

## ABSTRACT

Title of Document: CHARACTERIZATION OF NOVEL  
NUCLEIC ACID INHIBITORS (APTAMERS)  
OF HUMAN IMMUNODEFICIENCY VIRUS  
REVERSE TRANSCRIPTASE SELECTED  
USING A PRIMER-FREE SELEX APPROACH

Yi-Tak Lai, Doctor of Philosophy, 2011

Directed By: Dr. Jeffrey J. DeStefano, Professor  
Department of Cell Biology and Molecular  
Genetics

Aptamers are synthetic, single stranded nucleic acids that bind specifically to the target protein. Aptamers have many potential uses including therapeutic and diagnostic applications. Most aptamers are identified through standard SELEX (Systematic Evolution of Ligands by EXponential enrichment) procedures that include a starting pool of nucleic acids with a region of random nucleotides flanked by fixed sequences. A main disadvantage of the traditional SELEX is the potential for the fixed sequences to “bias” the selection by interacting with nucleotides in the random region of the oligonucleotide. I developed a novel primer-free SELEX method for isolating single strand 30 nt DNA aptamers from a random sequence pool for HIV reverse transcriptase (HIV-RT), in which selection occurs in the absence of any flanking, fixed nucleotides. Selected aptamers bound ~10-20 fold tighter than starting material to HIV-RT. The selected aptamer (PF1)

contained a motif of four diguanosine repeats. PF1 was compared to two classes of HIV-RT aptamers: a G-quadruplex aptamer (R1T family) from Michalowski *et al.*, Nuc. Acids Res. 2008, and a primer-template aptamer developed by our lab (37 NT SELEX, DeStefano and Nair, Oligonucleotides 2008). PF1 was an effective inhibitor of HIV-RT *in vitro* and bound to RT about as tightly as a full primer-template. Both the R1T and 37NT SELEX were more potent HIV-RT inhibitors. This may be due in part to their large size. Circular dichroism spectroscopy indicated that PF1 does not form a G-quadruplex, while the R1T, consistent with previous results, is a parallel G-quadruplex.

The effect of the primer-template aptamer on HIV inhibition in cell culture was also studied. Low micromolar concentrations of aptamer in the absence of a transfection agent inhibited replication in Jurkat cells without significant cellular toxicity. However constructs with similar structure also inhibit HIV replication leading us to conclude that aptamer blockage of replication may be due to alteration of cellular pathways.

In addition to the possible contribution in developing novel nucleic acid-based HIV inhibitors, aptamer selection experiments can also shed light on how RT interacts with nucleic acids both *in vitro* and in cellular models.

CHARACTERIZATION OF NOVEL NUCLEIC ACID INHIBITORS  
(APTAMERS) OF HUMAN IMMUNODEFICIENCY VIRUS REVERSE  
TRANSCRIPTASE SELECTED USING A PRIMER-FREE SELEX APPROACH

By

Yi-Tak Lai

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2011

Advisory Committee:  
Professor Jeffrey J. DeStefano, Chair  
Professor James N. Culver  
Professor Jonathan D. Dinman  
Professor Eric O. Freed  
Professor Siba K. Samal, Dean's Representative

© Copyright by  
Yi-Tak Lai  
2011

## Dedication

This dissertation is dedicated to  
my family, for their endless support and care;  
and to Josh, for all your patience and love.

## Acknowledgements

The journey to obtaining my Ph.D. degree required support from many others who have helped me along the way. First, I want to give my utmost thanks to my advisor, Dr. Jeff DeStefano, for all of his guidance and support during my graduate work at Maryland. Not only has he taught me to think critically about my research, he is one of the most knowledgeable scientists I have ever known. He is also the most patient mentor I have worked with. I benefit from his tireless efforts to gently correct my mistakes and to help me learn the best science I can. Beside research, Jeff has shown me the importance of work-life balance, and to be passionate about what one loves - whether it is politics, sports, or coffee - Jeff always manages to find time between benchwork and his interests beyond. I will cherish what I have learned with Jeff from our daily conversations.

I would like to thank my committee members: Drs. Jim Culver, Jon Dinman, Eric Freed, and Siba Samal for their time and advice. Their expertise and frankness are valuable for me to proceed with my studies and to help me to become a tougher person. I would also like to thank Dr. Janet Hartley for her generous support to me and the DeStefano lab towards the progress of our research.

I cannot express my gratitude to the DeStefano lab members: they are a second family to me. I would like to thank Dr. Gauri for her care, discipline, her skills on all things scientific, cloning and cooking, and being a great critic on my baking. To Dr. Deena, for her quiet grace, unexpected playfulness, and being a fantastic teacher to me on everything I need to know upon joining the DeStefano lab. To Divya, for her bubbly presence, steady patience and offering a shoulder for me to cry on: I am forever grateful for her support. To Jeff O., for his laughter, crazy ideas, and being a good companion while learning to become better teachers together. To Katherine, for her endless pop culture references and Google abilities, her Powerpoint expertise, and teaching the great art of humor to me. To Vasu, for his unique sense of humor and reminding me the importance of curiosity in science. To Manthan and Elizabeth: thank you for your great company and dedication. I would also like to thank Drs. Reshma Anthony and Niru Narayanan for showing me the ropes when I started in the lab.

I would also like to thank the staff at Center for Teaching Excellence, in particular Dr. Spencer Benson for his unique perspective and being a great teaching mentor, and the 2009-2010 CTE Lilly Graduate Fellows for educating each other on undergraduate teaching and learning as a cohort. I have gained many valuable skills through the exchanges between my colleagues and me on scientific research. I would like to thank the Fredericksen lab for providing antibodies for the immunoblot and access to the luminometer, the Kahn lab and Jen Alder for the CD spectrometer training, the McIver lab and Trey Belew for their generous support and creative ideas. To Danielle, Michele, Aprajita, Michael, and Katharina: thank you for your friendship and encouragement since the first day of class. I would not have been able to make it through without you all. To Dr. Thomas Man: thank you for inspiring me to appreciate the art of science. To Lydia: thank you for sharing our joys and sorrows together. You are my tower of strength and the source of wisdom. To my friends Cindy, Erica, Melissa, Susy, Phil and Richard: thank you for being with me regardless of the ups and downs in our lives. Your presence has really enriched my life here in the US.

My family has been a big component of my learning experience. My parents are the biggest supporters in my studies halfway around the globe. They have been extremely patient and caring as I have gone through the challenges in my graduate career. To my brother Bert, who understands the path to a Ph.D. in science, and cheers me along on my journey. I am extremely fortunate to have two surrogate families in the US: to Matt Bufano, for his faith in me, for being a close friend and a brother to me; and to the Lamianos, who have taken me in as a part of their family and made me feel truly at home.

Last but not least, I would like to express my gratitude to the person that has been my greatest support throughout my Ph.D. journey. To Josh, my best friend and teammate: thank you for believing in my abilities, and for encouraging me not to give up even at my darkest hours. Without your patience and prodding, I would not have mustered the courage and developed the endurance to break down the barriers. I look forward to running with you at the next stages of our lives. And to all the people I have met: thank you for teaching me various lessons on science, life, the universe, and everything (2).

# Table of Contents

Dedication .....	ii
Acknowledgements .....	iii
Table of Contents .....	v
List of Tables .....	vii
List of Figures .....	viii
List of Abbreviations .....	x
Chapter 1: Introduction .....	1
1.1 Discovery of HIV and AIDS.....	1
1.2 AIDS Epidemiology and HIV Classification.....	2
1.3 HIV Pathogenesis.....	4
1.4 HIV Structural Characteristics .....	7
1.5 HIV Genetic Organization and Protein Functions .....	9
1.6 HIV-1 Life Cycle .....	12
1.7 Reverse Transcription in HIV-1.....	17
1.8 Reverse Transcriptase .....	19
1.9 Genetic recombination in HIV .....	22
1.10 Antiretroviral Therapy .....	25
1.11 Aptamers and Systematic Evolution of Ligands by Exponential Enrichment (SELEX) .....	27
Chapter 2: A primer-free method that selects high affinity single-stranded DNA aptamers using Thermostable RNA Ligase.....	31
2.1 Introduction.....	31
2.2 Materials and Methods.....	33
2.2.1 Materials .....	33
2.2.2 Methods.....	34
2.3 Results and Discussion .....	40
2.3.1 The efficiency of the selection process and its effect on the recovery of tight binding aptamers.....	40
2.3.2 The analysis of the SELEX procedure by monitoring binding to HIV-RT, and the nature of the recovered aptamers .....	48
2.3.3 Aptamer sequences recovered from round 11 of the first SELEX and round 9 of the second were nearly identical and had little secondary structure.....	51
2.3.4 The ligation efficiency of Thermostable RNA Ligase improves with low structure and a T residue at the 3' terminus of the acceptor.....	52
2.4 Conclusions and significance.....	57
Chapter 3: Comparative study of HIV-RT aptamers from a primer-free SELEX approach with various classes of RT aptamers.....	58
3.1 Introduction.....	58
3.2 Materials and Methods.....	62
3.2.1 Materials .....	62
3.2.2 Methods.....	63
3.3 Results.....	66



3.3.1 Determination of equilibrium dissociation constant ( $K_d$ ) values for PF1 and the SELEX starting material .....	66
3.3.2 The middle guanosine doublets were required to maintain high affinity to HIV-RT .....	67
3.3.3 Truncations of the aptamer abolished its tight binding to HIV-RT .....	71
3.3.4 An RNA version of PF1 does not confer tight binding to HIV-RT .....	72
3.3.5 The diguanosine repeats require the specific sequences in PF1 to retain HIV-RT tight binding .....	72
3.3.6 The selected aptamer PF1 can inhibit HIV-RT primer extension, but not as effectively as some other previously characterized RT aptamers .....	73
3.3.7 Circular dichroism spectroscopy indicates the selected aptamer does not form a G-quadruplex .....	77
3.4 Discussion .....	83
3.4.1 Effects of the guanosine repeats on HIV-RT binding and possible structures of PF1 .....	83
3.4.2 The role of the aptamer length to HIV-RT tight binding .....	86
3.4.3 Comparison between various classes of aptamers in HIV-RT inhibition assays .....	87
3.4.4 Applications and future work .....	89
Chapter 4: Effects of primer-template aptamers to HIV-RT on inhibition of HIV replication in cell culture .....	91
4.1 Introduction .....	91
4.2 Materials and Methods .....	94
4.2.1 Materials .....	94
4.2.2 Methods .....	95
4.3 Results .....	99
4.3.1 The primer-template aptamer is a potent inhibitor to HIV-RT in vitro .....	99
4.3.2 The aptamers do not have a significant affect on cell growth .....	103
4.3.3 All structured aptamers show significant inhibition to HIV-1 infection .....	105
4.3.4 Effect of HIV-1 on aptamer entry into Jurkat cells .....	109
4.3.5 The aptamers did not induce BST-2 production in cell culture .....	112
4.4 Discussion .....	112
4.4.1 Effects of primer-template aptamers on cell proliferation .....	112
4.4.2 The possible roles of the structure of the primer-template aptamers to HIV-1 inhibition in cell culture .....	114
4.4.3 Possible relationships of the aptamer uptake and the presence of HIV-1 .....	117
4.4.4 The effects of aptamers on cellular pathways induction and HIV inhibition .....	118
4.4.5 Conclusion and future directions .....	118
Chapter 5: General discussion .....	121
Bibliography .....	126

## List of Tables

Table 1-1: A list of HIV proteins and the protein functions in the HIV life cycle .....	10
Table 2-1: Sequences recovered from SELEX experiments and $K_d$ values for binding HIV-RT .....	53
Table 3-1: $K_d$ values of HIV-RT binding for PF1 and various constructs .....	68
Table 3-2: $IC_{50}$ values for various groups of HIV-RT aptamers .....	74

## List of Figures

Figure 1-1: Classification of the human immunodeficiency virus .....	3
Figure 1-2: A generalized graph of the levels of HIV RNA copies (viral load) and CD4 <sup>+</sup> T cell counts during the different stages of HIV disease progression ...	5
Figure 1-3: A schematic diagram of human immunodeficiency virus (HIV) and structural characteristics .....	8
Figure 1-4: Organization of the HIV genome, and the encoded protein products .....	11
Figure 1-5: A schematic of the HIV-1 life cycle .....	13
Figure 1-6: Schematic representation of HIV-1 replication .....	17
Figure 1-7: Structural representation of the catalytic subunit p66 in HIV-1 reverse transcriptase (RT) and its various domains with the RNA-DNA hybrid .....	20
Figure 1-8: The SELEX process .....	25
Figure 2-1: Protocol for primer-free SELEX .....	41
Figure 2-2: Ligation reactions with T4 RNA ligase 1 and thermostable RNA ligase ....	44
Figure 2-3: Cleavage of PCR product with BbsI and MnlI to regenerate 30 nt selected material .....	49
Figure 2-4: Gel shift analysis of starting material and material from rounds 6, 10, and 11 of primer-free SELEX .....	50
Figure 2-5: Ligation of various 30 nucleotide oligos to oligo 1 with Thermostable RNA Ligase .....	56
Figure 3-1: Secondary structures of various classes of aptamers .....	61
Figure 3-2: Gel-shift analysis of the aptamer PF1 and its various constructs .....	70
Figure 3-3: A schematic representation of the primer-template used in the RT inhibition assays .....	75
Figure 3-4: A representative autoradiogram using various concentrations of PF1 .....	75
Figure 3-5: Graph showing the relative counts versus time in the presence of various amounts of PF1 or starting material (5'-N <sub>30</sub> -3') .....	76
Figure 3-6: Graph indicating the relative counts versus time in the presence of various amounts of R1T and 37 NT SELEX ddG .....	78

Figure 3-7: Graph indicating the relative counts versus time in the presence of various amounts of S4 and 37 NT SELEX ddG .....	79
Figure 3-8: A generalized circular dichroism (CD) spectrum of B-form DNA and parallel G-quadruplex in the presence of potassium ions .....	80
Figures 3-9 & 3-10: CD spectra of 4 $\mu$ M PF1, R1T, starting material (5'-N <sub>30</sub> -3') and TBA .....	82
Figure 3-11: CD spectrum of 4 $\mu$ M PF1, PF1-GA, PF1-GC and 37 nt Linear DNA control .....	84
Figure 3-12: CD spectrum of 4 $\mu$ M of PF1, T30, R1T, 37 NT SELEX, G-C Flipback and dA30 .....	84
Figure 4-1: Sequence, structure, K <sub>d</sub> , and $\Delta$ G values of the various single stranded loop-back constructs .....	100
Figure 4-2: Sequence and structure of the various linear single-stranded DNA constructs .....	102
Figure 4-3: Structures of MTS tetrazolium and its formazan product .....	104
Figure 4-4: Cell proliferation in Jurkat E6-1 cells in the presence of various concentrations of the 38 NT Thioaptamer and the control DNAs .....	106
Figure 4-5: Jurkat E6-1 cell count in the presence of various concentrations of the 38 NT Thioaptamer and 37 NT Arbitrary Control .....	107
Figure 4-6: TCID <sub>50</sub> /ml values of Jurkat E6-1 cells infected by HIV-1 LAI virus, with various concentrations of the 38 NT Thioaptamer and its controls .....	108
Figure 4-7: Entry of the aptamer into Jurkat cells visualized in the presence or absence of fluorescein-tagged 37 NT Thioaptamer .....	110
Figure 4-8: Graph showing the quantification of fluorescein-tagged 37 NT Thioaptamer cell entry in the absence or presence of HIV-1 .....	111
Figure 4-9: Effect of 38 NT Thioaptamer and 37 NT $\Delta$ G control on BST-2 expression in Jurkat cells .....	113

## List of Abbreviations

-sssDNA - Minus strand strong stop DNA  
+sssDNA - Plus strand strong stop DNA  
AIDS - Acquired Immunodeficiency syndrome  
BSA - Bovine serum albumin  
CD - Circular dichroism  
DAPI - 4'-6-diamidino-2-phenylindole  
ddG - Dideoxyguanosine  
DEAE - Diethylaminoethyl cellulose  
dG - Deoxyguanosine  
DMEM - Dulbecco's modified Eagle's medium  
dNTPs - Deoxyribonucleotide triphosphates  
dpi - days post infection  
ds - Double-stranded  
DTT - Dithiothreitol  
EDTA - Ethylenediaminetetraacetic acid  
ELISA - Enzyme-linked immunosorbent assay  
FBS - Fetal bovine serum  
HEK cells - Human embryonic kidney cells  
HIV - Human immunodeficiency virus  
HRP – Horseradish peroxidase  
IC<sub>50</sub> - Half-maximal inhibitory value  
IFN - Interferon  
IN - Integrase enzyme  
Kb – Kilobase  
K<sub>d</sub> – Dissociation constant  
kD - Kilo Daltons  
LTR - Long terminal repeat  
Min - Minutes  
NC - Nucleocapsid protein

NMR - Nuclear magnetic resonance  
nt - Nucleotides  
PAGE - polyacrylamide gel electrophoresis  
PBS - Primer binding site (Chapter 1), or phosphate-buffered saline (Chapter 4)  
PCR - Polymerase chain reaction  
Pen-Strep - Penicillin-streptomycin  
PNK - Polynucleotide kinase  
PPT - Polypurine tract  
PR - Protease enzyme  
RNase H - Ribonuclease H  
rNTPs - Ribonucleotide triphosphates  
RPMI - Roswell Park Memorial Institute medium  
RT - Reverse Transcriptase  
SDS - Sodium dodecyl sulfate  
SELEX - Systematic evolution of ligands by exponential enrichment  
ss - Single-stranded  
TBST - Tris-buffered saline with Tween 20  
TCID<sub>50</sub>/ml - 50% tissue culture infective dosage

# Chapter 1: Introduction

## 1.1 Discovery of HIV and AIDS

Human immunodeficiency virus (HIV) is a lentivirus (*lenti*, “slow (growth)” in Latin) in the *Retroviridae* family. The *Retroviridae* (retrovirus) family utilizes reverse transcriptase in a reverse flow of genetic information, which the enzyme converts the viral RNA into DNA during replication. HIV is the causative agent of Acquired Immunodeficiency Syndrome (AIDS), where the virus targets and decimates T helper cells and macrophages in humans. AIDS is defined by an individual testing positive for HIV antibodies or positive for HIV by PCR, and their CD4<sup>+</sup> T-lymphocyte count in blood is below 200 cells/ $\mu$ l. The infected individual also exhibits one or more of a collection of 26 known opportunistic infections and AIDS-related symptoms (150).

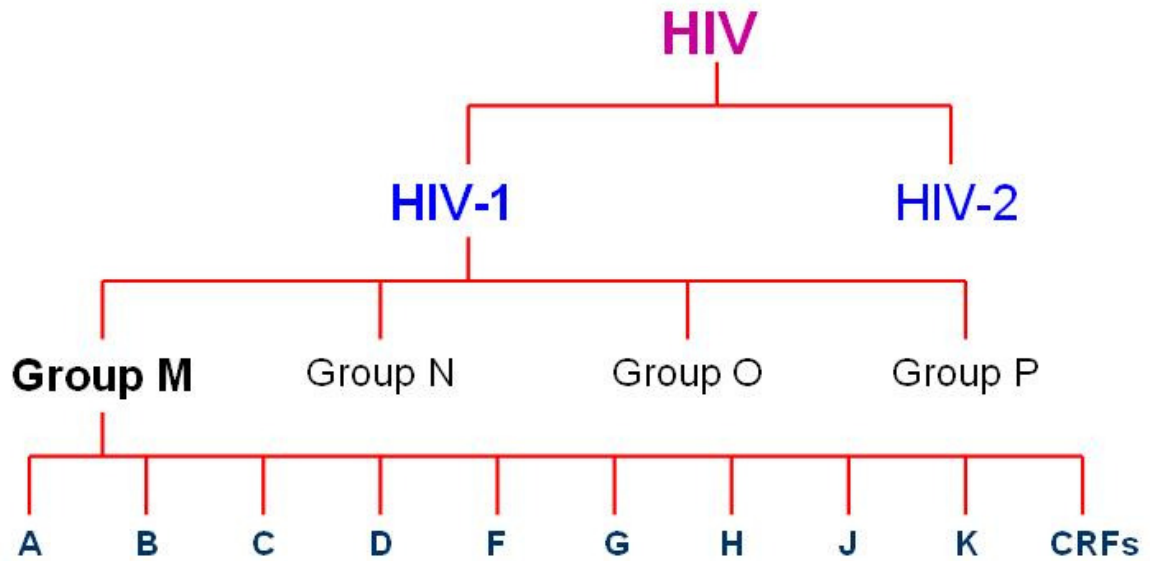
HIV was co-discovered by two lab groups in 1983. Françoise Barré-Sinoussi and Luc Montagnier *et al.* at the Pasteur Institute in Paris isolated the virus from the lymph nodes of a patient with opportunistic infections (pathogens which normally do not cause diseases in a healthy individual). The virus was named “lymphadenopathy-associated virus“, or LAV (14). Robert Gallo’s group at the U.S. National Institutes of Health used samples from Montagnier to purify and characterize the virus, which they named human T cell lymphotropic virus III (HTLV III) (54). In 1986, the Human Retrovirus Subcommittee of the International Committee on the Taxonomy of Viruses decided on the official name: Human Immunodeficiency Virus.

## **1.2 AIDS Epidemiology and HIV Classification**

Since the first recording of clinical cases of HIV/AIDS in 1981, AIDS has claimed more than 33 million lives. Even though the number of deaths each year is steadily declining, about 1.8 million people died from AIDS-related diseases in 2009, including 260,000 children (175). As of 2009, ~33.3 million people were living with HIV/AIDS, including 15.6 million women and 2.5 million children (175). HIV is a bloodborne pathogen which spreads through direct contact of bodily fluids via unprotected sex and contact with blood and/or blood products. The high-risk populations include workers in the sex industry (prostitutes, pornographic film actors, etc.) and intravenous drug users that practice needle sharing. Although blood transfusion patients were previously in this group, current blood screening practices have greatly decreased the risk of contracting HIV by transfusion. In the USA, homosexual men remain a high risk group due to the high level of HIV infection. HIV positive mothers can pass the disease to their newborn infants during childbirth or through breastfeeding. Health care and research workers are also at risk of coming into contact with HIV-contaminated samples, although the risk of contracting HIV remains low.

HIV has been defined as a quasispecies, which is a population of closely related yet genetically distinct viruses (182). The rapid divergence of HIV is due to several factors, including the large number of infected individuals, vigorous recombination, and an error-prone polymerase. These genetically distinct viruses are subjected to continuous genetic variation, selection, and competition (47). There are two types of HIV (Figure 1-1): HIV type 1 (HIV-1) and type 2 (HIV-2), which share 60% homology between their genomes (61). The majority of HIV infections worldwide are from HIV-1, while HIV-2





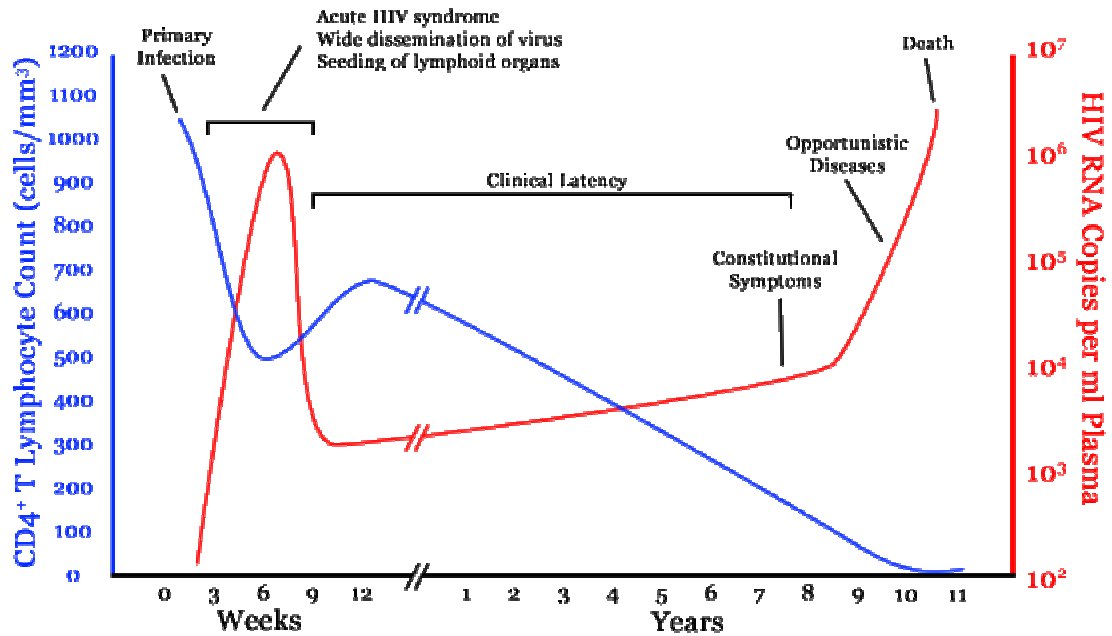
**Figure 1-1: Classification of the human immunodeficiency virus.** HIV-1 and Group M (Major) are the predominant types in each division respectively. Note that subtypes E and I are not included: subtype E are now known as CRF01\_AE (see above), while subtype I is a CRF consisting of subtypes A, G, H, K and other unclassified regions. Figure modified from (10).

cases are mainly found in West and Central Africa. HIV-1 is more infectious and easier to transmit than HIV-2. There are four groups of HIV-1: M (Major), O (Outlier), N (Non-M non-O) and the recently discovered P (Pending) (134). Group M is the predominant group which causes over 90% of the worldwide HIV-1 infections. Group M is further divided into ten subtypes, or clades (10). The different subtypes can undergo recombination during infection to produce “recombinant form” (RF) viruses. Infectious RFs that circulate and become persistent within a population are dubbed “circulating recombinant forms” (CRFs), which are considered one of the ten subtypes. The most epidemiologically abundant CRF, CRF01\_AE (formerly misclassified as subtype E), constitutes almost 5% of total HIV infections worldwide (138) .

### **1.3 HIV Pathogenesis**

HIV infection is marked by the detection of HIV antibodies, and a gradual decrease in total T cell counts - especially of CD4<sup>+</sup> T cells, which is the primary target of HIV. A healthy, uninfected individual has CD4<sup>+</sup> T cell counts between 800 to 1200/μl of blood. On average, the development of AIDS occurs about 8 to 12 years after initial infection with HIV in untreated hosts. Infection can be divided into three stages as listed below (Figure 1-2):

1. Acute (primary) infection: acute infection occurs 2-4 weeks after exposure to HIV, and lasts between a few days to ten weeks (91). The HIV viral load increase significantly, while the number of circulating CD4<sup>+</sup> T cells in the blood drops to about 50% of the original level (from ~ 1200 cells/μl to ~ 600 cells/μl) (60). HIV-infected patients may show signs such as fever, malaise, swollen lymph nodes



**Figure 1-2: A generalized graph of the levels of HIV RNA copies (viral load) and CD4<sup>+</sup> T cell counts during the different stages of HIV disease progression.** The viral load level is indicated in red, and the CD4<sup>+</sup> T cell count is in blue. The average clinical latency period is between 8-12 years without antiretroviral drug treatment, though the HIV disease progression pattern for each individual depends on their age and health. Figure adapted from (<http://en.wikipedia.org/wiki/File:Hiv-timecourse.png>) and (60).

among other non-specific symptoms. Some HIV-infected individual may be asymptomatic during the acute infection stage. The increase in viruses induces production of antibodies against HIV, and a rise in CD8<sup>+</sup> T cell mediated immune response. The slower progression rate to AIDS is linked to a stronger CD8<sup>+</sup> T cell response in individuals (128). The CD8<sup>+</sup> T cell response and the antibody-mediated response lower the amount of virus in the blood, and allow the CD4<sup>+</sup> T cells to rise back to near normal levels (~ 800 cells/ $\mu$ l). These host responses do not eliminate the virus completely.

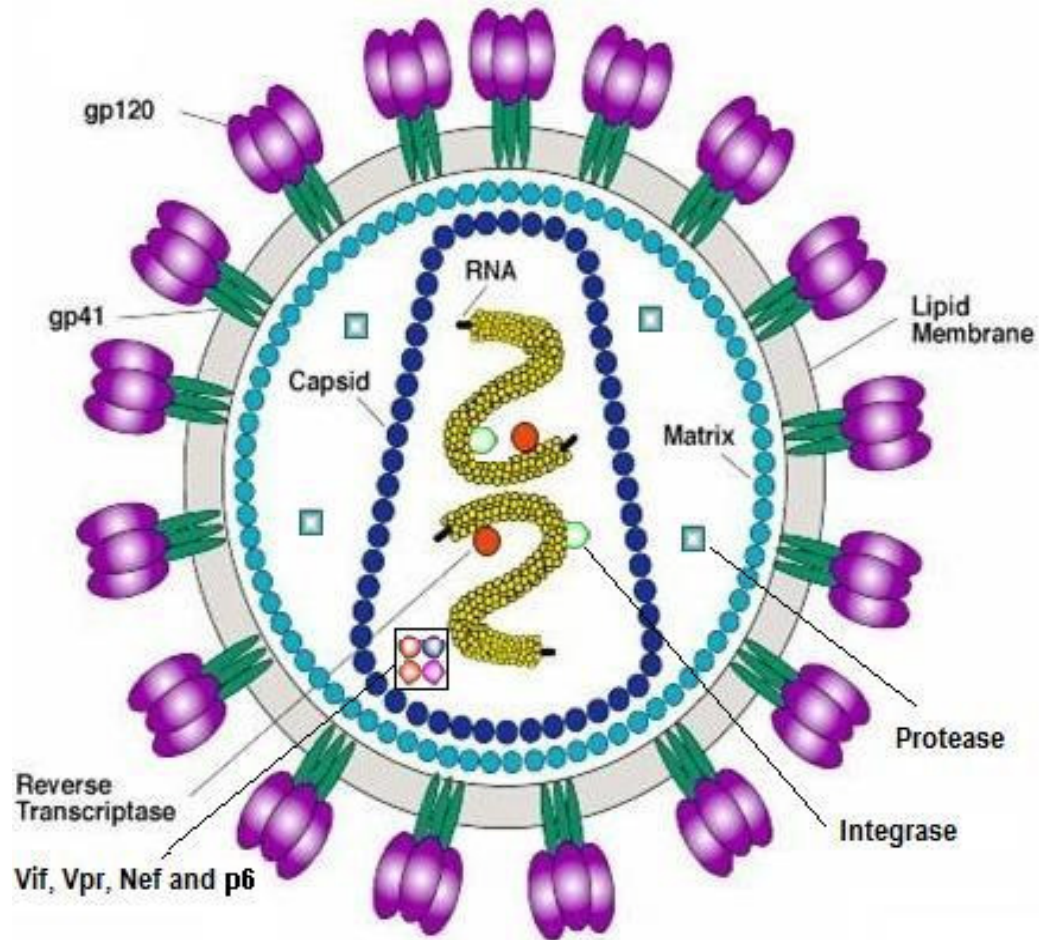
2. Chronic (secondary) infection: the CD4<sup>+</sup> T cell count in the individual's blood stream remains somewhat stable for several months to years without antiretroviral therapy intervention. Under the definition from the Centers for Disease Control and Prevention (CDC), clinical latency is referred as HIV infection stage 1, where the CD4<sup>+</sup> T cell count is between 500-1200 cells/ $\mu$ l in a patient's blood. During the chronic infection period, approximately 10 billion viruses are produced per day. HIV RNA levels remain low and constant, as the host immune system is still functioning to target the infected cells and viruses. Later in the chronic infection a patient may enter Stage 2 HIV infection where the CD4<sup>+</sup> T cell count drops to 200-499/ $\mu$ l, but no AIDS-defining conditions exist. However, other symptoms (exp. Herpes infections, increased cancer risk, papillomavirus infections etc.) usually appear in the HIV-positive individual at this stage.
3. AIDS: the stage where the HIV-infected patient shows a highly decimated CD4<sup>+</sup> T cell count and the appearance of various opportunistic infections. The onset of AIDS (Stage 3 HIV infection) is defined as when a patient's CD4<sup>+</sup> T cell count in

blood is below (200 cells/ $\mu$ l), and the patient is susceptible to AIDS-defining opportunistic infections (OIs) and neoplasms. A wide range of microorganisms that a healthy individual's immune system can normally combat become opportunistic pathogens in an AIDS patient. Some of the common OIs associated with AIDS are *Pneumocystis carinii* pneumonia (PCP), candidiasis, toxoplasma encephalitis, cytomegalovirus infections, as well as others. About 40% of AIDS patients develop neoplasms and cancers such as Kaposi's sarcoma, Burkitt's lymphoma, B cell lymphomas and cervical cancer (from human papillomavirus). AIDS also causes neurological and psychiatric effects: AIDS dementia complex (ADC), multifocal leukoencephalopathy and various myelopathies are observed in later stages of AIDS.

#### **1.4 HIV Structural Characteristics**

HIV is a complex, spherical virus about 80-100 nm in diameter (Figure 1-2). All retroviruses, including HIV, contain a lipid-based envelope, which is derived from the host cell outer membrane lipid bilayer. The virus encodes the surface glycoproteins gp120, and a transmembrane gp41 glycoprotein (13).

The inner membrane of the lipid envelope is lined by a ~17 kDa matrix protein (MA, p17). The genetic material of the virus is enclosed by a bullet-shaped core (32), made of the 24 kD capsid protein (CA, p24). The two copies of the single-stranded RNA genome are coated with the nucleocapsid protein (NC, p7), a nucleic acid chaperone that aids destabilization of secondary structures and facilitates template annealing, reverse transcription, and recombination (106).



**Figure 1-3: A schematic diagram of human immunodeficiency virus (HIV) and its structural characteristics.** HIV is an enveloped virus containing the surface glycoproteins gp120 and gp41. The lipid envelope is lined by the matrix protein (MA). The capsid (CA) protein forms a bullet-shaped core enclosing two copies of single-stranded plus sense RNA genome. Both copies of the RNA are coated with nucleocapsid protein (NC). The capsid core contains copies of the viral enzymes: such as reverse transcriptase (RT), protease (PR) and integrase (IN). Figure adapted from (<http://www.stanford.edu/group/virus/retro/2005gongishmail/hiv1.jpg>), ([http://en.wikipedia.org/wiki/HIV\\_structure\\_and\\_genome](http://en.wikipedia.org/wiki/HIV_structure_and_genome))

The capsid core also contains the three HIV key enzymes: reverse transcriptase, protease and integrase. Reverse transcriptase (RT) is a heterodimer with 51 and 66 kD subunits. Protease (PR, ~10 kD) is the enzyme responsible for cleavage of the HIV polyproteins. Integrase (IN, ~ 32 kD) facilitates proviral DNA insertion into the host DNA genome. Other HIV regulatory proteins, such as virion infectivity factor (Vif), viral protein R (Vpr), and Negative Factor (Nef) are also present in the virion (see Section 1.5 for details on their functions).

### **1.5 HIV Genetic Organization and Protein Functions**

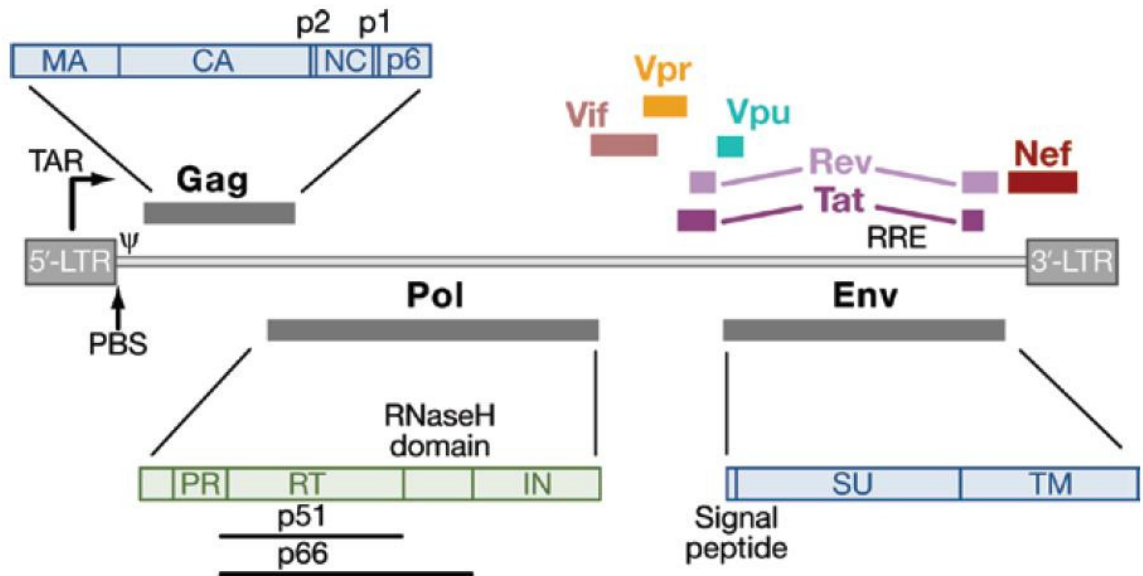
The HIV genome consists of two copies of plus sense single-stranded linear RNA, each of which is approximately 9.1 kb long. The two (nearly) identical copies of the RNA strands are linked together by a non-covalent bond at the 5' end to form a dimer. HIV is considered to be a pseudodiploid virus: only one of the two RNA genomes will be converted into the proviral double-stranded DNA for integration into the host DNA genome (74). The two genomes contribute to the genetic recombination in the virus, and enhance the survival rate of the virus if one of the RNA genome is damaged (74). The HIV three major genes are *gag*, *pol* and *env* (Figure 1-4). The *gag* codes for structural proteins (matrix, capsid, and nucleocapsid), while *pol* encodes the three key enzymes: reverse transcriptase (RT), integrase and protease; *env* is responsible for production of surface and transmembrane glycoproteins (gp120 and gp41, respectively). Table 1-1 lists the protein products of the nine genes in HIV, and the protein functions in the HIV life cycle (161). The protein products in HIV are generated by translation of full length genomic mRNA (*gag* and *pol* gene proteins), or mRNAs spliced from the full length

<b>Viral protein</b>	<b>No. of Protein copies / virion</b>	<b>Viral protein function in HIV replication</b>
Capsid, CA (p24 <sup>Gag</sup> )	~5000	- involves in core structure and assembly of the virion
Matrix, MA (p17 <sup>Gag</sup> )	~5000	- targets Gag to the plasma membrane for viral assembly - promotes Env incorporation - viral post-entry events
Nucleocapsid, NC (p7 <sup>Gag</sup> )	~5000	- genomic RNA packaging - facilitates reverse transcription - RNA chaperone activity - assembly of virion
p6 <sup>Gag</sup>	~5000	- promotes virion budding
Reverse Transcriptase, RT	~250	- proviral DNA synthesis from RNA or DNA templates - RNase H cleavage of RNA
Integrase, IN	~250	- insertion of proviral cDNA into host DNA
Protease, PR	~250	- proteolytic cleavage of Gag and Gag-Pol polyproteins
Surface Glycoprotein, SU (gp120 <sup>Env</sup> )	4-35 trimers	- host cell surface receptor binding - facilitates attachment and entry
Transmembrane Glycoprotein, TM (gp41 <sup>Env</sup> )	4-35 trimers	- fusion peptide aids membrane fusion and entry of the virion
Virion Infectivity Factor, Vif	1-150	- APOBEC3G/F suppression
Viral Protein R, Vpr	~700	- enhances post-entry infectivity - G2/M cell-cycle arrest
<i>trans</i> -Activator of Transcription, Tat	N/A	- activates viral transcription elongation by interacting with TAR
Regulator of Expression of Virion Proteins, Rev	N/A	- Enhances nuclear export of viral RNAs with introns in viral RNAs
Viral Protein U, Vpu	N/A	- CD4/MHC downregulation - viral release
Negative Factor, Nef	Yes, cleaved by PR	- CD4/MHC downregulation - T-cell activation - blocks apoptosis - pathogenicity determinant - enhances viral infectivity

**Table 1-1: A list of HIV proteins and the protein functions in the HIV life cycle.**

Table adapted from Swanson and Malim, Cell 2008 (161).





