

HIV-1 Reverse Transcriptase-associated Ribonuclease H Activity: Novel Mechanisms of
Inhibition

Gregory Beilhartz

Department of Microbiology and Immunology
McGill University, Montréal, Québec, Canada

January 2012

Thesis submitted to McGill University in partial fulfillment of the degree of Doctor of
Philosophy

© Gregory Beilhartz 2012

Abstract

The reverse transcriptase (RT) of HIV-1 contains two active sites, polymerase and ribonuclease (RNase) H that are both absolutely necessary for viral replication and disease progression. While several drugs currently used in the clinic to treat HIV-1 target the polymerase activity of RT, none currently target the RNase H activity specifically. The work contained in this thesis describes a potential intrinsic biochemical obstacle to the development of the most common class of RNase H inhibitors, namely steric competition with the natural substrate. Two active site RNase H inhibitors are studied in detail, β -thujaplicinol and GSK5750. Both bind to the RNase H active site through a metal ion-chelating mechanism; however neither is capable of accessing their binding sites in the presence of a pre-formed enzyme-substrate complex, and must access the RNase H active site through either the free enzyme, or the post-cleavage product complex. GSK5750 binds with much higher affinity than β -thujaplicinol, and therefore may represent progress toward a compound that can out-compete the nucleic acid substrate. This thesis also includes work that shows for the first time a potential for flexibility in enzyme-substrate contacts, especially in the vicinity of the RNase H primer grip structural motif. This may have implications for how RT binds to its substrate, in particular RNA vs DNA templates. In summary, this thesis contains significant advances in the field of HIV-1 drug development through novel mechanistic insights. It also provides a novel technique for the study of enzyme-substrate interactions in the vicinity of the RNase H primer grip.

Résumé

La transcriptase inverse (RT) du VIH-1 contient deux sites actifs: la polymérase et la ribonucléase (RNase) H. Cette enzyme est à la fois absolument nécessaire pour la réplication virale et pour la progression de la maladie. Alors que plusieurs médicaments utilisés présentement en clinique pour traiter le VIH-1 ciblent l'activité de la polymérase de la RT, aucun médicament ne cible l'activité de la RNase H de façon spécifique. Les travaux contenus dans cette thèse décrivent un obstacle biochimique potentiel pour le développement des inhibiteurs de la RNase H: le conflit stérique avec le substrat naturel. Deux inhibiteurs du site actif de la RNase H sont étudiés en détail: le β -thujaplicinol et le GSK5750. Les deux se lient au site actif de la RNase H par un mécanisme basé sur les ions de métal. Par contre, aucun des deux inhibiteurs n'est capable d'accéder à son site de liaison en présence d'un complexe pré-formé enzyme-substrat; les inhibiteurs doivent accéder au site actif de la RNase H soit dans l'enzyme libre, soit dans le complexe de produit. Le GSK5750 se lie avec une affinité beaucoup plus élevée que le β -thujaplicinol, et donc représente un composé qui peut supplanter le substrat d'acide nucléique. De plus, cette thèse montre pour la première fois un potentiel de flexibilité dans les contacts enzyme-substrat, particulièrement dans le environs du motif RNase H "primer grip". En résumé, cette thèse propose des avancées significatives dans le domaine du développement de médicaments contre le VIH-1, grâce à des idées mécanistiques novatrices. Elle fournit aussi une nouvelle technique pour l'étude des interactions enzyme-substrat dans les environs du motif de la RNase H "primer grip".

Table of Contents

Abstract	2
Résumé	3
List of Figures	7
Frequently Used Abbreviations	8
Preface	9
Contributions of Authors	9
Acknowledgements	12
Chapter 1	14
Human Immunodeficiency Virus Type 1	16
1.1 Virus Life Cycle	16
1.1.1 Entry of HIV-1 into cells	16
1.1.2 Capsid Uncoating	18
1.1.3 Reverse Transcription	20
Initiation and tRNA binding	20
(-)-strand DNA synthesis	20
(+) -strand DNA synthesis	21
1.1.4 Nuclear Import and Integration	23
1.1.5 Viral Gene Expression and Virus Assembly	24
1.2 HIV Gene Products	25
1.2.1 Structural Proteins – Gag, Pol and Env	26
Gag	26
Pol	28
Env	30
1.2.2 Regulatory Proteins – Tat and Rev	31
Transcriptional Transactivator (Tat)	31
Rev	32
1.2.3 Accessory Proteins: Nef, Vpu, Vpr (Vpx), and Vif	32
Nef	33
Vpu	33
Vpr and Vpx	34

Vif.....	35
1.3 HIV-1 Reverse Transcriptase Structure and Function.....	36
1.3.1 Substrate Binding to RT	38
1.3.2 RT Polymerase Structure and Function.....	41
1.3.3 RNase H Structure and Function.....	42
1.3.4 Role of RNase H in (+)-strand priming.....	48
1.3.5 Role of RNase H in strand transfer and (-)-strand primer removal	50
1.3.6 Role of RNase H activity in drug resistance	52
1.4 Inhibition of HIV-1 RT	55
1.4.1 Polymerase Inhibitors.....	55
NRTIs – Nucleos(t)ide-Analog Reverse Transcriptase Inhibitors.....	55
NNRTI – Non-Nucleoside Reverse Transcriptase Inhibitors.....	57
Foscarnet (PFA, or phosphonoformic acid)	57
1.4.2 RNase H Inhibitors.....	58
Active Site RNase H Inhibitors	58
Allosteric RNase H Inhibitors	61
Other compounds that affect RNase H activity.....	64
1.5 Viral Drug Resistance	65
1.5.1 NRTIs	66
NRTI Resistance-Confering Mutations – Excision.....	66
NRTI Resistance-Confering Mutations – Discrimination.....	68
1.5.2 NNRTIs	69
1.5.3 PFA/Foscarnet.....	71
1.6 Objectives.....	71
Chapter 2	90
2.1 Abstract	91
2.2 Introduction	92
2.3 Results	95
2.4 Discussion	102
2.5 Materials and Methods	108

2.6 Figures and Figure Legends	114
Chapter 3	124
3.1 Preface	125
3.2 Abstract	126
3.3 Introduction	126
3.4 Results	129
3.5 Discussion	133
3.6 Materials and Methods	139
3.7 Figures and Figure Legends	143
Chapter 4	152
4.1 Preface	153
4.2 Abstract	154
4.3 Introduction	154
4.4 Results	156
4.5 Discussion	159
4.6 Materials and Methods	163
4.7 Figures and Figure Legends	165
Chapter 5	171
References	183

List of Figures

Figure 1.1.....	71
Figure 1.2.....	72
Figure 1.3.....	73
Figure 1.4.....	74
Figure 1.5.....	75
Figure 1.6.....	78
Figure 1.7.....	79
Figure 1.8.....	80
Figure 1.9.....	81
Figure 1.10.....	82
Figure 1.11.....	83
Figure 1.12.....	84
Figure 1.13.....	85
Figure 2.1.....	111
Figure 2.2.....	112
Figure 2.3.....	113
Figure 2.4.....	115
Figure 2.5.....	116
Figure 2.6.....	117
Figure 2.7.....	118
Figure 3.1.....	138
Figure 3.2.....	139
Figure 3.3.....	140
Figure 3.4.....	142
Figure 3.5.....	144
Figure 3.6.....	146
Figure 4.1.....	159
Figure 4.2.....	161
Figure 4.3.....	163

Frequently Used Abbreviations

AIDS	- Acquired Immunodeficiency Syndrome
HIV-1	- Human Immunodeficiency Virus Type 1
RT	- Reverse Transcriptase
RNase H	- Ribonuclease H
NRTI	- Nucleoside Analog Reverse Transcriptase Inhibitor
NNRTI	- Non-Nucleoside Analog Reverse Transcriptase Inhibitor
PFA	- Phosphonoformic Acid (Foscarnet)
ASRI	- Active Site RNase H Inhibitor
dNTP	- Deoxynucleotide Triphosphate
ddNTP	- Dideoxynucleotide Triphosphate
PPi	- Pyrophosphate
k_D	- Dissociation constant
k_{obs}	- Observed rate of reaction
Y_{max}	- Maximum product of reaction formed
E	- Enzyme
S	- Substrate
I	- Inhibitor
nt	- nucleotide
bp	- base pair
KOONO	- potassium peroxyxynitrate
FRET	- Fluorescence Resonance Energy Transfer
AZT	- 3'-azido-2',3'-dideoxythymidine (zidovudine)

Preface

This thesis was written in accordance with McGill University's "Guidelines for Thesis Preparation". The format of this thesis conforms to the "Manuscript-based thesis" option, which states: "Candidates have the option of including, as part of the thesis, the text of one or more papers submitted, or to be submitted, for publication, or the clearly duplicated text (not the reprints) of one or more published papers. These texts must conform to the "Guidelines for Thesis Preparation" with respect to font size, line spacing and margin sizes and must be bound together as an integral part of the thesis."

Manuscripts included in this thesis:

Beilhartz GL, Wendeler M, Baichoo N, Rausch J, Le Grice S, Götte M. 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.* 2009 May 8;388(3):462-74. Epub 2009 Mar 13.

Greg L. Beilhartz, Brian A. Johns, Peter Gerondelis, Matthias Götte. GSK5750 is a novel active site inhibitor of HIV-1 RT-associated RNase H activity. *Manuscript in preparation.*

Greg L. Beilhartz, Egor P. Tchesnokov, Matthias Götte. Probing the effects of the RNase H primer grip using a novel site-specific footprinting method. *Manuscript in preparation.*

Contributions of Authors

For all manuscripts in this thesis, WT and mutant RT enzymes were expressed and purified by Suzanne McCormick.

Beilhartz GL, Wendeler M, Baichoo N, Rausch J, Le Grice S, Götte M. 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. 2009 May 8;388(3):462-74. Epub 2009 Mar 13.

This manuscript appears in Chapter two of this thesis and was written by GLB, MG. All experiments were performed by GLB with the exception of IC50 experiments with unnatural amino acid substitutions and molecular modelling which were performed by MW, NB and JR.

Greg L. Beilhartz, Brian A. Johns, Peter Gerondelis, Matthias Götte. GSK5750 is a novel active site inhibitor of HIV-1 RT-associated RNase H activity. *Manuscript in preparation.*

This manuscript appears in Chapter three of this thesis and was written by GLB and MG. All experiments were performed by GLB.

Greg L. Beilhartz, Egor P. Tchesnokov, Matthias Götte. Probing the effects of the RNase H primer grip using a novel site-specific footprinting method. *Manuscript in preparation.*

This manuscript appears in Chapter four of this thesis and was written by GLB. All experiments were performed by GLB except figure 1 which was performed by EPT.

Other manuscripts not included in this thesis but to which a significant contribution was made are as follows:

van Maarseveen NM, Wensing AM, de Jong D, **Beilhartz GL**, Obikhod A, Tao S, Pinggen M, Arends JE, Hoepelman AI, Schinazi RF, Götte M, Nijhuis M. Telbivudine exerts no antiviral activity against HIV-1 in vitro and in humans. Antivir Ther. 2011;16(7):1123-30.

Auger A, **Beilhartz GL**, Zhu S, Cauchon E, Falgout JP, Grobler JA, Ehteshami M, Götte M, Melnyk RA. Impact of primer-induced conformational dynamics of HIV-1 reverse transcriptase on polymerase translocation and inhibition. *J Biol Chem*. 2011 Aug 26;286(34):29575-83. Epub 2011 Jul 7.

Biondi MJ, **Beilhartz GL**, McCormick S, Götte M. N348I in HIV-1 reverse transcriptase can counteract the nevirapine-mediated bias toward RNase H cleavage during plus-strand initiation. *J Biol Chem*. 2010 Aug 27;285(35):26966-75. Epub 2010 Jun 8.

Chung S, Wendeler M, Rausch JW, **Beilhartz G**, Götte M, O'Keefe BR, Bermingham A, Beutler JA, Liu S, Zhuang X, Le Grice SF. Structure-activity analysis of vinylogous urea inhibitors of human immunodeficiency virus-encoded ribonuclease H. *Antimicrob Agents Chemother*. 2010 Sep;54(9):3913-21. Epub 2010 Jun 14.

Ehteshami M*, **Beilhartz GL***, Scarth BJ*, Tchesnokov EP, McCormick S, Wynhoven B, Harrigan PR, Götte M. Connection domain mutations N348I and A360V in HIV-1 reverse transcriptase enhance resistance to 3'-azido-3'-deoxythymidine through both RNase H-dependent and -independent mechanisms. *J Biol Chem*. 2008 Aug 8;283(32):22222-32. Epub 2008 Jun 10.

*Authors contributed equally to this work

Invited review articles to which a significant contribution by the author was made are as follows:

Brian J Scarth, Maryam Ehteshami, **Greg L Beilhartz**, Matthias Götte. HIV-1 reverse transcriptase inhibitors: beyond classic nucleosides and non-nucleosides. *Future Virology*, May 2011, Vol. 6, No. 5, Pages 581-598.

Greg L. Beilhartz and Matthias Götte. HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors. Review: *Viruses* **2010**, 2(4), 900-926

Michaela Wendeler, **Greg L Beilhartz**, John A Beutler, Matthias Götte, Stuart FJ Le Grice. HIV ribonuclease H: continuing the search for small molecule antagonists. *HIV Therapy*, Jan 2009, Vol. 3, No. 1, Pages 39-53.

Acknowledgements

I would like to acknowledge my PhD supervisor, Dr. Matthias Götte for his help and support over the many years of my studies in his lab. I thank him as well for the many opportunities he has provided me with during my PhD. As well, I would like to thank my PhD committee members, Dr. Arnim Pause and Dr. Jacques Archambault for their support and advice during the past six years. I would also like to thank all past and present members of the Götte lab for their friendship and support over the years, especially Dr. Bruno Marchand, Dr. Egor Tchesnokov, Dr. Stephen Barry and Dr. Jerome Deval for their mentoring support in the early days of my research.

I would like to extend my very sincere thanks to Suzanne McCormick for her invaluable help for the past six years. The work she put in to provide me with enzymes, materials, support and lunchtime conversation cannot be overstated and my time in the lab was made far more productive than it would otherwise be. If there are two words I've used more than any other in the past six years, there is no doubt they are: "Thanks Sue!"

I would also like to thank Dr. Maryam Ehteshami and the Rev. Dr. Brian Scarth for being such great friends over the years. I could not have finished this without them (although, since they've both moved away, it could be said that I have finished it without them) and their "fanatical" support and company, especially on road trips, will be sorely missed. I am truly lucky to have had colleagues that can make me think, as well as make me laugh.

Thanks also to Megan Powdrill for her friendship and support throughout my PhD studies. Having a real biochemist around was undeniably beneficial.

Thanks also to Anick Auger for her help and advice with the FRET assays.

I would also like to acknowledge my parents, Wendy and Lloyd, for their unwavering support through the years. They likely did not know what they were starting when they told me to stay in school, but they've always supported me in all my endeavours, and their good advice has bailed me out on innumerable occasions. Thanks Mom and Dad.

Finally, I'd like to thank Lindsey Banks for her love and support. You've sacrificed a great deal so that I could finish this, and you've always been so supportive. It's cliché, but I honestly could not have done this without you. Thank you so much, I love you.

Chapter 1

The human immunodeficiency virus (HIV) is the etiological agent of acquired immune deficiency syndrome (AIDS) and is a global pandemic and major public health issue (1). With the exception of a small percentage of people who carry natural genetic resistance to HIV, infection with HIV is fatal unless treated. According to the World Health Organization, 33.3 million people were infected with HIV worldwide in 2009, which represents a 27% increase from 1999 (2). The epidemic is at its worst in sub-Saharan Africa, which shares 68% of the global total of infected people (2)(Fig. 1.1). Globally, the number of new infections each year (incidence) is estimated at 2.6 million in 2009, which is a 19% decrease since 1999. While the prevalence of HIV globally remains at its highest point ever, the incidence of HIV is trending downward, due in part to large-scale prevention efforts, as well as expanded treatment and other factors such as the natural course of the epidemic (Fig. 1.2). Access to anti-retroviral treatment (ART) has been expanded to more than 5 million people worldwide. This represents a 13-fold increase from 2004, and has contributed to a decrease of 19% in deaths among people infected with HIV (2). However, millions of HIV positive people still do not have access to potentially life-saving treatment.

In June 2010, the World Health Organization (WHO) and the Joint United Nations Programme on HIV/AIDS (UNAIDS) launched a new program to combat the HIV epidemic called Treatment 2.0. Treatment 2.0 outlines 5 priorities that will be attempted by the year 2020. Firstly, the optimization of drug regimens should provide an “*effective*,

affordable, one-pill, once-daily potent ARV regimen with minimal toxicities or drug interactions and high barriers to resistance in low and moderate-income countries (LMICs)” (3). Secondly, the creation of “a package of simple, affordable, reliable, quality-assured point-of-care (POC) and other simplified diagnostics to be available and accessible in LMICs” is critical to proper monitoring and subsequently, proper response to the HIV epidemic (3). Currently many diagnostic tools are not well adapted to resource-limited settings, requiring either a high level of technical expertise, or high costs. Diagnostic tools not only for HIV, but for related pathogens such as tuberculosis (TB) and viral hepatitis are also greatly desired. The third priority of Treatment 2.0 is to provide “high quality HIV care and treatment programmes available at the lowest possible cost with optimal efficiency to all in need in LMICs” (3). Fourth, Treatment 2.0 will strive to ensure that “HIV care and treatment programs are decentralized and appropriately integrated with other HIV and non-HIV health services, with increased community engagement in service delivery and improved retention in care”(3). Finally, and perhaps most importantly, it is vital that “people living with HIV and key populations are fully involved in the demand-creation, planning, delivery and evaluation of quality-assured, rights-based HIV care and treatment programs in all LMICs”(3).

The development of potent, non-toxic drugs is a central aim of all anti-HIV prevention efforts. To that end, new viral targets not yet exploited for drug development are currently being explored, including the ribonuclease H active site of HIV-1 reverse transcriptase.

This chapter will provide an introduction to the HIV virus that will aid in the

understanding of my thesis work, with emphasis on HIV-1 reverse transcriptase, and ribonuclease H activity specifically.

Human Immunodeficiency Virus Type 1

HIV belongs to the genus *lentivirus* from the family *Retroviridae*. The name *lentivirus* comes from the Latin *lente*, meaning “slow”. As such, lentiviruses have a characteristically long incubation period and are unique among retroviruses in their ability to replicate in non-dividing cells, facilitating the formation of long-term, stable reservoirs as well as being an efficient gene transfer vector. Both of these topics are discussed in more detail below. HIV has two different types, called HIV-1 and HIV-2. HIV-1 is more prevalent globally and more virulent than HIV-2, the reasons for which are not fully understood. Since HIV-2 is less capable of transmission, it is mainly confined to West Africa (4). The majority of the differences between HIV-1 and HIV-2 sequences come from a variable region in the *Env* gene (5).

1.1 Virus Life Cycle

1.1.1 Entry of HIV-1 into cells

HIV is an enveloped virus containing a cone-shaped core that houses 2 copies of the (+)-strand RNA genome and associated proteins. It infects two main types of cells, T lymphocytes and macrophages (6) (Fig. 1.3). HIV is a budding virus that is assembled at the plasma membrane of infected cells, which is how it acquires its liquid-ordered

envelope (7). The envelope contains several host proteins, as well as approximately 10 trimeric viral envelope (Env) proteins (8-10). The Env gene product is synthesized as a protein precursor called glycoprotein 160 (gp160) which undergoes a complex maturation process involving cleavage by a cellular protease to form the surface unit (SU) gp120 and the transmembrane unit (TM) gp41 (11). During virus entry, gp120 associates with the glycoprotein CD4, expressed on the surface of T helper cells, monocytes, macrophages and dendritic cells (12-15). HIV must also use a coreceptor to CD4 in order to gain entry into cells. For this, the virus has two options: chemokine receptors CCR5 and CXCR4 (16,17). Previously, HIV viruses were classified as being M-tropic (macrophage) or T-tropic (T-cell), depending on which coreceptor they used, respectively. However, since all primary HIV isolates replicate in activated, primary CD4⁺ T-lymphocytes, a new nomenclature was adopted to identify HIV phenotypes based on coreceptor usage. Viruses using CCR5 as the main coreceptor are classified as R5 viruses, viruses using CXCR4 are classified as X4 viruses, and viruses that have the ability to use both are classified as R5X4 viruses (18). Once the virus is bound to CD4 and one or both coreceptors, a conformational change occurs in the gp120-gp41 complex that may include the dissociation of gp120 (19). Next, the hydrophobic N-terminal region of gp41 known as the fusion peptide is inserted into the cellular membrane. Finally, gp41 undergoes a second conformational change to bring the viral and cellular membranes into contact and promote fusion (19). However, recent evidence suggests that fusion might actually occur with the endosomal membrane after endocytosis (20). Either way, this results in the viral core particle being injected into the cytoplasm of the newly infected cell.

1.1.2 Capsid Uncoating

As mentioned above, viral core contains 2 copies of the (+)-strand RNA genome as well as various other viral proteins. One of these is the viral reverse transcriptase (RT), responsible for synthesizing double-stranded DNA from the single-stranded viral RNA genome. The RT-genome complex at this point is referred to as the reverse transcription complex (RTC) which is simply defined as an HIV-1 complex that is undergoing reverse transcription, to differentiate it from the pre-integration complex (PIC) (21,22). PICs are defined as integration-competent that can integrate efficiently into target DNA *in vitro* (23,24). At some point between the injection of the viral core into the cytoplasm and trafficking of the PIC through a nucleopore into the nucleus of the cell for proviral integration, uncoating must take place. Uncoating represents the dissociation of the individual capsid proteins that form the complete viral capsid or core. There is significant debate in the field on the timing and location of capsid uncoating, and as an extension, the significance of the role played by HIV-1 capsid (CA) at the early stages of infection. Briefly, there are three competing hypotheses that describe HIV-1 uncoating (25).

Firstly, complete or partial uncoating occurs immediately after plasma or endosomal membrane fusion. In this scenario, uncoating is required for the formation of the RTC, and is perhaps driven by a sudden drop in CA concentration outside of the viral envelope, which causes dissociation of the metastable viral cores (26-29). In support of this idea, very low concentrations of CA are found in intracellular HIV-1 complexes shortly after infection, as well as the absence of any CA in the cytoplasm of infected cells using

transmission electron microscopy, suggesting that the CA protein dissociates quickly and is degraded (30-34).

In a second scenario, uncoating takes place gradually, while the capsid is shuttled toward the nucleus during reverse transcription. Uncoating could be caused by sequential interactions with cellular factors and/or conformational changes of the capsid. This model is supported by studies using immunofluorescent microscopy that shows CA in complex with RTCs (35). Also, studies report that intracellular HIV-1 comes in a wide range of sizes, suggesting a stepwise uncoating process (31,35-37). Furthermore, capsids with modified stability (either increased or decreased) have a negative effect on reverse transcription (38).

Finally, a third model suggests that the capsid remains intact until the RTC reaches the nucleus, and uncoating occurs at the nucleopore after reverse transcription is completed. In this scenario, the capsid plays an important role in maintaining the concentration of viral proteins involved in reverse transcription, most notably RT. RT is a moderately processive enzyme and requires at least one template switching event during the reverse transcription process (see below), therefore a highly diluted RT concentration would lead to breakdowns in reverse transcription, and non-integrating complexes (39). However, the capsid is still a permeable membrane to small molecules such as dNTPs and RT inhibitors (40). In support of this hypothesis, recent studies have shown that successful transport of the HIV complex to the nucleus, nuclear import, and reverse transcription are highly dependent on both the integrity of the capsid, as well as uncoating at the right time (33,38,41-43).

1.1.3 Reverse Transcription

Initiation and tRNA binding

As mentioned earlier, the exact timing of the beginning of reverse transcription is still a topic of research and remains in question. However, the exact steps of reverse transcription are very well understood. The process begins with the annealing of a tRNA (specifically tRNA^{Lys3}) to the primer binding site (PBS) near the 5' end of the viral RNA genome. The tRNA primer is selectively packaged into the virion by Gag-Pol precursors (44). The tRNA binds along an 18 bp sequence that comprises the ultimate 3' end of the tRNA. In addition to the sequence complementarity, secondary structures in the viral genome just upstream of the PBS play an important role in tRNA specificity and binding. An A-rich loop just upstream of the PBS interacts with the U-rich anticodon loop of the tRNA, increasing the efficiency of the initiation of reverse transcription (45-47). Also, these interactions have been shown to prevent steric clashes between RT and genomic secondary structures, further enhancing the specificity of the (-)-strand primer (48,49).

(-)-strand DNA synthesis

RT binds to the RNA/RNA duplex formed by the binding of the tRNA primer to the PBS. RT is able to bind in two distinct conformations, namely polymerase-dependent and – independent. The former is characterized by the polymerase active site of RT being at the 3' primer terminus, in position for nucleotide incorporation. The latter is defined as any orientation of RT that is not in position for DNA synthesis. When bound to the (-)-strand

primer, RT flips between these two orientations frequently (50). However, it is not yet understood if this flipping is a cause or an effect of the slow-moving RT at the initiation of (-)-strand DNA synthesis (51,52). RT then synthesizes DNA until the 5' end of the genome, producing a DNA fragment called (-)-strand strong-stop DNA. At this point, the RNase H activity of RT degrades the RNA genome concomitantly as well as subsequently to DNA synthesis. The degradation of the RNA strand is necessary and sufficient for the first strand transfer event, which is mediated by the identical repeat (R) regions at both ends of the genome. This strand transfer event can be both inter- or intramolecular due to the presence of two copies of the viral genome per virion (53). After this strand transfer event, (-)-strand DNA synthesis proceeds while the RNase H activity concomitantly degrades the RNA genome.

(+)-strand DNA synthesis

During (-)-strand synthesis, certain purine-rich sequences of the viral RNA genome are resistant to degradation by the RNase H activity of RT, and these are referred to as polypurine tracts (PPT). There are two main PPTs, one located near the center of the genome and the other located near the 3' end and are therefore called the central PPT (cPPT) and 3' PPT, although the 3' PPT is often referred to as simply, PPT (54-56). These sequences are resistant to RNase H degradation due mainly to the structure of the nucleic acid, rather than the specific sequence. Structural aberrations such as uncommon rigidity of A-tracts, unusually narrow minor groove and an unusual departure from standard Watson-Crick base pairing referred to as the “unzipping” of the PPT are all features that may play a role in the PPT's resistance to RNase H degradation (57). The

PPT sequences are then used as RNA primers for (+)-strand DNA synthesis. RT extends the PPT primer approximately 12 bp before changing orientations and removing the RNA PPT by RNase H degradation (58). RT must then change orientations a second time to bind again as a polymerase, and continue (+)-strand DNA synthesis until it reaches the (-)-strand primer (tRNA) still bound at the end of the genome (58). RT copies the PBS sequence from the tRNA primer until it reaches a methylated adenosine residue located 19 bp into the tRNA sequence (59). Unlike the (+)-strand primer which is completely removed by RNase H cleavage at the RNA-DNA junction, the (-)-strand primer is removed with a single ribonucleotide on the end (59). This is a specific cleavage by RNase H and is critical for recognition later by the viral integrase. Once the (-)-strand primer is removed, a second strand transfer event takes place using the complimentary PBS sequences on the (-) and (+)-strand DNA sequences. This strand transfer is likely intramolecular the majority of the time, since it involves the circularization of the reverse transcription complex (60). After circularization, both strands complete DNA synthesis. RT carries a strand displacement activity that allows it to displace approximately 100 bp of DNA in order to reach the central termination sequence (CTS) that provides the signal to stop DNA synthesis (61). The CTS is extremely efficient at termination polymerization during strand displacement synthesis, but only causes slight pausing of RT during (-) or (+)-strand DNA synthesis (62). Analogous to the PPT sequences, the CTS sequences causes severe structural anomalies in the double helix structure, causing RT to pause and eventually dissociate from the complex (62,63). The strand displacement activity of HIV-1 RT causes the formation of a ~100 bp flap that remains attached to the complex. Called

the central flap, it is critical for nuclear import of the PIC and subsequent viral replication (64). Proviruses devoid of a central flap have been shown to accumulate in the cytoplasm at the nuclear membrane (64).

