ABSTRACT

Title of Document: **MODULATION OF HIV-1 REVERSE TRANSCRIPTASE AND FAMILY A DNA POLYMERASE PRIMER-TEMPLATE BINDING**

Katherine Joan Fenstermacher, Doctor of Philosophy 2014

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

Polymerases are enzymes used by all cellular and viral organisms to replicate their genomes. The human immunodeficiency virus (HIV) polymerase, reverse transcriptase (RT), uses a single-stranded RNA template to create double-stranded DNA during the course of the viral life cycle. Successful reverse transcription relies on the speed of catalysis and the ability of the enzyme to stay bound to the template during synthesis. I demonstrate that both of these properties can be modulated by the presence of different divalent cations, fundamentally altering the behavior of HIV RT. In the presence of 2 mM Mg^{2+}, the HIV RT primer-template complex has a half-life of 1.7±1.0 min, incorporating nucleotides at a maximum rate of 3.5 nucleotides (nt) per second (average speed 1.4±0.4 nt/sec). Substituting 2 mM Mg^{2+} with 400 μM Zn^{2+} dramatically slows the speed of catalysis (maximum 0.1 nt/sec, average

0.022±0.003 nt/sec) and promotes primer-template complexes that last hours (half-life of 220±60 min). These profound changes to the enzyme's function critically inhibit reverse transcription, even in the presence of optimal concentrations of Mg^{2+}. In addition to the cation composition during reverse transcription, previous studies have demonstrated that the sequence of the primer-template substrate can also affect the duration of a RT-primer-template complex. In light of this discovery, I investigated the tendency of two Family A DNA polymerases, the *Thermus aquaticus* DNA polymerase (*Taq* pol) and the Klenow fragment from *Escherichia coli* DNA polymerase I (Klenow), to selectively and tightly bind primer-template complexes. Using Primer-Template Systematic Evolution of Ligands by Exponential Enrichment (PT SELEX), I determined that both *Taq* pol and Klenow tightly bind to sequences containing regions that match the initiation and melting domains of promoters for the structurally similar bacteriophage T7-like RNA polymerases. This suggests a shared sequence preference that might be present in all Family A DNA polymerases, derived from a common ancestor. I plan to exploit this primer-template binding preference to advance biotechnologies utilizing these enzymes.

MODULATION OF HIV-1 REVERSE TRANSCRIPTASE AND FAMILY A DNA POLYMERASE PRIMER-TEMPLATE BINDING.

By

Katherine Joan Fenstermacher

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
2014

Advisory Committee:
Professor Jeffrey J. DeStefano, Chair
Professor Anne E. Simon
Professor James N. Culver
Professor Eric O. Freed
Professor Siba K. Samal, Dean's Representative

Dedication

This dissertation is dedicated to my mother, father, sister, and brother-in-law whose love and guidance made all of this possible.

Acknowledgements

When I applied to graduate school, I thought I knew what I was getting into: hard work, sleepless nights, and mountains of stress. What I couldn't predict were the amazing people I would meet along the way, who have kept me smiling through what has been both the most difficult and rewarding experience of my life.

First and foremost, I want to thank my advisor, Dr. Jeffrey DeStefano, for his time, encouragement, knowledge, and for providing me with the space to grow as an independent researcher. Jeff's immense patience has made it possible for me to continue through the periods of failed experiments and disappointing data that sometimes accompany research. In addition to discussing papers, grants, and data, I'll miss talking to him about music, philosophy, and (especially) politics. Joining his lab was one of the best decisions I've ever made.

I would also like to express my gratitude to my committee members, Drs. Anne Simon, Eric Freed, Jim Culver, and Siba Samal, for their feedback, support, and constructive criticism throughout this entire process. Your involvement has been invaluable.

In addition to the faculty who have played a role in my academic career, the members of my lab have shaped the day-to-day tenor of my research life. I can't thank my previous lab mates enough for their guidance and friendship: Deena Jacob, who helped me feel at home in the initially scary world of Real Science; Megan Lai, my cell culture buddy—Trogdor the Freezerburninator misses you!; Jeff Olimpo, for

a billion laughs and even more assists reaching high shelves; Gauri Nair, my lab mother, for her advice, food, and friendship; Divya Gangaramani, for Hindi lessons, conference entertainment, and being a wonderful and kind friend (I swear, I'll make those antibiotic stocks tomorrow. Really, I totally will). Of course, I can't forget my current lab mate, honorary little brother, and prank war adversary Vasu Achuthan: I'm going to miss you, Guppy. Thank you for making my last couple of years here so much fun.

Outside of the lab, I've been fortunate to meet an amazing group of people. To the former (and honorary) Pirate House roommates--Brian Kimble, Douglas Ishii, Erin (and Ben!) Anderson, and Ted Gibbons--you've kept me sane, or at least made me feel so by comparison. Thank you for the TV nights, group Halloween costumes, strangely elaborate board games, laughter, naps, five-hour discussions on what constitutes the *best* sandwich, and overall amazing friendship. Thanks as well to Jen Shipley, Susan Ochsner, and Shahnawaz Zaheer for countless meals, laughs, bad movie marathons, and general tomfoolery. I'm so incredibly lucky to have met you all of you.

A profound thank you to my parents, Bill and Ruth Fenstermacher, for always being supportive of my education. Dad, I guess this is the time to confess: when I was a kid, I used to ask you random questions at dinner hoping that you'd be so distracted answering me that I wouldn't have to finish my vegetables. It worked, but I ended up accidentally learning a lot on along the way. I still don't understand what 'offsides' means in football, but you taught me the importance of being inquisitive

and that there's a solution to every problem, if you just think hard enough. Mom, my fellow MBC chemistry alumna, thank you for giving me practice as a graduate student before I even knew such a thing existed—getting to hang out with you at work in the hospital pharmacy was secretly a pretty good approximation of research life: hard work, odd hours, no sun, unpronounceable chemicals, and the ever-present scent of alcohol. You had a much snazzier lab coat though!

And finally, thank you to my sister Ashley for blazing this Ph.D. path well before I did. I'm sure I got asked, "When are you graduating?" only a third as many times as I would have otherwise. You and your husband Rick have provided me with great advice, weekends away from the lab, jogging company, and—let's be honest—enough wine to make it through this whole process. Cheers!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

aa	-Amino acid(s)
AIDS	-Acquired Immune Deficiency Syndrome
AMV	-Avian myeloblastosis virus
ARV	-Anti-retroviral
bp	-Base pair(s)
BSA	-Bovine serum albumin
dNTPs	-Deoxyribonucleotide triphophates
ds	-Double-stranded
DTT	-Dithiothreitol
EDTA	-Ethylenediaminetetraacetic acid
exo	-Exonuclease
HAART	-Highly Active Anti-Retroviral Therapy
HIV	-Human Immunodeficiency Virus
hr	-Hour
Kb	-Kilobase(s)
K_d	-Dissociation constant
kDa	-Kilo Daltons
LTR	-Long terminal repeat
Min	-Minutes
MuLV	-Moloney murine leukemia virus
nt	-Nucleotide(s)
PBS	-Primer binding site

PCR	-Polymerase chain reaction
PNK	-Polynucleotide kinase
Pol	-Polymerase
PPT	-Polypurine tract
PR	-Protease
PT	-Primer-template
RNAP	-RNA polymerase
RNase H	-Ribonuclease H
RT	-Reverse Transcriptase
sec	-Second
SELEX	-Selective Enrichment of Ligands by Exponential Enrichment
SMF	-Small Molecular Fraction
SIV	-Simian Immunodeficiency Virus
ss	-Single-stranded
Taq	-*Thermus aquaticus*

Chapter 1: HIV and RT

1.1 HIV & AIDS

1.1.1 General history

During the late 1970s, doctors began to observe otherwise healthy patients suffering from diseases typically associated with immunocompromised individuals, such as transplant recipients (1, 2). Initially seen only in the gay community, the syndrome was first called Gay Related Immune Deficiency (GRID). However, new patient populations were soon discovered, such as blood transfusion recipients (3) and IV drug users (4), necessitating a more accurate name to encompass the full subset of individuals with symptoms: Acquired Immune Deficiency Syndrome (AIDS). Many theories were proposed to explain the disease, some of which included drug abuse, lifestyle factors, or an unknown pathogen. In 1983, Françoise Barré-Sinoussi and Luc Montagnier's group isolated the causative agent of AIDS from the lymph nodes of a patient and named it the Lymphadenopathy-Associated Virus (LAV) (5). Subsequently, Robert Gallo's team purified and characterized the same virus, naming it Human T-cell Lymphotrophic Virus III (HTLV-III) (6). Both groups share credit for the discovery of what was later renamed the Human Immunodeficiency Virus (HIV) and are responsible for determining that this novel complex retrovirus was responsible for the immune destruction that led to AIDS. While a death sentence in the 1980s, decades of research have yielded multiple therapeutics, which allow most HIV patients to control their infections and achieve lifespans approaching normal. Unfortunately, a cure remains elusive and treatment must be continued indefinitely.

Currently, an estimated 35 million are infected worldwide (0.8% of the global adult population), of whom 2 million died in 2012 alone (7). In countries that have promoted safe sex initiatives and have the infrastructure to provide treatment access, infection rates are tapering off. However, many areas of the world are unable to provide adequate treatment or prophylactic information, resulting in infection rates approaching 25% of the adult population in some parts of Sub-Saharan Africa (7).

1.1.2 Classification and diversity

HIV is a lentivirus ("slow virus") of the family *Retroviridae*. It is closely related to simian immunodeficiency virus (SIV), which infects many Old World primate species and is responsible for infections ranging from asymptomatic to causing AIDS-like illness. HIV itself consists of two discrete viruses, HIV-1 and HIV-2, that share a 60% nucleotide sequence identity (8). HIV-1 is largely responsible for the global pandemic and is broken down into four groups: M (major), O (outlier), N (non-M, non-O), and the most recently identified group P. Groups M, N, and O are most closely related to strains of SIV from chimpanzees while group P is most similar to SIV from gorillas (Figure 1) (9). Group M, which is responsible for >90% of all global HIV infections, can be further broken down into several subtypes, as well as intersubtype recombinants referred to as circulating recombinant forms (CRFs) (Figure 2) (10). Globally, subtype C is the most prevalent but the predominant subtype varies by country (for instance, American and European HIV infections are primarily subtype B). In contrast to HIV-1, HIV-2 is mostly limited to Central and West Africa and is likely derived from a zoonotic event with SIV from sooty mangabeys (11).

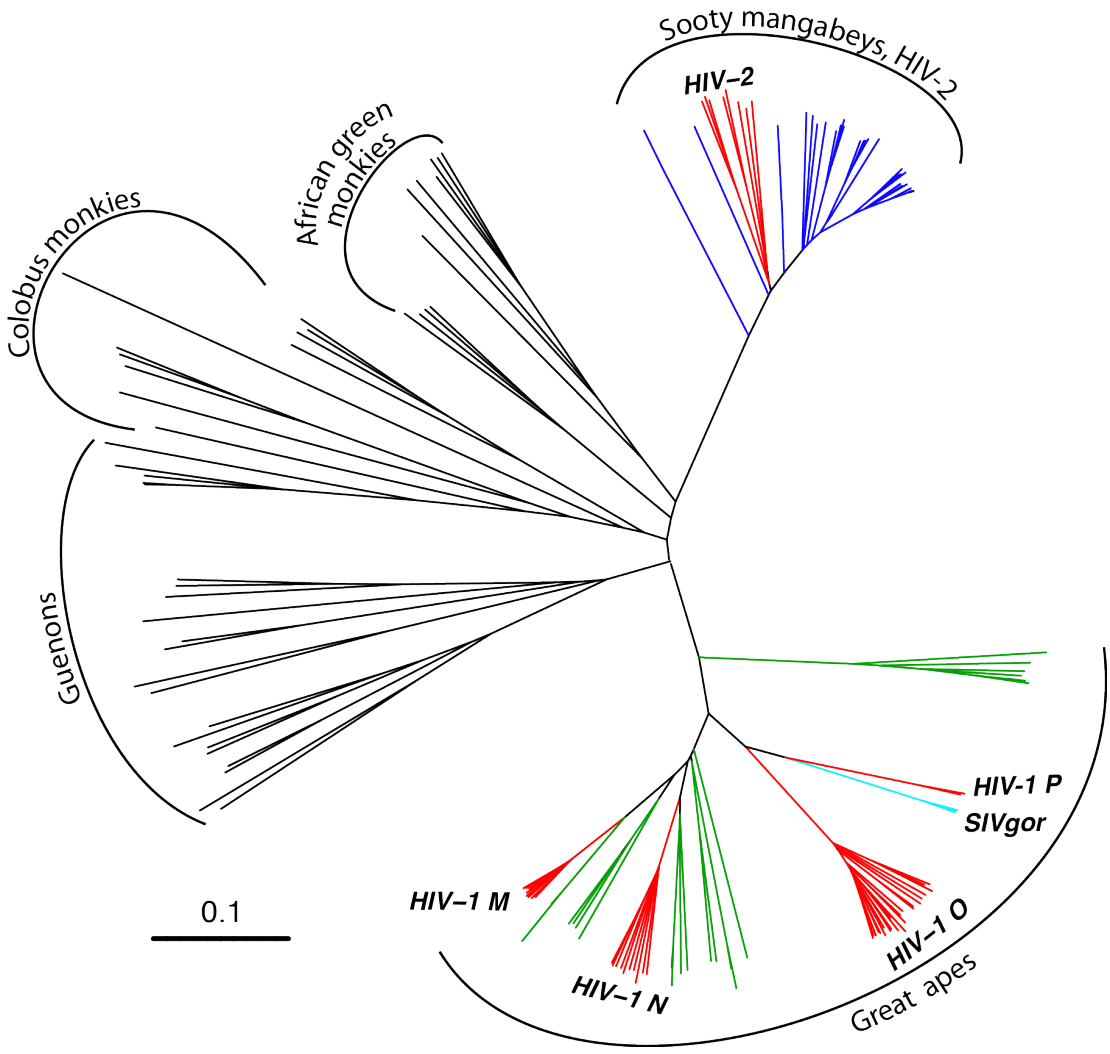


Figure 1: An unrooted phylogenetic tree showing the relationships between simian immunodeficiency viruses (SIV) and the two species of HIV (HIV-1 and HIV-2). 161 whole genome nucleotide sequences were acquired from the Los Alamos National Labs HIV Sequence Database curated web alignment and modeled using generalised time-reversible (GTR) paramaters in PhyML. The scale bar indicates the average number of substitutions per site for that branch length. HIV-1 is most similar to SIV of great apes and consists of four groups that most likely result from independent transmission events from chimpanzees (M, N, & O) and gorillas (P). HIV-2 is closely related to SIV from sooty mangabeys.

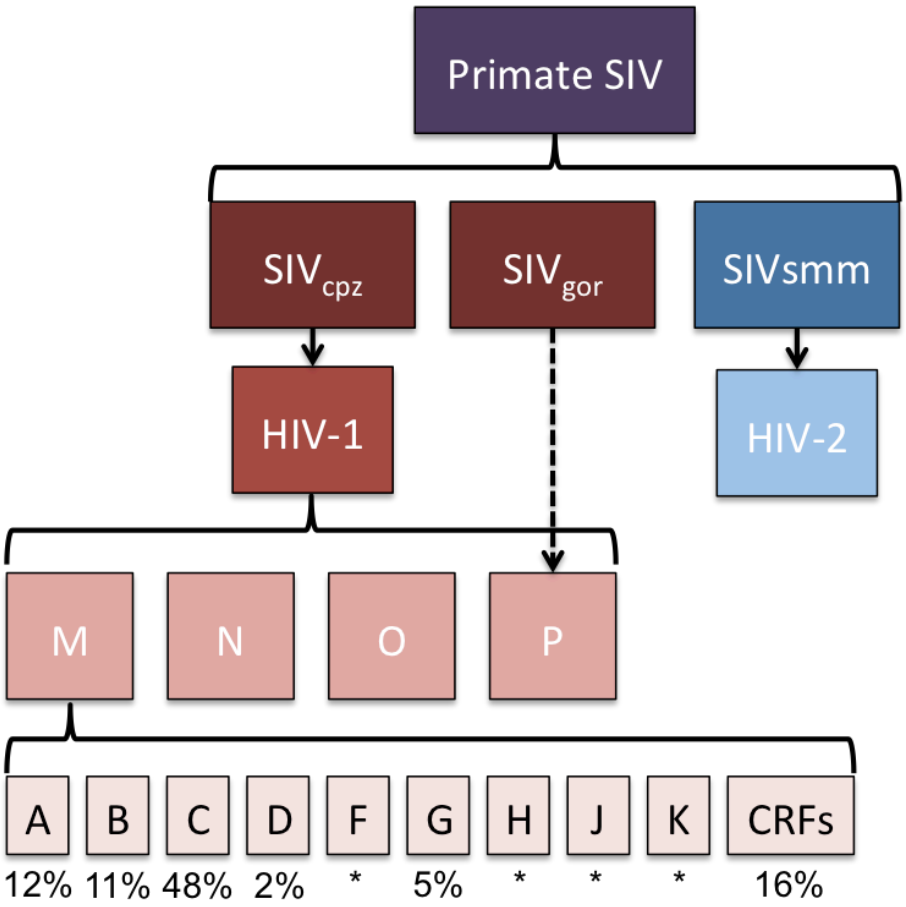


Figure 2: The subtypes of HIV-1 group M. There are several group M subtypes, as well as circulating recombinant forms that are derived from two or more parent subtypes. The percentage of group M infections by each subtype is included; an * indicates those accounting for <1% of all infections (12). Subtype B is the most common in America and Europe, while Subtype C is the most common in Africa.

1.1.3 HIV clinical progression and AIDS

The clinical progression of HIV infection is characterized by rising viral titers, $CD4^+$ T-cell depletion, immune disregulation and collapse. This process has roughly four phases (13):

1. **Infection and incubation:** Transmission occurs mainly through three mechanisms: sexual contact, intravenously (through drug use or blood transfusion), and vertical transmission (during birth or through breast feeding). The likelihood of transmission increases with higher viral titers in the blood and bodily fluids; patients on a successful anti-retroviral regimen are significantly less likely to infect others. During this period, which lasts two to four weeks, the typical tests performed by healthcare providers cannot detect HIV infection, although viral RNA and DNA might be assayable by qPCR.
2. **Acute viremia:** The virus begins to proliferate rapidly, rendering the patient highly infectious. Concurrently, $CD4^+$ T-cell counts decrease as infected cells are killed (from a normal value of ~1200 cells/µl in circulating blood down to ~600 cells/µl). This corresponds to immune suppression and, in some individuals, self-limiting flu-like symptoms. An adaptive immune response is mounted which depresses viral titers, although it cannot clear the infection entirely. After one to four weeks, viral load decreases and $CD4^+$ T-cell counts recover, but only to ~800 cells/µl. From this point, the most common clinical HIV test (anti-capsid/p24 antibody ELISA) can detect infection.

3. **Clinical latency:** After the partial recovery of the immune system, $CD4^+$ titers are maintained by accelerated production to account for virus-induced death. The immune response continues to suppress HIV titers for a period of months to years and patients receiving treatment or rare patients with a resistant genetic profile (long-term non-progressors) can maintain this stage indefinitely. However, without treatment, the $CD4^+$ T-cell count steadily decreases and viral titers increase.
4. **AIDS:** Eventually, the immune system decay reaches a critical point of collapse (usually when the $CD4^+$ T-cell count is < 200 cells/μl) and individuals become susceptible to opportunistic infections, such as Kaposi's sarcoma, *Pneumocystic carinii* pneumonia, some lymphomas, and other illnesses. The CDC considers HIV+ individuals with one of these AIDS-defining conditions to have progressed to AIDS, whereas the WHO also requires a $CD4^+$ T-cell count below 200 cells/μl. Without treatment, patients will succumb to one or more of these diseases, which are the most common causes of death for AIDS patients.

1.1.4 Treatments

In 1987, the FDA approved the first anti-retroviral (ARV) for treatment of HIV: zidovudine (AZT), which is a nucleoside-analog RT inhibitor (NRTI) (14). While initially often decreasing viral titers below detectable limits, it is unable to

completely clear infection due to some integrated virus remaining in long-lived quiescent cells, which can reactivate at any time to produce more viruses. Eventually, low levels of replicating virus permit drug-resistant mutants to flourish. Since then, several new classes of drugs (see Table 1: Classes and examples of HIV ARVs.) have been developed (15) and treatment has evolved to the current drug regime, Highly Active Antiretroviral Therapy (HAART), which utilizes combinations of drugs to reduce the likelihood of resistance. The general guidelines for HAART recommend two nucleoside-analog RT inhibitors and a third drug from a different class (for instance, a protease or integrase inhibitor). In areas where drug-resistant mutants are common, the patients beginning treatment will have their baseline resistance measured to determine the susceptibility of their HIV to specific ARVs. In cases where the HIV resistance profile is such that HAART fails to control infection, Mega-HAART/salvage therapy (HAART+additional drug combinations) is the final line of defense.

Early ARVs often had extremely unpleasant side effects, such as anemia, nausea, and vomiting. Newer therapeutics often have fewer problems, but adverse reactions are still common and can decrease the quality of life for those taking them. As ARVs fail to completely clear HIV, patients must maintain these regimes indefinitely (16).

An important factor in treating HIV infection is *when* to begin ARV therapy. Without ARVs, most patients will eventually progress to AIDS and develop opportunistic infections. As immune suppression of the virus fails, they will also have increasing viral titers and will be highly infectious. ARV therapy controls

viremia and both preserves the immune system from HIV-related destruction and lowers the risk of transmission. As any drug therapy creates a selective pressure for the enrichment of drug-resistant mutants relative to drug-susceptible viruses in the population, it is important to continue treatment one it is begun; cessation of treatment allows expansion of the viral pool, permitting new mutations and allowing the potential for recombination between partially- or fully-resistant viruses. Unfortunately, the unpleasant side effects associated with many ARVs can discourage compliance with a lifelong treatment regimen.

For a time, brief drug 'holidays' (where the patient discontinued their medication for a short duration) were encouraged by doctors to improve the quality of life for HIV patients. Ultimately, it was determined that these holidays increased the potential for drug resistance and the policy was discouraged. While the current treatment guideline is to 'treat early-treat hard' (17), patients and their doctors must weigh the risk of transmission and the consequences of unchecked viremia against the potential side effects, cost, and willingness of the patient to continue therapy indefinitely once it is begun. So far, there are no data to suggest that treating patients with $CD4^+$ cell counts >350 cells/ul has benefits that outweigh the potential issues (16). A clinical trial to address the risk:benefit ratio of beginning treatment in patients with $CD4^+$ counts >500 cell/ul is currently recruiting patients (18).

Class	Target protein	Function	Example
Nucleoside-analog RT inhibitors	Reverse Transcriptase	Acts as a chain-terminator during replication	Zidovudine (AZT)
Non-nucleoside RT inhibitor		Allosteric enzyme inhibition	Neviarpine
Integration inhibitor	Integrase	Blocks integration of viral DNA into the host genome	Raltegravir
Protease inhibitor	Protease	Prevents proteolytic cleavage of Gag and Gag-pol	Ritonavir
Maturation inhibitor	Gag	Blocks proteolytic cleavage of the Gag polyprotein at the capsid:sp1 junction	Bevirimat
Entry inhibitor	CCR5	Prevents binding of gp120 with CCR5	Maraviroc
	gp41	Prevents fusion with the membrane	Enfuvirtide

Table 1: Classes and examples of HIV ARVs.

In 2009, researchers announced the first known 'cure' of HIV in a man who became known as the Berlin Patient. After infection, he had developed leukemia and required a bone marrow transplant. His transplant came from a donor who had a CCR5Δ32 mutation, which results in a truncated CCR5 receptor; cells with this mutation are resistant to HIV infection (19). Six years after transplant and cessation of HAART, the Berlin Patient continues to have undetectable levels of viral replication and is considered functionally cured. Since then, two other cures were tentatively announced with a different protocol: conventional bone marrow transplant while continuing HAART (20). Sadly, this treatment failed and the patients experienced a return of viremia (21).

The treatment used by the Berlin Patient was expensive, difficult to obtain, and, due to the bone marrow transplant, dangerous. It is unlikely that this will ever be expanded into a global cure, although it does offer hope for treatments that utilize the resistance of CCR5Δ32 cells to infection. Potential avenues of achieving this are under investigation; one method uses zinc finger nucleases to target CCR5, resulting in $CCR5^-$ cells that are resistant to HIV infection (22).

Currently, the only way to prevent HIV transmission is through barriers (condoms or dental dams during sexual contact), avoidance (only using sterile needles, abstaining from unprotected sex, not breastfeeding if HIV^+). After exposure (such as during childbirth or following a contaminated needle stick), prophylactic ARV treatment can decrease the likelihood of contracting HIV. Attempts to develop a vaccine have so far been unsuccessful, in part because of HIV's genetic diversity.

1.2 The Viral Lifecycle

1.2.1 HIV genome

The diploid HIV genome packaged in the 80-100 nm virions is two synonymous copies of a ~9.3 kb single-stranded, positive-sense RNA. As a complex retrovirus, it contains the canonical retroviral genes *gag, pol,* and *env,* but also encodes several regulatory and accessory proteins (Tat and Rev; Vpu, Vpr, Vif, and Nef, respectively) (see Table 2 and). The gene products Gag, Pol, and Env are polyproteins. While Env is produced from its own mRNA, Gag contains a frameshift signal which is read through by the ribosome ~5-10% of the time, producing Gag-Pol (23). Pol is not synthesized independently. The regulatory and accessory proteins are also produced from individually spliced mRNAs. After reverse transcription and integration, the now 9.7 kb dsDNA genome is referred to as the provirus.

In addition to genes, the genome also contains several features that are essential to viral replication. These include:

In the RNA genome, each end is flanked by a unique region (termed U5 or U3 for the 5′ and 3′ ends of the genome, respectively) and a repeated region (R). During reverse transcription into proviral DNA, the U5 and U3 regions are duplicated to the opposite side of the genome, leaving each end of the provirus flanked with a U3-R-U5 sequence called the Long Terminal Repeat (LTR). The LTRs are important for both complete genome reverse transcription (1.2.3 Reverse Transcription) and mRNA transcription). From 5′ to 3′, the 5′ LTR contains:

- The target sequence for viral transactivation (TAR), which binds Tat.

    - A primer-binding site for a packaged host $tRNA_{lys}$, used as a primer for ssDNA synthesis.
    - The Psi element and dimer initiation site, which are necessary for co-packaging of the two copies of the RNA genome.
- The Rev response element (RRE), which is the Rev binding site.
- Two polypurine tracts (PPTs), which are used for plus-strand DNA synthesis priming (see section 1.2.3 Reverse Transcription).
- A 3′ polyadenylation signal.

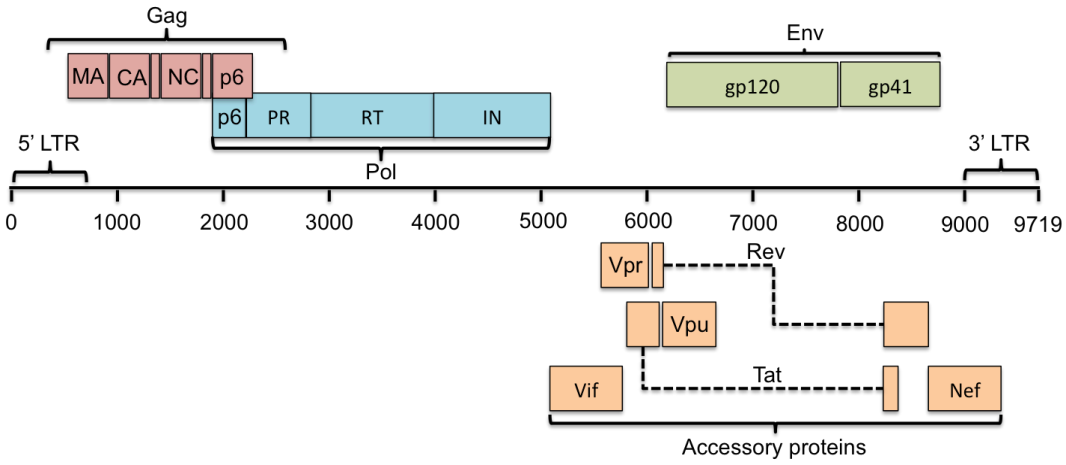


Figure 3: Organization of the HIV-1 provirus. The 9.7 kb provirus has the same gene arrangement as the 9.1 kb RNA genome but is slightly larger due to the U3/U5 duplication in the LTRs.

Human proteins necessary for infection		
CD4 Cluster of differentiation 4	N/A	Expressed on the surface of some components of the immune system ($CD4^+$ T-cells, macrophages, & dendritic cells) and bound by gp120 during HIV infection
CCR5 C-C chemokine receptor type 5	N/A	Expressed on the surface of leukocytes, acts as a co-receptor for HIV binding
CXCR4 C-X-C chemokine receptor type 4	N/A	Expressed on the surface of many cell types, including leukocytes, acts as a co-receptor for HIV binding
CypA Cyclophilin A	~500	Incorporated into virions, masks virus from host immune response
Viral proteins		
Gag Group-specific antigen	~5000	The polyprotein precursor to MA, CA, NC, & p6.
MA Matrix	~5000	Virion structural protein , while part of Gag, targets to plasma membrane for assembly and promotes incorporation of Env
CA p24, Capsid	~5000	Structural protein, encapsulates RNA
NC Nucleocapsid	~5000	Nucleic acid chaperone protein, coats and protects viral RNA in virion, promotes recombination, as part of Gag packages RNA into new virions
p6	~5000	As a part of Gag, incorporates Vpr into the virion and promotes budding. A transframe version of the protein is a part of Pol.
Pol Polymerase	~250	The polyprotein precursor to PR, RT, and IN
PR Protease	~250	Cleaves Gag and Pol polyproteins in a process called maturation
RT Reverse transcriptase	~250	Viral polymerase that converts ssRNA to dsDNA, RNaseH domain degrades RNA

IN Integrase	~250	Covalently inserts HIV into host genome
Env Envelope	N/A	The polyprotein precursor to gp120 and gp41
gp120 Glycoprotein 120	4-35 trimers	Viral surface protein that binds CD4 and chemokine co-receptors during entry
gp41 Glycoprotein 41	4-35 trimers	Viral transmembrane protein that induces fusion with the host membrane during entry
Vif Viral infectvity factor	1-150	Prevents incorporation of APOBEC3G, an innate immune component that causes hypermutagenesis during reverse transcription
Tat HIV transactivator	N/A	Increases transcription of viral mRNA
Rev Regulator of Expression	N/A	Prevents splicing and promotes nuclear export of full-length genomic RNA
Vpr Viral protein R	~700	Proteosomal degradation of host restriction factors, promotes G2/M cell-cycle arrest
Vpu Viral protein U	N/A	Degrades newly synthesized CD4 to allow transport of Env to the cell surface, blocks Tetherin from preventing viral release
Nef Negative factor	~60-200	Strips CD4 from the infected cell surface to prevent interactions with Env, blocks apoptosis

Table 2: Host and HIV-1 proteins necessary for the viral lifecycle.

1.2.2 Transmission, lifecycle, and disease

The viral life cycle is illustrated in Figure 4. HIV requires a CD4 receptor for entry, limiting infection to $CD4^+$ components of the immune system: T-cell lymphocytes and monocytes such as macrophages. A secondary entry requirement is a co-receptor such as the chemokine receptors CCR5 or CXCR4. Viruses that require CCR5 are considered R5-tropic, while CXCR4-requiring viruses are X4-tropic. Dual-tropism (R5/X4) and other rare receptor preferences also exist. The viral envelope protein gp120 docks with CD4 and then undergoes a conformational change, exposing regions that bind to the co-receptor, which allows gp41 to induce fusion with the host membrane (24).

