Aptamer Selection for Human Serum Albumin

By

Joanne Lee

Honours thesis submitted to the Faculty of Science Department of Chemistry
In partial fulfillment of the requirements of the B.Sc. Degree

Supervisor: Maxim V. Berezovski

Department of Chemistry
University of Ottawa
Ottawa, Canada

ABSTRACT

Human serum albumin (HSA) is a highly abundant protein in the human plasma, causing the discovery of new biomarkers challenging. Here we introduced a new method: aptamers as new affinity ligands for the depletion of HSA from serum. Aptamers are short RNA or DNA oligonucleotides (15-120 nucleotides) that bind to different targets, including proteins, peptides, and cells. Aptamers are like chemical antibodies, able to form 3-dimensional structures with targets and bind with high affinity and specificity. In this research, we selected aptamers against recombinant HSA (rHSA) using two different combinatorial DNA libraries: N40 and structured DNA. The method we used is called Nonequilibrium Capillary Electrophoresis of Equilibrium Mixture (NECEEM) with Systemic Evolution of Ligands in Exponential Enrichment (SELEX). This method uses capillary electrophoresis to separate target-bound DNA from non-target bound DNA in an equilibrium mixture containing both the DNA and the target. The collected target-bound nucleic acids are amplified with Polymerase Chain Reaction (PCR) and used for the next round of selection. By repeating this procedure, SELEX generates a pool of aptamers with the highest binding affinity to the target. The final aptamer pool is later cloned and sequenced. Prior to selection, we optimized both mastermix and PCR conditions for the amplification of the structured DNA library. Unfortunately, the selection of aptamers against rHSA with the two DNA libraries was unsuccessful. The structured DNA library did not generate high enough aptamer concentration in the first round for further selection rounds. The N40 DNA library showed no enrichment of aptamers with 5 rounds of selection.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. Berezovski, for his support during this work; Darija Muharemagic and Glebe Mironov for their help and guidance; Dr. Maria Derosa from Carleton University for supplying the DNA library and helping with amplifying the DNA library; Fanny Guillerme for her previous work on the project; and everyone in the Berezovski lab.

TABLE OF CONTENTS

LIST OF FIGURES

STATEMENT OF CONTRIBUTION

Darija Muharemagic guided the experiment procedures and analysis of results in this research project.

Glebe Mironov helped with capillary electrophoresis and the concept of NECEEM.

INTRODUCTION

Human blood plasma is the one of the most studied biological fluids, serving as the single most informative sample that can be collected from an individual (1,3). It is the primary sample for disease diagnosis with a high concentration of proteins, in the range of 60-80mg of protein per mL of blood plasma (2). The identification of the blood plasma proteome represents a powerful tool for clinical applications, with the potential to generate personalized "global-health profiles" (4).

Blood plays a vital role in the vascular system: in the production and degradation of blood cells; in the transportation of oxygen, carbon dioxide, and various nutrients; in the elimination of waste; and in thermal regulation (3). Blood is composed of cells and plasma, with the majority of proteins present in the plasma. As blood perfuse through different organs and tissues, proteins and protein fragments passively and actively enter the blood circulation, removing and modifying existing proteins. This is a process that varies according to physiological and pathological conditions of the individual (4, 6). Using this concept, each person's unique proteomic profile of blood plasma can serve as a reflection of their ongoing pathological and physiological condition (4, 5). Thus, plasma proteome possess high therapeutic potential as a diagnostic marker and as a therapeutic target for many human diseases (5, 7).

Since many potential biomarkers are likely to be present at low abundances in the plasma (1), the goal of plasma proteomics is to identify as many low-abundance proteins as possible. The identification of low-abundance proteins will establish a baseline for serum proteome, and allow a variety of other serum samples to compare against it and detect significant changes in biomarker levels (2). This will enable the identification of individuals at risk of developing certain diseases, the detection of existing diseases, and the monitoring

of therapeutic interventions in patients under treatment and subsequent determination of prognosis (1). Plasma proteome biomarker discovery, therefore, is the first step to personalized medicine.

Unfortunately, however, current understanding of plasma proteome is far from complete. Although there are more than 300 biomarkers identified in the human plasma, only 20 biomarkers are used clinically (5). The major obstacles in identifying new biomarkers are: the dynamic range of proteins in the plasma that exceeds ten orders of magnitude (1, 9), and the presence of high-abundance resident proteins, such as albumin (4). Human serum albumin (HSA) constitutes over half the plasma proteins and has a concentration of 40mg/ml in the blood. In contrast, most of the biomarkers of interest are secreted at very low concentrations, especially in the early onset of diseases; cytokines, for example, secrete at a concentration that is at less than 1ng/ml (5). Thus, the large abundance of HSA prohibits the complete analysis of serum proteome, interfering in the detection and assay of minor protein components containing new biomarkers (8). So far, proteins can be separated by charge, size, and isoelectric point, but no method is available to separate proteins by abundances (5).

The current proteomics technologies address this problem by depleting the high-abundant resident serum proteins, and then analyzing the low-abundance protein samples to find new biomarkers. They can be categorized as follows:

a. Isolation of low-abundance resident serum proteins through homogenization of cells followed by treatment with detergents, and reducing/ denaturing reagents (9, 15, 16).
b. Separation of proteins using 2-D gel electrophoresis and affinity chromatography (8, 9, 10, 11, 12, 14).

c. Separation and identification of proteins by Mass spectrometry (14).

a. Treatment with detergent and reducing agents

Conventional methods of purifying serum samples (to obtain low-abundance proteins) involve various treatments with denaturing and reducing agents. The cold ethanol precipitation method developed by Cohn *et al* (1946) uses the attraction between isoelectric serum albumin molecules through dehydration of ethanol to precipitate them out of the serum sample (15, 16). However, this technique still contains considerable quantities of albumin (15). Chemical precipitation of albumin with $(NH_4)_2SO_4$, and ion-exchange chromatography also yields the same problem of HSA contamination (15).

b. 2D gel electrophoresis and affinity chromatography

Two-dimensional polyacrylamide gel electrophoresis separates proteins based on their isoelectric points and molecular weights (9, 14). Visualization of the proteins after 2-D gel separation is a critical step in quantitative proteome analysis. Radioactive ^{35}S or ^{32}P and fluorescence dyes are used to visualize the proteins. Again, this method has disadvantages. The main problem is the inability to separate low-abundance serum proteins from high-abundant ones. Also, there is a crowding of molecular weights between 45kD and 80kD, and of the isoelectric points between 4.5 and 6 (10). This technique is also cumbersome, time-consuming, and cannot be automated (14).

Affinity chromatography is a liquid chromatography that makes use of biological interactions for the separation and analysis of analytes within a sample. The affinity ligands are mainly biological, but can also be synthetic (11). The first affinity ligand that was introduced in order to deplete HSA in the blood was, in fact, synthetic. It was a hydrophobic

dye called Cibacron blue that has high affinity for HSA (8, 15). This dye is conjugated to Sephorose and used as an affinity column for the separation of HSA from other serum proteins. Using the dye, however, has three major disadvantages: the efficiency of binding to HSA is low, the elution of the bound HSA required denaturing conditions (which affect the analysis of other less abundant proteins), and the dye tended to leach from the affinity columns (contaminating the sample) (15).

Affinity ligands can also be biological, such as antibodies. The antibody-based affinity chromatography has gained popularity primarily due to its increased specificity for HSA compared to the dye and ethanol precipitation, and its ease of handling (2, 16). [17]Steel *et al* (2003) proposed an improved affinity column for removing HSA through monoclonal antibodies. This newly-developed immunoaffinity resin is able to remove both HSA and HSA fragments present in serum (16). It can also be combined with protein G resin to remove Ig proteins in serum, another abundant resident serum protein. [12]Klooster *et al* (2007) developed a new antibody (Camelid antibody fragments (VHH)) for the depletion of HSA (12). Problems with affinity chromatography as a protein purification technique include: the high cost (antibodies are expensive to make), labour-intensity, and the inability to produce on a large scale (no high throughput) (10).

c. Mass spectrometry

Mass spectrometry is not really involved in removing HSA from plasma; instead, the technique is often combined with the above mentioned techniques (2-D gel electrophoresis) to further identify and characterize potential biomarkers in the plasma (10). Often, serum samples are already depleted of HSA when they are subjected to mass spectroscopy analysis. Surface-enhanced laser desorption/ionization time of flight mass spectrometry (SELDI-TOF-

MS) is perhaps one of the most powerful tools in proteomics. SELDI-TOF-MS is able to provide rapid protein profiles without the separation or purification of the protein sample (13, 14). However, so far, there have been no specific biomarkers identified using this technique. In addition, this technique has a low sensitivity, and cannot identify any protein biomarker at ng/mL level (13). The mass spectrum obtained does not enable identification of proteins analyzed (14). Mass spectrometry as a biomarker-discovery tool is somewhat limited: the serum sample may be too complex for direct MS analysis (difficulty in mining low abundance proteins (9). Thus, continued development of depletion and enrichment techniques coupled with improved MS separation is required in order to discover new biomarkers that mark different disease states (9).

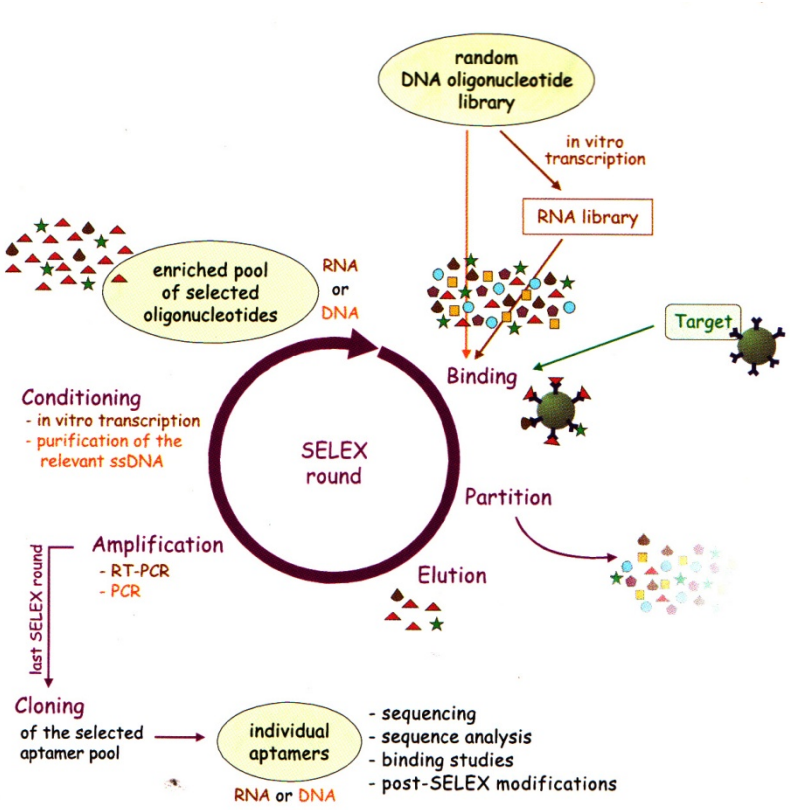
There is currently a great deal of interest in new methods to replace the above-mentioned proteomic techniques, which would enable efficient, cost-effective, and reliable purification for the removal of HSA from serum samples (12). As a result, this research project will attempt to break new ground in proposing aptamers as new affinity ligands for the depletion of HSA from serum.

Aptamers are short synthetic oligonucleic acids (15-120 nucleotides) that can bind to different targets, including proteins, peptides, and cells. They fold into 3-dimensional structures, and are able to bind to their targets with high affinity and specificity (18, 19, 20, 24). In contrast to previously-described proteomic technologies for depleting HSA, aptamers demonstrate advantages, in that they function like antibodies, but without the high labour intensity and cost (19). Aptamers are stable, unlike antibodies that may undergo modifications and structural perturbation during the eliciting process, and they can also be easily modified by adding on a reporter for imaging (21). Thus, aptamers are an attractive

option, because they are easily synthesized and involve simple purification steps, with the potential for high throughput screening.

The objective of this study is thus to select aptamers specific for HSA, with the goal of generating an affinity ligand (aptamer) that will bind to HSA with high affinity and specificity.

Aptamers are selected *in vitro* from large combinatorial libraries of DNA or RNA sequences in a process called the systematic evolution of ligands by exponential enrichment (SELEX) (18, 19). SELEX involves mixing the target (HSA) with a naive DNA library, in which the target-bound DNA is separated from the non-target-bound DNA to produce an aptamer-enriched library. The target-bound DNA is then amplified with polymerase chain reaction (PCR) and used for the next round of selection as the new DNA library. This procedure is repeated many times (10-15times) using the previously-obtained enriched library as the starting DNA library (18). The final aptamer pool is later cloned and sequenced. The evolution of the aptamer population is driven by the selection conditions and the repeated rounds of selection; so that only aptamers with the highest binding affinity are selected (21) (see fig. 1).



[21]**Figure 1. In-vitro illustration of SELEX technology.**

Partitioning methods used to separate target-bound DNA from non-target bound DNA range from conventional methods, such as filtration, gel electrophoresis, and chromatography, to the more newly-developed methods, such as different kinetic capillary electrophoresis methods (KCE methods). KCE methods are comprised of two types: non-equilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) and equilibrium capillary electrophoresis of equilibrium mixtures (ECEEM) (19). KCE methods have the advantage of requiring a small sample size, which improves efficiency (23).

NECEEM-based partitioning, unlike conventional methods, can improve the efficiency of aptamer selection by at least two orders of magnitude. Thus, fewer selection rounds are required, with as little as only one round of selection to obtain aptamers with high

affinity to the target, for example, the selection of aptamers for PFTase with a K_d in the nM range (22, 23).

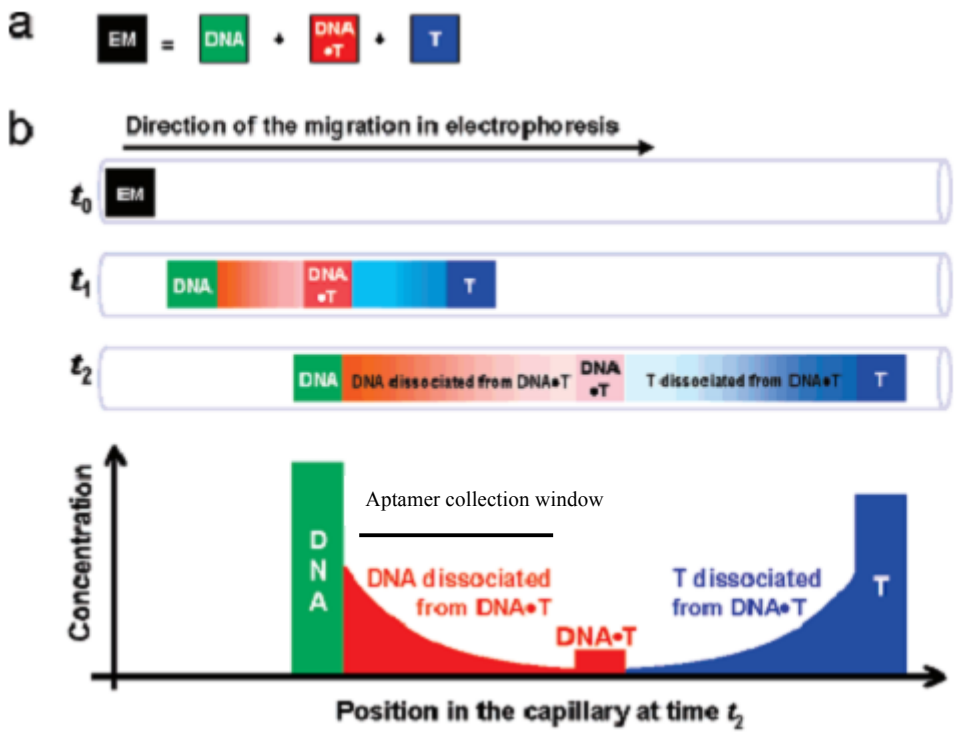
NECEEM separates target-bound DNA from non-target bound DNA by gel-free capillary electrophoresis under nonequilibrium conditions. When a high voltage is applied to the equilibrium mixture of the HSA with naive DNA library, the DNA and HSA are subjected to the electroosmotic flow (EOF), and move toward the negative electrode (18, 22,23). However, since DNA is negatively charged due to its phosphate backbone, its electrophoretic mobilities will be opposite to that of the EOF, and will migrate slower towards the negative electrode. Proteins (HSA), on the other hand, are slightly positively charged, so their electrophoretic mobilities will be in the same direction as the EOF (18, 22).