In summary, RT catalyzes the conversion of a single-stranded RNA genome into a double-stranded DNA provirus with extended ends (U3-R-U5) called long terminal repeats (LTRs) as well as a central flap. This structure then associates with multiple enzymes both viral and host to form the pre-integration complex or PIC which is imported into the nucleus for integration into the host chromosomes to produce a persistent infection and viral reservoirs.

1.1.4 Nuclear Import and Integration

Lentiviruses are unique among the family Retroviridae in that they can infect non-dividing cells such as macrophages and resting T cells. Unlike other retroviruses, HIV does not have to wait for dissolution of the nuclear membrane during mitosis in order to access the host chromosomes (65). After reverse transcription, the PIC is shuttled towards the nucleus, possibly by association with microtubules. Upon reaching the nucleus, the PIC is actively transported through the nucleopore complex (NPC) into the nucleus. At least 4 viral factors are involved in this process, namely integrase, Vpr, the central DNA flap, and MA. All of these proteins contain a sort of nuclear localization signal (NLS), although it appears that integrase plays the largest role, while MA, Vpr and the central flap play a supportive role in nuclear import of the PIC, while adding redundancy to the process as well (66).

Within the context of the PIC, the viral integrase enzyme catalyzes the integration of the double-stranded viral DNA genome. The mechanism of integration is discussed briefly in the section on integrase (see Table of Contents). HIV-1 tends to integrate near active genes, and the reasons for this are not clear. One possibility is that active transcription provides a favorable chromatin environment, i.e. one that is more “open” and provides physical access to the PIC (67). Another option is direct interactions between the PIC and local transcription factors bound near actively transcribing genes, as is the case with some yeast retrotransposons (68,69). Other research suggests that HIV has evolved to integrate into active genes because this allows for efficient expression of viral genes after integration (70). Integration of the viral cDNA into the host genome completes what is known as the “early phase” of HIV-1 infection, and the “late phase” begins with the transcription of the viral genes, starting with the regulatory genes, Tat and Rev.

1.1.5 Viral Gene Expression and Virus Assembly

As soon as the structural genes (Gag, Pol and Env) begin to be expressed along with the full length RNA genome and accessory proteins, the assembly of virus particles begins in the cytoplasm with the translation of the Gag protein. Gag is regarded as the molecular scaffold (possibly in tandem with the viral RNA) that guides virus assembly, since Gag is capable of self-assembling into non-infectious, virus-like particles (71). In the case of T cells, virus production takes place on the plasma membrane, however in macrophages, viruses have been observed to accumulate in large vacuoles and the membranes of these vacuoles are generally considered to be the site of virus assembly in macrophages

(72,73). It is generally believed that the NC domain of Gag is responsible for selecting the unpliced full-length RNA genome for packaging and transport to the plasma membrane (74), although the exact mechanism of genome selection is not well understood, since retroviruses are capable of packaging almost any RNA molecule, and some mutants are even able to package ribosomes (75). HIV-1 acquires its envelope by budding from the plasma membrane (or vacuole membrane in macrophages). The MA domain of the Gag polyprotein associates with the plasma membrane and deforms it as Gag oligomerizes into a sphere. In order to detach from the plasma membrane completely, HIV-1 engages a host pathway called ESCRT (Endosomal Sorting Complex Required for Transport) normally used for cellular membrane fission (76). In many cell types, newly budded HIV virions are still connected to the host cell by a host restriction factor called tetherin (77,78). In HIV-1, the Vpu accessory protein negates the effect of tetherin and the virus is allowed to detach from the host cell surface. Concomitant with virus release, the viral protease cleaves Gag into CA, MA, NC and p6 proteins, with MA remaining associated with the membrane, and CA condensing to form the mature viral capsid. This process is called maturation, and is now able to bind CD4 and infect a new cell, restarting the virus life cycle.

1.2 HIV Gene Products

The HIV-1 genome consists of a pseudo-double stranded RNA genome approximately 9.8 kilobases in length, and after reverse transcription into double-stranded DNA and

integration into the host genome, the provirus is flanked at both 5' and 3' ends by long terminal repeats (LTRs) (Fig. 1.4). The HIV genome encodes 9 proteins that belong to three distinct classes: structural proteins (Gag, Pol and Env), regulatory proteins (Tat and Rev) and accessory proteins (Vif, Nef, Vpu and Vpr). Another accessory protein, Vpx is found in primate lentiviruses such as SIV and HIV-2, but is not found in HIV-1. Some proteins encoded by the HIV genome are polyproteins that are processed further into even more proteins used by the virus.

1.2.1 Structural Proteins – Gag, Pol and Env

Gag

The structural protein Gag is expressed as a polyprotein precursor called Pr55^{Gag}. During viral maturation, Pr55^{Gag} is cleaved by the viral protease into capsid (CA), matrix (MA), nucleocapsid (NC) and p6 proteins. However, the uncleaved Pr55^{Gag} is sufficient for the assembly of non-infectious, virus-like particles (79,80). As such, the proper processing of Gag is necessary for viral maturation and the creation of infectious particles. Once Gag is cleaved, MA remains associated with the viral envelope, while CA condenses to form the viral capsid or core, surrounding the viral RNA genome which is associated with NC. During virus assembly however, these proteins play an important role while part of the Pr55^{Gag} polyprotein.

The MA protein is myristolated and is largely responsible for direct binding and assembly at the plasma membrane of an infected cell. Several domains in MA are responsible for

membrane binding, and are collectively referred to as M domains (81). Pr55^{Gag} is also responsible for mediating the incorporation of Env glycoproteins into plasma membrane (82). Studies have shown that the MA domain of Pr55^{Gag} is responsible for Env incorporation (83-86).

The CA protein is responsible for Gag-Gag interactions, as well as the incorporation of the Gal-Pol polyprotein into assembling viruses (87,88). CA is also responsible for the specific incorporation of the host protein cyclophilin A (CypA) into HIV-1 virions, although the role of CypA in virions is not yet fully understood (89,90). It is known that in Old World monkeys, CypA is required for viral restriction by TRIM5 α (91). In the mature virion, CA makes up the viral core that houses the ribonucleoprotein complex.

The NC protein is an RNA-binding protein containing two zinc-finger motifs that are highly conserved in all retroviruses except for the spumaviruses (92). NC is found to be tightly associated with viral RNA inside HIV-1 virions (93). The NC domain of Gag plays a role in several viral functions such as RNA dimerization (94), RNA packaging (95), Gag-Gag interactions (96,97), membrane binding (98), tRNA binding (initiation of reverse transcription) (99), strand transfer during reverse transcription (55), and stabilization of the pre-integration complex (100).

The final protein involved in Pr55^{Gag} is a small protein called p6. During virus budding, Gag recruits several host proteins such as Tsg101 and AIP1 via its p6 domain, in concert with other Gag domains (101). As such, p6 is crucial in the release of the virus from the infected cell (102).

Pol

The Pol gene in HIV-1 is expressed as a fusion protein with Gag (Pr160^{Gag-Pol}), which is due to a ribosomal frameshifting event that occurs about 5% of the time (103,104).

During viral maturation, Pr160Gag-Pol is cleaved from Gag by the virally-encoded protease, and is then further cleaved into 3 functional enzymes: protease (PR), reverse transcriptase (RT), and integrase (IN).

HIV-1 PR is an aspartyl protease and is active as a homodimer (105). Each monomer contributes one critical aspartic acid residue to form the active enzyme (106). PR is responsible for cleaving Gag into the substituent proteins described above, as well as self-cleaving from the Pol polyprotein. Furthermore, PR cleaves the RT enzyme (p66) to remove the RNase H domain (p15) on about half of the RT enzymes, allowing it to form its active heterodimer form (p66/p51). The substrates that are recognized by HIV-1 PR are diverse in sequence, and yet PR is a highly specific enzyme. The cleavage sites in Gag and Gag-Pol have a highly conserved, asymmetric shape, which is responsible for PR's high specificity, regardless of sequence or hydrogen bonding pattern (107,108). HIV-1 PR is a highly successful antiviral target for HIV, with many PR inhibitors (PIs) in clinical use as of 2011. Due to below average pharmacokinetic properties, many PIs are “boosted” with ritonavir, a protease inhibitor that is mainly used to increase bioavailability and cell penetration of other, more potent PIs (109,110).

HIV-1 IN mediates the integration of the reverse-transcribed viral DNA genome into the chromosome of an infected cell (111). Once inserted the viral genome is called a provirus, and persists in the cell while being transcribed by host transcription/translation machinery

to produce new virions. HIV-1 IN has two distinct enzymatic processes: 3'-processing and strand transfer. 3'-processing refers to the cleavage of a GT dinucleotide just 3' to a highly conserved CA dinucleotide at both ends of the genome (112,113). This reaction produces reactive 3'OH groups at the ends of the viral DNA, which are required for successful strand transfer. HIV-1 IN remains bound to both LTRs, as well as other viral proteins such as MA, Vpr, NC and RT as well as several host proteins such as LEDGF, Hsp60, BAF-1, HMG-A1 and others to form the pre-integration complex (PIC) (114-118). Of these, LEDGF has been most extensively studied, and has been co-crystallized with HIV-1 IN (119). A recent crystal structure of the related prototype foamy virus (PFV) integrase in complex with its substrate has also shed light on both the structure of HIV integrase and its catalytic mechanism (120,121). The strand transfer activity of IN occurs after nuclear import, when IN is bound to both viral and host DNA, and uses the 3'OH ends of the viral DNA for nucleophilic attack on the host DNA at the site of integration. Each viral DNA end attacks the host DNA along the major groove, and as such a 5 base pair, single stranded gap is introduced, as well as a 2 bp flap from the 5' overhang of the viral DNA (122,123). Cellular repair enzymes then fill in the gap and trim the flap to complete the integration step (124). The development of potent antivirals against HIV-1 IN have been quite successful, especially against the strand transfer activity of IN. Active site inhibitors of this function chelate metal cofactors in the active site in a manner very similar to the RNase H inhibitors described later in this thesis. In fact, many small molecule inhibitors are dual inhibitors of both IN strand transfer and RT-associated RNase H activity (125).

The third protein cleaved from the Gag-Pol precursor is reverse transcriptase (RT). RT and its associated activities as well as its inhibition is the topic of this thesis, and will be introduced in more detail elsewhere (see Section 1.3 for an introduction).

Env

The *env* gene in HIV-1 expresses a polyprotein precursor called gp160 from a bicistronic mRNA also containing *vpu*. This polyprotein is targeted to the membrane rough endoplasmic reticulum (RER) where it is glycosylated concomitant with translation (126,127). Monomers of gp160 oligomerize in the ER into trimers, which is thought to simplify trafficking of gp160 to the Golgi, where significant modifications the oligosaccharides, as well as proteolytic cleavage of gp160 into the surface unit (SU or gp120) and transmembrane unit (TM or gp41) takes place (128,129). After cleavage, gp120 and gp41 remain associated, as 3 molecules of each form a heterotrimeric HIV-1 glycoprotein spike. These spikes are then shuttled to the plasma membrane, where HIV assembly takes place. Env is rapidly recycled at the cellular membrane due to endocytosis, and gp120 is commonly shed due to the relatively weak interaction between gp120 and gp41 (130,131). Both of these factors contribute to a decreased presence of Env in viral particles than there otherwise would be. This is thought to help HIV-1 virions escape the host immune system as well as reduce virus-induced cytopathic effects (10).

Env is responsible for the spread of the virus, through one of two mechanisms: budding, where virus particles are released into the extracellular space to infect new cells, or through cell-cell contact at places called virological synapses, the latter being far more efficient (132). Contacts between HIV-1 Env and its receptors appear necessary for

virological synapse formation in T-cells, but not in macrophages (133,134). The primary host receptor for Env is CD4, which is required for virus entry and one of two co-receptors: CCR5 and CXCR4.

1.2.2 Regulatory Proteins – Tat and Rev

Transcriptional Transactivator (Tat)

HIV-1 encodes a small 14 kDa protein called Tat that is required for expression of viral genes after proviral integration. During initiation of transcription, Tat binds to a 59 bp RNA leader sequence called TAR (transactivation response element) located in the 5' LTR of the provirus and forms a part of the HIV-1 promoter sequence (135). This TAR RNA self-folds into a specific stem-loop structure that is recognized by Tat (136). Tat recruits RNA polymerase II (RNAPII) to TAR RNA and upregulates transcription of the HIV-1 genome (137). Tat also interacts with many host factors as well as upstream regulatory elements (138). After transcription, Tat also binds directly to the capping enzyme and stimulates the co-transcriptional capping of the nascent viral mRNA (139).

The HIV-1 provirus is organized into a compact chromatin structure after integration. This structure does not provide space for the recruitment of transcription factors and other enzymes, resulting in low-level transcription. Tat is responsible for recruiting chromatin remodeling enzymes including histone acetyltransferases (HATs) and others that result in transcriptional transactivation and increased expression of viral genes (140). Tat can also modify cellular gene expression through a similar mechanism, which allows Tat to control various cellular pathways and functions. One cellular pathway that Tat is directly

involved in is apoptosis. Tat upregulates a variety of pro-apoptotic and immunosuppressive factors such as TGF- β , TRAIL, Bcl-2 family proteins such as Bim and Bax, FasL, caspase-8 and others (141-146). This is critical to the depletion of CD4+ T cells and contributes significantly to HIV pathogenesis and progression to AIDS. Tat can also be secreted and taken up by other cells (147). Tat can affect many cellular processes when taken up by an uninfected cell, such as inducing CCR5 and CXCR4, which promotes infection by HIV-1 (148).

Rev

Rev is an 18 kDa protein that is responsible for the nuclear export of unspliced and singly-spliced viral mRNAs. There are 3 sizes of HIV-1 mRNAs. There is a fully spliced ~2 kb transcript that encodes the early genes such as Tat, Rev and Nef. This ~2 kb transcript is able to cross through the nucleopores into the cytoplasm unaided. There is also a ~4 kb transcript which includes Env, Vif, Vpu and Vpr in addition to Tat, Rev and Nef. Finally there is an unspliced ~9 kb transcript which finally contains Gag and Pol in addition to the others mentioned. The ~4 kb and ~9 kb transcripts require Rev in order to reach the translation machinery in the cytoplasm. Rev accomplishes this by binding to the Rev Response Element (RRE), a 351 bp sequence located in the Env encoding region (149). Rev functions as a multimer, as monomers have been shown to be defective in RNA export (150). Rev has also been shown to activate gene expression of hepatitis C virus (HCV) genes in co-infected individuals (151).

1.2.3 Accessory Proteins: Nef, Vpu, Vpr (Vpx), and Vif

Nef

Nef (negative factor) is a ~27 kDa myristolated protein that performs a wide variety of tasks within an infected cell. It could be said that Nef is responsible for generally making an infected cell a safer and more efficient place to produce new virions. In infected T cells, Nef downregulates CD4, MHC-1, CD28 and CXCR4 from the cell surface (152). This helps to hide infected T cells from cytotoxic T lymphocytes (CTLs) and may help the release of virions from the cell surface. In addition, Nef downregulates HLA-A and –B alleles from the cell surface, but not HLA-C and –E which helps to strike a balance between hiding from CTLs and attack by NK cells (153). Nef also modifies diverse cellular pathways that upregulate cellular transcription factors, some of which interact directly with Tat such as NF- κ B and AP-1 which increase transcription of viral genes (154,155). There is also evidence that Nef may cause the secretion of cellular factors that attract CD4⁺ T cells, providing more targets to infect (156,157). In fact, Nef itself may be secreted and may induce apoptosis in bystander T cells, providing a direct effect on HIV-1 disease progression and T cell depletion (158,159).

Vpu

Vpu (viral protein u) is an ~18 kDa protein that is expressed with Env. Viral Vpu has two main roles: promote the degradation of CD4 and antagonize tetherin (BST-2) to promote viral release. Firstly, in the ER, newly synthesized Env is capable of binding intracellular CD4, and these complexes can be trafficked to the cell surface and inserted into virions,

which would have Env molecules that would be unable to bind to CD4 on a new cell. In order to prevent this, Vpu mediates the polyubiquitinylation and proteasomal degradation of CD4 at the ER and reduces the formation of CD4-Env complexes (160,161). The second function of Vpu is to antagonize tetherin, a host restriction factor that is strongly induced by interferon- α responses. Tetherin is a dimeric protein that anchors nascent virions to the cell surface with one subunit anchored in the cell membrane, and the other in the viral membrane, preventing release of the virus (77,78). Vpu interacts with tetherin and directs it toward the trans-Golgi network, resulting in degradation of tetherin by a process that is not yet well understood (162). This has the effect of reducing the amount of tetherin at the cell surface and promoting virus release. Vpu has also been implicated in apoptosis by activation of the caspase pathway and by inhibiting NF- κ B, which regulates the expression of anti-apoptotic genes (163).

Vpr and Vpx

Vpr (viral protein R) is an ~18 kDa protein whose role appears to be arresting the cell cycle in the G2 phase through suppression of mitotic CDK (cyclin-dependent kinase) (164). Preventing the cell from dividing is believed to allow more time for viral protein synthesis and efficient virus production. Vpr can also be secreted and taken up by bystander cells that will also arrest in G2 phase. Prolonged arrest at G2 can result in apoptosis, in which Vpr has been implicated (165). In non-dividing cells such as macrophages, access to the nucleus is restricted to factors about 40 kDa and smaller (166). However, the PIC must localize to the nucleus, and it is much larger than 40 kDa.

Vpr plays an important role in the nuclear transport of the PIC, and as such allows lentiviruses such as HIV-1 the capacity to infect non-dividing cells (167). In fact, Vpr is responsible for allowing HIV-1 to replicate in macrophages, one of HIV-1's main targets (168). In HIV-2 and SIV_{SM} (sooty mangabey) the vpr gene appears to have duplicated, giving rise to the very similar Vpx protein (169). In these viruses, Vpr is responsible for G2 arrest, while Vpx is involved in nuclear transport of the PIC. As discussed above, both of these tasks are carried out by Vpr in HIV-1 (170). It was recently discovered that Vpx overcomes cellular restriction of HIV-1 in dendritic and myeloid cells by inducing proteasomal degradation of the newly discovered restriction factor SAMHD1 (171).

Vif

Vif (virion infectivity factor) is a ~ 23 kDa protein that has evolved as a solution to the host restriction factor APOBEC3G (A3G) (172). In the absence of Vif, A3G is packaged into virions via interactions with viral RNA or Gag. A3G is an RNA-editing enzyme, capable of deaminating C to U during (-)-strand DNA synthesis, which results in a G to A hypermutation in the HIV genome, resulting in error catastrophe of the genome and non-infectious viruses (173-175). Vif binds to A3G and recruits a ubiquitin ligase that marks both Vif and A3G for polyubiquitinylation and proteasomal degradation (176). There are several APOBEC genes on chromosome 22, of which 3G is the most active against HIV, although A3F is also quite active (175). Vif appears able to degrade both A3G and A3F, and Vif mutations have segregated those abilities to separate parts of Vif, suggesting that

the adaptation of HIV-1 to humans required Vif to independently degrade at least two different types of APOBEC molecules simultaneously (177).

1.3 HIV-1 Reverse Transcriptase Structure and Function

The discovery of a RNA-dependent DNA polymerase in 1970 in rous sarcoma virus was a major breakthrough in molecular biology and it was aptly named “reverse transcriptase” in reference to the central dogma of molecular biology, which stated that genetic information travelled in one direction only, from DNA to RNA (178,179). Reverse transcriptases (RT) occur in all members of the family Retroviridae, and catalyze the synthesis of double-stranded DNA (dsDNA) from the (+)-sense, single-stranded RNA (ssRNA) genome, called reverse transcription (Fig. 1.5) . HIV-1 RT contains two distinct active sites: a polymerase active site that catalyzes the incorporation of nucleotides into the nascent DNA strand, and a ribonuclease (RNase) H active site, that is responsible for the degradation of DNA/RNA hybrids that occur as intermediaries in the reverse transcription process. In addition to non-specifically degrading the RNA template strand during reverse transcription, the RNase H activity of HIV-1 RT must also make sequence-specific cleavages in order to create the (+)-strand or PPT primer, used to initiate (+)-strand DNA synthesis, as well as specifically removing that same primer once (+)-strand

synthesis has progressed far enough to no longer require the RNA portion. Also, the RNase H activity of HIV-1 RT must make a specific cleavage to remove the (-)-strand or PBS primer, in order for the integration complex to successfully recognize the product of reverse transcription. The two active sites are spatially separate on the enzyme. RT can accommodate 18 base pairs (bp) of DNA/DNA between the active sites, and 19 bp of DNA/RNA (180,181).

HIV-1 RT is heterodimeric enzyme consisting of a large subunit of 560 amino acids (p66) and a smaller subunit of 440 amino acids (p51) (Fig. 1.6). Both subunits are derived from the Gag-Pol polyprotein precursor by cleavage with the viral protease (PR). Despite their differences in length, the p66 and p51 subunits are identical in sequence, as the p51 subunit is derived from the cleavage of the p66 subunit by the viral PR. Both subunits contain polymerase (containing the fingers, palm and thumb subdomains) and connection domains, with the RNase H domain included only on the p66 subunit, absent from p51. Despite the similarities between the two subunits, in the active heterodimer they are folded quite differently, with the p66 subunit containing both active sites, and providing the majority of enzyme-nucleic acid contacts. The p51 subunit is folded more compactly, with the polymerase active site internalized and not accessible to solvent. This has led to the belief that p51 plays a solely structural role. The thumb subdomain of p51 appears to stabilize the position of the RNase H domain of p66, as deletions in the carboxy-terminus of p51 altered the positioning of the p51 thumb, and affected the cleavage specificity of the RNase H activity (182). Also, it has been demonstrated that only wildtype p66/p51 heterodimers are capable of initiating (-)-strand DNA synthesis from a tRNA^{Lys3}-RNA

complex, while neither p66/p66 homodimers nor p66/p51 heterodimers with C-terminal p51 deletions were capable of (-)-strand initiation (183).

1.3.1 Substrate Binding to RT

The nucleic acid binding cleft is approximately 60 Å in length and is formed primarily by the p66 subunit. Specifically the polymerase domain, which comprises the fingers (residues 1-85 and 118-155), palm (residues 86-117 and 156-236) and thumb (residues 237-321) subdomains, as well as the connection domain (residues 322-440) and RNase H domain (residues 441-560) (184,185) (Fig. 1.6). The p51 thumb subdomain also contributes to nucleic acid binding. In crystal structures showing RT bound to DNA/DNA substrates, the DNA double-helix adopts a structure resembling A-form DNA structure near the polymerase active site, while adopting a more B-form structure near the RNase H active site, which is more characteristic of DNA/DNA hybrids (184-188). The shift in DNA helix structure from A-form to B-form geometry takes place during a 40-45° bend in the vicinity of the p66 thumb, that is characteristic of a wide variety of polymerases (57). A structure referred to as the “primer grip” is composed of β 12- β 13 hairpin (M230 and G231) and interacts with the phosphate of the terminal primer base (184). Another structure motif is the “template grip”, which consists of β 4 and α B of the fingers subdomain, β 8- α E loop, and β 5a of the palm subdomain that interacts with the template strand within the first 4 base pairs of the polymerase active site (184). Both the primer grip and template grip optimally position the nucleic acid substrate for DNA synthesis.

Interactions between RT and a DNA/RNA primer/template are similar to DNA/DNA, with minor differences. RT makes extensive contacts with 2'-OH groups on the RNA template, that do not exist with a DNA template, supporting the observations that RT is both more processive and shows increased polymerase activity on DNA/RNA than DNA/DNA (57). Furthermore, the dissociation constant (K_d) which is a measure of equilibrium binding affinity is lower with DNA/RNA substrates than DNA/DNA, representing tighter binding (189-193). Ehteshami *et al.* showed that K_d values could be derived for both the polymerase and RNase H active sites on a DNA/RNA substrate, and observed tighter binding at the polymerase active site than at the RNase active site (194). This is supported by the observation that most enzyme-substrate contacts occur in the polymerase domain. The p51 subdomain has an increased role when RT is bound to DNA/RNA compared to DNA/DNA. P51 residues K395 and E396 are involved in binding both DNA/DNA and DNA/RNA substrates, making contacts with the DNA primer, while K22 and K390 are involved only in DNA/RNA, making contact with an RNA template, and not DNA (57).

The structure of the DNA/RNA hybrid bound to RT is quite similar to the DNA/DNA structure, with the 40° bend in the hybrid occurring at the same place regardless of the chemical nature of the template. However, instead of canonical B-form DNA upstream of the bend, a conformation closer to an intermediate A/B-form hybrid is observed (57). This is consistent with observations that show unliganded DNA/RNA conforms to a hybrid structure called H-form (195-197). Sarafianos *et al.* have published the only crystal structure of HIV-1 RT in complex with a PPT sequence DNA/RNA hybrid (57).

An unusual feature of this structure is what the authors refer to as the “unzipping of the PPT”. In this structure, this refers to the melting of two base pairs (primer bases 13 and 14, with 1 being the 3’ primer terminus), followed by an unpaired template base (template base 15), followed by a frameshifted A-T pair (template base 16, primer base 15) then a G-T mismatch (template 17, primer 16), followed by an unpaired primer base (primer 17) bringing the hybrid back into register and proper Watson-Crick base pairing. This is immediately followed by the scissile phosphate, then 4 bp of weakly paired bases, held together mostly by stacking interactions. Unfortunately, without another crystal structure of RT in complex with a non-PPT DNA/RNA hybrid, it is difficult to determine whether this “unzipping of the PPT” is a common feature of DNA/RNA hybrids bound to RT, or whether it is specific to the PPT sequence, and may even play a role in the resistance of the PPT sequence to RNase H degradation.

Another structural feature is the RNase H primer grip. Similar to the primer grip motif in the polymerase domain, the RNase H primer grip motif is responsible for aligning the bound DNA/RNA hybrid on the proper trajectory for RNase H cleavage at the active site. The RNase H primer grip is located near the RNase H domain in contact with the primer strand (positions -10 to -15), comprising residues T473, A360, G359, H361, I505, Y501, K476, and Q475 in the p66 subunit, as well as K395 and E396 in the p51 subunit. The RNase H primer grip is mostly conserved among other RNases H including *E. coli* and MLV (198,199). Mutations in the RNase H primer grip can cause deficiencies in RNase H activity as well as cleavage specificity, and in some cases decreases in viral titer (200-203).

1.3.2 RT Polymerase Structure and Function

The polymerase active site of HIV-1 RT is located in the palm subdomain, while the fingers and thumb subdomains wrap around the bound nucleotide. The active site is composed of 3 highly conserved catalytic aspartic acid residues (D185, 186 and 110) that coordinate two divalent metal ions (Mg^{2+} appears to be the relevant ion *in vivo*) (204).

D185 and D186 are part of the YXDD motif that is conserved among all reverse transcriptases (in HIV-1, X is methionine) (205). Other residues that form the dNTP binding site and are also conserved are R72 and K65 that bind the β and γ -phosphates of the bound dNTP, respectively (188). Also, Y115 which contacts the sugar ring of the dNTP and has been shown to act as a “steric gate” that allows for binding of deoxyribonucleotides and ribonucleotides by occupying the space that would be required for the 2'OH of NTPs (206). Finally, Q151 interacts directly with the 3'OH of the bound dNTP (188).

The biochemical mechanism of polymerization is quite well understood, and is reviewed in detail here (207,208). The polymerase reaction is started by the binding of a dNTP to the N site or nucleotide binding site. RT requires both a template and a primer with an exposed 3'OH group bound in the P or priming site. Once bound, a rate-limiting conformational change brings the fingers subdomain down towards the thumb, trapping the bound nucleotide and bringing the 3'OH, α -phosphate of the nucleotide and the polymerase active site aspartic acid residues in line for catalysis (191). Then a phosphodiester bond is formed between the 3' primer end and the incoming nucleotide,

increasing the length of the primer by one bp. The fingers subdomain then returns to its original position, releasing a molecule of pyrophosphate, the product of the polymerase reaction.