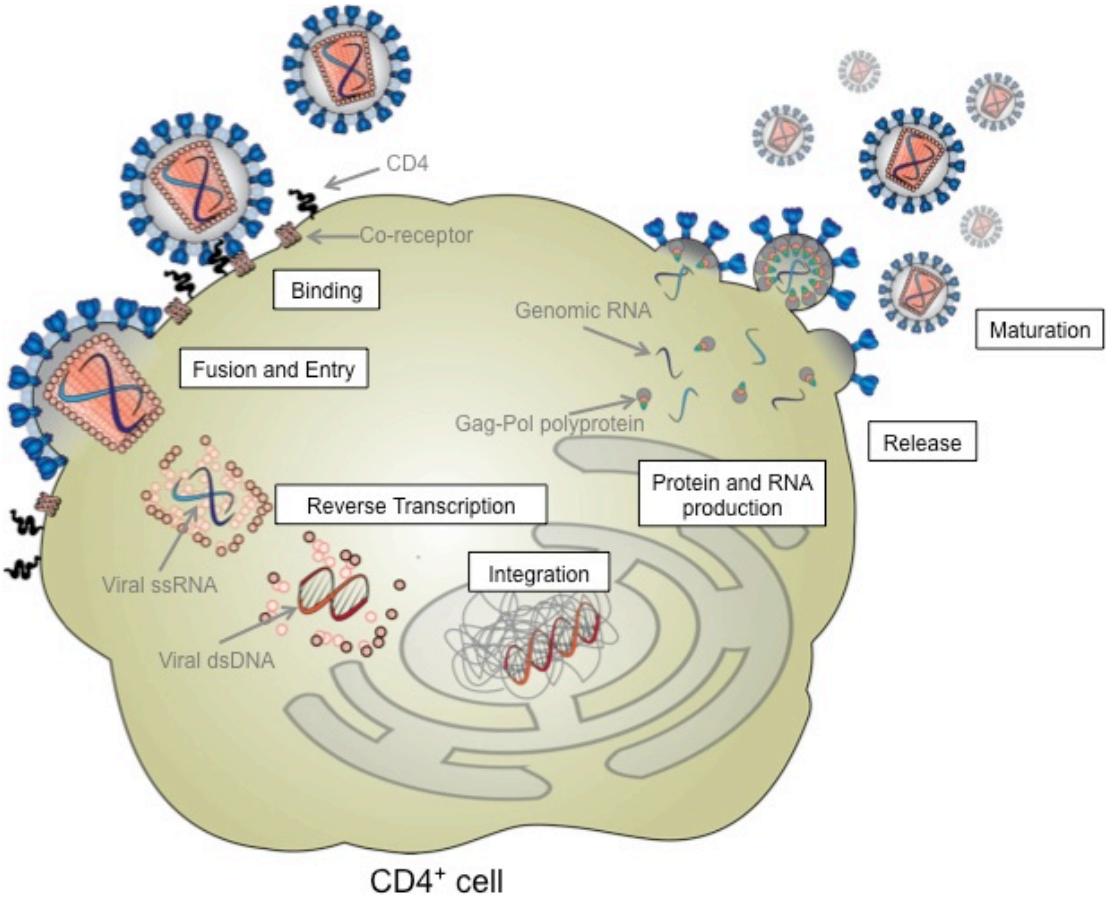


Figure 4. The life cycle of HIV. HIV infects $CD4^+$ cells, where it reverse transcribes its genome from ssRNA to dsDNA and integrates this into the host chromosome. By using cellular machinery, it synthesizes proteins and RNA genomes for packaging into new virions (24). See text for details.

After fusion, the viral capsid containing the genomic RNA is injected into the host cell. Progressive uncoating allows packaged RT to initiate reverse transcription and convert the single-stranded diploid RNA genome into double stranded DNA (see section 1.2.3 Reverse Transcription). The newly synthesized DNA, RT, IN, and Vpr, form the pre-integration complex that is transported to the nucleus. Lentiviruses, unlike simple retroviruses, can infect both dividing and non-dividing cells by entering through a nuclear pore complex. Once inside, IN removes the terminal two 3′ nucleotides from each strand, exposing the 3′ hydroxyl on the conserved terminal CA nucleotide pair. Integrase inserts these exposed ends into the host genome and the gaps are repaired by host cellular machinery, resulting in a permanent, covalently integrated provirus (24).

After integration, HIV relies on host machinery to produce new viral RNA for protein production and new genomes for nascent viral particles. Initially, the two LTRs act as weak promoters to induce RNA pol II to transcribe spliced viral mRNA. However, as viral proteins are synthesized, Tat is imported into the nucleus and recruits a transcription complex to the TAR that increases viral mRNA output. Later, Rev is also imported and, by binding the RRE, depresses splicing and allows export of full-length genomic RNA. Two copies of this viral RNA dimerize using the DIS and are bound by Gag at the Psi element for packaging into the nascent viral particles. If a cell is co-infected by two HIV viruses, it's possible for Gag to package an RNA from each parental sequence; the major requirement for co-packaging is a compatible DIS. This allows for potential recombination during the next round of replication and contributes to the incredible genetic diversity of HIV (24).

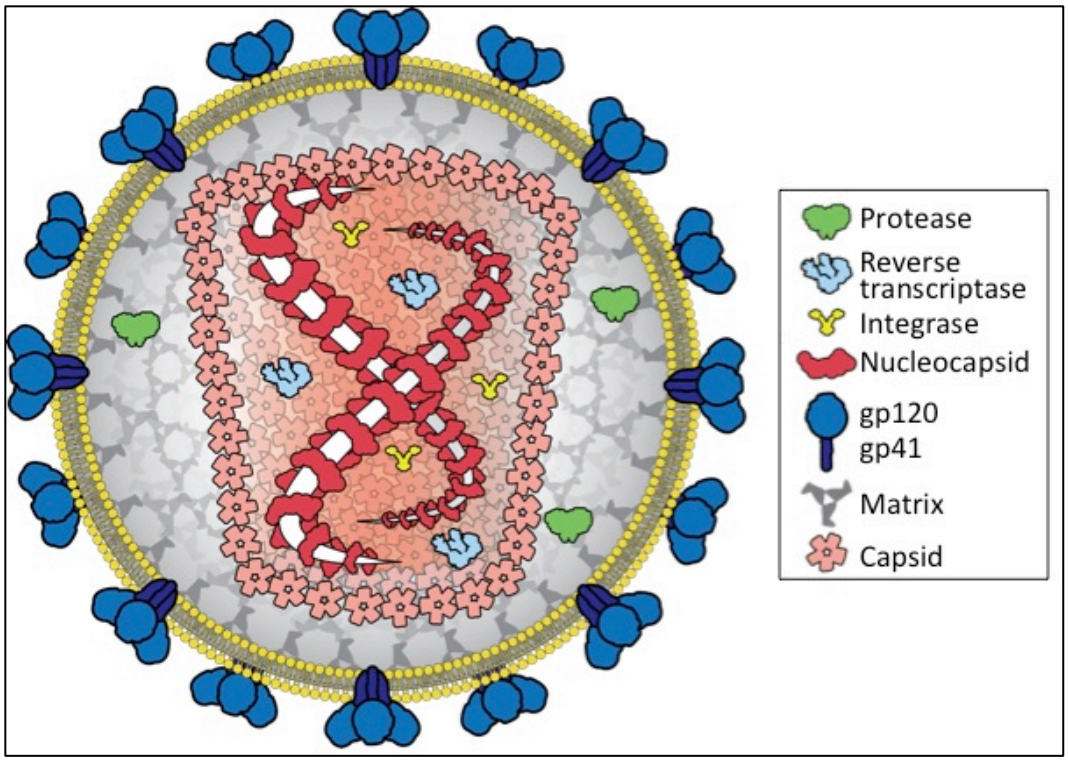


Figure 5: A Mature HIV virion. The diploid genomic RNA (white) is coated in nucleocapsid protein. Integrase and reverse transcriptase are packaged within the virion, for converting the RNA genome to dsDNA and integrating it into the host genome. The virion is enclosed in a host-derived lipid membrane, with gp41 and gp120 trimers present for attaching to permissive cells.

After packing of the viral genomes and the necessary viral/host proteins (see Table 2), new virions bud from the cell membrane, taking with them a lipid envelope derived from the host. Subsequently, protease cleaves the Gag and Gag-Pol polyproteins into the capsid, matrix, nucleocapsid, and p6 proteins (Gag) and RT, integrase, and protease enzymes (Pol). This process is called maturation and results in infectious viral particles (Figure 5) (24).

1.2.3 Reverse Transcription

HIV relies on RT for the complex process of reverse transcribing its ssRNA genome to dsDNA. The enzyme is a heterodimer, consisting of a 66 kDa subunit containing the enzymatic domains and a catalytically inactive 51 kDa subunit that acts as a facilitator for catalysis (25). Both are encoded by the same gene, but HIV protease converts a subset of p66 enzymes to the smaller p51 subunit. The enzyme has a right-handed structure, containing five regions: a connection domain (the major point of contact between the two subunits) and thumb, fingers, and palm domains that encircle the catalytic cleft in p66, which also contains an Ribonuclease H (RNase H) domain whose sequence is absent in the p51. Despite the subunits sharing the same amino acid sequence (excepting the p51 truncation) their structures and activities are vastly different. p66 is an open and flexible enzyme, capable of synthesizing ssDNA from ssRNA and dsDNA from ssDNA, as well as degrading RNA in RNA:DNA hybrids via the RNase H domain. In contrast, p51 is catalytically inactive and structurally rigid (26).

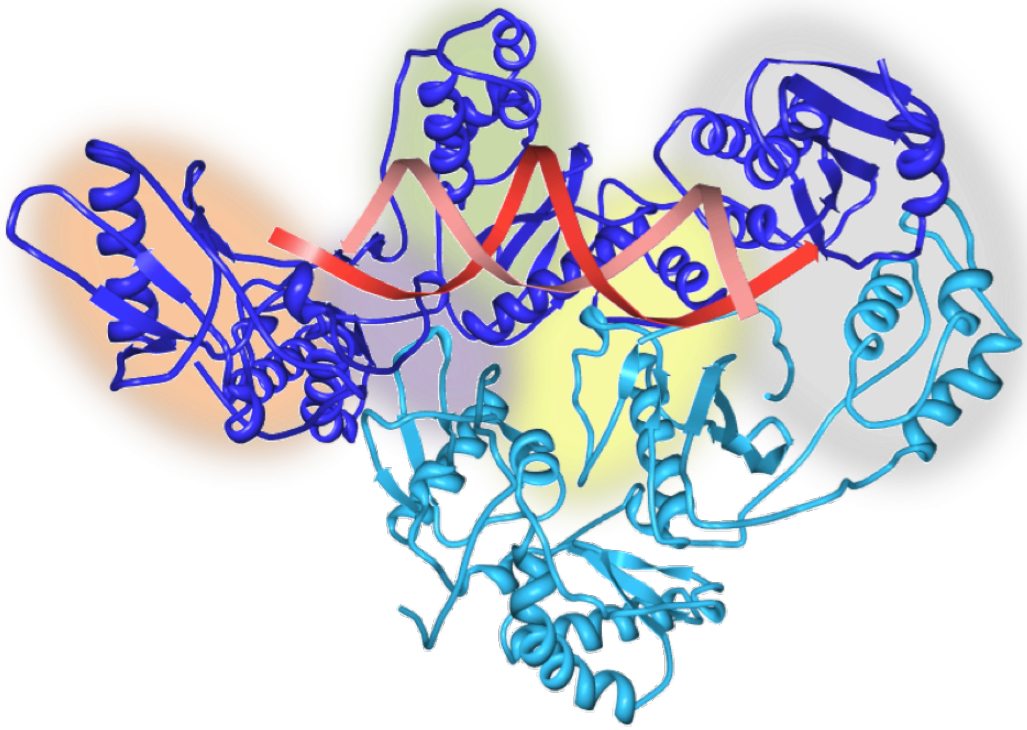


Figure 6. A crystal structure of HIV RT bound to dsDNA. The p66 fragment is represented by dark blue, p51 is light blue, and dsDNA is red. The domains of the enzyme are highlighted in color: palm (purple), thumb (green), fingers (orange), connection (yellow), and RNase H (gray). All of the catalytic functions of the enzyme are contained in the p66 subunit. Crystal structure data from (27).

The catalytic cleft in HIV RT is structurally similar to other DNA polymerases, particularly to Family A DNA polymerases like the Klenow fragment of *E. coli* pol I (discussed in chapter 3). Like these pols, it contains aspartic acid residues (Asp-110, Asp-185, and Asp-186) that coordinate the two divalent cations, which are required for nucleotide incorporation (reviewed in (28)). These cations act as stabilizers during the nucleophilic attack of the incoming dNTP by the 3′ OH of the primer. This role is most likely filled *in vivo* by the most readily abundant cation, magnesium (Mg^{2+}), which exists at an estimated free concentration of ~0.25 mM in the cytosol of lymphocytes (29, 30). The RNase H domain also coordinates divalent cations for RNA cleavage, although there is conflicting evidence for both one- and two-metal-ion binding sites (reviewed in (31)). These cations are coordinated by a catalytic triad (Asp-443, Asp-498, and Glu-478) in the active site (32).

As RT lacks proofreading activity, it is highly error-prone compared to DNA polymerases with exonuclease domains. On average, it makes approximately 3.4 x 10^{-5} mutations per bp per cycle *in vivo*, which is roughly one mutation per replication cycle (this does not include mutations generated by host DNA and RNA pols) (33). In contrast, replicative polymerases like *E. coli* pol I have error rates between 10^{-6} and 10^{-8} (34). This low fidelity results in enormous genetic diversity and is partially responsible for HIV's ability to rapidly evolve resistance to ARVs.

The process of reverse transcription (Figure 7) is complex and is primed by the $tRNA_{lys}$ that is packaged into the virion, which hybridizes to the PBS on the 5′ end of the RNA genome. Minus-strand DNA synthesis begins, truncating at the 5′

end of the RNA template, after which RNase H degradation allows dissociation of the newly formed DNA. As the 5′ and 3′ R regions are homologous, the DNA can hybridize with the 3′ R in a strand transfer event, which allows full minus-strand DNA completion. Meanwhile, RNase H cleaves the RNA template, leaving behind a small polypurine tract (PPT) located roughly 300 nt from the 3′ end of the genome. From this, plus-strand DNA synthesis is initiated and the 3′ U5-R-U3 LTR is completed. A second strand transfer event to the 5′ DNA PBS sequence allows completion of synthesis and results in a double stranded DNA copy of the genome (24). Because two copies of the RNA genome are packaged into each virion, each strand transfer event represents an opportunity for template switching between the two RNAs, called recombination. In addition to the two obligate strand transfers, secondary structures that slow synthesis by RT present opportunities for the nascent minus-strand DNA to switch to the second RNA genome and continue synthesis there. This switching during synthesis of the minus strand occurs on average 3-10 times during normal reverse transcription in a T-cell and even more in macrophages. This creates chimeric viruses that are composites of two parental sequences and, much like sexual recombination in cellular organisms, allows combinations of sequences between lineages (reviewed in (35)).

Due to its crucial role in viral replication, HIV RT has long been an area of study and many questions remain about the function of this fundamental enzyme. In this thesis, I will address the ability of divalent cations to profoundly alter the behavior of HIV RT during reverse transcription, as well as present research performed to translate RT inhibitors produced *in vitro* for use as therapeutics. I will

also investigate the ability of Family A DNA polymerases to preferentially bind specific DNA sequences, as has been previously demonstrated with HIV RT.

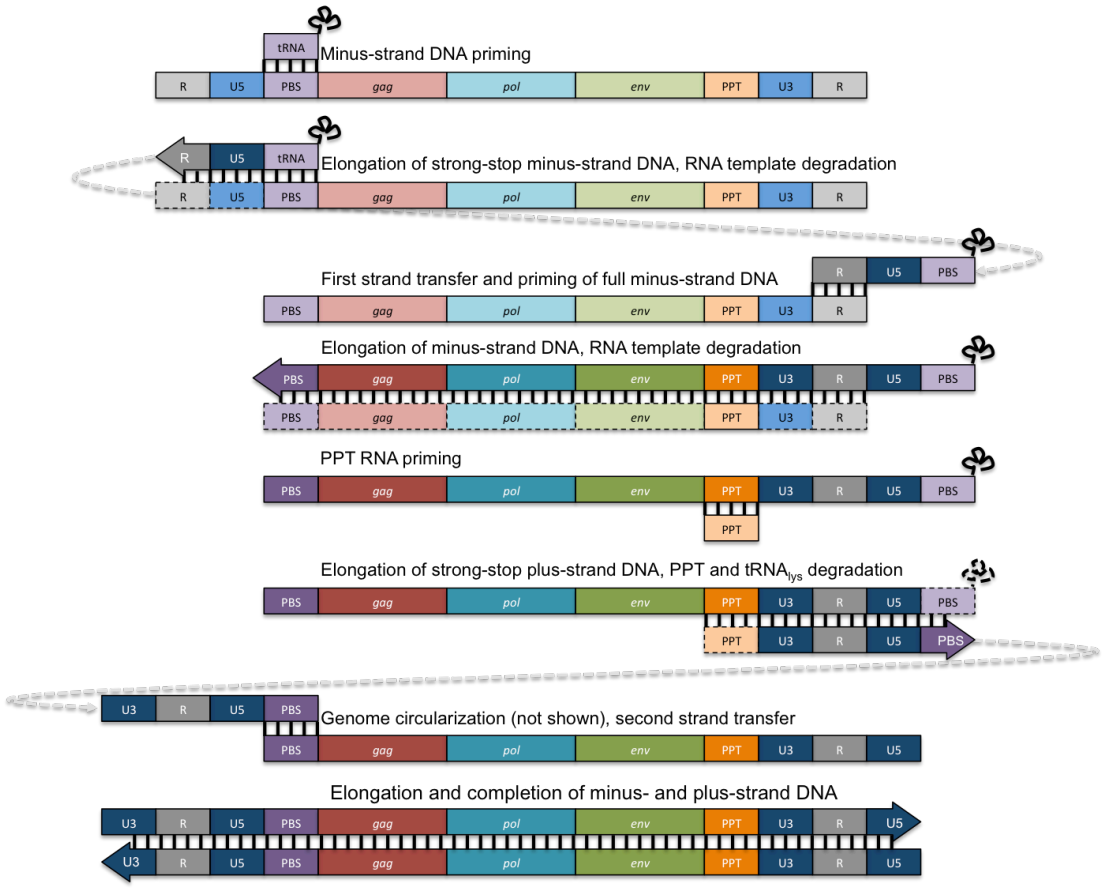


Figure 7. Reverse transcription of the HIV genome. HIV uses RT to convert the ssRNA (lighter colors) genome to dsDNA (darker colors). See text for details.

Chapter 2: Mechanism of HIV reverse transcriptase inhibition by zinc: formation of a highly stable enzyme-primer-template complex with profoundly diminished catalytic activity

2.1 Introduction

HIV RT, like other RTs, possesses both nucleotide polymerization and RNase H capabilities. Both activities require a divalent cation as an integral co-factor in the mechanism of catalysis (36, 37). Under physiological conditions Mg^{2+} presumably functions as the co-factor, however other divalent cations (including Mn^{2+}, $Ni^{2+,}$ and Cu^{2+}) can also be used for nucleotide catalysis (38). Some divalent cations, such as Pd^{+2}, Ca^{2+}, Mn^{2+}, and Zn^{2+}, have been shown to inhibit Mg^{2+}-dependent RT activity (38, 39, 40, 41, 42).

Magnesium binds to RT in the polymerase and RNase H domains where binding is coordinated by the acidic catalytic residues within the active sites. By analogy to other DNA polymerases, two Mg^{2+} ions participate in polymerization via a universal "two metal ion" mechanism (reviewed in (28)). The number of Mg^{2+} ions bound at the RNase H active site is more controversial, as both one and two metal mechanisms have been proposed (reviewed in (31)). In addition to the cation binding sites, other sites may also exist on the enzyme. For example, some reports suggest that as many as 21 Mg^{2+} ions are bound to a single *Escherichia coli* DNA polymerase I molecule (43). The role, if any, of these additional cations is unclear.

Many enzymes require cation co-factors that play an intricate role in the mechanism of catalysis. In humans these typically include magnesium, calcium, iron, manganese, and zinc, cobalt, copper, selenium, and molybdenum. Note that calcium,

though required to stimulate the activity of several enzymes, typically acts as an allosteric regulator involved in cell signaling rather than a co-factor in catalysis. Most enzymes that carry out catalysis on nucleic acids have evolved to use Mg^{2+} as a primary cofactor. Although the majority of the ~10 mM Mg^{2+} in cells is complexed with cellular components, the intracellular concentration of free available Mg^{2+}, typically reported to be between 0.25 mM and 2 mM, is still much greater than other divalent cations (44, 45, 46, 30, 47). Unlike Mg^{2+}, other cation co-factors are maintained at much lower levels in cells. The availability of these cations is carefully controlled by a combination of ion transporters, binding proteins that sequester the cations, and sequestration by cellular nucleotides and within specific cell organelles. For example, the concentration of free Zn^{2+} in cells is in the low nM range despite a total concentration (free and bound) of ~0.1-0.5 mM and a plasma concentration of ~15 μM (45, 48, 49, 50, 51). The tight regulation is required due to the ability of some cations to alter transcription and inhibit or alter the activity of several enzymes. Zinc is particularly notable because it is important to many processes in cells and is a component of several proteins, particularly zinc finger containing proteins (48, 51, 52). In addition, Zn^{2+} has recently been shown to modulate signal transduction, and like Ca^{2+}, serve a cellular second messenger role for specific processes (53, 54). Zinc is also a potent inhibitor of many enzymes, including several viral polymerases (40, 41, 55-56). The mechanism of inhibition may take different forms including direct displacement of Mg^{2+} due to higher affinity of Zn^{2+} for the cation binding site, allosteric inhibition through binding at secondary sites on the enzyme, or more

complex mechanisms that involve interactions with substrate or other cofactors (nucleotides or nucleic acids for example).

This chapter investigates the interaction of HIV-RT with Zn^{2+} (as well as Mn^{2+} and Ca^{2+}) to determine how it affects RT activity. The results showed that Zn^{2+} supported both polymerase and RNase H activity at concentrations much lower than Mg^{2+} and was a potent inhibitor of RT in the presence of Mg^{2+}. Surprisingly, the RT-Zn^{2+}-primer-template complex was much more stable than the same complex formed with Mg^{2+} while rate of nucleotide incorporation was much slower. Overall Zn^{2+} inhibits RT by forming what is essentially a "dead-end complex" that can remain bound to the primer-template for hours but has very slow incorporation kinetics. The possibility of using Zn^{2+} or other cations as inhibitors of virus replication is discussed.

2.2 Materials and Methods

Materials– The following reagent was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: pNL4-3 from Dr. Malcolm Martin (57). PCR primers and primers used to prime templates in reverse transcription assays were obtained from Integrated DNA Technologies, Inc. The HIV-RT clone was a gift from Dr. Michael Parniak (University of Pittsburg). HIV-RT was purified as described (58). Aliquots of HIV-RT were stored frozen at -80°C and fresh aliquots were used for each experiment. RNase H minus (E478>Q) RT was a gift from Dr. Stuart Le Grice, HIV Drug Resistance Program, National Cancer Institute, Frederick, MD. Taq polymerase, T4 polynucleotide kinase, MuLV-RT, and restriction enzymes were from New England Biolabs. T7 and SP6 RNA polymerases,

calf intestinal alkaline phosphatase (CIP), AMV-RT, and dNTPs were from Roche Diagnostics. DNase I-RNase-free and RNase H were from United States Biochemical Corporation. RNase inhibitor (RNasin) and the phiX174 HinfI digest DNA ladder was from Promega. Radiolabeled compounds were obtained from PerkinElmer. Sephadex G-25 spin columns were from Harvard Apparatus. QIAquick PCR Purification Kits and RNeasy Mini Kits for RNA cleanup were from Qiagen. All other chemicals were from Sigma or Fisher Scientific.

Generation of RNA transcripts– Extension assays were carried out on a 425 nt RNA template derived from a 420 nt section of the HIV-1 *gag-pol* genome region (bases 1924-2343 on pNL43) and 5 additional non-HIV nts at the 5′ end. PCR amplification was performed using forward primer 5′-GATTTAGGTGACACTATAG*TCAAA*CAGAAAGGCAATTTTAGGAAC-3′and reverse primer 5′-TATCATCTGCTCCTGTATCT-3′. Underlined bases contain an SP6 promoter and italicized bases correspond to the 5 non-HIV nts. After amplification (94°C/3 min followed by 25 cycles of: 94°C/1 min, 50°C/1 min, 72°C/1 min, followed by a final extension at 72°C for 5 min), template DNA was purified using the QIAquick PCR Purification Kit and 8 μls was used in RNA transcription reactions with SP6 polymerase according to the manufacturer instructions. Reactions were allowed to proceed for 3 hours, at which point they were treated with 0.4 units/μl of DNase I for 20 minutes at 37°C. RNA was purified using the RNeasy Mini Kit and quantified using UV spectroscopy. The 60 nt RNA for the RNase H assays was prepared by digesting 10 μg of pBSM13+ (Stratagene) plasmid with 20 units of HindIII in 20 μl using the manufacturer's protocol. Digested material was

used in a transcription reaction with T7 RNA polymerase as described above. Transcription reactions were extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with ethanol. The RNA was gel purified on an 8% denaturing polyacrylamide gel, located by ultra violet shadowing, and recovered by the crush and soak gel elution method (59). The amount of recovered RNA was determined spectrophotometrically from optical density. Gel-purified RNAs were dephosphorylated with CIP using the manufacturer's protocol and supplied buffer. The reactions were extracted and precipitated and the recovered RNA was end-labeled using T4 polynucleotide kinase as described below. The material was then run on a second denaturing polyacrylamide gel, located by autoradiography and recovered as described above. This material was hybridized to a DNA oligonucleotide as described below. The sequence of the recovered RNA was: 5′-GGGCGAAUUCGAGCUCGGUACCCGGGGAUCCUCUAGAGUCGACCUGCA GGCAUGCAAGCU-3′.

5′ end-labeling of DNA primers for reverse transcription and off-rate determinations, and RNA template for RNase H assays– The reverse DNA primer (see above), DNA primer for the binding half-life determination assays (33 nt: 5′-TCCCCGGGTACCGAGCTCGAATTCGCCCTATAG-3′), and the RNA template for the RNase H assays (see above) were 5′ end-labeled using T4 polynucleotide kinase and γ–^{32}P-ATP using the manufactures protocol and supplied buffer. The material was run through a G-25 sephedex column to remove excess ATP.

Nucleic acid hybridization– Hybrids were prepared by mixing RNA (425 nt product described above) and end-labeled DNA (reverse primer from above) for

reverse transcription assays at an approximately 1:1.5 ratio (template:primer) of 3′ termini in buffer containing 50 mM Tris-HCl (pH=8), 80 mM KCl and 1 mM DTT. For binding half-life determination assays the end-labeled 33 nt primer described above was hybridized to a 50 nt DNA (5′-TTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATC-3′) and for RNase H assays, the end-labeled 60 nt RNA was hybridized to a 23 nt DNA (5′-AGGATCCCCGGGTACCGAGCTCG-3′), each at a 1:1 ratio. For all hybrids, the mixtures were heated to 65°C for 5 min, and then cooled slowly to room temperature.

Reverse transcriptase primer extension assay– Hybridized primer-template was incubated in 21 μl of buffer containing 50 mM Tris-HCl (pH=8.0), 80 mM KCl, 1 mM DTT, 100 μM each dNTP (final concentration, unless otherwise noted), 0.2 units/μl RNase inhibitor, and varying concentrations of $ZnCl_2$, $MnCl_2$, $CaCl_2$, or $MgCl_2$ (as indicated in legends). Following a 3 minute pre-incubation at 37°C, 4 μls of RT (final concentration of 25 nM) was added to begin extension. The final template concentration was 5 nM in the 25 μl reactions. The reactions were stopped by addition of 2X loading buffer (90% formamide, 10 mM EDTA (pH=8.0), and 0.025% bromophenol blue and xylene cyanol) at the indicated times. Samples were separated on 6 or 8 % denaturing urea gels (7M urea, 19:1 acrylamide:bisacrylamide), dried, and imaged using a Fujifilm FLA-5100 or FLA-7000.

RNase H assays– Hybridized RNA-DNA from above (25 nM final concentration) was incubated in 21 μl of buffer containing 50 mM Tris-HCl (pH=8.0), 80 mM KCl, 1 mM DTT, 0.2 units/μl RNase inhibitor, and various

concentrations of $MgCl_2$ or $ZnCl_2$ (as indicated in legend to Figure 12). A 3 min preincubation at 37°C was followed by addition of 4 μls of HIV-RT (10 nM final concentration) and incubation was continued for 5 min (unless otherwise indicated). Products were resolved on 10% denaturing gels as described below.

Assays to determine RT-substrate binding half-lives– Half-life determination assays for reactions with RT and $ZnCl_2$, $MnCl_2$, or $MgCl_2$ were performed as described previously (34). Time points are indicated in the Fig. 7.

Processivity assay– Processivity assays were conducted using the conditions stated above under *Reverse transcriptase primer extension assay* with the following modifications: (1) the volume of the preincubation reaction was 50 μls; (2) divalent cations and dNTPs were not included in the preincubation step; (3) RT (50 nM final concentration) was included in the preincubation step; (4) reactions were initiated by the addition of a 10 μl solution containing a divalent cation (final concentration 400 μM or 2 mM for $ZnCl_2$ and $MgCl_2$, resp.); dNTPs (final concentration 100 μM each) and poly(rA)-oligo(dT) (final concentration 0.4 μg/ul) in 50 mM Tris-HCl (pH=8), 80 mM KCl, 1 mM DTT. Five μl aliquots were removed at the indicated times (Figure 13) and reactions were terminated with 5 μl of 2X loading buffer (see above). Samples were run on a 6% denaturing gel as stated above. The addition of poly(rA)-oligo(dT) limits synthesis to preformed complexes of RT-primer-template and further limits synthesis to a single processive cycle (35). Poly(rA)-oligo(dT) used in the reactions was prepared by mixing oligo (dT_{20}) with poly(rA) at a 1:8 ratio (w/w) in 10 mM Tris-HCl (pH 7.5). The mixture was incubated for 10 minutes at 37°C and slow cooled to room temperature.

Determination of average extension rate– Average extension rates for reactions with 400 μM $ZnCl_2$ or 2 mM $MgCl_2$ were determined using results from the reverse transcription primer extension assays and the Fuji MultiGauge software (version 3.0) on the FLA-5100 or FLA-7000 and the molecular weight determination program. Thirty and one min reaction times were used for $ZnCl_2$ and $MgCl_2$ reactions, respectively. At these time points, products had not reached the end of the template. The average extension rate was estimated by calculating the size in nts of each band on the gel (*s*) and subtracting the primer length (20 nt), then using the imager to determine the relative proportion (with the total being set to 1) of the total extended primers to which each band corresponded (*y*). The band's contribution to the average extension rate can be represented by the following equation: (*s*-20)**y*. The average extension rate can then be calculated using the following expression, where *t* is the reaction time in seconds:

$$\frac{\sum[(s\text{-}20)*y]}{t}$$

As an example, if a 60 second reaction produced 4 extension product bands of sizes 30, 40, 60, and 80 nts that were of equal intensity then the average extension rate would be [(30-20)*0.25+(40-20)*0.25+(60-20)*0.25+(80-20)*0.25]/60 which equals 32.5 nt/sec.

Gel electrophoresis– Denaturing polyacrylamide gels (19:1 acrylamide:bis-acrylamide, respectively), were prepared and subjected to electrophoresis as described (59).