The unbound DNA library will migrate more slowly than HSA, and the DNA-HSA complex will migrate at an intermediate speed. At this point, the equilibrium between the unbound DNA and DNA-HSA complex is no longer maintained: HSA starts disassociating and the order in which the components reach the end of the capillary is as follows (22):

1. Free HSA.
2. Free HSA formed by dissociation of DNA-HSA complex.
3. Remaining intact DNA-HSA complex.
4. Unbound DNA formed from the dissociation of DNA-HSA complex.
5. Unbound DNA.

Thus, one can selectively collect the dissociated DNA from the DNA-HSA complex in stage 4. This collected fraction will be the aptamer-enriched library used for subsequent rounds, and repeated rounds will generate aptamers with high affinity for HSA (fig. 2). This research

project will employ SELEX with NECEEM as a partitioning method for the selection of aptamers for HSA.



[21,22,23]**Figure 2. NEECEM-based partitioning of equilibrium mixtures.** (A) is the equilibrium mixture (EM) that contains unbound DNA, HSA (Target, abbrev. T), and DNA-HSA complex (DNA•T). (B) is an illustration of NECEEM-based separation of equilibrium mixtures. When the EM at t_0 is injected into the capillary, it is subjected to NECEEM separation (t_1). HSA (T) migrates the fastest, following with DNA-HSA complex and unbound DNA. The DNA•T continues to dissociate during separation, leaving smears of DNA dissociated from DNA•T, and T dissociated from DNA•T. The collection window of bound DNA to HSA would be the area of smear that is DNA dissociated from DNA•T.

RESULTS

Aptamer selection for HSA starts with a naive library of single stranded DNA. In this research, we are using recombinant HSA (rHSA) and two different DNA libraries: N40 and structured library. N40 DNA library has a randomized region of 40 nucleotides flanked by two 20 constant primer binding regions (overall, 80 bases). The structured library has a randomized region of 19 nucleotides flanked by one 19 and one 18 constant primer binding regions, and an overall length of 100 bases (fig 3). Both 5'ends of the libraries are fluorescently labelled with 56-FAM. The forward primers are labelled with Alexa 488, and the reverse primers are labelled with biotin.

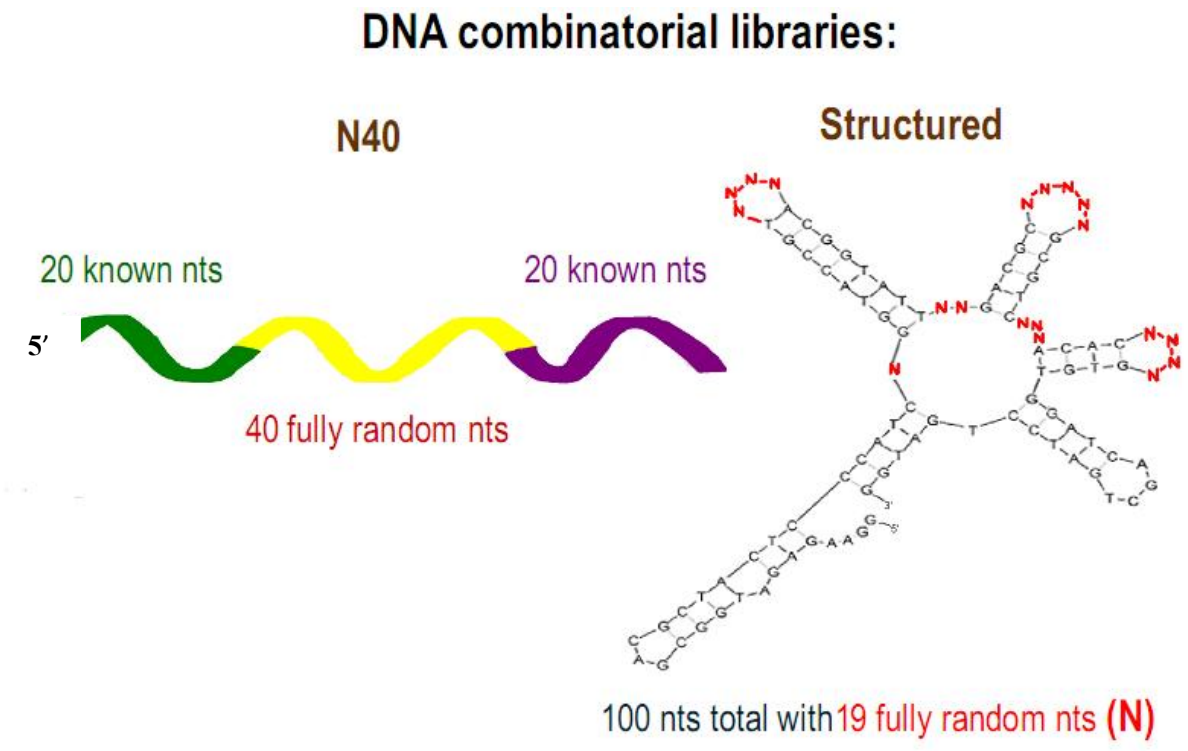


Figure 3. Illustration of two combinatorial DNA libraries used for rHSA aptamer selection. N40 library has 40 randomized base sequences, generating 4^{40} possible DNA sequences. Structured library has 19 randomized bases, and potentially holds the structure shown above with loops. Both libraries contain known sequences that are primer binding regions at the two terminals in order to amplify the collected aptamer pools in SELEX.

PCR optimization of structured DNA library (figs. 4-5)

Prior to aptamer selection for rHSA, PCR optimization must be performed with the amplification of the structured library. Based on previous negative results of HSA aptamer selection using the structured library (unpublished data), we suspect this is because the limit of detection (LOD) for the PCR product is low (can only visualize 10^6 DNA molecules as the lowest detection limit). Thus, the goal is to improve on the LOD to approximately 10^3 (LOD of N40 DNA library), so that, if aptamers are selected at low amounts, they can be detected after amplification.

The amplification of DNA aptamers involves two steps: symmetric PCR and asymmetric PCR. Symmetric PCR amplifies single-stranded DNA library to generate double stranded DNA through the addition of equal concentrations of forward and reverse primers. Forward and reverse primers bind to the primer binding regions of the antisense and sense strand respectively on the DNA libraries. This results in an abundant double-stranded pool of aptamers. Asymmetric PCR then amplifies the double-stranded aptamer pool to produce an abundance of one strand (sense strand), which corresponds to the original template in the DNA library. This is achieved by adding the forward primer in excess compared to the reverse primer (25).

To improve on the LOD of the structured library, we changed both the mastermix and PCR program according to the procedure proposed in [26]Luo *et al* that utilizes the same structural DNA library. One modification from the proposed mastermix recipe is the reduction of forward primer concentration in symmetric PCR from 1uM to 0.5uM. The old and new mastermix and the PCR program are compared in figures 4 and 5.

A

Master Mix Symmetric					Asymmetric
Reagents	**Stock concentration**	**1X concentration**	**1.1X**	**Volume of stock (uL)**	
Tris-HCl	0.6M	0.1M	0.11M	18.33	-
KCl	1M	50mM	55mM	5.5	-
Triton X-100	100%	1%	1.1%	1.1	-
MgCl2	25mM	1.9mM	2.09mM	8.36	
dNTPs Mix	10mM	0.3mM	0.33mM	3.3	
Forward primer (Alexa 488)	10uM	**0.5uM*** (1uM)	**0.55uM*** (1.1uM)	**5.5**	22uL (1X-1uM)
Reverse primer (biotin)	10uM	1uM	1.1uM	11	0.55uL (1X-50nM)
Taq polymerase	5U/mL	0.1U/uL	0.11U/uL	1.1	
ddH2O				45.81	51.31
Total volume				**100uL**	**100uL**

B

Master Mix Symmetric					Asymmetric
Reagents	**Stock concentration**	**1X concentration**	**1.1X**	**Volume of stock (uL)**	
5X green buffer	5X	1X	1.1X	22	-
MgCl2	25mM	2.5mM	2.75mM	11	-
dNTPs Mix	10mM	0.2mM	0.22mM	2.2	-
Forward primer (Alexa 488)	10uM	0.3uM	0.33uM	3.3	11uL (1X-1uM)
Reverse primer (biotin)	10uM	0.3uM	0.33uM	3.3	0.55uL (1X-50nM)
Taq polymerase	5U/mL	0.025U/uL	0.0275U/uL	0.55	-
ddH2O				57.65	52.7
Total volume				**100uL**	**100uL**

Figure 4. Recipes of mastermix for the amplification of structured DNA library. (A) is the proposed mastermix in [26]Luo *et al*. Differences include the replacement of 5X green buffer with the separate components of Tris-HCl and KCl in the mastermix. Also, triton X-100 is included in the mastermix. (B) is the old mastermix recipe that yields low LOD for the amplification of the structural library. *modification from the original protocol, a reduced concentration from 1uM to 0.5uM.

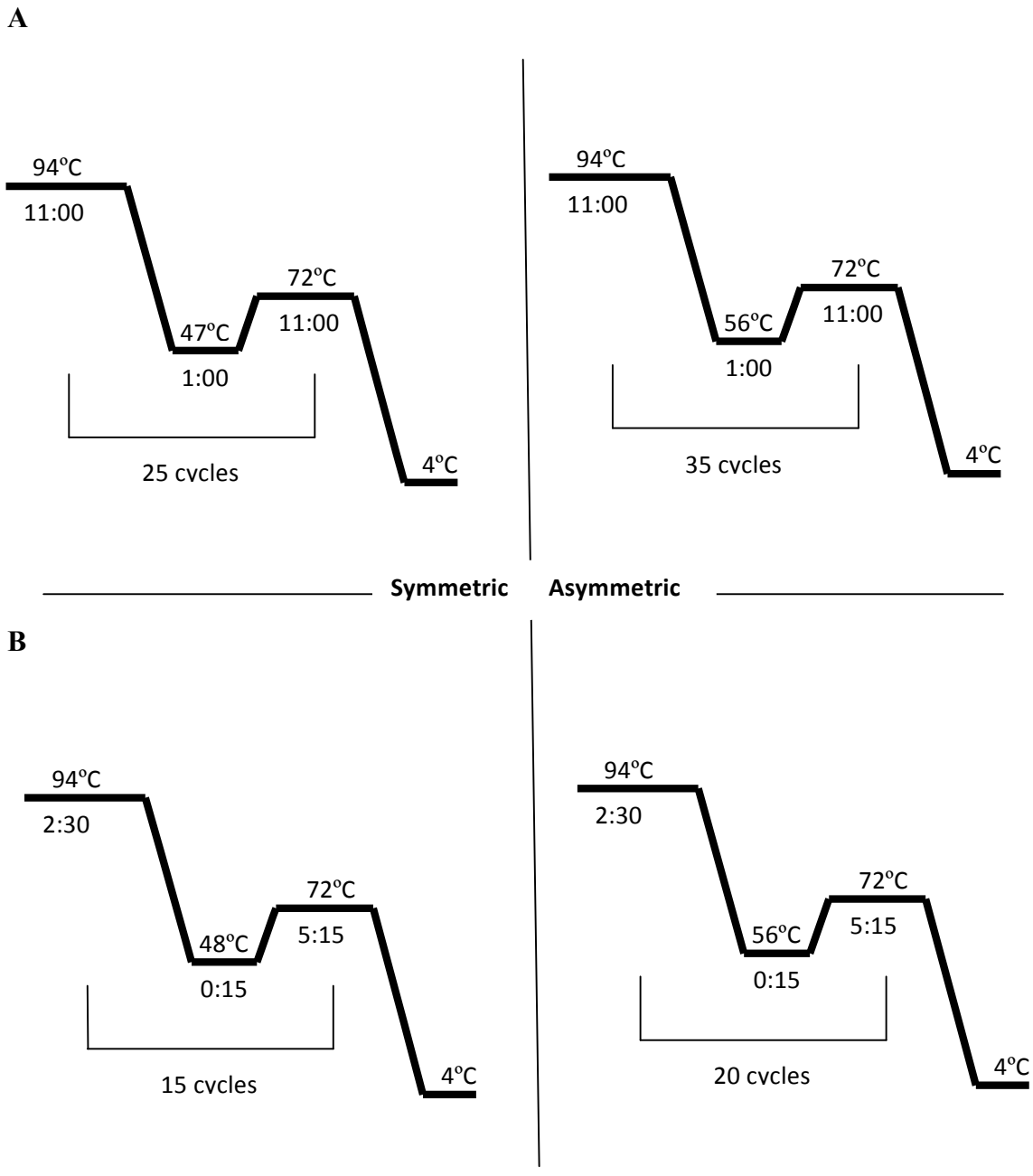


Figure 5. PCR temperature program for structured DNA library. (A) is the proposed PCR program for both symmetric and asymmetric from [26]Luo *et al*. There is an increase in number of cycles and time at each specific temperature. (B) is the original PCR program that yields low LOD.

PCR Optimization of Structured DNA Library (figs. 6 &7)

Using the mastermix and PCR program from [26]Luo *et al*, the LOD of the structured library was tested with concentrations of 10^2, 10^5, and 10^7 DNA molecules. The symmetric and asymmetric gel pictures are shown in figure 6. The 25 symmetric PCR cycles were checked to ensure no contamination was present. However, there was a DNA product-band present in the negative control after 35 asymmetric PCR cycles, so there was too much amplification, and a reduction of the number of asymmetric cycles was needed.

We then decided to reduce both symmetric and asymmetric cycles for the improvement of the LOD of the structured library to the following combinations (fig. 7):

A) 15 symmetric cycles with 15 asymmetric cycles.

B) 15 symmetric cycles with 20 asymmetric cycles.

C) 15 symmetric cycles with 25 asymmetric cycles.

The concentrations of the structured library tested were 10^2, 10^5, 10^6 DNA molecules. From figure 7, the combination B with 15 symmetric cycles and 20 asymmetric cycles appeared to improve the structural library's LOD to 10^2 DNA molecules (there was a faint DNA library band present) with no amplification of the negative control.

