The optimization of cell-SELEX based aptamer selection through masking DNA, PCR and non-SELEX.

Natalia Obrecht

A THESIS SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

SEPTEMBER, 2013

ABSTRACT

Aptamers are short oligonucleotide sequences that are capable of binding with high affinity and specificity to a wide variety of targets. The selection of aptamers from a random oligonucleotide library, by cell-SELEX, has lead to their *in vivo* application in both research and clinical settings. Cell-SELEX continues to be hindered by the non-specific binding of selection sequences to the complex cell surface as well as low specificity in the PCR amplification of DNA. In this work the cell-SELEX procedure was optimized to improve the efficiency of aptamer selection using three modifications: masking DNA characterization, non-SELEX selection and touchdown PCR. An established masking DNA model was tested against the MCF-7 and 4T1 cell lines and was shown to be capable of determining the masking DNA concentration needed for aptamer selection. The traditional SELEX approach was combined with a non-SELEX aptamer selection protocol to reduce the concentration and heterogeneity of sequences collected after the first round of selection thus leading to more accurate PCR amplification and a decrease in byproduct formation. Lastly, touchdown PCR was used to successfully eliminate the amplification of genomic DNA reducing the formation of genomic DNA byproducts.

DEDICATION

To my husband Sigfrido Salubro

&

My parents Michael & Nadejda Obrecht

ACKNOWLEDGEMENTS

I would like to start off with a very big thank you to my supervisor Dr. Sergey N. Krylov for giving me the opportunity to experience the world of research as a graduate student within his laboratory. Thank you for your support and guidance over the past two years. I feel very lucky to have had such an inspiring supervisor who always lead by example and demonstrated the value of hard work and dedication.

Thank you so much to my fellow laboratory members you have made my graduate experience an unforgettable one. Thank you for welcoming me so openly into the group and for helping me along the way. As someone without any previous research experience I was always grateful to have such supportive colleagues who were always willing to teach, demonstrate and discuss new ideas. Thank you for being so patient with me as I assimilated into the research setting and as I worked through my many research challenges down to my defense presentation. I will always feel privileged to have worked next to such intelligent but also kind and amazing people who made my graduate school experience so much more amazing than I ever expected.

I would like to thank Dr. Robert Tsushima for acting as my committee member. Thank you for your support and dedication and for helping me stay on track with both my project and career goals. Thank you to Dr. Vivian Saridakis and Dr. Philip Johnson for acting as examiners during my thesis defense.

I would like to thank my husband Sigfrido Salubro for supporting me during this challenging but exciting time. Thank you for getting me through the hard moments and being there to celebrate the happy ones. Last, but not least, thank you to my parents for their never-ending love and support.

LIST OF ABBREVIATIONS

bp	Base pair
cDNA	Complimentary DNA
CE	Capillary electrophoresis
DMEM	Dulbecco's modified eagle's medium
dsDNA	Double stranded DNA
EDTA	Ethylenediaminetetraacetic acid
EOF	Electro-osmotic flow
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
HCMV	Human cytomegalo virus
HIV	Human immunodeficiency virus
mDNA	Masking DNA
nt	Nucleotide
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PI	Propidium iodide
RBC	Red blood cells
RFU	Relative fluorescence units
RSV	Rous sarcoma virus

RT-PCR	Reverse transcription polymerase chain reaction
SELEX	Systematic evolution of ligands by exponential enrich
ssDNA	Single stranded DNA
T_m	Melting temperature
UV	Ultraviolet
VEGF	Vascular endothelial growth factor

CONTENTS

LIST OF FIGURES

Chapter 1: APTAMER SELECTION AND APPLICATION

1.1 Evolution of Aptamer Development

1.1.1 Discovery of Aptamers and the SELEX Procedure

Aptamers are oligonucleotide sequences typically up to 100 nucleotides in length, which can be selected to bind with high affinity and specificity to a wide range of targets [1]. Two separate groups simultaneously introduced the concept of aptamers, and process through which they are selected. Ellington and Szostak were the first to coin the term aptamer, which was based on the Greek word "aptus" meaning to fit. In their work they were able to successfully select a variety of RNA sequences that were able to bind with high affinity and varying levels of specificity to a variety of organic dyes immobilized on a chromatography column [1].

The process began with the synthesis of a ssDNA library which was transcribed into RNA. Each RNA sequence consisted of a random region approximately 100 nucleotides in length, flanked by constant primer regions to allow for polymerase chain reaction (PCR) amplification and transcription of target bound sequences. The targets of aptamer selection were various organic dyes, including Cibarcron Blue 3G1, Reactive Red 120, Reactive Yellow 86, Reactive Brown 10, Reactive Green 19 and Reactive Blue 4. Chromatography columns were prepared by immobilizing agarose beads with covalently bound target dyes. The RNA library was then applied to each column to allow any potential aptamers to bind to the immobilized target dyes. The column was washed with a high salt concentration buffer to remove any unbound or loosely bound RNA sequences. Lastly, tightly bound sequences were eluted with water and transcribed into cDNA via reverse-transcription PCR (RT-PCR) to allow PCR

amplification by Taq polymerase. The resulting pool of DNA was then converted back into RNA and used for the subsequent round of selection. The affinity of the aptamers was tested by measuring the percentage of RNA bound to the immobilized target column versus the percentage that eluted. With every round of selection the amount of RNA that adhered to the target increased, maxing out at approximately 80% of sequences bound. Once a significant increase in affinity was evident, the pool of RNA was cloned and sequenced to identify aptamers. The initial library consisted of a total of 10^{15} unique RNA sequences. Following four rounds of selection approximately 10 sequences, some unique and some with common elements, that could bind with high affinity and variable specificity to the range of organic dyes [7].

Gold and Tuerk performed a similar selection process, against the target T4 DNA polymerase. Gold and Tuerk were the first to name the process of aptamer selection as systemic evolution of ligands by exponential enrichment or SELEX [8]. Gold and Tuerk wanted to identify RNA sequences that were capable of binding to the ribosome binding site of the bacteriophage T4 DNA polymerase, which is typically bound to a 5 base pair helix with a unique 8 nucleotide loop. Instead of having a large variable region, their library consisted of a mere 8 nucleotide variable loop with the remainder of the sequence being made up of a fixed 5 base-pair helix and the constant primer regions to allow for reverse transcription and amplification of bound sequences. Starting with just over 65 000 different sequences, Tuerk and Gold were able to identify two high affinity sequences that could bind to the T4 DNA polymerase ribosome-binding site, one of which was the wild type target and the other a previously unidentified aptamer [8].

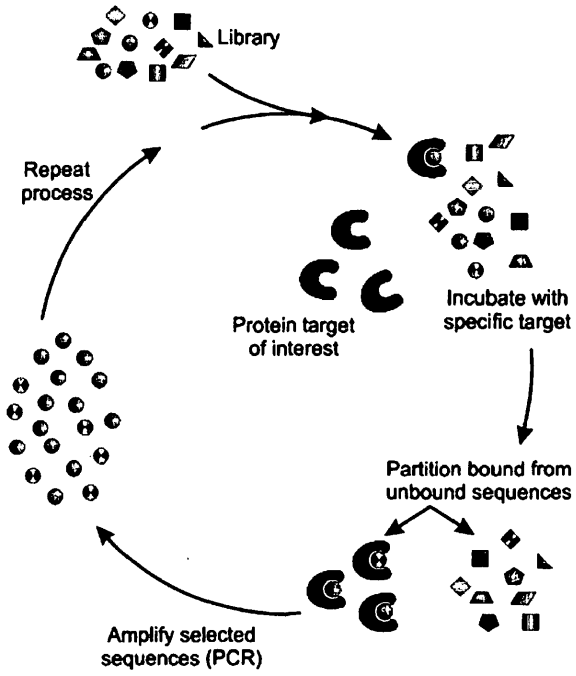

Figure 1.1 - An outline of the systemic evolution of ligands by exponential enrichment (SELEX) procedure used to identify aptamer against purified targets such as proteins. This figure was adapted from [3].

1.1.2 – The Overall SELEX Procedure

The SELEX process begins with the generation of a ssDNA or RNA library consisting of a random region, flanked by distinct primer regions used for the transcription and amplification of bound sequences between rounds. With the current advancement of technologies libraries can now be purchased commercially at a relatively low cost, consisting of various length of primer and randomized regions. Both ssDNA and RNA libraries have been successfully used for aptamer selection [1]. In general, ssDNA libraries are cheaper to work with and do not require the reverse transcription prior to amplification. On the other hand, RNA libraries are available

with a wider modification variety, which can be introduced both during the PCR amplification step of the SELEX procedure and following the sequencing of selected aptamers [1].

A broad outline of the SELEX procedure is present in **Figure 1.1** [3]. The process begins with the incubation of a randomly generated RNA or ssDNA library with the target of interest. The target of interest can be presented in a variety of ways including bound to agarose beads in suspension, purified protein in binding buffer, on a chromatography column or coated on a nitrocellulose filter [7, 8]. Sequences that bind tightly to the target are then separated from unbound and loosely bound sequences and are eluted from the target by thorough washing. The bound sequences are amplified via PCR and the resulting enriched pool is used for the subsequent round of selection [7, 8]. Since its initial introduction in the 1990s the SELEX process has evolved tremendously.

To date, numerous methods separating bound and unbound sequences have been introduced improving the efficiency of selection [9, 10]. Depending on the target of interest, specifically its ability to bind to oligonucleotides and its complexity, the SELEX procedures can take anywhere from 4 to 20 rounds of selection to reach a sufficiently high affinity aptamer pool [11]. In most cases, the final aptamer pool is then cloned and sequenced to identify the specific ssDNA or RNA sequences that bind with the highest affinity and specificity to the target of interest [7, 8]. Once specific sequences are obtained, their affinity for the target is determined by calculating the dissociation constant (K_d) which typically ranges in the low nM to pM range for strong binding aptamers [11].

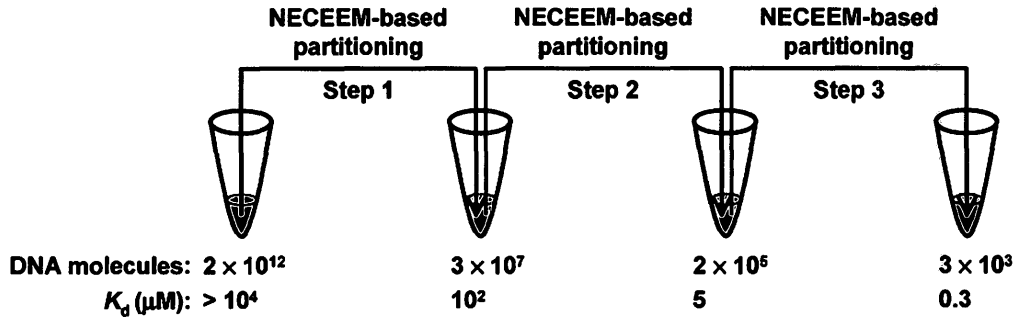

Figure 1.2 – Non-SELEX approach to aptamer selection against a purified product. The non-SELEX method consists of consecutive incubation and partitioning steps without PCR amplification of sequences in between [4]

1.1.3 Non-SELEX Selection of Aptamers

Previous work in our lab has shown that aptamer can also be selected using a modified non-SELEX approach [4]. In this approach the exponential enrichment of collected sequences is bypassed and consecutive incubations are performed (**Figure 1.2**). This provides the advantage of a faster aptamer selection protocol as well as a decreased consumption of reagents typically used at every round of PCR for sequence amplification and product clean-up. In addition this approach has the potential to be applied to libraries which are not easily amplifiable such as those composed of modified bases [4].

1.1.4 Evolution of SELEX targets

SELEX can theoretically be applied to an endless variety of targets ranging from simple small molecules all the way to membrane fragments, whole cells and whole tissue samples. The first two SELEX procedures were performed against purified targets, specifically organic dyes [8] and the bacteriophage protein T4 DNA polymerase [8]. Since then, SELEX targets have

expanded tremendously. Selecting against a purified target, such as a protein or small molecule, usually required few rounds, but can present the challenge of a having limited number of potential binding sites [1]. On the other hand, when selecting for unpurified targets the increased complexity is accompanied by the necessity for an increased number of rounds [6].

The selection of aptamers against unpurified targets was first done by Morris et. al. in 1997 against human red blood cell (RBC) ghosts [12]. RBC ghosts were created by lysing RBCs so that only the membrane remained, and all intracellular components were removed. Without having a specific epitope in mind Morris et. al. selected a heterogeneous pool of aptamers that were able to bind with high affinity and specificity to several proteins on the RBC membrane. Furthermore, by cross-linking the aptamers with their membrane targets the authors were able to confirm that individual sequences had unique targets [12].

Aptamers have been selected for a variety of viruses including human immunodeficiency virus (HIV) [13], human cytomegalovirus (HCMV) [14] and Rous sarcoma virus (RSV) [15]. Viruses were grown within a cell culture, collected and incubated with a starting SELEX library. Following anywhere up to 16 rounds of selection, aptamers were selected that, could not only bind to the target virus with high affinity and specificity, but that were also capable of reducing the virulence of viruses against the host cells [14, 15].

The first aptamer selected for a living target was performed against the protozoan pathogen *African trypanosomes* [16]. Previous attempts to select a functional antibody against the pathogen were unsuccessful since antibodies could only be selected against the quickly evolving main shell of the organism. On the other hand aptamers were successfully selected, following a mere ten rounds of SELEX, that were capable of binding the flagellar pocket of the

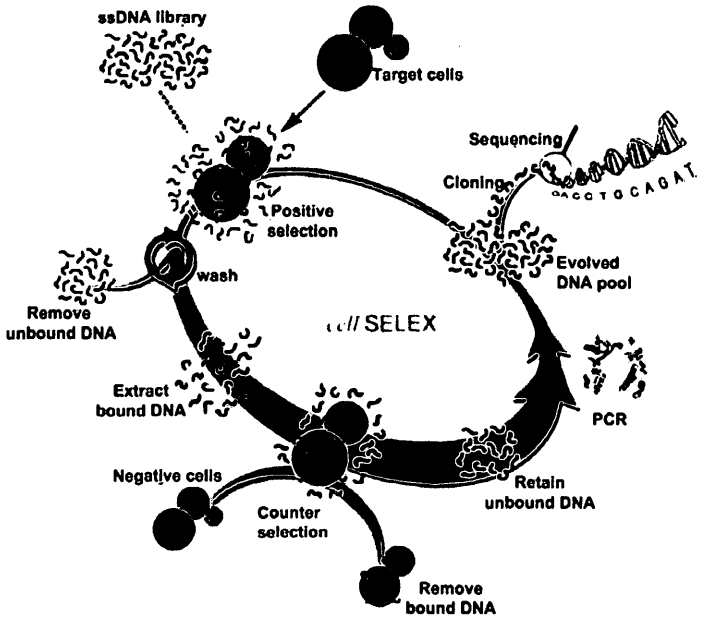

Figure 1.3 - Outline of the cell-SELEX procedure consisting of a counter selection using negative cells and a positive selection using target cells. This figure was adapted from [2].

organism that does not undergo mutations and is an ideal target for therapeutics. This demonstrated the fact that aptamers are so much smaller in size (10kDa) compared to antibodies (180kDa) allows them to bind to areas which are inaccessible to antibody selection [16].

Aptamer selection against eukaryotic cells was first performed by Blank et. al. in 2001 against mouse YPEN-1 endothelial cells. Due to the complexity of the eukaryotic cell surface 28 rounds of SELEX selection were necessary. The resulting aptamer was successfully used to differentiate between micro-vessels of rat glioblastoma and those of healthy rat brain [17]. Since then aptamers have been selected against a wide variety of cell lines, most commonly against cancer cells [1], through a modified procedure commonly referred to as cell-SELEX [6]. Most recently, segments of cancer tissue, embedded in paraffin, were used as a selection target

resulting in a heterogeneous pool of aptamers that could recognize biomarkers on the cell surface as well as within the extracellular matrix [17].

1.2 Cell-SELEX

1.2.1 Cell-SELEX procedure

Cell-SELEX is very similar to the original SELEX procedure with some additional steps to compensate for the increased complexity of the target. The diagram depicting the cell-SELEX procedure is presented in **Figure 1.3** [2]. Typically cell-SELEX consists of two different sets of selection: a positive selection against target cells and a negative or counter selection against non-target cells. Traditionally, positive selection is performed against a pathological tissue while the negative selection is performed using the healthy counterpart [6]. Positive selection commonly occurs against tumor cells resulting in aptamers that are capable of distinguishing between cancerous and non-cancerous tissue. However, other combinations have been used to select more specific aptamers such as those that distinguish between mature and immature dendritic cells [18].

