

## ABSTRACT

Title of Document: A COMPARATIVE ANALYSIS OF THE BINDING AFFINITY OF HIV-1 REVERSE TRANSCRIPTASE TO DNA vs. RNA SUBSTRATES

Jeffrey T. Olimpo, Jr., Master of Science, 2010

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

Human immunodeficiency virus reverse transcriptase (HIV-RT) binds more stably in binary complexes with RNA–DNA versus DNA–DNA. Current results indicate that only the -2 and -4 RNA nucleotides (-1 hybridized to the 3' recessed DNA base) are required for stable binding to RNA–DNA, and even a single RNA nucleotide conferred significantly greater stability than DNA–DNA. Replacing 2'-hydroxyls on pivotal RNA bases with 2'-O-methyls did not affect stability, indicating that interactions between hydroxyls and RT amino acids do not stabilize binding. Avian myeloblastosis and Moloney murine leukemia virus RTs also bound more stably to RNA–DNA, but the difference was less pronounced than with HIV-RT. We propose that the H- versus B-form structures of RNA–DNA and DNA–DNA, respectively, allow the former to conform more easily to HIV-RT's binding cleft, leading to more

stable binding. Biologically, this may aid in degradation of RNA fragments that remain after DNA synthesis.

A COMPARATIVE ANALYSIS OF THE BINDING AFFINITY OF HIV-1  
REVERSE TRANSCRIPTASE TO DNA vs. RNA SUBSTRATES

By

Jeffrey T. Olimpo, Jr.

Thesis 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  
Master of Science  
2010

Advisory Committee:  
Dr. Jeffrey DeStefano, Chair  
Dr. James Culver  
Dr. Brenda Frederickson

© Copyright by  
Jeffrey T. Olimpo, Jr.  
2010

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Jeffrey DeStefano, for his continued patience and support during my time in the lab. His mentorship was invaluable whether we were at the bench or simply discussing music, politics, or the news.

I would also like to thank the members of my advisory committee: Drs. James Culver and Brenda Frederickson, who have provided both advice and constructive criticism during the course of my research experience. Their input has greatly shaped how I approach and consider my own work.

To my fellow lab members, both past and present, you have been a source of inspiration and my backbone during this experience. I applaud Gauri, who has made many valiant attempts to convert me to Trader Joe's and Whole Foods, where they stock merchandise that is perhaps grown fresh on her farm. I thank Megan for her continuous supply of desserts and for being the human version of the Encyclopedia Britannica. I admire Katherine, who always manages to keep on top of any topic or trend. Divya, my unbiological twin sister, we both know words are not sufficient to express my thanks. Deena—you are also not to be forgotten; thank you for your humor and warm heart. Best of luck to you all in the future, and I love you.

Last, but certainly not least, I could not have completed this journey without the support of my family and friends. Thank you for your countless hours of support and motivation, for reassuring phone conversations, and for late nights spent convincing me that there was always a light at the end of the tunnel. I could not have done this without you.

# TABLE OF CONTENTS

Acknowledgements.....	ii
Table of Contents.....	iii
List of Tables.....	v
List of Figures.....	vi
List of Abbreviations.....	vii
<b>Chapter 1: HIV and the AIDS Epidemic</b>	
1.1 Structural and Genetic Properties of HIV-1.....	1
1.2 Life Cycle of HIV-1 in the Host.....	6
1.3 Epidemiological Aspects of Disease Acquired Immunodeficiency Virus (AIDS) and Therapeutic and Preventative Measures Against HIV-1 Infection.....	9
<b>Chapter 2: Analysis of Those Requirements Necessary and Sufficient for Tighter Binding of HIV-1 Reverse Transcriptase to RNA vs. DNA Substrates</b>	
2.1 Introduction	
2.1.1 The General Mechanism of HIV-1 Replication.....	15
2.1.2 Human Immunodeficiency Virus-1 Reverse Transcriptase (HIV-RT).....	17
2.1.3 Binding of HIV-RT to DNA and RNA Templates.....	20
2.2 Materials and Methods	
2.2.1 Materials.....	23
2.2.2 Preparation of Radiolabeled DNA Primer Strands.....	23
2.2.3 Preparation of RNA-DNA and DNA-DNA Hybrids.....	23
2.2.4 Determination of Dissociation Rate Constants by Primer Extension.....	24
2.2.5 Determination of Equilibrium Dissociation Constants ( $k_d$ ) by Primer Extension.....	25
2.2.6 Visualization and Quantification of Primer Extension.....	25
2.3 Results	
2.3.1 RNA-DNA and DNA-DNA Substrates Used to Test HIV-RT Binding.....	27
2.3.2 RT Dissociates from Templates with Just Two RNA Nucleotides at a Rate Similar to a Complete RNA Template and Even a Single RNA Nucleotide at the -4 Position Results in a Dissociation Rate Much Slower than a Complete DNA Template.....	31
2.3.3 A Uridine Residue at the -4 Position Promotes More Stable HIV-RT Binding than Other RNA Nucleotides.....	35
2.3.4 Moving the -4 Uridine Residue to Different Positions Relative to the DNA 3' Recessed Terminus Increased HIV-RT's Dissociation Rate, Confirming the Importance of the -4 Position in the RNA.....	36
2.3.5 The 2'-Hydroxyl Groups on the RNA Nucleotides are not Required to Stabilize Binding of HIV-RT.....	37
2.3.6 Slower Dissociation of HIV-RT from RNA-DNA is not Dependent upon Sequence.....	37
2.3.7 The Equilibrium Dissociation Constant ( $K_d$ ) for HIV-RT is Nearly the	

Same for RNA-DNA and DNA-DNA Despite Large Differences in $k_{\text{off}}$ .....	39
<b>Chapter 3: Analysis of the Preferential Binding of MuLV-RT and AMV-RT to RNA vs. DNA Substrates</b>	
<b>3.1 Introduction</b>	
<b>3.1.1</b> Moloney Murine Leukemia Virus Reverse Transcriptase (MuLV-RT) .....	41
<b>3.1.2</b> Avian Myeloblastosis Virus Reverse Transcriptase (AMV-RT) .....	42
<b>3.2</b> Materials and Methods.....	44
<b>3.3</b> AMV-RT and MuLV-RT Show More Modest Preference for Binding to RNA-DNA versus DNA-DNA.....	45
<b>Chapter 4: Discussion</b>	
<b>4.1</b> Slower Dissociation of HIV-RT from RNA-DNA versus DNA-DNA does not Require 2'-Hydroxyls and is Dependent on the RNA Nucleotides at the -2 and -4 Positions.....	46
<b>4.2</b> The H-Form versus B-Form Structure of RNA-DNA and DNA-DNA Hybrids, Respectively, is the Most Likely Explanation for HIV-RT's Slower Dissociation from RNA-DNA.....	47
<b>4.3</b> Both On- and Off-Rates Appear to be Different for RNA-DNA versus DNA-DNA Duplexes.....	49
<b>4.4</b> Structural Differences between HIV-, AMV- and MuLV-RT May Explain Differences in Duplex Binding.....	50
<b>Chapter 5: Future Directions</b>	
<b>5.1</b> Binding of Non-Retroviral RT Species to RNA versus DNA Substrates	
<b>5.1.1</b> Ty3 Reverse Transcriptase.....	52
<b>5.1.2</b> Biological Relevance for Different Binding States on DNA-DNA and RNA-DNA.....	53
<b>5.2</b> Dissociation Studies of HIV-RT	
<b>5.2.1</b> Background on Dissociation Studies of HIV-RT.....	54
<b>5.2.2</b> Proposed Experiments to Analyze those Requirements Necessary for Dissociation of HIV-RT from its Respective Template.....	55
<b>5.2.3</b> Discussion of Possible Outcomes.....	57
<b>References</b> .....	59

## LIST OF TABLES

Table 1: Major Classes of Antiretroviral Drugs.....	8
Table 2: Dissociation rate constants ( $k_{\text{off}}$ ) and equilibrium dissociation constants ( $K_d$ ) for RT-substrate complexes.....	22

## LIST OF FIGURES

Figure 1: Structural Representation of the human immunodeficiency virus.....	1
Figure 2: Schematic Representation of the HIV-1 Genome and HIV-1 Proteins and their Functions.....	3
Figure 3: Schematic Representation of the HIV-1 Life Cycle.....	4
Figure 4: Global Distribution of HIV-1 Subtypes and Recombinants.....	7
Figure 5: Schematic Representation of Reverse Transcription in HIV-1.....	10
Figure 6: Schematic Representation of HIV-1 Reverse Transcriptase.....	12
Figure 7: Schematic Representation of a DNA-RNA Primer:Template Hybrid.....	14
Figure 8: Autoradiograms of Half-Life Binding Experiments with Primer 33 bound to 50-mer DNA or RNA Templates.....	19
Figure 9: Sequence and Configuration of Nucleic Acid Substrates.....	20
Figure 10: Panels for Dissociation Rate Constants ( $k_{\text{off}}$ ) .....	23
Figure 11: Equilibrium Dissociation Constants.....	30
Figure 12: Model of HIV-RT Extension on a Template Containing a Secondary Structural Modification.....	46

## LIST OF ABBREVIATIONS

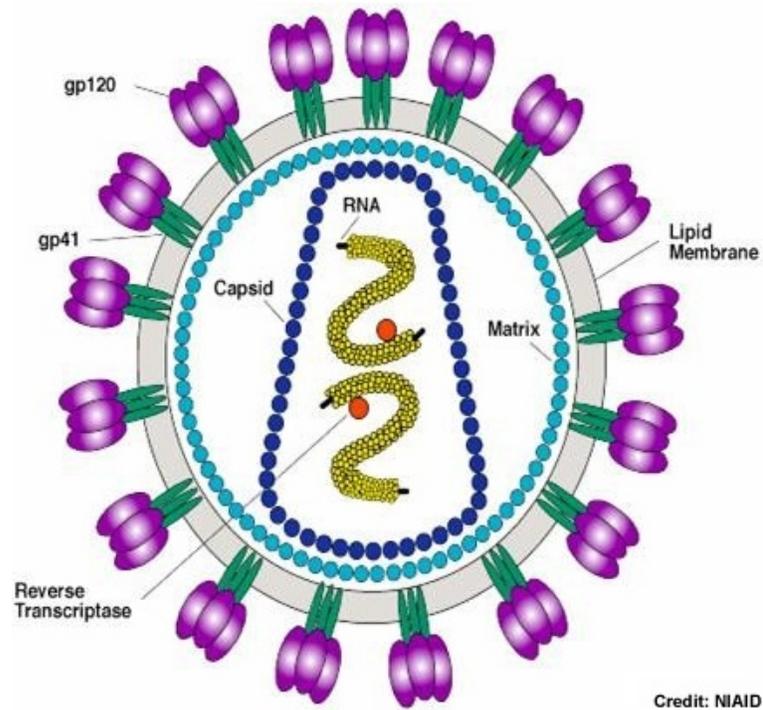
A:	Adenine
AA:	Amino Acids
AIDS:	Acquired Immunodeficiency Syndrome
AMV-RT:	Avian Myeloblastosis Virus Reverse Transcriptase
bp:	Base-Pair
C:	Cytosine
dNTP:	Deoxyribonucleotide Triphosphates
dsDNA:	Double-Stranded DNA
DTT:	Dithiothreitol
EDTA:	Ethylenediaminetetracetic Acid
G:	Guanine
HAART:	Highly Active Antiretroviral Therapy
HIV:	Human Immunodeficiency Virus
HIV-RT:	Human Immunodeficiency Virus Reverse Transcriptase
kb:	Kilobase
$K_D$ :	Equilibrium Dissociation Constant
kDa:	Kilodaltons
$k_{off}$ :	Dissociation Rate Constant
LTR:	Long Terminal Repeat
MAV:	Myeloblastosis-Associated Helper Virus
(m)M:	(Mili)Molar
m/min.:	Minutes

MuLV-RT:	Moloney Murine Leukemia Virus Reverse Transcriptase
nm:	Nanometers
nt:	Nucleotide
P33:	Primer 33
pmol:	Picomole
PNK:	Polynucleotide Kinase
PPT:	Polypurine Tract
RNase H:	Ribonuclease H
s:	Seconds
S2:	Sequence 2
siRNA:	Short interfering RNA
T:	Thymine
Ty3-RT:	Ty3 Reverse Transcriptase
U:	Uridine

# Chapter 1: HIV and the AIDS Epidemic

## 1.1 Structural and Genetic Properties of HIV-1

At 120 nm in diameter and with a genome approximately 9.3 kb in size, HIV-1 is structurally distinct from other members of the Retroviridae family to which it belongs (30). Though this is the case, it shares many common morphological and biological features with other members of the lentiviral genus, a group of viruses commonly associated with long-duration illnesses and prolonged periods of viral latency, such as a spheroid, enveloped structure and the possession of two copies of a positive-sense, single-stranded RNA genome (50; Figure 1).



**Figure 1.** Structural representation of the human immunodeficiency virus; adapted from <http://www.stanford.edu/group/virus/retro/2005gongishmail/hiv1.jpg>

Both the envelope proteins, as well as the structural and enzymatic proteins necessary for replication, integration, etc. are derived from this genome, which is composed primarily of the *env*, *gag*, and *pol* genes. In total, 15 distinct proteins are generated, largely via splicing of the full-length genomic mRNA. These proteins and their functions are illustrated in Figure 2 below.



