnology Innovation Fund grant. Continuing support is provided by Aptamer Technologies PTY LTD. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
1 Recipient of an Australian Postgraduate Award scholarship.
2 To whom correspondence should be addressed: Dept. of Gastroenterology and Hepatology, Flinders Medical Centre, Bedford Park, South Australia, Australia, 5042. Tel: 61-8-8204-5011; Fax: 61-8-8204-3943; E-mail: rob.james@flinders.edu.au.
the bound population from the target. The diversity of selected pools has been followed either by sequencing or some form of fingerprinting technique but only in the context of deciding when to terminate the screen rather than to understand in vitro selection dynamics. Much still needs to be done to more thoroughly understand the in vitro selection process, particularly when applied to complex mixtures.

It is well known that single-stranded nucleic acids adopt unique sequence-dependent three-dimensional shapes (27, 28). This property has been exploited to create shape libraries that have yielded ligands that bind with high affinity and specificity to a variety of molecules across a broad range of size and chemical composition (29, 30). A number of novel diagnostics and therapeutics have been isolated from these aptamer libraries, further indicating the significant amount of shape diversity they possess (31–36). We have initiated studies to establish whether there is enough shape diversity in a nucleic acid aptamer library to allow recognition of each component of a complex biological mixture. Human plasma was chosen as a model complex mixture because it contains a large number of proteins present at concentrations that span 6 orders of magnitude and because its composition can easily be changed by chromatography (37). The latter property is important because the high concentration of albumin has plagued other efforts to define the plasma proteome, and we needed to be able to remove it if required (38, 39). Using a straightforward in vitro selection protocol, we isolated aptamers that bound to several different plasma proteins. We followed the relative abundance of aptamer families with increasing rounds of in vitro selection to develop an understanding of the selection process particularly in terms of the number of targets identified. Insights gained from these initial results led to the design of a modified library screening protocol that yielded a different selected population profile and dynamic. Binding affinities were determined at different stages of the selection process for several aptamers that bound to prothrombin. High affinity aptamers were present in the population after only four cycles of selection.

**EXPERIMENTAL PROCEDURES**

**Oligonucleotide Library and Reagents**—DNA oligonucleotide template library 5′-AGCTCGAATACGCTCAGA-N96-TTCGACAT- GAGGGCGGATC-3′, unmodified and modified (5′-biotin or 5′-phosphate) primers (forward, 5′-GGAGGCTCAGATTAAGGCTCAA-3′; reverse, 5′-GATCCGGCTCATGTCGAA-3′); were produced by solid phase chemical synthesis (MWG Biotech, Ebersberg, Germany). Because of variability in the quality of oligonucleotide synthesis, full-length product was purified on 8% denaturing polyacrylamide gels. DNA was recovered from gels by electroelution and ethanol-precipitated. This single-stranded material was made double-stranded by a standard fill-in reaction using T7 DNA polymerase, and library complexity was estimated from the yield of double-stranded DNA (40). Larger amounts of this pool of double-stranded DNA were created by polymerase chain reaction. Typically, 2 × 10^14 double-stranded DNA molecules were amplified in a 34.8-ml reaction (for a 0.1-ml reaction: 20 pm forward and reverse primers, 200 μM dNTPs, 0.5 unit Taq polymerase (New England Biolabs, Beverly, MA)). A total of five cycles (94 °C 30 s, 58 °C 10 s, and 72 °C 30 s) were performed. Unless otherwise specified, all reagents were obtained from Sigma.

**Coupling Plasma Proteins to CH-Sepharose**—Human blood was obtained by venipuncture into vacuette tubes containing 5 mM EDTA (Greiner Bio-One, Frickenhausen, Germany). Platelet poor plasma was isolated by centrifugation at 2500 × g for 15 min and then dialyzed (Spectra/Por MWCO 3500; Spectrum Laboratories Inc., CA) against 137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 1.15 mM KH₂PO₄, 5 mM MgCl₂, pH 7.4 (1× BB). The plasma proteins were coupled to CH-Sepharose 4B according to the manufacturer’s instructions (Amersham Biosciences). The quantity of protein covalently bound to Sepharose was determined using a bicinchoninic acid binding assay (Pierce). Typically 70% of plasma proteins were coupled at a density of 17.5 μg protein/mg of Sepharose.

