MICROFLUIDIC SELECTION OF DNA APTAMERS USING CAPILLARY ELECTROPHORESIS AND MICRO FREE FLOW ELECTROPHORESIS SYSTEMATIC EVOLUTION OF LIGANDS BY EXPONENTIAL ENRICHMENT

A DISSERTATION
SUMBITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Meng Jing

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Michael T. Bowser, Adviser

August 2012

Acknowledgements

I want to thank Prof. Zude Zhang in the Department of Chemistry at the University of Science and Technology of China. It was in his Inorganic Chemistry class when I was first fascinated by the magic and beauty of Chemistry.

I want to thank Prof. Yi Xie and Dr. Benxia Li for mentoring me two years in my undergraduate research on hydrothermal synthesis of nano-materials. This precious experience not only gave me the first chance going from classic chemistry on book to cutting-edge research hotspots, but also built me a solid foundation of critical thinking and good laboratory practice.

I want to thank Prof. Hua Cui for encouraging me to go to graduate school. Her introduction to graduate school and advices during my application were very helpful.

Thank you to Prof. Michael Bowser, my research advisor for creating a great research environment where I had freedom and ample resources to follow my interest, and at the same time, gained encouragements and guidance from him. I am especially grateful for his help in improving my scientific writing and presentation, and for his supporting me to go to two conferences to present my work to the scientific community.

Thank you to Bowser Group for accompanying me on my graduate experience, during which I gained training on SELEX from Renee (Mosing) Cline and training on microfluidic fabrication from Ryan Turgeon and Nic Frost. I also enjoyed many

discussions and talks I had with you all, which gave me research sparkles and broadened my view in this country.

Thank you to my husband, Ligeng Yin for always being patient, supportive and encouraging during my graduate study. I appreciate his care and love, which make my life bright and beautiful; but most importantly, his enthusiasm in science, determined will of being an excellent chemist, and persistent hardworking set me a good example of conducting scientific research. I feel fortunate to have such a great life companion and fellow graduate student.

Finally I want to thank my parents for providing me everything I need to succeed in life and in school. Their unconditioned love, care, and education are very important for me to grow up and be a mature person to embrace life. I especially appreciate their understanding and support in my studying abroad with such a great geographical separation.

Dedication

To my parents,
who always supported my education.

Abstract

SELEX is the process used to generate aptamers, which are ssDNA or RNA molecules that can bind specific targets with high affinity. Although aptamers show great potential in clinical applications, the generation process is currently tedious and inefficient, being a limiting step. Thus, understanding, developing, and applying advanced partitioning platform are pivotal. CE is an advanced separation method in SELEX and has been successfully used to generate aptamers toward multiple targets. However, there are still interesting questions unanswered, making our understanding in CE-SELEX lag behind its applications. We applied high-throughput sequencing on CE-SELEX selected pools against rhVEGF and obtained sequencing information of more than 10^4 sequences per pool, which allowed characterization on diversity of individual pools. This study revealed the coexistence of high diversity and fast enrichment rate of CE-SELEX. To further improve the separation platform, we integrated a μFFE device into the SELEX process. Using this device, 10^{14}~10^{15} sequences were introduced and analyzed within 30 min, which was a 370 fold improvement compared to CE. As a proof of concept, four cycles of selection were performed to the target human IgE, and high affinity ligands were generated even after the first round of selection, proving the feasibility and high efficiency of μFFE in SELEX. Later, μFFE-SELEX was applied to generate aptamers for a membrane protein SERCA, whose selection has never been achieved in conventional SELEX due to the technical difficulty in target immobilization. High nM K_d pool was generated toward SERCA solubilized in $C_{12}E_8$ at the fifth round of selection, demonstrating

that this separation strategy is more compatible with membrane proteins due to the free solution based separation, lower electric field, faster separation speed, and straightforward fraction collection. The success of this application opens a door for high-throughput generation of aptamers toward complex targets in the future. Besides SELEX, the function and activity of the ATPase SERCA were also explored in this thesis. It was discovered that nonspecific ssDNA sequences can bind to the endogenous regulator of SERCA, PLN, in a length dependent way. A highly sensitive and reproducible SERCA activity assay, which cut the use of SERCA by 2,000 folds, was developed to directly detect the product ADP via TR-FRET.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

λ_{em}	Emission Wavelength
λ_{ex}	Excitation Wavelength
µFFE	Micro Free Flow Electrophoresis
µTAS	Micro Total Analysis Systems
ACE	Affinity Capillary Electrophoresis
ACS	Acute Coronary Syndrome
ADP	Adenosine Diphosphate
AMD	Age-related Macular Degeneration
AML	Acute Myeloid Leukemia
AMP	Adenosine Monophosphate
ATP	Adenosine Triphosphate
aSi	Amorphous Silicon
CE	Capillary Electrophoresis
CMACS	Continuous-flow Magnetic Activate Chip-based Separation
CRP	C-Reactive Protein
CTLA-4	Cytotoxic T Cell Antigen-4
CYT-18	Tyrosyl-tRNA synthetase from the mitochondria of the fungus *Neurospora crassa*
$C_{12}E_8$	Octaethylene glycol dodecyl ether
DNA	Deoxyribonucleic Acid
DOPC	Dioleoylphosphatidylcholine
DOPE	Dioleoylphosphatidylethanolamine
DTNB	Beta Dystrobrevin
DTT	DL-Dithiothreitol
EDTA	Ethylene Diamine Tetraacetic Acid
EGTA	Ethylene Glycol Tetraacetic Acid
EOF	Electroosmotic Flow

FA	Frontal Analysis
FACS	Fluorescence-Activated Cell Sorter
FFE	Free Flow Electrophoresis
FP	Fluorescence Polarization
FRET	Förster Resonance Energy Transfer
GE	Gel Electrophoresis
HD	Hummel-Dreyer
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER-3	Human Epithelial Growth Factor Receptor-3
HPLC	High-Performance Liquid Chromatography
i.d.	Inner diameter
IAEDANS	5-({2-[(iodoacetyl)amino]ethyl}amino) naphthalene-1-sulfonic acid
IEF	Isoelectric Focusing
IgE	Immunoglobulin E
ITC	Isothermal Titration Calorimetry
K_a	Association Constant
K_d	Dissociation Constant
kDa	Kilodalton
LDH	Lactate Dehydrogenase
LIF	Laser-Induced Fluorescence
LNA	Locked Nucleic Acid
MEK1	Human MAP kinase kinase
MMS	Micromagnetic Separation
MOPS	3-Morpholinopropanesul-fonic acid
MW	Molecular Weight
N_A	Aggregation Number
NAD^+	Nicotinamide Adenine Dinucleotide
PCR	Polymerase Chain Reaction

PDGF	Platelet Derived Growth Factor B
PDMS	Polydimethylsiloxane
PEO	Polyethylene glycol
PEP	Phosphoenolpyruvic acid
PfEMP1	Plasmodium falciparum Erythrocyte Membrane Protein 1
PK	Pyruvate Kinase
PKA	cAMP-dependent Protein Kinase
PLN	Phospholamban
PTP	PicoTiterPlate
R_h	Hydrodynamic Radius
Rho	Transcriptional termination protein from the Archaebacterium *Thermotoga maritima*
rhVEGF	Recombinant Human Vascular Endothelial Growth Factor
RNA	Ribonucleic Acid
SELEX	Systematic Evolution of Ligands by EXponential enrichment
SERCA	Sarco/Endoplasmic Reticulum Ca^{2+}-ATPase
SPR	Surface Plasmon Resonance
SR	Sarcoplasmic Reticulum
ssDNA	single stranded Deoxyribonucleic Acid
TGK	25 mM Tris, 192 mM Glycine, 5 mM KH_2PO_4, pH 8.3
TR-FRET	Time Resolved-Fluorescence Resonance Energy Transfer
VACE	Vacancy Affinity Capillary Electrophoresis
VDC-MSELEX	Volume Dilution Challenge-Microfluidic Systematic Evolution of Ligands by EXponential enrichment

VEGF	Vascular Endothelial Growth Factor
VP	Vacancy Peak
yTBP	Yeast TATA Binding Protein
yTFIIB	Yeast Transcription Factor IIB protein
ZE	Zonal Electrophoresis

Chapter 1 Introduction

1.1 Traditional SELEX and CE-SELEX on Generating Aptamers

1.1.1 Aptamers

Since the discovery of RNA molecules participating in the catalysis of living cells in 1970s,[1] various biochemical functions of nucleic acids have been identified beyond the commonly identified genetic storage and transfer.[2] Based on different functions, naturally occurring and artificially selected nucleic acids can be categorized into ribozymes/deoxyribozymes, allosteric ribozymes, riboswitches, and aptamers. Aptamers are *in vitro* selected single stranded DNA or RNA molecules that fold into unique structures to promote selective and high-affinity interactions with specific target molecules.[3]

Aptamers have been selected for a wide range of targets, from small molecules such as ions,[4, 5] amino acids,[6-8] antibiotics,[9-11] and biological cofactors,[12, 13] to large targets such as proteins,[14, 15] and even complex targets such as organelles[16] and cells.[17-21]Aptamers have many advantages over antibodies. Aptamers are produced via an in vitro process without using animals; binding is through adaptive recognition rather than strict paratope-to-epitope recognition; aptamers can work under any conditions they are isolated, not limited by physiological conditions; synthesis is straightforward without batch-to-batch difference; labeling is controllably at specific location so losing affinity is less of a concern. It is not surprising that applications of aptamers have grown rapidly into numerous areas of research. To name a few, sample purification in affinity chromatography[22-24] and microfluidic devices; quantification of targets in affinity

CE;[25, 26] high sensitivity detection of compound in complex matrixes in biosensors;[27, 28] "aptamer beacons"[28, 29] and aptamer arrays[30] in clinical diagnostics;[31, 32] Macugen,[33] ARC1779,[34] and other aptamer-based drugs,[35] are in market or in clinical phases for treating age-related macular degeneration (AMD), acute Myeloid Leukemia (AML), and other diseases.

1.1.2 SELEX

The in vitro process to generate aptamers is commonly referred to as Systematic Evolution of Ligands by EXponential enrichment (SELEX).[36, 37] The starting ssDNA or RNA library consists of 10^{13} to 10^{15} sequences that are completely random except for the flanking primer binding regions. As shown in Figure 1.1, after incubation with the target of interest, the mixture is subjected to separation. The bound sequences are then collected for PCR amplification, followed by purification to obtain the desired single stranded product for the next round of selection. Iterative rounds alternating between affinity enrichment of binding sequences and PCR amplification are performed, and the affinity of each generated pool is evaluated to monitor the selection process. Ideally, aptamers with low nM to high pM dissociation constants (K_d) can be selected.

Since the birth of SELEX, people have been working on advancing this technology. Great effort has been spent on improving isolation efficiency, or reducing the selection cycles to speed up aptamer generation, although attention has also been paid to improving bio-stability and affinity of aptamers as evidenced by the L-isomer aptamers, or spiegelmers,[38] and low pM affinity apta-

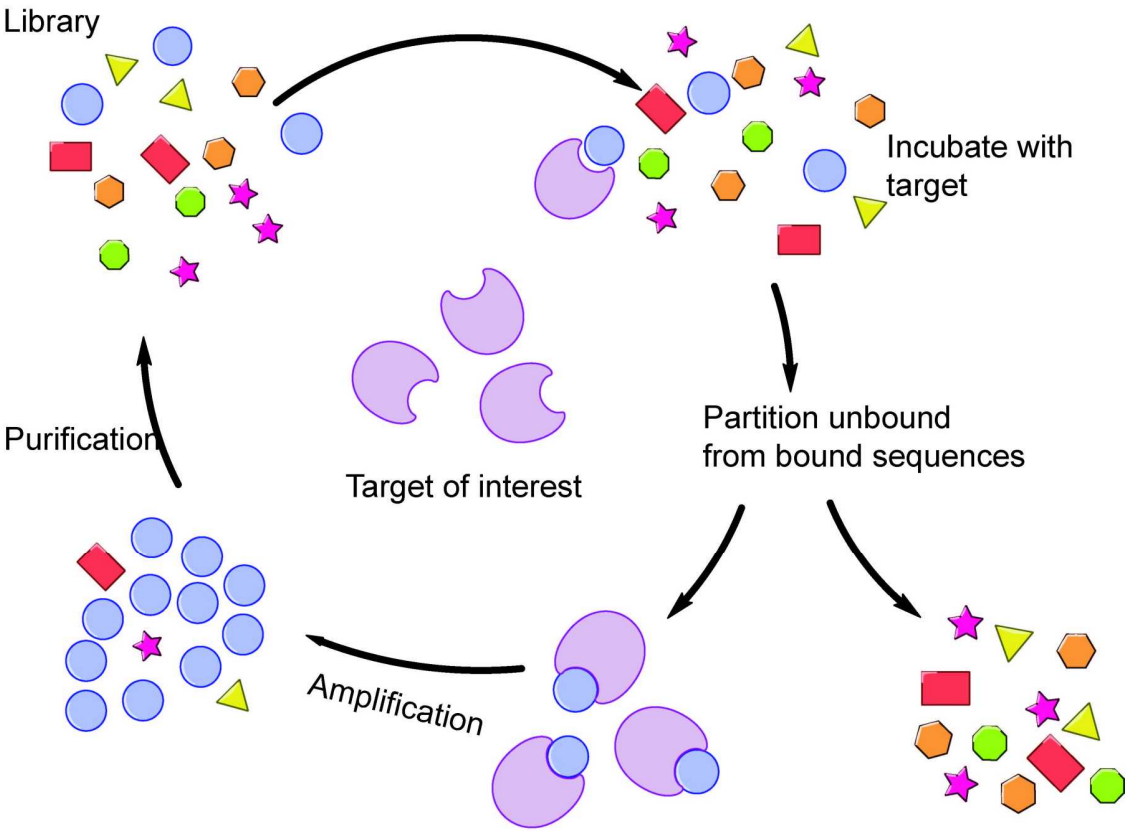

Figure 1.1 Schematic of SELEX process.

mers generated by PhotoSELEX.[39] Nitrocellulose membrane has been well accepted to perform isolation ever since Tuerk and Gold demonstrated its success in obtaining an RNA aptamer for T4 DNA polymerase.[37] However, the selection involves tedious processes consisted of target immobilization, incubation with nucleic acid library, rinse off non-binders, and elution of binders. Negative selections are needed to get rid of non-specific binders to the membrane; elution of tight binders is kinetically unfavorable; and not all target molecules are adsorbed onto membrane although most proteins are. Modifications to the partitioning method have been made to avoid the selection biases and eliminate the disadvantages, such as centrifugation,[40] magnetic

beads,[41] flow cytometry,[42] and surface plasmon resonance.[43] But none of them can accomplish all at one step. There are only ~150 targets[44] against which aptamers have been selected in the past two decades, mainly because selections are slow and labor intensive.

1.1.3 CE-SELEX

Capillary Electrophoresis (CE) is a well-developed and very powerful separation technique widely applied in Chemistry and Biology. As shown in Figure 1.2 (A), a capillary with i.d. of 20 μm~100 μm and length of 10 cm~75 cm is first conditioned with separation electrolyte, followed by injection of sample of several nL. The capillary ends and electrodes are then placed into the same vials containing the separation electrolyte. Upon applying high voltage, analytes are separated based on the electrophoretic mobility (μ), which is determined in Equation 1.1:

$$\mu = \frac{z}{6\pi\eta R_h} \qquad \textbf{Equation 1.1}$$

where z is the net charge of the analyte, η is the viscosity of the medium, and R_h is the hydrodynamic radius. For an uncoated silica capillary, the silanol groups (Si-OH) on the interior surface are ionized to silanoate groups ($Si\text{-}O^-$) in neutral or basic electrolyte, which causes the formation of a diffuse double layer. When high voltage is applied across the capillary, the solvated cations in the mobile layer drag the bulk electrolyte solution moving towards the cathode to form electroosmotic flow (EOF). In a CE separation, the apparent mobility of an analyte is the result of combined actions from both EOF and electrophoretic

mobility. ssDNA molecules are highly negatively charged, which makes their electrophoretic mobility opposite from the EOF, and thus a very small apparent mobility. This property is significantly different from most targets, whose electrophoretic mobilities are most often much smaller than that of ssDNA, leading to larger apparent mobilities. As a result, bound ssDNA molecules gain substantial mobility shift due to an obvious increase in size and less obvious change in charge. The ssDNA-target complexes migrate faster than the unbound ssDNA molecules as shown in Figure 1.2 (B).

Capillary electrophoresis-based selections (CE-SELEX) address many common issues encountered in traditional selections.[45-48] Selections are performed in free solution, eliminating complicated target immobilization strategies. Isolation is now reduced to two steps: incubation and separation. The free solution separation also reduces the opportunities for non-specific interactions which are common when sieving materials are involved. Most importantly, the high resolving power of CE increases the rate of enrichment and allows high affinity aptamers to be obtained in 2~4 rounds of selection, compared to 8~15 rounds in traditional SELEX. CE-SELEX has been successfully used to generate aptamers for over a dozen of protein and peptide targets.[49-54]

1.2 Microfluidic Devices in SELEX

1.2.1 Automation of SELEX

Laboratory automation is of revolutionary importance to biotechnology. A well-known example is the high-throughput sequencing technology in commercializing

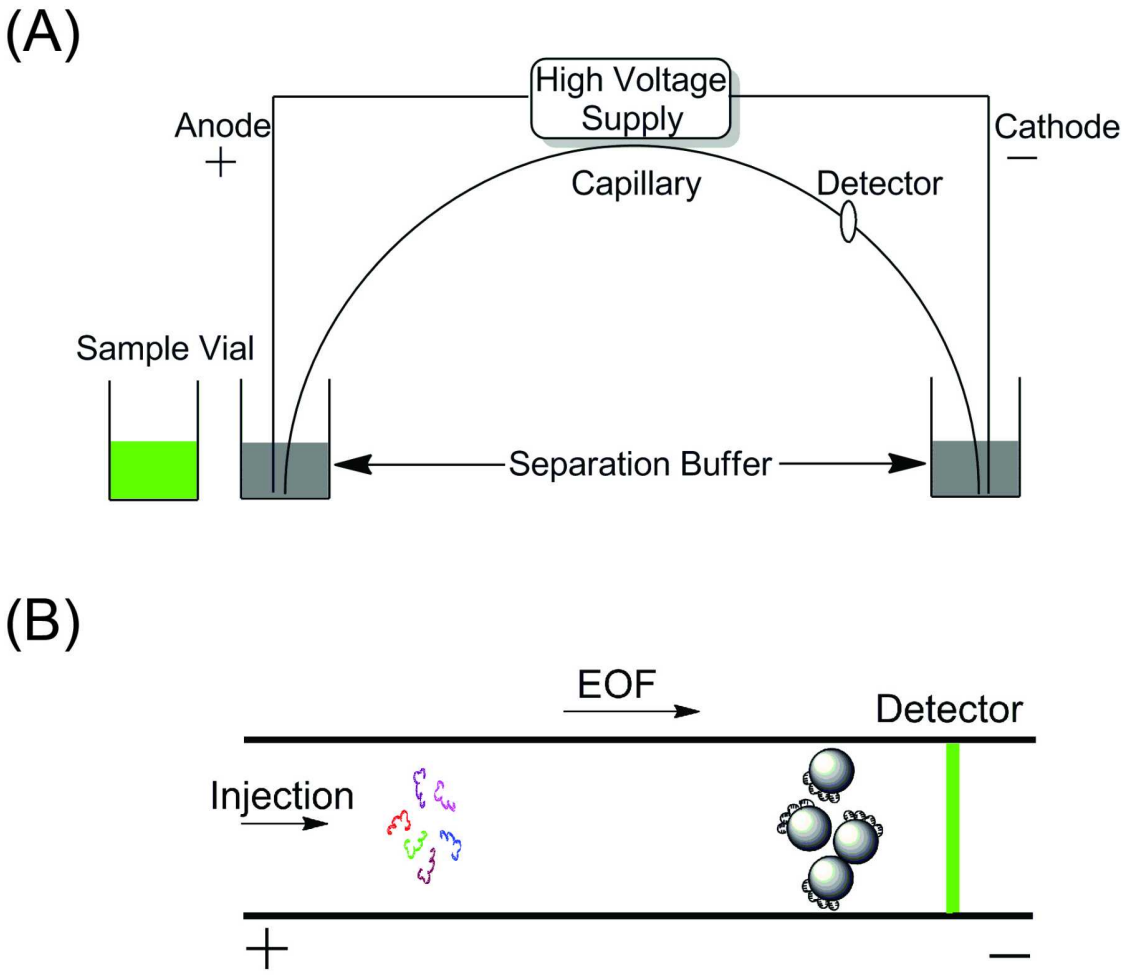

Figure 1.2 Schematics of (A) CE instrumental setup, and (B) CE separation of unbound ssDNA molecules (colored) and ssDNA-target complex molecules (grey).

human whole genome sequencing.[55-62] Automating SELEX is also of great interest for the rapid generation of aptamers. Unlike modifications on partitioning methods that aim at improving the separation efficiency to cut down selection cycles, automation targets at faster turnaround within one cycle in order to reduce the overall selection time.

Cox *et al.* first automated the selection procedure on a Beckman Biomek 2000 pipetting worksurface.[63] Pipette tip arrays and four devices, which included a

magnetic bead separator, a Dynal MPC-auto96 to perform isolation, a PCR thermal cycler, and a Peltier cooler for enzyme storage, were integrated on a platform on which the robot can transport liquid. RNA aptamers were generated toward a $(dT)_{25}$ oligonucleotide target after five rounds of selection. The overall time needed for one cycle of selection was reduced to 212 min, which allowed the selection to be finished within two days. Using slightly modified robotic pipetting worksurface, high affinity aptamers were later selected against protein targets, including lysozyme,[64] CYT-18, MEK1, and Rho.[65]

Automation not only speeds up the turnover rate of selection, but also holds many other important advantages. The lack of manual intervention avoids error in human manipulation and improves consistency; the robotic workface is easier to avoid contamination in sample handling and PCR process; optimization of incubation and PCR conditions and possibly simultaneous selections against different targets are achievable through this high-throughput format.

Later improvement of this technology has been demonstrated in a study isolating a spiegelmer aptamer toward L-substance P with mid-nanomolar K_d.[66] A RoboAmp 4200 E robotic configuration was utilized to allow online monitoring of the PCR process and more parameters to be adjusted, such as incubation time and separation stringency. Besides this study, however, there was little further progress on realizing the potential high-throughput generation of aptamers towards a proteome. This fact is mainly because a robotic platform only facilitates pipetting high-throughput without tackling the complexity in molecular biology. Despite the success of improved turnover rate, it is difficult to keep track of the

separation process, product purity after PCR, or yield after purification. It is only the success in each step and each round that will lead to a higher chance of getting aptamers. Thus, beyond rapid liquid transfer, the functionality, versatility, and flexibility of an automated system are crucial in order to achieve high-throughput SELEX.

1.2.2 Microfluidic SELEX

As appeared in the name, microfluidics precisely manipulate fluid for multifarious purposes with feature designs on sub-millimeter to micrometer scale, at least in one dimension. The concept of micro total analysis systems (μTAS) was introduced by Manz in 1990, predicting a faster separation, a dramatic reduction of sample and reagent, and the potential to perform multiple measurements simultaneously if miniaturizing chromatography or capillary electrophoresis.[67] During the past 20 years, rapid growth in the field of μTAS has been made by integrating multiple processes into miniaturized devices. More and more tasks that used to be fulfilled on bench top are now moved to microfluidics for the sake of cost, speed, and automation.[68-74]

The first automated microfluidic assembly for SELEX was built by connecting functional parts with actuatable valves and microlines.[75] One DNA aptamer against lysozyme, which was previously identified via manual SELEX, was successfully isolated on this platform. Rather than pipetting robots, this microfluidic assembly used pressure gradient to load and move samples through certain reaction loops for specific molecular biology processes. Compared to the

“macro” apparatus, this smaller grafted-microfluidic platform was called a “meso” structure before the advent of the self-contained microfabricated SELEX chip.

Although great progress has been made toward moving PCR from thermal cyclers to microfluidic chips in the past years,[76] it is still at the beginning stage for efficient separation and collection of bound aptamers from the unbound. The pioneering design for isolation is a sol-gel microfluidic device.[77, 78] By screening 100,000 formulations, the sol-gel was optimized to be an excellent sieving matrix for aptamer separation.[79] The nanoporous structure encapsulated large protein molecules and allowed smaller nucleic acid molecules to diffuse through. The optimized sol-gel improved the resistance of protein to chemical and heat denaturation, and also had very low background binding to nucleic acid molecules. Five protein-imbedded sol-gel micro-droplets with a volume of 7 nL were patterned on top of individual aluminum electrical heaters in a glass-PDMS microfluidic device as shown in Figure 1.3 (A).[78] Binders were collected after incubation and non-binders were rinsed off, by heating specific micro-droplet. Aptamers toward yTBP and yTFIIB were selected with nanomolar affinities.[77, 78] By scaling down to micro dimension, the volume for incubation was 0.22 μL, reducing the aptamer diffusion time to several ms. Also, the ratio of washing volume to binding volume was improved by 500 fold compared to conventional filter SELEX, which improved the separation efficiency and enrichment rate. 6 cycles of selection were performed to converge the pool to the best binders, compared with 11 cycles of positive selection and 3 cycles of negative selection by conventional SELEX.[80] Recently, a modification to this microfluidic device was

reported with a key component of a pneumatically-driven PDMS diaphragm valve included.[81] This valve eliminated cross-contamination of the five sol-gel reaction chamber, and the elution process was computationally and experimentally optimized, thus making it one step closer to a multiplex SELEX platform.

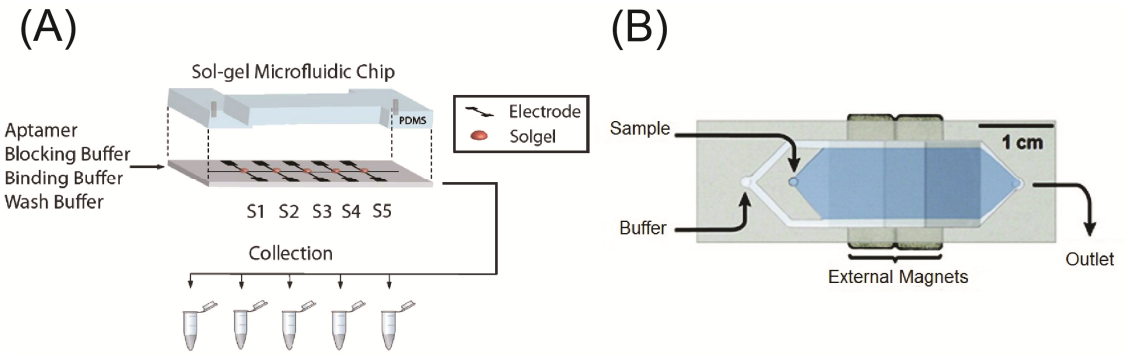

Figure 1.3 (A) Schematic of a sol-gel-based microfluidic device. S1 to S5 represent five sol-gel micro-droplets with targets encapsulated. Reprinted with permission from Park *et al*. **2009** Lab on a Chip. Copyright the Royal Society of Chemistry 2009. (B) Photograph of the MMS microfluidic device. The micro-channel is 12 mm in width and 25 μm in depth. 10 μm nickel strips were patterned on the device with decreasing grid pitch from 200 μm on the left to 100 μm in the middle to 50 μm on the right. Reprinted with permission from Qian *et al*. **2009** Analytical Chemistry. Copyright 2009 American Chemical Society.

Another thriving development of microfluidics in aptamer separation involves using micromagnetic beads. Although magnetic beads lag far behind other separation techniques on bench top due to the delicate manipulation and low separation efficiency,[82, 83] it gains popularity in SELEX when combined with microfluidics. The first continuous-flow magnetic activated chip-based separation (CMACS) device employed the multiple laminar flow architecture and well-designed imbedded Ni strips.[84] The multiple laminar flow architecture prevented the sample from mixing with the buffer, and the Ni strips guided the magnetic beads to a defined route to be separately collected when a highly reproducible magnetic gradient was generated at the interface of Ni strips and buffer solution.

Later this design was replaced with the micromagnetic separation (MMS) chip, in which Ni grids (Figure 1.3 (B)) rather than Ni strips were used.[85] In this design, magnetic beads were trapped on the Ni grids when external magnetic field was applied. Meanwhile the free nucleic acids or week binders were rinsed off the device with controlled buffer composition, flow rate, and time. The binders were later collected when external magnetic field was withdrawn. The elimination of specifically defined route for collecting magnetic beads made this device more robust than the previous design, as air bubbles or magnetic beads aggregation was not a concern. This improved MMS device had excellent beads recovery and separation efficiency was comparable to CE. Recently, this MMS device was successfully applied to phage selection,[86] the discovery of aptamers for thrombin and platelet derived growth factor B (PDGF-BB),[87] and the invention of Volume Dilution Challenge-Microfluidic SELEX (VDC-MSELEX).[88]

1.2.3 Micro Free Flow Electrophoresis (μFFE)

Free flow electrophoresis (FFE) is a continuous separation technique that was introduced in 1960's,[89, 90] and later commercialized in zonal electrophoresis (ZE) and Isoelectric focusing (IEF) modes as a preparative step for many proteomics[91] and cell experiments.[92] FFE is characteristic in that the separation occurs perpendicularly to the flow direction. As demonstrated in Figure 1.4 (A), a thin layer of pressure driven buffer flow is introduced into the device, which moves the analytes down along the chip. Meanwhile, an orthogonal electric field is applied, which triggers the separation of analytes into different sample streams.

Finally the analytes are fractionated into different collection vials at the exit, from which the sample complexity was reduced for further analysis.

Similar to magnetic beads discussed earlier, the application of FFE to sample preparation has been relatively slow due to the low separation efficiency. The reasons include: (i) convective mixing of separation buffer resulting from Joule heating due to the large cross-sectional area, (ii) hydrodynamic broadening gen-

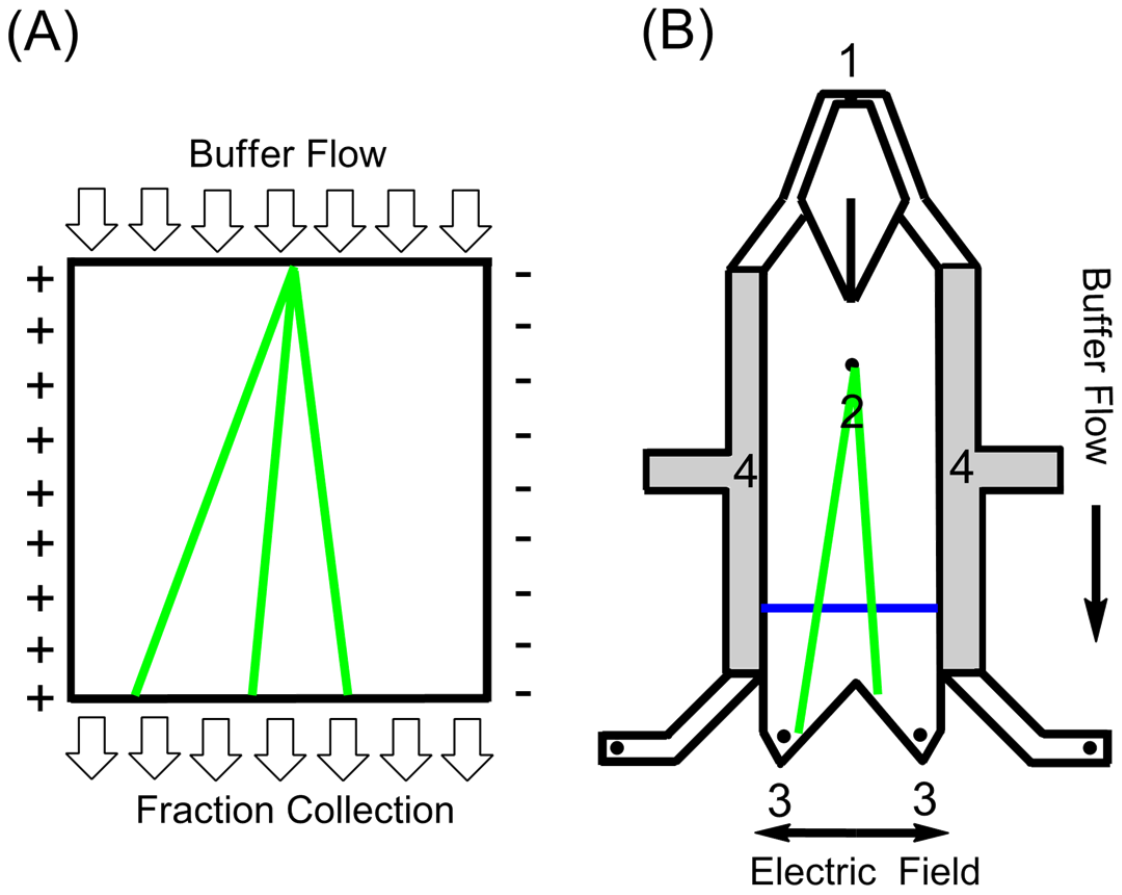

Figure 1.4 (A) An illustration of the separation mechanism of Free Flow Electrophoresis. (B) Schematic of a μFFE device showing the buffer inlet (1), sample inlet (2), fraction collection outlets (3), and electrode channels (4). The dimension of the separation channel is 2.5 cm×1cm×20 μm. The depth of the electrode channels is 100 μm. The blue line denotes the detection zone where the laser line is expanded across the separation channel.

erated from the laminar flow due to different residence time of analytes, and (iii) electrodynamic distortion broadening as EOF is generated in a closed system.[93]

The first micro FFE (μFFE) device made of silicon came into the world in 1994,[94] and the miniaturized chip showed significant improvement as the Joule heating was dramatically reduced. Later improvement in fabrication allowed a whole glass μFFE device to be anodically bonded, on which an electric field as high as 283 V/cm can be applied.[95] Online detection was readily compatible in μFFE, and the volume of buffer needed to perform separation was reduced from several L to mL. A challenge in the μFFE device is that electrolysis bubbles enter the separation channel, disrupting the separation. Inspired by the membrane spacers used in FFE, researchers employed side channel arrays in μFFE to separate the separation and electrode channels.[94-97] However, these long and narrow barriers possessed significant electrical resistance, and allowed only a very small portion of the electric field to be applied across the separation channel. Fonslow *et al.* replaced the physical barriers with a multiple-channel-depth design (Figure 1.4 (B)).[98] As predicted by lubrication theory, the linear flow rate (v) is proportional to the squared channel depth (H):

$$v = \frac{\Delta P H^2}{12 \eta L} \qquad \textbf{Equation 1.2}$$

where ΔP is pressure difference, η is viscosity, and L is the channel length. The authors used a ~20 μm separation channel and ~80 μm electrode channels. A 4-fold increase in the electrode channel depth resulted in a 16-fold increase in the linear velocity, which allowed bubbles to be effectively removed before they

entered the separation channel. This multiple-channel-depth design also reduced electrical resistance in the electrode channels, so that 91% of the applied voltage was on the separation channel.

1.3 Determination of Binding Affinity

1.3.1 Binding Isotherm

One important yet less mentioned procedure in SELEX is to understand the interactions between nucleic acid aptamers and their targets. It is crucial in monitoring the SELEX process and characterizing aptamer affinity. It also helps clarify binding mechanisms and improve rational design of aptamer structure. Although it is desired to have a comprehensive and detailed study of molecular recognition properties,[99] mapping of aptamer and target binding regions,[100] and kinetic studies of induced conformational changes of aptamers and/or targets upon binding,[10] it is fundamental and of great interest to first characterize the binding constant.

A simple binding equilibrium with a 1:1 stoichiometry is described by:

$$A + T \leftrightarrow C \qquad \textbf{Equation 1.3}$$

in which A is the aptamer, T is the target, and C is the aptamer-target complex. The equilibrium can be described using either a dissociation constant K_d (Equation 1.4), or association constant K_a (Equation 1.5).

$$K_d = \frac{[A][T]}{[C]} \qquad \textbf{Equation 1.4}$$

$$K_a = \frac{1}{K_d} = \frac{[C]}{[A][T]}$$ **Equation 1.5**

The most common strategy for measuring K_d or K_a is to titrate a constant concentration of one ligand with an increasing concentration of the other. For example, if the initial aptamer concentration remains constant, then the fraction of bound aptamer (f_a) is given by:

$$f_a = \frac{[T]}{K_d + [T]}$$ **Equation 1.6**

Equation 1.6 takes the general form of a rectangular hyperbola in the first quadrant with one asymptote of f_a equaling 1. K_d can be estimated directly from this binding curve using a nonlinear regression analysis. Although many equations have been introduced in various techniques, all are essentially variant forms of Equation 1.6. A semilogarithmic plot of f_a against log[T] is able to present data over three to six orders of magnitude of target concentration and K_d is easily estimated as the target concentration at half maximum binding fraction. Linear forms of the rectangular hyperbola, such as the double reciprocal plot (Equation 1.7), y reciprocal plot (Equation 1.8), and x reciprocal plot (Equation 1.9) are also useful for estimating K_d using linear regressions.

$$\frac{1}{f_a} = K_d \frac{1}{[T]} + 1$$ **Equation 1.7**

$$\frac{[T]}{f_a} = [T] + K_d$$ **Equation 1.8**

$$\frac{f_a}{[T]} = -\frac{1}{K_d} f_a + \frac{1}{K_d}$$ **Equation 1.9**

Although Equation 1.7 to 1.9 are mathematically equivalent, each equation transforms error differently, affecting K_d as well as the confidence in the estimate. Detailed discussions can be found in Kenneth Connor's book[101] and other articles.[102, 103] Generally speaking, rectangular hyperbola, semi-logarithmic plot, and the x reciprocal plot (also known as the Scatchard plot) are used most frequently. There are advantages and disadvantages associated with each approach and it is recommended to perform data analysis using at least two methods to compare the obtained K_d values. It is also recommended to collect data across the binding fraction range of 0.2~0.8.[103, 104] Plots made without data points in this range often give rise to large errors in K_d estimation.[105]

Direct nonlinear fitting using Equation 1.6 assumes that the concentration of aptamer is significantly smaller than K_d so that the free target concentration [T] at equilibrium does not change significantly compared with the total target concentration $[T]_t$. Figure 1.5 illustrates simulated plots for an RNA aptamer (K_d = 10^{-8} M) titrated with a broad concentration range of protein.[106] The binding curves are not affected by RNA concentration when below the value of K_d (i.e., [RNA] = 10^{-10} M or 10^{-9} M), and the K_d can be estimated as the protein concentration at F_B=0.5. When the RNA concentration approaches or exceeds K_d (10^{-8} M), the curve shifts and K_d can no longer be estimated directly from the plot since the abscissa is no longer representative of free protein concentration. Here the complex concentration should be accurately represented by replacing f_a with $[C]/[A]_t$, and [T] with $[T]_t$ – [C] in Equation 1.6. The accurate expression of free

target concentration [T] is now shown in Equation 1.10, and now Equation 1.6 is modified as Equation 1.11.

$$[T] = [T]_t - 0.5([A]_t + [T]_t + K_d - (([A]_t + [T]_t + K_d)^2 - 4[A]_t[T]_t)^{0.5}) \quad \textbf{Equation 1.10}$$

$$f_a = \frac{1}{1+(K_d/([T]_t - 0.5([A]_t + [T]_t + K_d - (([A]_t + [T]_t + K_d)^2 - 4[A]_t[T]_t)^{0.5})))} \quad \textbf{Equation 1.11}$$

Although Equation 1.11 is much more complicated, it gives a more accurate fitting in cases where the sensitivity of some techniques is not high enough or the K_d of the aptamer is extremely low. Concentration of the titrand aptamer could be comparable to or higher than K_d, and a good estimate of K_d can still be obtained.

Sometimes determining the equilibrium concentration of the protein target is difficult due to complications such as denaturation during purification of recombinant protein, surface adsorption during solution preparation, or degradation during storage. In any of these scenarios using Equation 1.6 or 1.11 would underestimate binding affinity. If protein loss is not negligible, the protein sample, whose concentration does not need to be known exactly if $<K_d$, should be titrated with increasing concentrations of aptamer.[107] Then K_d and total active target concentration $[T]_{ta}$ could be determined via nonlinear regression of Equation 1.12 or linear regression of Scatchard plot as shown in Equation 1.13 where f_p is the bound fraction of target protein.

$$[C] = f_p[T]_{ta} = \frac{[A]}{K_d + [A]}[T]_{ta} \quad \textbf{Equation 1.12}$$

$$\frac{[C]}{[A]} = -\frac{[C]}{K_d} + \frac{[T]_{ta}}{K_d}$$ **Equation 1.13**

Major analytical techniques developed to measure binding constants are summarized in Table 1.1, which includes separation-based techniques such as dialysis, nitrocellulose filter binding, gel electrophoresis, CE, and HPLC, and also

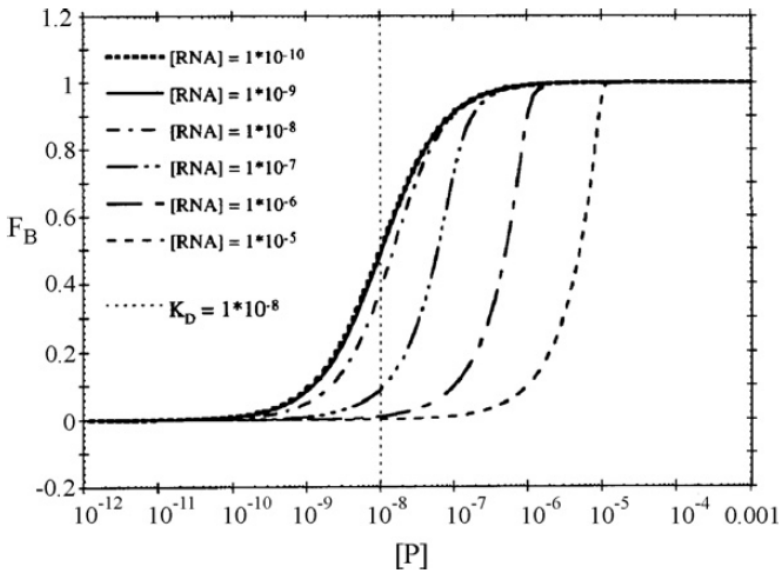

Figure 1.5 Simulated semilogarithmic binding plots of an RNA-protein interaction. In each plot RNA concentration is fixed. Binding stoichiometry is 1:1 and K_d is assumed to be 10^{-8} M. Figure is reprinted from Hall *et al.* **1999** Methods in Molecular Biology. Copyright 1999 Humana Press.

mixture-based techniques such as fluorescence intensity, fluorescence anisotropy/polarization, UV-vis absorption, surface plasmon resonance (SPR), and isothermal titration calorimetry. In the following up two sections, nitrocellulose filter binding, CE, SPR, and fluorescence techniques are reviewed, as they are the most popular techniques and most closely related to the research work covered by this dissertation.

