# HIGH-THROUGHPUT APTAMER DISCOVERY AND APTAMER INTEGRATION INTO MICROFLUIDIC DEVICES FOR RARE CELL ANALYSIS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sarah Jessica Reinholt

January 2017

# HIGH-THROUGHPUT APTAMER DISCOVERY AND APTAMER INTEGRATION IN MICROFLUIDIC DEVICES FOR RARE CELL ANALYSIS

Sarah Jessica Reinholt, Ph. D.

Cornell University 2017

Precision medicine is the idea where diagnostics and therapeutics are catered to each individual patient to provide personalized care that is optimally effective. For this to be achieved, technologies must exist that extensively examine samples and provide a highly detailed diagnosis for each patient, and processes must exist that can produce personalized drugs that specifically target the patient's illness.

Aptamers are short single stranded nucleic acids that bind to their targets with high affinity and specificity. Aptamers could make a substantial impact toward the goal of precision medicine. However, one of the main challenges preventing aptamers from reaching their potential is the efficient discovery of new high-affinity aptamers. Currently, aptamer selections are very time consuming and expensive, and often do not result in the discovery of a high-quality aptamer. The ability to reliably select aptamers with high affinity and specificity is paramount to the widespread use of aptamers. Consequently, there is great interest in improving selection technology to obtain high-quality aptamers much more rapidly.

Toward this effort, we have developed a Microplate-based Enrichment Device Used for the Selection of Aptamers (MEDUSA) that uses affinity microcolumn chromatography. Its versatile 96-well microplate-based design allows this device to be compatible with downstream plate-based processing in aptamer selections, and it lends itself to automation using existing microplate-based

liquid-handling systems. MEDUSA is also reconfigurable and is able to operate in serial and/or parallel mode with up to 96 microcolumns. We have demonstrated its use in high-throughput aptamer selections, characterization and optimization of the aptamer selection process, and characterization of previously selected aptamers. More specifically, MEDUSA was used to perform 96 simultaneous tests that determined the optimal target loading on resin to maximize aptamer enrichment for three target proteins, GFP, HSF, and NELF-E. These tests also verified the specificity of aptamers to these three proteins, as well as the non-specific binding of two suspected background binding aptamers. MEDUSA was also used to performed novel RNA aptamer selections to 19 different targets simultaneously. For these selections, a new, more efficient selection strategy was tested that greatly reduced the selection time and reagent consumption. Through the use of MEDUSA, aptamer selections can be optimized and performed in a high-throughput manner, and the success rate of novel aptamer discovery can be drastically improved.

In addition to the improvement of novel aptamer discovery, developing valuable applications that use aptamers is of equal importance. An area of study in which aptamers could be of great benefit is cancer. Cancer cells are extremely diverse and contain genetic mutations that allow them to escape the regulatory processes necessary for the healthy function of tissues and organs. Moreover, there are numerous mechanisms for malignancy each with different combinations of genetic mutations, and cancer cells are constantly evolving, which makes cancer treatment difficult with varying levels of efficacy. Aptamers can be selected that bind specifically to cancer cells, and this can be accomplished without any knowledge about the cancer cell surface composition. We have developed a diagnostic device that uses cancer cell-specific aptamers to capture and filter out cancer cells from complex samples, such as blood. Within this same device, the captured cancer

cells are lysed, and their genomic DNA (gDNA) is isolated using a micropillar array. The long strands of gDNA are physically entangled within the array, and remain in the microchannel even under flow. With this device, we have demonstrated the successful capture of human cervical and ovarian cancer cells. In addition, we have developed a custom isothermal amplification technique within the microchannel that amplifies a specific gene of interest for subsequent sequencing and analysis to determine the presence of any genetic mutations. Following the capture of cervical and ovarian cancer cells, we successfully amplified the *TP53* gene and sequenced a fragment from this gene. By comparing the sequencing results to the known human *TP53* gene sequence, we successfully detected a point mutation in the ovarian cancer cells, whereas the cervical cancer cells contained the wildtype version of this gene fragment. This device can be used to amplify multiple genes consecutively, since the gDNA is retained, so many genes of interest in cancer can be tested in the same small population of cancer cells. Using our device to test patients' cancer cells, a large amount of information can be provided to clinicians about each specific patient that will help them to prescribe the most effective treatment strategy.

From the research presented here, aptamers could have a great impact in many areas, and technologies like MEDUSA would help move the field forward by enabling novel aptamers to be successfully discovered rapidly and efficiently. Moreover, aptamers that bind specifically to cancer cells in diagnostics like the one presented here would enable highly informed decisions to be made by clinicians about treatment options. Furthermore, aptamers could also be used in targeted drug delivery and therapies. Therefore, technologies such as these are examples of key steps toward truly personalized and precision medicine.

## BIOGRAPHICAL SKETCH

Sarah Reinholt is originally from Sarnia, Ontario, Canada. She graduated with a Bachelor of Applied Science in Nanotechnology Engineering from the University of Waterloo in Waterloo, Ontario, Canada in 2011. In Fall 2011, she joined the Baeumner Lab in the Department of Biological and Environmental Engineering in the College of Agriculture and Life Sciences at Cornell University. She completed her Master of Science in December 2013, and her research involved the development of new techniques and materials to use in biosensors for point-of-care applications. She joined the Craighead Lab in the School of Applied and Engineering Physics in June 2013 to pursue her Ph.D. In her spare time, Sarah enjoys playing ice hockey and softball as well as other athletics and outdoor activities.

To Tim, Heather, Tom, Clark and Chase

# ACKNOWLEDGMENTS

I would like to express my most sincere gratitude to the Chair of my Special Committee, Dr. Harold Craighead, for his incredible support throughout my degree. I could not have completed this work without his terrific guidance. I would also like to sincerely thank my Special Committee Member, Dr. Brian Kirby, for the tremendous support and guidance he has offered me. I am extremely grateful for his time and invaluable advice. I wish to thank my Special Committee Member, Dr. Holger Sondermann, for his valuable academic and research guidance.

Additionally, I would like to express my great appreciation to Dr. John Lis for his excellent research guidance. I would also like to thank my laboratory colleagues, Kylan Szeto, Jaime Benitez, Harvey Tian, Abdullah Ozer, Jacob Tome, Fabiana Duarte, John Pagano, and Li Yao, for all of their help.

Finally, I would like to acknowledge the support provided by the US National Institutes of Health under Grants R01 GM090320 and DA030329. My work was also performed in part at the Cornell Nanoscale Facility, a member of the National Nanotechnology Coordinated Infrastructure network, which is supported by the National Science Foundation (Grant ECCS-15420819).

Six proteins included in the novel aptamer selection using MEDUSA in Chapter 3 were generously donated by the Marazzi Lab at Mount Sinai Graduate School for Biomedical Sciences (RTF1), the Tong Lab at Columbia University (DXO and Ag Rai1), the Freeman Lab at the University of Illinois (p23), and the Teitell Lab at the University of California, Los Angeles (wildtype and mutant PNPase).

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH ..... vi
ACKNOWLEDGMENTS ..... viii
TABLE OF CONTENTS ..... ix
LIST OF FIGURES ..... xii
LIST OF TABLES ..... xvii
LIST OF ABBREVIATIONS ..... xviii
CHAPTER 1 INTRODUCTION ..... 1
1 Aptamers and Aptamer Selection Technologies ..... 1
*1.1 Aptamer Selection Technologies* ..... 4
1.1.1 Nitrocellulose Filter Binding ..... 4
1.1.2 Affinity Chromatography ..... 5
1.1.3 Electrophoretic Approaches ..... 6
1.1.4 Magnetic Bead-based Approaches ..... 8
1.1.5 Microarrays, AFM, and SPR ..... 9
1.1.6 Fluorescence-activated Cell Sorting ..... 11
1.1.7 Microfluidic Approaches ..... 13
2 Cell Capture in Microfluidic Devices using Aptamers ..... 17
*2.1 Cell-SELEX* ..... 17
*2.2 Microfluidic Aptamer-based Cancer Cell Capture* ..... 18
3 Dissertation Overview ..... 20
4 References ..... 24
CHAPTER 2 HIGH-THROUGHPUT BINDING CHARACTERIZATION OF RNA APTAMER SELECTIONS USING A MICROPLATE-BASED MULTIPLEX MICROCOLUMN DEVICE ..... 35
1 Introduction ..... 35
2 Design and Fabrication of MEDUSA ..... 37
3 Characterization of selection parameters and previously-selected aptamers using MEDUSA ..... 41
*3.1 Recombinant protein target, RNA library, and aptamer preparation* ..... 42
*3.2 Performing RNA selections and characterizations* ..... 44
*3.3 Parallel selections reveal critical target concentration for maximal aptamer enrichments* ..... 46
*3.4 Multiplex serial selections show specificity of target aptamers and background binding sequences* ..... 49

4 Sample purification and concentration using microcolumns ..... 53
*4.1 Microcolumn Fabrication* ..... 53
*4.2 GFP RNA Aptamer Preparation* ..... 53
*4.3 Purification and Concentration of GFP using Aptamers in Microcolumns* ..... 54
5 Conclusions ..... 56
6 References ..... 58
CHAPTER 3 HIGHLY MULTIPLEXED RNA APTAMER SELECTION USING A MICROPLATE-BASED MICROCOLUMN DEVICE ..... 61
1 Introduction ..... 61
2 Materials and Methods ..... 64
*2.1 Target Protein Preparation* ..... 64
*2.2 Target Protein Immobilization on Affinity Resins* ..... 65
*2.3 Library and Primers* ..... 66
*2.4 MEDUSA Fabrication* ..... 66
*2.5 RNA Aptamer Selection Process* ..... 67
*2.6 High-Throughput Sequencing of Pools* ..... 69
*2.7 Motif Search using MEME* ..... 70
*2.8 Binding Affinity Determination via EMSA* ..... 70
*2.9 Analysis of DXO Aptamers Resisting Exoribonuclease Activity and Acting as Inhibitors.* ..... 71
3 Results and Discussion ..... 72
*3.1 MEDUSA as a Platform for Highly Multiplexed Aptamer Selections* ..... 72
*3.2 Incorporation of Multiple Non-Amplification Cycles to Improve Selection Efficiency* ..... 73
*3.3 qPCR and High-throughput Sequencing Results* ..... 74
*3.4 Aptamer Candidate Testing* ..... 77
*3.5 Studying Protein Function through Aptamer Selection* ..... 78
4 Conclusions ..... 83
5 References ..... 84
CHAPTER 4 MICROFLUIDIC DEVICE FOR APTAMER-BASED CANCER CELL CAPTURE AND GENETIC MUTATION DETECTION ..... 86
1 Introduction ..... 86
2 Materials and Methods ..... 88
*2.1 Cell Culture and Buffers* ..... 88
*2.2 Device Fabrication* ..... 89

*2.3 Surface Chemistry* ..... 90
*2.4 Cell Capture, Lysis, and gDNA Isolation* ..... 90
*2.5 Primers and Specific Gene Amplification from gDNA* ..... 91
*2.6 Gene Sequencing* ..... 92
3 Results and Discussion ..... 93
*3.1 Microchannel Design* ..... 93
*3.2 HeLa and CAOV-3 Cell Capture and Genomic DNA Isolation* ..... 95
*3.3 Specific Isothermal Amplification of the TP53 Gene from gDNA* ..... 97
*3.4 Sequencing of the TP53 Gene Fragment* ..... 100
4 Conclusions ..... 101
5 References ..... 103
CHAPTER 5 CONCLUSIONS AND FUTURE DIRECTIONS ..... 106
1 References ..... 111
APPENDIX A SUPPLEMENTARY FIGURES AND TABLES FOR CHAPTER 3 ..... 113

LIST OF FIGURES

**Figure 1.1:** A single-stranded nucleic acid molecule folds into a specific structure determined by its sequence, and this unique structure allows it to bind tightly and specifically to its target. .....2

**Figure 1.2:** The SELEX method. The process begins incubation of the target molecules with the random nucleic acid library in a binding step. A washing step then removes unbound nucleic acid sequences, and the bound sequences are eluted and amplified to produce an enriched aptamer pool. This process is repeated until the final enriched pool is dominated by the strongest binding aptamers. .....3

**Figure 2.1:** Image of MEDUSA when it is directly coupled to a 96-well microplate .....38

**Figure 2.2:** Schematic of the layers of MEDUSA in the order of assembly. (a) An exploded view of the customized device layers for configuring all 96 microcolumns to run in parallel. The flow path is shown in the lower boxed inset with no connections between microcolumns. The layers numbered 1 to 3 are the PMMA layers: the middle layer (1) containing the microcolumns, the next outer two layers being the caps (2) and washers (3). The outermost layers (4) consist of inlet and outlet ports that are bonded to the caps. The two layers numbered (5) are silicone layers, which are bonded to the microcolumn layer (1) to hold porous frits against both sides of the microcolumns to retain affinity resin and to make liquid-tight seals across the entire device. A photograph of MEDUSA assembled in parallel is shown in the upper inset. (b) The customized device layers for configuring 24 of the microcolumns to run in series. The two additional silicone layers (6) shown in blue, as well as the smaller complementary plastic layers (2 and 3), are specifically programmed to connect 3 sets of 8 microcolumns within the device. The flow path is shown in the lower boxed inset with microcolumns connected in series via a serpentine route through 8 microcolumns. MEDUSA assembled to run in series and parallel is shown in the upper inset. .....39

**Figure 2.3** Layout for the 96 samples in MEDUSA according to its analogous microplate position given by the rows A-H, and the columns 1-12. In section I, the 8 indicated targets were connected in series from A to H to test the specificity and partitioning efficiency of various RNA aptamers. This was tested in triplicate in columns 1 to 3. Sections II, III, and IV tested the effects of target surface concentration on aptamer enrichments. The colored triangles indicate decreasing concentrations of each protein from 10 µg/µL (row A) to 0.016 µg/µL (row H) in 2.5-fold dilutions. Section II (green triangle) aimed to confirm previous enrichment behaviors shown with GFP. Sections III and IV tested the same concentrations for the proteins hHSF1 (blue triangle) and NELF-E (red triangle) to assess the prevalence of target surface concentration effects on binding due to steric hindrance or other effects in aptamer selections. .....42

**Figure 2.4:** Recoveries and enrichments of specific RNA aptamers over the N70 library as a function of protein concentration. (a) The recovery of GFPapt and N70 library at various concentrations of GFP. Analogous data for the recovery of (b) HSFapt and N70 library from hHSF1, and (c) NELFapt and N70 library from NELF-E. (d–f) The calculated enrichments of the specific aptamers (GFPapt, HSFapt, NELFapt) over the random library. The error bars represent the standard deviation in recoveries or enrichments calculated for each condition performed in triplicate. .....47

**Figure 2.5:** The enrichment of RNA aptamers over the N70 library on various targets connected in series. The enrichment of each protein-specific aptamer, GFPapt (a), HSFapt (b), and NELFapt (c), and non-specific sequences, BBS1 (d), and BBS2 (e), on all eight microcolumns. The error bars represent the standard deviation in enrichments calculated for each target performed in triplicate. .....51

**Figure 2.6:** Schematic of aptamer-immobilized resin contained within a microcolumn for capturing and concentrating a target of interest in a sample.....55

**Figure 2.7:** Fluorescence analysis of aptamer-based GFP capture and concentration. (a) Fluorescent images showing of the microcolumn before elution containing GFP (upper photo), and after elution where the GFP signal has diminished (lower photo). (b) Measurement of the total GFP recovered by NanoDrop. (c) Measurement of the concentration change between the initial GFP solution pumped through the microcolumn and the recovered eluent by NanoDrop. The error bars indicate the standard deviation between triplicate NanoDrop measurements.....56

**Figure 3.1:** Schematics of the selection device and selection process. (a) The schematics describe the fluid path through the MEDUSA device in serial (left) and parallel (right) configurations, while the lower panels show images of the device in each configuration. (b) The list of targets in the order in which they saw the initial RNA library in a serial configuration. All the proteins in Lane 1 have an N-terminal MBP-tag, and in Lane 2 an N-terminal 6x His-tag, except RTF1, which has an N-terminal FLAG-tag for immobilization on amylose, Ni-NTA, and anti-FLAG resin, respectively. (c) The aptamer selection process beginning with a random RNA library, proceeding with 4 rounds of SELEX incorporating increasing numbers of non-amplification cycles prior to an amplification cycle, and ending with high-throughput sequencing of the enriched RNA pools. Ten cycles of selection were performed, starting at Round #1-Cycle #1 (R1C1) and ending at Round #4-Cycle #4 (R4C4).....73

**Figure 3.2:** qPCR results indicating the percent recovery of RNA after each binding cycle in Round 4 relative to the initial RNA concentration at the beginning of the round. The error bars represent the standard deviation for triplicate measurements.....75

**Figure 3.3:** Binding analysis of candidate aptamer sequences. (a) Measurement of dissociation constants for aptamers to 5 different protein targets by EMSA experiments. The data were fitted to the Hill Equation to determine the $K_D$. (b) A list of the dissociation constants for all of the aptamer candidates tested and their multiplicity ranking in their respective final (R4C4) pools. A complete list of the top 3,000 aptamer clusters for final pools of each target is provided in Supplementary Dataset 1. .....77

**Figure 3.4:** Characterization of DXO aptamers. (a) A highly specific DXO motif identified by MEME among the top 3,000 most abundant DXO aptamer clusters from R4C4. The E-value, representing the statistical significance of the identified motif as calculated by the MEME suite (http://meme-suite.org/), and number of aptamer clusters containing a match to Motif 1 are indicated. E-values lower than 0.05 are considered significant. (b) Secondary structure of DXO #2 aptamer, a representative aptamer of the second highest multiplicity cluster for DXO from the R4C4 pool, as predicted by the mFold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form). DXO Motif 1 is colored cyan. Predicted secondary structures of other DXO

aptamers are provided in Supplementary Fig. S3. (c) Stability of RNA aptamers against DXO exoribonuclease activity. Radiolabeled specific RNA aptamers or the N70 RNA library were incubated with DXO enzyme at 37°C for the indicated times and separated using 8% denaturing PAGE. Full-length RNA aptamers are indicated by ▻. (d) Inhibition of DXO exoribonuclease activity by RNA aptamers. Degradation of a 3'-Cy5-labeled 30-nt RNA substrate by DXO in the absence and presence of indicated amounts of RNA aptamers as determined by 10% denaturing PAGE analysis. Full-length, intact RNA substrate and degradation product are indicated by ▻ and ▻, respectively. The band indicated by * above the full-length RNA is likely incompletely deprotected RNA substrate (2'-ACE protected, Dharmacon). .....82

**Figure 4.1:** Microchannel design for capturing cancer cells and isolating their gDNA. The device contains two orthogonal channels, the cell channel and the DNA channel. It also contains two micropillar arrays, the cell capture array located at the intersection of the channels (red box, scale bar: 200 µm) and the DNA isolation array located downstream of the cell capture array in the DNA channel (blue box, scale bar: 30 µm). .....94

**Figure 4.2:** Cancer cell capture in microfluidic channels using aptamers. (a) Brightfield and fluorescent images of the cell capture region without aptamers after flowing HeLa cells through the microchannels. (b) Brightfield and fluorescent images of the cell capture region containing aptamers after flowing HeLa cells through the microchannels. The blue circles highlight the captured cells. (c) Brightfield and fluorescent images of the cell capture region without aptamers after flowing CAOV-3 cells through the microchannels. (d) Brightfield and fluorescent images of the cell capture region containing aptamers after flowing CAOV-3 cells through the microchannels. The HeLa cells fluoresce due to GFP bound to their histones, and the CAOV-3 cells were stained with calcein-AM. .....95

**Figure 4.3:** Cancer cell lysis and isolation of gDNA via physical entanglement within the micropillar array. (a) Brightfield and fluorescent images of HeLa cells bound to aptamers in the cell capture region. (b) Fluorescent image of gDNA from HeLa cells isolated by the micropillar array and stained with YOYO-1 dye. (c) Brightfield and fluorescent images of CAOV-3 cells bound to aptamers in the cell capture region. (b) Fluorescent image of gDNA from CAOV-3 cells isolated by the micropillar array and stained with YOYO-1 dye. .....96

**Figure 4.4:** Gel images of benchtop demonstration of specific MDA of the *TP53* gene and smaller gene fragment from PCR. (a) MDA product from reactions using 15 ng of purified gDNA. Different quantities of primers were tested, as well as negative controls containing no gDNA template. The MDA product was approximately 10 kb. (b) PCR product from reactions amplifying a shorter *TP53* gene fragment using 60% of the MDA product. The PCR product is 130 nt in length.....98

**Figure 4.5:** Gel images after running MDA and PCR products. (a) On-chip MDA product from isolated HeLa and CAOV-3 cell gDNA verified using a 1% agarose gel. The ladder indicates that the product is around 10 kb in length. (b) PCR product verified via 8% PAGE. The ladder confirmed a product length of 130 bp. .....99

**Figure 4.6:** Sanger sequencing results from HeLa and CAOV-3 cell *TP53* gene fragments. The wildtype sequence for this fragment is shown at the top of the figure and was obtained from the

International Agency for Research on Cancer. Plots from sequencing runs for both HeLa and CAOV-3 *TP53* gene fragments are shown with the corresponding bases indicated. The highlighted bases indicate the location of the known point mutation in CAOV-3 cells, and the specific base that is mutated is underlined. There are some bases that were not called, but the sequence can be discerned by looking at the peaks. .....100

**Figure A.1:** Analysis of TIP60-Chromo enriched aptamer clusters through selection rounds and cycles. (a) Median multiplicity of top 1,000 TIP60-Chromo aptamer clusters in a given round-cycle and their median multiplicity in subsequent round-cycles of selection. (b) Number of common aptamer clusters in top 1,000 clusters between different round-cycles of selection for TIP60-Chromo. Common aptamer clusters from R2C1 to R3C2 with the later round-cycles are mostly due to sequence artifacts that pass through our sequence filter criteria. (c) Correlation plots for rankings of common aptamers in top 1,000 clusters at different selection cycles for TIP60-Chromo. $R^2$ values for each correlation are indicated. Multiplicities (reads per million) for 1$^{st}$, 20$^{th}$, and 1000$^{th}$ cluster in each library are as follows: R2C1 – 130.4, 1.3, 0.5; R2C2 – 184.5, 1.7, 0.6; R3C2 – 145.1, 2.5, 0.6; R3C3 – 697.4, 2.4, 1.2; R4C2 – 3,164.8, 610.4, 14.8; R4C3 – 12,038.3, 1,958.4, 43.6; and R4C4 – 7,923.3, 1,520.1, 35.8. .....119

**Figure A.2:** Binding analysis of candidate aptamer sequences via EMSA. Results from EMSA experiments were quantified using ImageJ, and fitted to the Hill Equation using Igor Pro 5.04A.....120

**Figure A.3:** Specificity of DXO aptamers. Th**e s**pecificity of DXO aptamers was tested by nitrocellulose filter binding assay. Radiolabeled RNA aptamers, DXO #4 and #5, and the N70 library (negative control) were incubated with varying concentrations (3000-4 nM) of DXO, TIP60-Chromo, GCN5-GNAT, MOF-Chromo, or Trx-ZnFinger proteins. Protein bound aptamers were captured by a nitrocellulose filter and detected by phosphorimaging. DXO #4 and #5 aptamers showed very weak binding, indistinguishable from N70 library control, against all other target proteins tested. .....121

**Figure A.4:** Predicted secondary structures of DXO aptamers. MEME identified motif (DXO motif 1), which base-pairs with the BamHI site in the forward constant region of the N70 RNA library, is colored cyan. The predicted secondary structures and the deltaG values were obtained from mFold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form). .....122

**Figure A.5:** Stability of DXO aptamers against DXO exoribonuclease activity. (a) Each radiolabeled RNA aptamer was incubated with 1 μM DXO protein and samples collected at indicated time points were analyzed by running on an 8% denaturing PAGE gel. The gel image was analyzed, and a DXO exoribonuclease assay time course was plotted. (b) Aptamer half-lifes ($T_{1/2}$ min) are estimated from exponential decay function fit to measured % full-length aptamer remaining. Images were analyzed by ImageJ and data analysis was done in Microsoft Excel...122

**Figure A.6:** Mutant and truncated DXO aptamers. (a) Predicted secondary structure of mutant and truncated DXO aptamers. MEME identified motif (DXO motif 1) is colored cyan and mutations to motif and the corresponding mutations in the forward constant region are colored red. In mut1, the original motif, GGATCCC, was changed to CCTAGGC and the BamHI site, GGATCC, in the forward constant region was changed to CCTAGG. In mut2, the original motif, GGATCCC, was

changed to AATATTC and the BamHI site, GGATCC, in the forward constant region was changed to AATATT. Truncations of DXO #2 and #5 were made to retain the DXO motif 1 (cyan) and the first stem-loop structure for each aptamer. A sequence-shuffled version of DXO #5 trunc was used as a control. Secondary structures were predicted and the deltaG values were calculated by mFold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form). (b) Stability (deltaG values) of complete predicted structure and the first stem of original full-length, mutant, and truncated versions of DXO aptamers. (c) Stability of original, mutant and truncated DXO aptamers against DXO exoribonuclease activity.....123

**Figure A.7:** Inhibition of DXO exoribonuclease activity by RNA Aptamers. (a) Quantification of DXO exoribonuclease activity against 3'-Cy5 labeled 30-nt RNA substrate in the presence of indicated concentrations of N70 aptamer pool, or DXO #2, #4, or #5 RNA aptamers from gels images shown in Fig. 4d. Data was fitted to a 4-parameter dose-response curve using GraghPad Prism6 software, and the estimated IC50 values are listed (b). Errors represent the standard error. (c) Verification of equal loading of RNA aptamers, which remained largely intact even after two hours of incubation with DXO enzyme, in DXO exoribonuclease inhibition assays.....124

**Figure A.8:** Inhibition of DXO exoribonuclease activity by full-length or truncated DXO RNA aptamers. A) 3'-Cy5 labeled 30 nt RNA substrate was incubated with 1uM DXO exoribonuclease either in the absence of any aptamer (No Aptamer) or in the presence of 0.8 µM full-length or truncated DXO RNA aptamers. After 2 hours of incubation at 37°C, exoribonuclease reactions were separated by 10% denaturing PAGE, and the remaining intact RNA substrate, indicated by →, was visualized by Cy5 fluorescence scan. B) Percentage of full-length DXO substrate remaining under each condition as quantified by ImageJ.....125

LIST OF TABLES

**Table 2.1:** Properties of the target proteins .....41

**Table 2.2:** Aptamer nucleotide sequences.....44

**Table 2.3:** Primer sequences used for qPCR analysis.....44

**Table 2.4:** Frequencies of BBS1 and BBS2 in previous selections. Summary of the number of times BBS1 or BBS2 has been identified in all previous selections. The numbers indicate the instances in which BBS1 was more highly enriched than BBS2 (or vice versa) on each target, grouped according to the resin on which each target was immobilized. .....50

**Table 2.5:** Aptamer and primer nucleotide sequences. .....54

**Table 3.1:** Selection Parameters.....67

**Table 4.1:** Amplification Primer Sequences .....92

**Table A.1:** Summary of sequencing results and clustering.....113

**Table A.2:** Target Protein Information .....116

LIST OF ABBREVIATIONS

| | |
|---|---|
| 6xHis | Hexahistidine |
| AFM | Atomic Force Microscope |
| ASExp | Aptamer Selection Express |
| ATCC | American Type Cell Culture |
| BBS | Background Binding Sequence |
| BSA | Bovine Serum Albumin |
| cDNA | Complimentary DNA |
| CE-SELEX | Capillary Electrophoresis SELEX |
| CLADE | Closed Loop Aptameric Directed Evolution |
| CMACS | Continuous-flow Magnetic Activated Chip-based Separation |
| CTC | Circulating Tumor Cell |
| DEPC | Diethylpyrocarbonate |
| DNA | Deoxyribonucleic Acid |
| EDTA | Ethylenediaminetetraacetic Acid |
| EGFR | Epidermal Growth Factor Receptor |
| EMSA | Electrophoretic Mobility Shift Assay |
| EtBr | Ethidium Bromide |
| FACS | Fluorescence-Activated Cell Sorting |
| FISH | Fluorescence In-Situ Hybridization |
| FITC | Fluorescein Isothiocyanate |
| FOTS | (1H,1H,2H,2H-Perfluorooctyl)Trichlorosilane |
| gDNA | Genomic DNA |

| | |
|---|---|
| GFP | Green Fluorescent Protein |
| GSH | Glutathione |
| GST | Glutathione S-Transferase |
| HEPES | N-2-hydroxyethylpiperazine-N′-ethanesulfonic acid |
| hHSF1 | Human Heat Shock Factor 1 |
| HiTS-FLIP | High-Throughput Sequencing Fluorescent Ligand Interaction Profiling |
| HiTS-RAP | High-Throughput Sequencing RNA Affinity Profiling |
| I-SELEX | Inertial SELEX |
| kb | Kilobase |
| $K_D$ | Dissociation Constant |
| MARAS | Magnetic-Assisted Rapid Aptamer Selection |
| MBP | Maltose-Binding Protein |
| MDA | Multiple Displacement Amplification |
| MEDUSA | Microplate-based Enrichment Device Used for the Selection of Aptamers |
| MMLV-RT | Moloney Murine Leukemia Virus Reverse Transcriptase |
| MMS | MicroMagnet Separation |
| NELF-E | Negative ELongation Factor E |
| Ni-NTA | Nickel Nitrilotriacetic Acid |
| nt | Nucleotide |
| PAGE | Polyacrylamide Gel Electrophoresis |
| PBS | Phosphate-Buffered Saline |
| PCR | Polymerase Chain Reaction |
| PDMS | Polydimethylsiloxane |

| | |
|---|---|
| PectI | Polymer-enhanced capillary transient Isotachophoresis |
| PET | Polyethylene Terephthalate |
| PMMA | Poly(Methyl Methacrylate) |
| PSMA | Prostate-Specific Membrane Antigen |
| PTK7 | Protein Tyrosine Kinase 7 |
| QPASS | Quantitative Parallel Aptamer Selection System |
| qPCR | Quantitative PCR |
| QSAS | Quantitative Selection of Aptamer through Sequencing |
| RAPID | RNA APtamer Isolation via Dual-cycles |
| RCA | Rolling Circle Amplification |
| RNA | Ribonucleic Acid |
| RNA-MaP | RNA on a Massively Parallel array |
| SDS | Sodium Dodecyl Sulfate |
| SELEX | Systematic Evolution of Ligands by EXponential Enrichment |
| SPR | Surface Plasmon Resonance |
| ssDNA | Single-Stranded DNA |
| tRNA | Transfer RNA |
| µFFE | Micro Free Flow Electrophoresis |
| UTP | Uridine Triphosphate |
| VDC | Volume Dilution Challenge |

# CHAPTER 1

# INTRODUCTION

## 1 Aptamers and Aptamer Selection Technologies

Antibodies are a class of proteins that are specific and bind with high affinity to their target antigens that are foreign and potentially harmful to the body. These affinity ligands have become essential in the fields of research and medicine for a wide diversity of applications. However, despite their widespread use, they have several limitations.[1-3] Antibodies require a living host to produce them *in vivo*, which takes months and leads to variability between batches from different hosts. This also means antibodies cannot be acquired to targets that do not trigger an immune response or are toxic to living organisms. They are also proteins, which are highly sensitive to their surroundings, so they must be kept in an environment (temperature, pH, salt conditions, etc.) close to physiological conditions during storage and usage. However, there is another class of affinity ligands emerging called aptamers that can overcome many of the limitations of antibodies.

Aptamers are short, single-stranded nucleic acid molecules that bind tightly and specifically to their target, as illustrated in Figure 1.1. Nucleic acids (RNA or DNA) provide several advantages over their protein counterparts.[3] They are much less sensitive to their surroundings, and can be reversibly denatured and renatured, as well as lyophilized, which allows many more options for storage and usage. Researchers can even use modified nucleic acid bases to increase the durability and resistance of the aptamers to different conditions.[4-11] More importantly, aptamers are discovered and synthesized *in vitro*, which is highly advantageous. This eliminates the batch-to-batch variation, and enables aptamers to be selected to virtually any target from single metal ions to whole cells and tissue.[12-14] Also, once the aptamer sequence is known, it can be synthesized

reproducibly very quickly and inexpensively, compared to antibodies. Furthermore, the environment in which they will be used can be replicated during the selection to ensure the aptamer will function optimally in its ultimate application. Since the beginning of aptamer research over 25 years ago, significant work has been done in applying the use of aptamers in biosensors and diagnostics, imaging, and therapeutics, including targeted drug delivery.[15-28]

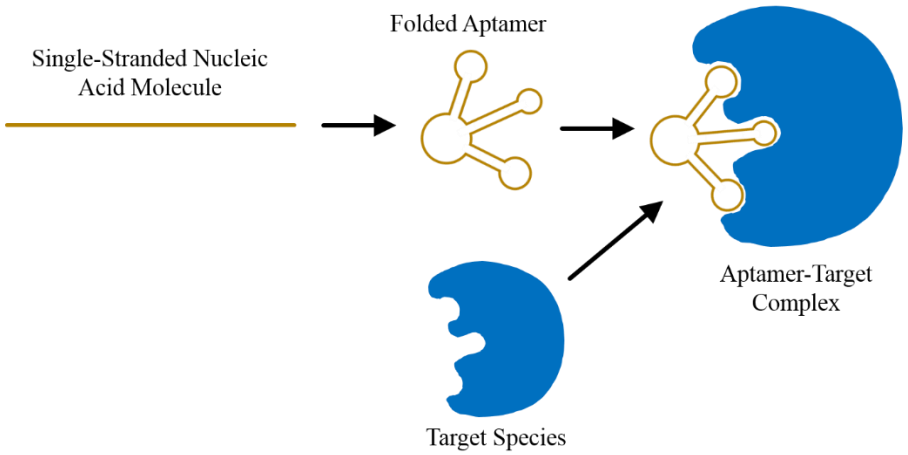



**Figure 1.1:** A single-stranded nucleic acid molecule folds into a specific structure determined by its sequence, and this unique structure allows it to bind tightly and specifically to its target.

The process by which aptamers are discovered is called SELEX, or Systematic Evolution of Ligands by EXponential enrichment.[29-31] This technique uses a large starting library of random nucleic acids ($10^{12}$-$10^{15}$ different sequences). These single-stranded nucleic acids have different primary sequences with the ability to base-pair with themselves, resulting in different three-dimensional secondary structures that provide the binding specificity to different target molecules. SELEX is an iterative process that begins with a binding step in which the starting library is incubated with the target of interest, followed by a washing step that removes unbound or weakly bound sequences. The target-bound sequences are then eluted and amplified to produce an enriched pool ready to undergo another SELEX cycle. After several cycles, the resulting final pool

is dominated by the strongest binding aptamers that are identified through sequencing of the final pool. The SELEX method is illustrated in Figure 1.2.

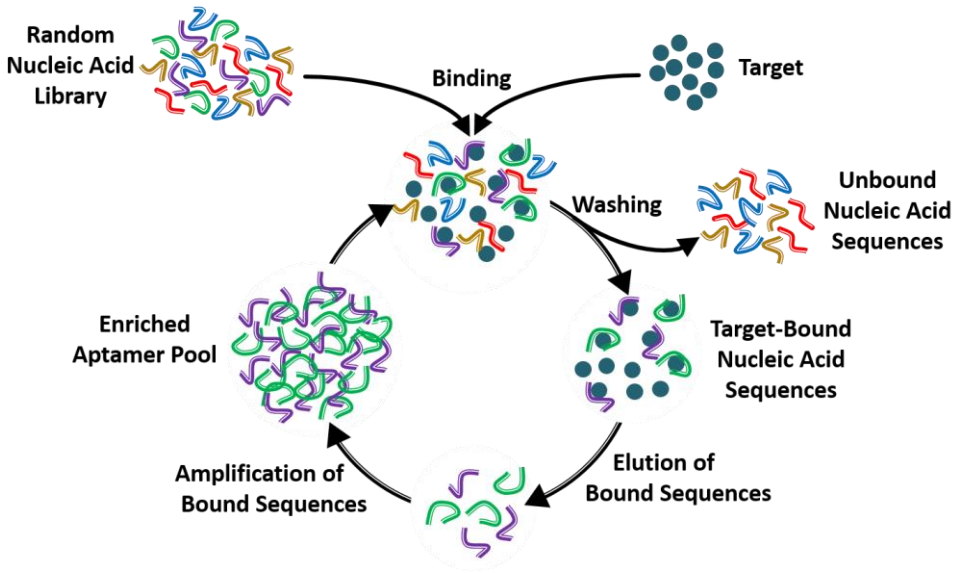



**Figure 1.2:** The SELEX method. The process begins incubation of the target molecules with the random nucleic acid library in a binding step. A washing step then removes unbound nucleic acid sequences, and the bound sequences are eluted and amplified to produce an enriched aptamer pool. This process is repeated until the final enriched pool is dominated by the strongest binding aptamers.

The potential benefit of aptamer technologies is immense, and it is the aptamer selection process that is the ultimate limitation hindering the quality of aptamers generated for these applications. Although the SELEX process is straightforward in theory, there are several challenges associated with it. Aptamer selections tend to be very time-consuming, use large quantities of reagents making them expensive,[13,14] and they suffer from a low rate of success in that aptamers identified often have low affinity and specificity, which drastically affects sensitivity and background effects of any resulting technologies.[4,14] Consequently, there has been a significant drive to develop aptamer selection technologies to improve the methods by which aptamers are discovered.