RT must then translocate one bp downstream, to move the new 3' primer end from the N-site to the P-site, in order to free the N-site for a new dNTP to bind. This process of translocation is a passive, thermodynamically driven event and RT is able to translocate back and forth with the 3' primer end in the N-site (pre-translocation) and the P-site (post-translocation) (209). The incoming dNTP can only bind to a post-translocated complex, trapping it in this conformation until it is catalytically added to the growing DNA chain, allowing the process to begin again. This is referred to as a Brownian ratchet mechanism of polymerase translocation (Fig. 1.7) (208).

1.3.3 RNase H Structure and Function

RNase H functions as an endonuclease that specifically cleaves the RNA moiety of RNA/DNA hybrids. In the case of HIV-1, RNase H non-specifically degrades the (+)-strand RNA genome, while specifically removing the (-)-strand tRNA primer and creating and removing the (+)-strand or PPT primer during reverse transcription. RNases H exist both as a free enzyme (*E. coli* RNase H1) and as a domain in a larger enzyme as in HIV-1 RT. The structure of both remains highly conserved, albeit with some crucial differences. In the case of HIV-1 RT, the RNase H domain constitutes the C-terminus of the p66 subunit, which is missing in p51, due to cleavage by the viral protease (57,184,186,188,210). The folded structure of the HIV-1 RNase H domain takes the form

of a 5-stranded mixed beta-sheet flanked by four alpha helices in an asymmetric distribution (210). The structure is homologous to other retroviral RNases H such as murine leukemia virus (MLV) (211) and avian sarcoma leukemia virus (ASLV), and both prokaryotic (*E. coli*, *B. halodurans*) (199,212,213) and eukaryotic (human) (214) RNases H, and is part of the superfamily of polynucleotidyl transferases. A notable difference between the various RNase H proteins is the presence or absence of the C-helix (present in *E. coli*, MLV and human RNases H, absent in HIV-1, *B. halodurans* and ASLV RNases H), a positively charged alpha helix also referred to as the basic loop, protrusion, or handle, and is thought to aid in substrate binding. Interestingly, deletion of the basic loop in *E. coli* inhibits but does not abolish activity (215), although the isolated RNase H domain of MLV with the basic loop deleted is not active (211). However, if the basic loop in *E. coli* RNase H is inserted into an isolated inactive HIV-1 RNase H domain, Mn^{2+} -dependent activity is partially restored (216). In HIV-1 RT, the connection subdomain contains a small loop (residues K353 to T365) that contains several basic residues and is structurally located at the exact position of the *E. coli* and MLV basic loops, and is thought to compensate for the lack of the C-helix (211,214) since RNase H activity was restored to an inactive isolated domain with the addition of the p66 connection domain *in trans* (217). Mutations to a family of human RNases H have been shown to cause a rare neurological disorder called Aicardi-Goutières Syndrome, and as such, it is important for potential HIV-1 RT-associated RNase H inhibitors to be highly selective for the viral, not human RNase H (3).

The active center of the HIV-1 RNase H domain contains a highly conserved DEDD motif (D443, E478, D498, D549), which coordinates two divalent cations required to hydrolyze the RNA substrate. Magnesium is likely the physiologically relevant ion; however, HIV-1 RNase H will tolerate manganese, cobalt, and other cations. Although crystal structures of HIV-1 show one Mg^{2+} ion in the RNase H active site (188), more recent structures of the more closely related *B. halodurans* and human RNases H show two magnesium ions (213,214), which is also supported by biochemical evidence of related enzymes (218). This has led to the general acceptance of a two-metal ion mechanism for retroviral RNase H hydrolysis (Fig. 1.8) (213,214,219). In brief, a two-metal ion mechanism requires that metal ion A activates a water molecule as a nucleophile and moves towards ion B, bringing the nucleophile in close proximity to the scissile bond, while metal ion B destabilizes the substrate-enzyme interaction and lowers the energy barrier to product formation (220). Ions A and B are involved in the stabilization of the transition state and product release. In order for hydrolysis to occur, the metals ions are likely coordinated at a distance of 3.5 to 4 Å from each other, possibly with some degree of flexibility (Fig. 1.8) (220). This scenario is exploited by small molecules used to inhibit RNase H activity.

The homologous HIV-2 RT shows markedly reduced RNase H activity (10-fold). This discrepancy in activity has been shown to be due to a single residue (Q294) in the catalytically inactive p54 subunit, which is the structural equivalent of the p51 subunit in HIV-1 RT (221,222). Mutagenesis of Q294 to P294 as in the WT HIV-1 RT shows an increase in RNase H activity comparable with that of

HIV-1 RT (221). Interestingly, all other mutations that have been tested at that position have also increased the RNase H activity of HIV-2 RT(221). Q294 is highly conserved in HIV-2 isolates, as well as the related SIV_{smm} model (221). P294 is seen in approximately 80% of sequenced isolates of HIV-1, while Q294 appears only about 1% of the time in HIV-1 and the recombinant mutant shows reduced RNase H activity (221). Although a structural or mechanistic explanation so far remains elusive, it is interesting to note that the p51 thumb is implicated in inhibitor binding in HIV-1.

HIV-1 RNase H has two major distinct modes of activity, namely polymerase-dependent and polymerase-independent. RT will position itself with the polymerase active site at the 3' end of the DNA primer, ready to accept an incoming nucleotide if given the opportunity. Cleavage in this situation will result in a cut on the RNA template, 18 base pairs (bp) upstream from the polymerase active site and the properly positioned 3' end of the primer. Per definition, this specific cut is known as polymerase-dependent RNase H activity (180,181,223-226). In the presence of deoxyribonucleotides (dNTPs) RT will advance and RNase H will eventually continue to cut the template; however, the rate of RNase H cleavage is approximately seven times slower than polymerization (191). In this sense, the polymerase and RNase H activities are spatially, but not temporally coordinated during polymerase-dependent activity. There are two structurally distinct modes for RT to bind a primer/template substrate that depend on the translocational status (208,209). Polymerase-dependent RNase H cleavages therefore will exist as a population of two cleavages, each one representing either the pre- or post-translocational cleavage(227). Binding of the next templated nucleotide will trap the enzyme in the post-

translocational conformation, and binding of the pyrophosphate (PPi) analogue PFA (phosphonoformic acid) traps the complex pre-translocation (228-230). By trapping the enzyme in a pre- or post-translocational position, it has recently been shown that RT can simultaneously engage the nucleic acid substrate at both the polymerase and RNase H active sites (227).

RNase H activity where the 3'-end of the DNA primer is not occupying the polymerase active site is referred to as polymerase-independent RNase H activity. Polymerase-independent cleavages can occur internally, or they depend on the precise positioning of the 5'-end of RNA fragments relative to the polymerase active site(231). The former refers to cleavages on long stretches of RNA/DNA that are too far from any terminus to be directed by nucleic acid ends, while the latter refers to cleavages that are directed by the 5'-end of the RNA template. It appears that these internal cleavages are affected by the nucleic acid sequence surrounding the cleavage site, as certain nucleotide positions in the vicinity of the scissile phosphate show a preference for specific nucleotides, although exactly what structural effect these sequence preferences may have and how it specifically directs RNase H cleavages is as yet unknown(232). The other form of polymerase-independent RNase H cleavage is called 5'-end directed cleavage, and occurs on fragments of RNA recessed on a longer DNA strand(233). The polymerase domain associates with the substrate near the 5' end of the RNA. Cleavages of this type occur in a "cleavage window" of approximately 13-19 nucleotides from the 5' termini of the RNA fragment(234).

Recent evidence suggests that the sequence specificity for internal RNase H cleavages are mostly shared by polymerase-dependent and –independent cleavages, including 3' DNA- and 5' RNA-end directed cleavages. Essentially all RNase H cuts show a preference for certain nucleotides at positions -4, -2 and +1 (where the scissile phosphate is between positions -1 and +1, and negative and positive positions proceed 5' and 3' from the scissile phosphate, respectively)(235). Whether this preference is specifically for the RNA base or DNA base is unknown, as is the mechanism by which this preference occurs.

RT makes the decision of whether to bind as a polymerase or an RNase H based on the structure and composition of the nucleic acid substrate itself(180,236), and on the presence and absence of ligands. As previously stated, a recessed DNA 3' primer end will be bound by the polymerase whether or not the template is DNA or RNA, whereas an RNA fragment recessed on a DNA template will be degraded with the polymerase domain associating with the substrate near the 5' RNA end. Chimeric primers *i.e.*, RNA-DNA/DNA hybrids are recognized differently by RT depending on the relative length of both the RNA and DNA portion of the primer(180). The part of the primer that is in contact with the C-terminal portion of RT near the RNase H domain is instrumental in determining the binding orientation of the enzyme, and thus the preferred activity *i.e.*, polymerization or RNA hydrolysis(236). The single most important nucleotide in determining the activity of RT on a chimeric RNA-DNA primer is the primer nucleotide in contact with the RNase H primer grip, in particular residues T473 and Q475(236).

1.3.4 Role of RNase H in (+)-strand priming

A purine-rich sequence (5' AAAAGAAAAGGGGGG 3') called the polypurine tract (PPT) located just 5' to the U3 sequence of the genome is resistant to RNase H cleavage and remains intact after the rest of the viral RNA has been removed (RNase H makes a specific cut following the 6th "G" residue to define the 3' end of the PPT primer) (58,237-240). The 3'-end of the newly formed RNA primer is extended by RT to position +12 downstream from the PPT RNA. RT then pauses, re-orientates itself and makes a specific cleavage at the PPT-U3 junction (Fig. 1.9) (58). This cleavage is important as it defines the end of the U3 LTR, which is used as a substrate by the viral integrase. PPT primers with aberrant cleavages or that improperly begin DNA synthesis are not used efficiently and reverse transcription does not proceed (241). In order for RT to successfully initiate (+)-strand DNA synthesis, RT must bind as a polymerase to extend the primer and as an RNase H to cleave at the RNA-DNA junction. As the active sites of RT are simultaneously positioned on opposite strands of the nucleic acid, RT must recognize its position on the genome and dynamically change its binding orientation appropriately for the task at hand. Thus, the primer removal reaction requires polymerase-independent RNase H activity.

In general, RT does not efficiently extend RNA primers with two exceptions: the (-)-strand or tRNA^{lys3} primer, and the (+)-strand or PPT primer. The PPT is structurally distinct in several ways that contribute to its resistance to RNase H cleavage. The PPT has an unusually narrow minor groove, and due to extensive contacts between the DNA

primer and the RNase H primer grip, only substrates with a certain minor groove width appear to have appropriate access to the RNase H active site (57).

A co-crystal of RT together with an RNA/DNA PPT substrate shows that the trajectory of the substrate is “missing” the catalytic residues of the active site by about 3 Å, the same as the apparent narrowing of the minor groove. The PPT has other unusual structural features such as an “unzipping” of the primer/template just downstream of the RNase H active site. An unpaired base on the template is followed by a frameshifted A:T pair and a mismatched G:T pair, followed by another unpaired base on the primer strand to restore the register. The next base pair is adjacent to the scissile phosphate, and together with the next 3 base pairs upstream are in frame, although the bases are more widely separated. This could also affect RNase H cleavage. It has also been hypothesized that during polymerization, the stiffness of the A-tracts in the PPT combined with the flexibility of the G:C base pair separating them could cause pausing of RT as it attempts to induce a 40° bend characteristic of nucleic acids bound to a wide variety of polymerases (including RT), which could allow time for RT to re-orient itself into an RNase H-competent mode and cleave at the PPT-U3 junction (57). However, more research is required to confirm this hypothesis. Recent evidence has suggested that the initiation of (+)-strand synthesis is preferentially inhibited by non-nucleoside analogue RT inhibitors (NNRTIs) that bind in close proximity to the nucleotide binding site, highlighting the importance of (+)-strand initiation in the development of potent antivirals (242).

1.3.5 Role of RNase H in strand transfer and (-)-strand primer removal

Reverse transcription in retroviruses is initiated by cellular tRNAs that are packaged into the virion during assembly, sometimes along with the corresponding tRNA synthetase.

For HIV-1, human tRNA^{lys3} base pairs with its 18 3'-terminal residues with the primer binding site (PBS) near the 5' end of the viral genome forming an RNA/RNA duplex.

This primer is extended to the 5' end of the genome producing the minus strand strong-stop DNA ((-)ssDNA). The (-)ssDNA must now dissociate from the genomic template and re-associate either with repeat (R) sequences at the 3' end of either the same genome (intramolecular strand transfer) or the other copy of the genome present in the virion (intermolecular transfer), with both types of transfers happening at approximately equal frequency (21,59). In order for the aforementioned dissociation of the (-)ssDNA occur, the RNA template must be degraded by RNase H cleavage. Models of the strand transfer reactions have been reviewed elsewhere (21).

The rates of strand transfer and RNase H activity are well correlated. Experiments with RNase H-deficient RT enzymes show that (-)ssDNA is formed, but strand transfer never occurs. Similarly if RNase H activity is moderately inhibited, strand transfer is also inhibited. The genomic template must be degraded to provide an available DNA sequence for an invading RNA template. If polymerization is arrested, with chain terminating nucleotides for example, strand transfer rates increase (243). This is likely due to the relatively slower RNase H activity having enough time to make more cleavages due to the pausing of the enzyme, and this in turn allows dissociation of cleaved genomic fragments and invasion by a new template (244). This type of pausing-induced RNase H activity is

polymerase-dependent. While other RT enzymes not involved in polymerization can bind and cleave RNA fragments in a 5'-end directed or internal cleavage mode, they are unlikely to be affected by pausing of the polymerizing RT. The question is then raised of whether strand transfer requires polymerase-dependent cleavage or polymerase-independent cleavage, or both? *In vivo* experiments with MLV have demonstrated that a mutant RT enzyme with RNase H activity eliminated and also diminished polymerase activity does not allow for strand transfer, as expected. However when that system is supplemented with a polymerase-negative mutant with a fully functional RNase H (pol(-)/rnh(+)), strand transfer is only marginally increased (245). It is assumed by the authors that the pol(-)/rnh(+) mutant is participating in mostly polymerase-independent RNase H activity, as the polymerizing RT will be occupying the 3'-end of the DNA primer, without being able to contribute any polymerase-dependent RNase H activity. The interpretation of these data by the authors suggests that in the absence of polymerase-dependent RNase H activity, polymerase-independent activity is inefficient. Also, that in the context of (-)ssDNA strand transfer, both polymerase-dependent and – independent activities are required for completion of reverse transcription (245). However, as noted by the authors, it is possible that a pol(-)/rnh(+) RT is binding occasionally at the 3'-end of the primer and providing polymerase-dependent cuts. Telesnitsky and Goff showed that MLV can successfully replicate when the polymerase and RNase H activities are provided on different enzymes, but again this system cannot completely discriminate between polymerase-dependent and –independent cuts, since polymerase-negative mutants could bind to the 3'-end of the primer and induce

polymerase-dependent cuts (246). It is therefore unclear whether both modes of RNase H activity are required for successful strand transfer, however they do appear to have slightly differing roles.

In HIV-1 the tRNA primer is incompletely removed, with one rA:dT base pair remaining attached to the U5 terminus of the double-stranded DNA (59,247). This rA is one of two bases removed during the 3'-end processing activity of the viral integrase. In most retroviruses, the sequence of the U5 or PBS regions do not affect cleavage specificity, however in HIV-1, the U5 sequence (G immediately 3' of the conserved CA dinucleotide) is pivotal in determining RNase H cleavage specificity. Integrase has been shown to recognize diverse structures as substrates for 3' processing and strand transfer (248). It can remove both one and three nucleotides adjacent to the conserved CA at the end of U5, however it is not efficient. This is perhaps not surprising since the related RNase H from HIV-2 normally removes 3 nucleotides before the CA dinucleotide (249). Studies suggest that mutation of the CA dinucleotide results in DNA ends that are inefficiently used by integrase and reduce viral titer *in vivo* (249). However, due to the flexibility of integrase in its ability to use substrates of various lengths for 3'-end processing, it remains unclear if the disruption of cleavage specificity by RNase H by small molecules or other techniques is practically useful as a therapeutic option.

1.3.6 Role of RNase H activity in drug resistance

All established drugs that target HIV-1 RT have binding sites either at the polymerase active site, or the non-nucleoside binding pocket, a hydrophobic depression created by the binding of non-nucleoside reverse transcriptase inhibitors (NNRTIs) and located near the base of the p66 thumb subdomain, about 10 Å from the polymerase active site (185,186,250). It is therefore not surprising that most of the major known resistance-conferring mutations are also located in the vicinity of the polymerase active site.

However, recent studies have shown that mutations in the connection and RNase H domains can affect the susceptibility of RT to NNRTIs and nucleos(t)ide reverse transcriptase inhibitors (NRTIs). Furthermore, RNase H activity itself is implicated in the mechanism of resistance to NRTIs such as

3'-azidodeoxythymidine (AZT), and also to NNRTIs such as nevirapine (NVP). For example, the NNRTI resistance mutation V106A showed a reduction in both 3'-end and 5'-end directed RNase H activity, which correlated with a decrease in viral fitness (251). In contrast, the NNRTI-resistance mutant Y181C demonstrates an increase in RNase H activity, while still demonstrating decreases in viral fitness levels despite no change in polymerization processivity compared to wildtype(251). AZT also selects for other connection/RNase H domain mutants such as A360I/V, A371V and Q509L (252-258).

The connection domain mutant N348I has been found in clinical isolates with treatment regimes that include both AZT and NVP (258), becoming the first resistance mutation selected outside of the polymerase domain that confers dual-class resistance. It was proposed that when RNase H activity degraded the RNA template of an AZT-terminated

primer, the chain-termination became permanent, as the template fragments would dissociate and the single-stranded AZT-terminated primer is not a substrate for the excision machinery. Therefore, decreased RNase H activity could act to increase the half-life of the template, allowing more time for RT to excise AZT at the primer terminus (252,259-263). This hypothesis was supported by a study that linked decreased RNase H activity to increased rates of excision (258), and Ehteshami *et al.* provided a plausible biochemical mechanism (264). As N348I and A360V often appear along with other thymidine analog mutations (TAMs) in the clinic, they both show resistance to AZT. RT enzymes containing N348I and A360V reduce RNase H activity by selectively dissociating from RNase H-competent (polymerase-independent) complexes (264). N348I had no effect on substrate binding of the polymerase domain, and A360V rescued a modest binding defect introduced by TAMs, which agrees with the observation that A360V appears later in therapy after the emergence of TAMs and N348I (264). The A360V and N348I mutations also appear to increase processive DNA synthesis, which points to an additional RNase H-independent contribution to resistance (264).

The RNase H mutant Q509L likely operates in a similar mechanism, although it appears only rarely in therapy-experienced patient (252,265,266). Also, G333D/E mutation is resistant to AZT and lamivudine (3TC) although it does not affect RNase H activity. This mutation as well as others such as E312Q, G335D, V365I, A371V and A376S that affect resistance to NRTIs and in some cases NNRTIs, have been reviewed (254,267).

1.4 Inhibition of HIV-1 RT

Anti-retrovirals (ARVs) comprise the frontline defense against HIV-1 infection since the discovery of the first ARV to get to market, AZT (zidovudine) in 1987. They are administered in a cocktail of at least 3 different ARVs in what is called HAART (Highly Active Anti-Retroviral Therapy) since 1996 and has significantly reduced mortality and morbidity in HIV infected individuals, turning a deadly disease with a life expectancy of 1 year to a manageable chronic illness (268). As of 2011 there are many ARV drugs against several viral targets such as RT, integrase, protease and Env. However RT remains the most popular drug target. As previously mentioned, RT carries two functions on two distinct active sites, both of which are absolutely required for viral replication. However, to date, only compounds that inhibit the polymerase activity of RT have been made available in the clinic. Although there are many inhibitory compounds described in the literature against the RNase H function of RT, none have been developed for clinical trials, or even have been able to select for drug resistance, an important marker for specific inhibition of a target. This section will discuss common ARVs that act against HIV-1 RT, with an emphasis on the RNase H inhibitors and the progress towards the first RNase H drug used in the clinic.

1.4.1 Polymerase Inhibitors

NRTIs – Nucleos(t)ide-Analog Reverse Transcriptase Inhibitors

NRTIs are molecular mimics of natural nucleotides that are incorporated by RT and act as chain terminators, in most cases due to the lack of a 3'OH required by RT to continue DNA synthesis. They are administered as unphosphorylated prodrugs to facilitate bioavailability and cellular uptake, and are subsequently phosphorylated into their triphosphate form by cellular kinases. First generation NRTIs include AZT (zidovudine, 3'-azido-2',3'-dideoxythymidine), 3TC (lamivudine, 2',3'-dideoxy-3'-thiacytidine), FTC (emtricitabine, β -L-3'-thia-2',3'-dideoxy-5-fluorocytidine), ABC (abacavir, (1*S*4*R*)-4-[2-amino-6-(cyclopropylamino)-9*H*-purin-9-yl]-2-cyclopentene-1-methanol sulfate), ddI (didanosine, 2',3'-dideoxyinosine), ddC (zalcitabine, 2',3'-dideoxycytidine) and d4T (stavudine, 2',3'-didehydro-2',3'-dideoxythymidine). (Fig. 1.10 and 1.11) Another drug classified as an NRTI but is a nucleotide analog is Tenofovir (TNF, tenofovir disoproxil fumarate) (Fig. 1.10). Some of these, such as ddC, is no longer used in the clinic, while tenofovir is one of the most frequently used drugs in first line therapy (269).

Due to the emergence of drug resistant mutations against all first-generation NRTIs, there is a need to develop new, second-generation NRTIs that are active against drug resistant strains. For example, elvucitabine is a novel NRTI with 10-20 fold greater activity than 3TC, and shows potent activity in patients harboring resistance to 3TC and other NRTIs (270). Racivir is another NRTI currently in phase II clinical trials that is active against 3TC resistant viruses (271). Apricitabine is a novel cytidine analog that is active against viruses containing mutations conferring resistance to multiple NRTIs, and resistance to

apricitabine develops slowly in vitro, and there is little evidence for the development of apricitabine resistance in clinical trial so far (272).

NNRTI – Non-Nucleoside Reverse Transcriptase Inhibitors

NNRTIs comprise a chemically diverse set of allosteric RT inhibitors that all bind to a hydrophobic pocket adjacent to the polymerase active site called the NNRTI binding pocket. Studies have shown that NNRTIs act predominantly at the step of (+)-strand initiation by affecting the ability of RT to bind to RNA/DNA primer/templates in a polymerase-dependent mode (242). First generation NNRTIs include nevirapine, efavirenz, and delavirdine. Newer, second generation NNRTIs include etravirine (TMC125) which is highly active against both WT and NNRTI-resistant viruses (273) and rilpivirine (TMC128) which has a higher potency, longer half-life and is better tolerated than earlier NNRTIs such as efavirenz (Fig. 1.12) (274) .

Foscarnet (PFA, or phosphonoformic acid)

PFA is a pyrophosphate (PPi) analog that inhibits the polymerase function of RT through a different mechanism than both NRTIs and NNRTIs, and as such will be discussed separately in this section (Fig. 2.1). Clinically, PFA is only used in salvage therapy with patients that are in late-stage HIV disease and have accrued multiple resistance mutations due to poor bioavailability and negative side effects (1,6,275). It is used in salvage therapy due to its resistance profile, which affects re-sensitizes resistant RT to AZT (14,15,78). For this reason, PPi analogs that are more bioavailable and less toxic are

highly sought after. PFA works by binding near the polymerase active site of RT and in a pre-translocated complex and stabilizes this complex, preventing RT from translocating, and thus, from binding the next nucleotide (228). For this reason, PFA appears to act as an RNase H inhibitor under steady state conditions (276). However, it was later shown that PFA stabilizes RT in the pre-translocational position, reducing enzyme turnover in steady-state systems, and has no direct effect on RNase H activity (277).

1.4.2 RNase H Inhibitors

Active Site RNase H Inhibitors

There are currently 25 compounds in clinical use to treat HIV (278). 12 of those 25 compounds target RT. Despite having two distinct enzymatic activities that are both absolutely necessary for successful replication of the virus, all of the 12 RT-targeting compounds block the polymerase activity of RT. This extreme bias of drugs towards the polymerase activity could be attributed in part to the relatively flat surface topography of the RNase H domain, unlike the polymerase domain which contains mobile subdomains and hydrophobic pockets providing a foothold for small molecules. In response to this, research has focused on the RNase H active site itself, in particular the two metal ions coordinated by the DEDD motif described earlier. This has led to several prototype compounds that chelate the Mg^{2+} ions through a 3-oxygen pharmacophore originally designed for inhibition of the influenza endonuclease (279), as well as several other compounds acting through a different mechanism. RNase H inhibitors that utilize a

cation-chelating mechanism as their mode of inhibition are referred to as active site inhibitors (218).

The first of such compounds designed for the inhibition of HIV-1 RNase H activity were N-hydroxyimides (Fig. 1.13) (218). These compounds were active against the isolated HIV RNase H domain and were highly selective as compared with *E. coli* RNase H (218,280). The metal ions were separated by a distance of 4-5 Å and chelated by three oxygen atoms, with the N-hydroxyl acting as a bridging oxygen contacting both metal ions (218,280). A crystal structure of the isolated RNase H domain with a bound N-hydroxyimide inhibitor in the presence of Mn^{2+} confirms the binding orientation and provides proof-of-concept for metal-chelating RNase H inhibitors (281).

Diketo acids also form a group of compounds initially developed against the structurally similar HIV-1 integrase (IN) (Fig. 1.13) (282,283). As such, the first diketo acid inhibitor developed showed similar inhibition of full-length RT, the isolated RNase H domain, and IN with IC_{50} values for all three in the low micromolar range, although it had no effect on the polymerase activity of RT (284). The diketo acid 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid (BTDBA), was shown to bind to the isolated HIV RNase H domain in a metal-dependent fashion by isothermal titration calorimetry (284). It also did not appear to require the presence of the nucleic acid substrate for efficient binding (284). The presence of the 3-oxygen pharmacophore suggests a similar mode of inhibition to the N-hydroxyimides. However, despite effectively inhibiting RNase H activity *in vitro*, this

compound did not show inhibition of viral replication in cell culture (284). Another diketo acid-based inhibitor, RDS1643, also specifically inhibited HIV RNase H activity while not affecting the HIV polymerase activity, or the RNase H activity of related enzymes such as *E. coli* and avian myeloblastosis virus (AMV) (285). Unlike other active site inhibitors, RDS1643 inhibited viral replication in MT-4 cells at essentially the same IC_{50} as the *in vitro* assays ($\sim 13 \mu M$) (285). Consistent with other diketo acids, binding of RDS1643 did not require the substrate for binding, but has a specific requirement for divalent cations. The mode of inhibition of RDS1643 has been determined to be reversible and non-competitive (285). Experiments with diketo acid inhibitors were instrumental in showing proof-of-concept that RNase H inhibitors could be synergistic with NNRTIs in a full reverse transcription assay. Curiously, the two inhibitors are antagonistic in an RNase H assay, and additive in an RNA-dependent DNA polymerase assay (286).

A third structural scaffold for the 3-oxygen pharmacophore is represented by the hydroxylated tropolones (Fig. 1.13). One of the first such compounds described, β -thujaplicinol is a natural product derived from the plant *Thuja plicata* and is active at submicromolar concentrations against HIV-1 and HIV-2 RNase H, precluding the NNRTI-binding pocket as the binding site (276). It is also a specific inhibitor of HIV RNase H, and is ~ 30 and ~ 250 -fold less effective at inhibiting human and *E. coli* RNase H1 respectively, and showed no inhibition of the polymerase activity of RT (276). However, β -thujaplicinol lacks antiviral activity in cell culture (276).