2.3 Results

2.3.1 Zn^{2+} potently inhibits HIV-RT catalysis in the presence of Mg^{2+}

Although Zn^{2+} has been shown to inhibit HIV-RT activity, the mechanism and nature of the inhibition have not been investigated. To investigate this, primer-directed reverse transcription was performed using a 425 nt RNA template derived from the *gag-pol* region of the HIV genome primed with a 20 nt 5′ ^{32}P end-labeled DNA primer (Figure 8A). Full extension of the primer to the end of the template produced a 425 nt DNA product (as indicated in Figure 8B). In these experiments reactions were performed using 100 μM of each dNTP. Control reactions performed in the absence of Mg^{2+} with $ZnCl_2$ (lanes 2 and 3, 1 and 30 min, resp.) and $ZnSO_4$ (lanes 4 and 5, 1 and 30 min, resp.) at 400 μM each showed synthesis with Zn^{2+} under optimal conditions (see Figure 10). Both $ZnCl_2$ and $ZnSO_4$ gave comparable results and the former was used for all other experiments. Other reactions were performed for 1 or 30 min (as indicated) with 2 mM $MgCl_2$ (optimal for extension rate (see Figure 11) in the absence (Ø) or in the presence of increasing amounts of $ZnCl_2$. Even very small amounts of Zn^{2+} (0.5-2 μM, first 3 lanes following the Zn^{2+} free reaction) had a detectable effect on the rate of nucleotide addition (based on results with 1 min reactions) to the primer and a more dramatic affect was observed at 5 μM, even thought the ratio of Mg^{2+}:Zn^{2+} was 400:1 at this concentration. Severe inhibition was observed between 25 and 800 μM Zn^{2+}. Both the maximum length of extension products and the extent of RT pausing were altered in the presence of Zn^{2+}

with the former decreasing and the latter increasing. Reactions performed for 30 min indicated that the inhibition caused by low amounts of Zn^{2+} (0.5-25 μM) did not prevent the production of fully extended products although the quantity of those products was reduced at the expense of more intense pausing. The profile of synthesis in the presence of 2 mM Mg^{2+} plus 400 μM Zn^{2+} compared to 400 μM Zn^{2+} alone (lanes 3 and 5) were similar, suggesting that these reactions were being driven almost solely by Zn^{2+}.

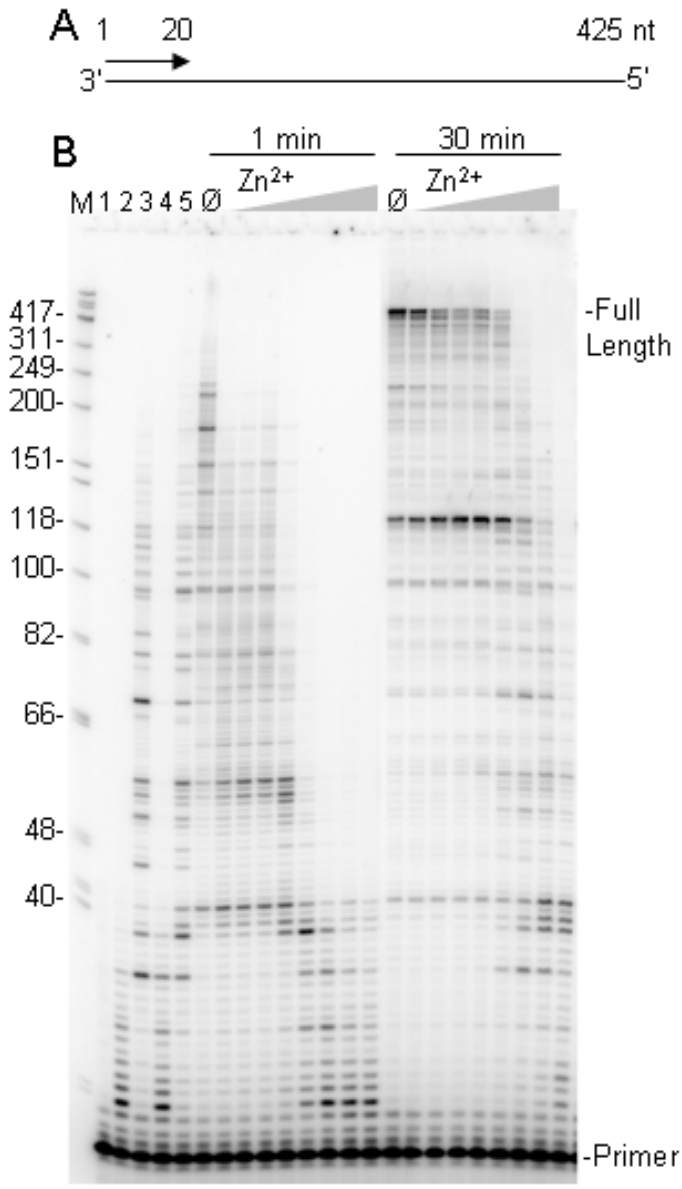


Figure 8: Reverse transcription with 2 mM Mg^{2+} and increasing Zn^{2+} concentrations. (A) Schematic representation of the primer-template used for assays. Numbers indicate size in nts. (B) Primers labeled at 5′ end were hybridized to the RNA template and reactions were performed in the presence of 2 mM $MgCl_2$ and increasing concentrations of $ZnCl_2$. Reactions were stopped after 1 min or 30 min as indicated and electrophoresed on a 8% denaturing gel. The concentration of

$ZnCl_2$ in reactions was 0.5, 1, 2, 5, 25, 100, 400, 800 μM. Lane M, size marker in nts as indicated; lane 1, no enzyme; lane 2 and 3, 400 μM $ZnCl_2$ with no $MgCl_2$ for 1 and 30 min, resp.; lane 4 and 5, 400 μM $ZnSO_4$ with no $MgCl_2$ for 1 and 30 min, resp.; lane Ø, 2 mM $MgCl_2$ with no $ZnCl_2$. Positions of the 20 nt primer and full extension products are indicated.

2.3.2 Zn^{2+} is a more potent inhibitor of HIV-RT than Ca^{2+} and Mn^{2}

Both Ca^{2+} and Mn^{2+} have also been shown to inhibit HIV-RT (see Introduction). These physiologically relevant divalent cations were compared to Zn^{2+} in an inhibition assay in the presence of 2 mM Mg^{2+} as described above but in 2 min reactions (Figure 9). The Zn^{2+}, Ca^{2+}, and Mn^{2+} were at 0, 7.8, 31, 125, 500, or 2000 μM. All three cations inhibited extension but Zn^{2+} was much more potent than the other two. The length of the longest extended products had decreased by ~50% in the presence of just 7.8 μM Zn^{2+} while ~500 μM Ca^{2+} or Mn^{2+} was required to observe the same inhibition. In general, inhibition of extension was accompanied by an increase in pausing rather than a decrease in the amount of total primers that were extended. At the higher concentrations some decrease in total primer extension was also observed. This was especially evident at 2000 μM in reactions with Zn^{2+} and Mn^{2+}.

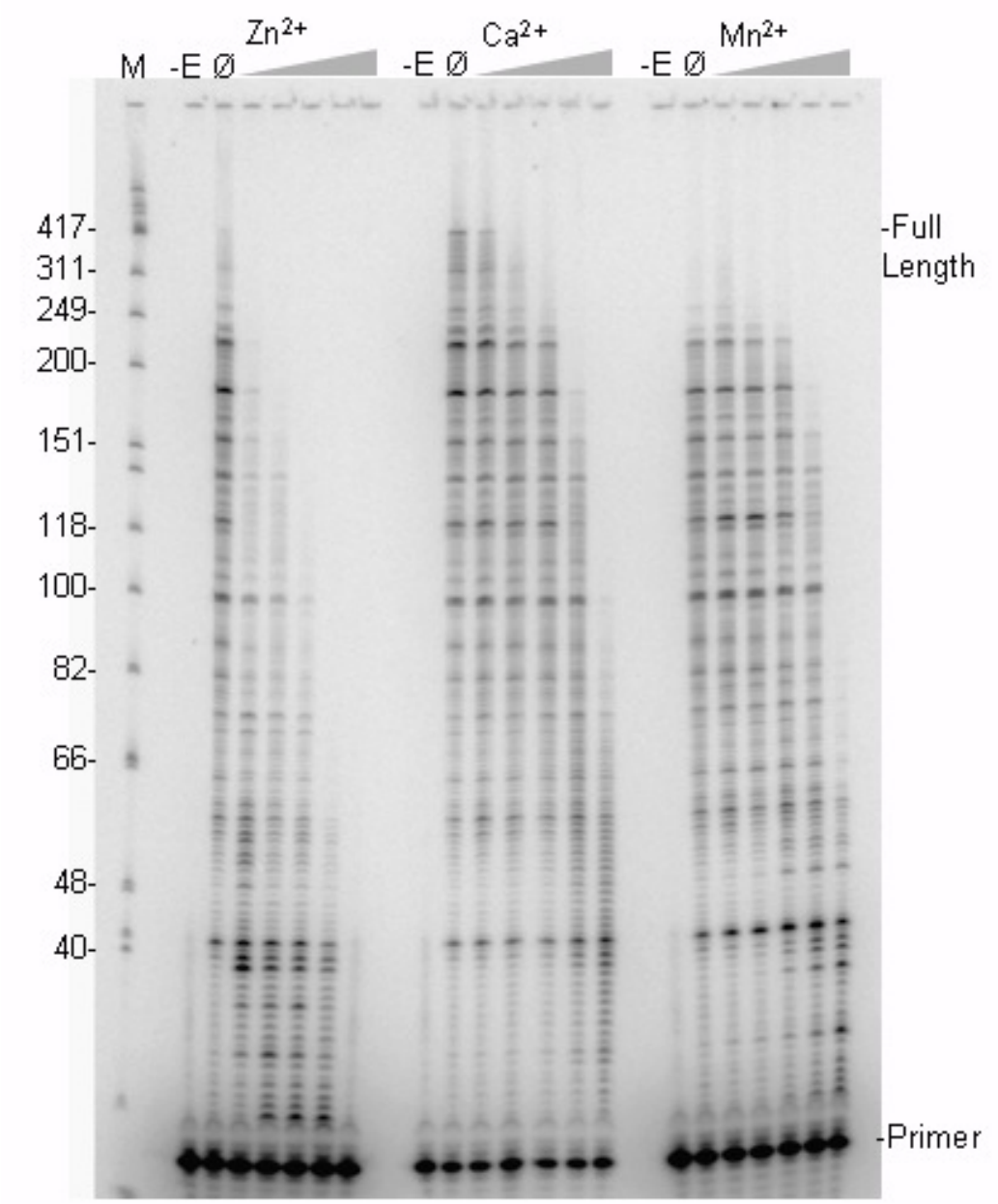


Figure 9: Inhibition of reverse transcription with 2 mM Mg^{2+} by increasing concentrations of Zn^{2+}, Ca^{2+} and Mn^{2+}. Reactions were performed on the primer-template shown in Figure 8A for 2 min with 2 mM $MgCl_2$ and increasing concentrations of the indicated divalent cation (all chloride salts) as follows: 7.8, 31, 125, 500, and 2000 μM. Lanes and markings are as indicated in Fig. 1B with the exception of the "-E" lane which corresponds to no enzyme added.

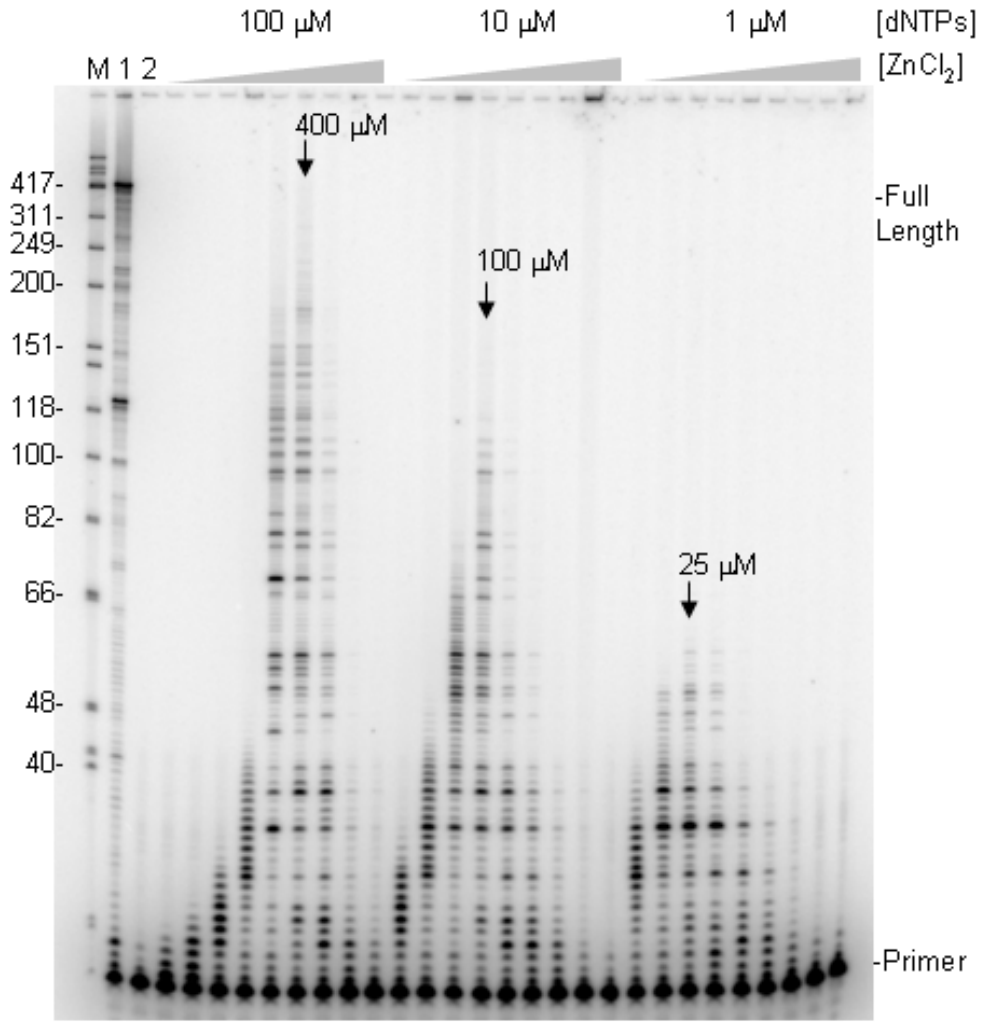


Figure 10: Reverse transcription with Zn^{2+} alone and various dNTP concentrations. Reactions were performed on the primer-template shown in Figure 8A for 30 min with increasing concentrations of $ZnCl_2$ as follows: 5, 10, 25, 100, 200, 400, 800, 1500, and 3000 μM. The concentration of each of the 4 equimolar dNTPs in the reactions is noted above the lanes. Lane M, refer to Figure 8B; lane 1, 30 min reaction with 2 mM $MgCl_2$ and 100 μM dNTPs; lane 2, reaction with no enzyme. Other markings are as indicated in Figure 8B.

2.3.3 Zn^{2+} supports the RNase H activity of HIV-RT at concentrations similar to those observed for nucleotide extension

Reactions performed in the absence of Mg^{2+} showed that Zn^{2+} supports RT synthesis at a highly reduced incorporation rate (Figure 10). Even at 5 μM in these 30 min reactions some extension of the primer was observed at all dNTP concentrations, indicating that Zn^{2+} binds with high affinity to RT. The maximum rate of nucleotide incorporation was observed at 400 μM Zn^{2+} when 100 μM equimolar dNTPs (400 μM total) were used. An assay performed with a homologous DNA template yielded the same optimal concentration and similar profile with respect to extension rate vs. the concentration of Zn^{2+} (data not shown). Since nucleotides can chelate divalent cations (36) and Zn^{2+} catalyzed synthesis peaks at a low concentration, the optimal determined with higher dNTP concentration is unlikely to reflect the actual affinity of Zn^{2+} for RT, as most Zn^{2+} would be complexed with nucleotides under these conditions. Reactions performed with 10 or 1 μM dNTP concentrations were optimal at ~100 and ~25 μM Zn^{2+}, respectively. There was a decline in extension length with lower dNTP concentrations indicating that the Zn^{2+} catalyzed reactions require high nucleotide concentrations to proceed at a maximal rate. Overall the results with the lowest dNTP concentration (1 μM) where very little Zn^{2+} is sequestered indicated that HIV-RT has a very high affinity for Zn^{2+}. For comparison purposes, a titration with Mg^{2+} and Mn^{2+} (Ca^{2+} showed no catalytic activity at any concentrations tested (data not shown)) at 100 μM dNTPs was conducted (3 min reactions in each case) (Figure 11). Like Zn^{2+}, Mn^{2+} showed optimal synthesis at a low concentration (~250 μM) indicating a very high affinity for

RT as most of the free Mn^{2+} would have been otherwise chelated under these conditions. This is consistent with previous results indicating that there is at least one high affinity binding site for Mn^{2+} on RT (39). In contrast, Mg^{2+} demonstrated a broad activity peak with an optimal at 2 mM. Some synthesis with Mn^{2+} was observed all the way down to 3.9 μM while synthesis with Mg^{2+} was first observed at 31 μM. At 2 mM Mg^{2+} and 400 μM total dNTPs the free concentration of Mg^{2+} in the reactions would be about ~1.6 mM (60) indicating a much higher concentration optimum than Zn^{2+} or Mn^{2+}. These results are consistent with Zn^{2+} showing strong inhibition of reactions with Mg^{2+} even when present at concentrations that were ~100-fold lower than Mg^{2+} (Figure 8 and Figure 9). Inhibition by Mn^{2+} appeared less potent even though it apparently binds RT as tightly as Zn^{2+}, probably due to the much more efficient synthesis of HIV-RT in the presence of Mn^{2+} compared to Zn^{2+}. For example, the maximum extension rate for reactions performed under optimal conditions (100 μM dNTPs with 2 mM, 400 μM, and 250 μM divalent cation, for Mg^{2+}, Zn^{2+}, and Mn^{2+}, respectively) was determined by calculating the lengths of the longest products in reactions that had not proceeded to the end of the template, and dividing by the reaction time in seconds. Values of 3.5, 1, and 0.1 nt/sec were determined for Mg^{2+}, Mn^{2+}, and Zn^{2+} respectively.

An additional control was also performed in these assays and is shown in lane 2 of Figure 11 and lane 2 of Figure 12B. In these assays RT was incubated with the substrate but divalent cation was omitted from the reactions. Depending on the purification procedures, small amounts of divalent cation can co-purify with enzymes. This could lead to low levels of activity in reactions performed in the absence of

additional divalent cation. No activity was observed when divalent cation was omitted from the reactions indicating that both polymerase (Figure 11) and RNase H (Figure 12B) activity are dependent on the addition of divalent cation.

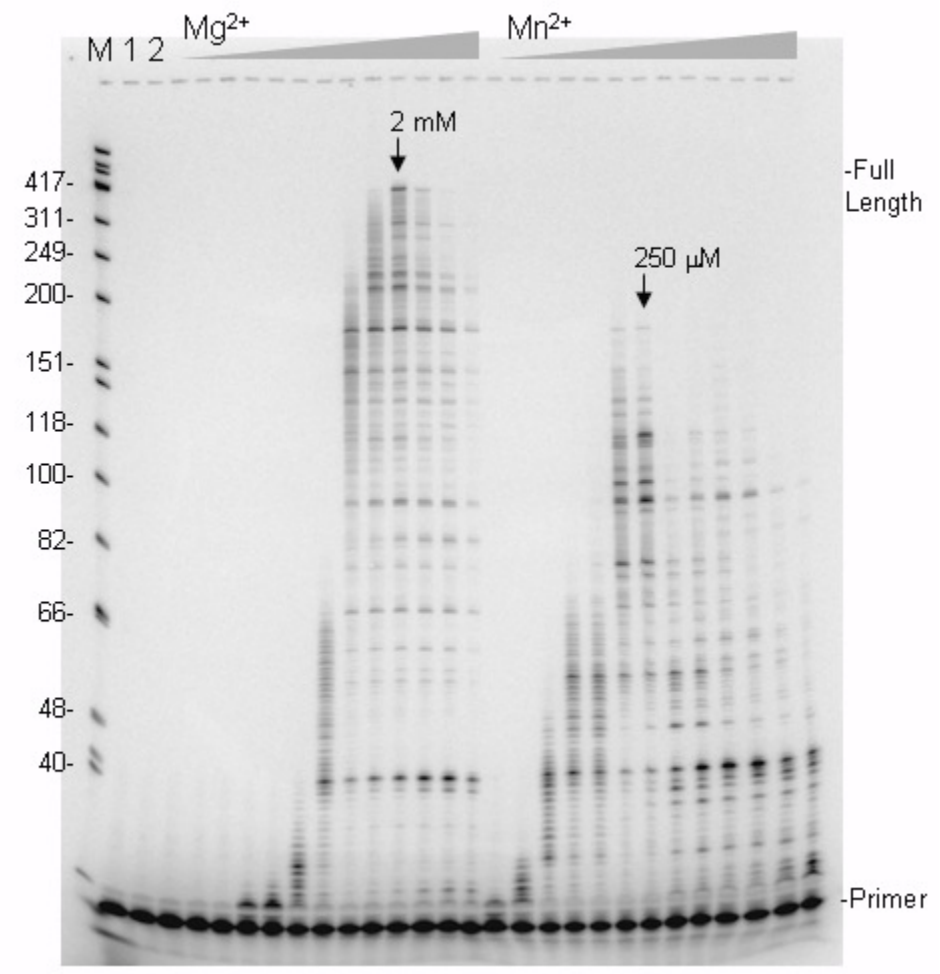


Figure 11: Reverse transcription with various concentrations of Mg^{2+} and Mn^{2+}. Reactions were performed on the primer-template shown in Fig. 1A for 3 min with 100 μM each dNTP and increasing concentrations of $MgCl_2$ or $MnCl_2$ as follows: 3.9, 7.8, 15.6, 31.3, 62.5, 125, 250, and 500 μM, followed by 1, 2, 4, 8, and 16 mM. Lane M, refer to Fig. 1B; lane 1, no enzyme; lane 2, enzyme without additional divalent cation. Lanes where optimal synthesis occurred are denoted on the figure with an arrow. Other markings are as indicated in Figure 8B.

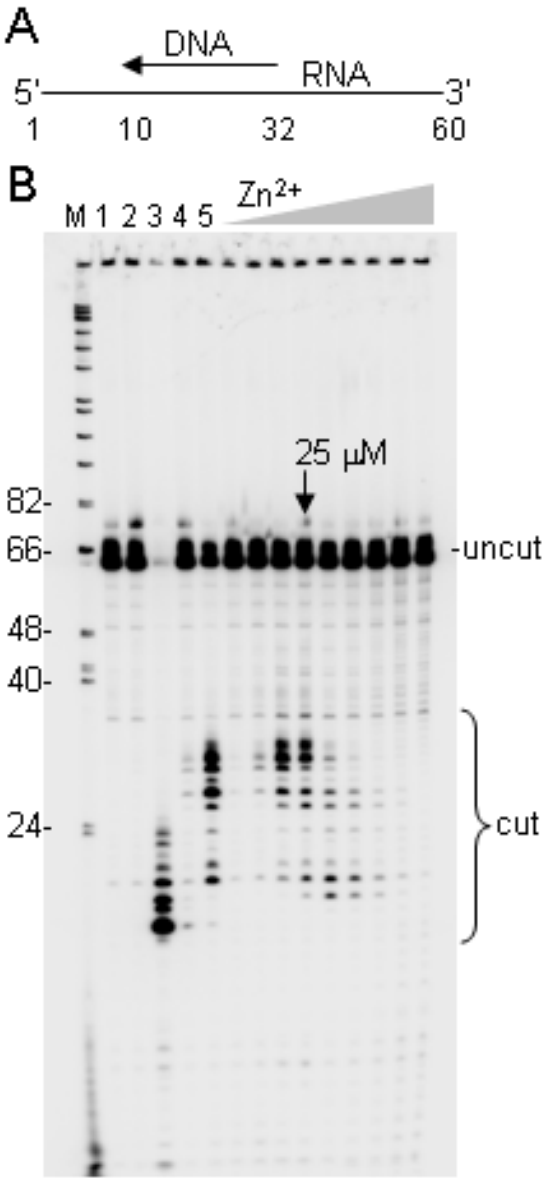


Figure 12: RNase H digestion with HIV-RT in the presence of varying concentrations of Zn^{2+}. (A) Schematic diagram of the primer-template. A 5′-labeled 60 nt RNA is hybridized to a 23 nt DNA primer. (B) In order, control lanes refer to: 1) no enzyme, 2) enzyme without additional divalent cation, and 3) *E. coli* RNase H digestion of substrate. Lanes labeled 4 and 5 contain 2 mM $MgCl_2$ and no $ZnCl_2$. Lane 4 was digested for 5 sec, all other lanes are 5 min digests. Zn^{2+}

concentrations were 2, 5, 10, 25, 100, 200, 400, 800, and 1200 μM. The concentration for optimal RNase H activity is noted with an arrow. Positions of uncut 60 nt RNA and cleavage products are indicated. Lane M, refer to Figure 8B.

2.3.4 HIV-RT forms a very stable catalytically competent complex with primer-template in the presence of Zn^{2+} that traverses the template slowly but can remain associated for hours

Since the extension rate of RT was ~35 times greater in the presence of Mg^{2+} compared to Zn^{2+} (see above), it was reasonable to assume that processivity (average numbers of nucleotides added to the primer in a single binding event between RT and the substrate) would be much greater with Mg^{2+}. Surprisingly, this was not the case. Reactions to test processivity were performed on the primer-template shown in Fig. 1A by pre-binding RT to the preformed primer-template then adding dNTPs and divalent cation along with an excess of poly(rA)-oligo(dT) to sequester enzymes that were not bound to or had dissociated from the primer-template. This limits synthesis to a single binding even between the enzyme and substrate (61). Reactions with 2 mM Mg^{2+} (left panel, Figure 13) showed the expected profile. Primers were extended at a maximum rate of ~3.5 nt/sec with some extension products reaching the end of the 425 nt template by 2 min and more by 4 min. The extension profile of reactions performed for 8 and 16 min looked very similar to those performed for 4 min indicating that the single round of enzyme extension was essentially complete in 4 min. In contrast, the first time point shown for the reactions with 400 μM Zn^{2+} was 15 min. In this case reactions proceeded at the expected ~0.1 nt/sec and some products reached the end of the template by about 1 hr. Fully extended products continued to accumulate even up to the final 4 hour time point, although there was just a small increase between 3 and 4 hours and those profiles looked very similar. It was remarkable that single enzyme molecules could stay associated with a single

template for hours. Although the maximum rate of synthesis was ~0.1 nt/sec, on average polymerases moved much more slowly, at a calculated rate of 0.022 ± 0.003 nt/sec for Zn^{2+} as compared to 1.7 ± 0.4 nt/sec for Mg^{2+} reactions (results are an average ± S.D. from 3 or more reactions using 30 min and 1 min time points for 400 μM Zn^{2+} and 2 mM Mg^{2+}, respectively; see Methods for calculation). Despite this, the final processivity on this template was only modestly higher for Mg^{2+}.

To test the stability of the enzyme-substrate complexes further, the half-life ($t_{½}$) of RT for binding to a DNA-DNA primer-template was determined in the presence of Zn^{2+}, Mn^{2+} and Mg^{2+} (Figure 14). This experiment required the use of DNA-DNA since off-rates in the presence of divalent cation cannot be easily measured with RNA-DNA except with RNase H mutant enzymes. Remarkably, complexes with Zn^{2+} and Mn^{2+} were ~130 ($t_{½}$=220±60 min) and ~45 ($t_{½}$=77±1 min) times more stable than those with Mg^{2+} ($t_{½}$=1.7±1.0 min). This is consistent with the ability of single enzyme molecules to remain bound to the template for such a long time in the processivity experiments (Figure 13). It is important to note that the dramatic differences in $t_{½}$ observed with DNA-DNA only provide a qualitative comparison between affinity differences that may occur during DNA synthesis on an RNA template or even a DNA template (see Discussion). Still the results clearly suggest that Zn^{2+} and Mn^{2+} lead to more stable RT-primer-template complexes in comparison to the natural divalent cation, Mg^{2+}.

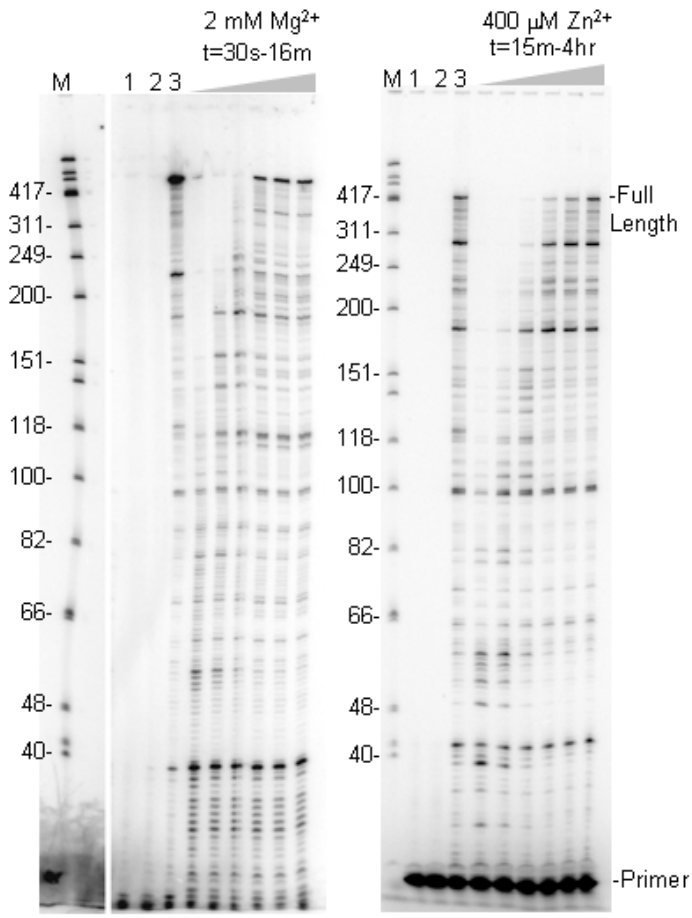
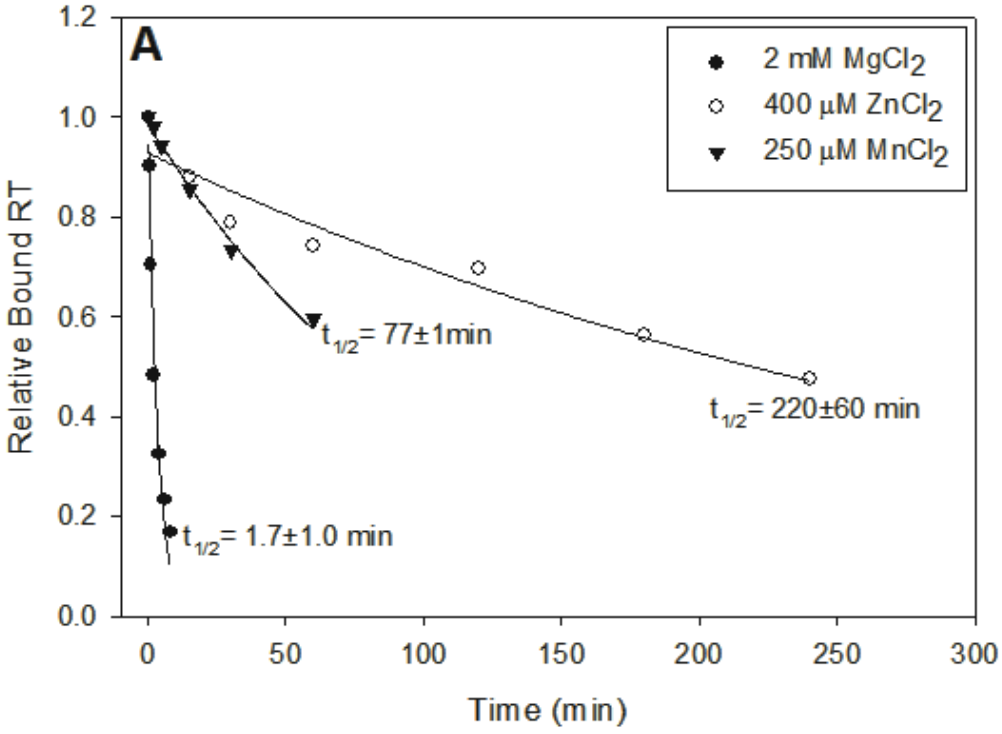


Figure 13: HIV-RT processivity with Mg^{2+} and Zn^{2+}. Extension with 2 mM $MgCl_2$ or 400 µM $ZnCl_2$ was performed with the primer-template in Fig. 1A and in the presence of a non-radiolabeled poly(rA)-oligo(dT) trap which limits primer extension to a single binding event between RT and the primer-template. Lane 1, no enzyme; lane 2, trap control: enzyme was mixed with the trap then incubated with primer-template in standard reaction conditions for 16 min or 4 hr for Mg^{2+} and Zn^{2+} reactions, resp.; lane 3, primer extension in the absence of trap for 16 min or 4 hours for Mg^{2+} and Zn^{2+} reactions, respectively. Time points for trap reactions with $MgCl_2$ were 30s, 1, 2, 4, 8, and 16 min, and for $ZnCl_2$, 15 and 30 min, followed by 1, 2, 3, and 4 hours. Refer to Figure 8B for other markings.