Accessory Protein	Nef, Negative Factor	CD4 downregulation and T-cell activation
-------------------	----------------------	--

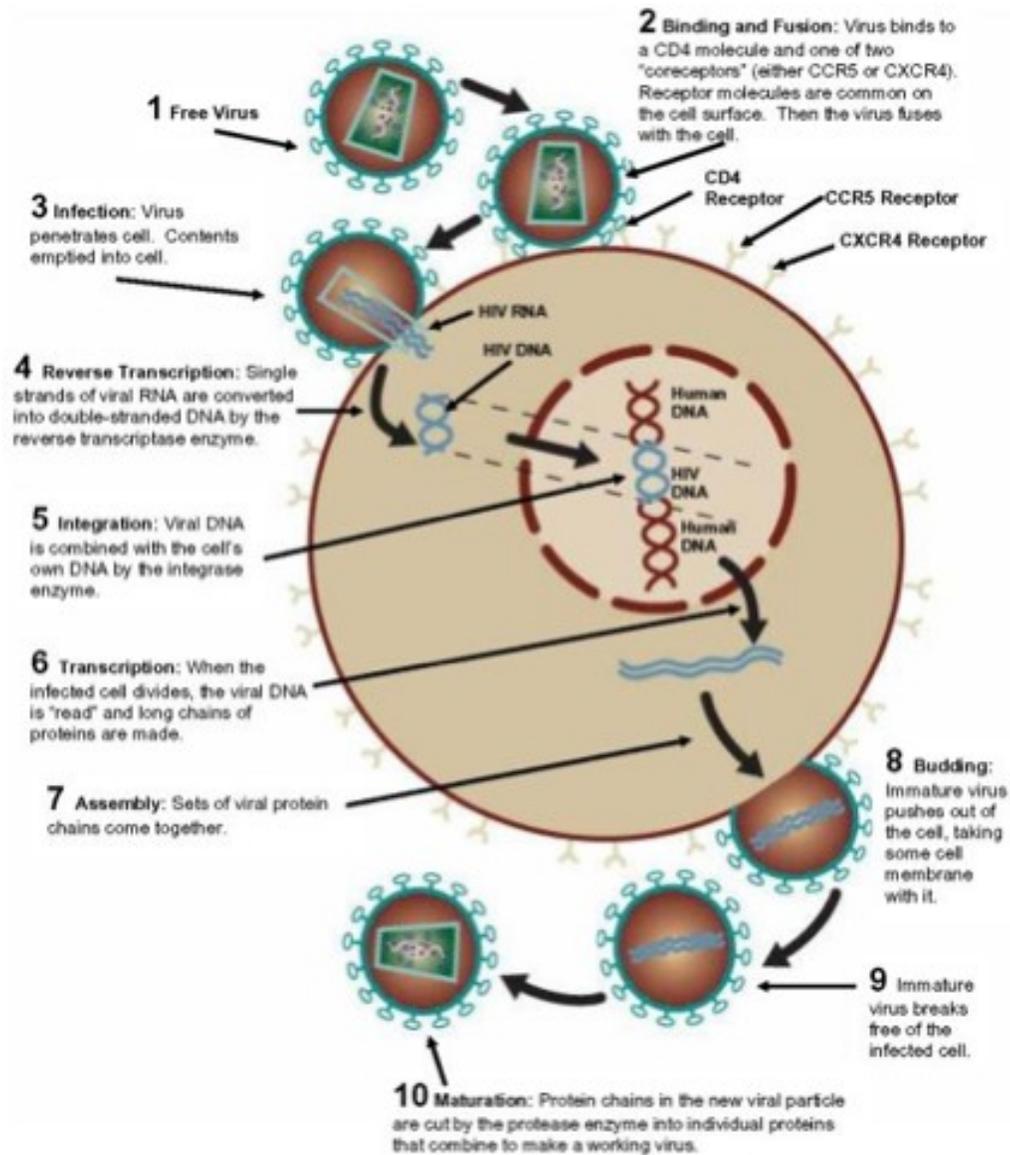
**Figure 2.** Schematic representation of the HIV-1 genome; adapted from <http://www.mimo.unige.ch/images/Genome.jpg>. HIV-1 proteins and their functions are adapted from (73).

It is also through proper transmission and replication of this genome that HIV-1 infection occurs.

### 1.2 Life Cycle of HIV-1 in the Host

In order for entry to occur, the HIV-1 surface glycoprotein, gp120, must interact with CD4 receptors on the surface of immune cells such as T cells and macrophages, the primary targets of HIV-1 infection. In addition to CD4 receptor binding, successful entry is dependent upon interaction of gp120 with co-receptors on the cell surface. The receptors commonly used for HIV entry are CCR5 (R5) and CXCR4 (X4), with the R5 and X4 moieties typically found on macrophages and T cells, respectively (12).

Once the virus enters the host cell, it uncoats and releases its contents into the cytoplasm (see Figure 3). It is here that the process of reverse transcription occurs. Reverse transcription, which is common to *Retroviridae* as well as select other viral and non-viral species (for instance, the *Saccharomyces Cerevisiae* Ty3 retrotransposon), involves conversion of single-stranded genomic RNA into a double-stranded DNA, which is subsequently transported to the host cell nucleus and integrated into the host genome. A more thorough description of this process can be found in section 2.1.1 below.



**Figure 3.** Schematic representation of the HIV-1 life cycle; adapted from (70).

The process of integration itself is facilitated by the enzyme integrase, which begins by cleaving two nucleotides from the 3' ends of both DNA strands, exposing the 3'- hydroxyl group on the terminal CA dinucleotide that is conserved among retroviruses and many transposons (15). The modified double-stranded DNA is then inserted into the target DNA sequence, and the host cell machinery is employed to ligate the sequence into place. While integration can potentially occur anywhere throughout the host genome, HIV-1 has been noted to exhibit a preference for integrating near sites of active gene transcription (15).

When the host cell divides, the newly integrated DNA is “read”<sup>1</sup> and RNA polymerase II carries out transcription, resulting in the production of viral RNA which, in the case of HIV-1, is subjected to the same processing steps as cellular mRNAs, including: a) 5' capping; b) 3' polyadenylation; c) splicing; etc. As partially illustrated in figure 2 above, RNAs are then utilized primarily to either act as the new viral genome or to encode the various proteins required for replication, viral maintenance, and virulence.

These proteins, which are detailed also in Figure 2, are synthesized externally of the nucleus and are largely derived from the proteolytic cleavage of the *Gag-Pol* and *env* polyproteins by viral protease or by cellular proteases, respectively (32). Proper maintenance of the ratio of the Gag and Pol proteins is critical for RNA dimerization and virion infectivity and is regulated by a -1 frameshifting event that

---

<sup>1</sup> Note that it is the viral long terminal repeats (LTRs) that act as promoters for transcription of HIV-1.

occurs due to the presence of a “slippery” sequence in the HIV-1 genome that allows for read-through of a stop codon at the end of the *gag* gene (66).

Following protein production, and during the final stages of viriogenesis, the complement of structural and enzymatic proteins necessary for the formation of a new, intact virus assemble at the surface of the host cell. This event is orchestrated largely by the association of Gag/Gag-Pol with viral genomic RNA dimers via the nucleocapsid portion of the precursor and by localization of Env proteins to the assembly site—a step that, in turn, causes migration of Gag-Pol/RNA constructs to the assembly site as well. Multimerization of Gag, along with interaction of a series of host proteins, leads to a curvature at the assembly site, allowing the new virion to bud and release (18). Protein maturation in the virion, which is thought to begin as early as during the late stages of assembly, is carried out by the viral protease and involves the cleavage of polyproteins into their active forms. The mature virus is then capable of infecting subsequent host cells.

### *1.3 Epidemiological Aspects of Acquired Immunodeficiency Virus (AIDS) and Therapeutic and Preventative Measures Against HIV-1 Infection*

From an epidemiological standpoint, reports estimate that more than 30 million people are currently living with HIV-1 infection worldwide, with the largest fraction of these individuals residing in sub-Saharan Africa (13). In this region, as well as North America and Eurasia, protective measures—such as male circumcision, sexual education and condom usage—have generated both an awareness of and

understanding towards the HIV-1 associated Acquired Immunodeficiency Syndrome (AIDS) and the impact of this disease on society (13, 60). While this is the case, the relatively high levels of recombination and genetic variation common to HIV-1 act as confounding factors in the development of an “ultimate” preventative which could potentially eradicate the disease.

HIV-1 is classified into three groups: M, N and O (main, new and outlier, respectively). Of these groups, M is comprised of a diversity of subtypes and is responsible for nearly 90% of all HIV-1 infections worldwide (44; Figure 4). Furthermore, the prevalence of a number of circulating recombinant forms, which are generated through intersubtype recombination, act to create additional diversity in regions such as Africa, South America, and Asia. Circulating recombinant form CRF01\_AE has, for instant, been shown to constitute over 3% of total HIV infections worldwide (57).



As indicated previously, several anti-retroviral drug regimens and therapeutics have been developed in an effort to curtail the incidence of infection, the most common of which is Highly Active Antiretroviral Therapy (HAART). In light of increased genetic diversity of the virus, and therefore an increased potential for emergence of drug resistant subtypes, patients undergoing HAART are typically prescribed a cocktail of three or more drugs (10). The major classes of antiretrovirals currently in use are illustrated in Table 1 below.

**Table 1.** Major classes of antiretroviral drugs

<b>DRUG CLASS</b>	<b>DESCRIPTION/FUNCTION</b>	<b>EXAMPLE</b>
Nucleoside Reverse Transcriptase Inhibitors	Mimic nucleosides and act as “chain terminators” once incorporated into viral DNA	Azidothymidine (AZT)
Non-nucleoside Reverse Transcriptase Inhibitors	Bind near RT active site, inhibiting polymerization	Nevarapine
Entry Inhibitors	Act to inhibit HIV-1 glycoproteins gp120 and gp41 from interacting with receptors on the host cell surface	Enfuviritide
Protease Inhibitors	Blocks protease, which is responsible for proteolytic cleavage of polyproteins	Ritonavir
Integrase Inhibitors	Prevents integration of viral DNA into host genome	Raltegravir

In addition to drug therapy, several studies have implicated the use of siRNA (small inhibitory RNA) and alternative natural products, such as alkaloids and flavonoids, as potentially viable means of treatment (6, 42; respectively). Vaccine development has also been proposed and tested, though given the extreme diversity of the virus, an effective vaccine has not yet been produced and there is no guarantee that such a vaccine can be developed in the future (3). For this reason, among others, it is imperative that individuals be exposed to and educated about the various preventative and protective measures against HIV-1 infection, as discussed above.

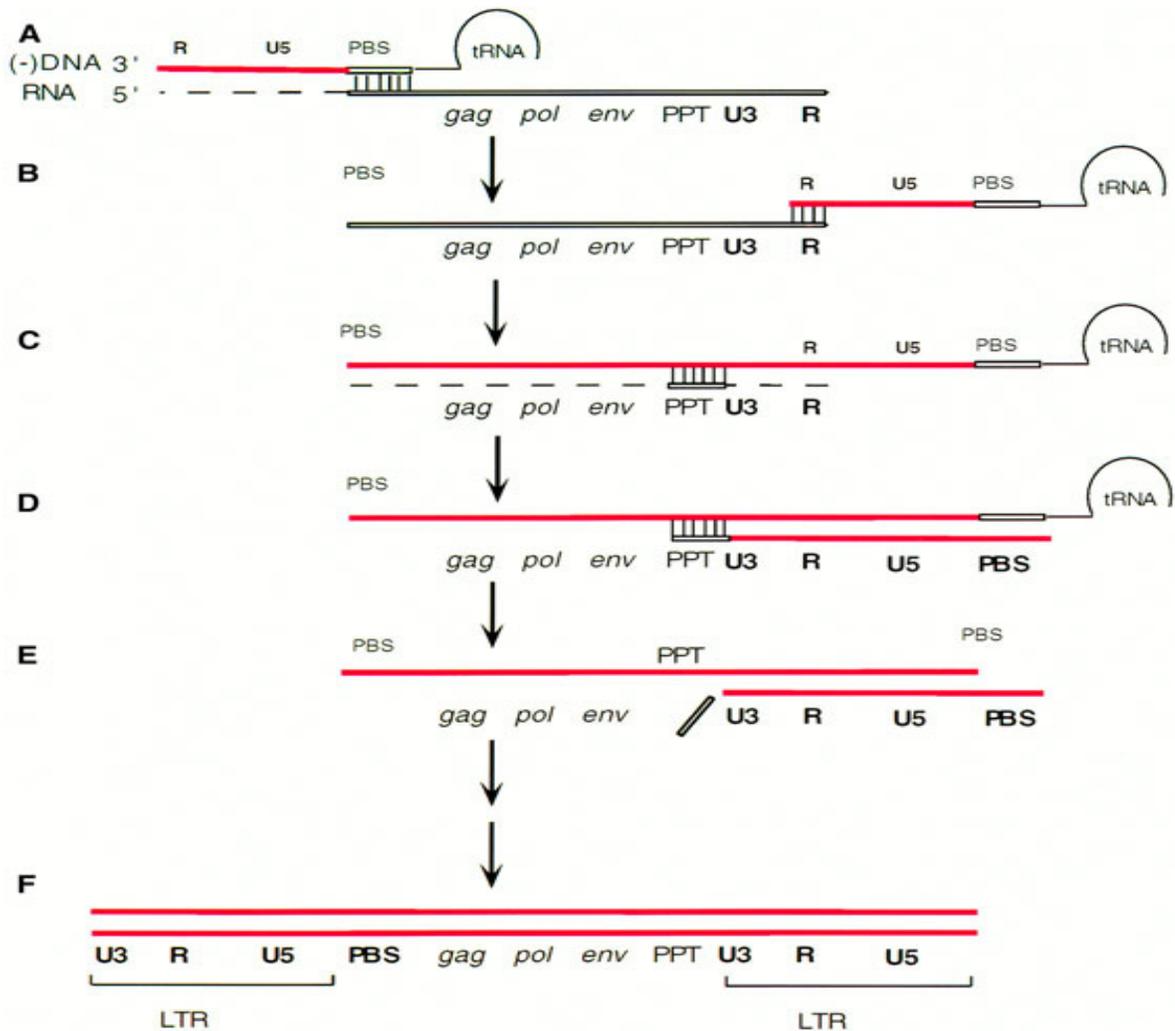
## Chapter 2: Analysis of Those Requirements Necessary and Sufficient for Tighter Binding of HIV-1 Reverse Transcriptase to RNA vs. DNA Substrates

### 2.1 Introduction

#### 2.1.1 The General Mechanism of HIV-1 Replication

Retroviruses, such as HIV-1, are unique among other positive-stranded RNA viruses in that they utilize the process of reverse transcription to convert their genomic RNA into double-stranded, proviral DNA. This process is achieved largely by the multifunctional reverse transcriptase, which possesses both RNA/DNA-dependent, DNA polymerization and RNase H functions (1, 14, 44). While such a description may paint a relatively simplistic view of this mechanism, it is important to note that reverse transcription is, in fact, a complex, multistep process.