A recent study has introduced a fourth scaffold of the 3-oxygen pharmacophore, pyrimidinol carboxylic acid derivatives (Fig. 1.13)(287). Compounds of this class show IC_{50} values in the low to submicromolar range against HIV RNase H activity, and were selective for the viral enzyme compared to human RNase H1 (287). Order-of-addition experiments support a metal-dependent binding mechanism for these compounds, in agreement with experiments using β -thujaplicinol (227,287). One of the pyrimidinol carboxylic acids was successfully crystallized with Mn^{2+} and the isolated HIV RNase H domain (287). The carboxyl group coordinates metal ion B, while metal ion A (responsible for activating a water molecule for nucleophilic attack) is coordinated by the two phenolic oxygen atoms of the bound inhibitor (287). Furthermore, possible π -interactions between the inhibitor and H539 are observed in the crystal structure, although what effects mutagenesis at position 539 may have on inhibitor binding is unknown (287).

Allosteric RNase H Inhibitors

Despite the effort involved in the development of active site inhibitors, they are not the only compounds that have been described as RNase H inhibitors. Several other classes of compounds have been shown to inhibit HIV-1 RNase H and are characterized as allosteric inhibitors, as they do not bind at the RNase H active site. Most notably among these are the hydrazones and vinylogous ureas (Fig. 1.13).

N-acyl hydrazones were originally described as dual inhibitors of HIV RNase H and polymerase activities. The first of such compounds (N-(4-tert-butylbenzoyl)-2-hydroxy-1-naphthaldehyde hydrazone) or BBNH was observed to have metal ion-chelating abilities, while lacking the 3-oxygen pharmacophore of the active site inhibitors (288). Also, BBNH inhibited *E. coli* RNase H1 which, in contrast to the 2-metal ion mechanism of HIV RNase H, uses a 1-metal ion activation/attenuation mechanism, further distancing itself from the active site inhibitors (218,289,290). BBNH also failed to inhibit the RNase H activity of HIV-2 RT, which does not possess the NNRTI-binding pocket of HIV-1 RT (288). However, BBNH was effective at inhibiting NNRTI-resistance mutants in HIV-1 RT, suggesting two distinct binding sites for BBNH: one in the RNase H domain and one that overlaps with the NNRTI-binding pocket (288). It was also shown that BBNH impacts the stability of the p66/p51 heterodimer (291). Modeling studies have suggested that Y501 is important for BBNH binding (292). Mutants Y501W and Y501R were both resistant to BBNH-mediated inhibition of the RNase H activity of RT, but remained sensitive to inhibition of the polymerase activity, which strongly supports the two-binding site hypothesis (292). Other variations of BBNH inhibit only the polymerase activity of RT (4-t-butylbenzoyl)-2-hydroxy-1-salicylyl hydrazone (BBSH) while another derivative inhibited only the RNase H activity (4,N,N-dimethylaminobenzoyl)-2-hydroxy-1-naphthyl hydrazone (DABNH). Of these, BBSH also affected dimer stability like its predecessor BBNH(291). A crystal structure of yet another hydrazone derivative, ((E)-3,4-dihydroxy-N'-((2-methoxynaphthalen-1-yl)methylene)benzohydrazide) (DHBNH), shows the inhibitor bound near the NNRTI-binding pocket, in contact with residues

W229 and D186 (293). This is in agreement with one of the proposed binding sites for BBNH. However, DHBNH also weakly inhibits the isolated HIV RNase H activity although two binding sites are not observed in the crystal structure (293). It is possible that DHBNH shares the same two binding sites as BBNH. It is not known whether DHBNH impairs dimer stability. More research is required to elicit the exact binding sites of this class of RT inhibitor.

Another class of RNase H inhibitors that bind in an allosteric manner is the vinylogous ureas. These compounds are thought to bind near the p51 thumb subdomain, based on protein footprinting/mass spectroscopy experiments (294). The p51 thumb is in contact with the p66 RNase H domain, and mutations in this region have been known to affect RNase H activity (295,296). Mutations in the RNase H primer grip residues Y501 and T473 show differing effects (294). The mutant Y501Bp-Phe which represents the mutation of residue 501 from tyrosine to benzophenyl-phenylalanine (an unnatural amino acid), confers resistance to the compound NSC727447 (IC_{50} values 6.6 μ M for Y501 to 196.7 μ M for Bp-Phe501) (294). However, the mutation T473C increases sensitivity of the enzyme for NSC727447 by 50-fold, and showed a significantly reduced affinity of T473C for its substrate in the presence of NSC727447 (294). The effects of the inhibitor on primer grip mutants support the hypothesis that NSC727447 binds near the p51 thumb.

The development of potent allosteric inhibitors provides new options for potential inhibitor binding sites, such as the p51 thumb subdomain. This combined with the

development of potent active site inhibitors could result in a powerful new combination of therapeutic options in the treatment of HIV infection. The field of HIV RNase H drug development has made significant advances. The emergence of crystal structures with bound inhibitors are an encouraging development and more are urgently needed, especially including the bound nucleic acid substrate. The field is also patiently awaiting the first RNase H inhibitor with (potent) antiviral effects. The selection of resistance in cell culture is viewed as an important milestone in this regard.

Other compounds that affect RNase H activity

There have been many other compounds described in the literature that show apparent inhibition of RNase H activity, including natural products such as various quinones and naphthoquinone derivatives, although many of these suffer from a lack of potency (297,298). DNA and RNA aptamers have recently been shown to be potent and selective inhibitors of HIV-1 RT, and its RNase H function, in cell-free assays (299-301). Their utility in cell-based assays is currently under investigation.

RNA-based RNase H inhibitors can potentially compete with the nucleic acid substrate for the same contacts; however, the mechanism of action of small molecule inhibitors remains often elusive and their effects on RNase H cleavage can be indirect. For instance, the pyrophosphate analog PFA that binds to the polymerase active site was also shown to reduce RNase H activity under steady-state conditions (286). However, inhibition of RNase H cleavage is not seen under pre-steady-state conditions (227). PFA traps and

stabilizes the pre-translocated complex, which, diminishes the turnover of the reaction, which, in turn, translates into RNase H inhibition. Thus, the effect of PFA on RNase H cleavage is not direct and most likely irrelevant in biological settings. There is an excess of HIV-1 RT in the virion and frequent dissociation of the enzyme will also cause dissociation of PFA. Unliganded RT molecules may then bind to the nucleic acid substrate and RNase H cleavage would occur uncompromised. This is a fundamental difference to the aforementioned RNase H active site inhibitors that bind to free enzyme in the presence of divalent metal ions. It is therefore difficult to discern whether a given compound with apparent inhibitory effects on RNase H activity is a *bona fide* RNase H inhibitor unless it has been assessed in more detailed mechanistic studies. Pre-steady-state kinetics, order-of-addition experiments, and the use of polymerase active site mutants and/or the use of the isolated RNase H domain have been proven successful with respect to the characterization of RNase H active site inhibitors. The characterization of allosteric RNase H inhibitors requires in addition to these experimental approaches the determination of the binding site for the inhibitor.

1.5 Viral Drug Resistance

The HIV-1 reverse transcriptase is a relatively error-prone enzyme, which misincorporates a nucleotide approximately every 7000 bases on an RNA template, and every 6000 bases on a DNA template (302). With a high rate a replication and ample opportunity for strand transfer events during reverse transcription, RT bestows a very high genetic diversity to HIV-1. This genetic diversity is the basis of viral resistance to

HIV drugs. Although HIV drugs are very potent and specific, there exists at least one mutation in the viral quasispecies that is resistant to drug action, and is therefore strongly selected for in the presence of the drug. This is simply artificial selection of naturally-occurring drug-resistant viral mutants. Any drug-resistant virus likely comes at a cost of viral fitness, therefore the more mutations required to provide drug resistance, the less fit the virus becomes. In theory, if the resistant virus accumulates enough resistance-providing yet harmful mutations, it will no longer be able to replicate, resulting in a form of error catastrophe (303). This is the final goal and most desirable outcome of HIV chemotherapeutics. However, due to latent viral reservoirs, the selection pressure provided by the drugs must be constant. If this selection pressure is removed, then the wildtype virus will quickly return from pre-integration and post-integration latent viruses.

1.5.1 NRTIs

There are generally two distinct biochemical mechanisms that describe drug resistance to NRTIs. The first is through excision of the chain-terminator through the reverse polymerization reaction, using pyrophosphate (PPi) (through ATP) as a substrate, resulting in the release of the NRTI as a dinucleotide tetraphosphate. The second mechanism of resistance is through increased discrimination between the NRTI and the natural nucleotide.

NRTI Resistance-Conferring Mutations – Excision

A series of mutations in the vicinity of the polymerase active site confer resistance to literally all known NRTIs are called thymidine-analog mutations (TAMs) due to their initial association with resistance to thymidine-analog NRTIs AZT and d4T (304-306). TAMs include M41L, D67N, K70R, L210W, T215F/Y, and K219Q/E. The incorporation of a nucleotide produces PPi as a reaction product, which can be used as a substrate for the reverse reaction i.e. excision. However, as mentioned above, ATP has been strongly suggested to be the physiologically relevant PPi donor in the excision reaction (307,308). It is important to note here that the translocational equilibrium plays an important role in nucleotide excision. During polymerization, incoming nucleotides can only bind to the N-site, which is only vacant when RT is in the post-translocational position. Conversely, an NRTI can only be excised if it is occupying the N-site, which is the pre-translocational position. Therefore, it is not surprising that TAMs has a bias towards pre-translocated complexes (194).

The level at which TAMs confers resistance to NRTIs varies depending on the specific NRTI. The variation in resistance levels is partially explained by the ability of different NRTIs to form dead-end complexes (DECs). A dead-end complex refers to the situation when the next templated nucleotide binds to the N-site after incorporation of an NRTI. The incoming nucleotide is not incorporated, but it stabilizes the ternary complex. The name DEC refers to the fact that polymerization cannot proceed forward due to the presence of the chain-terminating NRTI, and cannot proceed backward due to the presence of the incoming nucleotide in the N-site, protecting the NRTI from the excision machinery. For example, AZT-terminated primers will cause RT to sit in a predominantly

pre-translocated position (209). Therefore, TAMs provides a high level of resistance to AZT, since it is easily excised relative to other NRTIs (309). On the other hand, tenofovir does not have a pres-translocation bias and is easily excised *in vitro*, but forms DECes relatively easily at physiological dNTP concentrations (228). As such, TAMs do not provide a high level of resistance to tenofovir (310). In summary, differences in the translocational equilibrium can strongly affect the level of resistance conferred by TAMs.

NRTI Resistance-Confering Mutations – Discrimination

The discrimination mechanism of NRTI resistance is based on the increased ability of mutant enzymes to tell the difference between the NRTI and the natural nucleotide, and positively select for the natural nucleotide. There are various mutations that arise in response to different NRTIs.

M184I/V

The M184I mutation is rapidly selected under the pressure of 3TC or FTC, but is soon replaced by M184V (311). The M184V mutation is antagonistic to the excision reaction, making it a good candidate for use with other NRTIs that follow the excision resistance pathway, such as AZT (312). The mechanism of action of the increased discrimination with M184V is reasonably well understood as a steric clash between the sidechain of V184 and the oxathiolane rings of 3TC or FTC (313).

Q151M Complex

The Q151M complex is a group of associated mutations that confer resistance to a significant number of NRTIs, including AZT, ddC, ddI, ABC and d4T (314). The Q151M complex does not affect inhibitor binding, but does disrupt interactions between the incoming dNTP and the active site residues resulting in decreased rates of incorporation (315). As all NRTIs lack a 3'OH group, this further disrupts interactions with the active site resulting in the discrimination of NRTIs.

Other Resistance Mutations

K65R is selected under pressure by tenofovir and confers resistance to many NRTIs such as ABC, 3TC, FTC, ddC and ddI (316). K70E is selected by regimens containing tenofovir, ABC and 3TC and also antagonizes the excision reaction analogously to M184V (317,318). L74V is selected for my regimens containing ABC and/or ddI and also antagonizes the excision reaction (319,320). V75T is selected under the pressure of d4T, which is no longer widely used in the clinic (321).

1.5.2 NNRTIs

Structural studies of NNRTI-bound RT have shown that the NNRTI binding pocket only exists when an NNRTI is bound, i.e. the drug creates its own binding pocket (322). This is characterized by the movement of the side chains of Y181 and Y188 (323). NNRTIs are structural diverse compounds and act kinetically as non-competitive inhibitors.

NNRTIs inhibit RT by delaying the chemical step of the reaction, possibly by reducing the mobility of the various subdomains, including the fingers (185). Since Y188 and Y181 make significant contributions to not only the formation of the binding pocket, but

also binding of first generation NNRTIs, it is not surprising that they are among the first NNRTI-resistant mutations discovered (324). The interactions between first generation NNRTIs such as nevirapine are stacking interactions with the aromatic rings of both binding partners, so mutation to aliphatic residues such as cysteine provide drug resistance (324). The K103N mutation is the most commonly reported resistant mutation for treatment with NNRTIs. However, it remains unclear how K103N confers resistance, mostly to first generation NNRTIs such as nevirapine and delavirdine. Studies have suggested that N103 forms a hydrogen bond with Y188 in the unliganded position, stabilizing the unliganded form of RT, providing an energy barrier to the formation of the NNRTI binding pocket (322).

Other mutations such as L100I, V106A and V108I do not make direct contact with the bound inhibitor, but do impair the interactions with Y188 and Y181, especially for first generation inhibitors (322). The K101E mutation provides resistance against several NNRTIs and forms a salt-bridge with E138 from the p51 subunit, causing a less than optimal configuration for the inhibitor in the binding pocket (325). The mutations G190A and L234I provide resistance by increasing the side chain bulk which causes steric clashes in the binding pocket (326,327). Conversely, the F227C results in a decrease in the size of the side chain and subsequent loss of favorable contacts with bound inhibitors (328). Etravirine (TMC125) and rilpivirine (TMC278) were developed with inherent flexibility called “wiggling” (torsional flexibility) and “jiggling” (ability to reposition) to counter the steric clashes observed in NNRTI-resistant binding pockets (329). Rilpivirine has a higher genetic barrier to resistance as compared with efavirenz, is better tolerated in

patients, and is currently being formulated with tenofovir and FTC as a single-dose daily tablet (330).

1.5.3 PFA/Foscarnet

Since PFA and PPi share a binding site, resistance mutations to PFA cluster around the binding pocket for the β and γ phosphates of the incoming nucleotide. As such, mutations such as R72A and K65R are resistant to PFA, likely due to a change in electrostatic interactions at the binding site (77,91). Another foscarnet-resistance mutant of interest is E89K. This mutation is not located at the PPi binding site, like R72A and K65R, but at nucleotide position n-2 (2 nucleotides upstream from the polymerase active site). E89K affects the ability of RT to bind in a polymerase-dependent mode, the consequence of which is that RT will never sit in the pre-translocational position required for PFA binding (188,228). However, this also affects the ability of RT to excise a nucleotide from the primer terminus (such as AZT) since this process also requires RT to sit in the pre-translocated position. As such, E89K re-sensitizes AZT-resistant RT to AZT, while providing foscarnet resistance (14).

1.6 Objectives

When I started my doctoral research, there were many small molecule compounds described in the literature that inhibited the RNase H activity of HIV-1 RT. These compounds varied in specificity, potency, and cytotoxicity, however only one had ever been shown to be active in cell-based assays, the diketo acid derivative RDS1643.

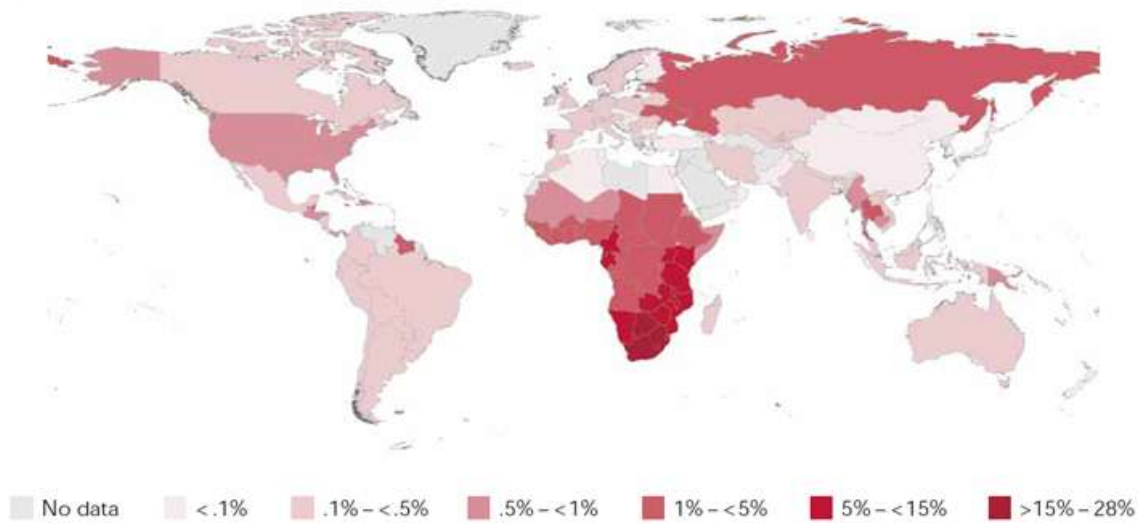
Despite this, many showed strong potency *in vitro*, and as RNase H activity is absolutely required for viral replication, it was somewhat perplexing that these compounds did not inhibit viral replication *in vivo*. We hypothesized that there must be a biochemical obstacle to inhibition that did not exist in the *in vitro* assays, but had been overlooked in cell-based assays. My first research objective was to find this obstacle, and determine if it could be overcome in order to develop the first clinically-relevant RNase H inhibitor. I observed that the active site RNase H inhibitor β -thujaplicinol exhibited a potency that was strictly dependent on the amount of enzyme turnover in the assay, i.e. the inhibitor potently inhibited RNase H activity under steady-state conditions, where each RT enzyme in the assay had to bind to many different substrates throughout the reaction, and was completely inactive under single-turnover conditions, where each RT enzyme was bound to a single substrate molecule for the duration of the experiment. Similarly, the anti-RNase H effects of the PPi analog foscarnet were shown to be related to enzyme turnover as well. Moreover, I concluded that the reason for this apparent dual nature of β -thujaplicinol was due to competition with the bound RNA substrate. This study is presented in detail in the next chapter of this thesis.

Following my work on β -thujaplicinol, I began to study a novel small molecule compound called GSK5750, which was several fold more potent than β -thujaplicinol. This increasing potency was due to increased binding affinity, and allowed me to do more detailed studies under single-turnover conditions than were possible with β -thujaplicinol. Using GSK5750, I further refined the mechanism of inhibition of small molecule active site RNase H inhibitors, showing that competition with the nucleic acid substrate was still

an obstacle to inhibition, but that compounds that bind with affinities similar to the substrate may stand a better chance of showing activity in cell-based assays. This providing strong support for our conclusions with β -thujaplicinol that single-turnover assays were necessary to properly characterize RNase H inhibitors, something that is not necessary with polymerase inhibitors of RT. This study is presented in detail in chapter 3 of this thesis.

The final study presented in this thesis builds on the work on site-specific footprinting that had been done in the lab prior to my doctoral studies, in an attempt to show a change in substrate trajectory in response to a ligand. Both Fe^{2+} - and KOONO-based site-specific footprinting techniques were used extensively to study the translocational equilibrium of HIV-1 RT, as well as the molecular determinants and effects of various ligands on this equilibrium. As KOONO-based footprinting requires a cysteine residue in close proximity to the nucleic acid (C280 in classical KOONO footprinting), we undertook a cysteine-scanning assay to determine if it was possible to develop a second footprinting site for RT. We found that the T473C mutant provided a specific cleavage of the DNA primer at positions -15 and/or -16. This provides a novel technique to study the interactions between the nucleic acid and RT in the area known as the RNase H primer grip, which is believed to be responsible for aligning the RNA template strand with the RNase H active site for RNA hydrolysis. Using this technique, we provide evidence that RNA/DNA hybrids bound to RT form a specific structure, and that inherent flexibility in the RT-nucleic acid interaction allows for this to occur. This study is provided in detail in chapter 4 of this thesis.

A



B

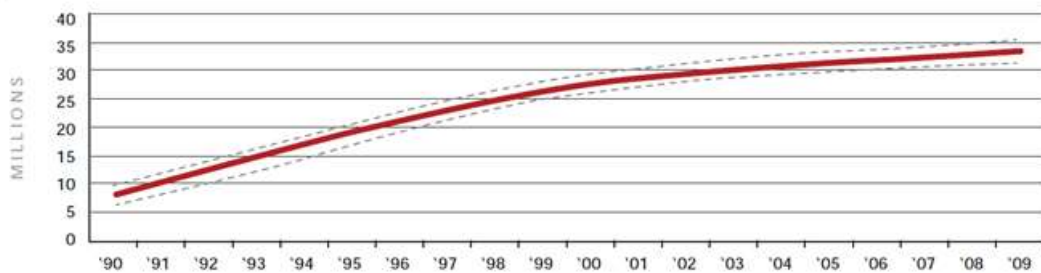
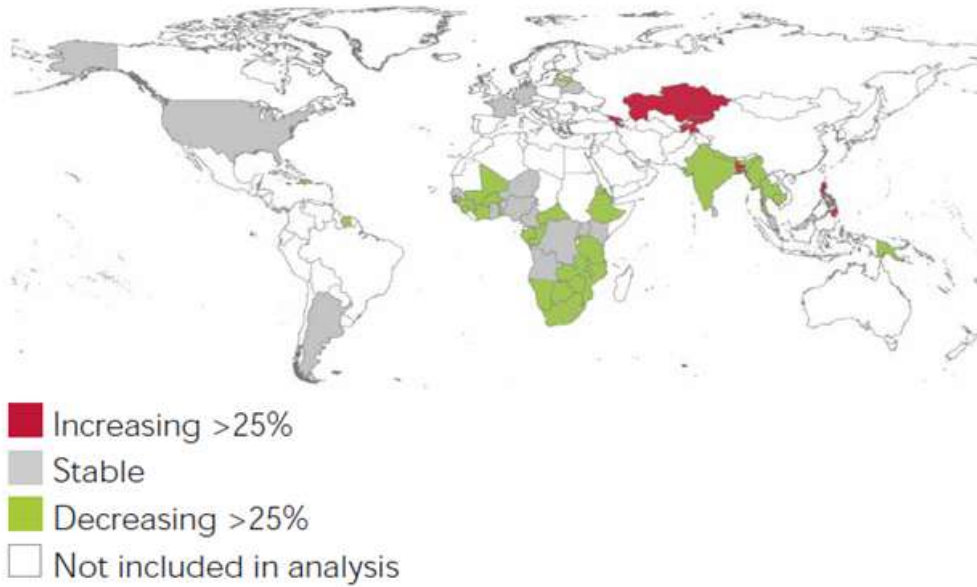


Figure 1.1: (A) Geographical distribution HIV-1 prevalence worldwide. (B) Graph showing the total number of people (in millions) living with HIV (prevalence). Adapted from (2).

A



B

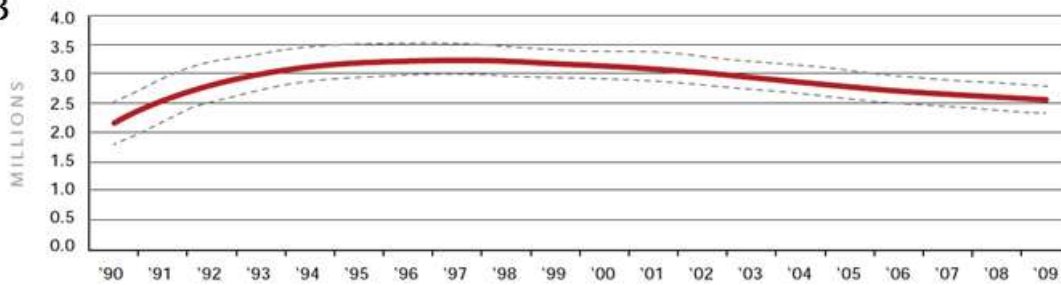


Figure 1.2 (A) Geographical distribution of HIV-1 incidence worldwide. Estimates are based on modeling data from 60 countries, 3 of which are the topic of peer-reviewed publications containing HIV-1 incidence trends. (B) Graph showing the number of new infections (in millions) per year (incidence). Adapted from (2).

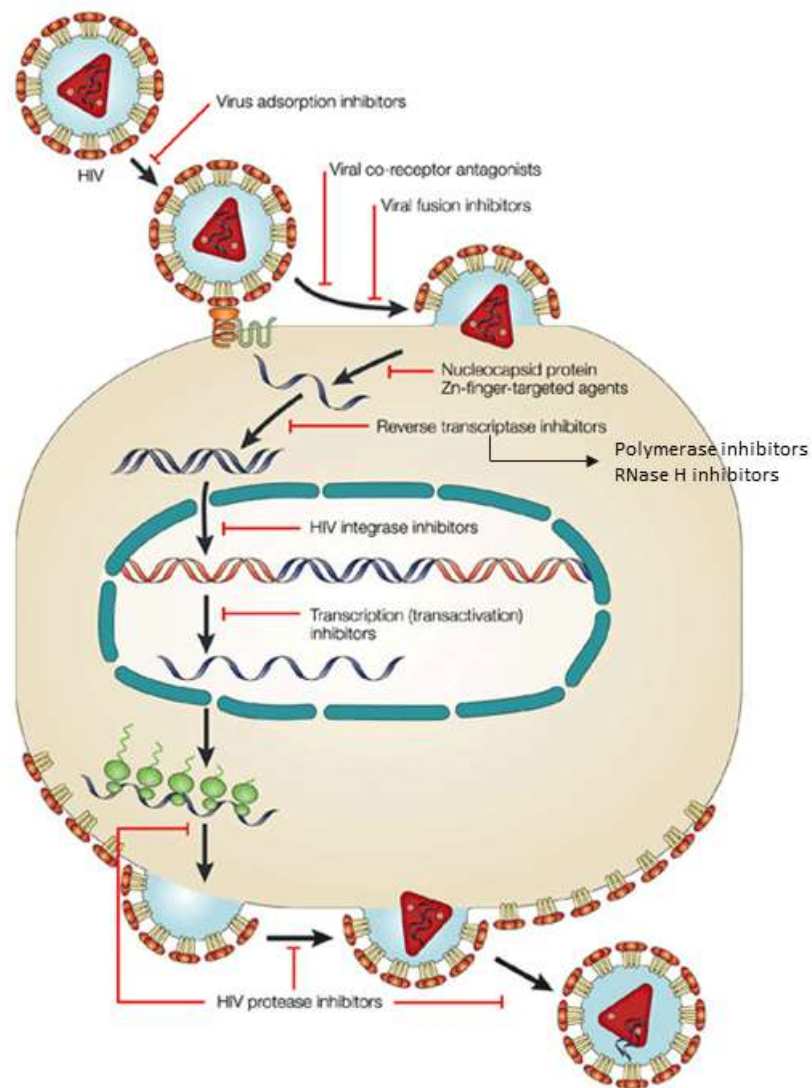


Figure 1.3. Schematic representing the HIV-1 life cycle. Adapted from (331).

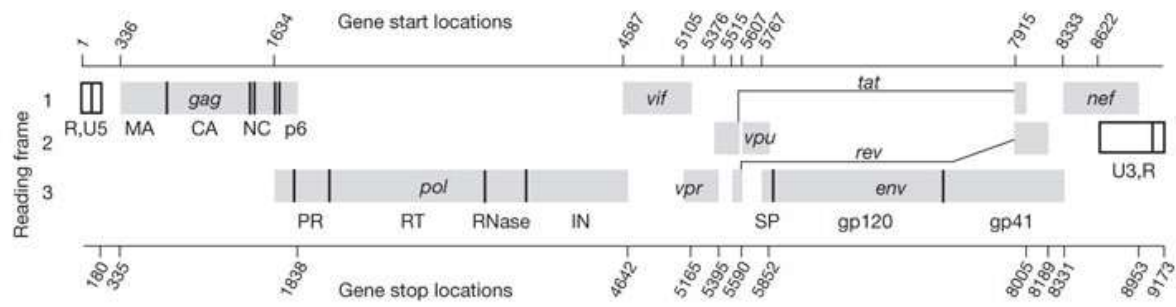


Figure 1.4: Schematic of the HIV-1 genome organization showing reading frames as well as repeat and untranslated regions. Adapted from (332).