**In Vitro Selection Protocol**—Library screens were initiated using 4.5 μg of single-stranded DNA (ssDNA) including 20 fmol of 5′-32P-labeled ssDNA (5 × 10⁶ CPM/fmol ssDNA) as tracer. Single-stranded DNA in 1× BB (1 μl) was heat-denatured for 10 min at 68 °C and cooled to room temperature. To remove matrix-binding aptamers the library was incubated for 90 min with 20 μg of CH-Sepharose that had been blocked with Tris-HCl. Unbound ssDNA was partitioned from the Sepharose by spin filtration (1000 × g for 30 s) with a SPIN-X spin filter (Corning Costar, Acton, MA). Recovered ssDNA was then incubated with 20 μg of plasma-coupled Sepharose for 90 min at room temperature. Unbound ssDNA was again removed by spin filtration, and the beads were washed five times with 0.5 ml of 1× BB. Bound ssDNA was recovered by incubating beads in 7 M urea, 5 mM EDTA, pH 8, at 80 °C for 5 min followed by ethanol precipitation. Recovered ssDNA was amplified (94 °C 30 s, 58 °C 10 s, and 72 °C 30 s) with forward and phosphorylated reverse primers using Taq polymerase (New England Biolabs). Excess primer was digested with exonuclease I (2 units/ml), and then ssDNA was generated using λ exonuclease (12.5 units/ml) (New England Biolabs). For subsequent cycles the amount of aptamer pool used for selections was reduced 5-fold.

**Cloning and DNA Sequence Analysis**—Amplified double-stranded DNA from selected pools was cloned into pGEM T-Easy (Promega) and sequenced using a Beckman CEQ™ 8000 automated DNA sequencer (Beckman Coulter, Fullarton, CA). For alignments, a software program written in Python (www.python.org/) and running the SPLASH pattern recognition algorithm (41) was used (Bilateral, Sydney, Australia). Sequences that conformed to a minimum pattern length of 8 bases with a minimum of 6 non-wild card bases in the pattern were obtained.

**Protein Purification and Identification**—Single-stranded DNA for individual aptamers was generated by PCR using a 5′-biotinylated forward primer. Biotinylated ssDNA (10 pmol) was thermally folded and incubated with 25 μl of dialyzed plasma in a final volume of 100 μl for 90 min at room temperature. Aptamer-protein complexes were then captured on streptavidin-Sepharose (Amersham Biosciences). The complexes were washed four times with 0.5 ml of 1× BB containing 0.05% Tween 20. The beads were resuspended in Laemmli loading buffer and heated at 90 °C for 3 min, and eluates were electrophoresed on 6% SDS-PAGE gels. The proteins were visualized by silver staining (42).

Matrix-assisted laser desorption ionization mass spectroscopy of aptamer head-bound protein and peptide mass fingerprinting analysis was performed by the Australian Proteomics Facility (Macquarie University, Sydney, Australia).

**Surface Plasmon Resonance Measurements**—Biomolecular interaction analysis was performed using a Biacore 2000 biosensor (Biacore, Uppsala, Sweden). Sensor chips (CM5 research grade) and reagents for covalent coupling were purchased from Biacore. The running buffer was 1× BB containing 0.005% Tween 20. For kinetic studies, streptavidin in 10 mM sodium acetate, pH 4.5, at 5 μg/ml was covalently coupled to the carboxymethylated dextran matrix on all four flow cells using amine coupling reagents N-ethyl-N′-(3-diethylamino propyl)carbodiimide and N-hydroxysuccinimide according to the preset software specifications (Biacore AB). Typically, 450–500 resonance units of

3 The abbreviation used is: ssDNA, single-stranded DNA.
streptavidin was immobilized. To flow cells 1 and 2, saturating amounts of linker oligonucleotide (1 μM in 1× BB), which is complementary to 5’ primer-binding site of the aptamer (5’-CGTTTATTCTGAGCTCCC-biotin-3’), were injected at 30 μl/min for 30 s. This process was repeated for flow cells 3 and 4 using a linker oligonucleotide that was complementary to the 3’ primer-binding site (5’-biotin-GATCCGGGCTCATG-3’). Aptamers (200 nM in 1× BB) were then injected over the biotinylated linker/streptavidin complex on flow cell 2 and 4 at a flow rate of 5 μl/min for 5 min. In some cases duplex formation was linker-specific in that some aptamers would only form a duplex with either the 5’ or 3’ linker oligonucleotide. The experiments were performed on flow cells where 20–50 resonance units of linker/aptamer partial duplex was formed.