In the case of HIV-1, reverse transcription initiates from the 3'OH of a cellular tRNA<sup>Lys,3</sup> that is hybridized to the primer-binding site (PBS) within the viral RNA genome, and extension of the tRNA primer to the 5' terminal repeat (R) occurs via RNA-dependent DNA polymerization (Figure 5; 63). This extension product is known as the minus-strand, strong-stop DNA intermediate (14).



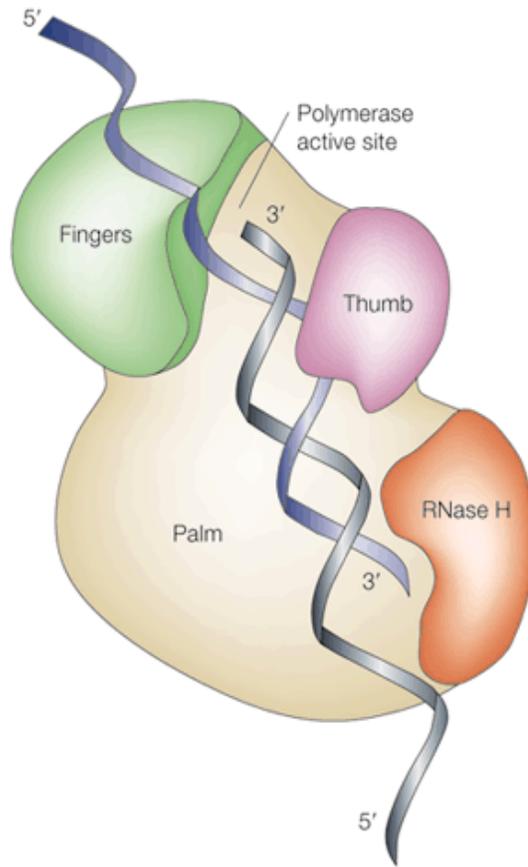
**Figure 5.** Schematic Representation of Reverse Transcription in HIV-1. **(A)** Minus strand synthesis (DNA strand in red) is initiated using a cellular tRNA annealed to the PBS. The RNA strand of the RNA:DNA hybrid is degraded by RNase H of HIV-RT. **(B)** First strand transfer event in which the newly formed DNA anneals to the 3' end of the viral genome. **(C)** Minus strand DNA synthesis resumes and complementary RNA (PPT withstanding) is degraded. **(D)** PPT is used as a primer for second strand DNA synthesis. **(E)** RNase H removes the tRNA and PPT. **(F)** Completion of second strand synthesis results in the production of double stranded DNA.

Following this step, RNase H digestion of the RNA portion (represented as a dotted line in part A of the above figure) of the newly created RNA-DNA hybrid region occurs, allowing for hybridization of both the minus-strand intermediate and viral genomic RNA 3' R regions (as seen in panel B). Once this first strand transfer event occurs, elongation of the nascent DNA and degradation of the complementary RNA proceeds in a similar fashion as previously described, though the hydrolysis-resistant polypurine tract (PPT) remains intact. This is critical, as subsequent second-strand DNA synthesis is dependent upon the PPT as a primer (14, 34). Upon completion of DNA synthesis, RNase H degrades both the tRNA primer and the PPT (see panel E), allowing for the second strand transfer event to occur via the interaction of the two, complementary PBS regions. This interaction signals for the completion of double-stranded DNA (dsDNA) synthesis via DNA-dependent DNA polymerization. The remaining dsDNA, with its long terminal repeat (LTR; see panel F) regions intact, is then ready to be processed by the downstream integrase protein for insertion into the host genome (12, 14).

### 2.1.2 Human Immunodeficiency Virus-1 Reverse Transcriptase (HIV-RT)

In addition to its role in replication, as detailed above, crystallographic data on HIV-RT has provided the first structural views of retroviral reverse transcriptases (RTs). As this data illustrates, the p66 catalytic subunit of the heterodimeric HIV-1 reverse transcriptase, like many other polymerases, resembles a right-handed structure whose subdomains have, appropriately, been identified as the palm, thumb and fingers (see Figure 6; 52). Mutational analysis of heterodimers has indicated that

both the polymerase and RNase H activities of the enzyme reside on the p66 subunit, and the polymerase site on p51 is inactive (31, 44, 61, 65).



Nature Reviews | Molecular Cell Biology

**Figure 6.** Schematic representation of HIV-1 Reverse Transcriptase indicating both the polymerase active site and RNase H domains of the enzyme. “Right-handed” structure of HIV-RT is illustrated by the fingers (green), palm (tan), and thumb (pink) subdomains.

In addition to studying the structure of HIV-RT, extensive research has been conducted on the interaction of HIV-1 reverse transcriptase with various RNA and DNA substrates. Generally, such studies (9, 21, 26, 69) have indicated a higher binding affinity of HIV-RT to RNA vs. DNA templates, possibly due to increased interactions of RT amino acids with 2'-OH groups of RNA nucleotides and a higher degree of steric flexibility in RNA vs. DNA substrates. A more comprehensive review of this literature is presented in section 2.1.3 below.

### 2.1.3 Binding of HIV-RT to DNA and RNA Templates

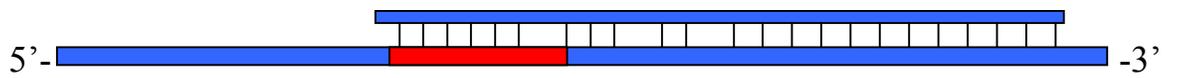
During HIV-1 replication, the role of its reverse transcriptase (HIV-RT) in binding both DNA and RNA substrates is critical, ultimately, for the production of double stranded viral DNA that can be integrated into the host genome. As mentioned previously, this process is largely achieved as a result of the multifunctional nature of the polymerase, specifically its ability to synthesize DNA as both a DNA- and RNA-dependent DNA polymerase, as well as its possession of an RNase H active site, which is critical for degradation of RNA regions of DNA-RNA hybrids (5, 44, 72).

Several studies have provided a more comprehensive examination of the interaction of HIV-RT with DNA and RNA substrates, concluding, among other findings, that the binding stability of the polymerase to RNA-DNA hybrids is much tighter than that of DNA-DNA or RNA-RNA duplexes (33, 37, 20, 21, 69, 81). It has been proposed that this phenomena is due, at least in part, to the numerous additional contacts HIV-RT makes with DNA-RNA primer-template hybrids and, more

specifically, with 2'-hydroxyl groups on the RNA nucleotide itself (3, 63). While this is the case, contacts with 2'-hydroxyl groups account for only a fraction of such interactions, suggesting that alternative explanations for the slower dissociation of RT from RNA-DNA substrates may exist (26, 69). For instance, DNA-DNA and RNA-DNA hybrids are known to take on different structures, with the former generally being B-form and the latter being H-form (hybrid form), an intermediate between A- (the form taken by most RNA-RNA hybrids) and B-form (11). These different structures could also possess a putative role in binding.

To better define the requirements for tighter binding of HIV-RT to RNA-DNA vs. DNA-DNA substrates, previous work by Bohlayer and DeStefano (2006) has demonstrated that tighter binding may be attributed to a short, 5 base pair RNA-DNA hybrid region, where positioning of the polymerase active site of HIV-RT over this region yields optimal binding that is ~20-fold more stable binding than for DNA-DNA (see Figure 7 for schematic of this interaction). This template with only a 5-nt region of RNA bound as tightly as a homogenous RNA template, suggesting that all interactions leading to tighter binding of RNA-DNA reside in this region. In this study, we attempt to expand upon these initial findings by modifying the short, 5 base pair RNA-DNA hybrid region in an effort to determine what conditions are both necessary and sufficient to confer tighter binding of HIV-RT to RNA-DNA vs. DNA-DNA hybrids.

D15R5D30



**Figure 7.** Schematic representation of a DNA-RNA primer:template hybrid. Illustrated is the DNA primer (top, blue rectangle) hybridized to the D15R5D30 template. DNA bases are indicated in blue; RNA bases are indicated in red. Refer to results section(s) for definitions and detailed descriptions of each template.

## 2.2 Materials and Methods

### 2.2.1 Materials

Custom oligonucleotides were from Integrated DNA Technologies. Plasmid clone for wild-type HIV-RT (type HXB2) was a kind gift from Dr Michael Parniak (University of Pittsburgh), and was prepared and purified as described (27). T4 polynucleotide kinase (PNK) was purchased from New England Biolabs. PCR grade dNTPs were purchased from Roche Applied Sciences. G-25 spin columns were from Harvard Apparatus. Radiolabeled nucleotides were from Perkin-Elmer. All additional reagents were purchased from Sigma Aldrich, Fisher Scientific or VWR.

### 2.2.2 Preparation of Radiolabeled DNA Primer Strands

Twenty-five picomoles (pmol) of DNA (31–35 nt in length) was 5'-<sup>32</sup>P-endlabeled using PNK. The labeling reaction was performed at 37°C for 30 min, as per the manufacturer's protocol. Reactions were shifted to 70°C for 15 min in order to heat inactivate the PNK. The DNA was then centrifuged on a Sephadex G-25 column in order to remove any excess radiolabeled nucleotide.

### 2.2.3 Preparation of RNA-DNA and DNA-DNA Hybrids

Hybrids were prepared by mixing 2 pmol of 5'-<sup>32</sup>P-labeled DNA from above and 2 pmol of 50-nt template (see 'Results' sections for full list of template sequences) in 15 µl of buffer containing 50 mM Tris-HCl, pH=8, 80 mM KCl, 1 mM dithiothreitol (DTT), and 0.1 mM EDTA pH=8. Reactions were placed at 80°C for 5

min and then allowed to slowly cool to room temperature prior to use.

#### 2.2.4 Determination of Dissociation Rate Constants by Primer Extension

HIV-RT (20 nM) was preincubated with hybrids (20 nM) in 42  $\mu$ l of buffer containing 50 mM Tris-HCl, pH=8, 80 mM KCl, 2 mM MgCl<sub>2</sub> and 1 mM DTT for 3 min at room temperature. After preincubation, 8  $\mu$ l of trap supplement in the same buffer containing heparin sulfate (final concentration 2  $\mu$ g/ $\mu$ l) was added to each reaction. The purpose of this supplement was to bind and sequester enzyme that had dissociated from substrates. Following trap addition, 5  $\mu$ l aliquots were removed at 15s (2s, 5s and 10s, time points were also taken of substrate S2P33-D50), 30s, 1m, 1.5m, 2m, 2.5m, 3m and 4m or 15s, 1m, 2m, 4m, 8m and 16m for DNA- or RNA-containing templates, respectively, and added to a tube containing 1 ml of dNTP supplement in reaction buffer plus dNTPs (100  $\mu$ M final concentration). Tubes were incubated at room temperature for 2 min., then terminated by the addition of 6  $\mu$ l of 2X formamide gel loading buffer (90% formamide, 20 mM EDTA, pH=8.0, 0.25% xylene cyanol and bromophenol blue).

A trap control in which RT was mixed with the heparin trap and dNTPs prior to substrate addition and a full extension control in which trap was excluded were also performed and incubated for 10 min. prior to 2X formamide gel loading buffer addition. Since the rate of dissociation for enzyme-nucleic acid substrate complexes is slow in comparison to the rate of polymerization (45,59), the amount of HIV-RT bound at the initiation of each reaction can be analyzed by measuring the amount of primer extension. A t=0 sample was performed by mixing 1  $\mu$ l of heparin trap

supplement and 1  $\mu$ l of dNTP supplement together and adding this to 4  $\mu$ l of the reaction + RT preincubation mix (see above). After 2 min., 6  $\mu$ l of 2X gel loading buffer was added to the sample. All samples were subsequently loaded onto a 12% polyacrylamide/7M urea sequencing gel and subjected to electrophoresis as previously described (62).

#### 2.2.5 Determination of Equilibrium Dissociation Constants ( $K_d$ ) by Primer Extension

HIV-RT (0, 0.5, 1, 2, 3, 4, 6 and 8 nM) was preincubated with hybrids (2 nM) in 8  $\mu$ l of buffer containing 50 mM Tris-HCl, pH=8, 80 mM KCl, 2 mM MgCl<sub>2</sub> and 1 mM DTT for 3 minutes at room temperature. After preincubation, 2  $\mu$ l of trap supplement in the same buffer containing dNTPs (final concentration 100  $\mu$ M) and heparin sulfate (final concentration 2  $\mu$ g/ $\mu$ l) was added to each reaction. The purpose of this supplement was to bind and sequester enzyme that had dissociated from substrate. Tubes were incubated at room temperature for 2 min, then terminated by the addition of 10 ml of 2X formamide gel loading buffer. A trap control and full extension control were performed as described above for the  $k_{off}$  determinations. Samples were processed as described for the  $k_{off}$  determinations.