## *1.1 Aptamer Selection Technologies*

A wide variety of aptamer selection techniques have been developed since the birth of the field in an effort to make the aptamer discovery process as optimal and effective as possible. Several methods are based on well-known analytical techniques, such as nitrocellulose filter binding, chromatography, and electrophoresis, which have been adapted to perform aptamer selections. Others use existing imaging technologies, such as AFM (atomic force microscopy), SPR (surface plasmon resonance), and FACS (fluorescence-activated cell sorting). There are techniques that use magnetic beads and microarrays, as well as a diversity of microfluidic selection techniques that have miniaturized some of these benchtop methods. Each of these techniques have both advantages and disadvantages, and the best technique to use depends on the target of interest and its ultimate application.

### 1.1.1 Nitrocellulose Filter Binding

Nitrocellulose filter binding is a simple method used in the first SELEX experiments.[29-31] This technique involves incubating the target and nucleic acid library, and allowing the solution to approach equilibrium. This solution is then filtered through a nitrocellulose filter, which binds the target protein and any nucleic acids that are bound to the target, while free nucleic acids pass through the filter.[32,33] Although this technique is straightforward and accessible to most labs, there are limitations. Nitrocellulose non-specifically binds proteins, so this method only works for protein targets, and the capture efficiency varies from protein to protein. Also, nucleic acids are known to non-specifically bind to nitrocellulose, especially those with multi-G motifs,[34] which could result in background nitrocellulose-binding aptamers dominating the final pool.

### 1.1.2 Affinity Chromatography

An affinity chromatography approach to aptamer selections is attractive because of its simplicity and accessibility, and it can be performed for virtually any type of target.[31,35-39] This technique uses affinity resins packed within a column on which the target molecule is immobilized. The random library is injected into the column, and the sequences with affinity to the target bind and are retained within the column, while unbound sequences flow out of the column as waste. The target-bound sequences are then eluted from the column. Although in theory affinity chromatography can be used for any type of target, the target does require an affinity tag or some other means of immobilizing it onto a resin. Immobilizing the target on a resin can also interfere with aptamer binding due to steric hindrance, and it can also change the conformation of the target such that the selected aptamers may not behave in the same manner when the target is in solution.[13] Other significant drawbacks of this technique are that it requires a large quantity of target, and non-specific binding of the library can occur to the resin itself.[35] A feature of this approach is that, aside from affinity, aptamers can also be selected based upon kinetic parameters, such as off-rate or on-rate, by adjusting the flowrates of the binding and washing steps.[14] An important consideration when performing an affinity chromatography-based selection is the density of target on the resin.[40,41] If the target is too sparse, this increases the surface available for non-specific binding of sequences to the resin. If the target is too dense, steric hindrance could affect aptamer binding, and sequences with artificially high affinity could be selected due to cooperative binding where the aptamer binds weakly to multiple target molecules.

This aptamer selection approach has been scaled down through the use of microcolumns, which provide several advantages over larger columns.[42] Most importantly, microcolumns allow the amount of target necessary to be reduced by several orders of magnitude, which allows this

technique to be used with targets that are difficult to purify in large quantities. They also decrease the amount of other reagents needed, such as resin and buffers. Decreasing the amount of resin drastically also reduces the surface area available for non-specific binding. Scaling down this technique also enables it to be multiplexed more easily. Optimizations have also been performed to maximize the enrichment of aptamers using the microcolumns. The density of the target on the resin, as well as the flowrates for each step of the selection, were optimized by monitoring the enrichment of a known aptamer.[42]

The affinity chromatography-based technique was miniaturized even further in a process called MonoLEX.[43] Here, a capillary is used as the affinity column, and the library flows through the capillary. Aptamers bind to the target along the capillary, and binders are separated into groups by cutting the capillary at different points along its length with the idea that higher affinity aptamers are retained closer to the beginning of the capillary. This selection process is advantageous in that the amount of reagents necessary is further reduced, and it is done in a single round, which greatly decreases the time needed.

### 1.1.3 Electrophoretic Approaches

Aptamer selections have been performed using Electrophoretic Mobility Shift Assays (EMSAs).[44,45] Here, the target and random library are incubated in solution and then separated electrophoretically on a gel. The separation is based on their electrophoretic mobility, which is affected by the size, charge and shape of the molecule. The gel is imaged to see the fluorescent or radioactive bands, and the target-bound aptamer band is cut out of the gel, crushed, and the nucleic acids diffuse back into solution. This technique can be performed using tools and equipment that are generally readily available in most laboratories, it is straightforward, and it uses relatively little

target; however, there are several limitations.[13,14] This approach can only accommodate targets that have significantly different mobilities than the nucleic acids, which means that small molecules and targets with similar charge to the library are difficult to resolve on a gel. The size of the library can be limited due to the small volumes used, and it must be labeled with fluorescence or radioactivity for visualization. The selection conditions are restricted to solutions that are compatible with electrophoresis, and the separation step tends to be quite long, so the bound complexes could potentially dissociate.

Capillary electrophoresis SELEX (CE-SELEX) has some advantages over conventional EMSAs.[46-51] There is no gel needed and everything is done in solution, which almost eliminates background binding. The process can be monitored using UV detectors, so fluorescent and radiolabeling is not necessary. Very small volumes are used in CE-SELEX, so the library size is very limited (~$10^{12}$ sequences) to prevent overloading, which can make it difficult to resolve the bound population from the free sequences. Fractions can also be collected, which can separate the sequences by affinity.[52]

A variation of CE-SELEX, called Non-SELEX, has been successfully used to generate aptamers.[53] This method uses capillary electrophoresis for separating bound from unbound sequences, but does not amplify the collected pool between binding cycles; it is simply re-injected into the capillary. This technique eliminates any amplification bias or non-specific amplification that can occur, but only a small fraction of the collected pool can be re-injected for additional binding and separation cycles, due to the significant dilution that occurs, and this severely limits the number of sequences that can be sampled. A more optimal version of this process that uses larger capillaries and multiple pool injections has been developed and sequence sampling was improved.[54] Another variation of CE-SELEX was recently developed called polymer-enhanced

capillary transient isotachophoresis (PectI).[55] Here, PectI was used to select high-affinity aptamers to cells by generating a single narrow zone of target cell-bound DNA that is very well-separated from the free DNA. This is in contrast to conventional CE-SELEX, which typically results in multiple peaks.

### 1.1.4 Magnetic Bead-based Approaches

Magnetic beads have been used in aptamer selections to provide several advantages over resin-based affinity chromatography techniques.[56-60] In this approach, the target is immobilized on magnetic beads and incubated with the random library in solution. A magnet is then applied, which rapidly separates and concentrates the beads, and allows the solution to be aspirated, thereby removing the unbound sequences. Aggressive washing can be done in the same manner, and amplification can actually be performed in the presence of the magnetic beads, which significantly reduces potential losses.[13] Similarly to affinity chromatography, aptamers can be selected to any target that can be immobilized onto magnetic beads, which has its limitations, but unlike affinity chromatography, a very small amount of target is needed. Magnetic beads have also provided a platform for single-cycle selections through repeated vigorous washes that dissociate weakly bound sequences, leaving only higher affinity binders,[61] which helps to overcome the drawback of magnetic beads being expensive.

Some techniques for selecting high-affinity aptamers have been developed using magnetic beads. Monitoring the progress of a selection is very advantageous, and a process called FluMag-SELEX was developed that uses a fluorescently labeled library to quantify binding to the target-immobilized magnetic beads.[62] The level of binding after each cycle of selection can be measured easily, since the magnetic beads can be concentrated, and this helps to monitor the progress and

success of the selection. A method called Aptamer Selection Express (ASExp) uses a starting library of double-stranded nucleic acids that must first dissociate before binding to the target.[63] This greatly increases the binding threshold, and only the strongest binders will favorably dissociate to bind to the target. To identify the bound sequences, magnetic beads coated with long random single-stranded nucleic acids hybridize with the anti-sense strand of the bound aptamers, and are removed from the solution. The selection stringency is so high that this technique only requires one cycle. Another single-cycle selection method called Magnetic-Assisted Rapid Aptamer Selection (MARAS) further exploits the benefit of bead manipulation using a magnetic field.[64,65] With this technique, the pre-incubated solution of target-immobilized magnetic beads and library is placed within a solenoid, and an alternating current is applied, which in turn creates an alternating magnetic field. This not only isolates the bead-bound sequences from the unbound, but also removes low-affinity binders from the target beads. The strength and frequency of the magnetic field dictate the stringency of the selection and therefore the affinity of the resulting aptamers.

### 1.1.5 Microarrays, AFM, and SPR

Aptamer selections have also been performed by immobilizing the target or library onto a surface. Microarrays have been used in which the library is immobilized on a surface in an array with specific sequences at precise locations on the surface.[66-70] The microarray is then exposed to fluorescently-labeled target, washed, and imaged to reveal the locations where the fluorescent target bound and identify the sequences. Using lithography, these arrays can contain up to $10^5$ distinct spots; however, this is an extremely small library compared to other SELEX techniques, so performing pre-selection cycles using a different method may be necessary. In fact, microarrays can be used to optimize selected aptamers by creating mutants of a specific sequence.[71,72] Using

microarrays also requires specialized optical equipment for imaging and reading. A major benefit of this technique is the ability to monitor the selection using fluorescence, which enables the direct observation of interactions between the nucleic acids and target, and assessment of background binding. One group even coupled gold nanoparticles to the immobilized library, and monitored the selection by measuring the generated current and/or voltage change upon binding with the target, which could remain label-free.[73] A selection method called Closed Loop Aptameric Directed Evolution (CLADE) was developed that uses a microarray to identify fluorescent target-bound sequences and controllably mutates these binding sequences to create a new microarray for the next round of selection.[74,75] This way, better binders can potentially be identified with each cycle, and the aptamer can be optimized.

Other optical selection techniques have been used where the target is immobilized on a surface and the library is introduced in solution, which eliminates the library size restriction. A single-step selection technique was developed using this approach, whereby the target was immobilized on a cover slip and a fluorescently-labeled library was introduced.[76] Binding was confirmed via fluorescence detection after the surface was washed. Two SELEX techniques have also been developed that use AFM. A method called NanoSelection uses beads each coated in one specific sequence that is fluorescently labeled, which are incubated with a target-immobilized surface.[77] After washing, the surface is imaged using a fluorescence microscope coupled to an AFM to identify the location of the bound sequences, and the AFM tip is used to pluck the beads from the surface and transfer them to a tube where amplification can be performed. In another technique called AFM-SELEX, a streptavidin-coated AFM tip is used to hold a biotinylated nucleic acid library.[78] The library interacts with a target-immobilized surface, and the streptavidin-biotin interaction must be overcome in order for the sequence to bind to the target.

This creates a high binding threshold as the interaction between streptavidin and biotin is very strong. The selection was monitored by measuring the force exerted on the tip during each round, and an increasing force indicated increasing affinity over time. The major drawback with these AFM-based techniques is the limited library size.

SPR is another technique that has been used for aptamer selections.[79-81] SPR is extremely sensitive and can detect very small surface changes, including binding events on a surface. A detector optically measures the change in refractive index of the layer very close (within ~300nm) to the surface of a thin metal film where surface plasmons are generated through excitement by an evanescent electromagnetic wave. This evanescent wave is produced by light undergoing total internal reflection. SPR was first used in the aptamer field as a later-round selection method after several rounds of nitrocellulose filter binding,[82] and has also been used as a method for determining aptamer binding affinity[83] and monitor binding to microarrays[84]. This ability can also be used to isolate aptamers based on off-rates by fractionation of the dissociation step. SPR has advantages over other surface-based techniques in that it is label-free, there is no library size restriction, and it enables monitoring of binding rates with time, which can indicate an increase in binding affinity in later selection rounds.

### 1.1.6 Fluorescence-activated Cell Sorting

Flow cytometry is an optical method used to count and quantitatively analyze properties of cells individually at very high rates. FACS uses flow cytometry to interrogate cells and sort them based on these properties. The cells flow in a stream to be analyzed, and this stream is broken up into droplets, which can then be deflected electrostatically into separate collection tubes. This technology has been used in aptamer selections where fluorescence measurements are used to

separate aptamer-bound cells from free nucleic acids and cells, as well as to separate different cell populations, such as live cells from dead cells, which is important since the surfaces that the aptamers will be binding to are different for live and dead cells.[85,86] This selection method is not limited to cells and has been used with targets immobilized onto beads. In these selections, each bead is coated with a different single sequence produced by emulsion PCR, and binding is determined by measuring the amount of fluorescently-labeled target on the bead.[87-89] The relative affinity of each sequence is also quantified directly during the selection by the relative brightness of the beads.

The library size is generally restricted with this technique due to the limited throughput of the FACS machine, so in one method called particle display, FACS was used in later rounds of selection after the library has been enriched and reduced by traditional SELEX procedures.[90] To further demonstrate the capabilities of FACS-based SELEX, a process was also developed to measure aptamer expression and function inside cells.[91] These aptamers were RNA mimics of green fluorescent protein (GFP), and the selection used a library of ~$5x10^8$ sequences transformed into *Escherichia coli* cells containing the target molecules (non-fluorescent in an unbound state), and transcription was induced. The fluorescent cells were sorted from the non-fluorescent cells and re-cultured, and after 6 rounds of selection, highly fluorescent aptamer-target pairs were identified. Using FACS as a platform for aptamer selections is highly advantageous because of the efficient and automatic separation of aptamer-bound targets from unbound nucleic acids and targets, as well as the ability to easily normalize background binding. This technique can be used with a wide variety of targets; however, it requires highly specialized and very expensive equipment.

### 1.1.7 Microfluidic Approaches

Aptamer selection techniques have been miniaturized to take advantage of the benefits of microfluidics. Unique selection approaches have also been developed that can only be done within microfluidic devices. There are many advantages to using microfluidics, including decreased reagent consumption, reduced contamination potential, increased surface area to volume ratios and therefore decreased diffusion times, and automation potential. The few drawbacks in some cases could be complicated fabrication and operation, expensive materials, and high fabrication costs.

A process called micro free flow electrophoresis (μFFE) was developed using a custom device, and this technology has enabled electrophoretic SELEX to be performed using lower electric fields, shorter separation times, and larger volumes.[92] The 300-fold volume increase compared to typical CE-SELEX techniques enables significantly larger libraries ($10^{14}$ sequences) to be used. Lower electric fields and shorter separation times greatly reduce the potential for dissociation of bound complexes, and this has also drastically decreased the dilution that occurs from ~6000-fold in typical CE-SELEX[46-48,93] to ~100-fold, which is beneficial for downstream amplification. A microfluidic chip incorporating both dielectrophoresis and electrophoresis was used to improve the cell-SELEX process specifically by helping to overcome the prominent issue of dead cells contaminating the selection and internalizing sequences non-specifically.[94] Here, dielectrophoresis was used to assemble only viable cells within the chip, and electrophoresis was used to separate the unbound and loosely bound sequences from the cell-bound sequences.

Magnetic beads have been used in several microfluidic SELEX platforms. M-SELEX was one of the initial techniques described, and it begins by incubating the library with target-functionalized magnetic beads.[95] These beads are then injected into a Continuous-flow Magnetic

Activated Chip-based Separation (CMACS) device that separates the magnetic beads from the unbound sequences using microfabricated nickel ferromagnetic structures. Using very low bead numbers and stringent selection conditions, high-affinity aptamers were identified in a single round of selection. A similar process is used in the MicroMagnetic Separation (MMS) chip, except here the magnetic beads are held stationary in the device while solutions flow through the channel, which enables a wider range of flowrates to be used to control the stringency.[96,97] To further increase stringency, a process called Volume Dilution Challenge (VDC) was used with the MMS chip during which lower affinity binders are removed from the magnetic beads.[98]

One group has even developed a unique selection process that must be done in a microfluidic channel because it exploits a property of microfluidics to perform stringent partitioning of cell-bound aptamers away from unbound sequences. This technique uses inertial microfluidic SELEX (I-SELEX) to focus cells into a separate fluid channel from the unbound sequences using inertial forces in curvilinear channels.[99] Sol-gel technology has also been incorporated into microfluidic devices as a matrix for target immobilization.[100-102] Sol-gels are very porous, allowing nucleic acids to diffuse into them and bind to the target, which is in its native state as opposed to bound to a surface or bead. Using microfluidics, the increased surface area-to-volume ratio allows more efficient sampling of the library, since binding is diffusion limited.

One of the major benefits of microfluidic-based aptamer selections is the ability to incorporate multiple process steps into a single device and automate the process.[103] There have been several demonstrations of automated and semi-automated SELEX experiments using microfluidic devices. MSELEX was the initial use of microfluidics that downscaled a large automated SELEX system that performed the whole process from binding to amplification in a simpler miniaturized system.[104] Another approach incorporated more than one benchtop

technique to bind, wash, and transfer the bound sequences to a separate chamber.[105-107] This chip contains two chambers separated by a gel-filled channel with the first chamber packed with target-immobilized beads. The nucleic acid library flows through these beads and the bound and unbound sequences are separated by affinity chromatography. Incorporated microheaters facilitate thermal elution of the bound sequences, and these sequences are electrophoretically driven across the gel-filled channel to the second chamber, thereby preventing any contamination from the initial steps, and microheaters in the second chamber would allow for on-chip PCR.[107] This device was also demonstrated with cell-SELEX where cell culture was incorporated directly on-chip, and all selection steps were fully integrated to enable multiple SELEX cycles to be performed completely on-chip.[108] Furthermore, a selection to cells was also performed entirely at 37°C using this process to select aptamers that function optimally at physiological temperature.[109]

There have been several demonstrations of microfluidic selections using magnetic beads that have integrated all SELEX steps on-chip and automated the device operation. These chips use integrated microvalves, micropumps, and micromixers to control fluid flow, and also contain microheaters and temperature sensors for performing on-chip amplification.[110,111] This chip design was further improved to enable multiple rounds of selection to be completely automated.[112,113] In addition to using these devices to select aptamers to proteins, they have also been used in selections to cells[114,115] and viruses[116].

Given the time-consuming nature of most aptamer selection processes, the ability to perform multiplexed and high-throughput selections is highly advantageous. Microfluidic approaches can often be easily multiplexed, since the process is contained within a single small chip. For example, the sol-gel-incorporating device described above was scaled up to perform selections to five different targets on one chip.[117,118] Scaling up processes to be compatible with multi-well

microplates is extremely beneficial not only for the parallelization, but also because these setups can be automated relatively easily using existing liquid-handling robots.[119,120] Nitrocellulose filter binding experiments can be performed in a high-throughput manner using 96-well filter devices.[121,122] Additionally, microplates have been used in selections where targets are immobilized on magnetic beads[119,120,123-127] or directly onto the surface of the microwell.[128,129] A high-throughput affinity chromatography approach using an array of microcolumns in a device called MEDUSA (Microplate-based Enrichment Device Used for the Selection of Aptamers) was recently used to perform a high-throughput aptamer selection.[130,131] This device was designed to directly couple to a 96-well microplate to enable high-throughput downstream processing.

The development of high-throughput sequencing technologies has greatly benefitted the aptamer selection community by providing substantially more sequencing data, and this has enabled SELEX experiments to be performed in fewer rounds, as seen in a method called Quantitative Selection of Aptamers through Sequencing (QSAS) where M-SELEX was coupled with high-throughput sequencing to monitor multiplicity (copy number) and enrichment of different aptamer candidates through multiple rounds of selection.[132] M-SELEX was subsequently coupled to next generation sequencing with in-situ-synthesized aptamer arrays that enabled binding affinity measurements for thousands of aptamer candidates in a process called Quantitative Parallel Aptamer Selection System (QPASS).[133] High-throughput sequencing platforms have also recently been used for high-throughput aptamer binding assays using fluorescently-labeled targets. Nucleic acids are assembled on a chip and their sequences are determined along with their location on the chip. Then the fluorescently-labeled target is introduced and allowed to bind. The fluorescence intensity is measured and associated with a specific sequence based on location to determine the binding affinity of each individual sequence.

There have been three methods developed using this type of approach: HiTS-FLIP (High-Throughput Sequencing-Fluorescent Ligand Interaction Profiling)[134] for DNA aptamers, and HiTS-RAP (High-Throughput Sequencing-RNA Affinity Profiling)[135] and RNA-MaP (RNA on a Massively Parallel array)[136] for RNA aptamers. These techniques are capable of performing hundreds of millions of independent binding assays, which completely eliminates the trial-and-error-based identification of high-affinity aptamers from sequencing results alone.

Scaling up the aptamer selection process to enable simultaneous selections to many different targets increases the likelihood that the selection will be successful, which is beneficial since selections tend to have a low rate of success.[4,14] However, despite the plethora of aptamer selection methods that exist, novel aptamer discovery is still the major obstacle preventing the widespread use of aptamers.

## 2 Cell Capture in Microfluidic Devices using Aptamers

### *2.1 Cell-SELEX*

Cell-SELEX is the process by which aptamers are selected to whole cells. Cancer cell lines are one of the most extensively studied targets for cell-SELEX. To date, many aptamers to cancer cells have been discovered.[137,138] One of the major benefits of cell-SELEX is that knowledge about the cell surface composition is not required to select specific and high-affinity aptamers. In fact, in many cases, the specific cell surface species to which the aptamer binds is unknown.[139] Consequently, cell-SELEX has been used for novel biomarker discovery.[140] Additionally, selecting aptamers to whole cells allows the proteins and other molecules on the cell surface to be in their native state, which is crucial, since it is difficult to purify and maintain recombinant membrane proteins in their native state.[140] Furthermore, techniques have been developed in which

membrane proteins are intentionally displayed in excess on live cells to ensure their native conformation, and cell-SELEX is performed. Aptamers discovered via cell-SELEX can be used in many applications, including cell labeling, cell capture, and targeted therapies.

### *2.2 Microfluidic Aptamer-based Cancer Cell Capture*

Aptamers have been used as affinity capture ligands in microfluidic devices designed to separate and/or detect target species.[141-143] Aptamers have been used over antibodies in these types of applications because of their increased robustness and ease of functionalization and oriented immobilization. With the emergence of cell-SELEX and cell-specific aptamers, microfluidic devices for specific cell capture using aptamers have been developed. [144-146]

Many aptamer-based cell capture devices use aptamers that were previously selected via cell-SELEX. The Tan group has developed several devices for capturing circulating tumor cells (CTCs) using aptamers they selected to CCRF-CEM cells (CCL-119 T-cell, human acute lymphoblastic leukemia), Ramos cells (CRL-1596 B-cell, human Burkitt's lymphoma), Toledo cells (CRL-2631, non-Hodgkin's B-cell lymphoma), DLD-1 cells (Duke's type C colorectal adenocarcinoma), and HCT 116 cells (colorectal carcinoma).[145,147] These aptamers have also been used by several other research groups to develop cell capture devices and surfaces. For most devices, the aptamers are immobilized on the surface simply by adsorbing avidin onto the surface and using biotinylated aptamers. Aptamer-based cell capture was demonstrated in a capillary for 3 different cell types individually, and achieved capture efficiencies ranging from 83.6 ± 5.8% for DLD-1 cells to 97.2 ± 2.8% for HCT 116 cells.[147] Using straight polydimethylsiloxane (PDMS) microchannels bonded to glass, CCRF-CEM cells were captured with >80% efficiency and >97% purity.[148] A simple PDMS channel incorporating microheaters allowed for thermal elution of

captured cells and a regenerable aptamer surface.[149] Furthermore, a simple straight channel coated with aptamers was developed into a sandwich assay-based sensor for the detection of Ramos cells using aptamers conjugated to gold nanoparticles.[150] Similarly, a paper-based lateral flow assay was developed for the point-of-care detection of CCRF-CEM and Ramos cells using gold nanoparticle reporters with a limit of detection of just 800 cells.[151] A multiplex cell capture device was designed using an S-shaped microchannel.[152] Here, three capture regions with different surface-immobilized aptamers enabled the capture and isolation of three different cancer cell types from the same sample.

Creating surfaces with greater surface area capable of increased aptamer immobilization efficiency can improve the efficacy of cell capture. For example, a device incorporating multivalent nanospheres, where many aptamers were immobilized onto gold nanoparticles, demonstrated an increased cell binding affinity by 39-fold.[153] Additionally, dendrimer-modified surfaces[154], silicon nanowire arrays[155,156], and hydrogels[157] have been used to capture cells using specific aptamers. A micropatterned surface was even developed to isolate single cells using aptamers immobilized in microwells.[158] A single-cell occupancy of 88.2% was achieved. For capturing CTCs in actual biological samples, an aptamer-coated device containing micropillars was designed for high-throughput capture of CTCs from whole blood.[159] This device was capable of processing 1 mL of whole blood in just 28 mins, and achieved 95% capture efficiency and 81% purity, while 93% of the captured cells remained viable.

Aptamers selected to surface proteins that are overexpressed on cancer cells can also be used for CTC capture. Anti-epidermal growth factor receptor (EGFR) aptamers have been used to isolate human glioblastoma cells.[160] Microchannels with optimized flow rates,[161] and increased surface area using glass beads[162] and nanotextured PDMS[163] were developed to increase cell

capture efficiency. Anti-prostate-specific membrane antigen (PSMA) aptamers were immobilized in 51 ultra-high aspect ratio curvilinear poly(methyl methacrylate) (PMMA) microchannels to achieve a 90% recovery of LNCaP cells (prostate cancer).[164] This device was designed for high throughput, and was able to process 1 mL of sample in ~29 mins. A unique device was developed by Zhao *et al.* featuring a three-dimensional network of repeating anti-protein tyrosine kinase 7 (PTK7) aptamers created by rolling circle amplification (RCA).[165] These long aptamer molecules extended tens of micrometers from the surface of the PDMS channel, which greatly increased the cell capture efficiency compared to monovalent aptamers, and also enabled higher throughput through the device.

Using aptamers as rare cell capture ligands within microfluidic devices has many advantages over immobilized antibodies, and aptamers will likely continue to be used increasingly in these applications. However, the largest barrier to their widespread use will be discovering specific high-affinity aptamers to new cell types. In fact, a rapid and reliable cell-SELEX process would revolutionize personalized and precision medicine by enabling specific targeting of each patient's unique cancer cells.

## 3 Dissertation Overview

Here, I present research we have performed and devices we have developed for carrying out high-throughput aptamer selections and characterizations of the SELEX process, as well as capturing cancer cells and detecting genetic mutations. Our aptamer selection technology provides a platform to help overcome some of the main challenges in aptamer discovery, which will help to advance the field toward the widespread use of aptamers over antibodies. Equally important is the development of the applications in which aptamers play an essential role in providing an

improvement over existing technologies. Therefore, we also developed a device that uses aptamers to capture cancer cells, and isolates their genomic DNA (gDNA) for the purpose of genetic mutation analysis.

Since aptamer selections are time consuming, expensive, and often unsuccessful, we developed a device that uses affinity microcolumn chromatography called MEDUSA (Microplate-based Enrichment Device Used for the Selection of Aptamers). MEDUSA, along with a modified version of SELEX we developed, greatly reduce the time and reagents needed for an aptamer selection. MEDUSA is a high-throughput device consisting of 96 microcolumns arranged with the same dimensions and spacing as a 96-well microplate. This design allows it to directly couple with a 96-well plate enabling downstream processes to be carried out in a high-throughput manner. This device is also reconfigurable and can switch between the microcolumns connected in series and the microcolumns in parallel, making it extremely versatile and capable of many different conformations for different uses and experiments. Using MEDUSA, we have demonstrated characterization of the SELEX process and previously-selected aptamers, as well as a highly multiplexed aptamer selection. We performed a high-throughput experiment in which 96 binding tests were performed simultaneously to i) determine the optimal target loading for three protein targets on resin to maximize aptamer enrichment; ii) verify the partitioning efficiency and selectivity of previously selected aptamers; and iii) test the non-specific binding of two suspected background binding sequences.[130] In addition, we used MEDUSA to perform novel RNA aptamer selections to 19 different targets simultaneously.[131] We demonstrated the ability to perform the equivalent of a 10-cycle selection to 19 targets in approximately the same time as 4 traditional SELEX cycles to one target. Therefore, this technology could be used to establish a rapid and

reliable method for high-throughput aptamer discovery, which would advance the aptamer field and promote the widespread use of aptamers.

Although contributing to the effort toward the efficient discovery of high affinity aptamers is important, developing innovative applications that incorporate them to improve upon existing technologies is also crucial. Therefore, our next goal was to create a platform in which aptamers could be used to make an impact in the challenging but critically important field, cancer. Cancer is a complex disease defined by genetic mutations in cells that confer growth advantage. The ability to identify the mutations present in a cancer patient would give clinicians information that would help in determining the most effective treatment strategy for that patient. Therefore, we developed a device that uses aptamers to capture cancer cells, or more specifically, CTCs. Following cell capture, the gDNA from the captured cells is isolated on-chip, gene amplification is performed, and this amplification product is sequenced to determine the presence of any genetic mutations. Using the HeLa cervical cancer cell line and the CAOV-3 ovarian cancer cell line, we have demonstrated the amplification and sequencing of a fragment of the *TP53* gene, which is commonly mutated in cancer, and therefore a gene of interest. Using aptamers over antibodies for capturing CTCs is beneficial because they do not vary from batch to batch, they are more robust, and they are much less expensive to produce, which enables this technology to be used for frequent, non-invasive cancer cell testing.

For the great potential impact of aptamers to be realized, novel aptamer discovery must be improved and valuable applications incorporating aptamers must be developed. Moreover, precision medicine would revolutionize healthcare, and I believe aptamers and aptamer-based technologies could bring us closer to enhanced diagnostics and targeted therapies, which are especially critical in the effective treatment of complex cancer. The research presented here

contributes to the goals of improved aptamer discovery and important application development, which are crucial steps progressing toward aptamer-based precision medicine.

## 4 References


[1] M. C. Hennion, D. Barcelo, *Anal. Chim. Acta* **1998**, *362*, 3–34.

[2] P. Chames, M. Van Regenmortel, E. Weiss, D. Baty, *Br. J. Pharmacol.* **2009**, *157*, 220–233.

[3] S. D. Jayasena, Clin. Chem. 1999, 45, 1628–1650.

[4] L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, et al., *PLoS One* **2010**, *5*, e15004.

[5] M. Kimoto, R. Yamashige, K. Matsunaga, S. Yokoyama, I. Hirao, *Nat. Biotechnol.* **2013**, *31*, 453–457.

[6] K. Sefah, Z. Yang, K. M. Bradley, S. Hoshika, E. Jimenez, L. Zhang, G. Zhu, S. Shanker, F. Yu, D. Turek, et al., *Proc Natl Acad Sci U S A* **2014**, *111*, 1449–1454.

[7] Z. Yang, M. Durante, L. G. Glushakova, N. Sharma, N. A. Leal, K. M. Bradley, F. Chen, S. A. Benner, *Anal. Chem.* **2013**, *85*, 4705–4712.

[8] S. Klussmann, A. Nolte, R. Bald, V. A. Erdmann, J. P. Furste, *Nat. Biotechnol.* **1996**, *14*, 1112–1115.

[9] A. D. Keefe, S. T. Cload, *Curr. Opin. Chem. Biol.* **2008**, *12*, 448–456.

[10] L. Crouzier, C. Dubois, S. L. Edwards, L. H. Lauridsen, J. Wengel, R. N. Veedu, *PLoS One* **2012**, *7*, e35990.

[11] V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S.-Y. Peak-Chew, S. H. McLaughlin, et al., *Science* **2012**, *336*, 341–344.

[12] R. Stoltenburg, C. Reinemann, B. Strehlitz, *Biomol. Eng.* **2007**, *24*, 381–403.

[13] K. Szeto, H. G. Craighead, *Appl. Phys. Rev.* **2014**, *1*, 031103–1–031103–17.

[14] A. Ozer, J. M. Pagano, J. T. Lis, *Mol. Ther. Nucleic Acids* **2014**, *3*, e183.

[15] K.-M. Song, S. Lee, C. Ban, *Sensors* **2012**, *12*, 612–631.

[16] A. B. Iliuk, L. Hu, W. A. Tao, *Anal. Chem.* **2011**, *83*, 4440–4452.

[17] E. J. Cho, J.-W. Lee, A. D. Ellington, *Annu. Rev. Anal. Chem.* **2009**, *2*, 241–264.

[18] S. Tombelli, M. Minunni, M. Mascini, *Biosens. Bioelectron.* **2005**, *20*, 2424–2434.

[19] K. Liu, B. Lin, X. Lan, *J. Cell. Biochem.* **2013**, *114*, 250–255.

[20] E. Dausse, S. Da Rocha Gomes, J. J. Toulme *Curr. Opin. Pharmacol.* **2009**, *9*, 602–607.

[21] K. W. Thiel, P. H. Giangrande, *Oligonucleotides* **2009**, *19*, 209–222.

[22] A. D. Keefe, S. Pai, A. Ellington, *Nat. Rev. Drug Discov.* **2010**, *9*, 537–550.

[23] A. Z. Wang, O. C. Farokhzad, *J Nucl Med* **2014**, *55*, 353–356.

[24] H. Xing, K. Hwang, J. Li, S. F. Torabi, Y. Lu, *Curr. Opin. Chem. Eng.* **2014**, *4*, 79–87.

[25] W. Zhou, P.-J. J. Huang, J. Ding, J. Liu, *Analyst* **2014**, *139*, 2627–40.

[26] P. Hong, W. Li, J. Li, *Sensors* **2012**, *12*, 1181–1193.

[27] M. Ilgu, M. Nilsen-Hamilton, *Analyst* **2016**, DOI 10.1039/C5AN01824B.

[28] H. Sun, W. Tan, Y. Zu, *Analyst* **2016**, *141*, 403–415.

[29] G. F. Joyce, *Gene* **1989**, *82*, 83–87.

[30] C. Tuerk, L. Gold, *Science* **1990**, *249*, 505–510.

[31] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818–822.

[32] R. C. Conrad, L. Giver, Y. Tian, a. D. Ellington, *Methods Enzymol.* **1996**, *267*, 336–366.

[33] L. Giver, D. P. Bartel, M. L. Zapp, M. R. Green, A. D. Ellington, *Gene* **1993**, *137*, 25–31.

[34] H. Shi, X. C. Fan, Z. Y. Ni, J. T. Lis, *RNA* **2002**, *8*, 1461–1470.

[35] A. Ellington, J. Szostak, *Nature* **1992**, *355*, 850–852.

[36] R. D. Jenison, S. C. Gill, A. Pardi, B. Polisky, *Science.* **1994**, *263*, 1425–1429.

[37] C. Berens, A. Thain, R. Schroeder, *Bioorganic Med. Chem.* **2001**, *9*, 2549–2556.

[38] M. N. Win, J. S. Klein, C. D. Smolke, *Nucleic Acids Res.* **2006**, *34*, 5670–5682.

[39] L. A. Holeman, S. L. Robinson, J. W. Szostak, C. Wilson, *Fold. Des.* **1998**, *3*, 423–431.

[40] A. Ozer, B. S. White, J. T. Lis, D. Shalloway, *Nucleic Acids Res.* **2013**, *41*, 7167–7175.

[41] J. Wang, J. F. Rudzinski, Q. Gong, H. T. Soh, P. J. Atzberger, *PLoS One* **2012**, *7*, 1–8.

[42] D. R. Latulippe, K. Szeto, A. Ozer, F. M. Duarte, C. V Kelly, J. M. Pagano, B. S. White, D. Shalloway, J. T. Lis, H. G. Craighead, *Anal. Chem.* **2013**, *85*, 3417–3424.

[43] A. Nitsche, A. Kurth, A. Dunkhorst, O. Pänke, H. Sielaff, W. Junge, D. Muth, F. Scheller, W. Stöcklein, C. Dahmen, et al., *BMC Biotechnol.* **2007**, *7*, 48.

[44] S. D. Goodman, N. J. Velten, Q. Gao, S. Robinson, A. M. Segall, *J Bacteriol* **1999**, *181*, 3246–3255.

[45] R. Y. Tsai, R. R. Reed, *Mol Cell Biol* **1998**, *18*, 6447–6456.

[46] S. D. Mendonsa, M. T. Bowser, *Anal Chem* **2004**, *76*, 5387–5392.

[47] S. D. Mendonsa, M. T. Bowser, *J. Am. Chem. Soc.* **2004**, *126*, 20–21.

[48] R. K. Mosing, S. D. Mendonsa, M. T. Bowser, *Anal. Chem.* **2005**, *77*, 6107–6112.

[49] M. Jing, M. T. Bowser, *Anal. Chem.* **2013**, *85*, 10761–10770.

[50] J. Yang, M. T. Bowser, *Anal. Chem.* **2013**, *85*, 1525–1530.

[51] B. Lou, E. Chen, X. Zhao, F. Qu, J. Yan, *J. Chromatogr. A* **2016**, DOI 10.1016/j.chroma.2016.01.073.

[52] Z. Luo, H. Zhou, H. Jiang, H. Ou, X. Li, L. Zhang, *Analyst* **2015**, *140*, 2664–2670.

[53] M. Berezovski, M. Musheev, A. Drabovich, S. N. Krylov, *J. Am. Chem. Soc.* **2006**, *128*, 1410–1411.

[54] J. Ashley, K. Ji, S. F. Y. Li, *Electrophoresis* **2012**, *33*, 2783–2789.

[55] S. Saito, K. Hirose, M. Tsuchida, K. Wakui, K. Yoshimoto, Y. Nishiyama, M. Shibukawa, *Chem. Commun.* **2015**, *52*, 461–464.

[56] J. G. Bruno, *Biochem. Biophys. Res. Commun.* **1997**, *234*, 117–20.

[57] J. G. Bruno, J. L. Kiel, *Biotechniques* **2002**, *32*, 178–183.

[58] T. Chaou, B. Vialet, L. Azéma, *Methods* **2016**, DOI 10.1016/j.ymeth.2016.01.002.