Table 1.1 A summary and comparison of different techniques used to measure aptamer-protein binding

Technique	Sample consumption	Experiment time and complexity	K_d limit	Equipment
Dialysis	~100 μL[a]	~48 h	10^{-8} M[108]	Microequilibrium dialyzer with fluorometer detector[108]
Nitrocellulose filter binding	200 μL~500 μL[a109, 110]	~5 min separation. Radioactive label and scintillation count are needed	10^{-11} M~ 10^{-12} M[111, 112]	Filters and centrifuge, or membranes and vacuum apparatus, and scintillation counter
Gel electrophoresis	10 μL~50 μL[a]	~3 h separation. Gel casting, radioactive labeling, autoradiography, gel cutting and counting are needed	10^{-13} M[113, 114]	GE power supply and autoradiography
CE	~10 μL[a] in zonal separation	~10 min per sample injection	10^{-9} M[115]	Automated CE system with LIF detector
HPLC	100 μL~8 mL[a] in FA, dependent on column type[116]	10~30 min per sample injection	10^{-6} M[23]	Automated HPLC system with UV detector
Fluorescence intensity	~150 μL[117]	~10 min per sample, waiting is needed after each titration	10^{-10} M[118]	Spectrofluorometer and cuvette, or Microplate reader
Fluorescence anisotropy/polarization	15 μL~25 μL[a],	~10 min	10^{-9} M[119]	Spectrofluorometer with polarizers and quartz cuvette, or Microplate reader
	~150 μL[117, 120]	~10 min per sample, waiting is needed after each titration		

UV-vis absorption	15 μL~25 μL[a],[121] ~2.5 mL	~10 min ~10 min per sample, waiting is needed after each titration	10^{-6} M[121]	Spectrophotometer, quartz cuvette
Surface plasmon resonance	10 μL~20 μL[a]	~20 min per sample, aptamer immobilization step is needed	10^{-12} M[122]	Biacore SPR instrument and sensor chip
Isothermal titration calorimetry	~1.5 mL in sample cell; 200 μL~ 500 μL for total injection volume[123]	1.5 h~3 h	10^{-8} M~10^{-9} M[121]	ITC calorimeter

8 to 25 samples are usually required to fit a binding isotherm and obtain a K_d value.

[a] Volume is for each sample.

1.3.2 Techniques to Obtain K_d: Separation Based

In ultrafiltration the binding mixture consisted of aptamer and protein target is driven through the membrane by pressure, vacuum, or centrifugal force. The most widely used membrane material is nitrocellulose, which possesses long pores allowing nucleic acids to pass through but retaining proteins as small as 2 kDa by hydrophobic adsorption.[124] Membranes are available in several traditional formats but "dot blot" apparatus are most commonly used due to the minimal material requirement. ^{32}P-labeled aptamer samples incubated with the protein target at various concentrations are pulled through the membrane using vacuum.[125, 126] Free protein and the aptamer-protein complex stayed on membrane so that the bound fraction is quantified in the membrane using a phosphor-imager. Alternatively, bound fraction can be obtained by counting scintillation from the free aptamer in the filtrate.

Nitrocellulose filter binding is very common and popular in measuring K_d due to the fast speed, high sensitivity and low cost. Radioactive labeling allows aptamer as low as 10^{-12} M to be used, convenient and accurate for aptamers whose K_d's are in the low pM range.[125] There are a few complications to consider before picking this method though. In some cases incomplete retention of aptamer-protein complex by the nitrocellulose membrane has been observed, giving rise to an over estimate of K_d. This could be due to a competitive binding between aptamer and membrane for the same binding site on protein;[127] or aptamer interfering with the interaction between the membrane and an unstructured

hydrophobic region of protein.[128] Another concern is nonspecific adsorption of nucleic acids to nitrocellulose membrane, which requires control experiments to be performed in the absence of protein and background signals to be subtracted before calculating the binding fraction. This is a particular concern for aptamers that have been isolated using nitrocellulose membrane. In addition, protein concentration should be lower than 10^{-5} M to prevent saturation of binding sites on nitrocellulose membrane.[128] Radioactive labeling introduces additional safety and regulatory concerns; however, alternative detection methods with non-radioactive labels have also been developed.[129, 130]

CE techniques have been developed to study protein and nucleic acids interactions.[116, 131-135] Various approaches have been demonstrated including free zone, frontal analysis (FA), Hummel-Dreyer (HD), affinity CE (ACE), vacancy peak (VP), and vacancy affinity CE (VACE).[136-140] Most of them are derived from similar techniques in HPLC analyses of protein-nucleic acid interactions, and covered in several excellent reviews.[131, 141] Here the discussion is focused on zonal CE techniques.

The dissociation rate of aptamer-protein complexes is often slower than the 5~20 min separation time period typical for CE. Protein and aptamer-protein complex generally have slower mobilities than the unbound aptamer, giving rise to distinct peaks. If a fluorescently labeled aptamer is titrated with increasing concentrations of non-labeled protein, two peaks representing the complex and unbound aptamer are observed in the electropherogram. The bound fraction can be calculated from the intensity of the free ligand peak when increasing the

amount of protein in titrant. This technique has been used to characterize aptamers selected via CE-SELEX, a process in which only nucleic acid sequences that stay bound to the target for approximately the same separation time would be collected to evolve aptamers.[45-47] Figure 1.6 inset shows an overlay of electropherograms in which the FAM-labeled free aptamer peak decreases with increasing human IgE.

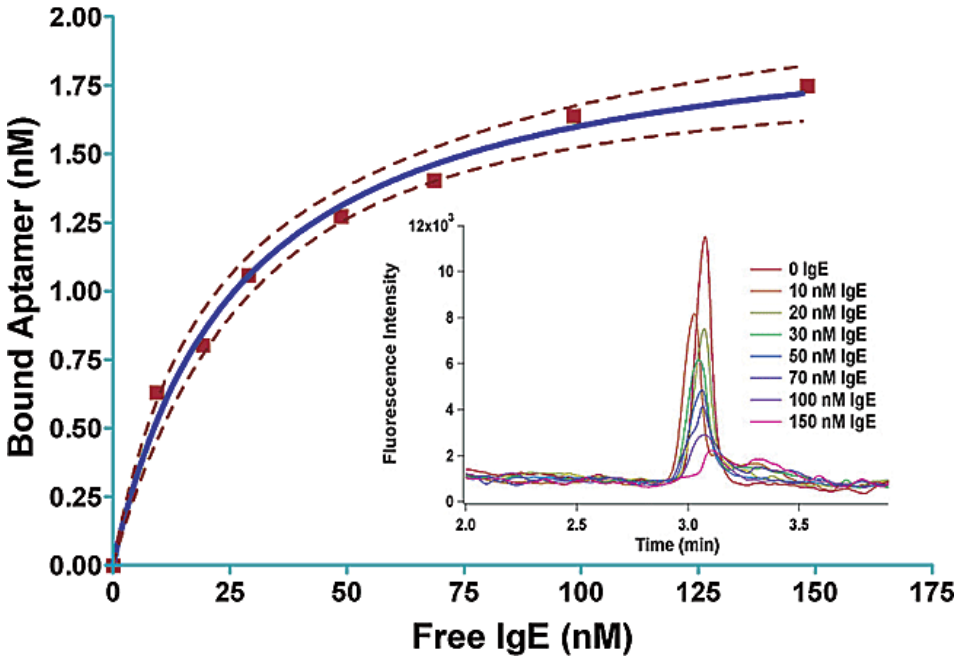

Figure 1.6 Binding Curve measured using zonal CE for an aptamer selected to bind human IgE. Dashed lines represent 95% confidence of the fit curve. The figure inset shows fluorescence intensities of free aptamer peaks at different human IgE concentrations. Reprinted with permission from Mendonsa *et al*. **2004** Journal of the American Chemical Society. Copyright 2004 American Chemical Society.

If the life time of aptamer-protein complex is on a time scale similar to that of the separation time, observed peaks will be distorted as the complex dissociates on the column during the separation. While this may be detrimental to a direct K_d measurement; the distortion of the peaks yields additional kinetic information, from which K_d could also be obtained.[142] To avoid assessing areas of unresolved

peaks, the first statistical moment of the aptamer signal across the entire electropherogram could be used to obtain K_d as well.[143, 144]

1.3.3 Techniques to Obtain K_d: Equilibrium Based

SPR is a phenomenon that occurs on a metal film when an incident light is polarized parallel to the plane of incidence. The evanescent electromagnetic field generated by electron-plasmon oscillation penetrates the metal film into sample solution near the film. When the light source and metal film are fixed, the change in reflected angle is only related to the refractive index of sample solution near the film. It is found that the signal linearly correlates with the immobilized analyte across a broad concentration range.[145, 146] Normally, aptamers are immobilized on the metal surface because large targets such as proteins can usually induce a more significant change in refractive index upon binding.[147-149] As shown in Figure 1.7 , when IgE was injected and flowed over the immobilized aptamer, binding changed the surface refractive index and the resonance signal increased; after the binding process reached the steady state a denaturing buffer was introduced at 900 s and dissociation occured.[150] Both the association and dissociation rate constants could be obtained from the exponential curves representing association and dissociation processes, and K_d could be further calculated by the ratio of the two. Meanwhile, the steady state signal at different protein concentrations can also be used to calculate binding fractions to obtain K_d.

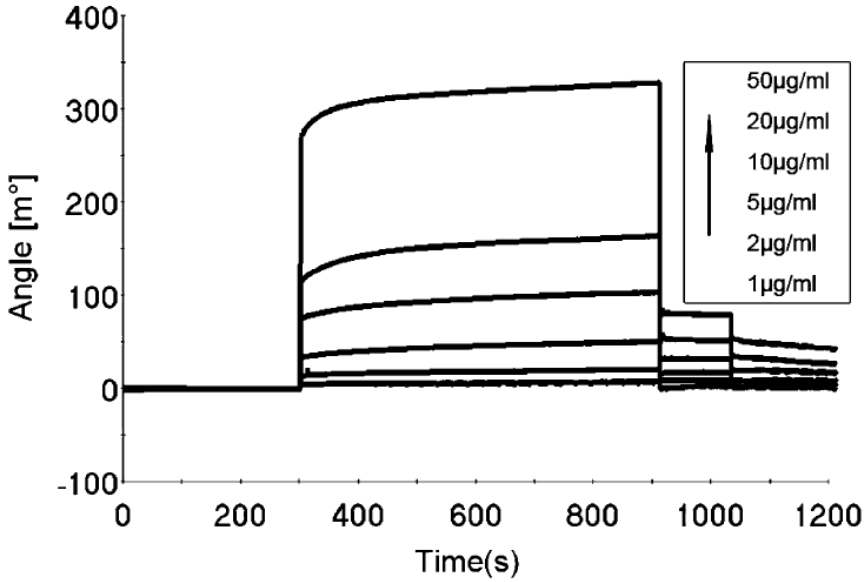

Figure 1.7 An overlay of SPR sensorgrams of an aptamer-IgE interaction. Thiol-derived aptamer was immobilized on a gold chip, titrated with IgE at different concentrations. Reprinted with permission from Wang *et al*. **2007** Analytical and Bioanalytical Chemistry. Copyright Springer-Verlag 2007.

SPR is a very robust technique to characterize aptamer-target binding. The detection is label free, both thermodynamic and kinetic information can be obtained simultaneously within 20 min, and K_d can be estimated by two methods.[145, 146] However, cautions should be taken to ensure that mass transfer of the target molecules from bulk solution to the surface of the SPR chip is not a limiting factor, which otherwise would slow down the observed association rate and give rise to an overestimate of K_d.[151-153] This can be conveniently checked by evaluating the association rate obtained at different flow rates. Faster mass transfer can be achieved by applying higher flow rate, using a thinner flow cell, or lowering aptamer surface density. In addition, immobilization methods should be optimized and K_d should be assessed by complementary techniques to confirm that aptamer affinity remains unchanged when immobilized onto a surface.

Fluorescence-based techniques are also very popular in K_d measurement, possessing many characteristics that are complementary to other techniques. Fluorescence signals are taken while the aptamer and target remain at their equilibrium concentrations. This not only avoids concerns regarding kinetic processes, but also allows K_d to be measured under various buffer conditions, which may be limited in certain separation techniques.[154] For example, increasing the ionic strength of binding buffer may affect the interaction between complex and nitrocellulose membrane, or increase Joule heat in ACE. Fluorescence based methods are generally much faster than separation based methods. There are hundreds of fluorescent probes available with wide ranges of lifetime and excitation/emission wavelength. Labeling procedures are well characterized, especially for nucleic acids.[155-159]

It is not uncommon to observe a change in fluorescence intensity when the labeled aptamer binds to its target. The microenvironment of the dye in aptamer-target complex could be dramatically different from that in free solution, resulting in a change in fluorescence quantum yield. Many dyes are sensitive to changes in buffer polarity. For example, the quantum yield of fluorescein increases in less polar environments.[160, 161] Quenching due to energy transfer to the binding partner is also possible.[162] In another scenario, binding may involve a change of aptamer conformation, which promotes the formation of a donor-acceptor complex at the ground state, resulting in a shift in the fluorescence emission profile.[163] Either fluorescence enhancement or quenching could be possibly observed depending on the specific system under study.[164-167] Practically, extra

care should be taken if the buffer composition of the target is different from that of the aptamer, since even a small change in pH or buffer additives could induce a change in fluorescence intensity, generating an artificial “binding curve”.

A derivative method employs aptamer beacons which incorporate both a fluorescent molecule and a quencher in close proximity in the native state.[168] Binding to the target induces a conformation change in the aptamer, separating the fluorescence dye and quencher and thus increasing fluorescence signal. This method requires pre-knowledge of the aptamer structure and is only applicable in cases where a suitable conformation change is observed. Similarly, if the quencher is replaced with another dye whose excitation spectrum overlaps the emission spectrum of the fluorescent dye, then a change in Förster Resonance Energy Transfer (FRET) could be used to monitor binding process.[169-171]

In Fluorescence Polarization (FP) only fluorescent molecules whose dipoles are in the same orientation as the polarized incident light are excited. If the dye rotates slowly in comparison to the emission kinetic, the polarity is retained when the dye emits. Intensities parallel and perpendicular to the incident light are used to assess polarization. Equation 1.13 describes how the lifetime and rotational diffusion of a fluorescent molecule affect polarization:[172]

$$\frac{1}{P}-\frac{1}{3}=\left(\frac{1}{P_o}-\frac{1}{3}\right)\left(1+\frac{RT}{\eta V}\tau\right) \qquad \textbf{Equation 1.13}$$

where P is the observed polarization, P_o is the limiting polarization when there is no molecular rotation, τ is fluorescence lifetime, R is universal gas constant, T is absolute temperature, η is buffer viscosity, and V is the molar volume of the dye.

$3\eta V/RT$ is Debye rotational relaxation time, which is directly proportional to V. The molar volume increases upon aptamer binding to its target, resulting in less rotation and an increase in polarization. The definition of anisotropy (A) and polarization (P) are given in Equation 1.14 and 1.15:[159]

$$A = \frac{I_{II} - I_{\perp}}{I_{II} + 2I_{\perp}} \qquad \textbf{Equation 1.14}$$

$$P = \frac{I_{II} - I_{\perp}}{I_{II} + I_{\perp}} \qquad \textbf{Equation 1.15}$$

where I_{II} and $I_{\perp}$ represents emission intensities parallel and perpendicular to the incident light, respectively. Anisotropy and polarization have similar physical meaning and are mathematically interchangeable.

It is important to match up the fluorescence lifetime of a dye with its rotation rate. If the lifetime is too short, polarization is high for both the free ligand and complex; if the lifetime is too long, even the large complex could depolarize and have similar polarization as the free ligand. It is recommended that fluorescein with a relatively fast lifetime of 4 ns be used for aptamer-protein complexes smaller than 100 kDa.[173, 174] Long linkers between dye and aptamer should be avoided, because the flexibility of this attachment adds rotational freedom and risks the change in polarization upon binding. G factor is an instrument parameter that measures the ratio of parallel and perpendicular signals recorded for non-polarized light. Ideally G should be 1, but polarization and anisotropy should be corrected when it is not.[175] Fluorescence enhancement or quenching

associated with binding may further complicate measurements since it directly affects the quantum yield and fluorescence lifetime of the dye. For fluorescence enhancement due to an increase in quantum yield upon binding, the complex and then the bound fraction are overestimated, leading to a smaller K_d. To correct this error, the maximal fluorescence enhancement factor Q_m is used in order to get the accurate bound fraction f_a from the apparent bound fraction F_a [163]

$$Q_m = \frac{I_m - I_o}{I_o} \qquad \textbf{Equation 1.16}$$

$$F_a = \frac{P - P_f}{P_b - P_f} \qquad \textbf{Equation 1.17}$$

$$f_a = \frac{F_a}{1 + Q_m(1 - F_a)} \qquad \textbf{Equation 1.18}$$

where I_m and I_o are the fluorescence intensities of the complex and free aptamer. P, P_f, and P_b are polarizations of the sample, free aptamer, and complex, respectively.

1.4 SELEX on Membrane Proteins

1.4.1 Membrane Proteins as Targets in SELEX

Membrane proteins play important roles in cellular functions. They maintain the structure and stability of cells, adhere cells to extracellular environment and other cells, transport mass in and out, and serve as ion channels that are prominent in signaling processes.[176] Mutations in membrane protein primary structure and

changes in expression level lead to misassembly, and subsequently abnormal trafficking and other functional impairments.[177-179] Membrane proteins are related to a variety of human diseases such as neurodegeneration, angiogenesis, and cancers. 70% of the therapeutic targets are membrane proteins.[180]

Chemotherapies have been in use for decades, whose toxic side effect to patients, however, is still a serious problem. This is mainly due to a lack of specifically targeted delivery of cytotoxic drug to the tumor cells without affecting normal healthy tissue. Recent progress in the field involves biochemical synthesis of monoclonal antibody (mAb)-linked drug conjugates with much improved specificity to certain receptors on cell surface and much reduced systemic toxicity.[181, 182] One example is the already approved humanized anti-CD33 antibody-alicheamicin conjugate Mylotarg, for curing acute myeloid leukemia (AML).[183] However, compared to aptamers antibodies have limitations as we have discussed in Chapter 1.1.1. The laborious production, batch to batch difference, large size, and immunogenicity make monoclonal antibodies not as promising as aptamers as drug carriers.

1.4.2 Current Approaches to Generating Aptamers for Membrane Proteins

Although often used in SELEX, nitrocellulose membrane filter is not suitable for generating aptamers for membrane proteins. This is because the lipids or surfactants used to solubilize membrane proteins interfere with the hydrophobic interaction between protein and filter, making target retention a big problem.

Similarly, other selection methods involving target immobilization, such as affinity chromatography, are not compatible with membrane protein targets either.

Aptamer selections against membrane proteins are much more difficult compared to those against soluble proteins, mainly because it is implausible to obtain pure and, at the same time, intact targets. The most popular and straightforward method is to use a peptide region or recombinant protein of the soluble ectodomain as the target. In this way, the selection has no difference from traditional SELEX. Aptamers have been selected for a great deal of targets on cell surface, such as human epithelial growth factor receptor-3 (HER-3), [184] plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1), [185] cytotoxic T cell antigen-4 (CTLA-4),[186] and a lot more.[187-189] An issue associated with this method is that aptamers selected against the truncated targets sometimes do not have similar affinities towards the intact ones. This discrepancy could be due to the post translational modifications which alter the surface chemistry or the structure of the target, or the lack of accessibility to a particular binding site.

Therefore, methods that are aimed at using targets in their native states start gaining more attention. Living cell lines,[190-192] or membranes containing the target of interest prepared by lysing cells[193-195] were directly used as the target. The main partitioning method in these selections was centrifugation,[191, 192] although other methods such as gel-shift selection have been applied as well.[194]

Cell-SELEX is uniquely advantageous in that prior knowledge of the identity of the target is not necessary. Aptamers have been selected against membrane protein biomarkers for specific cancer cell lines without knowing what to target

before selection.[17, 196-200] This selection method has been proved successful in novel biomarker discovery, and aptamers have been applied as excellent probes in diagnoses and therapeutics.[194, 195, 201-205] There are also complications associated with complex-target selection. Since the target is not pure, there is always a concern of nucleic acids binding to membrane skeleton or other proteins expressed on cell surface. One solution is to perform negative selections using the same cell membrane harboring no target protein;[190, 195] alternatively, counter selections were performed where inhibitors of target were used after incubation to replace nucleic acids that only had affinity to the target.[191, 192, 194] Another problem of this selection is the poor cell viability. Selections performed using centrifugation had a high risk of getting dead cells after separation. These dead cells usually have a high nonspecific background binding to nucleic acid library, which enriches more unwanted sequences and slows down aptamer generation. To overcome this problem, fluorescence-activated cell sorter (FACS) was used after binding to separate viable cells from dead ones, because the light scattering characteristic of the two are different.[206] Moreover, unlike pure protein target, target level in membrane structures is usually unknown and difficult to quantify. Inefficient target expression or existence of other abundant interferents would substantially affect selection, even making selection off track. To address this issue, cross-over SELEX was invented, combining selections using both complex target and the pure counterpart, in which the latter guaranteed the selection was in the right direction.[19, 207] This method, however, usually generates aptamers whose affinities are lower towards targets in native state.[207]

1.4.3 Sarco/ Endoplasmic Reticulum Ca^{2+} ATPase (SERCA)

Calcium is the most important ion in cardiac contraction and relaxation.[208] Ca^{2+} enters cardiac cell and is released from a subcellular organelle Sarcoplasmic Reticulum (SR). These two processes lead to an increase of free Ca^{2+} in cell cytosol, which switches on the contractile machinery, resulting in ventricular pressure and blood ejection. Shortly after this, Ca^{2+} is transported out of cytosol by four pathways, which include mitochondrial Ca^{2+} uniport, through sarcolemmal Ca^{2+}-ATPase, sarcolemmal Na^{+}/Ca^{2+} exchange, and through SR Ca^{2+}-ATPase, The lowered Ca^{2+} level allows the relaxation of cardiac myocytes. Amongst all four pathways, SR Ca^{2+}-ATPase plays the most important role, removing 70% to 90% Ca^{2+} varying in species.[209] The SR from rabbit skeletal muscle has a very similar functional behavior,[210] owing to the presence of a membrane bound protein SERCA in both cardiac and skeletal muscles.

Transport of Ca^{2+} by SERCA opposite to the ion gradient in SR membrane is a complicated process, which involves several structural changes and conformational transitions.[211, 212] The simplified version is shown in Figure 1.8. When in E1 state, the enzyme binds to Ca^{2+}, which triggers protruding cytoplasmic and partial transmembrane domains, allowing phosphorylation of the conserved Asp^{351} in the cytoplasmic domain by ATP. The enzyme then transforms to E2 state, which has the opened ion channel and a lower Ca^{2+} binding affinity. Ca^{2+} is then released into the lumen of SR, followed by hydrolysis of the phosphate to an unphosphorylated E2 state. A decrease in SERCA functionality is observed in almost all types of heart failure.[213] Besides, other

diseases associated with abnormal muscular conditions such as Brody's disease and Darier's disease have been recognized to be closely related with the genetic mutations in SERCA genes.[214]

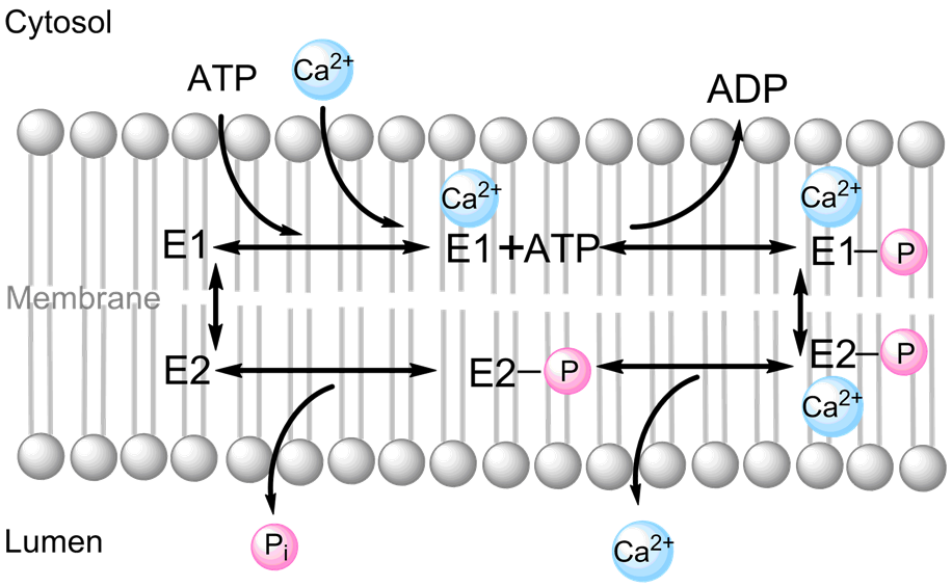

Figure 1.8 Schematic of the SERCA's enzymatic cycle for the transport of Ca^{2+} from cytosol to lumen, the inside of SR.

Phospholamban (PLN) is a small membrane protein co-localized with SERCA in cardiac SR. At low Ca^{2+} concentration, dephosphorylated PLN interacts with SERCA in both cytoplasmic[215] and transmembrane domains[216] when SERCA is in the E2 state, resulting in a prolonged time of conformational transition and a decrease in apparent Ca^{2+} binding affinity by SERCA.[217] The consequence is a decreased SERCA activity and Ca^{2+} store rate in SR. When β-agonists such as adrenaline are released into blood, the signal-transduction pathway in myocytes is initiated. cAMP-dependent protein kinase (PKA) is activated, which phosphorylates PLN and relieves its inhibition to SERCA. The consequence is an increased SERCA activity and heart contractility.

1.5 Scope of the Dissertation

Applications of aptamer in research and for clinical purposes have been growing explosively. Unmatched to this trend is the relatively slow and inefficient aptamer generation process. Recently, effort has been made to profile one advanced separation platform, capillary electrophoresis, in terms of target size, complexity, library injection size, and selection round by Mendonsa *et al*.[45-47] and Mosing *et al*.[48] However, a detailed characterization of the selection process that looks into the aptamer abundance change or the evolution profile is still yet to be probed, which prevents the fundamental understanding of the CE-SELEX process and improving of the SELEX platform. Chapter 2 of this thesis describes the application of 454 pyrosequencing on generating sufficient sequence information, and the subsequent analyses on desired selection products and PCR byproducts. Chapter 3 then presents the first successful integration of a free-solution-separation based microfluidics μFFE and SELEX, with mid-nM K_d aptamers generated right after one cycle of selection. Chapter 4 demonstrates the progress on this μFFE device being applied to isolate aptamers for a membrane protein target, SERCA. The discovery of ssDNA sequences binding to PLN and the factors affecting binding affinity is also discussed in Chapter 4. Finally, Chapter 5 describes the development of a novel activity assay that directly monitors ADP produced by SERCA using competitive immunobinding and Time Resolved-Förster Resonance Energy Transfer (TR-FRET) detection.

Chapter 2 Emergence of High-Affinity Aptamers for $rhVEGF_{165}$ during CE-SELEX Using High Throughput Sequencing

2.1 Summary

CE-SELEX is a powerful technique in developing aptamers toward various targets, from large proteins to small peptides with molecular weights of several kilodaltons. One of the unique characteristics of CE-SELEX is the relatively high heterogeneity in the ssDNA pools even after multiple rounds of selection. Enriched sequences or highly abundant oligonucleotide motifs were rarely reported in CE-SELEX studies. In this work, we employed 454 pyrosequencing to profile selected pools when performing CE-SELEX against the target $rhVEGF_{165}$. Information of up to 3×10^4 sequences were obtained from each selected pool and the control of the unselected library. Three contigs were identified with the highest abundance of 0.8% of contig 1 in pool 4. Analyses on contigs, top 1000 oligonucleotides, and eight original FASTA files indicated the lack of prevailing motifs in the desired products. Sequencing results also revealed the cause of getting pools with reduced affinities after pool 4 was the preferred amplification of one particular short byproduct in PCR. High-affinity aptamers with 10^{-8} M dissociation constants were identified. Both abundant and randomly picked aptamers had similar affinities as those of the selected pools, characterized by both ACE and FP. These sequences also had comparable affinities to VEGF DNA aptamers reported before.

2.2 Introduction

The merits of CE-SELEX over traditional SELEX have been discussed in Chapter 1.1.3. One distinctive property of CE-SELEX is the retained

heterogeneity of the selected pools. Although enriched sequences and conserved oligonucleotides (motifs) were identified in one report after 9 rounds of selection,[50] to the best of our knowledge, other studies have not shown enriched sequences or motifs at all.[45-49, 51-54] Typically, high picomolar to low nanomolar target concentrations were used in CE-SELEX, corresponding to 10^5~10^7 target molecules per injection. Theoretically, one nucleic acid molecule should not have an identical copy in the starting library, since the diversity of a library consisted of 40 nt random region surpasses the number of sequences used in selection substantially. In cases where aptamers are generated after one round, about the same amount of nucleic acid molecules are bound to the target. As a result, there are about 10^5~10^7 different aptamers generated in selections that are done after one cycle. Indeed, when selection round number goes up, the diversity in the selected pools decreases; however, tens of sequences from cloning are insufficient to characterize the heterogeneity. Another interesting but puzzling fact of CE-SELEX is that sometimes affinities of the selected pools started to decrease after a certain selection cycle.[45-48] Again, a massive sequencing data is of great importance to explain the phenomena. Additionally, although computational simulations on evolution of aptamers have been proposed,[218-221] few studies have been reported on tracking the real evolution profiles of individual aptamers. Thus, a comprehensive study on aptamer emergence with experimental sequencing data allows us to look into CE-SELEX process, helps us to develop in-depth understanding, and may shed light on prospective improvements and applications.

The revolutionary Next Generation Sequencing (NGS) technologies can produce enormous amount of data at low unit prices compared to the "first generation" Sanger sequencing.[222, 223] These new technologies triggered the explosively growing research in *de novo* assemblies of genome,[57, 224] transcriptome sequencing,[225] and gene discovery in metagenomics.[226] More interestingly to us, NGS has also been used in several SELEX studies, playing powerful roles in Genomic-SELEX[227] and Cell-SELEX[228] to identify aptamers, and helping generate higher affinity aptamers in earlier selection rounds by monitoring enrichment fold of individual sequences in microfluidic-SELEX.[229] 454 pyrosequencing is one of the NGS technologies allowing massively parallel sequencing of up to millions of sequences.[230, 231] Briefly, a single-stranded template DNA library is first prepared with the adaptor primer on and then immobilized on beads for an emulsion PCR (emPCR). Millions of identical copies of the same sequence are amplified on individual beads. These beads are layered onto a PicoTiterPlate (PTP) device with special designs to fit one bead per well. The sequencing-by-synthesizing (pyrosequencing) step is then performed and up to millions of pyrograms are generated at precisely recorded locations on the device.

Vascular endothelial growth factor (VEGF) is a signaling protein in angiogenesis. Over expression of VEGF is associated with many diseases such as Age-related Macular Degeneration (AMD), rheumatoid arthritis, and solid cancers.[232-235] The importance of VEGF and the advancement in aptamers being drug candidates can be evidenced by the first aptamer-based drug Macugen,

which was derived from concrete and abundant studies on the generation of anti-$VEGF_{165}$ aptamers. RNA aptamers[236, 237] and directly selected or post-selection modified 2'-O-methyl RNA aptamers[238, 239] have been developed using conventional SELEX. DNA aptamers were also selected against $VEGF_{165}$[240-243] by conventional SELEX and against a 32 aa peptide fragment[244] using CE-SELEX. From these studies, selected pools converged to several sequences and motifs were found when conventional SELEX methods were applied, but high diversity remained when CE was chosen as the partitioning method. With this knowledge in mind, we decided to use $VEGF_{165}$ as the target and 454 pyrosequencing as the characterization method in our study to probe the evolution process of CE-SELEX and compare our results with previous studies.

2.3 Experimental Methods

2.3.1 Materials and Chemicals

Recombinant Human Vascular Endothelial Growth Factor 165 ($rhVEGF_{165}$, MW 44 kDa) was purchased from ProSpec Protein Specialists, Inc. (East Brunswick, NJ). Nuclease free H_2O, forward primer 5'-FAM-AGC AGC ACA GAG GTC AGA TG-3', reverse primer 5'-Biotin-TTC ACG GTA GCA CGC ATA GG, fusion primer A 5'-GCC TCC CTC GCG CCA TCA GAG CAG CAC AGA GGT CAG ATG-3', fusion primer B 5'-GCC TTG CCA GCC CGC TCA GTT CAC GGT AGC ACG CAT AGG-3', random ssDNA library 5'-FAM-AGC AGC ACA GAG GTC AGA TG $(N)_{40}$ CCT ATG CGT GCT ACC GTG AA-3', and individual aptamers were from Integrated DNA Technologies, Inc. (Coralville, IA). 100 mM dNTP set and 25 bp

DNA ladder were purchased from Invitrogen, Inc. (Carlsbad, CA). *Taq* DNA polymerase and ThermoPol buffer were purchased from New England BioLabs, Inc. (Ipswich, MA). Blue/orange gel loading dye was purchased from Promega Corp. (Madison, WI). Streptavidin agarose resin was purchased from Thermo Scientific, Inc. (Waltham, MA). Acetic acid (CH_3COOH, 99.7%, Mallinckrodt Baker, Saint Louis, MI), boric acid for electrophoresis (H_3BO_3, Sigma-Aldrich, Saint Louis, MO), DL-Dithiothreitol (DTT, MP Biomedicals, Irvine, CA), ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA, 99.0%-101.0%, Sigma-Aldrich, Saint Louis, MO), glycine ($C_2H_5NO_2$, 99%, Sigma-Aldrich, Saint Louis, MO), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 99%, Alfa Aesar, Ward Hill, MA), magnesium chloride 6-hydrate ($MgCl_2 \cdot 6H_2O$, 99.8%, Mallinckrodt Baker, Saint Louis, MI), potassium phosphate, monobasic, (KH_2PO_4, 99.9%, Mallinckrodt Baker, Saint Louis, MI), sodium chloride (NaCl, 99.0%, Mallinckrodt Baker, Saint Louis, MI), sodium hydroxide (NaOH, 97%, Sigma-Aldrich, Saint Louis, MO), and tris(hydroxymethyl) aminomethane (tris, 99.8%, Sigma-Aldrich, Saint Louis, MO) were used without further purification. Binding buffer, CE separation buffer, and streptavidin buffer were prepared in nuclease free H_2O, and 0.5× TBE buffer were prepared in Milli-Q H_2O from a Milli-Q water purification system (Millipore Corp., Bedford, MA). Binding buffer consisted of 20 mM HEPES, 10 mM NaCl, and 0.25 mM DTT at pH 7.2 (pH was adjusted by 1M NaOH solution). CE separation buffer consisted of 25 mM tris, 192 mM glycine, and 5 mM KH_2PO_4 (TGK) at pH 8.3. Streptavidin binding buffer consisted of 10 mM tris, 50 mM NaCl, and 1mM EDTA at pH 7.5. 0.5× TBE buffer consisted of 45

mM tris, 45 mM boric acid, and 1 mM EDTA at pH 8.3. Each buffer was filtered through a 0.2 μm posidyne® membrane filter before use.

2.3.2 Capillary Electrophoresis Selections

All CE selections were performed on a P/ACE™ MDQ Capillary Electrophoresis System from Beckman Coulter, Inc. (Fullerton, CA). An argon ion laser (Beckman coulter, Inc., Fullerton, CA) with a 488 nm line for excitation was used and emission of fluorescence at 520 nm line was collected. A 50 cm long (40 cm to the detector) uncoated fused-silica capillary with 50 μm i.d. and 360 μm o.d. (Polymicro Technologies, Phoenix, AZ) was prepared each day before separation. The capillary was rinsed with 0.15 M NaOH and then CE separation buffer at 30 psi, each for 5 min before the first separation, and 1 min for the follow-up separations of the day. ssDNA library dissolved in binding buffer was heated to 72 °C for 5 min, let cool down to room temperature, and then incubated with VEGF for 30 min. In the first round of selection, the mixture contained 10 μL of the random ssDNA library at 100 μM and VEGF at 10 nM. The mixture was hydrodynamically injected at 1 psi for 4 seconds, corresponding to approximately 5×10^{11} ssDNA molecules and 5×10^{7} VEGF molecules. The separation was performed at 30 kV under normal separation condition at 25 °C. The cutoff time was calculated based on the apparent cutoff time from the electropherogram and the ratio of the effective capillary length to the total capillary length. Complexes migrating 45 sec before the rise of the free peak were collected at the capillary outlet into a 48 μL separation buffer vial for PCR amplification. In the subsequent

selections, ssDNA library was at the largest amount available after being aliquoted for binding experiments, and VEGF was kept at 10 nM.

2.3.3 PCR Amplification

943 µl of PCR master mixture was prepared with the recipe of 1 mM dNTPs, 7.5 mM $MgCl_2$, 500 nM forward and reverse primers, and 1× ThermoPol buffer. The mixture was aliquoted into nine PCR vials, of which 6 µL of binding buffer was added to the first one as a negative control and 6 µL of collected complexes were added to the other eight. All the PCR vials were then loaded onto a Mastercycler® (Eppendorf, Enfield, CT), heated to 94 °C for 1 min, and paused at 94 °C. 5 units of *Taq* polymerase was added into each of the 9 vials. 23 cycles of PCR were performed at 94 °C for 30 s for denaturation, 55 °C for 30 s for annealing, and 72 °C for 20 s for extension. A final extension at 72 °C for 5 min was carried out after the 23rd cycle. An agarose gel (2%, w/v) electrophoresis of the PCR product was performed to confirm the desired products at 80 bp and observe any possible contamination in the negative control. A purification step using streptavidin agarose resin and chromatography column was taken to collect the desired single strands with FAM labeling, after which the standard ethanol precipitation was used to get the final product as reported before.[45-48] The dry product was then dissolved in 10 µL binding buffer for the next round of selection and affinity characterization.

2.3.4 Sequencing and Bioinformatics

1 fmol of the random ssDNA library was PCR amplified as described above, and the product was purified and used as a control to monitor the diversity

change when the library was not exposed to the pressure of CE selection. 1,000 fold dilutions of this control and pools selected from CE-SELEX were used as the sample in an additional PCR, which was performed under exactly the same conditions, except that fusion primer A and B were used to add the adaptors for the emulsion PCR in 454 sequencing. PCR products were then loaded onto 2% (w/v) agarose gels and electrophoreses were carried out at 100 V for 3.5 hours. Gel images were taken to make sure there was no contamination in all eight reactions. The 80 bp bands were carefully cut out of the gel while being monitored on a DNA safe Visi-Blue^{TM} transilluminator (UVP, Inc., Upland, CA). These 80 bp products were extracted using QIAquick® gel extraction kit (Qiagen, Inc., Gaithersburg, MD) and stored in TE buffer of pH 7.0 at -20 °C. 454 sequencing was performed at the BioMedical Genomics Center of the University of Minnesota. Sequencing results were exported as eight FASTA files.

Since five of the eight FASTA files contain more than 20,000 sequences, it is very time consuming to perform multiple sequence alignment to find motifs or to obtain phylogenetic trees. Thus, a home-programmed code in Linux was used to identify enriched sequences/motifs quickly. Briefly, the desired products (80±2 nt) were first extracted for analysis. The primer regions of every sequence were first removed before analysis. The oligonucleotide with the length of n nt (n ranged from 5 to 15) was then extract from the 1st register position to the nth position. Then the second oligonucleotide was extracted from the 2nd to the (n+1)th register position, and the process continued to the last oligonucleotide from the (41–n)th to the 40th register position. The same analysis moved on to the second

sequence in the file and kept moving until finishing the last one. While extracting, every oligonucleotide was recorded, and the count number was added if a same one was found later in the analysis. When the program finished on one file, all the oligonucleotides were exported in abundance descending order. Enriched sequences were identified by piecing together abundant and overlapping oligonucleotides into complete 40 nt regions. The results were confirmed by CodonCode Aligner3.7.1 (CodonCode Corp., Bedham, MA), which identified the top three enriched sequences when analyzing each file divided into subgroups of ~3,000 sequences. Online software MEME4.7.0[245, 246] and ClustalW2.0[247, 248] were also used in order to find motifs both in the short byproducts and the desired products.

2.3.5 K_d measurements by Affinity Capillary Electrophoresis (ACE) and Fluorescence Polarization (FP)

ACE experiments were performed using the same MDQ CE system under the same separation conditions. Selected pools or aptamers with a fixed concentration in the range of 1 nM to 10 nM was heated to 72 °C for 5 min, and then let cool down to room temperature. 10 μL of these samples were then incubated with 10 μL of VEGF from 0 nM to 750 nM. The mixtures were then injected into the capillary from the lowest to the highest VEGF concentration. Electropherograms were collected with 32 Karat (Beckman Coulter, Inc., Fullerton, CA). The peak heights of the free aptamer were used to calculate the bound fractions, and the K_d value was obtained by nonlinear fitting of Equation 2.1[249, 250] in Origin 8.1 (OriginLab, Inc., Northampton, MA):

$$f_a = \frac{c}{1 + K_d / ([VEGF]_t - 0.5([A]_t + [VEGF]_t + K_d - (([A]_t + [VEGF]_t + K_d)^2 - 4[A]_t[VEGF]_t)^{0.5}))}$$ **Equation 2.1**

in which f_a, $[VEGF]_t$, $[A]_t$, and c are the bound fraction of aptamer, total VEGF concentration, total aptamer concentration, and the maximum bound fraction, respectively. $([VEGF]_t – 0.5([A]_t + [VEGF]_t + K_d – (([A]_t + [VEGF]_t + K_d)^2 – 4[A]_t[VEGF]_t)^{0.5}))$ is the concentration of unbound VEGF.

