SELECTION OF APTAMERS TO CD20 AND THEIR APPLICATION AS INHIBITORS OF COMPLEMENT DEPENDENT CYTOTOXICITY

By: Nadia Al Youssef

A Thesis Submitted to the Faculty of Graduate and Postdoctoral Studies in Partial Fulfillments of the Requirements for the M.Sc. Degree in Chemistry.

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa
DEDICATIONS

To Evah.
ACKNOWLEDGMENTS

First, I would like to thank my supervisor Maxim Berezovski for his gracious support in pursing my work. It’s been a privilege and I am grateful. I thank as well my fellow lab mates, who taught me much about critical assessments and objective thinking. And that the uniqueness of human experience is seldom that unique.

To Dr. Filion for his unsurpassed guidance throughout the years, his insights into novel avenues. And yes, of course, teaching me flow cytometry where the differences between something, anything and everything mean nothing except when seen in the light of proper controls.

Lastly and most fervently I thank my family and friends, my parents whose support and diligence continue to awe and inspire; and who accommodated late night plans way too often to be entirely normal.
ABSTRACT

CD20 is an important oncological B-cell marker. Immunotherapy using anti-CD20 antibodies have revolutionized the treatment of B-cell cancers. Aptamers are highly specific DNA ligands, raised to identify virtually any target molecule through an iterative process known as SELEX (systematic evolution of ligands by exponential amplification). Aptamers rival antibodies in both binding affinity and specificity. We developed a novel CD20 specific SELEX method, using a lentiviral system to transflect CD20 cDNA into HEK293 cells. Selection using CD20+HEK cells evolved pools of aptamers with potent and specific binding in both transfected and independent cell lines. Sequenced aptamer clones exhibited an antagonistic influence with anti-CD20 antibody and in a biological assay possessed a protective capacity, limiting the extent of antibody induced complement dependent cytotoxicity. Genetic transfection is a novel approach for the target specific selection of ligands within a biologically relevant system, and can produce aptamers endowed with both physical and biological actions.
# Table of Contents

LIST OF FIGURES ........................................................................................................... VII

ABBREVIATIONS ............................................................................................................. IX

1 GENERAL INTRODUCTION .......................................................................................... 1
   1.1 TARGETED MOLECULES FOR DIAGNOSTICS AND THERAPY .............................................. 1
   1.2 THERAPEUTIC MONOCLONALS ...................................................................................... 2
       1.2.1 Structure and development ......................................................................................... 3
       1.2.2 Effector Mechanisms of Monoclonals ...................................................................... 5
   1.3 THE CD20 MOLECULE .................................................................................................. 6
       1.3.1 Introduction and History of the CD20 Molecule .......................................................... 6
       1.3.2 Anti-CD20 Antibodies ............................................................................................... 8
   1.4 APTAMER INTRODUCTION ........................................................................................... 9
       1.4.1 Aptamers are created by SELEX ............................................................................... 11
       1.4.2 Cell-SELEX ............................................................................................................. 12
       1.4.3 Viral Transfection for Target Specific Cell SELEX .................................................... 14
       1.4.4 Target Positive Cell-SELEX using CD20 Transfected HEK Cells .............................. 16
   1.5 THESIS OVERVIEW .................................................................................................... 19

2 MATERIALS AND METHODS ....................................................................................... 20
   2.1 TRANSFECTION AND APTAMER POOL SELECTION OF CD20+HEK APTAMERS .................. 20
       2.1.1 Generation and Verification of CD20+HEK Cell line ................................................... 20
       2.1.2 Selection Protocol-Cell SELEX .................................................................................. 23
       2.1.3 Evaluating Aptamer Enrichment with CD20+HEK ...................................................... 25
       2.1.4 Evaluating Aptamer Pools Across Cell Lines ............................................................. 25
   2.2 NGS ANALYSIS AND EVALUATION OF APTAMER CLONES ............................................ 26
       2.2.1 Sample Preparation .................................................................................................. 26
       2.2.2 Aptamer Clone Assessment ...................................................................................... 26
       2.2.3 Assessment of specificity .......................................................................................... 27
       2.2.4 Co-stain experiments ............................................................................................. 27
   2.3 BIOLOGICAL ASSESSMENT ........................................................................................... 28
       2.3.1 Initial CDC Assessment of CD20+HEK and CCL-86 Cells ........................................... 28
       2.3.2 Binding of Aptamer Clones to the CCL-86 cells .......................................................... 28
       2.3.3 Aptamer Mediated Inhibition of CDC ...................................................................... 29

3 TRANSFECTION AND APTAMER POOL SELECTION .................................................... 30
   3.1 ABSTRACT ................................................................................................................... 30
   3.2 BACKGROUND ........................................................................................................... 30
       3.2.1 Characteristic of Lentiviral System .......................................................................... 30
       3.2.2 Stringency Measures ............................................................................................... 37
   3.3 RESULTS ..................................................................................................................... 39
       3.3.1 Evaluation of CD20 Transfection .............................................................................. 39
       3.3.2 SELEX generates pools of aptamer with progressively greater binding affinity to the CD20+HEK cells. 45
3.3.3 Pool 10 CD20 positively labels the naturally CD20 expressing CCL-86 Cells but not the naturally CD20 negative Cell line TIB-152. .......................................................... 47
3.4 DISCUSSION .................................................................................................................. 49
3.5 CONCLUSION ............................................................................................................... 51

4 NEXT GENERATION SEQUENCING AND THE EVALUATION OF APTAMER CLONES .......... 52
4.1 ABSTRACT ..................................................................................................................... 52
4.2 BACKGROUND .............................................................................................................. 53
4.2.1 Deconvolution of Aptamer Pools ............................................................................. 53
4.2.2 Aptamer Sequencing Past Methods: Introduction to Conventional Bacterial Cloning and Sanger Sequencing .............................................................................. 54
4.2.3 Introduction to Next-Generation-Sequencing (NGS) Technologies ......................... 57
4.3 RESULTS ....................................................................................................................... 62
4.3.1 NGS Reveals that Pool 10 Exhibits Significant Sequence Convergence and Motif Emergence ............................................................................................................. 62
4.3.2 HCN are lead aptamer candidates ......................................................................... 70
4.3.3 Aptamer Screening & $K_d$ Analysis: Sequenced Aptamers possess high affinity and selectivity. ................................................................................................. 73
4.4 DISCUSSION ............................................................................................................... 79
4.5 CONCLUSION .............................................................................................................. 81

5 BIOLOGICAL EFFICACY OF APTAMERS IN COMPLEMENT DEPENDENT CYTOTOXICITY ........ 82
5.1 ABSTRACT .................................................................................................................. 82
5.2 BACKGROUND ......................................................................................................... 82
5.2.1 Biological Action of Anti-CD20 Antibodies is Varied and Diverse ............................. 82
5.2.2 Effector Action of anti-CD20 Antibodies .............................................................. 83
5.2.3 Epitope Specificity ................................................................................................. 88
5.2.4 Aptamer to Better Elucidate Target- Antibody Dynamics ........................................ 90
5.3 RESULTS .................................................................................................................... 90
5.3.1 CDC is potently induced by anti-CD20 antibody in naturally CD20 expressing CCL-86 Cells but not in the transfected CD20+HEK cells .............................................. 90
5.3.2 CD20+HEK aptamers can limit extent of CDC in CCL-86 Cells .............................. 93
5.4 DISCUSSION ............................................................................................................ 102
5.5 CONCLUSION ......................................................................................................... 105

6 GENERAL CONCLUSION ............................................................................................. 107

7 REFERENCES ................................................................................................................ 110
**LIST OF FIGURES**

**Chapter 1**

Figure 1.1 General antibody structure

Figure 1.2 Effector mechanisms of therapeutic monoclonal antibodies

Figure 1.3 CD20 expression in B cell ontogeny

Figure 1.4. The CD20 molecule

Figure 1.5 Cell-SELEX

Figure 1.6 Complete selection protocol for CD20 target positive cell-SELEX

**Chapter 3**

Figure 3.1. The evolution of lentiviral vectors for the purpose of mammalian gene expression

Figure 3.2 The tet system for conditional gene expression

Figure 3.3. Generation of intact GOI-viruses by a packaging cell line

Figure 3.4. Complete selection protocol for CD20 target positive cell-SELEX

Figure 3.5. Assessment of CD20 expression in transfected HEK293 cells

Figure 3.6. Visual detection of CD20 expression in transfected HEK293 Cells

Figure 3.7. CD20 expression in different cell lines

Figure 3.8. Binding affinity of aptamer pools with CD20+HEK cells

Figure 3.9. Binding affinity of aptamer pools 1 and 10 with cell lines of varying CD20 expression

**Chapter 4**

Figure 4.1. Bacterial cloning

Figure 4.2. Illumina/Solexa overview

Figure 4.3. Pool 10 exhibits significant sequence convergence
Figure 4.4. Sequence comparison of HCNs and their closest related sequence

Figure 4.5. DREME motif analysis for Pool 10

Figure 4.6. Nucleotide sequence of lead aptamer candidates

Figure 4.7. Purported secondary structure of lead candidate aptamers

Figure 4.8. $K_d$ analysis of NLA aptamers

Figure 4.9. Evaluation of aptamer specificity

Figure 4.10. NLA aptamers inhibit the binding of anti-CD20 antibody

Chapter 5

Figure 5.1. Anti-CD20 effector actions.

Figure 5.2. Clinical status of anti-CD20 antibodies

Figure 5.3. Comparison of CDC induction in naturally CD20 expressing CCL-86 and transfected CD20+HEK cell lines

Figure 5.4 CD20+HEK aptamer show moderate and specific affinity with the CD20 positive CCL-86 and not the CD20 negative TIB-152

Figure 5.5. Complement dependent cytotoxicity in CCL-86: controls samples

Figure 5.6. NLA aptamers limit the extent of complement dependent cytotoxicity in CCL-86 cells

Figure 5.7. NLA-protected CCL-86 cells exhibit greater viability and have decreased staining of the pro-apoptotic marker annexin-V
## ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-AAD</td>
<td>7-Aminoactinomycin D</td>
</tr>
<tr>
<td>ADCC</td>
<td>Antibody dependant cell-mediated cytotoxicity</td>
</tr>
<tr>
<td>ADCP</td>
<td>Antibody dependent cellular phagocytosis</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>CCL-86</td>
<td>Small non-cleaved cell lymphoma (Burkitt's)</td>
</tr>
<tr>
<td>CD20</td>
<td>Cluster of differentiation -20</td>
</tr>
<tr>
<td>CDC</td>
<td>Complement dependent cytotoxicity</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic triphosphate</td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CRT</td>
<td>Cyclic reversible termination</td>
</tr>
<tr>
<td>CTL</td>
<td>Cytotoxic T lymphocytes</td>
</tr>
<tr>
<td>CTLA-4</td>
<td>Cytotoxic T lymphocyte-associated protein 4</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6'-diamidino-2-phenyindole</td>
</tr>
<tr>
<td>ddNTP</td>
<td>Dideoxynucleotide triphosphates</td>
</tr>
<tr>
<td>DREME</td>
<td>Discriminative regular expression motif elicitation</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
</tr>
<tr>
<td>Fab</td>
<td>Fragment of antigen binding</td>
</tr>
<tr>
<td>Fc</td>
<td>Fragment crystallisable</td>
</tr>
<tr>
<td>FcγR</td>
<td>Fragment crystallisable gamma receptor</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>GOI</td>
<td>Gene of interest</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>HCN</td>
<td>High copy number sequence</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic kidney 293</td>
</tr>
<tr>
<td>HiHS</td>
<td>Heat inactivated human serum</td>
</tr>
<tr>
<td>HS</td>
<td>Human serum</td>
</tr>
<tr>
<td>HTA</td>
<td>High throughput analysis</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>K_d</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>mAB</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>MAC</td>
<td>Membrane attack complex</td>
</tr>
<tr>
<td>MFI</td>
<td>Median fluorescence intensity</td>
</tr>
<tr>
<td>MS4A1</td>
<td>Membrane spanning 4 domains subfamily A member 1</td>
</tr>
<tr>
<td>NGS</td>
<td>Next generation sequencing</td>
</tr>
<tr>
<td>NTP</td>
<td>Nucleoside triphosphate</td>
</tr>
<tr>
<td>PD-1</td>
<td>Programmed cell death 1</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>tTA</td>
<td>Tetracycline trans-activator</td>
</tr>
<tr>
<td>rtTA</td>
<td>Reverse tetracycline controlled trans-activator</td>
</tr>
<tr>
<td>SELEX</td>
<td>Systematic evolution of ligands by exponential amplification</td>
</tr>
<tr>
<td>shRNA</td>
<td>Short hairpin ribonucleotide triphosphate</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single stranded deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>TIB-152</td>
<td>Acute T cell leukemia</td>
</tr>
<tr>
<td>VSV-G</td>
<td>Vesicular stomatitis virus envelope glycoprotein</td>
</tr>
</tbody>
</table>
1 GENERAL INTRODUCTION

1.1 TARGETED MOLECULES FOR DIAGNOSTICS AND THERAPY

The cornerstone of all biochemical interactions is molecular recognition. It is the foundation on which gene expression, catalytic enzymes, and all manners of cellular signaling cascades rely upon (1). Targeted molecules are agents that can specifically detect and augment critical biomarkers (2-4). Modern medicine has benefitted enormously from the advent of targeted ligands both as diagnostics as well as therapeutics. Targeted diagnostics reliably evaluate disease emergence and progression (5); while targeted drugs target virulence specific proteins or biochemical pathways (6-8). Together they have not only contributed significantly to the better understanding of disease but clinically targeted drugs are generally more efficacious, associated with fewer side effects and can significantly limit the extent of damage to surrounding cells and tissues (2, 9).

Targeted therapeutics include agents as diverse as antibiotics, anti-virals, small molecule inhibitors, shRNA and monoclonal antibodies (mABs). Antibodies are among the most prolific of targeted molecules and currently dominate both in diagnostics, best exemplified by ELISA kits (5, 10, 11), but also as highly successful therapies (2, 12, 13). Aptamers, which are targeted DNA molecules, rival antibodies in several key characteristics. They possess comparable specificity, in the nano to picomolar range (14-16), as well as a target specific mode of generation (17-19). Unlike proteinacious antibodies, aptamers are significantly more stable at ambient conditions, making them more amenable to diagnostic uses (20). They are non-
immunogenic and can be selected against virtually any target: protein, small molecules, and even whole cells (15, 16, 21-23). Aptamers have great potential and are becoming a rapidly growing consideration in the field of targeted molecular development.

The last 20 years have seen remarkable improvements in cancer treatments heralded by advancements in targeted therapies (13). Contributing significantly to this new age in oncology have been therapeutic monoclonal antibodies (24-26). CD20 is a B-cell specific biomarker with vital anti-cancer considerations (27-29). To its credit, there are 12 anti-CD20 antibodies approved for or still in clinical trials, which treat a variety of immunological malignancies. We endeavoured to use a biologically conscious selection method to generate CD20 specific aptamers and to evaluate and compare their role and functionality against native anti-CD20 antibodies.

1.2 Therapeutic Monoclonals

Monoclonal antibodies (mAbs) are an ideal marriage of target specificity and potent effector mechanisms, making them highly discriminative and biologically active drugs. Their prominence in oncology is significant. Targeted therapeutics make up 56% of all drugs used in cancer treatments (30). Of these 86% of all sales can be ascribed to just 5 drugs which—with the sole exception of Imatinib a small molecule inhibitor— are monoclonal antibodies (30). Unabated the global sales of anti-cancer medications are projected to reach $100 billion as early as 2016 (31). There is then great incentive to pursue targeted ligand development.
1.2.1 Structure and development

Therapeutic antibodies became technical possibilities after the advent of hybridoma technology, developed by Kohler and Milstein in 1975 (32). Hybridomas are immortalized antibody producing cells—the fusion product of antigen challenged and antibody producing B cells with a highly proliferative myeloma, a B-cell cancer. The end product is a long-lived immortalized cell line which produces consistently large amounts of a single antibody (33).

With antigen inoculation performed in-vivo and the requisite cellular assays in-vitro the actual production of antibodies is both time consuming and laborious, in discovery as well as in commercial practise. In the research stages, every hybridoma and the antibody it secretes must be evaluated for specificity, function and potency (34). Commercial antibodies produced in dedicated bioreactors require clarification, purification, ultrafiltration and concentration as well as a quality control validation procedure performed on a batch by batch basis, as variability is common (33, 35). Even then, as proteins, their shelf life is inherently limited. These constraints make antibodies among the most expensive of drugs. Leukemia treatment using alemtuzumab an anti-CD52 antibody is $54,000 yearly (36). Which pales to the $500,000 price of eculizumab (37) a monoclonal developed for paroxysmal nocturnal hemoglobinuria—a disease with orphan status.

The antibodies themselves are heterodimeric proteins, composed of two light chains (one $V_L$ and one $C_L$ unit) and two heavy chains (one $V_H$ and 3 $C_H$ units) linked by disulfide bonds, see Figure 1.1. Pepsin digestion cleaves antibodies into two functional fragments, the Fab or fragment of antigen binding, which determines target specificity, and the crystallisable constant domain or the Fc fragment, which controls the effector functions.
Figure 1.1 General antibody structure. Antibodies are composed of 2 functional domains; the Fab fragment determines antigen specificity while the Fc portion mediates immune effector actions (38).