[59] C. Y. Wang, B. L. Lin, C. H. Chen, *Int. J. Cancer* **2015**, *926*, 918–926.

[60] H.-F. Cui, Y.-J. Li, J. Wang, X.-J. Li, Q.-L. Wang, Y.-F. Bai, *RSC Adv.* **2016**, *6*, 1484–1490.

[61] R. Wilson, C. Bourne, R. R. Chaudhuri, R. Gregory, J. Kenny, A. Cossins, *PLoS One* **2014**, *9*, DOI 10.1371/journal.pone.0100572.

[62] R. Stoltenburg, C. Reinemann, B. Strehlitz, *Anal. Bioanal. Chem.* **2005**, *383*, 83–91.

[63] M. Fan, S. R. McBurnett, C. J. Andrews, A. M. Allman, J. G. Bruno, J. L. Kiel, *J. Biomol. Tech.* **2008**, *19*, 311–321.

[64] J. C. Lai, C. Y. Hong, *ACS Comb. Sci.* **2014**, *16*, 321–327.

[65] J. C. Lai, C. Y. Hong, *J. Mater. Chem. B* **2014**, *2*, 4114–4121.

[66] O. Aminova, M. D. Disney, in *Methods Mol. Biol.*, **2010**, pp. 209–224.

[67] O. Aminova, D. J. Paul, J. L. Childs-disney, M. D. Disney, *Biochemistry* **2008**, *47*, 12670–12679.

[68] J. L. Childs-Disney, M. Wu, A. Pushechnikov, O. Aminova, M. D. Disney, *ACS Chem. Biol.* **2007**, *2*, 745–754.

[69] M. D. Disney, L. P. Labuda, D. J. Paul, S. G. Poplawski, A. Pushechnikov, T. Tran, S. P. Velagapudi, M. Wu, J. L. Childs-Disney, *J. Am. Chem. Soc.* **2008**, *130*, 11185–11194.

[70] J. A. Martin, P. A. Mirau, Y. Chushak, J. L. Chávez, R. R. Naik, J. A. Hagen, N. Kelley-Loughnane, *J. Anal. Methods Chem.* **2015**, *2015*, 1-8.

[71] N. O. Fischer, J. B. H. Tok, T. M. Tarasow, *PLoS One* **2008**, *3*, 1–9.

[72] E. Katilius, C. Flores, N. W. Woodbury, *Nucleic Acids Res.* **2007**, *35*, 7626–7635.

[73] Y. Zhu, P. Chandra, C. Ban, Y. B. Shim, *Electroanalysis* **2012**, *24*, 1057–1064.

[74] C. G. Knight, M. Platt, W. Rowe, D. C. Wedge, F. Khan, P. J. R. Day, A. Mcshea, J. Knowles, D. B. Kell, *Nucleic Acids Res.* **2009**, *37*, 1–10.

[75] M. Platt, W. Rowe, D. C. Wedge, D. B. Kell, J. Knowles, P. J. R. Day, *Anal. Biochem.* **2009**, *390*, 203–205.

[76] L. H. Lauridsen, H. A. Shamaileh, S. L. Edwards, E. Taran, R. N. Veedu, *PLoS One* **2012**, *7*, 1–6.

[77] L. Peng, B. J. Stephens, K. Bonin, R. Cubicciotti, M. Guthold, *Microsc. Res. Tech.* **2007**, *70*, 372–381.

[78] Y. Miyachi, N. Shimizu, C. Ogino, A. Kondo, *Nucleic Acids Res.* **2009**, *38*, 1–8.

[79] T. S. Misono, P. K. R. Kumar, *Anal. Biochem.* **2005**, *342*, 312–317.

[80] N. A. C. Ngubane, L. Gresh, A. Pym, E. J. Rubin, M. Khati, *Biochem. Biophys. Res. Commun.* **2014**, *449*, 114–119.

[81] E. Dausse, A. Barré, A. Aimé, A. Groppi, A. Rico, C. Ainali, G. Salgado, W. Palau, E. Daguerre, M. Nikolski, et al., *Biosens. Bioelectron.* **2016**, DOI 10.1016/j.bios.2016.02.003.

[82] F. Pileur, M. L. Andreola, E. Dausse, J. Michel, S. Moreau, H. Yamada, S. A. Gaidamakov, R. J. Crouch, J. J. Toulmé, C. Cazenave, *Nucleic Acids Res.* **2003**, *31*, 5776–5788.

[83] M. Khati, M. Schüman, J. Ibrahim, Q. Sattentau, S. Gordon, W. James, J. F. Aptamers, M. Schu, *J. Virol.* **2003**, *77*, 12692–12698.

[84] Y. Chen, K. Nakamoto, O. Niwa, R. M. Corn, *Langmuir* **2012**, *28*, 8281–8285.

[85] M. S. L. Raddatz, A. Dolf, E. Endl, P. Knolle, M. Famulok, G. Mayer, *Angew. Chemie - Int. Ed.* **2008**, *47*, 5190–5193.

[86] G. Mayer, M.-S. L. Ahmed, A. Dolf, E. Endl, P. A. Knolle, M. Famulok, *Nat. Protoc.* **2010**, *5*, 1993–2004.

[87] X. Yang, S. E. Bassett, X. Li, B. a Luxon, N. K. Herzog, R. E. Shope, J. Aronson, T. W. Prow, J. F. Leary, R. Kirby, et al., *Nucleic Acids Res.* **2002**, *30*, e132.

[88] X. Yang, X. Li, T. W. Prow, L. M. Reece, S. E. Bassett, B. a Luxon, N. K. Herzog, J. Aronson, R. E. Shope, J. F. Leary, et al., *Nucleic Acids Res.* **2003**, *31*, e54.

[89] Z. Zhu, Y. Song, C. Li, Y. Zou, L. Zhu, Y. An, C. J. Yang, *Anal. Chem.* **2014**, *86*, 5881–5888.

[90] J. Wang, Q. Gong, N. Maheshwari, M. Eisenstein, M. L. Arcila, K. S. Kosik, H. T. Soh, *Angew. Chemie - Int. Ed.* **2014**, *53*, 4796–4801.

[91] J. Zou, X. Huang, L. Wu, G. Chen, J. Dong, X. Cui, Z. Tang, *J. Mol. Evol.* **2015**, *81*, 172–178.

[92] M. Jing, M. T. Bowser, *Lab Chip* **2011**, *11*, 3703–3709.

[93] S. D. Mendonsa, M. T. Bowser, *J. Am. Chem. Soc.* **2005**, *127*, 9382–9383.

[94] H. Stoll, H. Kiessling, M. Stelzle, H. P. Wendel, J. Schutte, B. Hagmeyer, M. Avci-Adali, *Biomicrofluidics* **2015**, *9*, 34111.

[95] X. Lou, J. Qian, Y. Xiao, L. Viel, A. E. Gerdon, E. T. Lagally, P. Atzberger, T. M. Tarasow, A. J. Heeger, H. T. Soh, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 2989–94.

[96] J. R. Qian, X. H. Lou, Y. T. Zhang, Y. Xiao, H. T. Soh, *Anal. Chem.* **2009**, *81*, 5490–5495.

[97] S. S. Oh, K. Plakos, X. Lou, Y. Xiao, H. T. Soh, *Proc. Natl. Acad. Sci.* **2010**, *107*, 14053–14058.

[98] S. S. Oh, K. M. Ahmad, M. Cho, S. Kim, Y. Xiao, H. T. Soh, *Anal. Chem.* **2011**, *83*, 6883–6889.

[99] C. M. Birch, H. W. Hou, J. Han, J. C. Niles, *Sci. Rep.* **2015**, *5*, 11347.

[100] S. M. Park, J. Y. Ahn, M. Jo, D. K. Lee, J. T. Lis, H. G. Craighead, S. Kim, *Lab Chip* **2009**, *9*, 1206–1212.

[101] J.-Y. Ahn, M. Jo, P. Dua, D.-K. Lee, S. Kim, *Oligonucleotides* **2011**, *21*, 93–100.

[102] H. Bae, S. Ren, J. Kang, M. Kim, Y. Jiang, M. M. Jin, I. M. Min, S. Kim, *Nucleic Acid Ther.* **2013**, *23*, 443–9.

[103] J. P. Hilton, T. Olsen, J. Kim, J. Zhu, T. H. Nguyen, M. Barbu, R. Pei, M. Stojanovic, Q. Lin, *Microfluid. Nanofluidics* **2015**, *19*, 795–804.

[104] G. Hybarger, J. Bynum, R. F. Williams, J. J. Valdes, J. P. Chambers, *Anal. Bioanal. Chem.* **2006**, *384*, 191–198.

[105] J. Kim, J. P. Hilton, K. A. Yang, R. Pei, K. Ennis, M. Stojanovic, Q. Lin, in *Proc. IEEE Int. Conf. Micro Electro Mech. Syst.*, **2012**, pp. 765–768.

[106] J. Kim, J. P. Hilton, K. A. Yang, R. Pei, M. Stojanovic, Q. Lin, *Sensors Actuators, A Phys.* **2013**, *195*, 183–190.

[107] J. Kim, J. P. Hilton, K. A. Yang, R. Pei, J. Zhu, M. Stojanovic, Q. Lin, in *Proc. IEEE Int. Conf. Micro Electro Mech. Syst.*, **2013**, pp. 1007–1010.

[108] J. Zhu, T. Olsen, R. Pei, M. Stojanovic, Q. Lin, in *2014 IEEE 27th Int. Conf. Micro Electro Mech. Syst.*, **2014**, pp. 242–245.

[109] J. P. Hilton, J. Kim, T. H. Nguyen, M. Barbu, R. Pei, M. Stojanovic, Q. Lin, in *Proc. IEEE Int. Conf. Micro Electro Mech. Syst.*, **2012**, pp. 100–103.

[110] C. J. Huang, H. I. Lin, S. C. Shiesh, G. Bin Lee, *Biosens. Bioelectron.* **2010**, *25*, 1761–1766.

[111] C. J. Huang, H. I. Lin, S. C. Shiesh, G. Bin Lee, *Biosens. Bioelectron.* **2012**, *35*, 50–55.

[112] Y. H. Chen, H. I. Lin, C. J. Huang, S. C. Shiesh, G. Bin Lee, *Microfluid. Nanofluidics* **2012**, *13*, 929–939.

[113] H.-I. Lin, C.-C. Wu, C.-H. Yang, K.-W. Chang, G.-B. Lee, S.-C. Shiesh, *Lab Chip* **2015**, *15*, 486–494.

[114] C.-H. Weng, I.-S. Hsieh, L.-Y. Hung, H.-I. Lin, S.-C. Shiesh, Y.-L. Chen, G.-B. Lee, *Microfluid. Nanofluidics* **2013**, *14*, 753–765.

[115] L. Hung, C.-H. Wang, Y. Che, C. Fu, H. Chang, K. Wang, G. Lee, *Sci. Rep.* **2015**, *5*, 10326.

[116] H.-C. Lai, C.-H. Wang, T.-M. Liou, G.-B. Lee, *Lab Chip* **2014**, *14*, 2002–13.

[117] T. K. Kim, S. W. Lee, J. Y. Ahn, T. Laurell, S. Y. Kim, O. C. Jeong, *Jpn. J. Appl. Phys.* **2011**, *50*, 2–6.

[118] S. W. Lee, J. Kang, S. Ren, T. Laurell, S. Kim, O. C. Jeong, *Biochip J.* **2013**, *7*, 38–45.

[119] J. C. Cox, P. Rudolph, A. D. Ellington, *Biotechnol. Prog.* **1998**, *14*, 845–850.

[120] J. C. Cox, A. D. Ellington, *Bioorganic Med. Chem.* **2001**, *9*, 2525–2531.

[121] K. B. Hall, J. K. Kranz, *Methods Mol. Biol.* **1999**, *118*, 105–14.

[122] I. Wong, T. M. Lohman, *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5428–32.

[123] A. Wochner, B. Cech, M. Menger, V. A. Erdmann, J. Glökler, *Biotechniques* **2007**, *43*, 344–353.

[124] T. Schütze, B. Wilhelm, N. Greiner, H. Braun, F. Peter, M. Mörl, V. A. Erdmann, H. Lehrach, Z. Konthur, M. Menger, et al., *PLoS One* **2011**, *6*, 1–10.

[125] J. C. Cox, A. Hayhurst, J. Hesselberth, T. S. Bayer, G. Georgiou, A. D. Ellington, *Nucleic Acids Res.* **2002**, *30*, e108.

[126] J. C. Cox, M. Rajendran, T. Riedel, E. A. Davidson, L. J. Sooter, T. S. Bayer, M. Schmitz-Brown, A. D. Ellington, *Comb. Chem. High Throughput Screen.* **2002**, 5, 289–299.

[127] P. W. Goertz, J. C. Cox, A. D. Ellington, *J. Assoc. Lab. Autom.* **2004**, *9*, 150–154.

[128] D. W. Drolet, R. D. Jenison, D. E. Smith, D. Pratt, B. J. Hicke, *Comb. Chem. High Throughput Screen.* **1999**, 2, 271–278.

[129] A. Jolma, T. Kivioja, J. Toivonen, L. Cheng, G. Wei, M. Enge, M. Taipale, J. M. Vaquerizas, J. Yan, M. J. Sillanpa, et al., *Genome Res.* **2010**, *20*, 861–873.

[130] K. Szeto, S. J. Reinholt, F. M. Duarte, J. M. Pagano, A. Ozer, L. Yao, J. T. Lis, H. G. Craighead, *Anal. Bioanal. Chem.* **2014**, *406*, 2727–2732.

[131] S. J. Reinholt, A. Ozer, J. T. Lis, H. G. Craighead, *Sci. Rep.* **2016**, *6*, 29771.

[132] M. Cho, Y. Xiao, J. Nie, R. Stewart, A. T. Csordas, S. S. Oh, J. A. Thomson, H. T. Soh, *Proc. Natl. Acad. Sci.* **2010**, *107*, 15373–15378.

[133] M. Cho, S. Soo Oh, J. Nie, R. Stewart, M. Eisenstein, J. Chambers, J. D. Marth, F. Walker, J. A. Thomson, H. T. Soh, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 18460–5.

[134] R. Nutiu, R. C. Friedman, S. Luo, I. Khrebtukova, D. Silva, R. Li, L. Zhang, G. P. Schroth, C. B. Burge, *Nat. Biotechnol.* **2011**, *29*, 659–664.

[135] J. M. Tome, A. Ozer, J. M. Pagano, D. Gheba, G. P. Schroth, J. T. Lis, *Nat. Methods* **2014**, *11*, 683–688.

[136] J. D. Buenrostro, C. L. Araya, L. M. Chircus, C. J. Layton, H. Y. Chang, M. P. Snyder, W. J. Greenleaf, *Nat. Biotechnol.* **2014**, *32*, 562–568.

[137] J. A. Phillips, D. Lopez-Colon, Z. Zhu, Y. Xu, W. Tan, *Anal. Chim. Acta* **2008**, *621*, 101–108.

[138] Y. Zhang, Y. Chen, D. Han, I. Ocsoy, W. Tan, *Bioanalysis* **2010**, *2*, 907–918.

[139] D. D. Dickey, P. H. Giangrande, *Methods* **2015**, DOI 10.1016/j.ymeth.2015.11.020.

[140] S. Ohuchi, *Biores. Open Access* **2012**, *1*, 265–272.

[141] Y. Xu, X. Yang, E. Wang, *Anal. Chim. Acta* **2010**, *683*, 12–20.

[142] F. Du, L. Guo, Q. Qin, X. Zheng, G. Ruan, J. Li, G. Li, *TrAC - Trends Anal. Chem.* **2015**, *67*, 134–146.

[143] B. Soontornworajit, Y. Wang, *Anal. Bioanal. Chem.* **2011**, *399*, 1591–1599.

[144] Y. Pu, Z. Zhu, H. Liu, J. Zhang, J. Liu, W. Tan, *Anal. Bioanal. Chem.* **2010**, *397*, 3225–3233.

[145] X. Fang, W. Tan, *Acc. Chem. Res.* **2010**, *43*, 48–57.

[146] D. D. Dickey, P. H. Giangrande, *Methods* **2016**, *97*, 94–103.

[147] J. A. Martin, J. A. Phillips, P. Parekh, K. Sefah, W. Tan, *Mol. Biosyst.* **2011**, *7*, 1720–7.

[148] J. A. Phillips, Y. Xu, Z. Xia, Z. H. Fan, W. Tan, *Anal. Chem.* **2009**, *81*, 1033–1039.

[149] J. Zhu, T. Nguyen, R. Pei, M. Stojanovic, Q. Lin, *Lab Chip* **2012**, *12*, 3504–3513.

[150] W. Liu, H. Wei, Z. Lin, S. Mao, J. Lin, *Biosens. Bioelectron.* **2011**, *28*, 438–442.

[151] G. Liu, X. Mao, J. A. Phillips, H. Xu, W. Tan, L. Zeng, *Anal. Chem.* **2009**, *81*, 10013–10018.

[152] Y. Xu, J. A. Phillips, J. Yan, Q. Li, Z. H. Fan, W. Tan, *Anal. Chem.* **2009**, *81*, 7436–7442.

[153] W. Sheng, T. Chen, W. Tan, Z. H. Fan, *ACS Nano* **2013**, *7*, 7067–7076.

[154] F. Zheng, Y. Cheng, J. Wang, J. Lu, B. Zhang, Y. Zhao, Z. Gu, *Adv. Mater.* **2014**, *26*, 7333–7338.

[155] L. Chen, X. Liu, B. Su, J. Li, L. Jiang, D. Han, S. Wang, *Adv. Mater.* **2011**, *23*, 4376–4380.

[156] Q. Shen, L. Xu, L. Zhao, D. Wu, Y. Fan, Y. Zhou, W. H. Ouyang, X. Xu, Z. Zhang, M. Song, et al., *Adv. Mater.* **2013**, *25*, 2368–2373.

[157] Z. Zhang, N. Chen, S. Li, M. R. Battig, Y. Wang, *J. Am. Chem. Soc.* **2012**, *134*, 15716–15719.

[158] Q. Chen, J. Wu, Y. Zhang, Z. Lin, J.-M. Lin, *Lab Chip* **2012**, *12*, 5180–5.

[159] W. Sheng, T. Chen, R. Kamath, X. Xiong, W. Tan, Z. H. Fan, *Anal Chem* **2012**, *84*, 4199–4206.

[160] Y. Wan, Y. T. Kim, N. Li, S. K. Cho, R. Bachoo, A. D. Ellington, S. M. Iqbal, *Cancer Res.* **2010**, *70*, 9371–9380.

[161] Y. Wan, J. Tan, W. Asghar, Y. T. Kim, Y. Liu, S. M. Iqbal, *J. Phys. Chem. B* **2011**, *115*, 13891–13896.

[162] Y. Wan, Y. L. Liu, P. B. Allen, W. Asghar, M. A. I. Mahmood, J. F. Tan, H. Duhon, Y. T. Kim, A. D. Ellington, S. M. Iqbal, *Lab Chip* **2012**, *12*, 4693–4701.

[163] Y. Wan, M. A. I. Mahmood, N. Li, P. B. Allen, Y. T. Kim, R. Bachoo, A. D. Ellington, S. M. Iqbal, *Cancer* **2012**, *118*, 1145–1154.

[164] U. Dharmasiri, S. Balamurugan, A. A. Adams, P. I. Okagbare, A. Obubuafo, S. A. Soper, *Electrophoresis* **2009**, *30*, 3289–3300.

[165] W. Zhao, C. H. Cui, S. Bose, D. Guo, C. Shen, W. P. Wong, K. Halvorsen, O. C. Farokhzad, G. S. L. Teo, J. A. Phillips, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 19626–31.

# CHAPTER 2

# HIGH-THROUGHPUT BINDING CHARACTERIZATION OF RNA APTAMER SELECTIONS USING A MICROPLATE-BASED MULTIPLEX MICROCOLUMN DEVICE[1]

## 1 Introduction

One of the main challenges in the aptamer field is the efficient discovery of new high-affinity aptamers. Consequently, there is interest in improving SELEX technology to obtain high-quality aptamers much more rapidly. New techniques include microfluidic devices that use dynamic behavior and flow to improve the efficiency of selections.[1-11] Furthermore, some designs incorporate several or all of the processing steps, and some have even been automated.[12-15] However, many technologies are difficult to scale for multiplexed or parallel selections. Highly multiplexed selections have been done using microplate technologies, which are particularly beneficial due to the availability of automated liquid-handling robots,[16,17] but are limited to traditional equilibrium solution binding[18] or interactions where the target species are bound to the surface of the well,[19,20] as opposed to flow and dynamic behavior achievable in microfluidic devices.

Furthermore, despite the development of improved SELEX processes, little effort has been made to characterize and optimize SELEX technologies. Although simple theory and simulations have been used in an effort to predict optimal SELEX conditions,[21-26] recent binding studies have shown significant discrepancies with existing theory.[27,28] This has resulted in the use of empirical optimization for selection conditions and the development of new models.[29] These results and

[1] This section has been adapted and reproduced with modifications to text and formatting with permission from K. Szeto, S. J. Reinholt, F. M. Duarte, J. M. Pagano, A. Ozer, L. Yao, J. T. Lis, H. G. Craighead, *Anal. Bioanal. Chem.* **2014**, *406*, 2727–2732.

techniques will be important as new high-throughput technologies are developed to determine the most effective and robust aptamer selection approaches under the available parameters.

Toward these goals, we have developed a high-throughput device called MEDUSA (Microplate-based Enrichment Device Used for the Selection of Aptamers). This device was designed to have the same dimensions as a 96-well microplate, which allows it to directly couple to a 96-well plate. This not only enables high-throughput aptamer selections, but also facilitates the use of existing plate-based methods for upstream and downstream processing, and the use of existing plate-based technologies for potential process automation. MEDUSA is a highly scaled-up version of our previously-developed modular and multiplexable microcolumns, which were capable of being configured in both parallel and serially-connected microcolumns. MEDUSA was also designed to be reconfigurable in this way. Using the microcolumns, selection conditions were also optimized to maximize the enrichment of high-affinity aptamers.[28] Here, we demonstrate using MEDUSA to characterize the binding of several RNA aptamers against various targets by performing 96 binding tests simultaneously. Four binding studies were carried out with 8 conditions each, and each condition was performed in triplicate to quantify the reproducibility, making a total of 96 binding tests. One study characterized the binding of an RNA aptamer to green fluorescent protein (GFP) (GFPapt)[30], over a random RNA library (N70) with increasing concentrations of GFP, aiming to reproduce previously-reported enrichment behavior indicating an optimal GFP concentration above which aptamer binding decreased likely due to steric hindrance. Additionally, analogous studies were performed for human heat shock factor 1 (hHSF1) and negative elongation factor E (NELF-E) and their respective RNA aptamers, HSFapt and NELFapt,[28,31] to determine optimal selection target concentrations and test the theory of steric hindrance. These tests showed similar optima for maximum aptamer enrichment, and

demonstrated the utility of MEDUSA for the characterization and optimization of aptamer selection parameters. The final binding study tested the specificity of the aptamers to their respective protein targets by connecting several of MEDUSA's microcolumns in series and exposing each aptamer to 8 different targets. As expected, the aptamers only significantly enriched on their specific targets. Two suspected background binding sequences were also tested in this way, and were shown to bind to the affinity resin and the device itself. Through the use of MEDUSA to perform these characterization tests, critical binding behaviors and background binding aptamers were determined that have a great impact on the performance and efficiency of aptamer selections.

Alternatively, this technology could be used for separations and sample preparation. By immobilizing aptamers on a resin within the microcolumns, the target of interest can be isolated and concentrated into a small volume. Aptamers have been used for sample preparation previously,[32-34] but by incorporating them into the microcolumns, very low concentrations of target in large volumes can be accommodated and the target can be concentrated by orders of magnitude into small volumes. We have demonstrated using the GFP RNA aptamer immobilized in microcolumns to significantly concentrate small quantities of GFP from large volumes. This could be easily scaled up using MEDUSA to achieve higher throughput or even multiple purification steps in series.

## 2 Design and Fabrication of MEDUSA

The motivation for designing MEDUSA was to construct a versatile device capable of performing high-throughput aptamer selections, as well as characterizations and optimizations of the aptamer selection process. Therefore, the basis for MEDUSA's design is the existing

microcolumn technology, which minimizes reagent consumption and has been shown to be successful in performing novel aptamer selections and characterizations of the selection process to maximize aptamer enrichment.[28] Potential automation of the selection process was also taken into consideration, and we determined that a design based on a 96-well microplate would be ideal due to the availability of existing microplate-based techniques and liquid-handling robots.[16,17] MEDUSA was designed to have the same dimensions as a microplate with 96 microcolumns spaced 9 mm apart in a 12x8 grid. This allowed easy coupling to a 96-well plate (Figure 2.1) to continue with high-throughput downstream processes. The microcolumns can be connected in series or used in parallel, which provides several benefits, including the ability to perform negative or counter selections simultaneously, the use of one aliquot of valuable starting library for selections to multiple different targets, and the prevention of cross-contamination through parallel use during certain steps of the process. Consequently, MEDUSA was designed to be reconfigurable with the ability to connect the microcolumns in series or in parallel. To accomplish

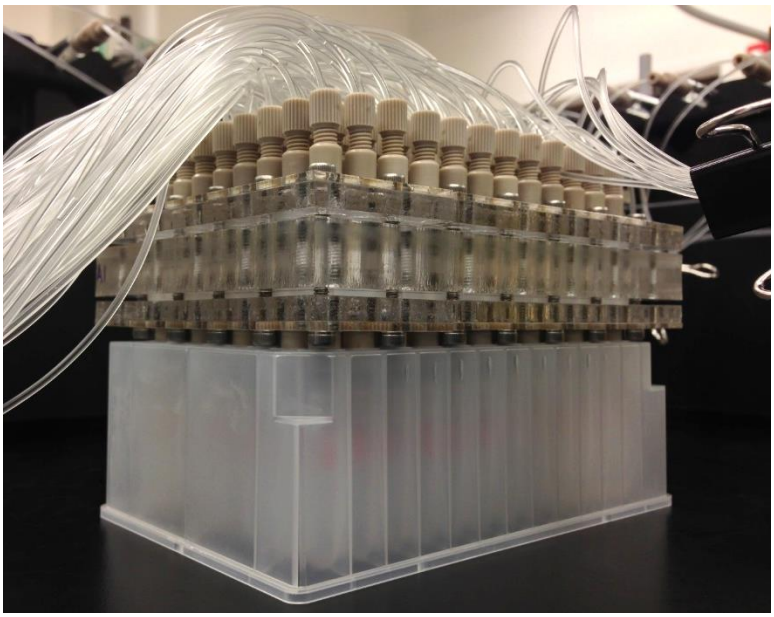


**Figure 2.1:** Image of MEDUSA when it is directly coupled to a 96-well microplate.

this, MEDUSA has a layered design consisting of layers that alter how the microcolumns are connected.

Figure 2.2 shows schematics of an exploded view of MEDUSA with microcolumns connected in series and in parallel, as well as images of the assembled device in both configurations. The layers of MEDUSA are made from 2 types of material, poly(methyl methacrylate) (PMMA) and silicone, and the fluidic connections are made using PEEK NanoPorts from IDEX. The middle layer of the device (labeled “1” in Figure 2.2) is 1/2" thick PMMA and this layer contains the microcolumns. There are 1/16” adhesive silicone layers (labeled “5” in Figure 2.2) on either side of the microcolumn layer, providing a liquid-tight seal between the layers

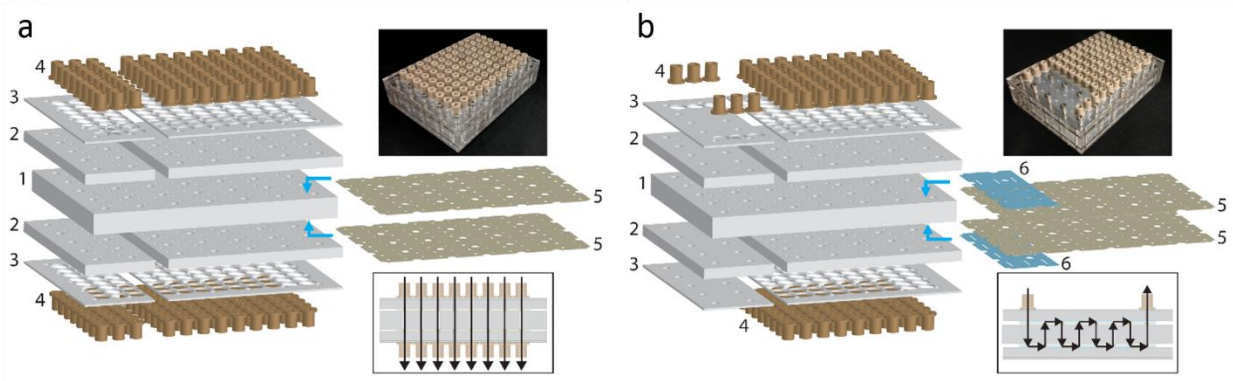



**Figure 2.2:** Schematic of the layers of MEDUSA in the order of assembly. (a) An exploded view of the customized device layers for configuring all 96 microcolumns to run in parallel. The flow path is shown in the lower boxed inset with no connections between microcolumns. The layers numbered 1 to 3 are the PMMA layers: the middle layer (1) containing the microcolumns, the next outer two layers being the caps (2) and washers (3). The outermost layers (4) consist of inlet and outlet ports that are bonded to the caps. The two layers numbered (5) are silicone layers, which are bonded to the microcolumn layer (1) to hold porous frits against both sides of the microcolumns to retain affinity resin and to make liquid-tight seals across the entire device. A photograph of MEDUSA assembled in parallel is shown in the upper inset. (b) The customized device layers for configuring 24 of the microcolumns to run in series. The two additional silicone layers (6) shown in blue, as well as the smaller complementary plastic layers (2 and 3), are specifically programmed to connect 3 sets of 8 microcolumns within the device. The flow path is shown in the lower boxed inset with microcolumns connected in series via a serpentine route through 8 microcolumns. MEDUSA assembled to run in series and parallel is shown in the upper inset.

of MEDUSA and housing 2 mm polyethylene frits that act to contain the affinity resins. Next there are 1/4" PMMA capping layers (labeled “2” in Figure 2.2) that contain small holes for fluid flow, and provide a surface on which the NanoPorts (labeled “4” in Figure 2.2) are bonded to allow standard connections at the inputs and outputs of the device. The outermost PMMA layers (labeled “3” in Figure 2.2) are 1 mm thick and contain large holes to facilitate precise placement of the NanoPorts. They also serve as a “washer” by distributing the forces required to hold the device together. For simple assembly, all layers contain 35 holes with just the center microcolumn layer having threaded holes to seal the device using screws (see the upper inset images in Figures 2.2a and b). For connecting the microcolumns in series, there are two additional non-adhesive 1/32” silicone layers (labeled “6” in Figure 2.2b), which contain small channels that connect adjacent microcolumns. There are also different capping layers to control the locations of the device inputs and outputs.

The layers of MEDUSA were inexpensively and rapidly machined using a fabrication method that took a two-dimensional CAD for each layer and cut the material with a $CO_2$ laser at 10.6 μm (Universal Laser Systems VersaLaser). The intensity, speed, and density of the laser pulses were optimized for each thickness and type of material to maximize the efficiency, quality, and reproducibility of each layer. This fabrication technique is ideal for this application because it is inexpensive and very fast, with one device taking only about an hour to machine. It also imparts fewer machining defects than the manual machining used to fabricate the individual microcolumns, and enables easy customization of the devices by simply altering the CAD. For example, in this demonstration of MEDUSA, we have 3 sets of 8 microcolumns connected in series, and 72 microcolumns in parallel (Figure 2.2). We chose to divide the capping and washer

layers to enable the reconfiguration of the serially-connected microcolumns without disrupting the other microcolumns.

## 3 Characterization of selection parameters and previously-selected aptamers using MEDUSA

To demonstrate MEDUSA's ability to characterize aspects of the aptamer selection process and selected aptamers in a high-throughput manner, four binding studies were designed. Three of these studies aimed to determine the optimal protein target loading on the affinity resins for 3 different targets, GFP, hHSF1, and NELF-E, that would produce the highest aptamer enrichment. Properties of these proteins can be found in Table 2.1. This type of study was previously carried out for GFP using the individual microcolumns,[28] so repeating it here would further validate the findings, which showed an optimal GFP concentration of 0.6 µg per µL of resin above which the aptamer enrichment decreased. This trend was attributed to steric hindrance, and studying other proteins in this way would test this hypothesis. The 4$^{th}$ binding study tested the specificity of each aptamer, as well as two suspected background binding sequences that were present in multiple aptamer selection pools,[28,31,35] against 8 different targets by measuring the enrichment of each aptamer. Enrichments were measured using quantitative polymerase chain reaction (qPCR) analysis, and all of the binding studies were performed in triplicate simultaneously using all of MEDUSA's 96 microcolumns. A layout of all the samples in MEDUSA is illustrated in Figure 2.3.

**Table 2.1:** Properties of the target proteins

| Protein | Molecular Weight (kDa) | Isoelectric Point | Affinity tag |
|---|---|---|---|
| GFP | 27 | 5.5 | Hexahistidine (n-terminus) |
| hHSF1 | 86 | 5.3 | GST* (n-terminus) |
| NELF-E | 36 | 8.9 | Hexahistidine (n-terminus) |
| UBLCP1 | 37 | 6.1 | Hexahistidine (c-terminus) |

*GST tag ~ 30 kDa

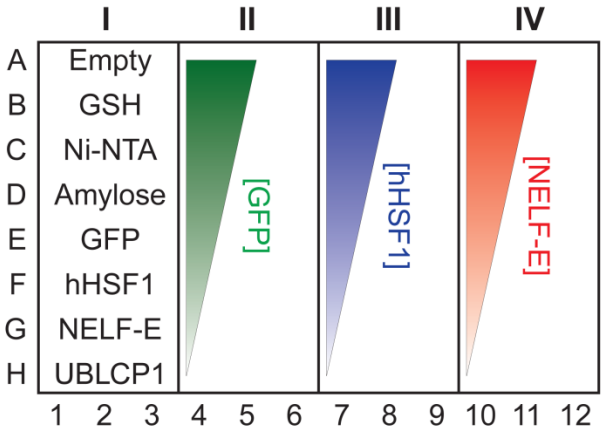



**Figure 2.3** Layout for the 96 samples in MEDUSA according to its analogous microplate position given by the rows A-H, and the columns 1-12. In section I, the 8 indicated targets were connected in series from A to H to test the specificity and partitioning efficiency of various RNA aptamers. This was tested in triplicate in columns 1 to 3. Sections II, III, and IV tested the effects of target surface concentration on aptamer enrichments. The colored triangles indicate decreasing concentrations of each protein from 10 µg/µL (row A) to 0.016 µg/µL (row H) in 2.5-fold dilutions. Section II (green triangle) aimed to confirm previous enrichment behaviors shown with GFP. Sections III and IV tested the same concentrations for the proteins hHSF1 (blue triangle) and NELF-E (red triangle) to assess the prevalence of target surface concentration effects on binding due to steric hindrance or other effects in aptamer selections.

### *3.1 Recombinant protein target, RNA library, and aptamer preparation*

Recombinant proteins were expressed in BL21(DE3)-RIPL *E. coli* cells (Agilent Technologies) transformed with plasmids that encode for hexahistidine (6xHis)-tagged GFP, NELF-E, and UBLCP1, or glutathione S-transferase (GST)-tagged hHSF1 (Table 2.1). Two- or four-liter LB cultures supplemented with 100 µg/mL ampicillin were inoculated with starter LB culture derived from a single colony and grown at 37°C until the OD600 reached approximately 0.6. Protein expression was induced by the addition of IPTG to a final concentration of 1 mM. After an additional incubation, bacteria were collected by centrifugation and the resulting pellet was processed using nickel-nitrilotriacetic acid (Ni-NTA) Superflow (Qiagen) or glutathione (GSH)-agarose (Thermo Scientific) resins according to the manufacturer's instructions. Sodium

dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to verify the purity and quality of the final protein product. Resulting protein preps were dialyzed against 1× PBS (supplemented with 5 mM 2-mercaptoethanol and 0.01% Triton X-100) and stored in small aliquots after addition of glycerol to a final concentration of 20%.

The random N70 RNA library used in the experiments contains ~$5\times10^{15}$ sequences that are 120-nucleotides (nt) in length and was prepared as described previously.[28] This library consists of a 70-nt random region flanked by two constant regions as described elsewhere.[36] hHSF1- and NELF-E-binding aptamers, HSFapt and NELFapt, as well as the background binding sequences, BBS1 and BBS2, were derived from previous multiplex SELEX experiments that used the N70 library, and therefore contain the same 5' and 3' constant regions. HSFapt was previously identified as hHSF2-R5-2 and characterized elsewhere.[28] NELFapt was previously identified as Napt1.[31] The background binding sequences were identified in several previous multiplex SELEX experiments for dozens of target proteins.[28,31,35]] The GFP-binding aptamer, GFPapt, was selected using a different library with a smaller random region and different constant regions, and was previously identified as AP3-1.[30] For these experiments, the sequence-verified DNA templates for the specific aptamers and background binding sequences were *in vitro* transcribed using T7 RNA Polymerase. The transcribed RNA was subsequently treated with DNase I (Ambion), PAGE-purified, phenol:chlrorform and chloroform extracted, isopropanol precipitated, and resuspended in diethylpyrocarbonate (DEPC)-treated water. The sequences of the RNA aptamers, background binding sequences, and library template can be found in Table 2.2. All of the primers used for qPCR analysis are listed in Table 2.3, and all of the oligomers used in this work were obtained from Integrated DNA Technologies.