Cell SELEX begins in a similar fashion to traditional SELEX with the incubation of target cells with a ssDNA or RNA library typically consisting of approximately 10^{15} unique sequences [6]. Selection conditions vary depending on the intended future application of the aptamer. The majority of aptamers selected via cell-SELEX are intended for *in vivo* applications, thus they should be selected for in physiological conditions (37°C and pH of 7-8). Previous work has shown that aptamers selected under non-physiological conditions have a decreased affinity at physiological conditions making them non-functional [19].

Following incubation, under the proper conditions, weakly bound and unbound sequences are separated from target bound sequences, commonly through centrifugation and multiple washes. The bound sequences are then eluted from the target cell by heat dissociation and amplified via PCR. If an RNA pool is used for selection, sequences are first converted to complementary DNA (cDNA) through reverse transcription PCR (RT-PCR) and then amplified by traditional PCR methods [6].

Subsequent rounds are preceded or followed by a round of negative selection. The cleaned-up PCR product from the previous round is incubated with the negative cells in identical conditions as with the positive selection incubation. In the negative round it is the unbound sequences that are collected and used for the subsequent round of positive selection. Negative selection reduces the amount of non-specific sequences collected during positive selection and ensures that the aptamers that are selected are unique to the pathogenic target and not to ubiquitous cellular epitopes [6].

The progress of selection during cell-SELEX is typically monitored via flow cytometry. DNA from each round is amplified using a fluorophore labeled primer such as fluorescein. The amplified pool of labeled sequences from each round is incubated with the target cell line. Unbound sequences are rinsed off, cells are re-suspended in binding buffer and analyzed via flow cytometry. Any sequences bound to the target cells will emit a fluorescent signal due to the addition of the fluorescent probe. This fluorescent signal is picked up by the flow cytometer and a histogram representing the number of events (i.e. cells) versus the relative fluorescence (given as relative fluorescence units or RFUs) emitted at a specific wavelength, is constructed. As selection progresses the number of events that result in a higher RFU signal will increase moving

the overall histogram right along the x-axis as shown in **Figure 1.4A**. This overall increase in the RFU signal indicates an increased number of sequences bound to target cell implying an increase in aptamer pool affinity to the target. At the same time, incubation of the selection pools with the negative or counter selection cell line should not show any increase in fluorescence intensity (**Figure 1.4B**). Selection typically continues until no significant increase is seen in RFUs from one round to the next [6].

1.2.2 Non-specific binding of sequences

The cellular surface is very complex, consisting of thousands of potential selection targets including proteins, glycoproteins and phospholipids [20]. While this means that numerous

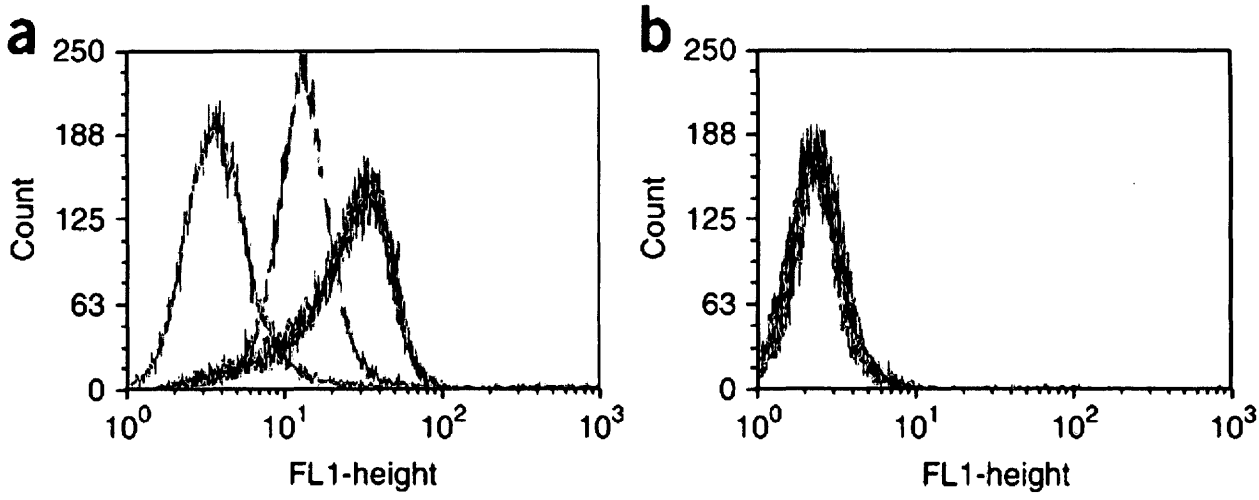

Figure 1.4 Flow cytometry histogram used to monitor the progress of aptamer selection. Both target (A) and negative cells (B) are incubated with sequence pools from several rounds and analyzed by flow cytometry creating a histogram of the number of events versus fluorescence signal. An improvement in the affinity is evident by the shift of the histogram peak right along the x-axis indicating an increase in the mean fluorescence of the cell sample analyzed. This figure was adapted from [6].

aptamers can be selected against a single cell type, it also creates the additional problem of a large amount of non-specific binding of oligonucleotides to the cell surface. This large amount of non-specifically bound sequences translates to an increase in the number of selection rounds necessary to obtain a high affinity aptamer pool. In addition, the collection of non-specific sequences along with potential aptamers means that the sequences will compete for PCR reagents resulting in a decreased number of high affinity sequences being amplified [6].

Non-specific sequences are also collected during selection due to dead cells taking up nucleic acids from the cell-SELEX suspension. Several studies have shown that dead cells become increasingly permeable to both RNA and ssDNA sequences and take up a large amount of sequences non-specifically [21]. In fact, the number of sequences collected from dead cells is exceedingly greater than the number of sequences that typically adhere to a living cell's membrane [10]. These sequences are then amplified along with the high affinity binders resulting in a decreased efficiency in the selection of high-affinity aptamers. Furthermore, the non-specific uptake of sequences by dead cells during negative selection may result in a loss of potential aptamers from the selection pool. If negative selection cells take up potential aptamer sequences non-specifically, these sequences will not be available for the subsequent round of positive selection [21].

To minimize the effect that dead cells may have on the selection process the SELEX cellular suspension is typically evaluated for the presence of dead cells prior to incubation with the library [6]. The number of dead cells can be quickly estimated using a propidium iodide (PI) stain and flow cytometry analysis. Dead cells will pick up the PI stain resulting in a fluorescent signal around 617 nm, while the living cell membrane is impermeable to PI [22]. The general

recommendation is that no more than 5% of dead cells should be present within the SELEX selection suspension. This stringency in dead cell count applies to both the positive and negative selection steps of the cell-SELEX procedure [20].

One approach for dealing with the high amount of uptake of non-specific sequence by dead cells is the use of fluorescence-activated cell sorting (FACS). This method relies on the separation of living and dead cells based on their light scattering characteristics. Following the incubation of the SELEX library with target cells, cells are sorted based on their light scattering characteristics using FACS. The living cells are then collected, and tightly bound sequences from these cells are eluted and amplified for future rounds. While this method is appealing, it still presents the challenge of potential aptamers being lost due to uptake by dead cells prior to sorting and sequences lost during the negative selection portion of SELEX. In addition, FACS requires expensive equipment, which is not required for traditional cell SELEX methods [10].

1.2.2.1 Masking DNA

To overcome the challenges of non-specific binding of sequences to the complex cell surface, as well as non-specific uptake of DNA by dead cells, masking DNA (mDNA) is used in the form of primer-free single stranded DNA (ssDNA) [18], yeast tRNA [17], salmon sperm DNA [10, 23], or genomic DNA[24]. Unlike the sequences present within the SELEX library, masking DNA lacks primer regions and as a result will not be amplified during the PCR amplification of sequences collected at the end of a round of selection [5].

Masking DNA present within the SELEX suspension will compete with low affinity binders and nonspecific binders for targets while allowing high affinity sequences to bind with

their unique targets. This results in and increased stringency in the selection process leading to improved efficiency demonstrated by a decrease number of necessary SELEX rounds. In addition the masking DNA in solution will act as a substrate for any nuclease enzymes and decrease the number of library sequences being digested [5].

To allow for maximum benefit of masking DNA present within the SELEX suspension, it must be as similar to the selection library as possible. This similarity increases the ability of the masking DNA to outcompete the library at non-specific sites. The optimal masking DNA is than an oligonucleotide sequence identical in length to the random region of the selection library without primer portions. It is selected at random, contains a 1:1 CG to AT ratio and cannot bind to the library primer regions to eliminate amplification during PCR [5].

Previous cell SELEX procedures in our lab, as well as others, have chosen the masking DNA concentration arbitrarily. To date no model existed that could be used to characterize the interaction of masking DNA and library sequences with target cells. Typically, masking DNA is present in significant excess to the library concentration. The concentrations of masking DNA vary anywhere from 5 to 100 times that of the library [10, 17, 18, 23, 24] .

1.2.3 PCR amplification

Once target bound sequences are collected they need to be amplified by PCR and prepared for the following round of selection [1]. PCR is a process, which allows for the exponential amplification of nucleic acids. In brief, PCR consists of three repetitive steps: denaturation of the target sequence, annealing of the primer regions to the single stranded

denatured template and the extension of a nascent strand starting from the double stranded primer site [25].

DNA polymerases cannot synthesize a complimentary strand purely from a single stranded template but instead requires a double stranded segment of DNA [25]. *In vivo* these double-stranded initiation points are provided by the creation of Okazaki fragments [26]. During *in vitro* PCR amplification, primer regions function as Okazaki fragments allowing for transcription to occur. Primers, short sequences complimentary to the terminal ends of the template, are added to the PCR mixture and hybridize with the terminal ends of template sequences providing the start-off point for new strand synthesis [25].

The accurate annealing of primer regions to the desired template is what controls the specificity of a PCR reaction. When primers bind with high specificity only to the terminal template regions, specific products are synthesized [26]. By products are created when hybridization occurs between primer-template (non-terminal area), template-template and primer-primer sequences [27].

During the amplification of potential aptamers it is important to maintain PCR specificity, reducing the creation of any potential byproducts. During cell-SELEX PCR, byproducts can result from two main sources: the non-specific amplification of library sequences [27] and the amplification of genomic DNA that has been released by target cells and collected along with potential aptamers [28].

1.2.3.1 Library byproduct formation

The amplification of a heterogeneous pool of oligonucleotides comes with the increased risk of byproduct formation as compared to a homogenous sequence [27]. Byproduct formation of heterogeneous pools of nucleic acids is due primarily to product-product hybridization, or template-template hybridization. When two ssDNA sequences hybridize to one another they provide the dsDNA starting point that is required by Taq polymerase to begin the extension process [25]. Therefore, when two template strands hybridize to one another the product that will form will not be a perfect copy of either of the initial templates but will be an undesired byproduct [27].

Previous work in our laboratory explored the process of byproduct formation within a heterogeneous nucleic acid pool. It was shown that the amplification of heterogeneous pools of DNA in the low pM range resulted in the formation of significant byproducts within 20 cycles and by 25 cycles all sequences had been converted to byproduct segments. This high production of byproducts occurred despite the fact that significant primers remained available within the PCR mixture, suggesting that as the amount of product increases so does the risk of product-product hybridization and subsequent byproduct formation. Furthermore, it was shown that as the complexity, with respect to the number of unique sequences, of the initial nucleic acid pool increases so did the risk of byproduct formation [27].

The template concentration is also an important factor that limits the specificity of a PCR reaction. The amplification of target sequences typically plateaus in the nM range after which byproducts begin to form and template sequences are lost from the suspension. While the exact

mechanism is unknown, the over-amplification of target sequences results in their degradation into smaller byproducts [29].

Byproduct formation is a major problem during the cell SELEX procedure since the stating pool of DNA is very complex, consisting of as many as 10^{15} unique sequences and the concentration of bound sequences collected during the first several rounds of selection is high [6]. At the start of selection the heterogeneity of the target bound sequences is very high and is greatly reduced in subsequent rounds as selection progresses. In order to ensure that all potential aptamer sequences collected in the initial SELEX rounds are accurately amplified and remain in the selection pool the number of byproducts formed must be minimized.

1.2.3.2 Amplification of genomic DNA

During the process of cell-SELEX, the collection of bound library sequences is often done through heating of the SELEX mixture, typically to 95°C [6]. Heating the SELEX suspension to such an extent causes cell lysis resulting in the ejection of genomic DNA from the cell [28]. The mixture is then spun down and the supernatant is collected. In the end, the supernatant will contain both library sequences that had bound to the target as well as genomic DNA. Library sequences contain primer regions, permitting them to be amplified during the PCR procedure [29]. Despite the fact that genomic DNA does not contain primer sequences, it still holds the potential to be amplified through various levels of annealing to either the primers or to library sequences present in suspension. Once genomic DNA anneals with another sequence the double stranded starting point for Taq polymerase extension is created. This in turn results in the production of non-specific sequences as well as in the consumption of PCR materials [28].

1.2.3.3 Annealing temperature

The hybridization of primer regions to their target occurs through complimentary base pairing between the nucleic acids and is dependent on the temperature of the reaction. The lower the annealing temperature of a PCR reaction mixture the easier base pairs will form between two ssDNA sequences that are not perfect base-pairing compliments of one another [25]. Using an annealing temperature that is too high will result in a very low number of primers being able to anneal to their target sequences causing a decrease in amplification efficiency. On the other hand, setting the annealing temperature too low will result in base pairing of primers to their targets (with a perfect base pair match) as well as potentially to other undesired sequenced. Furthermore, setting the annealing temperature too low will also result in the reformation of double stranded DNA templates making them inaccessible for primer hybridization and thus amplification [25].

Typically the annealing temperature is selected based on the melting temperature (T_m) of the primer sequences used. This melting temperature can be estimated using a variety of mathematical formulas [25]. The simplest formulas are based mainly on the length and GC content of a primer sequence. More complex formulas consist of several PCR reaction parameters including the Na^+ ion concentrations and base-stacking effects [30]. Since primers are typically purchased commercially, most companies provide an accurate T_m as part of the specification sheet eliminating the need to determine it experimentally or mathematically. For the majority of cell SELEX procedures the annealing temperature is set to just above (approximately 5^oC) the T_m of the primer sequences [25].

1.3 Aptamer applications

Aptamers are selected for a wide variety of applications ranging from *in vivo* to *in vitro* and from research to clinical settings. To date aptamers have been used for imaging, assays, drug targeting, and as drugs themselves [31, 32]. Aptamers provide several advantages compared to their closest counterparts, antibodies. Aptamers show low immunogenicity, are easy to synthesize, are relatively easy to amplify and can be selected against virtually any target using an established *in vitro* method of selection [1].

1.3.1 *In vitro*

In vitro application of aptamers ranges from assays to cell typing. Aptamers are ideal candidates for assay development due to the fact that they can be easily modified and are quite resilient in a variety of conditions [33]. Aptamer-based assays have been developed against a wide range of targets from anthrax spores [28] to ATP [34] to cocaine [35]. Aptamer based assays can be designed based on a variety of detection methods including fluorescence, electrochemistry, nano-labels and nano-constructs [33]. A fluorescence based aptamer assay for cocaine was able to detect levels as low as 1μM, making it a useful technique for cocaine detection in serum, plasma and urine [35]. Aptamers selected against anthrax spores were tested as potential candidates for assay detection of anthrax contamination and could be extremely useful in applications relating to biowarfare [28].

Recently, several aptamer based diagnostic assays have been published demonstrating the application of aptamer affinity and specificity for the detection of pathogens. In a paper published by Duan et. al. it was shown that an assay could be developed using aptamers

conjugated to magnetic nanoparticles to detect less than 10cfu/ml of either *Salmonella Typhimorium* or *Staphylococcus aureus* [36]. Such an assay would be beneficial in identifying contamination of water or food as well as potentially diagnosing patient infections [36].

Over the years both cell-SELEX and other forms of SELEX have generated a wide variety of aptamers against cell membrane proteins, especially against cancer cell lines [37]. Recently these aptamers have been used to develop an aptamer-based cell typing approach. Seven aptamers were used to try and evaluate the cell surface proteome of 14 different cell lines. While some aptamers worked well in identifying cell proteins accurately, others showed a less than ideal level of specificity for their target, suggesting that further research is needed in developing improved aptamers that show not only high affinity but also a high level of specificity for their target [38].