A
2 mM MgCl2
400 µM ZnCl2
250 µM MnCl2
Relative Bound RT
Time (min)
t1/2= 77±1min
t1/2= 220±60 min
t1/2= 1.7±1.0 min
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
50
100
150
200
250
300

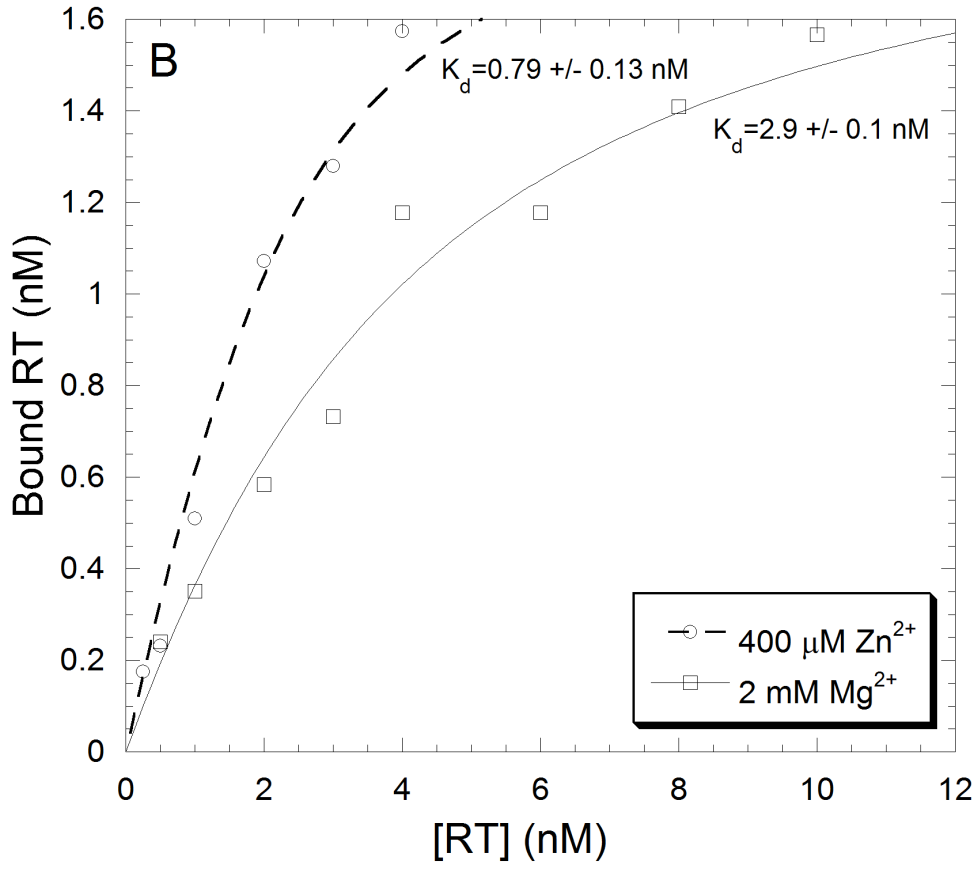

B
Kd=0.79 +/- 0.13 nM
Kd=2.9 +/- 0.1 nM
Bound RT (nM)
[RT] (nM)
400 µM Zn2+
2 mM Mg2+
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
0
2
4
6
8
10
12

Figure 14: Plot of RT-primer-template complex binding stability in the presence of Mg^{2+}, Mn^{2+}, and Zn^{2+}. Half-lives are indicated with their respective curve. Results are and average of 3 exp. ± standard deviations. Experiments were performed as described under "Experimental Procedures", using optimal concentrations for extension.

2.3.5 Other viral RTs from AMV and MuLV as well as an RNase H minus HIV mutant (HIV-RT$^{E478>Q}$) can also use Zn^{2+} for nucleotide incorporation

Nucleotide incorporation was also tested with 3 other RTs including an RNase H minus form of HIV-RT (HIV-RT$^{E478>Q}$) and two other viral RTs, AMV-, and MuLV-RT (Figure 15). Not only did all 3 RTs show activity, but peak activity occurred at ~400 μM Zn^{2+} in all cases (assay were with 400 μM total dNTPs). The similar results likely derive from common active site architecture in these closely related RTs.

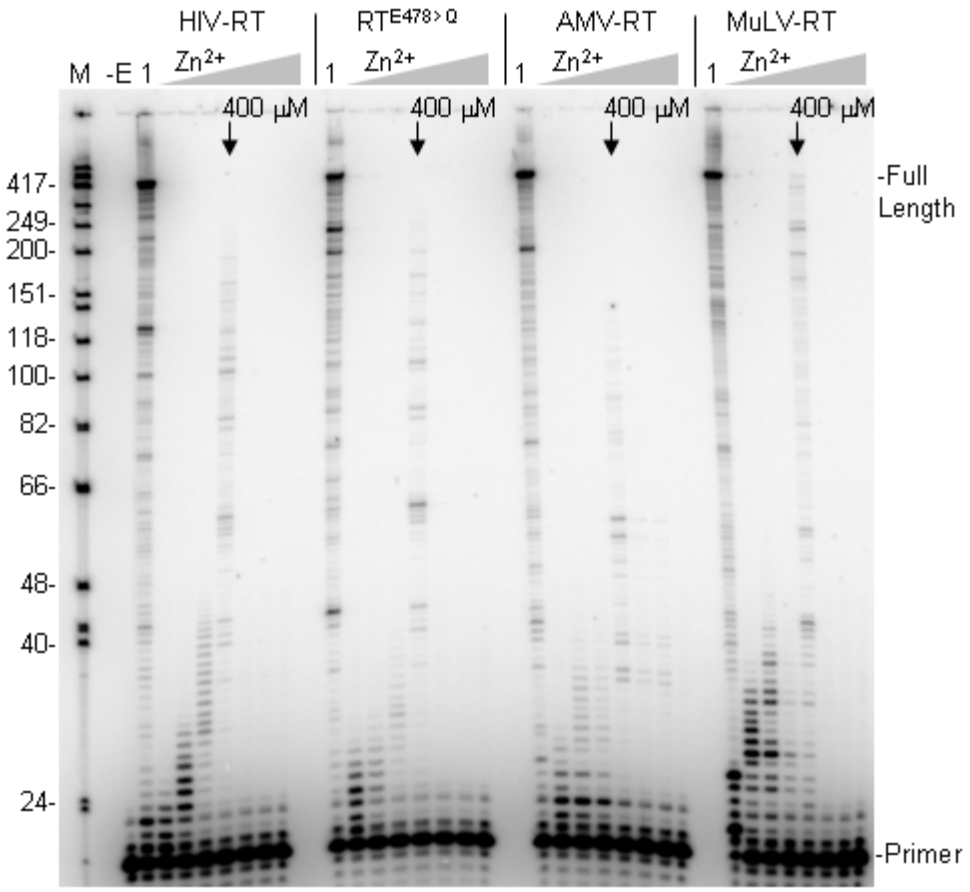


Figure 15: Reverse transcription with various RTs in the presence of Zn^{2+}. Extension was performed with the primer-template from Figure 8A with the indicated enzyme. Reactions were for 30 min and $ZnCl_2$ concentrations were as follows: 100, 200, 400, and 800 µM, followed by 1.5 and 3 mM. Lane M is as in Figure 8B. Optimal concentrations for extension are noted for each enzyme with an arrow. Lane –E, no enzyme control; lane 1, extension with the indicated enzyme for 30 min with 2 mM $MgCl_2$. Other markings are as indicated in Figure 1B.

2.4 Conclusions

As with Ca^{2+} and Mn^{2+}, previous reports have also shown that Zn^{2+} inhibits HIV-RT nucleotide incorporation in the presence of Mg^{2+} (see 2.1 Introduction). The mechanism of inhibition has been best studied using Mn^{2+}. Bolton and colleagues (39) have proposed that Mn^{2+} inhibits TY1 and HIV-RT by binding with high affinity and displacing Mg^{2+} at one of two proposed divalent cation binding sites near the polymerase active site. The authors propose that the two sites do not equally affect catalysis and the site that Mn^{2+} binds tightly to is dominant. Therefore, even if Mg^{2+} occupies the second site, RT reverts to a Mn^{2+} catalysis mode similar to what is observed in the presence of Mn^{2+} alone. This mode is much less efficient with respect to the rate of incorporation. The model is supported by mutational analysis as well as the observance of two divalent cations in the polyrmerase domain of HIV-RT in some crystals (62). Interestingly, TY1 retrotransposon mutants that could overcome Mn^{2+} inhibition mapped not to the polymerase but to the RNase H domain (63). This led the authors to propose a model where mutations in the RNase H domain could influence polymerization through allosteric effects. The fact that mutations in the polymerase domain conferring resistance were not found does not necessarily mean that this domain is uninvolved in the observed inhibition by Mn^{2+}. It is possible that mutations which alter Zn^{2+} binding may have a severe impact on RT activity and are therefore not selected.

The model proposed for Mn^{2+} is consistent with what was observed for Zn^{2+} in the current experiments, although our findings do not necessarily support this model vs. a simpler model where Zn^{2+} or Mn^{2+} (or even Ca^{2+}) simply displace Mg^{2+} from

one or more binding site in the polymerase domain that control catalysis. What was clear from the current results is that Zn^{2+} is a more potent inhibitor of Mg^{2+} catalysis than Mn^{2+} or Ca^{2+}, as it is required at a lower concentration to inhibit extension (Figure 9). This is probably not due to a higher affinity of Zn^{2+} for RT as the optimal concentrations for Zn^{2+} and Mn^{2+} catalysis in the absence of Mg^{2+} were comparable (~400 μM and 250 μM for Zn^{2+} and Mn^{2+}, resp., in the presence of 400 μM total dNTPs). Primer extension in the presence of Mn^{2+} was, however, much more efficient than with Zn^{2+} (Figure 8 and Figure 11, note time difference in reactions). The ability of polymerases to use Mn^{2+} in place of Mg^{2+} *in vitro* has been well documented and some polymerases function as well or even better with Mn^{2+}, although it tends towards lower polymerase fidelity (28). Since experiments reported here essentially evaluated polymerization efficiency by the rate of extension and the length of the products that were produced, the greater inhibition of Mg^{2+} catalysis by Zn^{2+} probably resulted from Zn^{2+} producing a more catalytically deficient enzyme. In contrast to Zn^{2+} and Mn^{2+}, no catalysis was observed with Ca^{2+}, while the level of inhibition of Mg^{2+} catalysis was similar to Mn^{2+}. RT's inability to use Ca^{2+} for catalysis is not surprising as it is not typically a co-factor in catalysis for enzymes (see 2.1 Introduction). Specific properties of Ca^{2+}, which differ markedly from Mg^{2+}, make it a poor candidate for nucleotide catalysis by polymerases (discussed in (64)). Ca^{2+} is also known to inhibit other DNA polymerases (65) and has been used to help stabilize crystal structures of some polymerases by preventing catalysis during crystal formation. Some crystals made with Ca^{2+} show this ion bound at the pol active site suggesting that it could inhibit polymerization by displacing Mg^{2+}. One polymerase,

Sulfolobus solfataricus DNA polymerase, has been reported to be able to use Ca^{2+} for nucleotide polymerization, though less efficiently than Mg^{2+} (66).

Since RNase H catalyzed cleavage of the RNA-DNA hybrid was observed in the presence of Zn^{2+} alone, it was clear that this cation can bind to the RNase H active site and stimulate cleavage in a manner similar to Mn^{2+} (63). Also notable was the ~10-25 μM Zn^{2+} optimal for cleavage (Figure 12), which was very close to the optimal concentration of Zn^{2+} for polymerization in the presence of low (1 μM) concentrations of dNTPs (Figure 10). This suggests that the affinity for Zn^{2+} at the catalytic binding sites in the polymerase and RNase H domains are similar, and that Zn^{2+} binds with greater affinity at these sites than Mg^{2+} which requires much higher concentrations for optimal activity (Figure 11). Interestingly, all the RTs that were tested including HIV, $HIV^{E478>Q}$, AMV and MuLV showed similar optimal Zn^{2+} concentrations for nucleotide incorporation. This suggests that the active sites of these RTs have similar affinity for Zn^{2+}.

HIV-$RT^{E478>Q}$ is a commonly used enzyme for studying polymerization properties in the absence of RNase H activity. The glutamic acid to glutamine mutation in the RNase H active site abrogates RNase H activity and also disrupts a divalent cation binding site in the RNase H domain. This enzyme has polymerization properties that essentially mimic wild type enzyme, although it does have altered binding characteristics to primer-template in the absence or at low concentrations of Mg^{2+} (67, 68). Polymerization with HIV-$RT^{E478>Q}$ occurred in the presence of Zn^{2+} with a concentration optimum and profile that was similar to wild type RT (Figure 15). Inhibition of catalysis in the presence of Mg^{2+} by Zn^{2+} essentially mimicked that

in the wild type enzyme (data not shown). Overall the results suggest that the mutation in the RNase H domain had no effect and an allosteric interaction of the type reported for TY1 RT (see above) was not observed. It is important to note that although the E478>Q mutation would presumably alter divalent cation binding in the RNase H domain, it is not clear how it would be changed and if the mutation would abrogate the binding of one or more cations. Therefore the results do not rule out potential allosteric interactions of the type proposed by others.

The dramatic increase (>100 fold) in the stability of RT-Zn^{2+}-primer-template compared to RT-Mg^{2+}-primer-template complexes was surprising. Importantly, the method for measuring dissociation did not directly reflect dissociation that occurs during active DNA synthesis. No dNTPs were present and the polymerase was not cycling thorough the various phases of nucleotide incorporation (primer-template binding, nucleotide binding, catalysis and translocation, then dissociation). Each of these phases may have different dissociation rates. Reports have also shown that binding of RT to nucleotides can stabilize binding (69). Finally, DNA-DNA primer-templates were used in the off-rate experiments as RNA-DNA cannot be used since it is degraded by RNase H activity in the presence of divalent cation. Experiments have shown that in the presence of Mg^{2+}, HIV-RT binds more stably to both RNA-DNA and DNA-DNA (67), so the results may under-represent the stability of Mg^{2+} complexes. Processivity measurements can be used as an indicator of dissociation during active synthesis although they cannot be used to directly calculate a dissociation rate. The processivity is dependent on several factors including the sequence/structure and type (RNA or DNA) of nucleic acid used, as well as the

polymerase dissociation and incorporation rates. Processivity was greater in the presence of Mg^{2+} than with Zn^{2+} as the average length of extended products was clearly greater with Mg^{2+}, but the difference was relatively small considering that RT incorporated nucleotides about 35 times more rapidly in the presence of Mg^{2+}. This large difference in incorporation rate is compensated by very stable binding in the presence of Zn^{2+}. With Zn^{2+}, RT molecules were able to remain bound to the primer-template for hours even though nucleotides were being added to the primer at a maximal rate of ~ 1 every 10 seconds and at an average rate that was much less than this (see Results and Figure 14). This essentially means that polymerases stalled for long periods of time after each nucleotide addition but did not dissociate from the template. In contrast, extension with Mg^{2+} was essentially complete after 4 min as paused products remained stable and the level of fully extended primers did not change beyond this time point. Overall the processivity experiment supports the results from the off-rate experiments and indicates that the RT-Zn^{2+}-primer-template complex is extremely stable. The reason for this high stability is unclear. Differing effects of the divalent cations on the template structure cannot be ignored as contributing factors and also the lower concentration of divalent cation that was used in the Zn^{2+} vs. Mg^{2+} experiments may have had some effect. Still, these would not be expected to account for the greater than 100-fold difference in off-rates. Interestingly, RT-Mn^{2+}-primer-template complexes were also more stable than those formed with Mg^{2+}, though to a lesser extent (Figure 14). Likewise, Mn^{2+} also reduced the rate of nucleotide incorporation, again to a lesser extent than Zn^{2+}. This brings about the possibility that these cations may stabilize a particular confirmation

of RT that is conducive to tight binding but not fast incorporation kinetics, or that they stabilize a tight binding confirmation and slower catalysis occurs because of relatively poor chemistry compared to Mg^{2+} (28). For example, they could promote the "productive" vs. "nonproductive" RT conformation postulated by Whörl *et al.* (70), the former of which is proposed to be more stable.

In conclusion, Zn^{2+} inhibits HIV-RT synthesis not by directly stopping catalysis, but by forming a highly stable complex that has very slow incorporation kinetics. The RT-Zn^{2+}-primer-template complex can be thought of as a "dead-end complex" since it ties up RT- potentially for hours - in a complex that is from a kinetic perspective, minimally productive. HIV-RT apparently has higher affinity for Zn^{2+} (and Mn^{2+} and Ca^{2+} also) than Mg^{2+} at one or more of the divalent cation binding sites. Therefore, relatively low concentrations of Zn^{2+} can inhibit RT extension in the presence of much higher concentrations of Mg^{2+}. Since Zn^{2+} has relatively low toxicity (71) and can inhibit RT extension at low concentrations (just a few μM was required to slow extension in the presence of 2 mM Mg^{2+}, Figure 8), it could, at least in theory, be useful for slowing viral replication and zinc-based compounds present an opportunity to exploit a weakness in HIV that might present a novel avenue for future drug development. However the possibly of using supplements or natural minerals including Zn^{2+} must be approached with caution. Even though Zn^{2+} can inhibit RT at low-μM concentrations that are lower than the level of total Zn^{2+} in cells and plasma, they are still ~2-3 orders of magnitude greater than the level of free available Zn^{2+} in cells (see 2.1 Introduction). In addition, obtaining low μM levels inside cells would likely require much higher levels in the

bloodstream unless there was a specific delivery mechanism. Finally, μM concentrations of free Zn^{2+} in cells would almost certainly have profound effects on the transcription of specific genes, the oxidation state of cells, and general heath (see 2.1 Introduction and (71)). Despite these hurdles, the work described in this report could help set the stage for further investigations with other cations or cation combinations that may be more specific for HIV-RT.

Chapter 3: Tightly binding DNA aptamers against Family A DNA polymerases

3.1 Introduction

3.1.1 SELEX and Aptamers

Single-stranded nucleic acids, like amino acid chains in proteins, have the ability to fold into secondary and tertiary structures based on their sequence. *In vivo*, this is crucial for the function of biologically active RNAs, such as ribosomes and tRNAs, but this trait has been exploited to design artificial oligonucleotide and peptide ligands that possess shapes necessary for targeted binding to other molecules. These are called aptamers, from the Latin word *apto* meaning 'to fit'.

Manually designing aptamers is difficult, but in 1990 the Gold and Szostak labs independently developed a technique to repeatedly enrich pools of nucleic acids for sequences that bound tightly to a target molecule, providing a robust method for aptamer generation. The Szostak lab called this *in vitro selection* and isolated ssRNA aptamers with tertiary structures that were able to bind selectively to organic dyes (72). The Gold lab's similar process, Systematic Evolution of Ligands by Exponential Enrichment (SELEX), was used to isolate ssRNAs that bound the T4 DNA polymerase (73).

SELEX relies on the concept that some sequences will naturally fold into secondary or tertiary structures that are able to bind targeted molecules. By creating a large pool of randomized sequences (sometimes as many as 10^{15} molecules), it is

possible that some members of the pool will be able to bind to the target in a complex that allows separation from those that cannot bind; common separation techniques are gel shift, capillary electrophoresis, nitrocellulose binding, or column chromatography, among others. These selected sequences can then be amplified to create a second enriched pool that only contains target-binding sequences. This process is repeated six to twelve times, resulting in an enriched pool of aptamers that bind with high affinity and specificity to the molecule of interest (often these aptamer:target complexes have a dissociation constant, K_d, in the low-nM range).

A general aptamer discovery SELEX protocol begins with a large initial pool of ssRNA, which all contain a random region (usually 20-80 nt) flanked by known fixed sequences (one of which carries an RNA polymerase promoter, usually T7 RNA polymerase). After selection with the target molecule, the ssRNA is amplified by RT-PCR into a DNA pool, using the fixed regions as primer binding sites. These sequences are then transcribed into ssRNA from the promoter, generating a pool with higher affinity for the target. The process is repeated until the pool's binding affinity no longer increases with subsequent rounds of selection (a variation on this is shown in Figure 19).

There are several versions of this protocol, each with advantages and disadvantages:

- Substrates
    - ssRNA SELEX requires an RT step and produces aptamers that are less stable than their DNA counterparts. This can be overcome with chemical modification to reduce autohydrolysis

and protect the aptamer from nuclease digestion, but that increases the cost of the aptamer and may affect binding affinity to the target.

- ssDNA SELEX, which eliminates the RT step but requires the generation of ssDNA, which is more difficult than producing ssRNA. An advantage of DNA aptamers is their inherent stability compared to RNA.
- Xenonucleic acid (XNA) SELEX, which uses synthetic nucleic acid alternatives to RNA or DNA. XNAs contain modifications of the naturally occurring deoxyribose (DNA) or ribose (RNA) sugar backbone, which can enhance stability (as in the case of the RNA analog locked nucleic acid, LNA, which has a methylene bridge connecting the 2′ oxygen atom to the 4′ carbon, preventing hydrolysis) or reduce recognition of the enzyme by nucleases (as seen with anhydrohexitol nucleic acid, HNA). XNA SELEX relies on modified enzymes that can transcribe between XNA and DNA. XNAs are attractive alternatives to DNA/RNA because they are able to avoid detection or destruction by most nucleic acid-recognizing enzymes (74). Unfortunately, SELEX with XNA is cumbersome and is limited by cost and availability of the necessary nucleotides.

- Method
    - Minimal primer SELEX uses shorter fixed sequences to reduce the impact that these regions might have on the selection outcome.
    - Primer-free SELEX eliminates the fixed sequences entirely during the selection process, but requires a ligation step to reintroduce fixed regions to allow for PCR.
    - Primer-template SELEX, a process developed in the DeStefano lab, uses restriction enzymes to create pools of dsDNA with a single recessed 3′ terminus, mimicking the natural substrate primer-template of DNA polymerases (described more fully in Figure 19).

Using these techniques, aptamers against several different targets have been produced: the HIV TAR (75), RT (76), and Env (77), CD4 (78), thrombin (79), prostate specific antigen (PSA) (80), and many others. Due to their sensitivity and specificity, aptamers can perform the same functions as antibodies and are both easier and cheaper to produce. DNA aptamers, in particular, have a longer shelf life than antibodies (RNA aptamers, while inherently less stable, can be protected by chemical modification to increase their resistance to hydrolytic cleavage). Another strength is their potential for reuse: as nucleic acids can be heated to denature their structure, aptamers can be dissociated from the targets and then cooled to regain their active conformation. Due to these benefits, aptamers have been incorporated into biomarker assays and protein precipitation kits, among other applications (reviewed in (81)).

Currently, only a single aptamer-based treatment has been approved for use by the FDA: Macugen, an aptamer targeting the vascular endothelial growth factor (VEGF), which is used to treat age-related macular degeneration (82). Other therapeutics are also being tested (Table 3).

Name	Type	Target	Condition	Phase	Company
Pegaptanib (Macugen)	RNA	VEGF	Macular Degeneration	FDA approved	Pfizer/Eyetech
AS1411	DNA	Nucleolin	Acute Myeloid Leukemia	Phase II	Antisoma Research
REG1	RNA	Coagulation Factor IX	Coronary Artery Disease	Phase II	Regado Biosciences
NOX-E36	RNA	CCL2	Type 2 Diabetes Mellitus	Phase I	NOXXON Pharma AG
ARC1779	DNA	vWF	Purpura, Thrombotic Thrombocytopenic	Phase II	ARCA biopharma

Table 3. Selected aptamers in clinical trials for the treatment of disease. Macugen is the only FDA approved aptamer medication, but other compounds are currently being tested. Table modified from (83).

3.1.2 PT-SELEX against HIV RT

Several aptamers have been designed against HIV RT for use as ARVs. These aptamers are able to inhibit HIV RT reverse transcription and are inhibitory to viral replication in cell culture (84, 85). The inhibition of reverse transcription is likely due to the structure of the single stranded nucleic acids, as well as the actual sequence, although structure probably plays a more significant role.

Like other DNA polymerases, the preferred substrate for RT is double-stranded nucleic acid with a recessed 3′ terminus (a primer-template complex). RT shows a preference for DNA-primer:RNA-template over DNA:DNA hybrids, but was assumed, like other DNA polymerases, to not have a strong preference for specific sequences. A paper published by our lab in 2006 demonstrated that HIV RT did, in fact, exhibit preferential binding to specific primer-template sequences (86). By using PT SELEX, it was discovered that HIV RT has a strong preference for sequences mimicking the PPT region of the HIV genome, specifically the six nucleotide run of Gs at the 3′ end. The sequences recovered from PT SELEX all had a 6-8 nucleotide run of Gs at the primer 3′ terminus, and this was all that was necessary for tight binding. Subsequent PT SELEX experiments determined that other viral RTs (Moloney murine leukemia virus RT, MuLV RT, and avian myeloblastosis virus RT, AMV RT) also showed a similar preference for their own PPTs and would select products that contained 3′ runs of Gs, as well (Table 4) (87).

Aptamer	Sequence	K_d (nM)
	HIV RT	
DNA PPT	5′(primer)CTTTTTAAAAGAAAAGGGGGG 3′(primer)GAAAAATTTTCTTTTCCCCCCTGAC	3.2 ± 0.2
SELEX product	5′(primer)AAACTGAAGAGGCACCGAAGGGGGGG 3′(primer)TTTGACTTCTCCGTGGCTTCCCCCCCATAA	5.0 ± 1.0
Random	5′(primer)NNNNNNNNNNNNNNNNNNNN 3′(primer)NNNNNNNNNNNNNNNNNNNNNNNN	106 ± 11
	MuLV RT	
DNA PPT	5′(primer)TAGTCTCCAGAAAAAGGGGGG 3′(primer)ATCAGAGGTCTTTTTCCCCCCTTAC	7.5 ± 2.5
SELEX product	5′(primer)ATAAGGAAGATCTGCGGGGGG 3′(primer)TATTCCTTCTAGACGCCCCCCTTTC	8.1 ± 2.0
Random	5′(primer)NNNNNNNNNNNNNNNNNNNN 3′(primer)NNNNNNNNNNNNNNNNNNNNNNNN	116 ± 23
	AMV RT	
DNA PPT	5′(primer)GCTTTTGCATAGGGAGGGGGA 3′(primer)CGAAAACGTATCCCTCCCCCTTTAC	2.1 ± 0.4
SELEX product	5′(primer)GATAAGGAAGATCTGCGGGGGG 3′(primer)CTATTCCTTCTAGACGCCCCCCTTTC	2.3 ± 0.7
Random	5′(primer)NNNNNNNNNNNNNNNNNNNN 3′(primer)NNNNNNNNNNNNNNNNNNNNNNNN	20 ± 0.3

Table 4. Binding affinities of various viral RTs for their PPT and SELEX products. PT SELEX with RTs from HIV, MuLV, and AMV yield tightly binding sequences that contain runs of Gs that match the 3′ end of the retroviral PPT. These aptamers bind to their respective enzymes with affinities similar to the full PPT and 1-2 orders of magnitude better than a random substrate (87).

The PPT is resistant to RNase H degradation, which allows its use as a primer for second strand DNA synthesis (88-92). RNase H-resistance alone might be the only reason it is used as a primer by RT, but it is also a more efficient primer than random RNA sequences, suggesting a unique interaction with RT (93, 94). The fact that RT binds this sequence more tightly than other, non-PPT sequences is therefore not surprising.

Given that RT not only has a sequence-binding preference, but also one that is biologically relevant, it is possible that other primer-template-utilizing polymerases also have preferences for specific sequences.

3.1.3 Family A DNA polymerases

RT belongs to the broader group of enzymes called DNA polymerases, which are used by all cellular organisms (and many viruses) for replication of their genetic material. To accomplish genome replication, these enzymes must be able to accurately incorporate nucleotides into new DNA. DNA polymerases can vary significantly in their error rates and processivity (the average number of nucleotides added by the enzyme during a single binding event), as well as in their structure and amino acid composition. While diverse, these enzymes can be grouped by sequence and structural homology into seven families (A, B, C, D, X, Y, and RT; Figure 16), all of which share several analogous regions that are necessary for their function (95).

- A catalytic palm domain, which directs the phosphoryl transfer reactions during nucleotide incorporation.
- The finger domain, which aids in recognizing and binding the incoming dNTP.

- A thumb domain that helps position the nucleic acid template.
- A two-metal-ion binding site in the catalytic cleft, described in Chapter 1.