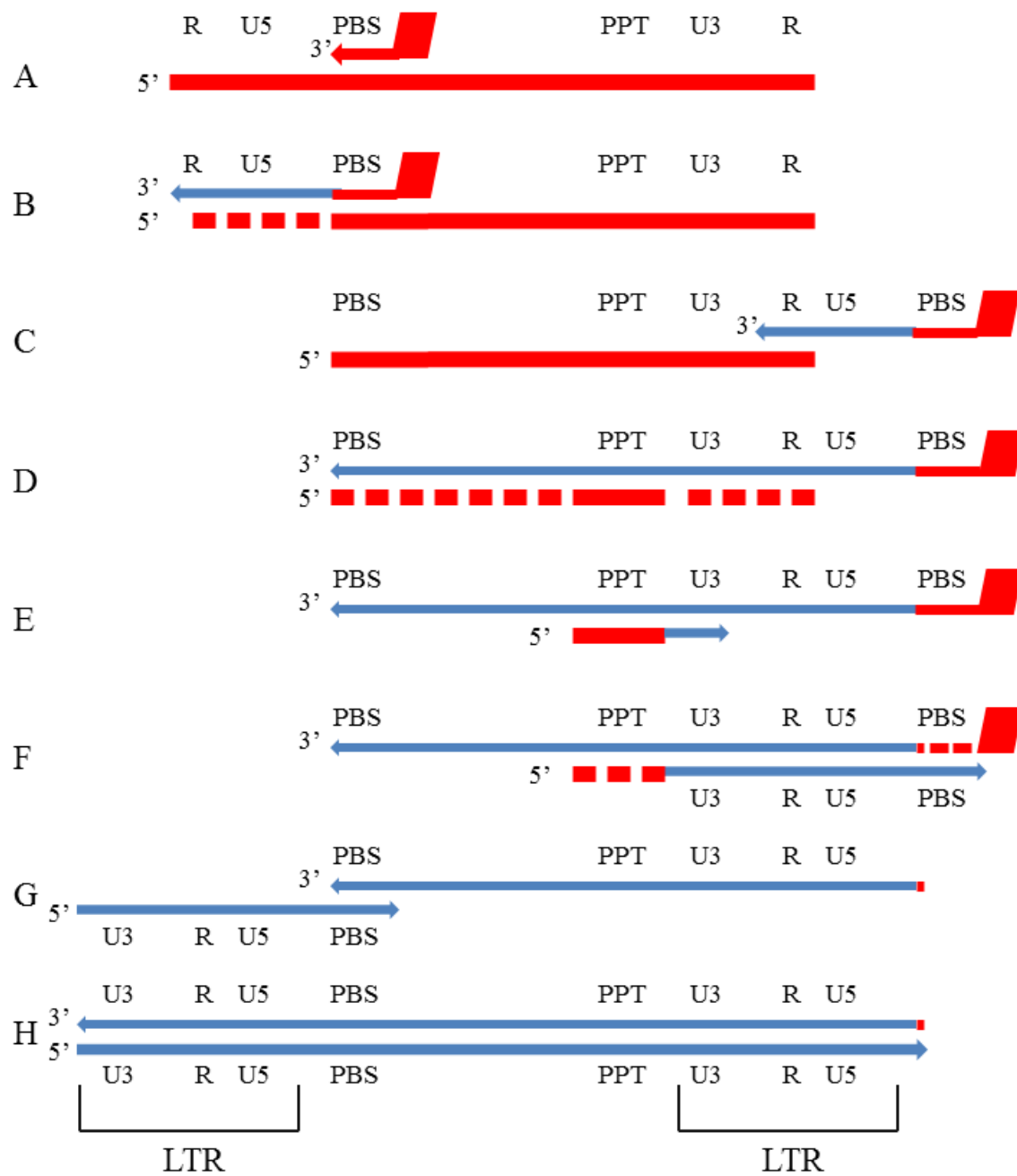


Figure 1.5.

Figure 1.5: Schematic of the process of reverse transcription in HIV-1. **(A)** The viral RNA genome is shown as a thick red line. Reverse transcription is initiated by the binding of an endogenous tRNA^{lys3} molecule to the primer binding site (PBS) on the genome. **(B)** RT elongates the tRNA primer to the 5' end of the genome, creating a fragment called (-)-strand strong stop DNA ((-)ssDNA). The RNase H activity of RT concomitantly degrades the RNA genome during DNA synthesis. The degradation of the 5' end of the genome is necessary for (-)-strand transfer, and failure to degrade the RNA at this point results in the arrest of reverse transcription. **(C)** First or (-)-strand transfer. The (-)ssDNA fragment dissociates from the PBS sequence and re-associates with the repeat (R) sequence at the 3' end of the genome. This step is capable of both intrastrand and interstrand transfer. **(D)** Continuation of (-)-strand DNA synthesis. RT extends the 3' end of the (-)ssDNA fragment toward the PBS sequence, while the RNase H activity concomitantly degrades the RNA genome, with the exception of the polypurine tract (PPT). **(E)** The PPT is used as the primer for the initiation of (+)-strand DNA synthesis. The PPT primer is extended by the RT polymerase activity. **(F)** After approximately 12 nucleotides have been added, the PPT primer is removed by RNase H activity. The nascent (+)-strand DNA is extended to the 5' end of the (-)-strand DNA, copying the PBS sequence from the tRNA that is still associated with the (-)-strand DNA. Here, the tRNA is removed by the RNase H activity, leaving a single ribonucleotide (rA) at the 3' end of the U5 sequence (shown in red). **(G)** In the second, or (+)-strand transfer, the PBS sequences on both strands associate. This step occurs predominantly in an intrastrand fashion. **(H)** Both DNA strands are extended

to the ends of their templates, forming the provirus that is ready to be integrated into the host genome by integrase. The long terminal repeats (LTRs) that are formed as a result of reverse transcription are shown.

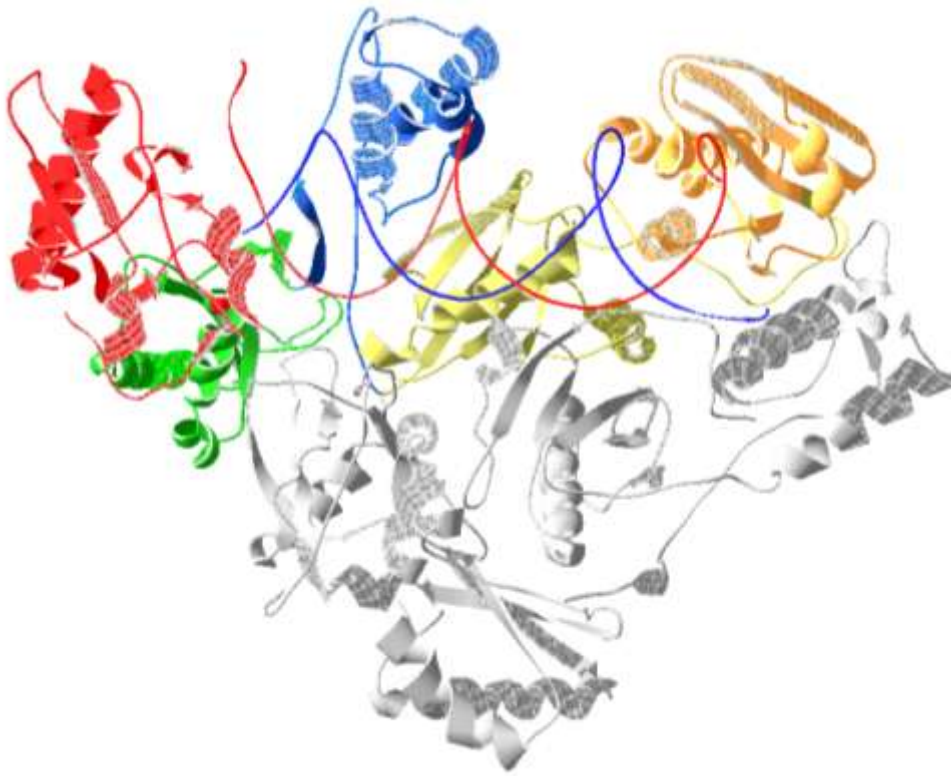


Figure 1.6: Structure of HIV-1 reverse transcriptase complexed with a DNA/DNA substrate. The p51 subunit is shown in grey. The polymerase domain consists of the fingers (red), palm (green), thumb (blue) and connection (gold) subdomains. The RNase H domain is shown in orange. The DNA primer is shown in blue, while the DNA template is shown in red. PDB code is 1RTD.

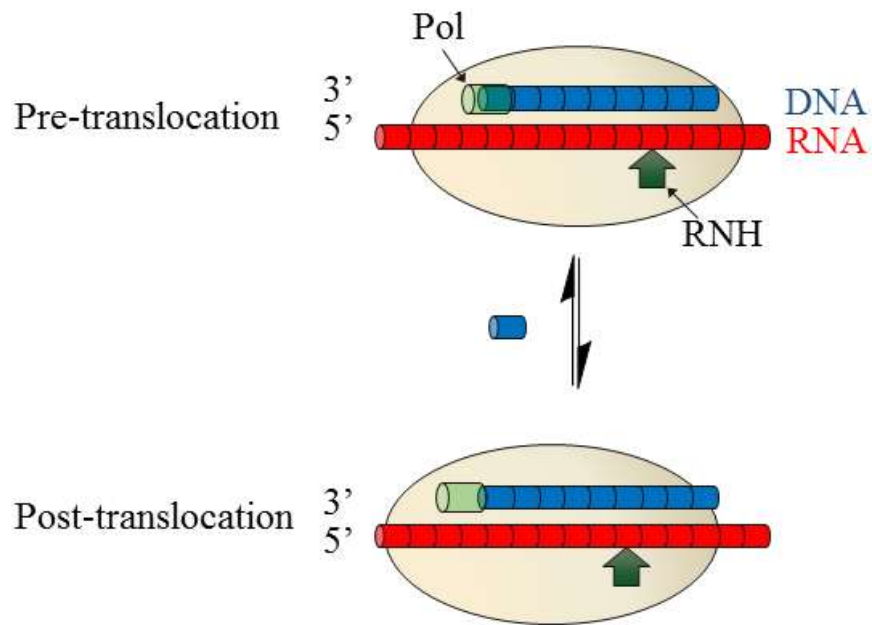


Figure 1.7: Schematic of RT translocation. Polymerase and RNase H active sites are marked by a green cylinder and green arrow respectively. Nucleotides can bind to post-translocated complexes, where they are incorporated and become pre-translocated complexes.

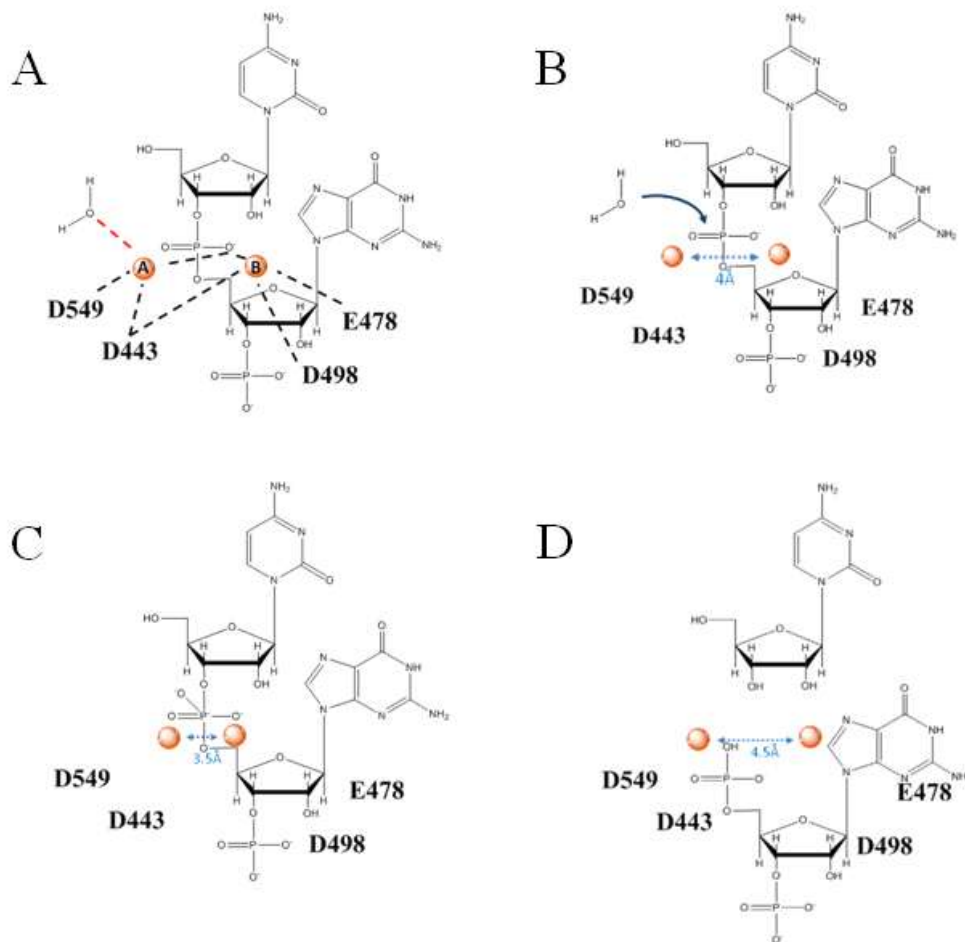


Figure 1.8: Schematic of the mechanism of RNA hydrolysis by RNase H. The chemistry of RNase H cleavage is believed to be a two-metal ion mechanism. **(A)** Two divalent metal ions (red spheres, marked A and B) are coordinated by the active site residues D549, D443, D498 and E478 approximately 4Å apart. Metal ion A activates a water molecule. **(B)** The activated water molecule carries out a nucleophilic attack (blue arrow) driving the phosphoryl transfer reaction. **(C)** In the putative transition state, the metal ions move toward each other to bring the nucleophile within range of the scissile phosphate. **(D)** The reaction products consist of a 3' OH group and a 5' phosphate group, and the metal ions are again likely to be re-positioned.

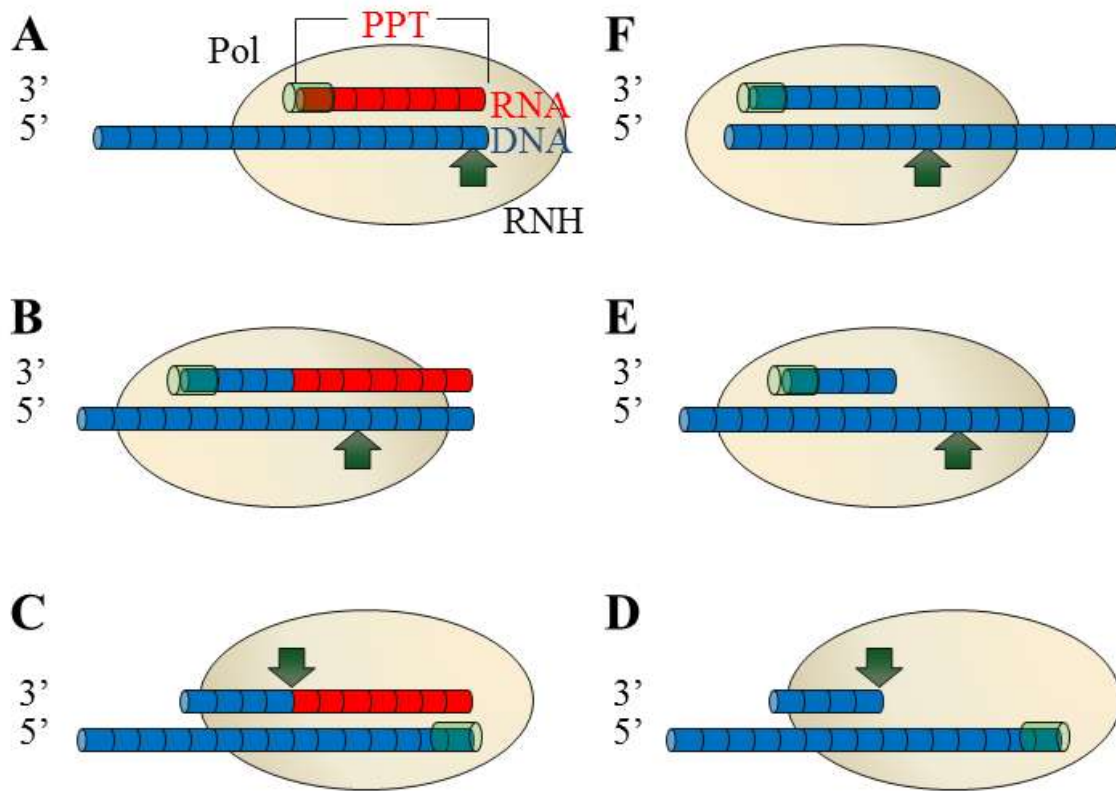


Figure 1.9: Schematic of the initiation of (+)-strand DNA synthesis. (A) RT binds to the PPT RNA sequence left over from RNase H degradation of the genomic RNA. (B) RT extends the (+)-strand primer (PPT) by 12 nucleotides, then pauses. (C) RT then re-positions itself so the RNase H active site is positioned over the DNA-RNA junction and cleaves. (D) RT then removes the PPT primer through RNase H activity. (E) RT then re-positions itself again as a polymerase. (F) (+)-strand DNA synthesis continues.

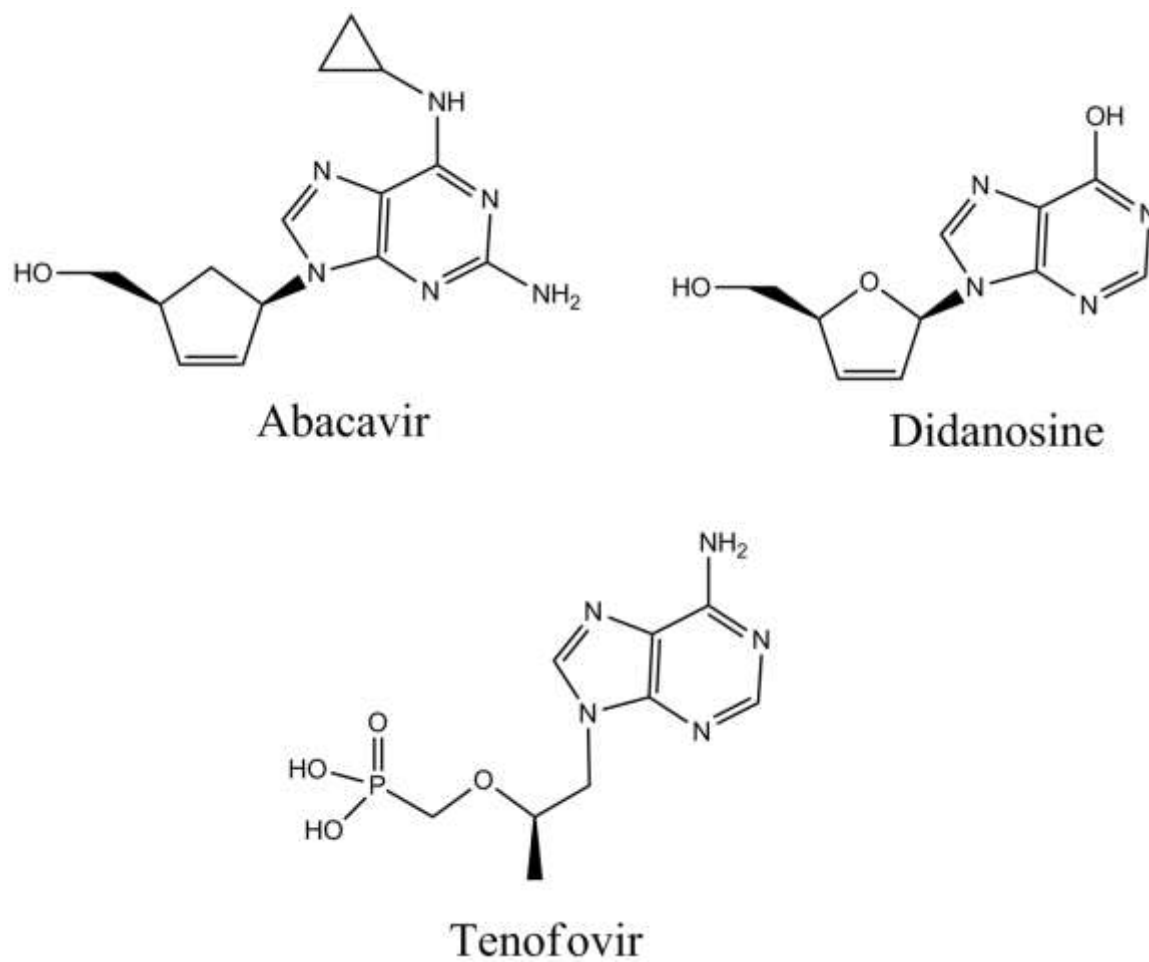
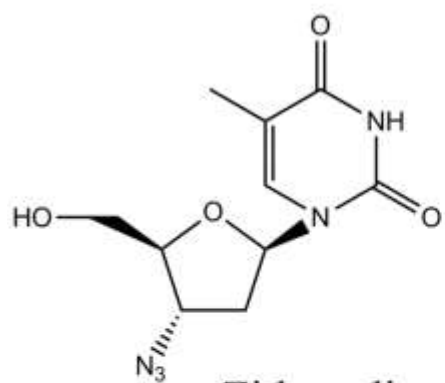
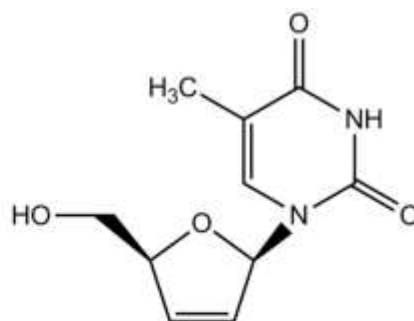


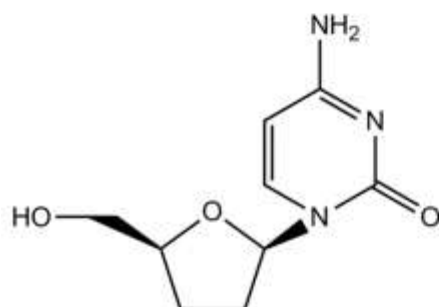
Figure 1.10: Chemical structures of purine-based NRTIs



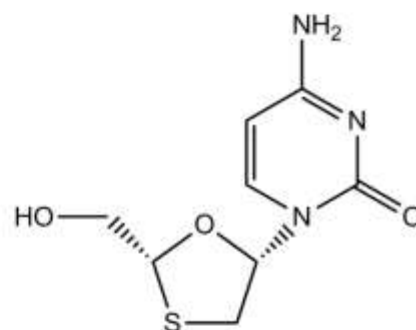
Zidovudine



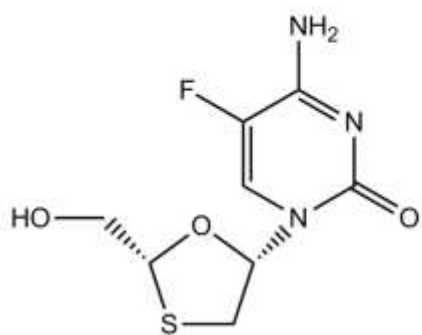
Stavudine



Zalcitabine



Lamivudine



Emtricitabine

Figure 1.11: Chemical structures of pyrimidine-based NRTIs

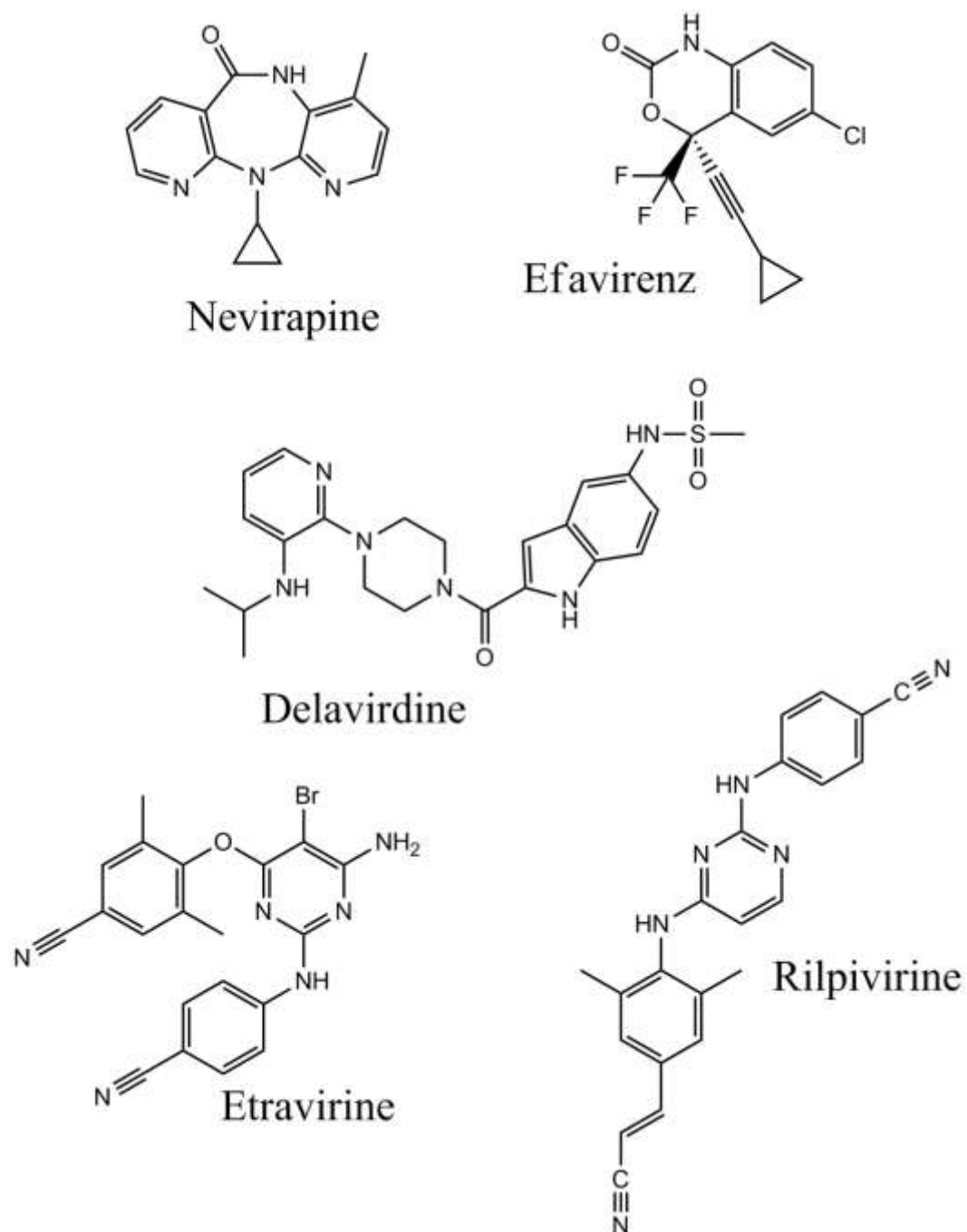
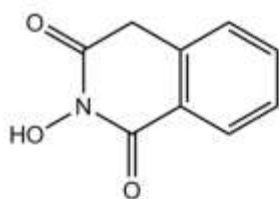
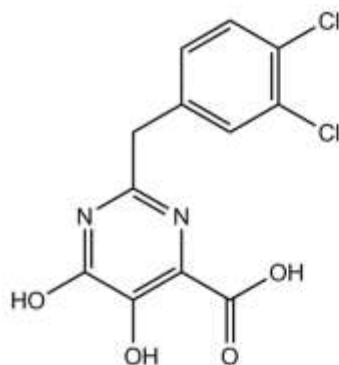


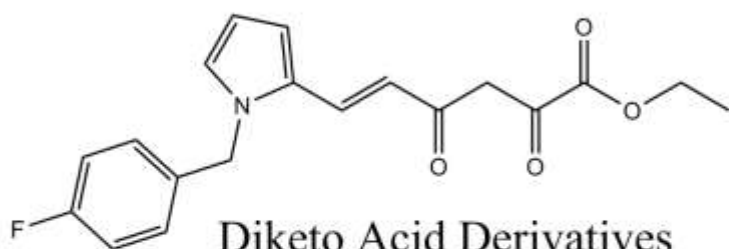
Figure 1.12: Chemical structures of select NNRTIs.