1.3.2 *In vivo*

Much like *in vitro* applications, *in vivo* applications of aptamers vary widely including imaging [39], drug delivery and therapeutics [40]. The idea that aptamers could be used as drugs was suggested in the 1990's with the selection of sequences against RSV. The RSV-specific aptamer showed not only high binding affinity and specificity but also caused a decreased level of infection of viral particle to their targets in cell culture [15]. Antiviral effects of aptamers were again demonstrated by Wang et. al. when SELEX generated sequences were able to diminish the infection rate of HCMV particles [14].

Several therapeutic aptamers have transitioned from the research world into clinical trials. The aptamer AS1411, which targets the membrane bound protein nucleolin was introduced into

clinical trials in 2005 [41]. Results of the clinical trials have varied and the progression of AS1411 as a therapeutic agent against human cancer has not been seen. On the other hand, aptamer pegaptanib, which targets vascular endothelial growth factor (VEGF), successfully underwent clinical trials for the treatment of macular degeneration [42]. Currently the aptamer is marketed as Macugen by Pfizer as a drug for the treatment of macular degeneration [1].

Aptamer-based *in vivo* imaging dates as far back as 1997 when Charlton et. al. successfully used aptamers to image sites of inflammation by linking the Tenascin-C aptamer with a radiation imaging probe [43]. Fluorescence based whole body imaging has been used in conjunction with aptamers most frequently within the field of cancer research. In 2006 Hicke et. al. used aptamers linked to a fluorescent imaging probe to allow for the visualization of glial blastoma within a mouse model [44]. Since then several works have been published showing the wide application of aptamers as fluorescent targeting probes to further research in tumor development, pathogenicity and drug delivery [45] .

The ability to combine aptamers with a wide variety of imaging probes allows researchers and clinicians to not only monitor disease processes directly in a living organism, but also to improve the treatment of diseases through improved pharmacokinetics [45]. In recent years several studies have shown that their ability to be conjugated to anything from a polyethylene glycol (PEG) molecule to a wide variety of nanoparticles makes aptamers very appealing tools in the field of theragnostics. As theragnostic targeting agents, aptamers are able to bind with high affinity and specificity to their target cells thereby directing imaging probes and/or the desired drugs to the cell/tissue of interest. To date, this has mainly been done in mice

model systems with promising results [46-48] Several studies have shown that using aptamers as targeting agents for imaging probes and drugs alike improves treatment outcomes [46-48].

1.4 Research objectives

This research project focused on two main objectives aimed at improving the process of aptamer selection against whole cells. The first objective of this project was to assist in building and subsequently testing a mathematical model that could be used to determine the appropriate masking DNA concentration needed for the selection of aptamers against living cells. The second objective was to modify the PCR amplification procedure for target-bound sequences collected during SELEX to decrease the effects of genomic DNA amplification and to reduce the formation of byproducts within the PCR mixture.

1.4.1 Masking DNA model

The advantages of using masking DNA, in aptamer selection, have been evident for several years [18]. To date, the concentration of masking DNA has been chosen arbitrarily. In response, our lab developed a mathematical model that could be used to determine the optimum concentration of masking DNA, within a cell SELEX mixture, that would outcompete library sequences by a certain number of times. This mathematical model was based on the experimental determination of the average EC_{50} of target cells for both the library and the masking DNA sequences [5].

I successfully tested this model using two different cell lines: MCF-7 human breast cancer and 4T1 mouse breast cancer. Using a commercially acquired ssDNA library and ssDNA

masking DNA sequence, it was shown that the mathematical model fit both the MCF-7 and 4T1 cell lines and could successfully be used to determine the concentration of masking DNA needed to outcompete the non-specific binding of library sequences at a desirable ratio [5].

1.4.2 Cell-SELEX optimization

1.4.2.1 Decreasing effects of genomic DNA amplification

As mentioned above (see section 1.2.1), heating the cell mixture to 95°C for several minutes, followed by centrifugation, achieves the elution of cell-bound library sequences. The supernatant is then collected and amplified via PCR. Unfortunately the resulting supernatant contains both library sequences and genomic DNA. This genomic DNA, despite lacking primer regions, can also be amplified during the PCR procedure.

Using the 4T1 murine breast cancer line I attempted several rounds of cell SELEX. Analysis of PCR products suggested that large (>300bp) sequences were contaminating the library pool. As selection progressed the amount of genomic DNA amplified increased and eventually surpassed the amplification of potential aptamer sequences. To reduce the effects of genomic amplification on cell-SELEX two main approaches were attempted. The first approach was the purification of the PCR products by slab-gel electrophoresis followed by gel extraction purification. This resulted in a significant decrease in PCR product and a potential for aptamer loss. As an alternative to gel purification, touch-down PCR was attempted to increase the specificity of the reaction by decreasing the likelihood of primers or library sequences annealing to genomic DNA and resulting in non-specific amplification. This resulted in a virtual

elimination of genomic DNA amplification while still resulted in a sufficient amount of library amplification.

1.4.2.2 Improving PCR specificity

Byproduct formation during the amplification of collected sequences is prevalent due to the heterogeneity and high concentration of the templates. Both of these factors contribute to a lower PCR specificity and result in the formation of byproducts. Byproduct formation is especially evident during the first several rounds of cell-SELEX when heterogeneity is at its peak and the number of sequences collected following target cell incubation is high.

To simultaneously decrease the heterogeneity and the concentration of target bound sequences a modified aptamer selection procedure was attempted. The traditional cell-SELEX approach was combined with a non-SELEX method for the first two rounds of selection. The consecutive incubation steps, without amplification of sequences in between, resulted in a decrease in the sequence concentration and heterogeneity. When the concentration of collected sequences reached a pM range it was amplified via PCR with an improved specificity as compared to the traditional cell-SELEX approach. Non-SELEX rounds were subsequently followed by the traditional cell-SELEX protocol.

CHAPTER 2: MASKING DNA IN CELL-SELEX

2.1. Introduction

The goal of cell-SELEX is to identify ssDNA sequences, within a random library, that are capable of binding with high affinity to a unique target on a specific cell's surface (in contrast to a ubiquitous target such as a glycolipid). By identifying sequences that are capable of binding strongly and exclusively to a target cell, one can use them as targeting agents for both clinical and research purposes. To efficiently isolate aptamers from other library sequences we need to be able to collect and preferentially amplify (through PCR) the aptamer sequences while minimizing the number of non-specific sequences collected.

Aptamer selection through cell-SELEX faces two major obstacles: the non-specific uptake of ssDNA sequences by dead cells and the non-specific binding of ssDNA sequences to the diverse cell surface [21]. Dead cells are always present, to a certain degree, within a cell suspension. It has been previously shown that the membrane of a dead cell becomes increasingly

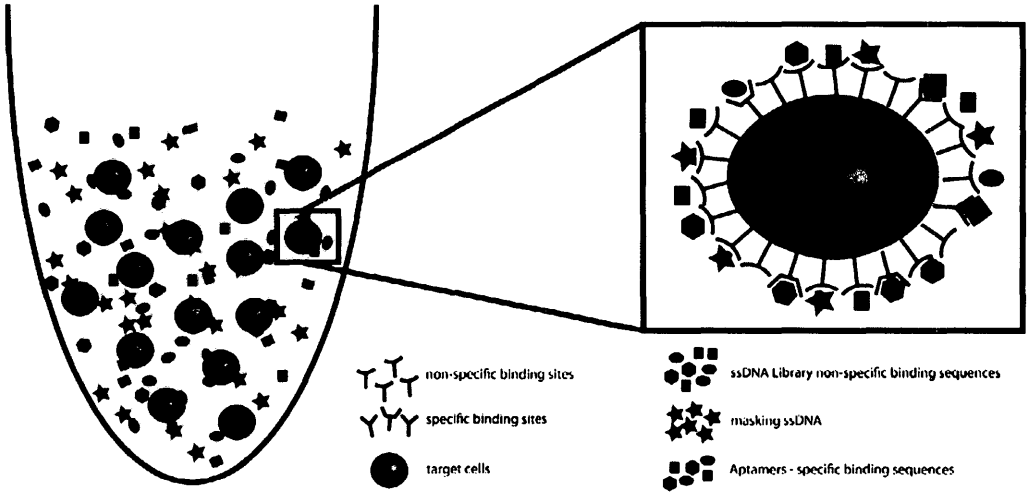

Figure 2.1 A representation of a cell-SELEX incubation mixture consisting of target cells, masking DNA and library sequences both specific and non-specific. The top right inset provides a zoomed in view of the cell surface made up of specific and non-specific binding sites.

permeable, resulting in non-specific uptake of oligonucleotide sequences [10, 21].

At the same time, the living cell surface contains a large variety of proteins and other molecules present in thousands of copies [20]. Many of these molecules are capable of binding ssDNA sequences, typically in a weak, non-specific manner. This large amount of non-specifically bound sequences translates to an increase in the number of selection rounds necessary to obtain a high affinity aptamer pool. In addition, the collection of non-specific sequences along with potential aptamers means that the sequences will compete for PCR reagents resulting in a decreased number of high affinity sequences being amplified successfully [5].

One way of improving the collection and amplification of aptamer sequences is through the use of masking DNA (mDNA). Masking DNA is a scrambled sequence of nucleotides, which lack primer regions and as a result will not be amplified during PCR. Various types of masking nucleic acids have been used in the past, ranging from ssDNA sequences [18] to salmon sperm [10] to yeast tRNA [17]. The masking sequence is added to the cell suspension along with the selection library as shown in **Figure 2.1**. The addition of mDNA along with the DNA library during the aptamer selection process leads to increased efficiency, sensitivity, stability and stringency of the cell-SELEX procedure.

It is important to be able to optimize the mDNA concentrations needed for the most efficient aptamer selection procedure. The addition of too much mDNA, early on in the selection process, could result in a loss of aptamers from the collected fraction (at the end of a round of selection). This loss of aptamer occurs if the mDNA sequences are present in such high concentrations that they are able to out-compete even medium to high affinity library sequences.

At the same time if too little mDNA is added a large amount of non-specific library sequences will be collected. This will result in an increase in the number of rounds needed to isolate aptamers and a decrease in PCR amplification, which works best using pM template concentrations.

2.2. Mathematical model

Previous cell-SELEX procedures have employed masking oligonucleotides but their concentration was chosen arbitrarily. We set out to create a mathematical model that could be used to describe the interactions within a cell-SELEX mixture, and to use this model to determine the optimal amount of masking DNA for cell-SELEX [5].

The number of aptamer sequences in a given library is small, since the odds of having sequences binding with high affinity and specificity to a unique target are very low. The large majority of sequences are non-specific. As a result, following an incubation of the library with target cells, the majority of DNA collected will be non-specific. The challenge is to be able to

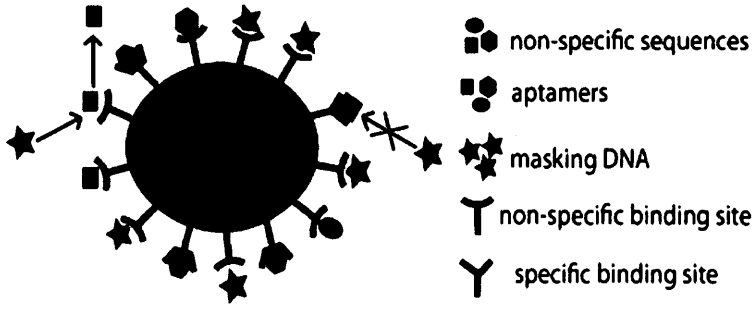

Figure 2.2 A schematic representation of the masking DNA model, where the masking DNA outcompetes non-specific library sequences at non-specific binding sites but is unable to outcompete specific sequences to specific sites due to their increased affinity [5]

separate specific sequences from non-specific sequences as efficiently as possible in the least number of rounds. When masking DNA is added into the suspension, it binds to the non-specific sites, displacing the non-specific library sequences present in the SELEX suspension (**Figure 2.2**). The aptamer sequences bind tightly to their unique targets and therefore are not displaced by the non-specific masking DNA. When sequences are collected at the end of a round the resulting pool contains a small number of aptamers, a large number of masking DNA sequences and decreased number of non-specific library sequences [5].

The cell-SELEX mixture can be simply described as consisting of the following components: aptamers (A), non-specific sequences within the selection library (L_{ns}), cells (n), specific binding sites (N_s) and non-specific binding sites (N_{ns}). The interactions that occur within a cell-SELEX mixture are complex and difficult to solve mathematically therefore, they were simplified based on two main assumptions: $A << L_{ns}$ and $N_s << N_{ns}$. Based on these assumptions the typical type of interaction that occurs in a cell SELEX mixture is non-specific ssDNA sequences from the library pool, or masking DNA (M), binding at non-specific binding sites. Therefore, the majority of interactions occurring on the cell surface within the cell-SELEX mixture can be characterized by two dissociation constants K_{nsL} and K_{nsM} which are defined as the ratio of non-specific sites occupied by non-specific library sequences and masking DNA respectively.

K_{nsL} and K_{nsM} can be combined using the relationship constant kappa (κ), which is defined as the ratio by which the masking DNA sequences outcompete the library sequences at non-specific sites. An equation was derived (**equation 1**) that allows for the calculation of the initial masking DNA concentration (M_0). The mDNA concentration is determined based on the

desired κ valued and other experimental parameters of the cell-SELEX procedure [5].

$$M_0 = \frac{\kappa K_{\text{nsM}} L_0}{K_{\text{nsL}}} + \left(1 - \frac{K_{\text{nsM}}}{K_{\text{nsL}}}\right) \frac{\kappa n S N L_0}{\kappa L_0 + K_{\text{nsL}}}$$
$$\kappa >> 1, \quad \kappa L_0 >> nSN \tag{1}$$

Majority of parameters within equation (1) are dictated by the experimental setup: L_0 (initial library concentration), n (number of cells in solution), S (estimated cell surface area in cm^2) and κ. However, K_{nsM}, N (number of binding sites) and K_{nsL} must be determined experimentally. The unknown parameters can be determined by measuring the binding affinity of library and masking DNA to target cells. The binding affinity of sequences to target cells cannot be described by a dissociation constant due to the complexity of interactions and unknown stoichiometry. Typically, when stoichiometry is unknown binding affinity can be estimated using EC_{50} values (equilibrium concentration at 50% of maximum binding). By measuring the EC_{50} at an increasing concentration of cells (for both the ssDNA library and masking DNA) we can plot the relationship of EC_{50} versus cell concentration and fit by linear regression analysis. The line of best fit will correspond to **equation 2**, from which the slope and y-intercept provide us with the values of N and K_d respectively [5].

$$EC_{50\text{M}} = \frac{1}{2} nSN + K_{\text{nsM}}, \quad EC_{50\text{L}} = \frac{1}{2} nSN + K_{\text{nsL}} \tag{2}$$

The masking DNA concentration required can be determined by introducing both of the K_d values and the average N value into equation 1 along with the experimentally defined parameters.

2.3. Choice of experimental model

The mathematical model was tested against two different cell lines (MCF-7 and 4T1) using the same random library and masking DNA sequence (see Materials and Methods).

2.3.1. MCF-7 human breast cancer

The first model system used to test our mathematical model was the human breast cancer cell line MCF-7. The MCF-7 cell line was established from a metastatic mammary carcinoma obtained from a pleural effusion clot in a female patient [49]. MCF-7 cells are mammary epithelial cells and have been used extensively in research since their initial establishment. The application of the MCF-7 cell line has ranged from *in vitro* studies where characteristics of cancer as well as cancer treatment modalities were explored [50]. In addition early on xenograft tumors of MCF-7 cells were established in athymic mice to allow for *in vivo* cancer research directly on a human cancer cell line within a model system [51].

This widespread use of MCF-7 cells makes them desirable targets for aptamer selection through cell-SELEX. While several aptamers have been developed to purified epitopes found on the MCF-7 membrane [38], the cell line itself has never been used for aptamer selection. Testing our masking DNA model with such a widely used cell line is of great potential value for any future cell-SELEX based aptamer selection that may be performed against MCF-7 cells in the future.

2.3.2. 4T1 mouse breast cancer

The second cell line tested was the murine breast cancer cell line 4T1. The 4T1 cell line is also a tumorigenic mammary epithelial cell line. Development of graft 4T1 tumors in mice closely mimics the development of tumors in humans. This similarity makes 4T1 a very useful and widely used research model system [52]. To date numerous pre-clinical studies have been done, paving the way for potential new treatments and metastasis prevention using 4T1 tumors [53].