Fluorescence polarization (FP) experiments were performed on a Microplate Reader of Synergy^{TM} 2 (BioTek Instruments, Inc., Winooski, VT). The samples were prepared the same manner as in ACE experiments. The 20 μL mixtures were loaded into a corning 3540 microplate (Corning Inc., Corning, NY). Gen 5^{TM} (BioTek^{R} Instruments, Inc., Winooski, VM) was used to collect parallel and perpendicular intensities (λ_{ex} = 485±20 nm, λ_{em} = 520±20 nm) and calculate polarizations. Each sample was measured three times and all data were used in K_d analysis. The same equation and software were used to determine K_d as in ACE, except for the determination of bound fraction, which was discussed in detail in Chapter 1.3.3., using Equation 1.16~1.18.

2.4 Results and Discussion

2.4.1 Monitoring Affinity of Selected Pools during CE-SELEX

Affinity of the initial ssDNA library was first evaluated before selections by two orthogonal methods, Affinity Capillary Electrophoresis (ACE) and Fluorescence Polarization (FP). The highest VEGF stock concentration was prepared at 1 μM, limited by the commercial availability. The random ssDNA library had no apparent affinity toward VEGF even at the highest VEGF concentration,

suggesting a K_d value at least above 1 μM. This is not surprising considering DNA aptamers for VEGF have been selected with a few hundred-nanomolar K_d values.[240, 241] Seven cycles of CE selection were performed. The incubated mixture in cycle one consisted of 100 μM library and 10 nM VEGF allowing ~4×10^{11} ssDNA sequences and ~4×10^{7} VEGF molecules to be injected per separation. Three CE separations were performed to increase the amount of complexes collected for PCR amplification. Most of the purified ssDNA product was used in the next round of selection except a small aliquot for K_d measurements. As shown in Figure 2.1, the affinity of the bulk pool was improved by at least one order of magnitude right after the first round of selection, with the K_d value of 50±60 nM from ACE experiment and 120±80 nM from FP experiment. The affinity was further slightly improved in the following selections until round 6 and 7, from which fluctuation was observed, a phenomenon that was also observed in previous CE-SELEX studies.

2.4.2 Overview of Sequencing Results

After selection, the same volume of sample from each selected pool was PCR amplified to incorporate the adaptor primers. The products were gel purified and recovered to the maximum amount for library preparation in 454 sequencing. The control sample generated as described in Chapter 2.3.4 sequencing and bioinformatics was also sequenced for a direct comparison to pool 1. Since the same volume of samples was used and samples were treated exactly the same way, sequencing results summarized in Figure 2.2 directly reflected variations of sample in quantity and quality after CE selection. 31,331 sequences were obtain-

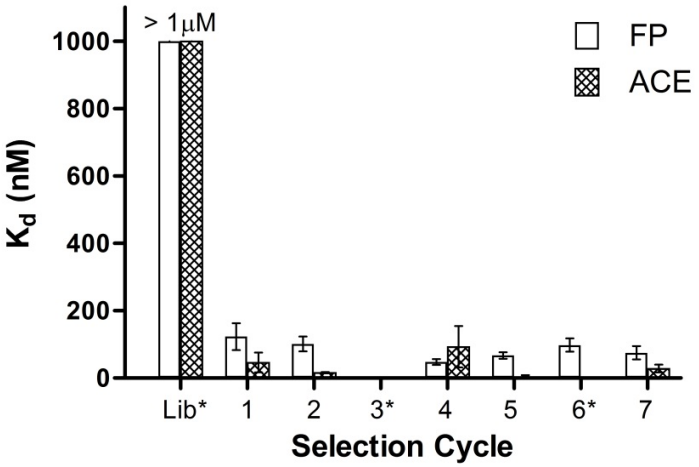

Figure 2.1 A histogram of binding affinities (K_d values) of the initial random ssDNA library and the pools after selection against VEGF. K_d values were obtained by both FP and ACE. Error bars represent 95% confidence interval. Affinity of the initial random ssDNA library was only roughly estimated, limited by the highest protein concentration available. K_d values of pool 3 (by FP and ACE) and pool 6 (by ACE) were not available due to limitation in DNA quantity.

ed for control sample, out of which 30,107 sequences had the lengths of 80±2 nt and were considered as desired products. For pool 1, both the total number of sequences and the number of desired sequences were very similar compared with the control. When more PCR were performed in the later selection rounds, both the total number of sequences and the number of desired sequences decreased, mildly in pool 2 through pool 5 and sharply in pool 6 and pool 7. The average sequence length followed the same trend as shown by the right Y axis in Figure 2.2. The value was 80 nt for control and cycle 1; it slightly decreased to 76, 76, 77, and 75 nt for pool 2, 3, 4, and 5, respectively; and it finally reached 70 nt in cycle 6 and 62 nt in cycle 7.

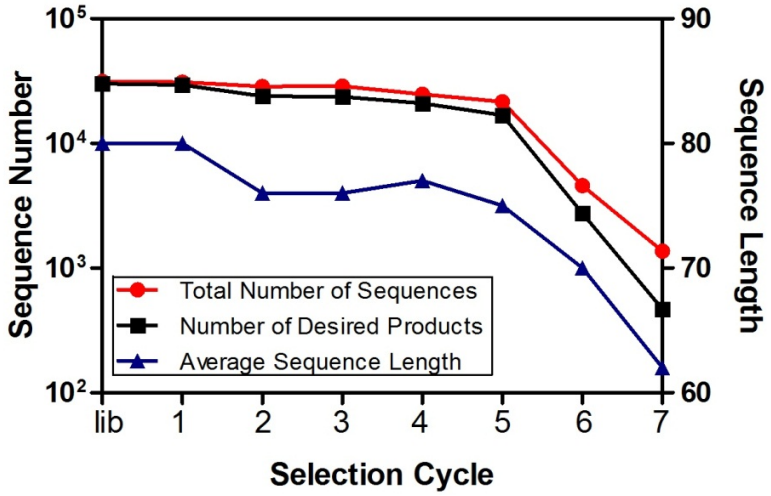

Figure 2.2 The summary of sequencing result in terms of total number of sequences (—●—), number of desired products (—■—), and average sequence length (—▲—). The left Y axis is in logarithm scale, and the right Y axis is in linear scale. Sequences with the lengths of 80±2 nt were considered as desired products.

2.4.3 Analysis of the Short Sequences

In order to probe into the sequencing results, I first summarized the frequency distributions over a wide range of sequence length for all the eight files. Shown in Figure 2.3, the highest frequency in every file was set to be 100%, and all the other frequencies were normalized. Results suggested that the most abundant length in all eight files was 80 nt, which was the length of the desired products. There were fewer sequences with lengths above 82 nt, indicated by the flat baselines in all the eight traces. However, for the short sequences, the relative abundance increased with selection cycle. Although there were no noticeable short byproducts in control and pool 1, there was an obvious peak centered at 47 nt from pool 2 through pool 5 with the normalized abundances around 10%. In pool 6, the length distribution of the short sequences stretched from 42 nt to

above 60 nt; in pool 7, this distribution even stretched to above 65 nt, and several peaks were observed with the normalized abundances above 20%.

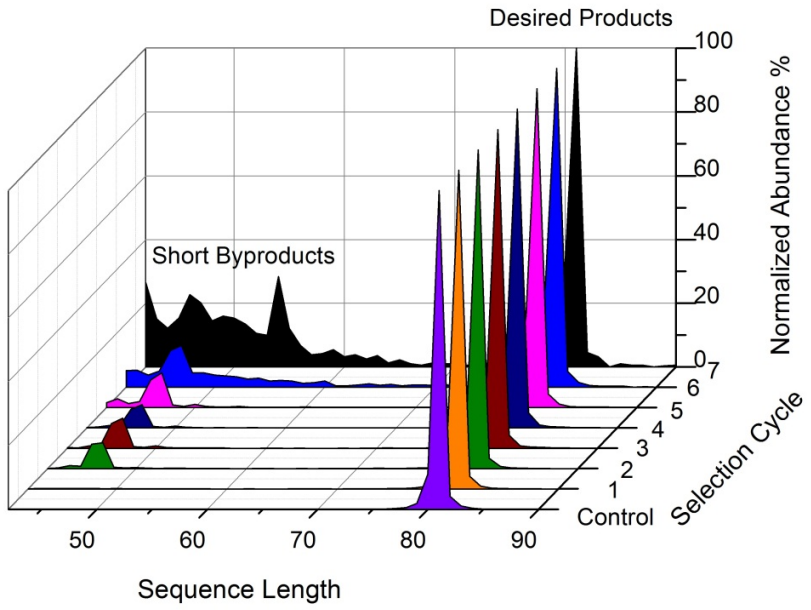

Figure 2.3 Normalized sequence abundance profiles of the control sample and selected pools. In each sample, the highest abundance at a specific length was set to be 100%, and abundances at other lengths were normalized. Only sequences with the length equal to or above 42 nt were exported after 454 sequencing, so 42 nt was chosen as one boundary. 90 nt was chosen as the other boundary because few examples existed above this length.

All the short byproducts at the peak lengths were further studied and motifs were identified by MEME. The normalized abundances of the short sequences at a specific length, the discovered motifs, and the normalized abundances of the motifs were summarized in Table 2.1. The control sample undergoing PCR amplification, purification, another PCR to incorporate adaptor primers, and 454 sequencing was treated exactly the same as pool 1 except that pool 1 was subjected to incubation with VEGF and separation using CE. The undesired short products only took a small portion in these two samples, below 0.3% quantity of

the 80 nt sequences. The motif discovered in control showed low consensus level (Figure 2.4, (I)); meanwhile this motif was found at different register positions in the short sequences without preference at a particular position; the normalized abundance was only 0.035% by MEME as compared to 0.157% for all 45 nt sequences. In pool 1, the 10 nt motif showed very high consensus level (Figure 2.4, (A)); 93% of the time the motif appeared at the same register position; most importantly this motif was found in two thirds of the 50 nt short sequences. Interestingly, although this motif also showed up in the short byproducts in the following pools, it was not the dominant. From pool 2 through pool 6, the most abundant short byproducts all appeared at 47 nt, shared the same motif and close abundances, and most of them were at the same register position. In pool 7, the three peaks with normalized abundances above 20% showed up at 54 nt, 46 nt, and 42 nt. The motif of the 46 nt was very similar to the motif found in pool 2 through pool 6. 42 nt sequences were primer dimers with a random 2 nt region in the middle, and 54 nt sequences shared a newly discovered motif. These motifs identified from pool 1 through pool 7 were also discovered in some short byproducts with other lengths and in some of the few long byproducts.

Although it is difficult to draw an explicit mechanism for the formation of these short products, these motifs share a common characteristic of containing more GC than AT, indicating a possible explanation that some GC-rich sequences were mistakenly amplified. It is also worth pointing out that a careful gel-electrophoresis separation and extraction were performed to get rid of undesired

products before 454 sequencing. A gel image taken after the purification clearly showed single bands at 80 bp (see Figure 2.5). However, the short products were up to 28% quantity of the desired product in sequencing results. Since our samples were not extremely long or short for 454 machines after the adaptor primers were incorporated, the sequencing error rate should be below 1%.[230] One possible reason is that the short sequences had much higher amplification efficiency and were preferably amplified during emPCR of 454 sequencing. In selections, due to the amplification preference of the short sequences and possibly the emergence of new short sequences in every PCR, the productivity of the desired products was greatly impaired in later selection rounds, and the bulk affinities of the pools were not further improved or even went worse.

2.4.4 Analysis of the Desired Products

454 sequencing generated up to ~3×10^4 sequences in individual selected pools, among which the most abundant sequences are the desired products whose lengths fall into the range of 78 nt to 82 nt, as shown in Figure 2.2 and Figure 2.3. Sequence information obtained from pyrosequencing is a few hundred folds more than cloning, which greatly expands our scope to characterize CE-SELEX. However, the massive data also bring challenges to our analysis. Most of the sequence-alignment or motif-finding programs are time consuming when dealing with this amount of sequences. For example, the running time of MEME is proportional to the data size squared, and it takes months to analyze a 30,000-

Table 2.1 Analysis of the most abundant short sequences

		Initial library (control)	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5	Pool 6	Pool 7
Length (nt)		45	50	47	47	47	47	47	54, 46, 42
Normalized abundance		0.157%	0.27%	8.08%	9.43%	7.28%	11.01%	13.13%	28.44%, 22.68%, 26.52%
Motif		CGTACTAC	CAGGTAGGGT	GGAGCCA	GGAGCCA	GGAGCCA	GGAGCCA	GGAGCCA	CGTGCTACCGTGAA, GAGCCA,
Normalized abundance of motif	**Perfect match**	0.004%	0.13%	5.04%	5.93%	4.82%	7.71%	6.57%	14.06%, 6.71%
	1 mismatch	0.024%	0.17%	6.75%	8.98%	6.95%	10.15%	8.40%	16.93%, 9.27%
	MEME*	0.035%	0.18%	6.79%	9.18%	7.02%	10.49%	8.61%	20.45%, 9.27%

* The p-value is equal to or smaller than 0.05 in MEME.

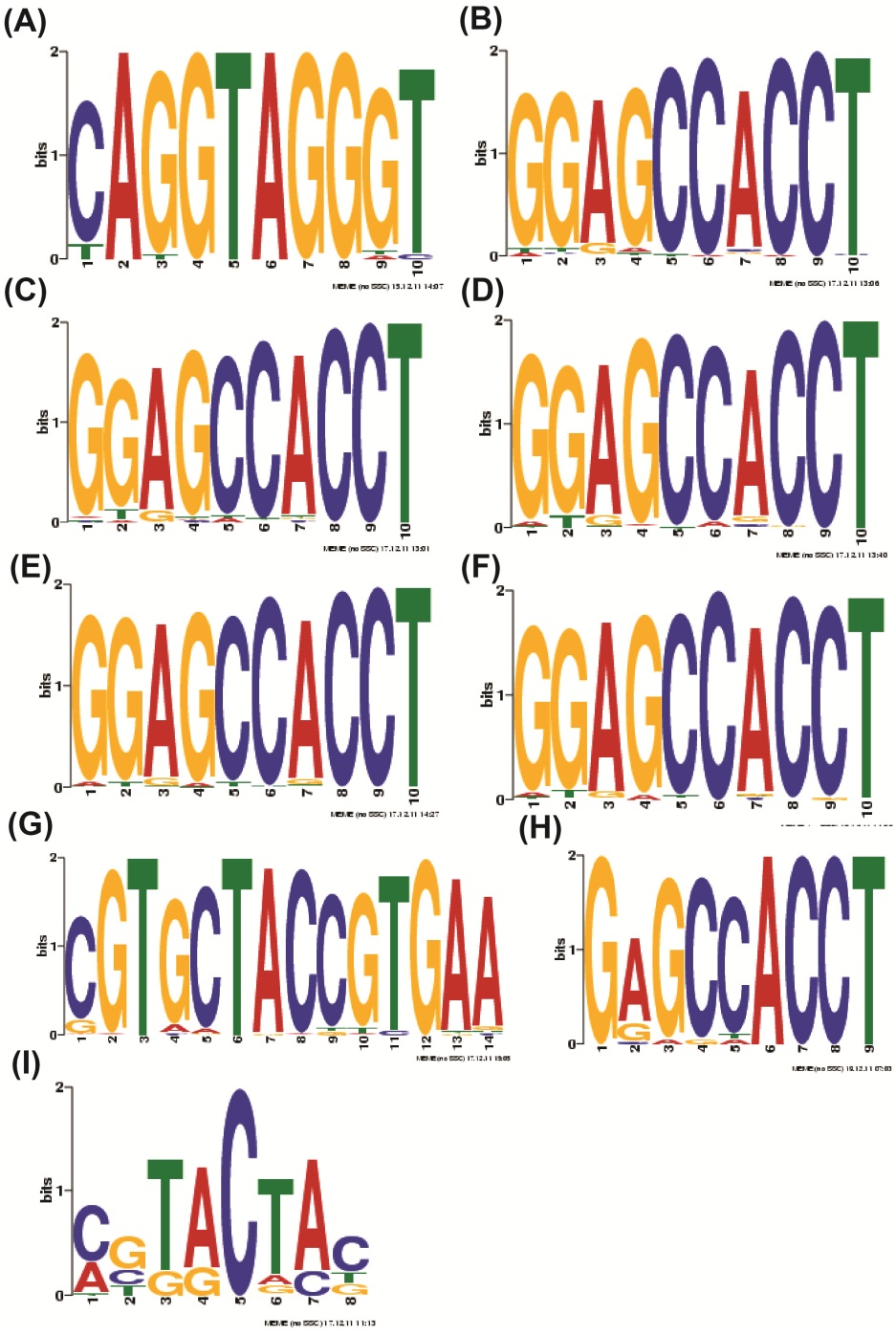

Figure 2.4 Motifs of the short byproducts at the peak lengths in Figure 2.3, identified by MEME: (A) motif found in pool 1, (B) motif found in pool 2, (C) motif found in pool 3, (D) motif found in pool 4, (E) motif found in pool 5, (F) motif found in pool 6, (G) motif found in pool 7 at the length of 54 nt, (H) motif found in pool 7 at the length of 46 nt, and (I) motif found in control.

sequence file.[245] As a result, I decided to employ several complementary methods to perform sequence analysis.

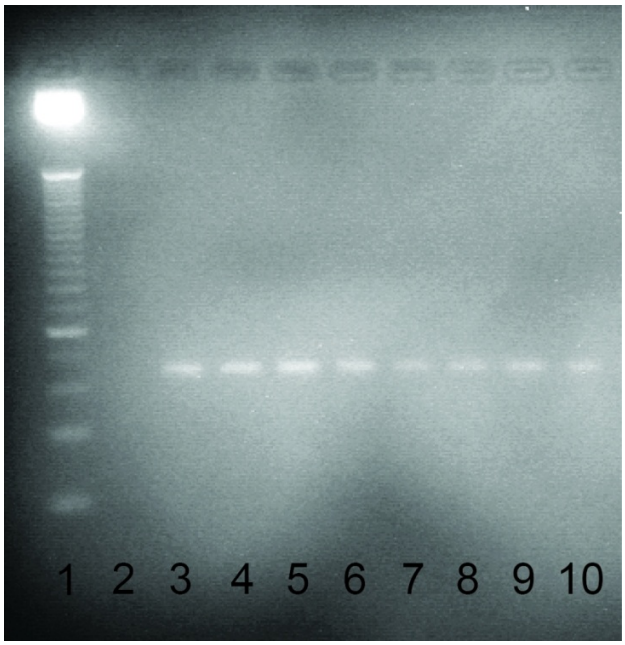

Figure 2.5 A gel image of the 8 samples (lane 3 through lane 10) after gel extraction to remove undesired short products. Lane 1 is the 25 bp DNA ladder, with the bottom band indicating 25 bp and an increment of 25 bp. Lane 2 is the negative control. Lane 3 through lane 10 correspond to the products of the initial library and round 1 through round 7 after gel electrophoresis purification, respectively. All the product lanes are between 75 and 100 bp, which agree with the desired length of 80 bp. There are no short products detectable below the desired product bands.

A home-programmed code in Linux was first used to search for the abundant oligonucleotides in the desired products (see Chapter 2.3.4 sequencing and bioinformatics for details) in all selected pools at once. The top 1,000 abundant oligonucleotides from 5 nt through 15 nt were exported in abundance descending order and the top 100 were shown in Table 2.2. The probability to find an n nt long random oligonucleotide in a ssDNA pool equals $0.25^n \times (41-n) \times 100\%$. These values, displayed as logarithms, were the theoretical values in Figure 2.6 (A). The evolution profiles of the most abundant oligonucleotides shown in Figure 2.6

(B) were obtained as below: the abundances were first normalized to theoretical abundances of random oligonucleotides with the same lengths, and then the logarithms were displayed. The most abundant 5 nt oligonucleotide TTTTG had similar abundances in selected pools as compared to that of a random 5 nt oligonucleotide, since the logarithm values were close to 0, indicating approximately 1:1 ratios. This suggests that there were no 5 nt oligonucleotides preferably selected. The most abundant 6 nt and 7 nt oligonucleotides, however, had the highest logarithm values of 0.52 and 0.91 at cycle 5, respectively, corresponding to 3.3:1 and 8.1:1 ratios. This indicates that they were preferably enriched. From 8 nt through 15 nt, the logarithm values were above one right from pool 1, and all enriched at pool 4. As a comparison, the abundances of top oligonucleotides found solely in control sample were very similar to the theoretical values from 5 nt to 9 nt (Figure 2.6 (A)). In other words, these oligonucleotides was not preferably amplified, indicating the remained randomness in the unselected control sample even after three multiple-round PCR reactions. The difference between the abundances in theory and in control only became significant when the length was greater than 9 nt. This is because the occurrence of a random oligonucleotide (larger than 9 nt) is less than one in a pool with sample size of ~30,000; however, practically the occurrence of the most abundant oligonucleotide (larger than 9 nt) has to be equal to or greater than one.

Table 2.2 Top 100 abundant oligonucleotides from 5 nt through 15 nt in selected pools

Abundance Ranking	5nt Oligo	6 nt Oligo	7 nt Oligo	8 nt Oligo	9 nt Oligo	10 nt Oligo	11 nt Oligo	12 nt Oligo	13 nt Oligo	14 nt Oligo	15 nt Oligo
1	TTTTG	GGA GCC	GGAG CCA	ATCTG GAG	GAAGC CCGC	GCCCG CTGTG	GAAGCC CGCTG	GAAGCC CGCTGT	GAAGCCC GCTGTG	GAAGCCC GCTGTGA	GCCCGCT GTGACATC
2	TTTGT	TTTG TT	TGGA GCC	TCTG GAGC	AGCCC GCTG	AGCCC GCTGT	AGCCCG CTGTG	AAGCCC GCTGTG	CTGTGAC ATCTGG	AGCCCGC TGTGACA	GAAGCCC GCTGTGAC
3	GTTTT	TTTT GT	TCTG GAG	TGTG ACAT	GGAGC CGGT	AAGCCC GCTG	AAGCCC GCTGT	AGCCCG CTGTGA	AGCCCGC TGTGAC	GCCCGCT GTGACAT	CGCTGTGA CATCTGG
4	GGTTT	TTTT TG	ATCTG GA	GGTT CCCG	TGGAG CCGG	CTGGA GCCGG	CTGGAG CCGGT	GCCCGC TGTGAC	GCCCGCT GTGACA	GCTGTGAC ATCTGG	AGCCCGCT GTGACAT
5	TTGTT	GTTT TT	GTGA CAT	ACATC TGG	GCCCG CTGT	GAAGC CCGCT	GCCCGC TGTGA	TCTGGA GCCGGT	AAGCCCG CTGTGA	AAGCCCG CTGTGAC	CCGCTGTG ACATCTG
6	TGTTT	TTTT GG	GCCG GTT	GGAG CCGG	CCCGC TGTG	TGGAG CCGGT	TCTGGA GCCGG	TGTGAC ATCTGG	CCGCTGT GACATC	CTGGAGC CGGTTCC	TCTGGAGC CGGTTCC
7	TTTGG	TTGT TT	CATCT GG	CATCT GGA	ATCTG GAGC	TCTGGA GCCG	GTGACA TCTGG	GACATC TGGAGC	CATCTGG AGCCGG	CTGTGACA TCTGGA	CTGGAGC CGGTTCCC
8	ATTTT	TTTG GT	TGTG ACA	CGCT GTGA	GGTTC CCGG	CATCTG GAGC	CCCGCT GTGAC	CTGTGA CATCTG	ATCTGGA GCCGGT	CCCGCTGT GACATC	AAGCCCG CTGTGACA
9	TTTTA	GGTT TT	GCTG TGA	CTGG AGCC	TCTGG AGCC	CCCGCT GTGA	GTTCCC GGAGC	GGTTCC CGGAGC	GCTGTGA CATCTG	CGCTGTGA CATCTG	CTGTGACA TCTGGAG
10	TTGGT	TTAT TT	CGCT GTG	GAGC CGGT	CTGGA GCCG	GGAAG CCCGC	CCGCTG TGACA	ATCTGG AGCCGG	GTGACAT CTGGAG	GTGACATC TGGAGC	GGAGCCG GTTCCCGG
11	TATTT	TGTT TT	GGAG CCG	TGGA GCCG	GACAT CTGG	CTGTGA CATC	CCGGTT CCCGG	CCCGCT GTGACA	CCCGCTG TGACAT	CCGCTGT GACATCT	TGTGACAT CTGGAGC
12	TTATT	ATTT TT	GGTT CCC	CTGT GACA	CCGCT GTGA	CCGCT GTGAC	GACATC TGGAG	CTGGAG CCGGTT	TGGAGCC GGTTCC	CCGGTTCC CGGAGC	CCCGCTGT GACATCT
13	TTTAT	GGGT TT	CTGT GAC	GAAG CCCG	AAGCC CGCT	TGACAT CTGG	ACATCT GGAGC	CGCTGT GACATC	TGACATC TGGAGC	CATCTGGA GCCGGT	GCTGTGAC ATCTGGA
14	TGGTT	TTTT TA	CTGG AGC	GTGA CATC	CGGTT CCCG	CGCTGT GACA	CATCTG GAGCC	CCGCTG TGACAT	TCTGGAG CCGGTT	TGGAGCC GGTTCCC	GCCGGTTC CCGGAGC
15	GGGT T	TATT TT	AGCC GGT	CCGC TGTG	CATCT GGAG	ATCTGG AGCC	CTGTGA CATCT	GGAAGC CCGCTG	GGAGCC GGTTCCC	GGAAGCC CGCTGTG	TGGAGCC GGTTCCCG

16	GTATT	TTGGTT	GGAAGCC	CGGAGCCA	CGCTGTGAC	GTGACATCTG	TGTGACATCTG	GGAGCCGGTTCC	CGGTTCCCGGAGC	GGAGCCGGTTCCCG	GACATCTGGAGCCGG
17	TTGGG	TTTATT	CCGGTTC	GCCCGCTG	CTGTGACAT	CGGTTCCCGG	ATCTGGAGCCG	GTGACAGCCGTCTGGA	GGAAGCCCGCTGT	GAGCCGGTTCCCGG	GTGACATCTGGAGCC
18	GTGTT	GTTTTG	ACATCTG	CGGTTCCC	GGAAGCCCG	GACATCTGGA	GCTGTGACATC	GCCGGTTCCCGG	CTGGAGCCGGTTC	TGTGACATCTGGAG	ACATCTGGAGCCGGT
19	GTTTG	GTTTGT	CGGTTCC	CCCGGAGC	TGTGACATC	GGTTCCCGGA	CGCTGTGACAT	CATCTGGAGCCG	CGCTGTGACATCT	TCTGGAGCCGGTTC	GAGCCGGTTCCCGGA
20	TTGTG	TTTGTG	GTTCCCG	GTTCCCGG	TCCCGGAGC	GTTCCCGGAG	GGTTCCCGGAG	TGACATCTGGAG	GACATCTGGAGCC	GACATCTGGAGCCG	TGACATCTGGAGCCG
21	GTTTA	TGGGTT	GAGCCGG	GCCGGTTC	CCGGACCA	CCGGTTCCCG	TGGAGCCGGTT	GCTGTGACATCT	TGTGACATCTGGA	ACATCTGGAGCCGG	GGAAGCCCGCTGTGA
22	GTTAT	TTTTAT	GACATCT	GCTGTGAC	ACATCTGGA	TTCCCGGAGC	GCCGGTTCCCG	ACATCTGGAGCC	CCGGTTCCCGGAG	GCCGGTTCCCGGAG	CCGGTTCCCGGAGCC
23	TATGT	TGGTTT	CCGGAGC	GGAAGCCC	GCTGTGACA	GGAGCCGGTT	GGAAGCCCGCT	CCGGTTCCCGGA	GAGCCGGTTCCCG	TGACATCTGGAGCC	CATCTGGAGCCGGTT
24	TGTGT	TTTGGG	TGACATC	CCGGTTCC	GCCGGTTCC	CCCGGAGCCA	TGACATCTGGA	TGGAGCCGGTTC	ACATCTGGAGCCG	ATCTGGAGCCGGTT	AGCCGGTTCCCGGAG
25	TTTAG	TTATGT	TTCCCGG	GACATCT	CCCGGAGCC	ACATCTGGAG	GAGCCGGTTCC	CGGTTCCCGGAG	GCCGGTTCCCGGA	CGGTTCCCGGAGCC	ATCTGGAGCCGGTTC
26	ATTTG	TTGGGT	CCGCTGT	CCGGAGCC	CCGGTTCCC	GCTGTGACAT	GGAGCCGGTTC	GAGCCGGTTCCCC	AGCCGGTCCCGGG	GGTTCCCGGAGCCA	CGGTTCCCGGAGCCA
27	TTATG	TATTTG	CCCGGAG	AGCCGGTT	GTGACATCT	TGTGACATCT	CGGTTCCCGGA	GTTCCCGGAGCC	GGTTCCCGGAGCC	AGCCGGTTCCCGGA	AGCACAGAGGTCAGA
28	TATTG	GTTATT	CGGAGCC	CCCGCTGT	GTTCCCGGA	GCCGGTTCCC	AGCCGGTTCCC	AGCCGGTTCCCG	GTTCCCGGAGCCA	CGTGCTACCGTGAA	GCACAGAGGTCAGAT
29	TGGGT	GTATTT	TACCGTG	AGCCCGCT	TTCCCGGAG	TCCCGGAGCC	TTCCCGGAGCC	TTCCCGGAGCCA	GTGCTACCGTGAA	GCACAGAGGTCAGA	CAGCACAGAGGTCAG
30	TGTAT	GTTTAT	TCCCGGA	TTCCCGGA	TGACATCTG	AGCCGGTTCC	TCCCGGAGCCA	GTGCTACCGTGA	CGTGCTACCGTGA	AGCACAGAGGTCAG	CACAGAGGTCAGATG
31	GGTAT	TTGTGT	GAAGCCC	TCCCGGAG	GAGCCGGTT	GAGCCGGTTC	GTGCTACCGTG	TGCTACCGTGAA	GCACAGAGGTCAG	CACAGAGGTCAGAT	ACAGAGGTCAGATGG
32	TTAGT	ATTTTG	AAGCCCG	TGACATCT	AGCCGGTTC	CTACCGTGAA	TGCTACCGTGA	CGTGCTACCGTG	CACAGAGGTCAGA	CAGCACAGAGGTCA	TATGCGTGCTACCGT

33	GGTTA	ATTTGT	TATGCGT	AAGCCCGC	CTACCGTGA	GTGCTACCGT	GCTACCGTGAA	CACAGAGGTCAG	CAGCACAGAGGTC	ACAGAGGTCAGATG	CTATGCGTGCTACCG
34	GTTGT	TGTTAT	GCCCGCT	TACCGTGA	TACCGTGAA	TGCTACCGTG	CGTGCTACCGT	AGCACAGAGGTC	CAGAGGTCAGATG	CAGAGGTCAGATGG	ATGCGTGCTACCGTG
35	TTAGG	TTGTGG	ATGCGTG	CTACCGTG	TGCTACCGT	GCTACCGTGA	GCACAGAGGTC	CAGCACAGAGGT	AGCACAGAGGTCA	TATGCGTGCTACCG	CCTATGCGTGCTACC
36	ATTGT	TTTTAG	CCCGCTG	ACCGTGAA	GTGCTACCG	CGTGCTACCG	CAGCACAGAGG	ACAGAGGTCAGAT	ACAGAGGTCAGAT	CTATGCGTGCTACC	TGCGTGCTACCGTGA
37	TGTTA	TATTGG	AGGTCAG	GCTACCGT	GCTACCGTG	CAGCACAGAG	CAGAGGTCAGA	GCACAGAGGTCA	AGAGGTCAGATGG	TGCGTGCTACCGTG	GCGTGCTACCGTGAA
38	GATTT	GTGTTT	CGTGCTA	GTGCTACC	CGTGCTACC	AGCACAGAGG	ACAGAGGTCAG	CAGAGGTCAGAT	GCGTGCTACCGTG	ATGCGTGCTACCGT	ACCTATGCGTGCTAC
39	TAGGT	TAGGGT	GTGCTAC	TGCTACCG	AGAGGTCAG	GCACAGAGGT	GAGGTCAGATG	AGAGGTCAGATG	CTATGCGTGCTAC	CCTATGCGTGCTAC	CACCTATGCGTGCTA
40	TTGTA	GTTTTA	ACCGTGA	CGTGCTAC	GCACAGAGG	CAGAGGTCAG	AGCACAGAGGT	GAGGTCAGATGG	TATGCGTGCTACC	GCGTGCTACCGTGA	GCAGCACAGAGGTCA
41	ATTGG	TATGTT	TCAGATG	AGGTCAGA	AGCACAGAG	CACAGAGGTC	AGAGGTCAGAT	TATGCGTGCTAC	ATGCGTGCTACCG	ACCTATGCGTGCTA	AGCAGCACAGAGGTC
42	ATGTT	TGTTTG	GTCAGAT	GAGGTCAG	ACAGAGGTC	GAGGTCAGAT	CACAGAGGTCA	GCGTGCTACCGT	TGCGTGCTACCGT	CACCTATGCGTGCT	GCCACCTATGCGTGC
43	TAGTT	TTATTG	AGCCCGC	GTCAGATG	GGTCAGATG	AGAGGTCAGA	AGGTCAGATGG	CCTATGCGTGCT	CCTATGCGTGCTA	GCAGCACAGAGGTC	CCACCTATGCGTGCT
44	AGTTT	TGTGTT	CAGATGG	ACAGAGGT	GAGGTCAGA	AGGTCAGATG	CCTATGCGTGC	CTATGCGTGCTA	ACCTATGCGTGCT	AGCAGCACAGAGGT	AGCCACCTATGCGTG
45	TGTTG	TGTATT	CTACCGT	TATGCGTG	AGGTCAGAT	ACAGAGGTCA	CTATGCGTGCT	ATGCGTGCTACC	CACCTATGCGTGC	CCACCTATGCGTGC	GAGCCACCTATGCGT
46	GTATG	GTTTGG	GGTCAGA	AGAGGTCA	CACAGAGGT	GGTCAGATGG	ATGCGTGCTAC	TGCGTGCTACCG	GCAGCACAGAGGT	GCCACCTATGCGTG	CAGAGGTCAGATGGG
47	TATGG	TTTAGG	CCGTGAA	GCACAGAG	CAGAGGTCA	CTATGCGTGC	GCGTGCTACCG	ACCTATGCGTGC	AGCAGCACAGAGG	AGCCACCTATGCGT	CGTGCTACCGTGAAG
48	GGTGT	GATTTT	AGAGGTC	GGTCAGAT	CAGCACAGA	CCTATGCGTG	TATGCGTGCTA	CACCTATGCGTG	CCACCTATGCGTG	GAGCCACCTATGCG	GTGCTACCGTGAAGC
49	GGGTA	GAGCCA	CAGAGGT	CAGAGGTC	GTCAGATGG	TATGCGTGCT	TGCGTGCTACC	GCAGCACAGAGG	GCCACCTATGCGT	AGAGGTCAGATGGG	AGAGGTCAGATGGGA

50	ATTAT	TTAGGT	GAGGTCA	TCAGATGG	CTATGCGTG	ATGCGTGCTA	ACCTATGCGTG	AGCAGCACAGAG	AGCCACCTATGCG	GTGCTACCGTGAAG	GAGGTCAGATGGGAG
51	GTTAG	TTTATG	ACAGAGG	AGCACAGA	TATGCGTGC	GCGTGCTACC	CACCTATGCGT	GCCACCTATGCG	GAGCCACCTATGC	GAGGTCAGATGGGA	CTACCGTGAAGCAGC
52	TGTGG	TTTAGT	TTTTTGT	CAGCACAG	ATGCGTGCT	TGCGTGCTAC	GCAGCACAGAG	CCACCTATGCGT	GAGGTCAGATGGG	TGCTACCGTGAAGC	TACCGTGAAGCAGCA
53	AGGGT	TTTGTA	TGCTACC	CACAGAGG	CCTATGCGT	ACCTATGCGT	AGCAGCACAGA	AGCCACCTATGC	TGCTACCGTGAAG	TACCGTGAAGCAGC	GCTACCGTGAAGCAG
54	GTTGG	GTTGTT	GCTACCG	CTATGCGT	GCGTGCTAC	CACCTATGCG	CCACCTATGCG	GAGCCACCTATG	AGGTCAGATGGGA	CTACCGTGAAGCAG	TGCTACCGTGAAGCA
55	AGGTT	GGTTTG	GTTATTT	ATGCGTGC	TGCGTGCTA	GCAGCACAGA	GCCACCTATGC	AGGTCAGATGGG	GCTACCGTGAAGC	GCTACCGTGAAGCA	AGGTCAGATGGGAGC
56	TATAT	TGTTTA	AGATGGG	CCTATGCG	ACCTATGCG	AGCAGCACAG	AGCCACCTATG	GCTACCGTGAAG	CCGTGAAGCAGCA	ACCGTGAAGCAGCA	ACCGTGAAGCAGCAC
57	TGGTA	GGTTTA	GGTTTTT	GCGTGCTA	CACCTATGC	GCCACCTATG	GAGCCACCTAT	CTACCGTGAAGC	CTACCGTGAAGCA	AGGTCAGATGGGAG	GGTCAGATGGGAGCC
58	TAGGG	ATGTTT	TTTTGTT	TGCGTGCT	GCAGCACAG	CCACCTATGC	GGTCAGATGGG	GGTCAGATGGGA	ACCGTGAAGCAGC	GGTCAGATGGGAGC	AAGCAGCACAGAGGT
59	GGGGT	GGTTAT	GTAGGGT	ACCTATGC	AGCAGCACA	AGCCACCTAT	CTACCGTGAAG	CCGTGAAGCAGC	TACCGTGAAGCAG	AAGCAGCACAGAGG	CCGTGAAGCAGCACA
60	GGATT	ATTTTA	TGGGTTT	CACCTATG	GCCACCTAT	GAGCCACCTA	TACCGTGAAGC	TACCGTGAAGCA	GGTCAGATGGGAG	CCGTGAAGCAGCAC	GAAGCAGCACAGAGG
61	TATTA	AGTTTT	TTTGTTT	GCAGCACA	CCACCTATG	GTCAGATGGG	GTCAGATGGGA	ACCGTGAAGCAG	GTCAGATGGGAGC	GTGAAGCAGCACAG	CGTGAAGCAGCACAG
62	TTATA	TAGGTT	GGTAGGG	AGCAGCAC	AGCCACCTA	TACCGTGAAG	GTGAAGCAGCA	CGTGAAGCAGCA	AAGCAGCACAGAG	GTCAGATGGGAGCC	GTGAAGCAGCACAGA
63	TGATT	ATTGGT	TGCGTGC	GCCACCTA	TCAGATGGG	TCAGATGGGA	CCGTGAAGCAG	GTCAGATGGGAG	CGTGAAGCAGCAC	GAAGCAGCACAGAG	TGAAGCAGCACAGAG
64	TTGGA	TTGTAT	TTGGGTT	CAGATGGG	GAGCCACCT	ACCGTGAAGC	ACCGTGAAGCA	TCAGATGGGAGC	TGAAGCAGCACAG	CGTGAAGCAGCACA	GGAGCCACCTATGCG
65	ATATT	TTGTTG	CTATGCG	CCACCTAT	ACCGTGAAG	CAGATGGGAG	CGTGAACGCAGC	CAGATGGGAGCC	GTGAAGCAGCACA	TGAAGCAGCACAGA	GTCAGATGGGAGCCA