#### 2.2.6 Visualization and Quantification of Primer Extension

In order to visualize and quantify primer extension, dried polyacrylamide gels were processed on an FLA-5100 or FLA-7000 phosphoimager from Fujifilm Life Sciences. Dissociation rate constants ( $k_{off}$  values) were obtained by fitting data to a

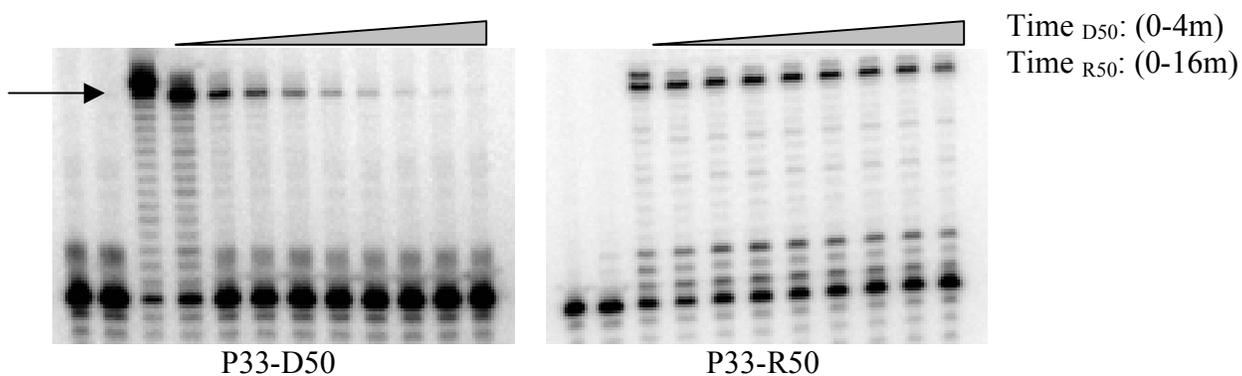
nonlinear least-squares equation for single, exponential decay [ $f(x) = ae^{-bx}$ , where  $a$  is the  $y$ -intercept at time zero and  $b$  is the dissociation rate] using Sigma Plot (Jandel Corp.). Data collected from time,  $t=0$  was not included in this calculation, as aberrant decreases in primer extension were routinely observed between this and the next measured time point for some substrates (8). Calculation of equilibrium dissociation constants ( $K_d$ ) was determined from graphs of the concentration of extended primer versus concentration of RT 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 hybrid concentration (40). Experiments were generally repeated two to four times and averages  $\pm$  standard deviations are reported in Table 1.

Note that this approach actually yields an ‘apparent  $K_d$ ’ value, as it is dependent on a secondary measurement (polymerase extension) to assess binding. This would be a concern if there were secondary binding sites on the substrates that could strongly compete with the 3’ recessed terminus for RT, or if a substantial proportion of RT interactions with the 3’ terminus were nonproductive with RT dissociating before incorporating nucleotides. Each of these concerns is of very low probability for these substrates.

## 2.3 Results

### 2.3.1 RNA-DNA and DNA-DNA Substrates Used to Test HIV-RT Binding

In previous work, we showed that RT dissociated ~20 times more slowly from a 33-nt DNA hybridized to a 50-nt RNA template (P33-R50, see Figure 8) as compared to the same DNA hybridized to a homologous DNA template (P33-D50). Progressively reducing the amount of RNA in the template strand showed that only 5 nt of RNA were required for strong binding, as a chimeric template composed of (5'–3') 15 bases of DNA followed by 5 bases of RNA then 30 of DNA (P33-D15R5D30) bound to RT with even greater stability than P33-R50 (8). In addition, positioning of the 3' terminus of the shorter DNA (P33) was shown to be pivotal to tight binding, as substrates in which the 5-nt RNA–DNA hybrid region were not positioned over the recessed 3' terminus of the DNA dissociated essentially like a homologous DNA template. We therefore concluded that positioning of the 3' terminus directly over the short, 5-bp region of RNA–DNA was both necessary and sufficient for mimicking the more stable interaction of RT with RNA templates.



**Figure 8.** Autoradiograms of half-life binding experiments with primer 33 bound to 50-mer DNA or RNA templates. Full extension products are indicated by an arrow and are quantified in order to determine  $k_{\text{off}}$  rates (see table 1). All experiments are completed using wild-type HIV-RT. Lanes 1 and 2 illustrate a ‘no enzyme’ and ‘trap’ control, respectively, which are designed to represent a negative control and a control to test the effectiveness of the heparin or poly(rA)-oligo(dT) trap supplement used in each assay (see section D.2.3). The latter control is prepared by first mixing enzyme, trap, and dNTPs and then introducing substrate. Samples are incubated for a time equal to the longest time point on the gel. Lane 3 is a full extension control in which enzyme was incubated with each substrate as in the trap control reaction except the trap was omitted to allow all of the bound primer to be extended. Time points for 4-minute time course experiments are: 0, 15s, 30s, 1m, 1.5m, 2m, 3m, 4m. Time points for 16-minute time course experiments are: 0, 30s, 1m, 2m, 4m, 8m, 16m.

Note that the previous experiments were conducted using the RNase H minus HIV-RT mutant, E478>Q, in the presence of  $Mg^{2+}$ , or wild-type HIV-RT in the absence of  $Mg^{2+}$ . This was required since it is not possible to accurately measure off-rates on an RNA–DNA hybrid with wild-type RT in the presence of  $Mg^{2+}$  due to RNase H activity. Based on its nearly equivalent binding to DNA–DNA substrates in the presence of 4-6 mM  $Mg^{2+}$  and DNA synthetic properties essentially identical to wild-type RT, E478>Q was considered a good model for the wild-type enzyme (16). Although RT substrate interactions in the absence of  $Mg^{2+}$  are considerably less stable than in its presence, the same trend was observed using wild-type HIV-RT in the absence of  $Mg^{2+}$ , with tight binding equivalent to a complete RNA template requiring just the properly positioned 5-bp hybrid region (8).

To examine this 5-bp region more closely, we proposed several changes. These included shortening the region and converting each base to DNA, as well as using 2'-O-methylated nucleotides to replace some or all of the bases (figure 9).

**A**

```

          3'-GATATCCCGCTTAAGCTCGAGCCATGGGCCCCCT-5' DNA Primer
5' D50-TTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATC-3'
5' D15R5D30-TTGTAATACGACTCACUAUAGGGCGAATTCGAGCTCGGTACCCGGGGATC-3'
5' D18R1 (-4) D31-TTGTAATACGACTCACTAUAGGGCGAATTCGAGCTCGGTACCCGGGGATC-3'

```

**B**

```

          3'-GTCTGCACACACGGGCAGACAACACACTGAGAC-5' DNA Primer
5' S2D50-TGCCTTGAGTGCTTCCAGACGTGTGTGCCCGTCTGTTGTGTGACTCTGGC-3'
5' S2D15R5D30-TGCCTTGAGTGCTTCCAGACGTGTGTGCCCGTCTGTTGTGTGACTCTGGC-3'
5' S2D18R1 (-4) D31-TGCCTTGAGTGCTTCCAGACGTGTGTGCCCGTCTGTTGTGTGACTCTGGC-3'

```

**Figure 9.** Sequence and configuration of nucleic acid substrates. **(A)** Representative sequences of the short DNA (top stand, denoted ‘P33’ for ‘Primer’ 33 nt) and long (DNA, or RNA–DNA chimera) strands are shown. Five different short DNA strands that shared a common 5′-end were used. Their lengths were 31, 32, 33, 34 and 35. Only the 33-nt strand that was used for most substrates is shown. In order to illustrate the nomenclature used in the text, the shorter DNA is placed over the complementary bases of three of the several different template strands that were used. The templates were all 50 nucleotides and consisted of either homogeneous DNA or RNA, or chimeras with both DNA and RNA. RNA nucleotides are underlined on the templates. Duplex substrates were named based on the short DNA and template used with the following nomenclature as an example: P33-D18R1(–4)D31. This substrate had the 33 nt DNA hybridized to a template strand where the first 18 5′ nt were DNA, followed by a single RNA nucleotide, then 31 DNA nucleotides. The -4 in parentheses indicates the position of the single RNA nucleotide relative to the 3′ terminus of P33 with the template nucleotide hybridized to the 3′ terminal base being designated as -1. This basic nomenclature was used for all substrates described in the text. **(B)** A second duplex with a different sequence that was used in experiments is shown. These substrates are designated ‘S2’ (sequence 2) in the text. Nomenclature is as stated above.

Given that these substrates are not cleaved by RT's RNase H activity, most experiments were conducted in the presence of 2 mM Mg<sup>2+</sup> and 80 mM KCl (as discussed above), using wild-type enzyme to more closely approximate physiological conditions. Results from these experiments are presented below.

2.3.2 RT Dissociates from Templates with Just Two RNA Nucleotides at a Rate Similar to a Complete RNA Template and Even a Single RNA Nucleotide at the -4 Position Results in a Dissociation Rate Much Slower than a Complete DNA Template

To determine if all five bases of the RNA–DNA hybrid region noted above were required for slow dissociation of RT, we generated a series of chimeric substrates containing one to five of the RNA nucleotides. These templates (D15R5D30; also D15R4D31, D15R3D32, D15R2D33 and D15R1D34) were hybridized to the P33 DNA and the  $k_{\text{off}}$  of RT from each duplex was measured. DNA P33 is designed to position the five-base RNA region of D15R5D30 in the polymerase domain of RT with the 5' most RNA base bound to the 3' terminal base of P33. Reverse transcriptase dissociated from the P33-D15R5D30 substrate ~38 times more slowly than to a homologous DNA template (P33-D50). Replacement of the fifth nucleotide (-5 position relative to the recessed 3' terminus of P33) with a DNA base resulted in a small but insignificant increase in the off-rate under the conditions used (Table 2, compare P33-D15R5D30 to P33-D15R4D31).

**Table 2.** Dissociation rate constants ( $k_{\text{off}}$ ) and equilibrium dissociation constants ( $K_d$ ) for RT-substrate complexes

<sup>a</sup> Duplex substrate	<sup>b</sup> Dissociation rate ( $k_{\text{off}}$ ) ( $\text{min}^{-1}$ )	<sup>c</sup> Fold decrease in $k_{\text{off}}$	$K_d$ (nM)
<b>Binding of RT to DNA–DNA duplex</b>			
P33-D50	0.90 ± 0.12	1	2.7 ± 1.3
<b>Effect of shortening the 5-bp RNA–DNA hybrid region in chimeric templates</b>			
P33-D15R5D30	0.024 ± 0.006	38	1.4 ± 0.6
P33-D15R4D31	0.030 ± 0.006	30	
P33-D15R3D32	0.13 ± 0.04	6.9	
P33-D15R2D33	0.16 ± 0.04	5.6	
P33-D15R1D34*	0.41 ± 0.07	2.2	2.6 ± 1.3
<b>Binding of RT to duplex with one or two RNA nucleotides in the template</b>			
P33-D15R1(-1)D34*	0.41 ± 0.07	2.2	2.6 ± 1.3
P33-D16R1(-2)D33	0.11 ± 0.02	8.2	1.5 ± 0.3
P-33-D17R1(-3)D32	0.25 ± 0.05	3.6	
P33-D18R1(-4)D31	0.066 ± 0.012	14	
P33-D16R1(-2)D1R1(-4)D31**	≤0.024	≥38	
<b>Effect of changing the nucleotide composition at the pivotal –4 template position</b>			
P33-D18R1(-4U>G)D31	0.19 ± 0.05	4.7	
P33-D18R1(-4U>C)D31	0.11 ± 0.02	8.2	
P33-D18r1(-4U>A)D31	0.18 ± 0.04	5	
<b>Effect of changing the composition at the –2 position</b>			
P33-D16R1(-2U>A)D33	0.090 ± 0.006	10	
P33-D16R1(-2U>C)D33	0.11 ± 0.01	8.2	
<b>Effect of repositioning the pivotal –4 template nucleotide</b>			
P35-D18R1(-4>-6)D31	0.39 ± 0.06	2.3	
P34-D18R1(-4>-5)D31	0.11 ± 0.01	8.2	
P32-D18R1(-4>-3)D31	0.66 ± 0.06	1.4	
P31-D18R1(-4>-2)D31	0.34 ± 0.09	2.6	
<b>Effect of replacing 2'-hydroxyls with 2'-O-methyls</b>			
P33-D18methylR1(-4)D31	0.042 ± 0.006	21	
P33-D15methylR5D30	0.030 ± 0.006	30	1.6 ± 1.0
<b>Analysis of binding to a duplex with an unrelated nucleotide sequence</b>			
S2P33-D50	9.3 ± 3.2	1	
S2P33-D15R5D30	0.10 ± 0.03	93	
S2P33-D18R1(-4)D31	0.54 ± 0.02	17	
S2P33-D18R1(-4A>U)D31	0.25 ± 0.01	37	
S2P33-D16R1(-2)D1R1(-4)D31	0.12 ± 0.01	78	
<b>Binding of AMV- and MuLV-RT to duplex substrates</b>			
AMV, P33-D50	0.10 ± 0.02	1	2.8 ± 1.0
AMV, P33-D15R5D30**	≤0.024	≥4.2	2.1 ± 0.8
MuLV-RT, P33-D50	0.16 ± 0.02	1	
MuLV-RT, P33-D15R5D30	0.078 ± 0.024	2.1	
<b>***Binding to 50 nucleotide DNA or RNA templates in the absence of Mg<sup>2+</sup></b>			
HIV-RT, P33-D50	1.3 ± 0.1	1	
HIV-RT, P33-R50	0.048 ± 0.012	27	
AMV, P33-D50	0.96 ± 0.06	1	
AMV, P33-R50	0.43 ± 0.18	2.2	
MuLV-RT, P33-D50	0.33 ± 0.02	1	
MuLV-RT, P33-R50	0.14 ± 0.05	2.4	

<sup>a</sup>For an illustration of the duplex sequences and configurations see Figure 1.

<sup>b</sup> $k_{\text{off}}$  or  $K_d$  for HIV-RT on the specific duplexes; see 'Materials and Methods' section for an explanation of how this was determined; results are an average of two to four independent experiments ± standard deviation.

<sup>c</sup>All numbers are relative to the specific DNA–DNA duplex of homologous sequence which was set equal to 1; larger numbers indicate slower dissociation of RT and tighter binding.