Aptamers have been selected through traditional SELEX that show high affinity and specificity for epitopes found on the 4T1 cell surface [54]. However, the cell line, despite its widespread use has not been subjected to cell-SELEX directly. The application of our masking DNA model against the 4T1 cell line not only improves any future aptamer selection against the cells but also demonstrates the robustness of our model for a range of cell lines across species.

2.4. – Materials and methods

2.4.1. DNA sequences

DNA sequences were purchased from IDT (Coralville, IA, USA). All DNA used was 6-carboxyfluorescein (FAM)-labeled at the 5′ end. The masking DNA sequence was generated at random and modified to ensure no complimentarity with the library primer regions (5′-FAM/ AA GGG TCC TGT GCT ATA ACT GTG GGT CTA GTG GTA TTT AG-3′). ssDNA library consisted of a 40-nt random region flanked by 20-nt-long primer regions (5′-FAM/AGC CTA ACG CAG AAC AAT GG-random region-CGA TGC CAG GTT AAA GCA CT-3′).

2.4.2. Cell cultures

MCF-7 cells were grown in the Dulbecco's Modified Eagle medium with high glucose content (DMEM/HG, catalog #D5796, Sigma- Aldrich, St. Louis, MO, USA) with 10% fetal bovine serum (FBS, catalog #SH30396.03, HyCyclone Laboratoried, Logan, UT, USA) and 1% 100 U/mL penicillin–streptomycin (penicillin–streptomycin solution catalog #P4333, Sigma-Aldrich, Oakville, ON, Canada) at 37 °C in humidified 5% atmosphere. Cells were sub-cultured 24 hours prior to cell-suspension preparation.

4T1 cells were grown in RPMI 1640 medium (catalog# 30-2001, ATCC, Manassas, VA, USA) with 10% fetal bovine serum (FBS, catalog #SH30396.03, HyCyclone Laboratories, Logan, UT, USA) and 1% 100 U/mL penicillin–streptomycin (penicillin–streptomycin solution catalog #P4333, Sigma- Aldrich, Oakville, ON, Canada) at 37 °C in humidified 5% atmosphere. Cell diameter of both cell lines was estimated using bright field microscopy calibrated by a stage micrometer grid and was in agreement with previously determined values for both cell lines.

2.4.3. Determining EC_{50}

Plates were rinsed twice with Dulbecco's phosphate buffered saline (PBS, catalog #D8537, Sigma-Aldrich, Oakville, ON, Canada), and cells were detached using 0.05% Trypsin (Trypsin-EDTA Solution 1X, catalog #59417C, Sigma-Aldrich, Oakville, ON, Canada) at 37 °C for no more than 1.5 minutes to minimize the degradation of cell surface proteins. Cells were then re-suspended in DMEM/HG or RPMI 1640 (for MCF-7 and 4T1 cells respectively), counted via hematocytometer, and spun down for 5 min at 300g and 4 °C. Medium was

discarded, and cellular pellet was re-suspended in PBS + 5 mM $MgCl_2$ to the desired concentration (1 to 5 cells per 1 nL).

Masking DNA and library sequences were diluted to 100μM concentration in PBS. Prior to incubation with target cells, masking DNA and library sequences were heated at 95°C for 5 minutes and then snap cooled on ice to allow for the formation of stable three dimensional structures.

Cell suspension was aliquoted into 500 μL samples. Prior to the addition of nucleic acids a single aliquot of target cells was tested for percent cell viability. Target cell suspension was incubated with 1 μL of 1mg/ml propidium iodide (PI) for 5 minutes at 37°C. Following incubation cells were analyzed via flow cytometry for 15 000 – 30 000 events (depending on cell concentration). Number of dead cells was determined by looking at the number of events with a fluorescence signal of 10^5 or higher with the PerCP-Cy5-5-A filter. Suspensions containing greater than 5% dead cells were discarded.

Suspensions with ≤ 5% dead cells were incubated with increasing concentrations of masking DNA or library DNA for 30 to 60 min at 37 °C while gently shaking. Cells were then analyzed by flow cytometry for 30 000 events (except 1 cell/nL suspension which was analyzed for 10 000 events to avoid a large variability in analysis time between different cell concentrations). Mean fluorescence signal was obtained for each masking DNA or library DNA concentration from 10nM until a saturation point was reached based on a lack of increase in fluorescence.

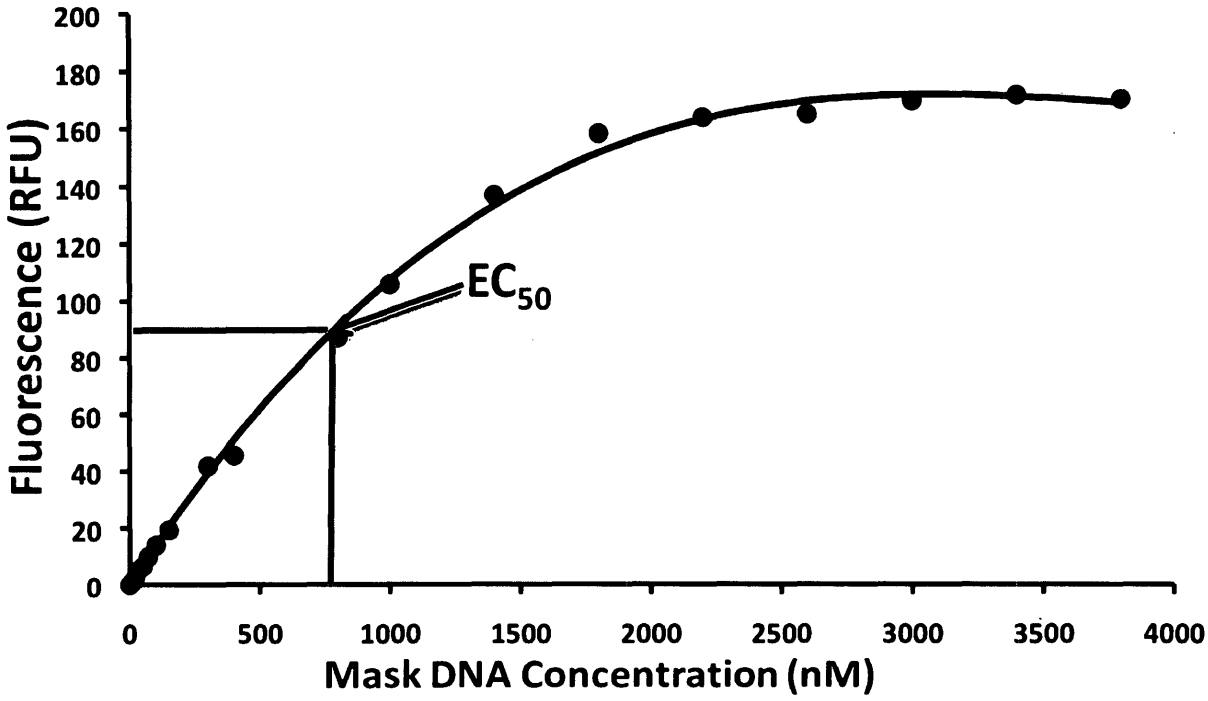

Figure 2.3 The binding affinity curve of 1 million MCF-7 cells incubated with increasing masking DNA concentrations and analyzed by flow cytometry. The EC_{50} value, demonstrated by the red solid line, is the point at which half the maximum fluorescence signal is reached.

A binding affinity curve was generated by plotting DNA concentration versus mean fluorescence signal, and the EC_{50} was determined by locating the DNA concentration at which half of the maximum fluorescence signal was reached. For each cell line EC_{50} values, for both the library and masking DNA experiments, were plotted versus cell concentration and a line of best fit and its representative equation were obtained through Excel, Microsoft Office based on the least squares model.

2.5. Results and discussion

2.5.1. MCF-7 cell line

The mathematical model was first tested using MCF-7 cells. The EC_{50} values were obtained for a range of cell concentrations from 1cell/nL to 5 cells/nL. To determine EC_{50} values at various cellular concentrations binding curves were obtained based on library or masking DNA fluorescence. The binding curves were created using excel and the EC_{50} value was determined manually as the point at which 50% of the maximum fluorescence was reached (**Figure 2.3**). The average cell diameter was measured by light microscopy and equal to 2.5×10^{-3}cm, resulting in an average surface area of 2×10^{-5} cm^2 based on the equation of the surface area of a sphere.

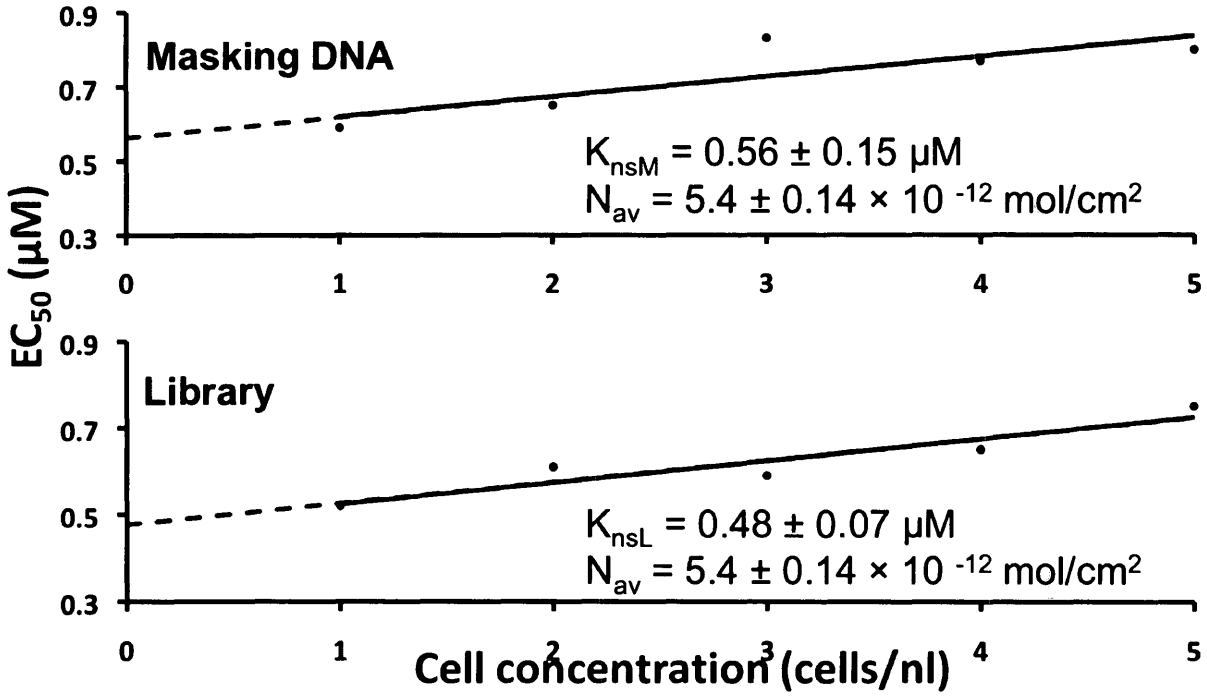

Figure 2.4 The dependence of EC_{50} on cell concentration for the masking DNA and library incubated with MCF-7 cell lines. With the masking DNA line of best fit represented by the equation $y = 0.05x + 0.48$ and library line of best fit $y = 0.054x + 0.566$.

As predicted by the mathematical model, and shown in **Figure 2.4**, the EC_{50} values increased with increasing cell concentration for both the masking DNA and library samples [5]. The line of best fit was calculated using the least squares model on Microsoft Office Excel. Overall the fit of the line, to the experimentally obtained data points, was much better for the library than the masking DNA with R^2 values of 0.87 and 0.42 respectively.

It is important to note that MCF-7 cells are known to aggregate in suspension [49]. Large cellular aggregates present a challenge in two different ways. The first challenge is in obtaining an accurate cell count during cell suspension preparation. Very large aggregates are formed making accurate cell counting difficult. Accurate counting was especially challenging with increased cell concentrations. To decrease the influence of a single inaccurate count, counts were done in triplicate and an average was taken.

The second challenge lay in the analysis of cells following incubation. The suspension is passed through a cellular filter, with a pore size of 45μm, to break up large aggregates prior to flow cytometry analysis. However, cells in small groups of two or three are still be able to pass through the filter. Two or three cells adhering to one another result in a decrease in the number of binding sites available to the nucleic acid sequences present in suspension. The combined effect of inaccurate cell counts and cell clusters resulted in a decrease in the accuracy of experimental data.

The average K_d values for both the masking DNA and the library to binding sites on the target cell surface, based on the y-intercept of the best fit line in **Figure 2.4**, were determined to be 0.56 ± 0.15 μM and 0.48 ± 0.07 μM respectively. The close proximity of the two values suggests that, on average, both the library sequences and the masking DNA sequence have a

similar affinity for non-specific sites on the MCF-7 cell surface. Specific aptamers, selected through cell-SELEX, typically have EC_{50} values in the high pM to low nM range. Since the library sequences and masking DNA should show predominantly non-specific binding, the obtained μM EC_{50} values are as expected [5].

The number of binding sites (N) was calculated based on the slope of the best-fit line, which was equal to 0.2nSN. N values were determined to be 5.7×10^{-12} mol/cm^2 and 5.1×10^{-12} mol/cm^2, for the masking DNA and library respectively. The average number of binding sites was determined to be $5.4 \pm 0.14 \times 10^{-12}$ mol/cm^2. The fact that the number of binding sites determined was similar, confirms that our model assumptions of non-specific binding to the same binding sites for both masking and library DNAs is valid. Furthermore, a crude estimate of the average binding site size was determined, using surface area (S) estimate and the experimentally determined number of binding sites, to be 3.1×10^3 Å^2 where we used the average value of $N = 5.4 \times 10^{-12}$ mol/cm^2 found from EC_{50} measurements for both library and masking DNAs. The obtained value of 3.1×10^3 Å^2 is sufficiently high to allow both library DNA and masking DNA to bind the cell surface with all their nucleotides. In this case, the 80-nt long library ssDNA would occupy ~ 3×10^3 Å^2 and the 40-nt long masking ssDNA would occupy half of this area [5]. Since aptamers fold into complex three-dimensional structures this type of binding is not required and likely does not occur during cell-SELEX. In addition, the target cell surface area is estimated using the cellular radius and the formula for the surface area of a sphere. While this provides a crude estimate sufficient to serve our purposes it is a gross underestimation of the actual available binding surface. Curvatures and folds in the membrane increase the binding surface area significantly.

2.5.2. 4T1 cell line

To confirm the accuracy of our model, experiments were repeated with the mouse cell line 4T1. All testing was done with identical DNA sequences and experimental procedures. The resulting EC_{50} values were plotted versus the cell concentration and are provided in **Figure 2.5**. The average dissociation constants and number of binding sites were determined based on **equation 2** for both the masking DNA and library sequences.

Unlike MCF-7 cells, 4T1 cells do not have a tendency to aggregate making them both easier to count and to accurately analyze via flow cytometry. This is reflected in the increased R^2

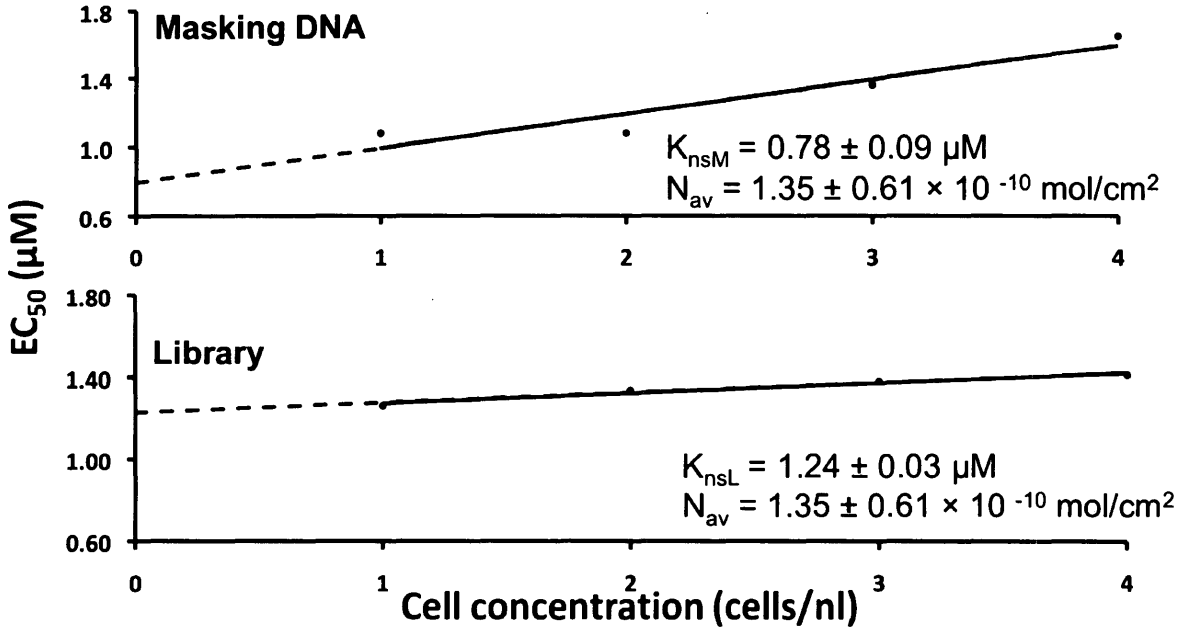

Figure 2.5 The dependence of EC_{50} on cell concentration for the masking DNA and library incubated with 4T1 cell lines. With the masking DNA line of best fit represented by the equation $y = 0.2x + 0.78$ and library line of best fit $y = 0.054x + 1.24$

values of 0.88 and 0.99 for the masking DNA and library respectively (Figure 3.3). The K_d values for the masking DNA and library were 0.78μM and 1.24μM respectively. The lower K_d

value of the masking DNA sequence, versus the library, suggests that the masking DNA sequence may have had a specific target to which it had a slightly higher affinity than the average library sequence.