66	TTTAA	GTATTG	GCACAGA	AGCCACCT	CAGATGGGA	GTGAAGCAGC	TCAGATGGGAG	GAAGCAGCACAG	TCAGATGGAGCC	GGAGCCACCTATGC	AGAGGTCAGATGGAG
67	GGTTG	GGGTAT	TTTGGTT	AGATGGGA	CCGTGAAGC	CCGTGAAGCA	CAGATGGGAGC	GTGAAGCAGCAC	GAAGCAGCACAGA	TCAGATGGGAGCCA	CAGAGGTCAGATGGA
68	TTTGA	ATTTGG	TTTTTTG	GGTAGGGT	AGATGGGAG	TGAAGCAGCA	AAGCAGCACAG	AAGCAGCACAGA	CAGATGGGAGCCA	GAGGTCAGATGGAG	GAGGTCAGATGGAGC
69	TGGGG	GGTATT	TTTTGTG	GAGCCACC	TGAAGCAGC	CGTGAAGCAG	AGATGGGAGCC	TGAAGCAGCACA	GGAGCCACCTATG	AGAGGTCAGATGGA	AGGTCAGATGGAGCC
70	ATGGT	TTTTGA	GCGTGCT	AGGTAGGG	GAAGCAGCA	AGATGGGAGC	TGAAGCAGCAC	GGAGCCACCTAT	AGGTCAGATGGAG	AGGTCAGATGGAGC	TATTGGACACCTTCC
71	GTGGT	GGATTT	TTTTGGT	GATGGGAG	GTGAAGCAG	GATGGGAGCC	GAAGCAGCACA	AGATGGGAGCCA	GAGGTCAGATGGA	GGTCAGATGGAGCC	CGATTGTGGGGACGT
72	AATTT	GTATGT	TGTTATT	CCGTGAAG	CGTGAAGCA	GAAGCAGCAC	GATGGGAGCCA	GGTCAGATGGAG	GGTCAGATGGAGC	GATTGTGGGGACGT	GATTGTGGGGACGTA
73	TTAAT	TATTGT	TTTTGGG	ATGGGAGC	GATGGGAGC	AAGCAGCACA	GGAGCCACCTA	GTCAGATGGAGC	GTCAGATGGAGCC	ATTGGACACCTTCC	GGTCAGATGGAGCCA
74	ATTTA	TAGTTT	GTTTTTT	TGAAGCAG	AGGTAGGGT	ATGGGAGCCA	GTCAGATGGAG	AGGTCAGATGGA	CGATTGTGGGGAC	CCGATTGTGGGGAC	GACGTATTGGACACC
75	GTGTA	TTAGTT	ATTTTTT	GAAGCAGC	TGGGAGCCA	GGAGCCACCT	GGTCAGATGGA	TCAGATGGAGCC	TTGGACACCTTCC	GTCAGATGGAGCCA	CCGATTGTGGGGACG
76	GTAGT	TATTAT	GTTTTTG	GTGAAGCA	ATGGGAGCC	CAGGTAGGGT	TCAGATGGAGC	TGGACACCTTCC	TATTGGACACCTT	TATTGGACACCTTC	CGTATTGGACACCTT
77	TAATT	TTGGTG	TATTTTT	AAGCAGCA	CAGGTAGGG	TCAGATGGAG	CAGATGGAGCC	GGACACCTTCCC	CGTATTGGACACC	ATTGTGGGGACGTA	ACGTATTGGACACCT
78	ATGGG	TTATGG	AGCACAG	CGTGAAGC	AAGCAGCAC	GTCAGATGGA	GGACACCTTCC	GATTGTGGGGAC	GATTGTGGGACG	GGGACGTATTGGAC	GGACGTATTGGACAC
79	TGTGA	TTGTTA	TTTTTGG	CAGGTAGG	GGAGCCACC	CAGATGGAGC	GTGGGGACGTA	TTGGACACCTTC	ATTGTGGGGACG	CGATTGTGGGGACG	GGGACGTATTGGACA
80	GTGTG	GTTGGT	TGTTTTT	GGAGCCAC	TTGTGGGGA	AGATGGAGCC	TCCGATTGTGG	TTGTGGGGACGTA	TTGTGGGGACGTA	ACGTATTGGACACC	GTGGGGACGTATTGG
81	GTAAT	GTTTAG	TGGTTTT	GGGAGCCA	TCAGATGGA	GTGGGGACGT	TGGACACCTTC	GGGACGTATTGG	ATCCGATTGTGGG	GACGTATTGGACAC	GTATTGGACACCTTC
82	TTGAT	AATTTT	GTTTGTT	TGGGAGCC	GATTGTGGG	GACACCTTCC	GACACCTTCCC	CGATTGTGGGGA	CCGATTGTGGGGA	CGTATTGGACACCT	GGGGACGTATTGGAC

83	TGGTG	TTAGGG	CACAGAG	TTGTGGGG	ACGTATTGG	GGACACCTTC	CCGATTGTGGG	GGGGACGTATTG	GGGACGTATTGGA	GTGGGGACGTATTG	GATATCCGATTGTGG
84	GTATA	TATGGT	AGGTAGG	TGTTATTT	AGATGGAGC	CGATTGTGGG	GGGACGTATTG	GTGGGGACGTAT	ATTGGACACCTTC	GTATTGGACACCTT	TCCGATTGTGGGGAC
85	GTAGG	ATTGTT	TTTGTGT	GATTGTGG	CAGATGGAG	GATTGTGGGG	ATTGTGGGGAC	TCCGATTGTGGG	TGGACACCTTCCC	GGACGTATTGGACA	TGGGGACGTATTGGA
86	TGGAT	TTTTAA	TTGTTTT	ATTGTGGG	ATTGTGGGG	CCGATTGTGG	GATTGTGGGGA	CAGATGGAGCCA	GAGGATATCCGAT	TTGTGGGGACGTAT	ATTGTGGGGACGTAT
87	GTGGG	TTGGTA	TTGTGGG	GTATTGGA	CGATTGTGG	TGGACACCTT	ATCCGATTGTG	ATCCGATTGTGG	GGACGTATTGGAC	GGGGACGTATTGGA	GGATATCCGATTGTG
88	TGGGA	AGGTTT	CCTATGC	TGTGGGGA	CGTATTGGA	ACGTATTGGA	CGATTGTGGGG	TATTGGACACCT	TCAGATGGAGCCA	TGGGGACGTATTGG	GAGGATATCCGATTG
89	TAGTA	TGGTTA	TTGGTTT	TTGGGTTT	GGGGACGTA	ATTGTGGGGA	TTGGACACCTT	TGTGGGGACGTA	GTGGGGACGTATT	GATATCCGATTGTG	ATATCCGATTGTGGG
90	AGTGT	GTTAGT	GTGTTTT	ACGTATTG	ATCCGATTG	TCCGATTGTG	GACGTATTGGA	ATTGGACACCTT	GGGGACGTATTGG	ATATCCGATTGTGG	ATCCGATTGTGGGGA
91	TGGAG	TCTGGA	GGGTTTT	CGTATTGG	TGGGGACGT	GGGGACGTAT	GAGGATATCCG	CCGATTGTGGGG	ACGTATTGGACAC	TATCCGATTGTGGG	TGTGGGGACGTATTG
92	ATGTG	ATTTAT	ATTTGTT	GGGGACGT	GGACACCTT	GGGACGTATT	GGGGACGTATT	GACGTATTGGAC	GACGTATTGGACA	TTGGACACCTTCCC	AGGATATCCGATTGT
93	GATTG	ATATTT	TTTGGGT	AGATGGAG	GACACCTTC	TGGGGACGTA	TGGGGACGTAT	ATTGTGGGACG	GTATTGGACACCT	TCCGATTGTGGGGA	TTGTGGGGACGTATT
94	ATATG	TGTTGT	TTTGTTG	GGGACGTA	GGGACGTAT	GACGTATTGG	TATTGGACACC	CGTATTGGACAC	TGGGGACGTATTG	GAGGATATCCGATT	ATTGGACACCTTCCC
95	ATGTA	ATCTGG	ATTTTGT	TATTGGAC	GACGTATTG	TTGTGGGGAC	TGTGGGGACGT	GTATTGGACACC	TATCCGATTGTGG	GGATATCCGATTGT	TATCCGATTGTGGGG
96	TATAG	AGGGTT	GTATTGG	GGACGTAT	GTGGGGACG	ACACCTTCCC	ACGTATTGGAC	GAGGATATCCGA	GGACACCTTCCCG	ATCCGATTGTGGGG	CGTGCTACCGTGAAA
97	GGTAA	GTTATG	TTTTATT	CGATTGTG	ACACCTTCC	ATCCGATTGT	GGACGTATTGG	GGACGTATTGGA	TGTGGGGACGTAT	AGGATATCCGATTG	TTGGACACCTTCCCG
98	GGGAT	GATTGT	TTATTTT	GAGGATAT	CCGATTGTG	GGACGTATTG	AGATGGAGCCA	AGGATATCCGAT	TCCGATTGTGGGG	TGTGGGGACGTATT	CGAGGATATCCGATT
99	TGTAG	GGGTTA	TTTTTAT	TTTTTGTG	GATGGAGCC	GAGGATATCC	ATTGGACACCT	GACACCTTCCCG	GATATCCGATTGT	TGGACACCTTCCCG	GTGCTACCGTGAAAC

100	GGTTC	ATGGTT	GTATTTT	TGTTTGTT	GTATTGGAC	CGTATTGGAC	TTGTGGGGACG	ACGTATTGGACA	ATATCCGATTGTG	GTGCTACCGTGAAA	TGCTACCGTGAAACA

Interestingly I found, in Figure 2.6 (B), all the oligonucleotides shared overlapped regions except “TTTTG”, and most of them appeared at the same register position in the desired products. For example, 65% of “GGAGCC” were found as the last six nucleotides in the 40 nt random region in pool 1, and the rest 35% were randomly distributed in the 40 nt random region with the abundance of 0.17%, compared to the theoretical abundance of 0.85% for a random 6 nt sequence and the abundance of 0.57% for the most abundant 6 nt

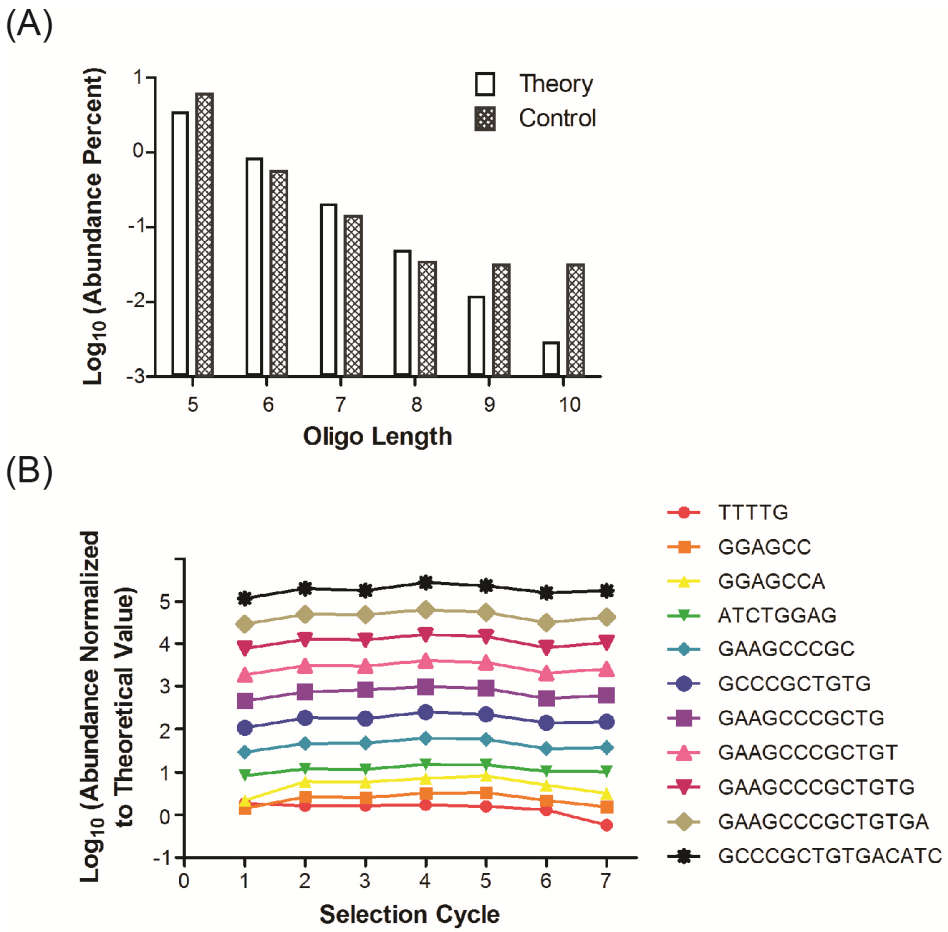

Figure 2.6 (A) Comparison between the theoretical abundances and the real abundances of the most abundant oligonucleotides in the control sample. (B) Evolution profiles of the most abundant 5 nt through 15 nt oligonucleotides.

sequence in control (see Figure 2.6 (A)). Another interesting phenomenon is that "GGAGCC" was the same motif found in the short byproducts from pool 2 through pool 6, confirming our conjecture that the short byproducts were formed as a result of mistakenly amplified GC-rich sequences.

I then analyzed the top 1,000 abundant oligonucleotides at each length from 5 nt to 15 nt. Similar as in genomic sequencing, contigs were identified by piecing together overlapping oligonucleotides into 40 nt regions. Although no meaningful results were obtained when aligning 5 nt oligonucleotides, all the other alignments gave the same results and contigs were discovered. I also obtained the top 1,000 abundant oligonucleotides at each length from 5 nt through 15 nt solely from the control file, and performed the same analysis. The most abundant contig in control showed up seven times, corresponding to an abundance of 2.33×10^{-2} % in the desired products. This slight enrichment was only due to inappreciable preference in multiple PCR reactions and purification process, and helped us identify contigs generated due to CE selections in the other seven files. Nine contigs were found to have abundances greater than 2.33×10^{-2} % at least in one pool. These contigs were summarized in Table 2.3 as C1 through C9 in abundance descending order. Not surprisingly, the most abundant oligonucleotides (6 nt to 15 nt) shown in Figure 2.6 (B) were all parts of C1. Within these nine contigs, only the top three had abundances that were significantly higher than 2.33×10^{-2} % (S/N= 3). The evolution profiles of the top three contigs were illustrated in Figure 2.7. None of the three contigs were present in control. C1 had a significant abundance right after one cycle of

selection, and the abundance kept increasing until pool 4, after which a slight decrease was observed in pool 5 and a sharp decrease was observed in pool 6. This trend is opposite to the trend of the most abundant short byproducts from pool 4 through pool 7 as shown in Table 2.1. The other two contigs had similar evolution profiles, with the highest abundances appearing at pool 5 for C 2 and pool 4 for C 3. On the other hand, the highest abundance in Figure 2.7 was still less than 1%, suggesting relatively high sequence diversity in the selected pools.

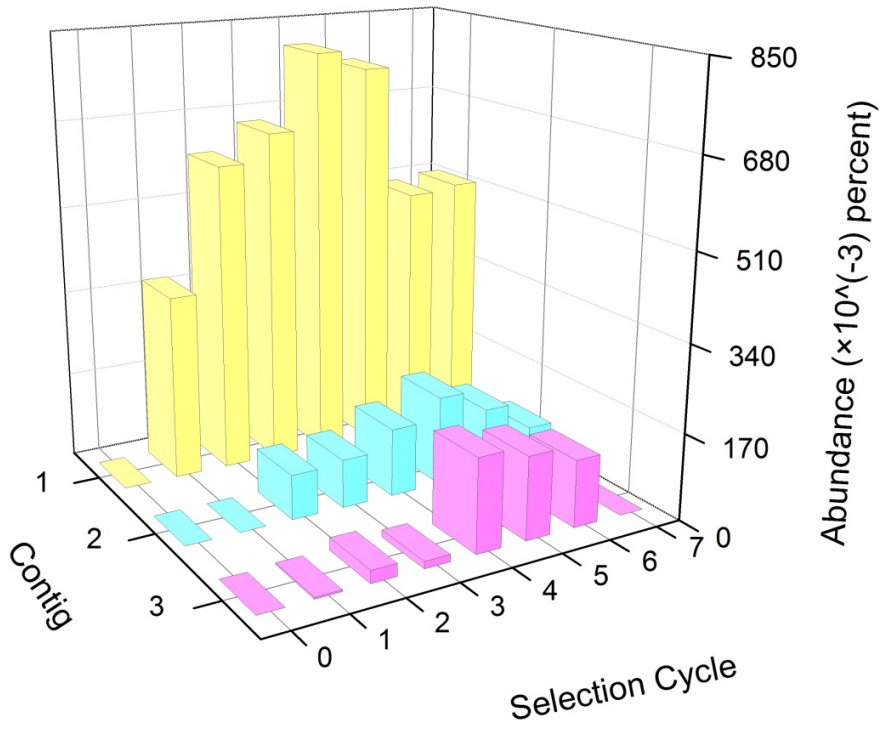

Figure 2.7 Evolution profiles of contigs. The abundances ($\times 10^{-3}$ percent) of the top three contigs were shown over the selection cycles, including the control sample. Yellow: C 1, Blue: C 2, Pink: C 3.

The top 1,000 abundant oligonucleotides at 6 nt through 15 nt were further analyzed to search for possible motifs in the desired products. Each oligonucleotide in the top 1,000 abundant oligonucleotides was searched in the

desired products; however, they either appeared at the same register position, for most of the time, or randomly distributed in sequence at abundances around the theoretical abundances of random oligonucleotides. Since the home-programmed code excludes the possibilities of finding mismatched or discontinuous motifs, I also divided the original eight files into subgroups containing ≤5,000 sequences and analyzed these subgroups using MEME. This size was chosen as a compromise between running time of the program and being representative of sample diversity in the original files. However, the results showed that the motifs with high and medium consensus levels were top oligonucleotides present in the top 1,000 abundant oligonucleotides (data not shown). CodonCode Aligner3.7.1 was also employed to perform sequence alignment on 3,000-sequence subgroups. Again, no new motifs or contigs were discovered other than the ones that were previously identified by home-programmed code. MEME and ClustalW2 were then used to perform alignment on the discovered contigs. Although several motifs were found with medium consensus level using MEME, the results by ClustalW2 clearly showed heterogeneity of these contigs, with the highest similarity score of 42 out of 100 between C 2 and C 6. The motif regions were highlighted in different colors in Table 2.3, and the motifs were given in Figure 2.8.

Table 2.3 Contigs and randomly picked aptamers (without primer regions) found in selected pools in abundance descending order

Contig (C #) and Randomly Picked Aptamers (R #)	Sequence (5' to 3')	Length (nt)	Abundance		
			Total Number	Number in Pool 4	$\times 10^{-2}$% in Pool 4
C1	GGAAGCCCGCTGTGACATCTGGAGCCGGTTCCCGAGCCA	40	869	199	80.3
C2	CGAGGATATCCGATTGTGGGTGACGTATTGGACACCTTCCCG	42	130	33	13.3
C3	TGGCTGTTATTTCTCTCCGGCCTCAGATCTTGGGCCCAACA	40	93	45	18.2
C4	ATCCGGTTAGACCACTTGTCAATGAATGAGATTGTCAACG	40	37	8	3.23
C5	TCTAGCGTTGTCTCGATCTATACCGTCGCGAATCGTAATG	40	29	6	2.42
C6	TGCCATTCCTGTTGAGTATGACGTAATGGTAGTTGCGAAG	40	23	8	3.23
C7	CGTTACCTGTAGGCGTATCTTTGTACAACCCTCTGGCGAG	40	17	7	2.82
C 8	AGTTTCTTGGGTTCTAAACGGATTTTATAAGGATTTTTAA	40	13	4	1.61

C 9	AAACTCAAAGACACTTTGGTTGAGTTGTGTGATACCCGTAG	41	11	0	0
R 1	GTCTGTACCAATCTAAAAGTGTATATTATATCTACGAGCCT	41	1	1	0
R 2	ATTCTGGGGAGTGGTACGTAAGCAGTGCCACCACGGTTAT	40	1	1	0.403
R 3	CCTTGTCATGATAGCTATGTTTCAGCCAACCATACCCGCT	40	1	1	0.403

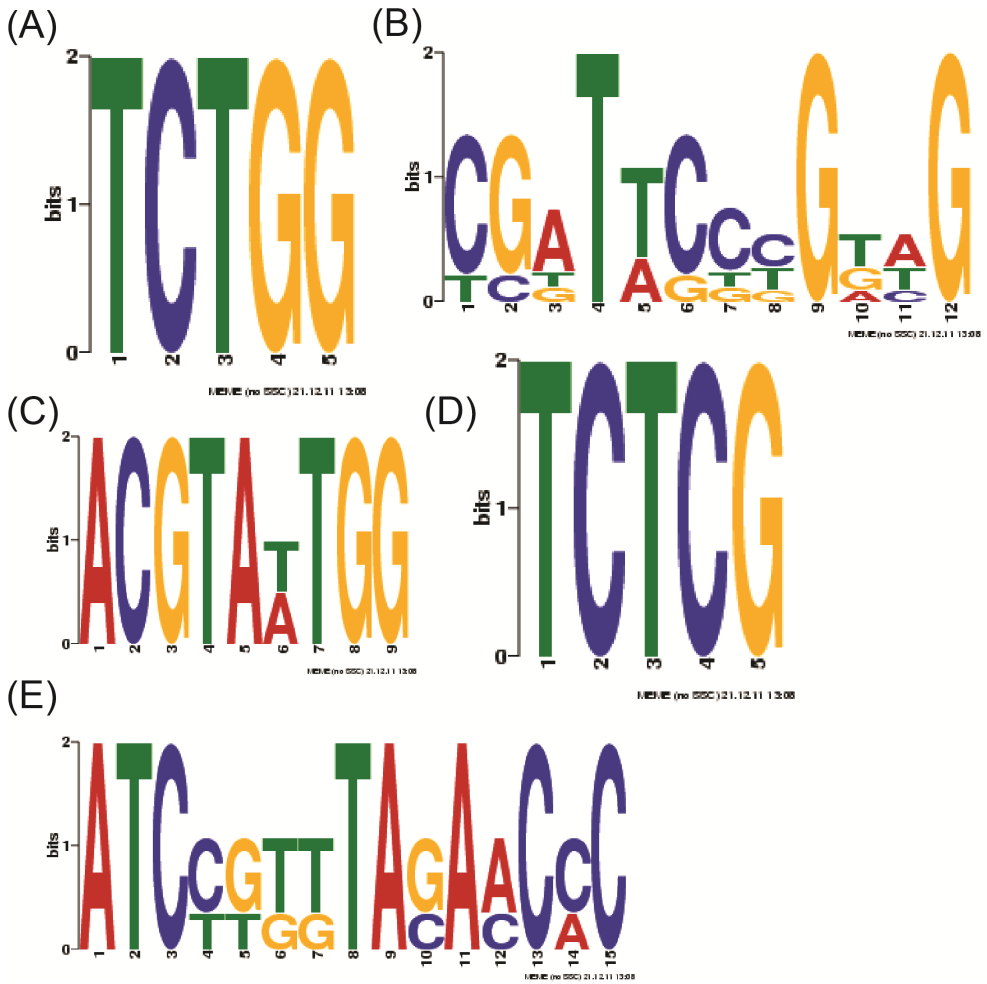

Figure 2.8 Motifs found in contigs. (A) Motif 1 was highlighted in pink in Table 2.3. (B) Motif 2 was highlighted in yellow in Table 2.3. (C) Motif 3 was highlighted in turquois in Table 2.3. (D) Motif 4 was highlighted in dark blue in Table 2.3. (E) Motif 5 was highlighted in green in Table 2.3.

2.4.5 Aptamer Affinity Characterization

To further characterize the aptamers, the top three contigs and another three sequences randomly picked from pool 4 (Table 2.3, R1~R3) were synthesized. As shown in Table 2.4, the affinity of each aptamer was characterized by both ACE and FP. Figure 2.9 (A) shows the electropherograms obtained by titrating 5

nM R3 with increasing amount of VEGF. The decrease of the free aptamer peak started right at 5.9 nM. Notablely, from 94 nM a small peak ahead of the free aptamer peak showed up and increased with VEGF concentration, indicating the aptamer-VEGF complex. Figure 2.9 (B) is the binding curve obtained from ACE experiment. K_d obtained from ACE experiment was 24±11nM, which was close to 50±11 nM from FP experiment in Figure 2.9 (C). Table 2.4 summarizes K_d values of the six sequences by both methods. The overall trend is that smaller K_d values were obtained from ACE experiments, even after modification to bound fractions in FP due to fluorescence quenching (see Chapter 1.3.3 for detailed calculations). However, most K_d ranges from ACE experiments overlapped with those from FP experiments. The affinities obtained for individual aptamers were also similar as those of selected pools and were comparable to or even better than DNA aptamers selected against $VEGF_{165}$ or VEGF peptide.[240-244]

Table 2.4 K_d values of selected aptamers measured by ACE and FP

Sequence	ACE*	FP*
C 1	49 ± 26 nM	101 ± 50 nM
C 2	21 ± 10 nM	162 ± 82 nM
C 3	66 ± 35 nM	164 ± 78 nM
R 1	53 ± 27 nM	100 ± 84 nM
R 2	72 ± 40 nM	88 ± 23 nM
R 3	22 ± 11 nM	46 ± 11 nM

***** Errors represent 95% confidence intervals.

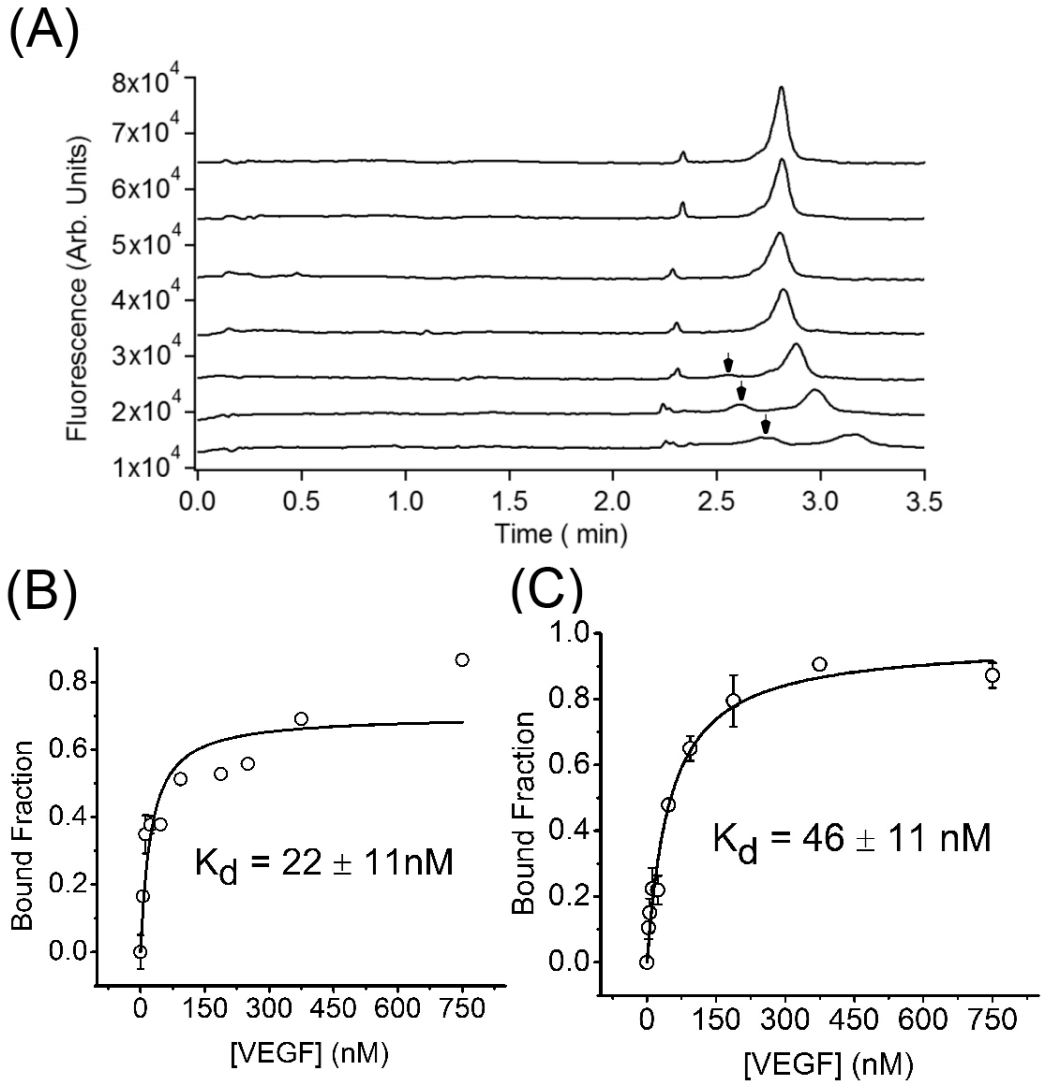

Figure 2.9 (A) An overlay of electropherograms from one ACE experiment, titrating 5 nM R3 by increasing amount of VEGF. The VEGF concentrations were 0 M, 5.9×10^{-9} M, 1.2×10^{-8} M, 2.3×10^{-8} M, 9.4×10^{-8} M, 3.8×10^{-7} M, and 7.5×10^{-7} M from the top to the bottom. The major peak was the unbound aptamer, and the minor peak indicated by arrow was the VEGF-R3 complex. (B) The binding curve obtained from ACE experiments. (B) The binding curve obtained from FP experiments performed with the same samples. In (B) and (C) error bars on data points represent standard deviation of three parallel measurements, and the errors of K_d represent 95% confidence level.

2.5 Conclusions

In summary, I successfully generated ssDNA aptamers for $hrVEGF_{165}$ using CE-SELEX for the first time. Affinities of the abundant and randomly picked

aptamers were characterized by two orthogonal methods, ACE and FP, with estimated mid-nanomolar dissociation constants. The high-throughput 454 pyrosequencing offers information of up to 3×10^4 sequences from each selected pool and the control sample, allowing an in-depth study of evolution profile of CE-SELEX. By using home-programmed code, and online and commercially available software, abundant oligonucleotides and three enriched contigs were identified. Instead of randomly distributed in the 40 nt random region, the top abundant oligonucleotides appeared at the same register position, and overlapped each other to extend to full 40 nt-long contigs. MEME and Codon-Code Aligner analyses on 5,000- or 3,000-sequence subgroups confirmed these abundant oligonucleotides and contigs, without discovery of new motifs. Although enriched sequences were identified, the highest abundance of 0.8% indicated the remained heterogeneity even in the pool with the highest binding affinity. The evolution profile of the undesired short products was also monitored. The sharp quantity increase in pool 6 and 7 was in contrast with the decrease of the three contigs in the same pools. A close look into the most abundant short byproducts revealed that they were mistakenly amplified C1. Thus, this study confirms one of the advantages of CE-SELEX, which is to generate aptamers beyond several binding motifs, benefitting from the free solution incubation and separation. On the other hand, the presence and increasing trend of the short products also impel us to develop novel technologies that can greatly increase the sample input and collection, in order to increase the possibility of including more high-affinity

ligands and decrease the error rate in PCR, to ultimately improve aptamer affinity and selection efficiency.[251]

Chapter 3 Isolation of DNA Aptamers Using Micro Free Flow Electrophoresis

3.1 Summary

A micro free flow electrophoresis (μFFE) device was used to select DNA aptamers for human immunoglobulin E (IgE). The continuous nature of μFFE allowed 1.8×10^{14} sequences to be introduced over a period of 30 min, a 300-fold improvement in library size over capillary electrophoresis based selections (CE-SELEX). Four rounds of selection were performed within four days. Aptamers with mid-nM dissociation constants for IgE were identified after a single round of μFFE selection.

3.2 Introduction

Conventionally, affinity selections in SELEX are performed using nitrocellulose membranes or affinity chromatography. The procedure and downsides were discussed in Chapter 1.1.2; as a result, SELEX process was labor intensive and time consuming. Since the earliest aptamer publications,[36, 37] great efforts have been made to develop protocols that minimize solution volumes, are simple and fast, and are compatible with automated process. Cox *et al.* were the first to automate the SELEX process by incorporating the separation, PCR and purification apparatus into a robotic pipetting worksurface.[63, 64] Twelve cycles of selection were completed within 42 h, reducing the selection time from several weeks to less than two days.

Recent developments in microfluidics provide opportunities for improving the SELEX process. It is anticipated that a high throughput lab-on-a-chip device that combines high separation efficiency, increased PCR speed, low sample volume,

minimal contamination, and automation will be possible.[75, 252-254] Towards this goal several microfluidic devices have been reported for isolating aptamers and were reviewed in Chapter 1.2.2, such as the microfluidic device patterned with sol-gel droplets on individual microheaters, the continuous-flow magnetic activate chip-based separation (CMACS) device[84] and the micromagnetic separation (MMS) chip.[85] However, there are limitations to each of these strategies. There are concerns regarding protein stability and integrity through the multiple selection cycles in the sol-gel-microheater device; complicated target immobilization procedures, elongated incubation, and the potential for non-specific interactions with surface structures are of concern in the CMACS and MMS devices.

Capillary electrophoresis based selections (CE-SELEX) address many common issues encountered in traditional selections.[45-48] With the advantages discussed in Chapter 1.1.3, aptamers were usually obtained in 2-4 rounds of selection; on the other side, however, CE-SELEX suffered from drawbacks such as small sample input and complex timing to determine complex collection window. Micro Free Flow Electrophoresis (μFFE) provides a unique solution to the challenges presented by CE-SELEX. Contrary to many separations, μFFE can be used to continuously introduce, separate and collect analytes.[93-95, 255] Analyte is continuously streamed into a planar separation chamber. An electric field is applied perpendicularly to the pressure driven flow, deflecting analytes laterally according to their mobility. μFFE designs and separation conditions have been optimized allowing long term operation of the device.[98, 256] μFFE separation

theory and sources of band broadening have been characterized.[257] Unique μFFE detection strategies have been demonstrated.[258] Most interestingly, a gradient technique has been developed to measure aptamer-target equilibria, demonstrating the feasibility of using μFFE to separate binding sequences from non-binding sequences.[167]

This chapter describes the first successful application of μFFE in a SELEX selection for DNA aptamers. IgE was chosen as a target since aptamers have previously been selected for this protein using both traditional SELEX[259] and CE-SELEX[45, 46] allowing a direct comparison with these approaches.

3.3 Experimental Methods

3.3.1 Materials and Chemicals

Human myeloma IgE protein was purchased from Athens Research and Technologies (Athens, GA). Nuclease free H_2O, forward primer 5'-FAM-AGC AGC ACA GAG GTC AGA TG-3', reverse primer 5'-Biotin-TTC ACG GTA GCA CGC ATA GG-3', the initial ssDNA library 5'-FAM-AGC AGC ACA GAG GTC AGA TG$(N)_{40}$ CCT ATG CGT GCT ACC GTG AA-3', and the selected aptamers were from Integrated DNA Technologies, Inc. (Coralville, IA). For PCR reactions, dNTPs and the 25 bp DNA ladder were from Invitrogen (Carlsbad, CA); *Taq* polymerase and ThermoPol buffer were from New England BioLabs (Ipswich, MA); the gel loading dye was from Promega (Madison, WI). Other chemicals were purchased from Sigma Aldrich (St. Louis, MO) at the highest grade available, except Acetic acid (CH_3COOH, 99.7%, Mallinckrodt Baker), 4-(2-

hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 99%, Alfa Aesar), $MgCl_2$ (99.8%, Mallinckrodt Baker), NaCl (99.0%, Spectrum), and KH_2PO_4 (99.9%, J. T. Baker). All buffers were prepared in nuclease free H_2O, and filtered through 0.2 μm membrane filters before use.

3.3.2 μFFE Fabrication

μFFE devices were fabricated using previously reported procedures.[167] Briefly, three photolithography steps were performed to define the electrode channels, the separation channel, and pattern the electrodes onto 1.1 mm borofloat wafer (Precision Glass & Optics, Santa Ana, CA) according to the design shown in Figure 1.4 (B). The patterned chip was anodically bonded to a second borofloat wafer that had previously been drilled with inlet and outlet holes. Channel depths were approximately 20 μm in the separation channel and 100 μm in the electrode channels. The sample inlet hole had a diameter of 355 μm.

3.3.3 μFFE Separation

Before separation, the μFFE chip was coated with PEO to suppress the electroosmotic flow.[260] Briefly, 1 M HCl was first pumped into chip at 6 mL min^{-1} for 10 min, followed by 0.2% PEO in 0.1 M HCl at 3 mL min^{-1} for 10 min. Finally, the separation buffer (25 mM HEPES, 300 μM Triton X-100, adjusted to pH 7.0 by 1M NaOH) was pumped at 6 mL min^{-1} for 10 min to remove HCl and unbound PEO. Before the selection, the ssDNA library was heated to 72 °C for 5 min and then gradually cooled to room temperature. For the 1^{st} round of selection, 100 μM random sequence ssDNA library was incubated with 10 nM IgE at room temperature for 20 min in the binding buffer (tris(hydroxyamino)methane-glycine-

potassium (TGK) buffer, composed of 25 mM tris(hydroxyamino)methane, 192 mM glycine, and 5 mM KH_2PO_4 at pH 8.3). The following IgE concentrations were used in subsequent selection rounds: 10 nM in the 2nd round, 1nM in the 3rd round, and 100 pM in the 4th round. The mixture (40 µL in total) was then loaded into a 100 µL syringe (Hamilton Company, Reno, NV), and pumped into the µFFE chip (Figure 3.1, hole (2)) using a syringe pump (PicoPlus, Harvard Apparatus, Holliston, MA) at 100 nL min^{-1}. The separation buffer was pumped into the device (Figure 3.1, hole 1) using a second syringe pump (pump 22, Harvard Apparatus) at 1 mL min^{-1}. An electric field of 150 V cm^{-1} was applied across the chip so that the nonbinding ssDNA was separated from aptamer-IgE complexes. A stable current of 0.32 mA was observed over the 30 min separation time. With EOF suppressed, the free ssDNA and the complex were split into two fractions (Figure 3.1, holes 3) and collected separately.

3.3.4 Laser Induced Fluorescence (LIF) Detection

The beam from a solid state laser (488 nm, 50 mW, Newport Corp, Irvine, CA) was expanded into a line and focused across the separation channel 1.5 cm downstream from the sample inlet. A microscope objective (3× zoom) was positioned above the detection zone. Fluorescence images were recorded every second using a Cascade 512B CCD camera (Photometrics, Tucson, AZ) through an AZ100 stereomicroscope (Nikon Corp., Tokyo, Japan). An Endow GFP bandpass emission filter cube (Nikon Corp., Tokyo, Japan) consisting of a dichroic mirror (495 nm cutoff) and two bandpass filters (450-490 nm and 500-550 nm) were used for wavelength selection. Images were processed using

MetaVue software (Downington, PA). Linescans were also recorded using MetaVue, and later analyzed in Cutter 7.0.[261]

3.3.5 PCR Amplification and Purification

Collected sequences were PCR amplified immediately after each selection round. In the final reaction vials, there were 1 mM each of the four dNTPs, 7.5 mM $MgCl_2$, 500 nM forward and reverse primers, 1× ThermoPol buffer and 0.05 units/μL *Taq* polymerase. A negative control with all PCR reagents but no ssDNA was also performed every round to verify the absence of background contamination. 23 cycles of denaturation (94 °C, 30 s), annealing (55 °C, 30 s), and extension (72 °C, 20 s) were performed, with a final extension at 72 °C for 5 min to ensure the protruding adenine nucleotide was incorporated in the newly synthesized template to facilitate cloning reaction. A 1.5% agarose gel with ethidium bromide staining was used to confirm the presence of the desired PCR products. Single stranded FAM labeled DNA sequences were obtained using an on column purification followed by the ethanol precipitation.[46]

3.3.6 Dissociation Constant K_d Measurements

Affinities of the selected pools and individual aptamer sequences were measured using both affinity capillary electrophoresis (ACE) and fluorescence polarization (FP). Approximately 2.5 nM ssDNA samples were titrated with increasing concentrations of human IgE. A commercial CE system (P/ACE MDQ, Beckman Coulter, Inc., Fullerton, CA) equipped with laser induced fluorescence detection (λ_{ex} = 488 nm, λ_{em} = 520 nm) was used to perform ACE experiments. Samples were injected into a 50 cm long, 50 μm i.d. fused-silica capillary

(Polymicro Technologies, Phoenix, AZ) at 1 psi for 4 s. The separations were then performed at 30 kV in TGK buffer. The peak heights of the unbound ssDNA were used to calculate the bound fractions and estimate K_d using the following equation:[262]

$$f_a = \frac{c}{1 + \frac{K_d}{([P]_t \quad 0.5([D]_t + [P]_t + K_d \quad (([D]_t + [P]_t + K_d)^2 \quad 4[D]_t[P]_t)^{0.5}))}} \qquad \textbf{Equation 3.1}$$

in which f_a is the bound fraction, $[P]_t$, $[D]_t$, and c are total IgE concentration, total DNA concentration, and maximum bound fraction, respectively. $[P]_t - 0.5([D]_t + [P]_t + K_d - (([D]_t + [P]_t + K_d)^2 - 4[D]_t[P]_t)^{0.5})$ represents the free IgE concentration.

In FP experiments, 15 µL of the same samples used in ACE were loaded into a corning 3540 microplate (Corning Incorporated, Corning, NY) and experiments were performed on a Synergy™ 2 Microplate Reader (BioTek Instruments, Inc., Winooski, VT). Parallel and perpendicular intensities (λ_{ex} = 485 ± 20 nm, λ_{em} = 520 ± 20 nm) were measured, and polarization values were calculated using Gen 5™ software (BioTek Instruments, Inc., Winooski, VT). The bound fraction was determined according to:

$$f_a = \frac{P - P_o}{P_m - P_o} \qquad \textbf{Equation 3.2}$$

In which P, P_o and P_m are the measured polarizations of a sample, free DNA, and the complex, respectively. Meanwhile, the overall fluorescence intensity was monitored and corrections to the bound fractions were made according to the intensity change as previously reported.[163]

3.3.7 DNA Cloning and Sequencing

A TOPO TA Cloning® Kit for Sequencing (Invitrogen, Carlsbad, CA) was used to perform cloning reactions and grow colonies. Briefly, 1 µL of the selected ssDNA from pool 1 through pool 4 was added to 99 µL PCR master mixes and amplified separately, as described above, except using unlabeled forward and reverse primers. 4 µL of the 100 µL fresh PCR products were incubated with TOPO® vector at room temperature for 5 min to allow the cloning reactions to take place. The inserted vectors were then chemically transformed into One Shot® TOP10 competent *E.coli* cells. The mixtures were incubated in a 37 °C shaking incubator for 30 min and then spread onto selective plates containing 50 µg mL^{-1} kanamycin. For each pool, two agar plates were used to grow colonies, which produced 300 to 400 colonies. Individual colonies were then randomly picked and cultivated in a 96 well plate with liquid LB containing 50 µg mL^{-1} kanamycin at 37 °C for another day. Cell pellets were sent to the Biomedical Genomics Center at the University of Minnesota for Sanger sequencing. Since some wells of the plate were used as controls, and some pellets were not sufficient to obtain confident sequencing information, 23, 28, 19, and 6 sequences were obtained from pools 1, 2, 3, and 4, respectively (see Table 1 for aptamer sequences).

3.4 Results and Discussion

3.4.1 µFFE Selection

Figure 3.1 (A) demonstrates the µFFE separation observed during the first round of selection. An electric field of 150 V cm^{-1} was applied across the flow

channel. Unbound ssDNA sequences were deflected toward the anode due to the suppressed electroosmotic flow (EOF). The aptamer-IgE complexes, which were only deflected minimally, were well separated from the unbound sequences. It should be noted that μFFE facilitates collection of analytes with low mobilities. Pressure rinses were necessary to observe or collect aptamer-IgE complexes in previously reported CE separations. The flow was split into two streams at the exit of the μFFE separation chamber. The IgE-DNA complexes were collected into a centrifuge tube through the tubing connected to the right nanoport. This continuous collection strategy was much simpler than that of CE-SELEX, eliminating the complicated timing associated with collecting complexes as they migrated off the end of a capillary. A linescan (Figure 3.1 (B)) recorded across the detection zone reinforces how well resolved the binding sequences are from the non-binding sequences. The high concentration of library (100 μM) used in the first round of selection caused the signal for the unbound sequences to go off scale but this was not a concern since the primary goal was to identify the position of the peaks for accurate fraction collection, not quantitation.