**Table 2.2:** Aptamer nucleotide sequences

| Sequence Name | Nucleotide Sequence (5'-3') |
|---|---|
| GFPapt | AGCUUCUGGACUGCGAUGGGAGCACGAAACGUCGUGGCGCAAUUGG GUGGGGAAAGUCCUUAAAAGAGGGCCACCACAGAAGCU (84 nt) |
| HSFapt | GGGAAUGGAUCCACAUCUACGAAUUCAAUCAAGUCCCCAGACUCAGC AACACUGGACAGCGAUAUGCAGAUAACCAAGACCAAUUCACUCCAGU UCACUGCAGACUUGACGAAGCUU (117 nt) |
| NELFapt | GGGAAUGGAUCCACAUCUACGAAUUCCCAACGACUGCCGAGCGAGAU UACGCUUGAGCGCCCCACUGAGGAUGCCCACGGGCGAUUGGGGCACG GCUUCACUGCAGACUUGACGAAGCUU (120 nt) |
| BBS1 | GGGAAUGGAUCCACAUCUACGAAUUCCGCAGGGCUAGCCGCAUGCUC AGGCCUGGCGGGUAGGGAGUUAGGGUAGGGAGACCAGGAGAGCUGG CUUCACUGCAGACUUGACGAAGCUU (118 nt) |
| BBS2 | GGGAAUGGAUCCACAUCUACGAAUUCCGAAGCUCGUGACGGUACCUC CUAAAAUGUCCAUGGGGAAGGGAGGGAAUGGGAAGGACAAUCGGAC ACCGUUCACUGCAGACUUGACGAAGCUU (121 nt) |
| N70 Library | GGGAAUGGAUCCACAUCUACGAAUUC---N70--- UUCACUGCAGACUUGACGAAGCUU |

*The underlined portions of the sequences denote the two constant regions.

**Table 2.3:** Primer sequences used for qPCR analysis

| Sequence Name | Forward Primer Sequence (5'-3') | Reverse Primer Sequence (5'-3') |
|---|---|---|
| GFPapt | GCTTCTGGACTGCGATGGGAGCA | GCTTCTGTGGTGGCCCTCTTTTAA GGACT |
| HSFapt | AATCAAGTCCCCAGACTCAGCAA CA | CTGGAGTGAATTGGTCTTGGTTAT C |
| NELFapt | CCAACGACTGCCGAGCGAGATTA C | GCCGTGCCCCAATCGCCCGTG |
| BBS1 | CGCAGGGCTAGCCGCATG | GCCAGCTCTCCTGGTCTCC |
| BBS2 | CGAAGCTCGTGACGGTACC | CGGTGTCCGATTGTCCTTC |
| N70 Library | GATAATACGACTCACTATAGGGA ATGGATCCACATCTACGA | AAGCTTCGTCAAGTCTGCAGTGA A |

### *3.2 Performing RNA selections and characterizations*

For the studies investigating the optimal protein target loading on the resin, MEDUSA was connected with the microcolumns in parallel, and 8 different concentrations of GFP, hHSF1, and NELF-E were immobilized, ranging from 10 µg of protein per µL of resin to 0.016 µg/µL in 2.5-fold dilutions. To study the specificity of each aptamer and the non-specific binding of two

background binding sequences, MEDUSA was connected in series such that 8 microcolumns would be exposed to a single solution.

Ni-NTA and GSH resins were washed extensively with binding buffer [10 mM N-2-hydroxyethylpiperazine-N′-ethanesulfonic acid (HEPES)−KOH pH 7.6, 125 mM NaCl, 25 mM KCl, 5 mM $MgCl_2$, and 0.02% Tween-20]. Hexahistidine-tagged GFP, NELF-E, and UBLCP1, and GST-tagged hHSF1, were immobilized on Ni-NTA and GSH resin, respectively. The washed resin was incubated with the protein at the desired concentration in a 10% slurry with binding buffer at 4°C with constant mixing for 1 hour. The protein-immobilized resin was then loaded into MEDUSA using a pipette, and MEDUSA was assembled.

For the protein loading studies, each microcolumn was exposed to a solution containing a mixture of 4.95 nM N70 library and 50 pM specific aptamer, and for the specificity studies, 8 targets in series were exposed to a mixed aptamer/N70 library solution in binding buffer [4.75 nM N70 library, 50 pM GFPapt, 50 pM HSFapt, 50 pM NELFapt, 50 pM BBS1, 50 pM BBS2, and 10µg/mL yeast tRNA (Invitrogen)]. After a 10-µL aliquot of each solution was taken to use as a standard for qPCR analysis, 990 µL was injected at 33 µL/min using programmable multichannel syringe pumps (Harvard Apparatus) with MEDUSA coupled to a 96-well formatted liquid waste reservoir. All buffers and solutions were degassed prior to being introduced into the microcolumns. Following aptamer/library binding, the serially-configured microcolumns were reconnected in parallel, and all 96 microcolumns were washed with 3 mL of binding buffer at 300 µL/min to eliminate unbound and loosely-bound RNA. MEDUSA was then placed onto a 96-well microplate with 2-mL wells to collect the eluent. The RNA/protein/RNA-protein complexes were eluted from the microcolumns by flowing elution buffer [binding buffer + 50 mM ethylenediaminetetraacetic acid (EDTA pH 8.0) for selections with Ni-NTA resin; binding buffer + 10 mM glutathione for

selections with GSH resin; binding buffer + 10 mM maltose for selections with amylose-resin] at 50 μL/min for 12 min.

The eluted RNA and input standards were phenol:chloroform extracted and ethanol-precipitated with 1 μL of GlycoBlue (Ambion) and 40 μg of yeast tRNA (Invitrogen), and the resulting pellets were resuspended in 20 μL of DEPC water. Four μL of the resuspended samples were then reverse transcribed using Moloney Murine Leukemia Virus Reverse Transcriptase (MMLV-RT) in two 96-well microplates. The N70 library, NELFapt and HSFapt were reverse transcribed using the N70 library reverse primer, and GFPapt using the GFPapt reverse primer (Table 2.3).

For qPCR analysis, 10 μL of each cDNA sample was used, and qPCR was performed in 384-well plates on a LightCycler 480 instrument (Roche) to measure the amount of RNA library and specific aptamer recovered from each microcolumn. The specific primers listed in Table 2.3 were used in this analysis.

### *3.3 Parallel selections reveal critical target concentration for maximal aptamer enrichments*

In previous work, we determined an optimal concentration of GFP that maximized GFPapt enrichment.[28] Above this concentration, there was a decrease in enrichment, and we attributed this decrease to steric hindrance as the protein density on the resin increased. To determine the prevalence of this limiting effect, we decided to use MEDUSA to reproduce the GFP results with additional data points, and test two other proteins, hHSF1 and NELF-E, and their respective aptamers. With MEDUSA connected in parallel, eight different concentrations of each protein,

starting at 10 µg per µL of resin and proceeding in 2.5-fold dilutions down to 0.016 µg/µL, were tested in triplicate (Figure 2.3, sections denoted II, III, and IV).

The qPCR results indicating the level of binding, as well as the aptamer enrichment for all three proteins are shown in Figure 2.4. The flowrates used for binding in this experiment (33 µL/min) were slightly lower than those used in the previous GFP experiment. Lower flowrates

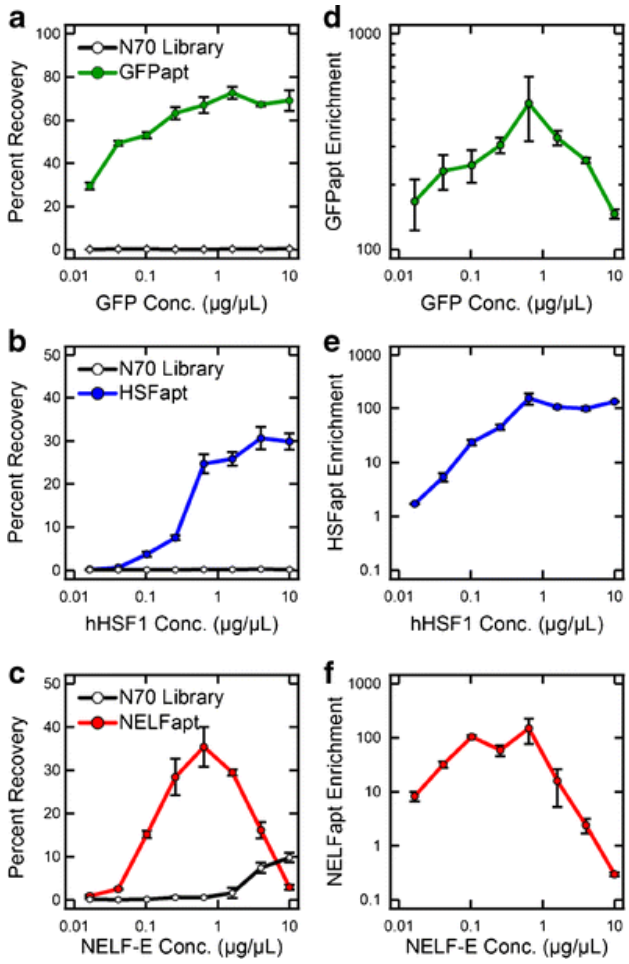



**Figure 2.4:** Recoveries and enrichments of specific RNA aptamers over the N70 library as a function of protein concentration. (a) The recovery of GFPapt and N70 library at various concentrations of GFP. Analogous data for the recovery of (b) HSFapt and N70 library from hHSF1, and (c) NELFapt and N70 library from NELF-E. (d–f) The calculated enrichments of the specific aptamers (GFPapt, HSFapt, NELFapt) over the random library. The error bars represent the standard deviation in recoveries or enrichments calculated for each condition performed in triplicate.

were used to increase the exposure times of the aptamers to the protein targets, since HSFapt and NELFapt have higher $K_D$s than GFPapt. This was meant to increase the percentage of bound aptamers to hHSF1 and NELF-E. Consequently, this also increased the amount of bound GFPapt and decreased the amount of N70 library recovered from the GFP microcolumns (Figure 2.4a). This resulted in an increased enrichment of GFPapt over the N70 library. However, the characteristic shape of the enrichment curve was the same as previously reported, and the optimal GFP loading concentration of 0.6 µg/µL was reproduced (Figure 2.4d). The recovery of HSFapt from the hHSF1 microcolumns followed a sigmoidal shape, where the binding and enrichment increased steadily and then saturated at higher hHSF1 concentrations (Figure 2.4b and e). Interestingly, the HSFapt enrichment plateaued at the same optimal protein concentration as for GFP. For NELF-E, there is a very clear optimal concentration for maximum NELFapt recovery, and it too is at 0.6 µg/µL (Figure 2.4c). Above this concentration, there is a drastic decrease in NELFapt binding. Additionally, N70 library binding increased significantly with NELF-E concentrations above 0.6 µg/µL, which is likely due to NELF-E's RNA Recognition Motif that causes it to bind RNA non-specifically.[31] These two binding trends not only result in a drastic decrease of NELFapt enrichment above the optimal concentration, but actually result in the de-enrichment of NELFapt at the highest concentration tested here, 10 µg/µL (Figure 2.4f).

The concentration studies performed here with GFP, hHSF1, and NELF-E present a strong case for the general steric hindrance argument of target molecules that are over-packed in solid-phase affinity selections. Although the GFP and hHSF1 aptamer recovery curves do not show substantial decreases at high concentrations as with NELF-E, this binding behavior is affected by the selection flowrates, which is clearly evident between our old and new GFP data, so it is possible that this phenomenon would occur for other protein targets using different selection parameters.

Importantly, aptamer recoveries saturated well below 100% loading of the protein targets on the resin, and this indicates the presence of a limiting effect. This is most clearly revealed by NELF-E, where the binding site for NELFapt appears to be significantly less accessible at higher concentrations resulting in a drastic decrease in total aptamer binding. Furthermore, a simple calculation assuming hard spheres for the resin predicts a critical surface protein density between 0.1 and 1 µg/µL (depending on protein size and shape, and the resin particle size). Therefore, since GFP, hHSF1, and NELF-E are similarly sized, it is not surprising that the same optimal concentration of 0.6 µg/µL for maximum aptamer enrichment was observed. These results suggest that the target concentration may be the most limiting parameter for enriching aptamers in solid-phase selections.

### *3.4 Multiplex serial selections show specificity of target aptamers and background binding sequences*

In previous selections, we connected several microcolumns together to perform multiple partitions to input pools and libraries.[28] We showed the highly specific and efficient partitioning of GFPapt to GFP over non-specific proteins and an empty microcolumn. This serial microcolumn configuration is useful for multitasking nucleic acid libraries for selections to multiple unrelated targets, or for separating enriched pools for aptamers that bind to different distinct sites on a complex target.[1] Here, we decided to demonstrate the similar multiplex selection capability of MEDUSA by expanding the experiment to include several protein targets and their respective RNA aptamers: GFP and GFPapt, hHSF1 and HSFapt, and NELF-E and NELFapt. To thoroughly characterize the specific, non-specific, and background binding of each RNA aptamer, we also included a non-specific protein, UBLCP1; three commonly used affinity resins, GSH, Ni-NTA, and amylose; and an empty microcolumn. The four protein targets were immobilized at 0.6 µg/µL

onto their respective resins. MEDUSA was assembled with the eight target microcolumns connected in series, and the multiplex selection was carried out in triplicate to quantify the reproducibility of each aptamer's partitioning efficiency and specificity. The order of the target microcolumns is shown in Figure 2.3, section I, and the flow was from top to bottom.

Two suspected BBSs, BBS1 and BBS2, were also included in the starting pool along with the random N70 library and the specific aptamers. These BBSs were chosen from previous multiplex selections in which the final pools underwent high-throughput sequencing to provide tremendous amounts of sequencing data and the sensitivity for early detection of aptamers.[28,31,35] Comparing the sequencing results from many different targets revealed several identical sequences that were frequently enriched in earlier cycles before the target-specific aptamers dominated the pool. Two sequences, BBS1 and BBS2, were present in many aptamer pools, and were actually generally among the most enriched candidates in selections where the targets were less aptagenic (see Table 2.4). Based on these data, we predicted that both BBS1 and BBS2 would enrich on all targets by binding to the device itself and the resins, but that BBS2 would enrich more strongly than BBS1, especially in microcolumns containing Ni-NTA resin (similar analyses have been used to identify NI-NTA resin-specific sequences).[37]

**Table 2.4:** Frequencies of BBS1 and BBS2 in previous selections. Summary of the number of times BBS1 or BBS2 has been identified in all previous selections. The numbers indicate the instances in which BBS1 was more highly enriched than BBS2 (or vice versa) on each target, grouped according to the resin on which each target was immobilized.

| Resin | BBS1 Dominant | BBS2 Dominant |
|---|---|---|
| Ni-NTA | 1 | 7 |
| GSH | 3 | 4 |
| Amylose | 3 | 4 |
| Empty (no Resin) | 0 | 2 |

The results from these experiments are presented in Figure 2.5 as enrichments over the random N70 library in each microcolumn for the three RNA aptamers and the BBSs. The specific aptamers show substantial enrichment only on their intended target. GFPapt enriched an average of 750-fold in the microcolumns containing GFP, and actually de-enriched on all other targets with an average enrichment of 0.6-fold (Figure 2.5a), which reflects its high specificity to GFP. Similarly, HSFapt enriched an average of 232-fold in hHSF1 microcolumns and only an average

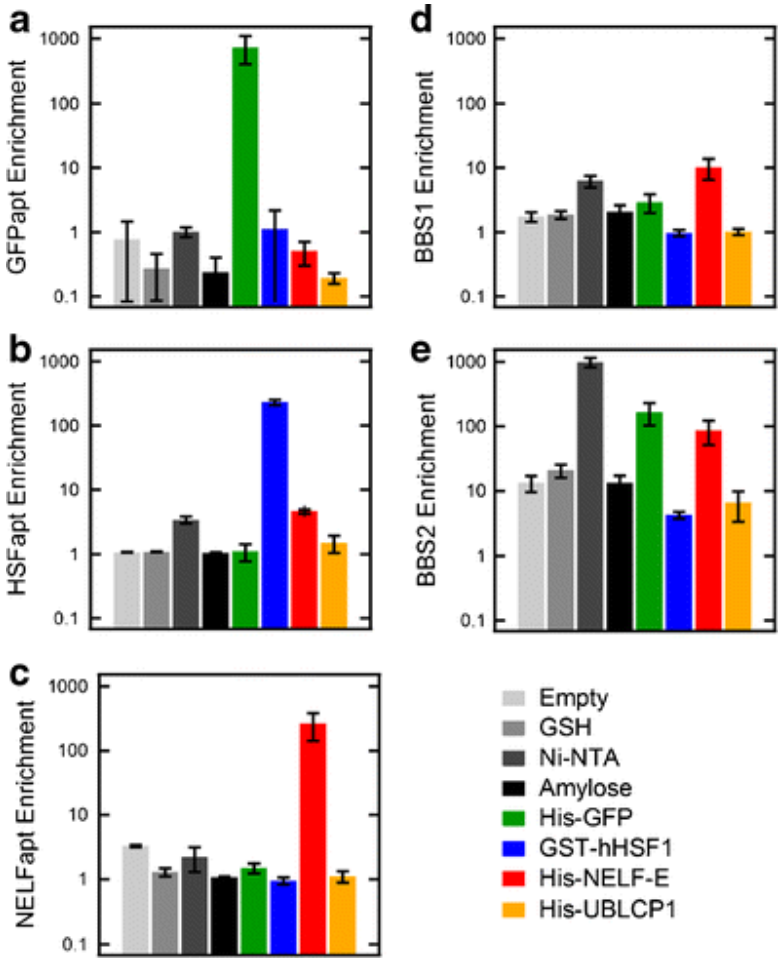



**Figure 2.5:** The enrichment of RNA aptamers over the N70 library on various targets connected in series. The enrichment of each protein-specific aptamer, GFPapt (a), HSFapt (b), and NELFapt (c), and non-specific sequences, BBS1 (d), and BBS2 (e), on all eight microcolumns. The error bars represent the standard deviation in enrichments calculated for each target performed in triplicate.

of 2-fold on all other targets (Figure 2.5b), and NELFapt enriched an average of 262-fold on NELF-E and only an average of 1.6-fold on other targets (Figure 2.5c).

The partitioning results for BBS1 and BBS2 showed agreement with the qualitative analysis of previous selection data.[28,31,35] BBS1 enriched on all targets as predicted with an average enrichment of 1.6-fold; however, enrichments were three times higher for Ni-NTA-containing microcolumns (Ni-NTA, GFP, NELF-E, and UBLCP1) with an average enrichment of 5.1-fold (Figure 2.5d). Similarly, BBS2 also enriched as predicted, with higher enrichments than BBS1 for all targets (Figure 2.5e). BBS2 also showed a high affinity to Ni-NTA, with enrichments of 13-fold on non-Ni-NTA microcolumns and a striking 311-fold average enrichment on Ni-NTA targets. In fact, in the first Ni-NTA-containing microcolumn (blank Ni-NTA resin), the average enrichment was almost 1000-fold, reflecting the specificity of BBS2 for Ni-NTA. In support of this hypothesis, we also noticed a monotonic decrease in enrichment of BBS2 across the Ni-NTA microcolumns, indicating a depletion of BBS2 from the pool as it was injected across Ni-NTA-containing microcolumns.

Negative selections are often used to eliminate sequences from the pool that bind to background sources; however, they are rarely able to completely remove the sequences that bind non-specifically. Thus, being able to identify these BBSs in different aptamer selections enables us to use them as indicators of selection progress. Furthermore, these BBSs could also be used to generate non-specific blocking reagents that are more effective than the yeast tRNAs that are typically used.

## 4 Sample purification and concentration using microcolumns

To effectively detect an analyte that exists in minute quantities, sample preparation is very important, and can greatly affect the limit of detection. It is highly advantageous and often necessary to purify and concentrate the analyte of interest prior to detecting it. This can be done using aptamers as the biorecognition element, and can be performed efficiently using our microcolumn technology. By immobilizing aptamers within the microcolumns, the target of interest that may exist at very low concentrations in large volumes can be accommodated and the target can be concentrated by orders of magnitude into small volumes. We have demonstrated this using an RNA aptamer to GFP immobilized on resin within a microcolumn, and concentrating a dilute solution of GFP into a small volume.

### *4.1 Microcolumn Fabrication*

The microcolumns consist of both custom-fabricated and commercially available parts, and a detailed description is available elsewhere.[28] Briefly, a polyethylene terephthalate (PET) rod is used to construct the microcolumn body. A hole 1 mm in diameter is drilled through the center of the plastic rod, and a shallow 1/16” hole is drilled at one end to house a polyethylene frit, which will contain the resin within the microcolumn. A NanoPort (IDEX Health and Science) is bonded to either end of the microcolumn to accommodate standard tubing connections. Microcolumns with a 10-µL volume were used in these studies.

### *4.2 GFP RNA Aptamer Preparation*

A DNA template for the GFP aptamer[30] that contained the T7 RNA polymerase promoter sequence at the 5’-end and an immobilization sequence at the 3’end was purchased as a gBlock

from Integrated DNA Technologies (see Table 2.5). The DNA template was PCR-amplified, purified using Zymo Clean and Concentrator columns, and verified by polyacrylamide gel electrophoresis (PAGE). The GFP RNA aptamer with the immobilization region was synthesized using T7 RNA polymerase (Ambion), and the DNA template was digested using DNase1. The RNA was purified by phenol-chloroform extraction and ethanol precipitation, verified by denaturing (7M urea) PAGE, and quantified using the Qubit BR RNA assay.

**Table 2.5:** Aptamer and primer nucleotide sequences.

| **Sequence Name** | **Nucleotide Sequence (5'-3')** |
|---|---|
| DNA Template: T7 promoter + GFPapt + *Immobilization region* | GATAATACGACTCACTATAAGCTTCTGGACTGCGATGGGAGCA CGAAACGTCGTGGCGCAATTGGGTGGGGAAAGTCCTTAAAAGA GGGCCACCACAGAAGCT*TTCACTGCAGACTTGACGAAGCTT* |
| GFP RNA aptamer + *Immobilization region* | AGCUUCUGGACUGCGAUGGGAGCACGAAACGUCGUGGCGCAA UUGGGUGGGGAAAGUCCUUAAAAGAGGGCCACCACAGAAGC*U UCACUGCAGACUUGACGAAGCUU* |
| Forward Primer | GATAATACGACTCACTATAAGCTTCTGGACTGCG |
| Reverse Primer | AAGCTTCGTCAAGTCTGCAGTGAAAGC |
| Immobilization Probe | Biotin—AAGCTTCGTCAAGTCTGCAGTGAA |

### *4.3 Purification and Concentration of GFP using Aptamers in Microcolumns*

Aptamer-immobilized resin was prepared to facilitate GFP capture and concentration (see Figure 2.6). Streptavidin resin was thoroughly washed with binding buffer and resuspended at a 10% slurry. GFP RNA aptamer and biotinylated immobilization probe were added to the resin slurry at concentrations of 30 pmol per µL of resin and 33 pmol per µL of resin, respectively, and this mixture was incubated at 4°C for 1 hour with constant mixing. To prepare the microcolumn, a 1.5 mm polyethylene frit was inserted into the hole at the outlet of the microcolumn. The aptamer-immobilized resin was added to the microcolumn using a pipette, and the fluidic connectors were connected.

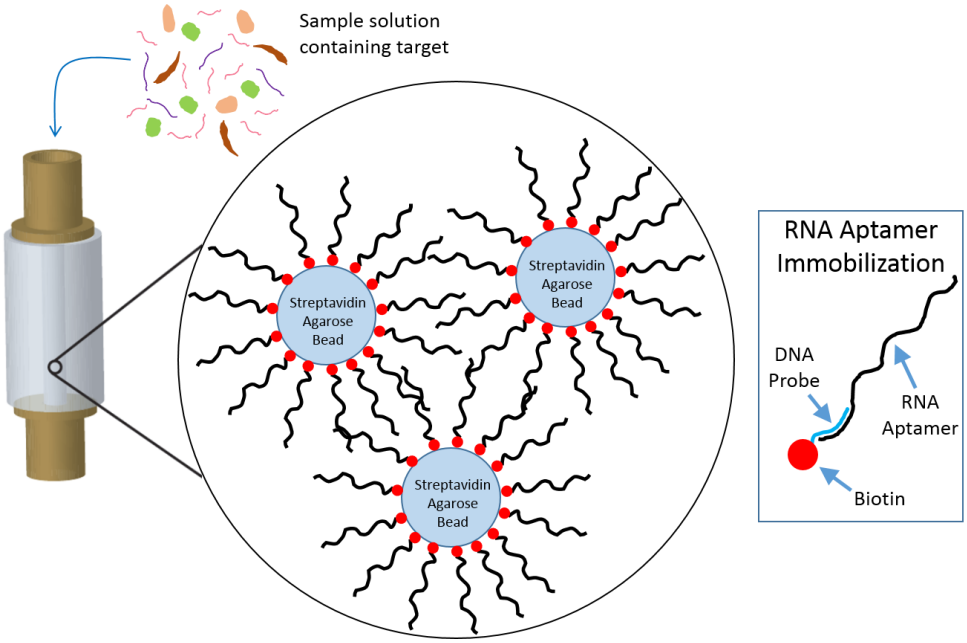



**Figure 2.6:** Schematic of aptamer-immobilized resin contained within a microcolumn for capturing and concentrating a target of interest in a sample.

All solutions were degassed prior to being pumped through the microcolumn. To prime the resin, 1 mL of binding buffer was pumped through at 50 µL/min. A solution containing 25 pmol (~700 ng) of GFP in 1 mL of binding buffer was prepared and pumped through the microcolumn at 10 µL/min. A washing step was performed by pumping 1 mL of binding buffer at 50 µL/min. For these experiments, the elution involved using enzymes to digest the aptamers, as heat elution would denature the GFP proteins and may render them non-fluorescent. A cocktail of RNases, including RNase A, T1, and H, in binding buffer was used to digest the RNA aptamer, thereby allowing the GFP protein to flow out of the microcolumn. Thirty µL of this solution was pumped at 50 µL/min, incubated for 30 min at room temperature, and then 70 µL was pumped through at 10 µL/min. The eluent was collected for fluorescence analysis and comparison to the initial GFP solution using a NanoDrop 3300 Fluorospectrometer. Figure 2.7a shows images of the microcolumn before and after elution. Fluorescence analysis indicated that all of the GFP was

recovered using this elution method (Figure 2.7b), and the GFP was successfully concentrated by a factor of 10 (Figure 2.7c).

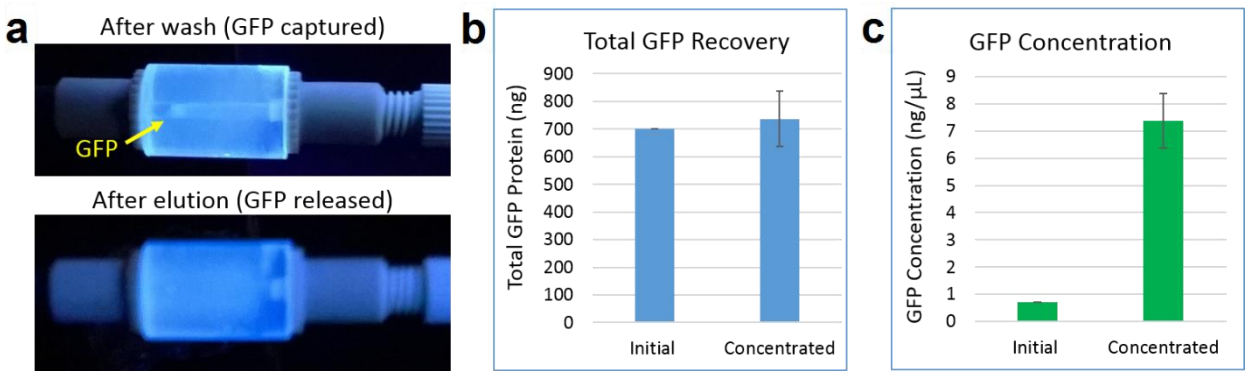



**Figure 2.7:** Fluorescence analysis of aptamer-based GFP capture and concentration. (a) Fluorescent images showing the microcolumn before elution containing GFP (upper photo), and after elution where the GFP signal has diminished (lower photo). (b) Measurement of the total GFP recovered by NanoDrop. (c) Measurement of the concentration change between the initial GFP solution pumped through the microcolumn and the recovered eluent by NanoDrop. The error bars indicate the standard deviation between triplicate NanoDrop measurements.

## 5 Conclusions

We believe MEDUSA could be used to improve high-throughput aptamer selections in productivity and efficiency. MEDUSA's design that allows for versatility and programmability enables a wide variety of SELEX designs to be explored in a large-scale manner. Here, we demonstrated how MEDUSA can be used to thoroughly characterize affinity chromatography-based aptamer selections. This enables this technique to be optimized to potentially outperform more traditional methods, while consuming fewer reagents, such as target protein, which is very advantageous because it allows aptamers to be selected to proteins that are difficult to purify or express in large quantities. Furthermore, not only can MEDUSA be used to optimize solid-phase affinity selections, but it can also be used to develop improved theories to better our understanding of these types of selections. Additionally, we demonstrated the use of microcolumns containing

aptamer-immobilized resin for capturing and concentrating a target of interest from a dilute sample. This could easily be scaled up and performed with MEDUSA, which would add sample preparation and separations to the list of uses for this versatile device.

## 6 References


[1] Q. Gong, J. Wang, K. M. Ahmad, A. T. Csordas, J. Zhou, J. Nie, R. Stewart, J. A. Thomson, J. J. Rossi, H. T. Soh, *Anal. Chem.* **2012**, *84*, 5365–5371.

[2] K. M. Ahmad, S. S. Oh, S. Kim, F. M. McClellen, Y. Xiao, H. T. Soh, *PLoS One* **2011**, *6*, DOI 10.1371/journal.pone.0027051.

[3] M. Cho, Y. Xiao, J. Nie, R. Stewart, A. T. Csordas, S. S. Oh, J. A. Thomson, H. T. Soh, *Proc. Natl. Acad. Sci.* **2010**, *107*, 15373–15378.

[4] M. Jing, M. T. Bowser, *Lab Chip* **2011**, *11*, 3703–3709.

[5] S. W. Lee, J. Kang, S. Ren, T. Laurell, S. Kim, O. C. Jeong, *Biochip J.* **2013**, *7*, 38–45.

[6] S. S. Oh, K. M. Ahmad, M. Cho, S. Kim, Y. Xiao, H. T. Soh, *Anal. Chem.* **2011**, *83*, 6883–6889.

[7] S. M. Park, J. Y. Ahn, M. Jo, D. K. Lee, J. T. Lis, H. G. Craighead, S. Kim, *Lab Chip* **2009**, *9*, 1206–1212.

[8] J. Kim, J. P. Hilton, K. A. Yang, R. Pei, M. Stojanovic, Q. Lin, *Sensors Actuators, A Phys.* **2013**, *195*, 183–190.

[9] X. Lou, J. Qian, Y. Xiao, L. Viel, A. E. Gerdon, E. T. Lagally, P. Atzberger, T. M. Tarasow, A. J. Heeger, H. T. Soh, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 2989–94.

[10] J. R. Qian, X. H. Lou, Y. T. Zhang, Y. Xiao, H. T. Soh, *Anal. Chem.* **2009**, *81*, 5490–5495.

[11] J.-Y. Ahn, M. Jo, P. Dua, D.-K. Lee, S. Kim, *Oligonucleotides* **2011**, *21*, 93–100.

[12] Y. H. Chen, H. I. Lin, C. J. Huang, S. C. Shiesh, G. Bin Lee, *Microfluid. Nanofluidics* **2012**, *13*, 929–939.

[13] C. J. Huang, H. I. Lin, S. C. Shiesh, G. Bin Lee, *Biosens. Bioelectron.* **2010**, *25*, 1761–1766.

[14] C. J. Huang, H. I. Lin, S. C. Shiesh, G. Bin Lee, *Biosens. Bioelectron.* **2012**, *35*, 50–55.

[15] G. Hybarger, J. Bynum, R. F. Williams, J. J. Valdes, J. P. Chambers, *Anal. Bioanal. Chem.* **2006**, *384*, 191–198.

[16] J. C. Cox, A. D. Ellington, *Bioorganic Med. Chem.* **2001**, *9*, 2525–2531.

[17] J. C. Cox, P. Rudolph, A. D. Ellington, *Biotechnol. Prog.* **1998**, *14*, 845–850.

[18] D. Eulberg, K. Buchner, C. Maasch, S. Klussmann, *Nucleic Acids Res.* **2005**, *33*, 1–10.

[19] D. W. Drolet, R. D. Jenison, D. E. Smith, D. Pratt, B. J. Hicke, *Comb. Chem. High Throughput Screen.* **1999**, 2, 271-278.

[20] A. Jolma, T. Kivioja, J. Toivonen, L. Cheng, G. Wei, M. Enge, M. Taipale, J. M. Vaquerizas, J. Yan, M. J. Sillanpa, et al., *Genome Res.* **2010**, *20*, 861–873.

[21] M. Djordjevic, A. M. Sengupta, *Phys. Biol.* **2006**, *3*, 13–28.

[22] D. Irvine, C. Tuerk, L. Gold, *J. Mol. Biol.* **1991**, *222*, 739–761.

[23] H. A. Levine, M. Nilsen-Hamilton, *Comput. Biol. Chem.* **2007**, *31*, 11–35.

[24] F. Sun, D. Galas, M. S. Waterman, *J. Mol. Biol.* **1996**, *258*, 650–60.

[25] B. Vant-Hull, L. Gold, D. A. Zichi, *Current protocols in nucleic acid chemistry / edited by Serge L. Beaucage* **2000**, Chapter 9, Unit 9 1.

[26] J. Wang, J. F. Rudzinski, Q. Gong, H. T. Soh, P. J. Atzberger, *PLoS One* **2012**, *7*, e43940.

[27] C. Daniel, Y. Roupioz, D. Gasparutto, T. Livache, A. Buhot, *PLoS One* **2013**, *8*, e75419.

[28] D. R. Latulippe, K. Szeto, A. Ozer, F. M. Duarte, C. V Kelly, J. M. Pagano, B. S. White, D. Shalloway, J. T. Lis, H. G. Craighead, *Anal. Chem.* **2013**, *85*, 3417–3424.

[29] A. Ozer, B. S. White, J. T. Lis, D. Shalloway, *Nucleic Acids Res.* **2013**, *41*, 7167–7175.

[30] B. Shui, A. Ozer, W. Zipfel, N. Sahu, A. Singh, J. T. Lis, H. Shi, M. I. Kotlikoff, *Nucleic Acids Res.* **2012**, *40*, e39.

[31] J. M. Pagano, H. Kwak, C. T. Waters, R. O. Sprouse, B. S. White, A. Ozer, K. Szeto, D. Shalloway, H. G. Craighead, J. T. Lis, *PLoS Genet.* **2014**, *10*, e1004090.

[32] F. Du, L. Guo, Q. Qin, X. Zheng, G. Ruan, J. Li, G. Li, *TrAC - Trends Anal. Chem.* **2015**, *67*, 134–146.

[33] Q. Liu, J. Shi, G. Jiang, *TrAC - Trends Anal. Chem.* **2012**, *37*, 1–11.

[34] L. Nováková, H. Vlčková, *Anal. Chim. Acta* **2009**, *656*, 8–35.

[35] K. Szeto, D. R. Latulippe, A. Ozer, J. M. Pagano, B. S. White, D. Shalloway, J. T. Lis, H. G. Craighead, *PLoS One* **2013**, *8*, e82667.

[36] J. C. Cox, P. Rudolph, A. D. Ellington, *Biotechnol. Prog.* **1998**, *14*, 845–850.

[37] B. Nastasijevic, N. A. Becker, S. E. Wurster, L. James Maher, *Biochem. Biophys. Res. Commun.* **2008**, *366*, 420–425.

# CHAPTER 3

# HIGHLY MULTIPLEXED RNA APTAMER SELECTION USING A MICROPLATE-BASED MICROCOLUMN DEVICE[2]

## 1 Introduction

The standard SELEX process tends to be time consuming and uses a large amount of reagents making aptamer selections costly.[1,2] To reduce time and reagent consumption, we developed a modified version of SELEX we called RNA APtamer Isolation via Dual-cycles (RAPID).[3] This method incorporates non-amplification cycles in which the eluted RNA is purified and used in a second binding cycle without prior amplification of the material. To compare the techniques, RAPID was performed in parallel with traditional SELEX using the same targets, and many of the same top aptamer candidates (~10% identical aptamers in the top 10,000 enriched sequences) emerged. Both selections using RAPID and traditional SELEX started with a library of $5x10^{15}$ unique sequences, yet the RAPID selection took approximately one third the time as our previously-optimized SELEX process and significantly reduced the amount of reagents used.[3] In this approach, only one non-amplification cycle was performed between amplification cycles, so it remained unknown whether multiple consecutive non-amplification cycles could be successful and further increase the number of total binding cycles possible within a given time.