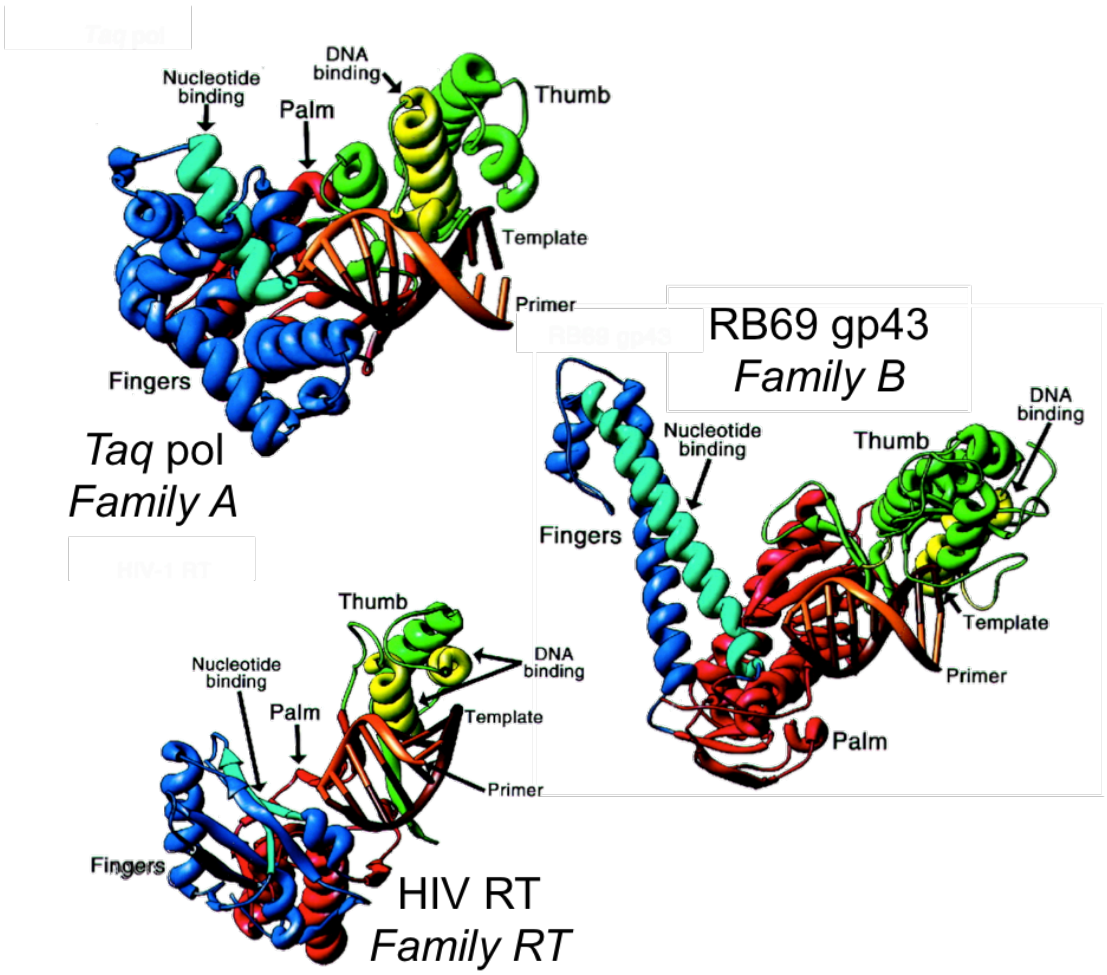


Figure 16 Crystal structures of DNA polymerases bound to DNA. *Taq* pol (Family A) from *Thermus aquaticus*, gp43 (Family B) from bacteriophage RB69, and RT (Family RT) from HIV have analogous finger, palm, and thumb domains which give the enzyme a right-handed conformation. Despite their superficial similarity, the finger and thumb domains between the three enzymes are significantly different in both sequence and structure. Adapted from (95).

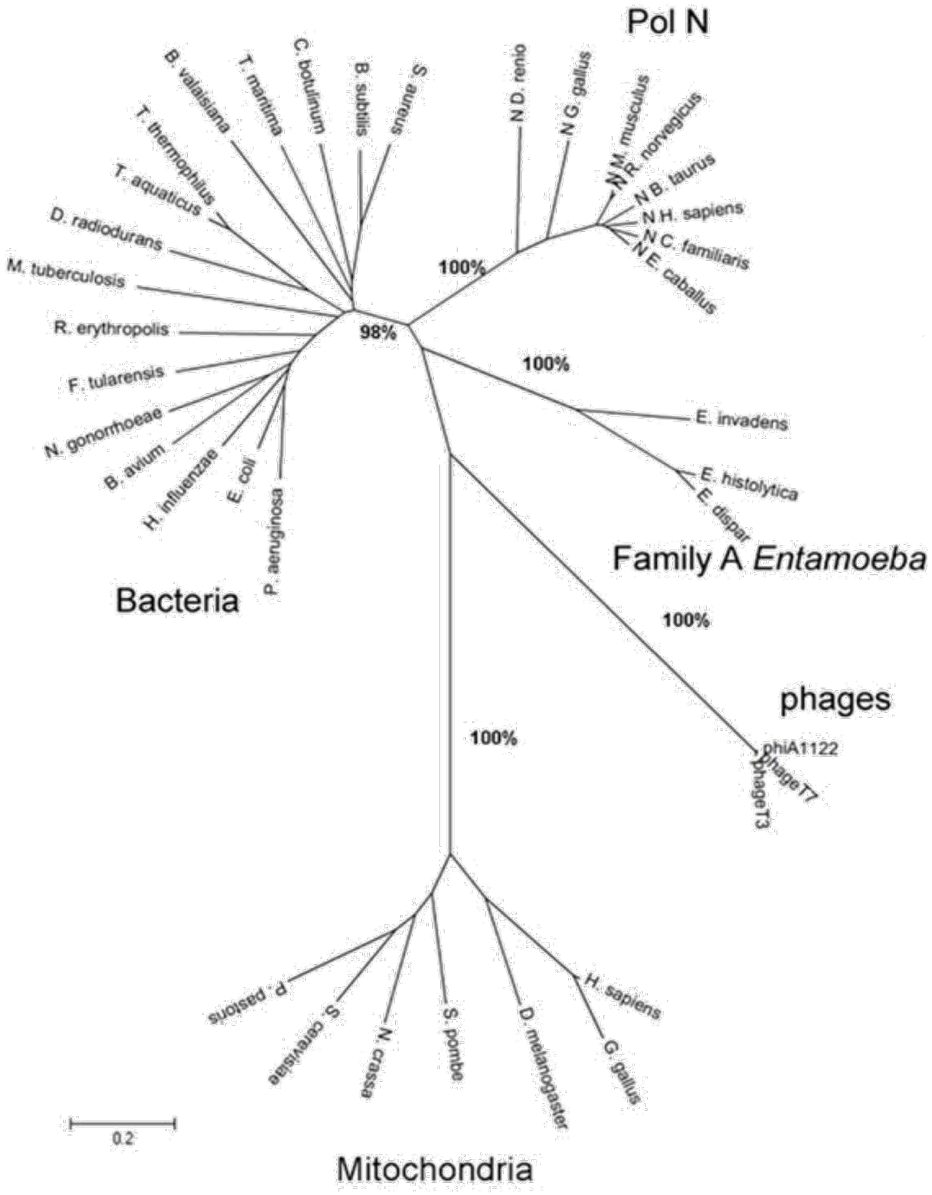


Figure 17: Sequence similarity of the Family A DNA polymerases. An unrooted phylogenetic tree shows the relationship between the amino acid sequences of various Family A DNA polymerases. The scale bar indicates average number of mutations at a single position for that branch length. Bacterial pols, Entamoeba pols, and pol Ns are primarily repair enzymes, while the phage and mitochondrial pols act as replicative enzymes. Adapted from (96).

The Family A DNA polymerases are a family that contains numerous historic and commercially important enzymes (Figure 17), with representatives found in prokaryotes, eukaryotes, and some viruses. These template-dependent polymerases are involved in both replication and repair:

- Replicative enzymes perform both leading- and lagging-strand synthesis, by recruiting helper proteins and molecules to modulate their processivity. Examples include the T3 and T7 bacteriophage DNA polymerases, as well as the mitochondrial DNA polymerase γ in eukaryotes (97). Alone, T7 DNA pol can incorporate only 1-15 bases before dissociating from the template, but by binding bacterial thioredoxin it can increase its processivity to incorporate thousands of bases in a single binding event (98).
- Repair enzymes are responsible for filling in the gaps generated during lagging-strand synthesis. An example is the Pol I enzyme (discussed further below), which is found in prokaryotes such as *Thermus aquaticus* and *Escherichia coli*. Homologous Pol I enzymes always contain a 5′→3′ domain that removes the RNA primer from Okazaki fragments; many isolates also contain a 3′→5′ proofreading domain to correct misincorporated nucleotides (97).

In 1958, Pol I was the first polymerase was discovered. Initially isolated from *E. coli*, this 928 aa enzyme accounts for >95% of polymerase activity in the organism (99). While Pol I is a weakly processive enzyme, it can only add ~20 nt before dissociating from the template and therefore is unable to act as the leading-strand

polymerase (this is carried out by Pol III) (100). In addition to the polymerase domain, the enzyme contains both 5′→3′ and 3′→5′ exonuclease activity, giving it an error rate of only 10^{-6} to 10^{-8} per nucleotide (101). Cleavage by subtilisin produces two fragments, the larger of which is derived from the C-terminal end and is called the Klenow fragment (or just Klenow) (Figure 18). This retains the polymerase and 3′→5′ exonuclease domains, but lacks the 5′→3′ domain. Klenow is a commercially important enzyme used for numerous applications, such as creating blunt-ended templates (either by removing 3′ overhangs or filling in 5′ overhangs), second strand DNA synthesis, or preparing radiolabeled probes. Before thermostable polymerases were discovered, it was also used as the enzyme for nucleotide incorporation in PCR.

Another commercially important Pol I is the 832 aa DNA polymerase from the bacteria *Themus aquaticus*, known as *Taq* pol (102) (Figure 18). As *Thermus aquaticus* lives in hydrothermal vents and hot springs, *Taq* pol is thermostable in temperatures that would denature many other proteins: its optimal temperature range is 75-80°C and it can withstand temperatures near boiling (at 97.5°C, the enzyme has a half-life of 9 min) (103). This has made it ideal for PCR-based applications, which repeatedly cycle through 90°C+ temperatures (104). Like other Pol I enzymes, it contains a 5′→3′ exonuclease domains, but lacks a functional 3′→5′ exonuclease proofreading domain. This lack of proofreading ability gives *Taq* pol an error rate of $1x10^{4}$, much higher than many other Pol I enzymes (105).

As RTs show preference for specific sequences, it is possible that other DNA polymerases do as well. Given the commercial and historic importance of Klenow and *Taq* pol, they are interesting targets for investigation.

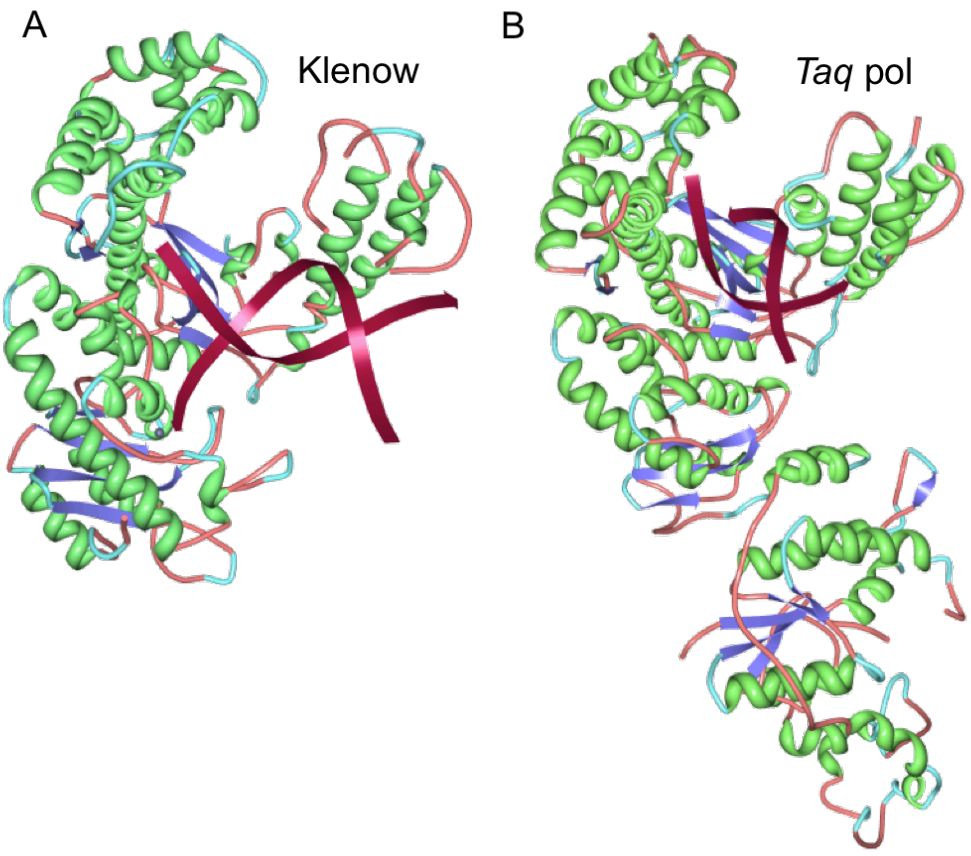


Figure 18. Crystal structures of Klenow and *Taq* pol bound to dsDNA. The two Family A DNA polymerases, *E. coli* Klenow and *Taq* pol, both perform lagging strand synthesis in their respective organisms. Klenow (A) is derived from the larger enzyme Pol I and lacks the 5′→3′ exonuclease activity of the full protein. *Taq* pol (B) contains 5′→3′ exonuclease activity but does not contain a functional 3′→5′ proofreading domain. The Klenow structure is from (106) and the Taq structure is from (107).

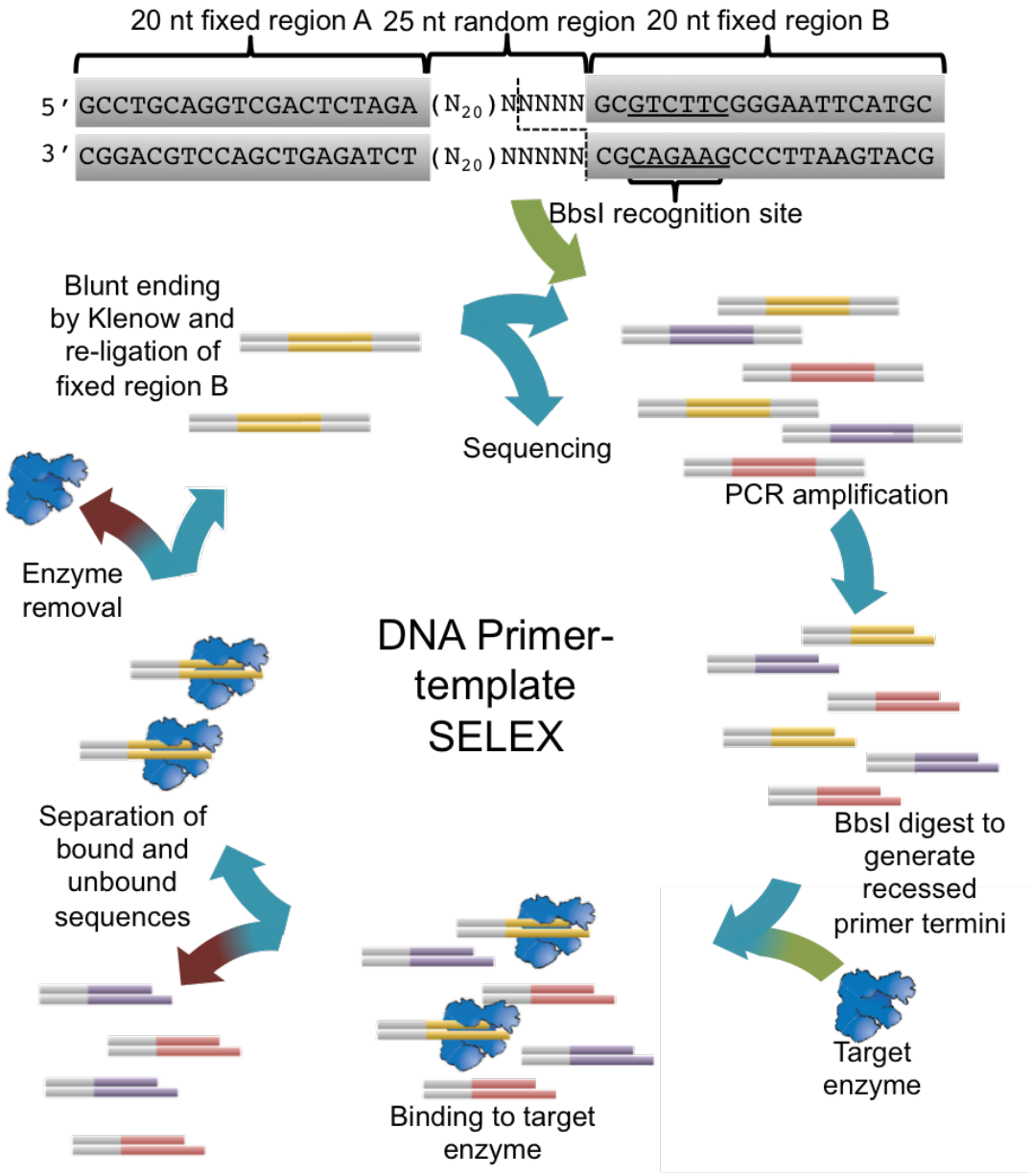


Figure 19. The PT SELEX method used to generate aptamers against Klenow and *Taq* pol. The primer is the top strand and the template is the bottom. The initial fixed-region-flanked random hybrid is indicated in gray with the staggered ends generated by the BbsI site shown by the dotted gray line. Fixed region B and four nucleotides of the random region from the primer are removed; fixed region A, 21 nt of the primer and 25 nt of the template random region are retained.

3.2 Materials and Methods

Materials

To avoid nuclease degradation during selection, a 3′→5′ exonuclease-minus mutant of Klenow (exo⁻Klenow, containing the mutations D355A and E357A) purchased from New England Biolabs was used for all binding experiments, as were wildtype Klenow, *Taq* DNA pol, BbsI, and T4 polynucleotide kinase (T4 PNK). dNTPs were purchased from Roche Applied Sciences, radiolabeled γ-P^{32} dATP from PerkinElmer, Sephadex G-25 spin columns from Harvard Apparatus, oligonucleotides from Integrated DNA Technologies, and all other reagents were obtained from Thermo Fischer Scientific, Inc., Sigma-Aldrich Co., or VWR.

Preparation of random DNA:DNA hybrids for the initial SELEX library.

5′ radiolabeled primers were prepared as described in 2.2 Materials and Methods. ~400 pmol each of radiolabeled primer (5′-GCCTGCAGGTCGACTCTAGA-3′) and template (5′-GCATGAATTCCCGAAGACGC$(N)_{25}$TCTAGAGTCGACCTGCAGGC-3′, where N is any base) were mixed in a 50 µl reaction containing 50 mM Tris–HCl (pH = 8), 1 mM DTT and 80 mM KCl. The reaction was heated to 65ºC for 5 minutes, then slow cooled to room temperature to form hybrids. The second strand was filled in by adding the hybrid to a 100 µl reaction containing 0.5 mM dNTPs, 6 mM $MgCl_2$, and 3 U wildtype Klenow and incubating for 45 minutes hours at 37°C. The product was purified by running on a 12% native polyacrylamide gel and cutting out bands corresponding to the double-stranded hybrid. The hybrid was eluted from the gel overnight in 500 µl of10 mM Tris-HCl (pH = 8) at 4°C and then passed through a

0.45 μm polyethersulfone membrane syringe filter. The material was precipitated overnight at -20°C in 2 volumes of 100% ethanol and 1/10th vol of 3 M sodium acetate (pH = 7) (final volume ~1500 μl), then spun for 30 min at ~15,000 *g* to pellet the DNA. This was washed with 500 μl of 70% ethanol then dried in a Savant DNA120 speed vacuum from Thermo Fisher.

Oligonucleotides (Tables 5-7) were purchased and hybridized similarly, except at a ratio of 1:1.5 primer:template (4 pmol ^{32}P-labeled primer:6 pmol template) in a 20 μl reaction.

Generating 3′ recessed primer:templates

Collected hybrids from the two reactions were incubated with 20 U of BbsI at 37°C for 3 hours, in a final volume of 100 μl of 50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, and100 μg/ml BSA. Cut products were separated at 220 V on a 12% native polyacrylamide gel, extracted and precipitated as described above. Each reaction typically yielded about 25-50 pm of final product.

Selection with exo$^-$Klenow and Taq *pol*

For the initial selection, ~200 pmol of cleaved hybrid was incubated with 20 pmol exo$^-$Klenow or *Taq* pol for 1 hour in 100 μl of 50 mM Tris-HCl (pH = 7), 50 mM KCl, 1 mM DTT, 10 mM $MgCl_2$, and 50 μg/ml BSA (Klenow buffer) or 10 mM Tris-HCl, 50 mM KCl, and 1.5 mM $MgCl_2$ (*Taq* pol buffer) at room temperature then run on a 12% native gel to separate bound and unbound hybrids. Bound hybrids were cut out of the gel, extracted and precipitated. This process is illustrated in Figure 19.

Generation of enriched pools from selected material

After selection, the hybrids' recessed ends were filled in by wildtype Klenow in reactions containing 0.1 U of enzyme, 50 mM Tris-HCl (pH=8) 100 µM dNTPs, 6 mM $MgCl_2$, 80 mM KCl and 1 mM DTT, total volume 50 µl. After 10 min at 37ºC, reactions were phenol extracted and precipitated with ethanol. Following this, the recovered material was ligated for 20 minutes to 5-10 fold excess of the hybrid duplex of 5′-ATAGCATGAATTCGCAGAAGACCC-3′ and 5′-GGGTCTTCTGCGAATTCATGC-3′, using the Promega rapid ligation kit. For the first round, the ligation was performed in a volume of 30 µl; subsequent rounds were performed in 10 µl following the manufacturer's protocol.

The entire ligation mixture was then PCR amplified by *Taq* pol in a 400 µl reaction containing 400 pmol each of the primers 5′-GCCTGCAGGTCGACTCTAGA-3′ (^{32}P end labeled) and 5′-GCATGAATTCGCAGAAGACCC-3′, in *Taq* buffer. The reactions were divided equally into 4 tubes and linear PCR amplified for 12-14 cycles of 94°C (1 min), 50°C (1 min) and 72°C (1min), followed by a final 5 min extension at 72°C. This material was then combined, BbsI digested and purified as above. This selection process was repeated 6 times for both *Taq* pol and exo$^-$Klenow, with subsequent rounds using ~1/20th the enzyme (mole:mole) of the recovered material, until no increase in binding was noticed between rounds. After round 2, 1/5 of the material was also saved from each round for use in K_d determinations. After reaching round 7, material from each selection pool was inserted into a vector and cloned into bacteria using the Strataclone Blunt Ended PCR kit from Agilent Technologies, following

manufacturer's instructions. Isolated colonies were grown overnight in 3 ml of 50 μg/ml ampicillin-LB broth. Insert-containing plasmids were purified using the Qiagen Spin Miniprep kit and sequenced by GeneWiz.

Determination of K_d by gel shift

Material from selection rounds and purchased oligonucleotides were 5' end-labeled on the primer and hybridized as described above. These were mixed with various amounts of either *Taq* pol or exo$^-$Klenow at a final concentration of 2 or 20 nM final substrate concentration, respectively, (binding affinities for the pools from each round and the random substrate were determined at a concentration of 20 nM; analysis of the constructures was performed at a final substrate concentration of 2 nM) in either Klenow or *Taq* pol buffer. The enzyme concentrations were 0, 1.1, 3.1, 6.3, 12.5, 25, and 50 nM *Taq* pol and 0, 25, 50, 100, 200, 400, and 800 nM exo$^-$ Klenow, unless otherwise indicated. The enzymes and hybrids were incubated for 5 min at room temperature and then run at 110 V on a 7% polyacrylamide native gel. The products were visualized on a FLA-7000 phosphoimager (Fujifilm) and the amount of product shifted vs. unshifted was quantified with Multi Gauge software. The K_d was determined by plotting the ratio of material shifted ([shifted]/[shifted+unshifted]) vs. the concentration of the enzyme 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.

Mutations of selected sequences

To test the importance of specific nucleotides for tight binding in the selected primer-template aptamers, mutational analysis was performed by substituting A→C, C→, G→T, and T→G. See Table 5 for construct details.

3.3 Results

3.3.1 SELEX results for exo⁻Klenow

After seven rounds of SELEX, the enriched pool bound the enzyme 5x better than a random pool (K_d of 90±10 nM vs. >500 nM), suggesting that it contained high-affinity sequences (Figure 20). The pool was clonal sequenced and twelve isolates were recovered, of which eleven showed significant similarity (Figure 21B). Interestingly, these eleven sequences contained a second BbsI site in an orientation opposite to the engineered site in the primer region (see Figure 21A). Due to incomplete cleavage by the restriction enzyme, this allowed a nick on each strand that resulted in a 9 nt overhang on the 5′ end of the template, instead of the intended 4 nt overhang (generating 36 nt primer/45 nt template pairs). Subsequent testing revealed that reducing the length of the overhang by more than 3 nt significantly disrupted binding (Table 5), suggesting a strong selective pressure for the location of this second BbsI site.

Intriguingly, seven of these sequences contained a region identical to T7 RNAP consensus promoter sequence (bases -6 to + 1, called the T7Pcore), while three others contained only a single nt difference in that region. In the phage RNAP promoters, this region is highly conserved between different RNAPs (a single base differentiates the T3 promoter from T7 and SP6); bases -4 to +3 are where a single-

stranded bubble forms for the initiation of transcription. The sequence selected by Klenow contained four of these bases in the duplex region and three bases in the single-stranded template overhang of the primer-template hybrid. A consensus primer-template sequence was selected for subsequent testing (named Exo$^-$Kl) to determine which bases contribute most to this selective binding. This construct bound with a K_d of ~150 nM, higher than the round 7 material, indicating that there are probably other, more tightly binding sequences present in the pool.

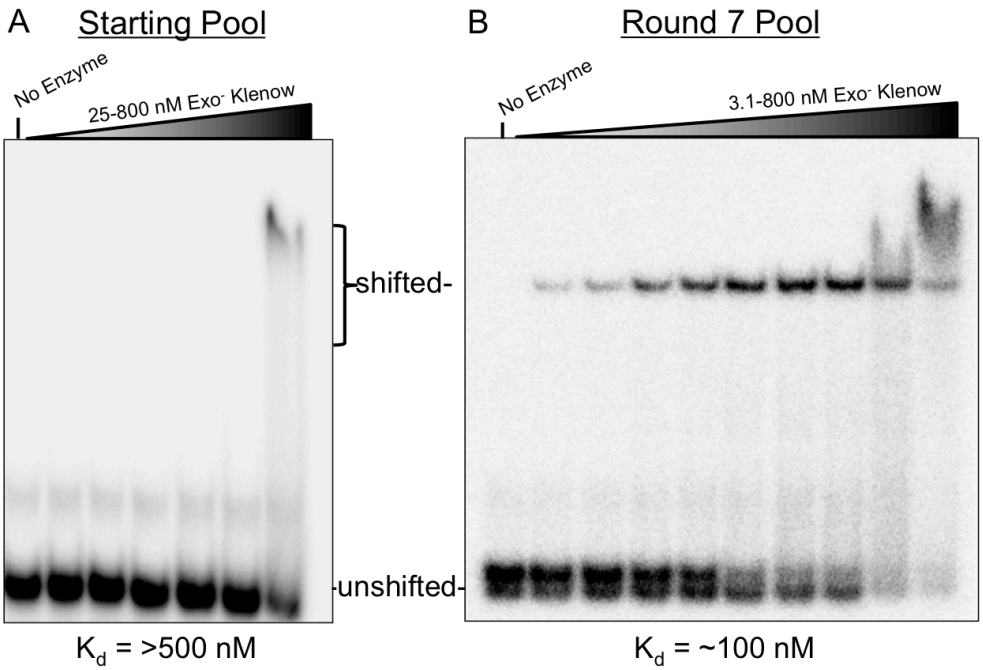


Figure 20. Gel shift comparison between the starting pool and round 7 pool from Exo$^-$Klenow SELEX. A) The starting pool of random hybrids was bound by exo$^-$ Klenow with a K_d greater than 500 nM. Klenow concentrations were 0, 25, 50, 100, 200, 400, and 800 nM. B) By round 7, the enzyme bound the pool with a K_d of 90±10 nM, indicated that the pool had been enriched for sequences that bound the enzyme. Concentrations used for this assay were 3.1, 6.25, 12.5, 25, 50, 100, 200, 400, and 800 nM exo$^-$Klenow.

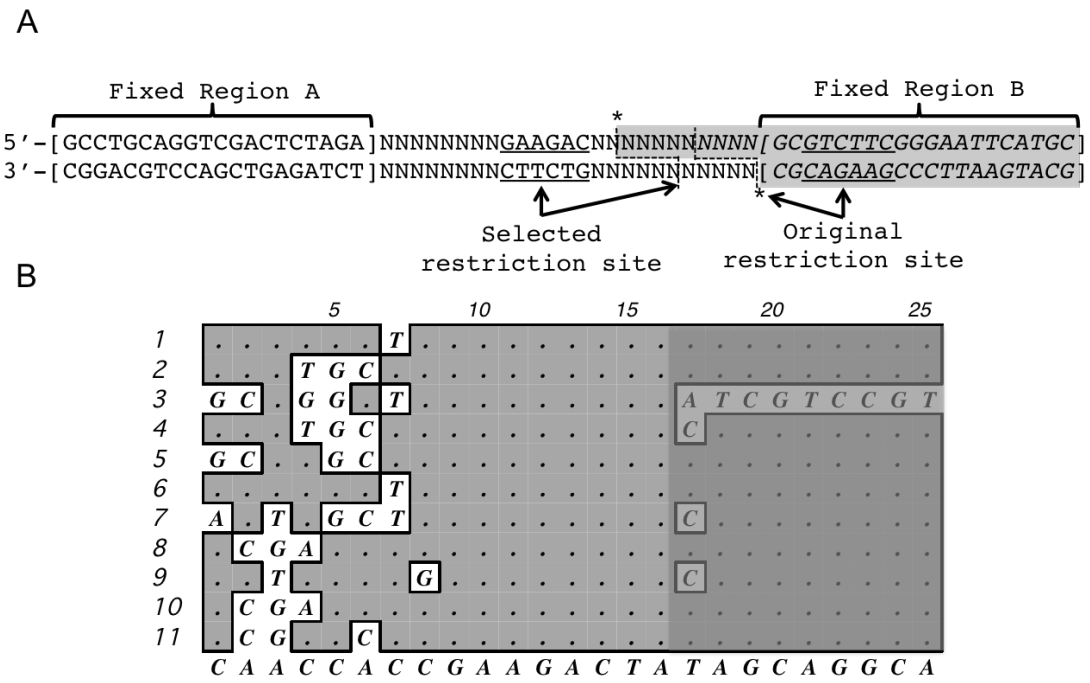


Figure 21. Primer-template sequences selected by Exo⁻Klenow SELEX. A) The original hybrids and intended BbsI cut site (which generates 41:45 nt primer:templates) and the second selected cut site (which yields 36:45 nt primer templates). B) The eleven sequences isolated from round seven containing a portion of the T7 promoter region, aligned using MacVector. Sequences are from the primer strand and the nine bases corresponding to the single-stranded template overhang are shaded darker gray.

3.3.2 Determination of sequences important for exo$^-$Klenow binding

Mutational analysis was performed on the representative sequence (see Table 5 for constructs and binding affinities) to determine which bases were most responsible for the specific binding of exo$^-$Klenow. In addition to constructs with template truncations (described above) which revealed that the enzyme's preference for nine nt overhangs, two other major types of mutations were made: base substitutions and shifted sequences within the context of the duplex:single-stranded primer-template.

Modifying the seven bases that made up the T7Pcore disrupted tight binding to exo$^-$Klenow; as this sequence was found in most of the recovered sequences, the result is not surprising and implies that that these bases contribute significantly to the specific interaction with the enzyme. Of the T7Pcore sequence, it appears that only the four bases in the duplex region are important: modifying the nine nt single-stranded template overhang sequence (which contains three nt of the T7Pcore) did not significantly alter enzyme binding, despite being highly conserved among recovered sequences. Modifying the bases 5' of the T7Pcore also reduced binding, although the defect was not as severe as when the T7Pcore was modified. Together, these results suggest that the four bases of the T7Pcore in the duplex region are the major contributors to tight binding but are aided by sequences upstream. This agrees with the observation that the region of conservation among these sequences is larger than just the T7Pcore.

To determine if the context of the sequence relative to the primer terminus was important, constructs which were created where the entire random region-derived

sequences were shifted further in (-1, -4, -8) or out (+1) of the duplex region. Only the -1 shift (which contained five bp of the T7Pcore in the duplex region and two nt in the single-stranded region) was tolerated; other shifts were deleterious to binding, demonstrating that sequence context was also integral to tight binding.