The number of binding sites (N) for masking DNA and library were determined to be $N = 2.2 \times 10^{-10}$ mol/cm^2 and $N = 4.8 \times 10^{-11}$ mol/cm^2 respectively. The average number of binding sites was determined to be $1.35 \pm 0.61 \times 10^{-10}$mol/cm^2. The 4T1 cell line showed an increase in the number of binding sites between the library and masking DNA versus the values obtained using MCF-7 cells. However, for masking DNA and library sequences the N values were within an order of magnitude. Furthermore, the fact that the masking DNA sequence is half the length of the library sequences (due to lack of primer regions) implies that it may be able to bind to a greater number of epitopes based on size alone. This variation between masking DNA and library sequences may not have been as evident with the MCF-7 cell line due to the variation in R^2 values between the two sequence types.

2.6. Conclusions and future directions

We introduced a simple mathematical model that describes the binding of library DNA (containing potential aptamers) and masking DNA to the cell surface. The model takes into account the high-affinity binding of aptamers to specific sites on the cell surface and the low-affinity binding of library DNA and masking DNA to non-specific sites on the cell surface. Based on this model we developed a simple method for estimating the concentration of masking DNA required to out-compete library DNA at any desirable ratio, κ, describing binding to non-specific sites. Parameters κ, L, n, and S are known from experimental conditions whereas

characteristics N, K_{nsL}, and K_{nsM} are not. We obtained theoretical relations allowing the determination of N, K_{nsL}, and K_{nsM} based on measurements of EC_{50} values for cells mixed separately with the library and masking DNAs [5].

The developed procedure was successfully applied to the MCF-7 breast cancer cell line and corresponding values of N, K_{nsL}, and K_{nsM} were established for the first time [5]. To demonstrate the robustness of the proposed model testing was repeated using the 4T1 cell line. Both cell lines resulted in the successful determination of all needed parameters to allow for the determination of the necessary masking DNA concentrations. The model was successfully applied to both cells despite the fact that the two cell lines vary significantly from their origin to their behavior within a cell-SELEX suspension.

In future experiments, it would be beneficial to further evolve our model by determining the optimum value that would result in the most efficient cell-SELX procedure. This could be done simply by monitoring the progress of selection of several simultaneous cell-SELEX procedures at the same time, which differ only in the value used.

The parameters determined for MCF-7 and 4T1 cell line can be applied to cell-SELEX for aptamer selection. Masking DNA concentrations can be calculated based on the experimental parameters and the desired kappa value. To increase the stringency of selection the kappa value can be increased as selection progresses.

CHAPTER 3: OPTIMIZATION OF 4T1 APTAMER SELECTION THROUGH INCREASED PCR SPECIFICITY

3.1. Introduction

3.1.1. Purpose of PCR in cell-SELEX

During cell-SELEX, sequences bound to the target cells are collected and need to be accurately and efficiently amplified to be used for the subsequent round of selection. Aptamer sequences in the early stages of selection are present in single to low copies. To ensure that the aptamer sequence is maintained within the selection pool it must be amplified accurately and efficiently. To obtain accurate amplicons from a PCR reaction byproduct formation needs to be minimized. The creation of byproducts consumes PCR reagents and results in the potential loss of aptamer sequences from the selection pool. An inefficient and non-specific amplification of aptamer sequences collected through SELEX could greatly hinder the selection process [1].

3.1.2. Amplification challenges

PCR amplification of cell-SELEX collected sequences faces two main challenges: the high level of heterogeneity of the template sequences, their high (nM) concentration [27] and the presence of genomic DNA [28]. Both of these challenges result in a need to optimize the PCR amplification of the desired sequences in order to ensure that accurate products are formed.

The starting library pool typically consists of 10^{15} unique sequences [1]. This type of variety is necessary to ensure that high affinity aptamers, for the targets of interest, are present at the onset of selection. As a result the heterogeneity of template sequences in cell-SELEX PCR step is very high. While this heterogeneity decreases as selection progresses it remains significantly higher, when compared to standard PCR protocols, throughout the selection procedure with potentially hundreds of sequences present in the final round of selection [1].

At the same time, due to the heating of cells during the elution process, genomic DNA is often ejected from the cell during lysis and is collected along with target bound sequences. Misannealing of primers to genomic DNA results in their amplification [28]. The amplification of genomic DNA results in a byproduct, decreasing the efficiency of selection. To decrease the amount of byproduct formation, from template sequences and from genomic DNA contamination, the specificity of the PCR reaction needs to be increased [25].

3.1.3. Traditional PCR approach

The polymerase chain reaction was introduced into the laboratory setting in 1971, but did not become mainstream until the 1980's [55]. The reaction consists of three general steps: denaturation, annealing and extension. The specificity of the reaction is dependent on the annealing of primers to the primer regions of the template sequence [25].

3.1.4. Components and procedure

Typically a PCR master mix is prepared consisting of four main components: buffer, forward primer sequence, reverse primer sequence, and dNTPs mixture. The forward primer initiates the synthesis of the sense sequence and the reverse primer initiates the antisense sequence. Symmetric PCR is performed when equal concentrations of reverse and forward primer are present and therefore dsDNA product is produced. Asymmetric PCR can be performed by adding a higher concentration of the forward primer compared to the reverse resulting in an increased production of the sense strand [29]. Master mix is then aliquoted into small vials and a thermostable polymerase, as well as a template sequences, are added.

Several repetitions, of the three PCR steps, results in the exponential amplification of the template sequence. The synthesized dsDNA strand consists of the sense (forward) strand identical to the template sequence(s) and antisense (reverse) strand complimentary to the template sequence(s). For the purposes of cell-SELEX only the forward sequence is required, therefore a post-amplification clean-up is necessary to obtain a ssDNA pool of sequences. Clean-up of the antisense strand is typically done by biotin tagging the reverse primers, which results in a biotin labeled reverse strand that can be removed by binding to streptavidin resin [25].

3.1.5. Specificity control

The stringency of the PCR reaction is guided by the accurate hybridization of the primer sequence to the primer regions of the template. Byproducts are created through the misannealing of primers-template, primer-primers, template-template and in the cell-SELEX mixture primers-genomic DNA. To ensure that all potential aptamer sequences are amplified sufficiently and accurately the synthesis of byproduct needs to be minimized by improving PCR specificity. The specificity of a PCR reaction can be regulated by three main factors: salt ($MgCl_2$) concentration, additives (eg. formamide, betaine) and annealing temperature.

3.1.5.1. Magnesium ion concentration

Commonly, $MgCl_2$ salt is added to PCR mixtures. Magnesium ions within the PCR mixture act as a concentration dependent co-factor for thermostable DNA polymerase. The presence of magnesium ions stimulates polymerase activity and effects the quality and quantity

of the products formed [56]. The concentration of magnesium within a PCR mixture ranges from 0.5mM to 5mM and is dependent on the presence of other chemicals (eg. ions, EDTA) and concentration of dNTPs. Since dNTPs are negatively charged they have the ability to bind the positively charge magnesium ion resulting in its overall depletion within the reaction mixture [25]. Other chemicals may be added to the reaction mixture through the addition of the template solution, which could contain EDTA or calcium ions added inadvertently during template DNA extraction. Much like dNTPs, such contaminants can affect the optimum magnesium concentration by either sequestering or competing with the Mg^{++} ions [56].

The concentration of salt can be carefully titrated for a given PCR master mix to improve the specificity of the reaction [57]. Commercial kits are currently available, which allow for the optimization of both salt and dNTP concentration for any given PCR reaction increasing the efficiency of the optimization process. However, even with the application of commercial kits optimization of salt and dNTPs concentration is still a fairly labor intensive and time-consuming procedure [25].

3.1.5.2. Additives

Several additives are commonly used to improve PCR stringency. Most commonly these are in the form of detergents, solvents, reducing compounds or enzyme stabilizing agents [25]. Enhancing additives can be used in a wide variety of combinations or on their own to improve both the efficiency and specificity of a PCR reaction. Numerous commercially available kits can be purchased, which contain a pre-mixed enhancing solution. Such kits are typically relatively expensive and the components are not disclosed [58]. Alternatively, enhancing mixtures can also

be synthesized in house in the interest of saving money as well as having a clear description of the mixture's components. Either way the introduction of enhancer(s), either separately or in combination, requires rigorous optimization for any PCR reaction. The way one template and primer combination reacts to an enhancer additive may vary significantly from another [58]. This type of optimization is incredibly time consuming especially since several components need to be optimized for a single PCR mixture [57].

3.1.5.3. Annealing temperature

Since the specificity of a PCR reaction is dependent on the annealing stringency of primer sequences exclusively to primer regions of the template, the annealing temperature is critical in controlling the quality of the PCR product. The annealing temperature is dependent on the melting temperature of the primer sequences with primer regions (T_m). T_m is defined as the temperature at which 50% of the primers in solution are annealed to complimentary primer regions and 50% are present as ssDNA [59]. This temperature can is dependent on several factors including primer length, GC content, salt and template concentrations. Several mathematical formulas exist that allow for the calculation of a T_m with varying levels of accuracy [25]. In addition Tm can be determined experimentally. Software is currently available, which is capable of quickly determining the T_m for a primer sequence taking into account a wide variety of variables [25].

In traditional PCR, the annealing temperature is set to 5°C above the T_m. For simple PCR reactions, using a known uncontaminated homogenous template mixture, this annealing temperature estimate results in a sufficiently stringent primer binding. On the other hand in the

case of a more complex reaction mixture, such as a heterogeneous pool of template sequences or the presence of other contaminants such as genomic DNA, the annealing temperature needs to be optimized experimentally. This is especially true when the template sequences are present in a complex mixture of unknown composition such as cell lysate. Without the optimization of the annealing temperature the risk of byproduct formation is very high [25].

3.1.6. Touchdown PCR

As a solution to the labor-intensive and time consuming optimization of PCR reactions Don et. al. were the first to suggest the use of a touchdown PCR protocol [60]. Instead of determining the optimum annealing temperature for a specific PCR set-up, the touchdown approach uses a range of annealing temperatures. The touchdown PCR protocol consists of two components: touchdown amplification and standard amplification. During touchdown amplification the starting annealing temperature is set to 10°C above the typical annealing temperature (i.e. 15°C above primer T_m) for the first cycle and is decreased by 0.5°C or 1°C at every cycle thereafter. This is typically done for ten to fifteen cycles followed by standard amplification at a typical annealing temperature [60].

The touchdown approach provides an amplification advantage to specific sequences present with the reaction mixture. The increased annealing temperature of touchdown PCR decreases the likelihood that non-specific annealing will occur within the reaction mixture. A decrease in non-specific annealing of sequences translates directly into the decrease of byproduct creation [57]. This means that during the touchdown procedure only target sequences will be amplified. As a result, by the time standard PCR amplification begins the number of specific

sequences will significantly outnumber any potential contaminants that may have been present within the reaction mixture [60, 61].

Increasing the annealing temperature by 10°C results in a significant decrease in PCR efficiency. To ensure sufficient product synthesis, touchdown amplification is followed by several cycles of standard amplification. During standard amplification the annealing temperature remains constant at approximately 5°C above T_m allowing for the exponential amplification of template sequences. The amount of byproducts that are formed as a result of the standard amplification will be insignificant compared to the concentration of target sequence amplicons [60].

3.1.7. Hot-Start PCR

Touchdown PCR is often coupled with hot-start PCR to ensure maximum reaction specificity. In hot-start PCR an essential component (polymerase, primers, dNTPs or magnesium) is kept separate from the remainder of the reaction mixture. Once the mixture reaches a sufficiently high temperature (>90°C) the missing component is added and the reaction is allowed to proceed. This type of approach ensures that all non-specific annealing that may have occurred during the preparation of the reaction mixture, at lower temperatures, is dissociated through heating before non-specific amplification can occur [25].

3.1.8. Product analysis

Following amplification, PCR products need to be analyzed to ensure the success of the reaction. Most commonly, during aptamer selection, DNA products are analyzed using

electrophoresis [25]. For the purpose of cell-SELEX, it is necessary to confirm that the length of the product formed matches the length of the initial library. The presence of dsDNA sequences smaller than the library sequences suggests primer-primer, template-template or primer-template byproduct formation [27]. On the other hand, the presence of dsDNA that is significantly longer than the selection library suggests the amplification of genomic DNA [28].

3.1.8.1.Slab gel electrophoresis

Gel electrophoresis is a commonly used technique in the analysis of nucleic acid. There are two commonly used gels in electrophoresis: agarose and polyacrylamide gel (PAGE). Gels are created from agarose/polyacrylamide dissolved in a buffer solution. The agarose/polyacrylamide creates a dense polymeric network within the gel. The density of the gel is directly proportional the amount of polymer added. Gels are then placed in a buffer solution and samples are loaded into wells. An electric current is applied across the width of the gel inducing movement of the sample components. The smaller components are able to pass easier through the polymeric network and therefore migrate faster within the gel. This variable migration is what causes separation of sample components according to their size [62].

When double stranded DNA (dsDNA) is separated by gel electrophoresis the migration time of different dsDNA sequences is based predominantly on the length of the DNA strand. Once an electric field is applied dsDNA sequences are separate based on their size. If distinct groups of dsDNA (in terms of their length) are included within the sample, they will be separated into different bands according to their size. To visualize the dsDNA within the gel a fluorescent intercalating agent can be included in the run buffer. The intercalating agent (eg. ethidium

bromide) binds the double stranded sequences and its fluorescence can then be visualized using ultraviolet (UV) excitation. To determine the corresponding dsDNA length of each band a base-pair ladder is often run alongside the sample in a separate well. Ladders are available commercially and will form pre-determined bands when analyzed by slab gel electrophoresis. Comparing the migration time of sample bands to those of the ladder allows for the estimation of the length of the dsDNA sequences present within the sample [62].

For slab gel electrophoresis large sample volumes can be applied to the gel for separation. As a result this separation can also be used for dsDNA cleanup. Using a ladder as a reference point, the bands of sample DNA corresponding to a desired length are carefully excised using a scalpel. The dsDNA can then be extracted from the gel using commercial gel extraction kits [63].