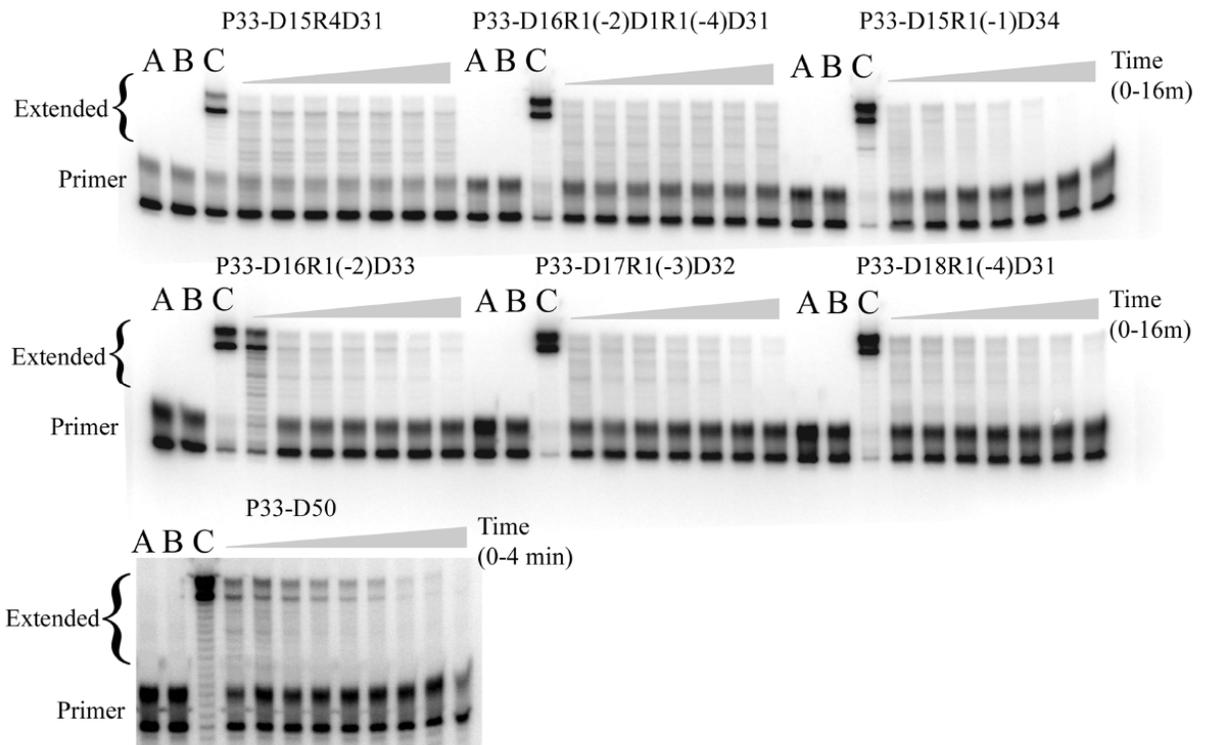
\*These two substrates are identical

\*\* $k_{\text{off}}$  was too low to be reliably determined under the conditions used; numbers provided indicate that it was at least as low as the lowest measured off-rate in the assays.

\*\*\*Assays were performed without Mg<sup>2+</sup> in dissociation phase and with 20 rather than 80 mM KCl in order to help stabilize binding, which is weaker in the absence of Mg<sup>2+</sup>.

In contrast, the  $k_{\text{off}}$  value increased ~4-fold when the -4 and -5 bases were replaced by DNA (P33-D15R3D32). A further, small but insignificant increase was observed when the -3, -4, and -5 nt were replaced (D15R2D33), while a substrate in which only the -1 nt was RNA (P33-D15R1D34) exhibited another 2–3-fold increase in  $k_{\text{off}}$ . From these results, it was clear that only a 4-bp RNA–DNA region was required for strong binding and that even a single RNA–DNA base pair at the -1 position conferred slower dissociation (~2–3-fold) as compared to a complete DNA–DNA duplex (P33-D50).

To further test the role of each of the four RNA bases at positions -1, -2, -3 and -4 in binding, chimeric templates containing a single RNA nucleotide at each of these positions were constructed. All were hybridized to the P33 DNA and  $k_{\text{off}}$  was measured for each. RT dissociated from the substrate with a single RNA base at position -3 (P33-D17R1(-3)D32) at approximately the same rate as the P33-D15R1(-1)D34 template, suggesting that both the -1 and -3 position nucleotides confer a modest increase in binding stability over DNA. In contrast, RT dissociated from the P33-D16R1(-2)D33 substrate ~3 times more slowly than the -1 and -3 substrates. Finally, RT dissociated from the P33-D18R1(-4)D31 substrate ~2 times more slowly than the -2 substrate and approximately twice as fast as the substrate containing all four RNA bases (P33-D15R4D31) (Figure 10 and Table 2).



**Figure 10.** Panels for dissociation rate constants ( $k_{\text{off}}$ ). Representative assays are shown to illustrate how dissociation rate constants were determined. Primer labeled with  $^{32}\text{P}$  at the 5' end was used in the assays. The level of primer extension over time was quantified using a phosphoimager as described in 'Materials and Methods' section and these values were plotted to determine  $k_{\text{off}}$ . Time points used for the P33-D50 assay were 0, 15 s, 30 s, 1 m, 1.5 m, 2 m, 2.5 m, 3m and 4 m, while for all other assays time points were 0, 30 s, 1 m, 2 m, 4 m, 8m and 16 m. Lane A in each panel shows a reaction in the absence of enzyme. Lane B shows a control reaction to test the effectiveness of the heparin trap (see 'Materials and Methods' section). In this reaction, the enzyme was mixed with the trap and the mixture was added to the substrate in the presence of dNTPs and divalent cation and incubated for 10 min. before termination. Lane C shows a full extension control in which enzyme was incubated with the substrate as in the trap control reaction except trap was omitted to allow all the bound primer to be extended.

Since the -2 and -4 positions appeared to be most important to the slow dissociation of RT from RNA–DNA, a template containing both RNA nucleotides (D16R1(–2)D1R1(–4)D31) was constructed and tested. Dissociation of RT from a substrate with this template (P33-D16R1(–2)D1R1(–4)D31) was very slow, such that the off-rate was difficult to measure under the conditions used (data not shown). Taken together, these results indicate that a single RNA nucleotide at the -4 position results in RT dissociating ~14 times more slowly than from a complete DNA strand (D50), while RT binds to templates with both the -2 and -4 RNA nucleotides even more stably than a complete RNA strand [based on D15R5D30 binding even more stably to RT than a complete RNA strand (see above)].

### 2.3.3 A Uridine Residue at the -4 Position Promotes More Stable HIV-RT Binding than Other RNA Nucleotides

It was interesting that, in addition to the importance of the -4 RNA nucleotide (P33-D18R1(–4)D31), the -2 position (P33-D16R1(–2)D33) also had a significant effect on RT dissociation. Further, both of these bases were uridines, suggesting that there may be some base preference for promoting stable RT binding. To test this theory, the U at the -4 position was changed to A, G and C with the corresponding base in the P33 DNA also changed to maintain complementarity. Dissociation rates from these substrates were measured, and RT was noted to dissociate more rapidly from all of them (Table 2).

The  $k_{\text{off}}$  was increased ~3-fold when the U was replaced by an A (P33-D18R1(-4U>A)D31) or a G (P33-D18R1(-4U>G)D31) and ~2-fold for a C (P33-

D18R1(-4U>C)D31) replacement. In contrast, changing the -2 position from a U to either an A (P33-D16R1(-2U>A)D33) or a C (P33-D16R1(-2U>C)D33) did not significantly change the  $k_{\text{off}}$  (data not shown). These results indicate that binding to RT is most stable if the single RNA nucleotide in the template is a uridine at the -4 position, though it is also important to note that other RNA residues at this position lead to  $k_{\text{off}}$  values that were several fold lower than a homogenous DNA.

2.3.4 Moving the -4 Uridine Residue to Different Positions Relative to the DNA 3' Recessed Terminus Increased HIV-RT's Dissociation Rate, Confirming the Importance of the -4 Position in the RNA

The above experiments do not rule out the possibility that the specific sequence context of the -4 nt in this particular substrate, rather than the position relative to the 3' recessed terminus, is the major determinant for slow dissociation. To test this further, the position of the -4 uridine relative to the DNA 3' terminus was shifted using DNAs of various lengths. Oligonucleotides P31, P32, P34 and P35 place the single RNA base in D18R1(-4)D31 at the -2, -3, -5 and -6 position, respectively, relative to the 3' terminus. These substrates all showed reduced binding to RT as compared to P33-D18R1(-4)D31. Placing the RNA base in the -5 position (P34-D18R1(-4>-5)D31) was least detrimental, resulting in a ~2-fold increase in  $k_{\text{off}}$ , while the -2 (P31-D18R1(-4>-2)D31), -3 (P32-D18R1(-4>-3)D31) and -6 (P35-D18R1(-4>-6)D31) positions increased  $k_{\text{off}}$  ~5-, 10- and 6-fold, respectively. The above results illustrate the importance of an RNA nucleotide at the -4 position for stable binding of RT. In addition, they show that moving this position further away

from the recessed 3' DNA terminus, as in P34-D18R1(-4>-5)D31, is less unfavorable than moving it nearer.

### 2.3.5 The 2'-Hydroxyl Groups on the RNA Nucleotides are not Required to Stabilize Binding of HIV-RT

As was noted in the introduction, several interactions between 2'-hydroxyl groups and RT amino acids have been proposed based on the crystal structure of HIV-RT bound to RNA-DNA (63). Many of these are hydrogen bonds involving the hydroxyl at the 2' position of the RNA nucleotide. To determine if the 2'-hydroxyl was pivotal for stabilizing RT binding on RNA-DNA, modified versions of D15R5D30 and D18R1(-4)D31 in which the RNA nucleotide hydroxyls were converted to 2'-*O*-methyls were tested (P33-D15methylR5D30 and P33-D18methylR1(-4)D31). In both cases, RT  $k_{\text{off}}$  values were similar compared to the templates with 2'-hydroxyl RNA bases (Table 2), indicating that the 2'-hydroxyls do not play a role in stabilizing binding.

### 2.3.6 Slower Dissociation of HIV-RT from RNA-DNA is not Dependent upon Sequence

Both the previous work (8) and the work described above were completed utilizing a single template sequence. It was possible, therefore, that some of the findings could be unique to the particular sequence used. To address this concern, a substrate with a different sequence was constructed and analyzed [Figure 9; sequence 2 (S2)]. Although the sequence design was essentially arbitrary, uridine was

intentionally excluded from the -2 and -4 positions. By subsequently converting the -4 base to uridine, we could then test whether the conclusions from the previous substrate, which indicated a strong preference for uridine at the -4 position in stable binding to RT, were upheld under these new conditions. Overall, results indicated that RT dissociated much faster from this new sequence (Table 2).

For instance, a ~10-fold increase in the  $k_{\text{off}}$  value was observed for the complete DNA template compared to the previous sequence (i.e. S2P33-D50 vs. P33-D50). Additionally, the five RNA nucleotide version of this template (S2P33-D15R5D30), while dissociating much more slowly from RT than the complete DNA-DNA duplex (93-fold), still showed a ~5-fold higher  $k_{\text{off}}$  than the previous sequence. A substrate with only one RNA nucleotide at the -4 position (S2P33-D18R1(-4)D31) also showed an ~4-fold increase in  $k_{\text{off}}$  compared to S2P33-D15R5D30, though it still bound much more stably than a complete DNA-DNA duplex (~17-fold).

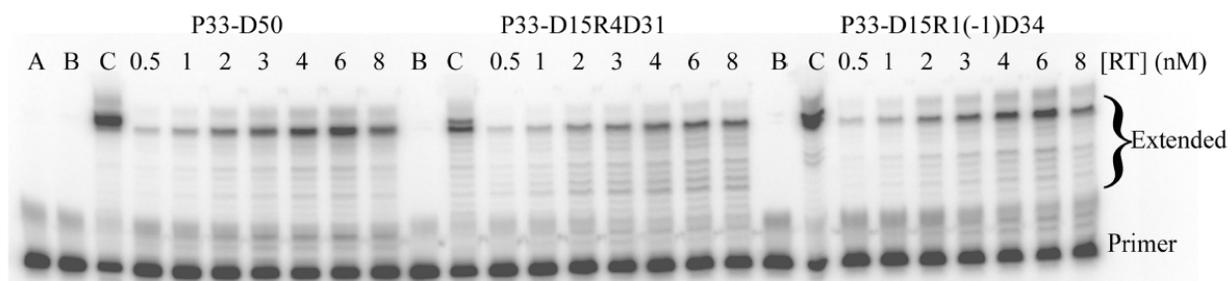
Consistent with a preference for U at the ~4 position, converting the -4 position from an adenine to a uridine and the corresponding base in the 33-mer DNA to an A resulted in ~2-fold decrease in  $k_{\text{off}}$  (S2P33-D18R1(-4A>U)D31 compared to S2P33-D18R1(-4)D31). Including RNA bases at both the -2 and the -4 positions resulted in a  $k_{\text{off}}$  value that was essentially the same as the substrate with 5 nt of RNA (compare S2P33-D15R5D30 and S2P33-D16R1(-2)D1R1(-4)D31), just as it did for the first sequence tested. Overall, results with the second template support a role for the -2 and -4 nucleotides, as well as the preference for U over other bases at the -4 position, in promoting stable binding of HIV-RT to RNA-DNA.

While this is the case, they also highlight distinct sequence-dependent

differences. This was the case in our previous work as well, where various primers were used to position RT at different locations along a 50-nt template identical to the first template used in this work. Off-rates differed several-fold and were dependent on the position of the primer relative to the template termini as well as the particular sequence to which the primer was bound. Specifically, the difference in off-rates for RNA–DNA versus DNA–DNA duplexes varied from as little as 7-fold to as much as 90-fold more stable binding to the former as compared to the latter (8).