Exo⁻Klenow Selex	Sequence	K_d (nM)
Exo⁻Kl	gcctgcaggtcgactctagaCAACCATCGAAGACTA cggacgtccagctgagatctGTTGGTAGCTTCTGATATCGTCCGT	150*
Random substrate	gcctgcaggtcgactctagaNNNNNNNNNNNNNNNN cggacgtccagctgagatctNNNNNNNNNNNNNNNNNNNNNNNN	>500
36nt primer:42nt Template	gcctgcaggtcgactctagaCAACCATCGAAGACTA cggacgtccagctgagatctGTTGGTAGCTTCTGATATCGTG	311 ± 70
36nt primer:44nt Template	gcctgcaggtcgactctagaCAACCATCGAAGACTA cggacgtccagctgagatctGTTGGTAGCTTCTGATATCGTGCC	178 ± 25
+1 shift	ggcctgcaggtcgactctagaCAACCATCGAAGACT ccggacgtccagctgagatctGTTGGTAGCTTCTGATATCGTCCG	472 ± 122
-1 shift	cctgcaggtcgactctagaCAACCATCGAAGACTAT ggacgtccagctgagatctGTTGGTAGCTTCTGATATCGTCCGTC	251 ± 131
-4 shift	gcaggtcgactctagaCAACCATCGAAGACTATAGC cgtccagctgagatctGTTGGTAGCTTCTGATATCGTCCGTCCGT	325 ± 86
-8 shift	gtcgactctagaCAACCATCGAAGACTATAGCAGGC cagctgagatctGTTGGTAGCTTCTGATATCGTCCGTCCGTCCGT	>800
Modified T7Pcore	gcctgcaggtcgactctagaCAACCATCGAAGCAGC cggacgtccagctgagatctGTTGGTAGCTTCGTCGCGAGTCCGT	571 ± 78
Modified Template Overhang	gcctgcaggtcgactctagaCAACCATCGAAGACTA cggacgtccagctgagatctGTTGGTAGCTTCTGATCGATGTAAG	201 ± 1
Exo⁻Kl 5′ modified non-T7 sequence	gcctgcaggtcgactctagaACCAACGATCCTACTA cggacgtccagctgagatctTGGTTGCTAGGATGATATCGTCCGT	266 ± 8

Table 5. K_d of exo⁻Klenow for the Exo⁻Kl sequence and modified constructs. Modified bases are underlined. An (*) indicates the average of two independent experiments. A K_d determined from a single experiment is prefixed with (~). Three types of mutations were tested for their effect on exo⁻Klenow binding: 1) the length of the template, 2) the location of the conserved T7Pcore relative to the primer terminus, and 3) mutations of the conserved nucleotide sequences. The only mutations that were relatively tolerated were a single nucleotide truncation from the template and shifting the sequence -1 from the primer terminus. All other mutations produced significant defects in binding to exo⁻Klenow.

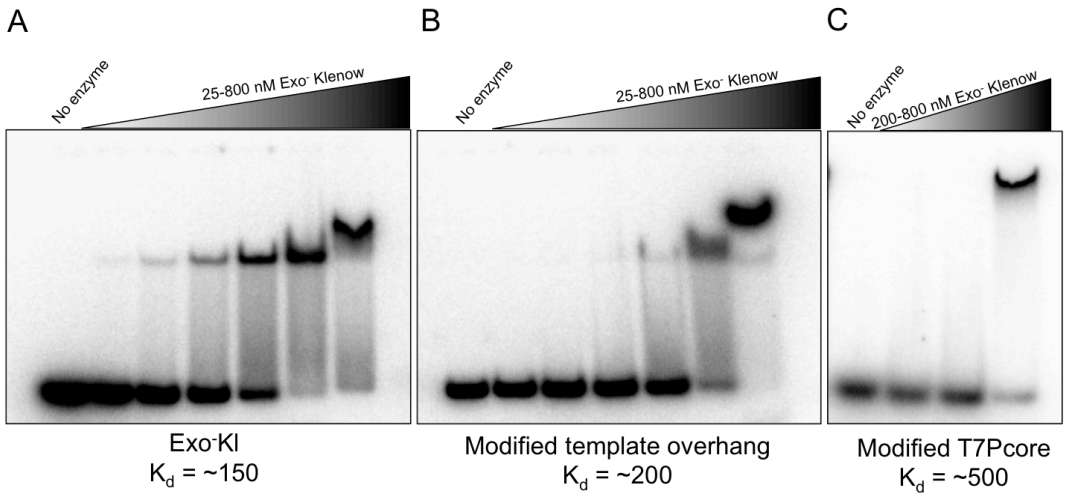


Figure 22. Gel shift with the Exo$^-$Kl construct and a modified version, which has the T7Pcore replaced. A) The Exo$^-$Kl construct is bound by exo$^-$Klenow with a K_d of ~150 nM. B) Modifying the template overhang diminishes binding but does not full disrupt it. C) Modifying the T7Pcore significantly prevents binding of the enzyme to the primer-template. In A and B, the concentration of exo$^-$Klenow is 0, 25, 50, 100, 200, 400, and 800 nM. In C, the concentration is 0, 200, 400, and 800. Sequences for each construct are in Table 5.

3.3.3 SELEX results for Taq pol

To determine if other Family A DNA polymerases preferentially bound primer:template sequences, SELEX was carried out for seven rounds using *Taq* pol. From this pool twelve sequences were isolated, out of which two sequence motifs were identified (see Figure 23). A consensus sequence from each motif group (named Taq1 and Taq2) was synthesized for further analysis.

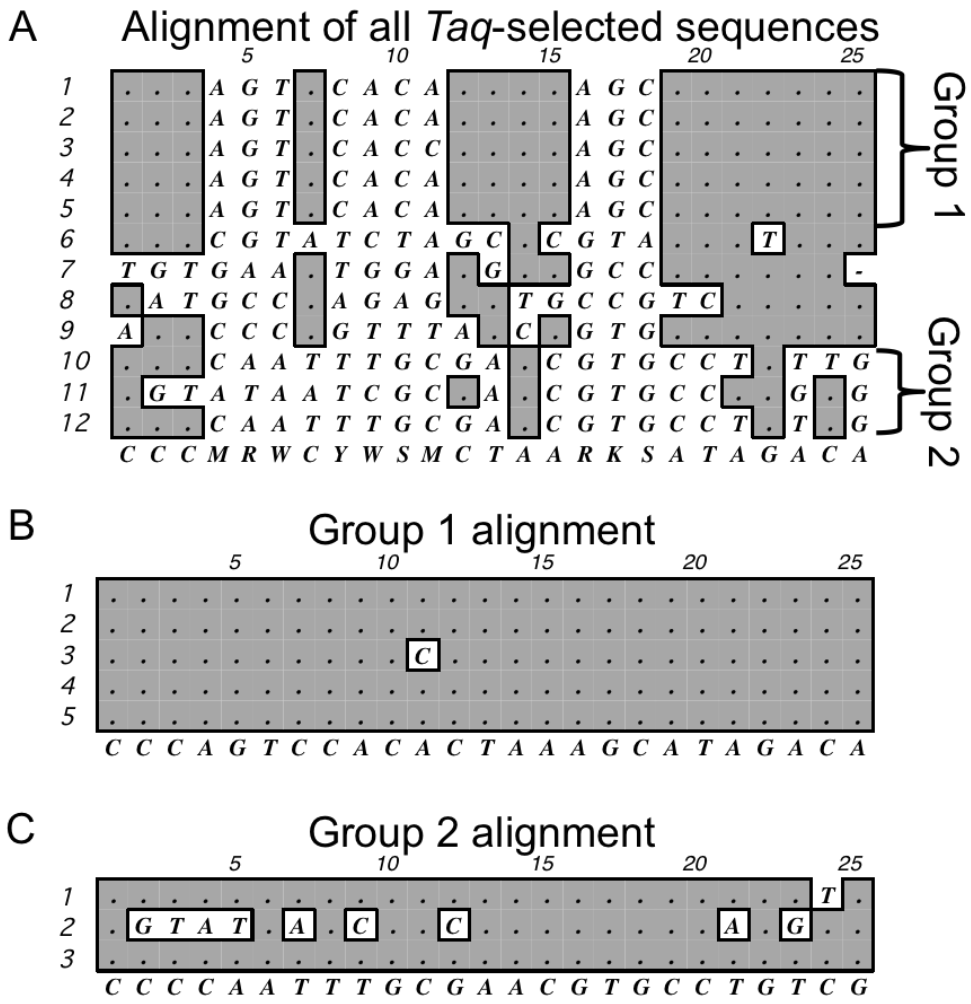


Figure 23. Sequences recovered from round seven of *Taq* pol PT SELEX. Nucleotides that match the consensus, located beneath each alignment, are represented by a dot. Of the twelve sequences isolated (A), two general motifs were observed. Group 1(B) contained a region that matched the initiation domain of the T3 RNAP promoter. Group 2 (C) did not contain a recognizable biologically relevant sequence. Sequences were aligned using MacVector.

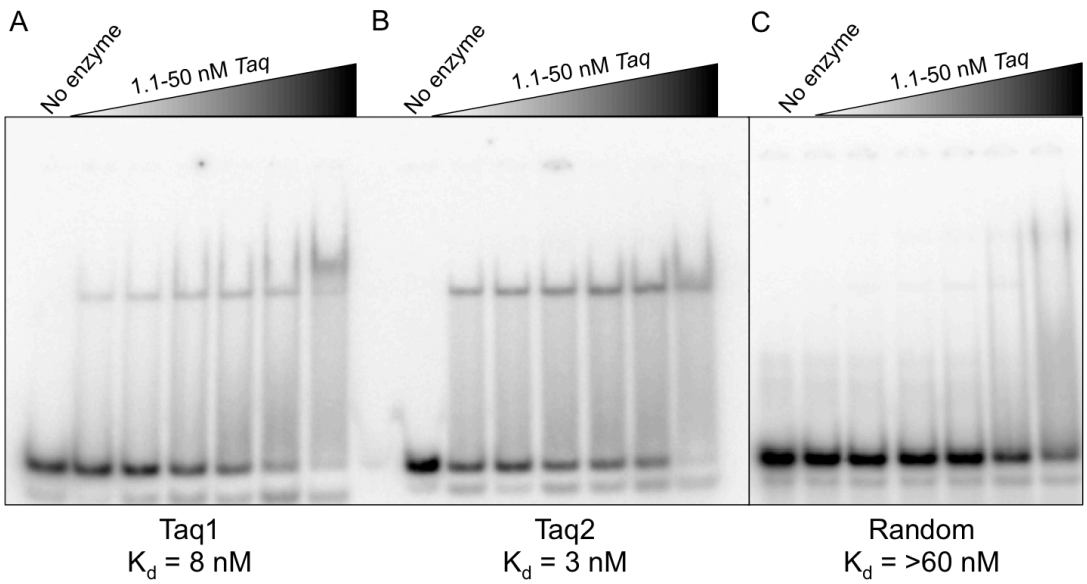


Figure 24. Gel shift of Taq1 and Taq2 constructs, as well as a random pool, by *Taq* pol. The constructs were derived from the consensus sequences for the two observed motifs isolated from round seven. Taq1 is bound by *Taq* pol ~8x better, while Taq2 is bound 20x more tightly. In each gel, the concentration of *Taq* pol is 0, 1.1, 3.1, 6.25, 12.5, 25, and 50 nM.

The five Taq1-like sequences were identical except for a single nt substitution in one isolate. All contained a region that, like the sequences selected by exo$^-$Klenow, contained bases -6 to +1 of the T3 RNAP promoter region (which is identical to the T7 promoter sequence at those positions except for a single nt at -2, see Table 8), called the T3Pcore. Unlike the exo$^-$Klenow-derived sequences, this region was contained fully within the duplex DNA (4 bp upstream from the primer terminus). Taq1 primer-template hybrids bound the enzyme with a K_d of 8 nM, 8x more tightly than the random pool.

The three non-T3-like sequences contained an eight bp conserved motif completely unlike that of the Taq1 sequences and with no known similarity to any biologically relevant motifs. A primer-template synthesized with the Taq2 consensus sequence bound the enzyme with a K_d of 3 nM, almost 3x more tightly than Taq1 and >20x that of the random pool.

3.3.4 Determination of sequences important for Taq pol binding

Much like with the sequences derived from exo$^-$Klenow SELEX, mutational analysis demonstrated that the location and context of each sequence was necessary for tight binding (see Table 6 for constructs). To see if the T3Pcore was responsible for the affinity of *Taq* pol to the Taq1 motif, a construct was created with substitutions at every nucleotide except the T3Pcore. This construct suffered from defects in binding (~13 nM, compared to 8 nM for Taq1) but still bound ~4x better than the random pool, indicating that the T3Pcore was an important factor in the enzyme's affinity for the substrate but that other sequences surrounding this region were also involved (interestingly, *Taq* pol bound a construct that contained the entire

T3 RNAP promoter as poorly as it did a random pool, suggesting that a construct with the entire promoter sequence is in a conformation that disrupts *Taq* pol binding, despite the presence of the T3Pcore). By mutating four bp pieces of the T3Pcore, it was learned that the bases matching from -3 to +1 in the T3 RNAP promoter were the largest contributors to tight binding (4 nt modification B in Table 8); mutation of these bases disrupted binding while mutating the -7 to -4 bases (4 nt modification A) did not have an effect. Intriguingly, sometimes small changes to the T3Pcore were more deleterious to binding than large ones: a single bp change of the -2 base (to mimic the T7 promoter sequence, Table 8) completely disrupted binding. Like exo$^-$ Klenow, the enzyme did not tolerate shifting the context of the sequence relative to the primer-terminus as a single -1 nt shift abolished tight binding.

Taq Selex sequence 1	Sequence	K_d (nM)
Taq1	gcctgcaggtcgactctagaCCCAGTCCACACTAAAGCATA cggacgtccagctgagatctGGGTCAGGTGTGATTTCGTATCTGT	9 ± 3
Random	gcctgcaggtcgactctagaNNNNNNNNNNNNNNNNNNNN cggacgtccagctgagatctNNNNNNNNNNNNNNNNNNNNNNNN	>50
Modified T3Pcore	gcctgcaggtcgactctagaCCCAGTCCACCAGCCCTCATA cggacgtccagctgagatctGGGTCAGGTGGTCGGGAGTATCTGT	>50
Modified sequences surrounding the T3Pcore	gcctgcaggtcgactctagaAAACTGAACAACTAAAGCATA cggacgtccagctgagatctTTTGACTTGTTGATTTCGTATCTGT	15*
Full T3 RNAP promoter sequence	gcctgcaggtcgactctagaAATTAACCCTCACTAAAGGGAG cggacgtccagctgagatctTTAATTGGGAGTGATTTCCCTCTGTC	>50
Modification of sequences 5' of the T3Pcore	gcctgcaggtcgactctagaAAACTGAACAACTAAAGCATA cggacgtccagctgagatctTTTGACTTGTTGATTTCGTATCTGT	17 ± 0.1
4 bp modification A	gcctgcaggtcgactctagaCCCAGTCCAACAGAAAGCATA cggacgtccagctgagatctGGGTCAGGTTGTCTTTCGTATCTGT	~5
4 bp modification B	gcctgcaggtcgactctagaCCCAGTCCACACTCCCTCATA cggacgtccagctgagatctGGGTCAGGTGTGAGGGAGTATCTGT	~25
-1 shift	cctgcaggtcgactctagaCCCAGTCCACACTAAAGCATAG ggacgtccagctgagatctGGGTCAGGTGTGATTTCGTATCTGTC	>50
T3→T7	gcctgcaggtcgactctagaCCCAGTCCACACTATAGCATA cggacgtccagctgagatctGGGTCAGGTGTGATATCGTATCTGT	16 ± 3.4

Table 6. Taq1 constructs generated for mutational analysis of the nucleotides important for *Taq* pol binding to primer-templates. Modified bases are underlined. An (*) indicates the average of two independent experiments. A K_d determined from a single experiment is prefixed with (~). Most modifications negatively impact binding; the best tolerated mutations were those which were outside of the bases underlined in '4nt modification B', suggesting that these bases are the most important determinants of tight binding to *Taq* pol.

Taq Selex sequence 2	Sequence	K_d
Taq2	gcctgcaggtcgactctagaCCCCAATTTGCGAACGTGCCT cggacgtccagctgagatctGGGGTTAAACGCTTGCACGGACAGC	3*
Random	gcctgcaggtcgactctagaNNNNNNNNNNNNNNNNNNNN cggacgtccagctgagatctNNNNNNNNNNNNNNNNNNNNNNNN	>50
Modified 8 bp core	gcctgcaggtcgactctagaCCCCAATTTGCGCCATGTAAT cggacgtccagctgagatctGGGGTTAAACGCGGTACATTACAGC	~25
Modified non-core nucleotides	gcctgcaggtcgactctagaAAAACCGGGTATAACGTGCCG cggacgtccagctgagatctTTTTGGCCCATATTGCACGGCACTA	>50
Modified non-core/ non-overhang nucleotides	gcctgcaggtcgactctagaAAAACCGGGTATAACGTGCCT cggacgtccagctgagatctTTTTGGCCCATATTGCACGGACAGC	~21
-1 shift	cctgcaggtcgactctagaCCCCAATTTGCGAACGTGCCTG ggacgtccagctgagatctGGGGTTAAACGCTTGCACGGACAGCC	>50
4 bp modification A	gcctgcaggtcgactctagaCCCCAATTTGCGCCATTGCCT cggacgtccagctgagatctGGGTCAGGTTGTCTTTCGTATCAGC	>50
4 bp modification B	gcctgcaggtcgactctagaCCCCAATTTGCGAACGGTAAT cggacgtccagctgagatctGGGTCAGGTGTGAGGGAGTATCAGC	>50

Table 7. Taq2 constructs generated for mutational analysis of the nucleotides important for *Taq* pol binding to primer-templates. Modified bases are underlined. An (*) indicates the average of two independent experiments. A K_d determined from a single experiment is prefixed with (~). All of the constructs generated for Taq2 negatively impacted binding of *Taq* pol to the substrate.

Unlike Taq1, all modifications of Taq2 disrupted binding. Modification of the 8 bp region found in all three sequences (called the 8 bp core) reduced binding ~8x, as did modification of the bases upstream of this region. In contrast to Taq1, the bases on the overhang had a substantial impact on binding ('Modified non-core nucleotides' vs. 'modified non-core/non-overhand nucleotides'); all constructs with modified overhangs showed binding similar to the random pool, suggesting that *Taq* pol interacts with this substrate differently than Taq1.

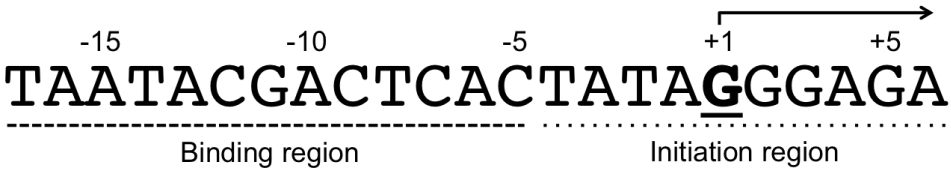
3.3.5 Enzyme specificity to selected sequences

Though the sequences selected by exo$^-$Klenow and some of those selected by *Taq* pol contained regions identical to the T7-like RNAP promoters, exo$^-$Klenow was unable to bind the Taq1 sequence with a higher affinity than for the random substrate (>800 nM). Likewise, *Taq* pol only bound the Exo$^-$Kl substrate with a K_d of ~50 nM. Whether this was due to sequence differences or the different location of the T7/T3 promoter core within the primer-template (spanning the primer terminus in Exo$^-$Kl vs. four bases upstream from the terminus in Taq1) was not determined.

3.4 Discussion

The single subunit RNA pols (RNAPs) are a group of enzymes found in bacteriophages (such as T3, T7, and SP6), chloroplasts, and mitochondria that are structurally homologous to Family A DNA pols. Like the Family A pols, they contain finger, thumb, and palm domains that use a two-metal ion cofactor for nucleotide incorporation off of a DNA template. However, unlike the template-dependent Family A polymerases, these RNAPs also require a specific promoter

sequence to bind to the DNA template and initiate RNA transcription. In the T7-like phage RNAPs, this sequence is sufficient for the enzyme to perform binding, melting, and initiation; mitochondrial and chloroplast RNAPs also require cellular factors for transcription. This promoter is highly specific for each enzyme and there is significant sequence variation within the group (Table 8). The phage RNAP promoter sequences from T3, T7, and SP6 are closely related with nearly identical binding and initiation domains. A single nucleotide triplet (-12 to -10 from the transcription start site) is what distinguishes between the sequences and is sufficient to prevent non-specific RNAP binding; this is detected by a specificity loop on the enzyme (absent in Family A DNA pols).



Organism	Promoter Sequence	Transcription factor required?
T7 phage	TAATACGACTCACTATAGGGAGA	No
T3 phage	AATTAACCCTCACTAAAGGGAGA	No
SP6 phage	ATTTAGGTGACACTATAGAAGRG	No
S. cervesiae mitochondria	ATATAAGTAR	Yes
Tobacco plastid	CRTNATAAT	Yes
Human mitochondria	TGGGGGGTGACTGTTAAAAGTGCATACCGCCAAAAGATAAAA	Yes

Table 8. Consensus sequences for the promoters of the single-subunit RNA polymerases. (Above) An annotated consensus sequence for the T7 RNAP. Bases -17 through -5 (dashed line) are important for T7 RNAP binding; bases -4 through +6 (dotted line) are involved in initiation of transcription and are melted open by the enzyme. The first base transcribed (+1) in bold and underlined. These regions are the same in other phage promoters (such as T3 and SP6) (108-111). (Below) Consensus sequences for RNAP promoters among phage and eukaryotes. The first base transcribed is underlined in each sequence. The phage RNAPs do not require a transcription factor for initiation of transcription, while the RNAPs from eukaryotes do. Table modified from (112).

It is likely that the regions of structural similarity between the T7-like RNAPs and the Family A DNA pols are responsible for the shared attraction to these promoter sequences. While the enzyme amino acid identities vary significantly (for example, Klenow and T7 RNAP share only 8 aa in the palm domain, corresponding to two conserved Asp-containing motifs A and C, which are found in most polymerases), the enzymes are all right-handed polymerases with homologous palm and finger architecture (though phage RNAPs contain additional structures not found in Klenow or *Taq* pol) (113).

Studies of how Klenow interacts with template stands suggest that the enzyme only has contact with the phosphate backbone of the nucleic acid (106), which has been thought to preclude sequence-specific attraction. Our results contradict the idea that Klenow lacks sequence specificity, as we have demonstrated that the sequence of a primer-template promotes significant differences in binding affinity. Still, it is possible that the structure of the selected primer-template is what is responsible for the tight binding and not base-specific interactions with the enzyme.

Previous reports have demonstrated that the T7 RNA promoter has an intrinsic bend of 10 degrees. When bound to the enzyme, this bend increases to 40-60 degrees and is centered around the -2 to +1 region of the promoter. It is possible that the intrinsic bend on the promoter aids in binding both to the T7 RNA promoter as well as to *Taq* pol and Klenow, which both have been observed to induce significant bends in their substrates upon binding (120 degrees (114) and 80 degrees, respectively). This seems most likely for the *Taq* pol-selected aptamers, as the promoter-like sequence is contained completely in the duplex region of the primer-template; it is

unlikely to explain the preferential binding of Klenow to these sequences, as they are partially single-stranded at this region.

This location disparity is possibly due to the differences between Klenow and Taq pol: despite their structural similarity, *Taq* pol and the full *E. coli* pol I enzyme contain only 49% protein sequence identity (115). A significant difference between the two enzymes is the presence of 3′->5′ exonuclease activity in Klenow. Structural analysis of Klenow bound to primer-template DNA reveals that the primer-template can adopt a 'frayed open' conformation that allows the primer strand to interact with the exonuclease domain while the template is in the pol domain (116). In this, the three terminal 3′ bases of the primer are single stranded and 'scrunched' into the exonuclease domain. A shuttling mechanism is proposed to move the primer terminus from the exonuclease region to the polymerase domain for catalysis. This allows the enzyme to perform both nucleotide addition and excision repair without dissociation from the template. It is tempting to suspect that the reason the Klenow-selected T7-like region is located partially on the overhang is to promote unwinding that allows the 3′ primer terminus bases to interact with the exonuclease domain in the frayed open conformation while the template binds to the polymerase domain. Reports have demonstrated that Klenow interacts with at least the first four bases past the polymerase site (117); this agrees with our results suggesting an affinity for single stranded templates that are at least six bases long. In contrast, *Taq* pol shows no specific affinity for single-stranded regions, explaining why the conserved T7-like domain is contained entirely within the duplex region.

To test this hypothesis, it would be possible to use circular dichroism (CD) to determine the conformation of the primer-template SELEX sequences when associated with their enzymes. Structural studies have found that primer-template DNA is in B-form when bound in the Klenow exonuclease domain; for both Klenow and *Taq* pol, duplex DNA is in A-form when bound in the polymerase domain. CD can distinguish between these forms and would allow an estimation of the nature of the selected primer-templates when bound to both Klenow and *Taq* pol.

3.5 Conclusions

Despite the divergent functions of the single subunit RNAPs and the Family A DNA polymerases, two Family A DNA polymerases have strong affinities for sequences containing portions of the RNAP promoters. It is possible that this affinity is from a common ancestor that has been retained in the divergent enzyme groups and was later exploited by RNAPs to use as their promoter sequence. It is also possible that the nature of the DNA pol domains contain this affinity as a byproduct of their catalytic nature. Unlike the PPT sequence selected by RT, which serves *in vivo* to promote binding of RT to the primer-template in the correct orientation for DNA synthesis (118-121), these sequences do not have a known biological significance for the DNA pols and determining if one exists will require more research. One possible method would be to measure over- or under-abundance of the SELEX motifs relative to a random sequence of the same length in the *E. coli* or *T. aquaticus* genomes, which would suggest a selective pressure for the number of times that sequence is present. The preferential location of the sequences on either the lagging or leading strand, as well as their context near genes, might also shed light on any possible

biological relevance. For example, if the T7Pcore sequence is only found on the leading strand in *E. coli* and is significantly underrepresented on the lagging strand (where Klenow/*E. coli* pol I performs DNA synthesis), then this would suggest that the T7Pcore sequence is somehow inhibitory to the enzyme's function and the genome is under selective pressure to not contain that sequence. Conversely, overabundance of the sequence would suggest that perhaps the enzyme is recruited to these locations for an unknown purpose.

It may be possible to exploit these preferences to develop new technologies based on the enzymes' binding affinities to specific primer-template sequences. *Taq* pol binds sequences with high affinity at temperatures between 40-50°C; this binding decreases when the enzyme reaches its optimal temperature for extension (115). Potentially, a primer-template containing one of the tight-binding sequences could be used as an inexpensive ligand to temporarily inactivate *Taq* pol before PCR, similar to the current antibody-based Hot Start PCR technology that is used to prevent extension of non-specifically annealed primers at low temperatures, such as during reaction preparation. *Taq* pol would stay bound to the primer-template through the first round of extension, whereupon it would extend the primer to create a fully duplex hybrid. As the specific affinity of *Taq* pol for these primer-template sequences is dependent on both single and double-stranded regions, this extension would essentially inactivate the primer-template and free the enzyme for subsequent amplification of the intended target.

Troublingly, if these sequences promote *Taq* pol extension (or conversely, if they inhibit it), this presents a potential problem with current quantitative-PCR and

next-generation sequencing data. In a situation where the occurrence of this T7 RNAP promoter-like sequence is limiting relative to the number of *Taq* pol molecules, the affinity of the enzyme for this sequence will likely have little effect on the efficiency of extension. However, if the relative abundances are reversed, then a preferential binding of *Taq* pol to promoter-like elements could result in artificially high quantities of these sequences being amplified. It is possible to test this by engineering a plasmid vector that contains both a tight-binding and poor-binding sequence then PCR amplifying using equimolar concentrations of both primers. Assuming no extension preference, both sequences will be amplified equally; however, if there is a preference then it would be detectable by measuring the relative abundance of the products.

Despite issues raised by this discovery, the knowledge that Klenow and *Taq* pol bind to T7 RNAP promoter-like sequences remains intriguing. This adds to the body of evidence that the Family A DNA polymerases share a common ancestor with single subunit RNAPs, despite their extreme evolutionary divergence among viruses, prokaryotes, and eukaryotes. The fact that this affinity has been retained through millions of years of evolution suggests a selective pressure for its retention, whether through biological relevance (as in the case of the RNAP promoter) or some byproduct of the nucleotide incorporation machinery remains to be discovered.

Chapter 4: Modulation of Reverse Transcription *in vivo*

4.1 The effect of cellular small molecules on reverse transcription

4.1.1 Introduction

Research into the mechanics of retroviral reverse transcription has been aided by the existence of cell-free *in vitro* assays, which facilitate replicable studies of individual steps in the reverse transcription process and are an attractive alternative to cell-based assays as they permit analysis of conditions that might otherwise be impossible to obtain *in vivo*. Reverse transcription assays using purified RT (described in 2.2 Materials and Methods) have been optimized for the speed of nucleotide incorporation by the enzyme; while these conditions promote efficient DNA synthesis by RT, they fail to completely mimic the natural environment of the cell where reverse transcription would occur during the viral lifecycle. For instance, many RT assays use Mg^{2+} and dNTP concentrations (4-6 mM and 50-100 μM, respectively) that are greater than the concentrations found in cells (0.25 mM-2 mM Mg^{2+} and 5 μM dNTPs in T cells (122)). These differences can significantly impact the behavior of HIV RT: for example, extensions performed with more physiological concentrations of Mg^{2+} have shown reduced RNase H cleavage and polymerase pausing, which led to increased enzyme processivity (60). In addition to Mg^{2+} and dNTPs, the cytosol also contains other small molecules (rNTPs, amino acids, cations, etc.), some of which have been implicated in affecting the behavior of HIV RT (e.g. polyamines such as spermidine can increase the fidelity of HIV RT (123)). While not

all of these components affect reverse transcription, it is clear that at least some do and their exclusion from cell-free RT assays could distort the results of these studies. It would be useful to evaluate these compounds for their role in modulating the behavior of HIV RT, which would permit researchers to create more life-like cell-free assays with components that are important for reverse transcription. Unfortunately, most protocols used to isolate various cell components require buffers or extraction steps that make the resulting supernatant unusable for direct testing in reverse transcription assays. To address this, I developed a minimally-invasive protocol for extracting the small molecular fraction (molecules and proteins <3 kDa, which includes the dNTPs and Mg^{2+} necessary for extension by RT) from cells for use in driving reverse transcription in cell-free assays.

4.1.2 Materials and Methods

Jurkat cells, an immortalized T cell leukemia line, were grown in RPMI supplemented with 10% heat-inactivated fetal bovine serum and 100 units of Penicillin/0.1 mg/ml of Streptomycin. 1 x 10^8 cells were harvested at a density of 5 x 10^5 cells/ml and pelleted for 5 min at 4°C and 250 *g* to remove propagation medium. To remove extracellular components, the cell pellet was washed by resuspension and then centrifugation (as before) twice in buffers with decreasing ionic strength: first, in 1 ml of ice cold 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4) then in 0.1x PBS. Following the removal of the supernatant, the cells were resuspended in 1 ml of Milli-Q water and allowed to swell on ice for 10 minutes, after which they were lysed by dounce homogenization (modified from (124)).