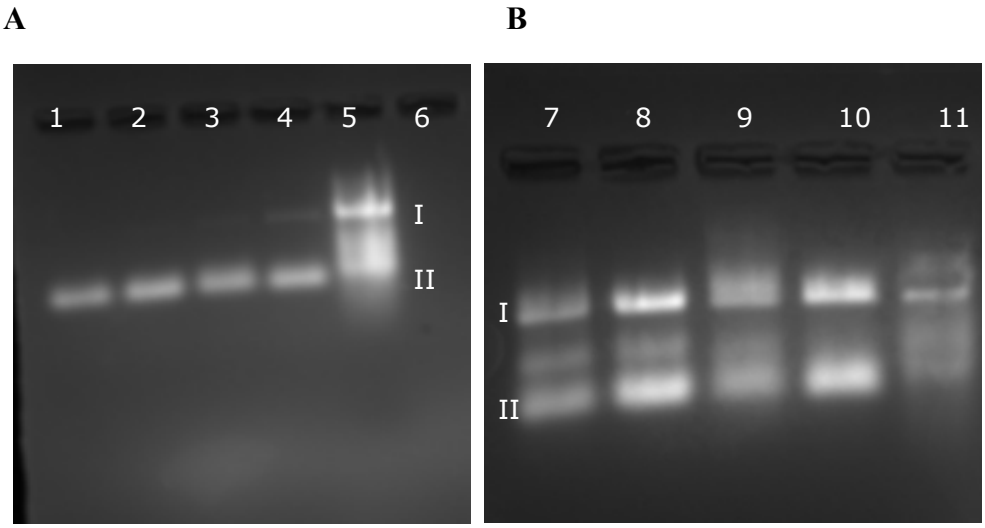


Figure 6. Symmetric and Asymmetric amplification of structured DNA library. (A) is after 25 symmetric cycles. (B) uses the product of (A) and shows the results after 35 asymmetric cycles. Columns 1 & 7 show the negative control that has no DNA library. Columns 2 & 8 are samples containing 10^2 DNA molecules. Columns 3 & 9 are samples containing 10^5 DNA molecules. Columns 4 & 10 are samples containing 10^7 DNA molecules. Columns 6 & 11 show the positive control, which amplifies 10nMof structured library. Row I is the DNA product band that is fluorescently labelled with 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples were run on 3% agar gel at 30 minutes at a voltage of 100V.

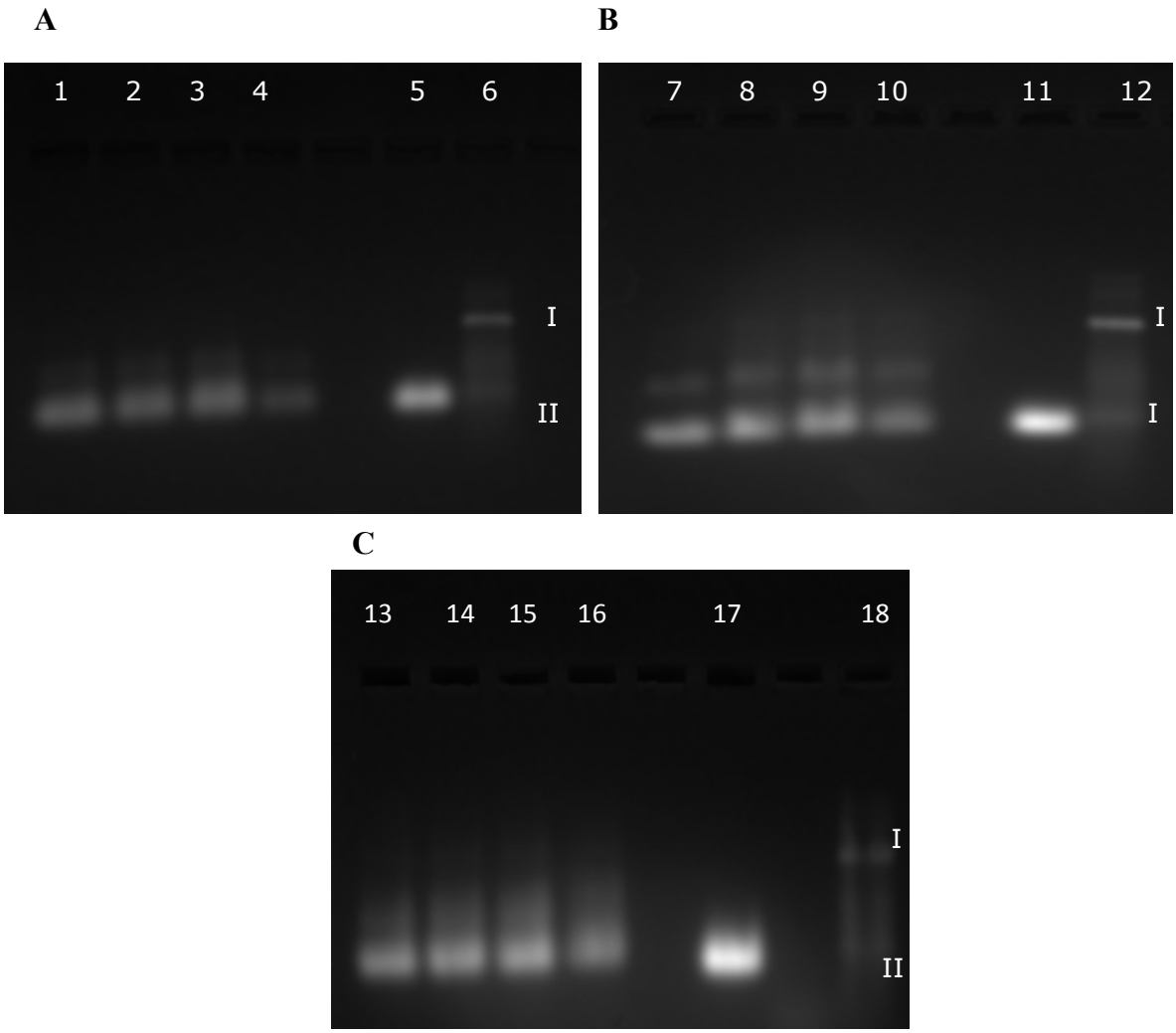


Figure 7. Improving LOD of structured DNA library with decreased symmetric cycles. All samples are amplified with 15 symmetric cycles, followed with the following numbers of asymmetric PCR cycles: (A) 15; (B) 20; and (C) 25. Columns 1, 7 & 13 are the negative control that has no DNA library. Column 2, 8 & 14 are samples containing 10^2 DNA molecules. Columns 3, 9 & 15 are samples containing 10^5 DNA molecules. Columns 4, 10 & 16 are samples containing 10^6 DNA molecules. Columns 5, 11, & 17 are the forward primer Alexa 488 only. Columns 6, 12 & 18 are the positive control, which amplifies 10nMof structured library. Row I is the DNA product band that is fluorescently labelled with 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples are run on 3% agar gel at 30 minutes at a voltage of 100V.

Removing Secondary Products present in the Amplification of Structured DNA Library (figs. 8-10)

The experiments just performed resulted in improvements the LOD of the structured library from 10^6 to 10^2 with 15 symmetric and 20 asymmetric PCR cycles. This improvement is important for selecting aptamers for HSA, as it will allow us to amplify aptamers at low concentrations. However, after the amplification there appears to be secondary products present in all the DNA samples, including the negative sample.

Our next goal is, therefore, to try and remove this secondary product from the amplification of the structured DNA library. To achieve this, we modified the PCR program to suit our taq polymerase used, while keeping the optimal number of cycles determined earlier. The new PCR program is shown in figure 8. We tested the structured DNA library with concentrations of 10^4 and 10^6 DNA molecules with the new PCR program to see if changing the PCR program worked (fig. 9).

Figure 9 indicates that the presence of secondary products in the amplification of the structured library persists with the new PCR program, but the improvement of the LOD is still maintained. We then changed the annealing temperature of the symmetric cycle from 47ºC to a range of 49ºC to 60ºC. We tested samples of 10^6 DNA molecules of the structured library in this temperature range (fig. 10). Although increasing the annealing temperature does appear to remove the secondary structures formed through amplification of the structured library, there is, however, also a loss of amplification of the structured library (no DNA band present).

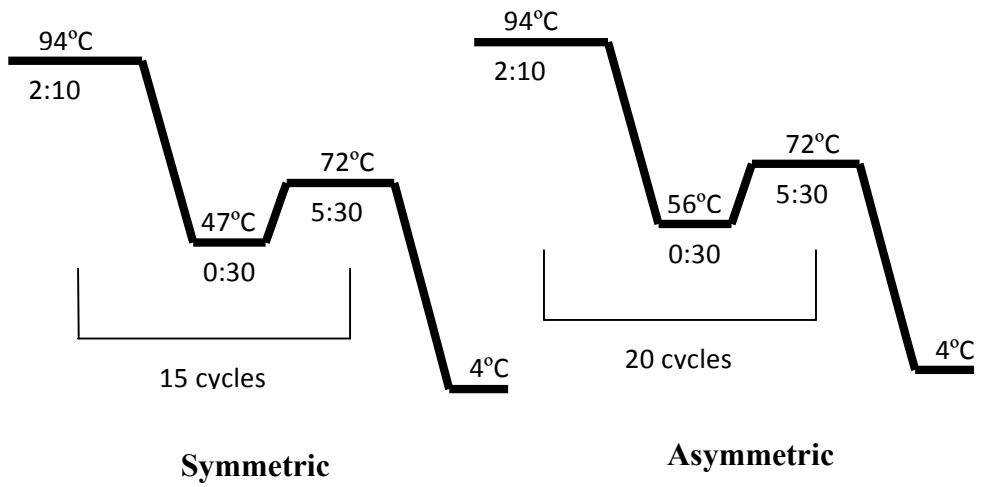


Figure 8. PCR program suitable for Taq Polymerase used in amplification of structured DNA library. This PCR program is determined using the instructions accompanying the [1]Taq polymerase used in this experiment.

[1]GoTaq Hot Start Polymerase instructions: the PCR program was made by following the general guidelines for amplification by PCR in the instructions, supplied by Promega.

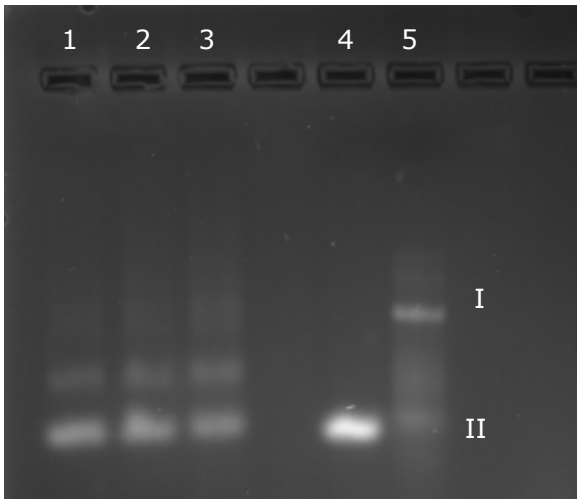
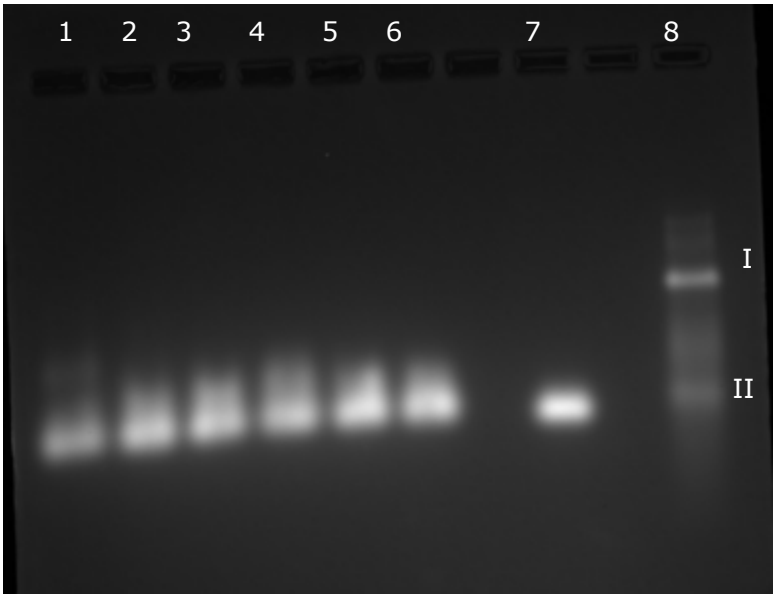


Figure 9. Amplification of structured DNA library with new PCR program (from fig. 8). The PCR program was designed to suit the Taq polymerase used in this experiment. The samples were amplified in 15 symmetric cycles and 20 asymmetric cycles. Column 1 is the negative control that has no DNA library. Column 2 is the sample containing 10^4 DNA molecules. Column 3 is the sample containing 10^6 DNA molecules. Column 4 is the forward primer Alexa 488 only. Column 5 is the positive control, which amplifies 10nMof structured library. Row I is the DNA product band that is fluorescently labelled with 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples are run on 3% agar gel at 30 minutes at a voltage of 100V.

A



B

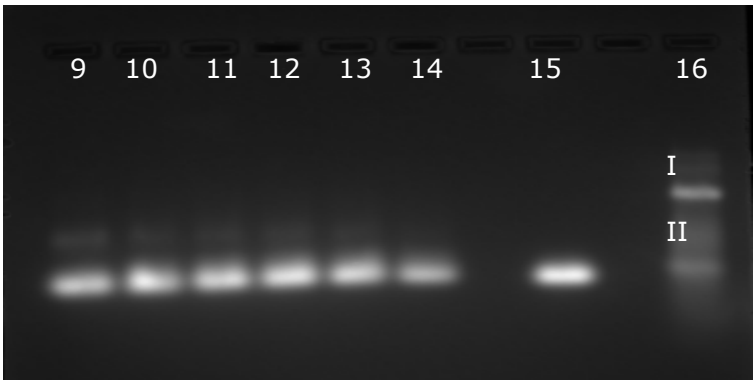


Figure 70. Amplification of 10^6 DNA molecules of the structured DNA library with different annealing temperatures. The samples were run in 15 symmetric and 20 asymmetric cycles. The annealing temperature was changed from 47ºC to 49ºC in column 2; 50ºC in column 3; 51ºC in column 4; 52 ºC in column 5; 54 ºC in column 6; 55 ºC in column 10; 57 ºC in column 11; 58 ºC in column 12; 59 ºC in column13; and 60 ºC in column 14. Columns 1 & 9 are the negative control that has no DNA library. Columns 7 & 15 are the forward primers Alexa 488 only. Columns 8 & 16 are the positive control, which amplifies 10nMof structured library. Row I is the DNA product band that is fluorescently labelled with 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples were run on 3% agar gel at 30 minutes at a voltage of 100V.

The LOD of the structured library with PCR improved to 10^2 DNA molecules when modified with 15 symmetric and 20 asymmetric cycles, but there were secondary products present. After changing the PCR cycles to suit the taq polymerase used (fig. 8) and increasing the annealing temperature of the PCR program (fig. 10), there were no improvements in removing the secondary products formed during the amplification of the structured DNA library.

We thus decided to follow exactly the protocol proposed in [26]Luo *et al* that utilizes the same structural DNA library, reverting the master-mix recipe of forward primer (Alexa 488) from 0.5uM to 1uM, and using the original 25 symmetric cycles followed with 35 asymmetric cycles for the amplification. The structured DNA library samples tested were 10^4 and 10^5 DNA molecules (fig. 11).