### 2.3.7 The Equilibrium Dissociation Constant ( $K_d$ ) for HIV-RT is Nearly the Same for RNA-DNA and DNA-DNA Despite Large Differences in $k_{off}$

Since the equilibrium dissociation constant is dependent on the on- and off-rate ( $K_d = k_{off}/k_{on}$ ), the lower  $k_{off}$  values for the RNA–DNA heteroduplex compared to DNA–DNA homoduplex would imply a higher affinity for the former. This would be reflected in a lower  $K_d$  for the heteroduplexes providing that no change in the on-rate had occurred. Assays were performed to measure the  $K_d$  of RT for P33-D50, P33-D15R5D30, P33-D15methylR5D30, P33-D16R1(–2)D33 and P33-D15R1(–1)D34. Surprisingly, RT had similar  $K_d$  values for all five substrates (Table 2 and Figure 11), even though dissociation was 38 and 30 times faster from P33-D50 than P33-D15R5D30 and P33-D15methylR5D30, respectively, and 8.2 and 2.2 times faster from P33D16R1(–2)D33 and P33-D15R1(–1)D34, respectively.



**Figure 11.** Representative assays are shown to illustrate how dissociation equilibrium constants were determined. Panels are labeled at the top with the primer-template that was used in the assay. The concentration of RT (nM) is noted above each lane. Lane A in each panel shows a reaction in the absence of enzyme. Lane B shows a control reaction to test the effectiveness of the heparin trap (see ‘Materials and Methods’ section). In this reaction, the enzyme was mixed with the trap and the mixture was added to the substrate in the presence of dNTPs and divalent cation and incubated for 10 min. before termination. Lane C shows a full extension control in which enzyme was incubated with the substrate as in the trap control reaction except trap was omitted to allow all the bound primer to be extended.

## Chapter 3: Analysis of the Preferential Binding of MuLV-RT and AMV-RT to RNA vs. DNA Substrates

### 3.1 Introduction

#### 3.1.1 Moloney Murine Leukemia Virus Reverse Transcriptase (MuLV-RT)

A member of the gammaretroviral genus, moloney murine leukemia virus (MuLV) is structurally similar to HIV-1 and, as its name suggests, is responsible for the induction of non-B, non-T cell leukemia in susceptible mice (49). Much like its HIV-1 counterpart, MuLV-RT consists of finger, palm, and thumb subdomains localized centrally to the polymerase active site, as well as an RNase H domain. While this is the case, the enzyme itself is a structurally distinct monomer comprised of approximately 671 amino acids (14). Studies, such as the ones conducted by Coté and Roth (2008) and Fuentes *et al.* (1996), have explored the effects of these structural variations on substrate binding, noting that such variations may be necessary in order to accommodate the full array of activities carried out by MuLV-RT and that they do not significantly affect RT-substrate interactions and extension. Fuentes and colleagues also demonstrated an increased preference of MuLV-RT for RNA rather than for DNA oligonucleotides bound to DNA templates. This result is interesting in light of previous evidence that implicated HIV-RT as having a higher affinity for RNA-DNA than for DNA-DNA (19).

### 3.1.2 Avian Myeloblastosis Virus Reverse Transcriptase (AMV-RT)

Avian myeloblastosis virus (AMV), a major etiological agent in the development of acute myeloblastic leukemia and other, related carcinomas, is unique, among the other viruses discussed here, in that it is defective in the absence of myeloblastosis-associated helper virus (MAV) (64). While this is the case, replication in this virus proceeds in similar fashion to that of other retroviral species, ultimately resulting in the production of proviral double-stranded DNA that can subsequently be integrated into the host genome. Though concerns regarding the purity of AMV-RT stocks continue to proliferate [given the necessity of a helper virus], this feature and the relative ease by which one can generate AMV-RT allows AMV to remain as one of the major sources of avian reverse transcriptase available (39, 56).

Upon closer examination of AMV-RT, it has been shown that, much like HIV-RT, this enzyme is a heterodimer comprised, in this case, of both an  $\alpha$  (65 kDa) and  $\beta$  (95 kDa) subunit that share a high degree of structural similarity. Interestingly, it has also been shown that the  $\alpha$  subunit possesses both the RNA-dependent DNA polymerase and RNase H activities common to retroviral reverse transcriptases, while the  $\beta$  subunit remains inactive. This arrangement does not seem to have any substantial effect on enzymatic activity (4, 56). Further studies have corroborated this claim, suggesting that the processivity [the average number of nucleotides added by the polymerase per association/dissociation with the template] of AMV-RT is higher than that of other reverse transcriptases, such as those derived from mice, and that, while incorporating a relatively large number of incorrectly paired bases while

carrying out the replication process, DNA synthesis on both RNA and DNA substrates occurs in an efficient manner (4, 35, 81).

### 3.2 Materials and Methods

Experimental and visualization procedures were as noted above in section 2.2, with the exception that  $k_{\text{off}}$  and  $K_d$  reactions with AMV- or MuLV-RT on both DNA–DNA and RNA–DNA duplexes were run over a 16-min. time course and the trap was poly(rA)-oligo(dT<sub>20</sub>) [8:1 rA:dT (w:w), final concentration 0.4 µg/µl] instead of heparin for AMV-RT. Additionally, magnesium was omitted from the preincubations in some experiments, which were conducted with 20 rather than 80 mM KCl. With respect to the  $K_d$  assay, the analyzed enzyme concentrations for AMV-RT were, in lieu of those presented in section 2.2, as follows: 1, 2, 4, 6, 8, 10 and 12 nM.

### 3.3 AMV-RT and MuLV-RT Show More Modest Preference for Binding to RNA-DNA versus DNA-DNA

In order to determine if the behavior of HIV-RT is typical for other reverse transcriptases, the binding of AMV- and MuLV-RTs on various substrates was also examined. With regard to P33-D50, both AMV-RT and MuLV-RT bound considerably more stably than HIV-RT (~9- and 6-fold, respectively, Table 2). Like HIV-RT, AMV-RT bound even more stably to P33-D15R5D30 [none of the enzyme used here showed significant RNase H-directed cleavage of the substrate (data not shown)]. Binding was so tight that no consistently measurable dissociation was detected under the conditions used. In contrast, MuLV-RT bound this substrate only ~2-fold more stably than the DNA–DNA. Also consistent with HIV-RT,  $K_d$  values for AMV-RT on P33-D50 and P33-D15R5D30 were similar despite the large  $k_{off}$  difference.

It is important to note that the D15R5D30 template was optimized for binding to HIV-RT using the HIV-RT E478>Q mutated RT (8). It is possible, therefore, that the 5-nt RNA region may not be ideal for the other enzymes tested here. To further evaluate duplex binding properties, each enzyme was additionally examined on P33-D50 and P33-R50. These experiments were carried out in the absence of  $Mg^{2+}$  in the dissociation phase in order to prevent cleavage of P33-R50. The KCl concentration was lowered to 20 mM to compensate for the lower binding stability observed in the absence of  $Mg^{2+}$  (20). Under these conditions, all the enzymes bound more stably to the RNA–DNA duplex, however, HIV-RT bound 27-fold more stably, while the other enzyme showed only ~2-fold better binding to RNA–DNA.

## Chapter 4: Discussion

### 4.1 Slower Dissociation of HIV-RT from RNA-DNA versus DNA-DNA does not require 2'-Hydroxyls and is Dependent on the RNA Nucleotides at the -2 and -4 Positions

Here, we show that HIV-RT's several fold slower dissociation from binary complexes with RNA–DNA versus DNA–DNA requires just two RNA nucleotides (–2 and –4 positions) of the RNA strand and does not require 2'-hydroxyl groups. Similar  $K_d$  values for all duplexes tested (see appropriate results sections) suggested that changes in  $k_{on}$  that were approximately proportional to the changes in  $k_{off}$  for the various substrates were occurring (see below). The differential stability of the binary complexes could be especially relevant for RT binding to RNA fragments for degradation (see below).

Replacement of 2'-hydroxyl groups on pivotal RNA nucleotides with 2'-*O*-methyl groups had no effect on RT's  $k_{off}$  or  $K_d$  for RNA-DNA hybrids (see section 2.3.5). Structural data suggest that the important -2 nucleotide interacts with glutamic acid 89 (E89) of RT by a hydrogen bond to the 2'-hydroxyl group, while an interaction between the pivotal -4 base and isoleucine 94 (I94) is also proposed (63). While the 2'-*O*-methyl modification would have abrogated the E89 interaction to -2, it is unclear if the I94 interaction would have been affected. There is always the possibility that novel interactions that promote stable binding to RT occur with 2'-*O*-methyl groups. However, the  $k_{off}$  values were essentially identical for the 2'-*O*-methyl and 2'-hydroxyl versions of the substrates containing the 5-bp RNA–DNA

hybrid region (compare P33-D15R5D30 and P33-D15methylR5D30) and the substrate with a single RNA at the -4 position (compare P33-D18R1(-4)D31 and P33-D18methylR1(-4)D31).

The  $K_d$  values were also the same for methylated and non-methylated versions of the 5-bp RNA–DNA substrates. It is therefore unlikely that novel interactions with 2'-*O*-methyls would have influenced RT binding to the same extent as 2'-hydroxyl interactions. Taken together, these results strongly suggest that amino acid interactions with RNA 2'-hydroxyls do not play a major role in stabilizing the binding of RT to RNA–DNA.

*4.2 The H-Form versus B-Form Structure of RNA-DNA and DNA-DNA Hybrids. Respectively, is the Most Likely Explanation for HIV-RT's Slower Dissociation from RNA-DNA*

The results presented above provide strong evidence in favor of an alternative explanation for why HIV-RT binds more stably to RNA templates; however, there is no clear indication as to what this alternative is. Some possibilities include: (i) the proposed H-form structure of RNA–DNA hybrids is more conducive to stable binding than the B-form structure of DNA–DNA; (2) differences in ‘flexibility’ between DNA–DNA and RNA–DNA hybrids (53) allow the latter to more easily conform to RT’s angled binding cleft; and (3) RT binds stably to the chimeric RNA–DNA substrates tested here not because they ‘mimic’ RNA–DNA but because they induce a bend or kink in the substrate that is favorable to RT binding. The last possibility was largely discredited in a previous report with chimeric nucleic acids

(8). Of the remaining two, we favor the first.

The first possibility would seem to be weakened by the fact that RT's slow dissociation from RNA can be mimicked by duplexes with only short stretches of RNA–DNA. Duplexes with 4- and 5-nt RNA–DNA hybrid regions may take on some H-form qualities over the short stretch, but this seems less likely for the duplexes that had only one or two RNA nucleotides. Still, these individual nucleotides may make it easier for the substrate to transition to a structure that conforms to RT's binding cleft. Pertinent to this was the preference for uridine at the -4 position, which may also make the transition easier.

Crystal structures of RT bound to DNA–DNA and RNA–DNA duplexes show that the ~4-bp hybrid region proximal to RT's polymerase active site has a structure similar, but not identical, to A-form nucleic acid (the structure formed by RNA-RNA hybrids), followed by a ~40° bend that occurs over 4 bp. The remaining duplex in the binding cleft is closer to B-form, similar to normal DNA–DNA (63,43). Interestingly, the A-form portion of the duplex is nearly the same size as the 4–5-bp RNA–DNA hybrid region required for maximal binding stability in these experiments and is positioned in the exact same region as the A-form DNA–DNA or RNA–DNA in the crystals. Although RNA–DNA is H-form rather than A-form, H-form, a 'hybrid' of A- and B-forms, is closer in structure to A-form than B-form is (38). Therefore, RT may have to contort DNA–DNA hybrids more in order to get them to fit properly into the active site, leading to less stable binding.

### 4.3 Both On- and Off-Rates appear to be Different for RNA-DNA versus DNA-DNA

#### Duplexes

Differences in  $k_{\text{off}}$  values between RNA–DNA and DNA–DNA bound to RT in binary complexes were not reflected in  $K_{\text{d}}$  values, which were similar for those duplexes that were tested (see results sections above). These findings are consistent with other reports that show similar  $K_{\text{d}}$  values, despite different off-rates, for RT binding to DNA–DNA and RNA–DNA (81,45) and minus strand initiation complexes with tRNA<sup>Lys,3</sup>-RNA versus the same RNA primed with an 18-nt DNA that is homologous to the last 18 3' nt of tRNA<sup>Lys,3</sup> (48). In that case, although  $K_{\text{d}}$  values differed by only ~2-fold, the off-rate of the DNA-primed complex was ~200 times slower. The  $K_{\text{d}}$  values reported for both duplex types analyzed here are also similar in magnitude to those reported by others in comparative studies, though the relative difference between  $k_{\text{off}}$  values for RNA–DNA and DNA–DNA are, in general, greater than previously reported (81,45, 80). It is possible that this difference could be sequence specific. Likewise, such a difference may also stem from the very tight binding of RT to the chimeric substrate (P33-D15R5D30), which bound RT about ~2.5-fold more stably than the substrate with a complete RNA strand in our previous experiments (8).

Overall the similar  $K_{\text{d}}$ , but vastly different  $k_{\text{off}}$ , values imply large differences in the  $k_{\text{on}}$  for RNA–DNA and DNA–DNA. However, for all the substrates tested in the current report to converge to a similar  $K_{\text{d}}$  value, approximately proportional changes in the on-rates, which compensate for the widely differing off-rates, must be invoked. This suggests that achieving a more stable binding state requires additional

time or steps that are reflected in a slower on-rate. Apparently, this also results in an approximately equal decrease in the off-rate. In comparing DNA-primed RNA and DNA templates, Whörl *et al.* (28) noted a slow isomerization step that occurred after initial binding of RT to the substrate. The RNA–DNA substrate showed a greater propensity to undergo the isomerization step and form a ‘productive’ complex. The authors also suggest, as we do, that structural differences between the duplexes make it easier for RNA–DNA to conform to the RT binding cleft.