The nuclear fraction removed by pelleting at 3300 *g* for 15 minutes in 4°C and the remaining cytosolic supernatant was passed through a 3 kDa molecular weight cutoff filter (Amicon Ultra-0.5 mL centrifugal filter, pre-washed with 500 µl 0.1 N NaOH, then 500 µl of Milli-Q water to remove glycerine). The collected filtrate, referred to as the small molecular fraction (SMF), was then volume-correct by vacuum drying on low heat to one half of the estimated volume of the cytosol of 1 x 10^8 Jurkat cells (0.8 picoliters/cell with the cytoplasm occupying about 80% of the total cell volume, calculated from (125)), creating a 2x concentration SMF. During volume correction, SMFs were never allowed to completely evaporate.

Reverse transcription assays were performed using a modified version of the protocol described in 2.2 Materials and Methods: extension was performed with 50 nM RT on a 100 nt ssDNA from the *gag-pol* region of the HIV-1(NL43) genome, which was hybridized 1:1 to a ^{32}P-radiolabeled 20 nt primer (4 nM final). Final reactions were carried out in 15 µl and contained 10 mM Tris-HCl (pH=8), 80 mM KCl, and varying concentrations of 2x SMF, dNTPs, and $MgCl_2$ as indicated.

4.1.3 Results and Discussion

In reactions containing only the SMF and no exogenous Mg^{2+} or dNTPs, RT was able to carryout efficient extension of the template, demonstrating that the SMF contained all of the components necessary for catalysis. While the speed of extension was qualitatively similar to control reactions that contained physiological concentrations of Mg^{2+} (1 mM) and dNTPs (1 or 5 µM, representing the lower and upper estimates for cellular dNTP concentrations), the SMF-driven reactions were slower even when a 1.5x concentration of SMF (10 µl) was added. Addition of 1 mM

$MgCl_2$ and 1 μM exogenous dNTPs to a 5 μl SMF reaction resulted in extension that was as robust as the control, suggesting that the slower extension speed was due to cation or dNTP-starvation and not an inhibitory compound in the SMF.

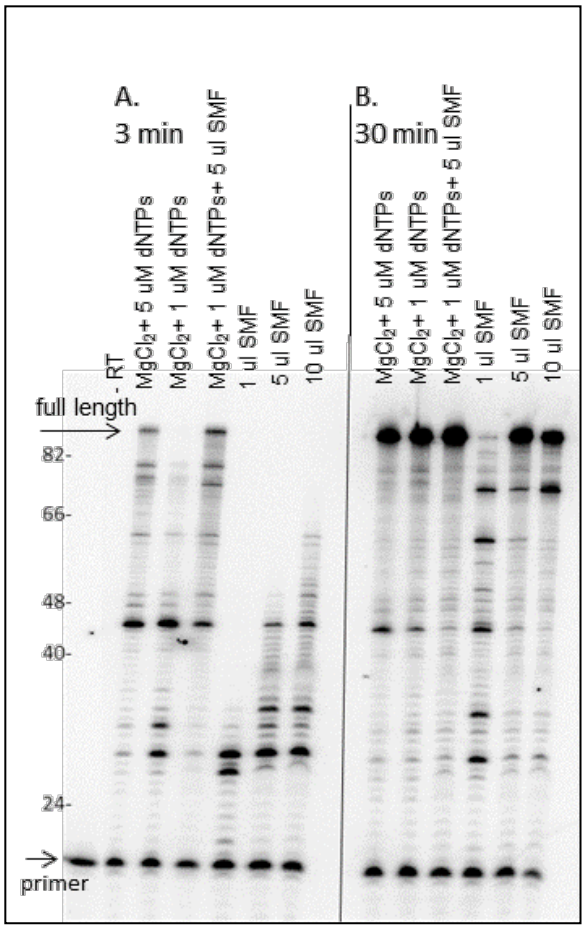


Figure 25. HIV RT extension of a ssDNA template stimulated by SMF. 3 (A) and 30 (B) minute extension reactions with known concentrations of $MgCl_2$ and dNTPs, SMF plus exogenous dNTPs (1 or 5 μM) and 1 mM $MgCl_2$, or SMF only. SMF contents are able to drive RT extension, although extension is slower than in reactions with estimated physiological concentrations of MgCl2 and dNTPs. Addition of 1 μM dNTPs to 5 μl of SMF restored extension speed.

It is most likely that the loss of cytosolic components occurs during the initial washes of the cell pellet, particularly during the hypotonic 0.1x PBS wash. Unfortunately, replacing the hypotonic wash with a second isotonic wash resulted in severe extension inhibition, most probably because excess PBS remaining on the pellet was excessively concentrated during the volume correction of the SMF.

Instead of extracting the SMF, it may be necessary to construct an artificial SMF using known concentrations of cellular components. This would allow for testing the compounds that affect characteristics of reverse transcription (speed of catalysis, fidelity, recombination, etc.) and could be used to improve existing RT assay protocols.

4.2 Inhibition of HIV replication in cell culture by aptamers against HIV RT

4.2.1 Introduction

As discussed in 3.1.2 PT-SELEX against HIV RT, numerous nucleic acid aptamers have been designed against HIV RT using SELEX and other methods. Unfortunately these aptamer inhibitors face many challenges for use as theraupeutics, namely in their ability to enter host cell and resist degradation. Uptake of nucleic acids by cells is limited and aptamers are easily degraded by host nucleases (reviewed in (126)). Off-target effects and cytotoxicity are additional concerns; the Toll pathway can be activated by foreign nucleic acids, inducing interferon (IFN) production (127).

Fortunately, many of these problems can be avoided. To protect against nucleases and degradation, aptamers can be made with modified sugar and thiophosphate backbones. Several teams are investigating methods to improve

aptamer entry, such as attaching aptamers to small protein anchors [128] or microsphere containers (129) which can more easily cross the cell membrane. Studies have also demonstrated that viral infection can promote cellular uptake of naked aptamers (130, 131), which suggests another promising avenue of directed anti-HIV aptamer entry.

Using DNA hairpin aptamers generated by PT SELEX against HIV RT (132), I began researching a new method of delivering aptamers using a DNA dendrimer container developed by the Muro lab (133) (Figure 26). These dendrimers contain multiple branched DNA modules made up of central double-stranded regions and four single-stranded arms, which can be coupled to antibodies. The antibodies are targeted against the cargo and also endocytic receptors on cells for delivery of cargo by cellular uptake. Inside the endosome, the low pH causes structural destabilization of the DNA, which releases the cargo and ultimately leads to escape. My research focused on determining if the dendrimers alone affected HIV replication or altered how the aptamers inhibited HIV; the dendrimers were not modified to carry the aptamer as cargo, although these experiments will be carried out in the future.

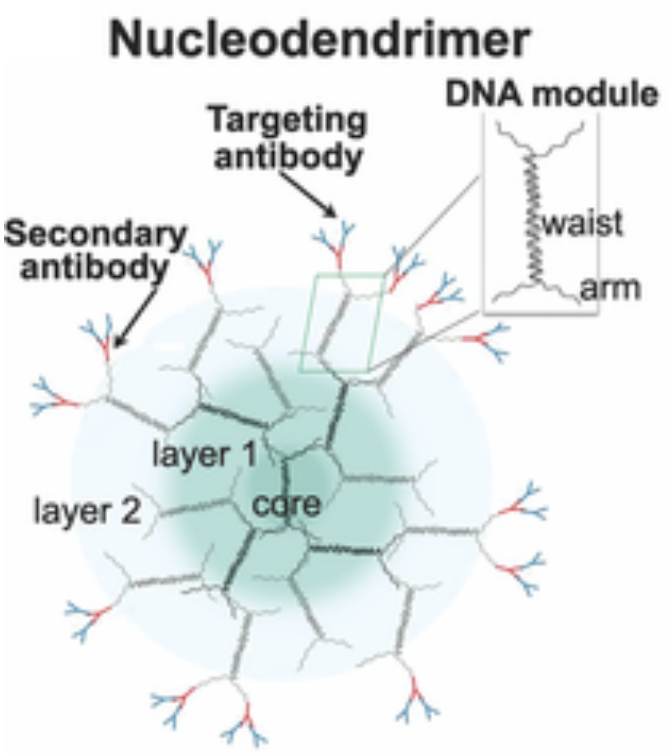


Figure 26. Structure of the nucleodendrimer. The dendrimer is made of DNA modules with branched ssDNA arms that can be attached to antibodies targeting the dendrimer to a cell and attaching to cargo. Figure modified from (133).

4.2.2 Materials and methods

Aptamers

The inhibitory hairpin aptamers (thioaptamer and 2,4 O-methyl aptamer, Figure 27) used in this study were initially generated using PT SELEX against HIV RT as described in (132). To protect against nuclease degradation, the 38 nt thioaptamer was engineered with three most 5' bases and the three bases predicted to be single stranded in the hairpin loop were substituted with thionucleosides. The 2,4 O-methyl aptamer is a second 38 nt hairpin aptamer (with a different hairpin sequence than the thioaptamer) which contains O-methyl bases at positions 2 and 4 bases from the 5' end of the hairpin. Aptamers were purchased from Integrated DNA Technologies.

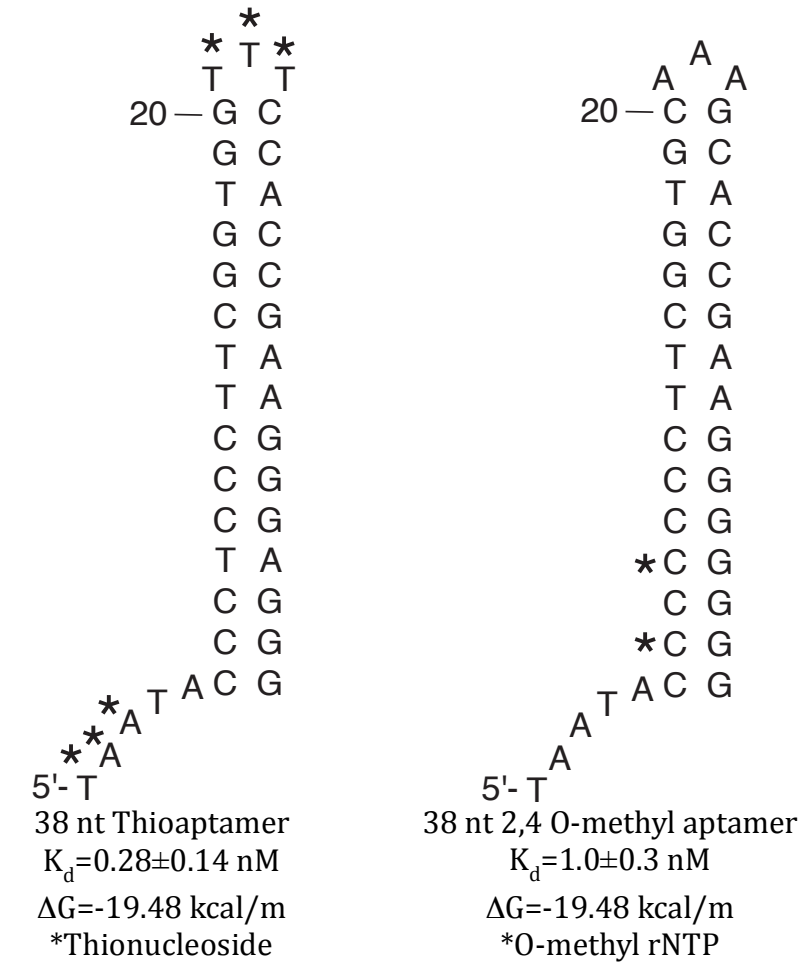


Figure 27. Aptamers used to test inhibition of HIV replication in cell culture. The 38 nt thioaptamer contains thionucleosides(*) to prevent nuclease degradation. The 2,4 O-methyl aptamer has O-methyl rNTPs (*) in the hairpin to take advantage of RT's increased affinity for substrates with ribonucleoside bases at those positions (134).

Cell culture and HIV infection

Infection studies were performed in TZM-bl cells, a HeLa-derived cell line which expresses the CD4 receptor and CCR5 coreceptor to permit HIV infection. This cell line contains a luciferase gene that is controlled by an upstream HIV-1-derived *tat* promoter, which is activated by HIV infection. Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum and 100 units of Penicillin/0.1 mg/ml of Streptomycin.

Wildtype HIV-1 was produced by transfection of 293T cells with the infectious clone pNL4-3. Pseudotyped HIV-1 was produced by cotransfection of pNL4-3/kfs (a plasmid which contains a frameshift mutation in the *env* gene, preventing production of gp120) and p-CMV-VSV-G (which produces the vesicular stomatitis virus G glycoprotein to serve as a replacement envelope protein that does not require a receptor for cellular entry).

24 hours before infection, cells were plated on opaque white 96-well plates at a $2x10^4$ cells/ml in 200 µl media. After 24 hours, the media was changed cells were infected with NL4-3-derived HIV at an MOI of 0.005 in the presence of aptamer (2 µM, unless otherwise stated), antibody-coupled dendrimers (1x concentration at100 ng/ml of DNA, 5 µg/ml antibody), antibody alone (5 µg/ml antibody), or antibody-coupled nanoparticles (6.8×10^{10} particles/ml, 5 µg/ml antibody). To measure infection, a luciferase assay was performed 24 hours post-infection using the britelite™ kit from Perkin Elmer and following manufacturer instructions. Relative Luciferase Units (RLUs) were determined using the Centro XS3LB 960 Microplate Luminometer from Berthold Technologies and MikroWin 2000 software. Results are

reported as a percentage of an HIV-only infection. Each condition was assayed in triplicate.

The dendrimers, antibodies, and nanoparticles were prepared and provided by the Muro lab, using the protocol in (133). Dendrimers targeted to TZM-bl cells were coupled to a monoclonal anti-transferrin receptor antibody, while non-targeted dendrimers were coupled to a rat anti-mouse IgG antibody. Nanoparticles were polystyrene coated with the anti-transferrin receptor antibody.

4.2.3 Results

Inhibition of HIV replication by the thioaptamer

The thioaptamer strongly inhibits HIV replication in a dose dependent manner. The highest concentration tested, 2 μM, reduces infection to ~10% of uninhibited control infections. This inhibition is unlikely to be due to a block at entry; both wildtype and pseudotyped HIV were inhibited by the thioaptamer, even when the aptamer was added 3 hours post infection.

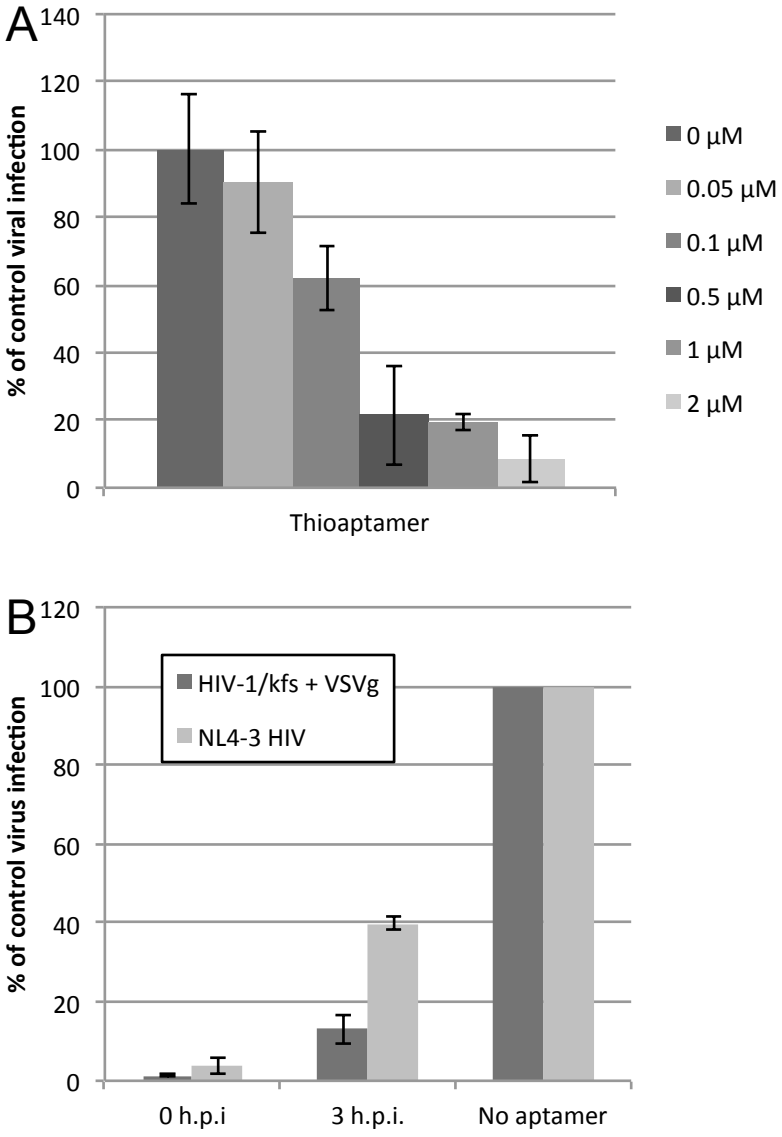


Figure 28. Inhibition of HIV replication by the thioaptamer. Only 0.5 μM is required to reduce HIV infection to ~20% of control and 2 μM reduces infection to ~10% of an uninhibited infection. This inhibition is not due to the aptamer preventing the virus from interacting with CD4 or entering the cell; receptor-independent pseudotyped HIV is also inhibited by the aptamer, as are both viruses when the aptamer is not added until 3 hours post infection (h.p.i.).

Co-infection of aptamers and dendrimers

The dendrimers were tested for their effect on cell growth and inhibition of HIV replication, both in the presence and absence of the thioaptamer (Figure 29). Cell density was measured 24 hours after the addition of the dendrimer; while there was variation in the cell counts, no significant changes were observed. Some inhibition of infection was observed with the addition of the dendrimers. Whether this is due to direct blocking of HIV or some other interaction is unknown. No effect was observed on the ability of the thioaptamer to inhibit HIV in the presence of the dendrimer.

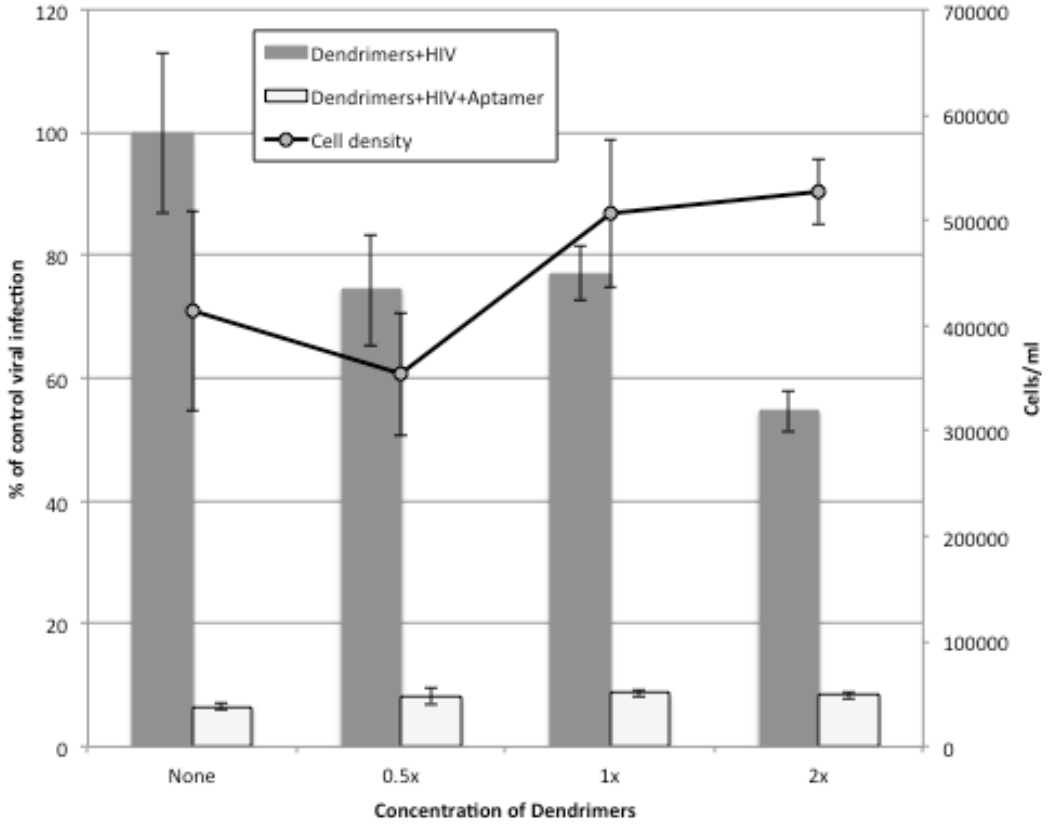


Figure 29. The effect of varying concentrations of dendrimers on HIV infection and cell growth in the presence and absence of thioaptamer. Dendrimers mildly inhibit HIV infection in a dose dependent manner and do not block the inhibitory effect of the thioaptamer (2 μM). At 24 hours post addition, the dendrimer had no significant effect on cell viability, which was measured using the trypan blue dye exclusion.

The thioaptamer's strong inhibitory effect on HIV replication makes observation of minor changes in inhibition difficult to observe. To test if the dendrimer could potentially increase the inhibitory effect of a non-cargo aptamer (perhaps reducing the concentration of aptamer necessary to reach therapeutic doses), we used the 2,4 O-methyl aptamer, which is only moderately inhibitory of infection and would allow the observation of both positive and negative effects by the dendrimers. Unfortunately, the dendrimers did not significantly alter the inhibitory effect of the aptamer and no synergistic effect was observed.

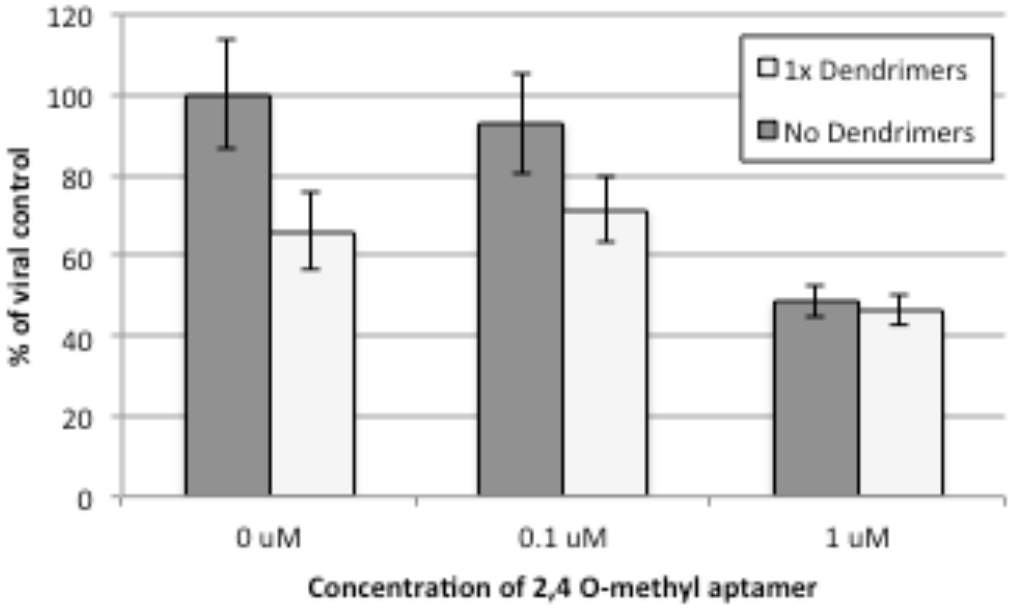


Figure 30. Inhibition of HIV replication by the 2,4 O-methyl aptamer in the presence and absence of the dendrimer. As the 2,4 O-methyl aptamer is less inhibitory than the thioaptamer, changes in inhibition are easier to visualize. Dendrimers are inhibitory of HIV infection on their own, but do not enhance or hinder the inhibitory effect of the 2,4 O-methyl aptamer.

As it was possible that the targeted nature of the dendrimers somehow blocked HIV entry into the cell, the individual components of the dendrimers were tested for their inhibitory effect on HIV replication (Figure 31). Dendrimers were prepared which contained an antibody (rat anti-mouse IgG) that was not targeted to the TZM-bl cells used in the assay. The anti-transferrin receptor antibody was also tested for its inhibitory effect, as were nanoparticles coated with the antibody (to determine if the DNA dendrimers themselves were inhibitory).

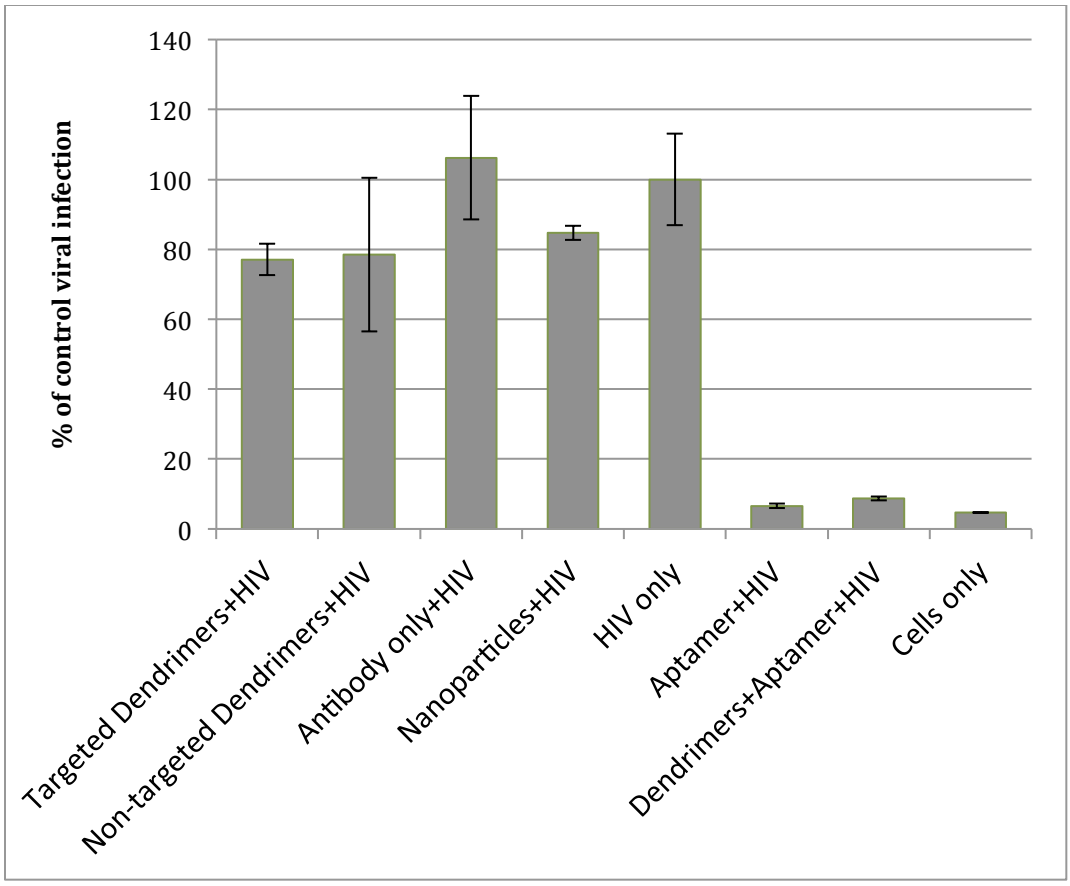


Figure 31. The effect of various dendrimer components on HIV infection. The inhibitory effect of the dendrimers on HIV infection is not due to interference by the anti-transferrin antibody or the composition of the dendrimers themselves. Non-targeted dendrimers (containing rat anti-mouse IgG) also inhibited HIV infection, as did nanoparticles complexed with the anti-transferrin antibody. The antibody alone did not inhibit HIV infection.

The antibody alone did not alter HIV infectivity, but the targeted and non-targeted dendrimers, as well as the targeted nanoparticles, inhibited the virus. This suggests that the dendrimers themselves are somehow inhibitory of infection but not due to their interaction with the targeted cell. Potentially, trace reagents used in the preparation of the dendrimers and nanoparticles remain present and inhibit HIV, but this is undetermined.

4.2.4 Discussion

The DNA dendrimers we tested do not appear to have any negative effects on cell growth or viability, nor do they block the action of the inhibitory aptamers. While they unfortunately do not have a synergistic effect with the aptamers (at least, when the aptamers are not attached as cargo), they are mildly inhibitory of HIV infection on their own. These results suggest that the dendrimers are a good candidate for use as cargo containers for anti-HIV aptamers. Whether they will increase the inhibitory effect of the aptamers (potentially lowering the amount necessary to reach therapeutic concentrations) remains to be seen through future research but these preliminary results are promising.

4.3 Conclusions

Often, translating *in vitro* research to a cellular infection encounters problems. RT assays with purified enzymes have been helpful in understanding the behavior of the enzyme during reverse transcription, but the unnatural conditions imposed by these reactions have produced results that, in some cases, vary wildly from the reality *in vivo*. Despite these deficits, *in vitro* assays remain a powerful tool for studying RT

under conditions that are impossible or impractical to achieve in cells. One way of reducing the discrepancies between *in vitro* and *in vivo* results would be to create a more lifelike assay, which incorporates cytosolic molecules that have an affect on HIV replication. I have attempted to create a protocol for extracting these molecules for use in RT assays, but unfortunately have been unable to prevent the loss of crucial components such as dNTPs.

SELEX is a versatile and powerful tool for generating inhibitory aptamers against viral protein targets such as HIV RT, however, they face severe difficulties in being used as cellular therapeutics: delivery of the aptamer into the cell, in particular, is a major hurdle. DNA dendrimers present an exciting opportunity to reduce the concentration of aptamers necessary for inhibition of HIV RT and could possibly be the next step in bringing this therapeutic further towards its use in patients.

Chapter 5: General Conclusions

Polymerases are integral to the lifecycle of all organisms and we have adapted them for use in numerous biotechnologies (102, 104, 135). Despite decades of research, fundamental information remains to be uncovered about their behavior.

In Chapter 2, I demonstrated that the behavior of HIV RT can be fundamentally altered by the cation composition during reverse transcription and dsDNA replication. Concentrations as low as 5 μM Zn^{2+} were able to slow the enzyme's nucleotide incorporation, even when the concentration of Mg^{2+} was 2 mM. Despite the potent inhibition by Zn^{2+} of reverse transcription, this divalent cation was also able to promote DNA synthesis from both RNA and DNA templates and at optimal concentration of 400 μM, which is $1/8^{th}$ the optimal concentration of Mg^{2+} and suggests a higher affinity of the enzyme for Zn^{2+}. In these Zn^{2+}-only reactions, HIV RT incorporated nucleotides at a maximum rate that was only $1/30^{th}$ the rate of a Mg^{2+}-driven reaction. The RNase H domain of RT remained functional in Zn^{2+} as well; in a titration performed in the absence of dNTPs, the RNase H domain reached optimal activity at 25 μM Zn^{2+}.

Despite the enzyme's poor catalytic speed in the presence of Zn^{2+}, its processivity (average number of nt added in a binding event) was relatively undiminished. This was due to the significantly increased stability of the enzyme-primer-template complex when Zn^{2+} was present: in 2 mM Mg^{2+}, the complex had a half-life of 1.7±1.0 min that increased to 220±60 min in the presence of 400 μM Zn^{2+}. The increased stability of the complex compensated for RT's slow catalysis, which

enabled some enzymes to complete reverse transcription of a 425 nt RNA template in a single binding event that lasted hours.

These results explain the mechanism by which Zn^{2+} inhibits HIV RT and, by extension, HIV replication. The enzyme is dramatically slowed down and synthesis is functionally inhibited, although the enzyme is still catalytically competent. I propose a model where Zn^{2+} displaces Mg^{2+} in the catalytic cleft. When bound to Zn^{2+}, the enzyme is unable to move between catalytic states, slowing both nucleotide incorporation and preventing dissociation from the template. To test if this is the case, it is possible to examine $k_{on/off}$ rates using the ^{65}Zn isotope to directly measure which cation is occupying the enzyme's metal ion binding sites under various conditions. This also presents an opportunity to determine if Zn^{2+} binds to more than just the metal ion binding sites in the catalytic cleft and RNase H domain, or if it is able to bind to other sites on the enzyme. If this is the case, it is possible that the restriction of the enzymes catalysis is due to an allosteric effect and not a direct action in the polymerase domain.