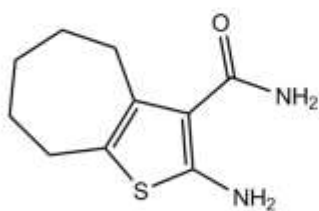
N-Hydroxyimides



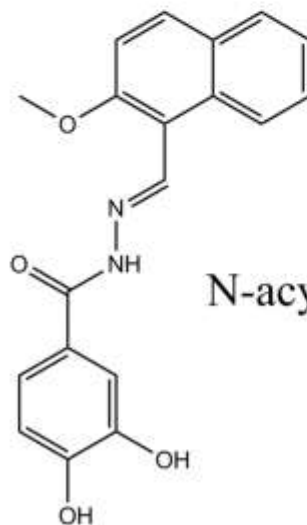
Pyrimidinol Carboxylic Acid Derivatives



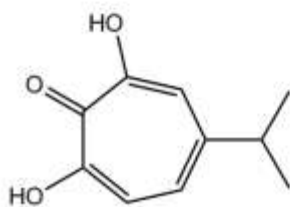
Diketo Acid Derivatives



Vinylogous Ureas



N-acyl Hydrazones



Tropolone Derivatives

Figure 1.13: Chemical structures of various inhibitors of HIV-1 RT-associated RNase H.

Chapter 2

HIV-1 Reverse Transcriptase Can Simultaneously Engage its DNA/RNA Substrate at both the DNA Polymerase and RNase H Active Sites: Implications for RNase H Inhibition

Greg L. Beilhartz¹, Michaela Wendeler², Noel Baichoo², Jason Rausch², Stuart Le Grice²,
and Matthias Götte^{1‡}

¹*Department of Microbiology & Immunology, McGill University, Montreal, QC, Canada,
H3A 2B4.*

²*RT Biochemistry Section, HIV Drug Resistance Program, National Cancer Institute,
Frederick, MA, USA*

2.1 Abstract

Reverse transcriptase (RT) of the human immunodeficiency virus (HIV) possesses DNA polymerase and ribonuclease (RNase) H activities. Although the nucleic acid binding cleft separating these domains can accommodate structurally-diverse duplexes, it is currently unknown whether regular DNA/RNA hybrids can simultaneously contact both active sites. In this study we demonstrate that ligands capable of trapping the 3'-end of the primer at the polymerase active site affect specificity of RNase H cleavage without altering the efficiency of the reaction. Experiments under single turnover conditions reveal that complexes with a bound nucleotide substrate show specific RNase H cleavage at template position -18, while complexes with the pyrophosphate analogue foscarnet show a specific cut at position -19. This pattern is indicative for post- and pre-translocated conformations. The data are inconsistent with models postulating that the substrate toggles between both active sites, such that the primer 3'-terminus is disengaged from the polymerase active site when the template is in contact with the RNase H active site. In contrast, our findings provide strong evidence to suggest that the nucleic acid substrate can engage both active sites at the same time. As a consequence, the bound and intact DNA/RNA hybrid can restrict access of RNase H active site inhibitors. We have mapped the binding site of the recently discovered inhibitor β -thujaplicinol between the RNase H active site and Y501 of the RNase H primer grip and show that the inhibitor is unable to bind to a pre-formed RT-DNA/RNA complex. In conclusion, the bound nucleic acid substrate, and in turn, active DNA synthesis can represent an obstacle to RNase H inhibition with compounds that bind to the RNase H active site.

2.2 Introduction

The human immunodeficiency virus type-1 (HIV-1) reverse transcriptase (RT) is a heterodimeric (p66/p51), multifunctional DNA polymerase that converts the single-stranded (ss) viral RNA genome into integration-competent double-stranded DNA. This process requires coordination of the RT-associated DNA polymerase and ribonuclease (RNase) H activities(333). Both active sites reside in the large subunit p66, and accommodate divalent metal ions required for catalyzing the nucleotide incorporation and strand scission events (57,184,188,210). The RNase H activity of HIV-1 RT degrades the RNA moiety of RNA/DNA replication intermediates during synthesis of the first, or minus DNA strand. The rate of nucleotide incorporation is approximately an order of magnitude faster than that of RNase H cleavage (191); however, specific RNase H cuts can be seen at a fixed distance of 18 base pairs (bp) upstream of the primer 3'-terminus in the absence of nucleoside triphosphates (dNTPs) or when DNA synthesis is arrested by chain-terminating nucleotides (180,181,223-226). This specific cleavage between template positions -18 and -19 is referred to as polymerase-dependent RNase H activity and reflects the distance between the polymerase and RNase H active sites. HIV-1 RNase H also functions in a polymerase-independent mode, which collectively refers to cleavage events that occur when the 3'-end of the primer is not located in the vicinity of the polymerase active site, although variations of this mechanism have been reported (233). This activity is required to remove viral RNA during and following minus-strand DNA synthesis, as well as to specifically remove tRNA^{Lys3} and the polypurine tracts (PPTs), the primers for minus- and plus-strand DNA synthesis, respectively (58,237-240,247).

Despite its crucial role in viral replication, the RT-associated RNase H activity has yet to be explored as a pharmaceutical target for drug development efforts. All currently approved RT inhibitors interfere with the polymerase activity of the enzyme (334). Some of these inhibitors also affect RT-associated RNase H activity. Non-nucleoside analogue RT inhibitors (NNRTIs), for example, have been shown to alter the specificity and efficiency of RNase H cleavage (335), while nucleoside analogue RT inhibitors (NRTIs) can alter the RNase H cleavage pattern, without affecting cleavage efficiency (336). Phosphonoformic acid (PFA, foscarnet), a pyrophosphate analogue, also inhibits both the polymerase and RNase H activities of RT in cell-free assays (286), as do N-acyl hydrazones, some of which have been shown to inhibit RT-associated RNase H activity exclusively (288,291,292). Some diketo acids (284), N-hydroxyimides (218,280), and dihydroxytropolones (276,337) show inhibition of RNase H cleavage in the submicromolar range and share common structural features capable of divalent metal ion chelation. These compounds may therefore interfere with positioning of Mg^{2+} ions at the RNase H active site (281). In contrast, vinylogous ureas show inhibition of RNase H cleavage through a different mechanism of action(294).

Crystallographic data of an RNase H-competent complex of HIV-1 RT are not available, which complicates mechanistic studies on RNase H inhibition. HIV-1 RT has been crystallized with duplex DNA (184), and a PPT-derived DNA/RNA hybrid (57). In each of these structures, the 3'-end of the primer is engaged at the polymerase active site, with most protein-nucleic acid contacts observed between the first 6 base pairs and the fingers, palm and thumb subdomains of the N-terminal polymerase domain of p66.

Residues that constitute the RNase H primer grip contact primarily the primer strand in a region of duplex 11 to 15 nucleotides upstream from the primer 3' terminus (57). The RNase H active site does not contact the scissile bond in any of the crystallized complexes, which is consistent with the expectation that the nucleic acids in these complexes would not be cleaved *in vivo*. However, recent modeling studies in which a polymerase-competent RT-DNA/RNA complex and a truncated RNase H-competent complex of human RNase H1 (containing the active site mutant D210N) are overlaid suggest that the RNA/DNA hybrid cannot simultaneously engage the DNA polymerase and RNase H active sites, regardless of the sequence context (214). The model predicts that positioning an RNA/DNA hybrid for cleavage at the RNase H domain of HIV RT requires that the substrate can shift positions such that the primer 3'-terminus is disengaged from the polymerase active site. As a consequence, ligands that are trapped at the polymerase active site and engage the 3'-end of the primer could indirectly inhibit RNase H cleavage by preventing the conformational change necessary for the DNA/RNA hybrid to contact the RNase H active site.

In the present study, we concurrently examined how engagement of substrate at the polymerase active site affects RNase H activity, and explored the mechanism of action of β -thujaplicinol, a dihydroxytropolone (Figure 2.1A) recently shown to directly inhibit RT-associated RNase H activity (276). We demonstrate that stabilizing the DNA polymerase active site at the 3'-end of the primer can affect the position of RNase H cleavage in a highly specific manner without inhibiting the activity. These findings are inconsistent with models postulating that the substrate can shift positions or toggles

between polymerase and RNase H active sites. Rather, these data show that the bound DNA/RNA hybrid can simultaneously engage both active sites. Primer/template binding restricts access of β -thujaplicinol to its target site that we map between the metal binding sites and Y501 of the RNase H primer grip, suggesting that the nucleic acid substrate can be an obstacle for RNase H active site inhibitors.

2.3 Results

Experimental design

In order to study whether the nucleic acid substrate can simultaneously engage both active sites of HIV-1 RT, we utilized RT-primer/template complexes that are stabilized through ligand binding at the DNA polymerase active site. We recently showed that the pyrophosphate-analogue PFA affects the translocation status of HIV-1 RT (228). Translocation of RT relative to its nucleic acid substrate follows a complete cycle of nucleotide incorporation to allow binding of the next complementary dNTP (208,230). This process frees the nucleotide binding site and moves the 3'-end of the primer into the adjacent priming site. The position of the RNase H active site is shifted in concert by a single nucleotide (Figure 2.1B). High resolution footprinting studies with HIV-1 RT bound to duplex DNA show that PFA binds to the DNA polymerase active site and traps the pre-translocated complex, while the nucleotide substrate traps the post-translocated complex (209). Both ternary complexes are stabilized, and resist challenge with heparin, relative to the binary RT-primer/template complex. Here we asked (i) whether ligand

binding at the DNA polymerase active site can affect specificity and efficiency of RNase H cleavage, and (ii) what role the bound nucleic acid substrate plays in RNase H inhibition by β -thujaplicinol.

Effects of ligands on polymerase-dependent RNase H cleavage

We devised a DNA/RNA substrate with a DNA primer (22-mer) recessed at both 5'- and 3'-ends on a 52-mer RNA template and monitored primary, polymerase-dependent cleavage events under steady-state conditions. Two major RNase H cuts are evident in the absence of small molecule ligands (Figure 2.2A, left). The first is located between residues -18 and -19, and the second is between residues -19 and -20 (defined here as positions -18 and -19, respectively). Saturating concentrations of β -thujaplicinol inhibited both cuts (Figure 2.2A, middle and C). The addition of PFA also showed considerable inhibition of RNase H activity (Figure 2.2A, right and C). However, in contrast to β -thujaplicinol, the pyrophosphate analogue shifted the cleavage pattern. The cut at position -18 is weaker than the cut at -19, which is reminiscent of our previous footprinting data (228).

We conducted the same experiment under single turnover conditions to exclude potential effects of repeated complex dissociation and association on the cleavage patterns. The RT-DNA/RNA complex was pre-formed and the reaction was initiated with divalent metal ions in the absence and presence of inhibitor. Under these conditions, β -thujaplicinol failed to inhibit RNase H activity (Figure 2.2B, middle and D). RNase H

cleavage in the presence of PFA is likewise not inhibited (Figure 2.2B, right and D). The intensity of the -19 cleavage product is almost identical to the sum of the two cuts seen in the absence of PFA. However, the cleavage pattern shows again a strong bias towards -19 cleavage, demonstrating that the DNA/RNA hybrid simultaneously contacts the DNA polymerase and RNase H active sites. The data suggest at the same time that the bound nucleic acid can represent an obstacle that limits access of β -thujaplicinol to its binding site.

To fully define the influence of the translocation status of HIV-1 RT on RNase H cleavage, we next monitored activity in the presence of increasing concentrations of PFA and in the presence of the next complementary dNTP, respectively (Figure 2.3A). The 3'-end of the primer contained a 2'-deoxy-thymidine-monophosphate (dTMP) in the former case, while a 2',3'-dideoxy-thymidine-MP (ddTMP) was used in the latter case to prevent nucleotide incorporation. Increasing concentrations of PFA biased cleavage to position -19, while increasing concentrations of the dNTP biased cleavage towards position -18. However, in spite of these alterations in cleavage patterns, binding of PFA or the nucleotide substrate, and in turn, formation of stable ternary complexes, does not diminish the efficiency of the RNase H activity under single turnover conditions (Figure 3B). We next studied whether β -thujaplicinol retains its ability to inhibit RNase H cleavage under steady-state conditions in the presence of PFA and the incoming dNTP, (Figure 2.3C). The two ligands diminish RNase H cleavage under these conditions; however, the addition of β -thujaplicinol does not cause further reductions in RNase H

activity, which is consistent with the notion that stable RT-DNA/RNA complexes restrict access of the dihydroxytropolone to its binding site.

Binding of β -thujaplicinol requires divalent metal ions

We performed order-of-addition experiments to analyze whether β -thujaplicinol can bind to the free enzyme. The following conditions were used: (i) pre-incubation of RT with inhibitor and Mg^{2+} , and initiating with primer/template, (ii) pre-incubation of RT with inhibitor only, and initiating the reaction with primer/template and Mg^{2+} , and (iii) pre-incubation with RT, primer/template, and β -thujaplicinol, and initiating the reaction with Mg^{2+} alone (Figure 2.4). Reactions were monitored under pre-steady-state conditions, prior to establishing equilibrium with all components. Differences in the order of addition do not significantly influence the rate of the reaction in the absence of inhibitor (Figure 2.4A). In keeping with the aforementioned data, β -thujaplicinol does not inhibit the reaction when the RT-DNA/RNA complex is pre-formed and the compound is added at the start of the reaction with Mg^{2+} (Figure 2.4B). Inhibition is also negligible when RT is pre-incubated with the compound in the absence of divalent metal ions. However, inhibition is seen when the RT-inhibitor complex is formed in the presence of Mg^{2+} . Thus, β -thujaplicinol appears to bind solely to the free enzyme, and binding is metal ion dependent.

The rate constants for the three reactions ($k_{RNase\ H}$) are $0.11\ s^{-1}$ with a pre-formed RT-inhibitor- Mg^{2+} complex, compared to $0.35\ s^{-1}$ when the enzyme-inhibitor complex was pre-incubated in the absence of metal ions, and $0.38\ s^{-1}$ with a pre-formed enzyme-

primer/template complex. The maximum product generated after 20 seconds remains unchanged under these conditions, suggesting that the components of the reaction equilibrate at longer reaction times. Thus, there is a modest ~3-fold inhibitory effect before the primer/template substrate can bind to the enzyme. The template is eventually cleaved, once the RT-DNA/RNA complex is formed, which in turn, suggests that the inhibitor is released from the complex within a single turnover. Measurements of the equilibrium dissociation constant (K_d) for the nucleic acid substrate revealed values of 3 nM and 15 nM, respectively, depending on whether nucleotide incorporation or RNase H cleavage was used as readout. Thus, although binding of RT to its DNA/RNA substrate is locally weakened at the RNase H domain, the inhibitor appears unable to compete with the intact nucleic acid substrate.

β -thujaplicinol inhibits predominantly secondary RNase H cleavage

The dependence on divalent metal ions for RNase H inhibition points to two possible binding sites for β -thujaplicinol: binding to the RNase H active site, which provides a mechanism for direct inhibition of RNase H cleavage, or alternatively, binding to the polymerase active site. To distinguish between the two possible scenarios, we used a chimeric DNA/DNA-RNA substrate that is cleaved in polymerase-independent fashion in the vicinity of the DNA-RNA junction (Figure 2.5A). This substrate mimics the tRNA-primer removal reaction. In agreement with previous data, the primary cut is seen a single residue upstream of the RNA-DNA junction, and ensuing secondary cuts follow with time (Figure 2.5B, left) (59,247,338). β -thujaplicinol inhibits ensuing secondary cleavages to a much greater degree than the primary cut (Figure 2.5B, middle). The

ensuing secondary cleavages showed 50% inhibitory concentrations (IC_{50}) of approximately 150 nM, while the primary cleavage displayed an IC_{50} of $> 25 \mu M$. The low value of 150 nM for the secondary cuts is in keeping with the value of 210 nM previously reported, based on a FRET- assay that does not distinguish between primary and ensuing secondary cuts (276,339). Due to its short length, the substrate is never in contact with the polymerase active site when RNase H cleavage takes place. The presence of PFA does not inhibit RNase H cleavage (Figure 2.5B, right), supporting previous findings that β -thujaplicinol appears to target divalent metal ions at the RNase H active site.

Benzoyl-L-Phe at position 501 confers resistance to β -thujaplicinol inhibition

Previous modelling studies have suggested that binding of N-acylhydrazone inhibitors of HIV RNase H requires the presence of divalent metal ions, as well as stacking interactions between the aromatic ring of a prototype inhibitor and the highly conserved Y501 of the RNase H primer grip (292). To study whether this residue likewise facilitates binding of β -thujaplicinol via its ring system, we constructed mutant enzymes with natural and unnatural amino acid substitutions at this position (Figure 2.6). All IC_{50} values reported in Figure 2.6 were obtained using a FRET-based RNase H assay, as described under Materials and Methods. Retention of RNase H activity despite replacement of a conserved RNase H primer grip residue allowed us to determine IC_{50} values in the context of the various mutant enzymes. Replacing Y501 with either W or F only marginally raised the IC_{50} , (470 and 356 nM, respectively), when compared with wild type RT (308 nM). This data demonstrates that the hydroxyl function of Y501 is not

critical for inhibitor binding. However, the Y501W mutant showed diminished RNase H activity per se when compared with wild type RT and Y501F, respectively, which is consistent with the aforementioned study. Of the two non-natural amino acid substitutions, introducing an azido function (p66^{AzF}/p51 RT) also had little to no effect on β -thujaplicinol sensitivity. In contrast, inserting a benzophenone into the RNase H primer grip (p66^{501BpF}/p51 RT) created an enzyme that was resistant to β -thujaplicinol at inhibitor concentrations as high as 300 μ M, i.e. three orders of magnitude greater than the IC₅₀ of the wild type enzyme. This substantial alteration in sensitivity to β -thujaplicinol suggests that the benzophenone moiety of p66^{BpF}/p51 RT restricts access of the inhibitor to its binding site, while the activity of this enzyme was comparable with the Y501W mutant that was fully sensitive to the inhibitor.

Modeling the binding site for β -thujaplicinol

The experimental observations could be reconciled in a model for binding of β -thujaplicinol, and mechanisms of inhibition and resistance (Figure 2.7). Inhibitor binding to wild type RT may be mediated by coordination of a catalytic Mg²⁺ ion by carbonyl and hydroxyl oxygens of β -thujaplicinol (Figure 2.7A). In addition, π -stacking between the central ring of the inhibitor and the side chain of Y501 may provide a second critical element in inhibitor binding (Figure 2.7B). The latter interaction requires rotation of the Y501 side chain around the C α -C β and C β -C χ bonds from the position observed in RT apoenzyme and co-crystal structures, as suggested for the N-acylhydrazone (292). Positioning of the inhibitor in this manner would be expected to be sterically incompatible with substrate binding, with the principal clash occurring between an

exocyclic hydroxyl group of β -thujaplicinol and C4' of the nucleotide immediately 5' to the scissile phosphate.

The structural basis for Y501BpF resistance to inhibition by β -thujaplicinol may be that the inhibitor is unable to recruit benzophenone to a position amenable for stacking (Figure 2.7B and 2.7C). The increased length of the unnatural side chain may allow for accommodation of the distal ring of benzophenone into a hydrophobic pocket flanked by L479 and the peptide backbone of K476. In addition, the ϵ -amino group of the latter may participate in hydrogen bonding with the benzophenone carbonyl moiety. Taken together, these interactions may prevent rotation of the side chain altogether, rendering it unavailable for stacking with β -thujaplicinol, and as a consequence, rendering the inhibitor incapable of stably binding RT. It is unlikely that benzophenone substitution at position 501 would itself preclude substrate binding, given the predicted trajectory of the unnatural side chain.

2.4 Discussion

Binding of DNA/RNA substrate by HIV-1 RT

Here we studied the coordination of interaction at the polymerase and RNase H active sites of HIV-1 RT with its nucleic acid substrate. Two conflicting models have been proposed for the HIV-1 RT-DNA/RNA complex. The first model suggests that both the polymerase active site and the RNase H active site can simultaneously engage the substrate. This model is supported by several biochemical studies that show RNase H

cleavage at a fixed distance of 18 base pairs upstream of the 3'-end of the primer (180,181,223-226). These findings imply that the polymerase active site can interact with the primer terminus at the same time when the template is cleaved. The second proposal suggests that the substrate cannot simultaneously interact with both active sites, based on modeling studies designed to reconcile the structure of the RT complex with a PPT-derived DNA/RNA hybrid and the structure of human RNase H1-substrate complex (214). The former provides the specific interaction at the polymerase active site, while the latter provides the specific interaction at the RNase H active site. With these constraints, it appears impossible to connect the RNA strands, which led the authors to propose that the nucleic acid substrate toggles between the two active sites. Such differences in substrate binding described by the two models, will impact on the binding properties and mechanisms of action of small molecule RNase H inhibitors. One would predict that substrate toggling increases access of small molecules to the RNase H active site, while substrate binding at both active sites restricts inhibitor accessibility. In this study, we demonstrate that a regular DNA/RNA hybrid occupies both active sites at the same time, which can be an obstacle in the development of RNase H active site inhibitors.

Previous studies under steady-state conditions have shown that binding of the PPi-analogue PFA or the dNTP substrate can diminish RNase H cleavage (224,286). At first glance, these findings are consistent with the model suggesting that the DNA/RNA substrate that is specifically poised at the polymerase active site cannot simultaneously interact with the RNase H active site. However, we show here that RNase H activity is not inhibited when the experiments were conducted under single turnover conditions.

Ligand binding is unambiguously shown through differences in the translocation status of RT. The presence of PFA yields a single RNase H cut at position -19 of the template, while the presence of the incoming dNTP yields a single cut further downstream at -18. The absence of ligands shows an even distribution of the two cleavage events, which points to the existence of two isoenergetic conformations in this particular sequence context. Consistent with a Brownian ratchet model for polymerase translocation (228,340), we suggest that the two complexes exist in a dynamic equilibrium, and represent pre- and post-translocational states. Nucleotide hydrolysis is not required for translocation and as such the enzyme can freely oscillate between the two positions. Interconversion of the two conformations in the absence of ligands is kinetically invisible (191). Ligands such as PFA or the nucleotide decrease the forward motion or the reverse, respectively. Thus, PFA traps the pre-translocated complex, which provides a mechanism of inhibition of DNA synthesis, while the nucleotide traps the post-translocated complex, which provides a mechanism for RT translocation associated with productive DNA synthesis (208,209,228,229,341). Thus, a single RNase H cut is indicative for specific interactions with the ligand, the 3'-end of the primer, and the DNA polymerase active site. Our observation that the efficiency of RNase H cleavage is not reduced when compared to the unliganded RT-DNA/RNA complex provides strong evidence for a model that allows simultaneous substrate binding to the DNA polymerase and RNase H active sites.

Coordination of polymerase and RNase H activities

Reduction in RNase H activity under steady-state conditions in the presence of dNTP substrate or PFA can be explained by the increased stability of ternary complexes (228,229). The increased stability of the complex with one or the other ligand diminishes the turnover of the reaction, and, consequently, overall efficiency of RNase H cleavage. However, our experiments under single turnover conditions show that the RNase H activity *per se* is not affected in the presence of PFA or dNTP.

The ability of the DNA polymerase and RNase H active sites to interact simultaneously with the DNA/RNA substrate does not imply that both activities are temporally coordinated. RNase H cleavage was shown to be slower than DNA synthesis (191). Here, we show that the affinity around the two active sites is likewise not identical. Equilibrium dissociation constants (K_d) were measured at both active sites, and values obtained at the RNase H active site are approximately 5-fold higher compared to measurements at the polymerase active site. This is consistent with crystallographic data showing that most interactions with the nucleic acid substrate are mediated through the polymerase domain (57,184,188). In light of the collective data discussed above, we propose a model that allows partial dissociation and re-association or “breathing” of enzyme-nucleic acid interaction around the RNase H active site, while contacts around the polymerase active site are retained. This model remains consistent with the observation that RNase H cleaves its substrate predominantly during RT pausing events (342). However, our data are inconsistent with a model that postulates mutually exclusive binding at each active site.

Consequences for RNase H inhibition

While different classes of DNA polymerase active site inhibitors, including nucleotide analogues, pyrophosphate-analogues, and nucleotide competing RT inhibitors require the primer/template for binding (343), our data show that the nucleic acid substrate can be an obstacle for RNase H active site inhibitors. Order-of-addition experiments suggest that β -thujaplicinol binds in close proximity to the RNase H active site of the free enzyme, while the nucleic acid substrate restricts access to the inhibitor binding site. Conversely, inhibition of ensuing secondary RNase H cuts is much more efficient than inhibition of primary RNase H cuts, which points to increased access following the first cut. The increased flexibility of the cleaved substrate may facilitate binding of the inhibitor. The various product complexes are structurally distinct from the original enzyme-substrate complex, which translates into non-competitive inhibition under steady-state conditions, as previously published (276). However, our data show that under pre-steady state conditions, binding of β -thujaplicinol and the intact DNA/RNA substrate is mutually exclusive and therefore competitive. We have observed a similar kinetic behavior with PFA, where binding of the inhibitor to the pre-translocated RT-substrate complex prevents the enzyme from shifting to the post-translocational state (228). Thus, PFA and the nucleotide substrate interact with the active site of conformationally distinct complexes. As with β -thujaplicinol, competitive inhibition is solely observed under single turnover conditions. Steady-state kinetics point to a non-competitive mode of inhibition, because PFA stabilizes the product complex in the pre-translocational state that prevents nucleotide binding. The nucleotide can only bind to the distinct post-translocated complex following enzyme dissociation (228).

Most importantly, our data show that the nucleic acid substrate engages the RNase H active site at the same time when the nucleotide substrate is bound at the polymerase active site and stabilizes the complex. These findings suggest that active DNA synthesis can represent an obstacle to RNase H inhibition. The model shown in Figure 2.7 is consistent with our data and provides a possible structural explanation for the competition with the nucleic acid substrate. The proposed conformational change of the side chain of Y501 appears to lock the inhibitor in a position that provokes a steric conflict with the template. In contrast, locking the side-chain in its natural orientation provides a plausible mechanism for resistance to β -thujaplicinol.

Implications for drug discovery and development efforts

Although high throughput screening of chemical libraries has led to the discovery of several compounds that inhibit the RT-associated RNase H activity, our data make clear that evaluation of RNase H inhibition under steady-state conditions has shortcomings. The design of more stringent secondary screening assays under single-turnover conditions may help to focus on the discovery of compounds that interfere directly with the RNase H activity rather than indirectly through turnover reduction. The use of stable ternary complexes with the bound nucleotide substrate or PFA provides simple versions of such an assay, and may facilitate screening for compounds capable of blocking RNase H activity during DNA synthesis. It remains to be seen whether the proposed steric clash with the intact nucleic acid substrate can be prevented with different classes of RNase H inhibitors, e.g. compounds that may not interact with Y501 of the RNase H primer grip as inferred for β -thujaplicinol and N-acylhydrazones (292). At the

same time, this work raises the question of whether potent inhibition of HIV RNase H in cell culture or *in vivo* requires blockage of primary/internal RNase H cuts.

2.5 Materials and Methods

Expression and Purification of HIV-1 RT Variants

Heterodimeric p66p51 HIV-1 RT was expressed and purified as described (344). The following protocol was used to introduce natural and unnatural amino acids in p66. The coding region of RT p66 was cloned into the vector pRSET to introduce a C-terminal H₆-affinity tag. The coding region of p51 RT was incorporated into the bacterial expression vector pPR-IBA2, introducing a Strep-tag at the N-terminus to facilitate purification. For incorporation of non-natural amino acids, the codon for Y501 was mutated to TAG, generating plasmid pRSET-p66His-501Stop. pRSET-p66His-501Stop was co-transformed with either pSup-BpaRS-6TRN or pSup-pAzPheRS-6TRN into bacterial strain BL21(DE3) (345). A single colony was used to inoculate in the presence of 50 µg ml⁻¹ ampicillin, 50 µg ml⁻¹ chloramphenicol and 1 mM of the unnatural amino acid. Mutations Y501W and Y501F were introduced using the QuikChange site-directed mutagenesis kit. Protein expression and purification was conducted essentially as described (346).