Since  $K_d$  values are similar for both duplex types, any bound state of RT, including those that have or have not undergone isomerization and those with more or less stable isomerized states, must quickly convert to a catalytically competent form upon ternary complex formation. This transition occurs without significant dissociation of RT from the substrate, leading to similar  $K_d$  values for each duplex type. This is consistent with biochemical analysis indicating that dNTP binding stabilizes the RT-substrate complex (45, 75, 47). In turn, the biochemical data are consistent with crystal structures of RT in binary and ternary complexes with DNA–DNA duplexes. Nucleotide binding leads to significant closure of the gap between the finger and thumb domains as they close down on the primer-template and stabilize substrate binding (41).

#### 4.4 Structural Differences between HIV-, AMV-, and MuLV-RT May Explain

##### Differences in Duplex Binding

Results indicated that all tested RTs exhibited a preference for binding RNA–DNA, although clear quantitative differences existed between each. A comparison of

RT binding to P33-D50 versus P33-D15R5D30 in the presence of  $Mg^{2+}$  and to P33-D50 versus P33-R50 in its absence revealed that HIV-RT showed the most difference between the different duplexes, followed by AMV-RT, and, finally, by MuLV-RT, which showed the least difference. It is notable that MuLV-RT is the only monomeric enzyme in this group, while HIV- and AMV-RTs are both heterodimers (74). Perhaps even more relevant are the extensive differences observed in crystal structures of MuLV-RT compared to HIV-RT (14,17). The single available crystal structure of the MuLV-RT complete monomer suggests that the duplex trajectory is significantly different compared to HIV-RT as are the positions of the finger and thumb domain. Computer modeling also suggests significant differences in how the template bends while traversing through the binding clefts of the two enzymes (14). Therefore, the structural explanation provided here, which relies on the B- to A-form transition of the duplex while bound to HIV-RT leading to more stable RNA–DNA binding, may not be relevant for MuLV-RT. A crystal structure of MuLV-RT with a clearly resolved nucleic acid duplex would help to answer these questions.

## Chapter 5: Future Directions

### 5.1 Binding of Non-Retroviral RT Species to RNA versus DNA Substrates

#### 5.1.1 Ty3 Reverse Transcriptase (Ty3-RT)

Here we provide an analysis of the requirements necessary and sufficient for tighter binding of various retroviral RTs to RNA-DNA vs. DNA-DNA duplexes. To provide a more comprehensive depiction of this interaction, we propose that future studies include data on non-retroviral species, such as the *Saccharomyces cerevisiae* Ty3 retrotransposon. Though not of retroviral origin, the p55 reverse transcriptase from the Ty3 retrotransposon (Ty3-RT) possesses both a polymerase and RNase H domain and requires both a host-derived tRNA primer and its own, conserved polypurine tract (PPT) sequence in order to initiate (-) strand and (+) strand synthesis, respectively (29, 58). Despite this apparent similarity, evidence suggests that Ty3-RT may be distinct in that (i) it is less active than its HIV-1 counterpart, suggesting an augmented level of strand displacement activity, (ii) RNase H-mediated selection and the release of the PPT primer is less precise than for retroviruses, and (iii) the interaction between the integrase protein and Ty3-RT may be more crucial for successful replication than in other, similar viruses (46, 54, 58, 76). Given these characteristics, subsequent studies have sought to explore the interactions of Ty3-RT with various DNA and RNA intermediates in an effort to provide a more descriptive picture of how retrotransposon-based reverse transcription takes place.

In one study by Bibillo and colleagues (2005), extension of duplex DNA and

DNA-RNA hybrids was examined in the presence of Ty3-RT which had undergone amino acid (AA) substitutions in both the polymerase and RNase H domains. Results suggested that the introduction of AA sequence substitutions in the enzyme had a dramatic effect on both the processivity of the enzyme on the aforementioned substrates, as well as proper functioning of the RNase H-mediated RNA decay mechanism. Interestingly, data also indicated that Ty3-RT might exploit similar structural motifs as HIV-RT during synthesis events. While it has been acknowledged that limited information currently exists on retrotransposons of this nature, such findings (see also: 25) continue to provide valuable, comparative knowledge regarding the reverse transcription process in this species. With regard to our own studies, determination of the  $k_{\text{off}}$  and  $K_d$  values and comparison of these values against those of the RTs studied herein may also aid in this effort.

#### 5.1.2 Biological Relevance for Different Binding States on DNA-DNA and RNA-DNA

In general, the ability of RT to form a stable, tight binding complex on RNA-DNA in the absence of dNTPs may aid in carrying out RNase H cleavage of RNA fragments that remain bound to the template after DNA synthesis. Since RT does not completely degrade the RNA genome during synthesis, secondary cleavage events are required to remove fragments that remain associated with the nascent DNA or to process important regions of the genome, such as the polypurine tract, so it can be used for second strand priming (24, 77-79). The orientation of RT during secondary cleavage places the polymerase domain at the 5'-recessed end of the RNA fragment

where extension cannot occur and dNTP binding is unlikely (19, 22, 23, 55).

Therefore, it is important for RT to be able to bind RNA–DNA stably even in the absence of dNTPs. Obtaining results with Ty3-RT that suggest tighter binding of the enzyme to RNA vs. DNA templates or, more specifically, to conserved regions such as the polypurine tract, would further support this argument.

## 5.2 Dissociation Studies of HIV-RT

### 5.2.1 Background on Dissociation Studies of HIV-RT

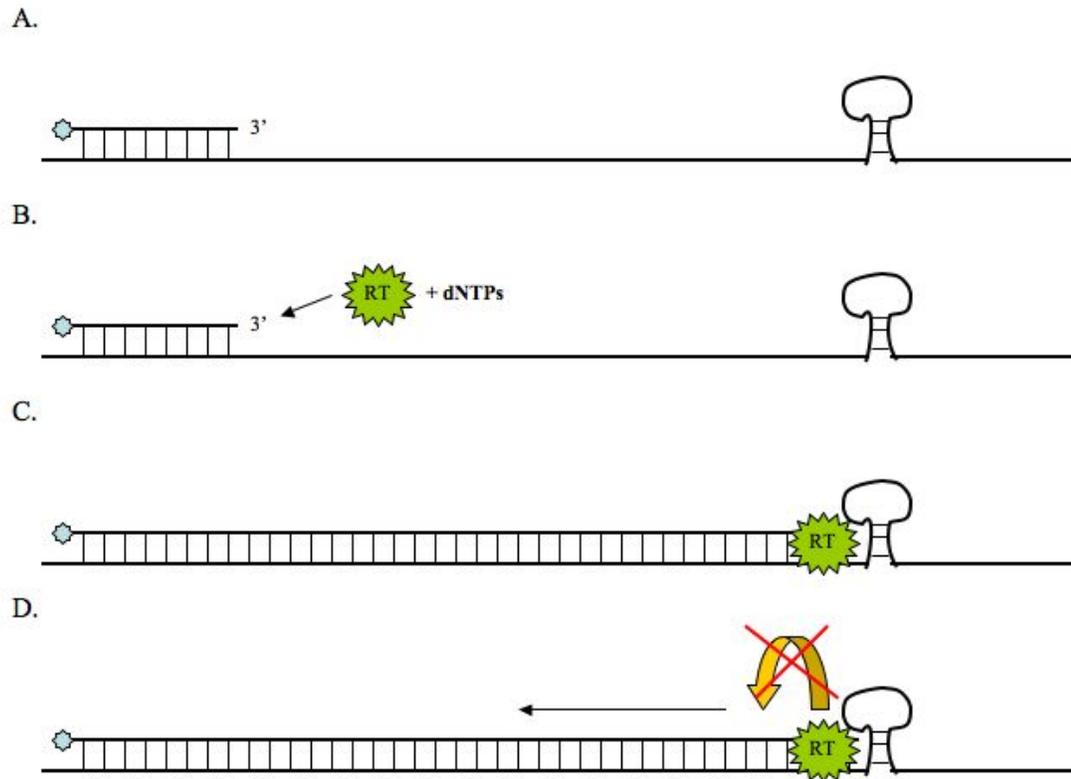
For well over a decade, researchers have analyzed the interactions of HIV-RT with various primer-template substrates in an effort to provide a detailed account of how this enzyme binds to, processes, and dissociates from such substrates during replication events. Such work has shown, for instance, that primer-template length and composition [RNA-DNA vs. DNA-DNA and the respective sequences of these hybrids], the concentration of the HIV-RT cofactor,  $Mg^{2+}$ , the presence of the HIV-1 nucleocapsid protein, and template secondary structure may all have a dramatic impact on the replication cycle and on dissociation of HIV-RT following DNA synthesis (33, 16, 36, 71, 72, 2; respectively). While this is the case, the exact mechanism of enzymatic dissociation is not entirely known.

Recently, Liu *et al.* (2008) showed that HIV-RT could shuttle between opposite termini of RNA-DNA and DNA-DNA duplexes, facilitating targeting of RT to the primer terminus for DNA polymerization and positioning of the template for RNA degradation, among other features. Furthermore, RT was able to flip into the

polymerization orientation once it had reached the primer terminus without first dissociating from the template. Such a finding is interesting, the authors argue, in light of HIV-RT's relatively low-processivity level and multifunctional nature.

### 5.2.2 Proposed Experiments to Analyze those Requirements Necessary for Dissociation of HIV-RT from its Respective Template

In an effort to block “sliding” events from occurring, custom oligonucleotides consisting of streptavidin-bound, biotinylated DNA and/or RNA nucleotides will be obtained from IDT. Modified bases will be located either upstream, downstream, or both upstream and downstream from the replication initiation site so as to help determine which direction, if any, RT must slide in order for dissociation to occur. Though previous reports have demonstrated the success of such templates in preventing translocation events in other model systems (67-68), it is unknown whether or not this same mechanism will be successful in blocking HIV-RT from “sliding” along its respective template. In the event that the above method is not sufficient to block RT trafficking, secondary structural modifications will instead be introduced into the template strand as a viable means by which to impede RT shuffling (see Figure 12).



**Figure 12.** Model of HIV-RT extension on a template containing a secondary structural modification. (A) Radiolabeled primer is hybridized to a template containing DNA, RNA or both DNA and RNA. (B) HIV-RT and dNTPs are introduced. (C) The introduction of these reagents allows for primer extension by RT. (D) Interaction of RT with template secondary structure causes the enzyme to pause, possibly preventing or slowing RT dissociation. RT potentially has the capacity to then “slide” in the 5’ direction in order to dissociate.

All other materials and methods are as in section 2.2.3 above.

### 5.2.3 Discussion of Possible Outcomes

Given previous reports (see, as example, Liu *et al.*, 2008), as well as our own findings regarding the effect of primer length on RT binding (see section 2.3.4 above), it is plausible that, in addition to its ability to traffic along primer-template substrates during extension and RNase H-mediated RNA degradation events, “sliding” along these same substrates may be necessary in order for HIV-1 reverse transcriptase to dissociate from the replication complex. We have proposed to explore this possibility by determining the dissociation rate of HIV-RT on various modified substrates, as described above in section 5.2.2. We anticipate that one of the following hypotheses is true:

- HIV-RT will require a “sliding” mechanism in order to dissociate from its template (and hence structural modifications to this template will prevent or slow dissociation from occurring)
- HIV-RT dissociates directly from the template without employing a “sliding” mechanism (in which case structural modifications should have no effect)

In the case of the former, it is expected that autoradiographic and quantitative evidence will indicate similar levels of primer extension at each time point, as RT would theoretically not be able to dissociate from the template nor continue extension given the structural modifications imposed. We would not expect this to happen in

the latter circumstance but rather would anticipate the dissociation rate to remain consistent with previously reported data (as discussed above). For instance, the P33-D50 dissociation value, as listed above, is  $0.90 \pm 0.12 \text{ min}^{-1}$ . If “sliding” events were not necessary in order for the polymerase to dissociate, we would expect a value similar to this. If, on the other hand, translocation between primer-template termini were required, we would expect no significant quantitative difference between those time points examined, which would likely result in a lower  $k_{\text{off}}$  value.

## References

1. **Abbink, T., and B. Berkhout.** 2008. HIV-1 reverse transcription initiation: a potential target for novel antivirals? *Virus Res.* **134**:4-18.
2. **Abbotts, J., Bebenek, K., Kunkel, T., and S. Wilson.** 1993. Mechanism of HIV-1 reverse transcriptase. *The Journal of Biol. Chem.* **14**:10312-23.
3. **Barouch, D.** 2008. Challenges in the development of an HIV-1 vaccine. *Nature.* **455**:613-619.
4. **Battula, N., and L. Loeb.** 1974. The infidelity of avian myeloblastosis virus deoxyribonucleic acid polymerase in polynucleotide replication. *The Journal of Biol. Chem.* **249**:4086-93.
5. **Beilhartz, G., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and M. Götte.** 2009. HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition. *J. Mol. Biol.* **388**:462-74.
6. **Berkhout, B.** 2009. Toward a durable anti-HIV gene therapy based on RNA interference. *Oligonucleotide Therapeutics.* **1175**:3-14.
7. **Bibillo, A, Lener, D., Tewari, A., and S. Le Grice.** 2005. Interaction of the Ty3 reverse transcriptase thumb subdomain with template-primer. *The Journal of Biol. Chem.* **280**:30282-90.
8. **Bohlayer, W., and J. DeStefano.** 2006. Tighter binding of HIV reverse transcriptase to RNA-DNA versus DNA-DNA results mostly from interaction in the polymerase domain and requires just a small stretch of RNA-DNA. *Biochemistry* **45**:7628-38.
9. **Boyer, J., Bebenek, K., and T. Kunkel.** 1992. Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. *Proc. Natl. Acad. Sci.* **89**:6919-23.
10. **Cane, P.** 2009. New developments in HIV drug resistance. *Journal of Antimicrobial Chemotherapy.* **64**:i37-i40.
11. **Cheatham III, T., and P. Kollman.** 1997. Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA, and DNA:RNA hybrid duplexes. *J. Am. Chem. Soc.* **119**:4805-25.
12. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. *Retroviruses.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. **Cohen, M., Hellmann, N., Levy, J., DeCock, K., and J. Lange.** 2008. The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. *The Journal of Clin. Invest.* **118**:1244-54.
14. **Coté, M., and M. Roth.** 2008. Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. *Virus Res.* **134**:186-202.
15. **Craigie, R.** 2001. HIV integrase, a brief overview from chemistry to therapeutics. *J. Biol. Chem.* **276**:23213-23216.
16. **Cristofaro, J., Rausch, J., Le Grice, S., and J. DeStefano.** 2002. Mutations in the ribonuclease H active site of HIV-RT reveal a role for this site in stabilizing enzyme-primer-template binding. *Biochemistry* **41**:10968-75.