**Figure 1-4: Organization of the HIV genome, and the encoded protein products.** The positions of the nine genes are relative. The products resulting from *gag*, *pol* and *env* are described. Note that the Gag-Pol polyprotein is produced as one single protein due to the -1 frameshifting mechanism for the *pol* gene to be translated after *gag* (23). *Tat* and *rev* are both shown upstream and downstream of *env*, since splicing within the *env* gene is required to complete the coding of the *tat* and *rev* proteins. See Table 1-1 for the HIV protein functions. Figure adapted from Swanson and Malim, Cell 2008 (161).

mRNA (all other proteins), and cleavage of the polyproteins after translation (Figure 1-4).

## 1.6 HIV-1 Life Cycle

The HIV-1 life cycle is divided into six main steps (Figure 1-5):

1. **Viral Entry:** HIV entry relies on the binding of the viral surface glycoprotein (SU, gp120) to the CD4 receptors, a type of glycoprotein expressed on the surface of T helper (CD4<sup>+</sup>) cells or macrophages. In addition to the CD4 receptor, the virus requires a chemokine co-receptor on the host cell to attain viral entry. HIV gp120 either attaches on the CXCR4 (X4) receptor of CD4<sup>+</sup> cells, or the CCR5 (R5) receptor found on macrophages. The binding of the gp120 to both the CD4 receptor and the chemokine receptor induces a conformational change in the HIV transmembrane glycoprotein (gp41) (147). The gp41 unfolds its hydrophobic domain into the host cell membrane (146), which facilitates the fusion of the viral and host cell membranes (97).
2. **Reverse Transcription:** After entry into the cell, the HIV capsid core undergoes structural changes and reverse transcription begins in this modified core structure. Although the viral RNA genome contains translatable information, HIV undergoes an intermediary double-stranded DNA stage by reverse transcription (166). Reverse transcription is the conversion of the single-stranded RNA genome into a double-stranded DNA. HIV-RT can both utilize RNA and DNA as templates for DNA synthesis. The minus sense DNA strand is first synthesized from the RNA genome. The ribonuclease H (RNase H) activity in RT cleaves the

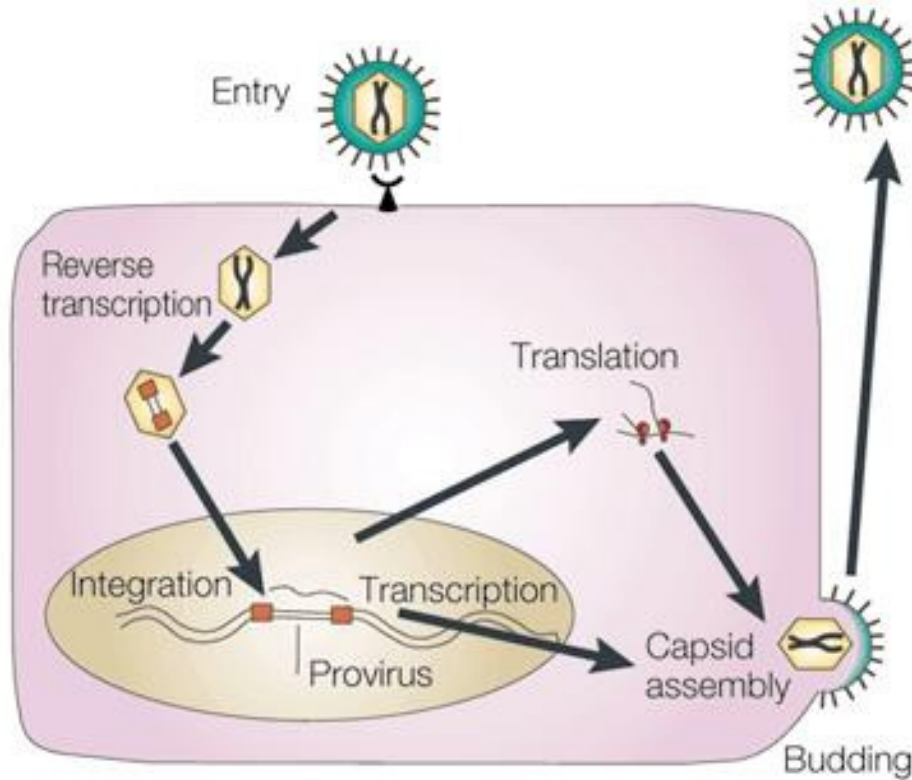
RNA template strand. The minus sense DNA is used as a template for the plus sense DNA strand (see section 1.7 for detailed steps).

3. **Proviral DNA Integration:** The enzyme integrase facilitates insertion of the proviral DNA into the host genome (25). The integration location on the host DNA is random, although active transcription sites seems to be favored (36, 152). A “central DNA flap” aids nuclear import of the viral double-stranded DNA into the host cell nucleus (9, 189). The proviral DNA is cleaved by HIV integrase. HIV integrase creates “sticky ends” by removing two nucleotides from the 3' ends of the viral DNA strands. The processed proviral DNA is inserted into the host DNA sequence. After integration, the host cell repairs the DNA, thus completing the process of incorporating the viral DNA into the host genome. The viral DNA is stable as a provirus within the host cell and the virus can utilize host RNA and protein synthesis machinery for transcription and splicing of the mRNA and translation of viral proteins.
4. **Viral mRNA synthesis and processing:** The proviral DNA initiates RNA transcription at the promoter site in the long terminal repeats (LTRs). Even though HIV contains a single-strand RNA genome, it cannot be directly used as a template for translation due to the lack of a 7-methylated-G cap at the 5' end (32). The newly transcribed viral mRNA undergoes the 5' capping, while the 3' end undergoes polyadenylation directed by the LTRs.
5. **Viral protein synthesis:** HIV protein synthesis occurs at the host cell cytoplasm, after the processed or full length viral mRNA is exported from the nucleus. Export of mRNAs that are spliced multiple times is favored early in infection.

Later, as the REV (which is made from a multiple spliced mRNA) concentration builds up, export of singly spliced and unspliced mRNAs increases (135). The *gag-pol* portion of the HIV-1 is transcribed as a single gene from unspliced mRNA. The Pol portion of the Gag-Pol polyprotein contains the viral enzymes (reverse transcriptase, protease and integrase) that are essential for HIV-1 replication and viral infectivity. Normally, only the *gag* portion is translated into the Gag protein product, however about 5-10% of the translation produces the Gag-Pol polyprotein (23, 79). The generation of the Gag-Pol relies on a programmed -1 ribosomal frameshift, which requires a “slippery site” which allows the simultaneous “slippage” of the two ribosome-bound tRNAs from the 0 reading frame to -1 frame. This step allows the tRNA to bypass the stop codon of the *gag* to continue translating the *pol* gene (79). The frequency of HIV-1 programmed -1 frameshifting is carefully controlled to achieve the precise ratio for the Gag:Gag-Pol proteins. Even a subtle modification of the Gag:Gag-Pol ratio can have significant effects on viral infectivity. Overexpression of the Gag protein produces non-infectious particles due to the lack of viral enzymes (26, 45). The overproduction of the Gag-Pol polyprotein can produce excess protease in the cell, which inhibits the HIV assembly of structural proteins into complete virions (29, 93, 129). When the Gag:Gag-Pol ratios in cells were changed from 20:1 to 20:21, the genomic RNA dimerization was progressively inhibited, and viral infectivity was reduced by about 1000-fold (153). The polyproteins translated from *gag*, *gag-pol* and *env* are cleaved by proteases. The Gag and Pol proteins are cleaved by the viral protease (PR), while the Env polyprotein gp160

is processed by host proteases to produce the surface gp120 and transmembrane gp41 proteins.

6. Assembly, budding and maturation of HIV: During assembly, the two single-stranded viral RNA genomes, structural proteins and viral enzymes are directed toward the cell surface of the host. The viral components are directed towards the lipid rafts at the host cell membrane, which provides a platform for viral assembly (181). The HIV-1 Gag or Gag-Pol precursor proteins ( $\text{Pr55}^{\text{Gag}}$ , and  $\text{Pr165}^{\text{Gag-Pol}}$ , respectively) associate with the host cell membrane by a myristic acid domain and a highly basic region in the matrix portion of the  $\text{Pr55}^{\text{Gag}}$ . Previous studies have showed that the matrix portion in Gag also interacts with the cytoplasmic tail in the gp41 transmembrane protein (181). The  $\text{Pr55}^{\text{Gag}}$  proteins also interact with each other (known as the Gag-Gag interaction) and multimerizes. During multimerization the Env protein and other host proteins are incorporated into the budding virion. The highly multimerized curvature on the cell surface, with the aid of the ESCRT facilitates the virus to “pinch off” from the host membrane and exit the host cell. The maturation of HIV occurs during the late assembly stage or during/immediately after viral budding. Maturation occurs when the HIV polyproteins (for example, Gag-Pol polyprotein) are cleaved by HIV protease to produce functional protein units. Thus, blocking polyprotein cleavage (maturation) provides a good target for HIV-1 drug therapy (193). The matured, infectious progeny can now infect and replicate in a new host cell.

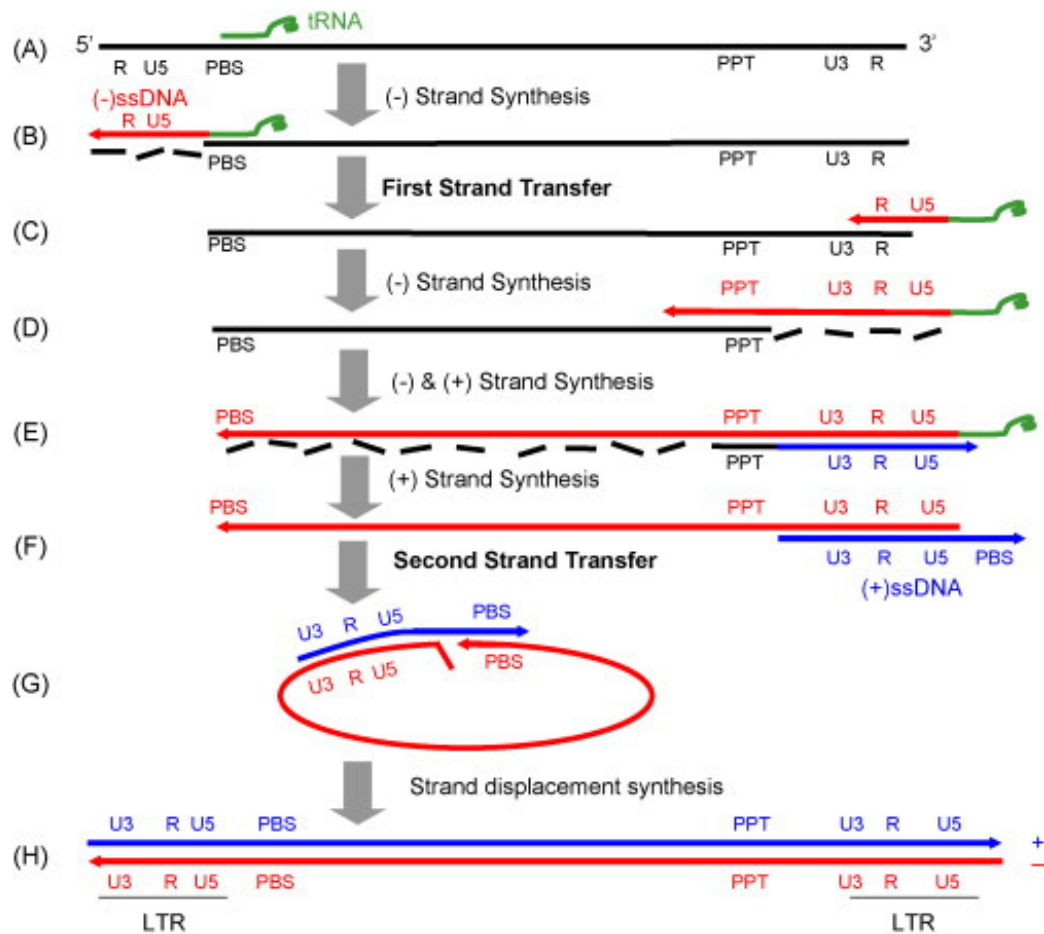


**Figure 1-5: A schematic of the HIV-1 life cycle.** HIV enters to the host cell by attaching onto the CD4 and CXCR4/CCR5 co-receptors on the host cell surface. After the fusion of the viral and host cell membrane, the virion enters into the cell and releases its core into the cytoplasm. Reverse transcription takes place to convert the single-stranded RNA genome into a double-stranded DNA. The proviral DNA then integrates into the host genome to form a provirus. The virus utilizes the host cell RNA polymerase II to transcribe the viral mRNA for protein synthesis, as well as new genomic RNA for new viruses. The genomic RNA, structural proteins and enzymatic polyproteins are assembled at the host cell membrane. The immature virions exit the cell surface and viral maturation completes after budding. Figure adapted from Negroni and Buc, *Nature Reviews Molecular Cell Biology* 2001 (119).

## 1.7 Reverse Transcription in HIV-1

One of the key steps in HIV replication involves copying the single-stranded plus sense RNA into double-stranded DNA (32). The conversion from ssRNA to dsDNA (reverse transcription) is catalyzed by HIV-RT in the following steps (Figure 1-6):

1. The RNA to DNA conversion begins with the synthesis of the minus sense DNA strand. The minus strand DNA synthesis initiates with a tRNA<sup>Lys3</sup> primer, packaged with the virus from the previous host cell. Instead of initiating DNA synthesis at the 3' end of the RNA genome, the HIV-RT starts DNA synthesis close to the 5' end at the primer binding site (PBS), located approximately 182 nucleotides downstream of the 5' end. The 3' end of the tRNA binds to the PBS (Figure 1-6A).
2. RT copies the genome up to the end of the 5' repeat region (R, Figure 1-6B). The newly synthesized DNA is ~164 bp long and is known as the minus strand strong-stop DNA (-sssDNA). The -sssDNA is translocated to the 3' end of the genomic RNA (Figure 1-6C). Besides its polymerase activity, HIV-RT also has a ribonuclease H (RNase H) component. The RNase H activity degrades the template RNA while the synthesis of the minus strand DNA takes place (Figure 1-6D). The minus strand DNA synthesis proceeds up to the PBS site, and the RNA template strand is almost completely degraded during this process.
3. Only a 15 nucleotide (nt) portion in the RNA genome called the polypurine tract (PPT) is not degraded by RNase H during minus strand synthesis. The PPT is the site for initiation of plus strand DNA synthesis. (Figure 1-6E). The tRNA primer at this point is still attached to the 3' end of the -sssDNA. Plus strand DNA



**Figure 1-6: Schematic representation of HIV-1 replication.** The HIV utilizes its reverse transcriptase (RT) to copy the single-stranded RNA into a double-stranded DNA (dsDNA). The minus strand DNA is initiated by a tRNA<sup>Lys</sup> primer, followed by transfer of the newly synthesized DNA strand to the 3' end of the RNA genome. RT produces the minus strand, while RT degrades the RNA template save the PPT. Plus strand DNA synthesis starts at the PPT. After the second strand transfer, the plus and minus DNA strand use each other as templates to produce a fully synthesized dsDNA with long terminal repeat (LTR) ends. The LTRs facilitate integration of the dsDNA into the host genome. Figure adapted from Basu *et al.*, *Virus. Res.* 2008 (15).



synthesis starts at the 3' end of the PPT. The plus sense strong stop DNA (+sss DNA) is produced when RT terminates at the 19th nucleotide from the 3' end of the tRNA primer.

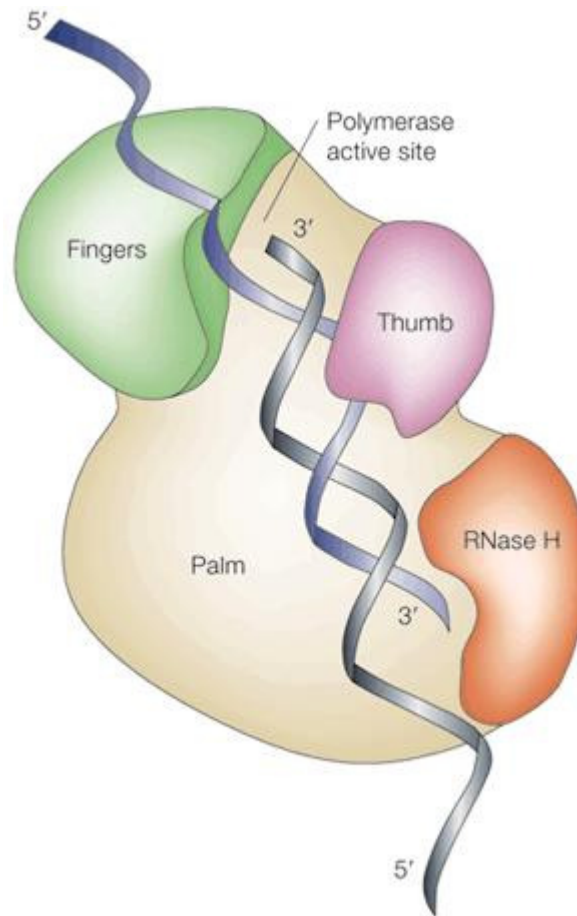
4. After +sssDNA synthesis, the tRNA is removed from the 5' end of the minus strand DNA by RNase H. This frees the 3' end of the newly synthesized +sssDNA. Synthesis of the plus strand involves another strand jump by the +sssDNA to the 3' end of the minus strand DNA, the PBS of the minus strand DNA and the 3' end of the +sssDNA bind due to complementary base pairing (Figure 1-6F).
5. The minus and plus strands utilize each other as templates for DNA synthesis until the double stranded DNA is fully made, creating long terminal repeat (LTR) ends (Figure 1-6H). The long terminal repeats (LTRs) are located at both the 5' and 3' ends of the double stranded DNA and are used during integration into the host genome (32).

## **1.8 Reverse Transcriptase**

Reverse transcriptase was first discovered independently in the early 1970s by David Baltimore and Howard Temin (11, 167). HIV-RT is encoded from the *pol* gene. The functional enzyme is a heterodimer with two subunit components of 66 kDa and 51 kDa. The RT enzyme is a DNA- and RNA-dependent DNA polymerase. HIV-RT also contains a Ribonuclease H (RNase H) domain, which degrades the RNA in the RNA-DNA intermediate during reverse transcription (12). The 66 kD subunit (p66) provides the active site for the DNA polymerase and the RNase H activities (Figure 1-7). The 51

kD subunit (p51) is derived from proteolytic cleavage of p66, 120 amino acids from the carboxy-terminus (C-terminus) (1). The RNase H domain lies in at the C-terminus of p66, while p51 only contains the DNA polymerase. Despite this, p51 shows little enzymatic activity, either on its own or in the dimerized form of RT (55, 105). The DNA polymerase activity and the RNase H activity in RT are found in p66. However, p66 attains optimal activity in the p66/p51 heterodimer, though p66 homodimers also have activity. Even though the exact role of p51 in RT is unclear, p51 may provide a scaffold for protecting and stabilizing the p66 subunit, interacting with the tRNA primer to initiate reverse transcription, and the binding of p66 to the primer-template (64, 164).

HIV-RT requires divalent cations as cofactors for both DNA polymerase and the RNase H enzyme activity (160). As with other polymerases,  $Mg^{2+}$  is the relevant physiological divalent cation although activity can be detected *in vitro* with others as well. The metal ions bind to the “catalytic triad” (Asp-110, Asp-185 and Asp-186) at the RT polymerase active site (73). Crystal structures of RT indicate that the enzyme assumes a right-hand conformation (80). Similar to other DNA polymerases, each subunit contains a finger, palm and thumb domain. The p66 palm and finger “grasps” the primer-template closer to the active site of the DNA polymerase (98) (Figure 1-7). The RNase H active site has four conserved residues that also coordinate  $Mg^{2+}$  binding (Asp-443, Glu-478, Asp-498 and Asp-549) (115, 139). The DNA polymerase and the RNase H active sites are about 18 bases apart, and primary RNA cleavage events occur ~18 bases downstream of the 3' end of the DNA primer in the RNA-DNA intermediate (80). Subsequent cleavage events that lead to shorter RNA cleavage products also occur after the primary event.



**Figure 1-7: Structural representation of the catalytic subunit p66 in HIV-1 reverse transcriptase (RT) and its various domains with the RNA-DNA hybrid.** The p51 subunit is situated behind the palm domain of p66. The RNA template and the DNA primer are shown as the strands in blue and grey respectively. Figure adapted from Negroni and Buc, Nature Reviews Molecular Cell Biology 2001 (119).

HIV-RT does not possess 3'-5' exonuclease activity to excise any misincorporated or mispaired nucleotides (16). Similar to the polymerases in many RNA virus, the lack of a proofreading activity in RT makes it a highly error-prone polymerase, especially when compared to the cellular DNA polymerase. RT has an error rate between  $10^{-4}$  to  $10^{-5}$  per incorporation event (1 error per 10,000 to 100,000 nucleotides incorporated) (17, 22, 188). RT's high error rate coupled with frequent recombination contributes to the genetic diversity of retroviruses, which has posed a problem in combating HIV/AIDS via drug treatment or vaccine developments due to escape mutants arising rapidly.

### **1.9 Genetic recombination in HIV**

Due to the high error rate in the DNA polymerase during HIV replication, HIV undergoes frequent recombination of its genetic material. HIV packages two copies of its viral RNA genome in each virion. The two RNA strands can be genetically identical or vary greatly. The genetic diversity stems from the errors from the viral DNA polymerase during reverse transcription or the host RNA polymerase during RNA transcription. If the host cell is co-infected by HIV of two different subtypes (30), the progeny may contain genomic RNA from either or both of the subtypes. HIV recombination differs from recombination in prokaryotic and eukaryotic cells which involves strand breaking and rejoining. In retroviruses, most recombination occurs during synthesis of the first DNA strand by a strand transfer/template switching mechanism referred to as "copy-choice" recombination. A second proposed mechanism for recombination occurs during second strand DNA synthesis and is termed "strand displacement-assimilation" (77). This mechanism presumably occurs with a much lower frequency. In retroviral copy-choice

recombination DNA being synthesized on one RNA genome (referred to as the “donor”) transfers (“switches”) to a homologous region on the second RNA genome (“acceptor”). DNA synthesis continues on the acceptor strand thus combining information from both strands. If recombination between two identical viral strands occurs, the progeny will remain the same as the parental strands (except for errors from the RT base incorporation). When the donor and acceptor are of a different genetic makeup, the recombined proviral DNA contains information from both the donor and acceptor. Progeny produced from the recombined provirus may be defective or may have a selective advantage under certain conditions or niches. It is by this mechanism that different viral subtypes as well as RFs and CRFs can originate (see section 1.2 for details).

In addition to recombination occurring at internal regions of the viral genome, strand transfer is involved in two essential steps during HIV replication, as discussed in section 1.8. The first strand transfer occurs after the minus strand strong-stop DNA (-sssDNA) “jumps” to the 3' end of the genome from the 5' end of the RNA template. The second strand transfer occurs when the plus sense strong stop DNA (+sssDNA) jumps to 3' end of the minus strand DNA template, where the primer binding sites of the DNA template and the +sssDNA overlap. Both of the strand transfer events are terminal transfer, since DNA elongation is completed at the end of the template.

Internal strand transfer events may occur at any position on the viral genome (85, 195) and at any point during reverse transcription. The vast majority of internal strand transfer events occur between homologous regions of two genomic RNAs (74, 190). HIV recombination rates vary in different human host cell types. In T lymphocytes on average

3-10 recombinations per HIV replication cycle were found (195), and about 30 recombinations per cycle in macrophages (107). Other simpler retroviruses have a much lower rate of recombination. For example, Moloney murine leukemia virus has recombination rates about 1/10 to 1/50 of the HIV (122), although more recent data has challenged these findings (196). Non-homologous recombination between RNA genomes has also been reported but occurs at a rate that is ~2-3 orders of magnitude lower (190).

In addition to serving a role in generating genetic diversity, internal strand transfer events may increase the chances of completing replication if one of the RNA genomes is damaged (31). The genetic diversity generated from the recombination also enhances the ability of HIV to evade other external factors, such as host immunity (75, 94, 165).

There are several restrictions on the HIV recombination during the replication process. After HIV infection, the viral protein U (Vpu) causes degradation of newly synthesized CD4 receptors in the host cell (185), while Nef removes the existing CD4 on the cell surface via endocytosis (4). This combined reduction of CD4 expression increases viral replication efficiency, however, it decreases the chances of another HIV virion entering the cell. Another restriction involves the dimer initiation sequences (DIS) of the different HIV subtypes. Although viruses from different subtypes (for example type A and type C virus) can undergo recombination (159), the discrepancy between the DIS sequences in different HIV-1 subtypes can prevent the formation of heterodimers and thus limit co-packaging of the virion (46). This leads to a much lower rate of recombination between viruses of different subtypes, especially when the DIS sequences are more highly divergent.

## 1.10 Antiretroviral Therapy

Due to the highly variable nature and the long latency of the virus, an effective vaccine has not been developed. The five main classes of antiretroviral drugs approved by the Food and Drug Administration (FDA) (50) are:

1. Nucleoside reverse transcriptase inhibitors (NRTIs), for example zidovudine (azidothymidine, or AZT), lamivudine (3TC), abacavir (ABC). AZT is the most commonly known and first anti-retroviral drug approved for HIV/AIDS treatment (52). NRTIs mimic their nucleoside counterparts and get incorporated into the elongating DNA strand during reverse transcription. The NRTIs act as chain terminators, as they lack the 3'-hydroxyl group, thus they cannot form the phosphoester bond with the next nucleoside and block further extension by HIV-RT (57).
2. Non-nucleoside reverse transcriptase inhibitors (NNRTIs), for example nevirapine, etravirine, delvaridine. The NNRTIs works by binding non-competitively at an allosteric "binding pocket" near the RT p66 subunit active site (78), thus preventing RT from transcribing the viral genome.
3. Protease inhibitors, for example saquinavir, indinavir, ritonavir. Protease inhibitors are made of a short sequence of amino acids or amino acid analogs that bind to the HIV protease active site. Protease inhibition prevents proteolytic cleavage of long HIV polyproteins, causing the released virions to be non-infectious (immature) due to the lack of functional viral proteins.
4. Integrase inhibitors block the integrase-mediated strand transfer (the process by which the processed 3' ends of the proviral DNA are inserted to the host

chromosomal DNA), therefore preventing the insertion of the viral DNA into the host DNA. The first and only FDA-approved integrase inhibitor is raltegravir (isentress, or MK-0518). Currently there are other integrase inhibitors in development. Examples include elvitegravir, a strand-transfer inhibitor to integrase currently in phase III clinical trial (154), and MK-2048, a second generation integrase inhibitor (6).

5. Entry (or fusion) inhibitors, for example enfuvirtide (fuzeon) and maraviroc. Enfuvirtide is a peptide-based inhibitor that blocks the formation of the six-helix bundle in the transmembrane gp41 fusion glycoprotein (48). Enfuvirtide is only used for HIV patients when all other antiretroviral drug treatments have failed. Enfuvirtide is administered with other antiretroviral drugs. Maraviroc blocks the interaction between the HIV surface protein (gp120) and the host cell chemokine receptor CCR5. However, maraviroc is ineffective against CXCR4-tropic HIV.

The combination of three or four drugs from two or more of the above classes of therapeutics is known as Highly Active Anti-Retroviral Therapy (HAART). Patients under HAART still face the problem of HIV drug resistance. This is due to HIV having an error-prone DNA polymerase and high rates of genetic recombination. Thus, recent treatment protocols suggest starting HAART only when the CD4<sup>+</sup> T cell counts are at or below (350 cells/ $\mu$ l). Alternative methods for combating HIV infection are currently in development. Viral release inhibitors, nucleic acid based inhibitors (aptamers) (87-89), small interfering RNA (siRNA) (151), and HIV Tat inhibitors (37) are just some of the examples of novel HIV inhibitors.



## **1.11 Aptamers and Systematic Evolution of Ligands by Exponential Enrichment (SELEX)**

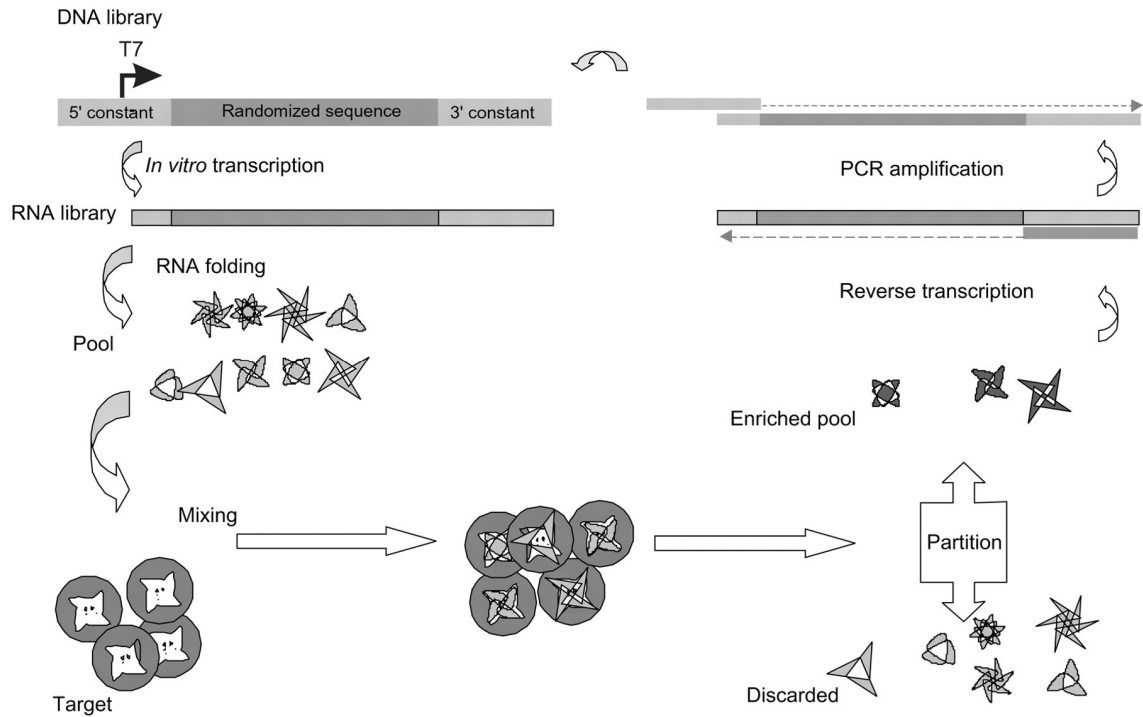
HIV-RT and several other proteins, even those that are not typical nucleic acid binding proteins, have been shown to bind strongly to particular RNA and DNA sequences, often based on the secondary and tertiary structures the sequences form (172, 173). These tight binding sequences - aptamers - have several potential uses as antiviral agents, monoclonal antibody complements, in diagnostics, and in disease therapy (56, 82, 88). Aptamers are mainly produced by a process called Systematic Evolution of Ligands by EXponential enrichment (SELEX). SELEX was invented in the early 1990s by the Gold, Joyce and Szostak labs independently (49, 90, 172) for selecting nucleic acid sequences that bind proteins with high affinity. SELEX selects from a library consisting of a large pool of nucleic acids (up to  $10^{15}$  random sequences) incubated with a limiting amount of target. These specific targets are mostly proteins, but they also include nucleic acids, cultured cells, viruses and other small molecules (27, 67, 68, 103, 113, 133, 163, 183, 184, 186, 191, 194). Nucleic acids that bind to the target (protein) with higher affinity will preferentially associate to form a complex. The nucleic acid-protein complex can be isolated by gel shift, nitrocellulose filter binding, column chromatography, or capillary electrophoresis (118). The nucleic acids that are bound to the target are amplified with reverse transcription Polymerase Chain Reaction (RT-PCR) for RNA aptamers, or PCR for DNA aptamers. New, enriched pools of nucleic acids are subjected to subsequent rounds of target binding. Approximately six to twelve cycles of the selection step are repeated until the sequences that show high affinity (usually with a dissociation constant ( $K_d$ ) in the nM range) towards the target are selected (Figure 1-8).

The first aptamers isolated for HIV-RT were RNAs with pseudoknot-type structures (173). RNA pseudoknot aptamers have been shown to be potent RT inhibitors by interfering with RT primer-template extension (28, 65, 81). Later, SELEX was also used to select for single stranded DNA aptamers with similar properties of RT inhibition (118, 149). Tight binding to RT is mainly determined by the folded structure of the aptamer instead of aptamer sequence. These aptamers typically bind several-fold more tightly than the natural RT substrate (RNA-DNA or DNA-DNA primer-template) (65, 96). Since aptamer binding depends on multiple contacts with RT amino acids, it is suggested that aptamers would produce fewer RT mutants that are resistant to the inhibitory effects of the aptamer, though more evidence is needed to support such a theory (65). Recent reports depict RT aptamer escape mutants (53, 89), however these escape mutants were replication defective in comparison to wild type virus.

SELEX has enabled researchers to select for aptamers that bind to several HIV proteins and targets for other diseases (24, 56, 65, 99, 109, 111). Many of these aptamers are currently being developed as potential treatments for diseases. Other non-therapeutic uses for aptamers are also being explored including their uses as alternatives to antibodies in biochemical assays, diagnostic biosensors (known as aptamer beacons) (117), and as tools to study the mechanisms of viral replication (40, 82, 136). Essentially, any technique that involves a sensitive and stringent recognition of protein or nucleic acids could benefit from aptamer-based approaches. Currently the only Food and Drug Administration (FDA) approved aptamer-based drug is Macugen, which was developed by EyeTech Pharmaceuticals and is used to treat age-related macular degeneration (AMD) by injecting the drug to the center of the retina (121). The reason why there are so

few aptamer-based drugs is due to the limited success in human therapeutics (130). The main disadvantage of DNA aptamer therapy is the lack of an effective method for delivering the inhibitor to target cells in animal models, especially since host nucleases may degrade the inhibitor before it is able to bind to the target protein. Modified sugar and thiophosphate backbones can help protect against degradation, and the potential use of gene vectors for delivery of RNAs is being explored (24, 56). A new approach where oligonucleotides are anchored to small proteins to cargo the protein-nucleic acid complex across the cell membrane is also promising for both RNA and DNA aptamers and other types of nucleic acid inhibitors (171). Recent research suggests that aptamers can enter cells during viral infection without the transfection or conjugation to cargo proteins (110, 111). Matzen *et al.* showed a DNA aptamer directed against the spleen focus-forming virus was able to inhibit viral infection not only in mouse cells, but also when administered to infected mice *in vivo* (110).

Though currently limited as therapeutics due to cost and a lack of delivery systems, aptamers still hold a great deal of promise. Potential uses include not only targeting specific pathogen proteins, but also host proteins involved in pathogenesis. In such cases, aptamers should be useful in defining the potential efficacy of both pathogen and host targets. Finally, the possibility of therapeutic uses should not be overlooked, especially considering that drug-delivery systems research is a highly active field.



**Figure 1-8: The SELEX process.** For RNA- or modified RNA-aptamer discovery, the library is typically transcribed by using T7 RNA polymerase. The RNA molecules are allowed to fold and then incubated with the target protein (or nucleic acid) before the unbound species are partitioned from the bound RNAs. The latter are recovered, reverse-transcribed and then amplified by PCR in order to produce a library of reduced complexity. This library is used to initiate a second cycle of SELEX. Binding of the enriched pool of RNAs to the target is monitored cycle by cycle and its RT-PCR products are cloned once it is judged that the proportion of binding sequences has risen to an adequate level, and they are analyzed individually. Figure adapted from James, J Gen Virol. 2007 (82).

## Chapter 2: A primer-free method that selects high affinity single-stranded DNA aptamers using Thermostable RNA Ligase

### 2.1 Introduction

Systematic evolution of ligands by exponential enrichment (SELEX) is a method for selecting small nucleic acid (RNA or DNA) molecules that bind with high affinity to specific targets. It was originally developed by three separate groups in the early 1990s (49, 90, 172). The selected nucleic acids are called “aptamers”. Though aptamer targets are most often proteins, they also include other small molecules, cultured cells, and viruses among others (27, 67, 68, 103, 113, 123, 133, 163, 183, 184, 186, 191, 194). Aptamers have many potential uses including therapeutic and diagnostic applications, as well as replacement of antibodies in ELISA and other technologies (24, 40, 56, 82, 88, 117, 121, 136, 191). Essentially, any technique that involves highly specific recognition of protein molecules could benefit from aptamer-based approaches. Although several compounds developed from SELEX-based methods are currently in clinical trials, only a single drug for macular degeneration, Macugen (EyeTech Pharmaceuticals Inc. New York, NY USA), has been approved by the Food and Drug Administration (FDA).

Most aptamers are identified through standard SELEX procedures that include a starting pool of nucleic acids with a region of random nucleotides of approximately 20-50 bases long. The random nucleotides are flanked on the 3' and 5' ends by fixed sequences of approximately 20 nucleotides. This results in oligonucleotides of ~60-90 nucleotides that are used for the selection process. After several rounds of selection tight binding aptamers that contain the flanking sequences are selected. Typically, these sequences are

modified to determine the minimal bases that are required for tight binding. A major advantage of this approach is that the flanking sequences provide a simple means to amplify oligonucleotides that are selected in each round of selection. A clear disadvantage is the potential for the fixed sequences to bias the selection by interacting with nucleotides in the random region of the oligonucleotide. Many, if not most selected aptamers require sequences from the fixed flanking regions to maintain tight binding.