Previously, we also developed a microcolumn technology designed for aptamer selections using an affinity chromatography approach. The target is immobilized on a resin packed within a microcolumn and the nucleic acid library is pumped through the column.[4] The scale of these

[2] This section has been reproduced, with modifications to conform to the required format, with permission from S. J. Reinholt, A. Ozer, J. T. Lis, H. G. Craighead, *Sci. Rep.* **2016**, *6*, 29771. Copyright and license information: http://creativecommons.org/licenses/by/4.0/.

devices (5-30 μL resin capacity) enabled a significant reduction in reagents used for a selection. Flow conditions were carefully optimized to yield the highest enrichment.[4] The microcolumns can also be connected in series allowing multiplex selections to be performed using a single aliquot of the valuable random library. In later cycles, selections are performed in parallel to avoid potential cross-contamination of enriched libraries.[4] One drawback of these devices is they are tedious to use in large-scale multiplex experiments. Consequently, we designed and fabricated a scaled-up version of the microcolumns for higher throughput. This device, named MEDUSA (Microplate-based Enrichment Device Used for the Selection of Aptamers), has all of the same benefits as the individual microcolumns, including reduced reagent consumption and the ability to be reconfigured between serial and parallel connections.[5] MEDUSA was also designed with specific dimensions and spacing that allow it to directly couple to a 96-well microplate. This enabled the continued use of high-throughput techniques in any downstream processing, which could be automated using existing liquid-handling robots. Similarly to individual microcolumn devices, we have optimized selection conditions, including protein density on the resin and flow-rates for aptamer library binding, washing, and elution, to yield the highest enrichment of previously-selected aptamers on their corresponding targets in a high-throughput manner.[5]

In this article, we describe the combination of the multi-well format MEDUSA device with the RAPID approach to simultaneously perform multiplexed aptamer selections to 19 different targets, including 4 background targets. The background targets, including maltose-binding protein (MBP), and amylose, nickel-nitrilotriacetic acid (Ni-NTA), and anti-FLAG resins, served as negative selections in the first round with the microcolumns connected in series. This allowed us to identify non-specific sequences that bound to these resins or the MBP-tag. The primary protein targets are associated with regulation of transcription, RNA stability, or localization.

Mouse DXO and its yeast homolog (Ag Rai1) are decapping exoribonucleases that possess pyrophosphohydrolase, decapping and 5′-to-3′ exoribonuclease activities, and are thus implicated in RNA surveillance.[6,7] PNPase is a mitochondrial polynucleotide phosphorylase with 3'-to-5' exoribonuclease and poly-A polymerase activities, and has recently been shown to be involved in the import of nuclear encoded RNAs into the mitochondrial matrix, thus indicating its involvement in mitochondrial homeostasis.[8,9] p23 is a molecular chaperone that initiates the disassembly of protein-DNA complexes, which impacts transcription factor activation potential and response time to environmental cues.[10] RTF1 is a subunit of the Paf1 complex (PAF1C), a multifunctional protein complex that associates with RNA Polymerase II and is implicated in histone modification, as well as transcriptional and posttranscriptional gene regulation.[11] NELF-E is a component of the NELF complex that negatively regulates elongation of transcription by RNA polymerase II.[12] The remaining 8 protein targets are various enzymes and enzymatic subunits involved in histone modification that are key features in epigenetic transcription regulation. These proteins are: TIP60-Chromo, GCN5-GNAT, MOF-Chromo, UTX-JMJC, JMJD2-Clav., JMJD2-JMJC, ASH1-BAH, and Trx-ZnFinger.

We exploited the reconfigurable design of the MEDUSA device to enable multiple selections to be performed using just 2 aliquots of library, and its microplate-based dimensions to perform downstream processes such as reverse transcription, PCR, quantitative PCR (qPCR), and transcription in a 96-well plate format. Encouraged by our earlier results with RAPID where amplification and non-amplification cycles are alternating, we decided to take our RAPID approach a step further by increasing the number of non-amplification cycles with later rounds of selection to drastically reduce the time and reagents needed for the selection. We performed the equivalent of a 10-cycle SELEX (with only 4 of these cycles requiring amplification) to 19 targets

in about the time it would take to perform 4 cycles of conventional SELEX to one target. Analysis of high-throughput sequencing results allowed us to identify potential aptamer candidates, and many of these candidate aptamers were verified to bind the corresponding target proteins with a dissociation constant ($K_D$) in the nanomolar range. In addition, further inspection of aptamer sequences enriched for DXO led us to identify aptamers that not only bind, but also resist DXO's exoribonuclease activity. Furthermore, in studies monitoring the degradation of a 30-nt RNA substrate by DXO, the aptamers demonstrated significant inhibition of DXO activity. These and the aptamers to other proteins could be used for inhibiting target protein function, and the sequencing data provides a rich source of information to understand the biology of these target proteins and perhaps uncover novel functions (e.g. RNA interaction) and specificity.

## 2 Materials and Methods

### *2.1 Target Protein Preparation*

MBP-tagged protein domains were purified from *E. coli* Bl21 (DE3) RIPL strain using amylose resin (NEB) according to the manufacturer's recommendations. Briefly, for each protein the bacteria were grown to an OD600 of 0.6 in 1 L of LB media supplemented with 100 μg/mL of ampicillin, and induced with 200 μM IPTG. Induction cultures were then incubated at 18-22°C for 12-14 hours with 250 rpm shaking. Bacteria pelleted at 3500 rpm and 4°C for 30 min were lysed in MBP buffer (20 mM Tris-Cl pH 7.5, 300 mM NaCl, 1 mM EDTA, 5 mM BME) supplemented with 0.1% NP-40, 1 mg/mL Lysozyme, and 1X Protease Inhibitor Cocktail (Roche). After a 15-min incubation on ice, lysates were sonicated on ice 3 times for 30 s each using a Branson Cell Distributor with a microtip, and cleared by centrifugation at 20,000 x g and 4°C for 45 min. Cleared lysates were applied to 1 mL of pre-washed amylose resin and allowed to pass through the column

under gravity flow. Protein-bound resin was washed 3 times with 10 mL of MBP buffer, and the proteins were eluted in five 1-mL fractions using MBP buffer supplemented with 10 mM maltose. Protein concentration was measured by Bradford Protein Assay (Bio-Rad) against a BSA standard, and the purity of each protein preparation was determined by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) and Coomassie Blue staining. One or two fractions that had the highest protein concentration (often >1 mg/ml) and were relatively pure (>80% pure estimated from Coomassie stained gel) were supplemented with 10% glycerol, aliquoted, and stored at -80°C. The maltose used for elution was not dialyzed, since most proteins precipitated out during dialysis, and re-binding of these proteins was done at very dilute conditions (>100-fold dilution, thus <0.1 mM maltose), which permitted efficient re-binding of the purified proteins to the amylose resin for the aptamer selections.

Hexahistidine (6xHis)-tagged NELF-E was purified as described previously.[12] All other target proteins were kindly provided by others: 6xHis-tagged DXO and Ag Rai1 by the Tong Lab at Columbia University, 6xHis-tagged wildtype and S448A mutant versions of PNPase by the Teitell Lab at the University of California, Los Angeles, 6xHis-tagged p23 by the Freeman Lab at the University of Illinois, and FLAG-tagged RTF1 by the Marazzi Lab at Mount Sinai Graduate School for Biomedical Sciences. Protein concentrations ranged between 0.3 mg/ml (FLAG-RTF1) and 34 mg/ml (PNPase). Supplementary Table S2 contains detailed information for all target proteins and tags used in the MEDUSA selections.

### *2.2 Target Protein Immobilization on Affinity Resins*

For each cycle of the selection, fresh protein-immobilized resins were prepared. Amylose, Ni-NTA, and anti-FLAG resins were washed thoroughly with binding buffer (25 mM Tris-Cl pH

7.5, 100 mM NaCl, 25 mM KCl, and 1 mM $MgCl_2$) prior to target immobilization. MBP-, 6xHis-, and FLAG-tagged protein targets were immobilized onto amylose, Ni-NTA, and anti-FLAG resins, respectively, by incubating the proteins at a concentration of 0.6 µg/µL of resin with the resin in a 10% slurry in binding buffer. The target resins were incubated for 1 h at 4°C with constant mixing. A list of targets and the order in which they were connected in series for the first round of selection can be found in Figure 3.1b.

### *2.3 Library and Primers*

The random RNA library (N70) used in this selection contained ~$5x10^{15}$ RNA sequences of 120 nt in length, which was transcribed *in vitro* from a DNA library that was chemically synthesized by GenScript. The RNA library consisted of a 70-nt random region flanked by two constant regions (5'-GGGAAUGGAUCCACAUCUACGAAUUC- $N_{70}$ -UUCACUGCAGACUUGACGAAGCUU-3') as described elsewhere.[13] The forward and reverse primers used for PCR, qPCR, and reverse transcription were T7pro-AptLibConsFOR, 5'-GATAATACGACTCACTATAGGGAATGGATCCACATCTACGA-3', and AptLibConsREV, 5'-AAGCTTCGTCAAGTCTGCAGTGAA-3', respectively. The forward oligo contains the T7 promoter sequence (underlined) used for transcription of the DNA template to RNA, and a shortened version of the original forward constant region to reduce primer dimer formation. These primers and other oligos, including the gBlock templates of the candidate aptamers, were purchased from Integrated DNA Technologies.

### *2.4 MEDUSA Fabrication*

A description of MEDUSA's components, as well as its fabrication, is described in detail elsewhere.[5] Briefly, a MEDUSA consisting of 48 microcolumns arranged in a 4x12 grid (Figure

3.1a) was fabricated using a $CO_2$ laser at 10.6μm (Universal Laser Systems, VersaLaser) to cut the poly(methyl methacrylate) (PMMA) and silicone layers. Each microcolumn had a 10-μL volume capacity, and fluidic connections were made using standard microfluidic connectors purchased from IDEX. This device was assembled such that it could be reversibly configured between the microcolumns connected in series or in parallel. Between rounds of selection, the device was washed extensively with 1 M KOH, 1 M HCl, and 1x binding buffer separated by DEPC water washes, and reused for all subsequent rounds.

### *2.5 RNA Aptamer Selection Process*

All solutions were degassed and introduced into the device via a standard syringe pump (Harvard Apparatus). The selection consisted of 4 rounds ending in pool amplification, with a total of 10 binding cycles by incorporating increasing numbers of non-amplification cycles in each round (Figure 3.1c). The flow parameters used in each round of the selection are summarized in Table 3.1.

**Table 3.1:** Selection Parameters

| Round | Number of Binding Cycles | Binding Step Parameters | Washing Step Parameters | Elution Step Parameters |
|---|---|---|---|---|
| 1 | 1 | 1mL at 1μL/min | 3mL at 1μL/min | 400μL at 50μL/min |
| 2 | 2 | 1mL at 10μL/min | 1mL at 1mL/min, 2mL at 70μL/min | 400μL at 50μL/min |
| 3 | 3 | 1mL at 10μL/min | 1mL at 1mL/min, 2mL at 70μL/min | 400μL at 50μL/min |
| 4 | 4 | 1mL at 10μL/min | 1mL at 1mL/min, 2mL at 70μL/min | 400μL at 50μL/min |

In the first round of selection, MEDUSA was configured with microcolumns connected in series, and two aliquots of the N70 RNA library were used for 19 total targets. The order of target proteins in each series is shown in Figure 3.1b. The target-immobilized resins were loaded into MEDUSA using a pipette, with 10 μL of resin in each microcolumn. With MEDUSA connected in parallel to prevent target contamination, the resin-packed microcolumns were washed and

primed by flowing 1 mL of binding buffer at 100 µL/min, which removed any unbound proteins from the microcolumns. The random RNA library was diluted to 1 mL in binding buffer, and heat denatured at 65°C for 5 min. The library was allowed to slowly cool to room temperature while degassing, and finally 200 units of RNase inhibitor (SUPERase•In, Ambion) were added to each library. A 10-µL sample of each library was collected and used as a standard for qPCR analysis. MEDUSA was reconnected in series for the binding step, and the libraries were pumped at 1 µL/min in the first round. MEDUSA was then reassembled into a parallel configuration for the remaining selection steps. After RNA library binding, each microcolumn was washed with 3 mL of binding buffer at 1 mL/min. To elute the protein-RNA complexes from each microcolumn, 400 µL of elution buffer [100 mM ethylenediaminetetraacetic acid (EDTA) (for Ni-NTA resin), 50 mM maltose (for amylose resin), and 100 µg/mL 3x FLAG peptide (for anti-FLAG resin) in binding buffer] were pumped at 50 µL/min, and the eluent was collected in a 96-well microplate coupled to MEDUSA. The eluted RNA was purified by phenol-chloroform extraction, and concentrated via ethanol precipitation with 1 µL of GlycoBlue (Ambion) and 40 µg of yeast tRNA (Invitrogen). The resulting pellet was resuspended in RNase-free water. This RNA pool subsequently underwent either another binding reaction (non-amplification cycle), or immediate reverse-transcription and PCR amplification.

The standards and resuspended pools were reverse transcribed using Moloney Murine Leukemia Virus Reverse Transcriptase (MMLV-RT). Residual RNA was digested using a cocktail of RNases, including RNase H, A, and T1 (Invitrogen). Approximately 3% of the cDNA from each pool was used to determine the RNA recovery compared to the initial quantity of RNA at the beginning of the round via qPCR. The remaining cDNA was PCR-amplified, and purified using DNA Clean & Concentrator spin columns (Zymo Research). A fraction of the purified DNA was

used for transcription with T7 RNA polymerase (Invitrogen) to produce the RNA pools for the next round of selection. Following transcription, the template DNA was digested using DNase I (Ambion), and the RNA was phenol-chloroform extracted and ethanol precipitated. The purified RNA pools were verified by denaturing (7M Urea) PAGE for purity and length, and quantified via Qubit BR RNA assay (Invitrogen).

In rounds 2, 3 and 4 of the selection, 50 μg of the enriched RNA pools were used, and an increasing number of non-amplification cycles were incorporated in which the eluted RNA was purified by phenol-chloroform extraction, ethanol precipitated, and directly subjected to the next binding cycle without amplification. The flowrates were increased slightly to allow multiple binding cycles to be performed within a single day (see Table 3.1).

### *2.6 High-Throughput Sequencing of Pools*

A small amount of cDNA from each final pool was PCR-amplified using primers containing unique 6-nt barcodes and the necessary adapters for the HiSeq 2500 (Illumina) sequencing platform operated in Rapid Run Mode. The sequencing data was analyzed with readily available software. Briefly, the sequencing data was quality filtered, barcode split, trimmed of constant regions and barcodes, and collapsed using programs included in the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The resulting library of sequences was first clustered using UCLUST software, part of the USEARCH software,[14] at a 90% sequence identity threshold. To identify the abundance of all aptamer clusters in every target library, the original aptamer clusters were then reclustered as before, and the total number of sequences in each cluster in every target library was tabulated. Only a representative aptamer sequence, the most abundant sequence in a given cluster, is listed for a cluster. Sequences with sufficient multiplicity (defined as: number

of sequences in a cluster / number of total sequences in the library), enrichment (defined as: abundance in the last round or cycle / abundance in an earlier round or cycle), and specificity (defined as: abundance in the target-specific library / abundance in all other libraries) were identified as probable aptamer candidates.

### *2.7 Motif Search using MEME*

To identify motifs, 5-20 nt RNA sequences that were enriched in a selected RNA pool and present in the top 3000 aptamer clusters were subjected to analysis by the MEME program of the MEME Suite (http://meme-suite.org/).[15,16] The E-value, representing the statistical significance of the identified motif, is estimated by MEME as the log likelihood ratio of finding an identical motif with the same width and site count in a similarly sized set of randomized sequences with identical base composition. Small E-values (up to 0.05) are considered significant, with smaller ones being the most significant. Motifs identified by MEME for target proteins were found to be specific and not identified in other target pools.

### *2.8 Binding Affinity Determination via EMSA*

To test the binding capability of the aptamer candidates and determine the dissociation constant, electrophoretic mobility shift assays (EMSAs) were performed for each candidate. The RNA aptamer candidate was fluorescently labeled at the 3' end with fluorescein isothiocyanate (FITC).[17] With the quantity of RNA constant, a series of protein concentrations were incubated with the labeled aptamer candidate for 60 minutes. These samples were then run at 4°C on an agarose gel (1.5% agarose, 0.5x TBE, 1 mM $MgCl_2$), and imaged using a fluorescence scanner. The gel images were quantified using ImageJ, and the resulting curves were fitted to the Hill

Equation using Igor Pro 5.04A to determine the $K_D$. For many of the aptamers, a second EMSA was performed, and the data from both experiments were merged and fitted to the Hill Equation.

### *2.9 Analysis of DXO Aptamers Resisting Exoribonuclease Activity and Acting as Inhibitors*

Body-labeled RNA aptamers were generated by *in vitro* transcription with T7 RNA polymerase and [α-$P^{32}$]UTP (Perkin Elmer), and purified by RNase-free Micro Bio-Spin P-30 columns (Bio-Rad). Exoribonuclease assays were performed as described elsewhere.[7] Briefly, radiolabeled RNA aptamers (10,000 cpm in RNase-free DXO buffer: 50 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-Cl pH 8.0, 0.5 mM DTT, 0.025 mg/mL UltraPure BSA) were first heat-denatured at 70°C for 5 min and renatured at room temperature for 15-30 min. They were then incubated with 1 μM recombinant DXO enzyme (in DXO buffer) for various durations at 37°C. For testing the inhibition of DXO 5'-3'exoribonuclease activity by RNA aptamers, DXO was first equilibrated with RNA aptamers at room temperature for 5 min. 3'-Cy5-labeled 30-nt RNA substrate (Dharmacon)[24] was then added, and the reaction was incubated at 37°C for 2 hours. The final reactions contained 1 μM DXO, 20 nM 3'-Cy5-labeled RNA substrate, and various concentrations of DXO RNA aptamers (0-1100 nM) in DXO buffer. After incubation, each reaction was mixed with an equal volume of 2X Gel Loading Buffer (Ambion), heat denatured at 75°C for 5 min, and run on 8 or 10% denaturing polyacrylamide gels. After gel running, the gel was either dried on Whatman paper, exposed to a Phosphorimager screen, and the screen was imaged using a Typhoon 9400 scanner, or the gel was EtBr stained and imaged with a Typhoon 9400 scanner. Quantification of these images was performed using ImageJ software, and the data analysis was carried out either in Microsoft Excel or in GraghPad Prism6 software.

# 3 Results and Discussion

## *3.1 MEDUSA as a Platform for Highly Multiplexed Aptamer Selections*

MEDUSA was designed for high-throughput aptamer selections and characterizations of the SELEX process using a microplate-based format allowing high-throughput downstream processes to be performed. This standardized format greatly increases the speed at which these processes are performed, and allows for potential automation of the aptamer selection process through the use of existing liquid-handling robots.[13,18] Multiplexing is especially beneficial with aptamer selections because it increases the likelihood of success per selection, where low success rates (30%) have been previously reported.[1,19] The SELEX process is also very time-consuming, often taking several weeks to months to complete,[2] so performing many selections simultaneously is extremely advantageous.

MEDUSA is capable of connecting interchangeably between parallel and serially-connected microcolumns during a selection (Figure 3.1a). This ability provides additional benefits, including the incorporation of in-line negative or counter selections and efficient use of the starting library. Here, just two aliquots of the random RNA library was used in selections to 19 different targets (Figure 3.1b), which saved significant cost as the library is very expensive to synthesize. Additionally, performing aptamer selections using a microcolumn approach significantly reduces reagent consumption, most importantly the amount of target needed. This is beneficial when the target might only be available in limited quantity, especially for proteins that cannot be generated recombinantly in bacteria, but instead are purified from a mammalian cell line (e.g. FLAG-RTF1). Each cycle in these selections required just 6µg of protein, which allowed selections to be carried out on protein targets that were difficult to purify in large quantities. Therefore, MEDUSA can

provide advantages over other techniques that help to overcome some of the limitations frequently observed during aptamer selections.

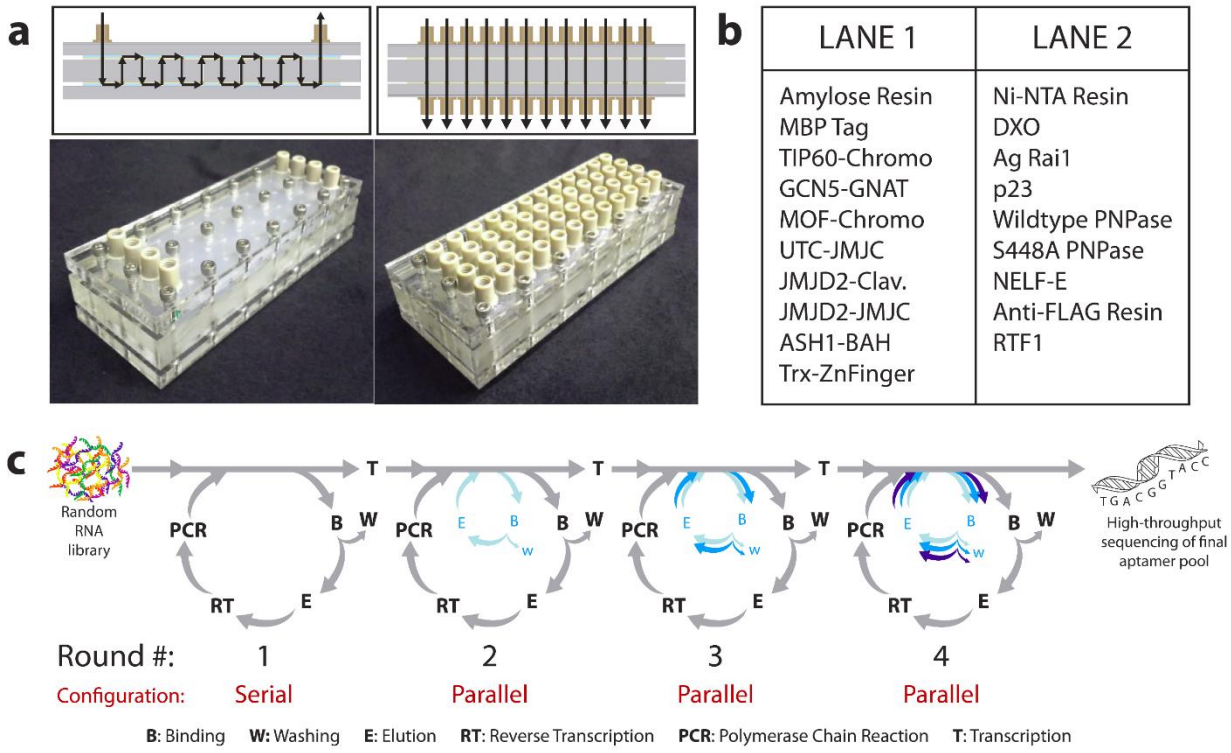



**Figure 3.1:** Schematics of the selection device and selection process. (a) The schematics describe the fluid path through the MEDUSA device in serial (left) and parallel (right) configurations, while the lower panels show images of the device in each configuration. (b) The list of targets in the order in which they saw the initial RNA library in a serial configuration. All the proteins in Lane 1 have an N-terminal MBP-tag, and in Lane 2 an N-terminal 6x His-tag, except RTF1, which has an N-terminal FLAG-tag for immobilization on amylose, Ni-NTA, and anti-FLAG resin, respectively. (c) The aptamer selection process beginning with a random RNA library, proceeding with 4 rounds of SELEX incorporating increasing numbers of non-amplification cycles prior to an amplification cycle, and ending with high-throughput sequencing of the enriched RNA pools. Ten cycles of selection were performed, starting at Round #1-Cycle #1 (R1C1) and ending at Round #4-Cycle #4 (R4C4).

### *3.2 Incorporation of Multiple Non-Amplification Cycles to Improve Selection Efficiency*

Because the conventional SELEX process is time consuming, many approaches have been proposed to decrease the time required for a selection.[1,2] RAPID-SELEX was previously

developed by our group, which incorporates a single non-amplification cycle between cycles that include amplification of the enriched aptamer pool.[3] Selections using RAPID-SELEX identified RNA aptamers in one third of the time compared to conventional SELEX. To further exploit the benefits of this process, additional non-amplification cycles were incorporated during the selections presented here. The selection consisted of 4 rounds ending in pool amplification, with a total of 10 binding cycles by incorporating increasing numbers of non-amplification cycles in each round (Figure 3.1c). This allowed 10 binding cycles to be carried out in the same amount of time as 4 cycles of conventional SELEX, and with our process this took less than 4 weeks. This also dramatically reduced the cost of the selections by significantly reducing the amount of reagents used for reverse transcription, PCR amplification, and *in vitro* transcription. Including non-amplification cycles was also found to lead to a faster sequence convergence over traditional SELEX.[3]

### *3.3 qPCR and High-throughput Sequencing Results*

Quantitative PCR was used to measure the amount of RNA recovered after each binding cycle relative to the amount of RNA in the starting pool for each round. Figure 3.2 shows qPCR results measuring the percent recovery of RNA after each of the 4 cycles in Round 4 relative to initial concentration of RNA at the beginning of the round for all of the targets in this selection. All 19 targets showed the same trend with decreasing RNA recovery after each cycle. These results indicate that despite multiple consecutive binding cycles without amplification, a sufficient amount of RNA was recovered to warrant continuation of the selections. Since this trend was observed for every target, it can be concluded that this modified SELEX process can likely be used for any target. Furthermore, the targets with the highest affinity aptamers identified from the

selection (DXO, TIP60-Chromo, and NELF-E) showed the highest RNA recovery during Round 4, suggesting that sequences with high affinity for these targets were being enriched.

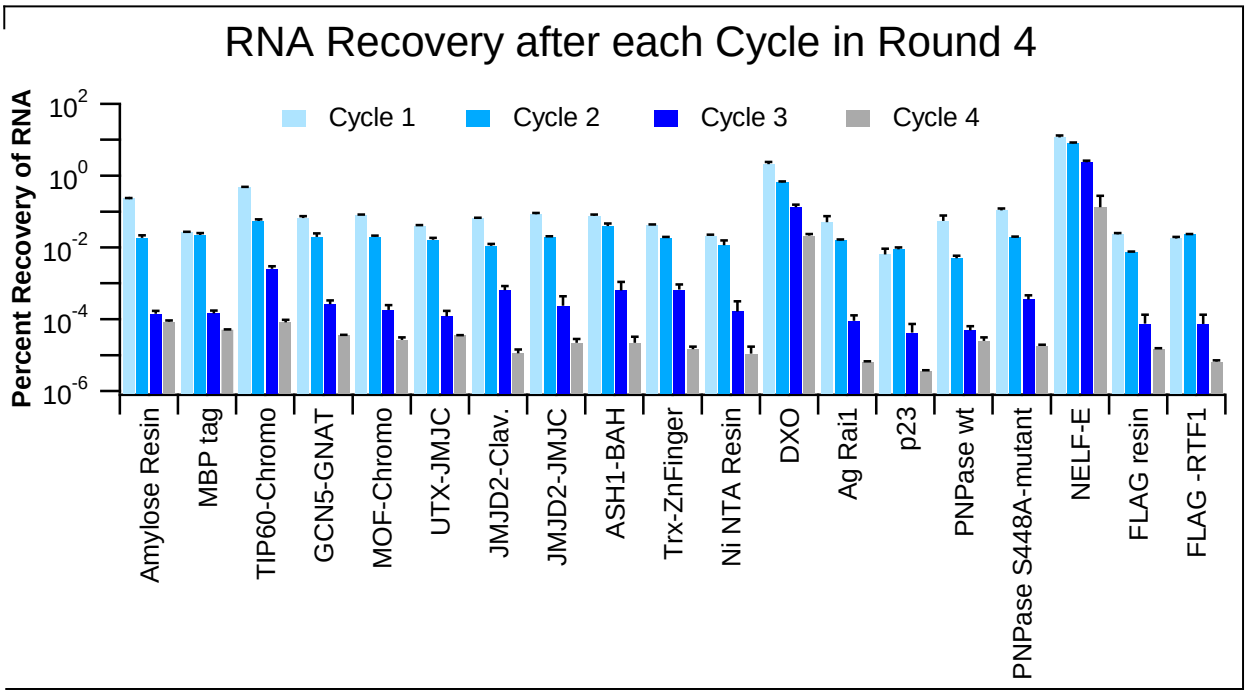



**Figure 3.2:** qPCR results indicating the percent recovery of RNA after each binding cycle in Round 4 relative to the initial RNA concentration at the beginning of the round. The error bars represent the standard deviation for triplicate measurements.

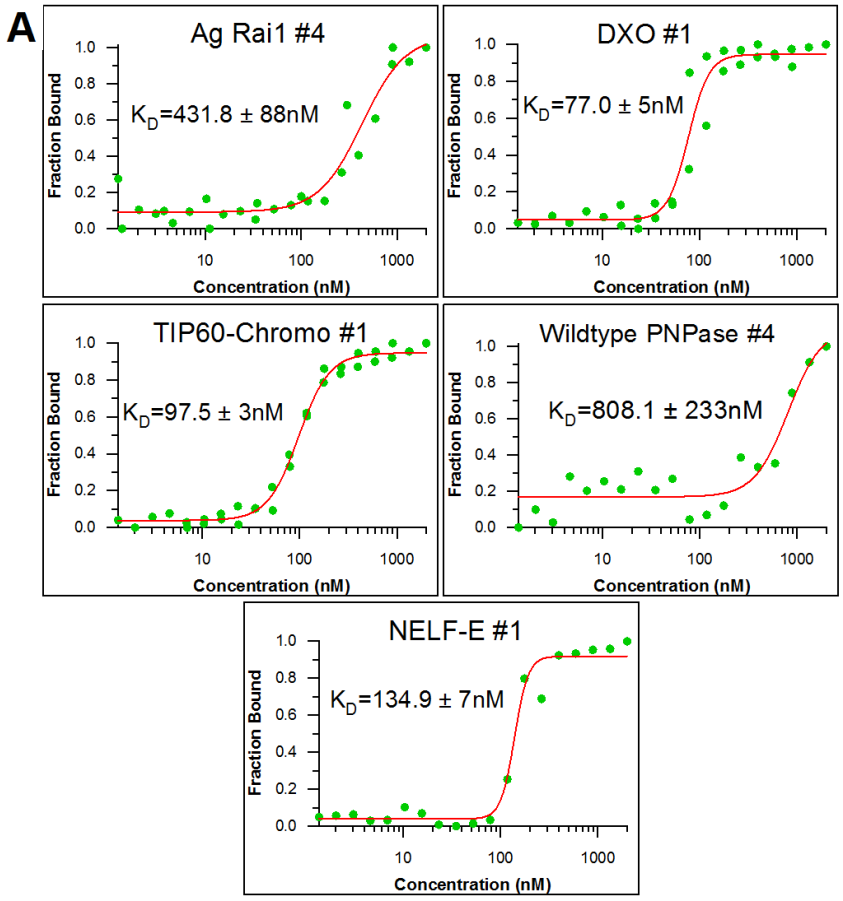
To identify candidate aptamers for each target protein/domain and assess their enrichment, the final RNA pools [Round #4-Cycle #4 (R4C4)] and the pools after Round #3-Cycle #3 (R3C3) for every target were analyzed via high-throughput sequencing. All intermediate pools for four of the targets were also sequenced (Appendix A, Table A1). This way, the presence of specific sequences can be monitored throughout the selection process, and help us better characterize the selection process with the MEDUSA device and modified RAPID-SELEX protocol. Unfortunately, of the four targets that were chosen for detailed analysis by high-throughput sequencing only one target, TIP60-Chromo, yielded sufficient aptamer convergence in the final and earlier rounds to be analyzed. Contrary to earlier reports, we did not observe that certain

aptamers enriched in earlier cycles are replaced by other aptamers in later cycles (Appendix A, Figure A1).[20] This difference could be a direct result of PCR-induced bias in conventional SELEX methods, where amplification of enriched pools is carried out after every cycle. Our RAPID-SELEX method greatly reduces the potential for such a bias, since amplification of enriched pools were carried out minimally, just 4 times in a 10-cycle selection.

Upon examination of the high-throughput sequencing results, sequences that could not be identified as an aptamer candidate (i.e. the constant regions were missing) were discarded, though a few still made it through our analysis pipeline (Appendix A, Figure A1). Sequences with 90% sequence similarity were grouped into clusters. A total of 40 aptamer candidates for 8 of the 15 protein targets were selected for further testing of specific binding by EMSA. Five candidates were chosen for each target protein/domain based on having i) the highest multiplicity in the final RNA pool, ii) the highest enrichment from the R3C3 to the R4C4 pools, and iii) the highest specificity, in that these sequences were not highly abundant in any other RNA pool. The other 7 protein targets did not yield any candidate aptamers that were deemed appropriate for further testing, which would be sequences that had both high multiplicity and high specificity after 10 cycles. The selection was arbitrarily stopped after 4 rounds and 10 cycles because this was a proof-of-concept selection to demonstrate MEDUSA and the new RAPID approach. Additional cycles could be quickly performed, and would likely yield aptamers to more of these targets, as well as aptamers with improved affinity. The sequencing results highlight another advantage of multiplexing in that background binding or non-specific sequences can be easily identified and discarded from the pool of potential aptamer candidates. Previously, sequences that bind to the device and resins have been identified,[5] so these non-specific and other cross-reacting sequences can be identified and eliminated to save time and cost.

### *3.4 Aptamer Candidate Testing*

EMSAs were performed to test the binding of candidate aptamers to their respective targets and determine the dissociation constants. Results for 5 of the aptamer candidates can be seen in Figure 3.3a, as well as dissociation constants for all of the aptamer candidates tested that showed binding using this technique (Figure 3.3b). A total of 25 aptamers were found for 5 different targets




B

| Aptamer | $K_D$ (nM) | Rank in Final Pool |
|---|---|---|
| Ag Rai1 #2 | 669.0 ± 56 | 2 |
| Ag Rai1 #3 | 519.1 ± 387 | 3 |
| Ag Rai1 #4 | 431.8 ± 88 | 4 |
| Ag Rai1 #5 | 643.5 ± 177 | 5 |
| DXO #1 | 77.0 ± 5 | 1 |
| DXO #2 | 96.4 ± 9 | 2 |
| DXO #3 | 113.1 ± 4 | 3 |
| DXO #4 | 86.9 ± 6 | 4 |
| DXO #5 | 104.5 ± 8 | 5 |
| NELF-E #1 | 134.9 ± 7 | 1 |
| NELF-E #2 | 254.7 ± 14 | 2 |
| NELF-E #3 | 215.6 ± 6 | 3 |
| NELF-E #4 | 209.3 ± 6 | 4 |
| NELF-E #5 | 214.5 ± 21 | 5 |
| TIP60 #1 | 97.5 ± 3 | 1 |
| TIP60 #2 | 127.0 ± 16 | 2 |
| TIP60 #3 | 102.5 ± 10 | 3 |
| TIP60 #4 | 131.6 ± 15 | 4 |
| TIP60 #5 | 97.6 ± 11 | 5 |
| TIP60 #6 | 107.1 ± 27 | 12 |
| TIP60 #7 | 82.2 ± 5 | 18 |
| TIP60 #8 | 53.8 ± 9 | 21 |
| TIP60 #9 | 40.3 ± 13 | 36 |
| TIP60 #10 | 104.2 ± 5 | 352 |
| wtPNPase #4 | 808.1 ± 233 | 4 |

**Figure 3.3:** Binding analysis of candidate aptamer sequences. (a) Measurement of dissociation constants for aptamers to 5 different protein targets by EMSA experiments. The data were fitted to the Hill Equation to determine the $K_D$. (b) A list of the dissociation constants for all of the aptamer candidates tested and their multiplicity ranking in their respective final (R4C4) pools. A complete list of the top 3,000 aptamer clusters for final pools of each target is provided in Supplementary Dataset 1.

with dissociation constants in the nanomolar range (EMSA analyses shown in Appendix A, Figure A2), and it is likely that further optimization of the aptamer sequence would improve their affinity.

The copy numbers (multiplicity) of the selected aptamer candidates varied greatly, and did not predict whether the aptamer would bind nor did they correlate with the determined $K_D$ (Figure 3.3b). There were highly abundant sequences that did not bind to their targets, and sequences with only a few copies that did bind, so it is possible that more aptamers exist in the selected pools. From these results, we conclude that sequencing will indicate what sequences are present in the pool and their abundance, but it will not necessarily predict binding or affinity. High-throughput binding assays of target proteins to all aptamers could be used to provide a more comprehensive identification of the highest affinity aptamers in a selected pool.[21,22]

### *3.5 Studying Protein Function through Aptamer Selection*

To gain further insight on the selection process and how it could be used to study the biological function of a target protein, as well as provide useful tools that do more than just bind, we turned our focus to the DXO protein, which has 5'-3' exoribonuclease and is implicated in RNA surveillance.[6,7] To identify short RNA motifs that are enriched in DXO candidate aptamers, we subjected the random regions of the top 3,000 most abundant aptamer clusters from the DXO R4C4 pool to MEME analysis.[15,16] This analysis led to identification of DXO Motif 1 (5'-GGATCCC-3'), which showed significant enrichment (E-value: 2.3e-7,360), was present on almost every candidate aptamer (2,932 out of 3,000), and showed high-specificity for DXO in that it was not identified in similar MEME analysis of other target pools (Figure 3.4a). The specificity of DXO aptamers containing this motif was also tested, and these aptamers showed no appreciable binding to non-specific targets (Appendix A, Figure A3). This motif is similar to the BamHI restriction site, which is included in the forward constant region of the N70 RNA library (see Materials and Methods). The motif is found to base-pair with the BamHI site in secondary structure

predictions giving rise to a stable stem structure close to the 5'-end of the RNA aptamer (Figure 3.4b and Appendix A, Figure A4). We reasoned this may hinder the 5'-3' exoribonuclease activity of DXO. We tested this hypothesis by incubating radiolabeled RNA aptamers (5 DXO candidate aptamers and the N70 RNA library) with DXO at 37°C in DXO exoribonuclease buffer, and monitoring the stability of full-length RNA via denaturing PAGE analysis. As shown in Figure 3.4c, all five candidate DXO aptamers showed resistance against DXO exoribonuclease activity, with DXO #3, #4, and #5 showing the strongest resistance, when compared to the N70 RNA library control. Using a finer time-course for these tests, we estimated the half-lives of the DXO candidate aptamers to be ~1.5-3.5-fold higher (28.9-69.3 min) than that of the N70 library (19.8 min) (Appendix A, Figure A5). DXO has significantly lower exoribonuclease activity in SELEX buffer compared to DXO exoribonuclease buffer (data not shown), perhaps explaining why these aptamers were enriched and survived incubations with DXO for as long as 17 hours.