17. **Das, D. and M. Georgiadis.** 2004. The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure* **12**:819-829.
18. **Demirov, D., and E. Freed.** 2004. Retrovirus budding. *Virus Res.* **106**:87-102.
19. **DeStefano, J.** 1995. The orientation of binding of human immunodeficiency virus reverse transcriptase on nucleic acid hybrids. *Nucleic Acids Res.* **23**:3901-8.
20. **DeStefano, J., Bambara, R., and P. Fay.** 1993. Parameters that influence the binding of human immunodeficiency virus reverse transcriptase to nucleic acid structures. *Biochemistry* **32**:6908-15.
21. **DeStefano, J., Buiser, R., Mallaber, L., Fay, P., and R. Bambara.** 1992. Parameters that influence processive synthesis and site-specific termination by human immunodeficiency virus reverse transcriptase on RNA and DNA templates. *Biochim. Biophys. Acta.* **1131**:270-80.
22. **DeStefano, J., Cristofaro, J., Derebail, S., Bohlayer, W., and M. Fitzgerald-Heath.** 2001. Physical mapping of HIV reverse transcriptase to the 5' end of RNA primers. *J. Biol. Chem.* **276**:32515-32521.
23. **DeStefano, J., Mallaber, L., Fay, P., and R. Bambara.** 1993. Determinants of the RNase H cleavage specificity of human immunodeficiency virus reverse transcriptase. *Nucleic Acids Res.* **21**:4330-4338.
24. **DeStefano, J., Mallaber, L., Fay, P., and R. Bambara.** 1994. Quantitative analysis of RNA cleavage during RNA-directed DNA synthesis by human immunodeficiency and avian myeloblastosis virus reverse transcriptases. *Nucleic Acids Res.* **22**:3793-3800.
25. **Eickbush, T., and V. Jamburuthugoda.** 2008. The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res.* **134**: 221-34.
26. **Fisher, T., Darden, T., and V. Prasad.** 2003. Mutations proximal to the minor groove-binding track of human immunodeficiency virus type 1 reverse transcriptase differentially affect utilization of RNA versus DNA as template. *Journal of Virol.* **77**:5837-45.
27. **Fletcher, R., Holleschak, G., Nagy, E., Arion, D., Borkow, G., Gu, Z., Wainberg, M., and M. Parniak.** 1996. Single-step purification of recombinant wild-type and mutant HIV-1 reverse transcriptase. *Protein Expression and Purification* **7**:27-32.
28. **Fuentes, G., Fay, P., and R. Bambara.** 1996. Relationship between plus strand DNA synthesis and removal of downstream segments of RNA by human immunodeficiency virus, murine leukemia virus and avian myeloblastoma virus reverse transcriptase. *Nucleic Acids Res.* **24**:1719-26.
29. **Gabus, C., Ficheux, D., Rau, M., Kieth, G., Sandmeyer, S., and J.L. Darlix.** 1998. The yeast Ty3 retrotransposon contains a 5'-3' bipartite primer-binding site and encodes nucleocapsid protein NCp9 functionally homologous to HIV-1 NCp7. *EMBO J.* **17**:4873-80.
30. **Gao., F., Bailes, E., Robertson, D., Chen, Y., Rodenburg, C., Michael, S., Cummins, L., Arthur, L., Peeters, M., Shaw, G., Sharp, P., and B. Hahn.**

1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**:436-41.
31. **Goff, S.** 1990. Retroviral reverse transcriptase: synthesis, structure, and function. *J Acquir. Immune Defic. Syndr.* **3**:817-31.
  32. **Gomez, C., and T. Hope.** 2005. The ins and outs of HIV replication. *Cell Microbiol.* **7**:621-626.
  33. **Gorshkova, I., Rausch, J., Le Grice, S., and R. Crouch.** 2001. HIV-1 reverse transcriptase interaction with model RNA-DNA duplexes. *Anal. Biochemistry* **291**:198-206.
  34. **Götte, M., Li, X., and M. Wainberg.** 1999. HIV-1 reverse transcription: a brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. *Arch. of Biochem. and Biophys.* **365**:199-210.
  35. **Gregerson, D., Albert, J., and T. Reid.** 1980. Processive nature of reverse transcription by avian myeloblastosis virus deoxyribonucleic acid polymerase. *Biochemistry* **19**:301-6.
  36. **Grohmann, D., Godet, J., Mély, Y., Darlix JL., T. Restle.** 2008. HIV-1 nucleocapsid traps reverse transcriptase on nucleic acid substrates. *Biochemistry* **47**:12230-40.
  37. **Hermann, T., and H. Heumann.** 1996. Strained template under the thumbs: how reverse transcriptase of human immunodeficiency virus type 1 moves along its template. *Eur. J. Biochem.* **242**:98-103.
  38. **Horton, N. and B. Finzel.** 1996. The structure of an RNA/DNA hybrid: A substrate of the ribonuclease activity of HIV-1 reverse transcriptase. *J. Mol. Biol.* **264**:521-533.
  39. **Houts, G., Miyagi, M., Ellis, C., Beard, D., and J. Beard.** 1979. Reverse transcriptase from avian myeloblastosis virus. *Journal of Virol.* **29**:517-22.
  40. **Hsieh, J., Zinnen, S., and P. Modrich.** 1993. Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J. Biol. Chem.* **268**:24607-24613.
  41. **Huang, H., Chopra, R., Verdine, G., and S. Harrison.** 1998. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance [see comments]. *Science.* **282**:1669-1675.
  42. **Hupfeld, J., and T. Efferth.** 2009. Drug resistance of human immunodeficiency virus and overcoming it by natural products. *in vivo.* **23**:1-6.
  43. **Jacobo-Molina, A., Ding, J., Nanni, R., Clark, A., Lu, X., Tantillo, C., Williams, R., Kamer, G., Ferris, A., Clark, P. et al.** 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl Acad. Sci. USA* **90**:6320-6324.
  44. **Jeang, K.** *HIV-1: Molecular Biology and Pathogenesis.* San Diego: Academic Press, 2007.
  45. **Kati, W., Johnson, K., Jerva, L., and K. Anderson.** 1992. Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.* **267**:25988-97.
  46. **Kirchner, J., and S. Sandmeyer.** 1996. Ty3 integrase mutants defective in reverse transcription or 3'-end processing of extrachromosomal Ty3 DNA. *J Virol.* **70**:4737-47.

47. **Kruhoffer, M., Urbanke, C., and F. Grosse.** 1993. Two step binding of HIV-1 reverse transcriptase to nucleic acid substrates. *Nucleic Acids Res.* **21**:3943-3949.
48. **Lanchy, J., Isel, C., Ehresmann, C., Marquet, R., and B. Ehresmann.** 1996. Structural and functional evidence that initiation and elongation of HIV-1 reverse transcription are distinct processes. *Biochimie.* **78**:1087-1096.
49. **Largaespada, D.** 2000. Genetic heterogeneity in acute myeloid leukemia: maximizing information flow from MuLV mutagenesis studies. *Leukemia* **14**:1174-84.
50. **Levy, J.** 1993. HIV pathogenesis and long-term survival. *AIDS* **7**:1401-10.
51. **Liu, S., Abbondanzieri, E., Rausch, J., Le Grice, S., and X. Zhuang.** 2008. Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science* **322**:1092-97.
52. **Negroni, M., and H. Buc.** 2001. Retroviral recombination: what drives the switch? *Nature Reviews* **2**:151-5.
53. **Noy, A., Perez, A., Marquez, M., Luque, F., and M. Orozco.** 2005. Structure recognition properties, and flexibility of the DNA:RNA hybrid. *J. Am. Chem. Soc.* **127**:4910-4920.
54. **Nymark-McMahon, M., and S. Sandmeyer.** 1999. Mutations in nonconserved domains of Ty3 integrase affect multiple stages of the Ty3 life cycle. *J Virol.* **73**:453-65.
55. **Palaniappan, C., Fuentes, G., Rodriguez-Rodriguez, L., Fay, P., and R. Bambara.** 1996. Helix structure and ends of RNA/DNA hybrids direct the cleavage specificity of HIV-1 reverse transcriptase RNase H. *J. Biol. Chem.* **271**:2063-2070.
56. **Perbal, B.** 2008. Avian myeloblastosis virus (AMV): only one side of the coin. *Retrovirology* **5**:49-53.
57. **Ramirez, B., Simon-Loriere, E., Galetto, R., and M. Negroni.** 2008. Implications of recombination for HIV diversity. *Virus Res.* **134**:64-73.
58. **Rausch, J., Bona-Le Grice, M., Henrietta, M., Nymark-McMahon, Miller, J., and S. Le Grice.** 2000. Interaction of p55 reverse transcriptase from the *Saccharomyces cerevisiae* retrotransposon Ty3 with conformationally distinct nucleic acid duplexes. *The Journal of Biol. Chem.* **275**:13879-87.
59. **Reardon, J.** 1992. Human immunodeficiency virus reverse transcriptase: steady-state and pre-steady-state kinetics of nucleotide incorporation. *Biochemistry* **31**:4473-79.
60. **Requejo, H.** 2006. Worldwide molecular epidemiology of HIV. *Rev. Saúde Pública.* **40**:331-45.
61. **Richter-Cook, N., Howard, K., Cirino, N., Wöhrl, B., and S. Le Grice.** 1992. Interaction of tRNA(Lys-3) with multiple forms of human immunodeficiency virus reverse transcriptase. *The Journal of Biol. Chem.* **267**:15952-7.
62. **Sambrook, J., and D. Russell.** *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.
63. **Sarafianos, S., Das, K., Tantillo, C., Clark, A., Jr., Ding, J., Whitcomb, J., Boyer, P., Hughes, S., and E. Arnold.** 2001. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**:1449-

- 1461.
64. **Schulz, R., Ghirikjian, J., and T. Papas.** 1981. Analysis of avian myeloblastosis viral RNA and *in vitro* synthesis of proviral DNA. Proc. Natl. Acad. Sci. **78**:2047-61.
  65. **Seckler, J., Howard, K., Barkley, M., and P. Wintrode.** 2009. Solution structural dynamics of HIV-1 reverse transcriptase heterodimer. Biochemistry **48**:7646-55.
  66. **Shehu-Xhilaga, M., Crowe, S., and J. Mak.** 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. J. Virol. **75**:1834-1841.
  67. **Shin, J and Z. Kelman.** 2006. The replicative helicases of bacteria, archaea, and eukarya can unwind RNA-DNA hybrid structures. The Journal of Biol. Chem. **281**:26914-21.
  68. **Shin, J.** 2008. Helicase translocation assay method using avidin and biotinylated nucleotides. Biotechnol Lett. **30**:2007-12.
  69. **Silverman, A., Garforth, S., Prasad, V., and E. Kool.** 2009. Probing the active site steric flexibility of HIV-1 reverse transcriptase: different constraints for DNA- versus RNA- templated synthesis. Biochemistry **47**:4800-07.
  70. **Subramani, P., Rajakannu, P., Sudhakar, P., and N. Jayaprakash.** 2005. Targeting the HIV-1 reverse transcriptase, integrase, P27, by expression of RNAi oligonucleotides from engineered human artificial chromosome. Journal of Young Investigators. **13**:published online at <<http://www.jyi.org/research/re.php?id=637>>.
  71. **Suo, Z., and K. Johnson.** 1997. Effect of RNA secondary structure on the kinetics of DNA synthesis catalyzed by HIV-1 reverse transcriptase. Biochemistry **36**:12459-67.
  72. **Suo, Z., and K. Johnson.** 1998. DNA secondary structure effects on DNA synthesis catalyzed by HIV-1 reverse transcriptase. The Journal of Biol. Chem. **42**:27259-67.
  73. **Swanson, C., and M. Malim.** 2008. SnapShot: HIV-1 proteins. Cell. **133**:742.
  74. **Telesnitsky, A. and S. Goff.** 1993. Reverse Transcriptase. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
  75. **Tong, W., Lu, C., Sharma, S., Matsuura, S., So, A., and W. Scott.** 1997. Nucleotide-induced stable complex formation by HIV-1 reverse transcriptase. Biochemistry. **36**:5749-5757.
  76. **Wilhelm, M., Heyman, T., Friant, S., and F. Wilhelm.** 1997. Heterogeneous terminal structure of Ty1 and Ty3 reverse transcripts. Nucleic Acids Res. **25**:2161-66.
  77. **Wisniewski, M., Balakrishnan, M., Palaniappan, C., Fay, P., and R. Bambara.** 2000. Unique progressive cleavage mechanism of HIV reverse transcriptase RNase H. Proc. Natl Acad. Sci. USA. **97**:11978-11983.
  78. **Wisniewski, M., Balakrishnan, M., Palaniappan, C., Fay, P., and R. Bambara.** 2000. The sequential mechanism of HIV reverse transcriptase RNase H. J. Biol. Chem. **275**:37664-37671.

79. **Wisniewski, M., Chen, Y., Balakrishnan, M., Palaniappan, C., Roques, B., Fay, P., and R. Bambara.** 2002. Substrate requirements for secondary cleavage by HIV-1 reverse transcriptase RNase H. *J. Biol. Chem.* **277**:28400-28410.
80. **Wöhrl, B., Krebs, R., Goody, R., and T. Restle.** 1999. 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. *J. Mol. Biol.* **292**: 333-344.
81. **Yu, H. and M. Goodman.** 1992. Comparison of HIV-1 and avian myeloblastosis virus reverse transcriptase fidelity on RNA and DNA templates. *The Journal of Biol. Chem.* **267**:10888-96.