3.1.8.2.Capillary electrophoresis

Capillary electrophoresis (CE) is an alternative method to PCR product analysis [64]. The CE set-up consists of a silica based narrow bore capillary which is filled with an electrolyte buffer, which varies depending on the desired conditions for separation. A small volume of the sample is injected into one end of the capillary by pressure. Both capillary ends are then placed into vials containing electrolyte solution as well as an electrode. An electric field is applied across the capillary resulting in the migration of sample components along the length of the capillary towards an electrode (**Figure 3.1**). An online detector interfaces with the capillary and is able to measure a signal (UV or fluorescence) from the sample components as they pass. The detector, in turn, interfaces with a computer, which records an electropherogram from the signal

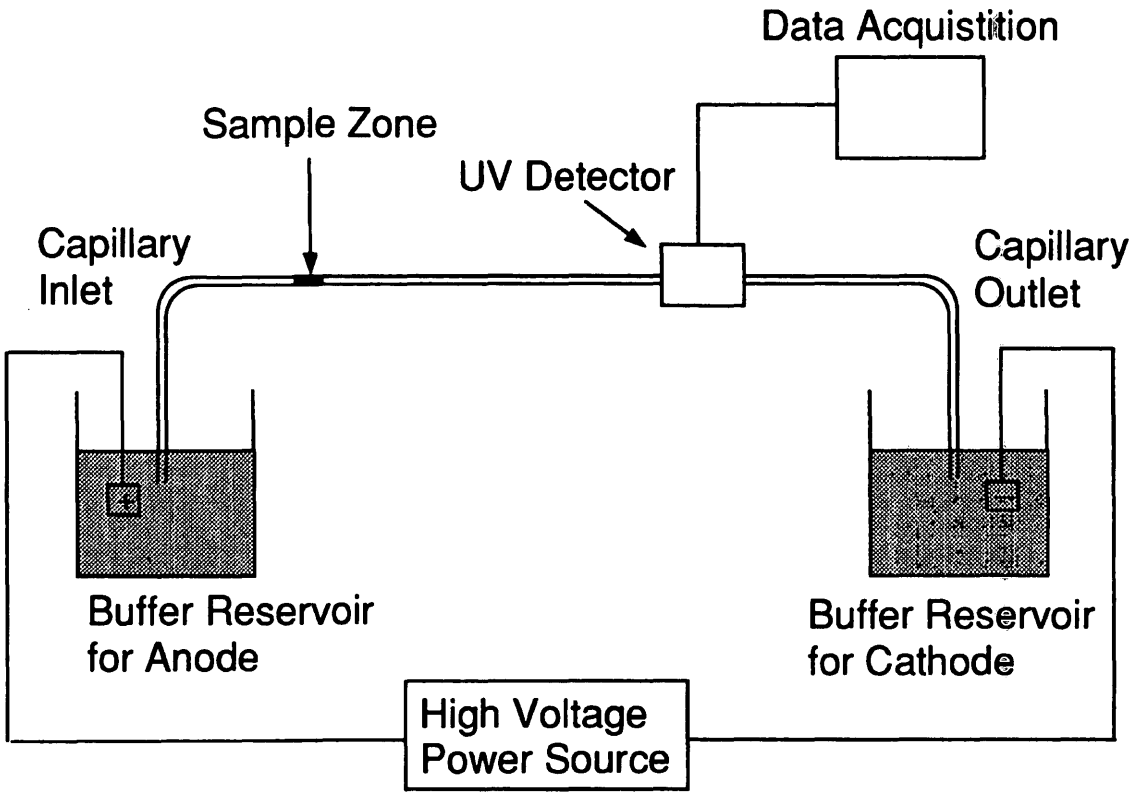

Figure 3.1 A schematic of the capillary electrophoresis set-up adapted from [67].

intensity over time. Migration within the capillary lumen is due to electrophoretic and electro-osmotic force-induced movement. Electrophoretic flow results from action of a uniform electric field along the capillary upon charged species. One buffer vial functions as the cathode and the other buffer vial functions as the anode. Based on their overall size to charge ratio, sample components will move within the capillary towards either the cathode or the anode at variable speeds. Small ions will move faster than large ions and negative ions will move towards the anode with the positive ions going towards the cathodes. This is the same mechanism, on which, nucleic acid separation within slab gel electrophoresis is based. However, due to a high surface area to volume ratio of the capillary used in CE a much larger current can be applied as

compared to slab gel electrophoresis. As a result, components can be separated directly within the capillary buffer and no gel is required [65].

Movement within the CE is also dependent on the electro-osmotic flow (EOF) within the capillary lumen. When a voltage is applied across the capillary buffer solution flows causing ionization of acidic silanol groups along the capillary wall. With a basic pH running buffer, silanol groups dissociate from the surface leaving behind a negatively charged surface, which attracts cations. During separation the cations will move towards the cathode and passively carry the water molecules in which they are dissolved. This creates an overall flow of solution within the capillary towards the cathode. In acidic conditions the reverse effect is seen with an overall EOF towards the anode [66].

The overall separation of species within a capillary will depend on the size to charge ratio. Typically the EOF is significantly stronger than the electrophoretic forces. As a result, all species in a sample are carried by the EOF in the same direction. However, the speed of each species will depend on their respective charge and size. The charge of a species will dictate the strength of the electrophoretic forces. The size and shape of a species will dictate the frictional forces experienced. Based on these distinctions numerous species within a single sample can be efficiently separated during CE [66].

PCR product analysis through CE provides several advantages. CE analysis requires a very small sample for accurate analysis as little as 7μL. Due to the presence of an on-line detector, separation and analysis occurs simultaneously resulting in faster analysis time as compared to slab gel electrophoresis. In addition CE has a lower limit of detection, which allows for the detection of lower DNA concentration than those seen with slab gel electrophoresis [65].

3.2. Choice of experimental model

An attempt was made to select aptamer against the 4T1 cell line. This model system is described in detail above (Chapter 2.3.1). The goal, of this portion, of the research project was to optimize the selection of aptamers for the 4T1 cell line by improving the specificity of the PCR amplification of collected sequences.

3.3. Materials and methods

3.3.1. Cell preparation

4T1 cells were grown in RPMI 1640 medium (catalog# 30-2001, ATCC, Manassas, VA, USA) with 10% fetal bovine serum (FBS, catalog #SH30396.03, HyCyclone Laboratories, Logan, UT, USA) and 1% 100 U/mL penicillin–streptomycin (penicillin–streptomycin solution catalog #P4333, Sigma- Aldrich, Oakville, ON, Canada) at 37 °C under humidified 5% CO_2. Cells were sub-cultured no more than 24 hours prior to selection to ensure that cells can be efficiently removed for selection.

Plates were rinsed twice with Dulbecco's phosphate buffered saline (PBS, catalog #D8537, Sigma-Aldrich, Oakville, ON, Canada), and cells were detached using 2 ml of 0.05% Trypsin (Trypsin-EDTA Solution 1X, catalog #59417C, Sigma-Aldrich, Oakville, ON, Canada) at 37 °C for no more than 1.5 minutes to minimize the degradation of cell surface proteins. Cells were then re-suspended 9 ml of RPMI 1640 and 1 ml of FBS counted via hematocytometer, and spun down for 3 min at 150g and 4 °C. Medium was discarded, and cellular pellet was rinsed twice by re-suspending in 10 ml of PBS and gentle pipeting. The final pellet was suspended in PBS + 5mM $MgCl_2$.

3.3.2. DNA preparation

All oligonucleotides were purchased from IDT (Coralville, IA, USA). All DNA used was 6-carboxyfluorescein (FAM)-labeled at the 5′ end. The masking DNA sequence was generated at random and modified to ensure no binding with the library primer regions (5′-FAM/ AA GGG TCC TGT GCT ATA ACT GTG GGT CTA GTG GTA TTT AG-3′). ssDNA library consisted of a 40-nt random region flanked by 20-nt-long primer regions (5′-FAM/AGC CTA ACG CAG AAC AAT GG-random region-CGA TGC CAG GTT AAA GCA CT-3′).

Prior to incubation with target cells, masking DNA and library sequences were heated at 95°C for 5 minutes and then snap cooled on ice to allow for the formation of stable three-dimensional structures.

3.3.3. Cell-SELEX against 4T1 cells

Selection suspension was analyzed for the percentage of dead cells. Target cell suspension was incubated with 1 μL of 1mg/ml propidium iodide (PI) for 5 minutes at 37°C. Following incubation cells were analyzed via flow cytometry for 30 000 events. Number of dead cells was determined by looking at the number of events with an RFU signal of 10^5 or higher with the PerCP-Cy5-5-A filter. Suspensions containing greater than 5% dead cells were discarded.

Suspensions containing ≤ 5%, containing approximately 5×10^5 4T1 cells were incubated with masking DNA (determined based on a kappa of 5) at 37°C for 5 minutes while shaking at 300rpm. Library was then added to the suspension and the cell-SELEX mixture was incubated

for 30 minutes at 37°C while shaking at 300 rpm. The suspension was spun down at 300g for 5 minutes at 4°C and rinsed twice with PBS + 5mM $MgCl_2$. The final pellet was then re-suspended in 50μL of TE buffer pH 7.5 and heated to 95°C for 5 minutes to remove any unbound sequences. Suspension was then spun down at 13 000 g for 20 minutes at 4°C and supernatant was collected.

3.3.4. PCR amplification

Symmetric PCR master mix consisted of 300nM forward primer, 300nM reverse primer, 1 × PCR Buffer (New England Biolabs Cat # M0273L, Mississauga, ON), 200 μM dNTPs and 1 U/μL Taq DNA Polymerase (New England Biolabs, Cat# M0273L, Mississauga, ON). Asymmetric PCR mix consisted of 1μM forward primer and 50nM reverse primer, 1 × PCR Buffer (New England Biolabs Cat # M0273L, Mississauga, ON), 200 μM dNTPs and 1 U/μL Taq DNA Polymerase (New England Biolabs, Cat# M0273L, Mississauga, ON). Standard PCR was performed by preheating at 94°C for 30 seconds followed by 10-15 cycles of template denaturation for 10 seconds at 94°C, annealing for 10 seconds at 56°C and polymerization for 10 seconds at 72°C. Touchdown PCR protocol consisted of preheating at 94°C for 30 seconds followed by 10-15 cycles of denaturation for 10 seconds at 94°C, annealing for 10 seconds at a range of 66 - 56°C (with -1°C per cycle) and extension for 10 seconds at 72°C.

3.3.5. Analysis of PCR product

PCR products were analyzed predominantly using capillary electrophoresis. To confirm size of produced sequences from genomic DNA amplification products were also analyzed using slab gel electrophoresis.

3.3.5.1. Slab gel electrophoresis

Agarose based gels were created using TAE (Tris base, acetic acid and EDTA) and a 2% agarose concentration. Products were visualized by adding 10μL of 10mg/ml ethidium bromide per 200 ml of running buffer (TAE). Product length was estimated by comparison to a 50bp DNA ladder (New England Biolabs, Cat# N3236S, Mississauga, ON).

3.3.5.2. PCR product purification

Following separation on a 2% agarose gel bands of interest (80bp) were excised using a scalpel and dsDNA was extracted using a commercial gel purification kit (Qiagen, Cat# 28704, Gaithersburg, MD).

3.3.5.3. Capillary electrophoresis

Capillary electrophoresis was performed in a capillary with an inner diameter of 75 μm and a total length of 50 cm. Running buffer consisted of 25mM Borax pH 9.3. Samples were injected by 0.5 psi pressure for 5 seconds and separated at 30 kV for 15 minutes.

3.3.5.4. Quantification of Library Sequences

The concentration of library sequences was estimated using the NanoDrop 3300 Fluorospectrometer from Thermo Scientific. A standard curve was established using fluorescein labeled forward primers (**Figure 3.2**). Measurement of fluorescence was taken at 519 nm

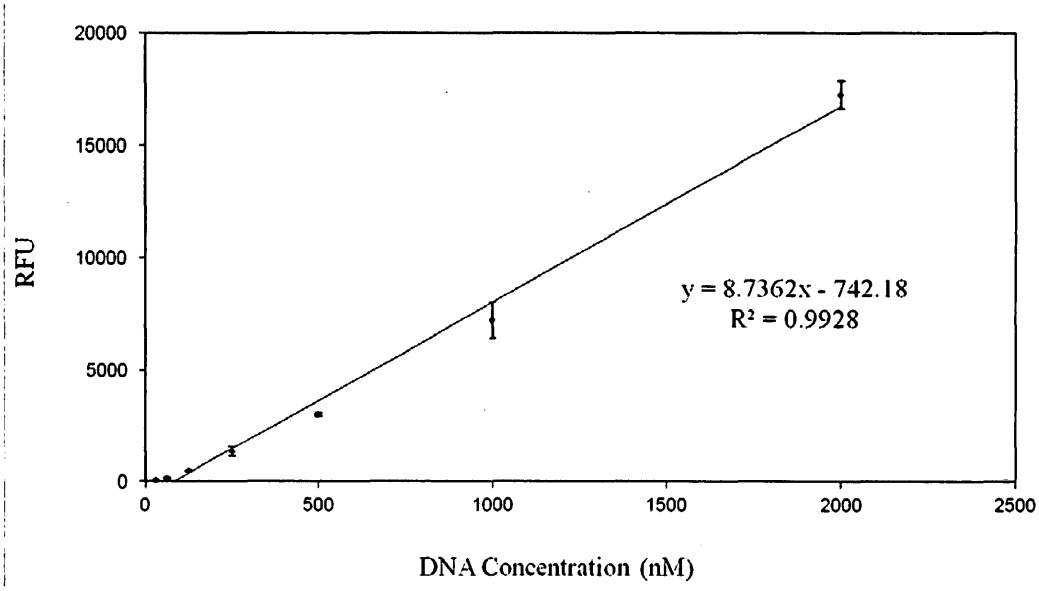

Figure 3.2 Primer calibration curve established using a fluorospectrometer with measurements taken at a wavelength of 519 nm for a range of primer concentrations from 31.25nM to 2000nM. No fluorescence was detected for primer concentrations below 1nM.

emission wavelength for a range of primer concentrations from 31.25 nM to 2000 nM. Measurements were taken in triplicate and plotted to obtain a line of best fit (**Figure 3.2**).

3.4 Results and discussion

3.4.1. Decreasing Byproduct Formation

Aptamer selection was initiated by the incubation of a high library concentration (~ 14μM) with approximately 10×10^6 4T1 cells in suspension. This high initial library concentration is necessary to ensure that potential aptamers are present to and selection is successful [1]. Following incubation, cells were rinsed and bound sequences were collected by

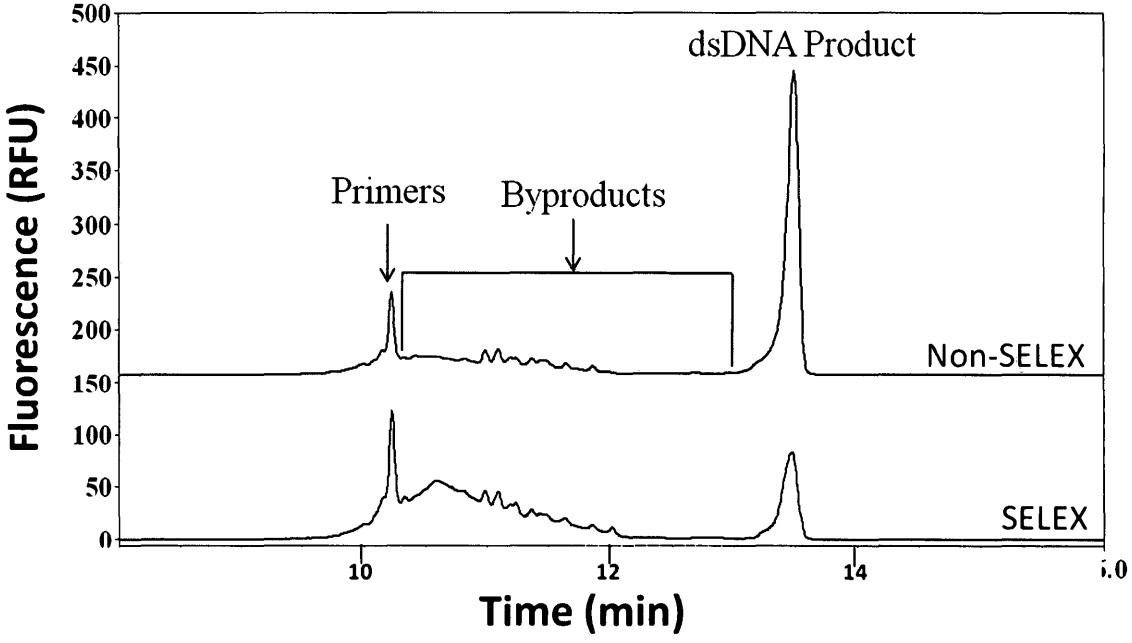

Figure 3.3 An electropherogram of PCR products following the amplification of target bounds sequences, for 10 rounds of symmetric PCR, obtained from a single round of cell-SELEX (blue) or two consecutive rounds of non-SELEX aptamer selection (green).

heating and centrifugation. The concentration of collected sequences was estimated, using a fluorospectrometer and a previously established ssDNA calibration curve (**Figure 3.2**), to be approximately 200nM. All bound sequences were then amplified using symmetrical PCR (with equal concentration of forward and reverse primers) for 10 rounds. PCR products were analyzed via CE and are presented in **Figure 3.3**.

Due to the high template concentration over-amplification was evident after only 10 rounds of PCR. The fact that the template was a heterogeneous mixture further increased the risk of byproduct production. Electropherogram of PCR products (**Figure 3.3**) showed a high byproduct yield. In fact the concentration of byproducts was comparable to the desired product concentration. The production of byproducts that are shorter than the target sequences is due to non-specific annealing between two template strands, which are then both amplified inaccurately. This results in the overall loss of template sequences from the PCR product pool. The loss of a high number of template sequences following a round of SELEX, especially so early in the selection process, can greatly hinder selection efficiency.