For kinetic studies, a concentration series of prothrombin in running buffer was injected at a flow rate of 100 μl/min over the linker and linker/aptamer flow cells at 20 °C. The association and dissociation phases were each monitored for 3 min, and the flow cells were regenerated with a pulse of 0.05% SDS. To determine error margins, two experiments were performed for each aptamer. This included stripping the aptamer from the flow cell with a pulse of 10 mM NaOH and then reforming the duplex and repeating the same concentration series of prothrombin.

When analyzing the data, the signals generated from injections of prothrombin over the linker only reference flow cell and those generated from injection of running buffer over the aptamer flow cell were subtracted from sensograms obtained for prothrombin injected over the aptamer flow cell. Rate constants (kₐ and kₐ) were obtained using global data fitting algorithms assuming a Langmuir (1:1) binding isotherm (BlAevaluation V3.1 software; Biacore). The equilibrium dissociation constant (KD) was determined from the ratio kₐ/kₐ.

Prothrombin Depletion—A 5’-biotinylated prothrombin aptamer (F1.06; see Fig. 3 for sequence) truncated to remove primer-binding sites was used. Plasma (60 μl), dialyzed against 1× BB, was mixed with 100 pmol of biotinylated oligonucleotide in a volume of 100 μl for 2 h at room temperature. Prothrombin-aptamer complexes were captured using streptavidin-Sepharose (Amersham Biosciences) and, recovered by spin filtration. Three cycles of aptamer mediated depletion were performed and the captured prothrombin together with known quantities of bovine serum albumin were analyzed on 8% SDS-PAGE. To quantify prothrombin extraction and check for the presence of other plasma proteins, the gels were stained with SYPRO Orange (Molecular Probes, Eugene, OR) and viewed on a Typhoon Fluorimager (Molecular Dynamics). Prothrombin-depleted plasma was subsequently coupled to CH-Sepharose, and an in vitro selection was performed as described above.

RESULTS

Aptamer Library Screening Strategy—In light of the large number of variables that could influence the outcome of an aptamer library screen, we deliberately chose the simplest selection protocol possible that included using a library with a moderately sized random sequence region. We anticipated that results from such a screen would indicate whether more sophisticated strategies were needed and what form they might take.

A library of ~10¹⁴ single-stranded DNA molecules was incubated with unfracionated plasma proteins coupled to activated Sepharose beads. The quantity of beads and therefore the concentration of target remained constant throughout the screen as did the washing and elution conditions. A trace amount of [32P]-labeled aptamer pool was included at each cycle of selection to monitor enrichment of specific binding sequences. Targets for individual aptamers were identified by incubating plasma with 5’-biotinylated aptamer, trapping the complex on streptavidin-Sepharose, and then analyzing the bound protein by SDS-polyacrylamide gel electrophoresis. Protein identification by this method was limited by the resolving capacity of the gel and the sensitivity of the silver staining method used. Despite these limitations, this format was considered the best option given the large number of aptamers to be assayed.

Dynamics of Aptamer Library Screening against a Complex Target—The DNA aptamer library contained a random sequence region of 45 nucleotides flanked by two 20-nucleotide regions for amplification. The amount of library used in a screen is a important variable, and we chose what corresponded to an average value (4.5 μg) obtained from the literature. It is common to quote a figure for the sequence complexity rather than a mass, but in our view the former is impossible to estimate accurately because of template bias that manifests itself at several steps during the creation of the library. A relatively modest amount of target was used (300 μg), especially considering the abundance of albumin. We reasoned that if albumin was going to dominate the screen, then increasing the amount of target would not significantly increase our chances of isolating aptamers that bound to other proteins. A total of 13 cycles of selection were performed, and the amount of bead-bound aptamer was measured after each cycle. Significant binding above background was detected after cycle five, and this increased to a total of 12% at the last cycle (Fig. 1).