Recently, new methods have been published for “primer-free” SELEX, in which selection occurs in the absence of any flanking, fixed nucleotides (126, 127, 184), and “minimal primer” SELEX, in which the fixed sequences are reduced to fewer nucleotides (83, 177), or other approaches in which primer ends are block or switched during the selection process (155). By eliminating fixed flanking regions (completely in the case of primer-free SELEX) during the selection step these methods abrogate the biases noted above. The trade-off is the methods are generally more times consuming than fixed sequence approaches, as several enzymatic and hybridization steps are involved to introduce flanking sequences after the selection step. These additional steps can introduce biases related to enzymatic reactions, and the additional steps can lead to loss of some material. The latter problem is especially important in the first round of selection where all or most sequences are unique.

In this report we introduce a new primer-free method for selecting DNA aptamers. The method takes advantage of a commercially available Thermostable RNA Ligase that functions at 60°C. Under the conditions as defined in this report, the ligase is capable of ligating DNAs with high efficiency (between ~50% and >90% depending on the sequence used), more efficiently than T4 RNA Ligase 1, which is commonly used for

this application. Performance of ligations at high temperature also minimizes the potential effects of second structures on ligation. This method, although not solving all the problems inherent in primer-free SELEX, provides another tool for selection of high affinity aptamers using an approach that does not bias the outcome with fixed sequences.

## **2.2 Materials and Methods**

### *2.2.1 Materials*

Human immunodeficiency virus reverse transcriptase (HIV-RT) was from Worthington Biochemical Corporation. *Taq* polymerase, restriction enzymes, Klenow DNA polymerase, T4 polynucleotide kinase (PNK) and T4 RNA Ligase 1 were from New England Biolabs. Thermostable RNA Ligase and CircLigase™ were from Epicentre Biotechnologies. Calf intestinal alkaline phosphatase (CIAP) and dNTPs were from Roche. Rapid DNA Ligation kits were from Promega. Nitrocellulose filters were from Whatman. *Pfu*Turbo® DNA polymerase and StrataClone™ Blunt PCR cloning kits were from Agilent Technologies. Miniprep DNA preparation kits were from Qiagen. Radiolabeled compounds were obtained from Perkin Elmer. Sephadex G-25 spin columns were from Harvard Apparatus. All oligonucleotides were from Integrated DNA Technologies. The 30 nucleotide random sequence oligomer (5'-N<sub>30</sub>-3') used as a starting pool for selection was synthesized using “hand-mixed” ratios of the four nucleotides in order to assure an equal probability of incorporation of each nucleotide at each position. This material was used directly in the SELEX protocol. All other chemicals were from Sigma, Fisher Scientific, or VWR Scientific.

### 2.2.2 Methods

*5' end-labeling of oligonucleotides:* Reactions for labeling of the starting material (5'-N<sub>30</sub>-3') were done in a 100 µl volume containing 1500 pm of 5'-N<sub>30</sub>-3', 70 mM Tris-HCl, pH 7.6, 10 mM MgCl<sub>2</sub>, 5 mM dithiothreitol (DTT), 10 µls of γ-<sup>32</sup>P ATP (3000 Ci/mmol, 10 µCi/µl) and 2 µls (20 units) of T4 polynucleotide kinase. Because of the large excess of oligonucleotide over label, only a small fraction of the primers have phosphate added to the 5' end. This limits circularization of this oligonucleotide and favors ligation to oligonucleotide 1 (see below) in the ligation reaction with Thermostable RNA Ligase. The reaction mixture was incubated for 30 minutes at 37°C and then the PNK was heat inactivated for 10 minutes at 70°C according to manufacturer's recommendation. The material was then run through a Sephadex G-25 spin column.

Labeling of oligo 1 (5'-/5'Phos/GTCGACGAGGTCTAGAATA/3'dideoxyC/- 3', MnlI site underlined) was performed as above, except that 200 pmoles of oligo 1 (without a phosphate at the 5' end) were incubated in a 50 µl volume as above for 15 minutes. After this non-radioactive ATP was added to a final concentration of 200 µM and the reaction was continued for 20 minutes then processed as described above.

*Selection of material with HIV-RT using nitrocellulose filters for rounds 1 and 2 and subsequent processing:* In the initial round, 3 separate reactions with approximately 500 pm each of a 5' <sup>32</sup>P end labeled 30 nucleotide oligonucleotide of random sequence (5'-N<sub>30</sub>-3') and 10 pm of HIV-RT were prepared in 50 µl of buffer containing 50 mM Tris-HCl (pH=8), 80 mM KCl, 6 mM MgCl<sub>2</sub>, and 1 mM DTT. The reactions were incubated at room temperature for 1 hour, then filtered over a nitrocellulose filter disk (25



mm, 0.2  $\mu\text{m}$  pore size) with suction. Disks were preincubated in the above buffer prior to use, then washed twice with 500  $\mu\text{l}$  of buffer containing 25 mM Tris-HCl (pH=8) and 10 mM KCl. Filter disks were then cut into small pieces and placed in 500  $\mu\text{l}$  of 50 mM Tris-HCl (pH=7) and 1 mM EDTA (pH=8). The material was then incubated at 90°C for 5 minutes then 70°C for 10 minutes. After allowing the material to cool, it was extracted with 500  $\mu\text{l}$  of 25:24:1 phenol: chloroform: isoamyl alcohol, followed by 500  $\mu\text{l}$  of chloroform. The aqueous phase was precipitated by adding 50  $\mu\text{l}$  of 3 M sodium acetate (pH=7), 40  $\mu\text{g}$  glycogen, and 1 ml of ethanol. Precipitated material was collected by centrifugation after incubation at -20°C for several hours. The pellet was washed with 70% ethanol, then dried in a DNA SpeedVac Concentrator. The precipitated material was resuspended in 75  $\mu\text{l}$  of water and passed over Sephadex G-25 spin columns equilibrated in water. Collected material was then dried. Nucleic acid from all 3 reactions (~20 pm total) was then combined and re-selected as above but in a 20  $\mu\text{l}$  final volume using 2 pm of HIV-RT. This step was necessary to reduce the total amount of nucleic acid for the ligation reactions. Nucleic acid (~0.3 pm total) was extracted as described above, except glycogen was omitted from the precipitation step and 30 pm of 5'  $^{32}\text{P}$  end labeled oligo 1 (see above) was added. The collected precipitate was once again run over a Sephadex G-25 spin column as described above and dried. The dried material was resuspended in 21  $\mu\text{l}$  of water and ligations were performed in 3 separate ligation reactions using Thermostable RNA Ligase. All reactions were in 20  $\mu\text{l}$  final volume containing: 50 units (0.5  $\mu\text{l}$ s) Thermostable RNA Ligase, ~10 pm 5'  $^{32}\text{P}$  end labeled oligo 1, ~0.1 pm 5'-N<sub>30</sub>-3', 50 mM MOPS (pH=7.5), 5 mM MgCl<sub>2</sub>, 10 mM KCl, 1 mM DTT, 50  $\mu\text{M}$  ATP, 1 mM hexamine cobalt chloride ([Co(NH<sub>3</sub>)<sub>6</sub>]Cl<sub>3</sub>), and 20% (w:v) PEG-8000. Reactions were

incubated 12-18 hours at 60°C in a thermocycler with a heated lid. After ligation reactions, 20 µl of 2X loading buffer (90% formamide, 20 mM EDTA (pH=8), 0.025% (w:v) each of bromophenol blue and xylene cyanol) was added and samples were run on a 12% polyacrylamide/7M urea denaturing gel (in 1X Tris/borate/EDTA buffer) with a molecular weight marker as described (145). Ligated material (50 nucleotides in length) was located using a phosphoimager (Fujifilm FLA-7000), excised, and then eluted by the crush-and-soak method (145) in 500 µl 50 mM Tris-HCl (pH=7), 1 mM EDTA (pH= 8). Forty µg of glycogen was added to the eluate and the nucleic acid was recovered by ethanol precipitation as described above. The 5' ends of the recovered nucleic acid were phosphorylated in a standard PNK reactions as described above, except the total volume was 25 µl, 20 mM KCl was included, radiolabeled ATP was replaced with 0.4 mM cold ATP and 10 units of PNK were used. After inactivating the PNK (see above), 10 µg of oligo 2 (5'-CTGGTATTCTAGACCTTCGTC-3', MnlI site underlined) was added. This oligo was complementary to a region on oligo 1 (see Fig. 1). Oligo 2 was hybridized to the ligated product by heating to 65°C, and slow cooling to room temperature. Then an extension reaction was performed with Klenow DNA polymerase (1 unit total) by adding dNTPs (final concentration 100 µM) and bringing the volume up to 50 µl, followed by incubation at 37°C for 10 minutes. The reaction volume was then increased to 100 µl with water and the material was extracted and precipitated. The recovered double stranded DNA was blunt-end ligated to a duplex formed using oligo 3 (5'-ATAGCATGAATTCCCGAAGACGC-3', BbsI site underlined) and oligo 4 (5'-CGTCTTCGGAATTCATGC-3', BbsI site underlined) using a rapid DNA ligation kit and T4 DNA ligase. Two pmoles of the oligo 3-oligo 4 duplex was used in the round 1

reactions, and 1 pmole was used in subsequent rounds. The manufacturer's recommended protocol was used in the rapid ligation reactions which were incubated for 30 minutes at room temperature. The entire 10  $\mu$ l ligation reaction (round 1) or 5  $\mu$ l of the reaction (all subsequent rounds) was used in a polymerase chain reaction (PCR). Reactions were performed in the manufacturer's supplied buffer in 100  $\mu$ l total volume, and included fifty pmoles each of oligo 2 and oligo 5 (5'-GCATGAATTCCCGAAGACGC-3'), 200  $\mu$ M each of dATP, dCTP, and dGTP, 25  $\mu$ M dTTP, 2  $\mu$ ls of  $\alpha$ -<sup>32</sup>P dTTP (800 Ci/mmol, 10  $\mu$ Ci/ $\mu$ l), and 2.5 units of *Taq* polymerase. A total of 15-18 cycles was performed using 94°C (30 seconds), 50°C (1 min), and 72°C (1 min) followed by an additional cycle of 5 min at 72°C. Native gel loading buffer (20  $\mu$ l: 40% sucrose, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) was added and reactions were run on a 12% native polyacrylamide (29:1 acrylamide:bisacrylamide) gel and products of the correct size (73 base pairs) were located and recovered as described above using a phosphoimager. Approximately 1/40<sup>th</sup> (for each reaction) of the recovered material was used in a second set of 4 PCRs. The PCRs were performed as described above for 6, 8, 10, or 12 cycles. A range of cycles was performed to insure that a sufficient quantity of the correct product was recovered. If product recovery was too low, another PCR was performed using an optimal cycle number based on the results with the variable cycle reactions. Because of the random 30 base region of the product, too many cycles results in the generation of multiple single strands that do not form complete duplexes. Products of the correct size from the 4 PCRs were recovered from 12% native gels as described above. The eluted material was precipitated and digested in 30  $\mu$ l total volume with 7.5 units each of BbsI and MnlI in the manufacturer's buffer for BbsI for 2 hours at 37°C. The sample was

heated to 90°C for 3 min, then cooled rapidly on ice. Four units of CIAP were added and incubation was continued at 37°C for an additional 30 min. Twenty-five µl of 2X loading buffer was added and the sample was run on a 12% denaturing gel as described above. A marker for the 30 nucleotide single stranded dephosphorylated product was included on the gel. The correct product was located with a phosphoimager and recovered using 40 µg of glycogen to aid precipitation as described above. The recovered material (~10-20 pmoles) was used for a second round of selection with nitrocellulose filters using the same protocol with the changes noted.

*Selection procedure for all rounds after round 2 performed by gel-shift:*

Following the second round, subsequent rounds were conducted using native 6% polyacrylamide gels to recover the selected material by gel-shift. The recovered 30 nucleotide product from the previous round was incubated with HIV-RT (see below) in 10 µl final volume using the conditions stated above for 30 minutes. Two µl of native gel loading buffer was then added and the material was run on a 6% native polyacrylamide gel at ~100 V until the blue dye marker was ~2-3 inches below the wells. Gel-shifted nucleic acid was located with a phosphoimager and eluted as described above. Ten pmoles of radiolabeled oligo 1 was added to the eluate and precipitation in the absence of glycogen was carried out as described above. The recovered material was processed as described above to regenerate the 30 nucleotide random nucleic acid for the next round of selection. After round 6, the KCl concentration in the selection step was increased from 80 to 150 mM. Various amounts of HIV-RT were used in the selection process. Our goal was to recover approximately 1/25<sup>th</sup> of the total nucleic acid material from each round as

shifted material on the gels. In round 3, the first round where the gel-shift was used, 5 pmoles of HIV-RT was added to the selection reaction. In subsequent rounds the level was decreased as the amount of recovered material from the previous round increased. By round 6, 0.5 pmoles of HIV-RT was being used and this amount was used in all subsequent rounds.

*Sequences analysis of products recovered from selection rounds 6, 11, and 12:*

The sequence of the 30 nucleotide random region from products selected in rounds 6, 11, and 12 was determined by performing a PCR with *PfuTurbo*® DNA polymerase, and then cloning the 73 base pair PCR product using StrataClone™ Blunt PCR cloning kit from Agilent Technologies and the manufacturer's suggested protocol. Mini-preps were prepared by the QIAprep Miniprep Kit from Qiagen. Products were sequenced by Genewiz Inc.

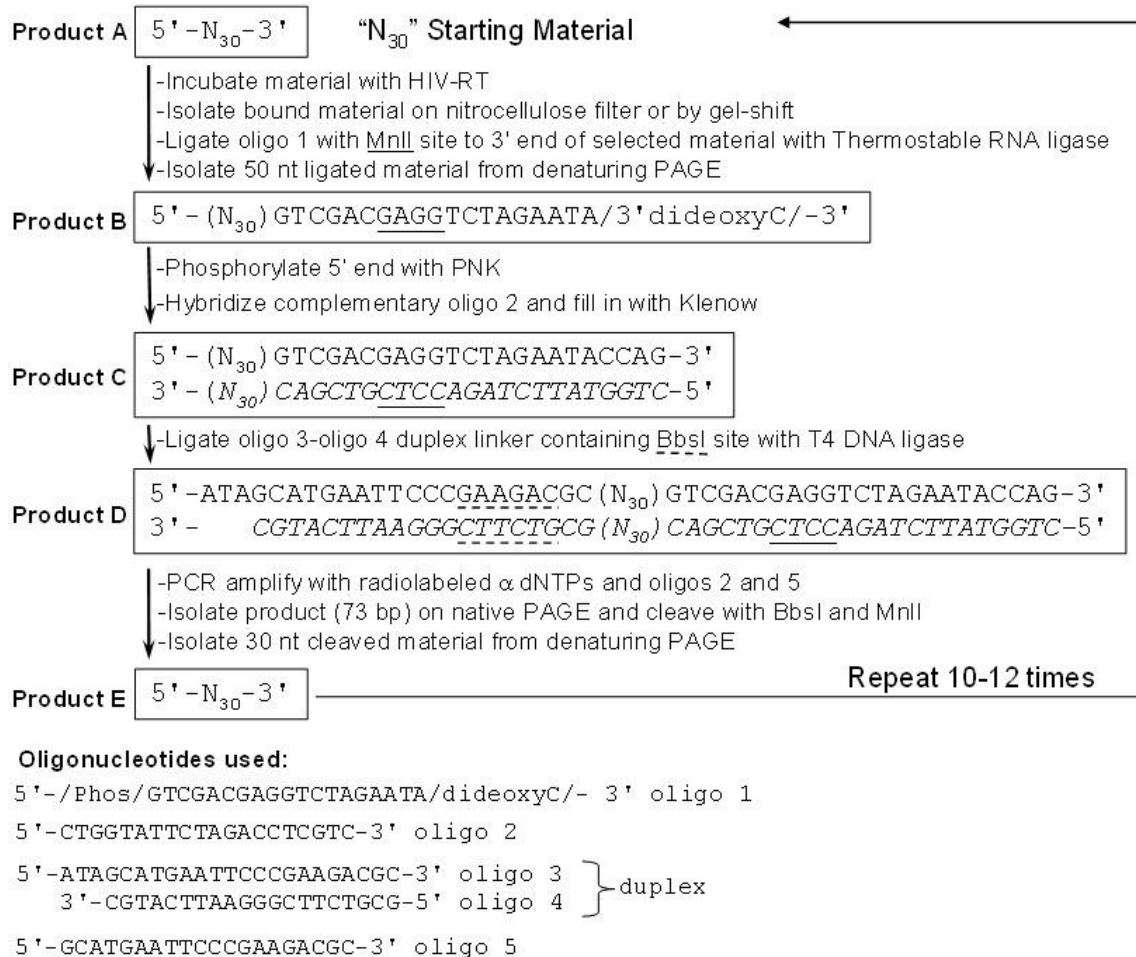
*Determination of equilibrium dissociation constants ( $K_d$ ):* Selected material (~2 nM) from various rounds of selection (see Results) or other designed sequences (see Fig. 5) were either labeled internally, or 5' end labeled with  $^{32}\text{P}$  and mixed with various amounts of HIV-RT (as stated in Fig. legends) in 10  $\mu\text{l}$  of buffer containing 50 mM Tris-HCl (pH=8), 1 mM DTT, 80 mM KCl, and 6 mM  $\text{MgCl}_2$ , for 10 min at room temperature. Two  $\mu\text{l}$  of 6X native loading buffer was added and the samples were run on a 6% native polyacrylamide gel as described above for the gel-shift experiments. The amount of shifted vs. non-shifted material was quantified using a phosphoimager. Values for  $K_d$  were determined by plotting the concentration of shifted product (nM) vs. the

concentration of HIV-RT and fitting the data by nonlinear least square fit to the quadratic equation:  $[ED] = 0.5([E]_t + [D]_t + K_d) - 0.5((([E]_t + [D]_t + K_d)^2 - 4[E]_t[D]_t)^{1/2})$ , where  $[E]_t$  is the total enzyme concentration and  $[D]_t$  is the total primer-template concentration (72). In cases where the affinity of HIV-RT for the nucleic acid was low ( $K_d \geq 1 \mu\text{M}$ ), a more precise  $K_d$  value was not determined.

## 2.3 Results and Discussion

### 2.3.1 *The efficiency of the selection process and its effect on the recovery of tight binding aptamers*

In the initial selection round with an oligonucleotide of 30 bases (5'-N<sub>30</sub>-3', see Figure 2-1), nearly all nucleic acid sequences are unique. With 500 pmoles ( $5 \times 10^{-10}$  moles) of starting material, only ~0.026% of the total number of possible sequences is represented ( $(((5 \times 10^{-10}) \times (6.023 \times 10^{23}))/4^{30}) \times 100$ ). On the surface this would suggest that many, if not most sequences that may bind very tightly to a protein, are unlikely to be present. However, proteins often recognize shorter nucleotide runs or specific structures that can form from a diverse array of sequences. Even if a particular protein bound tightly to a run of 20 specific nucleotides, such a run would be represented ~270 times in the 500 pmoles pool ( $(((5 \times 10^{-10}) \times (6.023 \times 10^{23}))/4^{20})$ ). Because of the way proteins recognize nucleic acids, it is highly likely that most of the possible recognition are motifs or sequences that could form with a nucleic acid that is 30 nucleotides or shorter would be present in a 500 pmole starting pool. More complex structures, for example duplexes resembling primer-templates with recessed 3' termini, or nucleic acids that can simultaneously interact with more than one domain of the protein, are less likely



**Figure 2-1: Protocol for primer-free SELEX.** A schematic diagram of the protocol for selecting 30 nucleotide aptamers is illustrated. Various intermediates are shown (Products A-E) and labeled for reference in the text. Primers used in the protocols are shown at the bottom. Details of individual steps are provided in the Methods section.

to occur in the starting material, mostly because of the short 30 nucleotide length.

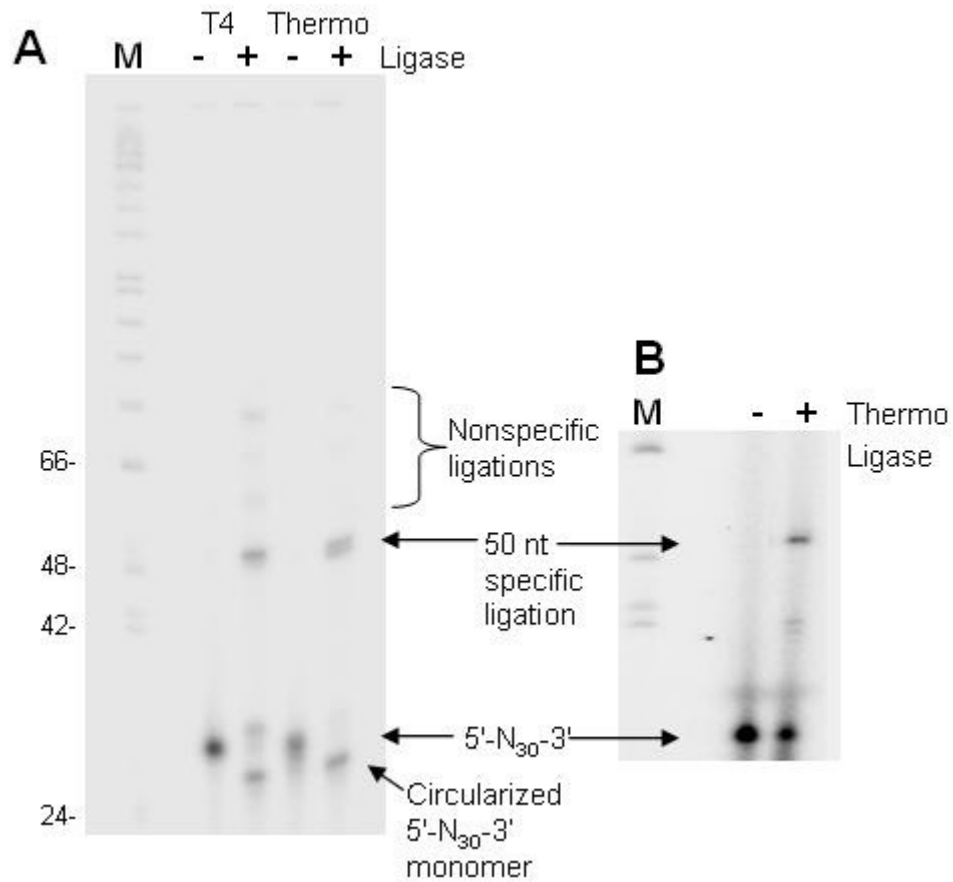
Assuming that the starting pool has reasonable but not complete diversity, it is important that those oligos in the pool that can bind with high affinity to the protein (HIV-RT in this case) are recovered. Since nearly all the sequences in the pool are unique, any sequence not recovered from round 1 is essentially lost. This is not a major issue in subsequent rounds since PCR amplification after round 1 generates many copies of each recovered sequence.

In the protocol used here the first and second rounds of selection were performed using nitrocellulose filters, while subsequent rounds were performed using gel-shift (Figure 2-1). Filter selection was faster, allowing rapid processing of larger amounts of material and high recovery. For these reasons it was advantageous to use this approach, at least for rounds 1 and 2. Later rounds used gel-shift because of preference to this approach due to lower background and higher selectivity. However, many proteins are unable to shift nucleic acids on gels. In such cases it is feasible to continue with filter selection, and we are currently doing this for some other proteins being tested in this protocol (data not shown).

Even with the high recovery efficiency of the nitrocellulose approach, it is impossible to achieve 100% recovery of tight binding sequences. To minimize the potential loss of high affinity sequences, for round 1 three separate selections with 500 pmoles each were conducted then combined. The combined material was reselected in order to decrease the total amount of material for the first ligation step. This was necessary to achieve the highest efficiency in the ligation reaction which required a high ratio of donor (5' phosphorylated oligo 1) to acceptor (selected nucleic acid). Since the



amount of recovered material was still relatively high after the two sequential selections, three separate ligations with Thermostable RNA Ligase were performed (see steps between products A and B in Figure 2-1). An experiment was performed to test the efficiency of the ligation process. A drawback of using T4 RNA Ligase 1 in a DNA selection is its relatively low efficiency at ligating DNA. The efficiency can be improved using hexaminecobalt chloride and polyethylene glycol (168). Even using these optimal conditions, T4 RNA Ligase 1 only ligated ~30% (see Figure 2-2 legend) of the 5'-N<sub>30</sub>-3' DNA to oligo 1 to produce the specific 50 nucleotide ligation product as shown in the experiment in Figure 2-2A. Thermostable RNA Ligase was more efficient ligating about 50% of the material. Thermostable RNA Ligase also produced less nonspecific ligation products than T4 RNA Ligase 1. A second thermostable ligase, CircLigase™, was about as efficient as T4 RNA Ligase 1, though as expected it did circularize the 5'-N<sub>30</sub>-3' DNA with higher efficiency (data not shown). We should note that in order to make detection and quantification of ligated products easier, the ligation tests shown in Figure 2-2A (and Figure 2-5 below) were conducted using 5' end-labeled (and thus 5' phosphorylated) 5'-N<sub>30</sub>-3' and unlabeled oligo 1, while in the ligation reactions used for selection the 5' end of 5'-N<sub>30</sub>-3' was not phosphorylated. Phosphorylating the 5' end of the acceptor increases the likelihood of circularization, which would decrease the level of acceptor available for ligation to oligo 1. This could in turn lead to an underrepresentation of the actual level of ligated product in the selection reactions. Although some circularization did occur (see Figure 2-2A), ligation to oligo 1 was more pronounced, probably because of the 100-fold excess of the oligo 1 donor to the 5'-N<sub>30</sub>-3' acceptor. Also, this approach did not seem to misrepresent ligation efficiency in the normal selection reactions. This was confirmed by



**Figure 2-2: Ligation reactions with T4 RNA Ligase 1 and Thermostable RNA**

**Ligase.** (A) In these reactions a small fraction of the 5'-N<sub>30</sub>-3' starting material was labeled with <sup>32</sup>P at the 5' end (see Methods) in order to follow the ligation process. The starting material was ligated to oligo 1 (with cold phosphate at the 5' end) in Figure 1 (oligo 1:starting material, 10 pm:0.1 pm) using either T4 RNA Ligase 1 (10 units) or Thermostable RNA Ligase (0.5 units) for ~ 8 hours as described under Materials and Methods. Reactions were run on 12 % denaturing polyacrylamide gels and dried gels were visualized using a phosphoimager. These reactions differ slightly from standard SELEX reactions where oligo 1 was radioactively labeled at the 5' end while the 5-N<sub>30</sub>-3' nucleic acid was dephosphorylated in all rounds except round 1, where a small fraction was end-labeled as described above. Positions of the starting material and 50 nucleotide

correctly ligated product (Product B in Figure 2-1) are indicated along with locations or aberrant ligations products which may include circularized products or products with more than one oligo 1 or 5'-N<sub>30</sub>-3'. The location of circularized 5'-N<sub>30</sub>-3' monomers, which occur due to the 5' <sup>32</sup>P labeling (see text), is also indicated. A DNA molecular size marker indicating positions (in nucleotides) of specific size standard is also shown (lane M). Ligation efficiency for the formation of the 50 nucleotide product was 31 ± 8 % (ave. of 2 exp. ± standard deviation) and 51 ± 7 % (ave. of 4 exp.) for T4 RNA Ligase 1 and Thermostable RNA Ligase, respectively. **(B)** Ligation experiment performed with unlabeled 5' dephosphorylated 5'-N<sub>30</sub>-3' and unlabeled 5' phosphorylated oligo 1, after the reaction products were radioactively labeled with T4 PNK using an exchange-reaction (145) and visualized as described above. Ligation efficiency was ~ 50 %.

conducting a ligation reaction with non-phosphorylated 5'-N<sub>30</sub>-3' and oligo 1, then end-labeling the resulting products in an exchange-reaction with PNK (Figure 2-2B). The proportion of correctly ligated (50 nucleotide product) vs. non-ligated material was similar (~50% efficiency) in this experiment in comparison to those with 5' end-labeled 5'-N<sub>30</sub>-3'. Although circularization did not significantly interfere with quantification of ligation using 5'-N<sub>30</sub>-3', it may interfere with oligos of a specific sequence or with longer oligos that could circularize more easily. Recently published primer-free or minimal primer (2 fixed nucleotides at the 3' and 5 ends flanking a random region) approaches that used a more complex ligation step with a “bridging” nucleic acid and T4 DNA ligase report >90% efficiency for the ligation step of the minimal primer approach, while the efficiency in the primer-free approach was not specified (127). A second method termed “primer-free genomic SELEX” reported somewhat lower efficiencies (184). Given the complexities of these various methods and the one used in this manuscript, it is difficult to compare their overall efficiency by comparing results from a single step or even multiple steps in the protocols. It would be interesting to compare the results of these approaches to see if they select similar aptamers.

Although the efficiency of the ligation step results in a significant loss of diversity in round 1, some of this loss is made up by conducting multiple selections and ligations in the first round. As noted above, this issue is less important in subsequent rounds. There is also the possible loss of diversity due to the Thermostable RNA Ligase having a bias for ligating specific sequences or nucleic acids with less structure. T4 RNA ligase has been shown to ligate specific DNA sequences preferentially (63). Although this is also the case with Thermostable RNA Ligase (see below), the higher ligation efficiency

suggests that the bias may be less pronounced. Inhibition of ligation because of high structure is also likely to be less pronounced because the reactions are performed at 60°C.

Following recovery of the ligated 50 nucleotide product (Figure 2-1, product B), a set of reactions was performed to produce a blunt-ended 53 base pair duplex, with a 5' phosphate at the end containing the random region (Figure 2-1, product C). The 3' end of the upper strand in product C is derived from Oligo 1, however, the 3' dideoxycytidine residue is replaced by a deoxycytidine and is followed by a 5'-CAG-3' sequence. This depiction was based on the assumption that the 3'-5' exonuclease activity of the Klenow polymerase would excise the dideoxy residue and add four template-directed nucleotides. Oligo 2, which was used to prime the step with Klenow, was not complementary to the first three 5' nucleotides of oligo 1. This is due to a previous version of this oligo with complete complementarity produced primer dimers in the subsequent PCRs. The 5' phosphate at the random end of product C was added by PNK before the Klenow step. It directed ligation by T4 DNA ligase of the oligo 3-oligo 4 duplex to this end. The oligo 3-oligo 4 duplex was staggered in order to force blunt end ligation in the correct orientation (Figure 2-1, product D). The blunt-end ligation step using the rapid ligation kit and a vast excess of the oligo 3-oligo 4 duplex typically showed a ligation efficiency of >60% (data not shown).

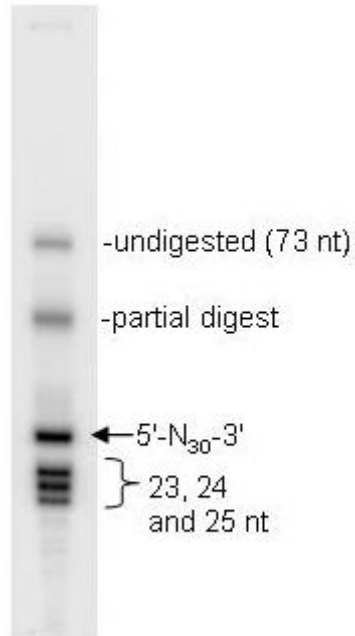
After producing product D, an initial PCR step was performed for 15-18 cycles (see Methods). Material from this step was isolated on a native gel. About 1/40<sup>th</sup> of the total recovered material was used in a second PCR step that was performed in 4 separate reactions using a range of cycles (see Methods). Amplification steps must always be performed carefully in SELEX protocols, as too many cycles produce products that

cannot form a complete duplex, due to the random region while too few cycles produce a low yield. This is easily followed on a native gel, where the desired product appears and builds up over a range of cycles then is lost as too many cycles are performed. Although the approach used here was effective, in many cases a single PCR step with multiple reactions conducted for 15-18 cycles (depending on the level of the starting material) would produce sufficient yield for the next round.

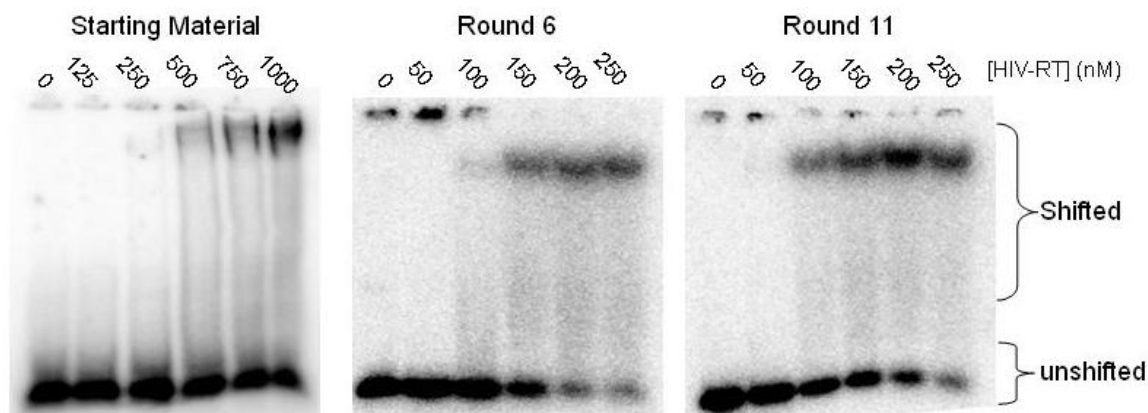
Cleavage of the PCR products to reproduce the 5'-N<sub>30</sub>-3' product (Figure 2-1, product E) was highly efficient (Figure 2-3). Note the system can be designed to select single stranded products in a wide range of sizes, provided that cleavage of the PCR product yields a product of unique size for the selected material. This may require using primers in the PCR step, or linkers that are different sizes than the ones employed here.

### *2.3.2 The analysis of the SELEX procedure by monitoring binding to HIV-RT, and the nature of the recovered aptamers*

The 30 nt material (Figure 2-1, product E) from select rounds was analyzed in a gel-shift assay to estimate binding affinity for HIV-RT (Figure 3-4). Results were compared to the starting material and  $K_d$  values were estimated. Note that this approach does not yield a true  $K_d$  value, as gel-shift is not an equilibrium binding method. However, the derived values are useful in comparing the binding affinity of material from the various rounds. The starting material bound weakly to HIV-RT with  $K_d$  value of  $\sim 1$   $\mu\text{M}$ . Attempts to measure the value more precisely were not performed as large amounts of enzyme would be required, and this resulted in a smearing effect in the gel-shift. The  $K_d$  values for 30 nucleotide material selected in rounds 2, 6, 10 and 11 were  $\sim 1$   $\mu\text{M}$ ,



**Figure 2-3: Cleavage of PCR product with BbsI and MnlI to regenerate 30 nucleotide selected material.** A cleavage reaction run on a 12% denaturing polyacrylamide gel is shown for cleavage of the 73 nucleotide double stranded PCR product (equivalent to Product D in Fig. 1 minus the 5' overhang) with BbsI and MnlI. Correct cleavage would produce 23, 24, 25 and 30 nucleotide products on a denaturing gel with the last being the selected material. Some undigested and partially digested material is also evident in this gel and was typically observed in experiments.



**Figure 2-4: Gel-shift analysis of starting material and material from rounds 6, 10, and 11 of primer-free SELEX.** An example of a gel-shift experiment run on a native 6% polyacrylamide gel is shown. Increasing amounts of HIV-RT (as indicated above wells, note higher amounts used for starting material) were mixed with 30 nucleotide 5'-N<sub>30</sub>-3' derived from the starting material or rounds 6, 10, and 11 as described under Methods. Starting material was 5' end-labeled with <sup>32</sup>P, while rounds 6, 10, and 11 were internally labeled during the SELEX procedure (see Methods). Positions of shifted and unshifted material are indicated. Refer to the Results and Discussion section for quantification of experiments.



170 ± 56 nM, 191 ± 42 nM, and 142 ± 31 nM (rounds 6, 10, and 11 are an average of 3 exp. ± standard deviation), respectively. Interestingly, the affinity did not increase appreciably after round 6, despite five more rounds of selection at higher salt concentrations (see Methods). A second independent SELEX procedure was also performed. The  $K_d$  value after 9 rounds was similar to the value for the round 11 material in the first SELEX and the sequences recovered were nearly identical (see below).

### *2.3.3 Aptamer sequences recovered from round 11 of the first SELEX and round 9 of the second were nearly identical and had little secondary structure*

Ten clones were sequences from the round 11 material. Only two different sequences (denoted R11-1, or PF1 in subsequent chapters and R11-2 in Table 2-1) were recovered and both were closely related, differing only by the first nucleotide. These same sequences were also the predominant oligos found in material sequenced from round 6, which explains the similar affinity of material from these rounds for HIV-RT. Gel-shift assays with each sequence revealed that they had similar affinity for HIV-RT. Interestingly, using the DNA mfold program (197), both sequences had a very low level of predicted secondary structure with the most stable predicted structures having  $\Delta G$  values of -0.61 kcal/mole for both sequences. A possible explanation for this is that Thermostable RNA Ligase preferentially ligates material with low structure (see below). This would seem to make sense as availability of the 3' end of the nucleic acid would be pivotal in the ligation step. However, this factor should be mitigated to some extent by the 60°C reaction temperature. Also notable was the regular arrangement of 4 separate G dinucleotides in the sequences, which suggests the possible formation of a higher ordered

structure. Several sequences closely related to R11-1 (PF1) and R11-2 were also recovered from round 9 of the second SELEX experiment (Table 2-1). Experiments were conducted to determine why the selected aptamers bind tightly to HIV-RT (Chapter 3).

#### *2.3.4 The ligation efficiency of Thermostable RNA Ligase improves with low structure and a T residue at the 3' terminus of the acceptor*

The significant increase in binding affinity of the round 11 selected material to HIV-RT indicated that the RT binding step likely played a major role in the selection process for this protocol. However, this does not rule out the possibility that Thermostable RNA Ligase may have also influenced selection. To test this further, the ligation efficiencies of the starting material and material selected from rounds 2, 6 and 11 were tested (data not shown). Starting material (see above) and material from round 2 each showed ~ 50% ligation efficiency in standard 8 hour ligation reactions, while efficiency was greater than 90% for round 6 and 11 materials.