To improve PCR amplification the template concentration needs to be decreased, ideally down to the pM range. The simplest way to do this is by simple dilution of collected sequences. However, because the concentration of the ssDNA collected after the first round is so high ($\approx$ 200nM) the sample would need to be diluted over 200 times. Such a large dilution would result in large sample volumes $\approx$ 10ml. Since the initial library contains as little as one copy of each unique sequence, the entire collected fraction would need to be amplified. PCR amplification of such large volumes would be tedious and extremely resource consuming. In addition, this would result in a final product volume of $\approx$ 100 ml of which, less than 1% would be used for subsequent rounds of selection. This would introduce a high level of sampling error. If the fraction of pool one used for subsequent rounds contains a low concentration of high affinity sequences, as compared to the entire PCR product, than an increased number of rounds will be required.

Previous work has demonstrated the successful selection of aptamers, against a purified target, using a non-SELEX approach [4]. In this modified aptamer selection procedure an

oligonucleotide library is incubated with the target of interest and bound and unbound sequences are collected. However, unlike traditional SELEX, non-SELEX aptamer selection does not perform an amplification step prior to the subsequent incubation (**Figure 1.2**) [4].

To decrease the number of target bound sequences, prior to PCR amplification, the non-SELEX approach was combined with the traditional cell-SELEX protocol. Following the initial incubation of library, with target cells, the concentration of the collected ssDNA was estimated using a fluorospectrometer and determined to be in the high nM range. Therefore, the collected fraction was not amplified but instead used directly in a second round of selection with a fresh target cell suspension. Following the second round of non-SELEX selection the collected fraction ssDNA was not detectable using a fluorospectrometer, suggesting a concentration in the pM range. Following the second round of incubation the collected sequences were PCR amplified and the products were analyzed through CE (**Figure 3.3**). The decrease in template concentration resulted in a decrease in the formation of byproducts and improved yield of double stranded target sequences. The total PCR products including percentage of byproduct and desired product formed was estimated based on the total area beneath the appropriate peaks in **Figure 3.3**. The percentage of byproduct formation decreased by 28% with the use of a non-SELEX approach, at the same time the relative percent yield of desired dsDNA product increased by 28%.

The use of two rounds of non-SELEX selection resulted in a sufficient decrease in the concentration of target bound sequences collected to allow for specific amplification via PCR resulting in significantly less byproduct formation. Such a non-SELEX approach allows for the decrease in non-specific sequences, while still maintaining specific sequences within the

collected fraction. Unlike the arbitrary dilution of ssDNA collected after incubation, the non-SELEX approach allows for non-specific sequences to be diluted out, by consecutive incubations with target cells. At the same time, high affinity sequences bind to the target cell and therefore remain within the selection pool. While this approach has been used previously for the selection of aptamers against a purified product [4], to the best of our knowledge this is the first time it has been applied during cellular aptamer selection.

3.4.2. Genomic DNA amplification

3.4.2.1. Identification of a genomic DNA byproduct

Following two rounds of non-SELEX selection, aptamer selection continued as per the typical cell-SELEX protocol described above (Chapter 1.2.1). Collected sequences were amplified using symmetric PCR followed by asymmetric PCR to increase the production of the sense target strand. All PCR products were analyzed via CE. Analysis of the third round PCR products revealed the presence of an additional peak following the expected dsDNA target peak (**Figure 3.4**).

The amplification of genomic DNA would produce long dsDNA sequences significantly longer that the initial library length. Genomic DNA is released by the target cells during the elution of bound sequences by heating to 95°C. If a portion of the genomic DNA is able to bind to either primer sequences or a specific library sequence within the PCR suspension, the result would be amplification of the genomic DNA sequence. Genomic DNA is longer than the 80 nt library and therefore its amplification will result in dsDNA byproduct that will appear after the library amplicons on the electropherogram (**Figure 3.4**). To confirm that the contamination peak

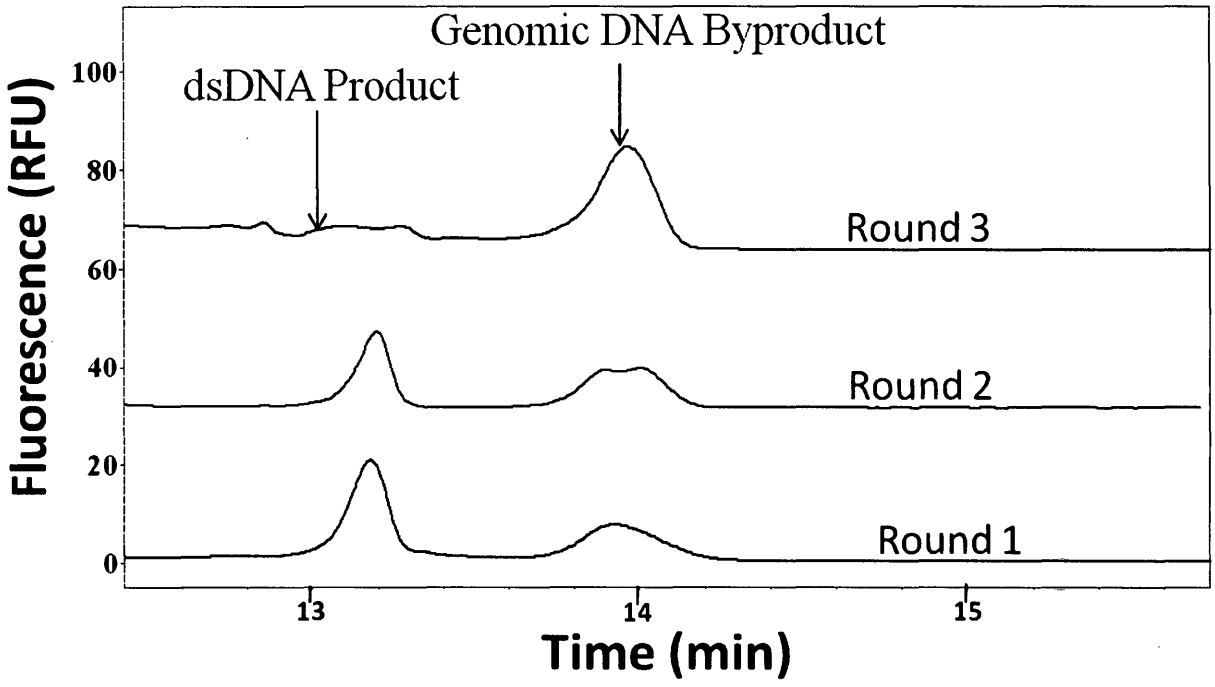

Figure 3.4 An electropharogram of asymmetric PCR products of three consecutive rounds of cell-SELEX showing the desired dsDNA product eluting shortly after 13 minutes and a genomic byproduct eluting at 14 minutes. Round 1, 2 and 3 are shown as red, blue and green traces respectively.

with dsDNA, longer than 80 nt in length agarose slab gel electrophoresis was performed. DNA bands were intercalated with ethidium bromide and visualized under UV light (**Figure 3.5**). Slab gel electrophoresis confirmed that the peak seen on CE was in fact a dsDNA sequence greater than 400 bp in length.

During the subsequent three rounds of cell-SELEX the peak gradually became larger than the desired dsDNA product peak. Eventually the genomic DNA was the predominant dsDNA product following asymmetric PCR. This suggested that the PCR amplification of the target pool resulted in the additional amplification of genomic DNA. Since the amount of genomic DNA

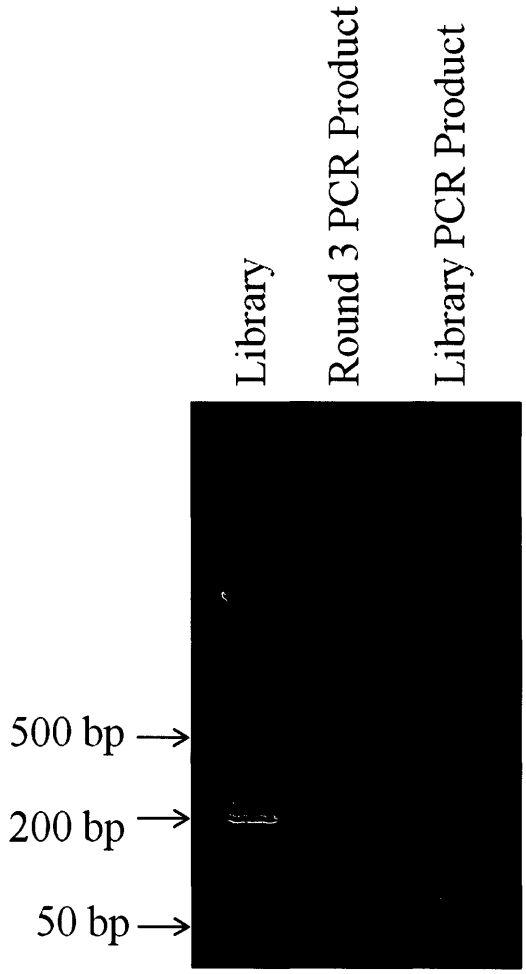

Figure 3.5 Agarose gel electrophoresis of the PCR products produced from the amplification of the third round of cell-SELEX (corresponding to the green trace on **Figure 3.3**) as well as the asymmetric amplification of a naïve library. The dsDNA product of each sample is outlined with a blue box showing a 80bp long product formed from library amplification and a broad band over 300bp long formed from cell-SELEX round 3 pool of sequences

PCR product increased as selection progressed, this suggests that the genomic DNA was not only amplified at each round but was able to bind to target cells with a sufficiently strong affinity to be collected from round to round (**Figure 3.4**).

3.4.2.2.Slab gel purification

In order to eliminate the contamination of the selected library sequences, with amplified genomic DNA, an attempt was made to purify PCR products following slab gel electrophoresis.

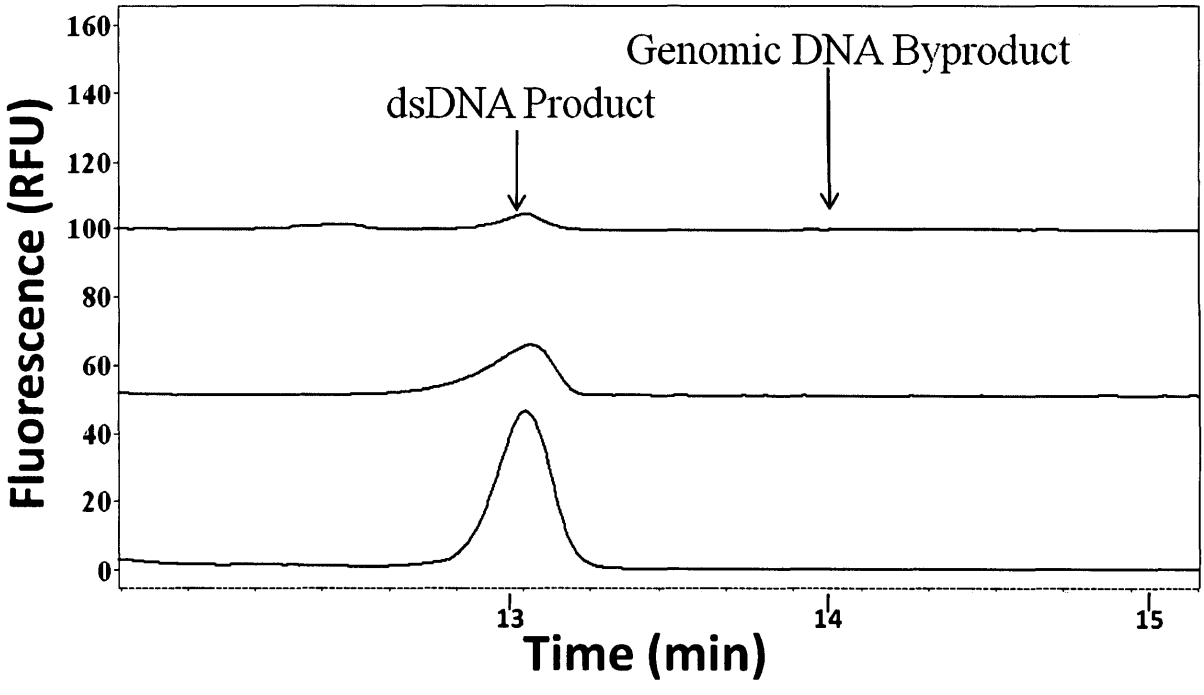

Figure 3.6 Electropharogram of PCR products of sequences collected following one round of cell-SELEX. Symmetric PCR products pre-clean-up (red) and following gel purification (blue) show the presence of the expected dsDNA product peak eluting shortly after 13 minutes. Gel purified products were amplified by asymmetric PCR (green) showing no genomic DNA amplification byproducts.

Following symmetric PCR products were separate through slab gel electrophoresis and the band of interest was excised. The dsDNA was extracted using a commercially acquired kit and was further amplified using asymmetric PCR. All products were analyzed by CE.

The purification process resulted in a significant decrease in the concentration of the target sequences (**Figure 3.6**). On a positive note, the purification step eliminated any visible genomic DNA amplification of the gel purified symmetric PCR product **(Figure 3.6)**.

Overall the gel purification of PCR products is significantly time consuming and added an extra 3 hours to each round of aptamer selection. In addition the decreased concentration of the amplified library sequences following purification implies a loss of sequences. Such a loss of sequences could greatly impede the aptamer selection process (as mentioned earlier) especially in the early rounds of selection when high affinity sequences may not necessarily be dominant within the selection pool.

3.4.2.3. Decreasing Genomic Amplification Through Touchdown PCR

Instead of attempting to purify PCR products, it would be ideal to decrease or eliminate genomic DNA amplification in the first place. Genomic DNA amplification occurs because of ssDNA sequences annealing to the genomic DNA. Since the genomic DNA is clearly visualized by fluorescence during CE, it is safe to assume that genomic amplification is likely due to the forward primer annealing to the genomic DNA, since the forward primer is the only one labeled with a fluororphore. This is further confirmed by the fact that genomic DNA amplification becomes noticeable following asymmetric PCR, in which the forward primer concentration is 20 times higher than that of the reverse primer. Thus, to decrease the amplification we need to be

able to increase the annealing stringency of primers to the primer regions of target ssDNA sequences during the annealing step of a PCR cycle.

The specificity of any PCR reaction is directly a reflection of the level of stringency involved in the primer sequence annealing to template strands [57]. When annealing stringency is high primer sequences anneal by perfect base pairing to their target. Consequently, as stringency decreases primers can bind to other sequences that are not perfect base pair matches. Overall, the primer binding stringency is determined by the annealing temperature, which can be determined mathematically or experimentally for a primer sequence and a specific PCR reaction [57]. Since the ideal annealing temperature is dependent on many factors, optimizing it is very

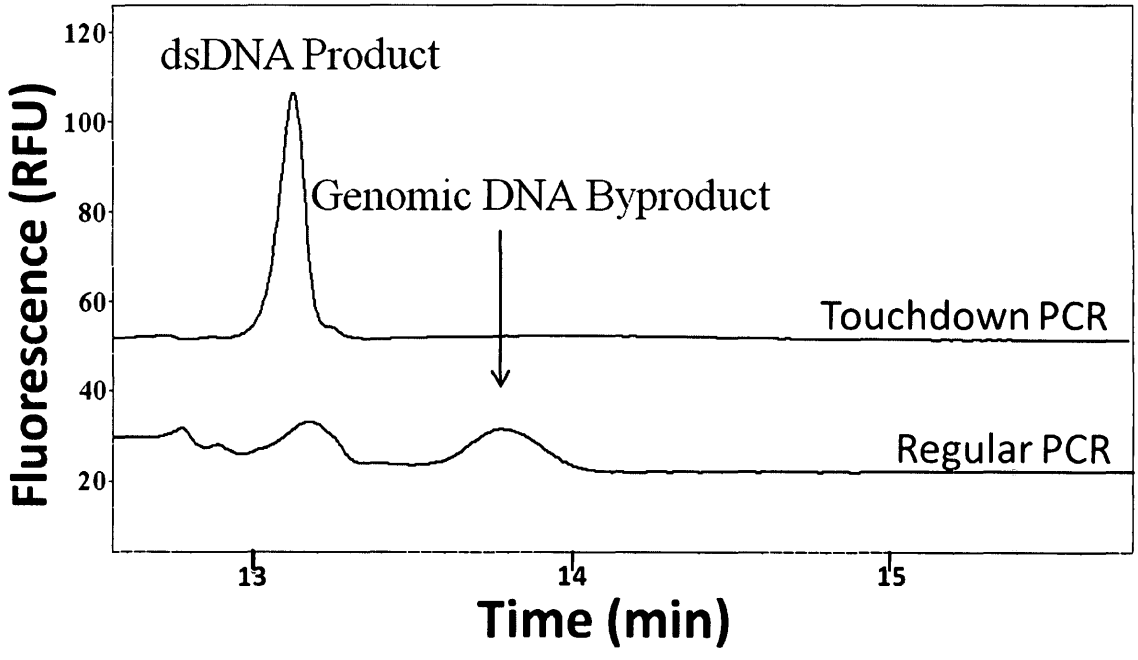

Figure 3.7 Electropharogram of touchdown asymmetric PCR product of round 7 (blue) and round 1 (green) using traditional PCR amplification of cell-SELEX against the 4T1 cell line. Double stranded desired product is shown eluting after 13 minutes and genomic DNA at approximately 14 minutes in the green trace only.

cumbersome and time consuming. Instead of optimizing the annealing temperature touchdown PCR can be used to increase stringency without the need of optimizing a single annealing temperature [59].