Aliquots of selected pools from cycles 3–12 were cloned, randomly chosen colonies were picked, and their inserts were sequenced. The resultant data were then analyzed with the SPLASH pattern algorithm to identify common sequence motifs. Homology between different aptamers suggests they might bind to the same target and is an indirect measure of the degree of enrichment obtained at any stage of the in vitro selection process. In addition to searching for multiple representations of a motif at any cycle of selection, we also followed individual sequences with increasing cycles of selection to see whether their relative frequency in the population changed.

The level of sequence diversity observed within aptamer families was highly variable, with some families exhibiting less than 5% sequence variation between clones, whereas others were characterized by small highly conserved consensus motifs (data not shown). Representative aptamer sequences for all aptamer families are included in TABLE ONE.

The first observation we make is that the number of aptamers that either show no homology to any other member of their pool or were...
In Vitro Selection Using Complex Targets

Summary of plasma protein targets identified for aptamer families

| Family | Molecular mass | Identity | Representative aptamer sequence
|--------|----------------|----------|-----------------------------|
| F1     | 72 kDa         | Prothrombin | See Fig. 4                  |
| F2     | 72 kDa         | Prothrombin | See Fig. 4                  |
| F3     | 180 kDa        | C3 complement | GGGCCACAAATTGATTACGCGTAATGGGGCGGACCCGGACGAGGCG |
| F4     | > 250          |            | CATCGCAGGCGCCAAAAATGAGTACGCGTGAGTACCAGAGGGAG |
| F5     | 140 and 190    |            | CACCGGCTGGGAGGGCGGCGAGGGAGGAGGGAGGGAGG |
| F6     | 68 kDa         |            | GTCCGGGACAGGAGGGCGGACAGGGAGGAGCTGAGGAGG |
| F7     | > 250          |            | AAGGCAAGCGCACAACCCTGCTGCAACGAGGGGAAAGG |
| F8     | 80 kDa         |            | CTGCGGCGGAGGGGCTTGTGAGTATGAAAGGGAGG |
| F9     | ND            |            | ACTAGTAACGCGACAGGGAGGAAAAAGGCGCTTACGAGG |
| F10    | ND            |            | CAGACGGAGACGCACAGACAGAGAAAAGGCAAGGAGG |
| F11    | 68 kDa         |            | GAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG |
| F12    | 68, 100, and > 300 | Prothrombin | GTCGGGAGAAATTAGTTAGGGGGCGGAGGAGGAGG |
| F13    | 68 kDa         |            | GGGGCAAGGGGACCCGAGGAGGAGGAGGAGGAGGAGGAGG |
| F14    | 68 kDa         |            | AAGCAGCGGCGAGGGGAGGAGGAGGAGGAGGAGGAGGAGG |
| F15    | 72 kDa         | Prothrombin | See Fig. 4                  |
| F16    | 72 kDa         | Prothrombin | See Fig. 4                  |
| F17    | 72 kDa         | Prothrombin | See Fig. 4                  |
| F18    | 72 kDa         | Prothrombin | See Fig. 4                  |
| F19    | NT            |            | AGGCCCGAAAATTGATTGGTAGATTAGGGGAGGAGG |
| F20    | ND            |            | AACCGCGGAAATGTTAGGGGGCAGGAGGAGG |
| F21    | 68 kDa         |            | CACGGGAAATCTGCGGTTGGGAGGAGGAGG |
| F22    | ND            |            | AAGGCAACGGCAGCCCAAGGCGAGGAGGAGG |
| F23    | ND            |            | GGGGCAAGGGGACCCGAGGAGGAGGAGGAGGAGGAGG |
| F24    | ND            |            | GGGGCAAGGGGACCCGAGGAGGAGGAGGAGGAGG |
| F25    | NT            |            | AAGGCAACGGCAGCCCAAGGCGAGGAGGAGG |
| F26    | 68 kDa         |            | TGGGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG |
| F30    | 80 kDa         |            | TCAGCGGTAGGTTATAGGGGGGAGGAGGAGG |
| F31    | 100 and 190    |            | CACCGTGGAGGGAGGACCCATTTACTGAGAGG |

* Further information regarding aptamer family sequence diversity can be obtained upon request.
* ND, not detected.
* NT, not tested.