Figure 1.2 Effector mechanisms of therapeutic monoclonal antibodies. The cytolytic mechanisms of oncolytic monoclonal antibodies include both direct (a) and indirect strategies (b). Direct methods are the result of antibody binding the target molecule stimulating pro-apoptotic cascades. Indirect methods are induced by the Fc portion and require the additional input of immune components in the form of serum proteins for CDC, and cells for ADCC, ADCP and cross presentation (13).
1.2.2 Effector Mechanisms of Monoclonals

Antibodies are effective drugs due to their biological actions (13). Cytotoxicity can be induced directly, the result of Fab binding, or indirectly using the Fc region. See Figure 1.2. Direct cytotoxicity occurs when the antibody binds its target; the agonistic or antagonistic effect depends directly on the nature of the receptor. For example, many cancers cells are associated with the up-regulation of growth factors like EGFR (39). Anti-EGFR antibodies exert an antagonistic influence, hindering signal transduction and leading to cell death (40). This is the same mechanism by which other clinically efficacious antibodies including anti-CTLA4 (41) and anti-PD1 operate (42).

Indirect actions are mediated through the Fc domain and require additional immune components in the form of proteins (for CDC or complement dependent cytotoxicity) or cells (for ADCC: antibody dependent cell mediated cytotoxicity, ADCP: antibody dependent cellular phagocytosis, and cross presentation) (13). There are 4 indirect effector mechanisms noted in Figure 1.2. CDC is the result of an Fc specific interaction with serum complement factors resulting in the formation of a cytolytic membrane attack complex (MAC), which by disrupting osmotic homeostasis leads to cell lysis. In ADCC, antibodies recruit leukocytes to destroy opsonized targets via the controlled release of granulocytic enzymes. In ADCP the Fc fragment acts as phagocytic signal resulting in cellular digestion. In cross presentation the antibodies prime dendritic cells to activate cytotoxic T cells against cancer specific markers. Antibody mediated cellular clearance is multifactorial and an immunologically complicated event.

The FDA approved the first antibody for cancer treatment in 1997. That antibody-rituximab-targets the CD20 molecule and has become the single most important treatment of B-cell
malignancies discovered within the last 20 years (12). This is in spite of the fact that the role and function of CD20 is not clear and that its natural ligand is also unknown. These factors aside it has not hindered drug development, CD20 is the target of 4 therapeutic mAbs currently approved and another 8 still in clinical development (43).

1.3 THE CD20 MOLECULE

1.3.1 Introduction and History of the CD20 Molecule

CD20 was the first B cell differentiation antigen identified (44) and is a characteristic B cell marker from the pre-B cell stage until plasma cell differentiation (45), see the Figure 1.3 below. It is a member of the MS4A1 gene family which encodes a 30-37kDa transmembrane phosphoprotein, depicted in Figure 1.4. CD20 spans the cellular membrane 4 times with 2 extracellular loops and both amino and carboxy termini located within the cytoplasm (46, 47).
Figure 1.3 CD20 expression in B cell ontogeny. CD20 is a marker of B cell development and maturation from pre-B cell to memory B-cell stages. Therapeutic anti-CD20 antibodies are restricted to CLL and B cell lymphomas, cancers that exhibit positive CD20 expression. (47).

Figure 1.4. The CD20 molecule. CD20 is 30-37kDa transmembrane phosphoprotein, with 2 extracellular domains. It spans the cellular membrane 4 times with both N and C termini located within the cellular interior. Mapping epitope studies have revealed the binding sites of several therapeutic antibodies. Rituximab binds the larger extracellular loop and exhibits an absolute requirement for the ANP residues of 170-172. The binding of ofatumumab however is disparate and localized to the smaller 7 amino acid loop (47).
Structurally, CD20 exhibits significant similarity to ion channels proteins and studies have suggested that it may be involved in the regulation of calcium transport. The transfection of CD20 into lymphoid and non-lymphoid cells resulted in significantly elevated transmembrane calcium conductance, the use of anti-CD20 antibodies not only further increased conductance but also resulted in alterations in cell cycle progression (48). In another paper CD20 ligation in tonsillar B-lymphocytes was found to induce $c\text{-}myc$ and $b\text{-}myb$ gene expression, a requirement for B cell proliferation (49). Interestingly CD20 knock-out mice retain normal B-cell number, development, and responses but the B-cells they produce exhibit 20-30% reduced IgM expression as well as lowered calcium responses when compared to the wildtype (50). In 2009 the case of a human patient with cryptic splicing in her MS4A1 gene was investigated. Though B-cell counts were normal, the patient exhibited a complete lack CD20 expression, persistently low IgG levels, a decreased frequency of somatic hyper-mutations and strongly reduced ability to mount T-cell independent immune responses (44). CD20 and its influence on immunity is still very much a matter of ongoing research.

### 1.3.2 Anti-CD20 Antibodies

Though the function of CD20 remains elusive, a consistent feature is that binding of anti-CD20 antibodies induce rampant and relatively widespread cellular depletion. Their use against CD20+ B-cell malignancies has been incredibly successful. There are currently 4 (Canada and US) approved anti-CD20 antibodies: rituximab, ofatumumab and the radio-conjugates bexxar and zevalin. Rituximab was originally licenced to treat aggressive and indolent Non-Hodgkin’s lymphoma (51), it is now widely used against many CD20+ B-cell cancers and even rheumatoid arthritis (52, 53). Bexxar and zevalin are the radio conjugates, bound respectively, to
radioactive yttrium-90 and iodine-131. Ofatumumab was the most recently approved, in 2009, to treat alemtuzumab and fludarabine resistant CLL (54, 55). Epitope mapping studies have shown that ofatumumab recognizes a distinctly different site on the CD20 molecule than rituximab. Ofatumab specifically binds the smaller loop proximal to the cell membrane while rituximab’s binding is restricted to the larger extracellular domain (29, 56). When compared to rituximab in in-vitro trials, ofatumumab is also a consistently more potent inducer of CDC (56). While the relationship between epitope specificity and effector mechanisms remain to be entirely elucidated; the growing interest in the development of anti-CD20 antibodies highlights the increased demand for targeted ligands. Especially to a molecule whose effects, if not well that well understood, are profound. Aptamers with targeted specificity to CD20 would serve as enhanced diagnostic and therapeutic agents with broader utility than antibodies. With a longer shelf life and structure invariability, they are more economical and a more widely applicable tool for diagnostic assays. As therapeutics they can be formulated as functional agents or carrier molecules delivering cytotoxic drugs in a target specific manner. Aptamers could greatly contribute to and would greatly further the current understanding of this molecule. With this in mind, we endeavoured to generate a CD20 specific aptamer selection protocol.

1.4 APOTMER INTRODUCTION

Aptamers are short DNA or RNA molecules, typically 60-120 nucleotides long, raised to specifically recognize a target of interest in a process known as selection (18, 57). They fold into stabilized 3D forms and have the capacity to label, detect and augment the biological
interactions of their target with a high degree of avidity and affinity (23, 58, 59). These characteristics have seen them commonly referred to as a “chemist’s antibody” (60).

Despite this moniker, aptamers possess significant advantages to antibodies. Aptamers are artificially generated and do not require the labour intensive and considerably expensive methods of in-vivo selection (59). Being wholly synthetic endows them with a capacity to withstand long-term storage conditions. DNA is significantly more stable across a variety of temperatures and can be denatured and re-natured without degradation (61). Antibodies, however, must be maintained in a temperature controlled environment to preserve structure and function, and even with preservative agents still possess an inherently limited shelf life. Moreover antibodies can vary batch by batch, and require routine performance validation (62) while aptamers are synthesized consistently and with great accuracy every time. Nucleotide modification can greatly broaden an aptamer’s utility, and can be used to protect aptamers from nucleases, increase their stability and in-vivo retention (63, 64). Unlike antibodies which can only be generated to antigens with immuno-stimulatory capacities (65, 66), aptamers can be raised to virtually any target- including ions, small molecules, proteins and agents which would not otherwise elicit a potent immune response. Aptamers have successfully been selected for zinc ions (21), anthrax spores (67), influenza virus (68), a multitude of proteins including alpha-thrombin and tenascin-C (15, 23), and even whole cells (16, 69-71).

Like antibodies, aptamers have the potential to be therapeutically efficacious. Macugen is an aptamer raised to VEGF (vascular endothelial growth factor) and is clinically approved for wet age-related macular degeneration (72). There are another 10 aptamers, for various therapeutic indications, in clinical trials (73).
1.4.1 Aptamers are created by SELEX

Aptamers are generated from an initial randomized library through an iterative selection process known as SELEX or the Systematic Evolution of Ligands through EXponential amplification. Discovered independently in 1990 by two groups of researchers, Golds (57) and Szostak (18), SELEX begins with a library composed of $1 \times 10^{15}$-20 different DNA or RNA sequences. Each round of selection isolates and retains increasingly higher affinity ligands while weak and non-binding sequences are discarded. This sequential enrichment process culminates with a pool containing high affinity aptamers.

Selection methods can vary extensively and include solid support strategies that immobile the target of interest onto beads (14, 74, 75), membranes affinity columns (21), and nitrocellulose filters (76, 77). There are also gel and liquid based selection techniques like electrophoretic (78) and microfluidic SELEX (79), in-vitro cell based strategies (19, 80) and relatively recently in-vivo selection (81). Solid support systems like beads or membranes represent the most controlled environment, with parameters like concentration and abundance determined at the users’ discretion. A major caveat is the lack of physiological context (80, 82). As target-specific as solid support systems are there is no guarantee that aptamers selected in this manner—to artificially immobilized purified protein extracts—will retain their specificity in an in-vitro setting. Their capacity to function in-vivo is equally uncertain. Thus, to better account for the in-vitro environment and the dynamic complexity of living samples cell-SELEX was developed (19, 60).
1.4.2 Cell-SELEX

Cell-SELEX can generate aptamers whose specificity is to whole cells, (Figure 1.5). Weak and non-binding aptamers are removed by washing or by using negative (counter) selection with a competing background cell group. By coupling positive and negative selection it enhances the discriminative capacity of a pool. In one of the first recorded instance of its use, Wang’s incorporation of what he called “subtractive” SELEX was able to generate aptamers specific only to nerve growth factor (NGF) differentiated PC12 cells, an embryonic cell line, and not their original parental form (71). Cell-SELEX has also successfully identified aptamers that discriminative label lung adenocarcinomas (16), liver cancer cells (70), and gastric carcinoma (83, 84), among others.
Figure 1.5 Cell-SELEX. Generalized selection scheme of cell-SELEX using positive and negative cells. The use of duelling cell lines enhances the discriminative capacity of the aptamers generated. Aptamers that bind positive cells are retained, while aptamer that associate with negative cells are discarded (19).
The major advantage of cell-SELEX is biomarker discovery (85). Aptamers that label and specifically detect cancer cells can be used to better inform on cancer specific characteristics including different phenotypes, the over or under expression of target proteins or mutations on the genetic level. However, cell-SELEX is hampered by the lack of target identity. Elucidating the target of a high affinity aptamer generated through conventional cell-SELEX is daunting work, given the plethora of possible binding sites on cells. These may include embedded and peripheral proteins, any of their distinct epitopes, branched glycosylated groups as well as the heterogenetic membrane lipids themselves. Thus, with cell-SELEX what is gained by the biological context is tempered by the lack of target identity.

To counter this, some SELEX strategies incorporate more than one selection method (86). Cross-SELEX uses target-specific solid selection initially and follows up with cell-SELEX for biological context. And while it has proven effective (15, 87, 88), it is inefficient and does not guarantee the truest high affinity binders but rather the one that best accommodates both techniques. Genetic transfection of cells, which couples target specificity with a cellular environment, is a solution to this dilemma.

1.4.3 Viral Transfection for Target Specific Cell SELEX

Transfection broadly refers to the transfer of genetic material into a cell for the purposes of modulating its expression profile. Current techniques are divided into transient transfection where expression is limited, or stable transfection—significantly the viral mediated strategies—where expression is more permanent (89).
Stable transfection is a technique that emerged after the discovery of DNA-integrating and transforming viruses (90, 91). Originally it was observed that Rous sarcoma virus, papovaviruses, and adenoviruses, could in the course of infection convert an otherwise normal cell into a cancerous neoplastic version. They do this by inserting proto-oncogenes into the host DNA thereby deregulating vital functions leading to malignancies. The chromosomal integration of these genes is heritable and passed down from mother to daughter cells (91). Modern day viral mediated gene transfer was founded on these principals, and by exploiting viral machinery can achieve the same goals, culminating in the site-specific insertion and expression of a gene of interest (GOI).

In contrast, in non-viral methods the gene of interest is carried on a plasmid and then subjected to microinjection, electroporation and/or calcium phosphate transfection into the host cell. However, without a strategy to ensure their genetic incorporation, the information encoded in the plasmid will be lost during the next subsequent cell division. Non-viral transfection is inherently time-sensitive resulting in transient, not stable, gene expression (92).

For stable transfection, retroviruses and lentiviruses make for compact and effective gene delivery vectors. They come packaged with all the necessary components to mediate cellular entry, localization signals for nuclear trafficking, integration factors to insert into the host genome, and mechanisms that even allow for modular gene of interest (GOI) expression. We used a lentiviral based gene transfer system to generate a CD20 specific cell-SELEX protocol. In this method of cell-SELEX, target positive cells will be expressing CD20; the target negative cells are their original untransfected counterparts. This strategy fuses the target specificity of immobilized SELEX methods but in the contextually relevant confines of a live cellular system.
1.4.4 Target Positive Cell-SELEX using CD20 Transfected HEK Cells

The selection protocol for generating CD20 specific aptamers is depicted in its entirety in Figure 1.6 below. The initial DNA library consisted of approximately $1 \times 10^{15}$ random single stranded sequences, 100 bases in length. As indicated in Figure 1.6A every DNA molecule consists of two 20 nucleotide long primer regions found at the 5’ and 3’ termini, which facilitate PCR amplification, and an internal randomized 60 nucleotide segment. Figure 1.6B pictorially represents the target positive and CD20 transfected cells (CD20+HEK) and target negative (CD20-HEK) cells, CD20 expression is indicated by the red star. Selection consisted of 10 rounds incorporating both positive and negative cells. Positive selection denotes the incubation and retention of the DNA ligands that associate with the CD20+HEK transfected cells whereas negative selection exerts a purifying influence by eliminating DNA ligands that associate with the native and untransfected HEK cells. In this way aptamers are evolved with heightened specificity to the key difference between the two cell lines—CD20 expression.
**Figure 1.6. Complete selection protocol for CD20 target positive cell-SELEX.** A) A schematic of a single DNA aptamer including both 20 nucleotide long forward and reverse primer regions and the internalized 60 nucleotide variable region. B) A depiction of both CD20+HEK, with CD20 pictorially represented as red star, and un-transfected HEK control cells. C) Generalized selection protocol featuring both positive and negative cells for generating CD20 specific aptamers. The preparatory clean-up of DNA, required to purify and amplify DNA pools in-between subsequent rounds of selection, is also included.
The selection regime begins first with heat denaturation and rapid snap cooling of the DNA to ensure that the sequences assume their unique energetically stabilized secondary structure, prior to incubation with the target CD20+HEK cells. Unbound or weakly bound sequences are washed away and CD20+ specific aptamer are eluted from the cells. For the first few rounds of selection it is important to enrich the pool in sequences specific to the positive cell line and so negative selection is not performed. When negative selection becomes applicable the eluted DNA pool is denatured and subsequently incubated with the untransfected HEK cells. In negative selection only the aptamers which do not bind HEK cells are desired. For that reason the supernatant is retained while both the negative cells and the sequences bound to them are discarded.

In order to purify, amplify and prepare each pool for the subsequent round of selection, an appropriate work up was necessary. Eluted DNA pools are typically collected in a volume of 1mL and may contain extraneous materials (proteins, lipids, and other cellular debris) acquired during selection. These contaminants complicate PCR amplification and will generally degrade a pool over time. To remedy this the pool is concentrated, with the use of a molecular weight cut-off filter, to a workable volume. Secondly the aptamer DNA is purified from impurities using agarose gel extraction. Next the sample is amplified for the next round of selection using PCR. The last step is the digestion of the 5’- 3’ strand of the amplified DNA using a T4 exonuclease. Single stranded DNA is more amenable to SELEX as it can take on secondary structure and bind with the target of interest.

In total, 10 rounds of aptamer selection were performed. Pools were evaluated with the transfected cells as well as against independent cells lines with varying CD20 expression to
validate specificity. Next generation sequencing was used to characteristically analyze the
sequences and to isolate for the most representative aptamers. These were then evaluated for
affinity, specificity and their impact on the biologically motivated and anti-CD20 antibody
induced complement dependent cytotoxicity.

1.5 Thesis Overview

The thesis has been separated into 4 sections. Chapter 2 is a cumulative overview of the
materials and methods used in the proceeding sections. Chapter 3 details the genetic
transfection, the establishment of CD20+HEK cells and the aptamer pool selection. Chapter 4
covers the next generation sequencing results and the binding and target specificity of the
aptamer clones. Chapter 5 provides the biological results, specifically CDC induction and the
protective effect of the aptamers. A general conclusion is also included which summarizes the
major findings.
2 MATERIALS AND METHODS

2.1 TRANSFECTION AND APTAMER POOL SELECTION OF CD20+HEK APTAMERS

2.1.1 Generation and Verification of CD20+HEK Cell line

2.1.1.1 Construction of CD20 containing pLVX-TRE3G Vector

Wildtype CD20 cDNA was a kind gift from Genmab. The circularized pLVC-TRE3G vector was linearized using MiuI and EcoRI restriction endonucleases, with incubation performed at 37°C for 3 hours. To purify the vector DNA the linearized DNA product was run through a 1% agarose gel with a molecular weight ladder and excised using the GeneJET gel extraction kit. To subclone, the cDNA was first amplified using Clontech Laboratories CloneAmp HiFi PCR premix using 15 base pair extended primers with homologous ends to the linearized pLVX-TRE3G vector. Ligation of the amplified PCR product with the linearized vector was performed using Clontech’s Fusion HD cloning kit; using 50ng of the purified CD20 cDNA, 100ng of the linearized vector, 2μL of 5x In-fusion HD enzyme premix and 7 μL of deionized water. The insertion of CD20 cDNA in the vector was verified using PCR and gel electrophoresis.