To test whether the determining factor for RNA aptamer resistance against DXO exoribonuclease activity is the sequence of the motif itself or just the stability of the first stem in the aptamer structure, we took the DXO #2 and #5 aptamers with perfect matches to DXO Motif 1, and generated mutant aptamers. In mut1 aptamers, the sequence of the motif was altered but the stability was maintained by changing the original motif (5'-GGATCCC-3' to 5'-CCTAGGC-3') and the BamHI site in the forward constant region (5'-GGATCC-3' to 5'-CCTAGG-3'). In mut2 aptamers, the stability of the first stem was reduced by changing the original motif to 5'-AATATTC-3' and the BamHI site in the forward constant region to 5'-AATATT-3'. In addition, we generated truncations of DXO #2 and #5 aptamers, which retained the DXO Motif 1 and the first stem-loop structure for each aptamer. A sequence-shuffled version of DXO #5 trunc was used as a control. As shown in Appendix A, Figure A6, these mutant and truncated aptamers are

predicted to adopt very similar, if not identical, secondary structures to the original aptamers. In the DXO exoribonuclease assay, the mut1 and truncated aptamers showed nearly identical stabilities compared to the original full-length RNA aptamers, whereas the mut2 aptamers showed enhanced susceptibility to degradation by DXO, similarly to that of the N70 aptamer library and DXO #5 shuffle control (Appendix A, Figure A6). These results highlight the importance of an RNA molecule's 5'-end stability in determining its susceptibility to DXO-mediated degradation, which sheds light on the *in vivo* function of DXO and can help to identify *in vivo* targets of DXO exoribonuclease activity.

Finally, we tested whether these RNA aptamers could inhibit the exoribonuclease activity of DXO. We measured the exoribonuclease activity on a 30-nt RNA substrate that was 3'-Cy5 labeled. DXO was pre-incubated with a range of concentrations (0-1100 nM) of full-length RNA aptamers, and the RNA substrate was subsequently added. As shown in Figure 3.4d, even in the presence of 1.1 μM N70 RNA library, the Cy5-labeled substrate was almost completely degraded. Conversely, a substantial amount of RNA substrate was still intact after 2 hours of incubation with 1 μM DXO in the presence of 730 nM DXO #2, 330 nM DXO #4, and 490 nM DXO #5 aptamer. By fitting these data to a dose-response curve, we calculated >>1100 nM, ~610 nM, ~290 nM, and ~570 nM IC50 values for the N70 RNA library, DXO #2, DXO #4, and DXO #5 aptamers, respectively (Appendix A, Figure A7). In these samples, we have verified equal loading of different aptamers and the library, and they remained largely intact (Appendix A, Figure A7). Truncated versions of DXO #2 and #5 aptamers did not inhibit DXO exoribonuclease activity (Appendix A, Figure A8). These results suggest that the full-length DXO aptamers, but not the truncated versions, inhibit the exoribonuclease activity directly, not simply by competing off the RNA substrate.

Given their moderately-high affinity and DXO exoribonuclease inhibitory effect, these full-length aptamers could potentially be used as DXO inhibitors for studying DXO function *in vivo*. In the absence of a highly specific small molecule inhibitor of DXO, DXO-specific inhibitory RNA aptamers, which can in principle be expressed transgenically at specified levels, time points, and durations,[23,24] enable the functional study of DXO exoribonuclease activity *in vivo*. Likewise, compared to RNAi knockdowns, inhibitory RNA aptamers offer unique advantages for the following *in vivo* functional studies: i) a specific function or interaction of the target protein (e.g. exoribonuclease activity of DXO) could be targeted rather than complete elimination of the entire protein as with RNAi, and ii) distinction between primary and secondary effects is possible, since inhibitory concentrations of RNA aptamers could be attained relatively quickly (i.e. minutes to tens of minutes) upon induction, whereas significant knockdown of most proteins by RNAi requires hours to days.

These studies have shown that different properties of the target proteins can also be discovered and studied through an aptamer selection. Furthermore, selecting sequences for more than just affinity, such as nuclease resistance and inhibition, can be accomplished. Some of the targets in this selection possess nuclease activity, which affected the affinity testing, and likely the selection process itself, as sequences with a higher resistance to the nuclease activity would be selected over those that may have higher affinity. DXO possesses 5'-3' exoribonuclease activity, and we have shown that the selected aptamers resist degradation significantly longer than the random RNA library (Figures 3.4a and 3.4b). We also demonstrated that this resistance is due to the physical structure of the aptamer with a stem loop at the 5' end. This type of structure was also shown to inhibit the exoribonuclease activity against an RNA substrate meant to represent nascent

RNA (Figure 3.4c). These results demonstrate an additional benefit and application of the SELEX process.

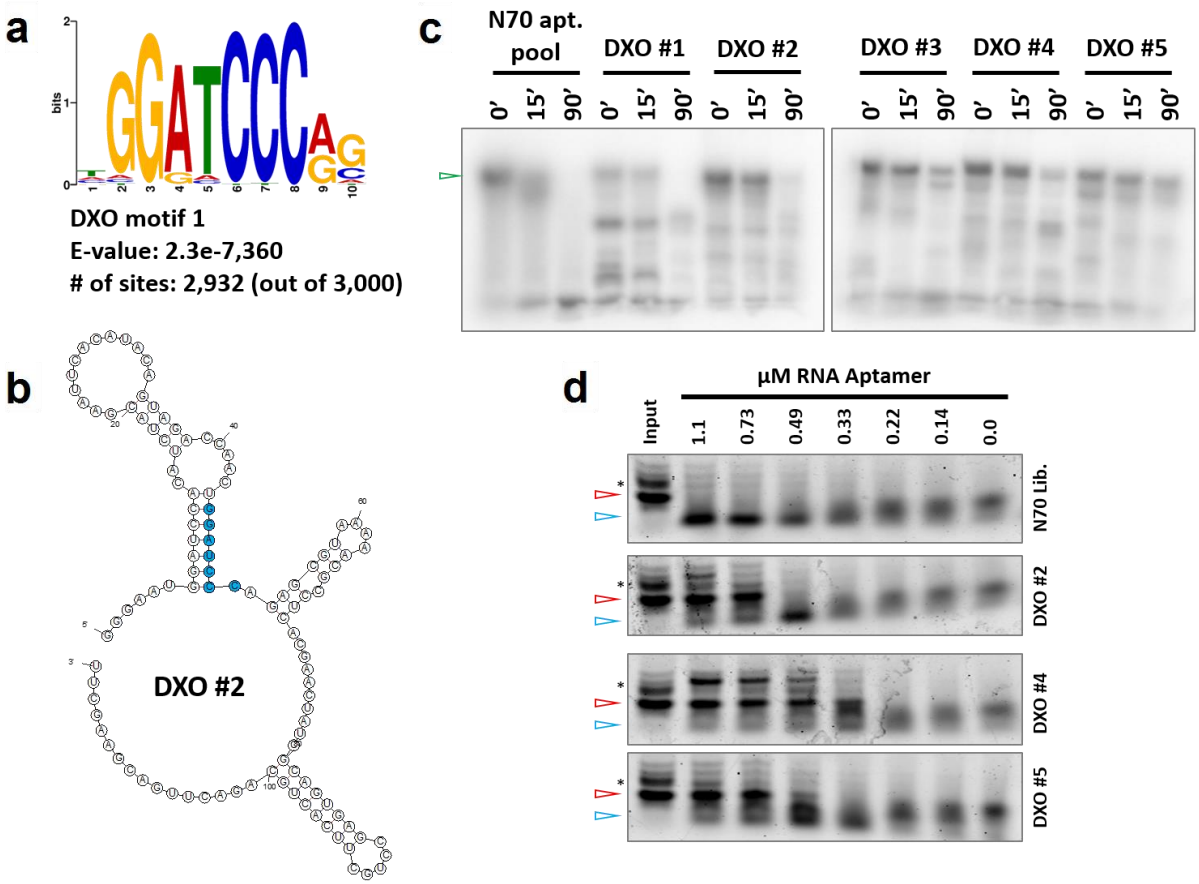



**Figure 3.4:** Characterization of DXO aptamers. (a) A highly specific DXO motif identified by MEME[15,16] among the top 3,000 most abundant DXO aptamer clusters from R4C4. The E-value, representing the statistical significance of the identified motif as calculated by the MEME suite (http://meme-suite.org/), and number of aptamer clusters containing a match to Motif 1 are indicated. E-values lower than 0.05 are considered significant. (b) Secondary structure of DXO #2 aptamer, a representative aptamer of the second highest multiplicity cluster for DXO from the R4C4 pool, as predicted by the mFold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form). DXO Motif 1 is colored cyan. Predicted secondary structures of other DXO aptamers are provided in Supplementary Fig. S3. (c) Stability of RNA aptamers against DXO exoribonuclease activity. Radiolabeled specific RNA aptamers or the N70 RNA library were incubated with DXO enzyme at 37°C for the indicated times and separated using 8% denaturing PAGE. Full-length RNA aptamers are indicated by ▻. (d) Inhibition of DXO exoribonuclease activity by RNA aptamers. Degradation of a 3'-Cy5-labeled 30-nt RNA substrate by DXO in the absence and presence of indicated amounts of RNA aptamers as determined by 10% denaturing PAGE analysis. Full-length, intact RNA substrate and degradation product are indicated by ▻ and ▻, respectively. The band indicated by * above the full-length RNA is likely incompletely deprotected RNA substrate (2'-ACE protected, Dharmacon).

# 4 Conclusions

We have demonstrated that a highly multiplexed aptamer selection can be performed using the MEDUSA device, and its reconfigurable microcolumn design allows for substantial reduction in reagent consumption. Most notably, a single aliquot of the starting random library can be used to select aptamers to multiple different targets, and just 6μg of protein target is needed for each cycle. Additionally, a modified SELEX process incorporating increasing numbers of consecutive non-amplification cycles significantly reduced the time required to perform the selections. Most importantly, high-throughput aptamer selections are particularly beneficial because they increase the rate of success per selection.

High-throughput sequencing allows for the identification of top sequences, as well as their multiplicity in the pool. Performing multiplex aptamer selections is also beneficial when analyzing sequencing results, since background binding and non-specific sequences can be recognized when they appear in multiple pools with significant abundance. However, after identifying potential aptamer candidates and performing binding experiments, it was obvious that sequence multiplicity could not predict binding or the affinity of the aptamer.

Finally, through these selections, we demonstrated that useful tools, other than high-affinity ligands, can be discovered. Aspects of a target protein's function were studied, and aptamers capable of resisting exoribonuclease activity and acting as inhibitors were determined. Therefore, aptamer selections can produce very useful tools that can be used in many different applications, and the device and selection process presented here provides advantages over other processes that improve the efficiency and success of aptamer selections.

## 5 References


[1] A. Ozer, J. M. Pagano, J. T. Lis, *Mol. Ther. Nucleic Acids* **2014**, *3*, e183.

[2] K. Szeto, H. G. Craighead, *Appl. Phys. Rev.* **2014**, *1*, 031103–1–031103–17.

[3] K. Szeto, D. R. Latulippe, A. Ozer, J. M. Pagano, B. S. White, D. Shalloway, J. T. Lis, H. G. Craighead, *PLoS One* **2013**, *8*, e82667.

[4] D. R. Latulippe, K. Szeto, A. Ozer, F. M. Duarte, C. V Kelly, J. M. Pagano, B. S. White, D. Shalloway, J. T. Lis, H. G. Craighead, *Anal. Chem.* **2013**, *85*, 3417–3424.

[5] K. Szeto, S. J. Reinholt, F. M. Duarte, J. M. Pagano, A. Ozer, L. Yao, J. T. Lis, H. G. Craighead, *Anal. Bioanal. Chem.* **2014**, *406*, 2727–2732.

[6] X. Jiao, J. H. Chang, T. Kilic, L. Tong, M. Kiledjian, *Mol. Cell* **2013**, *50*, 104–15.

[7] S. Xiang, A. Cooper-Morgan, X. Jiao, M. Kiledjian, J. L. Manley, L. Tong, *Nature* **2009**, *458*, 784–8.

[8] G. Wang, H. Chen, Y. Oktay, J. Zhang, E. L. Allen, M. Geoffrey, K. C. Fan, J. S. Hong, S. W. French, J. M. Mccaffery, et al., *Cell* **2010**, *142*, 456–467.

[9] G. Wang, E. Shimada, C. M. Koehler, M. A. Teitell, *Biochim. Biophys. Acta* **2012**, *1819*, 998–1007.

[10] E. Zelin, Y. Zhang, O. A. Toogun, S. Zhong, B. C. Freeman, *Mol. Cell* **2012**, *48*, 459–470.

[11] Q.-F. Cao, J. Yamamoto, T. Isobe, S. Tateno, Y. Murase, Y. Chen, H. Handa, Y. Yamaguchi, *Mol. Cell. Biol.* **2015**, *35*, 3459–3470.

[12] J. M. Pagano, H. Kwak, C. T. Waters, R. O. Sprouse, B. S. White, A. Ozer, K. Szeto, D. Shalloway, H. G. Craighead, J. T. Lis, *PLoS Genet.* **2014**, *10*, e1004090.

[13] J. C. Cox, P. Rudolph, A. D. Ellington, *Biotechnol. Prog.* **1998**, *14*, 845–850.

[14] R. C. Edgar, *Bioinformatics* **2010**, *26*, 2460–2461.

[15] T. L. Bailey, M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li, W. S. Noble, *Nucleic Acids Res.* **2009**, *37*, 202–208.

[16] T. Bailey, C. Elkan, *Proc. Second Int. Conf. Intell. Syst. Mol. Biol.* **1994**, 28–36.

[17] J. M. Pagano, C. C. Clingman, S. P. Ryder, *RNA* **2011**, *17*, 14–20.

[18] J. C. Cox, A. D. Ellington, *Bioorganic Med. Chem.* **2001**, *9*, 2525–2531.

[19] L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, et al., *PLoS One* **2010**, *5*, e15004.

[20] T. Schütze, B. Wilhelm, N. Greiner, H. Braun, F. Peter, M. Mörl, V. A. Erdmann, H. Lehrach, Z. Konthur, M. Menger, et al., *PLoS One* **2011**, *6*, 1–10.

[21] J. M. Tome, A. Ozer, J. M. Pagano, D. Gheba, G. P. Schroth, J. T. Lis, *Nat. Methods* **2014**, *11*, 683–688.

[22] J. D. Buenrostro, C. L. Araya, L. M. Chircus, C. J. Layton, H. Y. Chang, M. P. Snyder, W. J. Greenleaf, *Nat. Biotechnol.* **2014**, *32*, 562–568.

[23] H. Shi, B. E. Hoffman, J. T. Lis, *Proc Natl Acad Sci U S A* **1999**, *96*, 10033–10038.

[24] H. H. Salamanca, N. Fuda, H. Shi, J. T. Lis, *Nucleic Acids Res.* **2011**, *39*, 6729–6740.

# CHAPTER 4

# MICROFLUIDIC DEVICE FOR APTAMER-BASED CANCER CELL CAPTURE AND GENETIC MUTATION DETECTION

## 1 Introduction

Cancer cells contain genetic mutations that allow them to escape the regulatory processes necessary for the healthy function of tissues and organs.[1-3] Moreover, there are numerous mechanisms for malignancy each with different combinations of genetic mutations, and cancer cells are constantly evolving,[4] which makes cancer treatment difficult with varying levels of efficacy. Therefore, developing an assay that can identify a cancer patient's genetic mutations would provide a platform for more specific and detailed diagnoses, which would allow clinicians to better prescribe the most effective treatment strategy.

Many assays have been developed that detect specific mutations, while other methods have been developed that detect all mutations via sequencing.[5-10] Each of these approaches has advantages and disadvantages.[8] However, most of these assays also require significant sample preparation performed in a bulk solution where the initial amount of genetic material is limited, some is lost in processing, and the remaining material is used up quickly. Therefore, an assay that incorporates sample preparation and can perform several types of analysis on the same small population of cells would be highly advantageous.

To incorporate sample preparation into a cancer diagnostic, the first step is to selectively filter out cancer cells from a sample. Assays and devices are being developed that capture and isolate circulating tumor cells (CTCs), and although antibodies are traditionally used, aptamer-based CTC capture is also being performed.[11-16] Aptamers have been used over antibodies in these

types of applications because of their increased robustness and ease of functionalization and oriented immobilization. With the advancement of cell-SELEX, aptamers to cell surface species in their native state will continue to be discovered. This will further progress the use of aptamers in cell capture devices, and due to their reproducible and relatively inexpensive synthesis, aptamer-based diagnostic devices provide significant reliability and cost advantages over antibody-based systems.

One approach to enabling multiple genetic analyses to be performed on a single small population of cells is to retain the original genomic template between processes. We previously developed a device for extracting and purifying human chromosomal DNA from lysed cells.[17] This device incorporated a fine micropillar array that captured megabase-long genomic DNA (gDNA) strands via physical entanglement. The physical nature of this isolation enables the gDNA to be isolated without dependence on biochemical or electrostatic forces, making it available for any downstream reactions. This also allows the gDNA to remain on the micropillar array during flow, which allows multiple downstream analyses to be performed within the microdevice.

Here, we have designed and developed a device that incorporates both cancer cell capture and gDNA isolation and retention. With the gDNA isolated, specific genes are amplified for sequencing to determine any genetic mutations in those genes. The cancer cells are captured using aptamers immobilized on the microchannel surface, and the gDNA is isolated via physical entanglement within a micropillar array. In addition, we developed a modified version of multiple displacement amplification (MDA), an isothermal amplification technique, that amplifies a specific gene of interest. This amplification product undergoes sequencing, and the resulting sequence is compared to the known human genome to determine the presence of any genetic mutations. We have used our device to capture HeLa and CAOV-3 cancer cells using aptamers, as

well as subsequent isolation of their gDNA. We have also demonstrated on-chip amplification of the *TP53* gene, which is a gene of interest in cancer, from HeLa and CAOV-3 cells and sequenced a fragment of this gene for genetic mutation detection. We were able to identify the presence of a point mutation in the CAOV-3 *TP53* gene. This technology enables the identification of specific key mutations in a timely and cost-effective manner. Using this device to provide a more detailed diagnosis for each cancer patient would allow clinicians to better prescribe the most effective treatment options. Furthermore, since cancer cells are constantly evolving, regular testing of multiple important genes is beneficial for monitoring disease progression and determining future treatment, which could also be accomplished using our device.

## 2 Materials and Methods

### *2.1 Cell Culture and Buffers*

HeLa and CAOV-3 cells were purchased from American Type Cell Culture (ATCC). Both cell lines were cultured in DMEM media (Life Technologies) supplemented with non-essential amino acids, 110 mg/L sodium pyruvate, 200mM L-glutamine (Life Technologies), 1x Pen Strep (Life Technologies), 26.8 mM HEPES, betamercaptoethanol, and containing 10% fetal bovine serum. Prior to use, the cells were trypsinized and resuspended in phosphate-buffered saline (PBS) binding buffer [1x Dulbecco's PBS with calcium chloride and magnesium chloride (Fisher Scientific) containing 4.5 g/L glucose and 5 mM $MgCl_2$]. HeLa cells contained GFP-conjugated histones and CAOV-3 cells were stained with calcein-AM (Thermo Fisher Scientific) that allowed them to be observed using fluorescence microscopy.

## *2.2 Device Fabrication*

The devices used here were polydimethylsiloxane (PDMS) microchannels bonded to glass substrates. The devices consisted of two orthogonal microchannels that contained two micropillar arrays: a cell capture array at the intersection of the two microchannels and a gDNA isolation array downstream of the cell capture array (see Figure 4.1). The PDMS channels were made via soft lithography using a silicon master mold. The master mold was fabricated from a 4" silicon wafer via standard photolithography. Microposit S1813 photoresist (Shipley) was spun on silicon wafers and exposed to UV light using a contact mask aligner (ABM). The exposed wafer was developed using 726MIF developer (Microchemicals) and the pattern was etched 25 μm into the silicon via Bosch process using a Unaxis SLR 770 deep reactive ion etching system (Unaxis USA Inc.). A monolayer of (1H,1H,2H,2H-Perfluorooctyl)Trichlorosilane (FOTS) was deposited onto the etched silicon surface using an MVD100 molecular vapor deposition system (Applied Microstructures) to enable easy release of the PDMS from the silicon mold.

Sylgard 184 (Dow Corning) PDMS base resin was mixed at a 10:1 ratio with curing agent, and degassed in a vacuum oven at room temperature. The PDMS was poured onto the master mold and baked at 140°C for 1 hour. The PDMS was allowed to cool to room temperature, and it was carefully peeled off of the mold. Inlet and outlet holes were created using a 1.5mm biopsy punch (Sklar Instruments). To complete the formation of the channels, the patterned PDMS and a microscope slide were treated with oxygen plasma for 5 min, and then bonded together. An external 4-way L-type valve (IDEX Health and Science) was used at the DNA channel inlet and external 2-way valves were used at the cell channel inlet and both outlets to control the flow between the perpendicular channels.

### *2.3 Surface Chemistry*

DNA aptamers that were used to capture cancer cells were immobilized onto the surface of the microchannels via a streptavidin-biotin conjugation. The channels were initially primed and cleaned with ethanol and ultrapure water. A 1x PBS solution was flowed through the channels at 50 µL/min for 9 min. A 1 mg/mL solution of streptavidin (Life Technologies) in PBS was prepared, and 120 µL were pumped through the cell channel (see Figure 4.1) at 4 µL/min to immobilize streptavidin via adsorption to the channel surface. A 1% bovine serum albumin (BSA) solution was prepared in PBS binding buffer, and 500 µL of this solution was pumped through both the cell and DNA channels at 4 µL/min to block the surface of the channel and prevent non-specific adhesion of cells and other reagents. Finally, biotinylated aptamers (Integrated DNA Technologies), which were previously selected to bind to several different types of cancer cells,[18] were diluted to 10 µM in PBS binding buffer, and 180 µL were pumped through the DNA channel at 3 µL/min. This process enabled the aptamers to be concentrated in the cell capture region at the intersection of the two channels.

### *2.4 Cell Capture, Lysis, and gDNA Isolation*

Cells suspended in PBS binding buffer were pumped through the cell channel (see Figure 4.1) at 1 µL/min until several cells were captured, as observed via fluorescence microscopy. A solution of 6 M guanidinium isothiocyanate was pumped at 1 µL/min through the DNA channel until cell lysis was observed. By lysing the cells in the direction of the DNA micropillar array, the gDNA becomes physically entangled and remains in the channel even under flow. To verify the gDNA was captured by the micropillar array, YOYO-1 dye was used to stain the DNA in the channel by flowing it through the DNA channel at 1 µL/min. Otherwise, following lysis, the gDNA

was washed with 1x TE buffer to remove any remaining cellular debris and guanidinium isothiocyanate.

### *2.5 Primers and Specific Gene Amplification from gDNA*

To amplify a specific gene from gDNA isothermally, MDA was used with specific primers targeting a fragment of the *TP53* gene. Table 4.1 contains the 40 short primers used for this amplification. Initially, the isolated gDNA was chemically denatured using Buffer DLB from the Repli-g Mini Kit (Qiagen) prepared according to the manufacturer's instructions. This solution was pumped through the DNA channel at 1 µL/min for 25 min. The gDNA was then neutralized using the Stop Solution from the Repli-g Mini Kit with the addition of 12 pmol of each MDA primer. This solution was pumped at 1 µL/min for 25 min. An MDA reaction solution was prepared from the Repli-g Mitochondrial DNA Kit according to the manufacturer's instructions, and 30 pmol of each MDA primer was added to the reaction solution. The microchannel device was placed on a hotplate set to 31°C, and 50 µL of MDA reaction solution was pumped through at 0.05 µL/min to perform the MDA reaction for approximately 16 hours. The eluent was collected in an Eppendorf tube. For benchtop control experiments, gDNA from HeLa cells was extracted using the Blood and Cell Culture DNA Mini Kit (Qiagen). Buffer D1 and Buffer N1 from the Repli-g Mini Kit (Qiagen) were prepared, and 15 ng of extracted gDNA were denatured and neutralized according to the manufacturer's instructions. An MDA reaction master mix was prepared using the buffer and DNA polymerase from the Repli-g Mitochondrial DNA Kit (Qiagen) and various concentrations of specific MDA primers (Table 4.1), and this solution was added to the denatured gDNA. This reaction was incubated at 31°C for 16 hours.

**Table 4.1:** Amplification Primer Sequences

| **MDA Primers for the *TP53* Gene** | | | |
|---|---|---|---|
| **Primer Name** | **Sequence** | **Primer Name** | **Sequence** |
| *TP53* MDA FOR #1 | TTG TGC CCT G | *TP53* MDA REV #1 | GCC ATG GCG C |
| *TP53* MDA FOR #2 | CAG TTG CTT T | *TP53* MDA REV #2 | GCT GTG ACT G |
| *TP53* MDA FOR #3 | GTT TCT TTG C | *TP53* MDA REV #3 | CCT CAC AAC C |
| *TP53* MDA FOR #4 | GGA GGT GCT T | *TP53* MDA REV #4 | CGC TCA TGG T |
| *TP53* MDA FOR #5 | TAG CTC GCT A | *TP53* MDA REV #5 | CAC CAT CGC T |
| *TP53* MDA FOR #6 | GTG TAG ACG C | *TP53* MDA REV #6 | TCT CTC CAG C |
| *TP53* MDA FOR #7 | CCT ATC TCA A | *TP53* MDA REV #7 | CTG GGC AAC C |
| *TP53* MDA FOR #8 | CCT GAG TGA C | *TP53* MDA REV #8 | GAA TCA GAG G |
| *TP53* MDA FOR #9 | ATC ACA CCA C | *TP53* MDA REV #9 | ACC TAA GAG C |
| *TP53* MDA FOR #10 | GGA GGC TGC A | *TP53* MDA REV #10 | AGA TGC TGA G |
| *TP53* MDA FOR #11 | ATC ACT TGA G | *TP53* MDA REV #11 | TCC ACA CGC A |
| *TP53* MDA FOR #12 | TAG GAG GCT G | *TP53* MDA REV #12 | TGT TTC TGT C |
| *TP53* MDA FOR #13 | CAC CTA TAG T | *TP53* MDA REV #13 | CAC CAC ACT A |
| *TP53* MDA FOR #14 | TAG CCA GGC A | *TP53* MDA REV #14 | TCA GGC GGC T |
| *TP53* MDA FOR #15 | ACC TCG TCT C | *TP53* MDA REV #15 | CCC CAG TTG C |
| *TP53* MDA FOR #16 | AGC CTG GGT A | *TP53* MDA REV #16 | TTA ACC CCT C |
| *TP53* MDA FOR #17 | GAG CCC AGG A | *TP53* MDA REV #17 | AGG GCC ACT G |
| *TP53* MDA FOR #18 | CCA AGG CAG G | *TP53* MDA REV #18 | CCC CCC TAC T |
| *TP53* MDA FOR #19 | TCC CAG CAC T | *TP53* MDA REV #19 | TAG GGA GGT C |
| *TP53* MDA FOR #20 | TTG GTG GCT C | *TP53* MDA REV #20 | TTG CAC ATC T |
| **PCR Primers** | | | |
| **Primer Name** | **Sequence** | | |
| *TP53* PCR FOR | ACT TTC AAC TCT GTC TCC TTC CTC TTC CTA | | |
| *TP53* PCR REV | GGA CGC GGG TGC CG | | |

### *2.6 Gene Sequencing*

The MDA product was verified by running the product on a 1% agarose gel, and staining the gel with SYBR gold (Thermo Fisher Scientific). Since the product was around 10 kilobases (kb), a smaller product must be made to enable sequencing. Thus, a PCR amplification was performed using 10-50% of the MDA product. The PCR product was verified and purified using 8% PAGE stained with ethidium bromide. To purify the PCR product for sequencing, the 130-nt

band was cut from the gel, crushed, suspended in 3 M NaAc pH 6.2, and incubated at 37°C with mixing overnight. The gel pieces were removed from the sample, the sample underwent phenol-chloroform extraction and ethanol precipitation, and the DNA was resuspended in DEPC water. This purified PCR product was sequenced via Sanger Sequencing, and the sequence was compared to the known wildtype gene.

## 3 Results and Discussion

### *3.1 Microchannel Design*

The assay presented here involves two main steps, cancer cell capture using aptamers and genomic DNA analysis. The ability to perform these two steps easily all within one device would significantly reduce contamination and sample loss, which is important when developing an assay for rare cells. Therefore, the device was designed to have two intersecting orthogonal microchannels: the cell channel where cancer cells are captured, and the DNA channel where the gDNA from the captured cells is isolated (Figure 4.1). The cell channel was 1 mm wide, and the DNA channel was 250 µm, 500 µm, or 1 mm wide. The depth of the channels was ~25 µm. The device also contained two micropillar arrays, one for cancer cell capture and one for gDNA isolation. The cell capture array was located at the intersection of the two microchannels, and consisted of pillars 50 µm in diameter. This array was rotated by 4° to maximize the contact between the cells and the channel surface containing the aptamers, thereby improving the capture efficiency.[19] The DNA micropillar array was very similar to our previously-developed microarray for gDNA isolation.[17] The array presented here consisted of 4x4 µm pillars spaced in a gradient that started with the pillars 10 µm apart and finished with them 7 µm apart. These small pillars placed in such a fine array caused the gDNA to become physically entangled and remain on the

array even under flow, while allowing the cellular debris to flow out of the device. This is highly advantageous and enables multiple type of analysis to be performed on the same gDNA template.

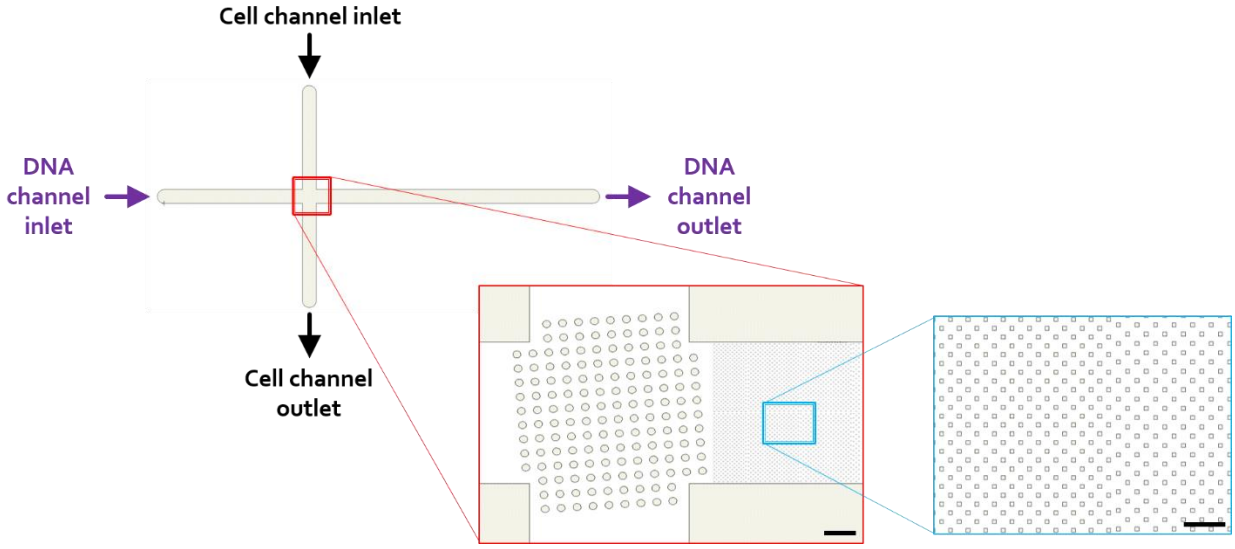



**Figure 4.1:** Microchannel design for capturing cancer cells and isolating their gDNA. The device contains two orthogonal channels, the cell channel and the DNA channel. It also contains two micropillar arrays, the cell capture array located at the intersection of the channels (red box, scale bar: 200 µm) and the DNA isolation array located downstream of the cell capture array in the DNA channel (blue box, scale bar: 30 µm).

To control the flow within the device, valves were needed at the inlets and outlets. An external 4-way L-type valve was used at the DNA channel inlet to not only control flow, but also eliminate most of the dead volume caused by the length of tubing needed to reach the syringe pump while the device is either observed under a microscope or is incubated using a hotplate. This is necessary because high initial flowrates cannot be used after the gDNA is isolated on the micropillar array, as there is a risk of losing some of the entangled gDNA under the high flow conditions. External 2-way shutoff valves were used at each of the outlets as well as the cell channel inlet to control the fluid flow. These types of valves were used because they do not disturb the fluid within the channels, whereas controlling the flow using tubing clamps causes significant

flow during the clamping process, which can disturb the captured cells and gDNA. This device design and setup enables this assay to be performed easily by the user.

### *3.2 HeLa and CAOV-3 Cell Capture and Genomic DNA Isolation*

The target cells were suspended in PBS binding buffer and pumped through the device with aptamers immobilized on the surface of the cell capture region containing a pillar array. The cell capture was most efficient at a flowrate of 1 µL/min when the cells were freshly trypsinized and the channel depth was ~25 µm. Several flowrates were tested ranging from 20 µL/min down to 0.1 µL/min. At high flowrates (>5 µL/min), the linear velocity of the cells was very high and the cells did not bind to the aptamers on the surface easily. At low flowrates (<1 µL/min), the cells did bind to the aptamers, but the throughput was very low and there was increased non-specific adhesion of the cells to the device. Cells used from frozen stocks did not bind as readily, likely due to different surface compositions caused by the shock of a freeze/thaw cycle. With channels deeper

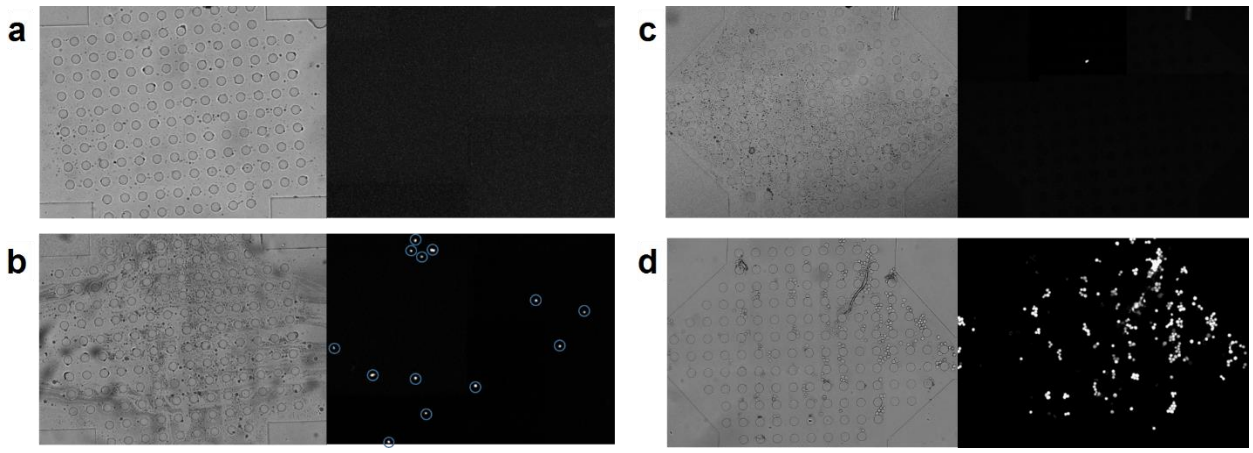



**Figure 4.2:** Cancer cell capture in microfluidic channels using aptamers. (a) Brightfield and fluorescent images of the cell capture region without aptamers after flowing HeLa cells through the microchannels. (b) Brightfield and fluorescent images of the cell capture region containing aptamers after flowing HeLa cells through the microchannels. The blue circles highlight the captured cells. (c) Brightfield and fluorescent images of the cell capture region without aptamers after flowing CAOV-3 cells through the microchannels. (d) Brightfield and fluorescent images of the cell capture region containing aptamers after flowing CAOV-3 cells through the microchannels. The HeLa cells fluoresce due to GFP bound to their histones, and the CAOV-3 cells were stained with calcein-AM.

than 25 μm, the cell capture took longer, since fewer collisions with the channel surface were likely occurring. To demonstrate using aptamers to capture cancer cells within the microchannels, cells were pumped through channels with and without aptamers immobilized on the surface. Figure 4.2a and 4.2b show the cell capture region without aptamers (Figure 4.2a) and with aptamers (Figure 4.2b) after HeLa cells were pumped through the channel. Figure 4.2c and 4.2d show the results from the same experiment performed using CAOV-3 cells. These results demonstrate the use of aptamers to successfully capture cancer cells.