**Protein Target Identification—**Computer analysis of nucleotide sequences predicts relationships between different aptamers. To see how informative this analysis was, we sought to identify protein targets for many of the aptamers isolated in the screen. The target assay was designed to allow rapid analysis of some 500 aptamers. We made an assumption that most of the aptamers isolated in the screen would bind to abundant or moderately abundant proteins. The silver staining protocol used could easily detect 10 ng of albumin, and so a single concentration of plasma (25 μl) was used in each binding reaction. If we assume that an aptamer traps the majority of its target present in the plasma sample, then we should detect any protein present at a concentration equal to or greater than 400 ng/ml. Results from these experiments showed that aptamer binding was very specific with little or no background (representative data in Fig. 3). In many cases aptamers bound single proteins, but in some cases two or more proteins were present, suggesting binding to one subunit of a complex. It should be noted that all of the proteins in our experiments should be in their native conformation except where attachment to the Sepharose beads alters their structure. In this first round of assays, protein targets were identified for 17 of the 23 aptamer families previously defined on the basis of unique sequence motifs (TABLE ONE). Six families that failed to yield a detectable target were then reassayed using five times more plasma. This resulted in the identification of protein targets for families F5 and F8.
In Vitro Selection Using Complex Targets

It became increasingly apparent when performing these experiments that aptamers bound to their targets with very high specificity. A single step of aptamer-mediated affinity chromatography typically yielded proteins that by gel analysis were 85–90% pure, which in the case of prothrombin represents a 400-fold purification from crude plasma. These results suggested to us that amino acid sequence information could be obtained directly from protein bound to aptamer bearing beads. In a proof of principle experiment, we used matrix-assisted laser adsorption ionization time-of-flight mass spectrometry and peptide mass fingerprinting to identifying protein targets for F1, prothrombin (72 kDa), and F3, C3 complement (180 kDa).

Several families (F2 and F15–18) appeared to bind a protein that comigrated on SDS-polyacrylamide gels with the protein recognized by F1. A Western blot was probed with a monoclonal antibody to prothrombin to establish whether these aptamers were binding to the same protein. All of the bands were positive, thereby confirming their identity (data not shown). This result indicates that greater than 70% of the aptamers present in the cycle 5 pool bound to prothrombin.

Aptamer Target Affinity and Cycles of Selection—Much of the in vitro selection method is based on the assumption that higher affinity aptamers should progressively replace lower affinity aptamers as the number of selection cycles increase. Our isolation of a large number of prothrombin binding aptamers provided us with an opportunity to determine whether this assumption was correct under the conditions of our experiment.

Aptamers assigned to F1 and F15–18 show a region of high homology characterized by a series of G-rich repeats. Sequences flanking this motif are variable, allowing classification of F1 into subfamilies. G-rich regions are also present in F2, but there is little homology to the more common motif present in the other families (Fig. 4).

The binding kinetics of a subset of these aptamers to prothrombin were determined by surface plasmon resonance. Sensor chips were prepared to allow multiple aptamers to be tested on the same sensor flow cell. This was achieved by immobilizing via a streptavidin linkage, a biotinylated oligonucleotide (18-mer) complementary to either the 5′ or 3′ primer-binding site of each aptamer. Individual aptamers were hybridized to the linker oligonucleotide forming a partial duplex that was stable upon prothrombin binding and chip regeneration. A 100-fold dilution series of pure prothrombin was injected over the partial duplex, and the resulting sensorgrams were fitted to kinetic models assuming a 1:1 binding interaction. Experimental and mathematically derived sensorgrams were essentially superimposable, supporting the 1:1 model (Fig. 5).