To generate sufficient quantities of the CD20 carrying vector for transfection 2.5ng of the cloned pLVX-TRE3G reaction mixture was mixed on ice with one reaction vessel (50μL) of Clontech’s Stellar Chemically Competent Cells. The mixture was heat shocked for 45 seconds at 42°C, and then cooled on ice for 2 minutes. SOC media was added to a final volume of 500 μL and the cellular suspension plated on 100μg/ml ampicillin fortified agar media. The cells were incubated at 37°C overnight. Individual bacterial cells were picked from the plate, and grown in ampicillin fortified LB media (100mg/mL) for 12 hours. Plasmids were extracted using
the GeneJET plasmid miniprep kit according to the manufacturer's instructions. Purified plasmids were sequenced at The Centre for Applied Genomics (TCAG) to validate CD20 insertion.

2.1.1.2 **Lentiviral Production and Transfection**

The CD20 containing pLVX-TRE3G plasmids were transfected into the Lenti-293 packaging cell line using Clontech’s Lenti-X HTX Packaging System, part of the Lenti-X Tet-on 3G Inducible Expression System, according to the manufacturer's instructions. Virus particles were harvested 48 hours after initial transfection and viral concentration measured using the Lenti-X Go-STIX. The freshly harvested viruses possess titre greater than 50,000 IFU/mL. CD20+Lentivirus was used to transfect wild type HEK293 cells. Cells were grown in 10% FBS fortified DMEM media for 48 hours. To select for and isolate successfully transfected cells the cellular media was fortified with G418 (1 mg/mL). Cells were re-cultured every 3-4 days as needed for 2 weeks. The pLVX-TRE-3G vector incorporates the tet-on operon upstream of the multiple cloning site. Tet-on is an inducible expression element, it permits the transcriptional control of the gene of interest in the presence of doxycycline. Once cell growth after selection had stabilized, the induction of CD20 expression was initiated using doxycycline (500ng/mL).

2.1.1.3 **Cellular Maintenance**

All cells were maintained in a humidified incubator at 37°C and 5% CO₂ an in 100mm² plates. Both untransfected and CD20+HEK cells were maintained in Dulbecco modified eagle medium (DMEM) supplemented with 10% fetal bovine serum. CD20+HEK cellular media was fortified with doxycycline (500 ng/mL) to ensure constant expression. TIB-152 and CCL-86 were purchased from American Type Culture Collection (ATCC) and grown in RPMI media supplemented with 10% FBS.
2.1.1.4 *Flow Cytometry Detection of CD20*

To evaluate CD20 expression cells both adherent CD20+HEK and HEK control cells were removed from the culture plates using non-enzymatic cell stripper. Cells were washed twice with PBS and initially assessed using the Muse count and viability reagent; only cells with greater than 90% viability were used. 500,000 cells were aliquoted into micro-centrifuge tubes and incubated with 10ng/μL FITC labelled anti-CD20 antibody for 30 minutes on ice. Cells were washed twice with PBS, re-suspended in the buffer and evaluated using a Beckman Coulter FC500 flow cytometer where 40,000 events were counted.

2.1.1.5 *Immunofluorescence Detection of CD20+HEK*

Requisite CD20+HEK and the un-transfected HEK control cells were grown in imaging chambers until 80% confluent in doxycycline fortified DMEM media. The spent media was aspirated and the monolayer washed with PBS. Cells were fixed using 4% PFA for 20 minutes and washed twice with PBS. For analysis the fixed cells were blocked using a solution of 10%FBS in PBS for 2 hours before the addition of primary FITC labelled anti-CD20 antibody applied using the recommended concentration of 50ng/μL. Incubation was performed for 1 hour in the dark. The cells were then washed using PBS, counterstained with DAPI (4’,6’-diamidino-2-phenylindole) and mounting on glass slides. Cells were visualized using a conventional fluorescence microscope.

2.1.1.6 *Contextual anti-CD20 Expression*

To contextualize CD20 expression the lymphocytic CCL-86 cells and TIB-152 cells were used. CCL-86 cells are a B-cell line that naturally express CD20 while TIB-152 T-cells do not. 500,000 cells were aliquoted into micro-centrifuge tubes, washed twice in PBS, and then re-suspended
in FITC labelled primary anti-CD20 antibody diluted at the following concentrations for 30 minutes on ice: 100ng/μL, 50ng/μL, 10ng/μL, 5ng/μL, 2.5ng/μL and 1ng/μL. Cells were washed in PBS twice and evaluated using a FC500 flow cytometer where 40,000 events were counted.

2.1.2 Selection Protocol-Cell SELEX

2.1.2.1 DNA Preparation And Aptamer Pool Selection

The initial library was purchased from IDT (Integrated DNA Technologies) and contains $10^{15}$ random sequences. It is 100 base lengths in size consisting of a 60 base length randomized internal region flanked on either side by constant primer regions as such: 5’CTCCTCTGACTGTAACCACG-N60-GCATAGGTAGTCCAGAAGCC 3’. The initial lyophilized stock was re-suspended in 1XTE buffer at a final concentration of 100μM as per the manufacturer’s instructions. Prior to incubation with cells the DNA sample, diluted to the requisite concentration (see Chapter 3 Table 1), was denatured at 95°C for 10 minutes and then snap cooled on ice for a minimum of 10 minutes prior to use in selection.

Positive selection denotes selection performed using CD20+ transfected cells; negative selection with the un-transfected HEK cells. Each round of selection consumes 1 100mm$^2$ polystyrene plate with cells grown to 90% confluency. Prior to each round of selection the spent media is aspirated and the monolayer of cells prepared by washed with PBS twice. The DNA solution denatured, snap cooled and diluted to a final volume of 1000μL is applied drop-wise to the CD20+HEK monolayer. The sample is incubated at 37°C on a heated shaker. For rounds 1 through 4 no negative selection was used. After incubation the cells are sloughed off the plate, collected into a 15mL Falcon tube and then washed twice with PBS. The cellular
pellet is heated at 95°C with gently vortexing for 10 minutes, centrifuged at 1000RCF for 5 minutes and the aptamers eluted from the supernatant.

For selection rounds incorporating negative selection, rounds 4-10 inclusive, after the eluted aptamers are collected they are applied to a 100mm² plate containing washed un-transfected HEK control cells. The aptamers are applied drop-wise to the cellular monolayer, and incubated for the allotted time indicated in Chapter 3 table 1. After which the supernatant, containing unbound aptamers, is carefully collected. This comprises the cell-based component of selection.

2.1.2.2 Selection Protocol-Work Up

To concentrate the pool into a more workable volume we used an Amicon 50kDa molecular weight cut-off filters. This device is bipartite; initially the DNA pool is applied to the upper filter and spun using a table top centrifuge at 2800RCF for 2 hours, the filtrate containing low molecular weight debris is discarded. In the second step the internal compartment is reversed and spun down again at 2800 RCF for 5 minutes. 1mL of solution was concentrated to a final volume typically around 50μL. The concentrated DNA pool is purified using agarose gel extraction. The DNA pools are run alongside a MW ladder on a 3% agarose gel for 30-45 min at 200V. The 100bp band is excised and purified with the GeneJET extraction kit. The collected and extracted DNA pool is PCR amplified with Thermoscientific Phire Hot Start DNA Polymerase kit according to the manufactures instructions. In each 50μL of the reaction mixture is 1x phire reaction buffer, 2%DMSO, 200 μM of dNTPs, 0.02U/μL of the Phire Hot Start II DNA Polymerase and 0.5 μM of both primers the 5’-Cy5-labeled forward primer the 5’-phosphorylated reverse primer. 30 cycles of PCR were performed according to the following
denaturation at 95 °C for 30 s, annealing at 56°C for 15 s, and extention at 72°C for 10 s. DNA amplification is verified on a 3% agarose gel alongside a molecular weight ladder. To generate a single stranded product Lambda exonuclease, purchased from New England Biolabs, was used according to the manufactures instruction. After PCR amplification the pool of aptamers is incubated both with the buffer and the exonuclease enzyme for 2 hours at 37°C.

2.1.3 Evaluating Aptamer Enrichment with CD20+HEK

Adherent CD20+HEK cells were removed from the plate using non-enzymatic cell stripper. Cells were washed twice in PBS, and counted using Muse cell count and viability reagent.

250,000 cells were aliquoted into each micro-centrifuge tube and incubated with relevant DNA pools derived from the unselected library and pools 1, 5 and 10 of selection at a concentration of 200nM. Cells were incubated at 37°C for 30 minutes and then washed twice with PBS. Cells were re-suspended in buffer and evaluate by a Beckman Coulter FC500 flow cytometer where 40,000 events were read.

2.1.4 Evaluating Aptamer Pools Across Cell Lines

In addition to CD20+HEK and the untransfected HEK control cells the lymphocytic cells lines, CCL-86 and TIB-152, were used. Adherent cells were treated as above, while lymphocytic cells were harvested from suspension. 500,000 cells were aliquoted into micro-centrifuge tubes and incubated for 30 minutes with aptamer pools 1 and 10 at 37°C for 30 minutes. Cells were washed with PBS, re-suspended in buffer and evaluated using a Beckman Coulter FC500 flow cytometer where 40,000 events were read. The MFI values of each curve were measured using Free Flowing Software and tabulated using GraphPad.
2.2 **NGS Analysis and Evaluation of Aptamer Clones**

2.2.1 **Sample Preparation**

Aptamer pool 10 was amplified using Illumina specific bar codes (5' ACACTGTC). The DNA was run on a 3% ultrapure agarose gel and purified using the GeneJET extraction kit. The final amount of 200ng was combined with other samples and sequenced in a single lane of Illumina Mi-Seq by the Eurofins Genomics Company. Data was received as a compiled fastaq file, it was and was uploaded, converted to fasta format and analyzed using the online Galaxy database (93-95). Data was clipped for length and collapsing to amalgamate sequences. To analyze sequence dependent relationships and build the phylogenetic trees the sequences were uploaded and analyzed using Clustal Omega (96). Their website is found at http://www.ebi.ac.uk/Tools/msa/clustalo/. Motif analysis was performed online using the DREME motif analyzer (97) part of the MEME NCBR database at http://meme.nbcr.net/meme/doc/dreme.html. The shape of the candidate sequences was analyzed using the RNAstructure platform (98), found online at http://rna.urmc.rochester.edu/RNAstructureWeb/.

2.2.2 **Aptamer Clone Assessment**

Candidate aptamers NLA-1 to NLA-4 were ordered as Cy-5 labelled clones from IDT (Integrated DNA Technologies). For clone assessment and \( K_d \) calculation the aptamers were diluted to the required concentrations, denatured at 95°C for 10 minutes and then snap cooled on ice. CD20+HEK cells were prepared as above by sloughing from plate, washing in PBS, and then re-suspending the cells in the diluted aptamer clone for 30 minutes at 37°C. Cells were
washed with PBS, re-suspended in buffer and analyzed using a FC500 flow cytometer where 40,000 events were read. For $K_d$ analysis the tabulated MFI were input into GraphPad software and the data fitted using a non-linear regression. $K_d$ value was measured per the formula $Y = B_{\text{max}} X / (K_d + X)$, where $B_{\text{max}}$ represents the maximal binding limit, $X$ is for the concentration of ligand and the $Y$ is the bound fraction.

### 2.2.3 Assessment of specificity

NLA aptamers were diluted to 300nM, heat denatured at 95°C for 10 minutes and then snap cooled on ice. The aptamer were then incubated separately with both CD20+HEK and the untransfected HEK cells for 30 minutes at 37°C. The samples were washed, re-suspended in buffer and analyzed using an FC500 flow cytometer. The MFI values of each curve were measured using free flowing software and tabulated using GraphPad.

### 2.2.4 Co-stain experiments

CD20+HEK cells were prepared as above. Cellular aliquots consisting of 500,000 cells were incubated with separately with initial DNA library, or a pool containing equimolar combination of the NLA aptamers and incubated for 30 minutes at 37°C. For the co-stain samples, FITC anti-CD20 antibody was added at a concentration of 10ng/µL and the cells incubated for an additional 30 minutes. The cells were washed and evaluated by a Beckman Coulter Gallios flow cytometer. The MFI readings in the FL1 channel, for FITC detection, were measured using Kaluza software and tabulated using GraphPad.
2.3 Biological Assessment

2.3.1 Initial CDC Assessment of CD20+HEK and CCL-86 Cells

Adherent transfected CD20+HEK cells were collected as above. CCL-86 cells were collected from cellular suspension. Both cells line were washed in PBS separately and re-equilibrated by incubating in un-supplemented media for 1 hour prior to use. Initial viability measures and cellular count were performed using Muse count and viability reagent. 150,000 cells were aliquoted into micro-centrifuge tubes and incubated with 10ng/µL of anti-CD20 antibody or 15 minutes. Cells were then topped off using either PBS, PBS with 50% fresh frozen human serum or PBS with 50% heat inactivated human serum. Heat inactivated human serum was generated by heating frozen human serum for 1 hour at 57°C as is referenced in Moore et al (99) and Lida et al (100). CDC was induced by incubated for an additional 4 hours at 37°C after which viability was analyzed.

2.3.2 Binding of Aptamer Clones to the CCL-86 cells

To evaluate the binding of the aptamers with CCL-86 cells, first the cells were collected from suspension and washed in PBS. After counting 200,000 cells were aliquoted into micro-centrifuge tubes and incubated with 2µM of either the DNA library or a pooled collection of the Cy5 labelled NLA aptamers. Cells were incubated for 30 minutes at 37°C, washed and then re-suspended in buffer. The cells were evaluated using a Beckman Coulter Gallios Flow Cytometer and analyzed using Kaluza Software. The tabulated data was performed with GraphPad.
2.3.3  Aptamer Mediated Inhibition of CDC

CCL-86 cells were collected from suspension and washed twice with PBS. Cells were re-equilibrated by incubating in un-supplemented RPMI-1640 media for 1 hour prior to use. Cells were counted and analyzed for initial viability using the Muse viability reagent, only cells with greater than 90% viability were used. Cell solutions containing 300,000 cells were aliquoted, spun down and re-suspended in 2µM aliquots of PBS diluted NLA aptamers or DNA library for 30 minutes at 37°C. Next CDC was induced using 10µg/mL of anti-CD20 antibody in a solution of PBS with either 50% fresh frozen human serum or, as a control, 50% heat inactivated human serum. CDC was induced for 4 hours at 37°C.

To assess viability in more depth we used the BD Pharmingen PE-Annexin V Apoptosis Detection Kit according to manufactures instruction. Samples from the cellular aliquots were added to 1X Annexin-V buffer and incubated with 7-AAD and PE labelled anti-Annexin-V antibody for 15 minutes in the dark. The cellular suspension was evaluated using a Beckman Coulter Gallios flow cytometer where 50,000 cells were measured. The data was analyzed using Kaluza software and the MFIs for FL2 (PE-annexin-V) and FL4 (7-AAD) were tabulated using GraphPad.
3 Transfection and Aptamer Pool Selection

3.1 Abstract

CD20 is a clinically important receptor and the target of several anti-cancer immunotherapeutics. HEK-293 cells, an ordinarily CD20 negative cell line, were stably transfected using a lentiviral system for the constitutive expression of this marker. Target positive (CD20+HEK) and target negative (CD20-HEK) were the basis for a CD20 specific cell-SELEX protocol. CD20 expression was verified using flow cytometry and fluorescence microscopy. 10 rounds of cell-SELEX with the CD20+HEK cells evolved a pool of high affinity and discriminative aptamers with increased binding associations to the transfected CD20+HEK cell line and not the original untransfected HEK cells. Pool 10, which contained the highest affinity binders, also showed heightened affinity to a naturally CD20 positive B-cell line, CCL-86, but not the naturally CD20 negative T-cell line, TIB-152. Therefore, SELEX using target positive transfected cells is an effective strategy for evolving high affinity discriminative aptamers.

3.2 Background

3.2.1 Characteristic of Lentiviral System

We employed a lentiviral based gene transfection system in order to generate CD20 target positive cells for aptamer selection. Lentiviral transfection methods are based on the prototypic lentivirus—HIV (101). But in the 20 plus years of their development, and in their current fourth generation, they have undergone significant genetic engineering designed to
increase safety, remove pathogenicity, broaden tropism, and even regulate GOI (gene of interest) expression (91).

3.2.1.1 Gene Expression is controlled by Modular Cassettes

Figure 3.1 summarizes some of the most significant modifications of lentiviral vectors within the last 20 years (91). Lentiviruses for gene expression must still operate as DNA transformation viruses, but within strict safety restrictions that all but eliminate the risk of creating replication competent lentiviruses. Doing so necessitated the modification and eventual segregation of the required genes onto modular cassettes. Cis-acting elements encode the genes required for RNA synthesis, packaging, reverse transcription and integration. These are wholly separated from the trans-encoding elements, which encode the structural and enzymatic proteins. The wide success of modern lentiviral vectors can be attribute to various innovations, some of which are briefly covered below.
Figure 3.1. The evolution of lentiviral vectors for the purposes of mammalian gene expression. Lentiviral transfection requires 3 key genetic components, here divided into cassettes. The envelope expression cassette determines tropism through the VSV-G capsid protein. The packaging and expression cassette carries the trans-elements and encodes the structural and enzymatic proteins. The vector expression cassette encodes the GOI and carries the cis-acting elements, which regulate its activity. Here, all 3 cassettes are shown transfecting HEK293T cells a human embryonic kidney cell line known for being highly susceptible and amenable to genetic engineering (91).
One of the first and most important developments was the substitution of the original HIV envelope genes to the VSV-G glycoprotein; see the envelope expression cassette in Figure 3.1. This was associated with notable improvements such as broadening tropism, directing endocytic entry, and stabilizing the vector during centrifugation (102).