In the first round of selection, 100 μM of fluorescently tagged ssDNA library was incubated with 10 nM IgE in TGK binding buffer for 20 min. After incubation, 3 μL of the library-IgE mixture was introduced into the μFFE device over a period of 30 min, which corresponded to 1.8×10^{14} sequences. A typical CE separation only allows a discrete injection of approximately 8 nL.[48] Over 300-fold more volume, and therefore sequences, could be introduced into the μFFE device. The

number of sequences could easily be increased further by increasing the analyte flow rate, library concentration, or collection time.

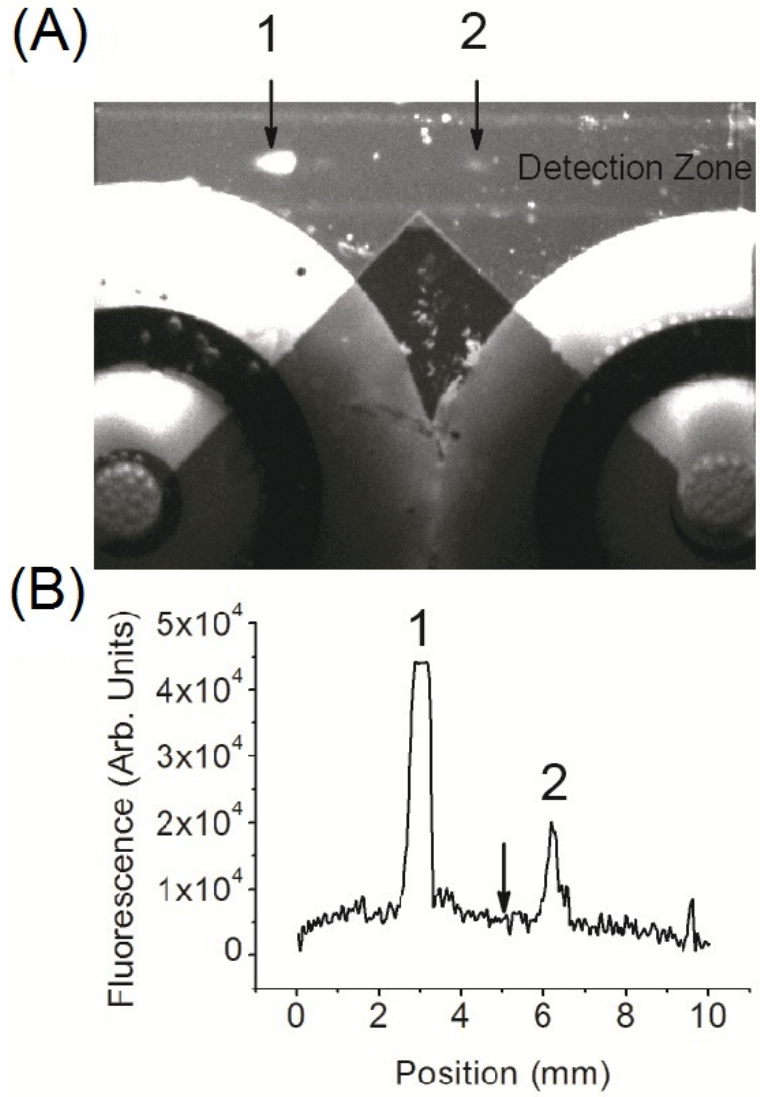

Figure 3.1 (A) An image of a μFFE separation of free (1) and bound ssDNA (2) recorded during selection cycle 1. The fraction collection channels and outlet ports are clearly visible in this image. (B) A linescan across the detection zone imaged in (A). The arrow indicates the fraction cutoff point at the exit of the μFFE channel demonstrating clear separation of free (1) from bound (2) ssDNA. The anode is on the left in all images.

The multiple-depth μFFE design[98] generated high flow over the electrodes to dissipate electrolysis bubbles while minimizing flow, and therefore dilution, in the

separation chamber. The dilution factor is defined as the ratio of the collected sample volume to the introduced sample volume. In a μFFE device, the planar flow profile can be described using lubrication theory according to the following equation:[263]

$$q = \frac{\Delta P H^3 w}{12 \eta L} \qquad \textbf{Equation 3.3}$$

In which q is the volumetric flow rate, ΔP is the pressure difference, H is the channel depth, w is the channel width, η is the buffer viscosity, and L is the channel length. Comparing the electrode and separation channels: the ratio of the channel depths is 20 μm/100 μm = 0.2 (separation channel/electrode channel), the ratio of channel widths is 10 mm/ (2×2 mm) = 2.5 (*i.e.* two electrode channels), and the ratio of channel lengths is approximately 1. The ratio of flow volume through the channels is $0.2^3 \times 2.5/1$ = 1: 50 (separation channel/electrode channel). Solution is only collected from half of the separation channel. As a result, approximately 1% of the total volume through the μFFE device is collected. Over a period of 30 min, 30 mL separation buffer and 3 μL sample are introduced into the chip, resulting in a collected volume of 300 μL. This calculation agrees very well with experimental results. Comparing with 3 μL of library injected yields a modest 100-fold dilution of the binding sequences during the μFFE separation. This compares favourably with dilution factors of 6,000 typically encountered in CE-SELEX.[45-48] The higher DNA concentration in the collected fraction and increased volume collected facilitates PCR and decreases the potential for contamination or non-specific amplification. Furthermore, the transit time through

the μFFE flow chamber is only 10-20 s, greatly decreasing the potential for dissociation during the separation when compared to CE, in which separations typically take 5-15 min. The electric field applied in μFFE (150 V cm^{-1}) is also lower than that typically used in CE-SELEX (~500 V cm^{-1}), further decreasing the potential for dissociation.

3.4.2 Aptamer Characterization.

Four rounds of μFFE selection were performed. The affinity of the selected pools for IgE was monitored using two orthogonal methods:[250] fluorescence polarization (FP)[163] and affinity capillary electrophoresis (ACE)[45] (see Experimental Methods, Dissociation Constant K_d Measurements). As shown in Figure 3.2, the initial library had a low affinity for IgE in TGK buffer, (K_d = 460±160 nM and 550±280 nM as measured by FP and ACE, respectively). A significant improvement in affinity was observed after a single round of μFFE selection, with K_d values of 29±15 nM and 58±55 nM measured by FP and ACE, respectively. No further improvement in affinity was observed after the second round of selection. IgE concentration was decreased by a factor of 10 (1 nM) in the third round of selection and by a factor of 100 (100 pM) in the fourth round of selection in an effort to improve the stringency of the selection. However, the enhanced stringency did not result in improved affinity for IgE, suggesting that the selection had converged after a single round of μFFE selection.

Aptamers were randomly cloned and sequenced from the DNA pools after each round of selection. 23 sequences were identified from round one, 28

sequences from round two, 19 sequences from round three, and 6 sequences from round four. Random regions of these sequences were listed in Table 3.1.

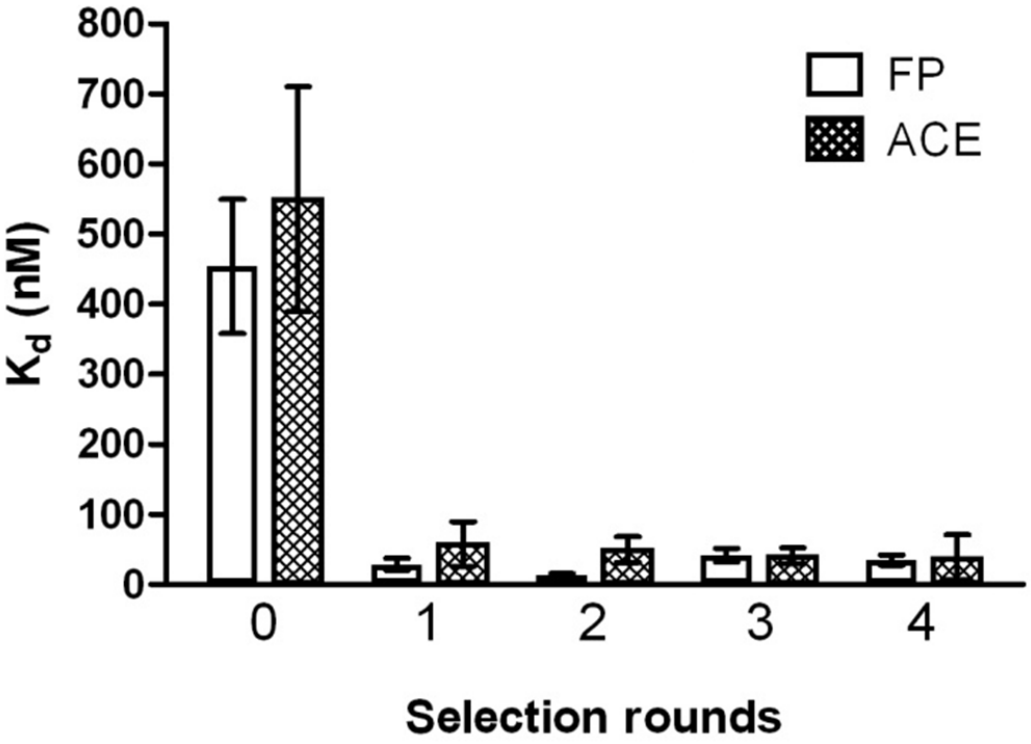

Figure 3.2 K_d of the starting library and the selected pools measured by fluorescence polarization (FP) and affinity capillary electrophoresis (ACE). Error bars represent 95% confidence interval.

Table 3.1 Aptamer sequences from round one through round four (random regions)

Clone	Sequence (5'→3')
1.1	TTATCAACGCGTCTGTCCTATAAAACTCATGTAAATTAAT
1.2	CCTGTGTCCATTATATGCTAAGGAAATTCATTACGTTCCT
1.3	GGACTAACTCATGATTCAGTACGATAGATCCTACCATTTC
1.4	ATCATAGGTCTGCTTGCTTGTATTATTTGAGCGGTCCATC
1.5	GCTATAACTCAGATTACAATATTTTGAATTGGTCCCAAT
1.6	TAGCAAGAGTACTGAGTAAATCTACAATTTCAGCCATTAA
1.7	ATTCCAAAGTTCCATCTAAAGCGATTATAGGGTACTT
1.8	GGCGCGGTTTCTGGTTCTGGATCGGCTCCACGGCTG

1.9	TTGAACGAAACACTAATTAAATGGTTTTCTCGTTAAGCGA
1.10	ACCTGACCTTTTATCTCTCAGCTATAATTTTGCCTTTGTT
1.11	GTTAACAACTTCGCATTAACGACTATAGCATCGGCCTGTA
1.12	AATTTCTATCGTCAGTTTATGAATCGGGAGGAAGCGCGG
1.13	ACTTAGTTCAGAATATGAGTACTATGGTATAAGGTCGATC
1.14	GATCATCCTTTTAAAGACGGAGCTCTTATATTATTAGAAT
1.15	GCGCATTATCCTAGGCGCTGGGAGGCTCTATTAAGGCCGGG
1.16	GCTTGGTTTTTAGTTATCAAAAGATCGTGTAGCCACATTT
1.17	ACAGTAAGTGTACTAGAATAGCGTGCTCAAACCGTTAGAG
1.18	GAGGGTATTAAGCTATAAATTATTGCGGGAACCCCGTTAT
1.19	CCCAGCACAGTTCTTGTCATTTAGAGGTATTATTGTGGTT
1.20	CTAGATATTATTGGCATGAAGACTCTACTGAAGGGACATG
1.21	CGAGTTTTAATATAAAGCTACAATTATTGAGCACGTTGAA
1.22	TGGAGGGTGTGTCTAAGTTCCGCGTGTGTGATGTGGACTA
1.23	CTTACCTCCCTCACGATAACCCGGAAGTAGTCGACTAACC
2.1	CCAGCGTTAGAGACTCCGACCTGGTGACATGTATTCACAA
2.2	AGTTGCCCAGCTTATTCCGGACAATAATAGTGGCGGTTTG
2.3	TCACATGCATAAATATACAGTCCTTTTTGTGCTCAACA
2.4	TGCTCCAATTTCAAGCTGTTGCATCTAGGCTAAAGTGCAT
2.5	CAGGGCAATGTTATGTGTTAGGAAAATATTAAGCACTG
2.6	AGGGGTTAGATATGAAACGTTTAGTTGGGGGGTCGGAAAA
2.7	ATCAACGTAAGCTTTTCTCCATTATTAAGTAGTCCTAAAG
2.8	AGCACATGGGATCAGTTCTCTTCTGTATGCGCATAGAGGT
2.9	TCGGGTATCCTTATAACATTAAGGCTAACTCATTAGTTTA
2.10	CATTTCCGTAGCACTCCTCTCTAATAGGGTAAAGATTTAG
2.11	CTTAAAAACGTTCAACCTGCGTGCTTATTTCATACTACTG
2.12	GAATACCGTACGGTGCTCGGTAGTGGTGGTCCCGGTTAGA
2.13	CTCCCGCATACAAAAATAATTTATTAGTCATTATTTCGCA
2.14	CTTCGGGCTATGTCATAGAACTGATAGGTAGGTGTTTAAG
2.15	ACGTTCGGGAAATTCTCAGAGTGATACGCTGTTTTTAATA
2.16	GCTTTGGGGATGAATCTCATGATATTATGTTCTCACC
2.17	AGGCTTAACACGTATCGCCCAGTCATATAGCCAATCACAG
2.18	CAAGTCTCGACGATAATTTAGGCTATAATAGTCCTACACA
2.19	CATTAGGTTGCGGCGGACTTTGATGGTTGCTGGGGCGG
2.20	TGGAGACTGCTGTGCTGGGGTCGGTTGAGATAGTGAGGGG

2.21	GTCAAAAGCTTATTCGTTGAGTAAGTGCGATCAACGACCA
2.22	TTCCATATAGGCGACCATGCTGTCAGGTCCGCATCTATA
2.23	AAGCATTGATCCTAGCGTTAGGCTATGGCTTTTACTTAAA
2.24	CTCCCGCATACAAAAATAATTTATTAGTCATTATTTCGCA
2.25	GTAGCTTGCGTGTAAATAAATTCACAAGACACTGGTGATC
2.26	TGCGTGTCGCTATTAAGTGACTCCTAGGGGCTAACTAGTG
2.27	GTCAAAAGGGCGCTTGTTTGGCGGGGTAATTCTCTACGAT
2.28	GATTAGTATCCGAGCAGTATGAAGTAACTTGCGTGGCGTA
3.1	TATACCACGTTTTATGGCTTTCATTCATTCTCATAATGGG
3.2	TAACTAGATCGTATATACGCGTATGAAGGTTCGTTAATAG
3.3	AGGCGGATATGGTTTCTCTTTCGAGGTGGTTTTCCAGGT
3.4	GGCTGAGTTTACATTTTATAGTAGTAATGAAACTTTCAAA
3.5	TATAGTTATTTTCAATGTTTCTAGGTACAGTGTGGTATAA
3.6	TATGCAAATGTTCTCAGTCGTGAAAACCTTGCTTATGTTC
3.7	AATGAGTACAGATTTGGTCCTAAGCTTCTCTATTTAAAGA
3.8	TTTAGGTCGCTCAGTACAGGCCATCTTATTTGTCTGAGAT
3.9	AAGTTAAAGAATGGGCGTTTTGTGTAAATCAATAAGTCAA
3.10	GTCAGAGGGGTAGCTTGATGGTGGGTGCGGGGGGGCGAGT
3.11	TACTTATTTTAGCCCTGTTCTCATTCTGCTTAGCGAACAA
3.12	TGATTGTTAACCGGCTTGTGTGGTCCATATTCCATCTCC
3.13	CATATAGTCTGCAAGTGGGAAACTGGTTATAGACCTCACCC
3.14	TATAGTTATTTTCAATGTTTCTAGGTACAGTGTGGTATAA
3.15	TCGCAGCTCGGTTGATCAGTGGCGGGGCTCCCGTCTTTGT
3.16	TTGATGTTCTCCGTTCTGCCGTTTAACTCAGTTTAGGCAC
3.17	ATACTCCAAAATCTATGAAAACGATGCGAACCAACTACTT
3.18	TATATATCGATACATTTTTACTGCAGCCTACCAAACGTGC
3.19	GAACGCTTCCGCATGTTAGCGTCTTAAGCTGCGTCTACTC
4.1	GTTCTAGAAGTGCACTTTATAAGTATCGGTTAAGTAACTT
4.2	ATTCTGGGGAGTGGTACCTACCAGTGCCACCCCCGTTAT
4.3	ATGTTAGGTGTGAACCGTAGTAGGTAAGGTTGAAATCACG
4.4	ATTCTGGGGAGTGGTACGTAAGCAGTGCCACCACGGTTAT
4.5	ATTCTGGGGAGTGGTACGTAAGCAGTGCCACCACGGTTAT
4.6	ATTCTGGGGAGTGGTACGTAAGCAGTGCCACCACGGTTAT

These sequences were analyzed to identify homologous sequences or motifs. As show in Table 3.2, no identical or similar sequences were found in round one; 2 identical sequences were found in rounds two and three; 3 identical sequences were found in round four with a 4th sequence demonstrating 89% similarity as revealed by software ClustalW2®. It should be noted that sequences that appeared multiple times in one round did not share any conserved regions or motifs when compared with sequences isolated form other rounds. Although the number of the clones was limited, identification of identical sequences in such a small sample size suggests a decrease in diversity as the pool converged. This is particularly evident in round four, where 4 out of 6 sequences isolated were very similar. The affinities of all sequences identified multiple times were measured by FP and ACE (see Table 3.2). The measured dissociation constants were remarkably similar and agreed well with the affinities measured for the bulk pools.

Finally, three aptamers were chosen from the cloned sequences obtained from rounds one and three using a random number generator. Dissociation constants for each of these sequences are shown in Table 3.2 (see Figure 3.3 for representative binding curves).

Table 3.2 K_d of selected aptamers

Aptamer[a)]	FP K_d[b)]	ACE K_d[b)]
Multiple copy sequences identified		
Clone 2.13 and 2.24	20 ± 4 nM	136 ± 58 nM
Clone 3.5 and 3.14	29 ± 7 nM	33 ± 21 nM

Clone 4.4, 4.5, and 4.6	39 ± 20 nM	58 ± 17 nM
Clone 4.2	44 ± 15 nM	62 ± 29 nM
Randomly chosen sequences from pool one and pool three		
Clone 1.4	23 ± 6 nM	58 ± 33 nM
Clone 1.13	22 ± 6 nM	63 ± 17 nM
Clone 1.18	32 ± 11 nM	66 ± 35 nM
Clone 3.2	20 ± 7 nM	50 ± 27 nM
Clone 3.9	29 ± 11 nM	62 ± 16 nM
Clone 3.11	28 ± 13 nM	47 ± 17 nM

[a)] Aptamers were given unique identifiers where the first number gives the selection round and the second number identifies a particular sequence cloned in that round. [b)] Errors represent 95% confidence interval.

No statistical difference was observed between sequences obtained after one or three rounds suggesting that the selection had converged after a single round of μFFE selection. K_d values for sequences chosen randomly were statistically indistinguishable from those that appeared multiple times in the cloning results. This result is similar to previous CE-SELEX selections that yielded seemingly diverse pools of aptamers with similar affinities.[45, 46] It should be noted that every sequence assessed demonstrated low nM affinity for IgE, even in the absence of negative selections, suggesting that similar to CE-SELEX, performing selections in free solution diminishes the opportunity for non-specific interactions. Dissociation constants measured using ACE were consistently higher than those measured using FP. Yang *et al.* have demonstrated that the high electric field of CE can affect observed half-lives and binding constants of protein complexes.[264]

This raises a concern regarding the high field typically used in CE-SELEX and whether this field modifies the selection environment, affecting the applicability of the selected aptamers. Buchanan *et al.* have demonstrated that a combination of low electric field and shorter time in the field minimize dissociation of aptamer-target complexes during a CE separation.[265] In practice this can be difficult to bal-

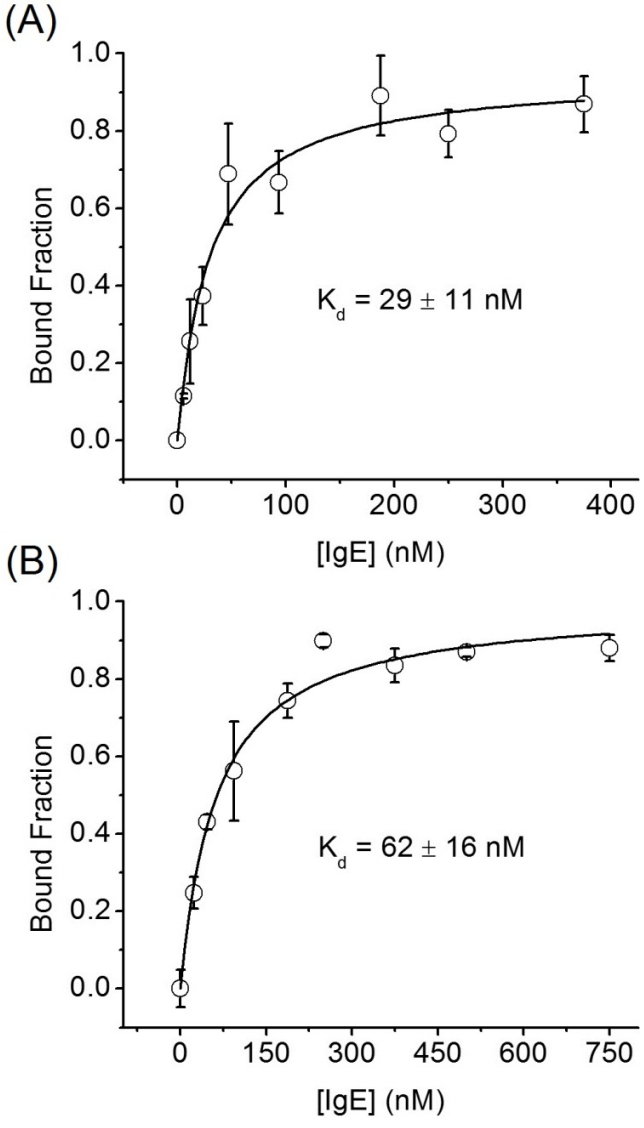

Figure 3.3 Binding curves for sequence 3.9, obtained using (A) FP and (B) ACE. Three measurements were taken at every IgE concentration. Error bars represent standard deviation. Errors of the K_d values represent 95% confidence interval.

ance in CE since separation time is inversely proportional to electric field. μFFE does not share this difficulty since electric field and separation time can be optimized independently. This unique property of μFFE is advantageous allowing more control over separation conditions when compared to CE.

Aptamers for IgE obtained using μFFE demonstrate similar dissociation constants to those obtained previously using CE-SELEX (~20 nM)[45, 46] and conventional SELEX (~10 nM).[259] It should be noted that these aptamers were obtained after a single round of μFFE selection, which compared favourably with the 2-4 rounds required by CE-SELEX[45, 46] or the 15 rounds required by conventional SELEX.[259]

3.5 Conclusions

In conclusion, we demonstrated the advantages of isolating aptamers using μFFE selections. Within 30 min of continuous μFFE separation and collection, 1.8×10^{14} sequences were assessed, a 300-fold improvement over CE-SELEX. μFFE also eliminated the complicated timing associated with fraction collection in CE-SELEX. Four selection cycles were completed within four days. Low nanomolar affinity sequences were identified after a single round of μFFE selection, suggesting that aptamers could be obtained even faster. Although the device was expected to have lower separation efficiency than CE, obtaining high affinity aptamers after a single selection round suggests a high rate of enrichment was achieved. The free solution μFFE separation simplifies the protocol, eliminating the need for target immobilization, elongated incubation, or

negative selections. With these advantages, we believe that μFFE device can be readily adopted to select aptamers for a wide range of targets.

Chapter 4 Tunable Control of PLN Binding to ssDNA and μFFE-SELEX Study on SERCA

4.1 Summary

PLN regulated SERCA plays a crucial role in normal cardiac functions. Heart failures are usually related to an abnormal change in the ratio of PLN to SERCA. Besides gene transfer therapy, aptamers for PLN or SERCA may tune the ratio of PLN to SERCA, and therefore are important in curing heart failure. In this chapter, the discovery of molecules that bind PLN with high pM dissociation constants is described. Factors affecting the binding, in terms of base order, sequence length, and sequence structure are discussed. In addition, attempts to generate aptamers toward SERCA using μFFE-SELEX are also described, with the selected ssDNA pool showing high nM binding affinity.

4.2 Introduction

Cardiac Ca^{2+} cycling has been reviewed in Chapter 1.4.3, along with the critical role of SERCA and the important regulation by the endogenous inhibitor PLN. In human heart failure, the SERCA level is often found to decrease while the level of PLN is unchanged.[266-268] The ratio of PLN to SERCA thus increases, which results in a prolonged cardiac relaxation and a decreased Ca^{2+} affinity by SERCA. At the same time, phosphorylation of PLN at Ser 16 and Thr 17 is decreased, which means more PLN molecules stay in the dephosphorylated state with inhibitory effect to SERCA.[266, 269]

Two strategies have been proposed for curing: one is to increase the SERCA: PLN ratio, and the other is to increase the phosphorylation of PLN. Recombinant adenovirus-mediated gene transfer has been employed to over express

SERCA2a for rats having heart failure.[270] Ca^{2+} uptake and cardiac contractility were observed to increase, and the improved cardiac function allowed rats to survive heart failure.[271, 272] Similarly, PLN antisense RNA and non-functional PLN mutants which mimic phosphorylated PLN were expressed using gene transfer to decrease the overall PLN inhibitory effect.[273-275]

Although these methods show good efficacy and promise in curing heat failure, gene transfer is indirect, inconvenient and expensive. Here in this chapter, I wish to report the discovery of ssDNA sequences' binding to PLN with high affinity. Importantly, the binding is irrelevant to the base order of the ssDNA sequences, but is dependent on length, varying from high nM K_d for 10 nt sequences to high pM K_d for sequences with the length equal to or above 80 nt. SERCA activity assays performed by Dr. Raffaello Verardi indicate that these ssDNA sequences can relieve the inhibitory effect of PLN to SERCA in reconstituted lipids, and the tunable recovery of SERCA activity is also length dependent (unpublished results). Our discovery shows that ssDNA sequences provide important structural information in drug design, and may open up a new path for curing heart failure.

Aptamers for membrane proteins have been generated using the peptide regions, and using membrane preparation or whole cell lines as discussed in Chapter 1.4.2. Both methods suffer from disadvantages associated with the target preparations: aptamers selected against peptides or truncated proteins have weak or no affinity towards the intact proteins,[276] and selections against membrane preparation or whole cell lines are long, tedious,[193] and limited by cell viability[206] and background binding.[190-192] In Chapter 3, I have demonstrated the

success of μFFE as an improved separation platform for SELEX. Similar to CE, μFFE keeps the benefits from free solution separation, such as more binding access to the target and no nonspecific binding to immobilizing linkers or supporting materials. Moreover, this technique offers more advantages than CE. 1.8×10^{14} sequences can be injected on chip within 30 min in the first round of selection, which was a 370-fold improvement.[251] The low electric field in μFFE is more compatible with complex targets consisted of more than one molecule. The fraction collection is based on spatial difference, which is more suitable to collect the low mobility aptamer-target complexes that barely move on capillary. SERCA can be solubilized in aqueous solutions by surfactant $C_{12}E_8$ protecting the hydrophobic transmembrane domain. An aptamer of SERCA would be an ideal drug candidate if it does not decrease SERCA activity but interferes the binding between PLN and SERCA. In addition, aptamers could be ideal molecular probes to quantify or track the membrane protein for various purposes both *in vitro* and *in vivo*. In this chapter, we report a further application of μFFE-SELEX platform, which is to isolate DNA aptamers for SERCA solubilized in $C_{12}E_8$. To our best knowledge, this is the first attempt to generate aptamers for a pure whole integral membrane bound protein using any form of SELEX technique.

4.3 Experimental Methods

4.3.1 Materials and Chemicals

SERCA1a isoform and recombinant PLN solubilized in surfactant $C_{12}E_8$ were gifts from Dr. Verardi and Dr. Veglia from Department of Biochemistry, Molecular

Biology, and Biophysics, at the University of Minnesota. All the ssDNA sequences for PLN binding assay listed in Table 4.1, the initial random ssDNA library 5'-FAM-AGC AGC ACA GAG GTC AGA TG$(N)_{40}$ CCT ATG CGT GCT ACC GTG AA-3' for aptamer selection, forward primer 5'-FAM-AGC AGC ACA GAG GTC AGA TG-3' and reverse primer 5'-Biotin-TTC ACG GTA GCA CGC ATA GG-3' for PCR reactions, and nuclease free H_2O were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). For PCR reactions, *Taq* polymerase and ThermoPol buffer were from New England BioLabs (Ipswich, MA); dNTPs and the 25 bp DNA ladder were from Invitrogen (Carlsbad, CA); gel loading dye was from Promega (Madison, WI). Other chemicals were purchased from Sigma Aldrich (St. Louis, MO) at the highest grade available, except Acetic acid (CH_3COOH, 99.7%, Mallinckrodt Baker), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 99%, Alfa Aesar), $MgCl_2$ (99.8%, Mallinckrodt Baker), KCl (99.5%, Mallinckrodt Baker), $CaCl_2$ (99.9%, Mallinckrodt Baker), NaCl (99.0%, Spectrum), 3-Morpholinopropanesul-fonic acid (MOPS, 99.5%, Fluka Analytical), and DL-Dithiothreitol (DTT, 99.0%, MP Biomedicals).

In ssDNA-PLN binding study, the binding buffer in both ACE and FP experiments contained 20 mM MOPS, 0.1% (w/v) Octaethylene glycol monododecyl ether ($C_{12}E_8$), 0.25 mM DTT, 1mM $MgCl_2$, 1mM KCl, and 5 mM $CaCl_2$ at pH 7.0 (pH adjusted by 1M NaOH). Separation buffer in ACE consisted of 20 mM MOPS and 0.1% (w/v) $C_{12}E_8$ at pH 7.0 (pH adjusted by 1M NaOH). In μFFE-SELEX, 28 μM SERCA aliquots were stored at -80 °C in the storage buffer composed of 20 mM MOPS, 0.1% (w/v) $C_{12}E_8$, 0.25 mM DTT, 1 mM $MgCl_2$, 1

mM $CaCl_2$, 1 mM KCl, and 20% (v/v) Glycerol at pH 7.2 (pH adjusted by 1M NaOH). The binding buffer had the same recipe as the storage buffer, except that no glycerol was included. Separation buffer was made of 25 mM HEPES and 300 μM Triton X-100 at pH 7.2 (pH adjusted by 1M NaOH). All buffers were prepared in nuclease free H_2O, and filtered through 0.2 μm membrane filters before use.

4.3.2 μFFE Fabrication

μFFE devices were fabricated in Nanofabrication Center at the University of Minnesota, following previously reported procedures.[167] The devices had exactly the same design as demonstrated in Chapter 3 by Figure 3.1. Briefly, two photolithography steps were performed to create the multiple-depth design with 100 μm deep electrode channels and a 20 μm deep separation channel on a 1.1 mm borofloat wafer (Precision Glass & Optics, Santa Ana, CA). Then a third photolithography step was performed to pattern the electrodes into the electrode channels. The chip was anodically bonded to a second borofloat wafer that had previously been drilled with inlet and outlet holes and deposited with amorphous silicon (aSi). The sample inlet hole had a diameter of 355 μm and was about 2 cm away from the split point close to the outlets. After Anodic bonding nanoports (Upchurch Scientific, Oak Harbor, WA) were attached to the inlet and outlet holes, and silver conductive epoxy was deposited to create the connection between electrodes and power supply.

4.3.3 μFFE Separation and online LIF Detection

μFFE chip was coated with PEO to suppress EOF before use.[260] Briefly, 1 M HCl, 0.2% PEO in 0.1 M HCl, and separation buffer were pumped into chip for 10

min at 6 mL min^{-1}, 3 mL min^{-1}, and 6 mL min^{-1}, respectively. FAM labeled ssDNA library was heated to 72 °C for 5 min and then naturally cooled down to room temperature. For the 1st round of selection, 125 μM random ssDNA library was incubated with 100 nM SERCA at room temperature for 20 min in binding buffer. SERCA concentration was reduced to 10 nM in the 2nd round, and further reduced to 1 nM for the rest of the selection (the 3rd to 5th round). 40 μL of ssDNA-SERCA mixture was loaded into a 100 μL syringe (Hamilton Company, Reno, NV), and pumped into the μFFE chip at 100~150 nL min^{-1} using a syringe pump (PicoPlus, Harvard Apparatus, Holliston, MA). Separation buffer was pumped into the device at 1 mL min^{-1} using another syringe pump (pump 22, Harvard Apparatus). An electric field of 150 V cm^{-1} was applied across the chip to separate the unbound ssDNA from the ssDNA-SERCA complex.

A solid state laser beam (488 nm, 50 mW, Newport Corp., Irvine, CA) was expanded into a line and focused across the separation channel 1.5 cm downstream from the sample inlet. Fluorescence images were recorded every second using a Cascade 512B CCD camera (Photometrics, Tucson, AZ) through an AZ100 stereomicroscope (Nikon Corp., Tokyo, Japan) with a microscope objective (3×zoom) positioned right above the detection zone. An Endow GFP bandpass emission filter cube (Nikon Corp., Tokyo, Japan) was equipped in the microscope, consisting of a dichroic mirror (495 nm cutoff) and two bandpass filters (450-490 nm and 500-550 nm) for wavelength selection. Images and linescans were processed using MetaVue software (Downington, PA).

4.3.4 PCR Amplification and Purification

SERCA bound sequences were collected from μFFE device and PCR amplified with the following recipe: 1 mM each of the four dNTPs, 80 mM KCl, 7.5 mM $MgCl_2$, 500 nM forward and reverse primers, 1× ThermoPol buffer, and 2.5 units of *Taq* polymerase in a final volume of 80 μL for each reaction vials. 20 reactions were performed for each selection, from which 2 were negative controls and the rest 18 were spiked with collections. 20 cycles were performed with 30 sec denaturation at 94 °C, 30 sec annealing at 55 °C, and 20 sec extension at 72 °C. Agarose gel electrophoresis with ethidium bromide staining and UV-vis imaging were applied to check the quality of the PCR products. Then a previously established column purification procedure was used to purify the products into the desired single stranded FAM labeled DNA sequences,[46] followed by ethanol precipitation, drying by CentriVap Concentrator (Labconco Corp., Kansas City, MO), and solubilization in binding buffer.

4.3.5 Binding Characterization

For PLN-ssDNA K_d measurements, both ACE and FP were employed. Approximately 2.5 nM ssDNA samples were titrated with increasing concentrations of PLN in MOPS binding buffer. All ACE experiments were performed on a commercial CE system (P/ACE MDQ, Beckman Coulther, Inc., Fullerton, CA) equipped with laser induced fluorescence detection (λ_{ex} = 488 nm, λ_{em} = 520 nm). Samples were hydrodynamically injected at 1 psi for 4s, into a 50 cm long fused silica capillary with 50 μm i.d. (Polymicro Technologies, Phoenix, AZ). Separations were then performed at 30 kV in MOPS separation buffer with reversed polarity. This is because EOF was greatly suppressed in the presence

of $C_{12}E_8$ and the unbound ssDNA moved toward the anode. Peak heights of the unbound ssDNA were used to calculate the bound fractions and estimate K_d using the following equation:[262]

$$f_a = \frac{c}{1+\frac{K_d}{([PLN]_t - 0.5([DNA]_t + [PLN]_t + K_d - (([DNA]_t + [PLN]_t + K_d)^2 - 4[DNA]_t[PLN]_t)^{0.5}))}} \quad \textbf{Equation 4.1}$$

in which f_a is the bound fraction, $[PLN]_t$, $[DNA]_t$, and c are total PLN concentration, total DNA concentration, and maximum bound fraction, respectively. $[PLN]_t - 0.5([DNA]_t + [PLN]_t + K_d - (([DNA]_t + [PLN]_t + K_d)^2 - 4[DNA]_t[PLN]_t)^{0.5})$ represents the free unbound PLN concentration.

The exactly same samples used in ACE were loaded into a corning 3540 microplate (Corning Incorporated, Corning, NY) for FP measurements. The experiments were performed on a Synergy™ 2 Microplate Reader (BioTek Instruments, Inc., Winooski, VT). Parallel and perpendicular intensities (λ_{ex} = 485 ± 20 nm, λ_{em} = 520 ± 20 nm) were measured, and polarization values were calculated using Gen 5™ software (BioTek Instruments, Inc., Winooski, VT). The bound fractions were determined as described in Chapter 1.3.3 and corrections to the bound fractions were made as previously reported when intensity changes were observed.[163]

The binding between SERCA and ssDNA was characterized by both μFFE and FP. For separations performed on μFFE, mixtures composed of ~50 nM ssDNA and 14 μM SERCA were introduced into chip. Electric fields changing from 0 V cm^{-1} to 150 V cm^{-1} were applied. For FP experiments, mixtures composed of ~5 nM ssDNA and varying concentrations of SERCA were preparation. The

measurement and data analysis were performed the same way as for PLN-ssDNA binding study.

4.4 Results and Discussion

4.4.1 ssDNA sequences binding to PLN

I first describe the observation of the strong interaction between PLN and ssDNA sequences. SERCA, as well as PLN, was incubated with the random ssDNA library to test the background binding before the μFFE selection. Specifically, 80 nt random ssDNA library composed of two 20 nt primer regions and one 40 nt random region was mixed with different amounts of PLN or SERCA in MOPS binding buffer for FP measurements. The strong binding between PLN and random ssDNA was identified with the K_d around 1 nM. On the other hand, SERCA showed no obvious influence on the polarization of the ssDNA in the same buffer regardless of mid-μM SERCA being used. This result is rather surprising, and is contrary to a previous report on the no binding affinity of the random library to the cytoplasmic region of PLN, tested before the selection of DNA aptamers.[277] Furthermore, a series of ssDNA sequences were chosen with different lengths, degrees of randomness, and stereochemistry to test the binding with PLN (see sequence information in Table 4.1).

Table 4.1 Random ssDNA sequences used in PLN binding assay

Sequence Length (nt)	Sequence 5'→3'
10	FAM-ATA GCT TGC A
15	FAM-AGT GAT AGC TAT GGT

20	FAM-AGC AGC ACA GAG GTC AGA TG
30	FAM-ACT GAG CAT GGG ATA ACC GTT CTC AGA CTT
50	FAM-AGC AGC ACA GAG GTC AGA TGC AGG TAG GGT CCT ATG CGT GCT ACC GTG AA
80*	FAM-$(N)_{80}$

* A mixture of completely randomized 80-mer ssDNA sequences

The binding measurements were performed in both ACE and FP, two orthogonal methods in principle. All of the sequences listed in Table 4.1 binded to PLN, with the histogram shown in Figure 4.1. The K_d values followed a decreasing trend from sub-μM for the 10 nt sequence, to mid-nM for the 15 nt sequence, and low nM or even high pM for 30, 50, and 80 nt sequences, suggesting the significant dependence of binding affinity on sequence length. The binding curves obtained from both FP and ACE for the 80 nt sequence are shown in Figure 4.2.

The K_d values obtained by ACE and FP were consistent within experimental error for sequences longer than 20 nt; however, the results had some discrepancies for 10 and 15 nt sequences. For example, the K_d values obtained for the 10 nt sequence were 36±5 nM by FP and 270±180 nM by ACE. The discrepancy is possibly due to the different principles of the two methods. Polarizations were allowed to occur in binding equilibrium, but complexes underwent a non-equilibrium process in high electric field, both of which may cause dissociation. This discrepancy is more obvious for weaker binders that have a higher k_{off} rate.

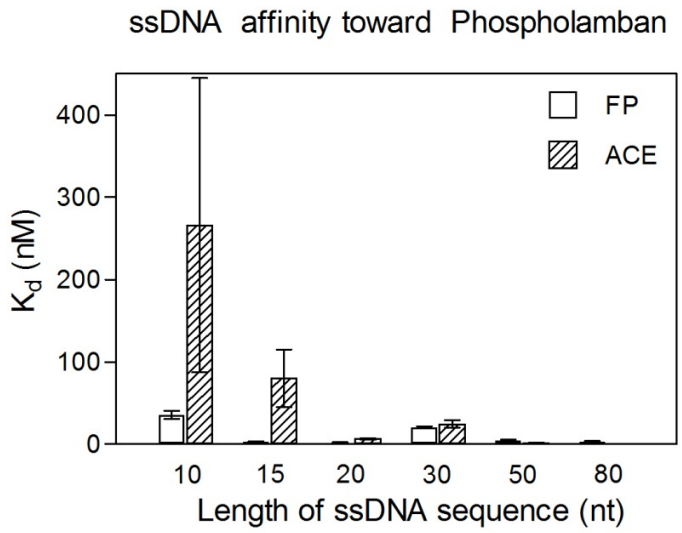

Figure 4.1 Affinity histogram of ssDNA sequences binding to PLN. Error bars represent standard deviation.