Rate constants ($k_a$ and $k_d$) and equilibrium dissociation constants for a number of aptamers together with their frequency in the selected populations are presented in TABLE TWO. To help interpret these data, we have assigned aptamers to one of four classes based on calculated rate constants and visual examination of sensorgrams. Aptamers that appeared at early cycles of selection and remained in the population until the last cycle had fast association rates and slow dissociation rates (class I). Aptamers present at early cycles of selection but then disappearing in later populations belonged to either class II or class III. Only one aptamer (F2) was assigned to class IV (very slow dissociation rate), and it appeared at the very end of the selection. In general we found that aptamer target association rates varied over a narrow range, whereas the
range of dissociation rates was significantly greater. It is noteworthy that aptamers with subnanomolar dissociation constants were present when this family first appeared in the selected population.

Selection Dynamics after Dominant Target Depletion—Our observation of the emergence of dominant aptamer families in a screen using a mixture of targets is consistent with what has been reported by others (18, 19, 21). How this occurs remains a puzzle given the concentration range and diversity of targets. Regardless of the mechanism, it is a problem that has to be overcome when trying to isolate aptamers against as many targets as possible.

In practice there are two obvious ways to prevent narrowing of the selected population because of dominant aptamer families. One is to focus on a more thorough analysis of early rounds of selection before the aptamer population becomes skewed. The problem with early round pools is the high frequency of low or nonspecific binding aptamers. The other is to deplete the complex mixture of targets recognized by dominant aptamer families. Our results showed that aptamers bind their targets with high affinity and specificity, making them ideal reagents for this purpose (Fig. 3).

To explore the effectiveness of specific target depletion and how this might influence the dynamics of aptamer isolation, we made use of the prothrombin-binding aptamer F1.06. Plasma was depleted of prothrombin by simple aptamer-mediated affinity chromatography and then coupled to activated Sepharose. The aptamer library screen was then performed as before, except that the selection was stopped after eight cycles. Individual aptamers from cycles 5 and 8 were sequenced, and their targets were identified by SDS gel electrophoresis.

The prothrombin-depleted plasma screen was performed twice to see whether the changes we observed were reproducible. The data show a dramatic decline in the number of prothrombin-binding aptamers isolated, such that they were absent in the cycle 8 population of both screens. It is important to note that in these experiments a low level of prothrombin was still present in the depleted plasma. It would appear from these results that there is no need to remove it completely.

The two prothrombin-depleted plasma screens gave qualitatively the same results (Fig. 6, B and C). The same aptamer families were isolated in both screens, but their relative representation in the selected populations differed. This is not surprising given the stochastic nature of the in vitro selection process. Interestingly, these families appeared earlier in the selection with the depleted plasma than they did in the whole plasma screen.

In the absence of prothrombin, it would appear that other aptamer families now begin to dominate the screen. In one experiment it was the C3 complement binding family (F3), and in the other it was F5. Whether this is solely due to the stochastic nature of the selection process or some other mechanism remains to be determined. It is also noteworthy that two new aptamer families (F30 and F31) not isolated in the whole plasma screen were present in the enriched populations of the depleted plasma screens.
In Vitro Selection Using Complex Targets

Our goal is to develop methods that extend aptamer library screening beyond the single target paradigm. Human plasma was chosen as a model for our study because of its molecular complexity and the practical challenges it presents (reviewed in Ref. 37). A number of examples have been reported where a complex target has been used for aptamer library screening (12–14, 18, 19, 21, 43–47). All of these studies followed traditional screening protocols developed for use with single pure targets where cycles of in vitro selection continue until some predetermined functional end point is reached. Because each study had a single end point there was no attempt and indeed no need to understand the dynamics of the selection process. In stark contrast, our study has as many end points as there are proteins in plasma. It was therefore critical that we obtain a detailed understanding of how the selection process works to establish whether our goal was achievable.

Using unfractionated human plasma as a target, we isolated and analyzed the binding characteristics of a large number of aptamers. Aptamers were assigned to families based on sequence motif homologies, and their protein targets were identified by SDS gel electrophoresis. Using these simple methods, we successfully identified aptamers that bound to 14 different plasma proteins. This appears a low number given the variety of potential targets in plasma, but a detailed analysis of aptamer populations at different times during the selection process revealed the reason for this result. A family of aptamers that bound prothrombin emerged midway through the screen and dominated all of the selected populations thereafter. The emergence of dominant aptamer families is the desired end point when screening libraries with single targets. However, for our purposes, narrowing the diversity of selected populations beyond a certain point is counterproductive, particularly when a family appears and then disappears with additional cycles of selection (e.g. F12).