Nucleic Acids

Oligodeoxynucleotides used in this study were chemically synthesized and purchased from Invitrogen. The following sequences were used: PBS-22Dpol, 5'

AGGTCCCTGTTTCGGGCGCCACT 3'; PBS-21Dpol 5'
AGGTCCCTGTTTCGGGCGCCAC 3'.

PBS-22Dchi, 5' CTAGCAGTGGCGCCCGAACAGG 3'. Both PBS-22Dpol and PBS-21Dpol are referred to as polymerase-dependent substrates, unless otherwise indicated.

PBS-52R oligoribonucleotide was synthesized by *in vitro* transcription using T7 polymerase, 5'

GGAAAUCUCUAGCAGUGGCGCCCGAACAGGGACCUGAAAGCGAAAGGGAA
AC 3'. The chimeric RNA-DNA oligonucleotide, PBS-14R8D, 5'
cuguucgggcgccaCTGCTAGA 3' was purchased from Trilink. Nucleotides,
deoxynucleotides and dideoxynucleotides were purchased at Fermentas Life Sciences,
phosphonoformic acid was purchased from Sigma.

Time-Course of RNase H Activity

3'-radiolabeled RNA template (or RNA/DNA chimera) was heat annealed to a 2-fold excess of DNA primer. The RNA/DNA hybrid was added to the reaction at a final concentration of 150 nM containing 25 nM RT in a buffer of 50 mM Tris-HCl, pH 7.8, 50 mM NaCl, 100 μ M EDTA and 6 mM MgCl₂, in the absence or presence of β -thujaplicinol (50 μ M) or PFA (200 μ M) unless otherwise indicated. Reactions in the absence of β -thujaplicinol contain 0.1% dimethyl sulphoxide (DMSO). The reaction is allowed to proceed at 37°C and stopped by the addition of 100% formamide containing 0.01% (w/v) xylene cyanol and bromophenol blue. RNA fragments were resolved on a 12% denaturing polyacrylamide gel and visualized by phosphorimaging.

Pre-Steady-State Single Turnover Reactions

Experiments conducted in pre-steady-state used a Kin-Tek RQF-3 rapid quench-flow apparatus (www.kintek-corp.com). RNA/DNA hybrids were prepared as described above to a final concentration of 50 nM with 500 nM RT in a buffer containing 50 mM Tris-HCl, pH 7.8, 50 mM NaCl, 200 μ M EDTA and 6 mM MgCl_2 , in the presence or absence of β -thujaplicinol or PFA. Reactions in the absence of β -thujaplicinol contain 0.1% DMSO. Single-turnover conditions were provided by a 10-fold excess of RT over primer/template, and confirmed by non-linear regression analysis and the addition of heparin at 4 mg/mL when possible. All reactions are terminated by addition of 0.5M EDTA, and 100 μ L of 100% formamide. Curves were fitted to a single-phase exponential association equation [$Y=Y_{\text{max}}*(1-\exp(-K*X))$] using the program GraphPad Prism 4.0. The rate constant $k_{(\text{RNase H})}$ was obtained from this function.

Kinetic Analysis

Kinetic parameters are obtained using the program GraphPad Prism 4. $K_{\text{d}(\text{RNase H})}$ values were obtained by adding variable amounts of heat annealed RNA/DNA hybrid (1.2 μ M to 0.01 μ M) to 1 μ M RT in a buffer containing 50 mM Tris-HCl, pH 7.8, 50 mM NaCl, and 500 μ M EDTA. MgCl_2 and heparin were added simultaneously at final concentrations of 6 mM and 4 mg/mL, respectively. $K_{\text{d}(\text{Pol})}$ values were obtained similarly, except the next templated nucleotide (dGTP) was added together with MgCl_2 and heparin at a final concentration of 25 μ M. The resulting points were plotted and fit to a quadratic equation

$[Y=0.5(K+E+X)-(0.25(K+E+X)^2-(E)X)^{0.5}]$ using the program GraphPad Prism 4.0 to obtain the equilibrium binding constant K_d .

Single-Turnover Dose-Responses

Reactions were prepared as described above except 300 μ M EDTA was used, and heparin was added at 4 mg/mL. Dideoxythymidine triphosphate (ddTTP) was added at a concentration of 10 μ M, while increasing concentrations of dGTP were added at 0, 0.039, 0.078, 0.156, 0.313, 0.625, 1.25, 2.5, 5 and 10 μ M when ddTTP was added. PFA was added at concentrations of 0, 0.049, 0.098, 0.195, 0.391, 0.781, 1.563, 3.125, 6.25, 12.5 μ M.

Ternary Complex Formation

DNA/RNA primer/template hybrids were prepared as described above. α -thujaplicinol was added at 100 μ M and PFA at 50 μ M. dGTP was added at a final concentration of (100 μ M). The DNA primer was PBS-22Dpol except for reactions containing dGTP, in which case the primer was PBS-21Dpol, and the reaction includes ddTTP at a final concentration of 20 μ M. RNA/DNA hybrid was added at a final concentration of 300 nM to 50 nM RT in a buffer containing 50 mM Tris-HCl, pH 7.8, 50 mM NaCl, 100 μ M EDTA and 6 mM $MgCl_2$. Reactions were allowed to proceed at 37°C and stopped with excess of 100% formamide containing 0.01% (w/v) xylene cyanol and bromophenol blue. Samples collected after 0.5, 1, 2, 3, 4, 5 and 6 minutes, were fractionated on a 12% polyacrylamide gel and visualized by phosphorimaging.

Order-of-Addition Experiments

Reactions were prepared essentially as described above. Reactions were started with either RNA/DNA hybrid, MgCl_2 , or RNA/DNA hybrid as well as MgCl_2 . Heparin trap was not added due to a conflict with the experimental design. Single turnover conditions were ensured by nonlinear regression analysis using the software GraphPad Prism 4.0.

RNase H IC_{50} determination for primary and secondary cleavages

Reactions were prepared essentially as described above. To determine IC_{50} values for secondary cleavages, 50 nM RNA-DNA/DNA chimeric hybrid (PBS-14r8d/PBS-22d) labelled at the 5' RNA end was added to 50 nM RT in increasing concentrations of β -thujaplicinol (0, 0.098, and doubling until 100 μM). Reactions were allowed to proceed at 37° for 8 minutes then stopped with 100% formamide containing 0.01% (w/v) xylene cyanol and bromophenol blue. For primary cleavages, 600 nM of chimeric hybrid (PBS-14r8d/PBS-22d) labelled at the 3' DNA end was added to 50 nM RT in increasing concentrations of β -thujaplicinol (same as above). Reactions were allowed to proceed and were stopped as described above.

FRET-Based RNase H Assay

RNaseH assays based on fluorescence-resonance energy transfer were performed as described (339).

Molecular Modeling

Molecular models were generated using Discovery Studio 7.0 (DS 7.0; Accelrys) using structural coordinates of HIV RT-RNA/DNA (1HYS; (57)) and human RNase H1-RNA/DNA (2KQ9; (214)) complexes downloaded from the protein data bank. The two complexes were overlaid by superposition of four carboxylate-containing catalytic residues common to the RNase H domains of the two enzymes. The panels of Figure 2.7 depict some or all of the following components of the overlaid complexes: The HIV-1 RNase H domain (from 1HYS); RNA/DNA (from 2KQ9); and site A and B Mg ions (2KQ9). β -thujaplicinol was constructed using the 'build' features of DS 7.0. Because it possesses structural features suggesting metal ion chelation, the inhibitor was manually docked into the RNase H active site at a position which permits contact with the B-site Mg^{2+} ion and stacking with Y501 (panel [A]) (292). The benzophenone side chain (panel [B]) was constructed by deleting the hydroxyl group of and extending the Y501 side chain. The trajectory of the unnatural side chain reflects original positioning of Y501 (as given in 1HYS), which otherwise has not been altered.

2.6 Figures and Figure Legends

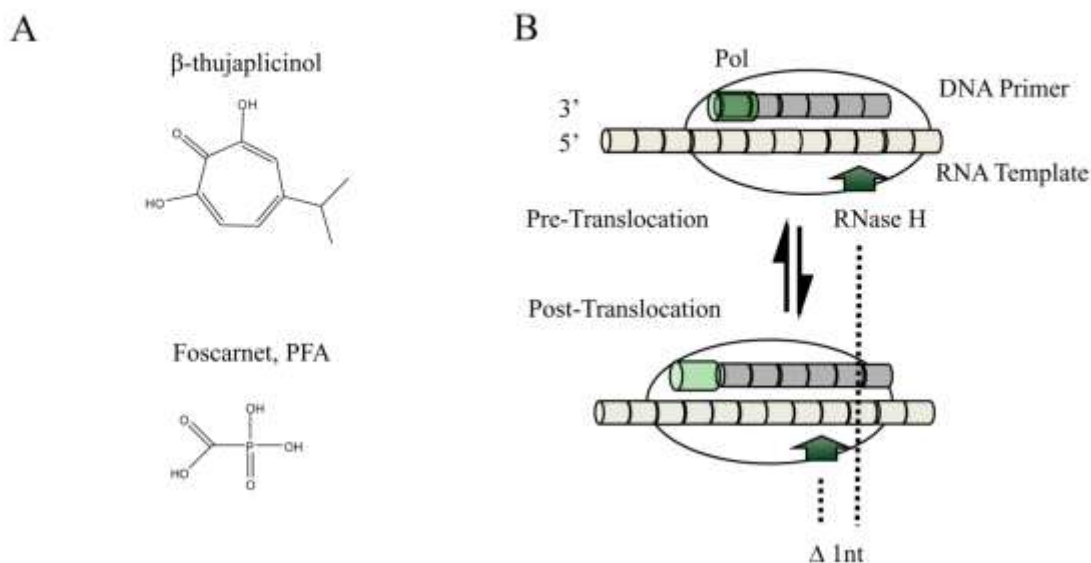


Figure 2.1. Structures of Small Molecules used in this Study and Schematic of the Polymerase-Dependent Binding Mode of HIV RT

(A) Structures of inhibitors used in this study: β -thujaplicinol and PFA.

(B) Schematic representation of pre- and post-translocated complexes of HIV-1 RT. The polymerase active center is represented by the green cylinder, which is either occupied by the 3' end of the primer (pre-translocation) or available for nucleotide binding (post-translocation). RT is shown here shifted relative to its nucleic acid substrate by a single nucleotide. This difference (Δ 1nt) is also reflected at the RNase H active site (green arrow).

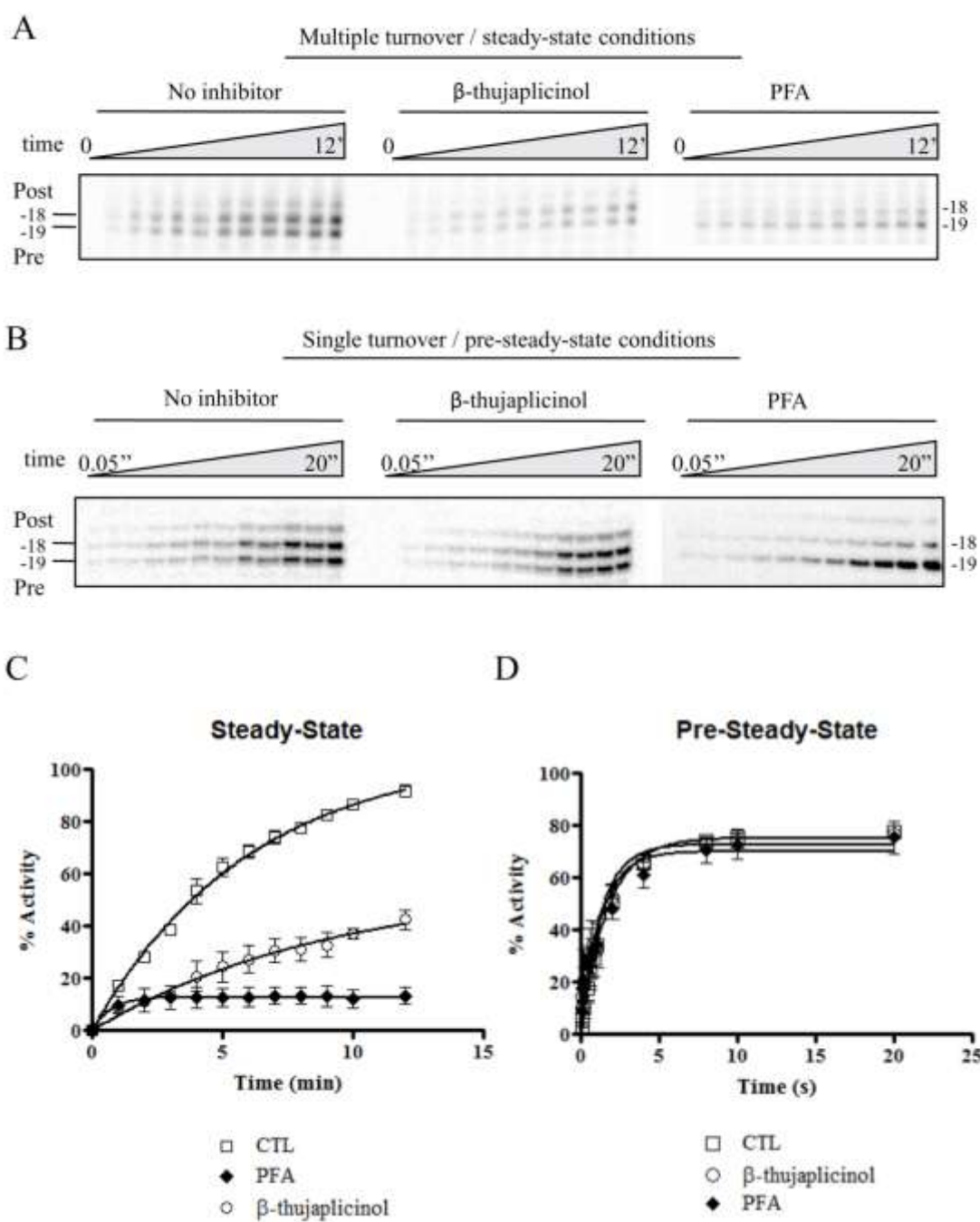


Figure 2.2

Figure 2.2: Inhibition of RNase H Activity by β -thujaplicinol and PFA

A) Inhibition of RNase H activity under steady-state conditions. Time-course reaction (0-12 mins) in the absence and presence of β -thujaplicinol (50 μ M) and PFA (200 μ M) under steady-state conditions. RNase H cleavages at positions -18 and -19 are marked post- and pre-translocation, respectively. The figure focuses on this part of the gel.

(B) Time-course experiment (0.05 – 20 sec) in the presence and absence of β -thujaplicinol (50 μ M) and PFA (200 μ M). The protein trap heparin was added to all reactions at 4 mg/mL.

(C) and (D) Results from (A) and (B) represent graphically, respectively. Error bars represent the standard deviation between three independent experimental replicates.

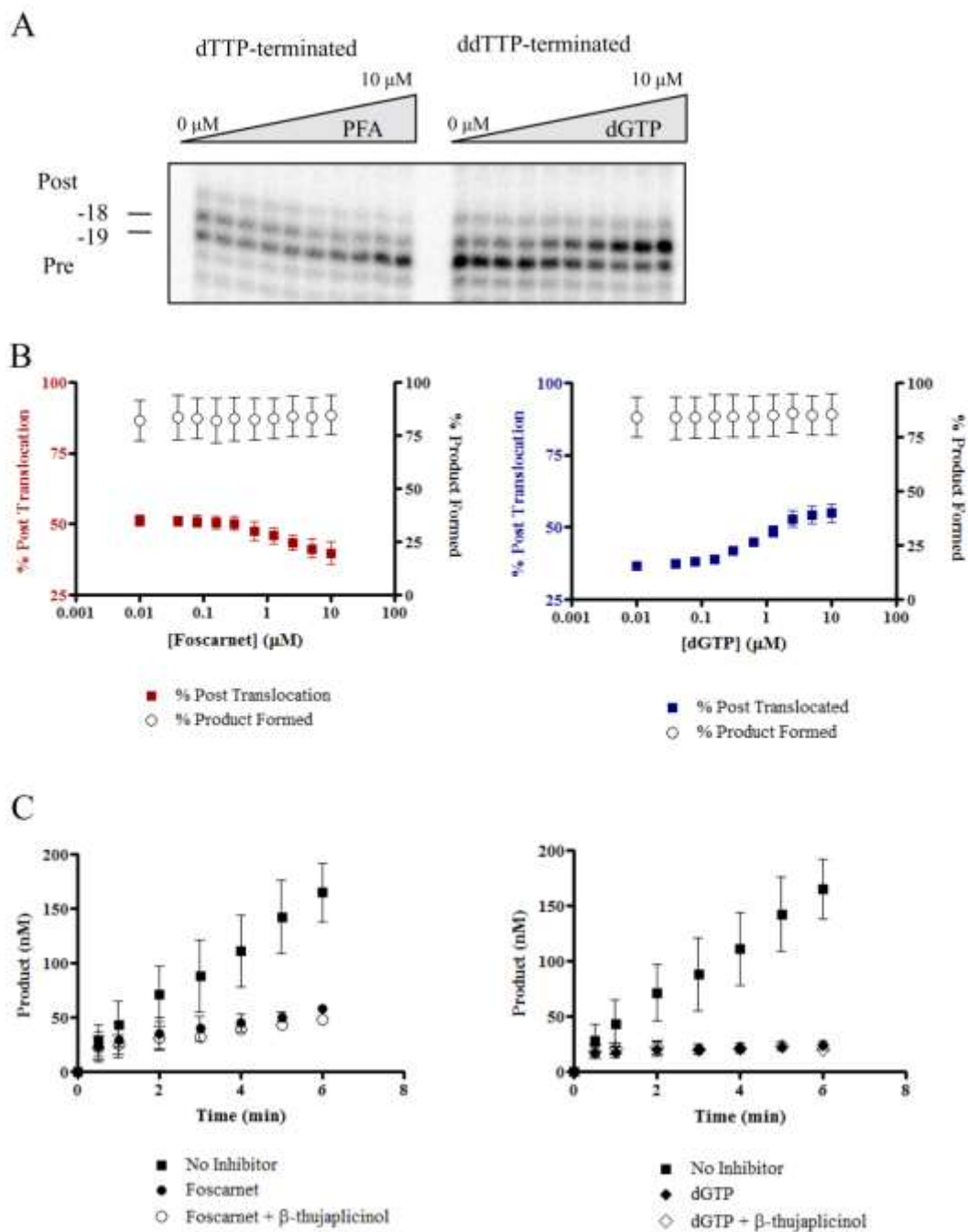


Figure 2.3

Figure 2.3: Effects of Ternary Complex Formation on RNase H Inhibition.

(A) Left: Increasing dose-response of PFA (0 – 10 μ M) on the polymerase-dependent substrate with a dTMP-terminated primer. Right: Increasing concentrations of the next template nucleotide (dGTP) were added from 0 to 10 μ M to reactions containing a polymerase-dependent substrate terminated with ddTMP.

(B) The data from (A) presented graphically showing both translocational trends and RNase H activity.

(C) Steady-state RNase H activity measured in the linear phase of the reaction in the absence of ligands, and with PFA or dGTP, in the absence and presence of β -thujaplicinol.

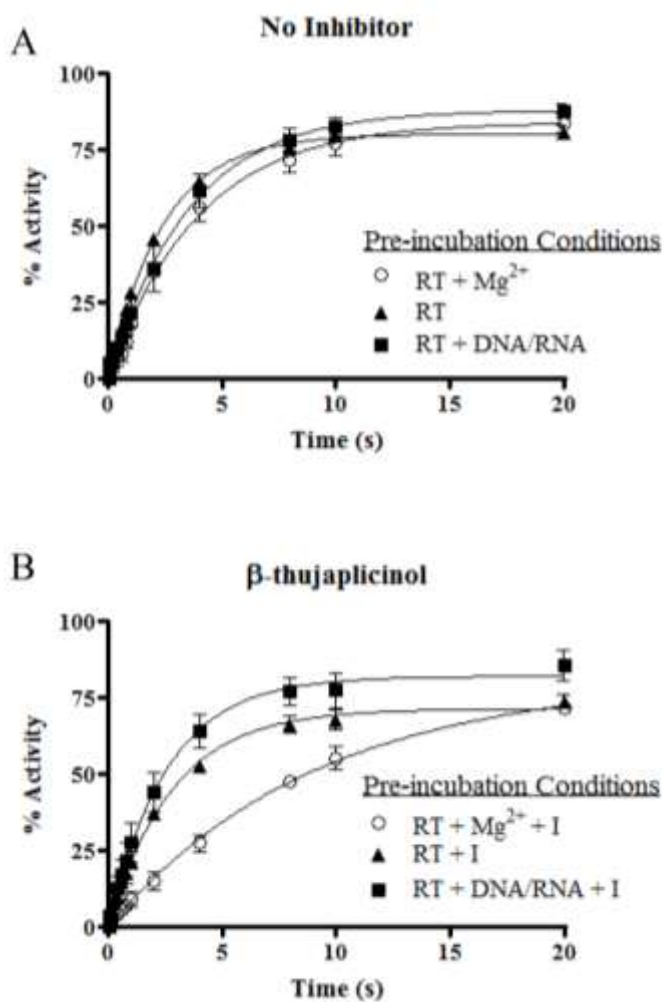


Figure 2.4: Effects of Order-of-Addition on RNase H Inhibition

Time-course assay under pre-steady-state conditions. % Activity refers to the percentage of the total substrate converted to the -18 and -19 major reaction products. (Top) Reactions were pre-incubated with RT, Mg²⁺ and the RNA/DNA substrate in the orders shown in the absence of inhibitor. (Bottom) Same as top but in the presence of 50 μM β-thujaplicinol.

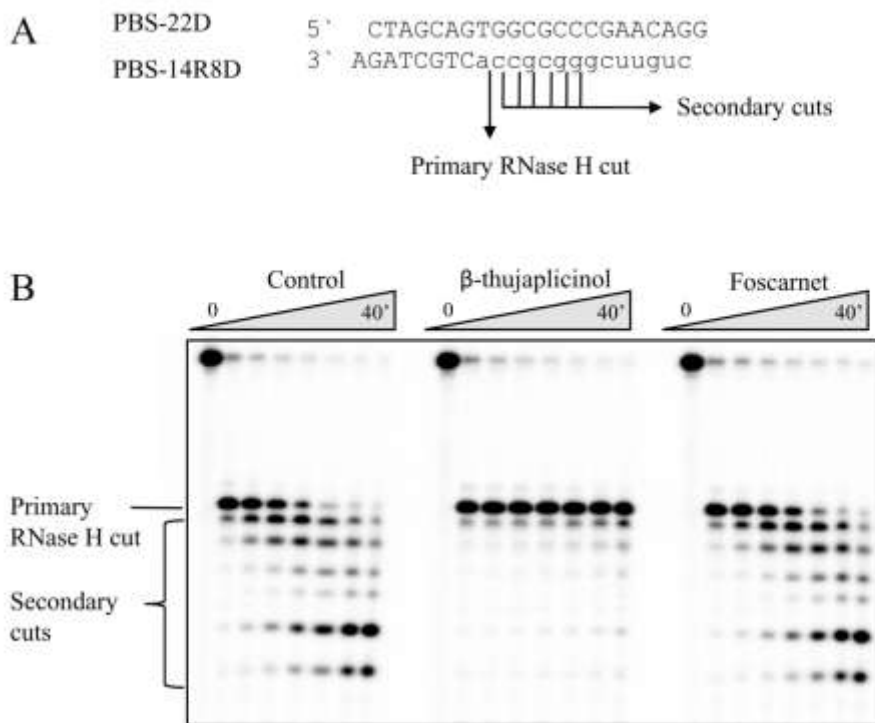


Figure 2.5: Effects of β-thujaplicinol on Polymerase-Independent RNase H activity.

(A) Sequence of the polymerase-independent substrate. A primary RNase H cut is expected at the RNA/DNA junction +1, while ensuing secondary cuts are expected to occur upstream of the primary cut.

(B) Time-course of RNase H activity (0 – 40') on the chimeric substrate PBS-14r8d. B-thujaplicinol was added at a concentration of 50 μM, while PFA was added at a concentration of 200 μM. Primary and secondary cuts are indicated.

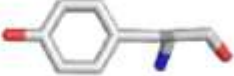

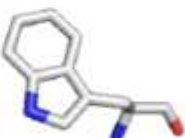
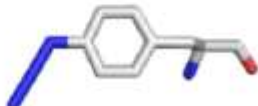
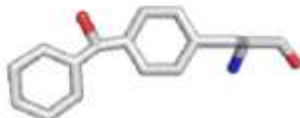
Amino acid at position 501		IC ₅₀ (nM) ± SD
 Y		308 ± 18 nM
 F		356 ± 49 nM
 W		470 ± 135 nM
 Az-F		386 ± 123 nM
 Bp-F		> 300 μM

Figure 2.6: Altering Y501 of the RNase H primer grip affects β-thujaplicinol sensitivity. Structures of natural (F and W) and unnatural amino acid insertions (AzF and BpF) for Y501 are illustrated in addition to the IC₅₀ for the mutant enzyme. IC₅₀ values are the average of triplicate analyses.

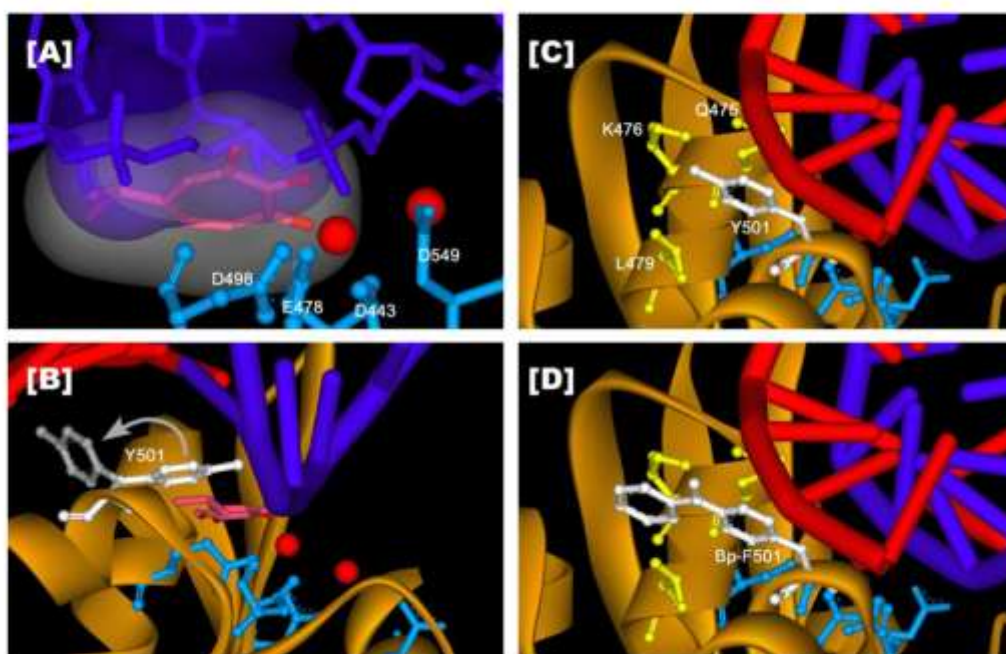


Figure 2.7.