The packaging/expression cassettes have undergone the most extensive modifications, note their numerous generations. They encode the trans-elements most critically the gag and pol genes which aid in the synthesis of the structural and enzymatic proteins required for functional vector particles. An early significant breakthrough, and one that helped to eliminate pathogenicity, was the removal of viral accessory genes vif, vpr, vpu and nef. This increased efficiency but, more importantly, without these proteins the viral vectors are rendered non-pathogenic (103).

The vector expression cassette encodes the cis-acting elements and, most importantly, the GOI; in Figure 3.1 it is indicated as the trans (transgene) cassette downstream of a promoter (promoter) element. The vector expression cassettes have also been modified to increase safety. SIN vectors, where SIN=self-inactivating, were generated by deleting the promoter and enhancer elements inherent to the 3’LTR. Because the 5’LTR is copied from the 3’LTR during reverse transcription, this deletion ensured that the integrated provirus remains replication incompetent (101). Another key modification, and a variation on SIN vectors referred to as cond-SIN, was the incorporation of the tetracycline-inducible element (Tet-Ind, see black arrow in Figure 3.1) to permit modular gene expression.
3.2.1.2 Conditional Gene Expression using Tet-System

Lentiviruses that incorporate the tetracycline inducible system permit the regulated control of GOI expression (104). The Tet system was derived originally from the tetracycline resistance operon in Escherichia coli (105). For the purposes of regulating gene expression, the original Tet-operon was reformulated and genetically engineered into two distinct systems. There is the original Tet-off using the tTA gene (tetracycline trans-activator) (105) or the Tet-on using the rtTA gene (reverse tetracycline controlled trans-activator) (106), they exert similar but entirely opposite actions, as depicted in the Figure 3.2. In Tet-off, transcription of tTA gene generates a TetR protein called tTA which will only bind the Tet-op promoter and transcribe the gene of interest in the absence of tetracyclines. The presence of tetracyclines, to which tTA affinity is very high, will turn expression of the GOI off. In the Tet-on or the rtTA system however, the opposite is true. Here transcription from the rtTA gene generates an rtTA protein which will only bind the Tet-op promoter and allow the expression of the GOI in the presence of tetracyclines.

Embedding TetO in close proximity to the GOI allows for modular gene expression. In either system, Tet-off or Tet-on, by titrating the amount of tetracycline, one can directly influence the degree to which the GOI protein is expressed (105, 106). As eponymous as it may be, expression of either system is most effective using doxycycline and not tetracycline. Doxycycline’s significant binding association (K_a=10^{10}) can activate tet-system elements at concentrations 100-fold lower than tetracycline (107).
Figure 3.2 The Tet system for conditional gene expression. The tetracycline inducible system permits modular gene expression. In tet-off transcription of tTA generates a tTA protein which enhances the transcription of GOI from the Tet-off site. In tet-off tetracyclines abrogate transcription. The affinity of tTA to the antibiotic is greater than its affinity for the promoter. In tet-on, transcription of rtTA generates an rtTA protein fused with a reverse transactivator element. Here, rtTA can only bind and promote transcription of GOI when tetracyclines are provided. ©Genoway

Figure 3.3. Generation of intact-GOI-viruses by a packaging cell line. Packaging cell lines produce infectious GOI- carrying virions. Here the packaging cell is shown as being transfected with all necessary gene-encoding cassettes for transcribing, producing and packaging live and infectious virions. These are harvested from the cellular supernatant for downstream application. ©Clontech Laboratories
3.2.1.3 Packaging Cells Produce the GOI-containing Lentivirus for Transfection

While the cassettes carry the vital genes, the generation of the actual virus must occur within a packaging cell (91). As Figure 3.3 illustrates, the packaging cell is where the archetypal viral gag and pol proteins precipitate the steps required for the recognition, transcription and packaging of competent viable virus particles. Infectious GOI-carrying viruses are harvested from the cellular media to transfect the choice target cells (101). Viral integration is one time event so target cells once transfected serve only to express, constitutively and stably, the GOI. To validate transfection, the transgene cassette often contains a selectable marker commonly for antibiotic resistance. Transfected target cells are grown in antibiotic fortified media in order to selectively isolate for cells that have been successfully transfected. Antibiotic selection is carried out for 1-4 weeks and is dependent on cellular turnover. Once cell growth has stabilized, conditional expression, using doxycycline, can be initiated.

Not all target cells are equally amenable to transfection. Characteristics including cellular type (adherent or suspension), growth rate, the composition of the cellular membrane, as well as internal regulatory pathways may all impact with what facultative ease or difficulty the target cell may be successfully transfected. HEK293 cells are human embryonic kidney cells and are one of the best established and most easily transflectable cell lines available (108).

Here, a lentiviral system was used to transfect CD20 cDNA into a HEK293 cells, a cell line normally absent for this marker. Cells were positively selected using ampicillin and CD20 expression induced using doxycycline. This resulted in the generation of CD20+HEK cells used for positive selection, and the untransfected HEK-293 cells for negative selection. The use of stable transfection of a known marker for the purposes of cell-SELEX has been suggested in
critical literature. However, to the best of our knowledge, this is the first time it has been
employed in a discrete fashion.

3.2.2 Stringency Measures

SELEX is only as effective as its conditions are stringent. The selection parameters exert a
direct influence on the quantity, quality, and specificity of the evolved ligands (109). The more
stringent the conditions the greater the binding affinity and discriminative capacity of the final
pool. During our selection regime stringency was increased incrementally every 2-3 rounds
until an appreciable increase in the binding affinity of the pool was achieved, see Table 3.
Stringency measures included increases in the number of washes, decreased incubation time
with the positive cells, increased incubation time with negative cells, and incrementally
decreasing the incubatory concentration of the DNA pools.
<table>
<thead>
<tr>
<th>Round Number</th>
<th>DNA Concentration (nM)</th>
<th>Negative Selection</th>
<th>Number of Washes</th>
<th>Incubation Time with Positive Cells (min)</th>
<th>Incubation Time with Negative Cells (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>N</td>
<td>1</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>2-3</td>
<td>500</td>
<td>N</td>
<td>2</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>3-4</td>
<td>250</td>
<td>N</td>
<td>2</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>4-6</td>
<td>250</td>
<td>Y</td>
<td>2</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>6-8</td>
<td>250</td>
<td>Y</td>
<td>3</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>8-10</td>
<td>100</td>
<td>Y</td>
<td>3</td>
<td>20</td>
<td>40</td>
</tr>
</tbody>
</table>

**Table 3. Selection regime.** Increases in the stringency of selection aid in evolving high affinity aptamer pools. Here stringency is increased incrementally over 10 rounds of selection. Positive and negative cell line denote, respectively, the CD20 transfected HEK cell line and the un-transfected HEK cell line. For round 1 the initial DNA pool is the unselected DNA library. Stringency measures included limiting concentration, increasing washes and varying incubatory times with cells.
3.3 RESULTS

3.3.1 Evaluation of CD20 Transfection

A lentiviral gene expression system was used to stably transfect HEK293 with CD20 cDNA. After selection using puromycin and induction of GOI using doxycycline the cells were evaluated for CD20 expression.

3.3.1.1 Flow Cytometry

Figure 3.5 panel A depicts the dot plots of untransfected control HEK and the transfected CD20+HEK cells before and after staining with fluorescently labelled anti-CD20 antibody. The X and Y axis represent respectively CD20 expression and internal complexity (SS). Cells exhibit consistent internal characteristics indicated by their occupying a relatively uniform distribution. Staining CD20+HEK with anti-CD20 antibody causes a shift in the cells population along the X-axis - this denotes greater antibody binding. A similar staining of the HEK control cells however did not induce a shift. Therefore anti-CD20 antibody positively label the CD20+HEK and not the untransfected control. This is also represented with the histogram in Figure 3.5 panel B which depicts an overlay composite comparing the antibody signal for both samples the HEK control (grey filled) and CD20+HEK cells (white filled). The significant shift confirms that CD20+HEK cells exhibit elevated amounts of anti-CD20 binding compared to the HEK control cells. Therefore, the lentiviral transfection of cells is successful with a CD20+HEK cell line established.
Figure 3.5. Assessment of CD20 expression in transfected HEK293 Cells. To evaluate and compare CD20 expression between untransfected and CD20+HEK transfected samples, live cells were incubated with 10ng/µL of FITC labelled anti-CD20 antibody for 30 minutes on ice. The cells were washed, re-suspended in buffer and fluorescence measured using a Beckman FC500 flow cytometer where 20,000 events were counted. A) Representative dot plots of cells before and after antibody incubation. B) Relevant histograms of the antibody labelled cells. Only the transfected CD20+HEK cells are positively labelled by the anti-CD20 antibody.
3.3.1.2  Fluorescence Microscopy

To further verify transfection we visualized antibody labelled cells using fluorescence microscopy. We stained cells with FITC-labelled anti-CD20 antibody (excitation 495nm; emission 519nm) and with DAPI (4',6'-diamidino-2-phenylindole) a DNA intercalating dye (excitation 358nm; emission 461nm). In live or fixed cells, DAPI fluoresces with strict localization in the nucleus. The results of the microscopy are presented in Figure 3.6. Only CD20+HEK cells show the distinct anti-CD20 labelling, the HEK control does not. CD20 is a transmembrane protein. As the overlay figure CD20+HEK cells exhibit anti-CD20 staining along the cellular membrane, clearly delineated from the DAPI stained nucleus. Together this confirms the successful expression of the CD20 molecule in the CD20+HEK cell line.
Figure 3.6. Visual detection of CD20 expression in transfected HEK293 Cells. Fluorescence microscopy for the visual detection of CD20 was performed 2 weeks after induction using doxycycline. Cells were fixed, stained with FITC anti-CD20 antibody at 50ng/µL for 4 hours. Cells were co-stained with DAPI for nuclear elucidation. Only the CD20+HEK cells exhibit CD20 positive labelling.
3.3.1.3 **Contextual Anti-CD20 Expression**

For context, we titrated the anti-CD20 antibody with CD20+HEK cells and compared it to the level expressed with CCL-86 a naturally CD20 expressing cell line and TIB-152 a naturally CD20 negative cell line. The results are depicted in Figure 3.7. Signal intensity is quantified by the median fluorescence intensity or MFI. The higher the MFI value the greater the signal strength.

Note the consistently low MFI values with the TIB-152 cells. This demonstrates that the antibody fails to bind to TIB-152 cells to any appreciable degree regardless of the concentration. TIB-152 cells therefore do not express CD20. Unlike the CCL-86 cell line which shows substantially higher MFI levels. And while the CD20+HEK cell line exhibits significant CD20 staining it is saturated at levels lower than that observed with the CCL-86 cells. This suggest that while the transfected cells do possess the CD20 molecule it is not as abundantly expressed as on the CCL-86 cells.
Figure 3.7. CD20 expression in different cell lines. To contextualize relative CD20 expression the anti-CD20 antibody was titrated using the CD20+HEK the transfected cell line, CCL-86 a naturally CD20 expressing cell line, and TIB-152 a naturally CD20 negative cell line. The antibody signal for CD20+HEK cells is saturated at much lower MFI intensities than with the naturally CCL-86 cells suggesting that CD20 is less abundantly expressed with the transfected cell line.
3.3.2 SELEX generates pools of aptamer with progressively greater binding affinity to the CD20+HEK cells.

10 rounds of selection using CD20+HEK cells were performed. To evaluate selection the CD20+HEK cells were incubated with fluorescently labelled aptamer pools derived from rounds 1, 5 and 10 which represent respectively the start, middle and end points of selection. As a control the cells were also stained with the unselected DNA library. The individual binding results are indicated in Figure 3.8A. Unstained cells are depicted by the grey filled histogram. Note that cells incubated with the library (grey curve) exhibit a shift in fluorescence; this represents the basal binding capacity of the library. Each subsequent round of selection, round 1 in red, round 5 in green and round 10 in purple, exhibit increasing rightward shifts, with pool 10 possessing the greatest movement. Therefore the selection regime is successful in generating pools of aptamers with cycle dependent increases in affinity with the CD20+HEK cells. For clarity, Figure 3.8B is an overlay histogram of all the samples. The MFI values of each curve are tabulated in Figure 3.8C. Unstained cells have a basal MFI value of 2, the library possess an MFI of 22, pool 1 at 26, pool 5 at 32 and Pool 10 with the strongest MFI value at 46. The signal intensity of pool 10 was found to be statistically significant when assessed by the T-test. This progressive increase in MFI, as a consequence of the selection round performed, shows that the SELEX protocol successful evolved stronger binding pools of aptamers with affinity to the CD20+HEK transfected cell line.
Figure 3.8. Binding affinity of aptamer pools with CD20+HEK cells. To evaluate the SELEX process CD20+HEK cells were incubated with 200nM aliquots of Cy5 labelled DNA library, aptamer pools 1, pool 5 and pool 10. A) Individual fluorescence curves for CD20+HEK incubated with library (dashed grey), aptamer pool 1 (red), aptamer pool 5 (green) and aptamer pool 10 (purple). Unstained control cells are represented by grey filled curve. B) Overlay composite of the histogram for all samples. C) The tabulated MFI values of each curve for all samples. Cumulatively this data demonstrates that selection using CD20+HEK cells resulted in the sequential increases in aptamer pool affinity for transfected cells. Aptamer pool 10 binds CD20+HEK cells with significantly more affinity than the unselected DNA library.
3.3.3 Pool 10 CD20 positively labels the naturally CD20 expressing CCL-86 Cells but not the naturally CD20 negative Cell line TIB-152.

To better evaluate the CD20 specificity of the aptamers we incubated pools 1 and 10 with independent cell lines of varying CD20 expression. TIB-152 is a T-cell line and naturally CD20 negative; CCL-86 cells are B cells (Burkett’s lymphoma) and are naturally CD20 positive. As can be seen from tabulated MFI data in Figure 3.9, pool 1 is a low affinity binder of all 4 cell lines with relatively consistent MFI values in all. As expected, Pool 10 is a substantially greater and therefore discriminative binder of CD20+HEK cells than the untransfected HEK.

In the independent cell lines, neither Pool 1 nor pool 10 possessed any significant binding with the CD20 negative cell line TIB-152. Therefore the aptamers selected are not for markers prevalent on TIB-152 cells. When evaluated with the naturally CD20 expressing CCL-86 cells only Pool 10, and not Pool 1, positively associated with CCL-86 cells. Pool 10’s MFI with CCI-86 is 38, and was a statistically significant association when evaluated with the student T test.

Pool 1’s low affinity with all the cell lines shows that selection has not yet cultivated sufficient target sensitivity. By round 10 the aptamer pool show a heightened affinity for CD20+HEK and not their target negative HEK counterparts. When compared to the independent CD20 positive cell line CCL-86, only pool 10 and not pool 1 possessed significant association. Together this data demonstrates that pool 10 appears to harbour aptamer sequences with specific binding affinity to the CD20 molecule. Therefore, cell-SELEX using transfected target positive CD20+HEK cells is an effective method for generating high affinity and discriminative pools of aptamers.
Figure 3.9. Binding affinity of aptamer pools 1 and 10 with cell lines of varying CD20 expression. To evaluate the specificity of the aptamers and investigate their selectivity for CD20, the binding affinity of pools 1 and 10 (200nM) were compared amongst the CD20+HEK, the untransfected HEK cells, and two unrelated independent cell lines: the naturally CD20 expressing cell line CCL and the naturally CD20 deficient TIB. By their tabulated MFI analysis both CD20 expressing cell lines, CD20+HEK and CCL-86, exhibited increased affinity with pool 10 and not with pool 1. Neither pool 1 nor pool 10 possessed significant binding with either TIB or the untransfected HEK cells. Therefore aptamer pool 10, derived by selection using CD20+HEK can positively identify and validate CD20 expression in independent cell lines.
3.4 Discussion

SELEX is a well-established method of generating high affinity DNA ligands. Solid-state selection where the target of interest is conjugated onto a membrane or bead may allow for the target specific generation of aptamers however it fails to replicate the complexity of a cellular setting. Cell-SELEX is a more physiological conscious method. Aptamers are selected in a more natural environment and are better adapted for future in-vitro and in-vivo analyses. However, a significant caveat of most cell SELEX is that while aptamers can be generated with high specificity and affinity for a particular cell type—the actual identity of the intended molecule to which the aptamer binds is not known. This represents a significant hurdle as without specific target identity, the functional role of an aptamer is limited.

Lentiviral transfection of genes represents one of greatest advancement in modern biological engineering. We generated live and infectious CD20 carrying lentiviruses. These were subsequently used to introduce the CD20 gene into HEK293 cells—a highly amenable cell line normally absent for this marker. The result was the formation of 2 cell lines, the negative control consisting of the untransfected HEK and the positively transfected CD20+HEK. These inherently identical cell lines but for the expression of CD20 represent a novel cell-SELEX strategy that couple target specificity and a physiologically relevant environment.