It is a widely held view that increasing rounds of in vitro selection yield aptamers with progressively higher target binding affinities. We tested this proposal by measuring the binding characteristics of prothrombin aptamers isolated at different times during the selection process. Surprisingly, aptamers with subnanomolar binding affinities were isolated after only four cycles of selection, that is, when they first appeared in the selected population. A single aptamer family (F2) with a higher binding affinity for the same protein only appeared toward the very end of the screen. Our data suggest that large numbers of selection cycles might not be needed to obtain aptamers with high binding affinity. We need to extend these binding measurements to other aptamer families isolated in this study to see how typical this observation is.

There are several reports of in vitro selections where aptamer families

| Table Two: Binding kinetics of prothrombin aptamers |
| --- |
| **Family** | **Cycle** | **$k_a$** ($\text{m}^{-1}\text{s}^{-1}$) | **$k_d$** ($\text{s}^{-1}$) | **$K_D$** | **$t_{1/2}$** |
| | 4 | 5 | 8 | 12 | |
| F1.01 | — | — | — | 5 | 2.46 $\times 10^6 \pm 0.1^c$ | 1.35 $\times 10^{-3} \pm 0.05$ | 0.55 $\pm 0.03$ | 8.50 | I |
| F1.05 | 2 | 4 | 8 | 8 | 2.32 $\times 10^6 \pm 0.05$ | 7.74 $\times 10^{-3} \pm 0.02$ | 0.35 $\pm 0.01$ | 14.90 | I |
| F1.06 | 2 | 6 | 2 | 1 | 5.12 $\times 10^6 \pm 0.5$ | 2.41 $\times 10^{-3} \pm 0.01$ | 0.48 $\pm 0.05$ | 4.80 | I |
| F1.10 | — | 7 | 3 | — | 1.28 $\times 10^6 \pm 0.12$ | 9.22 $\times 10^{-3} \pm 0.1$ | 7.25 $\pm 0.58$ | 1.26 | II |
| F1.11 | — | 6 | 5 | — | 9.49 $\times 10^6 \pm 0.15$ | 1.48 $\times 10^{-2} \pm 0.01$ | 1.55 $\pm 0.15$ | 0.80 | III |
| F1.12 | — | 4 | — | — | 2.09 $\times 10^6 \pm 0.8$ | 5.98 $\times 10^{-3} \pm 0.34$ | 3.61 $\pm 1.45$ | 19.30 | II |
| F15 | 2 | — | 1 | — | 4.73 $\times 10^6 \pm 0.3$ | 4.65 $\times 10^{-3} \pm 0.06$ | 9.85 $\pm 0.5$ | 2.50 | II |
| F16 | 1 | 6 | 1 | — | 7.57 $\times 10^6 \pm 0.5$ | 1.59 $\times 10^{-3} \pm 0.04$ | 2.11 $\pm 0.4$ | 0.74 | III |
| F17 | 1 | 1 | — | — | 6.01 $\times 10^6 \pm 0.4$ | 2.23 $\times 10^{-3} \pm 0.01$ | 3.75 $\pm 0.1$ | 0.52 | III |
| F2 | — | — | — | 18 | 6.88 $\times 10^6$ | 4.89 $\times 10^{-4}$ | 0.07 | 23.62 | IV |

$^a$ Representation of each aptamer in the population from each cycle.

$^b$ not detected.

$^c$ The standard errors were determined for two independent experiments.