The more efficient ligation of the selected material suggests that Thermostable RNA Ligase may have played some role in the selection process. However, it is also possible that material that binds tightly to HIV-RT just happens to ligate efficiently because of other properties. For example, low structure may favor binding to RT by allowing the aptamer to be flexible, while it also promotes higher ligation efficiency. Also, it was notable that all the aptamers recovered from round 11 (and round 6 (data not shown)) terminated with a T residue at the 3' end. Several oligonucleotides were designed to try and determine why the selected material ligated more efficiently. These included homopolymeric 30-mers composed of A (5'-dA<sub>30</sub>-3') or T (5'-T<sub>30</sub>-3') which presumably

Name <sup>1</sup>	Sequence <sup>2</sup>	# of times recovered in round	K <sub>d</sub> for binding HIV-RT (nM) <sup>3</sup>
Starting material	5'-N <sub>30</sub> -3'	NA	~1000
SELEX Experiment 1			
R11-1 (PF1)	5'-AGGAAGGCTTTAGGTCTGAGATCTCGGAAT-3'	9/10	82 ± 7
R11-2	5'- <u>GG</u> GAAGGCTTTAGGTCTGAGATCTCGGAAT-3'	1/10	102 ± 15
SELEX Experiment 2			
E2R10-1	5'-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGGAT-3'	1/10	ND
E2R10-2	5'-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGTAT-3'	2/10	ND
E2R10-3	5'-AGGAAGGCTTTAGGTCTGAGATCTCGGAAT-3'	3/10	ND
E2R10-4	5'-AGGAAGGCTT <u>C</u> AGGTCTGAGAT <u>TT</u> CGGAAT-3'	1/10	ND
E2R10-5	5'-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGT <u>G</u> T-3'	1/10	ND
E2R10-6	5'-AGGAAGGCTTTAGGTCTGAGATCTCGGTAT-3'	1/10	ND
E2R10-7	5'-AGGAAGGCTTTAGGTCTGAGAT <u>TT</u> CGGTAT-3'	1/10	ND
<sup>1</sup> For experiment 1 names indicate the round of selection from which the sequence was recovered followed by a designated number. For experiment 2, the name is preceded by an "E2" designation. <sup>2</sup> Underlined nucleotides in the recovered sequences indicate differences from the R11-1 sequence recovered in the first SELEX experiment. <sup>3</sup> Equilibrium dissociation constants (K <sub>d</sub> ) were determined using a gel shift assay as described under Material and Methods. Results for R11-1 (PF1) and R11-2 are an average of 3 experiments ± standard deviation. The result for "Starting material" was an estimate based on repeated experiments. ND- Not determined; NA-Not applicable.			

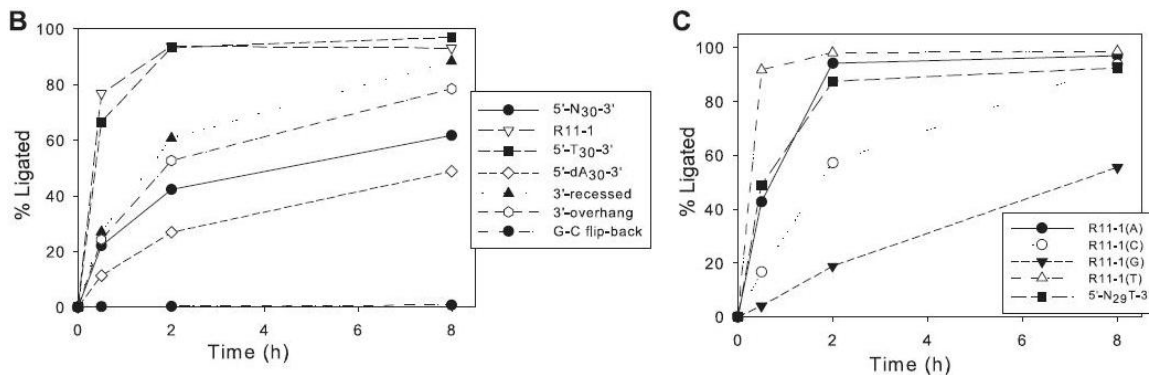
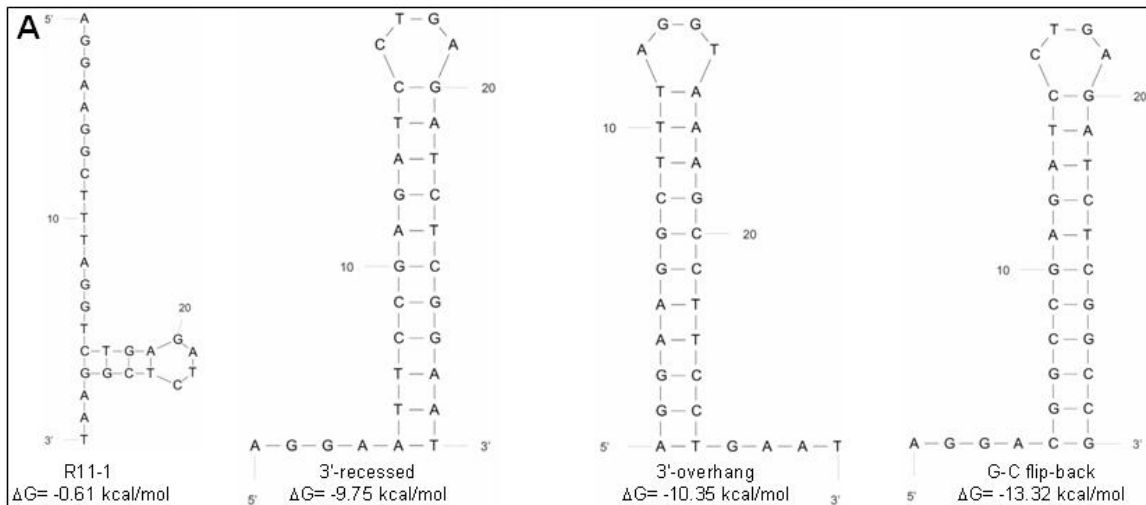
**Table 2-1: Sequences recovered from SELEX experiments and K<sub>d</sub> for binding HIV-RT**

contain little or no secondary structure (Figure 2-5A). A set of 30-mer oligos had sequences related to the round 11 selected material but with the 3' terminus either protruding or bound in a stem structure (3'-overhang and 3'-recessed, respectively); and finally, a highly structured sequence 30 nt G-C rich sequence (G-C flip-back). Ligation reactions for these oligos and the R11-1 (PF1) aptamer and 5'-N<sub>30</sub>-3' starting material were performed for 30 min, 2 hr and 8 hr and plotted as shown in Figure 2-5B. Both R11-1 (PF1) and 5'-T<sub>30</sub>-3' ligated with very high efficiency, reaching ~70% ligated product by 30 min and over 90% by 2 hours. The ligation advantage was not due solely to a lack of structure, as 5'-dA<sub>30</sub>-3' ligated with much lower efficiency that was comparable to 5'-N<sub>30</sub>-3'. Interestingly, the 3'-overhang and 3'-recessed oligos ligated better than starting material or 5'-dA<sub>30</sub>-3' and were comparable, despite the latter essentially “protecting” the 3' end in a secondary structure. Finally, the highly structured G-C flip-back ligated very poorly, although this may have been at least partially due to the 3' G residue (see below).

To test the effect of a T residue at the 3' end of the acceptor oligo, three derivatives of R11-1 (PF1) in which the 3' terminal nucleotide was replaced with A, G, or C were also tested in ligation reactions along with R11-1 (PF1) and a version of the random starting material with a T at the 3' end (5'-N<sub>29</sub>T-3') (Figure 2-5C). While oligos terminating in T, A, or C all showed greater than 90% ligation efficiency in 8 hour reactions, the T terminated oligo ligated far more efficiently at early time points followed by A then C in ligation efficiency. The version of R11-1 (PF1) with a 3' G residue ligated much more poorly than the others. The importance of a T residue at the 3' end is emphasized most strongly by ligation experiments with 5'-N<sub>29</sub>T-3' (Figure 2-5C, filled squares). This oligo differed from the starting material (5'-N<sub>30</sub>-3') by just the addition of a

single 3' T residue. Ligation with 5'-N<sub>29</sub>T-3' essentially mimicked that with R11-1(A) and reached a level of ~90% after 8 hours while 5'-N<sub>30</sub>-3' was only between 50-60% (Figure 2-5B, filled circles) at this time point. This finding suggests that it may be advantageous to start the selection process with random oligos that include a fixed T residue at the 3' end. This would decrease the loss of diversity in the early rounds of the protocol, though it would also create some bias in the results.

The above results indicate that Thermostable RNA Ligase showed a clear preference for a T residue at the 3' end of the acceptor, at least for ligation to the oligo 1 donor that was used here. Low structure was also advantageous, as modestly structured oligos closely related to R11-1 (3'-recessed and 3'-overhang) ligated with lower efficiency. Although strong binding to HIV-RT clearly played an important role in the selection process, it is not clear if, and to what extent Thermostable RNA Ligase sequence/structural preferences influenced the process. Selections using this method with other proteins are ongoing and should help in answering this question. The fact that only sequences with a 3' terminal T were recovered suggests that the ligase played at least some role, as deletion or replacement of the T does not affect binding to RT (data not shown). If this method preferentially selects low structure aptamers then it may be particularly useful for finding aptamers that could potentially be used as inhibitors in cells, as highly structured nucleic acids have a tendency to activate innate cellular responses (70).



**Figure 2-5: Ligation of various 30 nucleotide oligos to oligo 1 with Thermostable RNA Ligase.** Ligations were carried out for 0.5, 2, or 8 hours using conditions described in Figure 2-2A. Products were visualized as described in Figure 2-2 and quantified using a phosphoimager. **(A)** Predicted structures (using DNA mfold (197)) of the 30 nt oligos ligated to oligo 1 (see Figure 2-1), panel B. Only those oligos with a predicted secondary structure are shown. The predicted  $\Delta G$  value in kcal/mol using the conditions of 80 mM NaCl, 6 mM  $MgCl_2$  and  $37^\circ C$  is also listed for each structure. **(B)** Plot of the percent of 30 nt oligo ligated to oligo 1 vs. time. **(C)** Ligation plots for R11-1 (terminal 3' T) or R11-1 in which the 3' terminal T was changed to A, C, or G, and 5'-N<sub>29</sub>T-3'. All experiments were repeated at least once and similar results were obtained.

## **2.4 Conclusions and significance**

Just a few protocols for primer-free SELEX are currently described in the literature (see Introduction). Primer-free SELEX is advantageous because it allows selection of nucleic acid aptamers from random pools without interference from fixed flanking regions. Though all these methods are considerably more complex than SELEX approaches using flanking fixed sequences, the method proposed in this report takes advantage of a new Thermostable RNA ligase. In comparison to other primer-free methods, this simplifies the procedure for regenerating nucleic acids for subsequent rounds of selection. In practice each primer-free method may have specific advantages or biases and it would be interesting to compare the aptamers selected from the various methods using the same target protein.

## Chapter 3: Comparative study of HIV-RT aptamers from a primer-free SELEX approach with various classes of RT aptamers

### 3.1 Introduction

Since its invention by three separate groups in the early 1990s (49, 90, 172), Systematic evolution of ligands by exponential enrichment (SELEX) has been used for selecting small nucleic acids (RNA or DNA) which bind to specific targets with high affinity. SELEX starts with a large random pool of nucleic acid incubated with a limiting amount of target. These specific targets are mostly proteins, but they also include nucleic acids, cultured cells, viruses and other small molecules (27, 67, 68, 103, 113, 133, 163, 183, 184, 186, 191, 194). Nucleic acids that bind to the target with higher affinity will preferentially associate to form a complex. This complex can be isolated by gel shift, nitrocellulose filter binding, or capillary electrophoresis (118). The selected pool is expanded by PCR amplification, and the new, enriched pool of nucleic acids is subjected to subsequent rounds of target binding. After several rounds, nucleic acids with high affinity for the target called “aptamers” are isolated. Aptamers have many potential uses in diagnostic and therapeutic applications, such as replacements of antibodies in biochemical assays (e.g. ELISA), utilization as biosensors, as a tool for studying the molecular biology of virus replication, and development of antiviral drugs (24, 40, 56, 82, 88, 117, 121, 136, 191). Essentially, any technique that involves sensitive and stringent recognition of protein or nucleic acids could benefit from aptamer-based approaches.

Most aptamers are screened through standard SELEX procedures, which include a starting pool of random nucleic acids. These random oligonucleotides contain fixed



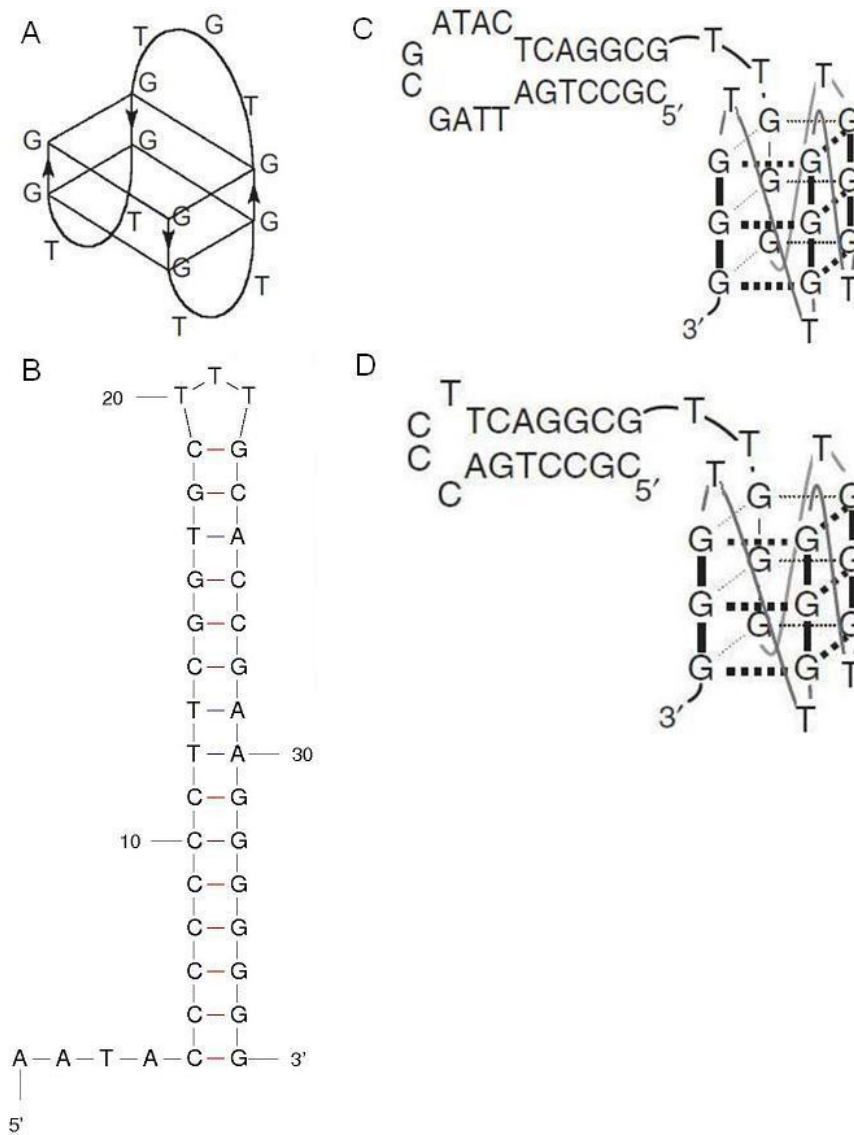
flanking regions of known sequence at the 5' and 3' termini. After selection, the selected aptamers often retain sequences from the fixed flanking regions which are required to achieve tight binding. Typically, at the end of the selection process, these sequences are modified to determine the minimal bases that are required for tight binding. This approach poses a disadvantage, where the fixed sequences create biases on the selection by interacting with nucleotides in the random region of the selected oligonucleotides. Recently, a new “primer-free” SELEX method was developed, where selection occurs in the absence of any fixed, flanking nucleotides (125, 127, 184), and a second “minimal primer” SELEX, which the fixed sequences are reduced to only a few nucleotides (83, 177). By eliminating most of the fixed flanking regions in minimal primer SELEX, or by removing them entirely in primer-free SELEX during the selection step, these methods reduce the biases from the fixed regions. A notable drawback of primer-free approaches is that they are more time consuming and include several additional steps. There are also steps where bias may be introduced.

Our lab has introduced a novel primer-free SELEX-based approach for selecting single stranded DNA sequences (102). Sequences that bound ~10-fold tighter than the starting random sequences to HIV-RT in gel-shift assays were isolated. All the recovered sequences had essentially the same sequence, and were characterized by the regular arrangement of four separate guanosine doublets in the sequences (5'-AGGAAGGCTTTAGGTCTGAGATCTCGGAAT-3', denoted PF1 in this thesis) and a predicted secondary structure ( $\Delta G = -0.61$  kcal/mole, predicted using DNA mfold program (197)). The repeated guanosine doublets were also observed in a 15 nt long aptamer that binds tightly to human thrombin (denoted as Thrombin Aptamer, or TBA,

Figure 3-1A) (21, 108). The TBA (5'-GGTTGGTGTGGTTGG-3') folds into an anti-parallel G-quadruplex, with two G-quadruplex layers connected by a TGT and two TT loops. The HIV nucleocapsid protein has been shown to destabilize the G-quadruplex formation of TBA (92). G-quadruplexes are formed by four guanine residues associated with one another by Hoogsteen bonds in a square planar configuration. Most monovalent cations stabilize G-quadruplex formation by the cation binding specifically to guanine O<sub>6</sub> carbonyl groups between the planes of the G-quartets. The stability is highly dependent on the cation size, with K<sup>+</sup> or Sr<sup>2+</sup> having the most favorable ionic radius at 1.3-1.4 Å, while Na<sup>+</sup> and Li<sup>+</sup> destabilize G-quadruplex formation (156, 169). These structures can assume both parallel and anti-parallel configurations and can form between 2 or more separate nucleic acid strands (intermolecular), or within the same strand (intramolecular). G-quadruplexes have been implicated in several cellular processes including telomerase regulation and gene expression (58, 66).

Several other aptamers to various HIV proteins also are predicted to form G-quadruplexes. These include aptamers that block HIV envelope glycoprotein gp120 by binding to the gp120 cationic V3 loop region (71). Various classes of G-quadruplex aptamers were found to inhibit RNase H activity (7, 39), and integrase activity (95). In previous studies, Schneider *et al.* and Michalowski *et al.* demonstrated a class of HIV-RT aptamer with a bimodular structure of a three-tier G-quadruplex and a stem-loop helical element (112, 149) (Figure 3-1C and D). In addition to HIV proteins, G-quadruplex aptamers to several other viral and non-viral proteins have been selected suggesting that G-quadruplexes are a common motif for high affinity protein-nucleic interactions (157).

In this report, we characterized the interaction between HIV-RT and the PF1



**Figure 3-1: Secondary structures of various classes of aptamers.** (A) Secondary structure of Thrombin Aptamer (TBA), model adapted from Kankia *et al.*, Nucleic Acids Res. 2005 (92). (B) Predicted structure of 37 NT SELEX from DNA mfold program. The run of 7Gs at the 5' end of the aptamer mimics the polypurine tract (PPT), and prevent the formation of other secondary structures, except for the flip-back hairpin loop (43). (C and D) Secondary structure of the G-quadruplex aptamers named “R1T” (C) and “S4” (D), respectively. Figures adapted from Michalowski *et al.*, Nucleic Acids Res. 2008 (112), Intraquadruplex loops are shown as curved lines.

single-stranded DNA derived from primer-free SELEX and denoted above. The central diguanosines were crucial for tight binding, and shortening the sequence by as few as three nucleotides significantly reduced binding to RT. We also compared the effects of several different classes HIV-RT DNA aptamers on RT inhibition. Besides PF1, two previously characterized aptamers were also studied: (1) a primer-template flipback aptamer developed in our lab that mimics the polypurine tract at the 3' end of the oligo (42, 43) (Figure 3-1B); (2) a G-quadruplex aptamer reported by Michalowski *et al.* (112). PF1 was an effective inhibitor of HIV-RT, however, G-quadruplex aptamer and primer-template aptamer, both were larger than PF1, were more potent inhibitors. Interestingly, based on circular dichroism (CD) analysis, PF1 did not form G-quadruplexes.

## **3.2 Materials and Methods**

### *3.2.1 Materials*

Human immunodeficiency virus reverse transcriptase (HIV-RT) was obtained from Worthington Biochemical Corporation. *Taq* polymerase, restriction enzymes and T4 polynucleotide kinase (PNK) were obtained from New England Biolabs. Deoxyribonucleotide triphosphates (dNTPs) were from Roche Applied Sciences. Klenow DNA polymerase and dideoxyguanosine triphosphate were from Affymetrix, Inc. Radiolabeled compounds were obtained from Perkin Elmer. Sephadex G-25 spin columns were from Harvard Apparatus. All oligonucleotides were commercially synthesized from Integrated DNA Technologies. All other chemicals were from Sigma-Aldrich Co., Thermo Fisher Scientific, Inc., or VWR Scientific, Inc.

### 3.2.2 Methods

*5' End-labeling of oligonucleotides with T4 PNK:* Reactions for labeling various oligos were done in a 50  $\mu$ l volume containing 25 pmole of the oligo, 70 mM Tris-HCl (pH=7.6), 10 mM MgCl<sub>2</sub>, 5 mM dithiothreitol (DTT), 5  $\mu$ l of  $\gamma$ -<sup>32</sup>P ATP (3000 Ci/mmol, 10  $\mu$ Ci/ $\mu$ l) and 2  $\mu$ l (20 units) of T4 polynucleotide kinase. The reaction mixture was incubated for 30 minutes at 37°C, and then the PNK was heat inactivated for 10 minutes at 70°C according to manufacturer's recommendation. The material was then run through a Sephadex G-25 spin column.

*Determination of equilibrium dissociation constants ( $K_d$ ) with 5' end-labeled substrates:* PF1 (~2 nM) from the primer-free SELEX or other designed sequences (Table 1) were either labeled internally with  $\alpha$ -<sup>32</sup>P dTTP, or 5' end labeled with  $\gamma$ -<sup>32</sup>P ATP and mixed with various amounts of HIV-RT in 10  $\mu$ l of buffer containing 50 mM Tris-HCl (pH=8), 1 mM DTT, 80 mM KCl, and 6 mM MgCl<sub>2</sub>, for 10 min at room temperature. Two microliters of 6X native gel loading buffer (40% sucrose, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol) was added and the samples were run on a 6% native polyacrylamide gel (29:1 acrylamide:bisacrylamide) at ~100 V until the blue dye marker was ~2-3 inches below the wells. The amount of shifted vs. non-shifted material was quantified using a Fujifilm FLA-5100 phosphoimager. Values for  $K_d$  were determined by plotting the concentration of shifted product (nM) vs. the concentration of HIV-RT and fitting the data by nonlinear least square fit to the quadratic equation:  $[ED] = 0.5([E]_t + [D]_t + K_d) - 0.5(([E]_t + [D]_t + K_d)^2 - 4[E]_t[D]_t)^{1/2}$ , where  $[E]_t$  is the total enzyme concentration and  $[D]_t$  is the total primer-template concentration (72). In cases

where the affinity of HIV-RT for the nucleic acid was low ( $K_d \geq 500$  nM), a more precise  $K_d$  value was not determined.

*Preparation of substrate for RT inhibition:* Sixty pmoles of 50 nt template (5'-TTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATC-3') and 50 pmoles of 33 nt 5'  $^{32}\text{P}$  end-labeled primer (5'-TTCCCCGGGTACCCGAGCTCGAATTCGCCCTATAG-3') were mixed in 20  $\mu\text{l}$  containing 50 mM Tris-HCl (pH=8), 1 mM DTT, and 80 mM KCl. The mixture was heated to 80°C for 2 minutes, then cooled at a rate of 1°C per minute to 30°C. This material was used directly in the assays.

*Preparation of ddG-terminated loop-back substrates:* Fifty pmoles of loop-back DNA phosphorylated at the 5' end was incubated with 5 units of Klenow DNA polymerase in 50  $\mu\text{l}$  of buffer containing 50 mM Tris-HCl (pH=8), 1 mM DTT, 50 mM KCl, 6 mM  $\text{MgCl}_2$ , and 25  $\mu\text{M}$  ddGTP for 30 minutes at 37°C. The material was extracted with phenol-chloroform and precipitated with ethanol. The precipitated material was run through a Sephadex G-25 spin column to remove any remaining unused ddGTP. The concentration of the recovered material was determined by comigration with the original substrate and was not extendable by HIV-RT in the presence of dNTPs (determined by 5' end-labeling a portion of the material at low specific activity with  $\gamma$ - $^{32}\text{P}$  ATP), indicating that it was 3' terminated with ddG.

*RT inhibition assay:* Reactions contained substrate (1:1.2 primer:template, final concentration in reactions was 50 nM in 5'  $^{32}\text{P}$  end-labeled primer) in 30  $\mu\text{l}$  of buffer

containing 50 mM Tris-HCl (pH=8), 1 mM DTT, 80 mM KCl, 6 mM MgCl<sub>2</sub>, and 0.1 µg/µl BSA. Various amounts (0.1, 0.5, 1.0 or 5.0 nM of inhibitors for 37 NT Flipback (terminated with ddG), R1T and S4, and 25, 50, 100 and 150 nM of inhibitors for PF1 were used. A 5 µl of a supplement containing 800 µM dNTPs (100 µM final) in the above buffer were added to the reactions. The mixture was placed at 37°C for 2 minutes. Primer extension was initiated by adding 5 µl of HIV-RT (0.25 nM final concentration in reactions). Five microliter aliquots were removed at 2, 5, 10, 15, and 20 minutes and add to 5 µl of 2X formamide gel loading buffer (90% formamide, 20mM EDTA (pH=8), 0.25% xylene cyanol, 0.25% bromophenol blue). Samples were run on a 10% denaturing gel as described below and quantified using a Fuji FLA-5100 phosphoimager as described above. A graph of the concentration of number of counts (photo-stimulated luminescence (PSL) radiation units) vs. time was plotted. The experiment was repeated with similar results. The RT inhibition was achieved by a decrease in the concentration of available enzyme, being taking up by the aptamer. The half-maximal inhibitory values (IC<sub>50</sub>) were calculated by the relationship between the aptamer concentration and percent of inhibition achieved at the particular inhibitor concentration. The concentrations used for measured were stated as above, at the 10 min timepoint. A four-parameter logistic equation was used for curve-fitting with Sigma Plot software to obtain IC<sub>50</sub> for the various aptamers:  $Y = \min + \frac{(\max - \min)}{[1 + 10^{(\log EC_{50} - x) \text{Hillslope}}]}$ , where Y is the percent of inhibition at a certain inhibitor concentration, and X is the log of the inhibitor concentration. Error terms for reported IC<sub>50</sub> values were standard deviations among triplicate assays.

*Circular dichroism spectroscopy:* The DNA aptamer samples (see Table 2) were heated to 90°C for 2 minutes, then cooled at a rate of 2°C per minute to 30°C. The aptamer concentration at 4 μM was adjusted in the corresponding potassium or sodium buffers (Potassium buffer final concentration: 50mM Tris–HCl (pH= 8), 80mM KCl, 6mM MgCl<sub>2</sub>, 1mM DTT; sodium buffer, final concentration: 50mM Tris–HCl (pH=8), 80mM KCl, 6mM MgCl<sub>2</sub>, 1mM DTT). Near-UV circular dichroism (CD) spectra were acquired at 25°C using a Jasco J-810 spectropolarimeter (Easton, MD) at 1.0 nm intervals between 200 nm and 300 nm in a 1.0 mm quartz cuvette, at a scan speed of 50 nm/min, response time at 8 second, and scan sensitivity at 100 mdegrees.

### **3.3 Results**

#### *3.3.1 Determination of equilibrium dissociation constant ( $K_d$ ) values for PF1 and the SELEX starting material*

In the previous study,  $K_d$  values were determined for the PF1 aptamer derived from the primer-free SELEX protocol (102) (Chapter 2). Gel-shift analysis was used for these determinations as described in Materials and Methods. Since gel-shifts do not necessarily reflect equilibrium binding, the determined  $K_d$  values are qualitative in nature and best used for comparison between the various tested constructs. In the current experiments, PF1 had a  $K_d$  value of  $82 \pm 7$  nM (Table 3-1, Figure 3-2B). In contrast, the random sequence material used to start the SELEX procedure (5'-N<sub>30</sub>-3') bound with much lower affinity ( $K_d > 1$  μM).



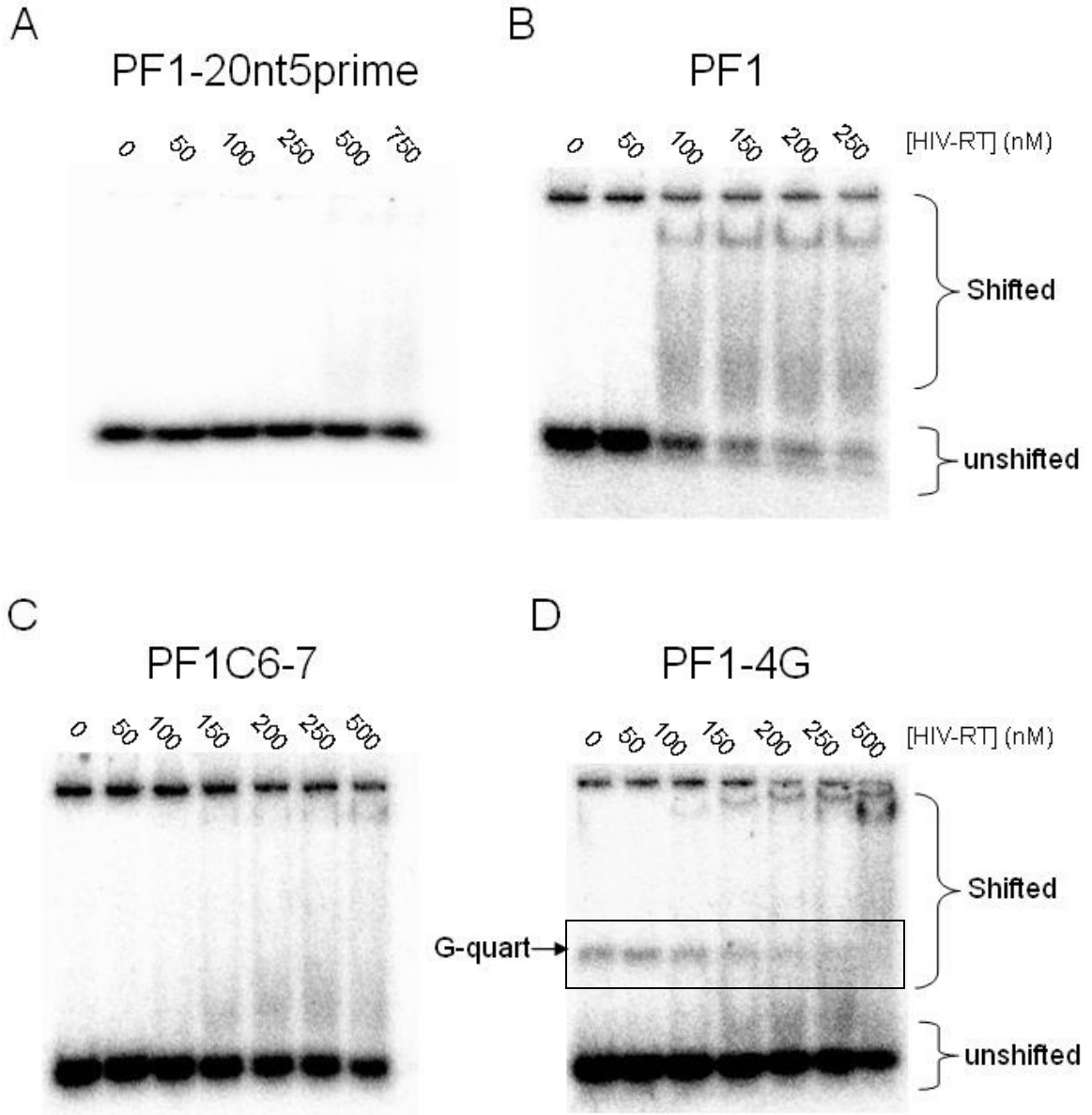
### 3.3.2 *The middle guanosine doublets were required to maintain high affinity to HIV-RT*

PF1 had a striking arrangement of four diguanosine repeats. To test the role of these repeats in binding to HIV-RT, several variations of this motif were made. Table 3-1 summarizes the modifications and lists the determined  $K_d$  values for binding to HIV-RT. Changing the four guanosine doublets to cytidine (PF1-GG>CC) or adenosine (PF1-GG>AA) increased the  $K_d$  values to > 500 nM, nearly equal to the starting material (5'-N<sub>30</sub>-3'). This demonstrated the importance of the diguanosine repeats for tight binding to RT. The effects of minor changes in the diguanosine runs were then tested. When the guanosine doublets at 5' and 3' termini were changed to dicytidine (PF1-terminal GG>CC) there was a small increase in the  $K_d$  (144 ±10 nM). In contrast, changing the two central diguanosines to cytidines led to a dramatic increase in the  $K_d$  values to near the value for 5'-N<sub>30</sub>-3'. Changing even one of the two central diguanosines to Cs, or even replacing one G nt with a C in one of the central doublets resulted in a significant loss in binding affinity (PF1C6-7, Figure 3-2C; PF1-C6 and PF1-C14). These results indicate the two central G doublets are pivotal for tight binding to RT while the terminal doublets play lesser roles. Interestingly, adding one or two extra guanosine(s) to the central doublet closer to the 3' end also increased the  $K_d$  by 2 to 5-fold (PF1-3G, PF1-3G2 (extra G added to 3' central and 3' terminal diguanosines) and PF1-4G in Table 3-1). The PF1-4G sequence showed G-quartet formation in gel-shift assay as was evidenced from a portion of the construct shifting upward on the gel (Figure 3-2D). This was presumably due to the formation of interstrand G-quartets. Finally, a construct that contained the two central diguanosine motifs but was truncated at the 5' and 3' ends such that the final length was only 15 nts, completely lost the strong affinity to RT (PF1-15ntmidG). This



PF1-18ntTlinkG	5'-AGGTTGGTTGGTTGGAAT-3'	>750
PF1-RNA	5'-AGGAAGGCUUUAGGUCUGAGAUCUCGGAAU-3'	>750
Structural controls		
PF1A	5'-AGGAAGGC AAAAGG AAAAAAAAAAAGGAAT-3'	> 750
PF1Flip	5'-AGGAAGGCTTTAGGTTTCCTAAAGCCTTCC-3'	667 ± 144
PF1Loop	5'-AGGAAGGCTTTAGGTCCGAGATCTCGGAAT-3'	558 ± 63
PF1N	5'-NGGNNGGNNNNNGGNNNNNNNNNNNGGNNN-3'	~ 500
PF1T	5'-TGGTTGGTTTTTGGTTTTTTTTTTGGTTT-3'	448 ± 25
<p><sup>1</sup>For experiment 1 names indicate the round of selection from which the sequence was recovered followed by a designated number.</p> <p><sup>2</sup>The repeated diguanosines were denoted in blue for PF1 and its variants. The modifications on the constructs were denoted in red. The RNA analog (PF1-RNA) of PF1 was also showed in the “Truncated constructs” section.</p> <p><sup>3</sup>Equilibrium dissociation constants (<math>K_d</math>) were determined using a gel shift assay as described under Material and Methods. Results for PF1 are an average of 3 experiments ± standard deviation. The result for the starting material (5'-N<sub>30</sub>-3') and the constructs were estimates based on 3 repeated experiments.</p>		

**Table 3-1:  $K_d$  values of HIV-RT binding for PF1 and various constructs.**



**Figure 3-2: Gel-shift analysis of the aptamer PF1 and its various constructs.**

Increasing amounts of HIV-RT (as indicated above wells) were mixed with 5' <sup>32</sup>P end-labeled PF1 (Figure 3-2B) or the constructs as described in Methods. Samples were run on a native 6% polyacrylamide gel. Positions of shifted and unshifted material are indicated. **Figure 3-2D:** G-quartets (G-quart, boxed lanes) were observed migrating above the unshifted DNA in the PF1-4G construct

indicated that the central diguanosines alone cannot confer high affinity binding to RT.

### *3.3.3 Truncations of the aptamer abolished its tight binding to HIV-RT*

The length of the selected aptamer was also tested as a potential factor for tight binding. Table 3-1 showed the various oligos that were designed by progressively eliminating nucleotides from the termini of the 30 nt PF1. As few as a three nucleotide deletion at either the 5' end increased the  $K_d$  approximately 6-fold. Deleting ten to fifteen nucleotides resulted in an even greater loss of affinity to essentially the level of 5'-N<sub>30</sub>-3' (Figure 3-2A). Similarly, when the nucleotides in between the guanosine doublets were eliminated (PF1- $\Delta$ (8-12), PF1- $\Delta$ (16-25), PF1- $\Delta$ (8-12/16-25), and PF1-18ntTlinkG (closely resembles TBA aptamer), the shortened oligo lost its ability to bind RT tightly. All of these constructs showed  $K_d$  comparable to 5'-N<sub>30</sub>-3'. These results indicate that PF1's high affinity for RT is highly dependent on the length of the oligo, as the binding of RT significantly decreases even when three nucleotides were trimmed off. However, it is important to point out that the truncations discussed above eliminate at least one of the 4 diguanosine motifs. Therefore they do not completely separate the role of the diguanosine repeats vs. nt length in achieving strong binding to HIV-RT. To test this further, a construct 1 and 3 nts eliminated from the 5' and 3' termini, respectively, was tested (PF1-26ntG). This construct retains all 4 diguanosines but was shortened by 4 nts. This resulted in about 3-fold increase in the  $K_d$  in comparison to PF1. The role of the terminal diguanosines was further evaluated by changing the diguanosine in PF1-26ntG to either dicytidines (PF1-26ntC) or dithymidines (PF1-26ntT). While dicytidine had a negative effect on binding in comparison to diguanosine, dithymidine had a modestly

positive effect. Overall, these results suggest a complex interaction between RT and the constructs and a critical role for the central diguanosines. Beyond this, length and nt sequence at the termini play some role in high affinity binding but are less vital.

### *3.3.4 An RNA version of PF1 does not confer tight binding to HIV-RT*

An RNA analog of the 30 nt PF1 aptamer (PF1-RNA) was tested for RT binding. PF1-RNA did not bind with high affinity to RT indicating that high affinity interaction was specific to DNA (Table 3-1).