Interestingly, we observed an exception of 30 nt in the general trend that the K_d values decreased as the length of ssDNA increases. The K_d values of the 30 nt was 25 ± 5 nM by ACE, and they were significantly higher than K_d values of the 20 nt and 50 nt, which were 6.5 ± 0.9 nM and 1.2 ± 0.4 nM, respectively. The 10 nt to 30 nt sequences as shown in Table 4.1 were all randomly generated sequences, but we found that the 30 nt sequence can fold into a special secondary structure with free energy of -5.41 kcal mol^{-1} under the experimental condition, as predicted by mfold (Figure 4.3 (A)). This is a stem-bulge-stem structure with 20 nucleotides serving in the double stranded region. The special structure makes the sequence more rigid than the 10, 15, and 20 sequences, which resembles dsDNA sequences. Sequences shorter than 30 nt (Table 4.1) were predicted to have several unfavored structures with positive free energy values. Although the 50 nt sequence also folds into a secondary structure as

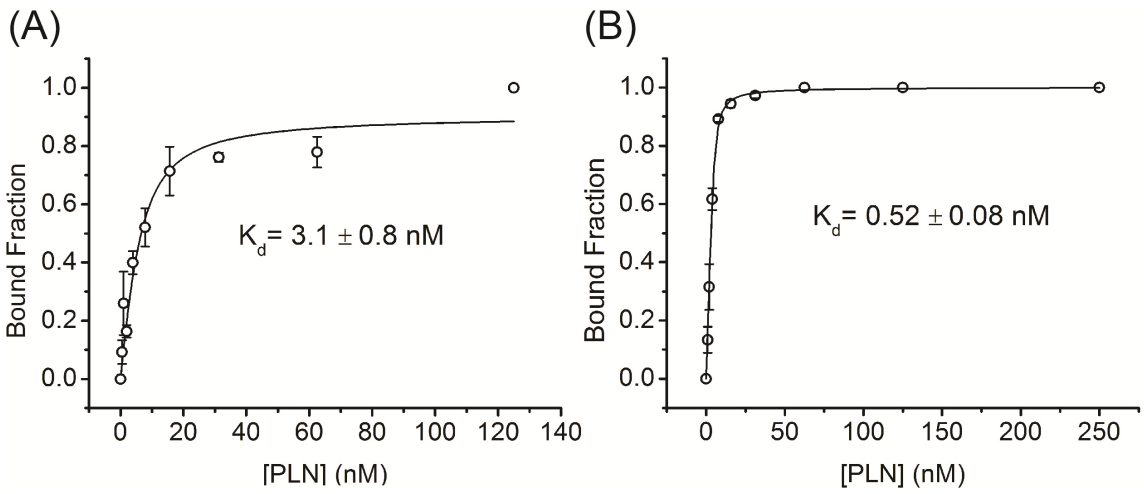

Figure 4.2 Binding curves obtained by titrating the completely randomized 80 nt ssDNA sequences in Table 4.1 by FP (A) and ACE (B). Three measurements were taken under each PLN concentration. Error bars on data points and errors for K_d values represent standard deviation.

shown in Figure 4.3 (B), the structure is less stable than the 30 nt sequence, with ΔG of -3.72 kcal mol^{-1}. There are fewer nucleotides involved in the double stranded region as well, making this sequence more flexible than the 30 nt sequence.

A completely randomized 80 nt sequence-mixture was commercially synthesized, and the K_d value was determined to be 0.52±0.08 nM by ACE. It is worthy pointing out that other sequences were either randomly generated (10 nt to 30 nt in Table 4.1) or used because it was within reach (50 nt in Tabe 4.1), and none of them was selected based on the binding affinity towards PLN. The results lead us to conclude that the binding between PLN and ssDNA sequences are nonspecific, regardless of the base order of the ssDNA sequences.

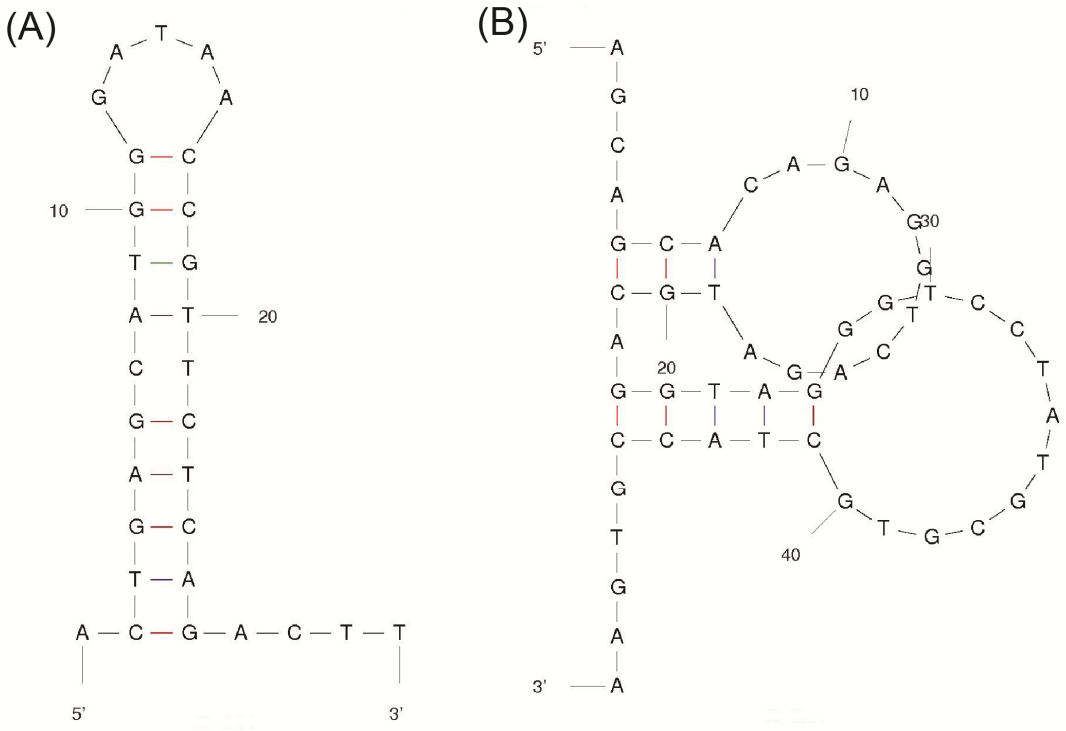

Figure 4.3 Possible secondary structures of the 30 nt sequence (A) and 50 nt sequence (B) in Table 4.1, predicted by mfold, http://mfold.rna.albany.edu/?q=mfold/DNA-Folding-Form.

4.4.2 Isolation of ssDNA aptamers that bind to SERCA1a

My first attempts of generating ssDNA aptamers against SERCA using CE were unsuccessful. The surfactant $C_{12}E_8$ present in the separation buffer hydrodynamically coats the inner surface of the bare fused-silica capillary. As a result, ssDNA peak was only observed when separation polarity was reversed, indicating that the ssDNA mobility became larger than EOF. $C_{12}E_8$ (MW: 539 g/mol, number of aggregation ≈ 95 at 25 °C)[278] wrapped SERCA (MW: 110,000 g/mol)[279] is significantly bigger than ssDNA molecules (MW: 25,000 g/mol). The small negative mobility of SERCA-ssDNA complex counterbalanced the small positive EOF, giving a net effect of barely moving complex in the electrical field.

Additional hydrodynamic rinses were necessary in order to collect the complex, which further diluted the complex and increased the chance of contamination. Low protein concentration in the mixture was employed in order to keep high selection stringency, making the online detection of complex impossible. The lack of effective online monitoring made the already complicated selection even more difficult.

Inspired by the success of generating IgE aptamers using μFFE-SELEX as described in Chapter 3, selections against SERCA were also moved to chip. Five rounds of selections were performed with the initial introduction of 2.25×10^{14} ssDNA sequences in the first round. Although complex was not identified in the first several rounds of selection, the unbound ssDNA was readily identified and discarded. In the 5^{th} round of selection, an analyte stream with a minimal mobility was identified (shown as (2) in Figure 4.4) beside the unbound ssDNA (shown as (1) in Figure 4.4). As expected, the separation pattern was very similar to the one for free and bound ssDNA to IgE as shown in Figure 3.2 (A), indicating that stream 2 was highly possible to be the aptamer-SERCA complex.

Stream 2 in Figure 4.4 was collected, PCR amplified, and purified into the desired ssDNA product. The affinity of the product from cycle 5 was first evaluated by FP. As shown in Figure 4.5 (A) as blue diamond, the FP value of the ssDNA increased as the concentration of SERCA in the mixture increased, indicating the formation of aptamer-SERCA complex. Figure 4.5 (B) shows the binding curve, from which a high nM K_d was obtained by nonlinear regression. As

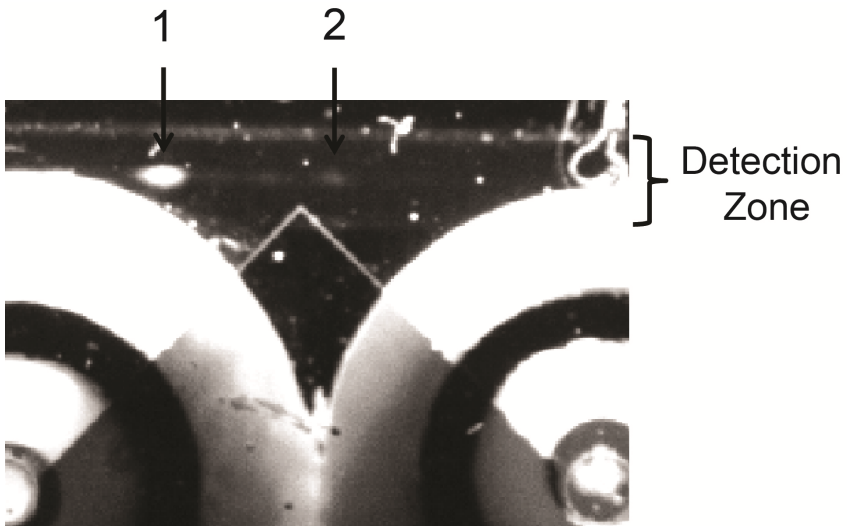

Figure 4.4 An image of μFFE separation of unbound (1) and bound (2) ssDNA sequences during selection cycle 5. The separation chamber becomes two collection channels below the vertex of the triangle design. The bound sequences were collected into a centrifuge tube connected to the right outlet port through tubing.

a comparison, ssDNA samples from other selection rounds (1 to 4) were also titrated against SERCA. However, the FP values (indicated by red circles, orange squares, yellow up triangles, and green down triangle in Figure 4.5 (A), respectively), did not change significantly when the SERCA concentration increased. This result agrees well with the on-chip observation, from which the aptamer-SERCA complex was only visually observed in cycle 5.

50 nM FAM labeled ssDNA from cycle 5 was then incubated with the non-labeled SERCA at the saturating concentration (14 μM), and separations at different electric field strength were performed on the μFFE device. As shown in Figure 4.6 (A), the fluorescent stream shifted towards the cathode instead of anode when the electric field was changed from 0 to 30 V cm^{-1}. The shift increased as the electric field was further increased to 40 and 50 V cm^{-1}, corroborating the band to be the aptamer-SERCA complex. Another interesting

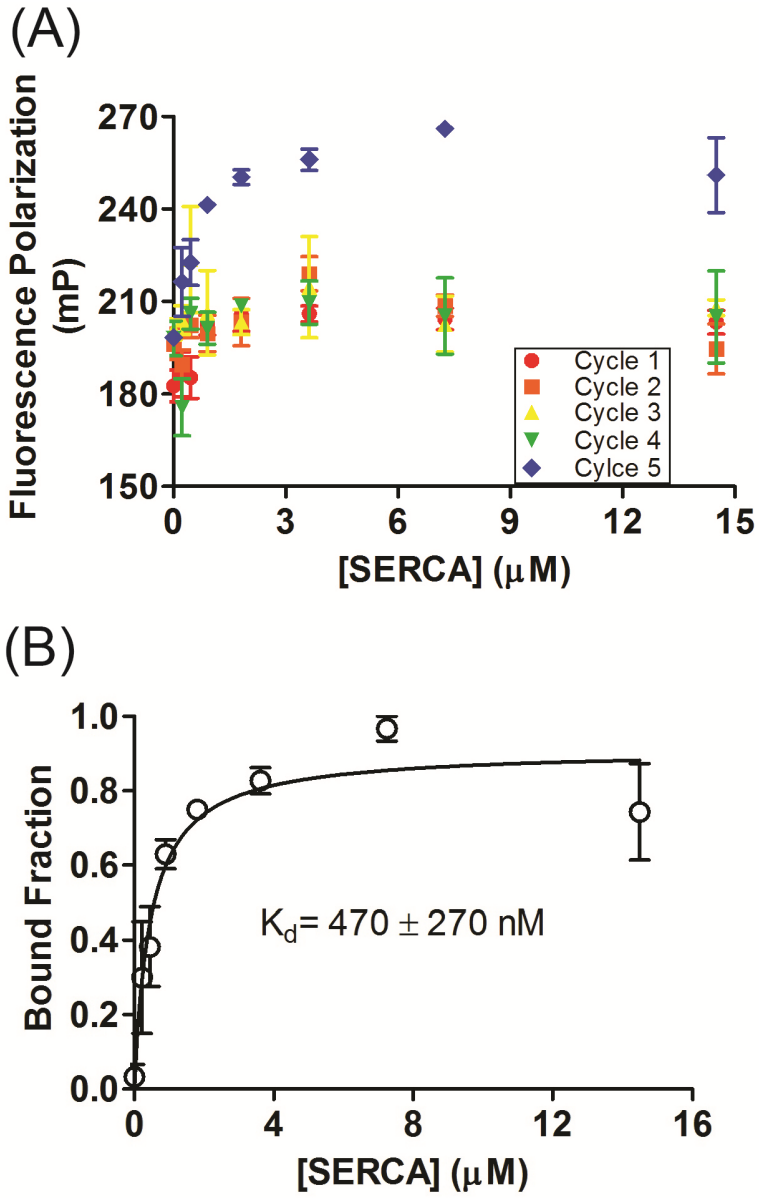

Figure 4.5 (A) FP values for mixtures composed of different amounts of SERCA and ~5 nM ssDNA samples from cycle 1 (●), cycle 2 (■), cycle 3 (▲), cycle 4 (▼), and cycle 5 (◆). SERCA concentrations were 0 M, 2.3×10^{-7} M, 4.5×10^{-7} M, 9.1×10^{-7} M, 1.8×10^{-6} M, 3.6×10^{-6} M, 7.3×10^{-6} M, and 1.4×10^{-5} M. (B) Binding curve for cycle 5. Error bars represent standard deviation of three measurements. The Error of K_d represents 95% confidence interval.

phenomenon was that the band broadened as the electric field increased. It was difficult to identify the band when electric field was equal to or above 50 V cm^{-1},

when the band became a smear. Moreover, the center of the smear shifted back toward anode at 80 V cm^{-1}, indicating dissociation of the complex. As a control, the unselected ssDNA library was also incubated with 14 μM SERCA and separations were also performed on μFFE chip. As shown in Figure 4.6 (B), the fluorescent stream shifted towards the anode as the electric field increased, and there was no band broadening associated with dissociation at higher electric field, thus indicating the band as the free unbound ssDNA. The on chip experiments further confirmed the binding between the product of cycle 5 and SERCA, and **the** estimated K_d value from FP experiments since almost all ssDNA molecules were in complex form, as evidenced by the shift of one band from 0 to 30 V cm^{-1} rather than two discrete bands.

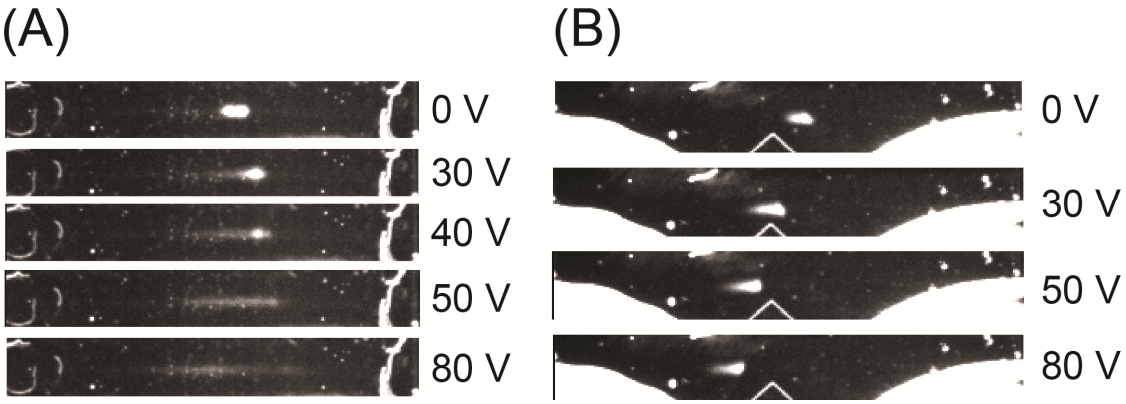

Figure 4.6 (A) μFFE analysis on the mixture composed of 50 nM ssDNA from cycle 5 and 14 μM SERCA. (B) μFFE analysis on the mixture composed of 50 nM ssDNA from unselected random library and 14 μM SERCA as a control. The anode is on the left side in both images.

Further attempts were made to perform the μFFE selection with a reduced electric field at 80 V cm^{-1} rather than the original 150 V cm^{-1}, in order to maintain and collect complex; however, no binding was observed for the eight selected

pools (data not shown). Further selection beyond eight cycles was unpractical, limited by the quantity of DNA product. This might indicate that the reduced electric field does not provide sufficient resolution to enrich aptamers within eight selection cycles. Another possibility is that the SERCA sample prepared in $C_{12}E_8$ may lose the desired structure within several days as we later discovered in developing a new SERCA activity assay.

4.5 Conclusions

In conclusion, this chapter demonstrated the high affinity binding between ssDNA molecules and PLN. The binding affinity is length dependent, and high pM K_d is obtained once the ssDNA length is equal to 80 nt. The binding affinity is also related to the flexibility of the sequence, but does not depend on the base order. This discovery offers rich information on anti-PLN drug design and may pave a new way for curing heart failure caused by PLN-related SERCA dysfunction. This chapter also demonstrated some promising results of μFFE device in evolving ssDNA pools against $C_{12}E_8$ stabilized SERCA. After five cycles of selection, a DNA pool was generated with high nM binding affinity toward SERCA. μFFE separation is superior to CE in terms of sample input, online detection, convenient complex collection, and maintaining of complex due to lower electric field and faster transit time. Although no individual aptamers were finally obtained, this work still showed the compatibility and unique advantages of the μFFE device to complex target SELEX.

Chapter 5 Development of a Highly Sensitive SERCA Activity Assay Based on the Direct Detection of ADP Product

5.1 Summary

SERCA is an important target for treating cardiac diseases due to its function in controlling cytoplasmic calcium concentrations during heart relaxation. The study of SERCA activity in the absence and presence of its endogenous inhibitor PLN is of great interest in understanding the cardiac function under both physiological and pathological conditions. Traditionally, SERCA activity assay was performed using a PK/LDH-coupled enzyme reaction, which suffers low sensitivity. We developed a new SERCA activity assay based on the direct detection of the product ADP via TR-FRET. Under the optimized condition, our assay reduced the amount of SERCA by 1,000 folds and showed good inter-day reproducibility in both $C_{12}E_8$ and reconstituted lipids. The inhibitory effect of PLN on SERCA under low-concentration conditions in our assay allowed us to re-evaluate the binding between PLN and SERCA, which should have a mid-nM K_d at the most.

5.2 Introduction

As discussed in Chapter 1.4.3, SERCA plays an important role in controlling cardiac muscle relaxation and therefore in maintaining normal cardiac function. Quantifications of SERCA activity and the regulatory effect by PLN are fundamental requirements in SERCA studies. In terms of enzyme activity assay, various types of detection system have been designed and applied for kinases, which compose around 2% of the human genome.[280] The kinases control multiple cellular functions, and have been targeted pharmaceutically for drug discovery and treatment of a wide range of diseases.[281-283] Although SERCA as

an ATPase does not produce phosphorylated substrate,[284-287] kinase assays monitoring the consumption of ATP or production of ADP would be referential for developing SERCA activity assays.

One method is to quantify ATP consumption by a chemiluminescent assay, in which luciferin and luciferase are added after a predetermined reaction time to consume all the rest of ATP and generate light signal.[288, 289] A major drawback of this method is that at least 20% to 30% ATP consumption is needed in order to have a desired detection window.[290] However, from Michaelis-Menten equation V_{max} is only obtained when the concentration of ATP is much higher than K_m, and therefore ATP consumption is very slow and difficult to be determined accurately by this method.[291]

Alternatively, ADP generated in assay was targeted to obtain the enzyme activity. Traditionally, a coupled enzyme assay involving a pyruvate kinase (PK) and lactate dehydrogenase (LDH) was used to detect ADP in a continuous mode.[292-297] Phosphoenolpyruvic acid (PEP) is first converted to pyruvate, which is then hydrogenated to form lactate, and at the same time NADH is dehydrogenated to NAD^+. The decrease of NADH at 340 nm is monitored to calculate the enzyme reaction rate. Although widely applied in kinase and ATPase assays, this method has several disadvantages. For example, the detection is based on UV which has relatively low sensitivity; the method requires large amounts of both enzyme and ATP; and sometimes it suffers from the emission overlap between NADH and other reagents or NADH decomposition.[298] Other assays have been developed to overcome the drawbacks. For example,

an improved coupled enzyme reaction used PK, pyruvate oxidase, and horseradish peroxidase to detect ADP via the fluorescent signal, which avoided UV detection interference.[290] Besides, a dithio-coupled kinase or ATPase assay was developed with ATPβS (sulfur on the β-phosphorous) substrate replacing ATP. Dithiol reagents such as DTNB were used to react with the ADPβS product, and absorbance or fluorescence was detected in the continuous mode.[299] The assay sensitivity was improved by using fluorescence or reagent with higher extinction coefficient; however, the labeled substrate ATPβS is not commercially available. Moreover, all these assays targeting produced ADP employed multiple reagents and steps of reactions. These indirect detection methods hold the possibility of nonspecific interactions between inhibitor/screening drug and detection mixtures. In addition, assay validation is necessary to confirm that the coupled enzyme reactions or dithio labeling reactions are not rate-limiting, in order to report enzyme reaction rate accurately in a continuous mode.

In this chapter, we report the development of a highly sensitive SERCA activity assay using TR-FRET detection. This immunoassay is based on the Transcreener® platform, whose key component, an ADP specific antibody, directly binds SERCA assay product, eliminating coupled enzyme reactions or extra labeling reactions.[300] The far red TR-FRET detection avoids interfering fluorescence or light scattering from the reaction mixture. The homogenous assay was performed in 384-well plates with only 20 μL total volume in each reaction, and the amount of SERCA was reduced by three orders of magnitude. Under the optimized conditions, SERCA activity performed in both surfactant

$C_{12}E_8$ and reconstituted lipids (DOPC: DOPE, 4:1) showed good reproducibility. With increased sensitivity, our assay also sheds light on quantifying the binding between two membrane proteins, SERCA and PLN.

5.3 Experimental Methods

5.3.1 Materials and Chemicals

All chemicals were purchased from Sigma Aldrich (St. Louis, MO) at the highest grade available, except 3-Morpholinopropanesul-fonic acid (MOPS, 99.5%, Fluka Analytical), $CaCl_2$ (99.9%, Mallinckrodt Baker), $MgCl_2$ (99.8%, Mallinckrodt Baker), KCl (99.5%, Mallinckrodt Baker), and NaCl (99.0%, Spectrum). All buffers were prepared in deionized water from a Milli-Q water purification system (Millipore Corp., Bedford, MA). Transcreener® ADP^2 Assay TR-FRET Red kit was purchased from BellBrook Labs (Madison, WI). SERCA1a extracted and purified from rabbit skeletal muscle was co-reconstituted in lipid membranes (DOPC: DOPE, 4:1) by Raffaello Verardi and given to me as a gift. This sample consisted of 8 μg of SERCA, 22 μg of PLN, 0.2 mg of lipids, 35 mM of Imidozole, 100 mM of KCl, 5 mM of $MgCl_2$, and 10% (v/v) of glycerol at pH 7.0 in a final volume of 200 μL, which gave a SERCA concentration of 364 nM and a PLN concentration of 18.3 μM (calculated based on the molecular weight of PLN monomer). Another two gifts were reconstituted 80 μg of SERCA, and 80 μg of SERCA and 22 μg of PLN, respectively, following the same preparation procedure. 2× Assay Buffer 1 composed of 100 mM MOPS, 200 mM KCl, 10 mM $MgCl_2$, and 2 mM EGTA at pH 7.0 (pH adjusted by 1 M NaOH). Assay Buffer 2

composed of 20 mM MOPS, 1 mM $MgCl_2$, 0.25 mM DTT, 0.1% (w/v) C12E8, and 40% (v/v) glycerol at pH 7.2 (pH adjusted by 1 M NaOH). Assay Buffer 3 composed of 1 μg/μL of lipids (DOPC: DOPE, 4:1), 35 mM of Imidozole, 100 mM of KCl, 5 mM of $MgCl_2$, and 10% (v/v) of glycerol at pH 7.0.

5.3.2 Instrument Settings

TR-FRET measurements were performed on a Synergy™ 2 Microplate Reader (BioTek Instruments, Inc., Winooski, VT) with filter settings of λ_{ex} = 360±20 nm, λ_{em1} = 620±5 nm, and λ_{em2} = 665±4 nm. A 400 nm dichroic mirror was used with the excitation range of 320-390 nm and the emission range of 410-800 nm. Detection sensitivity was set to be 135. A 20 sec shaking was performed before taking any readings to eliminate possible bubbles in sample wells. Xenon flash with high lamp energy was used as the light source. There was a 100 ms delay after plate movement and before taking measurements. 100 μs after the xenon flash the instrument started data collection for 200 μs; and each measured FRET value was an average of 40 data collections.

5.3.3 Optimizing Assay Detection Mixture

Since SERCA activity assays will be performed in our own reaction buffers containing surfactants or lipids, assay detection components including ATP and ADP tracer were first titrated to find the best working concentrations. For tracer titration, 20 μL mixtures composed of 4 nM antibody, 2 μM ADP/ATP, 10 mM EDTA, and a series of ADP tracer ranging from 0.001 nM to 1 μM were incubated in Assay Buffer 2 for 1 h before being loaded into a Corning® 384 plate (catalog

#3673) for TR-FRET measurements. The same experiments were also performed at 10 μM, and 20 μM ADP/ATP concentrations. After the best ADP tracer concentration was determined, ATP titration was performed. 20 μL mixtures consisted of 4 nM antibody, optimized ADP tracer, and a series of ATP and ADP ranging from 0.01 μM to 1 mM.

5.3.4 ATP to ADP Conversion Calibration Curve

ATP to ADP conversion calibration curve was taken whenever SERCA activity assay was performed, in order to translate the FRET signal into the amount of ATP consumed or ADP produced in reaction. 10 μL of ATP prepared in the desired assay buffer at twice the assay concentration was aliquoted into 20 vials. 20 μL of ADP at the same concentration was mixed with the 1st vial of ATP until even; 20 μL from the 1st vial was then transferred into the 2nd vial of ATP and mixed until even. The same procedure was performed till the last one. 10 μL of ATP-ADP mixtures were then incubated with 10 μL of the assay detection mixture at twice the optimized concentration. The final 20 μL mixtures were incubated at room temperature for 1 h before taking FRET measurements. Calibration curve was obtained based on three parallel experiments.

5.3.5 SERCA Titration

For SERCA titration performed in $C_{12}E_8$, serial dilutions of SERCA were prepared in Assay Buffer 2 with the 39 μM stock SERCA in 20 mM MOPS, 1 mM $MgCl_2$, 0.25 mM DTT, 0.1% (w/v) C12E8, 30% (v/v) glycerol, and 200 μM $CaCl_2$. 5 μL aliquots of 8 μM ATP and 10 μM $CaCl_2$ prepared in Assay Buffer 2 were added into the same volume of SERCA serial dilutions to start the reactions at

37 °C in a C24 Incubator Shaker (New Brunswick Scientific, Enfield, CT). The reactions were quenched 5 min and 18 min later by adding 10 µL aliquots of the assay detection mixture containing excess EDTA to reach a final volume of 20 µL. For SERCA activity assays performed in reconstituted lipids, serial dilutions of SERCA were prepared in Assay Buffer 3. 5 µL aliquots of 8 µM ATP, 2.5 mM $CaCl_2$, and 7 µg/mL Ca^{2+} ionophore A23187 prepared in 2×Assay Buffer 1 were added to the same volume of SERCA serial dilutions to initiate the reactions. Reactions were quenched by adding 10 µL aliquots of the assay detection mixture at twice the optimized concentration.

5.3.6 ATPase Activity Measurements

For activity assays performed in $C_{12}E_8$, 5 µL aliquots of 8 µM ATP and varying $CaCl_2$ concentrations of were prepared in Assay Buffer 2. Ca^{2+} concentration was controlled by direct addition of $CaCl_2$ stock solutions. SERCA at twice the desired working concentration was also prepared in Assay Buffer 2. Reactions were started by adding 5 µL of SERCA into ATP-$CaCl_2$ mixture, and quenched by adding 10 µL of assay detection mixture containing excess EDTA. Reactions were stopped at least at five different reaction times, and three parallel experiments were performed for each reaction time. Initial reaction rates at different Ca^{2+} concentrations were determined from the linear region of the FRET-reaction time plots. Activity assays were performed similarly for SERCA in reconstituted lipids, except that 8 µM ATP, 7 µg/mL Ca^{2+} ionophore A23187, and varying concentrations of $CaCl_2$ were prepared in 2× Assay Buffer 1; SERCA at twice the desired working concentration was prepared in Assay Buffer 3. All

reactions took place at 37 °C. Initial reaction rates at different Ca^{2+} concentrations were then used for Hill equation fitting to obtain pKca values:

$$V = \frac{V_{max}}{1 + 10^{-n(pK_{Ca} - pCa)}} \qquad \textbf{Equation 5.1}$$

where V is the initial reaction rate, V_{max} is the maximum initial reaction rate, n is the Hill coefficient, and pK_{Ca} is the pC_a value when the initial reaction rate is at half V_{max}.

5.4 Results and Discussion

5.4.1 Assay Principle

The detection principle of the assay is shown in Figure 5.1. The Terbium labeled ADP antibody selectively binds HiLyte647 labeled ADP tracer in the presence of ATP, with an approximate 100-fold selectivity of ADP over ATP.[301] The binding between ADP antibody and ADP tracer brings the Terbium dye and HiLyte647 dye in close proximity, allowing Förster Resonance Energy Transfer (FRET). Emissions were detected from both 665 nm and 620 nm channels. When SERCA reacts with ATP, the produced ADP molecules compete with HiLyte647 labeled ADP tracer molecules on binding sites of ADP antibody molecules. The more ADP produced, the more ADP tracer molecules are replaced, and the smaller the FRET signal would be since the emission from 665 nm channel decreases. The detection was time resolved, taking for 200 μs at 100 μs after the xenon flash, thus eliminating otherwise possible interference of

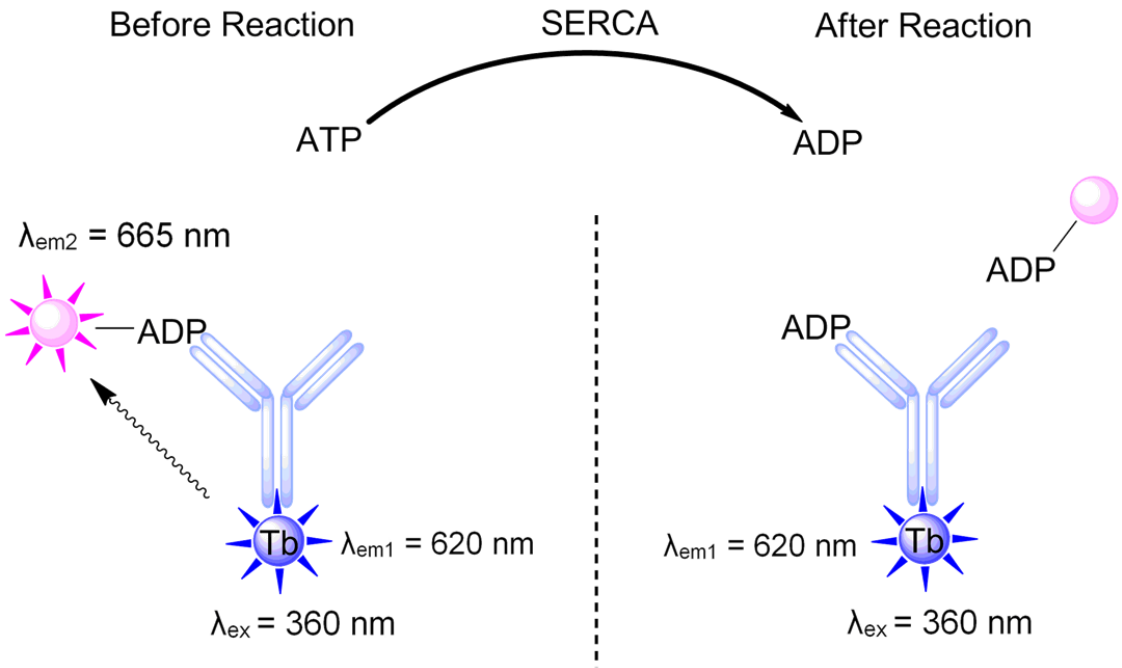

Figure 5.1 Schematic of the novel SERCA activity assay based on TR-FRET detection of ADP.

transient fluorescence from other molecules. Detecting emissions at wavelengths longer than 600 nm also eliminates light scattering.

5.4.2 Detection Window Optimization

In order to obtain SERCA activity curve, it is necessary to perform SERCA reactions at different Ca^{2+} concentrations, quench the reactions at different reaction times, quantify the ADP produced, and finally obtain the initial reaction rates by linear regression. To obtain initial reaction rates, reactions are quenched when substantial amounts of ATP are still remained in the reaction mixture. Accurate ADP quantification in the presence of excess ATP requires the detection window to be optimized to allow significant FRET change to occur when ATP to ADP conversion is low.

Initially, tracer concentrations were determined following a linear relationship between ATP and tracer provided by TRANSCREENER® ADP^2 TR-FRET Red Assay Technical Manual; however, the Z'-factors (Equation 5.2) were below 0.7 with six replicates of ATP and ADP samples. In order to obtain the maximum assay window and the best detection quality, tracer titration was performed in our own working buffer.

$$Z' = 1 - \frac{3 \times SD_{ATP} + 3 \times SD_{ADP}}{\overline{FRET_{ATP}} - \overline{FRET_{ADP}}} \qquad \textbf{Equation 5.2}$$

Figure 5.2 shows the tracer titration performed in Assay Buffer 2 with nonionic surfactant $C_{12}E_8$. The red circles in the left panel indicate the FRET signals when 2 μM, 10 μM, and 20 μM ATP were incubated with different amounts of tracer, and the black squares indicate FRET signals when the same amounts of ADP were incubated with tracer. Since the antibody has much higher binding affinity to ADP over ATP, more ADP tracer molecules were replaced by ADP rather than ATP when tracer concentrations are within a certain range (e.g., 1 nM to 1 μM in Figure 5.2 (A)). When tracer was below this range, both ATP and ADP were in excess to compete for binding sites, so that the FRET signals fell to baseline; when above this range, tracer was in excess and neither ATP nor ADP will trigger a significant decrease of FRET signal. By subtracting the ADP FRET signals from the ATP FRET signals for all tracer concentrations, the FRET signal differences were obtained. From the right panel, the maximum FRET signal differences occurred at ~63 nM, 250 nM, and 500 nM for 2 μM, 10 μM, and 20

μM adenosine, respectively. By using these tracer concentrations, the Z' factors were always above 0.7 from day-to-day experiments.

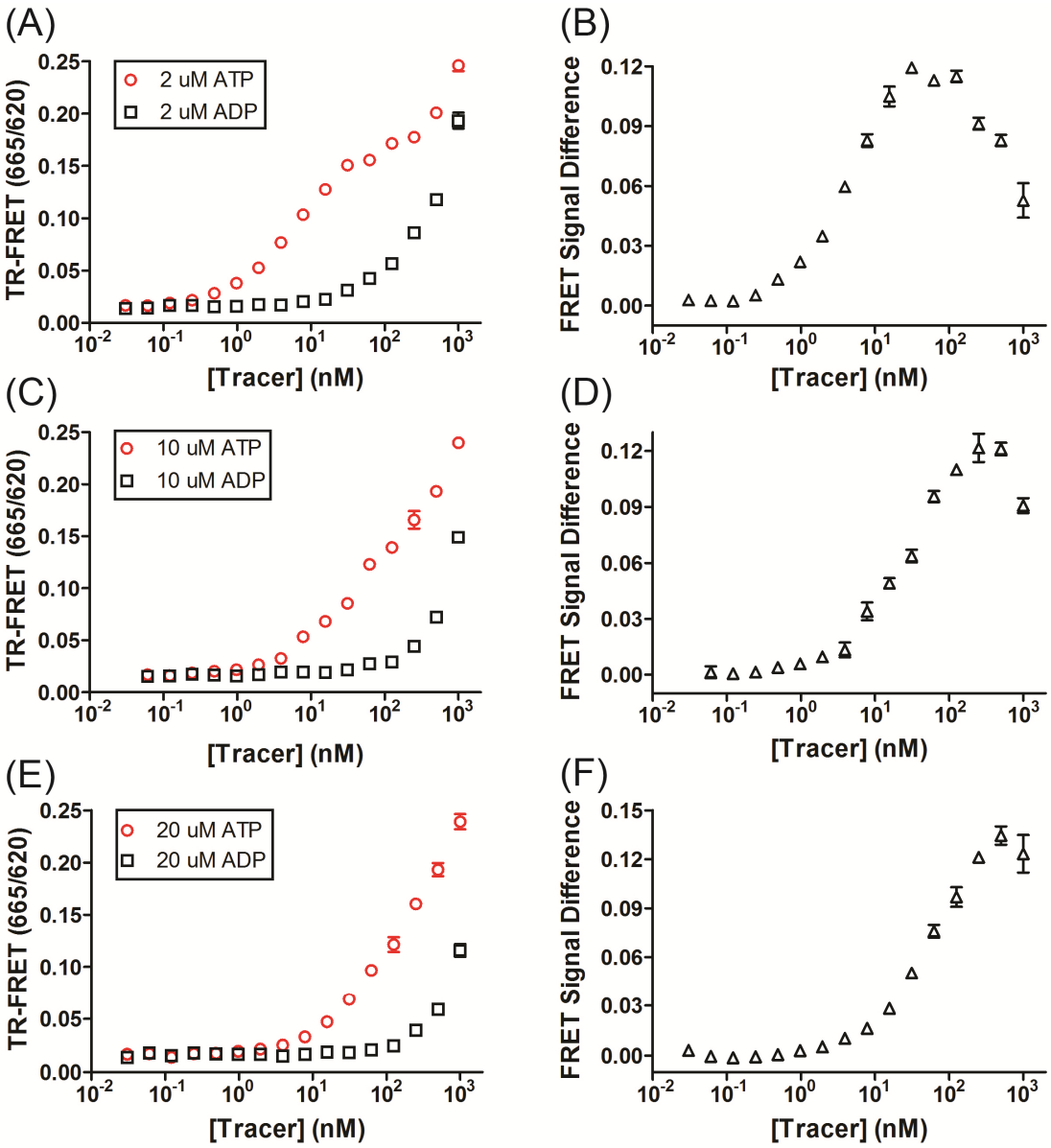

Figure 5.2 Tracer titration in Assay Buffer 2 with different concentrations of ATP and ADP: (A) 2 μM, (C) 10 μM, (E) 20 μM, using 4 nM ADP antibody. FRET signal differences between ATP and ADP at different tracer concentrations were shown in the right panel, with (B), (D), and (F) corresponding to (A), (C), and (E), respectively. Each point is the average FRET value of three parallel samples. Error bars represent standard deviation.

The right panel in Figure 5.2 presents the dependence of the difference between ATP and ADP on optimum tracer concentrations, from which the most sensitive region to distinguish ATP from ADP can be determined. Meanwhile, the maximum FRET differences also represent the maximum assay detection window, since ATP represents reaction at the beginning and ADP represents reaction at the point all ATP is consumed. From the left panel of Figure 5.2, the maximum assay window was approximately 0.140, 0.149, and 0.178 for assays starting with 4 μM, 20 μM, and 40 μM of ATP (the addition of detection mixture dilutes the ATP concentration to half). Although it was roughly a 20% decrease in assay window from 20 μM ATP (500 nM tracer) to 2 μM ATP (63 nM tracer), the latter reduced the use of tracer by 8 folds, and the Z' value was still above 0.7. Most importantly, less SERCA was required to fulfill ATP to ADP conversion within the same time, corresponding to improved assay sensitivity. Furthermore, titrations of ATP and ADP were performed with 63 nM tracer and 4 nM antibody, as shown in Figure 5.3 (A). The maximum FRET signal difference occurred at ~2 μM adenosine (Figure 5.3 (B)), showing good self-consistency of the detection mixture.