Currently, our lab is investigating the effect that Zn^{2+} and other cations have on RT fidelity. Given the profound effect that Zn^{2+} has on enzyme-primer-template complex stability and nucleotide incorporation, it is likely that the error rate of the enzyme will be impacted as well.

Based off our understanding that HIV reverse transcriptase's binding to primer-templates is impacted not just by the presence of different cations, but also the composition of the substrate, in Chapter 3 I investigated the ability of Family A DNA polymerases to selectively and preferentially bind specific sequences.

I determined that both Klenow and *Taq* pol had affinities for sequences that contained a portion of the single-subunit T7-like phage RNA polymerase promoter region. While this was not the only determinant of tight binding, it was an important factor for both enzymes. As the single-subunit RNAPs are structurally similar to Family A DNA polymerases, I hypothesize that this shared affinity is either retained from a common ancestor or a byproduct of the homologous structures in the polymerase domains of the single-subunit phage RNAPs and the Family A DNA pols.

A drawback to our method of determining the selected sequences from PT SELEX is that our final pool often only contains one or two sequence motifs and is a single snapshot in time of the evolution of sequences that lead to the final aptamers. A recent study by the Burke lab (136) demonstrated that using high throughput sequencing to monitor the populations in SELEX pools allows researchers to more readily identify motifs that bind tightly to the target. Rare sequences that bind the target but are of too low abundance to be detected by clonal sequencing can also be recovered using this method.

This is an attractive method that could potentially replace the method I described in Chapter 3, where I isolated sequences and performed mutational analysis to determine which nucleotides were important for tight binding. The volume of data generated by high throughput sequencing might narrow down which nucleotides are potentially involved in binding and which are not, saving time and money in analyzing the aptamer products of selection.

Finally, in Chapter 4 I presented two studies that began to translate research performed *in vitro* to the context of cellular infection. I presented a minimally

invasive method to recover the small molecular fraction of the cytosol for use in *in vitro* RT assays as a means to increase the accuracy of these studies relative to cellular RT activity. I also began research on a method of delivering aptamers to cells using DNA dendrimers, which addresses a significant barrier to the use of aptamers as anti-retroviral therapeutics and presents an exciting next step in aptamer research.

This work has attempted to uncover novel ways to modulate the behaviors of important biologically and commercially important proteins. Knowing that the behavior of HIV RT can be modulated so dramatically simply by the addition of different divalent cations during reverse transcription raises interesting questions about the effect these cations play during the viral lifecycle and, worryingly, about the validity of some *in vitro* measurements taken with purified enzymes in reactions containing Zn^{2+}. If RT is so profoundly affected by the addition of non-Mg^{2+} divalent cations, even at miniscule concentrations, it will be important in the future to determine what (if any) affect these molecules and others have on the behavior of the enzyme *in vivo*. Similarly, the understanding that the Family A DNA polymerases Klenow and *Taq* pol have sequence preferences for primer-templates, which they can bind up to 20x greater than a random substrate, suggests that these sequence preferences might be able to affect the behavior of these enzymes both in the context of the bacterial life cycle, as well as in the numerous commercial applications for which these enzymes are used. These results shed light on previous unknown traits of HIV RT and Family A DNA polymerases, underscoring the continued need for research into the behaviors of these enzymes.

Bibliography

1. **Gottlieb, M. S.** *Pneumocystis pneumonia--Los Angeles. 1981. Am J Public Health* **96,** 980–1– discussion 982–3 (2006).

2. **Centers for Disease Control (CDC)**. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb. Mortal. Wkly. Rep.* **30,** 305–308 (1981).

3. **Centers for Disease Control (CDC)**. Update on acquired immune deficiency syndrome (AIDS) among patients with hemophilia A. *MMWR Morb. Mortal. Wkly. Rep.* **31,** 644–6– 652 (1982).

4. **Masur, H., Michelis, M. A., Greene, J. B., Onorato, I., Stouwe, R. A., Holzman, R. S., Wormser, G., Brettman, L., Lange, M., Murray, H. W. & Cunningham-Rundles, S.** An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. *N Engl J Med* **305,** 1431–1438 (1981).

5. **Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L.** Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220,** 868–871 (1983).

6. **Gallo, R. C., Salahuddin, S. Z., Popovic, M. & Shearer, G. M.** Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* (1984).

7. **UNAIDS.** Global report: UNAIDS report on the global AIDS epidemic 2013. *unaids.org* (2013). at (http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf)

8. **Guyader, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L. & Alizon, M.** Genome organization and transactivation of the human immunodeficiency virus type 2. *Nature* **326,** 662–669 (1987).

9. **Plantier, J.-C., Leoz, M., Dickerson, J. E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Roberston, D. L. & Simon, F.** A new human immunodeficiency virus derived from gorillas. *Nat Med* **15,** 871–872 (2009).

10. **Taylor, B. S. & Hammer, S. M.** The challenge of HIV-1 subtype diversity. *N Engl J Med* **359,** 1965–1966 (2008).

11. **Sharp, P. M. & Hahn, B. H.** Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* **1,** a006841 (2011).

12. **Hemelaar, J., Gouws, E., Ghys, P. D., Osmanov, S.**WHO-UNAIDS Network for HIV Isolation and Characterisation. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* **25,** 679–689 (2011).

13. **Grossman, Z., Meier-Schellersheim, M., Paul, W. E. & Picker, L. J.** Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. *Nat Med* **12,** 289–295 (2006).

14. **Fischl, M. A., Richman, D. D., Grieco, M. H., Gottlieb, M. S., Volberding, P. A., Laskin, O. L., Leedom, J. M., Groopman, J. E., Mildvan, D. & Schooley, R. T.** 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,** 185–191 (1987).

15. *Antiretroviral drugs used in the treatment of HIV infection.* (FDA). at (http://www.fda.gov/ForConsumers/byAudience/ForPatientAdvocates/HIVandAIDSActivities/ucm118915.htm)

16. *Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents.* (AIDSinfo). at (http://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv-guidelines/10/initiating-art-in-treatment-na%C3%AFve-patients)

17. **Sungkanuparph, S., Anekthananon, T., Hiransuthikul, N., Supparatpinyo, K., Mootsikapun, P., Chetchotisakd, P., Kiertiburanakul, S., Tansuphaswadikul, S., Buppanharun, W. & Techasathit, W.** Guidelines for antiretroviral therapy in HIV-1 infected adults and adolescents: the recommendations of the Thai AIDS Society (TAS) 2008. *J Med Assoc Thai* **91,** 1925–1936 (2008).

18. *Strategic Timing of Antiretroviral Treatment (START)*. (ClinicalTrials.gov). at (http://clinicaltrials.gov/ct2/show/NCT00867048?term=NCT00867048&rank=1)

19. **Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müssig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O., Blau, I. W., Hofmann, W. K. & Thiel, E.** Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* **360,** 692–698 (2009).

20. *Stem-cell transplants may purge HIV*. (Nature News). at

(http://www.nature.com/news/stem-cell-transplants-may-purge-hiv-1.13297)

21. *Marrow Transplants Fail to Cure Two H.I.V. Patients*. (The New York Times). at (http://www.nytimes.com/2013/12/07/health/marrow-transplants-fail-to-cure-two-hiv-patients.html?_r=0)

22. **Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G. M., Kohn, D. B., Gregory, P. D., Holmes, M. C. & Cannon, P. M.** Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat. Biotechnol.* **28,** 839–847 (2010).

23. **Cassan, M., Delaunay, N., Vaquero, C. & Rousset, J. P.** Translational frameshifting at the gag-pol junction of human immunodeficiency virus type 1 is not increased in infected T-lymphoid cells. *J Virol* **68,** 1501–1508 (1994).

24. **Coffin, J. M., Hughes, S. H. & Varmus, H. E.** *Retroviruses*. (Cold Spring Harbor Laboratory Press, 1997).

25. **di Marzo Veronese, F., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C. & Sarngadharan, M. G.** Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. *Science* **231,** 1289–1291 (1986).

26. **Le Grice, S. F. J.** Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise. *Journal of Biological Chemistry* **287,** 40850–40857 (2012).

27. **Brautigam, C. A. & Steitz, T. A.** Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Current Opinion in Structural Biology* **8,** 54–63 (1998).

28. **Johnson, K. A.** The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim. Biophys. Acta* **1804,** 1041–1048 (2010).

29. **Delva, P., Pastori, C., Montesi, G., Degan, M., Micciolo, R., Paluani, F. & Lechi, A.** Intralymphocyte free magnesium and calcium and insulin tolerance test in a group of essential hypertensive patients. *Life Sci.* **63,** 1405–1415 (1998).

30. **Delva, P., Pastori, C., Degan, M., Montesi, G. & Lechi, A.** Catecholamine-induced regulation in vitro and ex vivo of intralymphocyte ionized magnesium. *J. Membr. Biol.* **199,** 163–171 (2004).

31. **Schultz, S. J. & Champoux, J. J.** RNase H activity: structure, specificity, and function in reverse transcription. *Virus Res* **134,** 86–103 (2008).

32. **Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A.** Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256,** 1783–1790 (1992).

33. **Mansky, L. M. & Temin, H. M.** Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol* **69,** 5087–5094 (1995).

34. **Kunkel, T. A.** DNA replication fidelity. *J Biol Chem* **279,** 16895–16898 (2004).

35. **Basu, V. P.**, **Song, M., Gao, L., Rigby, S. T., Hanson, M. N. & Bambara, R. A.** Strand transfer events during HIV-1 reverse transcription. *Virus Res* **134,** 19–38 (2008).

36. **Herschhorn, A. & Hizi, A.** Retroviral reverse transcriptases. *Cellular and Molecular Life Sciences* **67,** 2717–2747 (2010).

37. **Goff, S. P.** Retroviral reverse transcriptase: synthesis, structure, and function. *J. Acquir. Immune Defic. Syndr.* **3,** 817–831 (1990).

38. **Filler, A. G. & Lever, A. M.** Effects of cation substitutions on reverse transcriptase and on human immunodeficiency virus production. *AIDS Res Hum Retroviruses* **13,** 291–299 (1997).

39. **Bolton, E. C., Mildvan, A. S. & Boeke, J. D.** Inhibition of reverse transcription in vivo by elevated manganese ion concentration. *Mol. Cell* **9,** 879–889 (2002).

40. **Levinson, W., Faras, A., Woodson, B., Jackson, J. & Bishop, J. M.** Inhibition of RNA-dependent DNA polymerase of Rous sarcoma virus by thiosemicarbazones and several cations. *Proc Natl Acad Sci USA* **70,** 164–168 (1973).

41. **Palan, P. R. & Eidinoff, M. L.** Specific effect of zinc ions on DNA polymerase activity of avian myeloblastosis virus. *Mol. Cell. Biochem.* **21,** 67–69 (1978).

42. **Tan, C. K.**, **Zhang, J., Li, Z. Y., Tarpley, W. G., Downey, K. M. & So, A. G.** Functional characterization of RNA-dependent DNA polymerase and RNase H activities of a recombinant HIV reverse

transcriptase. *Biochemistry* **30,** 2651–2655 (1991).

43. **Mullen, G. P., Serpersu, E. H., Ferrin, L. J., Loeb, L. A. & Mildvan, A. S.** Metal binding to DNA polymerase I, its large fragment, and two 3‘,5’-exonuclease mutants of the large fragment. *J Biol Chem* **265,** 14327–14334 (1990).

44. **Maguire, M. E. & Cowan, J. A.** Magnesium chemistry and biochemistry. *Biometals* **15,** 203–210 (2002).

45. **Moomaw, A. S. & Maguire, M. E.** The unique nature of mg2+ channels. *Physiology (Bethesda)* **23,** 275–285 (2008).

46. **Traut, T. W.** Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140,** 1–22 (1994).

47. **Wang, S., McDonnell, E. H., Sedor, F. A. & Toffaletti, J. G.** pH effects on measurements of ionized calcium and ionized magnesium in blood. *Arch. Pathol. Lab. Med.* **126,** 947–950 (2002).

48. **Cousins, R. J., Liuzzi, J. P. & Lichten, L. A.** Mammalian zinc transport, trafficking, and signals. *J Biol Chem* **281,** 24085–24089 (2006).

49. **Dubben, S., Hönscheid, A., Winkler, K., Rink, L. & Haase, H.** Cellular zinc homeostasis is a regulator in monocyte differentiation of HL-60 cells by 1 alpha,25-dihydroxyvitamin D3. *J. Leukoc. Biol.* **87,** 833–844 (2010).

50. **Vallee, B. L. & Falchuk, K. H.** The biochemical basis of zinc physiology. *Physiol. Rev.* **73,** 79–118 (1993).

51. **Eide, D. J.** Zinc transporters and the cellular trafficking of zinc. *Biochim. Biophys. Acta* **1763,** 711–722 (2006).

52. **Coleman, J. E.** Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annu Rev Biochem* **61,** 897–946 (1992).

53. **Kitamura, H., Morikawa, H., Kamon, H., Iguchi, M., Hojyo, S., Fukada, T., Yamashita, S., Kaisho, T., Akira, S., Murakami, M. & Hirano, T.** Toll-like receptor-mediated regulation of zinc homeostasis influences dendritic cell function. *Nat. Immunol.* **7,** 971–977 (2006).

54. **Yamasaki, S.**, **Sakata-Sogawa, K., Hasegawa, A., Suzuki, T., Kabu, K., Sato, E., Kurosaki, T., Yamashita, S., Tokunaga, M. & Nishida, K.** Zinc is a novel intracellular second messenger. *J. Cell Biol.* **177,**

637–645 (2007).

55. **Hönscheid, A., Dubben, S., Rink, L. & Haase, H.** Zinc differentially regulates mitogen-activated protein kinases in human T cells. *J. Nutr. Biochem.* **23,** 18–26 (2012).

56. **Walther, U. I., Wilhelm, B., Walther, S. C., Mückter, H. & Forth, W.** Effect of zinc chloride on GSH synthesis rates in various lung cell lines. *In Vitr Mol Toxicol* **13,** 145–152 (2000).

57. **Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R., Rabson, A. & Martin, M. A.** Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol* **59,** 284–291 (1986).

58. **Hou, E. W., Prasad, R., Beard, W. A. & Wilson, S. H.** High-level expression and purification of untagged and histidine-tagged HIV-1 reverse transcriptase. *Protein Expression and Purification* **34,** 75–86 (2004).

59. **Sambrook, J. & Russell, D. W.** In vitro mutagenesis using double-stranded DNA templates: selection of mutants with DpnI. *Molecular Cloning* (2001).

60. **Goldschmidt, V., Didierjean, J., Ehresmann, B., Ehresmann, C., Isel, C. & Marquet, R.** Mg^{2+} dependency of HIV-1 reverse transcription, inhibition by nucleoside analogues and resistance. *Nucleic Acids Research* **34,** 42–52 (2006).

61. **DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Fay, P. J. & Bambara, R. A.** Parameters that influence processive synthesis and site-specific termination by human immunodeficiency virus reverse transcriptase on RNA and DNA templates. *Biochim. Biophys. Acta* **1131,** 270–280 (1992).

62. **Huang, H., Chopra, R., Verdine, G. L. & Harrison, S. C.** Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282,** 1669–1675 (1998).

63. **Yarrington, R. M., Chen, J., Bolton, E. C. & Boeke, J. D.** Mn^{2+} suppressor mutations and biochemical communication between Ty1 reverse transcriptase and RNase H domains. *J Virol* **81,** 9004–9012 (2007).

64. **Wang, M., Lee, H. R. & Konigsberg, W.** Effect of A and B metal ion site occupancy on conformational changes in an RB69 DNA polymerase

ternary complex. *Biochemistry* **48,** 2075–2086 (2009).

65. **Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H. & Kraut, J.** A structural basis for metal ion mutagenicity and nucleotide selectivity in human DNA polymerase beta. *Biochemistry* **35,** 12762–12777 (1996).

66. **Irimia, A., Zang, H., Loukachevitch, L. V., Eoff, R. L., Guengerich, F. P. & Egli, M.** Calcium is a cofactor of polymerization but inhibits pyrophosphorolysis by the Sulfolobus solfataricus DNA polymerase Dpo4. *Biochemistry* **45,** 5949–5956 (2006).

67. **Cristofaro, J. V., Rausch, J. W., Le Grice, S. F. J. & Destefano, J. J.** Mutations in the Ribonuclease H Active Site of HIV−RT Reveal a Role for This Site in Stabilizing Enzyme−Primer−Template Binding. *Biochemistry* **41,** 10968–10975 (2002).

68. **Cromme, F. V., Grüninger-Leitch, F. & Le Grice, S.** Point mutations in conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function. *FEBS Lett.* (1989).

69. **Tong, W., Lu, C.,-D., Sharma, S. K., Matsuura, S., So, A. G. & Scott, W. A.** Nucleotide-Induced Stable Complex Formation by HIV-1 Reverse Transcriptase †. *Biochemistry* **36,** 5749–5757 (1997).

70. **Wöhrl, B. M., Krebs, R., Goody, R. S. & Restle, T.** Refined model for primer/template binding by HIV-1 reverse transcriptase: pre-steady-state kinetic analyses of primer/template binding and nucleotide incorporation events distinguish between different binding modes depending on the nature of the nucleic acid substrate. *Journal of Molecular Biology* **292,** 333–344 (1999).

71. **Plum, L. M., Rink, L. & Haase, H.** The essential toxin: impact of zinc on human health. *Int J Environ Res Public Health* **7,** 1342–1365 (2010).

72. **Ellington, A. D. & Szostak, J. W.** In vitro selection of RNA molecules that bind specific ligands. *Nature* **346,** 818–822 (1990).

73. **Tuerk, C. & Gold, L.** Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249,** 505–510 (1990).

74. **Pinheiro, V. B., Taylor, A. I., Cozens, C., Abramov, M., Renders, M., Zhang, S., Chaput, J. C., Wengel, J., Peak-Chew, S.,-Y., McLaughlin, S. H., Herdewijn, P. & Holliger, P.** Synthetic genetic polymers capable of heredity and evolution. *Science* **336,** 341–344

(2012).

75. **Darfeuille, F.**, **Reigadas, S., Hansen, J. B., Orum, H., Di Primo, C., Toulmé, J.,-J.** Aptamers targeted to an RNA hairpin show improved specificity compared to that of complementary oligonucleotides. *Biochemistry* **45,** 12076–12082 (2006).

76. **Kensch, O.**, **Connolly, B. A., Steinhoff, H. J., McGregor, A., Goody, R. S. & Restle, T.** 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 (2000).

77. **Zhou, J., Swiderski, P., Li, H., Zhang, J., Neff, C. P., Akkina, R. & Rossi, J. J.** Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells. *Nucleic Acids Research* **37,** 3094–3109 (2009).

78. **Kraus, E., James, W. & Barclay, A. N.** Cutting edge: novel RNA ligands able to bind CD4 antigen and inhibit CD4+ T lymphocyte function. *J. Immunol.* **160,** 5209–5212 (1998).

79. **Long, S. B., Long, M. B., White, R. R. & Sullenger, B. A.** Crystal structure of an RNA aptamer bound to thrombin. *RNA* **14,** 2504–2512 (2008).

80. **Savory, N., Abe, K., Sode, K. & Ikebukuro, K.** Selection of DNA aptamer against prostate specific antigen using a genetic algorithm and application to sensing. *Biosens Bioelectron* **26,** 1386–1391 (2010).

81. **Germer, K., Leonard, M. & Zhang, X.** RNA aptamers and their therapeutic and diagnostic applications. *Int J Biochem Mol Biol* **4,** 27–40 (2013).

82. **Gragoudas, E. S., Adamis, A. P., Cunningham, E. T., Feinsod, M. & Guyer, D. R.** Pegaptanib for neovascular age-related macular degeneration. *N Engl J Med* **351,** 2805–2816 (2004).

83. **Ni, X., Castanares, M., Mukherjee, A. & Lupold, S. E.** Nucleic acid aptamers: clinical applications and promising new horizons. *Curr. Med. Chem.* **18,** 4206–4214 (2011).

84. **Joshi, P. & Prasad, V. R.** 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 (2002).

85. **Joshi, P. J., North, T. W. & Prasad, V. R.** 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,** 677–686 (2005).

86. **Destefano, J. J. & Cristofaro, J. V.** Selection of primer-template sequences that bind human immunodeficiency virus reverse transcriptase with high affinity. *Nucleic Acids Research* **34,** 130–139 (2006).

87. **Nair, G. R., Dash, C., Le Grice, S. F. J. & Destefano, J. J.** Viral reverse transcriptases show selective high affinity binding to DNA-DNA primer-templates that resemble the polypurine tract. *PLoS ONE* **7,** e41712 (2012).

88. **Kvaratskhelia, M., Budihas, S. R. & Le Grice, S. F. J.** Pre-existing distortions in nucleic acid structure aid polypurine tract selection by HIV-1 reverse transcriptase. *J Biol Chem* **277,** 16689–16696 (2002).

89. **Rausch, J. & Le Grice, S.** Binding, bending and bonding': polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus. *The International Journal of Biochemistry & Cell Biology* **36,** 1752–1766 (2004).

90. **Coté, M. L., Pflomm, M. & Georgiadis, M. M.** Staying straight with A-tracts: a DNA analog of the HIV-1 polypurine tract. *Journal of Molecular Biology* **330,** 57–74 (2003).

91. **Jones, F. D. & Hughes, S. H.** In vitro analysis of the effects of mutations in the G-tract of the human immunodeficiency virus type 1 polypurine tract on RNase H cleavage specificity. *Virology* **360,** 341–349 (2007).

92. **Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Alvord, W. G., Arnold, E. & Hughes, S. H.** Effects of mutations in the G tract of the human immunodeficiency virus type 1 polypurine tract on virus replication and RNase H cleavage. *J Virol* **78,** 13315–13324 (2004).

93. **Jacob, D. T. & Destefano, J. J.** A new role for HIV nucleocapsid protein in modulating the specificity of plus strand priming. *Virology* **378,** 385–396 (2008).

94. **Post, K., Kankia, B., Gopalakrishnan, S., Yang, V., Cramer, E., Saladores, P., Gorelick, R. J., Guo, J., Musier-Forsyth, K. & Levin, J. G.** Fidelity of plus-strand priming requires the nucleic acid chaperone activity of HIV-1 nucleocapsid protein. *Nucleic Acids Research* **37,** 1755–1766 (2009).

95. **Steitz, T. A.** DNA polymerases: structural diversity and common mechanisms. *J Biol Chem* **274,** 17395–17398 (1999).

96. **Pastor-Palacios, G., Azuara-Liceaga, E. & Brieba, L. G.** A Nuclear Family A DNA Polymerase from Entamoeba histolytica Bypasses Thymine Glycol. *PLoS Negl Trop Dis* **4,** e786 (2010).

97. **Rothwell, P. J. & Waksman, G.** Structure and mechanism of DNA polymerases. *Adv. Protein Chem.* **71,** 401–440 (2005).

98. **Tabor, S., Huber, H. E. & Richardson, C. C.** Escherichia coli thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J Biol Chem* **262,** 16212–16223 (1987).

99. **Lehman, I. R., Bessman, M. J., Simms, E. S. & Kornberg, A.** Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. *J Biol Chem* **233,** 163–170 (1958).

100. **De Lucia, P. & Cairns, J.** Isolation of an E. coli strain with a mutation affecting DNA polymerase. *Nature* **224,** 1164–1166 (1969).

101. **Englisch, U., Gauss, D. & Freist, W.** Error rates of the replication and expression of genetic information. ... *Edition in English* (1985).

102. **Chien, A., Edgar, D. B. & Trela, J. M.** Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *J. Bacteriol.* **127,** 1550–1557 (1976).

103. **Lawyer, C., Stoffel, S., Saiki, R. K., Chang, S. Y., Landre, P. A., Abramson, R. D., and Gelfand, D. H.** High-level expression, purification, and enzymatic characterization of full-length Thermus aquaticus DNA polymerase and a truncated form deficient in 5‘ to 3’ exonuclease activity. *PCR Methods Appl.* **2,** 275–287 (1993).

104. **Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A.** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239,** 487–491 (1988).

105. **Tindall, K. R. & Kunkel, T. A.** Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry* **27,** 6008–6013 (1988).

106. **Beese, L. S., Derbyshire, V. & Steitz, T. A.** Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* (1993).

107. **Eom, S. H., Wang, J. & Steitz, T. A.** Structure of Taq polymerase with DNA at the polymerase active site. 1–4 (2013).

108. **Chapman, K. A. & Burgess, R. R.** Construction of bacteriophage T7 late promoters with point mutations and characterization by in vitro transcription properties. *Nucleic Acids Research* **15,** 5413–5432 (1987).

109. **Chapman, K. A., Gunderson, S. I., Anello, M., Wells, R. D. & Burgess, R. R.** Bacteriophage T7 late promoters with point mutations: quantitative footprinting and in vivo expression. *Nucleic Acids Research* **16,** 4511–4524 (1988).

110. **Li, T., Ho, H. H., Maslak, M., Schick, C. & Martin, C. T.** Major groove recognition elements in the middle of the T7 RNA polymerase promoter. *Biochemistry* **35,** 3722–3727 (1996).

111. **Újvári, A. & Martin, C. T.** Identification of a minimal binding element within the T7 RNA polymerase promoter. *Journal of Molecular Biology* **273,** 775–781 (1997).

112. **Nayak, D., Guo, Q. & Sousa, R.** A Promoter Recognition Mechanism Common to Yeast Mitochondrial and Phage T7 RNA Polymerases. *Journal of Biological Chemistry* **284,** 13641–13647 (2009).

113. **Bonner, G., Patra, D., Lafer, E. M. & Sousa, R.** Mutations in T7 RNA polymerase that support the proposal for a common polymerase active site structure. *EMBO J* **11,** 3767–3775 (1992).

114. **Ho, D. L., Byrnes, W. M., Ma, W.-P., Shi, Y., Callaway, D. J. E. & Bu, Z.** Structure-specific DNA-induced conformational changes in Taq polymerase revealed by small angle neutron scattering. *J Biol Chem* **279,** 39146–39154 (2004).

115. **Datta, K. & LiCata, V. J.** Thermodynamics of the binding of Thermus aquaticus DNA polymerase to primed-template DNA. *Nucleic Acids Research* **31,** 5590–5597 (2003).

116. **Datta, K., Johnson, N. P., LiCata, V. J. & Hippel, von, P. H.** Local conformations and competitive binding affinities of single- and double-stranded primer-template DNA at the polymerization and editing active sites of DNA polymerases. *Journal of Biological Chemistry* **284,** 17180–17193 (2009).

117. **Turner, R. M., Grindley, N. D. F. & Joyce, C. M.** Interaction of DNA polymerase I (Klenow fragment) with the single-stranded template beyond the site of synthesis. *Biochemistry* **42,** 2373–2385 (2003).

118. **DeStefano, J.** The orientation of binding of human immunodeficiency virus reverse transcriptase on nucleic acid hybrids. *Nucleic Acids Research* **23,** 3901 (1995).

119. **DeStefano, J. J., Bambara, R. A. & Fay, P. J.** Parameters that influence the binding of human immunodeficiency virus reverse transcriptase to nucleic acid structures. *Biochemistry* **32,** 6908–6915 (1993).

120. **DeStefano, J., Cristofaro, J., Derebail, S., Bohlayer, W. & Fitzgerald-Heath, M.** Physical mapping of HIV reverse transcriptase to the 5′ end of RNA primers. *Journal of Biological Chemistry* **276,** 32515 (2001).

121. **Palaniappan, C., Kim, J. K., Wisniewski, M., Fay, P. J. & Bambara, R. A.** Control of initiation of viral plus strand DNA synthesis by HIV reverse transcriptase. *J Biol Chem* **273,** 3808–3816 (1998).

122. **Diamond, T. L., Roshal, M., Jamburuthugoda, V. K., Reynolds, H. M., Merriam, A. R., Lee, K. Y., Balakrishnan, M., Bambara, R. A. & Planelles, V.** Macrophage tropism of HIV-1 depends on efficient cellular dNTP utilization by reverse transcriptase. *J Biol Chem* **279,** 51545–51553 (2004).

123. **Bakhanashvili, M., Novitsky, E., Levy, I. & Rahav, G.** The fidelity of DNA synthesis by human immunodeficiency virus type 1 reverse transcriptase increases in the presence of polyamines. *FEBS Lett.* **579,** 1435–1440 (2005).

124. **Abmayr, S. M., Yao, T., Parmely, T. & Workman, J. L.** Preparation of nuclear and cytoplasmic extracts from mammalian cells. *Curr Protoc Pharmacol* **Chapter 12,** Unit12.3 (2006).

125. **Rosenbluth, M. J., Lam, W. A. & Fletcher, D. A.** Force Microscopy of Nonadherent Cells: A Comparison of Leukemia Cell Deformability. *Biophysical Journal* **90,** 2994–3003 (2006).

126. **Patil, S. D., Rhodes, D. G. & Burgess, D. J.** DNA-based therapeutics and DNA delivery systems: a comprehensive review. *AAPS J* **7,** E61–77 (2005).

127. **Akira, S., Uematsu, S. & Takeuchi, O.** Pathogen recognition and innate immunity. *Cell* (2006).

128. **Tripathi, S., Chaubey, B., Barton, B. E. & Pandey, V. N.** Anti HIV-1 virucidal activity of polyamide nucleic acid-membrane transducing peptide conjugates targeted to primer binding site of HIV-1 genome. *Virology* **363,** 91–103 (2007).

129. **Carrasquillo, K. G., Ricker, J. A., Rigas, I. K. & J. W. Miller, J. W.** Controlled delivery of the anti-VEGF aptamer EYE001 with poly (lactic-co-glycolic) acid microspheres. ... *ophthalmology & visual* ... (2003).

130. **Matzen, K., Elzaouk, L., Matskevich, A. A., Nitzche, A., Heinrich, J. & Moelling, K.** RNase H-mediated retrovirus destruction in vivo triggered by oligodeoxynucleotides. *Nat. Biotechnol.* **25,** 669–674 (2007).

131. **Métifiot, M., Leon, O., Tarrago-Litvak, L., Litvak, S. & Andréola, M.-L.** Targeting HIV-1 integrase with aptamers selected against the purified RNase H domain of HIV-1 RT. *Biochimie* **87,** 911–919 (2005).

132. **Destefano, J. J. & Nair, G. R.** Novel aptamer inhibitors of human immunodeficiency virus reverse transcriptase. *Oligonucleotides* **18,** 133–144 (2008).

133. **Muro, S. A** DNA Device that Mediates Selective Endosomal Escape and Intracellular Delivery of Drugs and Biologicals. *Adv. Funct. Mater.* n/a–n/a (2014). doi:10.1002/adfm.201303188

134. **Olimpo, J. T. & Destefano, J. J.** Duplex structural differences and not 2'-hydroxyls explain the more stable binding of HIV-reverse transcriptase to RNA-DNA versus DNA-DNA. *Nucleic Acids Research* **38,** 4426–4435 (2010).

135. **Weis, J. H., Tan, S. S., Martin, B. K. & Wittwer, C. T.** Detection of rare mRNAs via quantitative RT-PCR. *Trends Genet.* **8,** 263–264 (1992).

136. **Ditzler, M. A., Lange, M. J., Bose, D., Bottoms, C. A., Virkler, K. F., Sawyer, A. W., Whatley, A. S., Spollen, W. S., Givan, S. A. & Burke, D. H.** High-throughput sequence analysis reveals structural diversity and improved potency among RNA inhibitors of HIV reverse transcriptase. *Nucleic Acids Research* **41,** 1873–1884 (2013).