Figure 11 reveals the presence of a faint band in the negative control, but we concluded that this band was negligible. The band is much fainter than the bands showing 10^4 and 10^5 DNA molecules. Furthermore, the concentration of the DNA library to be used to select for aptamers will be much higher than these samples of 10^4 and 10^5 DNA molecules. As a result, the presence of the non-specific amplification in the negative control will be insignificant.

Figure 11 also reveals the presence of secondary products, but in this experiment we cut off the aptamer bands and used them for the next stage of the selection of aptamers for rHSA. Thus, secondary products were not a problem. The high LOD of the structured DNA library using the procedure in [26]Luo *et al* appeared to be maintained with an LOD of 10^4 DNA molecules (an improvement from the original low LOD of 10^6 DNA molecules).

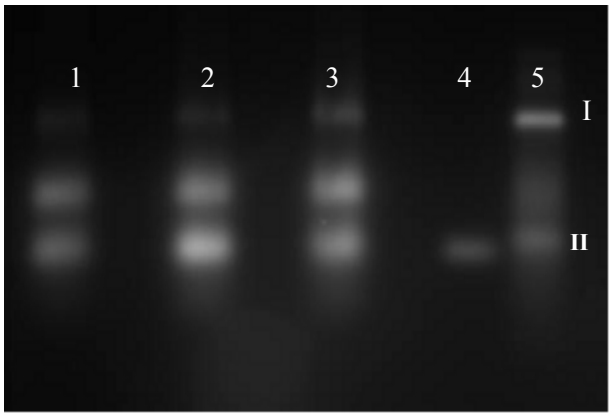


Figure 11. Amplification of structured DNA library with protocol proposed in [26]Luo *et al*. The samples were amplified with 20 symmetric cycles and 35 asymmetric cycles. Column 1 is the negative control that has no DNA library. Column 2 is the sample containing 10^4 DNA molecules. Column 3 is the sample containing 10^5 DNA molecules. Column 4 is the forward primer Alexa 488 only. Column 5 is the positive control, which amplifies 10nMof structured library. Row I is the DNA product band that is fluorescently labelled with Alexa 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples were run on 3% agar gel at 30 minutes at a voltage of 100V

rHSA-Aptamer Selection with Structured DNA Library (figs. 12-14)

After optimizing the PCR conditions for the amplification of the structured library, using [26]Luo *et al*'s protocol, the selection of aptamers for rHSA using this structured DNA library could now be carried out.

Aptamer selection using SELEX-NECEEM starts with determining the migration times of the rHSA and naive structured library in the capillary electrophoresis. This is crucial in finding the aptamer collection window for NECEEM-based selection of aptamers (18). The migrations times of the structured DNA library and rHSA are shown in fig.12 (see Appendix A for more details). The migration time of rHSA is approximately 10 minutes, and the migration time of structured DNA library is approximately 15 minutes. Since in the first few rounds, we were not able to observe a complex of the rHSA and DNA library, the aptamer-collection window was between the left-hand boundaries of the peaks of rHSA and structured library. Thus, the aptamer collection window was between 9 and 15 minutes (see fig.12).

Once the aptamer collection window was determined, the next step was the preparation of an equilibrium mixture subjected to NECEEM separation. We incubated the structured DNA library with the rHSA. We could now collect the bound-DNA through the aptamer-collection window, isolating the first pool of potential aptamers (fig 13). Figure 13 shows the collection of the bound-DNA and aptamer collection window.

We amplified the first pool of aptamers through the PCR cycles according to [26]Luo *et al*'s protocol. Figure 14 shows the result of the first round of selection of DNA binding to rHSA with the structured DNA library. There were, indeed, aptamers selected as figure 14 showed the presence of the structured DNA library amplified in the first pool of aptamers. The aptamers were then purified and used for the next round of selection.

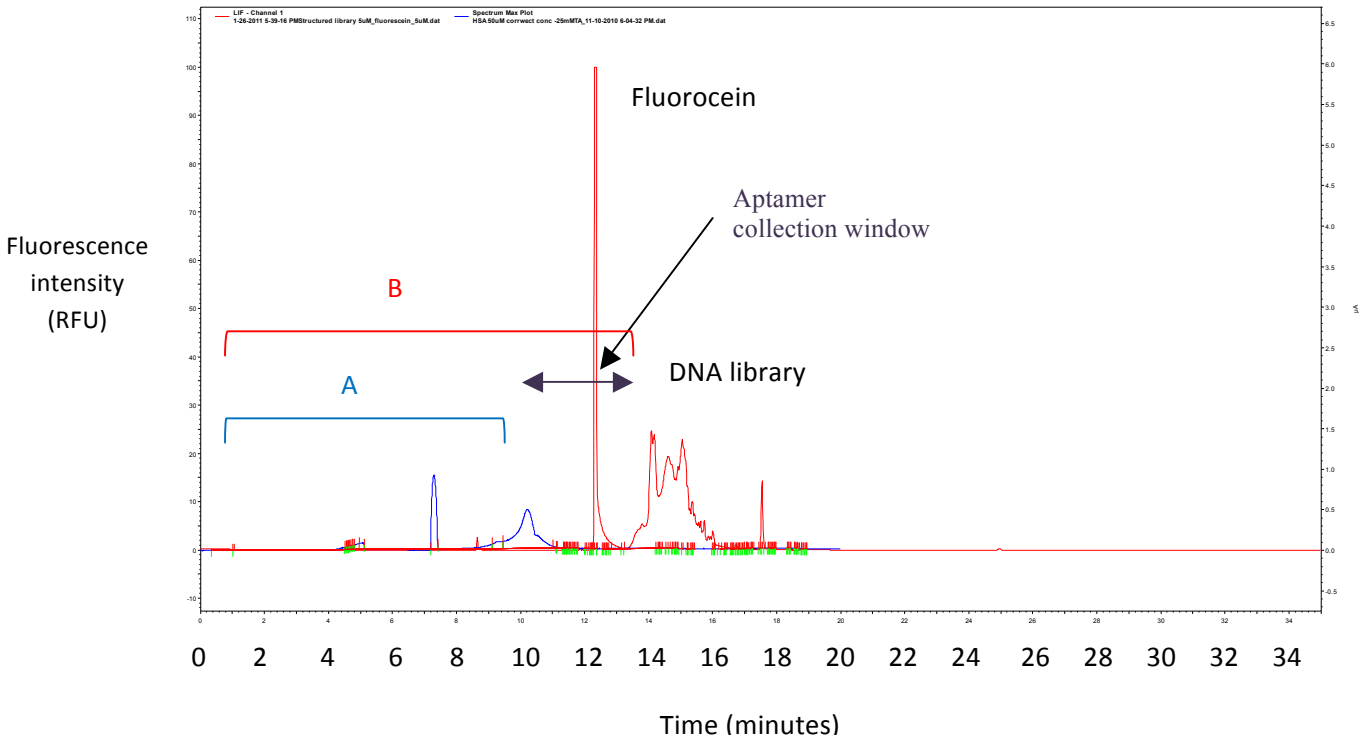


Figure 12. Aptamer collection window of rHSA and structured DNA library. (A) is the migration time of 50uM of rHSA protein injected in 80cm of capillary. 50uM of rHSA was diluted in 2X PBS solution. The capillary was rinsed with 25mM tris-acetate, pH 8. The detection device is UV-based photo diode array detector (PDA at 280nm. (B) is the migration time of 5uM structured DNA library with 5uM fluorocein as a control. The capillary length is 80cm to the detection window with 25mM tris-acetate running through the capillary. The detection device is laser-induced fluorescence (LIF). The migration time to the detector was multiplied by the conversion factor f, f= length total/ length to detector= 90cm/80cm=1.12.

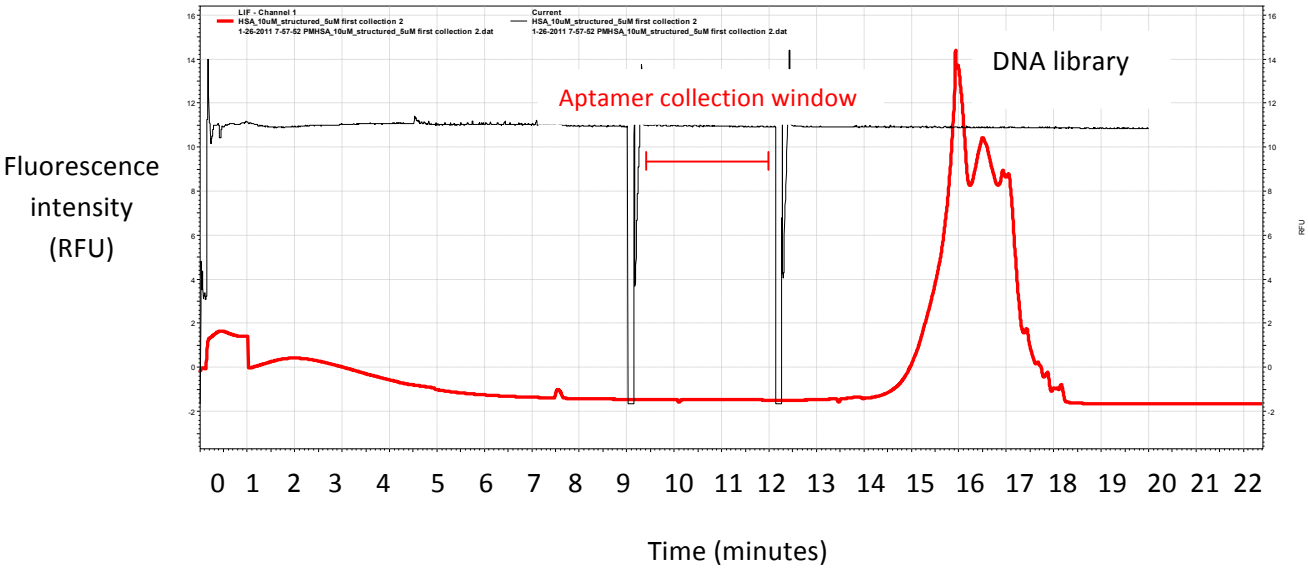


Figure 13. Collection of first round of bound structured DNA library to rHSA in NECEEM-based separation. The equilibrium mixture of rHSA (10uM) and structured DNA library (5uM) was injected into the capillary and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer. The collection window is between 9 and 12 minutes of migration time. The migration time to the detector was multiplied by the conversion factor f, f= length total/ length to detector= 90cm/80cm=1.12. The detector is laser-induced fluorescence (LIF).

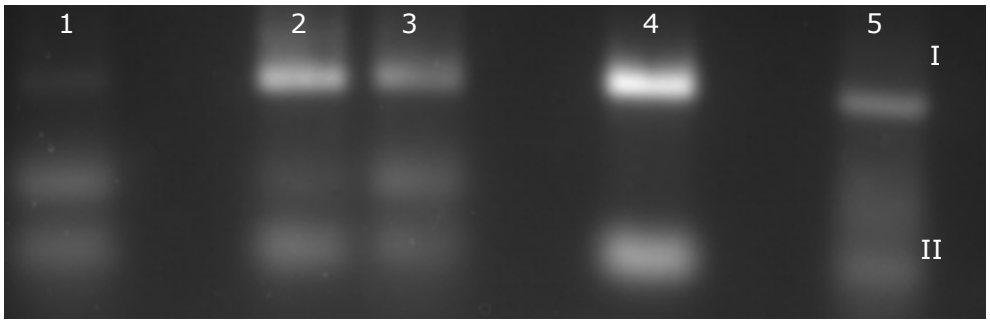


Figure 84. Amplification of first collected rHSA-Aptamer fraction with structured DNA library using SELEX-NECEEM. The aptamers isolated were amplified with 20 symmetric cycles and 35 asymmetric cycles. Column 1 is the negative control that has no DNA library. Columns 2 & 3 are the collected bound-DNA to rHSA (1st round of aptamers). Columns 4 & 5 are the positive control, which amplifies 10nMof structured library. Row I is the DNA aptamer that is fluorescently labelled with 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples were run on 3% agar gel at 30 minutes at a voltage of 100V

Thus, with the first pool of aptamers confirmed in the first selectionround of rHSA using the structured DNA library, we then purified the DNA aptamer by cutting the aptamer band off of the agar gel. This eliminates secondary products in the subsequent selection rounds. By following the accompanying DNA gel extraction spin protocol, the PCR product was now purified from double stranded DNA, primers, polymerase, and dNTPs. The first aptamer pool was purified and isolated.

In accordance with the SELEX method, we wanted to repeat the same selection process using this first-collected pool of aptamers as the initial naive structured DNA library to continue selection. However, there appeared to be insufficient DNA aptamers for the second round of selection, as the fluorescent intensity was 0.050 RFU (fig. 23 Appendix A), and the concentration determined through nanodrop was 2.9ug/uL (0.1uM).

Since we had already amplified a large amount of the first pool of aptamers, and this still resulted in low concentrations of aptamers for the second round, we decided to abandon the selection of rHSA using the structured DNA library. Instead, we focused on selecting aptamers for rHSA using the N40 DNA library.

rHSA-Aptamer Selection with N40 DNA Library (figs. 15-17)

N40 DNA library has a randomized region of 40 nucleotides flanked by two 20 constant primer binding regions (overall, 80 bases). Unlike the structured DNA library, there is no need to improve the LOD of the amplification procedure. The LOD of the N40 library is 10^2 DNA molecules, which is sufficient for amplification of collected aptamers.

Selection of aptamers with N40 DNA library was carried out in the same manner as with the structured DNA library. The procedure started with determining the migration times of the rHSA and naive N40 library in the capillary electrophoresis. Since we already had found the migration time for rHSA, that did not need to be repeated. The migration time for 1mM of N40 DNA library is approximately 19 minutes and the aptamer collection window is between 9 minutes and 19 minutes (fig. 15). However, in the actual aptamer selection, we used 5uM of the N40 DNA library, thus reducing the migration time to 16 minutes, so the aptamer collection window is actually between 9 minutes and 16 minutes.

The first selection round was carried out with the equilibrium mixture of N40 DNA library and rHSA subjected to NECEEM separation (fig. 16). After the bound DNA was collected in the aptamer-collection window, it was then amplified with PCR and purified with 30kDa cut-off filters. This produced the first pool of enriched aptamers, which was used as the starting DNA library for the second round of selection (see experimental section for detailed method). This SELEX method involves repeating the aptamer selection many times, using the previous DNA aptamer-enriched pool for successive rounds.

Figure 17 shows (1) the DNA product band of the first three selection rounds of rHSA aptamer with N40 library, and (2) a negative selection which selects against the rHSA-free equilibrium mixture.