5.4.3 Standard Calibration Curve

After the assay detection mixture was optimized, standard calibration curve was obtained for SERCA reactions starting with ATP concentration of 2 μM. Figure 5.4 (A) shows the FRET values of ATP and ADP mixtures, which had a

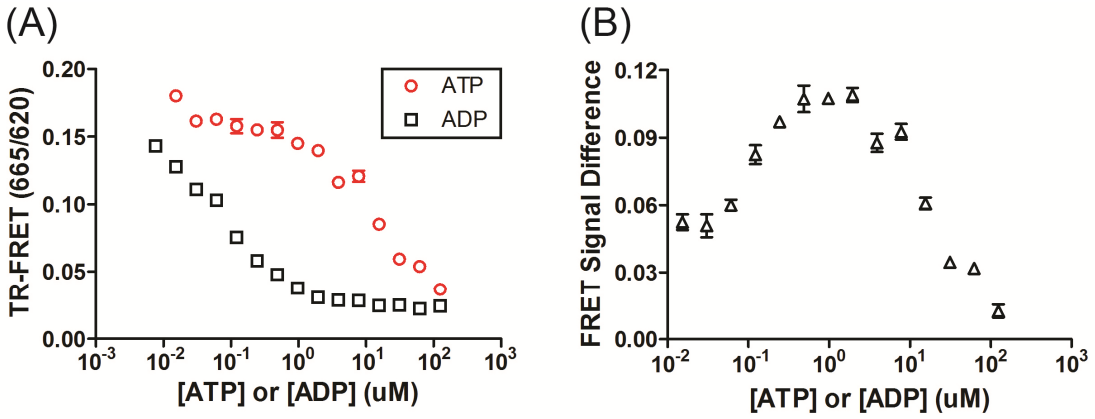

Figure 5.3 (A) ATP and ADP titrations in Assay Buffer 2 with 63 nM tracer and 4 nM ADP antibody. (B) FRET signal differences between ATP and ADP under the same adenosine concentrations. Each point is the average FRET value of three parallel samples. Error bars represent standard deviation.

total adenosine concentration of 2 μM but different compositions of ATP and ADP, simulating SERCA reactions taking place for different reaction times. As shown in Figure 5.4 (A), although no significant signal change occurred when ATP to ADP conversion was within 1%, FRET signal was linearly correlated to the conversion in the range of 1% to 100% in this semilog plot. The linear portion gave a regression equation of Y= - (0.051±0.001) × lgX + (0.140±0.001), with R^2 of 0.9861. It is obvious from this calibration curve that 50% of the maximum FRET change occurs when only 10% ATP is converted to ADP, demonstrating the excellent selectivity of the ADP antibody and the compatibility of the detection mixture in determining initial reaction rates of SERCA. Figure 5.4 (B) shows the calibration curves under different ATP and Ca^{2+} concentrations with the optimized detection mixture that contained 4 nM antibody and 63 nM tracer. When the ATP concentration was above 4 μM or below 1 μM, the maximum assay window was not optimum; in the range of 4 μM to 1 μM, the curves had the

steepest slope and the assay window was optimized around 0.1 units of FRET. The three curves obtained at 4 μM ATP and different concentrations of Ca^{2+} overlapped with each other, demonstrating that the detection mixture is not sensitive to Ca^{2+} change, which would be the only variable besides ATP to ADP ratio in SERCA activity assay. These results lead us to conclude that the FRET

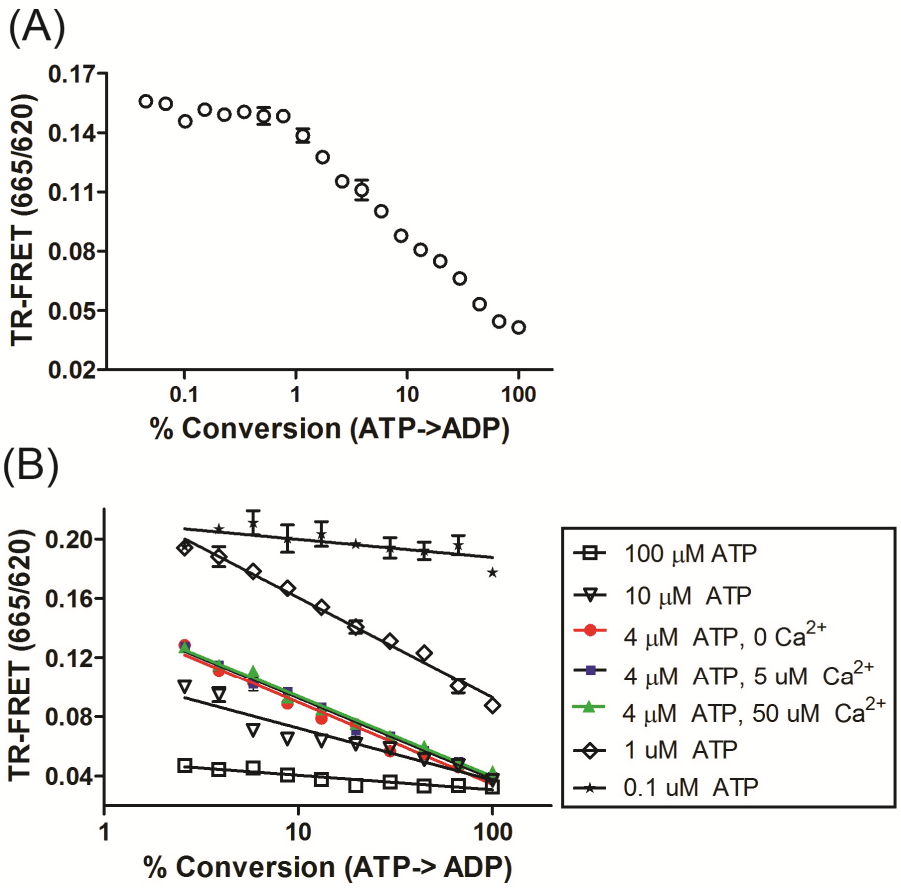

Figure 5.4 (A) FRET signals at different ATP to ADP conversion percentages. The total concentration of adenosine in each sample was 2.0 μM. Detection mixture contained 4 nM ADP antibody and 63 nM ADP tracer. (B) Calibration curves showing FRET signals at different ATP to ADP conversion percentages. Curves with black symbols and lines were obtained with different starting ATP concentrations in the absence of Ca^{2+}; curves with colored symbols and lines were obtained with varying concentrations of Ca^{2+} in the presence of 4 μM ATP. Each point is the average FRET value of three parallel samples. Error bars represent standard deviation.

signal only depends on the ATP to ADP conversion, and has no dependence on the Ca^{2+} concentration.

5.4.4 SERCA Activity Assay

SERCA titration was performed prior to activity assay in order to have a robust assay that was applicable to the lowest SERCA concentrations. As shown in Figure 5.5 (A) and (C), SERCA samples with the concentration spanning over 4 orders of magnitude were allowed to react with 4 μM of ATP at saturating Ca^{2+} concentrations in surfactant and reconstituted lipids, respectively. At desired SERCA concentrations, the FRET signal would at least decrease by ~80% of the maximum assay window within a reasonable reaction time. 5 nM and 500 pM SERCA were chosen to perform SERCA activity assay in surfactant and reconstituted lipids separately, as I observed that they gave ~80% decrease for a 5 min reaction and ~95% decrease for a 18 min or 20 min reaction, respectively. Compared with the coupled enzyme assay, our method reduced the concentration of SERCA by 10 folds in $C_{12}E_8$ and 100 folds in reconstituted lipids.

SERCA activity assays were finally performed in 4 μM ATP and 5 nM SERCA in $C_{12}E_8$ or 500 pM SERCA in reconstituted lipids. Reactions were quenched with the same volume of the optimized detection mixture to reach final concentrations of 2 μM adenosine, 4 nM antibody, and 63 nM tracer. Five reaction times were chosen with three experiments for each reaction time, and initial reaction rates were determined from linear regression. Shown in Figure 5.5 (B) and (D) are SERCA activity curves obtained by varying [Ca^{2+}]. Although the maximum initial

reaction rate (V_{max}) varied from day to day in both curves, these normalized curves shown very reproducible K_{Ca} values. The curve in Figure 5.5 (B) has a K_{Ca} value of 47±6 μM, compared to K_{Ca} values of 29±13 μM and 32±7 μM obtained from another two batches of SERCA sample (errors represent standard deviation). SERCA curve in Figure 5.5 (D) has a K_{Ca} value of 3±0.5 μM, compared to K_{Ca} values of 4.1±0.5 μM and 4±1 μM from another two reconstituted samples (errors represent standard deviation). Once PLN was

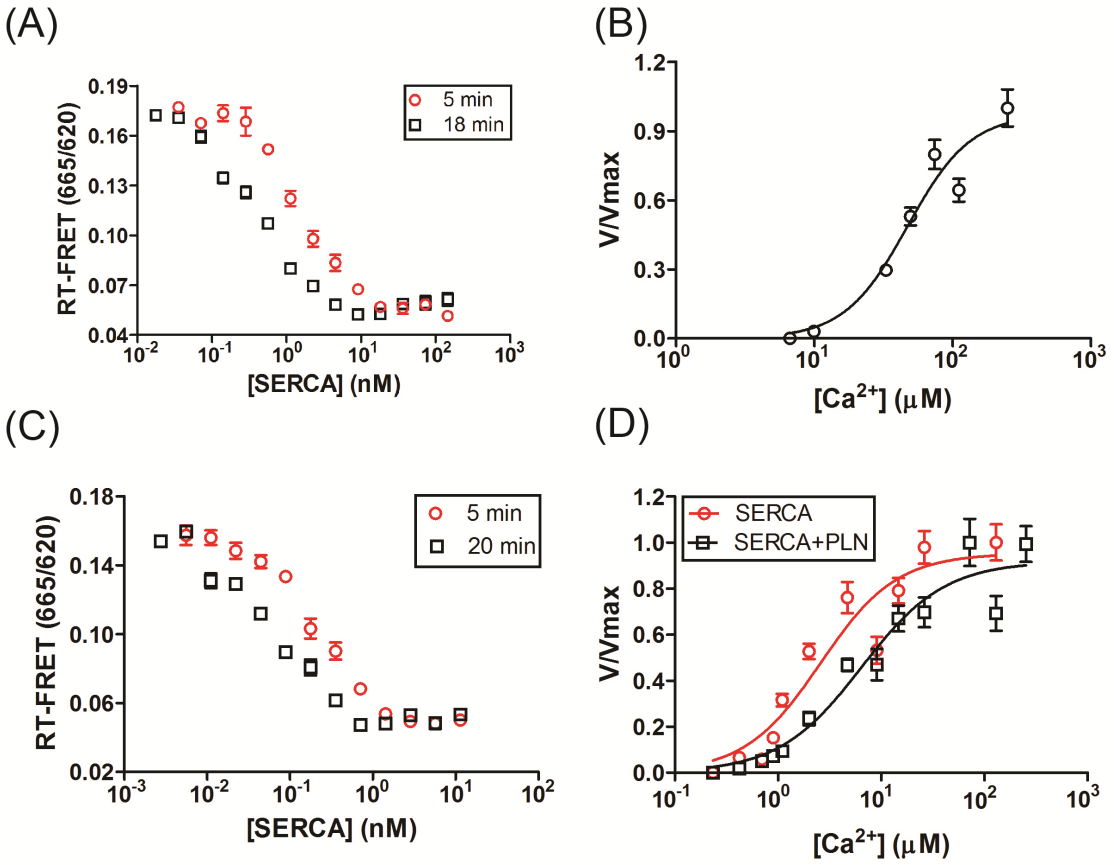

Figure 5.5 (A) SERCA titration curves in $C_{12}E_8$. Reactions were quenched at 5 min and 18 min. (B) Normalized SERCA activity curve in $C_{12}E_8$. (C) SERCA titration curves in reconstituted lipids. Reactions were quenched at 5 min and 20 min. (D) Normalized SERCA and SERCA+PLN activity curves in reconstituted lipids. Each data point shows the average FRET value of three parallel samples, and error bars represent standard deviation.

added, the curve shifted to high Ca^{2+} concentration, indicating the inhibition of PLN on SERCA activity. The shift of pK_{Ca} (ΔpK_{Ca}) in Figure 5.5 (D) is 0.40±0.17, compared to ΔpK_{Ca} values of 0.48±0.7 and 0.45±0.13 from another two assays in reconstituted lipids performed on different days (errors represent standard deviation).

Interestingly, the SERCA+PLN sample in Figure 5.5 (D) had the concentration of 500 pM SERCA and 25 nM PLN. It is obvious from Figure 5.5 (D) that this concentration of PLN already shifted the SERCA activity curve, suggesting a significant fraction of SERCA being bound by PLN. Conservatively, if we assume 50% of SERCA was in the SERCA-PLN complex, the dissociation constant (K_d) was approximately 25 nM calculated from the chemical equilibrium. The K_d is even smaller if the ratio of SERCA in the complex form is higher. The unusually low SERCA and PLN concentrations used in the assay opened up a promising way to estimate K_d of these two binding membrane proteins, which is conventionally very difficult.[302]

5.5 Conclusions

In conclusion, we developed a sensitive SERCA activity assay with the Transcreener® ADP^2 Assay TR-FRET Red kit. The optimized detection mixture contains 4 nM ADP antibody and 63 nM ADP tracer, allowing assays to be performed at 4 μM ATP with the maximum assay detection window. Our assays reduced the SERCA concentration by 10 folds and 100 folds compared to the coupled enzyme assays. Considering the decrease in reaction volume from 200

μL to 10 μL, the amount of SERCA used in each reaction was cut by 200 folds and 2000 folds for assays in $C_{12}E_8$ and reconstituted lipids separately. Moreover, SERCA K_{Ca} and pK_{Ca} shift in the presence of PLN were very reproducible from day to day experiments. Finally, the greatly improved sensitivity offers us additional information on SERCA-PLN binding. The K_d is at most 25 nM in reconstituted lipids, three orders of magnitude smaller than previous estimation.[302]

Chapter 6 Conclusions and Future Directions

6. 1 Conclusions

Aptamers have been gaining much attention and their applications have been growing explosively in many different areas, including detection, quantification, purification, and diagnostics and drug candidates. Selecting aptamers, however, is one major bottleneck that impedes their broader application and recognition, in sharp contrast to antibodies. Although rational designs have been used to generate aptamers based on prior knowledge on structure,[303, 304] *in vitro* selection from a combinatorial library is the most general and popular way, because the capability of discovering novel binding sites and binding mechanisms is crucial in generating new binding motifs for existing targets, or selecting aptamers for novel targets. The priority then falls onto improving *in vitro* selection efficiency, including improving separation efficiency to cut down selection cycles and speeding up an individual cycle. This thesis has highlighted progress I have made toward achieving this goal during my PhD study.

CE-SELEX is an advanced selection method with excellent separation efficiency, but few efforts have been reported toward the fundamental understanding. Chapter 2 described the application of 454 pyrosequencing to profile the CE-SELEX selection process against a protein target, rhVEGF$_{165}$. High affinity ssDNA ligands with mid-nM affinity were selected. At the same time, sequence information of up to 30,000 sequences per selection cycle was obtained by the NGS technique. Using home-programmed code, MEME, and CodonCode aligner, enriched sequences were identified with the highest abundance of 0.8%, indicating tremendous heterogeneity in the selected pools.

With the ample sequence information, the evolution profiles of the most abundant and desired PCR products were obtained, which allowed us to visualize the enrichment of aptamers. The comparison of evolution profiles of the desired and undesired sequences also revealed the correlation between the decrease of the most abundant aptamer and the increase of the most abundant short byproduct in selection cycle 6 and 7, as well as the structural relevance.

Although CE-SELEX is an advantageous platform to generate aptamers, it has limitations: only a small amount of the initial library is injected and subjected to separation, and it is thus difficult to achieve automation in order to increase the turnover of one selection cycle. Chapter 3 described the first successful application of a microfluidic free-solution-based separation platform, μFFE, in generating DNA aptamers for a protein, human IgE. 1.8×10^{14} sequences were continuously introduced into the device and analyzed, which improved the sample input more than 300 folds compared to CE. Mid-nM affinity aptamers were generated right after one cycle of selection. Although it took one day to complete one cycle of selection, the same as CE-SELEX, μFFE-SELEX has the potential of ultimate automation once integrated with other microfluidic devices fulfilling the rest of duties in SELEX, which will greatly reduce the time of one selection.

Besides the larger sample input and greater potential of automation, μFFE device also has many advantages over CE in terms of separation, and thus there is much promise for generating aptamers for targets that are conventionally problematic. Chapter 4 described the first attempt to generate aptamers toward a

whole integral membrane protein, SERCA, using a μFFE device. High-nM affinity pools were obtained after five rounds of selection as characterized by both FP and μFFE, which was an at least two orders of magnitude improvement over the unselected library. Fraction collection of SERCA-aptamer complex was based on the spatial difference, which was more straightforward on μFFE than on CE. This is because CE required additional hydrodynamic rinsing, since the complex barely moved with $C_{12}E_8$ coating the inner wall of the capillary. The decoupled electric field and transit time on μFFE device made the separation power more controllable and adjustable; however, electric field and migration time of analytes are inversely correlated on CE.

Besides improvements on SELEX, this thesis also covers studies performed on SERCA, an important membrane protein in normal cardiac muscle functions. Chapter 4 described an interesting finding that ssDNA sequences bind to one of the endogenous inhibitors of SERCA, PLN, with as low as high-pMK_d. The interaction was extensively characterized in terms of base order, sequence length, and structure. It was found that binding affinity is not related to the base order, which means the random library are naturally great drug candidates in inhibiting the activity of PLN. Moreover, the binding can be improved by increasing sequence length and structural flexibility. More importantly, these ssDNA sequences have been demonstrated to release the inhibition of PLN to SERCA *in vitro*, the degree of which also depends on the sequence length (by Dr. Raffaello Verardi and Prof. Gianluigi Veglia, unpublished results). In Chapter 5, the development of a novel SERCA activity assay was described, which was

based on directly detecting the product ADP in a competitive immunobinding format. This assay is compatible and reproducible with SERCA prepared in both detergent and reconstituted lipids. Under the optimized condition, the amount of SERCA used in assay was reduced up to 1000 folds compared to the traditional PK/LDH-coupled enzyme assay. Another important conclusion from this sensitive assay is the binding of PLN and SERCA has at most a mid-nM K_d.

6. 2 Future Directions

6.2.1 SELEX Studies on SERCA

As promising results showed binding of SERCA and selected pool 5 in the first attempt using μFFE, selection was not reproducible when a lower electric field (80 Vcm^{-1}) was applied to keep more aptamers in the complex form, as discussed in Chapter 4. Other than sample aging, it was also possible that the separation of the bound from the unbound was insufficient. One way is to improve the separation power of the μFFE device, defined as P = Et,[98, 305] by tuning electric field and separation flow rate. As the electric field goes down from 150 Vcm^{-1} to 80 Vcm^{-1} or even lower, the buffer flow rate needs to go down proportionally to keep the transit time long enough to achieve good separation. On the other hand, stream stability of μFFE, which is highly correlated to buffer flow rate, is crucial to SELEX application, since successful collection is critical in the progress of selection. Although our group has demonstrated surface tension and formation of air bubbles can be significantly reduced by adding buffer additives such as Triton-X 100,[256] one should keep in mind that lowering flow rate

always trades off stream stability. Alternatively, device design could be changed by increasing the length of the separation chamber so that sufficient transit time is achieved under original flow rate.

Although SERCA structures and activities have been broadly characterized in surfactants such as $C_{12}E_8$, it is well accepted that there are interactions between hydrophilic region of $C_{12}E_8$ and cytoplasmic domain of SERCA, which may prevent aptamers from binding to the desired sites. Moreover, from my experience of developing the new SERCA activity assay as described in Chapter 5, sample storage and long term stability may be other concerns in SELEX study. Therefore, another option to generate aptamers for SERCA is to adopt a more reliable sample preparation. SERCA has been successfully prepared in lipid bilayers on thin glass slides for NMR experiments, which showed good stability and similar structure as that in SR.[306, 307] Preliminary experiments were performed to test the compatibility of this sample preparation to SELEX. 166 μg of IAEDANS-labeled SERCA in DOPC-DOPE bilayers (DOPC:DOPE 4:1,) was spread on a thin glass slide for evaporation at 37 °C, followed by rehydration in a 100% humidity chamber. This slide was then rinsed with 40 μL of MOPS buffer with gently pipetting for 5 min, imitating incubation with ssDNA library in SELEX. The rinsing buffer was collected for fluorescence intensity measurements with a 360 nm excitation line and a 490 nm emission line. The rinse was repeated for four more times. As shown in Figure 6.1, only small amounts of SERCA were rinsed off chip from Rinse 1 through 5, very close to the buffer background. Only when the glass slide was finally incubated with the 6th rinsing buffer and

subjected to sonication and vortexing, most SERCA came off. This experiment demonstrates that this sample preparation is very compatible with SELEX incubation, rinsing, and collection procedures. On the other hand, it should be noted that longer incubation time is needed to allow aptamers to diffuse in order to be in contact with SERCA; the low separation efficiency means that selection rounds could go up to 20 cycles.

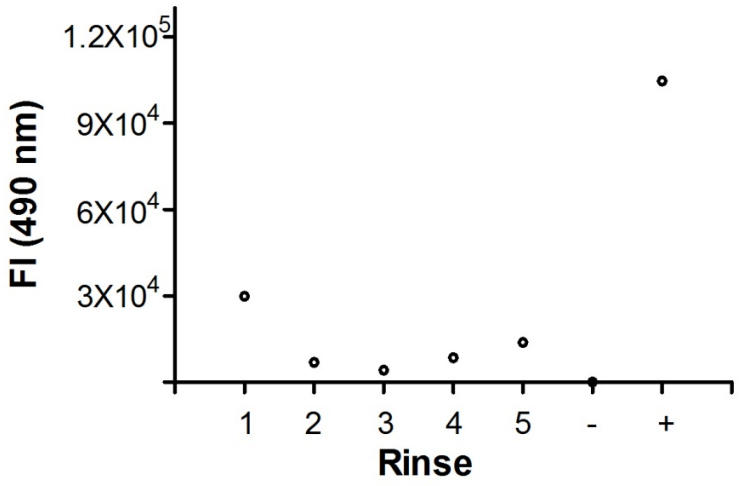

Figure 6.1 A SERCA-SELEX control of SERCA prepared in lipid bilayers on 10 mm×8 mm×60 μm glass slides. The amount of free IAEDANS-labeled SERCA in the same volume of rinses was monitored from Rinse 1 to 5. In negative control, the rinse was performed on a glass slide with only lipid bilayers. In positive control, the glass slide that was rinsed five times before was soaked in the same volume of rinsing buffer for vortexing and sonication. Three measurements were taken for each sample, and data are shown here with error bars representing standard deviations.

6.2.2 Integration of Microfluidic Devices for SELEX

The ultimate goal of developing microfluidics for SELEX is to integrate individual microfluidic devices that are capable of performing one or more step(s) into the whole machinery for the final automation of SELEX. As one example, Huang *et al.* reported a microfluidic system that combined incubation, magnetic

separation, and fast PCR amplification for CRP aptamer generation, which saved 60% of sample consumption, and cut 100 min for one selection.[254] Bowser group is also dedicated to the development of a microfluidic DNA purification device. Sheng *et al.* fabricated a counter current dialysis device with a nanoporous polycarbonate membrane separating two PDMS channels.[308] This device showed selective transport of fluorescein and ssDNA molecules with different length. In the most recent work, 80 mer ssDNA molecules were transported through the membrane using streptavidin beads in co-current mode with much improved recovery. Miniaturization of PCR process has been well developed in the past decades with many designs compatible with a range of sample volume and flow rate.[309, 310] Integration of μFFE separation and microfluidic PCR and DNA purification devices would be a huge step moving forward, not only because of the advantages associated with μFFE separation, but also due to the fact of moving almost all SELEX work onto this machinery, thus reducing time needed to perform one cycle, minimizing contamination, making it more close to the final automation.

6.2.3 Screening Drug Candidates that Release PLN Inhibition to SERCA

SERCA activity is a focus in heart diseases because of its important role in the cycling and control of Ca^{2+} in a heartbeat. Releasing the inhibition of PLN to SERCA is one important way to activate SERCA and regain the functionality. Small molecules such as quercetin[311] and plant phenol tannin,[312] heparin[313] and heparin-derived oligosaccharides[314] have been identified as inhibitors to PLN;

their stimulation effect to cardiac SERCA have been characterized; and parameters such as concentration[311] and electrostatic phenomenon[315] have been studied. However, to our best knowledge, it is the first time that an inhibitor with high picomolar K_d was identified; the structural-affinity relationship in terms of base order, sequence length and flexibility were systematically characterized, offering us precious knowledge on exploring potential drug candidates for curing heart failure caused by SERCA dysfunction. One direction of future work is to screen compounds whose syntheses are straightforward and structures are more bio-stable than DNA. Binding experiments can be performed on modified DNA and RNA molecules such as methylated nucleic acids, L-nucleic acids, and LNA, to characterize binding affinity and structure-binding affinity relationship. One other direction is to expand SERCA activity assay, which is now performed in reconstituted lipids, to sarcoplasmic reticulum, cells, and finally *in vivo*.

Bibliography

(1) Stark, B. C.; Kole, R.; Bowman, E. J.; Altman, S. *Proceedings of the National Academy of Sciences of the United States of America* **1978**, *75*, 3717-3721.
(2) Wilson, D. S.; Szostak, J. W. *Annual Review of Biochemistry* **1999**, *68*, 611-647.
(3) Syed, M. A.; Pervaiz, S. *Oligonucleotides* **2010**, *20*, 215-224.
(4) Ciesiolka, J.; Yarus, M. *Rna-a Publication of the Rna Society* **1996**, *2*, 785-793.
(5) Hofmann, H. P.; Limmer, S.; Hornung, V.; Sprinzl, M. *Rna-a Publication of the Rna Society* **1997**, *3*, 1289-1300.
(6) Famulok, M. *Journal of the American Chemical Society* **1994**, *116*, 1698-1706.
(7) Geiger, A.; Burgstaller, P.; vonderEltz, H.; Roeder, A.; Famulok, M. *Nucleic Acids Research* **1996**, *24*, 1029-1036.
(8) Yarus, M. *Journal of Molecular Evolution* **1998**, *47*, 109-117.
(9) Wallis, M. G.; Streicher, B.; Wank, H.; vonAhsen, U.; Clodi, E.; Wallace, S. T.; Famulok, M.; Schroeder, R. *Chemistry & Biology* **1997**, *4*, 357-366.
(10) Patel, D. J.; Suri, A. K.; Jiang, F.; Jiang, L. C.; Fan, P.; Kumar, R. A.; Nonin, S. *Journal of Molecular Biology* **1997**, *272*, 645-664.
(11) Berens, C.; Thain, A.; Schroeder, R. *Bioorganic & Medicinal Chemistry* **2001**, *9*, 2549-2556.
(12) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550-553.
(13) Burgstaller, P.; Famulok, M. *Angewandte Chemie-International Edition in English* **1994**, *33*, 1084-1087.
(14) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, *355*, 564-566.
(15) Tuerk, C.; Macdougal, S.; Gold, L. *Proceedings of the National Academy of Sciences of the United States of America* **1992**, *89*, 6988-6992.
(16) Ringquist, S.; Jones, T.; Snyder, E. E.; Gibson, T.; Boni, I.; Gold, L. *Biochemistry* **1995**, *34*, 3640-3648.
(17) Shangguan, D.; Li, Y.; Tang, Z. W.; Cao, Z. H. C.; Chen, H. W.; Mallikaratchy, P.; Sefah, K.; Yang, C. Y. J.; Tan, W. H. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 11838-11843.
(18) Tang, Z. W.; Shangguan, D.; Wang, K. M.; Shi, H.; Sefah, K.; Mallikratchy, P.; Chen, H. W.; Li, Y.; Tan, W. H. *Analytical Chemistry* **2007**, *79*, 4900-4907.
(19) Hicke, B. J.; Marion, C.; Chang, Y. F.; Gould, T.; Lynott, C. K.; Parma, D.; Schmidt, P. G.; Warren, S. *Journal of Biological Chemistry* **2001**, *276*, 48644-48654.
(20) Shangguan, D. H.; Meng, L.; Cao, Z. H. C.; Xiao, Z. Y.; Fang, X. H.; Li, Y.; Cardona, D.; Witek, R. P.; Liu, C.; Tan, W. H. *Analytical Chemistry* **2008**, *80*, 721-728.

(21) Ulrich, H.; Magdesian, M. H.; Alves, M. J. M.; Colli, W. *Journal of Biological Chemistry* **2002**, *277*, 20756-20762.
(22) Romig, T. S.; Bell, C.; Drolet, D. W. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **1999**, *731*, 275-284.
(23) Deng, Q.; German, I.; Buchanan, D.; Kennedy, R. T. *Analytical Chemistry* **2001**, *73*, 5415-5421.
(24) Hutanu, D.; Remcho, V. T. *Advances in Chromatography, Vol 45* **2007**, *45*, 173-196.
(25) Huang, C. C.; Cao, Z. H.; Chang, H. T.; Tan, W. H. *Analytical Chemistry* **2004**, *76*, 6973-6981.
(26) German, I.; Buchanan, D. D.; Kennedy, R. T. *Analytical Chemistry* **1998**, *70*, 4540-4545.
(27) Pavlov, V.; Xiao, Y.; Shlyahovsky, B.; Willner, I. *Journal of the American Chemical Society* **2004**, *126*, 11768-11769.
(28) Mairal, T.; Ozalp, V. C.; Sanchez, P. L.; Mir, M.; Katakis, I.; O'Sullivan, C. K. *Analytical and Bioanalytical Chemistry* **2008**, *390*, 989-1007.
(29) Hamaguchi, N.; Ellington, A.; Stanton, M. *Analytical Biochemistry* **2001**, *294*, 126-131.
(30) McCauley, T. G.; Hamaguchi, N.; Stanton, M. *Analytical Biochemistry* **2003**, *319*, 244-250.
(31) Proske, D.; Blank, M.; Buhmann, R.; Resch, A. *Applied Microbiology and Biotechnology* **2005**, *69*, 367-374.
(32) Tombelli, S.; Minunni, M.; Mascini, M. *Biomolecular Engineering* **2007**, *24*, 191-200.
(33) Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R.; Adamis, A. P. *Nature Reviews Drug Discovery* **2006**, *5*, 123-132.
(34) Siller-Matula, J. M.; Merhi, Y.; Tanguay, J. F.; Duerschmied, D.; Wagner, D. D.; McGinness, K. E.; Pendergrast, P. S.; Chung, J. K.; Tian, X. B.; Schaub, R. G.; Jilma, B. *Arteriosclerosis Thrombosis and Vascular Biology* **2012**, *32*, 902-U100.
(35) Ni, X.; Castanares, M.; Mukherjee, A.; Lupold, S. E. *Current Medicinal Chemistry* **2011**, *18*, 4206-4214.
(36) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818-822.
(37) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510.
(38) Eulberg, D.; Klussmann, S. *Chembiochem* **2003**, *4*, 979-983.
(39) Jensen, K. B.; Atkinson, B. L.; Willis, M. C.; Koch, T. H.; Gold, L. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 12220-12224.
(40) Rhie, A.; Kirby, L.; Sayer, N.; Wellesley, R.; Disterer, P.; Sylvester, I.; Gill, A.; Hope, J.; James, W.; Tahiri-Alaoui, A. *Journal of Biological Chemistry* **2003**, *278*, 39697-39705.
(41) Bruno, J. G.; Kiel, J. L. *Biosensors & Bioelectronics* **1999**, *14*, 457-464.
(42) Blank, M.; Weinschenk, T.; Priemer, M.; Schluesener, H. *Journal of Biological Chemistry* **2001**, *276*, 16464-16468.

(43) Misono, T. S.; Kumar, P. K. R. *Analytical Biochemistry* **2005**, *342*, 312-317.
(44) Shamah, S. M.; Healy, J. M.; Cload, S. T. *Accounts of Chemical Research* **2008**, *41*, 130-138.
(45) Mendonsa, S. D.; Bowser, M. T. *Journal of the American Chemical Society* **2004**, *126*, 20-21.
(46) Mendonsa, S. D.; Bowser, M. T. *Analytical Chemistry* **2004**, *76*, 5387-5392.
(47) Mendonsa, S. D.; Bowser, M. T. *Journal of the American Chemical Society* **2005**, *127*, 9382-9383.
(48) Mosing, R. K.; Mendonsa, S. D.; Bowser, M. T. *Analytical Chemistry* **2005**, *77*, 6107-6112.
(49) Tang, J. J.; Xie, J. W.; Shao, N. S.; Yan, Y. *Electrophoresis* **2006**, *27*, 1303-1311.
(50) Mallikaratchy, P.; Stahelin, R. V.; Cao, Z. H.; Cho, W. H.; Tan, W. H. *Chemical Communications* **2006**, 3229-3231.
(51) Cella, L. N.; Sanchez, P.; Zhong, W. W.; Myung, N. V.; Chen, W.; Mulchandani, A. *Analytical Chemistry* **2010**, *82*, 2042-2047.
(52) Tok, J.; Lai, J.; Leung, T.; Li, S. F. Y. *Electrophoresis* **2010**, *31*, 2055-2062.
(53) Williams, B. A. R.; Lin, L. Y.; Lindsay, S. M.; Chaput, J. C. *Journal of the American Chemical Society* **2009**, *131*, 6330-6331.
(54) Lin, L.; Hom, D.; Lindsay, S. M.; Chaput, J. C. *Journal of the American Chemical Society* **2007**, *129*, 14568-14569.
(55) Drmanac, R.; Sparks, A. B.; al., e. *Science* **2010**, *327*, 78-81.
(56) Levy, S.; Sutton, G.; Ng, P. C.; Feuk, L.; Halpern, A. L.; Walenz, B. P.; Axelrod, N.; Huang, J.; Kirkness, E. F.; Denisov, G.; Lin, Y.; MacDonald, J. R.; Pang, A. W. C.; Shago, M.; Stockwell, T. B.; Tsiamouri, A.; Bafna, V.; Bansal, V.; Kravitz, S. A.; Busam, D. A.; Beeson, K. Y.; McLntosh, T. C.; Remington, K. A.; Abril, J. F.; Gill, J.; Borman, J.; Rogers, Y. H.; Frazier, M. E.; Scherer, S. W.; Strausberg, R. L.; Venter, J. C. *Plos Biology* **2007**, *5*, 2113-2144.
(57) Wheeler, D. A.; Srinivasan, M.; Egholm, M.; Shen, Y.; Chen, L.; McGuire, A.; He, W.; Chen, Y. J.; Makhijani, V.; Roth, G. T.; Gomes, X.; Tartaro, K.; Niazi, F.; Turcotte, C. L.; Irzyk, G. P.; Lupski, J. R.; Chinault, C.; Song, X. Z.; Liu, Y.; Yuan, Y.; Nazareth, L.; Qin, X.; Muzny, D. M.; Margulies, M.; Weinstock, G. M.; Gibbs, R. A.; Rothberg, J. M. *Nature* **2008**, *452*, 872-U875.
(58) Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R.; Boutell, J. M.; Bryant, J.; Carter, R. J.; Cheetham, R. K.; Cox, A. J.; Ellis, D. J.; Flatbush, M. R.; Gormley, N. A.; Humphray, S. J.; Irving, L. J.; Karbelashvili, M. S.; Kirk, S. M.; Li, H.; Liu, X. H.; Maisinger, K. S.; Murray, L. J.; Obradovic, B.; Ost, T.; Parkinson, M. L.; Pratt, M. R.; Rasolonjatovo, I. M. J.; Reed, M. T.; Rigatti, R.; Rodighiero, C.; Ross, M. T.; Sabot, A.; Sankar, S. V.; Scally, A.; Schroth, G. P.; Smith, M. E.; Smith, V. P.; Spiridou, A.; Torrance, P. E.; Tzonev, S. S.; Vermaas, E. H.; Walter, K.;

Wu, X. L.; Zhang, L.; Alam, M. D.; Anastasi, C.; Aniebo, I. C.; Bailey, D. M. D.; Bancarz, I. R.; Banerjee, S.; Barbour, S. G.; Baybayan, P. A.; Benoit, V. A.; Benson, K. F.; Bevis, C.; Black, P. J.; Boodhun, A.; Brennan, J. S.; Bridgham, J. A.; Brown, R. C.; Brown, A. A.; Buermann, D. H.; Bundu, A. A.; Burrows, J. C.; Carter, N. P.; Castillo, N.; Catenazzi, M. C. E.; Chang, S.; Cooley, R. N.; Crake, N. R.; Dada, O. O.; Diakoumakos, K. D.; Dominguez-Fernandez, B.; Earnshaw, D. J.; Egbujor, U. C.; Elmore, D. W.; Etchin, S. S.; Ewan, M. R.; Fedurco, M.; Fraser, L. J.; Fajardo, K. V. F.; Furey, W. S.; George, D.; Gietzen, K. J.; Goddard, C. P.; Golda, G. S.; Granieri, P. A.; Green, D. E.; Gustafson, D. L.; Hansen, N. F.; Harnish, K.; Haudenschild, C. D.; Heyer, N. I.; Hims, M. M.; Ho, J. T.; Horgan, A. M.; Hoschler, K.; Hurwitz, S.; Ivanov, D. V.; Johnson, M. Q.; James, T.; Jones, T. A. H.; Kang, G. D.; Kerelska, T. H.; Kersey, A. D.; Khrebtukova, I.; Kindwall, A. P.; Kingsbury, Z.; Kokko-Gonzales, P. I.; Kumar, A.; Laurent, M. A.; Lawley, C. T.; Lee, S. E.; Lee, X.; Liao, A. K.; Loch, J. A.; Lok, M.; Luo, S. J.; Mammen, R. M.; Martin, J. W.; McCauley, P. G.; McNitt, P.; Mehta, P.; Moon, K. W.; Mullens, J. W.; Newington, T.; Ning, Z. M.; Ng, B. L.; Novo, S. M.; O'Neill, M. J.; Osborne, M. A.; Osnowski, A.; Ostadan, O.; Paraschos, L. L.; Pickering, L.; Pike, A. C.; Pinkard, D. C.; Pliskin, D. P.; Podhasky, J.; Quijano, V. J.; Raczy, C.; Rae, V. H.; Rawlings, S. R.; Rodriguez, A. C.; Roe, P. M.; Rogers, J.; Bacigalupo, M. C. R.; Romanov, N.; Romieu, A.; Roth, R. K.; Rourke, N. J.; Ruediger, S. T.; Rusman, E.; Sanches-Kuiper, R. M.; Schenker, M. R.; Seoane, J. M.; Shaw, R. J.; Shiver, M. K.; Short, S. W.; Sizto, N. L.; Sluis, J. P.; Smith, M. A.; Sohna, J. E. S.; Spence, E. J.; Stevens, K.; Sutton, N.; Szajkowski, L.; Tregidgo, C. L.; Turcatti, G.; vandeVondele, S.; Verhovsky, Y.; Virk, S. M.; Wakelin, S.; Walcott, G. C.; Wang, J. W.; Worsley, G. J.; Yan, J. Y.; Yau, L.; Zuerlein, M.; Mullikin, J. C.; Hurles, M. E.; McCooke, N. J.; West, J. S.; Oaks, F. L.; Lundberg, P. L.; Klenerman, D.; Durbin, R.; Smith, A. J. *Nature* **2008**, *456*, 53-59.

(59) Pushkarev, D.; Neff, N. F.; Quake, S. R. *Nature Biotechnology* **2009**, *27*, 847-U101.

(60) McKernan, K. J.; Peckham, H. E.; Costa, G. L.; McLaughlin, S. F.; Fu, Y. T.; Tsung, E. F.; Clouser, C. R.; Duncan, C.; Ichikawa, J. K.; Lee, C. C.; Zhang, Z.; Ranade, S. S.; Dimalanta, E. T.; Hyland, F. C.; Sokolsky, T. D.; Zhang, L.; Sheridan, A.; Fu, H. N.; Hendrickson, C. L.; Li, B.; Kotler, L.; Stuart, J. R.; Malek, J. A.; Manning, J. M.; Antipova, A. A.; Perez, D. S.; Moore, M. P.; Hayashibara, K. C.; Lyons, M. R.; Beaudoin, R. E.; Coleman, B. E.; Laptewicz, M. W.; Sannicandro, A. E.; Rhodes, M. D.; Gottimukkala, R. K.; Yang, S.; Bafna, V.; Bashir, A.; MacBride, A.; Alkan, C.; Kidd, J. M.; Eichler, E. E.; Reese, M. G.; De la Vega, F. M.; Blanchard, A. P. *Genome Research* **2009**, *19*, 1527-1541.

(61) Kim, J. I.; Ju, Y. S.; Park, H.; Kim, S.; Lee, S.; Yi, J. H.; Mudge, J.; Miller, N. A.; Hong, D.; Bell, C. J.; Kim, H. S.; Chung, I. S.; Lee, W. C.; Lee, J. S.; Seo, S. H.; Yun, J. Y.; Woo, H. N.; Lee, H.; Suh, D.; Kim, H. J.;

Yavartanoo, M.; Kwak, M.; Zheng, Y.; Lee, M. K.; Kim, J. Y.; Gokcumen, O.; Mills, R. E.; Zaranek, A. W.; Thakuria, J.; Wu, X.; Kim, R. W.; Huntley, J. J.; Luo, S.; Schroth, G. P.; Wu, T. D.; Kim, H.; Yang, K. S.; Park, W. Y.; Church, G. M.; Lee, C.; Kingsmore, S. F.; Seo, J. S. *Nature* **2009**, *460*, 1011-U1096.

(62) Drmanac, R.; Sparks, A. B.; Callow, M. J.; Halpern, A. L.; Burns, N. L.; Kermani, B. G.; Carnevali, P.; Nazarenko, I.; Nilsen, G. B.; Yeung, G.; Dahl, F.; Fernandez, A.; Staker, B.; Pant, K. P.; Baccash, J.; Borcherding, A. P.; Brownley, A.; Cedeno, R.; Chen, L. S.; Chernikoff, D.; Cheung, A.; Chirita, R.; Curson, B.; Ebert, J. C.; Hacker, C. R.; Hartlage, R.; Hauser, B.; Huang, S.; Jiang, Y.; Karpinchyk, V.; Koenig, M.; Kong, C.; Landers, T.; Le, C.; Liu, J.; McBride, C. E.; Morenzoni, M.; Morey, R. E.; Mutch, K.; Perazich, H.; Perry, K.; Peters, B. A.; Peterson, J.; Pethiyagoda, C. L.; Pothuraju, K.; Richter, C.; Rosenbaum, A. M.; Roy, S.; Shafto, J.; Sharanhovich, U.; Shannon, K. W.; Sheppy, C. G.; Sun, M.; Thakuria, J. V.; Tran, A.; Vu, D.; Zaranek, A. W.; Wu, X. D.; Drmanac, S.; Oliphant, A. R.; Banyai, W. C.; Martin, B.; Ballinger, D. G.; Church, G. M.; Reid, C. A. *Science*, *327*, 78-81.

(63) Cox, J. C.; Rudolph, P.; Ellington, A. D. *Biotechnology Progress* **1998**, *14*, 845-850.

(64) Cox, J. C.; Ellington, A. D. *Bioorganic & Medicinal Chemistry* **2001**, *9*, 2525-2531.