![Figure 6: In vitro selection dynamics using prothrombin depleted plasma. Prothrombin was removed (>80%) from plasma by successive rounds of exposure to biotinylated aptamer F1.06. Eight cycles of in vitro selection were performed, and randomly chosen aptamers from cycle 5 (open bars) and cycle 8 (filled bars) were sequenced and analyzed for motif homologies. Plasma protein target identification assays were also performed to confirm assignment of aptamers to particular families. The number of aptamers identified for each family as a percentage of the total selected pool analyzed at cycles 5 and 8 are shown.](image-url)
have emerged to dominate a population that upon further characterization has been shown to recognize sites not present on the intended target. In many cases these aptamers bind to the support used to immobilize the desired target. In others, the intended target is present as part of a mixture that contains molecules that are preferred targets. Attempts to overcome this problem have focused on depleting the aptamer population of sequences that rapidly accumulate. In practice, solutions involve some form of negative selection and range from simply passing aptamer populations over the undervatized target support to more complicated sequence specific enzymatic methods (48). Alternatively, attempts have been made to block ligand-binding sites with nonamplifiable sequences isolated in an initial in vitro selection experiment. The success of this approach depends on a number of factors including the binding affinity of the blocking ligand and the amount of target in the mixture. In addition it has been found that blocking one site on a target does not prevent the isolation of ligands to other sites on the same protein (16). In our case we considered it more efficient to deplete the target population than to focus on the aptamer population. Our data showed that this strategy was effective as selection dynamics changed significantly as did the representation of different aptamer families within selected pools. Of note, new aptamer families were identified, and other aptamer families that previously represented only a small fraction of the selected pool now began to dominate.

Although the target depletion strategy employed in this study has some advantages, it also has some limitations. First, it can only be applied to soluble molecules and not to complex systems such as cell surfaces. Second, if the dominant target is part of a protein complex, then other proteins could be depleted indirectly via their association with the complex. Based on current information we are unable to say why dominant aptamer families arise. From our work and that of others it would appear that target concentration is not a major factor. It is possible there is something special about the shape or surface of a protein analogous to the concept of immunogenicity that might differentiate one protein from another. Alternatively, the appearance of dominant families might have nothing to do with the target but might be determined by the relative abundance of a particular shape in the uns-elected aptamer library.

Ours is the first study we are aware of that has examined whether the traditional in vitro selection strategy is appropriate for deconvoluting a complex mixture. We conclude from the data reported in this manuscript that it probably is not. Our results have identified a number of areas where further work is required to better understand how certain variables influence the outcome of an in vitro selection experiment. We need to establish why certain aptamer families come to dominate selected populations and how their influence on subsequent cycles of selection can be minimized. What role does target concentration within a complex mixture play in successful aptamer isolation? How many high affinity binding aptamers are present in populations obtained in the early rounds of selection? Answers to these and several other questions arising from this study should lead to the development of more efficient and effective library screening protocols.

Acknowledgments—We thank Prof. Ralph Bradshaw for suggestions throughout this study and for critical review of this manuscript. We are also grateful to Prof. Graeme Young for support and encouragement.

REFERENCES
1. Baliga, N. S., Pan, M., Goo, Y. A., Yi, E. C., Goodlett, D. R., Dimitrov, K., Shannon, P., Aebersold, R., Ng, W. V., and Hood, L. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 14913–14918
2. Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Cabestrani, C., Yuh, C. H., Mino-

In Vitro Selection Using Complex Targets

VOLUME 280 • NUMBER 40 • OCTOBER 7, 2005

34200 JOURNAL OF BIOLOGICAL CHEMISTRY
41. Califano, A. (2000) *Bioinformatics* 16, 341–357
42. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* 68, 850–858
43. Blank, M., Weinschenk, T., Priemer, M., and Schluesener, H. (2001) *J. Biol. Chem.* 276, 16464–16468
44. Hicke, B. J., Marion, C., Chang, Y. F., Gould, T., Lynott, C. K., Parma, D., Schmidt, P. G., and Warren, S. (2001) *J. Biol. Chem.* 276, 48644–48654
45. Daniels, D. A., Chen, H., Hicke, B. J., Swiderek, K. M., and Gold, L. (2003) *Proc. Natl. Acad. Sci. U. S. A.* 100, 15416–15421
46. Lorger, M., Engstler, M., Homann, M., and Goring, H. U. (2003) *Eukaryot. Cell* 2, 84–94
47. Thomas, M., Chedin, S., Carles, C., Riva, M., Famulok, M., and Sentenac, A. (1997) *J. Biol. Chem.* 272, 27980–27986
48. Shi, H., Fan, X., Ni, Z., and Lis, J. T. (2002) *RNA* 8, 1461–1470
49. Mutuberria, R., Hoogenboom, H. R., van der Linden, E., de Bruijn, A. P., and Roovers, R. C. (1999) *J. Immunol. Methods* 231, 65–81