The transfection of HEK cells was evaluated using both flow cytometry and fluorescence microscopy. After the requisite negative screening with puromycin and gene induction using doxycycline we evaluated the cells for CD20 expression. In Figure 3.5 flow cytometry using anti-CD20 antibodies verified the expression of CD20 only on CD20+HEK and not the HEK
control. Fluorescence microscopy, in Figure 3.6, further validates the expression of CD20 on the cell membrane of CD20+HEK. Together these verify CD20 transfection and the production of a CD20+HEK cell line. By titrating anti-CD20 antibody in Figure 3.7 we were better able to contextualize CD20 expression. We found that CD20 expression in the transfected cell line is less abundant than CD20 expression with CCL-86 cells. This is understandable as CCL-86 are a lymphocytic B-cell line, and CD20 a characteristic B-cell marker. This also explains its significant absence in TIB-152 cells, a T-cell line.

Using both positive and negative cell lines we performed 10 rounds of alternating cell-SELEX, selecting for aptamers that associate with the CD20+HEK cells and eliminating those whose specificity was for markers on HEK cells. The hallmark of SELEX is the iterative enrichment and the stepwise generation of target specific pools. As shown in Figure 3.8 pools 1, 5 and 10 show gradual increases in affinity for the CD20+HEK cell lines. This demonstrates that the stringency measures were effectual. By increasing stringency and therefore the selection pressure it encouraged the retention of progressively higher affinity ligands. Pool 10 was the highest affinity pool with significantly greater to the CD20+HEK than the initial library. This confirms that the selection of aptamers specific to CD20+HEK was successful.

Equally important is to show target, not cellular, specificity. In Figure 3.9 we evaluated the specificity of the aptamer pools taken from the very first and very last rounds of selection against independent cell lines with differential CD20 expression. As expected, Pool 10 showed the most substantial increases in binding with the CD20+HEK. There was also a small increase in the binding of pool 10 with the control HEK cells. This increase from MFI$_1$=25, MFI$_{10}$=30, is not statistically significant but it does suggest that some of the aptamers in pool 10 could be
general binders of HEK cells and not specific for the CD20 molecule. That these aptamer were able to persist regardless of stringency suggests the existence of a basal threshold limit.

When evaluated with the independent cells, we showed that only sequences in aptamer pool 10 and not pool 1, exhibited enhanced binding to CCL-86 cells. Neither pool 1 nor pool 10 appear to associate with TIB-152 cells to any appreciable degree. As CCL-86 cells are CD20 positive and TIB-152 cells CD20 negative this appears to corroborate that target positive cell-SELEX can raise aptamers specific only to the transfected marker. Pool 10 validates CD20 expression in 2 entirely distinct cell lines. It positively labelled the naturally CD20 expressing CCL-86 but not the CD20-absent TIB-152. Cumulatively this strongly suggests that lentiviral transfection of CD20 and the subsequent targeted cell-SELEX method were successful in generating a pool of aptamers possessing high affinity and specificity to the CD20 molecule.

3.5 CONCLUSION

A lentiviral CD20-specific cell-SELEX strategy was used to evolve aptamer pools with heightened affinity and specificity to the transfected marker. Here, cell-SELEX using CD20+HEK cells evolved aptamers pools that showed incremental increases in binding affinity with the CD20+HEK cell line and not to the untransfected control. Pool 10 the most significant binder of CD20+HEK, and not Pool 1, showed increased affinity with CCL-86, a naturally CD20 positive cell line and not TIB-152 a naturally CD20 negative cell line. This suggests that selection using transfected CD20+HEK evolved an aptamer pool specific to the CD20 marker.
4 NEXT GENERATION SEQUENCING AND THE EVALUATION OF APTAMER CLONES

4.1 ABSTRACT

Pool 10 was sent for next generation sequencing (NGS) to elucidate and quantify individual aptamers. Sequences were analysed using the web based bioinformatic platform Galaxy. Additional analysis included phylogenetic examination using Clustal Omega, common motif evaluation performed using DREME, and secondary structure derived using RNAstructure. NGS showed that Pool 10 exhibited significant sequence convergence with over 70% of all sequences grouped into 5 clusters. Enrichment was demonstrated by the presence of exceptionally high copy number sequences (HCN) found in each cluster. DREME detected 3 common motifs, the most pervasive of which was identified in 84% of all sequences. This demonstrates that selection was successful in cultivating a consolidated pool of aptamers.

The highest copy number (HCN) sequence of each cluster were synthesized as aptamers NLA-1 through NLA-4. All aptamers exhibited a high binding affinity, with $K_d$ values less than 100nM. Clones NLA-3 and NLA-4 had $K_d$s less than 60nM, and possessed the most specific binding to CD20+HEK cells. Co-staining cells with labelled antibody and aptamers resulted in a decrease of antibody signal intensity, suggesting that the aptamers and antibody compete for binding at mutual sites. This data confirms that selection using target positive transfected cells and the analysis of the resultant data using NGS was successful in isolating for discrete aptamer sequences that possessed heightened affinity and specificity to the CD20+HEK cell line.
4.2 BACKGROUND

4.2.1 Deconvolution of Aptamer Pools

Aptamer pools are by definition heterogeneous (110). The goal of selection is to condense a massive combinatorial library into a highly specific aptamer pool containing sequences with heightened affinity to the target of interest. The goal of sequencing is to transform that pool into discrete and individual aptamers to critically evaluate their attributes including structure, shape, interaction dynamics, as well as any biological considerations. Once the initial sequence is known its characteristics can be enhanced with modifications. Macugen- a clinically licenced aptamer for the treatment of wet age-related macular degeneration (72, 111), Pegnivacogin- an anti-coagulative aptamer developed for acute coronary syndrome (112), and Spinach a fluorescence activating RNA aptamer for live cellular tracking (113) are all aptamers whose optimization required extensive post-selection modifications. These included nucleotide substitution to resist nucleases, truncation to fold into more stable secondary structures and pegylation to increase retention. Selection may yield a high affinity pool but it is the modifications of individual sequences that truly refine an aptamer potency and efficacy.

The acquisition of high quality aptamers is significantly hampered by the inherent sequence complexity that persists even in an evolved pool (114). Though selection may eliminate the majority of sequences in the library, the final pool could still contain thousands of different species. Together the sequencing of aptamer pools and the requisite screening assays to evaluate functionality have represented, until very recently, the two largest hindrances in efficient aptamer development (86).
Next Generation Sequencing (NGS) coupled with bioinformatics is a very powerful high throughput method of analysis; and has dramatically changed how large-scale genomic studies are performed (115-117). It’s also proven equally illuminating with respect to aptamer sequencing where NGS can identify, count, and categorize each discrete DNA molecule. This information can then be used to evaluate all the aptamers contained in the pool, to assess the convergence of popular motifs, and chart the evolution of high affinity sequences (110). It has also been used to better understand the factors, both inherent and obscure, that operate on SELEX, the effect that negative selection can have (79, 118), how stringency affects enrichment (114), and how PCR amplification may exert undue influence on ligands (23, 119). Its speed, output and applicative potential have seen it become the preeminent strategy for aptamer sequencing.

4.2.2 Aptamer Sequencing Past Methods: Introduction to Conventional Bacterial Cloning and Sanger Sequencing

SELEX was developed in the 1990s with most aptamers of that time sequenced using bacterial cloning and Sanger sequencing (59, 111, 120). Both are labour intensive and low-throughputs and are considered by todays standards wanting.

Transformation is the acquisition and stable replication of exogenous pieces of DNA. A common phenomenon in bacteria, it was the pioneering work of Stanley Cohen et al in 1972 (121) which brought it to the forefront of molecular cloning. In bacterial transformation the DNA of interest is subcloned into circularized vectors, which are introduced into viable and competent bacterial cells either by a chemical method like calcium chloride exposure or a physical method like electroporation. See Figure 4.1 for details. Transformed bacterial cells are
plated on antibiotic fortified agar at a concentration low enough to ensure single colonies are the result of only one progenitor. The multiple cloning site usually lies within a reporter gene, like LacZ, which makes it easy to phenotypically evaluate between bacterial colonies carrying a recombinant vector positive for the target DNA, and bacterial cells carrying an empty vector which does not (122). Functional LacZ encodes a protein that activates β-galactosidase which has the capacity to cleave the molecule X-gal generating a blue pigment (122). The successful incorporation of insert DNA at the MCS however, will not yield active β-galactosidase, as a result X-gal remain uncleaved, and the bacterial colony is white.

Bacterial cloning and amplification remain a mainstay of modern molecular biology, even well after the advent of PCR. Bacterial logarithmic growth coupled with a high copy number plasmid can, in the space of less than 12 hours, massively amplify plasmid DNA from the initial nanogram amount used for transformation to microgram quantities after purification. What differentiates this from PCR, is that genes amplified within a bacterial system have the benefit of proof reading and repair exonucleases. PCRs taq polymerase is notoriously error prone with an error rate among the highest of all polymerases at $2 \times 10^{-4}$ errors per nucleotide synthesized (123, 124); it also lacks repair mechanisms to correct mis-incorporated nucleotides (125). As a result, whenever there is a requirement for high fidelity gene synthesis bacterial cloning is still widely employed.
**Figure 4.1. Bacterial cloning.** In bacterial cloning the DNA of interest is first introduced into a circularized vector typically within the MCS of a reporter gene like LacZ. Restriction endonucleases cleave double stranded DNA at palindromic sequences. Through reciprocal end ligation and annealing a recombinant vector carrying the gene of interest is generated. Ligated plasmid DNA is used to transform bacterial cells which uptake and maintain the vector, replicating it with each and every cellular division. Bacteria are plated on antibiotic fortified plates where successful recombinant colonies are easily distinguished by their white colour. Copyright Commons.
Bacterial cloning and Sanger sequencing were techniques that evolved together. And up until the early 1990s the latter was a mainstay of all sequencing technologies. In Sanger sequencing, DNA is amplified with a mastermix spiked with chain terminating dideoxynucleotide triphosphates (ddNTPs). Unlike natural deoxynucleotide triphosphates (dNTPs), the incorporation of the 3’OH lacking ddNTPs abrogate any further extension of the complementary strand (126). Due to the stochastic incorporation of ddNTPs the result is the amplification of DNA strands prematurely terminated at every possible nucleotide position. Using gel or capillary electrophoresis the DNA strands are sorted according to molecular weight (116), and differential labelling of each ddNTPs permits the ordered reading of the DNA sequence in a nucleotide specific manner.

High fidelity bacterial cloning and Sanger sequencing are highly precise techniques, genes and genome sequenced in this manner have a significantly reduced error rate achieving 99.999% fidelity (126). And for this reason, the Sanger method was the gold standard for nucleic acid sequencing for more than two and half decades (117) but its limited approach, restricted automation, and prohibitive costs have seen it replaced with the cheaper, faster and massively parallel techniques collectively identified as “next generation sequencing” or NGS.

4.2.3 Introduction to Next-Generation-Sequencing (NGS) Technologies

NGSs are high throughput approaches, allowing for the simultaneous sequencing of millions of DNA fragments (117) at a fraction of the costs and time of conventional Sanger sequencing. And although only commercially available since 2004 (127) the high throughout nature of these techniques, their amenability to automation, and their capacity to uncover and identify previously undiscovered genes, regulatory elements and even expression trends have shown
them to be critical and important tools for all aspects of genome analysis. All of which has prompted their use in the exploration of aptamer pools as well. Of the 4 major platforms for high throughput analysis (HTA), the Illumina/Solexa Genome Analyzer is the most economical and accessible to researchers.

4.2.3.1 **Illumina Overview and Use of Bioinformatics to Extract Meaningful Information**

Like all HTA methods, a typical Illumina procedure incorporates 3 elements, template preparation, sequencing coupled imaging and data analysis (115). These are indicated in Figure 4.2. First the template DNA is amplified using specific 5’ and 3’ adapter sequences generating an Illumina specific adapter library. Solid synthesis immobilizes the same adapter oligomers randomly on the surface of a glass flow cell (115), see panel 4.2A. The adapter-amplified template library is denatured and distributed at a low concentration over the immobilized adapter sequences to which it will hybridize (116). Unlabelled nucleotides and DNA polymerase are added to build the complementary strand via bridge amplification, the process by which the original strand will bend to hybridize with its cognate primer. This ensures the clonal amplification of the target sequence DNA culminating in dense clusters containing several thousand to a million copies of one original template strand. This amplification is required to ensure adequate signal detection in subsequent steps.

For sequencing and imaging, see Figure 4.2B, Illumina uses cyclic reversible termination (CRT), an approach similar to Sanger’s chain terminating ddNTPs (115). In lieu of irreversibly removing the 3’OH, in CRT the 3’OH group is simply blocked. This ensures that the strand can only be synthesized one fluorescently labelled nucleotide at a time. This type of approach is known as “sequencing by synthesis”. Initially all four nucleotides are bathed onto the surface of the flow
cell, only the complementary nucleotide to the template will bind. The laser light is shone and
detection of the fluorescently labelled bound nucleotide is measured, because the 3’OH is
blocked no further extension of the strand is possible. To initiate a new cycle, the remnant
nucleotides and enzymes are washed away, the 3’OH chemically unblocked, new labelled
nucleotides added with the detection of the next newly incorporated base. This cycle is
repeated until the strand is fully sequenced.

Modern NGS strategies outperform the older Sanger sequencing methods by factors of 100-
1000 (116). The rapid acquisition of sequences—millions of nucleotides in length—is now possible
within weeks. In order to extract meaningful information bioinformatics software is required.
To align sequences against reference genomes there is BLAST and its many permutations
including blastn for genomes, blastp for proteins, blastx for cDNA which are publically
licenced, freely available over the internet, and updated by researchers. Multiple alignment
software, like clustalw and clustal omega, are pairwise sequence alignment tools and can
inform on the homology of related sequences, protein or DNA. Motif analysis tools for
consensus sequences discovery include MEME and gibbs motif sampler. DNA and RNA folding
software like mfold and RNAstructure are also available to predict shape, structure and
stability of sequences. Software and tools like these are important and pivotal for research;
both their current applications and future endeavours will continue to be advanced by next
generation sequencing.
Figure 4.2. Illumina/Solexa Overview. A) Clonal amplification of sample. Target DNA is amplified using Illumina specific adapter primers and bathed onto a flow cell bearing identical immobilized primers. Bridge amplification ensures the clonal amplification of sequences resulting in high density template DNA clusters. B) Sequencing by synthesis. Blocked dNTPs are provided so that DNA polymerase can catalyze their singular hybridization to the template strand. The incorporation of a nucleotide coincides with a unique fluorescence signal which is detected and logged by computers. Every nucleotide addition is an isolated event requiring molecular cleavage for further extension. In this way the complementary DNA strand is sequenced as it is being synthesized- one base pair at a time. Michael L. Metzker. Sequencing Technologies- the next generation. (115)
4.2.3.2  **NGS as applied to Aptamers**

Until relatively recently (2010-onwards) most aptamers were derived only after bacterial cloning and Sanger sequencing. As slow and expensive as Sanger sequencing is, it was not the biggest hurdle to overcome. Practically speaking, bacterial cloning is inherently limited. In a pool of, conservatively estimating, hundreds of different species, the requirement to grow thousands of bacterial clones to ensure ample coverage and sequence each one individually is prohibitive. In fact most publications estimate an average of 50 clones are sequenced (17). There is then no guarantee that the sequences acquired with bacterial cloning are true representative of the pool as a whole nor that they’ll be target specific and efficacious. The risk of sequencing artifacts, aptamers inadvertently raised against a procedural reagent or those that persisted in the pool because they exhibit a PCR advantage, is a considerably more significant concern when the total breadth of analysis is limited to only 50 clones.

The advent of NGS has proven significantly beneficial to aptamer sequencing for two main reasons. Firstly, it serves as a highly effective analytical tool capable of thoroughly reporting on all the sequences contained. Bioinformatics can refine the data by describing their abundance, length, copy number and even motifs. By evaluating all the sequences based on objective criteria it greatly simplifies the analysis and aids in the acquisition of high quality aptamers. Secondly, NGS has probative considerations. It can critically evaluate selection parameters and their influence on cultivating a high affinity pool. By sequencing an aptamer pool by NGS, it makes for a more comprehensive study of the trends and patterns evolved during selection and can allow for the independent evaluation of sequences in a quantitative and quantitative manner.
4.3 RESULTS

4.3.1 NGS Reveals that Pool 10 Exhibits Significant Sequence Convergence and Motif Emergence

4.3.1.1 Phylogenetic Assessment of Sequences Identifies 5 Distinct Clusters

Pool 10, the greatest binder of the CD20+HEK cells was sent for NGS analysis. Illumina read a total 17,613,460 discrete sequences; of these a total 627,826 were clipped to the barcode specific to CD20 Pool 10. These sequences were clipped by length, isolating for aptamers 100 nucleotides in size- the original length of the library. After clipping, a total 541,258 total reads were retained. This corresponds to 86% of total sequences, therefore 13% or 86,268 sequences were discarded based on length alone. To evaluate copy number or the total frequency of specific sequences the data was collapsed. This removes redundancies by merging and identical sequence together. Collapsing the data reduced the total number from 541,258 individual sequences to 62,737 collapsed reads.

Figure 4.3 is a phylogenetic representation of the top 29 collapsed sequences. This was performed using Clustal omega, a pairwise alignment software. Every sequence is represented numerically with both an identification number as well a copy number indicating its frequency. For example the first sequence in the analysis below is #31-438, where 31 is the identification number and 438 is its copy number. Therefore, in pool 10 sequence 31 is present 438 times. Clustal-Omega is ordinarily confined to the comparative analysis of genes and proteins whereby constructing a phylogenetic tree mutational trends can be charted and measured. In much the same way here, Clustal-Omega was used to analyze the similarity and differences
between aptamer sequences. Clustered groupings indicate closely related sequences while outliers represent isolated and therefore unrelated sequences.