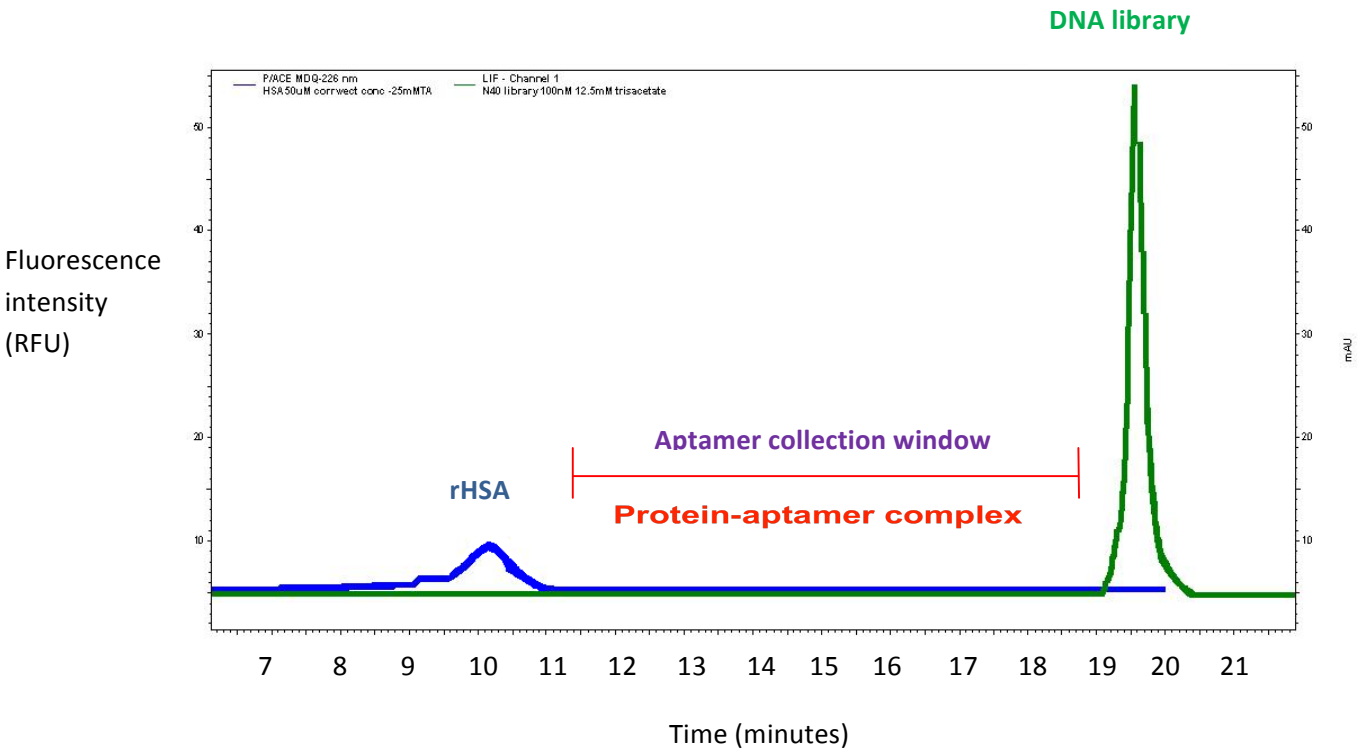


Figure 15. Migration times of N40 DNA library and rHSA with illustration of aptamer collection window. The migration time of 50uM of rHSA protein is around 9 minutes (injected in 80cm of capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer). 50uM of rHSA was diluted in 2X PBS solution. The detection device of rHSA is UV-based PDA at 230nm. N40 DNA library (1uM) has a migration time of 19.5minutes. The capillary length is 80cm with 25mM tris-acetate running through the capillary. The detection device is LIF. The collection window is between 9 and 19 minutes of migration time. The migration time to the detector was multiplied by the conversion factor f, f= length total/ length to detector= 90cm/80cm=1.12. The detector is laser-induced fluorescence (LIF).

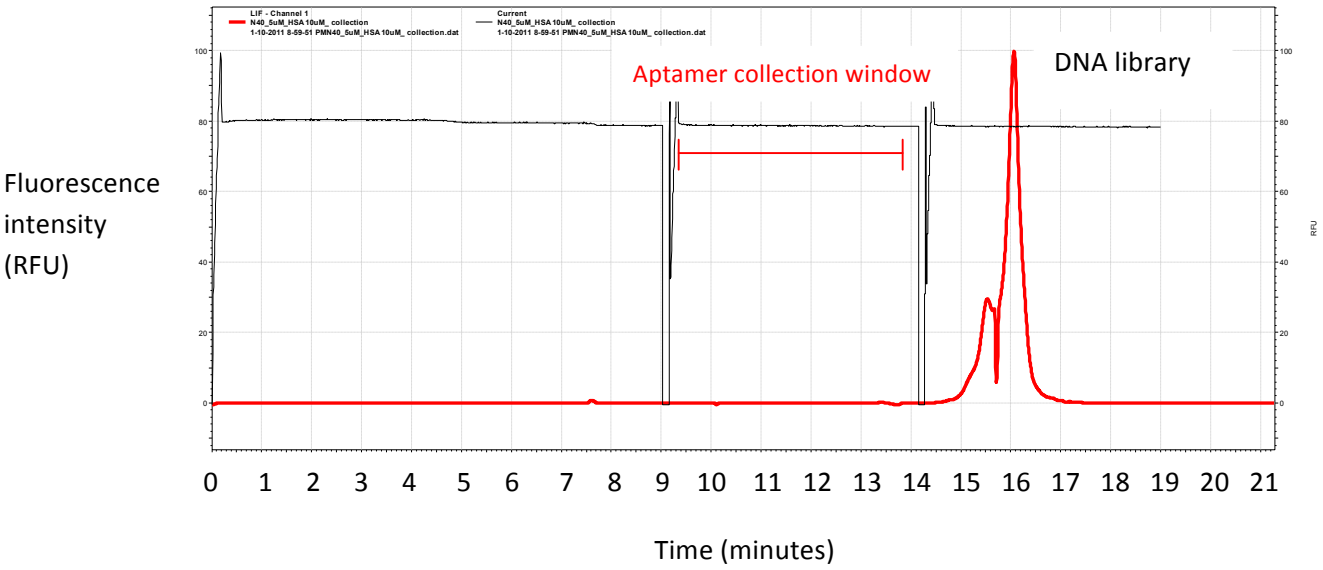


Figure 16. Aptamer collection window for the first selection round of rHSA and N40 DNA library in NECEEM-based separation of aptamers. The equilibrium mixture of rHSA (10uM) and structured DNA library (5uM) was injected into the capillary and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer. The collection window is between 9 and 14 minutes of migration time. The migration time to the detector was multiplied by the conversion factor f, f= length total/ length to detector= 90cm/80cm=1.12. The detector is laser-induced fluorescence (LIF).

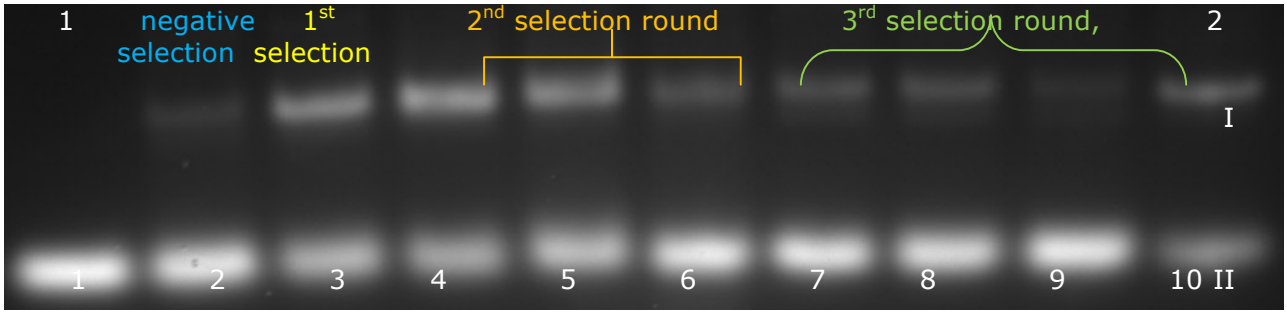


Figure 17. Amplification of the first three rounds of rHSA aptamer selection using N40 DNA library. The columns from left to right correspond to: (1) negative control in which no collected aptamer is added in the PCR mastermix; (2) negative selection in which aptamers are selected in incubation buffer containing no rHSA; (3) first round of aptamer selection against rHSA using naive N40 library; (4-6) second round of aptamer selection using the aptamer pool from 1st round as the starting DNA library; (7 -9) third round of aptamer selection using the aptamer pool from the second round of selection as the starting DNA library; (10) positive control which has 250nM of N40 DNA library. Each selection round was carried out according to the SELEX-NECEEM method described in the experimental section. The samples were amplified with 15 symmetric followed with 20 asymmetric PCR cycles. Row I is the DNA aptamer that is fluorescently labelled with 56-FAM. Row II is the fluorescently-labelled Alexa 488 forward primer. The samples were run on 3% agar gel at 30 minutes at a voltage of 100V

A negative selection against the rHSA-free sample was introduced in order to eliminate nonspecific binders to rHSA, and obtain rHSA-specific aptamers. However, as figure 17 shows, there was no enrichment of N40 DNA library towards the rHSA-free sample. Thus, there is no reason to perform further negative selection rounds. By contrast, the first three rounds of aptamer selection for rHSA showed an enrichment of library towards rHSA: there was amplification of the DNA library for all three selections.

rHSA- N40 DNA library Complex (figs. 18 & 19)

It is usually possible to detect target-oligonucleotide complexes using NECEEM in as few as three rounds. Indeed, [27]Krylov demonstrated that three rounds of NECEEM-based selection have been sufficient to reach a level of affinity which cannot be improved upon. However, in the 1st or 2nd round of selection, the bulk affinity of potential aptamers may be too low to detect complexes of rHSA-DNA library (21).

In our 3rd round of selection (c), we began to see a complex forming between rHSA and the 2nd round of aptamer pool (fig. 18). To confirm that this was indeed a complex, we ran samples of the 2nd round of aptamer pool both on its own (a) and with rHSA (b) (fig 19). Upon close examination, there appeared to be no complex forming since (b) did not have the additional complex peaks initially observed during (c). Additionally, using (a) as a control, we still did not observe a difference between the fluorescence signals of (a) and (b). Thus, we can conclude that the initial complex peak observed was probably due to some particles present in the DNA aptamer pool.

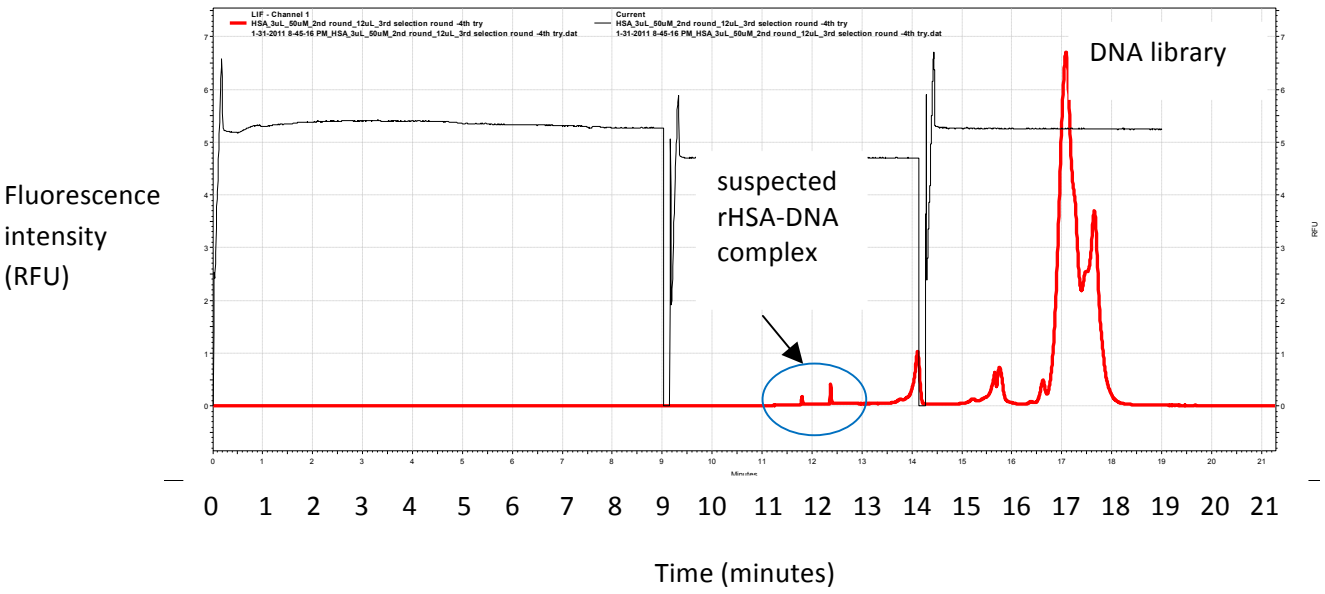


Figure 18. Third aptamer selection round for rHSA using N40 DNA library. The equilibrium mixture of rHSA (10uM) and 2nd round of aptamer pool was injected into the capillary and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer. In the collection window, a complex of rHSA and DNA library can be seen. The capillary length to the detector is 80 cm. The detector is laser-induced fluorescence (LIF).

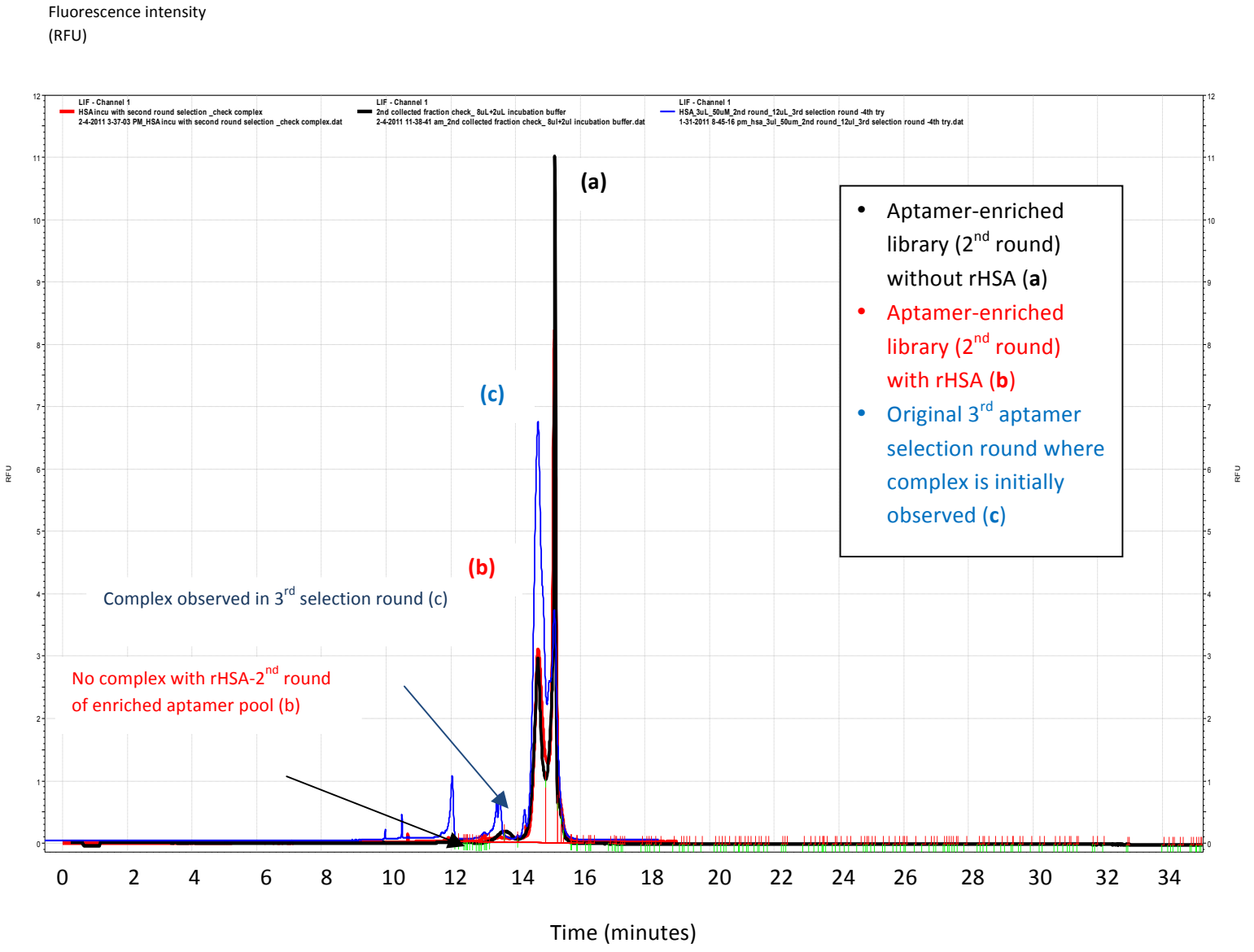


Figure 9. rHSA-N40 DNA complex formation. (a) is the 2nd round of aptamer pool on its own. **(b)** is 2nd round of aptamer pool with 10uM of rHSA. **(c)** is the initial 3rd round of selection, where the complex was first observed. There appears to be no complex forming in (b), corresponding to results in (c). (b) and (a) are identical, meaning that no complex was formed between DNA and rHSA. (a) and (b) were injected into the capillary and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer.

rHSA- N40 DNA library complex (figs. 20 & 21)

We then continued the 4th round of rHSA-aptamer selection. The lack of any complex forming confirms our conclusion on complex formation above (fig. 19). Despite this lack of complex formation, oligonucleotides from the library continued to be selected, and amplification of the 4th selection round showed DNA product band (fig. 20). Although there remained the possibility of an enrichment of DNA library towards rHSA, our 5th round of aptamer selection revealed no complex (fig.24 appendix A).