After successfully capturing cancer cells within the cell capture region, the cells were lysed through the DNA channel. The cell contents flowed through the DNA micropillar array, and the gDNA was isolated on the pillars via physical entanglement. To verify that the gDNA was isolated and retained within the micropillar array, the gDNA was stained with YOYO-1 dye, and the results are shown in Figure 4.3 for both HeLa and CAOV-3 cells. The gDNA from both types of cells was

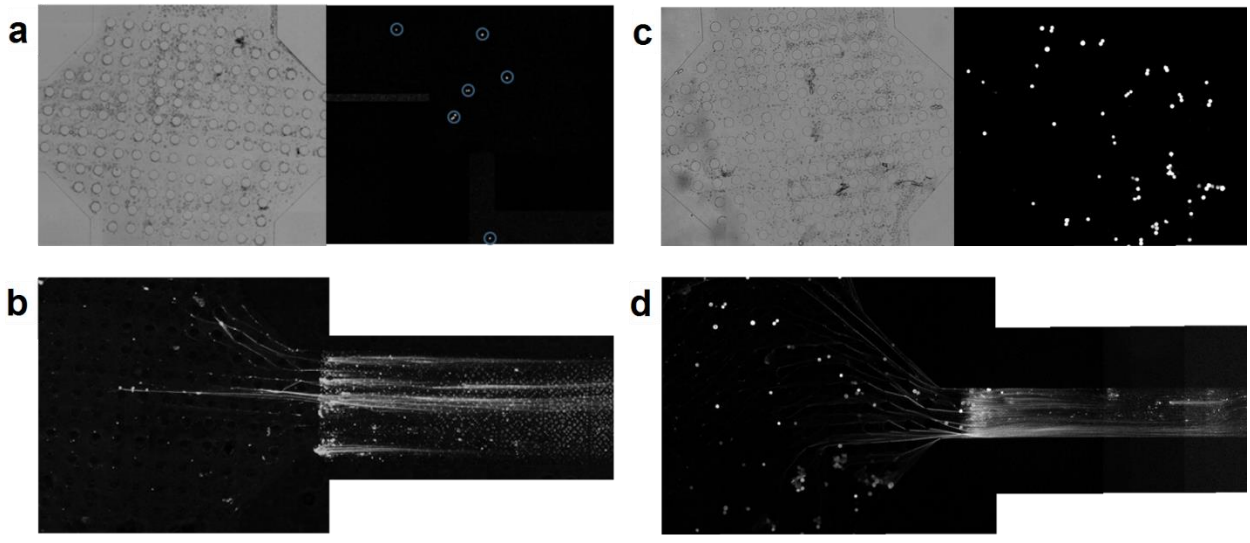



**Figure 4.3:** Cancer cell lysis and isolation of gDNA via physical entanglement within the micropillar array. (a) Brightfield and fluorescent images of HeLa cells bound to aptamers in the cell capture region. (b) Fluorescent image of gDNA from HeLa cells isolated by the micropillar array and stained with YOYO-1 dye. (c) Brightfield and fluorescent images of CAOV-3 cells bound to aptamers in the cell capture region. (b) Fluorescent image of gDNA from CAOV-3 cells isolated by the micropillar array and stained with YOYO-1 dye.

successfully isolated by the micropillar array and remained within the microchannel after multiple hours of flow. This gDNA isolation technique has several advantages, including allowing the gDNA to be completely chemically available for any downstream reaction, and more importantly it enables the gDNA to be retained within the device even under flow. This enables multiple types of analyses to be performed on the same template gDNA, so a large amount of information can be gathered from a single small population of cells without necessitating substantial whole genome amplification, which can be biased and information can be lost.[20-22]

### *3.3 Specific Isothermal Amplification of the TP53 Gene from gDNA*

With the gDNA isolated and retained within the micropillar array, several types of genetic analyses can be performed. Here, we demonstrated the amplification of a specific gene to determine its nucleotide sequence and discover any genetic mutations. MDA, an isothermal amplification technique, was used to reduce the complexity and cost of the device and enable straightforward device operation, since thermocycling is not necessary. MDA uses many short primers and a strand-displacing DNA polymerase (φ29) to make numerous copies of the template DNA. Whereas MDA is commonly used as a method for whole genome amplification,[23-26] in this work we use MDA to amplify a specific gene from gDNA. To our knowledge, this is the first demonstration of using MDA to amplify a specific gene from gDNA. Forty 10-nt primers (20 forward and 20 reverse primers) were designed around a region of the *TP53* gene, which is a commonly mutated gene in cancer whose resulting protein (p53) is responsible for monitoring genetic integrity and preventing cell survival and proliferation in the presence of genetic damage.[27] This modified MDA technique for amplifying the *TP53* gene from gDNA was initially verified in benchtop experiments using purified HeLa gDNA. Figure 4.4a shows the successful

amplification of *TP53* from purified gDNA. Shorter gene fragments were also produced via PCR from these MDA products (Figure 4.4b).

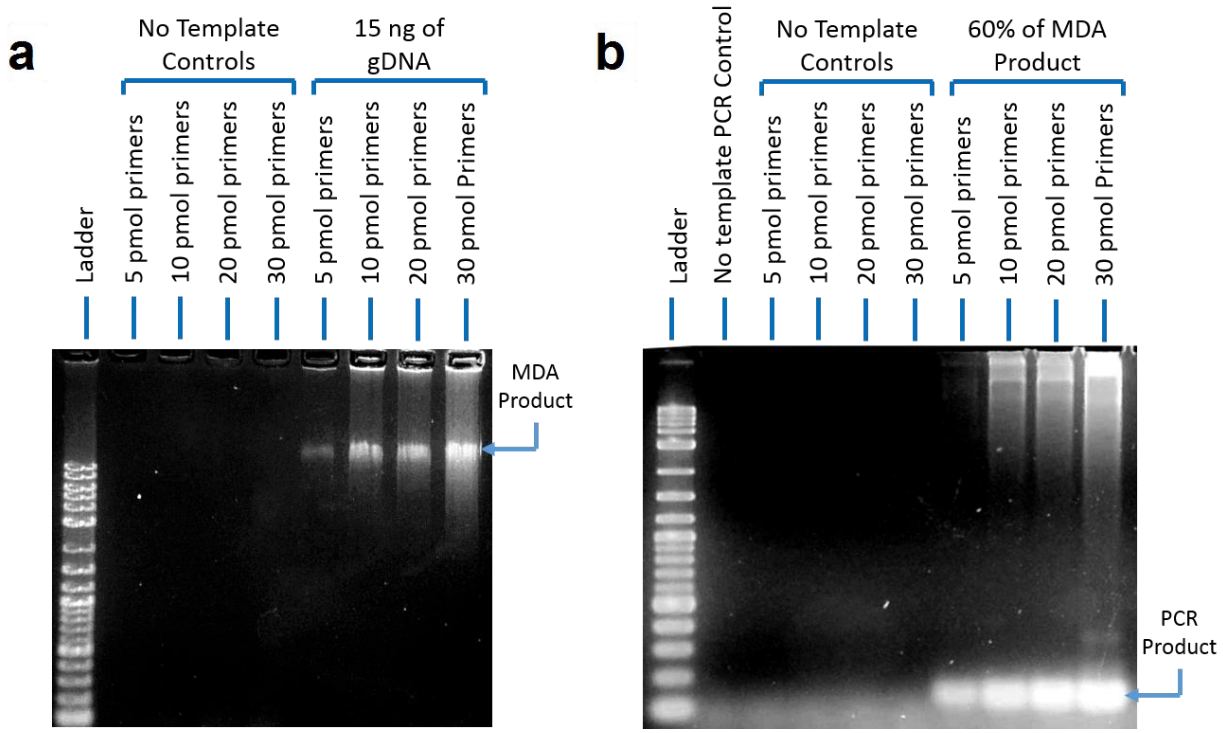



**Figure 4.4:** Gel images of benchtop demonstration of specific MDA of the *TP53* gene and smaller gene fragment from PCR. (a) MDA product from reactions using 15 ng of purified gDNA. Different quantities of primers were tested, as well as negative controls containing no gDNA template. The MDA product was approximately 10 kb. (b) PCR product from reactions amplifying a shorter *TP53* gene fragment using 60% of the MDA product. The PCR product is 130 nt in length.

The isolated gDNA from HeLa and CAOV-3 cells was first chemically denatured to create single stranded DNA (ssDNA). The DNA was then neutralized prior to the MDA reaction. Since the gDNA is entangled in the micropillar array, upon denaturation the two complimentary strands likely cannot diffuse and migrate away from each other as readily as they can in a bulk solution, so it is possible that some may rehybridized upon neutralization. This would prevent the MDA reaction from occurring efficiently. To reduce the likelihood of this occurring, the neutralization buffer was spiked with the MDA primers allowing the primers to bind to the ssDNA before it

rehybridizes. The MDA reaction solution containing additional primers and the DNA polymerase was pumped through the device and the amplification product was collected. The MDA product from HeLa and CAOV-3 cell gDNA was analyzed on an agarose gel, and Figure 4.5a shows images of the gels after a successful MDA reaction. The MDA product DNA strands were around 10 kb in length, and this is because the φ29 polymerase is known to extend for approximately 10 kb.[28,29] To prepare the amplified gene sample for sequencing, it must be shortened to accommodate the sequencing technology. For Sanger sequencing, up to a few hundred bases can be sequenced. To create a shorter gene fragment from the HeLa and CAOV-3 MDA product, PCR was used to create 130-nt DNA strands, which were verified via PAGE (Figure 4.5b). MDA is an advantageous technique to use not only because it is isothermal, but also because amplifying 10-kb strands allows whole genes to be amplified from the gDNA template. These genes can be

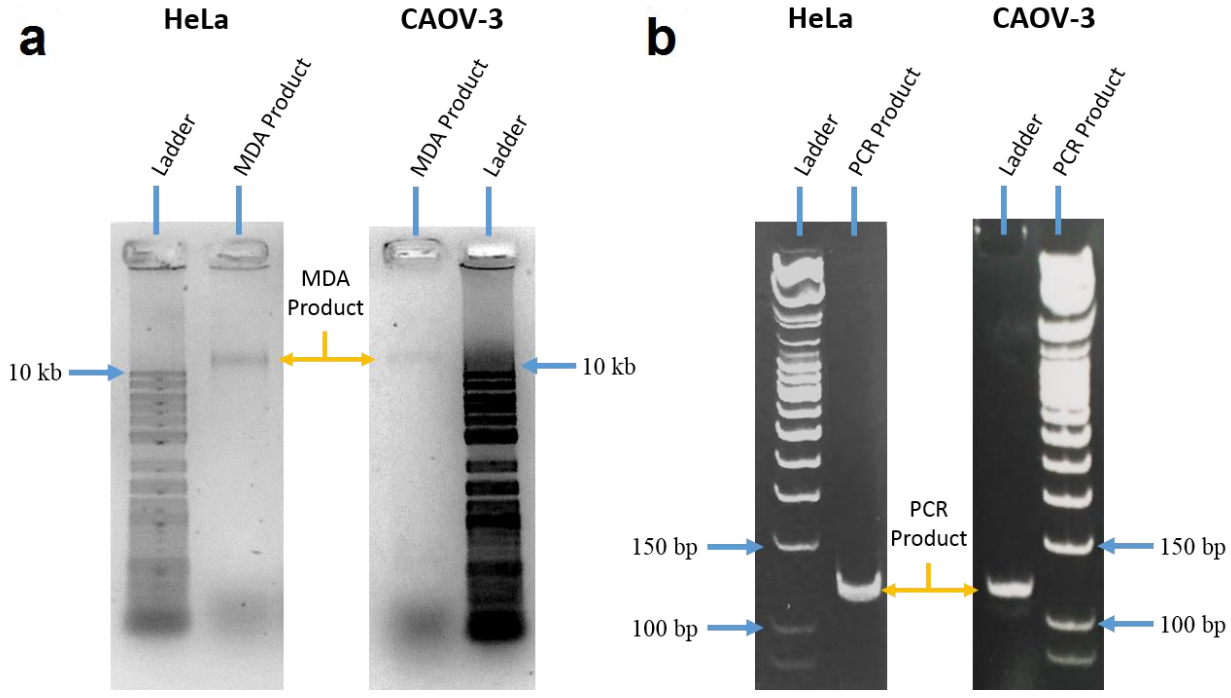



**Figure 4.5:** Gel images after running MDA and PCR products. (a) On-chip MDA product from isolated HeLa and CAOV-3 cell gDNA verified using a 1% agarose gel. The ladder indicates that the product is around 10 kb in length. (b) PCR product verified via 8% PAGE. The ladder confirmed a product length of 130 bp.

completely sequenced inexpensively by performing a few PCRs to isolate different regions of the gene, and performing Sanger sequencing on these pure samples. Alternatively, the MDA product can be digested into shorter fragments and targeted next-generation sequencing can be performed.

### *3.4 Sequencing of the TP53 Gene Fragment*

The purified *TP53* gene fragment from HeLa and CAOV-3 cells was sequenced via Sanger sequencing. This sequencing technique was chosen because it is fast, inexpensive, and can perform longer reads than other techniques. It also requires the sample to be pure, as only one read is performed. The sequencing results from the *TP53* gene fragments from both HeLa and CAOV-3 cells are shown in Figure 4.6. The wildtype sequence for *TP53* was obtained from the International Agency for Research on Cancer, and the wildtype nucleotide sequence for the short fragment tested here is shown in Figure 4.6. From these results, we were able to detect a point mutation in the CAOV-3 *TP53* gene, while HeLa cells contained the wildtype sequence of this gene fragment. This confirms that CAOV-3 cells contain a nonsense mutation in the *TP53* gene leading to an

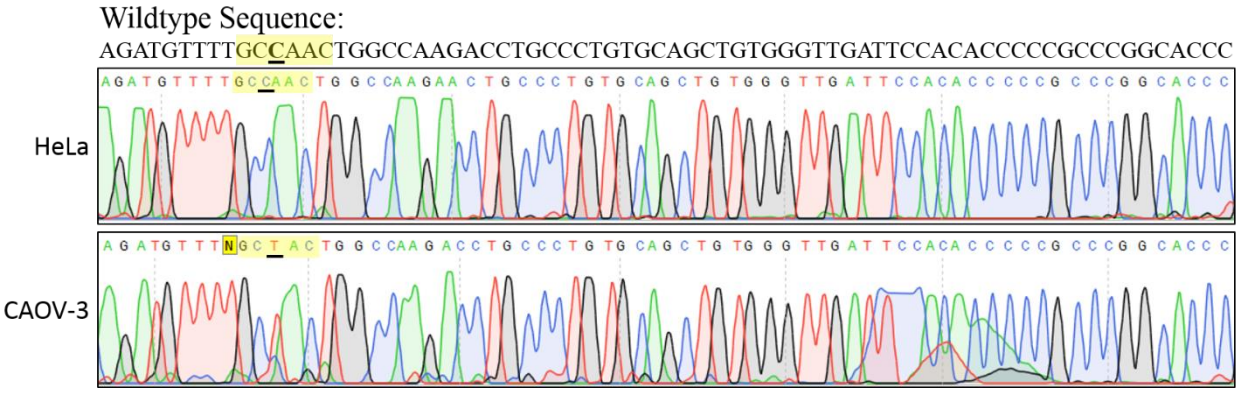



**Figure 4.6:** Sanger sequencing results from HeLa and CAOV-3 cell *TP53* gene fragments. The wildtype sequence for this fragment is shown at the top of the figure and was obtained from the International Agency for Research on Cancer. Plots from sequencing runs for both HeLa and CAOV-3 *TP53* gene fragments are shown with the corresponding bases indicated. The highlighted bases indicate the location of the known point mutation in CAOV-3 cells, and the specific base that is mutated is underlined. There are some bases that were not called, but the sequence can be discerned by looking at the peaks.

aberrantly-functioning p53 protein. Analogous testing of all locations on this gene would allow us to determine additional genetic information about this gene in these two cancer cell lines. Additional genes that commonly contain mutations in cancer could also be tested in this way, since the gDNA is retained on the micropillars and is available for additional amplification reactions and analysis.

## 4 Conclusions

We have developed a device and technique for capturing cancer cells using aptamers, and isolating their gDNA for genetic analysis. The device was designed to enable both the cell capture and gDNA isolation to occur within a single device to significantly reduce contamination and sample loss. Preventing sample loss is particularly important, since most cells of interest, such as circulating tumor cells, are rare and are generally found in low numbers. The device was also designed to isolate the gDNA via physical entanglement within a micropillar array. This isolation technique allows the gDNA to be chemically available for downstream reactions. It also enables the gDNA to be retained within the microchannel even under flow, which enables multiple types of analyses to be performed on the same template gDNA.

We demonstrated on-chip isothermal amplification of a specific gene and subsequent sequencing of this gene to determine any genetic mutations. We successfully detected a point mutation in the *TP53* gene in CAOV-3 cells. Specifically, we developed a method using MDA to amplify a specific gene from whole gDNA. This technique allows for long strands of DNA to be produced that can cover an entire large gene up to 10 kb. Here, we used PCR to produce pure samples of smaller gene fragments that could be sequenced using Sanger sequencing, but other sequencing methods can be used where pure samples are not necessary as multiple reads are taken.

An assay like this can be extremely impactful when it comes to cancer diagnosis and treatment. If the mutations present in a specific patient's cancer cells can be identified, clinicians will be able to use this information to better diagnose the patient and better determine the most effective treatment options. This would truly enable precision medicine when it comes to cancer treatment, which is necessary to effectively treat patients considering the substantial diversity observed in cancer.

## 5 References


[1] K. Vermeulen, D. R. Van Bockstaele, Z. N. Berneman, *Cell Prolif.* **2003**, *36*, 131–149.

[2] B. Vogelstein, K. W. Kinzler, *Nat. Med.* **2004**, *10*, 789–799.

[3] L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, M. R. Stratton, *Nat. Rev. Cancer* **2004**, *4*, 177–183.

[4] P. C. Nowell, *Science.* **1976**, *194*, 23–28.

[5] C. G. Mullighan, S. Goorha, I. Radtke, C. B. Miller, E. Coustan-Smith, J. D. Dalton, K. Girtman, S. Mathew, J. Ma, S. B. Pounds, et al., *Nature* **2007**, *446*, 758–64.

[6] A. Didelot, D. Le Corre, A. Luscan, A. Cazes, K. Pallier, J. F. Emile, P. Laurent-Puig, H. Blons, *Exp. Mol. Pathol.* **2012**, *92*, 275–280.

[7] M. N. Nikiforova, A. I. Wald, S. Roy, M. B. Durso, Y. E. Nikiforov, *J. Clin. Endocrinol. Metab.* **2013**, *98*, 1852–1860.

[8] G. Ellison, G. Zhu, A. Moulis, S. Dearden, G. Speake, R. McCormack, *J. Clin. Pathol.* **2013**, *66*, 79–89.

[9] P. N. Gray, C. L. M. Dunlop, A. M. Elliott, *Cancers (Basel).* **2015**, *7*, 1313–1332.

[10] D. T. Cheng, T. N. Mitchell, A. Zehir, R. H. Shah, R. Benayed, A. Syed, R. Chandramohan, Z. Y. Liu, H. H. Won, S. N. Scott, et al., *J. Mol. Diagnostics* **2015**, *17*, 251–264.

[11] Y. Wan, Y. Liu, P. B. Allen, W. Asghar, M. A. I. Mahmood, J. Tan, H. Duhon, Y. Kim, A. D. Ellington, S. M. Iqbal, *Lab Chip* **2012**, *12*, 4693–4701.

[12] J. A. Phillips, Y. Xu, Z. Xia, Z. H. Fan, W. Tan, *Anal. Chem.* **2009**, *81*, 1033–1039.

[13] F. Zheng, Y. Cheng, J. Wang, J. Lu, B. Zhang, Y. Zhao, Z. Gu, *Adv. Mater.* **2014**, *26*, 7333–7338.

[14] Y. Xu, J. A. Phillips, J. Yan, Q. Li, Z. H. Fan, W. Tan, *Anal. Chem.* **2009**, *81*, 7436–7442.

[15] W. Sheng, T. Chen, R. Kamath, X. Xiong, W. Tan, Z. H. Fan, *Anal Chem* **2012**, *84*, 4199–4206.

[16] Q. Shen, L. Xu, L. Zhao, D. Wu, Y. Fan, Y. Zhou, W. H. Ouyang, X. Xu, Z. Zhang, M. Song, et al., *Adv. Mater.* **2013**, *25*, 2368–2373.

[17] J. J. Benítez, J. Topolancik, H. C. Tian, C. B. Wallin, D. R. Latulippe, K. Szeto, P. J. Murphy, B. R. Cipriany, S. L. Levy, P. D. Soloway, et al., *Lab Chip* **2012**, *12*, 4848.

[18] D. Van Simaeys, D. López-Colón, K. Sefah, R. Sutphen, E. Jimenez, W. Tan, *PLoS One* **2010**, *5*, e13770.

[19] J. P. Gleghorn, E. D. Pratt, D. Denning, H. Liu, N. H. Bander, S. T. Tagawa, D. M. Nanus, P. A. Giannakakou, B. J. Kirby, *Lab Chip* **2010**, *10*, 27–29.

[20] R. Pinard, A. de Winter, G. J. Sarkis, M. B. Gerstein, K. R. Tartaro, R. N. Plant, M. Egholm, J. M. Rothberg, J. H. Leamon, *BMC Genomics* **2006**, *7*, 216.

[21] C.-Z. Zhang, V. A. Adalsteinsson, J. Francis, H. Cornils, J. Jung, C. Maire, K. L. Ligon, M. Meyerson, J. C. Love, *Nat. Commun.* **2015**, *6*, 1–10.

[22] J. Sabina, J. H. Leamon, in *Whole Genome Amplifcation,* **2015**, pp. 15–41.

[23] F. B. Dean, S. Hosono, L. Fang, X. Wu, a F. Faruqi, P. Bray-Ward, Z. Sun, Q. Zong, Y. Du, J. Du, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5261–6.

[24] C. Spits, C. Le Caignec, M. De Rycke, L. Van Haute, A. Van Steirteghem, I. Liebaers, K. Sermon, *Nat. Protoc.* **2006**, *1*, 1965–1970.

[25] J. G. Paez, M. Lin, R. Beroukhim, J. C. Lee, X. Zhao, D. J. Richter, S. Gabriel, P. Herman, H. Sasaki, D. Altshuler, et al., *Nucleic Acids Res.* **2004**, *32*, e71.

[26] S. H. L. Mandey, M. S. Schneiders, J. Ã. Koster, H. R. Waterham, *Hum. Mutat.* **2006**, *27*, 796–802.

[27] M. Hollstein, K. Rice, M. S. Greenblatt, T. Soussi, R. Fuchs, T. Sørlie, E. Hovig, B. Smith-Sørensen, R. Montesano, C. C. Harris, *Nucleic Acids Res.* **1994**, *22*, 3551–5.

[28] S. J. Foster, B. J. Monahan, *Fungal Genet. Biol.* **2005**, *42*, 367–375.

[29] O. Alsmadi, F. Alkayal, D. Monies, B. F. Meyer, *BMC Res. Notes* **2009**, *2*, 48.

# CHAPTER 5

# CONCLUSIONS AND FUTURE DIRECTIONS

Aptamers have the potential to rival antibodies as the affinity ligand of choice in many applications, including biosensors and diagnostics, imaging, and therapeutics. They are capable of binding to their targets with the same strength and specificity, but provide several benefits over antibodies.[1] However, despite the past 26 years of aptamer research, antibodies are still the gold standard. One of the main reasons for this is the difficulty in selecting aptamers to new targets. The aptamer selection process is very time consuming and expensive, and the majority of selections are unsuccessful in discovering an aptamer that binds with high affinity and specificity. I believe that this is the most important hurdle to overcome before aptamers will be used regularly. To contribute to the effort of improving aptamer discovery, we developed a Microplate-based Enrichment Device Used for the Selection of Aptamers (MEDUSA). This device is a scaled-up version of our microcolumn technology in which 96 microcolumns are contained within one device with the same dimensions as a 96-well microplate. This allows MEDUSA to directly couple to a 96-well plate to facilitate high-throughput downstream processes. MEDUSA is also very versatile in that it can be reconfigured in many ways between microcolumns in parallel and microcolumns connected in series. As we have demonstrated, MEDUSA can not only perform high-throughput parallelized selections, but can also be used to characterize and optimize aptamer selection processes, so improved, more efficient selection techniques can be developed that greatly increase success rates resulting in the increased production of high affinity aptamers. Another important feature of this device is that its microplate-based design lends itself to process automation using existing liquid-handling robots that are compatible with 96-well microplates. This would further improve the efficiency and throughput of aptamer selections. Establishing technologies that

provide rapid and reliable methods for high-throughput aptamer discovery would advance the aptamer field and promote the widespread use of aptamers.

Equally important to efficient aptamer discovery is the development of applications that use aptamers to improve upon existing technologies. Therefore, we have also worked toward creating novel assays that incorporate aptamers. We have shown that MEDUSA's versatility enables it to be used for more than just aptamer selections. We demonstrated sample preparation techniques in the microcolumns where dilute protein samples were purified and concentrated using aptamer-immobilized resin. A step further would be to separate, purify, and concentrate several species from complex mixtures using MEDUSA by connecting several microcolumns in series, and again, these processes could be automated using existing systems that handle 96-well plates.

One area of research that could greatly benefit from aptamer-based technologies is that of cancer. Cancer is a disease that has been extensively studied by researchers in an attempt to understand its underlying causes and mechanisms, and more importantly, to treat it effectively. However, despite many decades of research, it is still not well understood, and is still one of the leading causes of death in the world.[2] The specific characteristics of the disease often vary significantly between patients, and this is fundamentally due to the genetic variations present in the cancer cells that allow them to ultimately escape the regulatory processes necessary for the healthy function of tissues and organs.[3-5] There are numerous mechanisms for malignancy each with different combinations of genetic mutations and cancer cells are constantly evolving, which makes cancer treatment difficult with varying levels of efficacy. Therefore, improved cancer diagnostics would make a significant impact in how cancer patients are treated. More specifically, regular testing of patient's cancerous cells to monitor their physical characteristics and genetic information would be quite beneficial. This would allow clinicians to use substantially more

information about the patient's specific cancer cells to better prescribe the most effective treatment strategy. For this to be possible, the testing would need to be rapid and reasonably inexpensive.

Toward these goals, we have developed a technology capable of capturing rare cancer cells from a complex sample, such as blood, and isolating their gDNA for analysis all within a single device. By combining these steps into a single device, the risk of contamination and sample loss are eliminated. To capture cancer cells, aptamers that bind specifically to cancer cells are immobilized within a microchannel containing a micropillar array to increase capture efficiency. The captured cells are lysed, and a secondary, finer micropillar array isolates the gDNA via physical entanglement. This technology provides a platform where different types of testing can be performed on the same population of cells. Initially, once the cells are captured, testing can be performed on whole cells to determine their surface composition. For example, fluorescent labeling and staining can be done to identify the presence and abundance of different surface species. The cells can then be lysed, their gDNA captured, and then several types of analysis can then be performed on the gDNA to gather information about the genetics of this specific population of cancer cells. We have demonstrated how genetic mutations in specific genes of interest can be identified through amplification of a specific gene and subsequent sequencing of that gene product. Since the gDNA is retained within the microchannel, several amplification reactions could be performed consecutively for different genes, or a larger amplification reaction targeting many cancer genes could be performed to prepare a sample for targeted next generation sequencing.[6-10] Different fluorescent techniques, such as fluorescence in situ hybridization (FISH),[11-14] could also be performed to analyze the cancer genetics. Furthermore, with the analysis occurring within a single device, these processes could be automated to increase the speed and efficiency of the testing. Therefore, with many different types of analysis possible, this technology can provide

clinicians with a plethora of information about each specific population of cancer cells, which helps them prescribe the most effective treatment strategies.

Aptamers that bind to cancers cells are not only useful for diagnostic applications like our technology where they are used to capture and filter out cancer cells from complex mixtures containing several different types of cells, but these aptamers can also be used in biomarker discovery for early detection and cancer therapeutics. Aptamers that can bind specifically to cancer-related cell surface molecules can be used for delivering targeted drugs and precisely directed therapies. The process by which aptamers to cells are discovered is known as cell systematic evolution of ligand by exponential enrichment (Cell-SELEX).[15,16] This approach allows for specific ligands to be identified that bind to external biomarkers on cancer cells without requiring any prior knowledge about the cells.[17] Currently, the aptamer selection process to cellular targets is very time consuming and often does not result in a high-quality aptamer.[17-19] One seemingly universal problem is dead cell contamination causing nonspecific binding and resulting in aptamers with low specificity and limited affinity to the live target cells.[20-22] These limitations could potentially be overcome using our microcolumn technology for aptamer selections to cells, since it has been shown to reduce the number of selection rounds needed, and the dynamic flow ensures that non-adherent dead cells are expelled into waste. This technique could subsequently be scaled up and automated using MEDUSA. An optimal cell-SELEX technique would allow for the discovery of novel cancer cell-specific aptamers and cancer biomarkers more rapidly, giving rise to the possibility of targeted therapies customized for each patient, and with higher affinity, enabling earlier detection and diagnosis due to a lower detection limit.

Precision medicine, where diagnostics and therapeutics are catered to each individual patient, would revolutionize the healthcare industry, and greatly improve lives and life expectancies. I believe that aptamers could be key to achieving precision medicine. In order to bring the aptamer field forward toward these goals, technologies that improve the aptamer discovery process must continue to be developed. Establishing efficient and reliable methods for selecting specific high affinity aptamers to targets and cells associated with disease is paramount to developing diagnostics for early detection and targeted therapeutics. In addition, developing technologies and treatments that effectively use these aptamers is also very important. Therefore, I believe through continued research and collaboration in the aptamer, diagnostic, and therapeutic fields we have the potential to achieve truly personalized and precision medicine.

# 1 References


[1] S. D. Jayasena, Clin. Chem. 1999, 45, 1628–1650.

[2] R. L. Siegel, K. D. Miller, A. Jemal, *CA. Cancer J. Clin.* **2016**, *66*, 7–30.

[3] K. Vermeulen, D. R. Van Bockstaele, Z. N. Berneman, *Cell Prolif.* **2003**, *36*, 131–149.

[4] B. Vogelstein, K. W. Kinzler, *Nat. Med.* **2004**, *10*, 789–799.

[5] L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, M. R. Stratton, *Nat. Rev. Cancer* **2004**, *4*, 177–183.

[6] N. D'Haene, M. Le Mercier, N. De Nève, O. Blanchard, M. Delaunoy, H. El Housni, B. Dessars, P. Heimann, M. Remmelink, P. Demetter, et al., *PLoS One* **2015**, *10*, 1–13.

[7] D. M. Hyman, D. B. Solit, M. E. Arcila, D. T. Cheng, P. Sabbatini, J. Baselga, M. F. Berger, M. Ladanyi, *Drug Discov. Today* **2015**, *20*, 1422–1428.

[8] H. Beltran, R. Yelensky, G. M. Frampton, K. Park, S. R. Downing, T. Y. MacDonald, M. Jarosz, D. Lipson, S. T. Tagawa, D. M. Nanus, et al., *Eur. Urol.* **2013**, *63*, 920–926.

[9] O. Harismendy, P. C. Ng, R. L. Strausberg, X. Wang, T. B. Stockwell, K. Y. Beeson, N. J. Schork, S. S. Murray, E. J. Topol, S. Levy, et al., *Genome Biol.* **2009**, *10*, R32.

[10] L. Mamanova, A. J. Coffey, C. E. Scott, I. Kozarewa, E. H. Turner, A. Kumar, E. Howard, J. Shendure, D. J. Turner, *Nat. Methods* **2010**, *7*, 111–118.

[11] P. Gerami, S. S. Jewell, L. E. Morrison, B. Blondin, J. Schulz, T. Ruffalo, P. th Matushek, M. Legator, K. Jacobson, S. R. Dalton, et al., *Am J Surg Pathol* **2009**, *33*, 1146–1156.

[12] J. M. Levsky, R. H. Singer, *J Cell Sci* **2003**, *116*, 2833–2838.

[13] R. Swiger, J. Tucker, *Environ. Mol. Mutagen.* **1996**, *27*, 245–254.

[14] D. G. Hicks, R. R. Tubbs, *Hum. Pathol.* **2005**, *36*, 250–261.

[15] K. T. Guo, A. Paul, C. Schichor, G. Ziemer, H. P. Wendel, *Int. J. Mol. Sci.* **2008**, *9*, 668–678.

[16] X. Fang, W. Tan, *Acc. Chem. Res.* **2010**, *43*, 48–57.

[17] H. Ma, J. Liu, M. M. Ali, M. A. I. Mahmood, L. Labanieh, M. Lu, S. M. Iqbal, Q. Zhang, W. Zhao, Y. Wan, *Chem. Soc. Rev.* **2015**, *44*, 1240–1256.

[18] W. Y. Zhang, W. Zhang, Z. Liu, C. Li, Z. Zhu, C. J. Yang, *Anal. Chem.* **2012**, *84*, 350–355.

[19] M. You, Y. Chen, L. Peng, D. Han, B. Yin, B. Ye, W. Tan, *Chem. Sci.* **2011**, *2*, 1003.

[20] M. Avci-Adali, M. Metzger, N. Perle, G. Ziemer, H. P. Wendel, *Oligonucleotides* **2010**, *20*, 317–323.

[21] C. Meyer, U. Hahn, A. Rentmeister, *J. Nucleic Acids* **2011**, *2011*, 904750.

[22] A. Ozer, J. M. Pagano, J. T. Lis, *Mol. Ther. Nucleic Acids* **2014**, *3*, e183.