### *3.3.5 The diguanosine repeats require the specific sequences in PF1 to retain HIV-RT tight binding*

The structural component of the sequences between the diguanosine repeats in PF1 was also tested. Table 3-1 showed the various constructs designed with a low structure between the diguanosine repeats, where the sequence is entirely made of As and Ts (PF1A, PF1T) to prevent any formation of a higher structure. A random sequence retaining the G repeats (PF1N) was tested as well. Two higher structure constructs were made: PF1Flip contained a flipback structure in the presence of three of the diguanosine repeats; PF1 loop had a hairpin-loop structure in between the middle diguanosines. Despite the structural differences, all of the constructs had  $K_d$  values increasing approximately 6-fold or even more in comparison to PF1 (Table 3-1). The structure of the constructs in fact played little role in the tighter binding; for example PF1A had a  $K_d$  value over 750 nM, while the  $K_d$  for PF1T was ~ 450 nM. Both of these are predicted to

have essentially no secondary structure, similar to the very low structure of PF1, yet they bind poorly.

### *3.3.6 The selected aptamer PF1 can inhibit HIV-RT primer extension, but not as effectively as some other previously characterized RT aptamers*

An HIV-RT primer extension assay was performed to determine whether the selected aptamer PF1 could effectively inhibit HIV-RT *in vitro*. A 33 nt 5' <sup>32</sup>P end-labeled DNA primer was hybridized to a 50 nt template (50 nM final concentration, complex shown in Figure 3-3). HIV-RT utilizes this primer-template to extend the primer portion towards the 5' end of the template. Assays were performed using 50 nM primer-template and 0.25 nM HIV-RT and the amount of extended primer was measured over a time course from 2-20 min (Figure 3-4). Results were plotted (relative counts vs. time) and used to calculate the half-maximal inhibitory value (IC<sub>50</sub>) as described Material and Methods. The IC<sub>50</sub> of PF1 was 63 ± 24 nM. This concentration is comparable to the primer-template concentration used in the experiment, suggesting that PF1 binds approximately as well as the primer-template. The IC<sub>50</sub> for 5'-N<sub>30</sub>-3' was > 500 nM (Figure 3-5, Table 3-2). Interestingly, even those oligos that showed only modestly lower affinities for RT in the gel-shift assays (PF1-26ntT and PF1-terminal GG>CC) were poor inhibitors of RT extension relative to PF1.

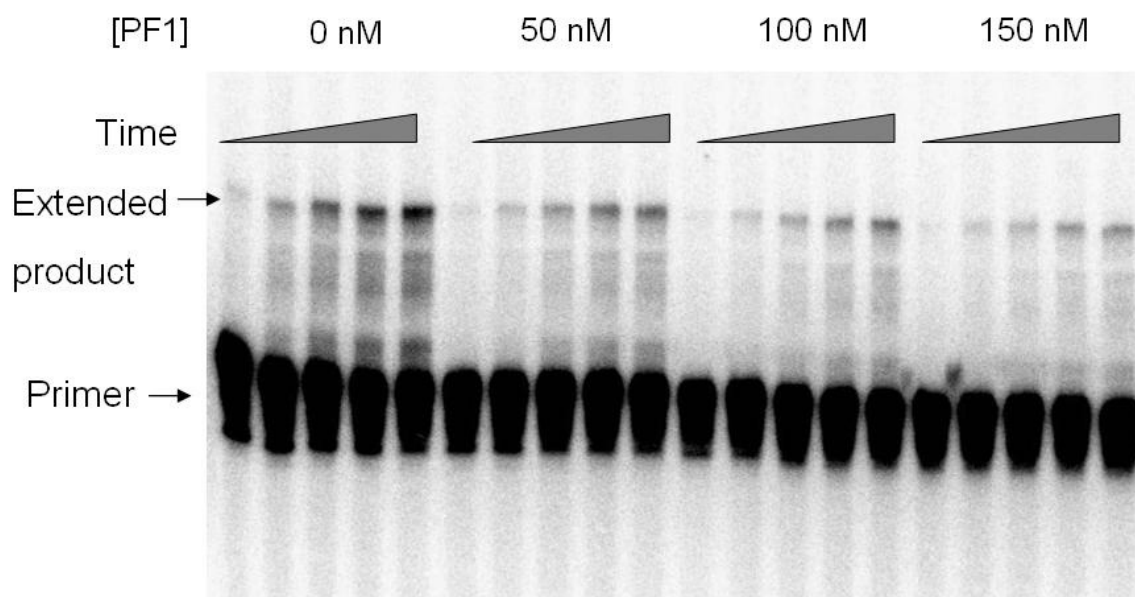
As a comparison, we used a class of G-quadruplex RT aptamers studied by the Michalowski *et al.* that were potent inhibitors of HIV-RT (112). The main G-quadruplex aptamer, known as “R1T”, was a 41 nt oligo with a three tier parallel G-quadruplex (Figure 3-1B). We also evaluated “S4”, a 35 nt DNA similar to R1T with minor





3' – **GATATCCCGCTTAAGCTCGAGCCATGGGCCCT** – 5'  
 5' – TTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATC – 3'

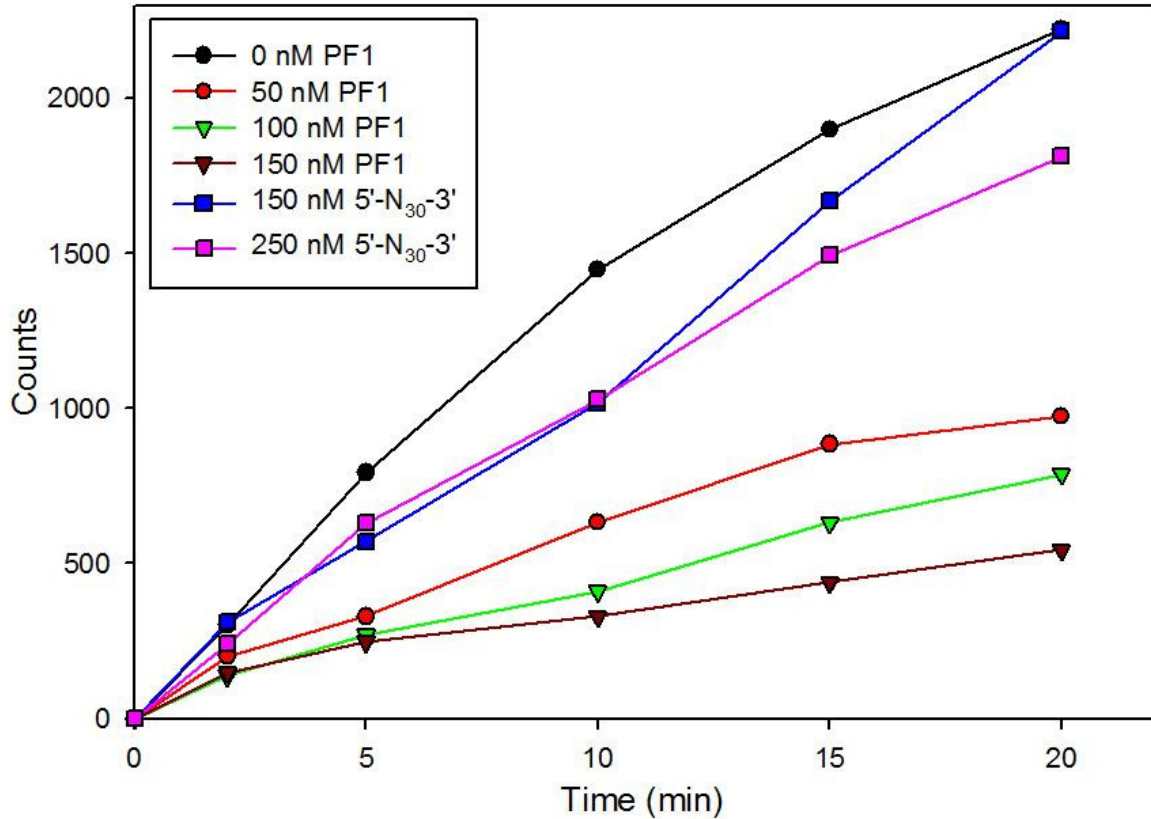
**Figure 3-3: A schematic representation of the primer-template used in the RT inhibition assays. The primer was indicated in bold.**



**Figure 3-4: A representative autoradiogram using various concentrations of PF1.**

The assay was conducted by incubating the primer-template (50 nM 5'  $\gamma$ -<sup>32</sup>P end-labeled primer) with 100  $\mu$ M dNTPs and various concentrations of aptamers as indicated.

Reactions were initiated with HIV-RT (0.25 nM), and aliquots were removed at 2, 5, 10, 15 and 20 minutes. The samples were run on a denaturing 10% polyacrylamide gel and quantified as described in Materials and Methods. The IC<sub>50</sub> determination of PF1 was performed as described in Materials and Methods.



**Figure 3-5: Graph showing the relative counts versus time in the presence of various amounts of PF1 or starting material (5'-N<sub>30</sub>-3').** The number of counts is defined as photo-stimulated luminescence (PSL) radiation unit measured. The assay was conducted by incubating the primer-template (50 nM 5'  $\gamma$ -<sup>32</sup>P end-labeled primer) with various concentrations of aptamers and dNTPs. Reactions were initiated with HIV-RT (0.25 nM), and aliquots were removed at the indicated times. The samples were run on a denaturing 10% polyacrylamide gel and quantified as described in Materials and Methods.

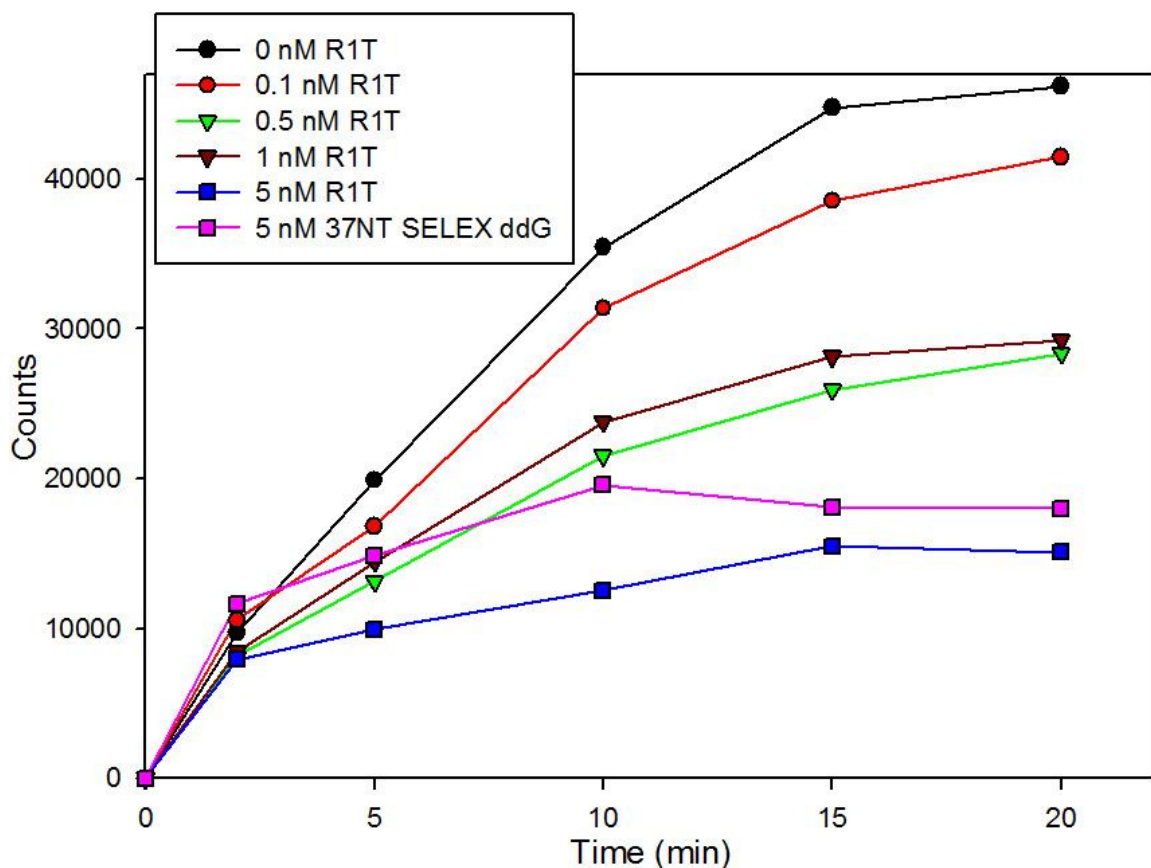
modifications to the stem-loop component, while retaining the G-quadruplex structure (Figure 3-1D). Consistent with the previous results, both the R1T and S4 aptamers were potent inhibitors in the RT extension assay used here with  $IC_{50}$ s of  $0.90 \pm 0.33$  nM and  $0.54 \pm 0.12$  nM, respectively (Table 3-2; Figures 3-6 and 3-7 respectively).

Our lab developed a unique primer-template flipback aptamer, “37 NT SELEX”, which is a 37 nt long single strand DNA, and a potent inhibitor of HIV-RT (42, 43) (see Introduction, Figure 3-1C). For the RT inhibition experiments, the primer-template aptamer had the terminal 3' guanosine residue replaced with a dideoxyguanosine (ddG) to prevent extension by HIV-RT. Consistent with previous results, 37 NT SELEX was an effective inhibitor with an  $IC_{50}$  of  $3.7 \pm 1.1$  nM (Table 3-2, Figures 3-6 and 3-7).

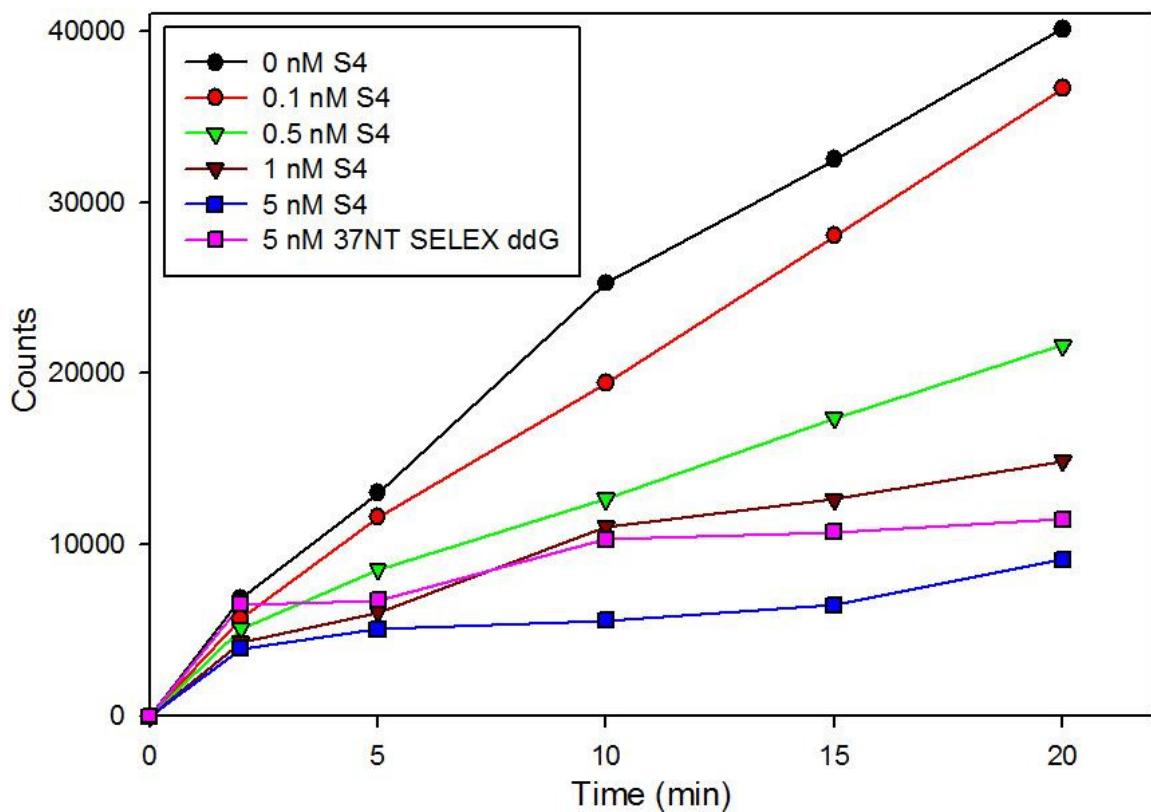
### *3.3.7 Circular dichroism spectroscopy indicates the selected aptamer does not form a G-quadruplex*

Circular dichroism (CD) spectroscopy is an analytical tool for detecting secondary structures and G-quadruplexes in DNA. The characteristic feature of a G-quadruplex is a distinctive positive band at 210 nm, between 260-265 nm, and a negative peak at ~240 nm. For antiparallel quadruplexes, the negative peak is at ~260 nm, and a positive peak at ~290 nm. A regular B-DNA has a negative peak at 210 nm and ~240 nm, plus a positive peak at 220 nm and between 270-280 nm (101) (Figure 3-8). For single-stranded DNA sequences, the CD spectrum is similar to the spectrum for the double-stranded DNA with the same coding strand sequence (76). A hairpin loop DNA “flips-back” onto itself to form an intramolecular B-DNA (101).

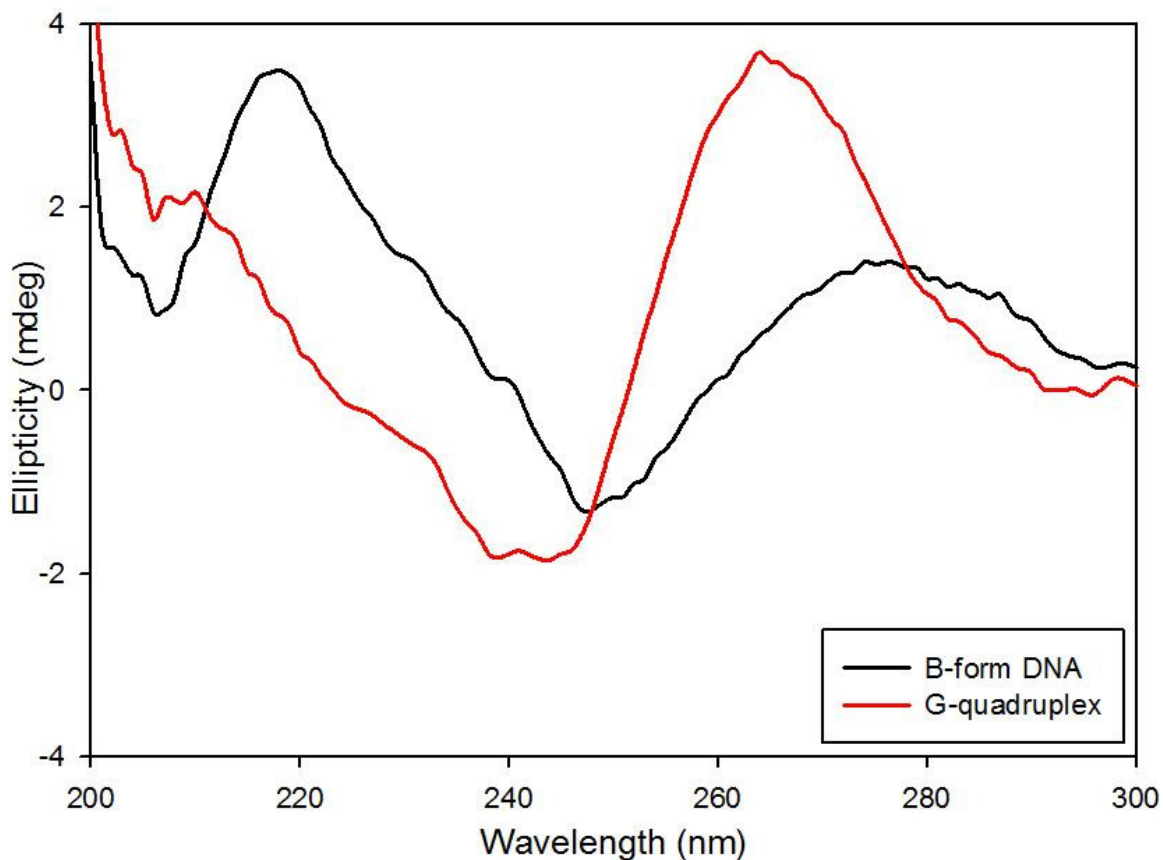
Consistent with the previous results, the CD spectrum for the TBA aptamer



**Figure 3-6: Graph indicating the relative counts versus time in the presence of various amounts of R1T and 37 NT SELEX ddG.** The number of counts is defined as photo-stimulated luminescence (PSL) radiation unit measured. The 37 NT SELEX aptamer contained a 3' terminal dideoxy G residue in place of the normal dG to prevent extension of the aptamer itself. The assay was conducted by incubating the primer-template (5'  $\gamma$ - $^{32}\text{P}$  end-labeled primer, 50 nM) with various concentrations of aptamers and dNTPs. Reactions were initiated with HIV-RT (0.25 nM), and aliquots were removed at the indicated times. The samples were run on a denaturing 10% polyacrylamide gel and quantified as described in Materials and Methods.



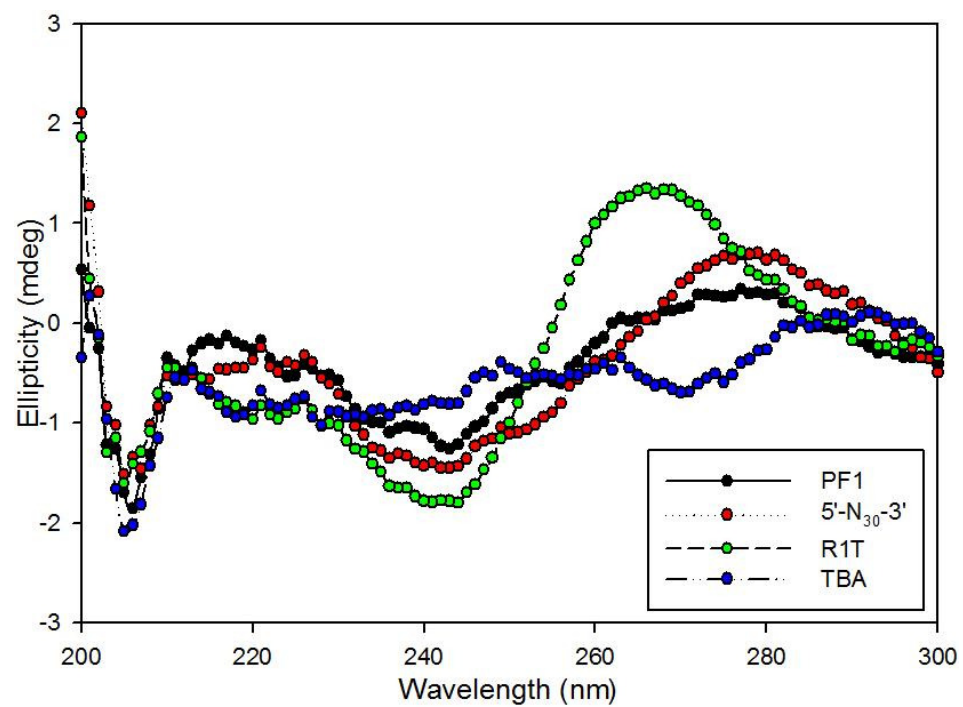
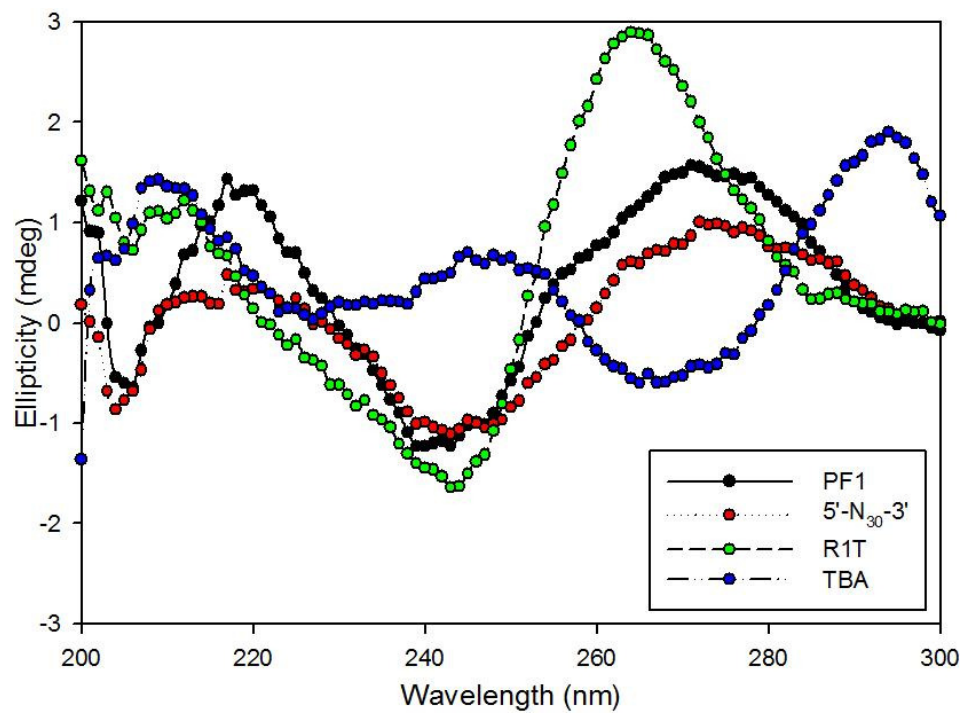
**Figure 3-7: Graph indicating the relative counts versus time in the presence of various amounts of S4 and 37 NT SELEX ddG.** The number of counts is defined as photo-stimulated luminescence (PSL) radiation unit measured. The 37 NT SELEX aptamer contained a 3' terminal dideoxy G residue in place of the normal dG to prevent extension of the aptamer itself. The assay was conducted by incubating the primer-template (5'  $\gamma$ - $^{32}\text{P}$  end-labeled primer, 50 nM) with various concentrations of aptamers and dNTPs. Reactions were initiated with HIV-RT (0.25 nM), and aliquots were removed at the indicated times. The samples were run on a denaturing 10% polyacrylamide gel and quantified as described in Materials and Methods.



**Figure 3-8: A generalized circular dichroism (CD) spectrum of B-DNA and parallel G-quadruplex with the presence of potassium ions.** The B-DNA spectrum was shown in black, while the spectrum for the parallel G-quadruplex was shown in red. The anti-parallel G-quadruplex has a negative peak at ~265 nm and positive peaks at ~210 and 295 nm (data not shown) (101).

(Figure 3-1A) contained positive peaks at 210 and 295 nm, and a negative peak at 265 nm in Figure 3-9, confirming its antiparallel G-quadruplex nature (108). R1T showed positive peaks at 210 and 260 nm, and a negative peak at 240 nm at room temperature in the presence of potassium, indicating R1T is a parallel G-quadruplex. While potassium stabilizes G-quadruplex formation, sodium disfavors formation (156, 169). Figure 3-10 showed R1T retained its G-quadruplex structure (even the spectrum is slightly flattened) in the presence of 80 mM NaCl, while the TBA G-quadruplex structure was disrupted under these conditions (142, 156, 169). The retention of quadruplex structure by R1T in the presence of Na<sup>+</sup> likely reflects the high stability of this particular quadruplex (112). PF1 showed a positive peak at ~270 nm and a negative peak at ~240 nm as well as a positive peak at 220 nm, and negative peak at 210 nm, indicating PF1 did not form a G-quadruplex. The constructs with the guanosine doublets substituted (PF1-GA, PF1-GC, see table 3-1) were compared with PF1 by CD measurement and yield a spectrum similar to PF1. A 37 nt Linear DNA Control was also tested to observe the CD spectrum of a linear single-stranded DNA in the presence of potassium ions. All modified constructs and the 37 nt Linear DNA Control had patterns close to that of the PF1 in the CD spectroscopy, indicating that the PF1 and the various constructs did not form intramolecular G-quadruplexes (Figure 3-11).

To study the CD spectra of single-stranded DNA sequences with different levels of structure, several single-stranded DNA constructs were measured. These included PF1, the G-quadruplex aptamer R1T, homopolymeric 30-mers composed of A (5'-dA<sub>30</sub>-3' in Chapter 2, dA<sub>30</sub> in Chapter 3) or T (5'-T<sub>30</sub>-3' in Chapter 2, T<sub>30</sub> in Chapter 3) which presumably contain little or no secondary structure, the flipback 37 NT SELEX aptamer,



**Figures 3-9 & 3-10: CD spectra of 4  $\mu$ M of PF1, R1T, starting material (5'-N<sub>30</sub>-3') and TBA.** Aptamers in Figures 3-9 and 3-10 were measured in potassium and sodium buffer respectively. Buffer conditions were described in Methods and Materials.

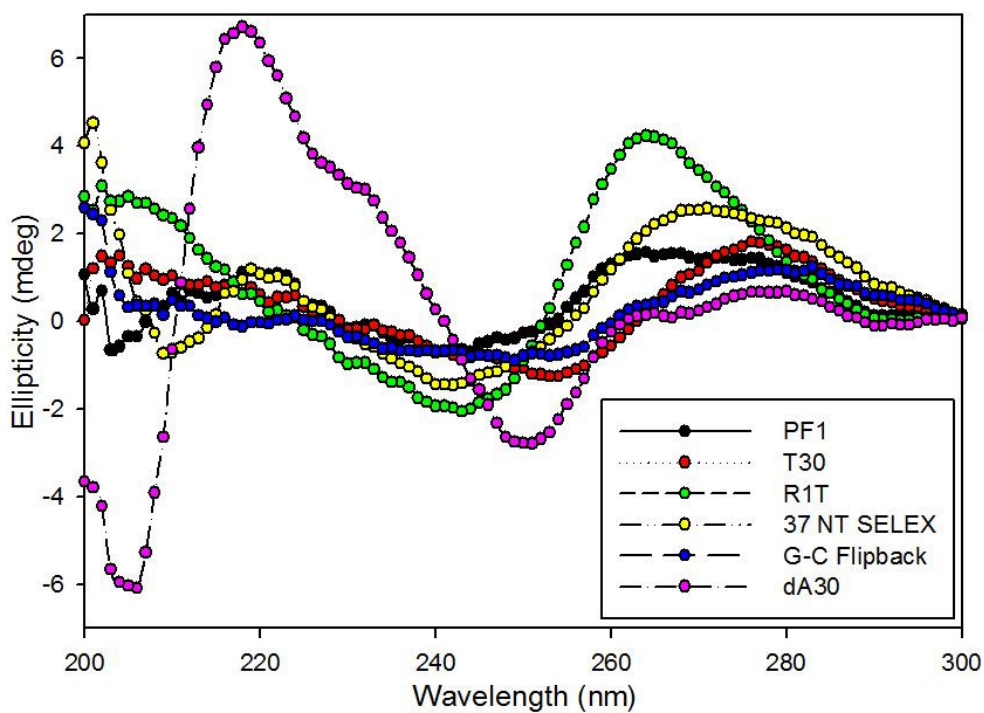
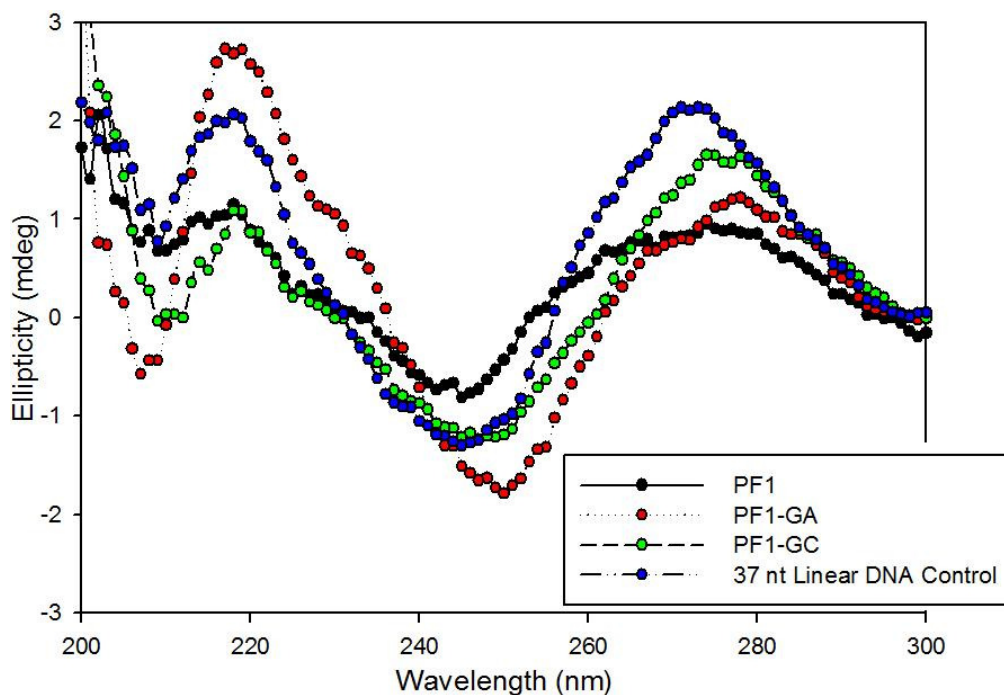


and a highly structured 30 nt G-C rich sequence (G-C flip-back, see Figure 2-5A in Chapter 2). The PF1 aptamer and the various DNA constructs (with the exception of dA30, see below) showed similar CD spectra (Figure 3-12). Note that dA30 (5'-dA<sub>30</sub>-3'), showed a sharp positive and negative peak at 220 and 210 nm, respectively, which is a characteristic trait of poly [d(A)] DNA sequence (101). Consistent with previous results, R1T retained a CD spectrum for DNA forming a parallel G-quadruplex.

### **3.4 Discussion**

#### *3.4.1 Effects of the guanosine repeats on HIV-RT binding and possible structures of PF1*

In this study, we characterized a group of single-strand DNA aptamers to HIV-RT that were isolated using a novel primer-free SELEX technique (102), and defined whether the aptamer affinity to RT was dependent on its length, sequence or structure. We speculated that the regular arrangement of four separate diguanosines in the sequences affect the tighter binding of the selected aptamer to HIV-RT. This prediction was based on this arrangement in aptamers isolated from two separate SELEX experiments (Chapter 2). The likelihood of such a specific sequence occurring arbitrarily in two separate SELEX experiments let alone in a single selection is highly unlikely. Although it is possible that the diguanosine array resulted from the unique SELEX protocol used to select the aptamers (see Chapter 2), the fact that the selected aptamers bind much more tightly to HIV-RT than the SELEX experiment starting material strongly suggests that their selection was based on affinity to RT. It could also be argued that the low predicted secondary structure of PF1 was at least in part due to the Thermostable RNA ligase step in the SELEX protocol favoring low structure and therefore excluding



**Figure 3-11: CD spectrum of 4  $\mu$ M of PF1, PF1-GA, PF1-GC and 37 nt Linear DNA Control, and Figure 3-12: CD spectrum of 4  $\mu$ M of PF1, T30, R1T, 37 NT SELEX, G-C Flipback and dA30. The various constructs were measured in potassium buffer, as described in Methods and Materials.**

structured aptamers from being selected. However, results showed that only strong secondary structure led to inefficient ligation (102) (Chapter 2). The low secondary structure of PF1 may permit it to more easily conform to RT and maximize binding. The low structure could provide flexibility that allows the diguanosines to be positioned for optimal binding to RT or formation of a G-quadruplex. However, low structure alone clearly cannot account for PF1's affinity to RT as oligo's with even lower predicted structure and identical diguanosine spacing bound RT poorly (PF1A, and PF1T). Therefore, PF1's strong binding results at least in part from the specific sequence inbetween the diguanosine repeats, even though those sequences presumably generate little structure. Other results suggested that low structure was a factor in tight binding to RT. When a loop structure was placed between the middle diguanosines (PF1Loop), or PF1 is modified into a flip-back DNA (PF1Flip), both constructs did not bind tightly to HIV-RT, even though their sequences are close to the PF1. Overall the results suggest the interaction between PF1 and HIV-RT is complex and may be driven by diguanosine repeats optimally spaced by specific sequences that generate little structure.

The diguanosine repeats in PF1 were similar to those found in intramolecular G-quadruplex forming aptamers that bind tightly to other proteins (108, 112). In addition, G-quadruplex forming aptamers have been shown to bind HIV-RT and many other proteins with high affinity (see Introduction). Surprisingly, CD spectroscopy indicated PF1 did not form a G-quadruplex (Figure 3-9). This does not necessarily rule out the possibility of PF1 forming G-quadruplexes or other higher order structures, since both the low and high structure constructs all showed similar CD spectra (Figure 3-12). The interaction between HIV-RT and PF1 may induce a change in the conformation of the

aptamer leading to G-quadruplex or other structural formation. If this were the case, and assuming quadruplex formation is required for tight binding, then changes in any of the four guanosine doublets would be expected to be highly detrimental to RT binding. Changing any doublets did decrease binding but the two central doublets were clearly more important as changes in these had a profound affect on RT affinity (Table 3-1). PF1-terminal GG>CC, which changes the diguanosines at the termini to Cs showed just a modest decrease in binding. PF1-26ntT which changed the terminal Gs to Ts and also deleted the flanking terminal nucleotides similarly showed a small change in binding. It is important to reiterate, however, that gel-shift is not a highly quantitative method for measuring affinity. Consistent with this, PF1-26ntT and PF1-terminal GG>CC were much weaker inhibitors than PF1 in the RT extension assay (Table 3-2) despite the small changes observed by gel-shift. To clearly determine whether G-quadruplex formation is important to PF1's interaction with RT, structural analysis using nuclear magnetic resonance (NMR) or X-ray crystallography would be required.

#### *3.4.2 The role of the aptamer length to HIV-RT tight binding*

Although PF1 demonstrated ~10-fold tighter binding to RT than the starting material, deletion of just a few nucleotides caused a significant decrease in binding (Table 3-1). Even PF1-26ntG which retains all the diguanosines and intervening sequences, but lacked 1 and 3 nts from the 5' and 3' ends, respectively, showed a modest loss in binding. Any changes that eliminated or changed the spacing between diguanosine runs essentially eliminated tight binding. From these results it appears that there may be length limitations for tight binding to RT as well as specific sequence spacing

requirement for the diguanosine. The second point is strengthened by the two separate SELEX experiments that selected aptamers with identical diguanosine spacing (Chapter 2) (102). In fact the entire sequence of the selected material from each experiment was very similar with typically just 1-2 nucleotide differences in each sequence in comparison to PF1. This also suggests that there is sequence specificity for non-diguanosine bases as well. With respect to length, DNA aptamers to HIV-RT reported in the literature are typically larger than PF1. The 37 nt and 35 nt aptamers (37 NT SELEX and S4, respectively) used in the inhibition studies here are among the shortest reported for aptamers that bind in the nM range (Table 3-2). Perhaps 30 nts is approaching the limit required to obtain low nM binding for a DNA aptamer. In contrast RNA aptamers of ~ 30 nts that form pseudoknots can bind RT with very high affinity (pM binding constants are reported) (96). We are currently attempting to select larger and smaller DNA aptamers than PF1 using the primer-free SELEX method. This should help determine the importance and limitations of aptamer length.