The top 29 sequences of pool 10 range in copy number from 84 at the scarcest to 11,019 at their most plentiful. Cumulatively, the total copy number of all 29 sequences is 43,700 which reflects about 70% of all sequences. The residual sequences were excluded from analysis due to their low copy number. As can be seen from Figure 4.3 these sequences can be broadly divided into 5 distinct groups indicated as Groups A-E. Each cluster is defined by a different common ancestor (black arrow) which spawned a series of closely related descendants. With the exception of only 1 cluster, group E, each group possesses a significantly high copy number sequence (HCN) indicated by grey stars. SELEX relies on an iterative selection process therefore, the greater the copy number of a sequence, the more persistent or enriched it was in selection, and so presumably the more specific it is to the target CD20+HEK cells.

Groups A and C possess only a single sequence with a dominant high copy number (HCN). While groups B and D both possess 2 HCNs. In group A the HCN sequence is #7 with a copy number 2826, in group C it is sequence #6 with the highest of all copy numbers 11,019. Group B possess 2 HCNs, as indicated they are sequences 12-5810 and 13-7036. Group D’s HCNs are sequence #23-2958, and sequence #5 with copy number 4983. There does not appear to be any relationship between the number of HCNs and the size of each cluster.
**Figure 4.3. Pool 10 exhibits significant sequence convergence.** Indicated above is the phylogenetic assessment of the top 29 clipped, merged and collapsed sequences in Pool 10 performed using Clustal omega. In each instance the identification number precedes the copy number. For example in 31-438, 31 is the arbitrary identification number of the sequence and 438 is its copy number. Sequences were resolved into 5 distinct clusters indicated as groups A through E. Black arrows denote the a hypothetical common ancestral sequence preceding each group. In every grouping is one or more significantly high copy number sequences (HCNs, indicated by grey stars). As the most over-represented species of each cluster, HCNs best reflect the inherent characteristics of pool 10.
4.3.1.2 *HCN Sequences are Significantly Preferred to their Closest Relation*

Though sequences are grouped into different clusters the extent of variation within each grouping is not great. And yet in spite of the sequence similarity all HCN sequences exhibit significantly higher frequencies when compared to their closest relation. A list of all the HCN derived from groups A-D and their closest related sequence are shown in Figure 4.4. The first panel for Group A depicts its HCN #7-2826 and its closest relation #32-291. For simplicity both the forward and the reverse primers have been omitted, only the randomized N60 portion is represented. Because they are clustered within the same family the two sequences are virtually identical. The differences being that sequence #7-2826 possess an A at position 3 and G at position 9, while sequence #32-291 at the same sites bears a C and T respectively. These are indicated in bold. When comparing copy numbers, sequence #7 is favoured at a ratio of 9:1 to sequence number #32. Therefore though the sequences are highly similar, there is a marked preference for sequence #7. This trend is typical of all of the high copy number sequences, most impressively by sequence #6 in group C where the copy number ratio between it and its closest relative, sequence # 48, is more than 131:1. Here a guanosine residue is favoured to a thymine.

While groups A and C contain only one dominant HCN, the case is slightly more complicated with Groups B and D which both possess 2 HCN per cluster. The observed trend here remains the same. HCNs persist in being the favoured sequences when compared to their closest relation; however, there are no similar biases between intra-HCN sequences. Group B’s dominant HCN sequences are #12 and #13. Considering for the moment only sequence #12 and its closest relation #45, the sole difference is a single nucleotide change a C-T transition,
with sequence #12 is favoured by a ratio of 47:1 to sequence #45. An intra-HCN comparison between sequences #13 and #12, identifies a sole change- a G-T modification. Recall that a G-T modification was also present in group C and gave rise to the largest frequency variation (between sequence #6 and #48). Here, a G-T modification does not appear to significantly impact the frequency of either sequence as the frequency difference between sequence #12 and #13 is a paltry 1.2:1. Whatever impact on enrichment the G-T modification exerted for group C and sequence #6 it does not appear to have been extended here. Group D also possesses 2 HCN sequences #5 and #23. Sequence #5 is preferred at a ratio of 6:1 to sequence #41- the lowest difference for any HCN encountered. While sequence #23 is favoured at a ratio of 21:1 to sequence #42. Like group B, an intra-HCN comparison between #5 and #23, identified an A-C modification at position 3 and a G-T modification at position 8. The difference in frequency between HCN #5 and #23 is a modest 1.6:1.

The phylogenetic assessment grouped the top 70% of all sequences into 5 distinct groups with the HCN sequence being the most dominant of each cluster. All groups exhibited significant sequence similarity with the difference between an HCN sequence and its closet relative sequence limited to 1-2 nucleotides. This demonstrates the aptamer pool 10 is enriched for very particular sequences. In spite of extreme sequence identity, however, all HCN exhibited a significant copy number bias in their favour. Therefore though sequences are similar, selection exerted a particular preference resulting in the over-representation and retention of highly specific sequences.
Figure 4.4. Sequence comparison of HCNs and closest related sequence. Depicted above are the 60-nucleotide internal sequences of each HCN(s) and its closest related sequence. Groups A and C possess only 1 HCN (sequence #7 and #6 respectively). Both Group B and D possess each 2 HCN sequence (group B sequence #12 and #13; group D sequence #5 and #23). Nucleotide differences are shown in bold if they are between the HCN and its closest related sequence, and underlined when they represent intra-HCN nucleotide differences.
4.3.1.3 DREME identifies 3 Distinct Motifs

To further refine analysis the top 77 collapsed sequences were subjected to motif analysis using DREME. DREME is the discriminative regular expression motif elicitation and can be used to discover short, ungapped motifs within DNA. The information is consolidated into a picture showing each motif, with the frequency of each nucleotide indicated by its size. Aptamer pool 10’s motifs are presented in Figure 4.5. The first motif is GGRCAR where R is any purine residues, this was by far the most prevalent motif found in 66 out of 77 total sequences. Of both variations, the GGACAR variant was present in 36 of the collapsed sequences, while the GGGCRA was found in 29. The second most common motif, but the longest in size, was CAMTCA where M represents either cytosine or adenine. This sequence was less common, found in a total of 40 of 77 sequences, however here the CACTCA variant is more than twice as represented then the CAATCA. The third most common motif and the shortest is GWARA, where W can represent either thymine or adenine. This motif was found in 52 of 77 of all sequences, with the GTARR variant heavily favoured, as is also indicated pictorially. The abundance of common motifs is another indication that selection was successful in generating a highly specific enriched pool of aptamers.
Table 4.5. DREME motif analysis for Pool 10. DREME motif analysis identified 3 common motifs. GGRCA was found in 65 of 77 collapsed sequences or 84% of all sequences read by NGS with the GGACA variant slightly preferred. CAMTCA was the second most common motif in 52% of all sequences; the CACTCA was twofold more common than the CAATCA motif. The shortest motif, and least plentiful of the three, was GWAA. The prevalence of common motifs shows that selection was stringent enough to favour the discrete evolution of specific sequences.

<table>
<thead>
<tr>
<th></th>
<th>Motif</th>
<th>Sequence</th>
<th>Variant</th>
<th>Positive (/77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GGRCA</td>
<td>GGACA</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGGCA</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>CAMTCA</td>
<td>CACTCA</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAATCA</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>GWAA</td>
<td>GTAA</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAAA</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>
4.3.2 HCN are lead aptamer candidates.

The nucleotide sequence of the top HCN sequences from each of the 4 dominant clusters is represented in Figure 4.6. The presence of the DREME motifs is indicated in the coloured stretch of nucleotides. Motif 1 is illustrated in red, motif 2 in yellow and motif 3 in blue. All of the top HCN possessed at least one of these motifs. Some like #6-11019 and #5-4983 possessed 2 of the motifs. The abundance of DREME motifs expressed in the HCN sequences is yet another indicator of sequence convergence. With both phylogenetic data and DREME analysis in their favour the HCN sequences depicted above can be considered the best representatives of selection. These sequences, referred to as NLA-1 through NLA-4, were ordered as synthetic clones for affinity and binding analysis. Their putative secondary structure, derived from the RNAstructure software, is illustrated in Figure 4.7.
<table>
<thead>
<tr>
<th>Group</th>
<th>Sequence</th>
<th>Nucleotide</th>
<th>NLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7-2826</td>
<td>GCACGTACGAAACGCATGAGTGCGGACATCCAGCGGCCTCACATGGCTATGTGTAC</td>
<td>NLA1</td>
</tr>
<tr>
<td>B</td>
<td>13-7036</td>
<td>CTGCCCACTCCACATGCCTGCGCCGTCAATCATTTGCATGCACGCTCGCTCCTAACCCTAT</td>
<td>NLA2</td>
</tr>
<tr>
<td>C</td>
<td>6-11019</td>
<td>CCGTATGTCCGAATACGAGAAGCAGCACTCATTGAAAGCCATACGCGGAAGGATGCACGC</td>
<td>NLA3</td>
</tr>
<tr>
<td>D</td>
<td>5-4983</td>
<td>ACCACGGAGGGCATGTGCAGAGATAGGCGCGATCAGCTTGCCTCCGCATCGCGCCT</td>
<td>NLA4</td>
</tr>
</tbody>
</table>

**Figure 4.6. Nucleotide sequence of lead aptamer candidates.** The internal 60-nucleotide region for the top HCN sequences derived from each cluster is depicted above. For simplicity the forward and reverse primer sequences are omitted. Lead aptamer candidate were renamed NLA-1 through NLA-4. DREME motifs are represented by colour with motif 1 GGRCA is shown in red, CAMTCA in yellow and GWAA in red. Sequences NLA-3 and NLA-4 each possessed two DREME motifs. As the top HCN sequences of each cluster, and in possession of one or more DREME motifs, these NLA sequences best represent the inherent characteristics of pool 10.
Figure 4.7. Purported secondary structure of lead candidate aptamers. RNAstructure is an online predictive module that can estimate, based of free energy calculations, the secondary structure of DNA or RNA strands. Lead aptamer candidates were evaluated using this software and, represented above, are the lowest minimum free energy depictions of each.
4.3.3  **Aptamer Screening & K₈ Analysis**: Sequenced Aptamers possess high affinity and selectivity.

4.3.3.1  **NLA-3 and NLA-4 as the most specific and potent binders of CD20+HEK cells.**

To determine affinity and calculate $K_d$ the fluorescently labelled aptamers were titrated and incubated with the transfected CD20+ cells. The resultant fluorescence was measured using flow cytometry. The signal intensity is charted as the MFI (median fluorescence intensity) and graphically represented in Figure 4.8. The data was also fitted using a non-linear regression to measure $K_d$. $K_d$ is the equilibrium binding constant and represents the rate of disassociation. Graphically it is the concentration of aptamers necessary to bind half the receptor sites. Here, $K_d$ was calculated according to the following formula using the software GraphPad. $B_{max}$ refers to the saturation limit, $X$ and $Y$ are calculated values on the graph.

$$y = \frac{B_{max} \times X}{K_d + X}$$

All of the candidate aptamer show signal saturation at concentration exceeding 200nM. This corresponds with the final rounds of selection which were performed with DNA pools at 100nM. The calculation of the disassociation constant $K_d$ revealed that all the aptamers exhibited appreciable affinity with $K_d$s significantly less than 75nM. In particular aptamers NLA-3 and NLA-4 possess the smallest $K_d$ values at 58.4nM and 49.3nM respectively. It is interesting the aptamers NLA-3 and NLA-4 which both possess 2 DREME motifs also possessed lower $K_d$ values than aptamers NLA-1 and NLA-2 which each contained only one motif. This suggests that the possession of highly common motifs evolved during selection could be used to better predict affinity.
Figure 4.8. $K_d$ analysis of NLA aptamers. CD20+HEK cells were titrated with fluorescently tagged aptamer at indicated concentration (10nM, 50nM, 100nM, 200nM and 500nM) and binding measured as the MFI signal intensity. The data was fitted using a non-linear regression and $K_d$ measured from the graph. All the aptamers exhibited relatively appreciable binding with $K_d$s less than 100nM. In particular NLA-3 and NLA-4 were the best binder of CD20+HEK cells possessing the smallest $K_d$ at 58.4nM and 49.3nM respectively.
4.3.3.2 Aptamer show selective binding to CD20+ HEK cells

To evaluate specificity saturating concentrations of aptamers were incubated with both CD20+HEK cells and, separately, the untransfected HEK counterparts. After the cells were washed they were re-suspended in appropriate buffer and the resultant fluorescence measured. The MFI values are depicted in Figure 4.9. In all cases, the aptamers exhibited elevated MFI values when incubated with the transfected CD20+HEK cells relative to the original and untransfected HEK cells. This difference in binding was found to be statistically significant when assessed by the student T test. The most discriminative aptamers were NLA-3 and NLA-4 which exhibited the greatest difference in MFI signals between the two cell lines. In particular NLA-3 bound to the CD20+HEK cells three times more avidly than it bound the untransfected cells. That the sequenced aptamer clones possess increased binding association with the transfected cell and not with the untransfected HEK shows that selection was successful in cultivating a pool of aptamer whose specificity is to CD20+HEK. That all of the candidate aptamers possessed this characteristic illustrates the utility of using NGS to characteristically evaluate sequences.
Figure 4.9. Evaluation of aptamer specificity. To appraise the original selection regime, the binding intensity of the lead aptamer candidates were compared between the CD20+HEK and the un-transfected cells. In all cases the aptamers (300nM) possessed greater MFI values when incubated with the CD20+HEK cells and not the original untransfected cells. NLA-3 and NLA-4 were the most specific binders of the transfected cell line. Therefore the NGS analysis of pool 10 successfully returned sequences that possessed heightened affinity to the CD20+HEK cells.
4.3.3.3 Pooled NLA Aptamers Exhibit Competitive Binding with anti-CD20 Antibody

To further evaluate the specificity of the aptamers, a co-staining experiment was performed. Transfected cells were stained alternatively with either the unselected DNA library or the pooled NLA aptamers as well as with the anti-CD20 antibody. These co-stained cells were compared to CD20+HEK cells stained with only the antibody. The results are depicted in Figure 4.10. The solid black bar represent the antibody signal intensity of the CD20+HEK cells single stained with the anti-CD20 antibody, MFI=50.3. Alongside it are the MFI values for the same channel but where cells have been co-stained with either the initial DNA library or the pooled NLA aptamers. Note that incubating the CD20+HEK cells with the pooled NLA aptamers substantially and significantly decreases the antibody signal. Incubation of the cells with the DNA library however does not alter, to any significant degree, the amount of anti-CD20 staining. This effect appears to be concentration dependent. Cells incubated with 2µM of the pooled aptamers exhibit a much more significant reduction in the antibody signal, down to MFI=14 a reduction of more than a third. Cells incubated with concentration of pooled aptamers at 1 µM exhibit a decrease in the MFI=23, roughly half the original.

The co-staining experiments show that there are significantly greater antibody signal reductions when cells are incubated the aptamers than when the cells are incubated with the DNA library. This would appear to suggest that the aptamers compete with the anti-CD20 antibody for binding and therefore may recognize and bind identical epitopes on the CD20 molecule. This position is strengthened by the observations that the library whose sequences are more diverse, and not target specific, fails to induce the same reduction.
Figure 4.10. NLA aptamers inhibit the binding of anti-CD20 antibody. CD20+HEK cells were co-stained using anti-CD20 antibody (10ng/µL) and DNA derived from the unselected DNA library, or a pooled mixture of all the candidate aptamers NLA-1 to NLA-4. Co-stained samples are compared to the singly antibody stained anti-CD20 control (solid black bar). Incubating the cells with the aptamers dose dependently decreased the binding of anti-CD20 antibody. This reduction in signal intensity was significantly greater with the aptamers than with the unselected DNA library. This antagonism of antibody binding using the aptamers suggests that NLA aptamers and the antibody compete for binding site at mutual sites on the transfected CD20 molecule.
4.4 **Discussion**

Next generation sequencing is a powerful method of evaluating millions of discrete DNA species efficiently and robustly. Aptamer Pool 10 was the strongest and most specific binder of CD20+HEK cells. As such, it was important to elucidate the key aptamer sequences responsible for this association and to evaluate their characteristics. NGS analysis confirmed that pool 10 exhibited significant sequence convergence - a hallmark of a successful selection. After collapsing the data the total number of sequences was reduced from 541,258 to 62,737, an 8-fold reduction. The top 29 sequences accounted for nearly 70% of all sequences and as the phylogenetic assessment showed exhibited significant sequence similarity. In total there were 5 dominant clusters each of which possessed 1 or more sequences with a significantly high copy number. These high copy number sequence (HCN) reflect the most over-represented sequences in pool 10 and in turn became the lead aptamer candidates NLA-1 to NLA-4. Their retention and subsequent over-expression indicates that they are the sequences that best reflect the intrinsic characteristics of pool 10. DREME analysis further refined the pools characteristics. It identified 3 candidate motifs, the most prevalent of which GGRCA, was found in 84% of all sequences. Additionally all the HCNs and the candidate aptamers possessed one or more of these motifs. This further confirms the sequence consolidation of selection.