At this point, there appear to be no aptamers binding to rHSA or the affinity is too low for a complex to be detectable. Although the SELEX-NECEEM method theoretically should give an enriched aptamer pool with high affinity with a complex formed within five selection rounds, unfortunately we did not observe this. Thus, we decided to cease the selection of rHSA aptamers using this method.

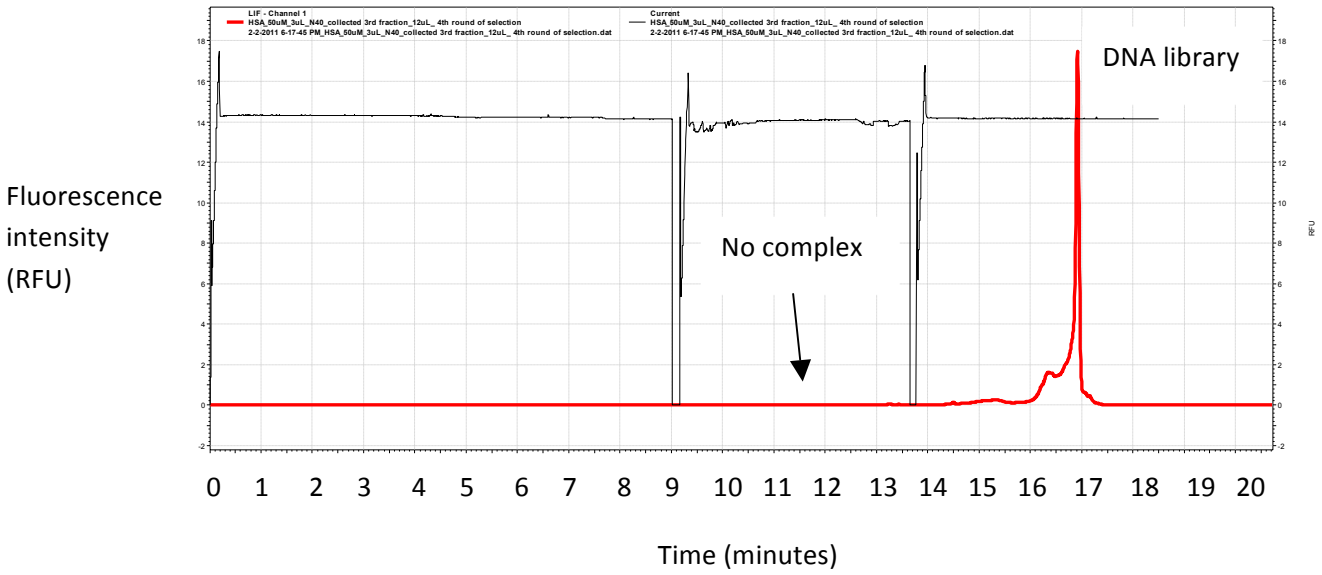


Figure 100. Forth aptamer selection round for rHSA using N40 DNA library. The equilibrium mixture of rHSA (10uM) and 3nd round of aptamer pool was injected into the capillary and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer. In the collection window, no complex of rHSA and DNA library can be seen. The detector is laser-induced fluorescence (LIF)

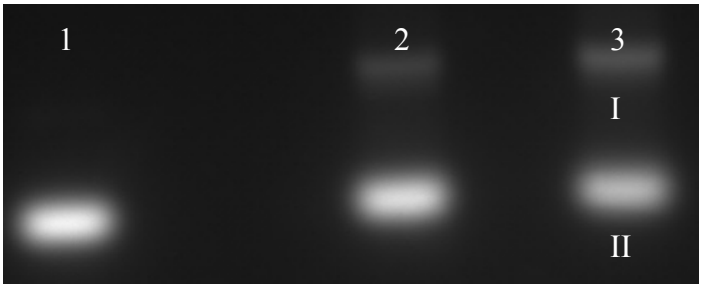


Figure 111. Amplification of the forth round collected rHSA-Aptamer fraction with N40 DNA library using SELEX-NECEEM. The aptamers isolated were amplified with 15 symmetric cycles and 20 asymmetric cycles. Column 1 is the negative control that has no DNA library. Column 2 is the collected bound-DNA to rHSA (4th round of aptamer selection). Column 3 is the positive control, which is 250nMof N40 DNA library. Row I is the DNA aptamer that is fluorescently labelled with 56-FAM. Row II is the Alexa 488 fluorescently-labelled forward primer. The samples are run on 3% agar gel at 30 minutes at a voltage of 100V.

DISCUSSION

PCR Optimization of Structured DNA Library

To improve the LOD of the amplification of the structured DNA library, we followed the mastermix and PCR program proposed in [26]Luo *et al*. However, we modified the concentration of the forward primer in symmetric PCR from 1uM to 0.5uM. In figure 4, we replaced the 5X green buffer in the original mastermix with tris-HCl and KCl, and added Triton X-100 in the new proposed mastermix. The addition of triton X-100, a non-ionic detergent, would improve the LOD as the detergent will stabilize taq polymerase and suppress unspecific amplification (28).

There was also a one-fold increase in Taq polymerase and primers, which explains why there was an increase in the amplification of product. In figure 5, the PCR program was changed so that there was an increased time in each temperature step of the PCR cycle. PCR cycle involves three steps: denaturing of the template by heat, annealing of the primers to the template DNA, and extension from the primers. Increasing the time period of each of these steps, especially the denaturing step, would improve product yield (LOD of the DNA library) because the DNA strand will be maintained in a linear state (28). This is especially important as the structured DNA library contains many loops (fig. 3), so the chances of the DNA library "snapping back" are very high.

According to [26]Luo *et al,* there was also an increase in cycle numbers for both symmetric and asymmetric from 15 to 25, and 20 to 35, respectively (fig. 5). Increasing cycle numbers does, indeed, improve LOD, but as figure 6 shows, there is amplification in the negative control. Thus, there was nonspecific amplification in the background. [28]Innis and Gelfan stated that the best way to avoid amplifying background products is to optimize the number of PCR cycles.

As a result, when we reduced both symmetric and asymmetric cycles to 15 symmetric and 20 asymmetric cycles (fig. 7), the nonspecific amplification in the negative control was removed and the LOD of the amplified structured DNA library improved to 10^2 DNA molecules, but it was very faint band.

With the improved LOD for the structured DNA library, we wanted to eliminate secondary products present in all samples (fig.7). Secondary products occur when only a fragment of the DNA template is amplified, generating unwanted segments of amplified DNA. The main cause can be attributed to premature termination of the taq polymerase. Since the taq polymerase we used is different from the one used in [26]Luo *et al*, its activity is not optimal, which may explain presence of secondary products. Thus, we changed the PCR program to suit out taq polymerase used (fig 8), while keeping the previously-determined optimal cycle number. The results, however, were not satisfactory (fig. 9); we did not observe the elimination of secondary products.

Another way to remove secondary products is to increase the annealing temperature, as secondary products may form from nonspecific annealing. Increasing annealing temperature will enhance discrimination against incorrectly-annealed primers and reduce mid-extension of incorrect nucleotides at the 3' end (28). When we increased the annealing temperature of the symmetrical cycles, however, the secondary products were removed, but there was no amplification of the structured DNA library (fig. 11). This may be due to the high selectivity that eliminates all DNA from being amplified, and explains why we did not observe a secondary product.

Having expended considerable effort to improve the LOD of the structured DNA library and remove secondary products, we thus decided to follow exactly the protocol

proposed in [26]Luo *et al.* We changed the master-mix recipe of the forward primer (Alexa 488) back from 0.5uM to 1uM, and used the original 25 symmetric cycles followed with 35 asymmetric cycles for the amplification.

We decided to use [26]Luo *et al*'s protocol because Luo *et al* has been using the same structured DNA library for selection of aptamers, and manually cutting off the DNA band, removing the secondary products. We similarly decided to cut the aptamer bands out of the gel, eliminating secondary products for future selection rounds. Thus, secondary products were not a problem. Furthermore, the high LOD of the structured DNA library using the procedure in [26]Luo *et al* appeared to be maintained with an LOD of 10^4 DNA molecules. Although we did initially improve the LOD to 10^2 DNA molecules, the DNA product band was fairly faint (fig. 7) compared to the DNA product band with [26]Luo *et al*'s protocol (fig.11). Thus, with all factors considered, [26]Luo *et al*'s protocol seemed to balance all factors: improved LOD, more amplification of DNA library, and secondary products still removable.

rHSA-Aptamer Selection with Structured DNA Library

Although we obtained the first pool of enriched aptamers (fig. 14) for rHSA using the structured DNA library, there appeared to be insufficient DNA aptamers for the second round of selection, with a low concentration of approximately 0.1uM for the first generated aptamer pool. This low concentration was after a total volume of 250uL amplification.

The loss of DNA product (first pool of aptamers) may occur during the purification phase. The purification of the structured DNA library involved cutting off the DNA-product band present on the gel under fluorescence to eliminate secondary product observed during the amplification of this particular DNA library. The cut-off bands were then purified using standard gel purification kit (AxyPrep DNAgel Extraction Spin Protocol). DNA might have been lost at any stage in this process.

Future research could consider changing the purification method. However, the problem of secondary product formation may not be eliminated even with a new purification method. According to previous research performed on rHSA using this structured DNA library (unpublished data) with standard purification procedures, purifying entire PCR product with secondary structures and using it for subsequent selection rounds, resulted in unsuccessful selection of aptamers for rHSA.

The purification technique could be improved by using polyacrylamide gel electrophoresis to separate the secondary ssDNA product from the ssDNA product band, which would give a higher resolution and potentially higher recovery.

rHSA-Aptamer Selection with N40 DNA Library

We carried out, in total, five rounds of aptamer selection for rHSA using N40 DNA library, with an additional negative selection round (figs.17 & 21). All five rounds of selection showed an enrichment of library towards rHSA: there was amplification of the DNA library for all five selections. The negative selection against the rHSA-free sample was introduced in order to eliminate nonspecific binders to rHSA (non-target specific aptamers), and to obtain rHSA-specific aptamers. We had already shown that the N40 DNA library does not bind to the rHSA-free sample, so we did not have to carry out more negative selection rounds (fig. 17).

In NECEEM-based separation, [27]Krylov demonstrated that three rounds of NECEEM-based selection are sufficient to reach a level of affinity which cannot be improved upon. Theoretically, we should see a complex formation by the third selection round, as the enrichment of aptamer pools would give a bulk affinity high enough to bind to rHSA and observe a complex. However, this was not the case for our rHSA selection. We initially observed a complex forming during the third round of selection, but were later confirmed as false.

However, the complex observed may not be false-positive. Instead, it may be explained by the high "off rates" of the selected aptamers. The aptamer-rHSA complex is very unstable, and so the complex completely decays during separation (29). As a result, no peak was later observed in a second try of the third selection round and in the fourth and fifth selection rounds. To improve on the bulk affinity of the aptamer pool to the rHSA, we could increase the concentration of rHSA, incubate the equilibrium mixture with rHSA and DNA library for longer time, or decrease the incubation temperature (19).

Another explanation could be that there were simply no rHSA-aptamers selected. The N40 library does not contain combinations of sequences that are binding to rHSA, and the complex peak may have been an accident caused by some random fluorescent signals from particles present in the DNA library. This is perhaps more in line with the results, as we did check the presence of the complex a second time and verified it with a negative control (with no rHSA). If this scenario is true, then further research projects could try to use a different DNA library for the selection of rHSA (hopefully a new DNA library would contain sequences that have high affinity for rHSA).

Another possible explanation could lie in the aggregation of rHSA in the capillary, causing no DNA aptamers binding to rHSA. HSA is proven to form aggregates. In a commercial sample of HSA tested, it was found that monomer, dimer, tetramer, and hexamers of HSA formed (30). To solve this problem, the partitioning method would have to be changed.

Dynabeads might be effective in separating bound aptamers from non-bound aptamers for the selection of rHSA using N40 DNA library. In brief, the process involves the amine group on the rHSA, conjugated to the dynabeads through surface-activated carboxylic acid groups on the beads. This forms rHSA-coupled dynabeads. The N40 DNA library is then added to the Dynabeads and potential aptamers are isolated by magnetic separation and subjected to subsequent purification. In accordance to the SELEX method, the first-isolated pool of aptamers is used as the starting library for more rounds of selection. In contrast to NECEEM-based separation, this separation method usually takes 10-15 rounds.

Thus further research projects could either change the DNA library used for the rHSA aptamer selection, still employing the SELEX-NECEEM method, or revert back to the

conventional partitioning methods used to separate target-bound DNA from non-target bound DNA, such as the proposed Dynabeads.

CONCLUSION

The ability to select rHSA aptamers would significantly advance the discovery of new biomarkers in human plasma. This would have potential medical applications in the matching of biomarker levels to different diseases. Our research in selecting aptamers for rHSA using the SELEX-NECEEM method unfortunately did not produce encouraging results. There is a history of slow progress in aptamer selection for protein targets, and we have shown that the SELEX-NECEEM method may not be effective for this particular protein target. Thus, further work may involve investigating new selection methods for rHSA.

REFERENCES

[1] Devine, D.D., Schubert, P. (2011). Proteomic applications in blood transfusion: working the jigsaw puzzle. *Proteomics and transfusion Medicine Vox Sanguinis , 100*, 84-91.