#### *3.4.3 Comparison between various classes of aptamers in HIV-RT inhibition assays*

HIV-RT and other DNA polymerases utilize a primer-template duplex as their natural substrate. In an RT extension assay, the HIV-RT aptamers can compete with the primer-template complex as has been shown for 37 NT SELEX and the R1T and S4 G-quadruplex aptamers used here (42, 43, 112). The G-quadruplex inhibitors were the most potent with IC<sub>50</sub> values that were ~ 5 fold lower than 37 NT SELEX, which was also a potent inhibitor. It is notable that the G-quadruplex inhibitors seem to bind even more tightly than an “optimized” primer-template, the natural substrate for HIV-RT. Even

though PF1 did not inhibit as well as the other aptamers, the determined  $IC_{50}$  value ( $63 \pm 24$  nM) indicated that it still bound approximately as tightly as the primer-template used in the assays. Such tight binding would not be predicted for a single stranded B-form 30-mer with very low predicted structure. The possibility that PF1 was forming a loop-back type structure or dimerizing to form a primer-template was investigated. Folding and hybridization programs did not predict any of these structures and no extension of PF1 was observed in assays with HIV-RT and dNTPs (data not shown). This suggests that no 3' recessed termini was formed.

The stronger inhibition observed here and tighter binding reported for 37 NT SELEX and the G-quadruplex aptamers indicates that these aptamers bind with significantly greater affinity to RT than PF1. There are several possible reasons for this including the different SELEX protocols used, the length of the aptamers, and the extensive post-SELEX modifications of the 37 NT SELEX and the G-quadruplex aptamers. For example, the 37 NT SELEX binds ~10 fold better than the aptamers selected from the original SELEX (112) while the G-quadruplex aptamers used here are about twice as potent in inhibition assays as the parent aptamers from which they were derived (112). With respect to length, when 37 NT SELEX was reduced to only 27 nt, its  $K_d$  increased by a 100-fold, even though the run of 7Gs was still present (43). Similarly, the shortest G-quadruplex aptamer was 35 nt, and it required the stem-loop component to achieve tight binding to HIV-RT (112, 149). It will be interesting to see if PF1 can also be modified for tighter binding, for example, by increasing its length.

#### *3.4.4 Applications and future work*

The HIV-RT aptamers tested here were derived from various SELEX protocols including a traditional protocol for the G-quadruplex aptamers, and two protocols developed in this lab: one that selects “primer-templates” for 37 NT SELEX (42), and a second primer-free protocol for PF1 (102). Aptamers derived from each protocol had unique properties with respect to sequence and structure. In particular the primer-free SELEX generated low structure aptamers although it was not entirely clear if this is a general property of this protocol or was specific to HIV-RT. The protocol includes a step with Thermostable RNA Ligase that was sensitive to structure, but only when the substrates for ligation were highly structured. Presumably this step could be modulated to help regulate the structure of the aptamers that are selected. This could be done by lowering the ligation temperature (60 °C in the standard protocol) or shorting the time allowed for ligation, both which would favor aptamers with lower structure. The low structure aptamers may not bind as tightly to their target as aptamer binding is often structure-dependent, however, they may have other advantages, particularly as models for the development of cellular inhibitors. It is known that highly structured nucleic acids tend to elicit innate immune responses in cells (70). Low structure aptamers may overcome such cellular response, and also facilitate possible couplings of delivery systems for entry into the cell. Both G-rich oligonucleotides and those that can form G-quartets have effects on cells ranging from induction of senescence and aging (141), to inhibition of proliferation (33, 34, 148). G-quadruplexes are also associated with inhibitory cellular properties, via the oligos presumably binding to nucleolin (41, 62), elongation factor 1A (eIF1A) (19, 38, 59) and STAT3 (170). Classes of aptamers without

these properties could be valuable tools for the development of oligonucleotide therapies in the future.

The primer-free SELEX I developed is a powerful technique for isolating unique aptamers for HIV-RT; the significance of this primer-free SELEX technique lies in the reduction of biases, the relative ease in the selection process, and the selection of a unique set of aptamers for HIV-RT. However, it does not replace the traditional SELEX or other SELEX techniques. It will still be important to use multiple SELEX approaches to generate different classes of aptamers, since there are still values on the other methods for aptamers screening. For example, aptamers with lower structure may be more useful as models for cellular inhibitors, while those that bind very tightly to the target protein may be more useful in diagnostic applications. Nevertheless, the primer-free SELEX technique I developed is a powerful tool for potential small-molecule based drug and diagnostic development on essentially any protein or even nucleic acid targets.



## Chapter 4: Effects of primer-template aptamers to HIV-RT on inhibition of HIV replication in cell culture

### 4.1 Introduction

HIV-RT and several other proteins, even those that are not typical nucleic acid binding proteins, have been shown to bind strongly to particular RNA and DNA sequences, often based on the secondary and tertiary structures the sequences form (172, 173). These tight binding sequences - aptamers - are synthetic, single stranded nucleic acids with unique three-dimensional structures, allowing the folded sequence to bind specifically to the target protein (137). Aptamers are mainly produced by a process called Systematic Evolution of Ligands by EXponential enrichment (SELEX), which was discovered independently by the Gold, Szostak, and Joyce labs (49, 90, 172). The Gold lab used SELEX to isolate short RNA sequences as high-affinity ligands to HIV-RT (172). The first aptamers isolated for HIV-RT were RNAs with pseudoknot-type structures (173). RNA pseudoknot aptamers have been shown to be potent RT inhibitors by interfering with primer-template extension by RT (28, 65, 81). Later SELEX was also used to select for single stranded DNA aptamers with similar properties of RT inhibition (118, 149). Tight binding to RT is mainly determined by the folded structure of the aptamer instead of aptamer sequence. These aptamers typically bind several-fold more tightly than the natural RT substrate (RNA-DNA or DNA-DNA primer-template) and show high affinity for RT (65, 96). Since aptamer binding depends on multiple contacts with RT amino acids, it is suggested that aptamers would induce fewer mutants that are resistant to RT, though more research is needed to prove this theory (65). Relevant to this

are recent reports of RT aptamer escape mutants (53, 89), however these escape mutants were replication defective in comparison to wild type virus.

SELEX has enabled researchers to select for aptamers that binds to several HIV proteins and targets for other diseases (24, 56, 65, 82, 88, 99, 109, 111). Many of these aptamers are currently being developed as potential treatments for diseases. Other non-therapeutic uses for aptamers are also being explored including their uses as alternatives to antibodies, diagnostic biosensors (known as aptamer beacons), and as a tool to study the mechanisms of viral replication (40, 82, 136). Currently the only Food and Drug Administration (FDA) approved aptamer-based drug is Macugen, which was developed by EyeTech Pharmaceuticals and is used to treat age-related macular degeneration (AMD) by injecting the drug to the center of the retina (121). The reason why there are so few aptamer-based drugs is due to the limited success in human therapeutics (130). The main disadvantage of DNA aptamer therapy is the lack of an effective method for delivering the inhibitor to target cells in animal models, especially since host nucleases may degrade the inhibitor before it is able to bind to the target protein. Modified sugar and thiophosphate backbones can help protect against degradation, and the potential use of gene vectors for delivery of RNAs is being explored (24, 56). A new approach where oligonucleotides are anchored to small proteins to cargo the protein-nucleic acid complex across the cell membrane is also promising for both RNA and DNA aptamers and other types of nucleic acid inhibitors (171). Recent research suggests that aptamers can enter cells during viral infection without transfection or conjugation to cargo proteins (110, 111). Matzen *et al.* showed a DNA aptamer directed against the spleen focus-forming

virus was able to inhibit viral infection not only in mouse cells, but also when administered to infected mice *in vivo* (110).

Though currently limited as therapeutics due to cost and a lack of delivery systems, aptamers still hold a great deal of promise. Potential uses include not only targeting specific pathogen proteins, but also host proteins involved in pathogenesis. In such, aptamers should be useful in helping to define the potential efficacy of both pathogen and host targets. The possibility of therapeutic use should not be overlooked, especially considering that drug-delivery systems research is a highly active field. Moreover, the study of aptamers could help us to understand the interaction between RT and the different nucleic acid substrates. The structural information gained from this could set the stage for developing small molecule inhibitors that would be useful in HIV treatment.

Our lab has recently used a unique SELEX technique to show that HIV-RT can bind with high affinity to specific primer-template sequences (42). Based on these sequences, which resembled the HIV polypurine tract (PPT), DNA aptamers that demonstrate tight binding and inhibition of HIV RT *in vitro* were developed (43). Our results showed that the aptamers and controls with the flipback structures caused HIV replication inhibition, while less structured controls did not effect replication. The flipback structure controls bound RT *in vitro* with much lower affinity than the flipback aptamer while inhibition of replication in cells was essentially equal. This suggests inhibition may not have resulted from binding to RT but was due to other factors such as activation of cellular pathways. Previous studies have shown that nucleic acids can induce interferon (IFN) production in cells by activating Toll pathways (5). Interferon in

turn induces tetherin (BST-2) expression in cells which can inhibit virus replication (8). BST-2 is a 30-36 kD transmembrane protein that traps fully mature HIV-1 virus particles on the surface of infected cells (120). HIV counteracts BST-2 inhibition with Vpu (176). The effect of the aptamer inducing interferon, and subsequently BST-2 expression was an aspect we considered. We did not find any evidence of a correlation between tetherin induction and the presence of aptamers or flipback controls used in this study. This does not rule out the possibility that interferon or other cellular pathways may cause HIV-1 inhibition in cell culture.

## **4.2 Materials and Methods**

### *4.2.1 Materials*

Human immunodeficiency virus reverse transcriptase (HIV-RT) was from Worthington Biochemical Corporation. dNTPs were from Roche Applied Sciences. Klenow DNA polymerase and dideoxyguanosine triphosphate were from Affymetrix, Inc. Radiolabeled compounds and britelite plus Reporter Gene Assay System kits were obtained from PerkinElmer, Inc. All oligonucleotides were commercially synthesized from Integrated DNA Technologies. Sephadex G-25 spin columns were from Harvard Apparatus. Fetal bovine serum (FBS) was from Aleken Biologicals. Jurkat E6-1 clone cell line was obtained from American Type Culture Collection (ATCC). TZM-bl cell line and anti-human BST-2 antibodies (diluted at 1:5000) (114) were obtained from the AIDS Research and Reference Reagent Program, National Institutes of Health (NIH). Horseradish Peroxidase (HRP)-conjugated goat anti-rabbit IgG antibodies (diluted at 1:10000) and rabbit anti-human GAPDH antibodies (diluted at 1:5000) were obtained

from Thomas Scientific, Inc. Enhanced chemiluminescence (ECL) Western blotting detection system and ECL DualVue Western blotting markers were obtained from GE Healthcare Bio-Sciences Corp. CellTiter 96 AQueous One Solution Reagent was from Promega, Inc. P24 antigen capture enzyme-linked immunosorbent assay (ELISA) kits were provided by SAIC-Frederick, Inc. Vectashield mounting medium was from Vector Laboratories, Inc. All other chemicals and media were from Sigma-Aldrich Co., Thermo Fisher Scientific, Inc., Lonza Group, Ltd., or VWR Scientific.

#### 4.2.2 Methods

*Effects of various aptamers on Jurkat cell proliferation:* Jurkat clone E6-1 cells were washed with RPMI complete media (10% FBS, 1x Pen-Strep (Penicillin: 50 IU/mL, Streptomycin: 50 µg/mL) in RPMI-1640). Jurkat cells (50 µl of  $5 \times 10^4$  cells/ml per well) were plated into a 96-well plate. Various aptamers at their respective concentrations (50 µl each) were added into the cells. The cells were incubated for 48 hours at 37°C, with 5% CO<sub>2</sub> and 95% humidity. After the 48 hours incubation, 20µl of CellTiter 96 AQueous One Solution Reagent was added to each well. The cells were incubated at 37°C for 45-60 minutes in 5% CO<sub>2</sub>, 95% humidity. The absorbance of the samples was recorded using a 96-well plate reader at 450 nm. For the trypan blue dye exclusion method, Jurkat cells (100 µl of  $1 \times 10^5$  cells/ml per well) were plated into a 96-well plate. Various aptamers at their respective concentrations (100 µl each) were added into the cells. The cells were incubated at 37°C, with 5% CO<sub>2</sub> and 95% humidity. Cell growth were monitored each day for 10 days by mixing equal volumes (15 µl) of each cell sample and trypan blue dye. The number of live cells that excluded the trypan blue dye was counted.

*HIV-1 inhibition with various aptamers on Jurkat cells:* Jurkat clone E6-1 cells were washed with RPMI complete media as described above. Jurkat cells (at  $1 \times 10^5$  cells/well) were infected with HIV-1 LAI virus at MOI = 0.01, or mock-infected with 1x phosphate-buffered saline (PBS). Cells were plated in a 24-well plate, with a total volume of 150  $\mu$ l in a 24-well plate. Incubation conditions were at 37°C, 5% CO<sub>2</sub> and 95% humidity for 2 hours. The infected cells were pelleted and washed with the RPMI complete media (see above) for two times. The cells were resuspended in 300  $\mu$ l of the RPMI complete media. Various concentrations of the each aptamer and the controls (0, 1, 2 and 4  $\mu$ M) were added to the cells, to make up a total of 600  $\mu$ l in a 24-well plate. The cells were incubated at 37°C, 5% CO<sub>2</sub> and 95% humidity. Aliquots (20  $\mu$ l per sample) were taken on 5, 6 and 7 days post infection (dpi) and replaced with an equivalent concentration of aptamer.

*Limiting-dilution assay to quantify infectious virus produced from various aptamers:* Limiting-dilution assay were performed using five-fold serial dilutions from aliquots obtained above, with the initial viral dilution at 1/500. Fresh Jurkat cells (2.5 x 10<sup>4</sup>/cells per well, 200  $\mu$ l total volume) were infected with the diluted virus, with 3 separate infections for each dilution. The cells were incubated on a 96-well plate, at 37°C, 5% CO<sub>2</sub> and 95% humidity for 5 days. P24 antigen capture enzyme-linked immunosorbent assays (ELISA) were performed on the infected cell supernatant (methods are described in the manual provided by SAIC-Frederick Inc.). The plates were read at 450 nm with a 96-well microplate reader. The TCID<sub>50</sub>/ml was calculated by determining the lowest concentration required to infect the Jurkat cells, using a specific

mathematical equation, provided by the Mullins lab, University of Washington School of Medicine. Limit-dilution assays were also performed by infecting TZM-bl cells (obtained from the NIH AIDS Research and Reference Reagent Program) to test the aptamer inhibition during HIV replication. TZM-bl cell line is a HeLa cell derivative with a HIV-1 *tat* promoter upstream of a luciferase reporter and a beta-galactosidase reporter. The cells were treated with 20 µg/ml of DEAE-dextran prior adsorption to increase sensitivity of the cells to HIV-1. The cells will be infected with viral aliquots diluted in DMEM supplemented with 10% FBS and 1x Pen/Strep. At 48 hours post infection, the cells were washed once with 1x phosphate-buffered saline (PBS). The Britelite plus Reporter Gene Assay System was added to the samples according to the manufacturer's instructions. The plates were read with the Centro XS<sup>3</sup> LB 960 Microplate Luminometer from Berthold Technologies, and results were recorded with the MikroWin 2000 software.

*Determining aptamer cell entry with the presence or absence of virus:* Jurkat clone E6-1 cells were washed with RPMI complete media as described above. Jurkat cells (at  $1 \times 10^5$  cells/well) were infected with HIV-1 LAI virus at MOI = 0.1, or mock-infected with 1x PBS. Cells were plated, incubated, and washed as above. The cells were resuspended in 300 µl of the RPMI complete media. Various concentrations of the fluorescein-tagged aptamer (0, 0.5, 1, 2 and 4 µM) were added to the cells (total volume 600 µl) in a 24-well plate. The cells were incubated on a 96-well plate, at 37°C, 5% CO<sub>2</sub> and 95% humidity for 48 hours. The cells were washed once in 1x PBS, then fixed in 4% formaldehyde, 10 minutes at room temperature. The fixed solution was removed, and cells were washed three times in 1x PBS. DAPI (4'-6-diamidino-2-phenylindole) stain

was applied to the cells at 1 µg/ml, and incubated at 37°C for 10 minutes. The cells were washed three times, and resuspended in 10 µl Vectashield mounting medium. The resuspended cells were placed on a glass slide, and sealed with a cover slip. The cells were analyzed using an epifluorescence microscope at 400x total magnification.

*Immunoblot analysis on BST-2 expression in cells with the presence of aptamer:*

Jurkat clone E6-1, CEM-SS and 293 cells were washed with RPMI complete media (or DMEM complete media for 293) as described above. Cells ( $5 \times 10^5$  cells/sample) were resuspended in 300 µl of the RPMI or DMEM complete media. Various concentrations of the aptamer and its controls (0 and 2 µM) were added to the cells, to make up a total of 600 µl in a 24-well plate. The cells were incubated at 37°C, 5% CO<sub>2</sub> and 95% humidity for 48 hours. The cells were washed once in 1x PBS, then resuspended in 80 µl PBS ( $\sim 2 \times 10^6$  cells/sample). The 80 µl cells were mixed with an equal volume of 2X Laemmli buffer [4% SDS, 125 mM Tris-HCl (pH=6.8), 10% 2-mercaptoethanol, 10% glycerol, 0.01% bromophenol blue]. The mixture was incubated at 95 °C for 10–15 min. The denatured cell lysates (15 µl), along with 5 µl of the DualVue molecular marker were run on a SDS-PAGE (5% stacking gel and 10% separating gel). The protein gel was transferred onto a nitrocellulose membrane with a transfer apparatus with 30 V at 4°C overnight. A blocking solution [5% nonfat milk in 1x TBST (50 mM Tris (pH=7.5), 150 mM NaCl, 0.05% Tween)] was applied to the membrane at room temperature for 1 hour. The diluted primary antibodies (rabbit anti-human BST-2, diluted 1:5000 in 10% FBS in 1x TBST) were added to the blot. The blot was incubated at room temperature for 1 hour on a rocking platform. The blot was washed two times at 5 minutes with 1x TBST

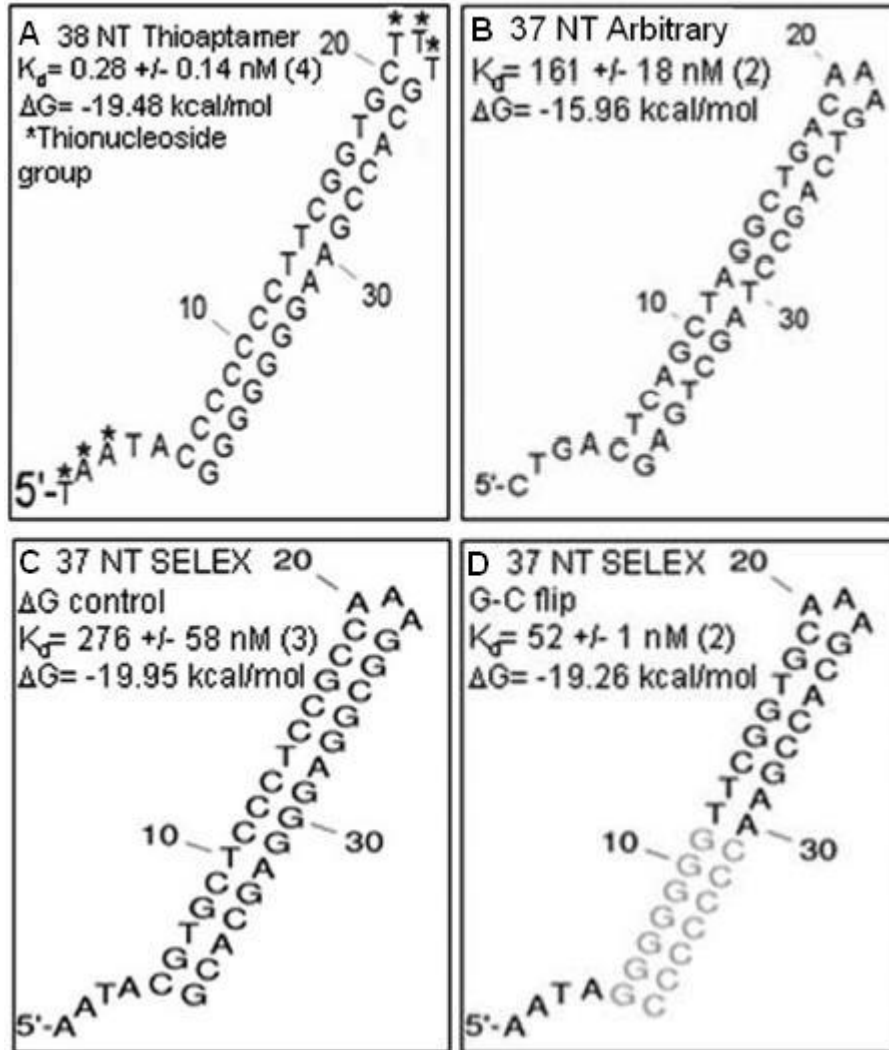


in room temperature. The diluted secondary antibodies (goat anti-rabbit HRP-conjugated antibodies, 1:10000 dilution in 5% nonfat milk and 1x TBST) were added to the blot, and incubated at room temperature for 1 hour on a rocking platform. The blot was washed briefly with 1x TBST, and then washed with 1x TBST for 15 minutes in room temperature. The enhanced chemiluminescence (ECL) Western blotting detection system was added to the blot and visualized with the Fujifilm LAS-3000 Imager for chemiluminescence detection. For the GAPDH detection, the blot was stripped of the bound BST-2 antibodies. After the ECL detection, the blot was incubated in a stripping buffer (100 mM 2-Mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl (pH=6.7) at 50°C for 30 minutes. The blot was washed two times at 10 minutes with 1x TBST in room temperature. The blot was then blocked and probed with rabbit anti-human GAPDH antibodies (diluted 1:5000 in 10% FBS in 1x TBST) as described above.

## **4.3 Results**

### *4.3.1 The primer-template aptamer is a potent inhibitor to HIV-RT in vitro*

Our lab developed a unique SELEX-based technique for selecting primer-template sequences that bound HIV-RT with high affinity (42). The selected sequences bound RT with at least 10-fold tighter binding than the starting material. This was the first time a strong sequence preference for specific primer-templates with a DNA polymerase was demonstrated. Each of the selected DNA-DNA primer-templates resembled the HIV polypurine tract (PPT) in that a run of 6-8 G's was present at the 3' end of the aptamer isolated, and the run of Gs are critical for tight binding to RT. The selected primer-template aptamer was further modified and designed to be as small as

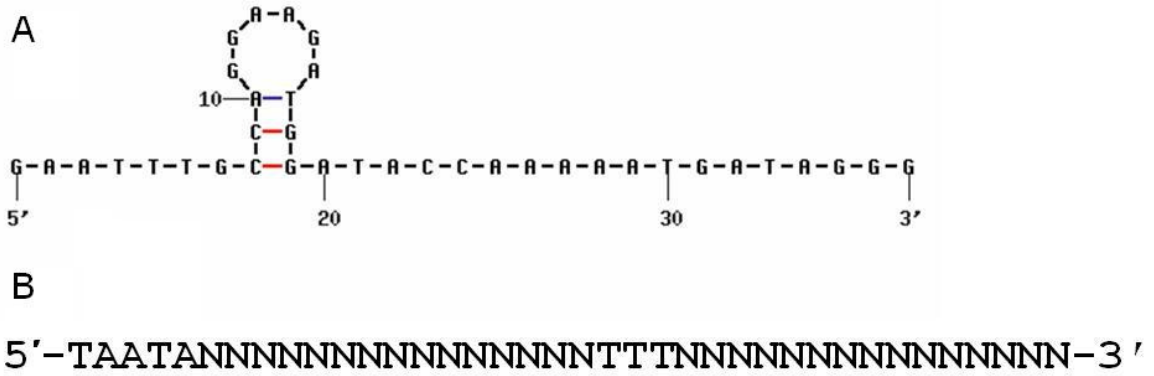


**Figure 4-1: Sequence, structure,  $K_d$ , and  $\Delta G$  values of the various single stranded loop-back constructs.** Structures were predicted using mfold (80 mM NaCl, 6 mM  $Mg^{2+}$  at 37°C) (197). Figure 4-1A showed a modified version of the selected aptamer with phosphothioate nt, denoted in asterisks (\*). The DNA sequences in Figures 4-1B to D were controls with a similar DNA hairpin loop structure. Nucleotides in gray represent important changes to 37 NT SELEX (Figure 3-1). The  $K_d$  values are averages (number of experiments in parentheses) from independent experiments  $\pm$  standard deviation values (43). Figure adapted from DeStefano and Nair, Oligonucleotides 2008 (43).

possible while retaining tight binding and potentially inhibiting RT (43). The designed primer-template like inhibitor, 37 NT SELEX had a  $K_d$  value of  $0.66 \pm 0.16$  nM (see Chapter 3, Figure 3-1B for details) (43).

A potential problem with nucleic acid inhibitors as therapeutics is their susceptibility to nuclease attack in the cell. Phosphothioate nucleosides are modified with the 5'-phosphate replaced by a sulfur atom to resist cellular degradation. The modified aptamers have been used successfully by others with DNA oligonucleotides designed as HIV inhibitors (51, 104, 110, 116). Since single stranded regions of nucleic acid are particularly susceptible to nucleases, our lab designed a 38 nt aptamer with its three 5' nucleotides and three loop nucleotides replaced by phosphothionucleosides (Figure 4-1A). The double-stranded loop portion of this aptamer is identical to the 37 NT SELEX aptamer (Figure 3-1B) while the loop is composed of phosphothioate thymidines rather than adenosine residues. An additional phosphothioate thymine was also added to the 5' end. The  $K_d$  value ( $0.28 \pm 0.14$  nM) indicated tight binding of the aptamer to HIV-RT.

To test for structural importance of the SELEX aptamers, three control substrates with a similar flipback structure were also tested (Figure 4-1B, C and D). Figure 4-1B represents a control in which the nucleotides in the stem portion of 37 NT SELEX were scrambled, while Figure 4-1C shows a similar sequence, while retaining a similar  $\Delta G$  (folding energy) to the designed aptamer. A ~400-fold increase in  $K_d$  was observed for both controls. These controls varied in the degree of decrease in affinity, but both showed weak binding compared to 37 NT SELEX. Thus, there was no correlation between structural stability ( $\Delta G$ ) and affinity ( $K_d$ ), therefore the binding to RT was driven by the specific sequence. In Figure 4-1D the positions of the 7G and C runs in 37 NT SELEX



**Figure 4-2: Sequence and structure of the various linear single-stranded DNA**

**constructs.** The structures shown were predicted with the DNA mfold program using default conditions (80 mM NaCl, 6 mM Mg<sup>2+</sup> at 37°C) (197). Names include the length

in nts of each DNA and the nature of the random sequence. The 37 NT Linear control is a single-strand DNA with low structure ( $\Delta G = -1.17$  kcal/mol). The 37 NT Random

Control is a single-strand DNA, where N denotes random nucleotides.

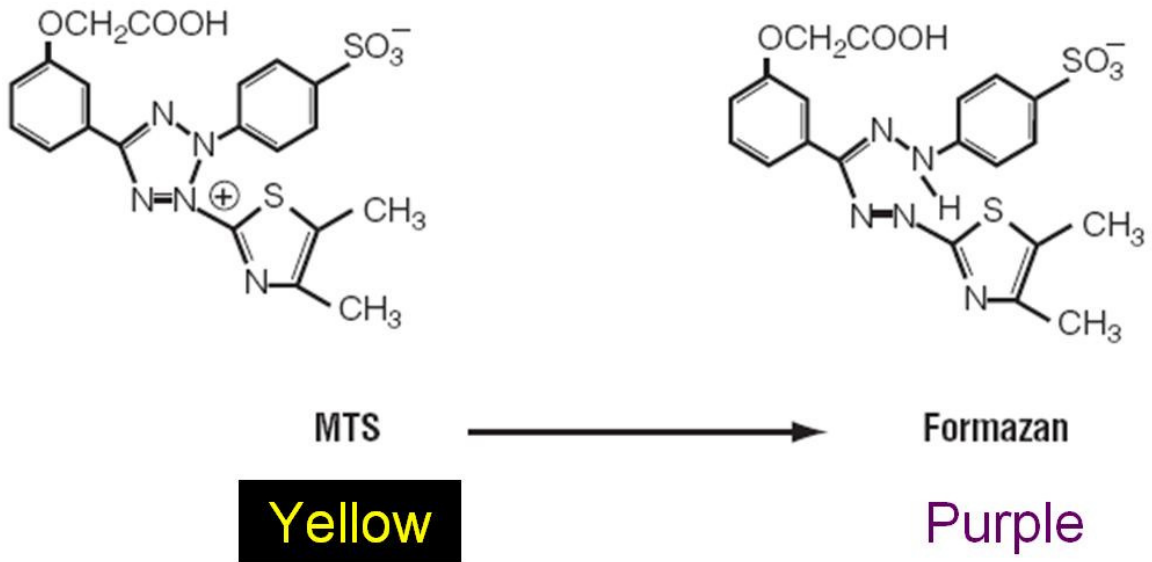
were flipped without significant alterations to the  $\Delta G$  values ( $\sim -19$  kcal/mol). The G-C Flip control showed a  $\sim 80$ -fold decrease in affinity for RT compared to 37 NT SELEX.

To investigate if the structure or the sequence of the designed aptamer played a role in the HIV-1 inhibition, two control DNA oligos with low structures were constructed. The 37 NT Linear control is a single-strand DNA with a linear structure and a weak stem loop (Figure 4-2A). The 37 NT Random Control is a single-stranded DNA that retain the first five nts at the 5' end and the thymidine loop (Figure 4-2B). However, the 37 NT Random Control contains a completely random sequence in the helical region, thus the sequence is unlikely to form a primer-template structure with a hairpin loop.

#### *4.3.2 The aptamers do not have a significant affect on cell growth*

Any cellular insult, physical or chemical, most often results in activation of a cellular pathway, sometimes resulting in either necrosis or programmed cell death. The ability of an antimicrobial agent to cause this effect in the very cells it is supposed to protect is loosely termed cell toxicity. Nucleic acids, especially CG-rich sequences have been known to activate Toll-like receptors in host cells leading to the production of proinflammatory cytokines (178, 179), and inhibition of cell proliferation (33, 34, 148).

To study the effects of the aptamer and controls (Figures 4-1 and 4-2) in cell proliferation, the CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega Inc.) was used. This kit is a colorimetric assay to measure the amount of NADPH or NADH produced in proliferating cells. The NADPH or NADH turns the MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium), which is yellow in color, to a purple formazan compound (Figure 4-3). The purple formazan compound can be directly read by a 96-well plate reader at 450 nm. All



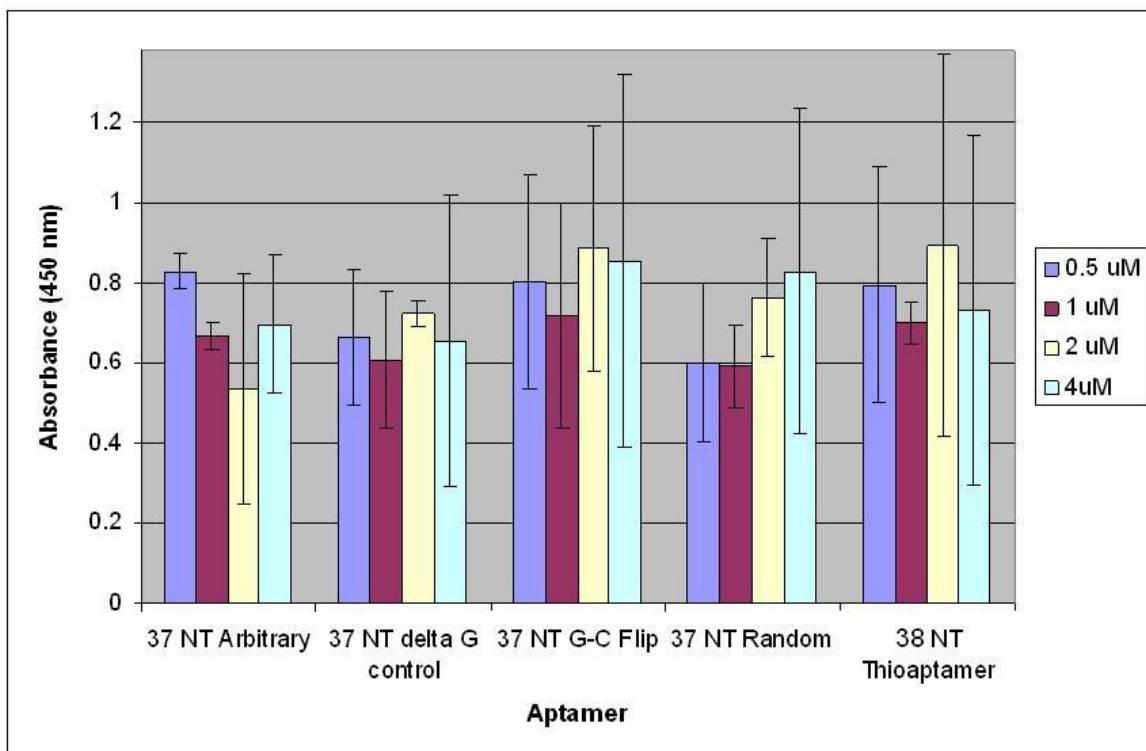
**Figure 4-3: Structures of MTS tetrazolium and its formazan product.** The MTS tetrazolium compound, which is yellow in color, is reduced by cells into a purple-colored formazan product that is soluble in tissue culture medium (35). This conversion is presumably accomplished by NADPH or NADH produced by dehydrogenase enzymes in metabolically active cells (20).

the aptamer and controls at any of the concentrations used (0, 0.5, 1, 2, and 4  $\mu\text{M}$ ) did not have a significant affect on Jurkat E6-1 cell growth over a 48 hour growth period (Figure 4-4). Similarly, the 38 NT Thioaptamer and 37 NT Arbitrary Control both did not have a significant effect on cell growth when monitored with the trypan blue dye exclusion method (Figure 4-5).

#### 4.3.3 All structured aptamers show significant inhibition to HIV-1 infection

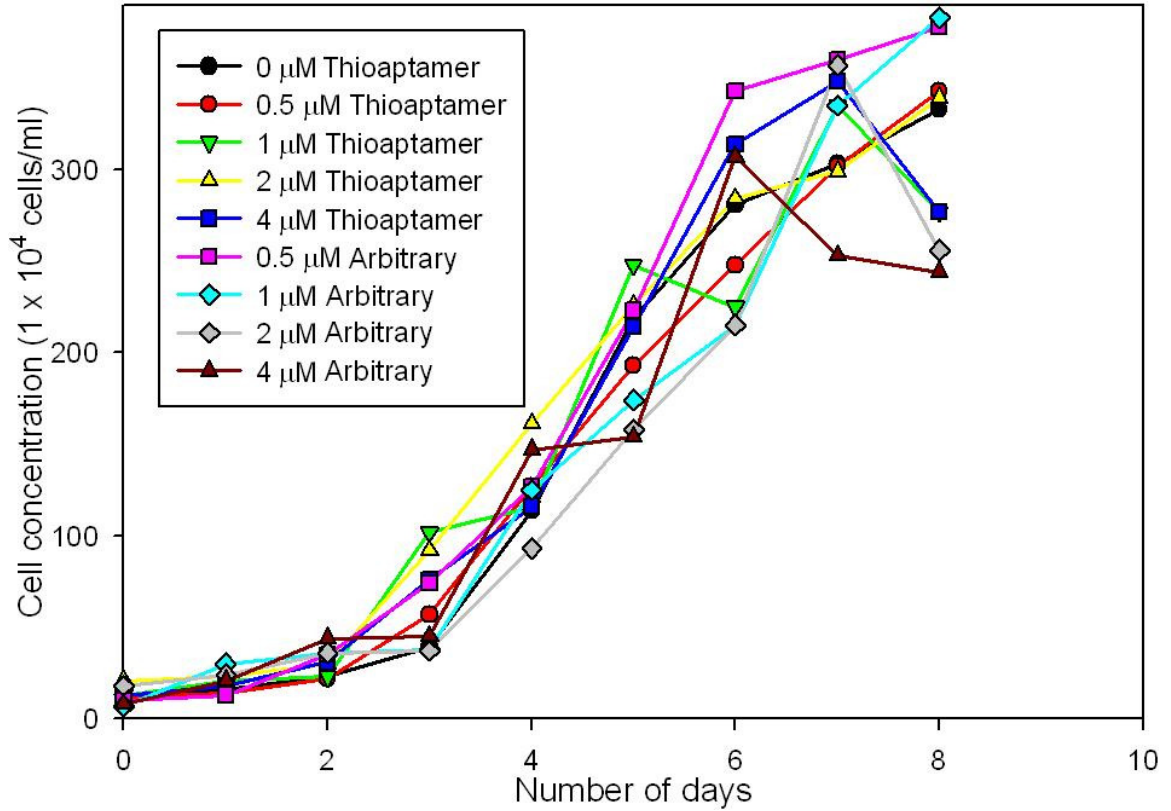
HIV-1 viral replication was examined in Jurkat E6-1 cells with various concentrations of the Thioaptamer and several aptamer controls (see Figure 3A-F and (43)). For these experiments, all the flipback controls were also modified with phosphothioates in the loop and 5' tail region. It is important to note that for the *in vitro*  $K_d$  assays, which required nucleotide addition by HIV-RT, the flipback aptamers had a dG at the 3' end. In the cellular inhibition assays the 3' dG was replaced with ddG using Klenow polymerase (see Methods). This was done to prevent extension of the aptamers which would have resulted in blunt-ends and a loss of affinity for RT (43).