(65) Cox, J. C.; Rajendran, M.; Riedel, T.; Davidson, E. A.; Sooter, L. J.; Bayer, T. S.; Schmitz-Brown, M.; Ellington, A. D. *Combinatorial Chemistry & High Throughput Screening* **2002**, *5*, 289-299.

(66) Eulberg, D.; Buchner, K.; Maasch, C.; Klussmann, S. *Nucleic Acids Research* **2005**, *33*.

(67) Manz, A.; Graber, N.; Widmer, H. M. *Sensors and Actuators B-Chemical* **1990**, *1*, 244-248.

(68) Smejkal, P.; Foret, F. *Chemicke Listy* **2012**, *106*, 104-112.

(69) Jokerst, J. C.; Emory, J. M.; Henry, C. S. *Analyst* **2012**, *137*, 24-34.

(70) Beebe, D. J.; Mensing, G. A.; Walker, G. M. *Annual Review of Biomedical Engineering* **2002**, *4*, 261-286.

(71) Dittrich, P. S.; Manz, A. *Nature Reviews Drug Discovery* **2006**, *5*, 210-218.

(72) Weibel, D. B.; Whitesides, G. M. *Current Opinion in Chemical Biology* **2006**, *10*, 584-591.

(73) Skurtys, O.; Aguilera, J. M. *Food Biophysics* **2008**, *3*, 1-15.

(74) Kraly, J. R.; Holcomb, R. E.; Guan, Q.; Henry, C. S. *Analytica Chimica Acta* **2009**, *653*, 23-35.

(75) Hybarger, G.; Bynum, J.; Williams, R. F.; Valdes, J. J.; Chambers, J. P. *Analytical and Bioanalytical Chemistry* **2006**, *384*, 191-198.

(76) Zhang, C. S.; Xu, J. L.; Ma, W. L.; Zheng, W. L. *Biotechnology Advances* **2006**, *24*, 243-284.

(77) Ahn, J. Y.; Jo, M.; Dua, P.; Lee, D. K.; Kim, S. *Oligonucleotides* **2011**, *21*, 93-100.

(78) Park, S. M.; Ahn, J. Y.; Jo, M.; Lee, D. K.; Lis, J. T.; Craighead, H. G.; Kim, S. *Lab on a Chip* **2009**, *9*, 1206-1212.
(79) Kim, S.; Kim, Y.; Kim, P.; Ha, J. M.; Kim, K.; Sohn, M.; Yoo, J. S.; Lee, J.; Kwon, J. A.; Lee, K. N. *Analytical Chemistry* **2006**, *78*, 7392-7396.
(80) Fan, X. C.; Shi, H.; Adelman, K.; List, J. T. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 6934-6939.
(81) Kim, T. K.; Lee, S. W.; Ahn, J. Y.; Laurell, T.; Kim, S. Y.; Jeong, O. C. *Japanese Journal of Applied Physics* **2011**, *50*, 06GL05-01-06GL05-04.
(82) Bruno, J. G.; Kiel, J. L. *Biotechniques* **2002**, *32*, 178.
(83) Tok, J. B. H.; Fischer, N. O. *Chemical Communications* **2008**, 1883-1885.
(84) Lou, X. H.; Qian, J. R.; Xiao, Y.; Viel, L.; Gerdon, A. E.; Lagally, E. T.; Atzberger, P.; Tarasow, T. M.; Heeger, A. J.; Soh, H. T. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 2989-2994.
(85) Qian, J. R.; Lou, X. H.; Zhang, Y. T.; Xiao, Y.; Soh, H. T. *Analytical Chemistry* **2009**, *81*, 5490-5495.
(86) Liu, Y. L.; Adams, J. D.; Turner, K.; Cochran, F. V.; Gambhir, S. S.; Soh, H. T. *Lab on a Chip* **2009**, *9*, 1033-1036.
(87) Ahmad, K. M.; Oh, S. S.; Kim, S.; McClellen, F. M.; Xiao, Y.; Soh, H. T. *Plos One* **2011**, *6*.
(88) Oh, S. S.; Ahmad, K. M.; Cho, M.; Kim, S.; Xiao, Y.; Soh, H. T. *Analytical Chemistry* **2011**, *83*, 6883-6889.
(89) Barrolier, V. J.; Watzke, E.; Gibian, H. Z. *Naturforshung* **1958**, *13B*, 754.
(90) Hannig, K. *Zeitschrift fuer Analytische Chemie* **1961**, *181*, 244-254.
(91) Lion, N.; Rohner, T. C.; Dayon, L.; Arnaud, I. L.; Damoc, E.; Youhnovski, N.; Wu, Z. Y.; Roussel, C.; Josserand, J.; Jensen, H.; Rossier, J. S.; Przybylski, M.; Girault, H. H. *Electrophoresis* **2003**, *24*, 3533-3562.
(92) Heidrich, H. G.; Hannig, K. *Methods in Enzymology* **1989**, *171*, 513-531.
(93) Turgeon, R. T.; Bowser, M. T. *Analytical and Bioanalytical Chemistry* **2009**, *394*, 187-198.
(94) Raymond, D. E.; Manz, A.; Widmer, H. M. *Analytical Chemistry* **1994**, *66*, 2858-2865.
(95) Fonslow, B. R.; Bowser, M. T. *Analytical Chemistry* **2005**, *77*, 5706-5710.
(96) Zhang, C. X.; Manz, A. *Analytical Chemistry* **2003**, *75*, 5759-5766.
(97) Raymond, D. E.; Manz, A.; Widmer, H. M. *Analytical Chemistry* **1996**, *68*, 2515-2522.
(98) Fonslow, B. R.; Barocas, V. H.; Bowser, M. T. *Analytical Chemistry* **2006**, *78*, 5369-5374.
(99) Iqbal, S. S.; Mayo, M. W.; Bruno, J. G.; Bronk, B. V.; Batt, C. A.; Chambers, J. P. *Biosensors & Bioelectronics* **2000**, *15*, 549-578.
(100) Gopinath, S. C. B. *Analytica Chimica Acta* **2009**, *636*, 117-128.
(101) Connors, K. A. *Binding Constants*; John Wiley: New York, 1987.
(102) Harrison, F.; Katti, S. K. *Chemometrics and Intelligent Laboratory Systems* **1990**, *9*, 249-255.

(103) Larsson, A. *Journal of Immunological Methods* **1997**, *206*, 135-142.
(104) Deranleau, D. A. *Journal of the American Chemical Society* **1969**, *91*, 4044.
(105) Weber, G. *Biochemistry* **1965**, *4*, 1942.
(106) Hall, K. B.; Kranz, J. K. *Methods in Molecular Biology* **1999**, *118*, 105.
(107) Setzer, D. R. *Methods in Molecular Biology* **1999**, *118*, 115.
(108) Cruz-Aguado, J. A.; Penner, G. *Journal of Agricultural and Food Chemistry* **2008**, *56*, 10456-10461.
(109) Cary, P. D.; Kneale, G. G. *Methods in Molecular Biology* **2009**, *543*, 613.
(110) Romaniuk, P. J. *Journal of Biological Chemistry* **1990**, *265*, 17593-17600.
(111) Bock, C.; Coleman, M.; Collins, B.; Davis, J.; Foulds, G.; Gold, L.; Greef, C.; Heil, J.; Heilig, J. S.; Hicke, B.; Hurst, M. N.; Husar, G. M.; Miller, D.; Ostroff, R.; Petach, H.; Schneider, D.; Vant-Hull, B.; Waugh, S.; Weiss, A.; Wilcox, S. K.; Zichi, D. *Proteomics* **2004**, *4*, 609-618.
(112) Kensch, O.; Connolly, B. A.; Steinhoff, H. J.; McGregor, A.; Goody, R. S.; Restle, T. *Journal of Biological Chemistry* **2000**, *275*, 18271-18278.
(113) Gaillard, C.; Strauss, F. *BMC Molecular Biology* **2000**, *1*, 1.
(114) Jaouen, S.; de Koning, L.; Gaillard, C.; Muselikova-Polanska, E.; Stros, M.; Strauss, F. *Journal of Molecular Biology* **2005**, *353*, 822-837.
(115) Drabovich, A. P.; Berezovski, M.; Okhonin, V.; Krylov, S. N. *Analytical Chemistry* **2006**, *78*, 3171-3178.
(116) Oravcova, J.; Bohs, B.; Lindner, W. *Journal of Chromatography B-Biomedical Applications* **1996**, *677*, 1-28.
(117) Li, Y.; Lee, H. J.; Corn, R. M. *Nucleic Acids Research* **2006**, *34*, 6416-6424.
(118) Chang, C. C.; Wu, J. Y.; Chien, C. W.; Wu, W. S.; Liu, H.; Kang, C. C.; Yu, L. J.; Chang, T. C. *Analytical Chemistry* **2003**, *75*, 6177-6183.
(119) Gokulrangan, G.; Unruh, J. R.; Holub, D. F.; Ingram, B.; Johnson, C. K.; Wilson, G. S. *Analytical Chemistry* **2005**, *77*, 1963-1970.
(120) Peterson, K. J.; Sadowsky, J. D.; Scheef, E. A.; Pal, S.; Kourentzi, K. D.; Willson, R. C.; Bresnick, E. H.; Sheibani, N.; Gellman, S. H. *Analytical Biochemistry* **2008**, *378*, 8-14.
(121) del Toro, M.; Gargallo, R.; Eritja, R.; Jaumot, J. *Analytical Biochemistry* **2008**, *379*, 8-15.
(122) Tsuji, S.; Tanaka, T.; Hirabayashi, N.; Kato, S.; Akitomi, J.; Egashira, H.; Waga, I.; Ohtsu, T. *Biochemical and Biophysical Research Communications* **2009**, *386*, 227-231.
(123) Muller, M.; Weigand, J. E.; Weichenrieder, O.; Suess, B. *Nucleic Acids Research* **2006**, *34*, 2607-2617.
(124) Ryan, P. C.; Lu, M.; Draper, D. E. *Journal of Molecular Biology* **1991**, *221*, 1257-1268.
(125) Carey, J.; Cameron, V.; Dehaseth, P. L.; Uhlenbeck, O. C. *Biochemistry* **1983**, *22*, 2601-2610.
(126) Wong, I.; Lohman, T. M. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90*, 5428-5432.

(127) Oehler, S.; Alex, R.; Barker, A. *Analytical Biochemistry* **1999**, *268*, 330-336.
(128) Hall, K. K. J. In *RNA-Protein Interaction Protocols: Nitrocellulsoe Filter Binding for Determination of Dissociation Constants*; Humana Press, 1999.
(129) Christopoulos, T. K.; Diamandis, E. P.; Wilson, G. *Nucleic Acids Research* **1991**, *19*, 6015-6019.
(130) Czerwinski, J. D.; Hovan, S. C.; Mascotti, D. P. *Analytical Biochemistry* **2005**, *336*, 300-304.
(131) Busch, M. H. A.; Carels, L. B.; Boelens, H. F. M.; Kraak, J. C.; Poppe, H. *Journal of Chromatography A* **1997**, *777*, 311-328.
(132) Heegaard, N. H. H.; Nilsson, S.; Guzman, N. A. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **1998**, *715*, 29-54.
(133) Tanaka, Y.; Terabe, S. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002**, *768*, 81-92.
(134) He, X. Y.; Ding, Y. S.; Li, D. Z.; Lin, B. C. *Electrophoresis* **2004**, *25*, 697-711.
(135) Chen, Z.; Weber, S. G. *Trac-Trends in Analytical Chemistry* **2008**, *27*, 738-748.
(136) Erim, F. B.; Kraak, J. C. *Journal of Chromatography B* **1998**, *710*, 205-210.
(137) Mammen, M.; Gomez, F. A.; Whitesides, G. M. *Analytical Chemistry* **1995**, *67*, 3526-3535.
(138) Sebille, B.; Thuaud, N.; Tillement, J. P. *Journal of Chromatography* **1979**, *180*, 103-110.
(139) Busch, M. H. A.; Boelens, H. F. M.; Kraak, J. C.; Poppe, H. *Journal of Chromatography A* **1997**, *775*, 313-326.
(140) Busch, M. H. A.; Boelens, H. F. M.; Kraak, J. C.; Poppe, H.; Meekel, A. A. P.; Resmini, M. *Journal of Chromatography A* **1996**, *744*, 195-203.
(141) Sebille, B.; Zini, R.; Madjar, C. V.; Thuaud, N.; Tillement, J. P. *Journal of Chromatography-Biomedical Applications* **1990**, *531*, 51-77.
(142) Krylov, S. N. *Electrophoresis* **2007**, *28*, 69-88.
(143) Shimura, K.; Uchiyama, N.; Enomoto, M.; Matsumoto, H.; Kasai, K. *Analytical Chemistry* **2005**, *77*, 564-572.
(144) Shimura, K.; Waki, T.; Okada, M.; Toda, T.; Kimoto, I.; Kasai, K. *Electrophoresis* **2006**, *27*, 1886-1894.
(145) Fagerstam, L. G.; Frostellkarlsson, A.; Karlsson, R.; Persson, B.; Ronnberg, I. *Journal of Chromatography* **1992**, *597*, 397-410.
(146) Jonsson, U.; Fagerstam, L.; Ivarsson, B.; Johnsson, B.; Karlsson, R.; Lundh, K.; Lofas, S.; Persson, B.; Roos, H.; Ronnberg, I.; Sjolander, S.; Stenberg, E.; Stahlberg, R.; Urbaniczky, C.; Ostlin, H.; Malmqvist, M. *Biotechniques* **1991**, *11*, 620-&.
(147) Di Primo, C.; Lebars, I. *Analytical Biochemistry* **2007**, *368*, 148-155.
(148) Balamurugan, S.; Obubuafo, A.; Soper, S. A.; McCarley, R. L.; Spivak, D. A. *Langmuir* **2006**, *22*, 6446-6453.

(149) Tang, Q. J.; Su, X. D.; Loh, K. P. *Journal of Colloid and Interface Science* **2007**, *315*, 99-106.
(150) Wang, J. L.; Lv, R. J.; Xu, J. J.; Xu, D. K.; Chen, H. Y. *Analytical and Bioanalytical Chemistry* **2008**, *390*, 1059-1065.
(151) Schuck, P.; Minton, A. P. *Trends in Biochemical Sciences* **1996**, *21*, 458-460.
(152) Schuck, P.; Minton, A. P. *Analytical Biochemistry* **1996**, *240*, 262-272.
(153) Schuck, P. *Biophysical Journal* **1996**, *70*, 1230-1249.
(154) Anderson, B. J.; Larkin, C.; Guja, K.; Schildbach, J. F. *Fluorescence Spectroscopy* **2008**, 253.
(155) Heyduk, T.; Lee, J. C. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *87*, 1744-1748.
(156) Haralambidis, J.; Duncan, L.; Angus, K.; Tregear, G. W. *Nucleic Acids Research* **1990**, *18*, 493-499.
(157) Haralambidis, J.; Angus, K.; Pownall, S.; Duncan, L.; Chai, M.; Tregear, G. W. *Nucleic Acids Research* **1990**, *18*, 501-505.
(158) Allen, D. J.; Benkovic, S. J. *Biochemistry* **1989**, *28*, 9586-9593.
(159) Jameson, D. M.; Sawyer, W. H. *Biochemical Spectroscopy* **1995**, *246*, 283-300.
(160) Nag, A.; Bhattacharyya, K. *Journal of Photochemistry and Photobiology a-Chemistry* **1989**, *47*, 97-102.
(161) Nakayama, K.; Endo, M.; Fujitsuka, M.; Majima, T. *Journal of Physical Chemistry B* **2006**, *110*, 21311-21318.
(162) Livak, K. J.; Flood, S. J. A.; Marmaro, J.; Giusti, W.; Deetz, K. *Pcr-Methods and Applications* **1995**, *4*, 357-362.
(163) Wei, A. P.; Herron, J. N. *Analytical Chemistry* **1993**, *65*, 3372-3377.
(164) Drees, B. L.; Rye, H. S.; Glazer, A. N.; Nelson, H. C. M. *Journal of Biological Chemistry* **1996**, *271*, 32168-32173.
(165) McAfee, J. G.; Edmondson, S. P.; Zegar, I.; Shriver, J. W. *Biochemistry* **1996**, *35*, 4034-4045.
(166) Rachofsky, E. L.; Osman, R.; Ross, J. B. A. *Biochemistry* **2001**, *40*, 946-956.
(167) Turgeon, R. T.; Fonslow, B. R.; Jing, M.; Bowser, M. T. *Analytical Chemistry* **2010**, *82*, 3636-3641.
(168) Tan, W. H.; Wang, K. M.; Drake, T. J. *Current Opinion in Chemical Biology* **2004**, *8*, 547-553.
(169) Lee, W.; Obubuafo, A.; Lee, Y. I.; Davis, L. M.; Soper, S. A. *Journal of Fluorescence*, *20*, 203-213.
(170) Endoh, T.; Funabashi, H.; Mie, M.; Kobatake, E. *Analytical Chemistry* **2005**, *77*, 4308-4314.
(171) Juskowiak, B. *Analytica Chimica Acta* **2006**, *568*, 171-180.
(172) Perrin, F. *Journal de Physique et le Radium* **1926**, *7*, 390.
(173) Chen, R. F.; Scott, C. H. *Analytical Letters Part a-Chemical Analysis* **1985**, *18*, 393-421.

(174) Heyduk, T.; Ma, Y. X.; Tang, H.; Ebright, R. H. *RNA Polymerase and Associated Factors (Pt. B)* **1996**, 495.
(175) Lakowicz, J. R.; Gryczynski, I.; Gryczynski, Z.; Dattelbaum, J. D. *Analytical Biochemistry* **1999**, *267*, 397-405.
(176) Josic, D.; Clifton, J. G.; Kovac, S.; Hixson, D. C. *Current Opinion in Molecular Therapeutics* **2008**, *10*, 116-123.
(177) Sanders, C. R.; Myers, J. K. *Annual Review of Biophysics and Biomolecular Structure* **2004**, *33*, 25-51.
(178) Sanders, C. R.; Nagy, J. K. *Current Opinion in Structural Biology* **2000**, *10*, 438-442.
(179) Kirkbride, K. C.; Ray, B. N.; Blobe, G. C. *Trends in Biochemical Sciences* **2005**, *30*, 611-621.
(180) Wu, C. C.; Yates, J. R. *Nature Biotechnology* **2003**, *21*, 262-267.
(181) Carter, P. J.; Senter, P. D. *Cancer Journal* **2008**, *14*, 154-169.
(182) Law, C. L.; Gordon, K. A.; Toki, B. E.; Yamane, A. K.; Hering, M. A.; Cerveny, C. G.; Petroziello, J. M.; Ryan, M. C.; Smith, L.; Simon, R.; Sauter, G.; Oflazoglu, E.; Doronina, S. O.; Meyer, D. L.; Francisco, J. A.; Carter, P.; Senter, P. D.; Copland, J. A.; Wood, C. G.; Wahl, A. F. *Cancer Research* **2006**, *66*, 2328-2337.
(183) Huang, Y. F.; Shangguan, D. H.; Liu, H. P.; Phillips, J. A.; Zhang, X. L.; Chen, Y.; Tan, W. H. *Chembiochem* **2009**, *10*, 862-868.
(184) Chen, C. H. B.; Chernis, G. A.; Hoang, V. Q.; Landgraf, R. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 9226-9231.
(185) Barfod, A.; Persson, T.; Lindh, J. *Parasitology Research* **2009**, *105*, 1557-1566.
(186) Santulli-Marotto, S.; Nair, S. K.; Rusconi, C.; Sullenger, B.; Gilboa, E. *Cancer Research* **2003**, *63*, 7483-7489.
(187) Dassie, J. P.; Liu, X. Y.; Thomas, G. S.; Whitaker, R. M.; Thiel, K. W.; Stockdale, K. R.; Meyerholz, D. K.; McCaffrey, A. P.; McNamara, J. O.; Giangrande, P. H. *Nature Biotechnology* **2009**, *27*, 839-U895.
(188) Mori, T.; Oguro, A.; Ohtsu, T.; Nakamura, Y. *Nucleic Acids Research* **2004**, *32*, 6120-6128.
(189) Chu, T. C.; Twu, K. Y.; Ellington, A. D.; Levy, M. *Nucleic Acids Research* **2006**, *34*.
(190) Ohuchi, S. P.; Ohtsu, T.; Nakamura, Y. *Biochimie* **2006**, *88*, 897-904.
(191) Cerchia, L.; Duconge, F.; Pestourie, C.; Boulay, J.; Aissouni, Y.; Gombert, K.; Tavitian, B.; de Franciscis, V.; Libri, D. *Plos Biology* **2005**, *3*, 697-704.
(192) Chen, F.; Hu, Y. L.; Li, D. Q.; Chen, H. D.; Zhang, X. L. *Plos One* **2009**, *4*.
(193) Morris, K. N.; Jensen, K. B.; Julin, C. M.; Weil, M.; Gold, L. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 2902-2907.
(194) Ulrich, H.; Ippolito, J. E.; Pagan, O. R.; Eterovic, V. A.; Hann, R. M.; Shi, H.; Lis, J. T.; Eldefrawi, M. E.; Hess, G. P. *Proceedings of the National*

Academy of Sciences of the United States of America **1998**, *95*, 14051-14056.
(195) Huang, Z.; Pei, W.; Jayaseelan, S.; Shi, H.; Niu, L. *Biochemistry* **2007**, *46*, 12648-12655.
(196) Wang, C. L.; Zhang, M.; Yang, G. A.; Zhang, D. J.; Ding, H. M.; Wang, H. X.; Fan, M.; Shen, B. F.; Shao, N. S. *Journal of Biotechnology* **2003**, *102*, 15-22.
(197) Daniels, D. A.; Chen, H.; Hicke, B. J.; Swiderek, K. M.; Gold, L. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 15416-15421.
(198) Shangguan, D. H.; Cao, Z. H. C.; Li, Y.; Tan, W. H. *Clinical Chemistry* **2007**, *53*, 1153-1155.
(199) Chen, H. W.; Medley, C. D.; Sefah, K.; Shangguan, D.; Tang, Z. W.; Meng, L.; Smith, J. E.; Tan, W. H. *Chemmedchem* **2008**, *3*, 991-1001.
(200) Li, S. H.; Xu, H.; Ding, H. M.; Huang, Y. P.; Cao, X. X.; Yang, G.; Li, J.; Xie, Z. G.; Meng, Y. H.; Li, X. B.; Zhao, Q.; Shen, B. F.; Shao, N. S. *Journal of Pathology* **2009**, *218*, 327-336.
(201) Cui, Y.; Urich, H.; Hess, G. P. *Journal of Membrane Biology* **2004**, *202*, 137-149.
(202) Taghdisi, S. M.; Abnous, K.; Mosaffa, F.; Behravan, J. *Journal of Drug Targeting* **2010**, *18*, 277-281.
(203) Chen, Y.; Munteanu, A. C.; Huang, Y. F.; Phillips, J.; Zhu, Z.; Mavros, M.; Tan, W. H. *Chemistry-a European Journal* **2009**, *15*, 5327-5336.
(204) Smith, J. E.; Medley, C. D.; Tang, Z. W.; Shangguan, D.; Lofton, C.; Tan, W. H. *Analytical Chemistry* **2007**, *79*, 3075-3082.
(205) Mallikaratchy, P.; Tang, Z. W.; Kwame, S.; Meng, L.; Shangguan, D. H.; Tan, W. H. *Molecular & Cellular Proteomics* **2007**, *6*, 2230-2238.
(206) Raddatz, M. S. L.; Dolf, A.; Endl, E.; Knolle, P.; Famulok, M.; Mayer, G. *Angewandte Chemie-International Edition* **2008**, *47*, 5190-5193.
(207) Pestourie, C.; Cerchia, L.; Gombert, K.; Aissouni, Y.; Boulay, J.; De Franciscis, V.; Libri, D.; Tavitian, B.; Duconge, F. *Oligonucleotides* **2006**, *16*, 323-335.
(208) Bers, D. M. *Nature* **2002**, *415*, 198-205.
(209) Bassani, J. W. M.; Bassani, R. A.; Bers, D. M. *Journal of Physiology-London* **1994**, *476*, 279-293.
(210) Ebashi, S.; Lipmann, F. *Journal of Cell Biology* **1962**, *14*, 389-400.
(211) Toyoshima, C.; Nomura, H. *Nature* **2002**, *418*, 605-611.
(212) MacLennan, D. H.; Rice, W. J.; Green, N. M. *Journal of Biological Chemistry* **1997**, *272*, 28815-28818.
(213) Kranias, E. G.; Bers, D. M. *Sub-cellular biochemistry* **2007**, *45*, 523-537.
(214) Brini, M.; Carafoli, E. *Physiological Reviews* **2009**, *89*, 1341-1378.
(215) Toyofuku, T.; Kurzydlowski, K.; Tada, M.; Maclennan, D. H. *Journal of Biological Chemistry* **1994**, *269*, 3088-3094.
(216) Kimura, Y.; Kurzydlowski, K.; Tada, M.; MacLennan, D. H. *Journal of Biological Chemistry* **1996**, *271*, 21726-21731.

(217) Cantilina, T.; Sagara, Y.; Inesi, G.; Jones, L. R. *Journal of Biological Chemistry* **1993**, *268*, 17018-17025.
(218) Irvine, D.; Tuerk, C.; Gold, L. *Journal of Molecular Biology* **1991**, *222*, 739-761.
(219) Sun, F. Z.; Galas, D.; Waterman, M. S. *Journal of Molecular Biology* **1996**, *258*, 650-660.
(220) Djordjevic, M.; Sengupta, A. M. *Physical Biology* **2006**, *3*, 13-28.
(221) Levine, H. A.; Nilsen-Hamilton, M. *Computational Biology and Chemistry* **2007**, *31*, 11-35.
(222) Shendure, J.; Ji, H. L. *Nature Biotechnology* **2008**, *26*, 1135-1145.
(223) Metzker, M. L. *Nature Reviews Genetics* **2010**, *11*, 31-46.
(224) Li, R. Q.; Zhu, H. M.; Ruan, J.; Qian, W. B.; Fang, X. D.; Shi, Z. B.; Li, Y. R.; Li, S. T.; Shan, G.; Kristiansen, K.; Li, S. G.; Yang, H. M.; Wang, J. *Genome Research* **2010**, *20*, 265-272.
(225) Cloonan, N.; Forrest, A. R. R.; Kolle, G.; Gardiner, B. B. A.; Faulkner, G. J.; Brown, M. K.; Taylor, D. F.; Steptoe, A. L.; Wani, S.; Bethel, G.; Robertson, A. J.; Perkins, A. C.; Bruce, S. J.; Lee, C. C.; Ranade, S. S.; Peckham, H. E.; Manning, J. M.; McKernan, K. J.; Grimmond, S. M. *Nature Methods* **2008**, *5*, 613-619.
(226) Petrosino, J. F.; Highlander, S.; Luna, R. A.; Gibbs, R. A.; Versalovic, J. *Clinical Chemistry* **2009**, *55*, 856-866.
(227) Zimmermann, B.; Gesell, T.; Chen, D.; Lorenz, C.; Schroeder, R. *Plos One* **2010**, *5*, e9169.
(228) Bayrac, A. T.; Sefah, K.; Parekh, P.; Bayrac, C.; Gulbakan, B.; Oktem, H. A.; Tan, W. H. *Acs Chemical Neuroscience* **2011**, *2*, 175-181.
(229) Cho, M.; Xiao, Y.; Nie, J.; Stewart, R.; Csordas, A. T.; Oh, S. S.; Thomson, J. A.; Soh, H. T. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 15373-15378.
(230) Huse, S. M.; Huber, J. A.; Morrison, H. G.; Sogin, M. L.; Mark Welch, D. *Genome Biology* **2007**, *8*.
(231) Ronaghi, M. *Genome Research* **2001**, *11*, 3-11.
(232) Kim, K. J.; Li, B.; Winer, J.; Armanini, M.; Gillett, N.; Phillips, H. S.; Ferrara, N. *Nature* **1993**, *362*, 841-844.
(233) Leung, D. W.; Cachianes, G.; Kuang, W. J.; Goeddel, D. V.; Ferrara, N. *Science* **1989**, *246*, 1306-1309.
(234) Myoken, Y.; Kayada, Y.; Okamoto, T.; Kan, M.; Sato, G. H.; Sato, J. D. *Proceedings of the National Academy of Sciences of the United States of America* **1991**, *88*, 5819-5823.
(235) Rosen, L.; Gordon, D.; Kalman, L. A.; Kaminski, M.; Munoz, M.; Ballester, O.; Pawl, L. E.; Yunus, F.; Chen, B.; Seaman, J. *Breast Cancer Research and Treatment* **2001**, *69*, 305-305.
(236) Ruckman, J.; Green, L. S.; Beeson, J.; Waugh, S.; Gillette, W. L.; Henninger, D. D.; Claesson-Welsh, L.; Janjic, N. *Journal of Biological Chemistry* **1998**, *273*, 20556-20567.

(237) Jellinek, D.; Green, L. S.; Bell, C.; Janjic, N. *Biochemistry* **1994**, *33*, 10450-10456.
(238) Green, L. S.; Jellinek, D.; Bell, C.; Beebe, L. A.; Feistner, B. D.; Gill, S. C.; Jucker, F. M.; Janjic, N. *Chemistry & Biology* **1996**, *3*, 960-960.
(239) Burmeister, P. E.; Lewis, S. D.; Silva, R. F.; Preiss, J. R.; Horwitz, L. R.; Pendergrast, P. S.; McCauley, T. G.; Kurz, J. C.; Epstein, D. M.; Wilson, C.; Keefe, A. D. *Chemistry & Biology* **2005**, *12*, 25-33.
(240) Hasegawa, H.; Sode, K.; Ikebukuro, K. *Biotechnology Letters* **2008**, *30*, 829-834.
(241) Nonaka, Y.; Sode, K.; Ikebukuro, K. *Molecules* **2010**, *15*, 215-225.
(242) Potty, A. S. R.; Kourentzi, K.; Fang, H.; Jackson, G. W.; Zhang, X.; Legge, G. B.; Willson, R. C. *Biopolymers* **2009**, *91*, 145-156.
(243) Taylor, J. N.; Darugar, Q.; Kourentzi, K.; Willson, R. C.; Landes, C. F. *Biochemical and Biophysical Research Communications* **2008**, *373*, 213-218.
(244) Rose, C. M.; Hayes, M. J.; Stettler, G. R.; Hickey, S. F.; Axelrod, T. M.; Giustini, N. P.; Suljak, S. W. *Analyst* **2010**, *135*, 2945-2951.
(245) Bailey, T. L.; Elkan, C. *Machine Learning* **1995**, *21*, 51-80.
(246) Bailey, T. L.; Williams, N.; Misleh, C.; Li, W. W. *Nucleic Acids Research* **2006**, *34*, W369-W373.
(247) Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. *Nucleic Acids Research* **2003**, *31*, 3497-3500.
(248) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Research* **1994**, *22*, 4673-4680.
(249) Hall, K. B., Kranz, J.K., ; Humana Press: Totowa, New Jersy, 1999; Vol. 118, pp 109-110.
(250) Jing, M.; Bowser, M. T. *Analytica Chimica Acta* **2011**, *686*, 9-18.
(251) Jing, M.; Bowser, M. T. *Lab on a Chip* **2011**, *11*, 3703-3709.
(252) Gopinath, S. C. B. *Analytical and Bioanalytical Chemistry* **2007**, *387*, 171-182.
(253) Xu, Y. H.; Yang, X. R.; Wang, E. K. *Analytica Chimica Acta* **2010**, *683*, 12-20.
(254) Huang, C. J.; Lin, H. I.; Shiesh, S. C.; Lee, G. B. *Biosensors & Bioelectronics* **2010**, *25*, 1761-1766.
(255) Xu, Y.; Zhang, C. X.; Janasek, D.; Manz, A. *Lab on a Chip* **2003**, *3*, 224-227.
(256) Frost, N. W.; Bowser, M. T. *Lab on a Chip* **2010**, *10*, 1231-1236.
(257) Fonslow, B. R.; Bowser, M. T. *Analytical Chemistry* **2006**, *78*, 8236-8244.
(258) Turgeon, R. T.; Bowser, M. T. *Electrophoresis* **2009**, *30*, 1342-1348.
(259) Wiegand, T. W.; Williams, P. B.; Dreskin, S. C.; Jouvin, M. H.; Kinet, J. P.; Tasset, D. *Journal of Immunology* **1996**, *157*, 221-230.
(260) Iki, N.; Yeung, E. S. *Journal of Chromatography A* **1996**, *731*, 273-282.
(261) Shackman, J. G.; Watson, C. J.; Kennedy, R. T. *Journal of Chromatography A* **2004**, *1040*, 273-282.
(262) Hall, K. B., Kranz, J.K.,, Ed.; Humana Press: Totowa, New Jersy, 1999.

(263) Bird, R. B.; Stewart, W. E.; Lightfoot, E. N., Eds. *Transport Phenomena*, 2 ed.; John Wiley and Sons: New York, 2002.
(264) Yang, P.; Ma, Y.; Lee, A. W. M.; Kennedy, R. T. *Electrophoresis* **2009**, *30*, 457-464.
(265) Buchanan, D. D.; Jameson, E. E.; Perlette, J.; Malik, A.; Kennedy, R. T. *Electrophoresis* **2003**, *24*, 1375-1382.
(266) Dash, R.; Frank, K. F.; Carr, A. N.; Moravec, C. S.; Kranias, E. G. *Journal of Molecular and Cellular Cardiology* **2001**, *33*, 1345-1353.
(267) Hasenfuss, G.; Reinecke, H.; Studer, R.; Meyer, M.; Pieske, B.; Holtz, J.; Holubarsch, C.; Posival, H.; Just, H.; Drexler, H. *Circulation Research* **1994**, *75*, 434-442.
(268) Meyer, M.; Schillinger, W.; Pieske, B.; Holubarsch, C.; Heilmann, C.; Posival, H.; Kuwajima, G.; Mikoshiba, K.; Just, H.; Hasenfuss, G. *Circulation* **1995**, *92*, 778-784.
(269) Schwinger, R. H. G.; Munch, G.; Bolck, B.; Karczewski, P.; Krause, E. G.; Erdmann, E. *Journal of Molecular and Cellular Cardiology* **1999**, *31*, 479-491.
(270) Giordano, F. J.; He, H. P.; McDonough, P.; Meyer, M.; Sayen, M. R.; Dillmann, W. H. *Circulation* **1997**, *96*, 400-403.
(271) del Monte, F.; Williams, E.; Lebeche, D.; Schmidt, U.; Rosenzweig, A.; Gwathmey, J. K.; Lewandowski, E. D.; Hajjar, R. J. *Circulation* **2001**, *104*, 1424-1429.
(272) He, H. P.; Giordano, F. J.; HilalDandan, R.; Choi, D. J.; Rockman, H. A.; McDonough, P. M.; Bluhm, W. F.; Meyer, M.; Sayen, M. R.; Swanson, E.; Dillmann, W. H. *Journal of Clinical Investigation* **1997**, *100*, 380-389.
(273) Eizema, K.; Fechner, H.; Bezstarosti, K.; Schneider-Rasp, S.; van der Laarse, A.; Wang, H.; Schultheiss, H. P.; Poller, W. C.; Lamers, J. M. J. *Circulation* **2000**, *101*, 2193-2199.
(274) del Monte, F.; Harding, S. E.; Dec, G. W.; Gwathmey, J. K.; Hajjar, R. J. *Circulation* **2002**, *105*, 904-907.
(275) Hoshijima, M.; Ikeda, Y.; Iwanaga, Y.; Minamisawa, S.; Date, M. O.; Gu, Y. S.; Iwatate, M.; Li, M. X.; Wang, L. L.; Wilson, J. M.; Wang, Y. B.; Ross, J.; Chien, K. R. *Nature Medicine* **2002**, *8*, 864-871.
(276) Dua, P.; Kim, S.; Lee, D. K. *Methods* **2011**, *54*, 215-225.
(277) Tanaka, Y.; Honda, T.; Matsuura, K.; Kimura, Y.; Inui, M. *Journal of Pharmacology and Experimental Therapeutics* **2009**, *329*, 57-63.
(278) Danino, D.; Talmon, Y.; Zana, R. *Journal of Colloid and Interface Science* **1997**, *186*, 170-179.
(279) Brandl, C. J.; Green, N. M.; Korczak, B.; Maclennan, D. H. *Cell* **1986**, *44*, 597-607.
(280) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912.
(281) Sridhar, R.; Hanson-Painton, O.; Cooper, D. R. *Pharmaceutical Research* **2000**, *17*, 1345-1353.

(282) Dancey, J.; Sausville, E. A. *Nature Reviews Drug Discovery* **2003**, *2*, 296-313.
(283) Cohen, P. *European Journal of Biochemistry* **2001**, *268*, 5001-5010.
(284) Park, Y. W.; Cummings, R. T.; Wu, L.; Zheng, S.; Cameron, P. M.; Woods, A.; Zaller, D. M.; Marcy, A. I.; Hermes, J. D. *Analytical Biochemistry* **1999**, *269*, 94-104.
(285) Li, Y.; Cummings, R. T.; Cunningham, B. R.; Chen, Y. L.; Zhou, G. C. *Analytical Biochemistry* **2003**, *321*, 151-156.
(286) Turek-Etienne, T. C.; Lei, M.; Terracciano, J. S.; Langsdorf, E. F.; Bryant, R. W.; Hart, R. F.; Horan, A. C. *Journal of Biomolecular Screening* **2004**, *9*, 52-61.
(287) Loomans, E.; van Doommalen, A. M.; Wat, J. W. Y.; Zaman, G. J. R. *Assay and Drug Development Technologies* **2003**, *1*, 445-453.
(288) Kupcho, K.; Somberg, R.; Bulleit, B.; Goueli, S. A. *Analytical Biochemistry* **2003**, *317*, 210-217.
(289) Pommereau, A.; Pap, E.; Kannt, A. *Journal of Biomolecular Screening* **2004**, *9*, 409-416.
(290) Charter, N. W.; Kauffman, L.; Singh, R.; Eglen, R. M. *Journal of Biomolecular Screening* **2006**, *11*, 390-399.
(291) Lundin, A.; Eriksson, J. *Assay and Drug Development Technologies* **2008**, *6*, 531-541.
(292) Anderson, K. W.; Murphy, A. J. *Journal of Biological Chemistry* **1983**, *258*, 4276-4278.
(293) Coll, R. J.; Murphy, A. J. *Journal of Biological Chemistry* **1984**, *259*, 4249-4254.
(294) Shivanna, B. D.; Rowe, E. S. *Biochemical Journal* **1997**, *325*, 533-542.
(295) Karim, C. B.; Paterlini, M. G.; Reddy, L. G.; Hunter, G. W.; Barany, G.; Thomas, D. D. *Journal of Biological Chemistry* **2001**, *276*, 38814-38819.
(296) Reddy, L. G.; Cornea, R. L.; Winters, D. L.; McKenna, E.; Thomas, D. D. *Biochemistry* **2003**, *42*, 4585-4592.
(297) Gustavsson, M.; Traaseth, N. J.; Veglia, G. *Biochemistry* **2011**, *50*, 10367-10374.
(298) Kiianitsa, K.; Solinger, J. A.; Heyer, W. D. *Analytical Biochemistry* **2003**, *321*, 266-271.
(299) Chiku, T.; Pullela, P. K.; Sem, D. S. *Journal of Biomolecular Screening* **2006**, *11*, 844-853.
(300) Huss, K. L.; Blonigen, P. E.; Campbell, R. M. *Journal of Biomolecular Screening* **2007**, *12*, 578-584.
(301) Kleman-Leyer, K. M.; Klink, T. A.; Kopp, A. L.; Westermeyer, T. A.; Koeff, M. D.; Larson, B. R.; Worzella, T. J.; Pinchard, C. A.; van de Kar, S. A. T.; Zaman, G. J. R.; Hornberg, J. J.; Lowery, R. G. *Assay and Drug Development Technologies* **2009**, *7*, 56-67.
(302) Traaseth, N. J.; Thomas, D. D.; Veglia, G. *Journal of Molecular Biology* **2006**, *358*, 1041-1050.
(303) Barash, D.; Gabdank, I. *Rna Biology* **2010**, *7*, 90-97.

(304) Gilch, S.; Schatzl, H. M. *Cellular and Molecular Life Sciences* **2009**, *66*, 2445-2455.
(305) Afonso, J. L.; Clifton, M. J. *Electrophoresis* **1999**, *20*, 2801-2809.
(306) De Angelis, A. A.; Jones, D. H.; Grant, C. V.; Park, S. H.; Mesleh, M. F.; Opella, S. J. In *Nuclear Magnetic Resonance of Biological Macromolecules, Part C*; James, T. L., Ed., 2005; Vol. 394, pp 350-382.
(307) Opella, S. J.; Nevzorov, A.; Mesleh, M. F.; Marassi, F. M. *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* **2002**, *80*, 597-604.
(308) Sheng, Y. X.; Bowser, M. T. *Analyst* **2012**, *137*, 1144-1151.
(309) Kopp, M. U.; de Mello, A. J.; Manz, A. *Science* **1998**, *280*, 1046-1048.
(310) Marcus, J. S.; Anderson, W. F.; Quake, S. R. *Analytical Chemistry* **2006**, *78*, 956-958.
(311) McKenna, E.; Smith, J. S.; Coll, K. E.; Mazack, E. K.; Mayer, E. J.; Antanavage, J.; Wiedmann, R. T.; Johnson, R. G. *Journal of Biological Chemistry* **1996**, *271*, 24517-24525.
(312) Chiesi, M.; Schwaller, R. *Biochemical and Biophysical Research Communications* **1994**, *202*, 1668-1673.
(313) Xu, Z. C.; Kirchberger, M. A. *Journal of Biological Chemistry* **1989**, *264*, 16644-16651.
(314) Hughes, E.; Edwards, R.; Middleton, D. A. *Biochemical and Biophysical Research Communications*, *401*, 370-375.
(315) Chiesi, M.; Schwaller, R. *Febs Letters* **1989**, *244*, 241-244.