[2] www.sigma-aldrich.com. (2010). *Protein Depletion for Plasma and Serum Proteimic Analysis* . Retrieved from Sigma.

[3]Wintrobe, M. M. (2010). Blood proteomica. *Journal of proteomics , 73*, 466-467.

[4]Liotta, Lance A, Ferrai, Mauro, Petrocoin, Emanuel. (2003). Written in Blood. *Nature , 425*, 905.

[5]Anderson, N.Leigh, Polanski, Malu.,Pieper, Rembert., Gatlin, Tina.,Tirumalai, Radhakrishna S., Conrads, Thomas P., Veenstra, Timothy D., Adkins, Joshua N., Pounds., Joel G., Fagan, Richard., Lobley, Anna. (2004). The Human Plasma Proteome. *Molecular and Cellular proteomics* , 311-325.

[6]Tammen, H. Specimen Collection and Handling: standardization of blood sample collection. In A. Vlahou, *Methods in Molecular Biology: Clinical proteomics, methods and protocols* (Vol. 428, pp. 35-40). Totowa: Humana Press.

[7]Lathrop, Julia Tait., Anderson, N Leigh.,Anderson, Norman G., Hammond, David J. (2003). Therapeutic potential of the plasma proteome. *Current opinion in Molecular Therapeutics , 5* (3), 250-256.

[8] Travis, James., Pannell, Ralph. (1973). Selective removal of albumin from plasma by affinity chromatography. *Clinca Chimia Acta , 49*, 49-52

[9] Schubert, P., Devine, D.V. (2008). The application of proteomics to blood banking. *Proteomics in transfusion medicine , 3*, 30-36.

[10] Pieper, Rember., Su, Qin., Gatlin, Christine., Huang, Shih-Ting., Anderson, N.Leigh., Steiner, Sandra. (2003). Multi-compoment immunoaffinity subtraction chromatography: An innovative step toward comprehensive surbey of the human plasma proteome. *Proteomics , 3*, 422-432.

[11] Hage, D. S. (1999). Affinity Chromatography: A review of clinical applications. *Clinical Chemistry , 45* (5), 593-615.

[12] Klooster, Rinse., Maassen, Bram T.H., Stam, Jord C., Hermans, Pim W., Haaft, Mark R. ten., Detmers, Frank J.M., deHaard, Hans J., Post, Jan A., Verrips, C. Theo.. (2007). Improved anti-IgG and HSA affinity ligands: Clinical application of VHH antibody technology. *Journal of Immunological Methods , 324*, 1-12.

[13] De Bock, Muriel., Seny, Dominique de Seny., Meuwis, Marie-Alice., Chapelle, Jean-Paul., Louis, Edouard., Malaise, Michel., Merville, Marie-Paule., Fillet, Marianne. (2010). Challenges for biomarker disocery in body fluids using SELDI-TOF-MS. *Journal of Biomedicine and Biotecnology* , 1-10.

[14] Cristea, Illeana M., Gaskell, Simon J., Whetton, ANthony D. (2004). Proteomics techniques and their application to hematology. *Blood , 103*, 3624-3632.

[15] Travis, James., Bowen, Jean., Tewksbury, Duane., Johnson., David., Pannell, Ralph. (1976). Isolation of albumin from Whole human plasma and fractionation of albumin-depleted plasma. *Biochemistry journal , 157*, 301-306.

[16] Mateijtschuk, P., Dash, C.H., and Gascoigne, E.W. (2000). Production of uman albumin solution: a continually developing colloid. *Birtish Journal of Anaesthesia , 85* (6), 887-895.

[17] Steel, Laura F., Trottert, Michael G., Nakajima, Pamela B., Mattu, Taj S.,Gonye, Gregory., Block, Timothy. (2003). Efficient and Specific removal of albumin from human serum samples. *Molecular and cellular proteomics , 2*, 262-270.

[18]Javaherian, Sahar., Musheev, Michael U., Kanoatov, Mirzo., Berezovski, Maxim V., Krylov, Sergey. (2009). Selection of aptamers for protein target in cell lysate and their application to protein purifrication. *Nucleic Acids Research , 37(8)*,1-10.

[19]Beresovski, Maxim., Musheev, Michael U., Drabovich, Anderi p., Jitkova, Julia V., Krylov, Sergey. (2006). Non-SELEX: selection of aptamers without intermediate amplification of candidate oligonucleotides. *Nature Protocols , 1* (3), 1359-1368.

[20]Mayer, G. (2009). The CHemical Biology of Aptamers. *Angew. CHem. Int. Ed , 48*, 2672-2689.

[21]Mascini, M. (2009). *Aptamers in Bioanalysis.* New Jersey: John Wiley & Sons. Inc.

[22] Berezovski, Maxim., Drabovich., Krylova, Svetlana M., Musheev, Michael., Okhonin, Victor., Petrov, Alexander., Krylov, Sergev. (2005). Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures: A Universal Tool for Development of Aptamers. *J.Am. CHEM.SOC , 127*, 3165-3171.

[23]Krylov, S. N. (2006). Nonequilibrium Capillary Electroporesis of Equilibrium Mixtures (NECEEM): A Novel Method for Biomolecular Screening. *Journal of Biomolecular Screening , 11* (2), 115-121.

[24] Berezovski, Maxim V., Lchmann, Matthias., Musheev, Michael U., Mak, Tak W., Krylov, Sergey. (2008). Aptamer-Facilitated Biomarker Discovery (AptaBiD). *Jounrnal of American Chemical Society , 130*, 9137-9143.

[25]Poddar, S. (2000). Symmetric vs Asymmetric PCR and molecular beason probe in the detection of a target gene of adenovirus. *Molecular and Cellular Probes , 14*, 25-32.

[26] Luo, Xuemei, Mckeague, Maureen., Pitre, Sylvain et al. (2010). Computational approaches toward the design of pools for the in vitro selection of complex aptamers. *RNA , 16*, 2252-2262.

[27] Krylov, S. N. (2005). NECEEM for development, characterisation, and analytical utilisation of aptamers. *LabPlus international* .

[28] Innis, Michael A., Gelfand, David H. (1990). Optimization of PCRs. In M. A. Innis, *PCR Protocols: a guide to methods and applications* (pp. 3-11). New York: Academic Press Inc: Harcourt Brace Jovanovic Publishhers.

[29] Berezovski, Maxim., Nutio, Razvan., Li, Yingfu., Krylov, Sergey N. (2003). Affinity Analysis of a protein-aptamer complex using nonequilibrium capillary electrophoresis of equilibrium mixtures. *Analytical Chemistry , 75* (6), 1382-1386.

[30] Atmeh, Ragheb F., Arafa, Isam M., Ai-Khateeb, Maysoon. (2007). Albumin Aggregates: Hydrodynamic Shape and physio-chemical properties. *Jordan Journal of Chemistry , 2* (2), 169-182.

EXPERIMENTAL SECTION

Chemicals and materials

1G of recombinant HSA was ordered from Sigma Aldrich (product #A9731). Tris-(hydroxymethyl) amminomethane HCl (TRIS, > 99.9% purity) was purchased from BioBasic Inc. Acetic acid (glacial, >99.7% purity with assay) was purchased from BioBasic Inc. EDTA Disodium salt (ethylenediamine tetracetic acid), dehydrate (>99% purity) was purchased from EMD chamicals. 500g of Boric acid (>99.0% purity) was purchased from Bio Basic Inc. 500g of sodium borate decahydrate was purchased from EMD. 200mM of Tris-acetate (pH 8.2), NaCl, and MgCl2 was supplied by the Berezovski lab. 2X PBS solution (without Mg) was supplied by the Berezovski lab. 100mM of HCl solution (12M HCl supplied by the University of Ottawa). 100mM of NaOH solution (NaOH supplied by University of Ottawa). Purified agar powder was purchased from Himedia Laboratories Pvt. Ltd.PCR mastermix reagents were supplied by the GoTaq Hot Start Polymerase kit (catalog # M5005). The kit included 5X green biffer, $MgCl_2$ solution (25mM), PCR nucleotide mix (10mM), GoTaq Hot start polymerase (5u/ul). Fused-silica capillaries (diameter 75um) were purchased from Polymicro technologies. 3% agar gel was made by mixing 1.5g of agar powder in 50mL of TAE buffer. The mixture was shaken and micro-waved for one minute, poured on the gel cast and the agar solution was left for 20-30 minutes to solidify.

Buffers

All buffers were prepared using Milli-Q-quality deionized water. Buffers used in the capillary electrophoresis were filtered through minipore before use (100mM HCl, 100mM NaOH, deionized water, 25mM tris-acetate buffer).

100mL of 50X TAE buffer (500mM, pH 8.2) was made by adding 48.46g of tris-(hydroxymethyl) amminomethane HCl in 100mL of deionized water. Acetic acid (glacial) was added until the solution had a pH of 8.2, checked by pH meter. 20mL of EDTA was added in the mixture, and finally deionized water was added so the total volume was 100mL.

100mL of Tris-acetate buffer (25mM, pH 8.2) was prepared from 200mM tris-acetate buffer: 12.5mL of 200mM of tris-acetate buffer was added into 87.5mL of de-ionized water. 500mL of 100mM HCl solution was made by adding 4.17mL of 12M HCl in 495.83mL of de-ionized water. 500mL of 100mM NaOH solution was made by adding 2.0g of NaOH crystals in 500mL of de-ionized water. 100mL of incubation buffer (50mM Tris-Acetate, 50uM NaCl, 5uM $MgCl_2$) was made by adding 25mL of 200mM of tris-acetate, 0.292g of NaCl, and 0.102g of $MgCl_2$ together with 75mL of de-ionized water.

DNA library, primers

The DNA libraries and primers were ordered from Integrated DNA technologies (IDT). The structured DNA library contained 19 randomized nucleotides flanked one 19 and one 18 constant primer binding regions, with an overall length of 100 bases. The 5' end was labelled with 56-FAM. The library sequence is 5'-/56-FAM/ATA CCA GCT TAT TCA ATT GC(N:252525252) (N) (N) (N) GCA ATT (N) (N) (N) GTC (N)GG AC (N) (N) (N) (N) GTT C(N)G AC(N) (N)TC GGC G(N)(N) (N)CG CCG A(N)C TAT CT(N) (N)(N)(N) (N)AG ATA GTA AGT GCA ATC T-3'. The forward primer (fp) was labelled with Alexa

488. The fp sequence was 5'- /5Alex488N/ ATA CCA GCT TAT TCA ATT G-3'. The reverse primer (rp) was labelled with biotin; the sequence is 5'-/5Biosg/ AGA TTG CAC TTA CTA TCT-3'. The N40 DNA library contained a randomized region of 40 nucleotides flanked by two 20 constant primer-binding regions (overall, 80 bases). The 5'end of the DNA sequence is labelled with 56-FAM. The library sequence is 5'-/56-FAM/CTC CTC TGA CTG TAA CCA CG (N:25252525) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) (N)(N)(N) GCA TAG GTA GTC CAG AAG CC-3'. The forward primer (fp) was labelled with Alexa 488. The fp sequence is 5'- /5Alex488N/ CTC CTC TGA CTG TAA CCA CG-3'. The reverse primer (rp) was labelled with biotin; the sequence was 5'-/5Biosg/ GCC TTC TGG ACT ACC TAT GC-3'. The library (5 nmol) was dissolved in 50 uL of water to obtain 100 uM stock solution; the primers were dissolved in 50uL of water to get 100uM of stock solution. Primers and DNA library were stored separately at 20 °C.

Instruments

NECEEM procedures were performed using either a Beckman Coulter P/ACE MDQ apparatus or Beckman Coulter proteome lab PA 800; both were equipped with fraction collection and photodiode array detector (PDA) or laser-induced fluorescence (LIF). The thermocycler used for PCR amplification, both symmetric and asymmetric, was Eppendorf Mastercycler pro S. The centrifuge used for the purification of aptamers was Labnet prism R frin MBI Lab equipment. The imaging device used to visualize DNA product bands on agar gel was the Alpha Innotech Fluorochem Q imaging system. The pH meter used to make solutions was purchased from Fisher Scientific; model is accumet basic AB15.

Equipment

Centrifugal Devices/ cut off filters were purchased from Pall; the model is nanosep 30K omega. Gel Extraction Kit used for the purification of DNA product bands directly from the gel was purchased from Axygen Bioscience. The procedure of gel extraction followed the AxyPrep DNAgel Extraction Spin Protocol accompanying the kit.

Methods

Non-equilibrium capillary electrophoresis of equilibrium mixtures (NECEEM)

Capillary electrophoresis procedures were performed using the above stated instrumentation (Beckman Coulter models). A 90-cm long (80 cm to a detection window) uncoated fused silica capillary with an inner diameter of 75um and outer diameter of 360 um was used. Both the inlet and the outlet reservoirs contained the electrophoresis run buffers25 mM tris-acetate pH 8.2. Prior to NECEEM, the capillary was pressure-rinsed pressure for 20psi, 2 min with each of the following reagents: RNAse away,100mM HCl, 100mM NaOH, de-ionized water, tris-acetate 25mM *3. The sample (equilibrium mixture of rHSA and the selected DNA library) was injected into the capillary by pressure at 0.5psi for 5 seconds. Electrophoresis was then carried out for 25 minutes at 30KV. The direction of the electroosmotic flow was from the inlet to the outlet, positive electrode to negative electrode. The fraction between the rHSA migration time, and the DNA library migration time was collected (the aptamer collection window). The fraction collections were automated by replacing the regular outlet

reservoir with a fraction collection vial containing 5 uL of 25mM tris-acetate. At the end of each run, the capillary was rinsed with RNAse away for 2 minutes, 100 mM HCl for 2 min and 100 mM NaOH for 2 min, followed by a rinse with de-ionized water for 2 min.

CE migration time:
A new capillary was inserted (total length 90cm, with 80cm to the detection window) into the cartilage. Prior to each run, the capillary was pressure-rinsed for 20psi, 2 min with each of the following reagents: RNAse away,100mM HCl, 100mM NaOH, de-ionized water, tris-acetate (25mM) *3. The temperature of the capillary is set at 20ºC.
For the **structured** and **N40 DNA library**, the CE is equipped with LIF detector. 10uL of the 5uM structured DNA library with 5uM fluorocein was made and this solution was injected by pressure at 0.5psi for 5 seconds, waited for 0.10minutes, and electrophoresis was carried out for 20 minutes at 30kV. The migration time to the detector was observed when a fluorescence signal (peak) was present. This time was multiplied by the conversion factor f, f= length total/ length to detector= 90cm/80cm=1.12.

For **rHSA**, the CE was equipped with photodiode array detector (PDA). 10uL of 50uM HSA was prepared in 2X PBS solution (no MgCl), and this solution was injected by pressure at 0.5psi for 5 seconds, waited for 0.10minutes, and electrophoresis was carried out for 20 minutes at 30kV. The migration time to the detector was observed of a peak at an absorbance of 280nm. This time was multiplied by the conversion factor f, f= length total/ length to detector = 90cm/80cm=1.12.

In the first few selections when no rHSA-DNA complex was observed, the aptamer collection window spans from the peak of the rHSA to the peak of the DNA library used. For the structured DNA library, it was between 9 and 15 minutes. For N40 library, the collection window was between 9 and 14 minutes of NECEEM-based separation.