In the cellular assays, no carrier for transfection was used to aid entry of the aptamer into cells; previous studies had shown that some DNA aptamers at high concentrations can enter cells without transfection (110, 111, 116). To measure the infectivity of the virus in the presence of various concentrations of different aptamers, the  $\text{TCID}_{50}/\text{ml}$  (tissue culture infectious dosage, of 50% of the cell population) of the Jurkat cells were measured using a p24 ELISA antigen capture assay or TZM-bl reporter cells (Figure 4-6). The measurement of  $\text{TCID}_{50}/\text{ml}$  was based on a limit-dilution assay, which quantified the infectious virus produced. The 0  $\mu\text{M}$  aptamer control showed a significant increase of

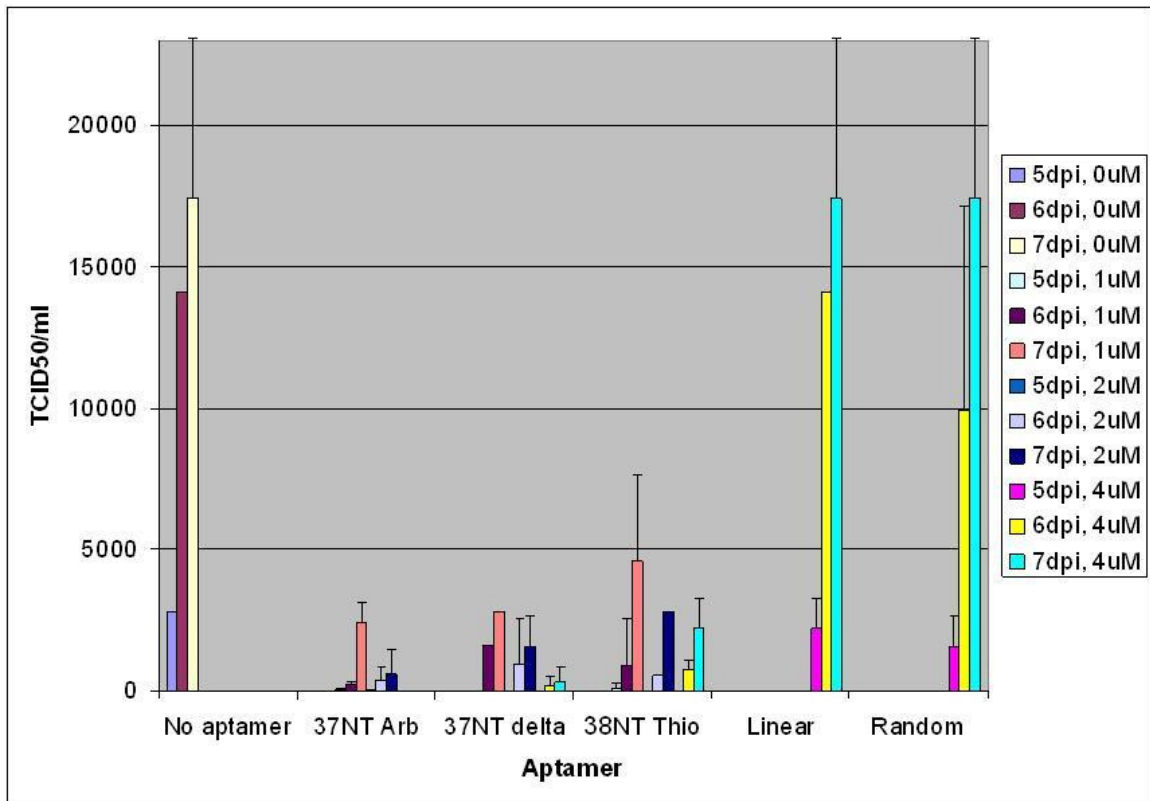


**Figure 4-4: Cell proliferation in Jurkat E6-1 cells in the presence of various concentrations of the 38 NT Thioaptamer and the control DNAs.** Cells were incubated with 0.5, 1, 2 and 4  $\mu\text{M}$  of the 38 NT Thioaptamer and various aptamer controls for 48 hours at  $37^\circ\text{C}$ , 5%  $\text{CO}_2$ . The MTS tetrazolium compound (from Promega) was added to the Jurkat cells and incubated for 48 hours at  $37^\circ\text{C}$ , 5%  $\text{CO}_2$ . The absorbance values (at 450 nm) were averages from three independent experiments. The error bars denoted the standard deviation values.





**Figure 4-5: Jurkat E6-1 cell count in the presence of various concentrations of the 38 NT Thioaptamer and 37 NT Arbitrary Control.** Cells were incubated with 0.5, 1, 2 and 4  $\mu\text{M}$  of the 38 NT Thioaptamer and 37 NT Arbitrary Control for 10 days at  $37^\circ\text{C}$ , 5%  $\text{CO}_2$ . The starting concentration of the Jurkat cells were at  $1 \times 10^5$  cells/ml. Typically Jurkat cells double in  $\sim 24$  hours and can reach densities of  $\sim 3 \times 10^6$  cells/ml. Cell growth were monitored each day for 10 days using the trypan blue dye exclusion method.

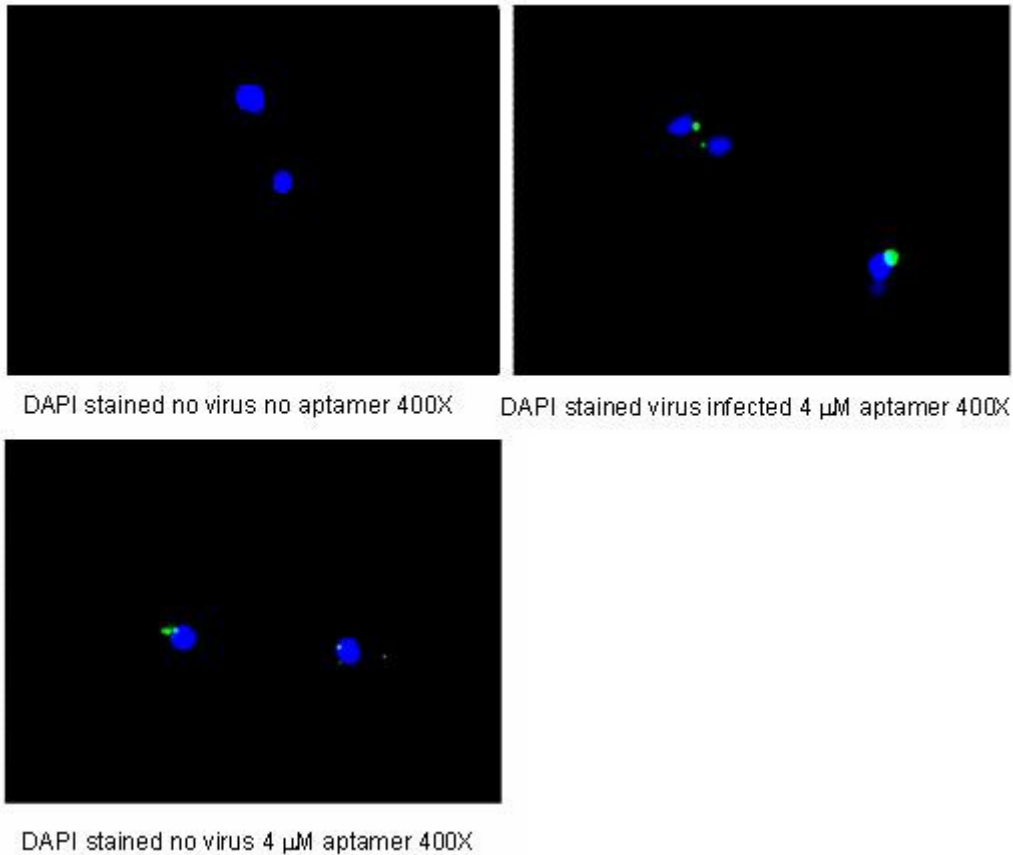


**Figure 4-6: TCID<sub>50</sub>/ml values of Jurkat E6-1 cells infected by HIV-1 LAI virus, with various concentrations of the 38 NT Thioaptamer and its controls.** Cells were incubated with 0, 1, 2 and 4  $\mu$ M of the 38 NT Thioaptamer and various aptamer controls after the adsorption of HIV-1 at 37°C, 5% CO<sub>2</sub>. Time points used in the infection were 5, 6 and 7 days post infection. Cell-free media collected from the various times were used to infect Jurkat cells for 5 days in a limit-dilution assay. The TCID<sub>50</sub>/ml values were based on the calculations with the p24 antigen capture assays for evaluating the limit-dilutions from the various aptamers. The TCID<sub>50</sub>/ml values were averages from two independent experiments. The error bars denoted the standard deviation values.

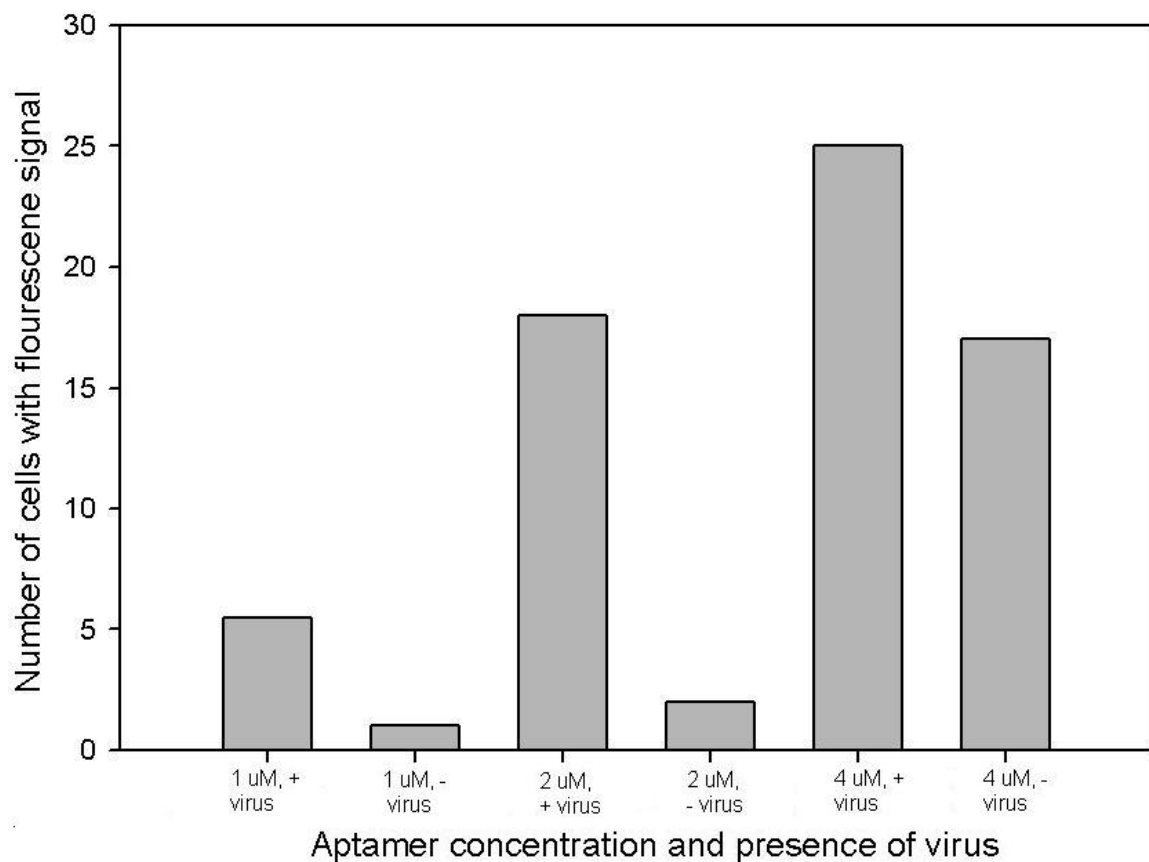
TCID<sub>50</sub>/ml level between days 5 and 6 post infection, and further increase of TCID<sub>50</sub>/ml to a peak level between days 6 and 7, indicating the virus was infecting the Jurkat cells and producing infectious virus. All the hairpin loop aptamers (Figure 4-1) resulted in a significance decrease in infectious virus production, even when only 1 μM of aptamer was present, and a further decrease with 4 μM aptamer. In contrast, both 37 NT Linear Control and 37 Random Control (Figure 4-2) did not show any decrease in the TCID<sub>50</sub>/ml level, indicating the low structure DNA did not have an effect in HIV replication in the Jurkat cells.

#### 4.3.4 Effect of HIV-1 on aptamer entry into Jurkat cells

Since no carrier was used to aid entry of the aptamer into cells, we examined entry of the aptamer into Jurkat cells in the presence or absence of fluorescently-tagged (with fluorescein) Thioaptamer. Figure 4-7 shows representative panels generated by epifluorescence microscopy, which demonstrates the increase in the amount of aptamer in cells in the presence versus the absence of HIV virus infection. Results quantifying cell entry of the aptamer are shown graphically in Figure 4-8. No fluorescent label was observed at 0 or 0.5 μM aptamer (data not shown). At 2 μM aptamer approximately 50% of the cells showed the green fluorescence signal in the presence of virus, while very few cells showed the fluorescein signal in the absence of virus. This is consistent with previous studies by Metifiot *et al.* (111). At 4 μM about 70% of the cells showed the presence of the fluorescein signal in the presence of virus. Note that over 50% of the cells were fluorescently labeled even without the virus. A general observation was that the fluorescence intensity in a given cell increased at the higher amount of aptamer and in the



**Figure 4-7: Entry of the aptamer into Jurkat cells visualized in the presence or absence of fluorescein-tagged 37 NT Thioaptamer.** Jurkat cells were infected at MOI = 0.1 or mock infected with PBS. At 48 hours post infection, cells were washed three times with 1x PBS, fixed in 4% formaldehyde, and then analyzed using an epifluorescence microscope. Cells were stained with DAPI to identify the nucleus. The panels with aptamer were not meant to reflect the proportion of the cells that contained the fluorescein signal, but as a representation of fluorescently labeled cells.



**Figure 4-8: Graph showing the quantification of fluorescein-tagged 37 NT**

**Thioaptamer cell entry in the absence or presence of HIV-1.** A total of 40 cells in each sample were counted, and each cell that contained the fluorescein signal was counted as a positive. Cells were infected, fixed, and stained as described in Methods and Figure 4-6. The values in the graph were averages from two independent experiments.

presence of virus.

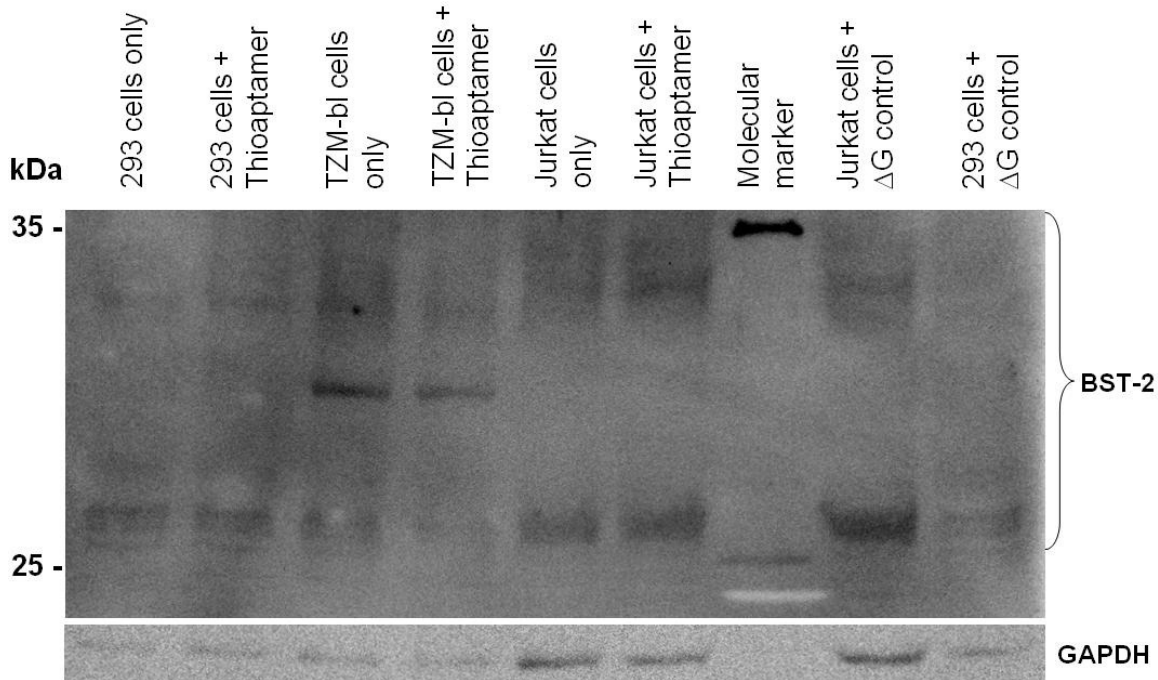
#### *4.3.5 The aptamers did not induce BST-2 production in cell culture*

To investigate if 38 NT Thioaptamer induces interferon production which in turn upregulates BST-2 expression, the 38 NT Thioaptamer and a primer-template construct (37 NT  $\Delta$ G control) were used in cells for analysis of BST-2 expression. The aptamers were incubated with Jurkat cells, HEK 293 cells (both Jurkat and 293 cells are BST-2 inducible) and TZM-bl cells (a HeLa cell derivative; HeLa cells constitutively express BST-2). The cell lysates were subjected to SDS-PAGE, and BST-2 was detected by western blotting probed with an anti-BST-2 primary antibody and a horseradish peroxidase-conjugated secondary antibody. The various cell lines showed very similar expression of BST-2 when detected on an immunoblot, regardless of the presence or absence of the aptamers (Figure 4-9). Note that BST-2 is expressed as multiple bands between 25–35 kDa on an SDS-PAGE due to protein being heterogeneous glycosylated (100). The TZM-bl cell samples had an extra glycosylated band at ~30 kDa, however it is still somewhat surprising that the overall BST-2 expression level was similar to the BST-2 levels in 293 and Jurkat cells and not higher, due to the constitutive expression of BST-2 in this cell line.

## **4.4 Discussion**

### *4.4.1 Effects of primer-template aptamers on cell proliferation*

Previous studies have shown that the DNA primer-template aptamer developed by our lab inhibits HIV-RT *in vitro* (43). It was of interest to determine whether the designed



**Figure 4-9: Effect of 38 NT Thioaptamer and 37 NT ΔG control on BST-2 expression in various cell lines.** Jurkat E6-1 clone cells, HEK 293 cells and TZM-bl cells were pretreated with 0 or 2 μM 38 NT Thioaptamer and 37 NT ΔG control for 48 hours at 37°C. Cells were lysed in 2X Laemmli buffer [4% SDS, 125 mM Tris-HCl (pH=6.8), 10% 2-mercaptoethanol, 10% glycerol, 0.01% bromophenol blue] at 95°C. Levels of BST-2 expression were determined by western blotting using anti-BST-2 antibody. The expression of the GAPDH was also measured by probing the blot to anti-GAPDH antibodies. The different cell lines have various level of GAPDH expression while maintaining the same cell concentration per sample.

aptamer could inhibit HIV replication in cells. To show that the aptamer can effectively inhibit HIV replication, it was important to establish that there was little or no cellular toxicity at concentrations where viral inhibition occurred. Other studies had shown nucleic acids and aptamers in cell media at concentrations in the low micromolar range had little or no toxicity (131). The designed thioaptamer and controls did not alter cell proliferation in a 48-hour period (Figure 4-4), and the trypan blue cell exclusion monitoring over a 10-day period (Figure 4-5) both indicated the cells had a normal cell growth rate. Note that even the cells were dividing normally in the presence of the aptamers, a cell proliferation assays should be performed in the same time course (up to 7 days) as the infection assay. Also, even the cells can proliferate at a normal rate, even when cellular responses are elicited by external factors. Therefore the level of cell proliferation does not necessarily reflect specific changes in the internal cellular environment, some of which could alter virus replication. Some oligonucleotides, most notably those that can form G-quartets and other GC-rich oligos, have been reported to have effects on cells ranging from induction of senescence and aging (141) to inhibition of proliferation (33, 34, 148).

#### *4.4.2 The possible roles of the structure of the primer-template aptamers to HIV-1 inhibition in cell culture*

The thioaptamer and the flipback control aptamers (Figure 4-1) showed comparable inhibition of HIV replication in Jurkat cells. In contrast, the low-structure DNA controls (Figure 4-2) had no effect on HIV replication. This suggests that the structure of the aptamer may have caused the inhibition in HIV production. Assuming



there were no effects on cell proliferation and no cell toxicity, the structured oligos most likely either induced activation of specific cellular pathways, or directly interacted with cellular proteins required for HIV replication, or interacted with viral proteins to inhibit replication. Since all the structured oligos including the thioaptamer formed primer-template configurations, it is possible that they bound to RT and inhibited replication. Since the flipback controls used here bound with low affinity to RT *in vitro* (43), it seems unlikely that they would have comparable inhibition to the tight binding thioaptamer in cells. However, a ddG terminated version of the 37 NT  $\Delta$ G Control showed much greater inhibition in the *in vitro* RT extension assays than was expected from the  $K_d$  measurement (43). A possible reason for this is formation of a “dead-end complex” when the chain-terminator is present. Dideoxy terminated primer-templates bind with a very long half-life to HIV-RT in the presence of the template-directed dNTP that follows the dideoxy moiety (170). This effect may be observed in the cells where dNTPs are present. In effect, the presence of a dideoxy nucleotide at the 3' terminus of all the flipback inhibitors used here could have made them more effective RT inhibitors in the cell. In order to test the possibility of the flipback oligos inhibiting RT in cells we attempted unsuccessfully to isolate revertants that could grow in the presence of the thioaptamer. Presumably this could have occurred through RT mutations that allowed the ddG to be excised so that the oligos could then be extended. In this case changes in RT that decrease binding to the oligos, which may not have been possible, would not have been necessary to overcome inhibition.

In the future, it will be more advantageous to utilize the TZM-bl cell line to directly perform the HIV infection with various aptamers. The TZM-bl cell line is

designed to allow single-cycle HIV infection, which single-cycle infections would allow more precise measurement on the step(s) in HIV replication the aptamers have an inhibition effect. The quantification process will also be more direct, which the viral levels will be measured by the luciferase activity in the TZM-bl cells.

It is known that highly structured nucleic acids can elicit innate immune responses in cells (70). In addition to activation of cellular pathways through Toll-like receptors (132), presumed binding to cellular proteins including ribosomal protein L7a (143), nucleolin (41, 62), elongation factor 1A (eIF1A) (19, 38, 59), STAT3 (170), and growth factors (18) have been reported. It is notable that GC-rich oligonucleotides and those that can form G-quartets are most associated with inhibitory cellular properties. None of the flipback aptamers used here inhibited proliferation at the concentrations used. However, proliferation assays were done in the absence of viral infection and under these conditions entry of the oligos into the cells was reduced (Figures 4-4 and 4-7). Attempts to investigate toxicity by transfecting the oligos into cells were inconclusive because Jurkat and other T cell lines transfected poorly (data not shown). Transfection reagents that improved delivery efficiency (specifically N-TER Nanoparticle siRNA Transfection System from Sigma-Aldrich, Inc.), also inhibited virus infection even in absence of oligos and were not pursued for the purpose of delivery. Therefore it is unclear if the flipback, or even the low structure controls would alter cell proliferation if a higher concentration had entered the cells. With respect to the nature of the oligos used here, the Thioaptamer, though not particularly G rich (12 out of 38 of the nucleotides are Gs), was highly structured and had a high GC content (24 out of 38). Other flipback controls were also highly structured with high GC content. Based on previous results (see above) the

possibility of interference or alterations of cellular functions that lead to HIV inhibition cannot be dismissed.

#### 4.4.3 Possible relationships of the aptamer uptake and the presence of HIV-1

Since the assays to measure cellular uptake were conducted over a short 48 hour period, a higher MOI (0.1 vs. 0.01 for assays measuring aptamer inhibition) was used to provide more viral particles for more pronounced effect on visualization of the cells. Even in the presence of virus it took at least 2  $\mu\text{M}$  of the labeled thioaptamer to visualize aptamer uptake. Aptamer entry in the absence of virus did occur, however it took twice the amount of oligo (at 4  $\mu\text{M}$ ) for  $\sim 50\%$  of the cells to show the fluorescent signal. Since the threshold for observing a signal is not known it is possible that lower levels of aptamer entered cells even in the absence of virus and below the 2  $\mu\text{M}$  oligo concentration. Some inhibition of virus replication was observed at 1  $\mu\text{M}$  oligo (Figure 4-6) despite no apparent entry of the aptamers into cells based on the cell uptake assay (data not shown). This suggests that the aptamer may enter the cell but does not overcome the threshold for visualization. In the future, bright field microscopy will be performed on the Jurkat cells to locate the cell nucleus, and its relation to where the fluorescent signal is located in the Jurkat cell. It is possible that inhibition could be at the level of virus entry. The role of the aptamers in viral entry could be investigated using VSV-G pseudotyped HIV virus (44, 140). These viruses, which use the VSV-G protein to gain entry, are typically much less susceptible to entry blocks. If the virus does facilitate aptamer entry into cells, the virus chaperoned uptake may be size specific; less aptamer was required for uptake the uptake observed by Metifiot *et al.* where aptamers of  $\sim 20$  nts were used.

Aptamers used by Metifiot *et al.* also formed G-quartet structures which could have influenced uptake. Other groups have observed virus chaperoned uptake of unrelated larger nucleic acids (>50 nt) suggesting that this may be a general phenomena (110).

#### *4.4.4 The effects of aptamers on cellular pathways induction and HIV inhibition*

BST-2 was identified as an IFN-inducible host factor in cells (120), where interferon production was induced by viral infection or by nucleic acids. Even though the aptamer at low micromolar concentrations did not induce a higher level of BST-2 production, there are a couple of possibilities concerning the role of the aptamers in cellular pathway induction. BST-2 is relatively stable on the surface of cells, with a half-life of ~8 hours (158), however the aptamer may play a role in blocking the synthesis of the new BST-2 molecules, especially if the aptamer also induce other cellular factors. Even though the level of BST-2 expression was similar with or without the presence of the aptamers, other host cell pathways (such as innate immunity responses) could be activated. Some of the future experiments to determine what factor(s) may play a role in the aptamer inhibition include ELISAs for interferon alpha (IFN- $\alpha$ ) detection, caspase-3/7 assay to study possible effects of programmed cell death, or TLR3 / TLR9 expression to observe if the aptamer is inducing any toll pathways.

#### *4.4.5 Conclusion and future directions*

Overall the above results demonstrate inhibition of HIV replication using high concentrations of flipback nucleic acids in Jurkat cells. Replication inhibition did not correlate with the affinity of the oligonucleotides for HIV-RT as the thioaptamer which

was a better inhibitor of RT *in vitro* and bound to RT with much greater affinity than the other flipback oligos (43), was not a more effective inhibitor (Figure 4-6). Many questions, in particular, how the oligos inhibited replication, were left unanswered. Some relatively straight forward experiments could have been performed to rule out some of the possibilities. For virus entry could be investigated using VSV-G pseudotyped HIV virus (see section 4.4.3 for details). Specific steps in the replication process (1<sup>st</sup> and 2<sup>nd</sup> strand DNA synthesis) could be evaluated using real-time PCR (124), integration using Alu-PCR (144), processing of viral proteins using radioimmunoassays (3), and budding of the virus from the cells using scanning electron microscopy (86). These experiments would have helped to home in on the step(s) of virus replication that were affected by the oligos. Based on previous results, the non-specific nature of the inhibition, and the high concentration of oligos required for inhibition, there was a high likelihood that inhibition was caused by activation of a cellular response, or non-specific interference with virus entry. Such a finding would add little to the current understanding how cells respond to nucleic acid treatment or whether the aptamer was a good model for further development as an HIV inhibitor. Because of this and time limitations, the above experiments were not pursued. These experiments would be much more worthwhile for an aptamer with demonstrated specificity.

One clear problem with this work was the inability of the oligos to gain entry into the cells without adding large concentrations to the medium. There are several methods available to transfect cells with oligonucleotides. In addition to the N-TER reagent described above, vector-assisted delivery methods using liposome vesicles are another alternative method for aptamer delivery. Liposome vesicles contain an aqueous interior

surrounded by a phospholipid bilayer. The aptamer is either enclosed inside the aqueous core or complexed to the phospholipids of the liposome (51, 104, 110, 116). The aptamer enters the cells by fusion of the liposome vesicle and the cell membrane. Adding a lipophilic moiety or cell penetrating protein (HIV Tat protein or Antennapedia penetrin protein for example) to the oligos is another possibility (84). These have been shown to improve the entry of nucleic acids including aptamers, siRNAs and antisense RNAs (171, 174). Attaching the oligos to HIV envelope-specific aptamers that bind to the virus and can cargo other nucleic acids into the cell is another method which has shown some success for aiding siRNA entry (192). To be effective, the delivered aptamer must not only enter the cell and remain intact, it must reach a high enough concentration to inhibit replication and gain access to the replication complex. Aptamer based therapies that target extracellular proteins have already been demonstrated in practice (see “Macugen” in Introduction) and will likely expand rapidly. Those targeted against intracellular proteins must be coupled with effective delivery systems and this will undoubtedly be more challenging.

## Chapter 5: General discussion

Aptamers are a powerful tool in developing various biosciences applications, as well as understanding the mechanistic of protein-nucleic acid interaction (82). There are already various classes of DNA and RNA aptamers for HIV being developed and studied (Chapter 3) (7, 39, 71, 95, 112, 149), however to gain a better understanding on HIV replication, and to develop novel techniques for facilitating the continuing discovery for alternative drug therapy.

The main goal of this dissertation focuses on the study of two different classes of HIV-RT aptamers: (1) to develop and characterize a unique primer-free SELEX aptamer and (2) to investigate the effects of a previously isolated primer-template HIV-RT aptamer on HIV-1 infection in cell culture. In chapter 2, I introduced a novel primer-free SELEX technique on selecting aptamer targeting HIV-RT. The only other HIV-RT aptamer selected from a similar method (minimal primer SELEX) was developed by Joshi and Prasad (87). In primer-free SELEX (or minimal primer SELEX), the main goal is to eliminate the biases from the flanked regions present in the random sequences selected from traditional SELEX. Primer-free SELEX techniques usually requires hybridization and enzymatic steps to regenerate the flanked regions for PCR amplification (126, 127, 184), which required more time and resources for each round of the SELEX process. However, the reduction of biases may outweigh the added effort to select for random sequences directly interacting with their targets.

In my primer-free SELEX technique, Thermostable RNA ligase was used for streamlining the primer-SELEX process. However, there were some limitations posed by the use of this RNA ligase: the enzyme favored a T at the 3' end, and in particular

disfavored a G at the 3' end (102). In future selections, a random pool of sequences with a T at the 3' end will be used to increase the efficiency of the selection process.

Thermostable RNA ligase showed reduced efficiency in the ligation of DNA sequences with very high structures, but oligos with some secondary structures ligated with efficiency comparable to low structure DNA (Chapter 2, Figure 2-5) (102). Surprisingly, with two separate SELEX experiments a set of sequences almost identical to each other and with arrangement of four guanosine doublet was selected (102). The selected aptamer (PF1) bound ~10 fold tighter than the starting material. Chapter 3 focused on understanding on PF1 and the factors contributing to PF1's high affinity to HIV-RT. Two different classes of HIV-RT aptamers were compared with PF1 on their ability to inhibit RT extension: (1) G-quadruplex aptamers (R1T and S4) (112), and (2) a primer-template aptamer developed from our lab (37 NT SELEX) (42, 43).

PF1's binding ability to HIV-RT was affected by the central diguanosine repeats, length, and the nucleotide arrangement. Even though the guanosine motifs are found in DNA sequences that form an intramolecular G-quadruplex (Chapter 3) (108), CD analysis indicated that PF1 did not form a G-quadruplex (Figure 3-9). It is still unclear yet how exactly the diguanosine repeats and the length played a role in terms of the conformation of the aptamer. The DNA mfold analysis of PF1 indicated that it has very low structure (Chapter 2), and RT tended to have a lower affinity towards unstructured nucleic acids (see below). However, it is possible that interaction between HIV-RT and PF1 may induce a change in the conformation of the aptamer leading to a higher secondary structure, or G-quadruplex formation. To evaluate if PF1 forms a more complex secondary structure (or a G-quadruplex) in the presence of HIV-RT, structural



analysis using NMR or X-ray crystallography would be needed. It is possible to measure the CD spectrum of PF1 with HIV-RT; however the presence of the protein would interfere with the spectrum, since both the nucleic acid and protein are measured in a similar range in terms of wavelength. This would make interpretations more difficult.

Consistent with previous results, both the G-quadruplex aptamers and the primer-template aptamers are potent inhibitors to HIV-RT extension *in vitro*. The inhibition is due to their ability to directly bind to the RT active site and block the access of the HIV-RT to the substrate. Even though PF1 had a relatively lower inhibitory effect compared to the other aptamers, it still showed 50% inhibition at a concentration approximately equal the substrate concentration in the inhibition reaction. This suggests that PF1 binds about as tightly as the primer-template used in the assays. Such tight binding would not normally be seen on a 30 nt single stranded DNA with very low predicted structure. It has been speculated that PF1 does contain some secondary structure, and the length of PF1 might be at the limit of the minimal length where a single-stranded DNA can bind tightly to HIV-RT. It would be interesting to use the same primer-free SELEX method and select for DNA aptamers to HIV-RT with different lengths (both longer and shorter than 30 nt). These aptamers could then be analyzed to see if they enhance binding to HIV-RT, and/or whether the same diguanosine repeats would still be selected.

The selection of single-stranded DNAs from this unique primer-free SELEX technique can be utilized in development of aptamers for other HIV enzymes, or for any protein targets. It has potential for development of small molecule-based therapy, especially with the coupling of a delivery system (see below). Aptamers developed using this method could also be used as diagnostic tools and in bioassays.

One main consideration however is that the *in vitro* analysis of the aptamer does not necessarily translate to the interaction between the oligo and the virus in a cellular system. In chapter 4 I focused on studying the effects of the primer-template aptamer developed by our lab (37 NT SELEX) on HIV-1 replication in cell culture. The 37 NT SELEX was modified with thiophosphate at the 5' end and the loop structure to prevent degradation of the oligo in the cell ("thioaptamer" in chapter 4). Previous studies have shown that highly structured nucleic acids have a tendency to activate innate cellular responses (70). Even though the designed thioaptamer and its structured controls did not decrease cell proliferation, the effects of the oligos on the host cell response might interfere with the HIV infection process (see below) (69, 162).

The thioaptamer and control aptamers with higher structure (Figure 4-1) all showed HIV inhibition in cell culture. The thioaptamer, while effective at curbing HIV replication, did probably did not achieve inhibition through the aptamer binding to HIV-RT as other structured controls that bound RT with much lower affinity were equally inhibitory. One of the possible steps that the aptamer inhibits HIV replication is at the entry stage, and there are compounds that block HIV-1 cell entry (180). The inhibition by structured DNA also pointed to the possibility that the thioaptamer blocks viral replication via the induction of specific cellular pathways, or other direct interactions with cellular proteins required for HIV replication. Double stranded nucleic acids are known to activate cellular pathways through Toll-like receptors (132), or to induce interferon production, which downregulates host factors induction, such as tetherin (BST-2) which prevents HIV-1 from budding out of the cell (8). Even though the aptamer at 2  $\mu\text{M}$  did not increase BST-2 induction, it did not rule out other cellular factors

contributing to the HIV inhibition. Some of the possible candidates for investigation included a more direct measurement of interferon levels (via ELISA against IFN- $\alpha$  or IFN- $\gamma$ ), or the caspase-3/7 assay to study on the effect of programmed cell death.

Cell-based analysis on PF1 and its variants will be the next logical step for investigating how PF1 affects HIV replication in cell culture. The lower structure and shorter oligo length provide some advantage for PF1 in future cell culture studies. For both PF1 and 37 NT SELEX aptamers, one of the barriers to overcome is the delivery of the oligo into the cell. An effective aptamer delivery into the cell lowers the aptamer dosage required for inhibition, and may reduce the cytotoxic effects and host responses. Vector-assisted delivery methods using liposome vesicles are one of the alternative methods for aptamer delivery (51, 104, 110, 116), where the nucleic acid enters the cells by fusion of the liposome vesicle and the cell membrane. Note that such delivery methods must not elicit any host response or be toxic to host cells. Recent reports have showed that DNA aptamers can be modified to form a self-assembled aptamer-micelle nanostructures, and have low cytotoxicity (187). However, the ability of the modified aptamer-micelle particles to inhibit HIV replications in cells has not been assessed.

Aptamers facilitate the understanding of HIV-RT properties, and its interaction with different types of nucleic acids. A diverse range of SELEX approaches provide the tools for generating different classes of aptamers, where each has their own unique features for various applications. With the *in vitro* and cell-based analysis, aptamer development aids the discovery of new and exciting ways to learn more about HIV functions and other diagnostic functions.

## Bibliography

1. **Abram, M. E., S. G. Sarafianos, and M. A. Parniak.** 2010. The mutation T477A in HIV-1 reverse transcriptase (RT) restores normal proteolytic processing of RT in virus with Gag-Pol mutated in the p51-RNH cleavage site. *Retrovirology* **7**:6.
2. **Adams, D.** 1979. *The Hitchhiker's Guide to the Galaxy*.
3. **Adamson, C. S., S. D. Ablan, I. Boeras, R. Goila-Gaur, F. Soheilian, K. Nagashima, F. Li, K. Salzwedel, M. Sakalian, C. T. Wild, and E. O. Freed.** 2006. In vitro resistance to the human immunodeficiency virus type 1 maturation inhibitor PA-457 (Bevirimat). *J Virol* **80**:10957-71.
4. **Aiken, C., J. Konner, N. R. Landau, M. E. Lenburg, and D. Trono.** 1994. Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. *Cell* **76**:853-64.
5. **Akira, S., S. Uematsu, and O. Takeuchi.** 2006. Pathogen recognition and innate immunity. *Cell* **124**:783-801.
6. **Alcorn, K.** 2009. Raltegravir shows potential for use as PrEP drug.
7. **Andreola, M. L., F. Pileur, C. Calmels, M. Ventura, L. Tarrago-Litvak, J. J. Toulme, and S. Litvak.** 2001. DNA aptamers selected against the HIV-1 RNase H display in vitro antiviral activity. *Biochemistry* **40**:10087-94.
8. **Andrew, A., and K. Strebel.** 2011. The interferon-inducible host factor bone marrow stromal antigen 2/tetherin restricts virion release, but is it actually a viral restriction factor? *J Interferon Cytokine Res* **31**:137-44.
9. **Ao, Z., X. Yao, and E. A. Cohen.** 2004. Assessment of the role of the central DNA flap in human immunodeficiency virus type 1 replication by using a single-cycle replication system. *J Virol* **78**:3170-7.
10. **AVERT.** 2011. HIV types, subtypes groups and strains.
11. **Baltimore, D.** 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **226**:1209-11.
12. **Baltimore, D., and D. F. Smoler.** 1972. Association of an endoribonuclease with the avian myeloblastosis virus deoxyribonucleic acid polymerase. *J Biol Chem* **247**:7282-7.
13. **Barin, F., M. F. McLane, J. S. Allan, T. H. Lee, J. E. Groopman, and M. Essex.** 1985. Virus envelope protein of HTLV-III represents major target antigen for antibodies in AIDS patients. *Science* **228**:1094-6.
14. **Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dautuet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**:868-71.
15. **Basu, V. P., M. Song, L. Gao, S. T. Rigby, M. N. Hanson, and R. A. Bambara.** 2008. Strand transfer events during HIV-1 reverse transcription. *Virus Res.*
16. **Battula, N., and L. A. Loeb.** 1976. On the fidelity of DNA replication. Lack of exodeoxyribonuclease activity and error-correcting function in avian myeloblastosis virus DNA polymerase. *J Biol Chem* **251**:982-6.