It is important that HCN sequences were preferred even to closely related and virtually identical sequences. More interesting still that in the most skewed of instances the biased nucleotides were G or C. For example between sequences #6 and #38, copy number bias favours sequence #6 by 131:1. Sequence #6 and #48 differ by only one nucleotide, the #6’s G
to sequence #48’s T. This was also observed between sequence #12 and #45 which are the second most heavily skewed. Here the ratio favoured sequence #12’s C to sequence #45’s T by a factor of 47:1. It is known that Taq polymerase is biased against GC rich DNA (128, 129). That there is the specific retention of sequences containing G and C nucleotides in spite of this polymerase bias is further evidence that selection was specifically isolating distinct and specific sequences and not random species. Relevantly, sequence # 6 would become NLA-3 the most potent and specific of all the aptamer candidates.

Cumulatively, NGS analysis was able to potently describe, quantify and evaluate aptamer candidates with a high degree of resolution. To its testament, all of the lead aptamer candidates bound CD20+HEK cells with $K_d$s less than 100nM. Among them NLA-3 and NLA-4 possessed the strongest association with the cells with $K_d$s of 58.4nM and 49.3nM respectively. When compared to the untransfected cell line all of the aptamers possessed heightened specificity for the CD20+HEK cells than to the untransfected HEK control. Here, NLA3 was especially potent, binding the CD20+HEK almost three times better than the HEK control. Therefore NGS helped to elucidate key sequences in pool 10 responsible for heightened affinity and specificity to the transfected cell line.

To further evaluate specificity we performed a co-staining experiment to gauge what mutual interaction, if any, the aptamers and the anti-CD20 antibody had. We compared co-stained cells to a singly antibody stained control. The incubation of the CD20+HEK cells with the pooled NLA aptamers resulted in a much more significant reduction in antibody binding. Incubating the cells with the unselected DNA library had no significant impact on antibody binding. The aptamers ability to limit the binding of anti-CD20 antibody was concentration
dependent, and the greater the concentration of the aptamer the more significant the reduction in the antibody signal. That this was more appreciable only with the NLA sequences and not the library shows that it is an intrinsic characteristic of the aptamers themselves and not the arbitrary influence of DNA. This antagonistic action suggests that both aptamers and antibody recognize mutual epitopes on the CD20 molecule. Cumulatively the results of NGS confirm that an in-vitro CD20 target positive selection strategy is an effective way to generate highly specific CD20 aptamers.

4.5 Conclusion

NGS analysis represents an extremely effective method of sequence discovery and organization. We showed that pool 10 exhibits significant sequence convergence, and the possession and retention of highly common motifs. Lead HCN aptamer exhibited high frequency ratios and were preferred in spite of PCR bias. All NLA aptamers exhibited potent binding, with \( K_d \)s less than 100nM and possessed specific binding to the CD20+HEK cells, relative to the untransfected control. Aptamer NLA-3 is a testament to the resolving capacity of NGS, it was the most abundant aptamer in pool 10, with the highest overall copy number of 11,019 and the possessor of 2 common motifs, it was also the strongest affinity aptamer \( (K_D=49.3\text{nM}) \) and the most discriminative binder of CD20+HEK. Incubating CD20+HEK cells with NLA aptamers, and not the DNA library, significantly reduced the binding of the anti-CD20 antibody. This co-staining experiment suggest that NLA aptamers and antibodies recognize mutual binding sites. In summary, NGS analysis of pool 10 helped to identify highly potent and specific aptamer sequences with appreciable CD20 target sensitivity.
5  **Biological Efficacy of Aptamers in Complement Dependent Cytotoxicity**

### 5.1 Abstract

One effector mechanism of anti-CD20 antibodies is complement dependent cytotoxicity (CDC). Here, we evaluated CDC induction in transfected CD20+HEK and the naturally CD20 expressing cell line CCL-86. We found that CDC could only be induced in the CCL-86 cells, while CD20+HEK cells were refractive to stimulation. Interestingly, aptamers selected against CD20+HEK were found to exert a protective effect against CDC induction with the CCL-86 cells, limiting the total amount of cell death by 10%, and significantly decreasing the intensity of pro-apoptotic markers 7-AAD and annexin-V. The unselected DNA library exerted no protective influence. These findings show that CDC induction relies on factors which the expression of CD20 alone on HEK293 cells cannot facilitate. Aptamers selected against transfected CD20, however, will exhibit an antagonistic action, limiting the extent of CDC induced cell death in naturally CD20 expressing cells. This finding further corroborates the specificity of the aptamers and helps to define their mechanism of action.

### 5.2 Background

#### 5.2.1 Biological Action of Anti-CD20 Antibodies is Varied and Diverse

Therapeutic antibodies have revolutionized the treatment of disease. By combining target specificity with effective clearance mechanisms they represent a highly specific and extremely
potent class of drugs for ailments as diverse as inflammation, autoimmune conditions and cancer. Though efficacy is apparent the immunological mechanism by which it is achieved can be elusive. In oncology, few targets are as avidly pursued as CD20. Indeed the first FDA approved therapeutic antibody, rituximab, was raised against the CD20 molecule. Its wide success prompted the clinical approval of 3 anti-CD20 antibodies, with another 8 still in clinical development. One interesting implication of this longevity has been a more thorough investigation of both anti-CD20 epitopes and their effector mechanisms.

5.2.2 Effector Action of anti-CD20 Antibodies

The therapeutic action of all monoclonals is currently limited to 4 mechanisms. They are, as depicted in Figure 5.1, PCD or programmed cell death, ADCC or antibody dependent cell mediated cytotoxicity, ADCP antibody dependent phagocytosis and CDC complement dependent cytotoxicity. To a significant extent, research into the effector action of anti-CD20 antibodies has been primarily focused on ADCC and CDC.
Figure 5.1. Anti-CD20 effector actions. CD20 ligation induces, through direct or indirect means, rapid and potent B cell depletion. Direct cytotoxicity, or PCD programmed cell death, is facilitated by the antibody-mediated activation of pro-apoptotic pathways. Indirect methods rely on an external stimulus in the form of effector cells or serum proteins. Effector cell stimulus relies on Fc and FcδR interactions. These may result in either ADCC: antibody dependent cell mediated cytotoxicity and the release of granulocytic enzymes or alternatively ADCP: antibody dependent cellular phagocytosis. CDC, another indirect effector mechanism, relies on the Fc mediated cleavage of serum complement proteins. These will self-assemble resulting in the formation of cytolytic transmembrane pores known as the membrane attack complex (MAC). From (43).
5.2.2.1 **PCD: Programmed Cell Death**

PCD refers to programmed cell death, when the binding of an antibody turns on or off signalling pathways resulting in the induction of apoptosis. Rituximab is a potent PCD agent, in-vitro, when crosslinked using a secondary antibody. Although how this could operate in-vivo is not well defined (43). It is interesting that the clinical side effect of all anti-CD20 antibodies is the rapid and conserved depletion of CD20 positive cells from the blood. So while the clinical mechanism of PCD may not be entirely clear its manifestation is obvious.

5.2.2.2 **ADCC: Antibody Dependent Cell mediated Cytotoxicity**

ADCC is mediated through the Fc domain of bound antibodies. Exposed IgG Fc domains are recognized by cognate Fc\(\delta\) receptors (Fc\(\delta\)R), which are expressed on a variety of effector leukocytes including natural killer cells, monocytes, macrophages and neutrophils (32, 43).

There are different isoforms of Fc\(\delta\)R with varying affinities and activities. The successful union between Fc and active Fc\(\delta\)R stimulates the release of cytotoxic granules containing enzymes like perforin, granulysin and granzymes from the effector cells resulting in the destruction of the Fc bound cell (43, 130). ADCC is frequently associated with anti-CD20 antibodies. Corroborative studies have shown the failure of anti-CD20 therapy in Fc\(\delta\)R knock out mice (131-133). In human studies, it’s been shown allelic variation in the genes encoding Fc\(\delta\)R can impact the outcome of treatment in patients with certain types of cancer. For example, NHL patients with high affinity variants of Fc\(\delta\)RIIIA exhibit better response rates (134), this correlation however could not be extended to patients with CLL (135). The appreciation of the importance of ADCC is evident in the design of GA101, one of the newest anti-CD20
antibodies. Its Fc domain was glyco-engineered to improve binding to FcδRIII receptors, a modification that resulted in a 100-fold enhancement of ADCC activity (136).

5.2.2.3 **ADCP: Antibody Dependent Cell mediated Phagocytosis**

FcδR interactions may alternatively result in phagocytosis if the effector cell is a monocyte or macrophage. Detailed information regarding ADCP is significantly lacking due to the fact that it is often highly variable in in-vitro studies and notoriously difficult to assess in-vivo (137); neither of which is reason to discount its influence. One interesting implication of ADCP, and a potential rationale for observed long-lived remission in cancer patients, is cross-presentation (38). Phagocytes can acquire and present antigens through phagocytosis, which are in turn used to stimulate and activate subsets of T-cells, including highly potent CTL (cytotoxic T lymphocytes) (138). Antibodies like anti-CD20 may act as signals for phagocytes, and if the antigens they acquire are cancer specific they may give rise to tumour specific subsets of CTLs, thereby establishing long lived anti-tumour immunity.

5.2.2.4 **CDC: Complement Dependent Cytotoxicity**

Complement dependent cytotoxicity also relies on the input of additional components from the immune system, in this case not cells but complement proteins. The complement system was discovered more than 100 years ago (139) and predates antibody elucidation. CDC is initiated by the antibodies Fc domain and requires stepwise cleavage of an additional 30 proteins—both blood borne and membrane bound. These free factors are synthesised in the liver in an inactive “pro-enzyme state” (139) and are found circulating in the blood, lymph and interstitial fluid. Some of these complement proteins like C1a, C3 and C3b are indicated in Figure 5.1. Their ordered deposition on the afflicted cell surface results in the formation of a
membrane attack complex (MAC), a transmembrane pore that disrupts membrane integrity, resulting in the loss of homoeostasis and eventual cell death (140).

CDC is well implicated with anti-CD20 antibodies. Rituximab has been shown to bind C1Q and induce potent CDC in both lymphoma cell lines and primary tumours (141). Studies have shown that the up-regulation of cellular factors known to provide resistance to complement — notably the expression of CD55 and CD59 — are associated with a resistance to anti-CD20 treatment (142). Neutralization of these factors is enough to re-sensitive cells to treatment (143). In-vivo studies in mice and humans have shown that complement is rapidly depleted after rituximab infusion (144). And, that the injection of fresh complement proteins can rescue this effect (145, 146). Other studies, though, have questioned the impact and extent of CDC. For example mice devoid of critical complement factors C1q, C3 or C4 were still susceptible to B-cell induced depletion using anti-CD20 antibodies (132). And in humans with follicular lymphoma and CLL no correlation was found between the expression of complement resistant factors and clinical outcome (147).

Elucidating the collective impact of these pathways has also been complicated. Although the antibody can initiate one or all of these effector mechanisms, the overall effect may not be synergistic. Indeed, studies have shown that the addition of viable serum complement protein inhibited rituximab initiated ADCC; while heat inactivated human serum, whose proteins are denatured, had no such effect (148). Another study showed that the cellular deposition of complement protein C3B promoted trogocytosis, the removal of antibody bound complexes from the cellular surface by other immune cells (149, 150) an outcome that would all but curtail
any FcδR mechanisms. The effector mechanism of clinical antibodies remains a complicated matter.

5.2.3 Epitope Specificity

One suggested explanation for this variability in effector mechanisms is epitope specificity. Figure 5.2 is a current list of all anti-CD20 antibodies approved or in clinical development, their epitope (binding site) and their effector actions (comments). Mapped epitopes are categorized as type 1 (rituximab based) or type 2 (non-rituximab). Of those approved, Ofatumumab recognizes a distinctly different epitope, Its epitope was mapped to the smaller extracellular loop, which places it in closer proximity to the cell membrane. Relevantly for CDC induction, which relies on the cellular deposition of proteins, it may be that ofatumumab is better placed to recruit more of these factors for MAC formation than rituximab. (56, 151, 152). Significantly, and perhaps partly due to the limited nature of epitope analysis or the inherent variability in biological assays, these generalization can’t be applied to other antibodies. For example ocrelizumab, veltuzumab, and zevalin are all antibodies that possess the same epitope as rituximab and yet all favour different effector mechanisms.
**Figure 5.2. Clinical Status of anti-CD20 antibodies.** Indicated above is the generic and brand name of the specific antibody, its structural class under format, clinical indication, epitope binding site and the most commonly associated biological effector actions under comments. In comments, PCD= programmed cell death, ADCC= antibody dependent cell mediated cytotoxicity, ADCP= antibody dependent cellular phagocytosis and CDC=complement dependent cytotoxicity. From (43)
There remains still much to understand regarding anti-CD20 antibodies and their effector actions. This understanding is complicated significantly by the staggering complexity of the immune system; where neither the cancer, nor the antibody, nor its clearance mechanism can be relied upon to be predicative indicators of efficacy.

5.2.4 Aptamer to Better Elucidate Target-Antibody Dynamics

A more comprehensive understanding of these dynamics is required. Aptamers have proven to be an incredibly useful probative tool to help elucidate, report on and in some cases even modify the interaction of antibodies and receptors. Aptamers selected against the immunogenic region of the acetylcholine receptor acts as decoys inhibiting the damaging effect of autoantibodies implicated in Myasthenia Gravis Disease (153). High affinity and highly stable aptamers selected against prostate specific antigen were found to inhibit 60% of its activity while avoiding the immunogenic nature of anti-PSA antibodies (14). And aptamers selected against ErbB-2/HER2, a growth factor receptor and a well-established cancer biomarker, exhibited two fold greater anti-tumour effects when compared to Erb-2/HER2 specific antibodies (58). Therefore, aptamers occupy an important innovative niche and can significantly aid in the elucidation and manipulation of complex antibody based mechanisms.

5.3 RESULTS

5.3.1 CDC is potently induced by anti-CD20 antibody in naturally CD20 expressing CCL-86 Cells but not in the transfected CD20+HEK cells.

We evaluated CDC induction in CD20+HEK cells and compared it to the naturally CD20 expressing CCL-86 cells. CDC can only be induced by antibodies in the presence of viable
human serum (HS). As a relevant control we also incubated cells in heat inactivated human serum (HiHS), whose proteins are denatured and incapable of stimulating CDC.

Figure 5.3 provides the results of the initial CDC assessment. Naturally CD20 expressing CCL-86 cells incubated with anti-CD20 antibody in either PBS or with heat inactivated human serum will maintain high levels of cellular viability at over 80%. This shows that anti-CD20 antibodies cannot, on their own or when deprived of viable complement proteins, induce CDC. Only when CCL-86 cells are incubated with both anti-CD20 antibodies and viable human serum does the viability drop significantly to 33%. Note that disruption in the shape of the cell population. Unexpectedly, CD20+HEK transfected cells failed to undergo CDC. Here, cellular viability remained constant at over 90% regardless of reaction conditions - in PBS, with viable serum or with heat-inactivated serum. CD20+HEK cells are therefore not susceptible to CDC induction.
Figure 5.3. Comparison of CDC induction in naturally CD20 expressing CCL-86 and transfected CD20+HEK cell lines. A) Representative plots of CDC using anti-CD20 antibody (10ng/µL). In CCL-86 cells provided viable HS undergo CDC evidenced by the dramatic decrease in cellular viability (82.9% to 32.9% of cells). Heat inactivated human serum exerts no action. CD20+HEK cells were refractive to CDC induction and persisted with high cellular viability regardless of reaction conditions. B) Tabulated results.
5.3.2 CD20+HEK aptamers can limit extent of CDC in CCL-86 Cells.

The antagonistic action of aptamers selected to receptors is widely known. We were curious to assess the activity and impact our selected aptamer sequences could exert with CCL-86 cells.