APPENDIX A

SUPPLEMENTARY FIGURES AND TABLES FOR CHAPTER 3

**Table A.1:** Summary of sequencing results and clustering

| | **Total # of reads** | **Quality filtered (>Q25 for 80%)** | **Percent passed quality filter:** | | | |
|---|---|---|---|---|---|---|
| Lane1 | 210774447 | 132780375 | 63.00 | | | |
| Lane2 | 205393613 | 68306006 | 33.26 | | | |
| **Sequencing Lane** | **Target protein** | **SELEX Round-Cycle** | **Barcode** | **Barcode matched** | **Clustered Reads** | **% Clustered** |
| Lane1 | Amylose | R3C3 | 1 | 3751360 | 3090917 | 82.4 |
| | | R4C4 | 2 | 4275552 | 3860899 | 90.3 |
| | GCN5_GNAT | R2C2 | 3 | 3632234 | 3335381 | 91.8 |
| | | R3C2 | 4 | 3315571 | 2993167 | 90.3 |
| | | R3C3 | 5 | 3347685 | 3025012 | 90.4 |
| | | R4C2 | 6 | 3454226 | 3127458 | 90.5 |
| | | R4C3 | 7 | 4888895 | 4364730 | 89.3 |
| | | R4C4 | 8 | 4086807 | 3535763 | 86.5 |
| | JMJD2_Clav | R3C3 | 9 | 2444505 | 2246627 | 91.9 |
| | | R4C4 | 10 | 5013347 | 4200451 | 83.8 |
| | MBP | R2C1 | 11 | 2264022 | 2092324 | 92.4 |
| | | R2C2 | 12 | 4830747 | 4401192 | 91.1 |
| | | R3C2 | 13 | 4505792 | 4081211 | 90.6 |
| | | R3C3 | 14 | 4782592 | 4367501 | 91.3 |
| | | R4C2 | 15 | 4108004 | 3648420 | 88.8 |
| | | R4C3 | 16 | 3453035 | 3103763 | 89.9 |
| | | R4C4 | 17 | 2194247 | 1922155 | 87.6 |
| | MOF_Chromo | R3C3 | 18 | 3983548 | 3664643 | 92.0 |
| | | R4C4 | 19 | 4043852 | 3598491 | 89.0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | TIP60_Chromo | R2C1 | 20 | 4093407 | 3735109 | 91.2 |
| | | R2C2 | 21 | 3930809 | 3544290 | 90.2 |
| | | R3C2 | 22 | 3575983 | 3266581 | 91.3 |
| | | R3C3 | 23 | 2406272 | 1641927 | 68.2 |
| | | R4C2 | 24 | 3729753 | 3104503 | 83.2 |
| | | R4C3 | 25 | 3144953 | 2892011 | 92.0 |
| | | R4C4 | 26 | 4582331 | 4132911 | 90.2 |
| | UTX_JMJC | R3C3 | 27 | 5184360 | 4734991 | 91.3 |
| | | R4C4 | 28 | 3681639 | 3297236 | 89.6 |
| | Unmatched | | | 18190321 | | |
| Lane2 | AgRai1 | R3C3 | 1 | 2324032 | 2146812 | 92.4 |
| | | R4C4 | 2 | 84290 | 76828 | 91.1 |
| | Ash1_BAH | R3C3 | 3 | 2269338 | 2102474 | 92.6 |
| | | R4C4 | 4 | 1884865 | 1611908 | 85.5 |
| | DXO | R3C3 | 5 | 2211899 | 2040984 | 92.3 |
| | | R4C4 | 6 | 2090182 | 1951136 | 93.3 |
| | FLAG-RTF1 | R3C3 | 7 | 2131522 | 1961767 | 92.0 |
| | | R4C4 | 8 | 1289219 | 1124605 | 87.2 |
| | FLAG | R3C3 | 9 | 2365489 | 2185911 | 92.4 |
| | | R4C4 | 10 | 2173980 | 1925590 | 88.6 |
| | JMJD2_JMJC | R3C3 | 11 | 2575224 | 2383706 | 92.6 |
| | | R4C4 | 12 | 1102417 | 983567 | 89.2 |
| | NELF-E | R2C1 | 13 | 2486073 | 2318500 | 93.3 |
| | | R2C2 | 14 | 2347505 | 2172774 | 92.6 |
| | | R3C2 | 15 | 1731767 | 1590004 | 91.8 |
| | | R3C3 | 16 | 1135320 | 1038583 | 91.5 |
| | | R4C2 | 17 | 2132048 | 1973489 | 92.6 |
| | | R4C4 | 18 | 2194432 | 2035697 | 92.8 |
| | NiNTA | R3C3 | 19 | 1584669 | 1432636 | 90.4 |
| | | R4C4 | 20 | 5926827 | 5306374 | 89.5 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Trx_ZnF | R3C3 | 21 | 3008680 | 2753132 | 91.5 |
| | | R4C4 | 22 | 2844793 | 2529447 | 88.9 |
| | mutPNPase | R3C3 | 23 | 1756627 | 1622867 | 92.4 |
| | | R4C4 | 24 | 1830710 | 1683166 | 91.9 |
| | wtPNPase | R3C3 | 25 | 481140 | 338727 | 70.4 |
| | | R4C4 | 26 | 919315 | 749263 | 81.5 |
| | p23 | R3C3 | 27 | 2672349 | 2467321 | 92.3 |
| | | R4C4 | 28 | 2517795 | 2308017 | 91.7 |
| | unmatched | | | 7046111 | | |

**Table A.2:** Target Protein Information

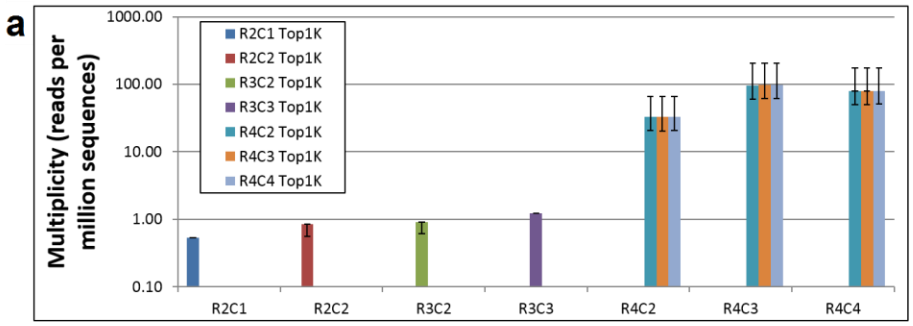
| Protein Name | GenBank ID | Protein Function | Specific Domain | Domain Boundaries (a.acids - a.acids) | Domain Function | Molecular Weight (kDa) | Isoelectric point (pI) | Affinity Tag/Resin used for SELEX |
|---|---|---|---|---|---|---|---|---|
| TIP60 | NM_131923 | H4 and H2A histone acetyltrasnferase. Catalytic subunit of NuA4 HAT complex. | Chromodomain | 11-300 | Involved in protein-protein and/or protein-nucleic acid interactions. Many chromodomains act as methyl-specific histone binding module, and some chromodomains were found to associate with RNA. | 31.5 | 8.25 | MBP/ Amylose |
| GCN5 | AF029776 | A major histone acetyltransferase, which promotes transcriptional activation. Catalytic subunit of SAGA, SALSA, and ADA complexes. | GNAT domain | 481-627 | Gcn5-related N-acetyltransferases (GNAT) catalyze the transfer of the acetyl from the CoA donor to a primary amine of the acceptor. | 16.8 | 8.64 | MBP/ Amylose |
| MOF | NM_078496 | A major histone acetyltransferase, which promotes transcriptional activation. Catalytic subunit of MSL and NSL complexes. | Chromodomain | 376-539 | Involved in protein-protein and/or protein-nucleic acid interactions. Many chromodomains act as methyl-specific histone binding module, and some chromodomains were found to associate with RNA. | 18.7 | 7.00 | MBP/ Amylose |
| UTX | NM_135524 | Histone demethylase that specifically demethylates 'Lys-27' of histone H3. | JMJC domain | 832-995 | Catalytic domain responsible for histone demethylation. Predicted metalloenzyme domain with cupin fold that binds Fe(III) and alphaKG. | 18.7 | 7.15 | MBP/ Amylose |
| JMJD2 | NM_136487 | Histone demethylase that specifically demethylates 'Lys-9' and 'Lys-36' residues of histone H3. | Clavaminate synthase domain | 95-336 | Catalytic domain responsible for histone demethylation. Predicted metalloenzyme domain with cupin fold that binds Fe(III) and alphaKG. | 28.7 | 8.33 | MBP/ Amylose |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| JMJD2 | NM_136487 | Histone demethylase that specifically demethylates 'Lys-9' and 'Lys-36' residues of histone H3. | JMJC domain | 149-315 | Catalytic domain responsible for histone demethylation. Predicted metalloenzyme domain with cupin fold that binds Fe(III) and alphaKG. | 19.5 | 6.32 | MBP/ Amylose |
| ASH1 | NM_079436 | Trithorax group (TrxG) protein with histone methyltransferase activity, which specifically trimethylates 'Lys-4' of histone H3 (a specific tag for epigenetic transcriptional activation) and to a lesser extend trimethylate H3 'Lys-9' and H4 'Lys-20'. | BAH domain | 1943-2063 | BAH (bromo-adjacent homology) is commonly found in chromatin-associated proteins implicated in transcriptional regulation. Acts as protein-protein interaction modules. | 14.5 | 8.26 | MBP/ Amylose |
| TrX | NM_134282 | Histone methyltransferase responsible for methylating 'Lys-4' of histone H3 (a specific tag for epigenetic transcriptional activation). Functions in segment determination through interaction with genes of bithorax (BX-C) and antennapedia (ANT-C) complexes. | Zinc Finger | 1266-1482 | Plant homeodomain (PHD) zinc finger domain has a C4HC3-type motif, found in many chromatin regulatory factors. They can bind zinc and other metal ions. Involved in protein-protein and protein-DNA interactions, and impliaceted in transcriptional regulation, translation, chromatin remodelling and etc. | 24.3 | 8.39 | MBP/ Amylose |
| MBP tag (MalE) | AF097412 | Affinity tag | N/A | 1-400 | Maltose binding / Affinity tag | 44.0 | 4.88 | N/A |

| | **GenBank ID** | **Protein Function** | **Molecular Weight (kDa)** | **Isoelectric point (pI)** | **Affinity Tag / Resin used for SELEX** | **Reference PubMed ID** |
|---|---|---|---|---|---|---|
| NELF-E | NM_139984 | Regulation of RNA Pol II elongation, maintenance of paused RNA Pol II. | 31.8 | 9.47 | 6xHis / Ni-NTA | 24453987 |
| DXO | NM_033613 | Exoribonuclease that specifically degrades pre-mRNAs with a defective 5'-cap and is part of a pre-mRNA capping quality control. Has decapping, pyrophosphohydrolase and 5'-3' exonuclease activities. | 45.3 | 8.29 | 6xHis / Ni-NTA | 23523372 |
| Ag Rai1 | NM_001181112 | Exoribonuclease that specifically degrades pre-mRNAs with a defective 5'-cap and is part of a pre-mRNA capping quality control. Has decapping, pyrophosphohydrolase and 5'-3' exonuclease activities. | 44.5 | 6.14 | 6xHis / Ni-NTA | 19194460 |

| | | | | | | |
|---|---|---|---|---|---|---|
| p23 | NM_001179683 | Molecular chaperone implicated in gene regulation by disassembly of protein-DNA complexes, which in turn allows GCN5 acetyltransferase to prolong the dissociated state through lysine acetylation. | 24.1 | 4.46 | 6xHis / Ni-NTA | 23022381 |
| PNPase (wildtype and S484A mutant) | NM_033109 | Polyribonucleotide nucleotidyltransferase 1 catalyzes the phosphorolysis of RNA processively in the 3'-to-5' direction as part of the mitochondrial degradosome (mtEXO) complex, also implicated in translocatio of nuclear-encoded RNA into mitochindria matrix. | 86.0 | 7.87 | 6xHis / Ni-NTA | 20691904 |
| 6xHis tag | AF097413 | Ni+2 binding / Affinity tag | 3.7 | 5.42 | | |

| | **GenBank ID** | **Protein Function** | **Molecular Weight (kDa)** | **Isoelectric point (pI)** | **Affinity Tag / Resin used for SELEX** | **Reference PubMed ID** |
|---|---|---|---|---|---|---|
| RTF1 | NM_015138 | Component of the PAF1 complex (PAF1C) which has multiple functions during transcription by RNA polymerase II and is implicated in regulation of development and maintenance of embryonic stem cell pluripotency. | 80.3 | 8.21 | FLAG peptide / FLAG M2 resin | 26217014 |
| Flag peptide tag | N/A | FLAG M2 antibody binding / Affinity tag | 2.7 | 3.97 | N/A | N/A |




b

| Common aptamers between Tip60-Chromodomain selected pools | | | | | | | |
|---|---|---|---|---|---|---|---|
| | **R2C1** | **R2C2** | **R3C2** | **R3C3** | **R4C2** | **R4C3** | **R4C4** |
| **R2C1** | 1000 | 4 | 4 | 3 | 1 | 1 | 2 |
| **R2C2** | | 1000 | 6 | 5 | 3 | 4 | 5 |
| **R3C2** | | | 1000 | 3 | 3 | 4 | 4 |
| **R3C3** | | | | 1000 | 21 | 21 | 22 |
| **R4C2** | | | | | 1000 | 906 | 905 |
| **R4C3** | | | | | | 1000 | 933 |
| **R4C4** | | | | | | | 1000 |

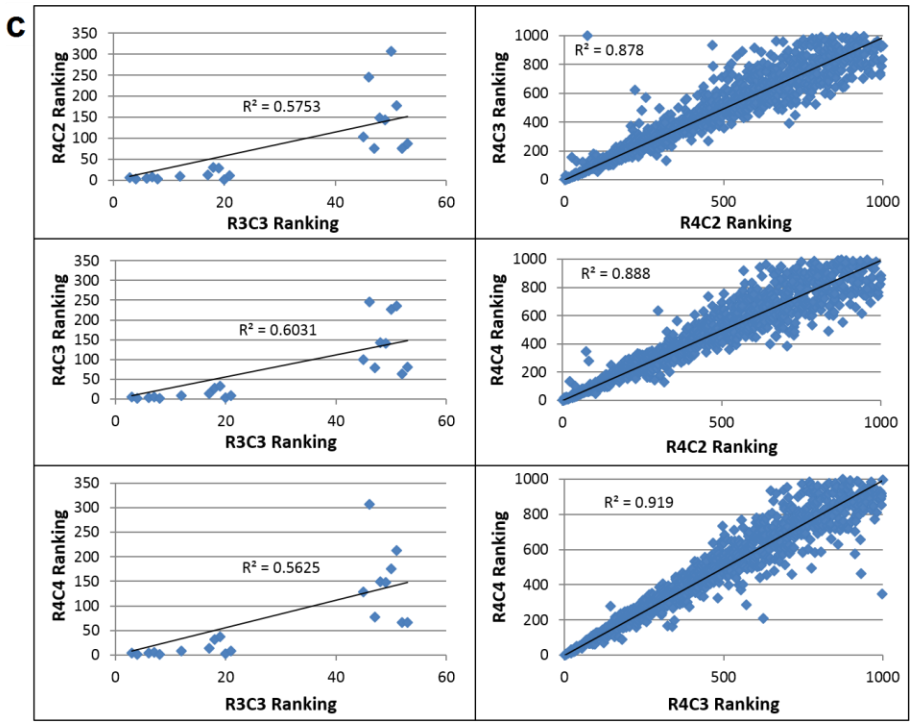
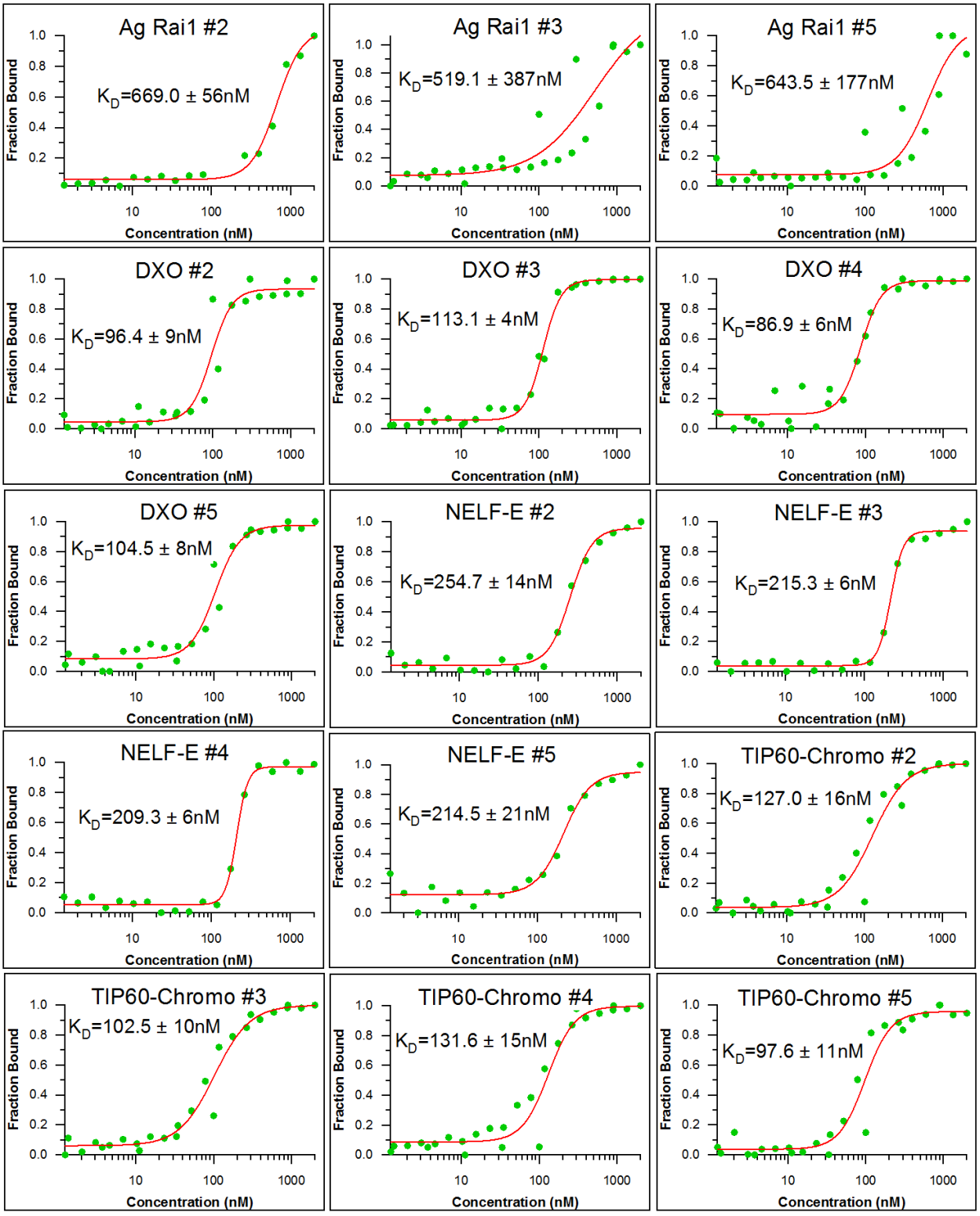



**Figure A.3:** Analysis of TIP60-Chromo enriched aptamer clusters through selection rounds and cycles. (a) Median multiplicity of top 1,000 TIP60-Chromo aptamer clusters in a given round-cycle and their median multiplicity in subsequent round-cycles of selection. (b) Number of common aptamer clusters in top 1,000 clusters between different round-cycles of selection for TIP60-Chromo. Common aptamer clusters from R2C1 to R3C2 with the later round-cycles are mostly due to sequence artifacts that pass through our sequence filter criteria. (c) Correlation plots for rankings of common aptamers in top 1,000 clusters at different selection cycles for TIP60-Chromo. $R^2$ values for each correlation are indicated. Multiplicities (reads per million) for 1$^{st}$, 20$^{th}$, and 1000$^{th}$ cluster in each library are as follows: R2C1 – 130.4, 1.3, 0.5; R2C2 – 184.5, 1.7, 0.6; R3C2 – 145.1, 2.5, 0.6; R3C3 – 697.4, 2.4, 1.2; R4C2 – 3,164.8, 610.4, 14.8; R4C3 – 12,038.3, 1,958.4, 43.6; and R4C4 – 7,923.3, 1,520.1, 35.8.


Ag Rai1 #2
$K_D$=669.0 ± 56nM
Ag Rai1 #3
$K_D$=519.1 ± 387nM
Ag Rai1 #5
$K_D$=643.5 ± 177nM
DXO #2
$K_D$=96.4 ± 9nM
DXO #3
$K_D$=113.1 ± 4nM
DXO #4
$K_D$=86.9 ± 6nM
DXO #5
$K_D$=104.5 ± 8nM
NELF-E #2
$K_D$=254.7 ± 14nM
NELF-E #3
$K_D$=215.3 ± 6nM
NELF-E #4
$K_D$=209.3 ± 6nM
NELF-E #5
$K_D$=214.5 ± 21nM
TIP60-Chromo #2
$K_D$=127.0 ± 16nM
TIP60-Chromo #3
$K_D$=102.5 ± 10nM
TIP60-Chromo #4
$K_D$=131.6 ± 15nM
TIP60-Chromo #5
$K_D$=97.6 ± 11nM
Fraction Bound
Concentration (nM)

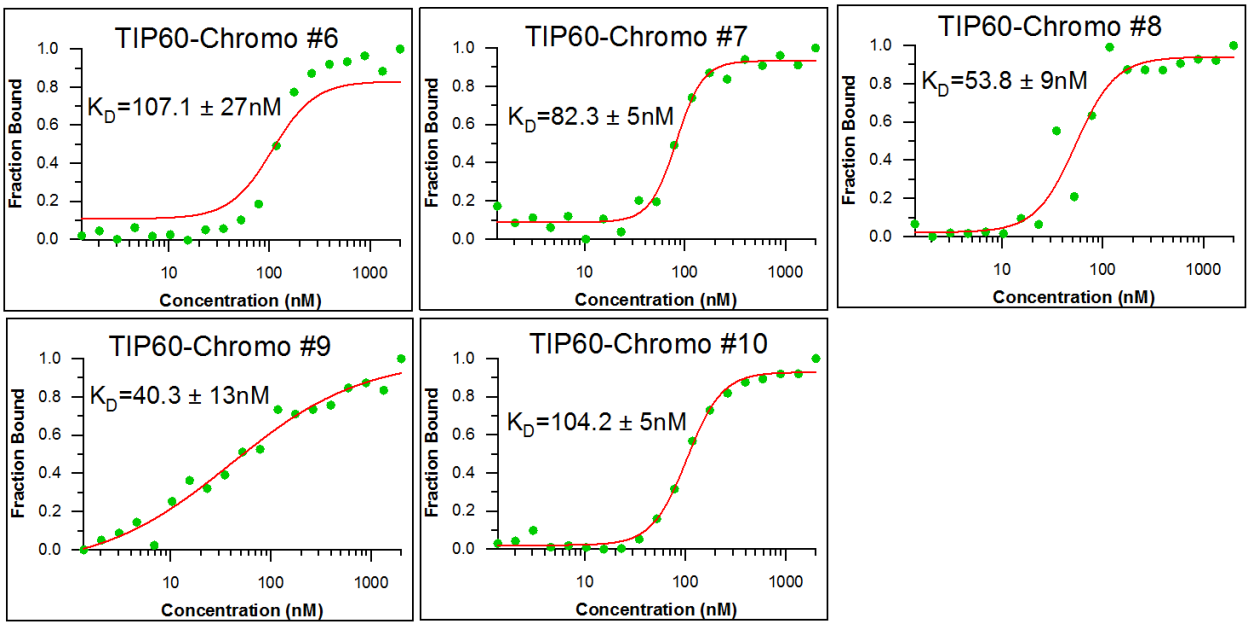



**Figure A.4:** Binding analysis of candidate aptamer sequences via EMSA. Results from EMSA experiments were quantified using ImageJ, and fitted to the Hill Equation using Igor Pro 5.04A.

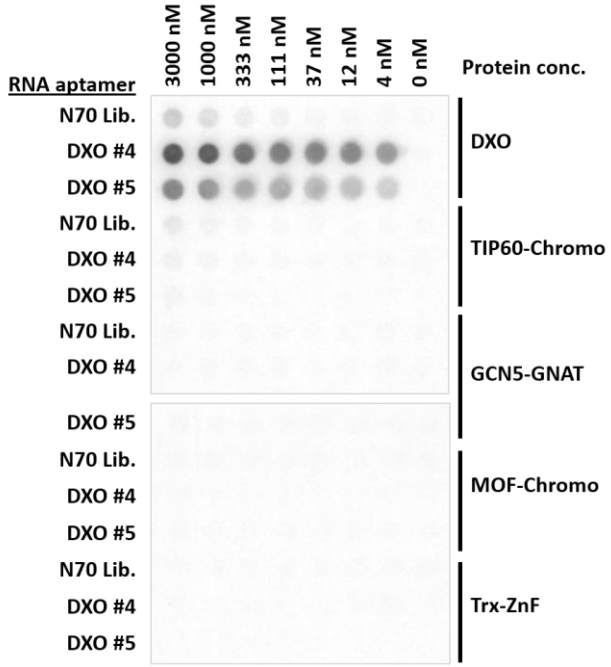



**Figure A.3:** Specificity of DXO aptamers. The specificity of DXO aptamers was tested by nitrocellulose filter binding assay. Radiolabeled RNA aptamers, DXO #4 and #5, and the N70 library (negative control) were incubated with varying concentrations (3000-4 nM) of DXO, TIP60-Chromo, GCN5-GNAT, MOF-Chromo, or Trx-ZnFinger proteins. Protein bound aptamers were captured by a nitrocellulose filter and detected by phosphorimaging. DXO #4 and #5 aptamers showed very weak binding, indistinguishable from N70 library control, against all other target proteins tested.

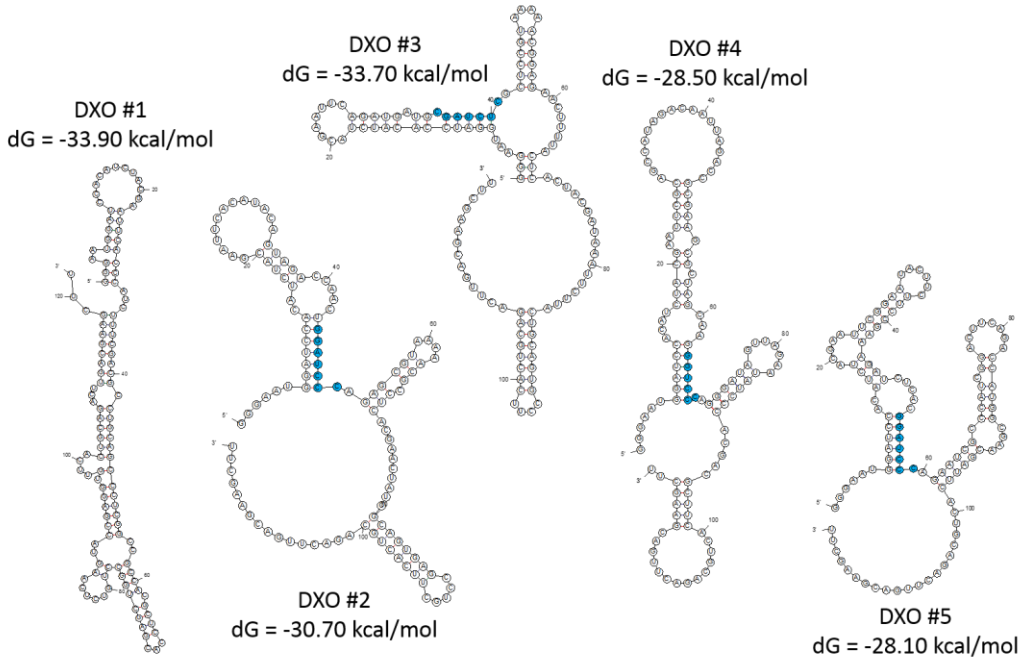
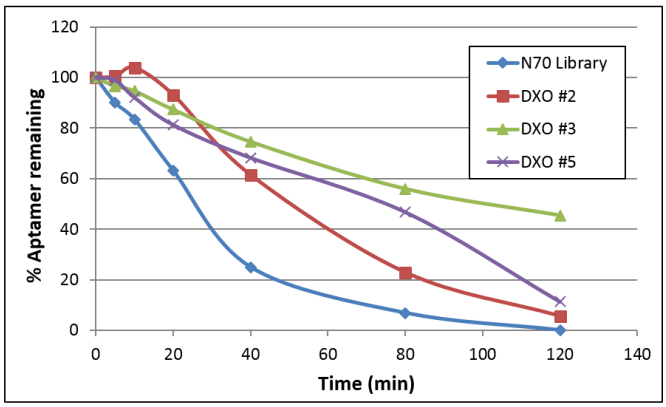



**Figure A.4:** Predicted secondary structures of DXO aptamers. MEME identified motif (DXO motif 1), which base-pairs with the BamHI site in the forward constant region of the N70 RNA library, is colored cyan. The predicted secondary structures and the deltaG values were obtained from mFold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form).

**a**

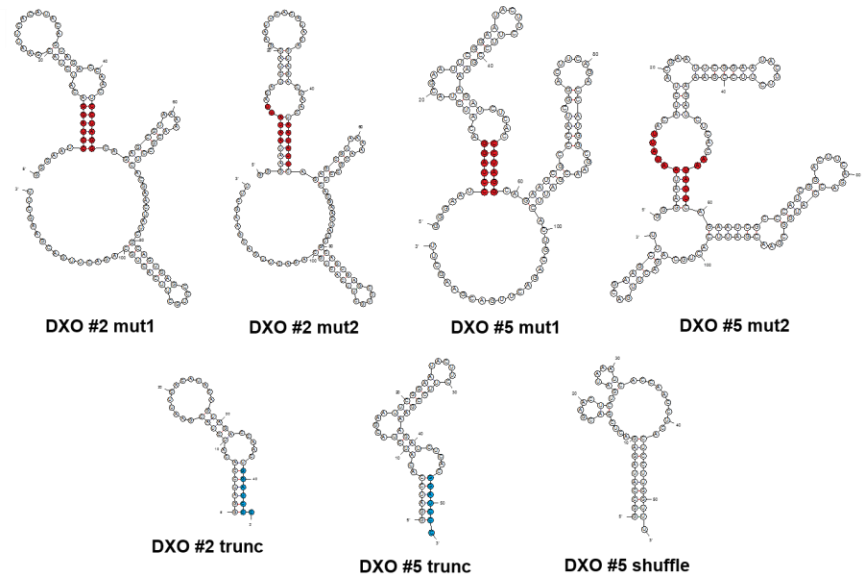



**b**

| | $T_{1/2}$ (min) |
|---|---|
| **N70 Library** | 19.8 |
| **DXO #2** | 28.9 |
| **DXO #3** | 43.3 |
| **DXO #5** | 69.3 |

**Figure A.5:** Stability of DXO aptamers against DXO exoribonuclease activity. (a) Each radiolabeled RNA aptamer was incubated with 1 µM DXO protein and samples collected at indicated time points were analyzed by running on an 8% denaturing PAGE gel. The gel image was analyzed, and a DXO exoribonuclease assay time course was plotted. (b) Aptamer half-lifes ($T_{1/2}$ min) are estimated from exponential decay function fit to measured % full-length aptamer remaining. Images were analyzed by ImageJ and data analysis was done in Microsoft Excel.

a




b

| Aptamer | DXO #2 | DXO #2 mut1 | DXO #2 mut2 | DXO #2 trunc | DXO #5 | DXO #5 mut1 | DXO #5 mut2 | DXO #5 trunc | DXO #5 shuffle |
|---|---|---|---|---|---|---|---|---|---|
| $dG_{complete}$ | -30.70 kcal/mol | -30.30 kcal/mol | -23.30 kcal/mol | -13.70 kcal/mol | -28.10 kcal/mol | -27.40 kcal/mol | -21.80 kcal/mol | -16.70 kcal/mol | -6.00 kcal/mol |
| $dG_{stem\#1}$ | -8.40 kcal/mol | -8.30 kcal/mol | -0.20 kcal/mol | -8.40 kcal/mol | -7.60 kcal/mol | -7.20 kcal/mol | 1.00 kcal/mol | -7.60 kcal/mol | -5.50 kcal/mol |

c

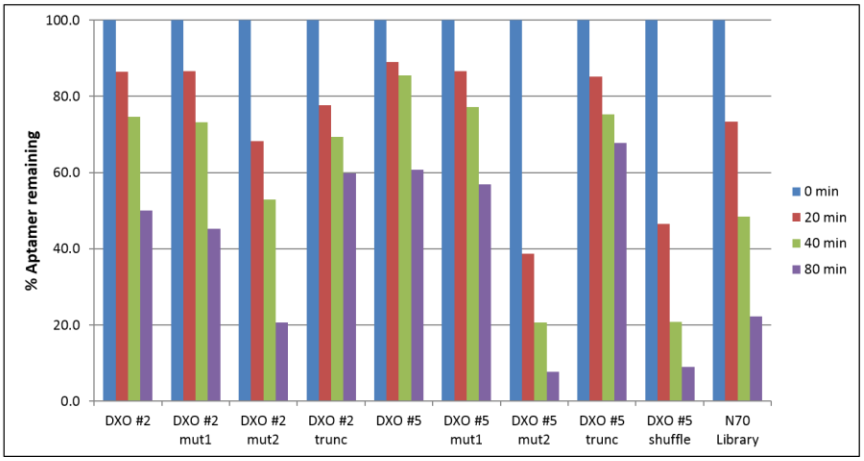
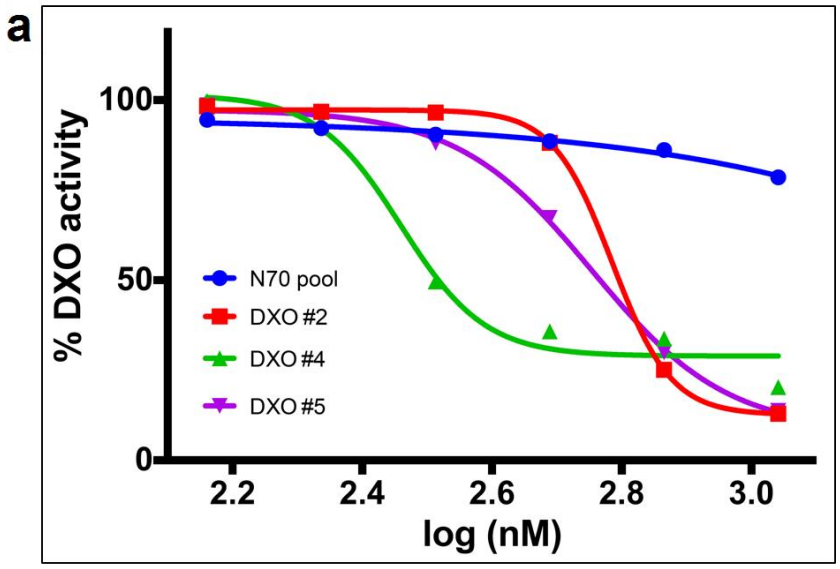



**Figure A.6:** Mutant and truncated DXO aptamers. (a) Predicted secondary structure of mutant and truncated DXO aptamers. MEME identified motif (DXO motif 1) is colored cyan and mutations to motif and the corresponding mutations in the forward constant region are colored red. In mut1, the original motif, GGATCCC, was changed to CCTAGGC and the BamHI site, GGATCC, in the forward constant region was changed to CCTAGG. In mut2, the original motif, GGATCCC, was changed to AATATTC and the BamHI site, GGATCC, in the forward constant region was changed to AATATT. Truncations of DXO #2 and #5 were made to retain the DXO motif 1 (cyan) and the first stem-loop structure for each aptamer. A sequence-shuffled version of DXO #5 trunc was used as a control. Secondary structures were predicted and the deltaG values were calculated by mFold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form). (b) Stability (deltaG values) of complete predicted structure and the first stem of original full-length, mutant, and truncated versions of DXO aptamers. (c) Stability of original, mutant and truncated DXO aptamers against DXO exoribonuclease activity.




b

| | IC50 (nM) |
|---|---|
| **N70 apt. pool** | >> 1100 |
| **DXO #2** | 609.5 ± 5.3 |
| **DXO #3** | 288.4 ± 29.8 |
| **DXO #5** | 567.5 ± 17.2 |

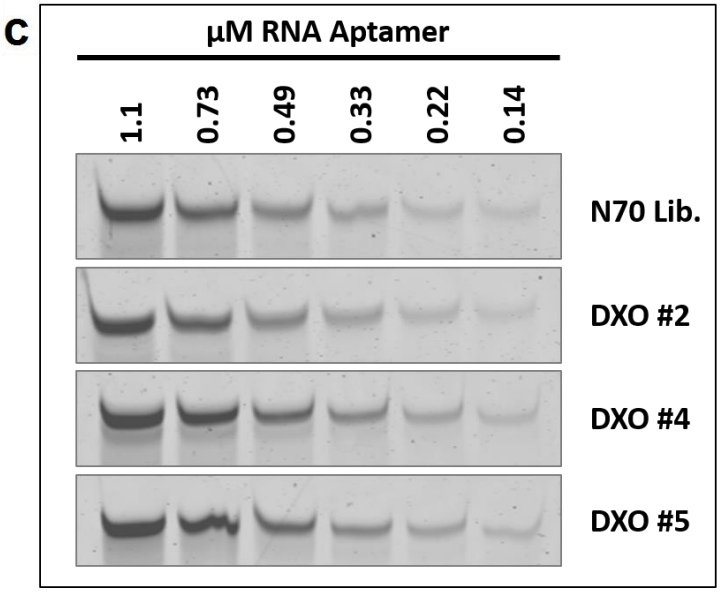



**Figure A.7:** Inhibition of DXO exoribonuclease activity by RNA Aptamers. (a) Quantification of DXO exoribonuclease activity against 3'-Cy5 labeled 30-nt RNA substrate in the presence of indicated concentrations of N70 aptamer pool, or DXO #2, #4, or #5 RNA aptamers from gels images shown in Fig. 4d. Data was fitted to a 4-parameter dose-response curve using GraghPad Prism6 software, and the estimated IC50 values are listed (b). Errors represent the standard error. (c) Verification of equal loading of RNA aptamers, which remained largely intact even after two hours of incubation with DXO enzyme, in DXO exoribonuclease inhibition assays.

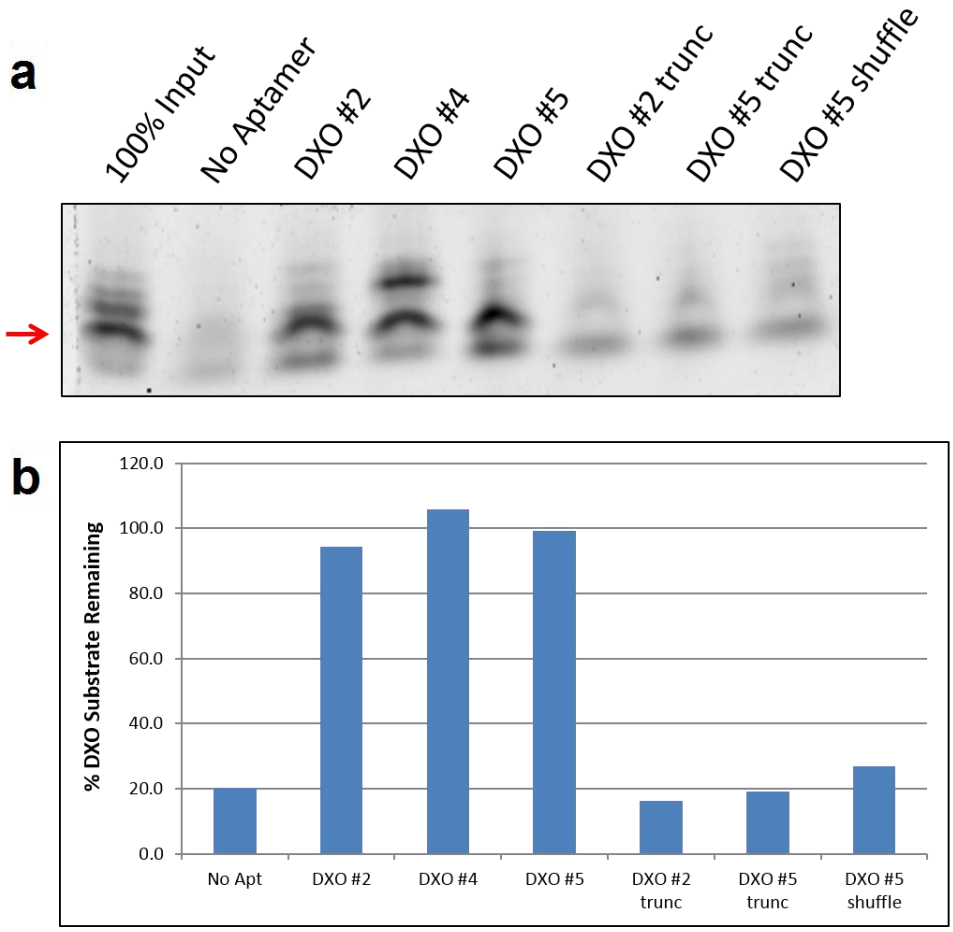



**Figure A.8:** Inhibition of DXO exoribonuclease activity by full-length or truncated DXO RNA aptamers. A) 3'-Cy5 labeled 30 nt RNA substrate was incubated with 1uM DXO exoribonuclease either in the absence of any aptamer (No Aptamer) or in the presence of 0.8 µM full-length or truncated DXO RNA aptamers. After 2 hours of incubation at 37°C, exoribonuclease reactions were separated by 10% denaturing PAGE, and the remaining intact RNA substrate, indicated by →, was visualized by Cy5 fluorescence scan. B) Percentage of full-length DXO substrate remaining under each condition as quantified by ImageJ.