17. **Bebenek, K., J. Abbotts, S. H. Wilson, and T. A. Kunkel.** 1993. Error-prone polymerization by HIV-1 reverse transcriptase. *Journal of Biological Chemistry* **268**:10324-10334.
18. **Benimetskaya, L., J. L. Tonkinson, M. Koziolkiewicz, B. Karwowski, P. Guga, R. Zeltser, W. Stec, and C. A. Stein.** 1995. Binding of phosphorothioate oligodeoxynucleotides to basic fibroblast growth factor, recombinant soluble CD4, laminin and fibronectin is P-chirality independent. *Nucleic Acids Res* **23**:4239-45.
19. **Bernardi, F., E. Gaggelli, E. Molteni, E. Porciatti, D. Valensin, and G. Valensin.** 2006. <sup>1</sup>H and <sup>13</sup>C-NMR and molecular dynamics studies of cyclosporin a interacting with magnesium(II) or cerium(III) in acetonitrile. Conformational changes and cis-trans conversion of peptide bonds. *Biophys J* **90**:1350-61.
20. **Berridge, M. V., and A. S. Tan.** 1993. Characterization of the cellular reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): subcellular localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction. *Arch Biochem Biophys* **303**:474-82.
21. **Bock, L. C., L. C. Griffin, J. A. Latham, E. H. Vermaas, and J. J. Toole.** 1992. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **355**:564-6.
22. **Boyer, J. C., K. Bebenek, and T. A. Kunkel.** 1992. Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. *Proc Natl Acad Sci U S A* **89**:6919-23.
23. **Brierley, I., and F. J. Dos Ramos.** 2006. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* **119**:29-42.
24. **Brody, E. N., and L. Gold.** 2000. Aptamers as therapeutic and diagnostic agents. *J. Biotechnol.* **74**:5-13.
25. **Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop.** 1989. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc Natl Acad Sci U S A* **86**:2525-9.
26. **Cen, S., F. Guo, M. Niu, J. Saadatmand, J. Deflassieux, and L. Kleiman.** 2004. The interaction between HIV-1 Gag and APOBEC3G. *J Biol Chem.*
27. **Cerchia, L., F. Duconge, C. Pestourie, J. Boulay, Y. Aissouni, K. Gombert, B. Tavitian, V. de Franciscis, and D. Libri.** 2005. Neutralizing aptamers from whole-cell SELEX inhibit the RET receptor tyrosine kinase. *PLoS Biol* **3**:e123.
28. **Chen, H., and L. Gold.** 1994. Selection of high-affinity RNA ligands to reverse transcriptase: inhibition of cDNA synthesis and RNase H activity. *Biochemistry* **33**:8746-56.
29. **Cherry, E., C. Liang, L. Rong, Y. Quan, P. Inouye, X. Li, N. Morin, M. Kotler, and M. A. Wainberg.** 1998. Characterization of human immunodeficiency virus type-1 (HIV-1) particles that express protease-reverse transcriptase fusion proteins. *J Mol Biol* **284**:43-56.
30. **Clavel, F., H. M. D., R. L. Wiley, K. Strebel, M. A. Martin, and R. R.** 1989. Genetic Recombination of human immunodeficiency virus. *Journal of Virology* **63**:1455-1459.

31. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *Journal of General Virology* **42**:1-26.
32. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
33. **Cogoi, S., M. Ballico, G. M. Bonora, and L. E. Xodo.** 2004. Antiproliferative activity of a triplex-forming oligonucleotide recognizing a Ki-ras polypurine/polypyrimidine motif correlates with protein binding. *Cancer Gene Ther* **11**:465-76.
34. **Cogoi, S., F. Quadrifoglio, and L. E. Xodo.** 2004. G-rich oligonucleotide inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells. *Biochemistry* **43**:2512-23.
35. **Cory, A. H., T. C. Owen, J. A. Barltrop, and J. G. Cory.** 1991. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer Commun* **3**:207-12.
36. **Craigie, R.** 2001. HIV integrase, a brief overview from chemistry to therapeutics. *J Biol Chem* **276**:23213-6.
37. **D'Orso, I., J. R. Grunwell, R. L. Nakamura, C. Das, and A. D. Frankel.** 2008. Targeting tat inhibitors in the assembly of human immunodeficiency virus type 1 transcription complexes. *J Virol* **82**:9492-504.
38. **Dapas, B., G. Tell, A. Scaloni, A. Pines, L. Ferrara, F. Quadrifoglio, and B. Scaggiante.** 2003. Identification of different isoforms of eEF1A in the nuclear fraction of human T-lymphoblastic cancer cell line specifically binding to aptameric cytotoxic GT oligomers. *Eur J Biochem* **270**:3251-62.
39. **de Soultrait, V. R., P. Y. Lozach, R. Altmeyer, L. Tarrago-Litvak, S. Litvak, and M. L. Andreola.** 2002. DNA aptamers derived from HIV-1 RNase H inhibitors are strong anti-integrase agents. *J Mol Biol* **324**:195-203.
40. **Deisingh, A. K.** 2006. Aptamer-based biosensors: biomedical applications. *Handb Exp Pharmacol*:341-57.
41. **Dempsey, L. A., H. Sun, L. A. Hanakahi, and N. Maizels.** 1999. G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G-G pairing in immunoglobulin switch recombination. *J Biol Chem* **274**:1066-71.
42. **DeStefano, J. J., and J. V. Cristofaro.** 2006. Selection of primer-template sequences that bind human immunodeficiency virus reverse transcriptase with high affinity. *Nucleic Acids Res.* **34**:130-9.
43. **DeStefano, J. J., and G. R. Nair.** 2008. Novel aptamer inhibitors of human immunodeficiency virus reverse transcriptase. *Oligonucleotides* **18**:133-44.
44. **Diaz, R. M., A. Bateman, L. Emiliusen, A. Fielding, D. Trono, S. J. Russell, and R. G. Vile.** 2000. A lentiviral vector expressing a fusogenic glycoprotein for cancer gene therapy. *Gene Ther* **7**:1656-63.
45. **Dinman, J. D., S. Richter, E. P. Plant, R. C. Taylor, A. B. Hammell, and T. M. Rana.** 2002. The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. *Proc Natl Acad Sci U S A* **99**:5331-6.
46. **Dirac, A. M., H. Huthoff, J. Kjems, and B. Berkhout.** 2002. Requirements for RNA heterodimerization of the human immunodeficiency virus type 1 (HIV-1) and HIV-2 genomes. *J Gen Virol* **83**:2533-42.

47. **Domingo, E., E. Baranowski, C. M. Ruiz-Jarabo, A. M. Martin-Hernandez, J. C. Saiz, and C. Escarmis.** 1998. Quasispecies structure and persistence of RNA viruses. *Emerg Infect Dis* **4**:521-7.
48. **Duffalo, M. L., and C. W. James.** 2003. Enfuvirtide: a novel agent for the treatment of HIV-1 infection. *Ann Pharmacother* **37**:1448-56.
49. **Ellington, A. D., and J. W. Szostak.** 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**:818-22.
50. **FDA.** 2008. Drugs Used in the Treatment of HIV Infection.
51. **Ferguson, M. R., D. R. Rojo, A. Somasunderam, V. Thiviyanathan, B. D. Ridley, X. Yang, and D. G. Gorenstein.** 2006. Delivery of double-stranded DNA thioaptamers into HIV-1 infected cells for antiviral activity. *Biochem Biophys Res Commun* **344**:792-7.
52. **Fischl, M. A., D. D. Richman, M. H. Grieco, M. S. Gottlieb, P. A. Volberding, O. L. Laskin, J. M. Leedom, J. E. Groopman, D. Mildvan, R. T. Schooley, and et al.** 1987. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N Engl J Med* **317**:185-91.
53. **Fisher, T. S., P. Joshi, and V. R. Prasad.** 2002. Mutations that confer resistance to template-analog inhibitors of human immunodeficiency virus (HIV) type 1 reverse transcriptase lead to severe defects in HIV replication. *J Virol* **76**:4068-72.
54. **Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, and et al.** 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**:500-3.
55. **Goff, S. P.** 1990. Retroviral reverse transcriptase: synthesis, structure and function. *Acquired Immune Deficiency Syndromes* **3**:817-831.
56. **Gold, L.** 1995. The SELEX process: a surprising source of therapeutic and diagnostic compounds. *Harvey Lect* **91**:47-57.
57. **Goldschmidt, V., and R. Marquet.** 2004. Primer unblocking by HIV-1 reverse transcriptase and resistance to nucleoside RT inhibitors (NRTIs). *Int J Biochem Cell Biol* **36**:1687-705.
58. **Gonzalez, V., and L. H. Hurley.** 2010. The c-MYC NHE III(1): function and regulation. *Annu Rev Pharmacol Toxicol* **50**:111-29.
59. **Goodchild, A., A. King, M. M. Gozar, T. Passioura, C. Tucker, and L. Rivory.** 2007. Cytotoxic G-rich oligodeoxynucleotides: putative protein targets and required sequence motif. *Nucleic Acids Res* **35**:4562-72.
60. **Grossman, Z., M. Meier-Schellersheim, W. E. Paul, and L. J. Picker.** 2006. Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. *Nat Med* **12**:289-95.
61. **Guyader, M., M. Emerman, P. Sonigo, F. Clavel, L. Montagnier, and M. Alizon.** 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. *Nature* **326**:662-9.
62. **Hanakahi, L. A., H. Sun, and N. Maizels.** 1999. High affinity interactions of nucleolin with G-G-paired rDNA. *J Biol Chem* **274**:15908-12.

63. **Harada, K., and L. E. Orgel.** 1993. In vitro selection of optimal DNA substrates for T4 RNA ligase. *Proc Natl Acad Sci U S A* **90**:1576-9.
64. **Harris, D., R. L., H. S. Misra, P. K. Pandey, and V. N. Pandey.** 1998. The p51 subunit of human immunodeficiency virus type 1 reverse transcriptase is essential in loading the p66 subunit on the template. *Biochemistry* **37**:5903-5908.
65. **Held, D. M., J. D. Kissel, J. T. Patterson, D. G. Nickens, and D. H. Burke.** 2006. HIV-1 inactivation by nucleic acid aptamers. *Front. Biosci.* **11**:89-112.
66. **Hershman, S. G., Q. Chen, J. Y. Lee, M. L. Kozak, P. Yue, L. S. Wang, and F. B. Johnson.** 2008. Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **36**:144-56.
67. **Hicke, B. J., C. Marion, Y. F. Chang, T. Gould, C. K. Lynott, D. Parma, P. G. Schmidt, and S. Warren.** 2001. Tenascin-C aptamers are generated using tumor cells and purified protein. *J Biol Chem* **276**:48644-54.
68. **Hicke, B. J., A. W. Stephens, T. Gould, Y. F. Chang, C. K. Lynott, J. Heil, S. Borkowski, C. S. Hilger, G. Cook, S. Warren, and P. G. Schmidt.** 2006. Tumor targeting by an aptamer. *J Nucl Med* **47**:668-78.
69. **Hornung, V., A. Ablasser, M. Charrel-Dennis, F. Bauernfeind, G. Horvath, D. R. Caffrey, E. Latz, and K. A. Fitzgerald.** 2009. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458**:514-8.
70. **Hornung, V., and E. Latz.** Intracellular DNA recognition. *Nat Rev Immunol* **10**:123-30.
71. **Hotoda, H., M. Koizumi, R. Koga, M. Kaneko, K. Momota, T. Ohmine, H. Furukawa, T. Agatsuma, T. Nishigaki, J. Sone, S. Tsutsumi, T. Kosaka, K. Abe, S. Kimura, and K. Shimada.** 1998. Biologically active oligodeoxyribonucleotides. 5. 5'-End-substituted d(TGGGAG) possesses anti-human immunodeficiency virus type 1 activity by forming a G-quadruplex structure. *J Med Chem* **41**:3655-63.
72. **Hsieh, J. C., S. Zinnen, and P. Modrich.** 1993. Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J Biol Chem* **268**:24607-13.
73. **Hsiou, Y., J. Ding, K. Das, A. D. Clark, Jr., S. H. Hughes, and E. Arnold.** 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure* **4**:853-60.
74. **Hu, W.-S., and H. M. Temin.** 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci* **87**:1556-1560.
75. **Hu, W. S., and H. M. Temin.** 1990. Retroviral recombination and reverse transcription. *Science* **250**:1227-33.
76. **Hung, S. H., Q. Yu, D. M. Gray, and R. L. Ratliff.** 1994. Evidence from CD spectra that d(purine).r(pyrimidine) and r(purine).d(pyrimidine) hybrids are in different structural classes. *Nucleic Acids Res* **22**:4326-34.
77. **Hunter, E.** 1978. The mechanism for genetic recombination in the avian retroviruses. *Curr Top Microbiol Immunol* **79**:295-309.



78. **Ivetac, A., and J. A. McCammon.** 2009. Elucidating the inhibition mechanism of HIV-1 non-nucleoside reverse transcriptase inhibitors through multicopy molecular dynamics simulations. *J Mol Biol* **388**:644-58.
79. **Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus.** 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell* **55**:447-58.
80. **Jacobo-Molina, A., J. Ding, R. G. Nanni, A. D. Clark, Jr., X. Lu, C. Tantillo, R. L. Williams, G. Kamer, A. L. Ferris, P. Clark, and et al.** 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc Natl Acad Sci U S A* **90**:6320-4.
81. **Jaeger, J., T. Restle, and T. A. Steitz.** 1998. The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. *The EMBO Journal* **17**:4535-4542.
82. **James, W.** 2007. Aptamers in the virologists' toolkit. *J Gen Virol* **88**:351-64.
83. **Jarosch, F., K. Buchner, and S. Klussmann.** 2006. In vitro selection using a dual RNA library that allows primerless selection. *Nucleic Acids Res* **34**:e86.
84. **Jarver, P., and U. Langel.** 2004. The use of cell-penetrating peptides as a tool for gene regulation. *Drug Discov Today* **9**:395-402.
85. **Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J Virol* **74**:1234-1240.
86. **Jorgenson, R. L., V. M. Vogt, and M. C. Johnson.** 2009. Foreign glycoproteins can be actively recruited to virus assembly sites during pseudotyping. *J Virol* **83**:4060-7.
87. **Joshi, P., and V. R. Prasad.** 2002. Potent inhibition of human immunodeficiency virus type 1 replication by template analog reverse transcriptase inhibitors derived by SELEX (systematic evolution of ligands by exponential enrichment). *J. Virol.* **76**:6545-6557.
88. **Joshi, P. J., T. S. Fisher, and V. R. Prasad.** 2003. Anti-HIV inhibitors based on nucleic acids: emergence of aptamers as potent antivirals. *Curr Drug Targets Infect Disord* **3**:383-400.
89. **Joshi, P. J., T. W. North, and V. R. Prasad.** 2005. Aptamers directed to HIV-1 reverse transcriptase display greater efficacy over small hairpin RNAs targeted to viral RNA in blocking HIV-1 replication. *Mol Ther* **11**:677-86.
90. **Joyce, G. F.** 1989. Amplification, mutation and selection of catalytic RNA. *Gene* **82**:83-7.
91. **Kahn, J. O., and B. D. Walker.** 1998. Acute human immunodeficiency virus type 1 infection. *N Engl J Med* **339**:33-9.
92. **Kankia, B. I., G. Barany, and K. Musier-Forsyth.** 2005. Unfolding of DNA quadruplexes induced by HIV-1 nucleocapsid protein. *Nucleic Acids Res* **33**:4395-403.
93. **Karacostas, V., E. J. Wolffe, K. Nagashima, M. A. Gonda, and B. Moss.** 1993. Overexpression of the HIV-1 gag-pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. *Virology* **193**:661-71.

94. **Katz, R. A., and A. M. Skalka.** 1990. Generation of diversity in retroviruses. *Annu Rev Genet* **24**:409-45.
95. **Kelley, S., S. Boroda, K. Musier-Forsyth, and B. I. Kankia.** HIV-integrase aptamer folds into a parallel quadruplex: A thermodynamic study. *Biophys Chem.*
96. **Kensch, O., B. A. Connolly, H. J. Steinhoff, A. McGregor, R. S. Goody, and T. Restle.** 2000. HIV-1 reverse transcriptase-pseudoknot RNA aptamer interaction has a binding affinity in the low picomolar range coupled with high specificity. *J. Biol. Chem.* **275**:18271-18278.
97. **Kielian, M., and S. Jungerwirth.** 1990. Mechanisms of enveloped virus entry into cells. *Mol Biol Med* **7**:17-31.
98. **Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz.** 1992. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**:1783-90.
99. **Kolb, G., S. Reigadas, D. Castanotto, A. Faure, M. Ventura, J. J. Rossi, and J. J. Toulme.** 2006. Endogenous expression of an anti-TAR aptamer reduces HIV-1 replication. *RNA Biol* **3**:150-6.
100. **Kupzig, S., V. Korolchuk, R. Rollason, A. Sugden, A. Wilde, and G. Banting.** 2003. Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. *Traffic* **4**:694-709.
101. **Kypr, J., I. Kejnovska, D. Renciuik, and M. Vorlickova.** 2009. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res* **37**:1713-25.
102. **Lai, Y. T., and J. J. Destefano.** 2011. A primer-free method that selects high-affinity single-stranded DNA aptamers using thermostable RNA ligase. *Anal Biochem* **414**:246-53.
103. **Lang, W., H. Perkins, R. E. Anderson, R. Royce, N. Jewell, and W. Winkelstein, Jr.** 1989. Patterns of T lymphocyte changes with human immunodeficiency virus infection: from seroconversion to the development of AIDS. *J Acquir Immune Defic Syndr* **2**:63-9.
104. **Lavigne, C., J. Yelle, G. Sauve, and A. G. Thierry.** 2001. Lipid-based delivery of combinations of antisense oligodeoxynucleotides for the in vitro inhibition of HIV-1 replication. *AAPS PharmSci* **3**:E7.
105. **Le Grice, S. F., T. Naas, B. Wohlgensinger, and O. Schatz.** 1991. Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase. *Embo J* **10**:3905-11.
106. **Levin, J. G., J. Guo, I. Rouzina, and K. Musier-Forsyth.** 2005. Nucleic Acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acids Res. Mol. Biol.* **80**:217-286.
107. **Levy, D. N., G. M. Aldrovandi, O. Kutsch, and G. M. Shaw.** 2004. Dynamics of HIV-1 recombination in its natural target cells. *Proc Natl Acad Sci U S A* **101**:4204-9.
108. **Macaya, R. F., P. Schultze, F. W. Smith, J. A. Roe, and J. Feigon.** 1993. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc Natl Acad Sci U S A* **90**:3745-9.

109. **Matsugami, A., H. Tochio, E. Niyada, Y. Tamura, M. Kudo, T. S. Misono, P. Kumar, and M. Katahira.** 2005. Structure of RNA aptamer complexed with an RNA-binding peptide of Tat with aid of residue-specific <sup>13</sup>C, <sup>15</sup>N labeling. *Nucleic Acids Symp Ser (Oxf)*:69-70.
110. **Matzen, K., L. Elzaouk, A. A. Matskevich, A. Nitzsche, J. Heinrich, and K. Moelling.** 2007. RNase H-mediated retrovirus destruction in vivo triggered by oligodeoxynucleotides. *Nat Biotechnol* **25**:669-674.
111. **Metifiot, M., O. Leon, L. Tarrago-Litvak, S. Litvak, and M. L. Andreola.** 2005. Targeting HIV-1 integrase with aptamers selected against the purified RNase H domain of HIV-1 RT. *Biochimie* **87**:911-9.
112. **Michalowski, D., R. Chitima-Matsiga, D. M. Held, and D. H. Burke.** 2008. Novel bimodular DNA aptamers with guanosine quadruplexes inhibit phylogenetically diverse HIV-1 reverse transcriptases. *Nucleic Acids Res* **36**:7124-35.
113. **Missailidis, S., and A. Perkins.** 2007. Update: aptamers as novel radiopharmaceuticals: their applications and future prospects in diagnosis and therapy. *Cancer Biother Radiopharm* **22**:453-68.
114. **Miyagi, E., A. J. Andrew, S. Kao, and K. Strebel.** 2009. Vpu enhances HIV-1 virus release in the absence of Bst-2 cell surface down-modulation and intracellular depletion. *Proc Natl Acad Sci U S A* **106**:2868-73.
115. **Mizrahi, V.** 1989. Analysis of the ribonuclease H activity of HIV-1 reverse transcriptase using RNA-DNA hybrid substrates derived from the gag region of HIV-1. *Biochemistry* **28**:9088-94.
116. **Moelling, K., S. Abels, J. Jendis, A. Matskevich, and J. Heinrich.** 2006. Silencing of HIV by hairpin-loop-structured DNA oligonucleotide. *FEBS Lett* **580**:3545-50.
117. **Mok, W., and Y. Li.** 2008. Recent progress in nucleic acid aptamer-based biosensors and bioassays. *Sensors* **8**:7050-7084.
118. **Mosing, R. K., S. D. Mendonsa, and M. T. Bowser.** 2005. Capillary electrophoresis-SELEX selection of aptamers with affinity for HIV-1 reverse transcriptase. *Anal Chem* **77**:6107-12.
119. **Negrone, M., and H. Buc.** 2001. Retroviral recombination: what drives the switch? *Nat. Rev. Mol. Cell Biol.* **2**:151-5.
120. **Neil, S. J., T. Zang, and P. D. Bieniasz.** 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**:425-30.
121. **Nimjee, S. M., C. P. Rusconi, and B. A. Sullenger.** 2005. Aptamers: an emerging class of therapeutics. *Annu Rev Med* **56**:555-83.
122. **Onafuwa, A., W. An, N. D. Robson, and A. Telesnitsky.** 2003. Human Immunodeficiency Virus Type 1 Genetic Recombination Is More Frequent Than That of Moloney Murine Leukemia Virus despite Similar Template Switching Rates. *J. Virol.* **77**:4577-4587.
123. **Palaniappan, C., P. J. Fay, and R. A. Bambara.** 1995. Nevirapine alters the cleavage specificity of ribonuclease H of human immunodeficiency virus 1 reverse transcriptase. *J Biol Chem* **270**:4861-9.
124. **Palmer, S., A. P. Wiegand, F. Maldarelli, H. Bazmi, J. M. Mican, M. Polis, R. L. Dewar, A. Planta, S. Liu, J. A. Metcalf, J. W. Mellors, and J. M. Coffin.**

2003. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol* **41**:4531-6.
125. **Pan, W., and G. A. Clawson.** Primer-free aptamer selection using a random DNA library. *Methods Mol Biol* **629**:369-85.
126. **Pan, W., and G. A. Clawson.** 2010. Primer-free aptamer selection using a random DNA library. *Methods Mol Biol* **629**:369-85.
127. **Pan, W., P. Xin, and G. A. Clawson.** 2008. Minimal primer and primer-free SELEX protocols for selection of aptamers from random DNA libraries. *Biotechniques* **44**:351-60.
128. **Pantaleo, G., J. F. Demarest, T. Schacker, M. Vaccarezza, O. J. Cohen, M. Daucher, C. Graziosi, S. S. Schnittman, T. C. Quinn, G. M. Shaw, L. Perrin, G. Tambussi, A. Lazzarin, R. P. Sekaly, H. Soudeyns, L. Corey, and A. S. Fauci.** 1997. The qualitative nature of the primary immune response to HIV infection is a prognosticator of disease progression independent of the initial level of plasma viremia. *Proc Natl Acad Sci U S A* **94**:254-8.
129. **Park, J., and C. D. Morrow.** 1991. Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. *J Virol* **65**:5111-7.
130. **Patil, S. D., D. G. Rhodes, and D. J. Burgess.** 2005. DNA-based therapeutics and DNA delivery systems: a comprehensive review. *Aaps J* **7**:E61-77.
131. **Pendergrast, P. S., H. N. Marsh, D. Grate, J. M. Healy, and M. Stanton.** 2005. Nucleic acid aptamers for target validation and therapeutic applications. *J Biomol Tech* **16**:224-34.
132. **Peter, M., K. Bode, G. B. Lipford, F. Eberle, K. Heeg, and A. H. Dalpke.** 2008. Characterization of suppressive oligodeoxynucleotides that inhibit Toll-like receptor-9-mediated activation of innate immunity. *Immunology* **123**:118-28.
133. **Pieve, C. D., A. C. Perkins, and S. Missailidis.** 2009. Anti-MUC1 aptamers: radiolabelling with (99m)Tc and biodistribution in MCF-7 tumour-bearing mice. *Nucl Med Biol* **36**:703-10.
134. **Plantier, J. C., M. Leoz, J. E. Dickerson, F. De Oliveira, F. Cordonnier, V. Leme, F. Damond, D. L. Robertson, and F. Simon.** 2009. A new human immunodeficiency virus derived from gorillas. *Nat Med* **15**:871-2.
135. **Pollard, V. W., and M. H. Malim.** 1998. The HIV-1 Rev protein. *Annu Rev Microbiol* **52**:491-532.
136. **Porschewski, P., M. A. Grattinger, K. Klenzke, A. Erpenbach, M. R. Blind, and F. Schafer.** 2006. Using aptamers as capture reagents in bead-based assay systems for diagnostics and hit identification. *J Biomol Screen* **11**:773-81.
137. **Proske, D., M. Blank, R. Buhmann, and A. Resch.** 2005. Aptamers--basic research, drug development, and clinical applications. *Appl Microbiol Biotechnol* **69**:367-74.
138. **Ramirez, B. C., E. Simon-Loriere, R. Galetto, and M. Negroni.** 2008. Implications of recombination for HIV diversity. *Virus Res* **134**:64-73.
139. **Repaske, R., J. W. Hartley, M. F. Kavlick, R. R. O'Neill, and J. B. Austin.** 1989. Inhibition of RNase H activity and viral replication by single mutations in

- the 3' region of Moloney murine leukemia virus reverse transcriptase. *J Virol* **63**:1460-4.
140. **Richard, E., M. Mendez, F. Mazurier, C. Morel, P. Costet, P. Xia, A. Fontanellas, F. Geronimi, M. Cario-Andre, L. Taine, C. Ged, P. Malik, H. de Verneuill, and F. Moreau-Gaudry.** 2001. Gene therapy of a mouse model of protoporphyria with a self-inactivating erythroid-specific lentiviral vector without preselection. *Mol Ther* **4**:331-8.
  141. **Riou, J. F., L. Guittat, P. Mailliet, A. Laoui, E. Renou, O. Petitgenet, F. Megnin-Chanet, C. Helene, and J. L. Mergny.** 2002. Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proc Natl Acad Sci U S A* **99**:2672-7.
  142. **Risitano, A., and K. R. Fox.** 2004. Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res* **32**:2598-606.
  143. **Russo, G., M. Cuccurese, G. Monti, A. Russo, A. Amoresano, P. Pucci, and C. Pietropaolo.** 2005. Ribosomal protein L7a binds RNA through two distinct RNA-binding domains. *Biochem J* **385**:289-99.
  144. **Sakurai, Y., K. Komatsu, K. Agematsu, and M. Matsuoka.** 2009. DNA double strand break repair enzymes function at multiple steps in retroviral infection. *Retrovirology* **6**:114.
  145. **Sambrook, J., and D. W. Russell.** 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
  146. **Sattentau, Q. J., and J. P. Moore.** 1993. The role of CD4 in HIV binding and entry. *Philos Trans R Soc Lond B Biol Sci* **342**:59-66.
  147. **Sattentau, Q. J., J. P. Moore, F. Vignaux, F. Traincard, and P. Poignard.** 1993. Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. *J Virol* **67**:7383-93.
  148. **Scaggiante, B., C. Morassutti, B. Dapas, G. Tolazzi, F. Ustulin, and F. Quadrifoglio.** 1998. Human cancer cell lines growth inhibition by GTn oligodeoxyribonucleotides recognizing single-stranded DNA-binding proteins. *Eur J Biochem* **252**:207-15.
  149. **Schneider, D. J., J. Feigon, Z. Hostomsky, and L. Gold.** 1995. High-affinity ssDNA inhibitors of the reverse transcriptase of type 1 human immunodeficiency virus. *Biochemistry* **34**:9599-610.
  150. **Schneider, E., S. Whitmore, K. M. Glynn, K. Dominguez, A. Mitsch, and M. T. McKenna.** 2008. Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged <18 months and for HIV infection and AIDS among children aged 18 months to <13 years--United States, 2008. *MMWR Recomm Rep* **57**:1-12.
  151. **Schopman, N. C., O. ter Brake, and B. Berkhout.** 2010. Anticipating and blocking HIV-1 escape by second generation antiviral shRNAs. *Retrovirology* **7**:52.
  152. **Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**:521-9.

153. **Shehu-Xhilaga, M., H. G. Kraeusslich, S. Pettit, R. Swanstrom, J. Y. Lee, J. A. Marshall, S. M. Crowe, and J. Mak.** 2001. Proteolytic processing of the p2/nucleocapsid cleavage site is critical for human immunodeficiency virus type 1 RNA dimer maturation. *J Virol* **75**:9156-64.
154. **Shimura, K., E. Kodama, Y. Sakagami, Y. Matsuzaki, W. Watanabe, K. Yamataka, Y. Watanabe, Y. Ohata, S. Doi, M. Sato, M. Kano, S. Ikeda, and M. Matsuoka.** 2008. Broad antiretroviral activity and resistance profile of the novel human immunodeficiency virus integrase inhibitor elvitegravir (JTK-303/GS-9137). *J Virol* **82**:764-74.
155. **Shtatland, T., S. C. Gill, B. E. Javornik, H. E. Johansson, B. S. Singer, O. C. Uhlenbeck, D. A. Zichi, and L. Gold.** 2000. Interactions of Escherichia coli RNA with bacteriophage MS2 coat protein: genomic SELEX. *Nucleic Acids Res* **28**:E93.
156. **Simonsson, T.** 2001. G-quadruplex DNA structures--variations on a theme. *Biol Chem* **382**:621-8.
157. **Sissi, C., B. Gatto, and M. Palumbo.** 2011. The evolving world of protein-G-quadruplex recognition: A medicinal chemist's perspective. *Biochimie*.
158. **Skasko, M., A. Tokarev, C. C. Chen, W. B. Fischer, S. K. Pillai, and J. Guatelli.** 2011. BST-2 is rapidly down-regulated from the cell surface by the HIV-1 protein Vpu: evidence for a post-ER mechanism of Vpu-action. *Virology* **411**:65-77.
159. **St Louis, D. C., D. Gotte, E. Sanders-Buell, D. W. Ritchey, M. O. Salminen, J. K. Carr, and F. E. McCutchan.** 1998. Infectious molecular clones with the nonhomologous dimer initiation sequences found in different subtypes of human immunodeficiency virus type 1 can recombine and initiate a spreading infection in vitro. *J Virol* **72**:3991-8.
160. **Starnes, M. C., and Y. C. Cheng.** 1989. Human immunodeficiency virus reverse transcriptase-associated RNase H activity. *J Biol Chem* **264**:7073-7.
161. **Swanson, C. M., and M. H. Malim.** 2008. SnapShot: HIV-1 proteins. *Cell* **133**:742, 742 e1.
162. **Takeshita, F., and K. J. Ishii.** 2008. Intracellular DNA sensors in immunity. *Curr Opin Immunol* **20**:383-8.
163. **Tang, Z., D. Shangguan, K. Wang, H. Shi, K. Sefah, P. Mallikratchy, H. W. Chen, Y. Li, and W. Tan.** 2007. Selection of aptamers for molecular recognition and characterization of cancer cells. *Anal Chem* **79**:4900-7.
164. **Tasara, T., M. Amacker, and U. Hubscher.** 1999. Intramolecular chimeras of the p51 subunit between HIV-1 and FIV reverse transcriptases suggest a stabilizing function for the p66 subunit in the heterodimeric enzyme. *Biochemistry* **38**:1633-42.
165. **Temin, H. M.** 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc Natl Acad Sci U S A* **90**:6900-3.
166. **Temin, H. M., C. Y. Kang, and S. Mizutani.** 1973. Endogenous RNA-directed DNA polymerase activity in normal cells. *Johns Hopkins Med J Suppl* **2**:141-56.
167. **Temin, H. M., and S. Mizutani.** 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226**:1211-3.

168. **Tessier, D. C., R. Brousseau, and T. Vernet.** 1986. Ligation of single-stranded oligodeoxyribonucleotides by T4 RNA ligase. *Analytical Biochemistry* **158**:171-178.
169. **Tohl, J., and W. Eimer.** 1997. Interaction of a G-DNA quadruplex with mono- and divalent cations. A force field calculation. *Biophys Chem* **67**:177-86.
170. **Tong, W., C.-D. Lu, S. K. Sharma, S. Matsuura, A. G. So, and W. A. Scott.** 1997. Nucleotide-Induced Stable Complex Formation by HIV-1 Reverse Transcriptase. *Biochemistry* **36**:5749-5757.
171. **Tripathi, S., B. Chaubey, B. E. Barton, and V. N. Pandey.** 2007. Anti HIV-1 virucidal activity of polyamide nucleic acid-membrane transducing peptide conjugates targeted to primer binding site of HIV-1 genome. *Virology*.
172. **Tuerk, C., and L. Gold.** 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**:505-10.
173. **Tuerk, C., S. MacDougall, and L. Gold.** 1992. RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proc Natl Acad Sci U S A* **89**:6988-92.
174. **Turner, J. J., S. Jones, M. M. Fabani, G. Ivanova, A. A. Arzumanov, and M. J. Gait.** 2007. RNA targeting with peptide conjugates of oligonucleotides, siRNA and PNA. *Blood Cells Mol Dis* **38**:1-7.
175. **UNAIDS.** 2010. Global report: UNAIDS report on the global AIDS epidemic 2010.
176. **Van Damme, N., D. Goff, C. Katsura, R. L. Jorgenson, R. Mitchell, M. C. Johnson, E. B. Stephens, and J. Guatelli.** 2008. The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* **3**:245-52.
177. **Vater, A., and S. Klussmann.** 2003. Toward third-generation aptamers: Spiegelmers and their therapeutic prospects. *Curr Opin Drug Discov Devel* **6**:253-61.
178. **Wagner, H.** 2008. The sweetness of the DNA backbone drives Toll-like receptor 9. *Curr Opin Immunol* **20**:396-400.
179. **Wagner, H., and S. Bauer.** 2006. All is not Toll: new pathways in DNA recognition. *J Exp Med* **203**:265-8.
180. **Waheed, A. A., S. D. Ablan, M. K. Mankowski, J. E. Cummins, R. G. Ptak, C. P. Schaffner, and E. O. Freed.** 2006. Inhibition of HIV-1 replication by amphotericin B methyl ester: selection for resistant variants. *J Biol Chem* **281**:28699-711.
181. **Waheed, A. A., and E. O. Freed.** 2009. Lipids and membrane microdomains in HIV-1 replication. *Virus Res* **143**:162-76.
182. **Wain-Hobson, S.** 1992. Human immunodeficiency virus type 1 quasispecies in vivo and ex vivo. *Curr Top Microbiol Immunol* **176**:181-93.
183. **Wang, J., H. Jiang, and F. Liu.** 2000. In vitro selection of novel RNA ligands that bind human cytomegalovirus and block viral infection. *RNA* **6**:571-83.
184. **Wen, J. D., and D. M. Gray.** 2004. Ff gene 5 single-stranded DNA-binding protein assembles on nucleotides constrained by a DNA hairpin. *Biochemistry* **43**:2622-34.

185. **Willey, R. L., F. Maldarelli, M. A. Martin, and K. Strebel.** 1992. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J Virol* **66**:7193-200.
186. **Wolinsky, S. M., B. T. Korber, A. U. Neumann, M. Daniels, K. J. Kunstman, A. J. Whetsell, M. R. Furtado, Y. Cao, D. D. Ho, and J. T. Safrit.** 1996. Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection [see comments]. *Science* **272**:537-42.
187. **Wu, Y., K. Sefah, H. Liu, R. Wang, and W. Tan.** 2010. DNA aptamer-micelle as an efficient detection/delivery vehicle toward cancer cells. *Proc Natl Acad Sci U S A* **107**:5-10.
188. **Yu, H., and M. F. Goodman.** 1992. Comparison of HIV-1 and Avian Myeloblastosis Virus Reverse Transcriptase Fidelity on RNA and DNA Templates. *J. Biol. Chem.* **267**:10888-10896.
189. **Zennou, V., C. Petit, D. Guetard, U. Nerhbass, L. Montagnier, and P. Charneau.** 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101**:173-85.
190. **Zhang, J., and H. M. Temin.** 1993. Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science* **259**:234-238.
191. **Zhang, Z., M. Blank, and H. J. Schluesener.** 2004. Nucleic acid aptamers in human viral disease. *Arch Immunol Ther Exp (Warsz)* **52**:307-15.
192. **Zhou, J., P. Swiderski, H. Li, J. Zhang, C. P. Neff, R. Akkina, and J. J. Rossi.** 2009. Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells. *Nucleic Acids Res* **37**:3094-109.
193. **Zhou, J., X. Yuan, D. Dismuke, B. M. Forshey, C. Lundquist, K. H. Lee, C. Aiken, and C. H. Chen.** 2004. Small-molecule inhibition of human immunodeficiency virus type 1 replication by specific targeting of the final step of virion maturation. *J Virol* **78**:922-9.
194. **Zhu, T., L. Corey, Y. Hwangbo, J. M. Lee, G. H. Learn, J. I. Mullins, and M. J. McElrath.** 2003. Persistence of extraordinarily low levels of genetically homogeneous human immunodeficiency virus type 1 in exposed seronegative individuals. *J Virol* **77**:6108-16.
195. **Zhuang, J., A. E. Jetzt, G. Sun, H. Yu, G. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2002. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J Virol* **76**:11273-82.
196. **Zhuang, J., S. Mukherjee, Y. Ron, and J. P. Dougherty.** 2006. High rate of genetic recombination in murine leukemia virus: implications for influencing proviral ploidy. *J Virol* **80**:6706-11.
197. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**:3406-15.