5.3.2.1 CD20+HEK aptamer exhibit modest binding affinity to CCL-86 Cells

We assessed the binding affinity of our CD20+HEK aptamers with both the naturally CD20 expressing cells CCL-86 and the naturally CD20 negative TIB-152. As Figure 5.4 illustrates when aptamers NLA-1 through NLA-4 are pooled together they exhibit significantly greater binding affinity, when compared to the library, with the CCL-86 cells. CCL-86 cells incubated with the library possess an MFI of 5, and when incubated with the pooled clones it is nearly 3 times greater with MFI of 15. Only a slight increase was seen with the MFI of the TIB-152 cells, from 4 to 7 when incubated with the library and pooled aptamers respectively. As the aptamer were selected from CD20+HEK cells and exhibit specific binding to the naturally CD20 expressing cells CCL-86 and not CD20 negative cells lines TIB-152, this would appear to suggest that NLA aptamers are CD20 specific.
Figure 5.4 CD20+HEK aptamer show moderate and specific affinity with the CD20 positive CCL-86 and not the CD20 negative TIB-152. CCL-86 and TIB-152 cells were stained with 2μM of either the DNA library or pooled NLA aptamers sequences. CCL-86 cells, which are naturally CD20 positive, exhibit significantly greater binding affinity with the pooled aptamers, relative to the unselected DNA library, than the TIB-152 cells, which are naturally CD20 negative. This suggests that NLA aptamers positively label and associate with the CD20 molecule on CCL-86 cells.
5.3.2.2 CDC in CCL-86 cells is associated with significant increases in 7-AAD and Annexin-V

Staining

To better evaluate CDC in CCL-86 cells and the potential impact aptamers have, we measured both total cell death, in the form of the vital stain 7-AAD, as well as entry into apoptosis using the PE labelled annexin-V. 7-AAD is a DNA intercalating dye. Its permeability into cellular nuclei is barred by the presence of intact cellular membranes. Dead cells whose membranes offer no such protection, will be positively stained by 7-AAD. Annexin-V binds phosphatidylserine (PS) a phospholipid whose expression is restricted by flippases to the inner leaflet of the cellular membrane. In apoptotic cells ATP dependent flippases are no longer maintained leading to extracellular PS expression. It is the extracellular PS which is the target of annexin-V antibodies. Together 7-AAD and annexin-V are dual dyes that can be used to differentiate not only between live and dead cells but also live cells committed to undergoing apoptosis. Figures 5.5 (control) and Figure 5.6 (experimental samples) are histograms representing the MFI of 7-AAD and annexin-V as well as a dot plot depicting total events captured by the flow cytometer, assessed by both SS (internal complexity) and FL4 (for 7-AAD detection). High 7-AAD fluorescence indicates cell death; therefore in the dot plots the % total value of the I^+ quadrant represents the total percentage of dead cells.
**Figure 5.5. Complement dependent cytotoxicity in CCL-86: control Samples.** 7-AAD and annexin-V staining for the control samples of CCL-86 cells. Represented are CCL-86 cells incubated in RPMI media (unsupplemented), in 10ng/µL of anti-CD20 antibody diluted in PBS, in PBS with 50% heat inactivated human serum (HiHS), in PBS with 50% normal human serum (HS), in 2µM of DNA library diluted in PBS, in 2µM of pooled CD20+HEK aptamers diluted in PBS. Incubating CCL-86 cells in unfortified conditions was associated with slightly elevated increases in 7-AAD and annexin-V relative to the cells incubated in HiHS or HS.
Figure 5.6. NLA aptamers limit the extent of complement ependnet cytotoxicity in CCL-86 cells. CDC was initiated by incubating cells with 10ng/µL of anti-CD20 in the presence of viable human serum. CDC was associated with increases in total cell death evidenced by the elevated % of events in the I++ quadrant of the dot plot and by the significant increases in total 7-AAD and annexin-V staining. Incubating cells with pooled CD20+HEK aptamers resulted in decreased levels of both markers, as well as an overall 10% reduction in the amount of dead cells. Incubating the cells with the unselected DNA library did not provide any protective impact. Therefore aptamer selected against CD20+HEK protect CCL-86 cells from CDC.
Figure 5.5 represents the controls samples: the cells incubated in media without serum, with antibody in PBS, with PBS supplemented with HiHS or HS, as well and DNA library or pooled aptamer in PBS. Note the relative similarity of MFI with all samples. This proves that neither the media, the antibody alone, DNA library, pooled aptamers nor some agent inherent in HS or HiHS can significantly impact cellular viability. We did notice a modest increase in the MFI for 7-AAD and for annexin-V when samples were deprived of serum. For example, in CCL-86+HiHS and CCL-86+HS the total percentage of dead cells in the I++ quadrant is limited to 4.51-6.60%, MFI for 7-AAD is in the range of 1.46-1.78, the MFI for annexin-V 1.95-2.19. These values are slightly lower than then that observed with the other controls, where the total percentage of cell death was 5.18-6.98%, the MFI for 7-AAD was 2.41-2.86 and MFI for annexin-V 5.59-6.78. This suggest that incubating cells wholly without serum for the duration of this experiment is associated with slightly elevated levels of cell death and PS expression.

CDC was evaluated by incubating cells with antibody and HS (CDC positive control) or alternatively with antibody and HiHS (CDC negative control), refer to the first two columns of Figure 5.6. Incubating cells with anti-CD20 antibody and HS induces morphological changes in cellular distribution, significantly increases the percentage of cells in the I++ quadrant from 6.77% to 68.81%, and is associated with significant increases in the MFI for both 7-AAD and annexin-V. Note that in the CDC positive control there are two population for 7-AAD, indicated as K and L. When gating each peak—population K has an MFI of 5, which is associated with the living cell population, and is roughly comparable to the values from the control samples in Figure 5.5. Population L, depicts the dead cells whose MFI is higher at 82. CDC is also associated with increased Annexin-V staining from and MFI of 2.40 in the CDC negative control
to MFI signal of 63.61 with the CDC positive control. This dramatic increase is also significantly higher than that observed with the serum deprived controls in Figure 5.5. Note that only staining with 7-AAD, and not annexin-V, can resolve the cell into two discrete populations. Therefore all the cells detected by the flow cytometer, both 7-AAD positive and 7-AAD negative, possess a significant increase in PS expression. In this assay the CDC positive control is associated with significant increases in cell death (up to 70% of total population), and in elevated levels of 7-AAD and annexin-V staining.

5.3.2.3 CD20+HEK aptamers protect CCL-86 cells from CDC

To evaluate the effect of CD20+HEK aptamers we incubated cells with 2µM of either DNA library or the pooled aptamers for one hour before initiating CDC using viable human serum and anti-CD20 antibodies. Representative flow diagrams are depicted in Figure 5.6, with the relevant tabulated MFIs in Figure 5.7. When compared to the CDC positive control where the total cell death=68.81%, we found that cells pre-incubated with the aptamers, and not the library, were more viable with the percentage of dead cells in the I++ quadrant amounting to 59.04%. Therefore incubating CCL-86 cells with NLA aptamer can increase total cellular viability by nearly 10%. This is also reflected in the lower MFI reading for both 7-AAD and annexin-V. There are still two peaks in the NLA treated 7-AAD histogram but the peak associated with the lower MFI is noticeably greater. Population M contains 40.80% of all events, this is contrasted to the low 7-AAD population of the CDC control which contains 26.79% and the DNA library treated sample 27.71%. These results are also shown in Figure 5.7. Therefore incubating CCL-86 cells with the aptamers and not the library will significantly increase the amount of cells that exhibit low 7-AAD staining.
These protective effects were noticed in the extent of annexin-V staining as well. The aptamer pre-incubated sample shows a slightly bimodal distribution whose MFI of 45.63 is notably lower than the MFI for annexin-V from either the CDC control (MFI=63.61) or the DNA library treated sample (MFI=59.22). Therefore cells incubated with aptamers, and not the DNA library, exhibited less superficial PS staining. Cumulatively, incubating cells with NLA aptamer resulted in pronouncedly lower 7-AAD and extracellular PS staining-characteristics consistent with greater viability.

Figure 5.7 is the tabulated MFI results from triplicate data. In the CDC positive control as well the library incubated sample total cell death (% of cells with high 7-AAD staining) is 72-74%, in these samples the percentage of cells with low 7-AAD staining is 22-23%, and level of annexin-V staining is MFI=62-64. The NLA aptamers treated cells differ, here the % total percentage of dead cells is 61%, leaving the percentage of cells with low 7-AAD staining at 38%, and the staining of annexin-V is also markedly reduced down to MFI-48. This shows that incubating CCL-86 cells with NLA aptamers appears to limit the extent of anti-CD20 induced antibody CDC.
Figure 5.7. NLA-protected CCL-86 cells exhibit greater viability and have decreased staining of the pro-apoptotic marker annexin-V. CDC was induced in CCL-86 cells using 10ng/µl of anti-CD20 antibody and viable human serum. To evaluate the effect of aptamers, CCL-86 cells were incubated with 2µM of either the DNA library (CDC+DNA library) or the pooled NLA aptamers (CDC+Pooled aptamers) and compared to the CDC positive control. In the CDC control total cell death was 74%, leaving 23% of cells considered living by exhibiting low 7-AAD staining; the PE–annexin-V staining is maximal at an MFI of 64. Cells incubated with the DNA library had values that fall within range of the CDC positive control (% dead cells=72, low 7-AAD staining population=22%, and the MFI of annexin-V=62). CCL-86 cells incubated with the NLA aptamers however show higher viability with the total percentage of dead cells at 61%, the % of cells with low-7-AAD staining at 38, and the MFI of annexin-V is decreased to 48. Therefore incubating CCL-86 cells with aptamers derived from CD20+HEK selection protects and limits the extent of damage induced by CDC (p<0.05).
The results from the biological analysis validate that aptamers selected to CD20 on transfected cells appear to exert a positive protective effect against CDC induction in naturally CD20 expressing CCL-86 cells. NLA aptamers significantly decreased the total number of dead cells by almost 10%, and resulted in lower MFI values of the apoptotic markers 7-AAD and annexin-V. Importantly, this effect was restricted to the usage of selected DNA aptamers. The unselected DNA library exhibited no such protective effects and possessed 7-AAD and annexin-V staining consistent with the CDC positive control. This shows that CD20+HEK aptamers appear to have the capacity to limit the extent of CDC induced cellular damage even in naturally CD20 expressing cells.

5.4 DISCUSSION

Transfection permits the expression of a target gene in a cell line that would ordinarily lack it. Expression does not necessarily guarantee, however, that the gene or its protein product will function in a manner consistent with its original state. Protein function depends on a series of various factors, whose interaction can be obscure and complex- and which transfection alone may not satisfy. In what was a much-publicized event at the time, the retroviral transfection of gamma-c gene as part of a gene therapy clinical trial resulted in some patients developing leukemia 3 years after the initial infusion. It was later found that the gamma-c gene had inserted proximal to the LMo2 proto-oncogene and activated its malignant expression (154). This is just one example, admittedly exceptional, of how transfection and its subsequent expression can often have unintended and off-target effects.
CD20+HEK cells failed to undergo CDC. This is in spite of the fact that CDC is a well-established anti-CD20 mediated mechanism. Their imperviousness, indicated in Figure 5.3, may be the result of various factors. It is possible that CD20 expression on HEK293T cells was not at the threshold levels required to appreciably induce CDC. Some papers have shown that low expression levels of CD20 abrogates the action of anti-CD20 antibodies, while increases in total expression can reverse this (155). In a clinical setting susceptibility to anti-CD20 therapy closely mirrors its expression level. For example rituximab is a much more potent CDC inducer in follicular lymphoma cells which are high CD20 expressing, than small lymphocytic lymphoma which express low levels of CD20 (156). Indeed, 29% of patients with diffuse large B cell lymphoma relapsed due to a CD20-negative transformation (157). As our data from Chapter 2 transfection shows (pg=44), CD20+HEK cells do possess a significantly weaker anti-CD20 MFI signal when compared to CCL-86 cells. The low level expression of CD20 in the transfected cells represent one plausible explanation for the failure of CDC.

We also cannot eliminate the possibility that the inherent cellular composition of HEK293 cells may have had some influence. The expression of complement resistant factors like CD55 and CD59 will inhibit CDC (147). It stands to reason that molecules likes these or others yet undiscovered may have also had some role in inhibiting CDC induction in the CD20+HEK cells. Unfortunately the expression of complement resistant factors CD55 and CD59 were not assessed with HEK293 cells nor could any relevant information be found in published literature. Another important consideration comes from papers that have evaluated effector actions using CD20 transfected cells. Perhaps not surprisingly in these papers the transfected CD20 only elicited positive effector actions when the transfected cell line was of immune origin.
Teeling et al paper (56) they transfected CD20 into both HEK 293T and CEM cells (a T cell line) to elucidate epitope sites. While binding assays were performed on the CD20 transfected HEK293 cells, the biological analysis of CDC was only evaluated, positively, with the CEM T cells. In Introna et al (158) CD20 was successfully transfected and CDC positively induced in human derived and blood borne T cells. Indeed the work of Griffioen et al (159) suggests that the transfection of CD20 into T cells may be one way to regulate their control for adoptive T cell therapy. CD20 is a receptor whose natural distribution among cells is remarkably narrow; naturally its expression is restricted to B-cells. The work of Teeling, Introna and Griffioen—all of whom transfected CD20 into T-cells—strongly suggest that the cellular composition of T-cells, which share a common hematopoietic origin with B-cells, may possess a uniquely primed microenvironment for the expression and elicitation of CDC. Perhaps in ways that the composition of HEK293, which originate from a different precursor entirely, fail to reproduce.

The lack of CDC induction in CD20+HEK cells abrogated their use in biological analysis and so we focused on CCL-86 which are naturally CD20 expressing. In Figure 5.4, we show that when pooled NLA aptamer do possess significantly greater binding affinity with CCL-86 cells than when compared to the unselected library. TIB-152 cells, which are not CD20 expressing, are not as officiously labelled with the aptamers. Therefore the aptamer selected against CD20+HEK cells appear to positively associate with the CD20 molecule expressed on CCL-86 cells and not on markers inherent to TIB-152 cells.

It is worth mentioning that this binding is noticeably lower than what was observed initially with pool 10 (pg=48). Indeed binding required significantly higher amounts of aptamers before appreciable detection, after washing, became obvious. Aptamer pool 10 positively labelled
CCL-86 cells at a 200nM concentration, while nearly 10 times the amount of aptamers was necessary to see appreciable detection with the clones. Whether as a result of the inherent selection strategy or the NGS selection of clones is not known and will remain to be elucidated.

We were curious to evaluate the impact that CD20 specific aptamer had on antibody mediated CDC induction in CCL-86 cells. CDC is potently induced in CCL-86 cells provided anti-CD20 antibody and viable human serum. CDC is associated with increases in total cell death, evidenced by greater 7-AAD staining but also significantly elevated levels of annexin-V staining, a well-established apoptosis marker. Heat inactivated human serum whose proteins are denatured will fail to induce CDC. Incubating CCL-86 cells with the aptamers, and not the DNA library, was able to protect the cells from the total extent of CDC, reducing the total burden of cell death by nearly 10%. Cells treated with protective aptamers had significantly lower 7-AAD as well as lower annexin-V staining. This shows that CD20+HEK selected NLA aptamers may actively be involved in discouraging CDC, slowing down cell death and increasing overall cellular viability. Importantly, since treatment with the unselected DNA library was ineffective it shows that these effects are inherent to the NLA aptamers themselves and is not a by-product of the DNA itself. Cumulatively, this data suggests that aptamers selected to CD20+HEK cells appear to be specific to the CD20 molecule as it is expressed on native CCL-86 cells and can modestly limit the total extent of anti-CD20 induced CDC.

5.5 Conclusion

Aptamers are important tools to elucidate, probe and evaluate complex chemical and biological interactions. Here, NLA aptamers generated from selection using transfected
CD20+HEK showed appreciable and specific binding affinity to positively label the CD20 expressing cells CCL-86, and not the CD20 deficient TIB-152. In a testament to their specificity, incubating CCL-86 cells with NLA aptamers protected them from the total extent of CDC induced cell death, leading to increased cellular viability as evidenced by lower 7-AAD and annexin-V staining. Cumulatively, this demonstrates the validity of using target positive cell-SELEX to generate aptamers with both specific binding affinity and biological efficacy.
6 GENERAL CONCLUSION

Conventional selection strategies are intrinsically imperfect. Conventional solid state SELEX is ideal for the direct selection of ligands to known targets. However, it fails to adequately appreciate how these targets may be expressed in live cellular systems. Cell-SELEX can isolate for aptamers in their functionally active state but the identification of the intended target remains a complex and exhaustive task. Viral transfection of genes into cells incorporates the best attributes of both methods. By transfecting a gene and establishing target positive and target negative cell lines we were able to select aptamers in a target specific and a biologically conscious context.

We transfected CD20 into HEK293 cells and successfully evolved CD20+HEK cells. We used CD20+HEK and their untransfected HEK counterparts in a novel SELEX strategy and generated pools of aptamers that showed increased affinity to the CD20+HEK cells. Pool 10, the greatest binder of CD20+HEK, also positively validated CD20 expression in the CD20 expressing CCL-86 cells and not in the CD20-negative TIB-152. Pool 10 was a discriminative pool and contains aptamers that recognize the CD20 molecule regardless of cell line.

We employed NGS analyses to analyze the sequences in pool 10 with a greater depth and sensitivity than what would be afforded with general bacterial cloning and Sanger sequencing. We showed that pool 10 exhibited dramatic sequence consolidation, as evidenced by the phylogenetic assessment, HCN sequences and the MEME motifs. The most abundantly
expressed sequences, the HCNs, showed a copy number bias even against very closely related aptamers. These traits demonstrate that the evolution of sequences in SELEX was directed and not random.

All the HCN sequences, which became aptamers NLA-1 through NLA-4 exhibited potent and specific binding to the CD20+HEK cells, with $K_d$s less than 100nM and, on average, capable of binding CD20+HEK two-three times as avidly as the untransfected cells. In co-staining experiments incubating CD20+HEK cells with NLA aptamers inhibited the binding of anti-CD20 antibody by more than half. This demonstrates that NLA aptamers may bind mutually recognizable sites on the CD20 molecule as the anti-CD20 antibody itself.

Anti-CD20 antibodies are clinically established immunotherapeutics. By binding CD20 they elicit rapid and sustained cell death. One method by which they do this is through complement dependent cytotoxicity. The Fc domain on anti-CD20 antibodies elicits a proteolytic cascade of complement factors, found in viable human serum, which results in the formation of a pore forming membrane attack complex. We showed that CD20+HEK cells, the ones used for selection, were refractive to CDC induction. Aptamers selected against CD20+HEK however, not only positively bound the CCL-86 cells, but in a CDC assay exerted a protective effect. Incubating CCL-86 cells with NLA aptamer decreased total cell death by 10% and also significantly reduced the signals of the pro-apoptotic markers 7-AAD and annexin-V. This shows that aptamers selected using target positive SELEX possess not only physical binding capacities but have also a biological merit. By binding to the CD20 molecule they exert an antagonistic effect against CDC induction.
In summary, the selection of aptamers using a lentiviral mediated cell-SELEX method was successful in evolving sequences that are both strong and specific binder of the CD20 molecule as it is expressed in both transfected CD20+HEK and on native CD20 expressing cells like CCL-86. Aptamers selected using CD20+HEK cells were also biologically effective limiting the extent of CDC induction in CCL-86 cells. This demonstrates the utilitarian value of lentiviral specific cell-SELEX. It represents a strategy that is simultaneously specific and universal with the capacity to generate targeted aptamers which not only physically associate with but actively augment the biological mechanism of their intended molecule.