Selection of aptamers for recombinant HSA by NECEEM
Prior to each round of selection, 10uM of DNA library was heated to 50ºC and cooled down to 20ºC at a rate of 0.5ºC/s in a thermocycler. For the first round of selection, 5uM of DNA library after temperature treatment (N40 or structured DNA) was incubated with 10uM of rHSA in a total volume of 10uL (with incubation buffer) at room temperature for 10 minutes. For the second and later selection rounds, the equilibrium mixture was changed into 10uM of rHSA in a total volume of 15uL with 12uL of the previous round of collected fraction. While this equilibrium mixture was incubating, the capillary was rinsed with the 4 different solutions described previously in the NECEEM section above. After incubation, the equilibrium mixture was subjected to NECEEM separation. The detailed program parameters for injection and separation of the individual DNA libraries are described below:
N40 injection pressure is 0.5psi for 5 seconds>-wait for 0.10minute->separate for 9minutes at 30kV->autozero->collection fraction at 9 minutes to 14 minutes with 30kV for 5 minutes-> continue separation at 30kV for 10 minutes.
Structured injection pressure is 0.5psi for 5 seconds>-wait for 0.10minute->separate for 9minutes at 30kV->autozero->collection fraction at 9 minutes to 13 minutes with 30kV for 4 minutes-> continue separation at 30kV for 8 minutes.

Amplification of structured DNA and N40 DNA library
The collected fraction from the aptamer collection window was the pool of aptamers (so the first collected fraction was the first pool of aptamers, and so on). This collected fraction was stored at -20ºC until used.

N40: collected fraction (aptamer pool) for the N40 DNA library was amplified through 15 symmetric and 20 asymmetric PCR cycles. 2uL of collected fraction was added to 18uL of symmetric mastermix, and ran for 15 symmetric cycles. Then 2uL of the symmetric PCR product was added into 18uL of asymmetric mastermix and underwent 20 asymmetric PCR cycles. For mastermix ingredients, see appendix B.
Structured: collected fraction (aptamer pool) for the structured DNA library was amplified through 25 symmetric and 35 asymmetric PCR cycles. 2uL of collected fraction is added to 18uL of symmetric mastermix, ran for 15 symmetric cycles. Then 2uL of the symmetric PCR product was added into 18uL of asymmetric mastermix and underwent 20 asymmetric PCR cycles (see [26]Luo *et al* for details)
PCR mastermix regaents (GoTaq Hot Start Polymerase kit) were bought from Promega.

DNA purification
Structured DNA library purification after sym/asym amplification followed the procedure of gel extraction by AxyPrep DNAgel Extraction Spin Protocol. **N40** DNA library purification after sym/asym amplification used nanosep 30K omega cut off filters: streptavadin magnetic beads were added to the PCR mixture in a ratio of 3:10 (3uL of streptavin for every 10uL of PCR mixture). This mixture was then incubated for 20 minutes at room temperature. Magnetic beads was then used to pull down the streptavin-coated magnetic beads. The supernatant of the PCR mixture was transferred into the 30kDa cut off filters. The purification step from this step onwards followed the instructions accompanying the nanosep 30K omega cut-off filters.

APPENDICES

Appendix A

Migration times of rHSA and structured DNA library

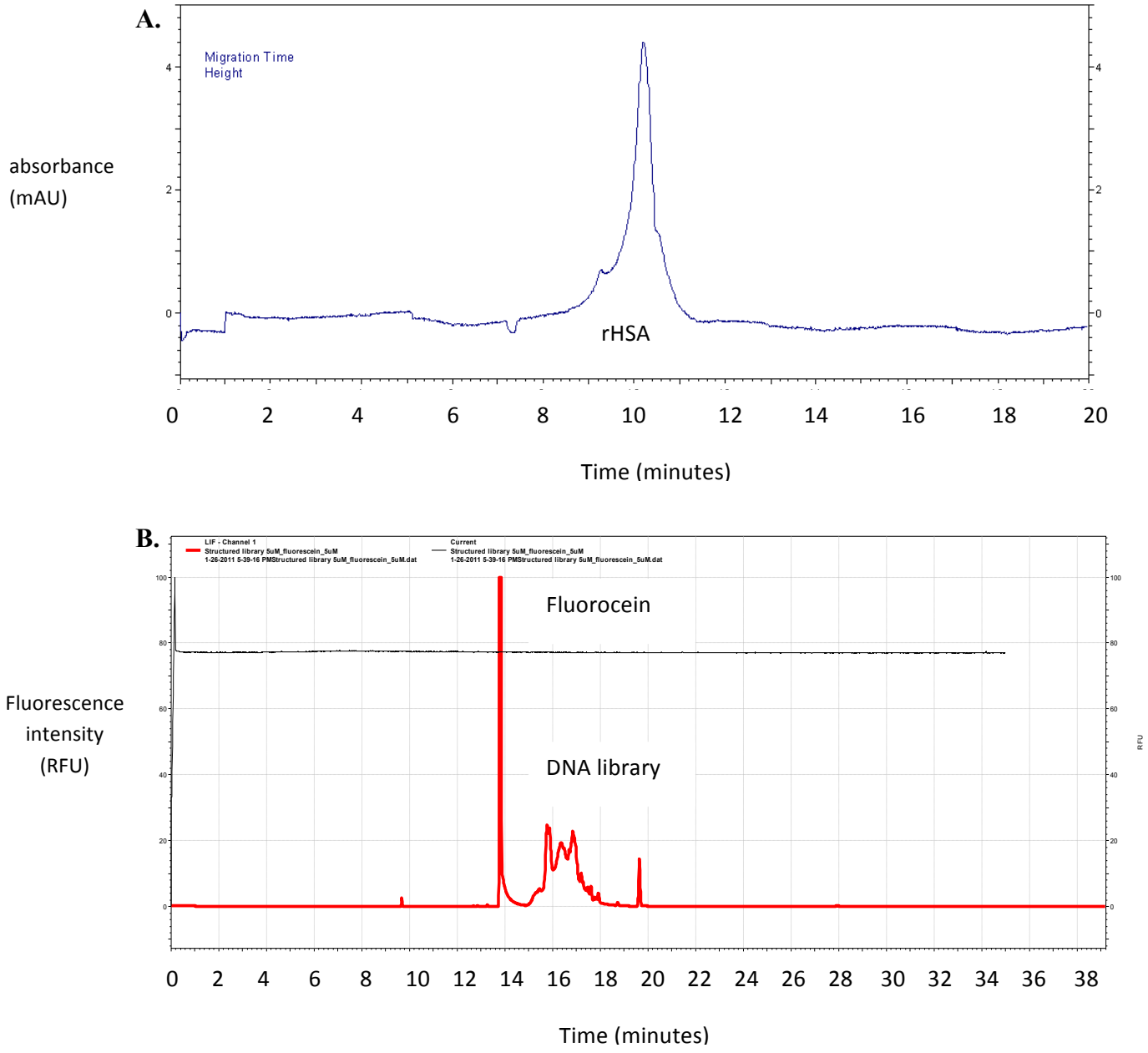


Figure 22. rHSA and structured DNA library migration times on capillary electrophoresis. (A) is the migration time of 50uM of rHSA protein of approximately 10 minutes. The detection device is UV-based PDA at 280nm. (B) is the migration time of 5uM structured DNA library with 5uM fluorocein as a control is approximately 15 minutes. The detection device is laser-induced fluorescence (LIF).

Concentration check of 1st round of selection with structured library

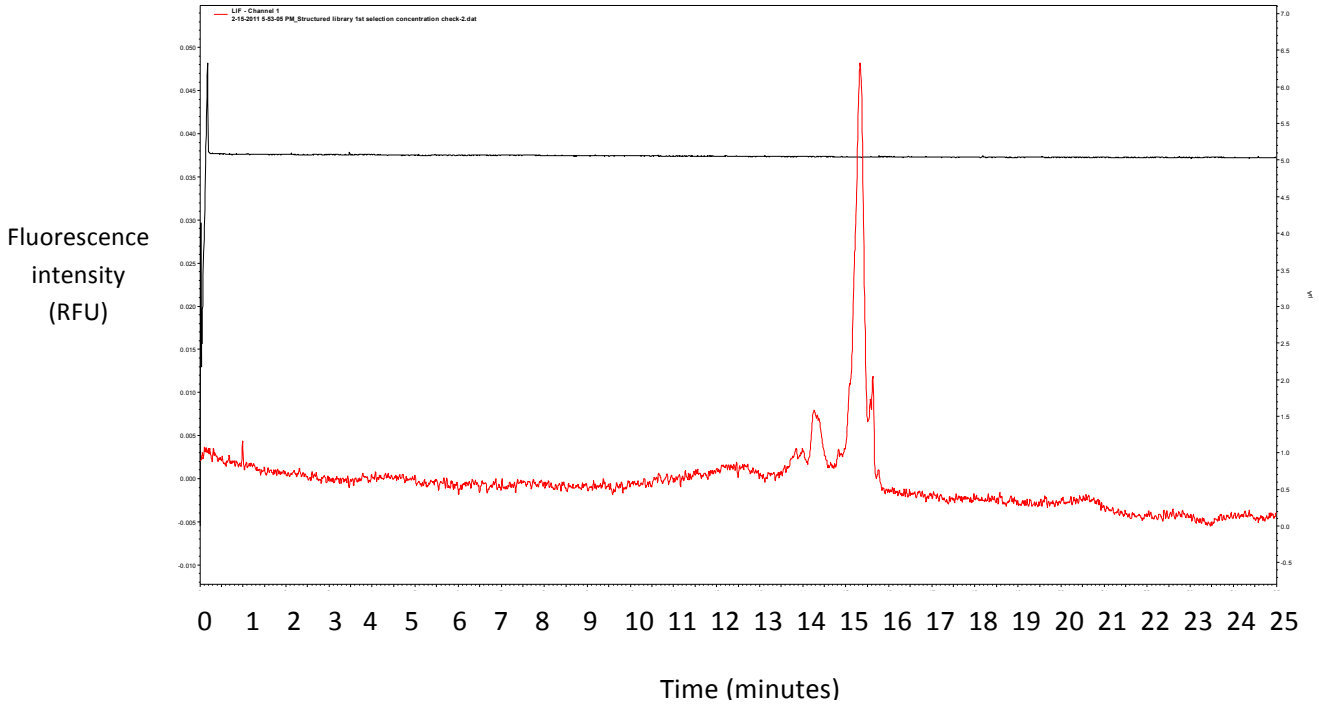


Figure23. Concentration of first pool of rHSA aptamers generated with structured DNA library using SELEX-NECEEM. The first pool of aptamers after amplification and purification was injected into the capillary, and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer. The fluorescence intensity was 0.05RFU. The detector was laser-induced fluorescence (LIF).

5th selection round with rHSA and N40 DNA library

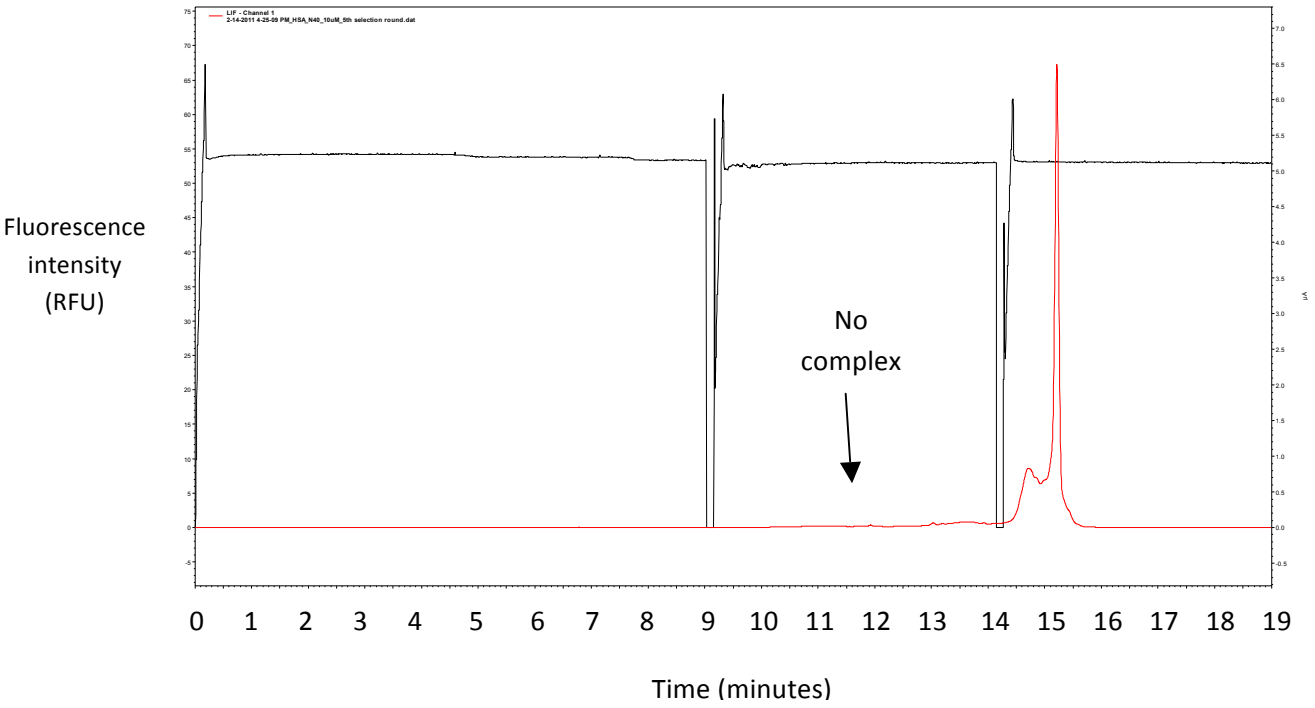


Figure 24. Fifth aptamer selection round for rHSA using N40 DNA library. The equilibrium mixture of rHSA (10uM) and 4th round of aptamer pool was injected into the capillary, and electrophoresis was carried out in an 80cm long capillary at 300 V/cm electric field with 25mM tris-acetate at pH 8 as running buffer. In the collection window, no complex of rHSA and DNA library can be seen. The detector is laser-induced fluorescence (LIF)

Appendix B

Mastermix ingredients for N40 DNA library amplification

Master Mix Symmetric					Asymmetric
Reagents	**Stock concentration**	**1X concentration**	**1.1X**	**Volume of stock (uL)**	
5X green buffer	5X	1X	1.1X	22	-
MgCl2	25mM	2.5mM	2.75mM	11	-
dNTPs Mix	10mM	0.2mM	0.22mM	2.2	-
Forward primer (Alexa 488)	10uM	0.3uM	0.33uM	3.3	11uL (1X-1uM)
Reverse primer (biotin)	10uM	0.3uM	0.33uM	3.3	0.55uL (1X-50nM)
Taq polymerase	5U/mL	0.025U/uL	0.0275U/uL	0.55	-
ddH2O				57.65	52.7
Total volume				**100uL**	**100uL**

Figure 25. Mastermix recipe for the amplification of N40 DNA library. 2uL of DNA product is added to 18uL of mastermix (total volume 20uL). Symmetric mastermix has equal amounts of forward and reverse primer, whereas asymmetric mastermix has 20X more forward primer than reverse primer. The mastermix is made under the fume-hood, previously cleaned with RNAse Away.