Touchdown PCR was set-up by using a range of annealing temperatures starting at 10°C above the primer T_m. The touchdown protocol consisted of a gradual decrease in annealing temperature by 1°C per round of amplification. The touchdown method was applied to both the symmetric and asymmetric amplification of DNA collected at the end of a round of cell-SELEX. In total nine rounds of selection were performed and no genomic DNA byproduct was observed during CE analysis of asymmetric PCR products (**Figure 3.7**).

Overall this suggests that touchdown PCR is an effective way of decreasing/eliminating the amplification of genomic DNA within the cell-SELEX fraction. It is a simple protocol to incorporate into pre-existing PCR methodology. In addition, touchdown PCR does not require any additional equipment, additional steps or optimization [61].

3.5. Conclusions and future directions

The above experiments offer two simple, yet efficient ways of improving the traditional cell-SELEX protocol. Introducing a non-SELEX approach, at the start of aptamer selection against a cell line, lead to a significant decrease in the formation of small byproducts during PCR amplification. While non-SELEX aptamer selection has been previously used to identify aptamers against a purified target [4], it had not previously been applied in combination with the traditional cell-SELEX protocol.

Genomic DNA amplification was virtually eliminated by modifying the existing amplification protocols with a touchdown annealing temperature. This adjustment in the PCR protocol resulted in elimination of genomic DNA amplification product after as many as nine rounds of aptamer selection. The removal of genomic DNA contamination resulted in a modest improvement in the binding affinity as selection progressed. Both techniques suggested above were able to not only improve the cell-SELEX protocol but were also able to do so in a simple manner without relying on excessive reagents, equipment, time or optimization. The above experiments have been able to obtain some improvement in the binding affinity following the seventh round of selection against the 4T1 cell line. Several more rounds of cell-SELEX, modified with the above specifications, need to be completed before potential sequences can be cloned and screened for aptamer applications.

To further improve the binding affinity of the selection pool, better separation techniques need to be introduced to separate out aptamers from the non-specifically binding sequences. Improved separation techniques of bound and unbound sequences would further improve the efficiency of the cell-SELEX procedure. If such separation techniques are combined with the above techniques including the masking DNA model to decrease non-specific binding, aptamer selection against target cell line should show significant improvement.

Obtaining aptamers against the 4T1 cell line would be a valuable research tool for the *in vivo* study of breast cancer metastasis with in a mouse model system. The 4T1 cell line is commonly used in the study of tumor progression and having aptamers against such a commonly used cell line would allow for improvements in research. Specifically, aptamers could be used for targeting imaging probes, allowing for the observation of metastatic tumor progression. At

the same time, conjugation of cancer drugs to the aptamer along with an imaging probe could be used to gain better insight into treatment efficiency of specific therapies. On a broader scope, the cell-SELEX optimizations presented in my work can be used to improve aptamer selection against any cell line. Improving the aptamer selection procedure will lead to selection of a greater number and variety of aptamer sequences, which can subsequently be applied within the research and/or clinical settings.

REFERENCES

1. Mascini, M., *Aptamers in bioanalysis.* 2009: Wiley.
2. Donovan, M.J., et al., *Aptamer–Drug Conjugation for Targeted Tumor Cell Therapy*, in *Therapeutic Oligonucleotides*2011, Springer. p. 141-152.
3. Guthrie, J.W., et al., *Assays for cytokines using aptamers.* Methods, 2006. **38**(4): p. 324-330.
4. Berezovski, M., et al., *Non-SELEX selection of aptamers.* Journal of the American Chemical Society, 2006. **128**(5): p. 1410-1411.
5. Cherney, L.T., N.M. Obrecht, and S.N. Krylov, *Theoretical Modeling of Masking DNA Application in Aptamer-Facilitated Biomarker Discovery.* Analytical chemistry, 2013. **85**(8): p. 4157-4164.
6. Sefah, K., et al., *Development of DNA aptamers using Cell-SELEX.* Nature protocols, 2010. **5**(6): p. 1169-1185.
7. Ellington, A.D. and J.W. Szostak, *In vitro selection of RNA molecules that bind specific ligands.* Nature, 1990. **346**(6287): p. 818-822.
8. Tuerk, C. and L. Gold, *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.* Science, 1990. **249**(4968): p. 505-510.
9. Drabovich, A., M. Berezovski, and S.N. Krylov, *Selection of smart aptamers by equilibrium capillary electrophoresis of equilibrium mixtures (ECEEM).* Journal of the American Chemical Society, 2005. **127**(32): p. 11224-11225.
10. Raddatz, M.S.L., et al., *Enrichment of Cell-Targeting and Population-Specific Aptamers by Fluorescence-Activated Cell Sorting.* Angewandte Chemie, 2008. **120**(28): p. 5268-5271.
11. Stoltenburg, R., C. Reinemann, and B. Strehlitz, *SELEX—a (r) evolutionary method to generate high-affinity nucleic acid ligands.* Biomolecular engineering, 2007. **24**(4): p. 381-403.
12. Morris, K.N., et al., *High affinity ligands from in vitro selection: complex targets.* Proceedings of the National Academy of Sciences, 1998. **95**(6): p. 2902-2907.
13. Zhou, J., et al., *Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells.* Nucleic acids research, 2009. **37**(9): p. 3094-3109.
14. Wang, J., H. Jiang, and F. Liu, *In vitro selection of novel RNA ligands that bind human cytomegalovirus and block viral infection.* Rna, 2000. **6**(4): p. 571-583.
15. Pan, W., et al., *Isolation of virus-neutralizing RNAs from a large pool of random sequences.* Proceedings of the National Academy of Sciences, 1995. **92**(25): p. 11509-11513.
16. Homann, M. and H.U. Göringer, *Combinatorial selection of high affinity RNA ligands to live African trypanosomes.* Nucleic acids research, 1999. **27**(9): p. 2006-2014.
17. Blank, M., et al., *Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels selective targeting of endothelial regulatory protein pigpen.* Journal of Biological Chemistry, 2001. **276**(19): p. 16464-16468.

18. Berezovski, M.V., et al., *Aptamer-facilitated biomarker discovery (AptaBiD).* Journal of the American Chemical Society, 2008. **130**(28): p. 9137-9143.
19. Hicke, B.J., et al., *DNA aptamers block L-selectin function in vivo. Inhibition of human lymphocyte trafficking in SCID mice.* Journal of Clinical Investigation, 1996. **98**(12): p. 2688.
20. Wollscheid, B., et al., *Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins.* Nature biotechnology, 2009. **27**(4): p. 378-386.
21. Avci-Adali, M., et al., *Pitfalls of cell-systematic evolution of ligands by exponential enrichment (SELEX): existing dead cells during in vitro selection anticipate the enrichment of specific aptamers.* Oligonucleotides, 2010. **20**(6): p. 317-323.
22. Jones, K.H. and J.A. Senft, *An improved method to determine cell viability by simultaneous staining with fluorescein diacetate-propidium iodide.* Journal of Histochemistry & Cytochemistry, 1985. **33**(1): p. 77-79.
23. Tang, J., et al., *The DNA aptamers that specifically recognize ricin toxin are selected by two in vitro selection methods.* Electrophoresis, 2006. **27**(7): p. 1303-1311.
24. Shangguan, D., et al., *Aptamers evolved from live cells as effective molecular probes for cancer study.* Proceedings of the National Academy of Sciences, 2006. **103**(32): p. 11838-11843.
25. van Pelt-Verkuil, E., A. Van Belkum, and J.P. Hays, *Principles and technical aspects of PCR amplification*2008: Springer.
26. Marinus, M.G., *Adenine methylation of Okazaki fragments in Escherichia coli.* Journal of bacteriology, 1976. **128**(3): p. 853.
27. Musheev, M.U. and S.N. Krylov, *Selection of aptamers by systematic evolution of ligands by exponential enrichment: addressing the polymerase chain reaction issue.* Analytica chimica acta, 2006. **564**(1): p. 91-96.
28. Bruno, J.G. and J.L. Kiel, *In vitro selection of DNA aptamers to anthrax spores with electrochemiluminescence detection.* Biosensors and Bioelectronics, 1999. **14**(5): p. 457-464.
29. Bartlett, J.M. and D. Stirling, *PCR protocols*. Vol. 226. 2003: Springer.
30. Rychlik, W., W. Spencer, and R. Rhoads, *Optimization of the annealing temperature for DNA amplification in vitro.* Nucleic acids research, 1990. **18**(21): p. 6409-6412.
31. Iliuk, A.B., L. Hu, and W.A. Tao, *Aptamer in bioanalytical applications.* Analytical chemistry, 2011. **83**(12): p. 4440-4452.
32. Barbas, A.S., et al., *Aptamer applications for targeted cancer therapy.* Future Oncology, 2010. **6**(7): p. 1117-1126.
33. Citartan, M., et al., *Assays for aptamer-based platforms.* Biosensors and Bioelectronics, 2012. **34**(1): p. 1-11.
34. Wang, J., et al., *Aptamer-based ATP assay using a luminescent light switching complex.* Analytical chemistry, 2005. **77**(11): p. 3542-3546.

35. Xie, S.-J., et al., *In situ amplification signaling-based autonomous aptameric machine for the sensitive fluorescence detection ofcocaine.* Biosensors and Bioelectronics, 2013. **44**: p. 95-100.
36. Duan, N., et al., *Dual-color upconversion fluorescence and aptamer-functionalized magnetic nanoparticles-based bioassay for the simultaneous detection of< i> Salmonella</i> Typhimurium and< i> Staphylococcus aureus</i>.* Analytica chimica acta, 2012. **723**: p. 1-6.
37. López-Colón, D., et al., *Aptamers: turning the spotlight on cells.* Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 2011. **3**(3): p. 328-340.
38. Li, N., et al., *Technical and biological issues relevant to cell typing with aptamers.* Journal of proteome research, 2009. **8**(5): p. 2438-2448.
39. Smith, A.M., et al., *Bioconjugated quantum dots for< i> in vivo</i> molecular and cellular imaging.* Advanced drug delivery reviews, 2008. **60**(11): p. 1226-1240.
40. Thiel, K.W. and P.H. Giangrande, *Therapeutic applications of DNA and RNA aptamers.* Oligonucleotides, 2009. **19**(3): p. 209-222.
41. Bates, P.J., et al., *Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer.* Experimental and molecular pathology, 2009. **86**(3): p. 151-164.
42. Cunningham Jr, E., et al., *A phase II randomized double-masked trial of pegaptanib, an anti-vascular endothelial growth factor aptamer, for diabetic macular edema.* Ophthalmology, 2005. **112**(10): p. 1747-1757.
43. Bless, N.M., et al., *Protective effects of an aptamer inhibitor of neutrophil elastase in lung inflammatory injury.* Current Biology, 1997. **7**(11): p. 877-880.
44. Hicke, B.J., et al., *Tumor targeting by an aptamer.* Journal of Nuclear Medicine, 2006. **47**(4): p. 668-678.
45. Hong, H., et al., *Molecular imaging with nucleic acid aptamers.* Current medicinal chemistry, 2011. **18**(27): p. 4195.
46. Savla, R., et al., *Tumor targeted quantum dot-mucin 1 aptamer-doxorubicin conjugate for imaging and treatment of cancer.* Journal of Controlled Release, 2011. **153**(1): p. 16-22.
47. Yu, M.K., et al., *Image-Guided Prostate Cancer Therapy Using Aptamer-Functionalized Thermally Cross-Linked Superparamagnetic Iron Oxide Nanoparticles.* Small, 2011. **7**(15): p. 2241-2249.
48. Gao, H., et al., *Whole-cell SELEX aptamer-functionalised poly (ethyleneglycol)-poly (ε-caprolactone) nanoparticles for enhanced targeted glioblastoma therapy.* Biomaterials, 2012. **33**(26): p. 6264-6272.
49. Brooks, S.C., E.R. Locke, and H.D. Soule, *Estrogen receptor in a human cell line (MCF-7) from breast carcinoma.* Journal of Biological Chemistry, 1973. **248**(17): p. 6251-6253.
50. Yang, J., et al., *Curcumin reduces the expression of Bcl-2 by upregulating miR-15a and miR-16 in MCF-7 cells.* Medical Oncology, 2010. **27**(4): p. 1114-1118.

51. Roomi, M.W., et al., *In vitro and in vivo antitumorigenic activity of a mixture of lysine, proline, ascorbic acid, and green tea extract on human breast cancer lines MDA-MB-231 and MCF-7.* Medical Oncology, 2005. **22**(2): p. 129-138.
52. Aslakson, C.J. and F.R. Miller, *Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor.* Cancer research, 1992. **52**(6): p. 1399-1405.
53. Gao, Z.-G., et al., *Prevention of metastasis in a 4T1 murine breast cancer model by doxorubicin carried by folate conjugated pH sensitive polymeric micelles.* Journal of Controlled Release, 2011. **152**(1): p. 84-89.
54. Roth, F., et al., *Aptamer-mediated blockade of IL4Rα triggers apoptosis of MDSCs and limits tumor progression.* Cancer research, 2012. **72**(6): p. 1373-1383.
55. Kleppe, K., et al., *Studies on polynucleotides: XCVI. Repair replication of short synthetic DNA's as catalyzed by DNA polymerases.* Journal of molecular biology, 1971. **56**(2): p. 341-361.
56. Park, Y.H. and R. Kohel, *Effect of concentration of MgCl2 on random-amplified DNA polymorphism.* Biotechniques, 1994. **16**(4): p. 652-656.
57. Kramer, M.F. and D.M. Coen, *Enzymatic amplification of DNA by PCR: Standard procedures and optimization.* Current Protocols in Toxicology, 2001: p. A. 3C. 1-A. 3C. 14.
58. Ralser, M., et al., *An efficient and economic enhancer mix for PCR.* Biochemical and biophysical research communications, 2006. **347**(3): p. 747-751.
59. Korbie, D.J. and J.S. Mattick, *Touchdown PCR for increased specificity and sensitivity in PCR amplification.* Nature protocols, 2008. **3**(9): p. 1452-1456.
60. Don, R., et al., *'Touchdown'PCR to circumvent spurious priming during gene amplification.* Nucleic acids research, 1991. **19**(14): p. 4008.
61. Wu, W.-M., et al., *Touchdown thermocycling program enables a robust single nucleotide polymorphism typing method based on allele-specific real-time polymerase chain reaction.* Analytical biochemistry, 2005. **339**(2): p. 290-296.
62. Barron, A.E. and H.W. Blanch, *DNA Separations by Slab Gel, and Capillary Electrophoresis: Theory and Practice.* Separation and Purification Reviews, 1995. **24**(1): p. 1-118.
63. Kaneta, T., et al., *Band broadening of DNA fragments isolated by polyacrylamide gel electrophoresis in capillary electrophoresis.* Journal of separation science, 2012. **35**(3): p. 431-435.
64. Drabovich, A. and S.N. Krylov, *Single-stranded DNA-binding protein facilitates gel-free analysis of polymerase chain reaction products in capillary electrophoresis.* Journal of Chromatography A, 2004. **1051**(1): p. 171-175.
65. Schmitt-Kopplin, P., *Capillary electrophoresis: methods and protocols*. Vol. 384. 2008: Humana PressInc.
66. Altria, K.D., *Capillary electrophoresis guidebook: principles, operation, and applications*. Vol. 52. 1996: Springer.