Figure 2.7: Model of β -thujaplicinol binding site. An RT-substrate complex generated by superposition of HIV RT-RNA/DNA and human RNase H-RNA/DNA co-crystal structures (57,214) is depicted in all panels. (A) Surface representation of β -thujaplicinol (pink, with red oxygen atoms) and the RNA nucleotide 17 bp from the primer 3' terminus (blue). Steric interference between the inhibitor and the RNA is evident at the junction between the two surfaces, as well as in the stick representations near the ribose C4' atom. RNA (dark blue ball and stick), active site Mg ions (red spheres), and active site residues D443, E478, D498, D549 (light blue ball-and-stick) are also highlighted. (B) Y501 rotation from the position observed in published crystal structures (grey) to stack with β -thujaplicinol (rotated Y501, white; β -thujaplicinol, pink). For illustrative purposes, RNA and DNA strands are shown as blue and red tubes, respectively; however, binding of substrate and β -thujaplicinol at the RNase H active site is thought to be mutually exclusive. (C) Same as B, except the complex is rotated to highlight the positioning of Y501 relative to L479, K476, and Q475 (yellow ball-and-stick), and β -thujaplicinol is not shown. (D) Same as C, with benzophenone substituted for the Y501 side chain. Note that while the extended side chain permits potential hydrophobic interactions with L479 and K476 that cannot occur with Y501, it is not likely to directly affect substrate binding.

Chapter 3

GSK5750 is a novel active site inhibitor of HIV-1 RT-associated RNase H activity

Greg L. Beilhartz¹, Brian A. Johns², Peter Gerondelis³, Matthias Götte¹

¹Dept. of Microbiology and Immunology, McGill University, Quebec, Canada

²Dept. of Medicinal Chemistry and ³Dept. of Virology, Infectious Diseases Therapeutic Area Unit, GlaxoSmithKline, 5 Moore Drive, Research Triangle Park, NC 27709

3.1 Preface

The characterization of the active site RNase H inhibitor β -thujaplicinol provided us with a model of inhibitor binding. This was supplemented by the publication of the crystal structure of RT in complex with β -thujaplicinol shortly thereafter. It seemed that both our data and the crystal structure agreed on a model where β -thujaplicinol bound to the free enzyme (and/or potentially the product complex) and this resulted in a steric clash with the natural nucleic acid substrate. However, the main failure of β -thujaplicinol as an RNase H inhibitor was that it did not compete well with the natural substrate, and as such did not inhibit RNase H activity under relevant conditions. We identified this as a potential intrinsic biochemical obstacle to the development of an RNase H inhibitor that would be active in cell-based assays. We hypothesized that better competition with the nucleic acid would result in a more potent RNase H inhibitor. To that end, we performed a biochemical characterization of the active site inhibitor GSK5750, presented in Chapter 3. Due to its tighter binding characteristics, we were able to perform experiments that were not possible with β -thujaplicinol.

3.2 Abstract

Highly Active Anti-Retroviral Therapy (HAART) has been very successful in reducing morbidity and mortality among HIV-infected patients. The rapid emergence of drug resistance emphasizes the need for more drugs against different viral targets. The HIV-1 RT-associated RNase H activity is strongly inhibited by the novel compound 4-((benzo[b]thiophen-2-ylmethyl)amino)-1-hydroxypyrido[2,3-d]pyrimidin-2(1H)-one, referred to in this manuscript as GSK5750 which binds to the RNase H active center with a K_D of 386.5 ± 177.5 nM. Order-of-addition experiments show that GSK5750 binds to the free enzyme in a Mg^{2+} -dependent fashion, but is not observed to bind to a pre-formed enzyme-substrate (E-S) complex. We also provide evidence that the substrate can bind over the inhibitor in the active site, and does not provide a barrier to dissociation of the inhibitor. As we have previously shown that competition with the substrate is deleterious in inhibitory potential, GSK5750 represents an important step in this regard to the development of potent RNase H inhibitors that may soon be able to enter clinical development.

3.3 Introduction

Human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) is a heterodimeric (p66/p51) multifunctional enzyme that catalyzes the transformation of the single-stranded viral RNA genome into a double-stranded DNA provirus through two spatially distinct active sites (22). Both active sites reside on the larger, catalytic p66 subunit. The N-terminal polymerase active site catalyzes nucleotidyl transfer reactions,

while the C-terminal ribonuclease (RNase) H active site hydrolyzes the RNA strand of DNA/RNA hybrids that are generated during reverse transcription (57,184,188,210). Both active sites are necessary for successful reverse transcription, however presently all clinically available drugs that target HIV-1 RT inhibit DNA synthesis, while specific RNase H inhibitors that show antiviral effects have yet to be developed (334).

The distance between the polymerase and RNase H active sites is approximately 60 Å and can accommodate 18 base pairs (bp) of DNA/DNA or 19 bp of DNA/RNA primer/template (180,181). A ternary complex with a chain-terminated primer and a bound deoxynucleoside triphosphate (dNTP) yields a specific cut at position -19; the scissile bond being located between residues -19 and -20. Here the complex exists in its post-translocated conformation, in which the nucleotide binding site is accessible. In contrast, a pre-translocated complex that exists immediately following the nucleotidyl transfer yields a cut at position -18. Such a complex can be trapped in the presence of the pyrophosphate (PPi) analogue phosphonoformic acid (PFA) (209). In general, these cuts are referred to as polymerase-dependent RNase H activity. Conversely, polymerase-independent cuts collectively refer to RNase H cleavage events that occur while the polymerase active site is not positioned at the 3' primer terminus, and variations of this mechanism have been proposed (233). This activity is required for non-specifically degrading the transcribed viral RNA genome during (-)-strand DNA synthesis, as well as for specifically creating and removing the (+)-strand primer or polypurine tract (PPT), and for removing the (-)-strand or tRNA^{lys3} primer (58,237-240,247).

The RNase H active site is comprised of a DEDD catalytic motif (D443, E478, D498, D549) which coordinates two divalent metal ions as essential cofactors (213,214,219). This architecture has also driven the design of RNase H small molecule antagonists. Several distinct chemical compounds have been published that all contain a 3-oxygen pharmacophore, or equivalent of which, that is capable of chelating the catalytic metal ions. As such, these compounds are referred to as active site RNase H inhibitors (ASRIs). Several crystal structures of HIV-1 RT RNase H with a bound ASRI have confirmed that these compounds are anchored at the RNase H active site via the 3-oxygen pharmacophore (281,287,293,347,348). Other inhibitors have been shown to bind allosterically in either a binding site overlapping the NNRTI binding pocket in the case of N-acyl hydrazones (292,293) or in the vicinity of the p51 thumb subdomain in the case of the vinylogous ureas (294,349).

Order-of-addition experiments revealed that the tropolone ASRI β -thujaplicinol does not bind to a pre-formed RT-nucleic acid complex (277). Inhibition of a primary, endonucleolytic RNase H cut is only seen if the inhibitor is pre-incubated with RT and Mg^{2+} ions. The added DNA/RNA primer/template is completely cleaved within a single turnover event (30 seconds), suggesting that inhibitor and substrate compete for the same binding site. However, β -thujaplicinol appears to be able to bind to the nicked product of the primary cleavage reaction as subsequent secondary cuts are effectively inhibited (277). We proposed that the inhibitor was competing with the nucleic acid substrate for access to the RNase H active site (277). The structure of RT in complex with β -thujaplicinol, along with a modeled RNA/DNA hybrid, point to a steric conflict between

the inhibitor and the substrate (347). In this model, inhibition of a primary cleavage event may occur if the substrate binds to a pre-formed RT-inhibitor complex. Thus, the bound substrate may even trap the inhibitor, provided that contacts between the RT-associated polymerase domain and RNA/DNA substrate are retained under these circumstances. Regardless of the precise mechanism, the reversible nature of binding of both inhibitor and substrate may ultimately lead to RNase H cleavage unless the affinity between RT and ASRI can be significantly improved.

In this study, we show that a novel ASRI, GSK5750 (4-((benzo[b]thiophen-2-ylmethyl)amino)-1-hydroxypyrido[2,3-d]pyrimidin-2(1H)-one), is a potent inhibitor of the RT-associated RNase H activity. In contrast to β -thujaplicinol, GSK5750 can form long lasting complexes with the enzyme.

3.4 Results

GSK5750 is a specific RNase H inhibitor

The two RT inhibitors used in this study are shown in Fig. 3.1. In order to determine the most likely binding site for GSK5750 we performed a series of experiments to discount other potential binding sites on HIV-1 RT. We employed a FRET-based assay to directly study the effects of GSK5750 on the polymerase activity of RT (Fig. 3.2A). We do not observe any inhibition until the highest concentration ($IC_{50} = >30 \mu M$). As a control we performed a titration of the polymerase inhibitor PFA (Fig. 3.2B). The curve produced an IC_{50} value of $0.244 \pm 0.019 \mu M$, in agreement with previously published results (350). Finally, GSK5750 inhibited RNase H activity against the polymerase active site double-

mutant D185N-D186N which prevents Mg^{2+} binding to the polymerase active site. These data collectively demonstrate that GSK5750 binds to Mg^{2+} ions in the RNase H active site of HIV-1 RT. Since GSK5750 fails to inhibit polymerase activity at concentrations where 100% of RNase H inhibition is observed strongly suggests that GSK5750 can be bound to RT simultaneously with the substrate, forming an enzyme (E) –substrate (S) –inhibitor (I), or E-I-S complex.

Product inhibition by GSK5750

To assess how GSK5750 interferes with primary and secondary cleavage events, we used a chimeric DNA-RNA/DNA primer/template system that mimics the tRNA (-)-strand primer removal reaction (Fig 3.2). In this experiment, we observe a primary RNase H cleavage one bp downstream of the DNA-RNA junction, followed by downstream secondary cleavages, as previously shown (277). GSK5750 strongly inhibits secondary cleavages, but only weakly inhibits the primary cleavage with IC_{50} values of approximately 600 nM and $>50 \mu M$ for the secondary and primary cleavages respectively (Fig. 3.3). The efficiency of inhibition of secondary cuts is higher as seen with β -thujaplicinol, while both compounds show insignificant inhibition of primary cuts under these conditions.

Pre-Steady-State analysis

Next, we performed order-of-addition experiments to determine if GSK5750 binds preferentially to the free enzyme, in the absence and presence for divalent metal ions. The experimental setup was as follows: (i) pre-incubation of RT with DNA/RNA and

GSK5750, starting the reaction with MgCl_2 , (ii) pre-incubation of RT with GSK5750 alone, starting the reaction with DNA/RNA and MgCl_2 , (iii) pre-incubation of RT with GSK5750 and MgCl_2 , and starting the reaction with DNA/RNA. A polymerase-dependent substrate (recessed 3' primer end) was chosen to easily observed the primary cleavage reaction, which is the focus of pre-steady-state studies (Fig. 3.4). The order-of-addition in the absence of inhibitor had no significant effect on the rate of RNase H cleavage nor the amount of product formed (data not shown). Inhibition was observed solely when GSK5750 was pre-incubated with enzyme and catalytic metal ions, which provides strong evidence for active site binding (Fig. 3.4). Intriguingly, inhibition of RNase H activity is sustained over 30 seconds at a point when substrate cleavage rebounds in the presence of β -thujaplicinol (277). The data suggests that binding of GSK5750 to the RNase H active site is improved, while both compounds are unable to bind to a pre-formed enzyme-substrate (E-S) complex.

Effects of complex stability on dissociation and slow association kinetics of GSK5750

As mentioned above, Figures 3.2 and 3.6 provide compelling evidence that GSK5750 can form an E-I-S complex, and as such we next investigated the effect of stable complex formation on the binding of GSK5750. To test the hypothesis that the nucleic acid substrate can bind to a pre-formed E-I complex, we designed an experiment to see whether or not the substrate must dissociate in order for GSK5750 to dissociate, or if GSK5750 can dissociate even in the presence of bound substrate. We generated an E-I complex which consisted of GSK5750 bound to free enzyme in the presence of MgCl_2 . Then the substrate was added in the presence and absence of PFA (Fig. 3.5A). We

followed this reaction for 6 hours at which time all the substrate was cleaved, implying that all the GSK5750 had dissociated from the complex. The substrate was cleaved very quickly (within 1 minute) in the absence of inhibitor or in the presence of PFA. When GSK5750 was present, the RNase H activity progress curves were similar in the presence of PFA as in its absence, suggesting that GSK5750 dissociation from RT is unaffected by the stable E-I-S-PFA complex. The experiment was designed so that all enzymes are bound with GSK5750, and therefore all substrate is bound by E-I complexes. Therefore, the rate of RNase H cleavage is approximately equal to the rate of GSK5750 dissociation. This rate was found to be 0.013 min^{-1} for GSK5750 alone, and 0.010 min^{-1} in the presence of GSK5750 and PFA (Fig 3.5B).

To determine how GSK5750 associated with RT, we investigated the effect of the pre-incubation period on RNase H inhibition by GSK5750 (Fig. 3.5C). In this assay, we varied the time from formation of the pre-incubation complex as described in Fig. 3.5A to the time of addition of the substrate (reaction start). We observed that maximum inhibition by GSK5750 occurred only after approximately 8 minutes of incubation at 37° . This shows that GSK5750 can be qualified as a slow-binding inhibitor.

Determination of k_D

To further assess binding of GSK5750 to HIV-1 RT, we were able to determine the equilibrium dissociation constant (k_D) of GSK5750 (Fig. 3.6A and B). RNase H activity was determined on a DNA/RNA primer/template with a recessed 3' primer end (same as Fig. 3.4) in the presence of varying concentrations of GSK5750. The resulting curves

show a change in the maximum product (Y_{\max}) but not in the observed rate (k_{obs}).

Plotting the values for Y_{\max} against inhibitor concentration gives a K_D value for GSK5750 of 386.5 ± 177.5 nM.

3.5 Discussion

The development of potent and specific inhibitors of HIV-1 RT-associated RNase H activity has been hindered back a lack of potency in cell-based assays. Despite the development of compounds that are quite potent and selective *in vitro*, not a single small molecule that targets RNase H has entered clinical development. This is at least partly due to the lack of a deep binding pocket for small molecules to exploit. However, the recent approval of boceprevir and telaprevir for the hepatitis C virus NS3-4A serine protease demonstrate that effective small molecules can be designed for extremely shallow binding sites (351-353). Other problems could include low cell penetration or cellular toxicity. Recently, we provided evidence that another problem in designing effective RNase H inhibitors is the nucleic acid substrate itself. We showed that the tropolone derivative β -thujaplicinol is unable to bind to a pre-formed E-S complex, but only to the free enzyme (E) or the product complex (P) (277). This competition with the bound substrate could be partly responsible for the lack of activity observed *in vivo*. GSK5750 appears to follow the same pattern, as shown in figure 3.2. On a chimeric RNA-DNA/DNA substrate, RT makes a primary cleavage at the RNA-DNA junction +1 (Fig. 3.3A). This cut is very weakly inhibited by GSK5750, with an IC_{50} of >50 μ M. After this cleavage, RT makes secondary cleavages downstream of the primary cut. These

secondary cleavages use the product of the first cleavage as a substrate, and are efficiently inhibited by GSK5750, with an IC_{50} of 600 nM, a difference of at least 100-fold compared to the primary cut (Fig. 3.3B and C). This suggests that GSK5750 can bind to the product complex after an initial cleavage has been made. We hypothesize that this cleavage may provide more physical space for the inhibitor to bind to the active site.

Possible E-I-S complex formation by GSK5750

Order-of-addition experiments show that with a canonical DNA/RNA substrate, GSK5750 is unable to bind to a pre-formed E-S complex (Fig. 3.4). These experiments also provide strong evidence that GSK5750 is bound at the RNase H active site through a metal ion-chelation mechanism, as inhibition is $MgCl_2$ -dependent. When GSK5750 is allowed to bind to the free enzyme before the reaction is started, no RNase H activity is observed over the time course of a single turnover event. Therefore, like β -thujaplicinol, GSK5750 is denied access to the Mg^{2+} ions in the RNase H active site by the bound substrate. However, once bound GSK5750 can tolerate the presence of the substrate, unlike β -thujaplicinol which dissociated from the complex within the timeframe of a single enzyme turnover. Whereas a dissociation constant (k_D) for GSK5750 was shown to be 386.5 ± 177.5 nM, the k_D for β -thujaplicinol was not measureable by this method due to lack of inhibitory potency under single turnover conditions. We interpret these data to mean that the E-I-S complex formed with GSK5750 is significantly more stable than with β -thujaplicinol. Furthermore, timecourse experiments with various concentrations of GSK5750 show changes in Y_{max} , but not in k_{obs} , which is consistent with non-competitive or mixed type inhibition, in agreement with previous studies on β -

thujaplicinol (Fig. 3.6) (276,347). In the case of GSK5750, evidence for the formation of an E-I-S complex is provided by figure 3.2. In figure 3.2, GSK5750 does not inhibit the polymerase activity at concentrations where 100% inhibition of RNase H activity is observed. If the substrate could not bind to E-I complexes, we would expect to see inhibition of nucleotide incorporation as well as RNase H activity. Further evidence for the existence of an E-I-S complex is provided by figure 3.6. Single turnover conditions are assured in our pre-steady-state assay at an enzyme:substrate ratio of 5:1 or greater (data not shown). In figure 3.6A, that ratio is 10:1 (1000 nM:100 nM), and we observed inhibition at GSK5750 concentrations of 250 nM and 500 nM. Even if we assume 100% inhibitor binding to free enzyme, at an inhibitor concentration of 500 nM the free enzyme:substrate ratio would be reduced to 5:1, which is still defined as single turnover conditions and therefore the substrate would be completely cleaved. Since inhibition is observed at these concentrations, this suggests that substrate molecules are being protected by the formation of an E-I-S complex. The K_D of the substrate was previously shown to be 15 nM as measured from the RNase H active site. Therefore, the substrate still binds tighter to RT as compared with GSK5750. However, a realistic goal for active site RNase H inhibitors should be the development of a compound with a lower K_D than the substrate (sub-nanomolar). At this point, the substrate would dissociate before the inhibitor, resulting in potentially very potent inhibition. GSK5750 represents an important step in that direction.

Dissociation of GSK5750 is unaffected by stable complex formation

We have shown that GSK5750 must bind to the free enzyme in a Mg^{2+} -dependent fashion, and that the nucleic acid substrate can then bind to form an E-I-S complex. The reverse is not possible however, as GSK5750 will not bind productively to a E-S complex. It is known that PFA binds at the polymerase active site and stabilizes the pre-translocational complex, thus reducing dissociation of the substrate. We wanted to investigate whether the binding of PFA to the E-I-S complex would prevent dissociation of GSK5750 by preventing dissociation of the substrate from the E-I-S complex. We designed a timecourse experiment where GSK5750 was in excess of RT, and in turn RT was in excess of substrate. This experimental design ensured that there is no free RT enzymes in the reaction, and therefore all substrate molecules would be bound by E-I complexes, forming E-I-S complexes. In the absence of inhibitor, all of the substrate is cleaved within 1 minute, which was expected since all substrate molecules would be immediately bound by a large excess of RT (10-fold molar excess). The presence of PFA did not affect the rate of RNase H cleavage, as previously demonstrated, although only the -19 pre-translocational cleavage is visible due to the stabilization of the pre-translocational complex, in agreement with previously published results (Fig. 3.5A) (277). In the presence of GSK5750, the rate of RNase H cleavage is reduced and complete cleavage of the substrate is achieved after approximately 6 hours (Fig. 3.5A). RNase H cleavage therefore represents an E-I-S complex that has dissociated to form an E-S complex. Since GSK5750 cannot re-associate to an E-S complex, after inhibitor dissociation, the substrate is cleaved. Therefore, we can interpret the rate of RNase H cleavage of the E-I-S complex to be the minimum rate of dissociation of GSK5750. As

mentioned above, PFA reduces dissociation of the substrate from RT. We wanted to see if PFA bound at the polymerase active site affected the dissociation of GSK5750. If PFA is added with substrate to form an E-I-S-PFA complex, the rate of RNase H cleavage (and therefore GSK5750 dissociation) is unaffected. The presence of PFA in the complex is confirmed by the RNase H cleavage pattern i.e. only the -19 cut is visible. This strongly suggests that GSK5750 can dissociate from the E-I-S-PFA complex without dissociation of the substrate.

GSK5750 is a slow-binding inhibitor

Many of our experiments require the construction of a “pre-incubation complex” that exists before the reaction is started. We wanted to see if the length of time of such a pre-incubation was relevant to the inhibition by GSK5750. By varying the time between construction of the pre-incubation complex and the start of the reaction, we discovered that maximum inhibition by GSK5750 was only achieved after a pre-incubation of approximately 8 minutes. Therefore, we concluded that GSK5750 represents a slow-binding inhibitor.

Implications for RNase H Inhibition

We have previously demonstrated that the nucleic acid substrate presents an obstacle to binding for the active site RNase H inhibitor β -thujaplicinol. Similarly, the substrate blocks binding of GSK5750 to the RNase H active site. However, GSK5750 is a potent inhibitor of RNase H activity in stringent pre-steady-state assays, whereas β -thujaplicinol was not (277). This is likely due to tighter binding in the active site, or perhaps positive

interactions with the substrate, as suggested by Himmel *et al.* for β -thujaplicinol (347). Since the substrate is an impediment to inhibitor binding, it may also be an impediment to inhibitor dissociation. We have shown here that GSK5750 can dissociate from an E-I-S complex without prior dissociation of the substrate, by stabilizing the E-I-S complex with PFA. This suggests that the substrate does not block dissociation of the bound inhibitor, although it does block binding of free inhibitor. We have previously presented a model of substrate binding that allows for partial dissociation and re-association of the substrate around the RNase H active site, while it remains fixed at the polymerase active site. We have termed this the “breathing” of the substrate around the RNase H active site (277). Our data here support this hypothesis. GSK5750 may have transient access to the RNase H active site on an E-S complex while the substrate “breathes”, however since GSK5750 is a slow-binding inhibitor, this may not provide enough time for productive binding. Therefore, GSK5750 must bind to free enzyme, but during dissociation has transient access to the solvent that allows for inhibitor dissociation.

GSK5750 is a slow-binding, tight-binding novel inhibitor of HIV-1 RT-associated RNase H activity. Although the bound substrate is an obstacle to inhibitor binding, it can bind tightly to the free enzyme, and form long-lasting E-I-S complexes. The ability of GSK5750 to bind tightly to RT may circumvent the obstacle presented by the nucleic acid substrate. As mentioned above, the key to successful active site RNase H inhibitors may be a binding affinity that is higher than that of the nucleic acid. Inhibitors based on this scaffold may help to further the development of the first clinically-relevant RNase H-specific drug.

3.6 Materials and Methods

Enzymes and Nucleic Acids - Heterodimeric HIV-1 RT (p66/p51) was expressed and purified as previously described (344). All nucleic acids used in this study were synthesized by IDT DNA Technologies. 5'-radiolabeling was performed essentially as described previously. Briefly, 5'-radiolabeling was performed with [γ -P³²]ATP (PerkinElmer Life Sciences) and T4 polynucleotide kinase (Fermentas). Reactions were allowed to proceed for 1 hour at 37°C. Labeled DNA or RNA was subjected to phenol-chloroform purification, and further purified using P-30 size exclusion columns (Bio-Rad).

FRET Assay and IC₅₀ determination – In order to determine whether GSK5750 inhibits the polymerase activity of HIV-1 RT we employed a FRET-based assay to determine IC₅₀ values for GSK5750 and PFA. The assay was performed essentially as described (354). Ten different GSK5750 or PFA was added to the FRET mixture at concentrations of 30 μ M, then descending by a factor of 1/3. Reactions were prepared in a 96-well plate format and reactions were visualized on a SpectraMax Plus microplate reader. DMSO concentrations in all GSK5750 wells including DMSO control is 1%.

IC₅₀ determination for primary and secondary RNase H cleavages – 100 nM of 5' radiolabeled PBS-14r8d (5' cuguucgggcgccaCTGCTAGA 3') was heat annealed to a 3 fold molar excess of PBS-22D (5' CTAGCAGTGGCGCCCGAACAGG 3') and allowed to cool to room temperature for 45 minutes. The resulting DNA-RNA/DNA hybrid was added to a reaction containing 500 nM HIV-1 RT in a buffer of 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 0.2 mM EDTA and varying concentrations of GSK5750. Reactions

were started with the addition of 6 mM MgCl₂, and allowed to proceed at 37°C for 6 minutes. Reactions were stopped with the addition of a 2 fold volumetric excess of 100% formamide with traces of bromophenol blue and xylene cyanol. Samples were resolved on a 12% denaturing polyacrylamide gel and visualized by PhosphorImaging. Bands were quantified by QuantityOne software (Bio-Rad). Results were graphed using GraphPad Prism 5.0.

Order-of-Addition Experiments – Experiments were conducted on a Kin-Tek RQF-3 rapid quench-flow apparatus. DNA/RNA hybrids were prepared as described above with PBS-22dpol primer (5' AGGTCCCTGTTCGGGCGCCACT 3') and 5'-radiolabeled PBS-50r template (5'

AAAUCUCUAGCAGUGGCGCCCGAACAGGGACCUGAAAGCGAAAGGGAAAC 3'). 100 nM of DNA/RNA hybrid was added to 1 μM HIV-1 RT in a buffer containing 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 1 mM EDTA and 50 μM GSK570, as well as 6 mM MgCl₂. Components in the pre-incubation mixes were incubated at 37°C for 10 minutes. Reactions were allowed to proceed for 0.05, 0.1, 0.25, 0.5, 0.75, 1, 2, 4, 8, 16, 24 and 30 seconds. Reactions were rapidly quenched with 100 μL of 0.5 M EDTA. All concentrations are after 1:1 mixing. Order-of-addition of the components varies and is reported in the text. Samples were resolved on a 12% denaturing polyacrylamide gel and visualized by PhosphorImaging. Bands were quantified by QuantityOne software (Bio-Rad). Results were graphed using GraphPad Prism 5.0.

Determination of the dissociation constant (k_D) for GSK5750 – Multiple time-course experiments were performed on a Kintek Quench-Flow apparatus as described for order-

of-addition experiments. In this case, 1000 nM HIV-1 RT was added to 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 6 mM MgCl₂ and varying concentrations of GSK5750. This mixture was rapidly mixed with 100 nM DNA/RNA hybrid (PBS-22dpol/PBS-50r) and 1 mM EDTA. Concentrations of GSK5750 were 0, 250, 500, 1000, 2000 and 5000 nM. The resulting curves were fit to a one-phase exponential equation ($Y=Y_{\max}*(1-\exp(-K*X))$). The rates of reaction (k_{obs}) were unchanged, while Y_{\max} decreased with increasing concentrations of inhibitor. This is characteristic of non-competitive or mixed inhibition. The burst amplitudes were then plotted against GSK5750 concentration and the resulting curve was fit to a quadratic equation ($Y=E-0.5((K+E+X)-[(K+E+X)^2-4(E)X]^{0.5})$) to give the k_D (inhibitor) value. The entire experiment was performed in triplicate and the reported values represent the mean $k_D \pm$ the standard deviation.

Complex Stabilization Timecourse – DNA/RNA hybrids were prepared as described above with PBS-22dpol primer (5' AGGTCCCTGTTCGGGCGCCACT 3') and 5'-radiolabeled PBS-50r template (5' AAAUCUCUAGCAGUGGCGCCCGAACAGGGACCUGAAAGCGAAAGGGAAAC 3'). 2 μ M GSK5750 or 1% DMSO was added to 500 nM HIV-1 RT in a buffer containing 50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 1 mM EDTA and, as well as 6 mM MgCl₂. Components in the pre-incubation mixes were incubated at 37°C for 10 minutes. Reactions were started with 50 nM DNA/RNA hybrid and/or 500 μ M PFA, and were allowed to continue for 0, 1, 5, 10, 30, 60, 120, 240, and 360 minutes. Samples were resolved on a 12% denaturing polyacrylamide gel and visualized by PhosphorImaging.

Bands were quantified by QuantityOne software (Bio-Rad). Results were graphed using GraphPad Prism 5.0.

Pre-incubation timecourse – Experiments were performed essentially as described above under Complex Stabilization Timecourse except GSK5750 and β -thujaplicinol concentrations are 0.5 μ M and 50 μ M, respectively. Time = 0 is defined as the addition of $MgCl_2$ to the reaction. Samples were taken at pre-incubation times of 1, 2, 4, 8, 16, 24, 32, 60 and 90 minutes. Reactions were started by the addition of DNA/RNA hybrid and allowed to proceed for 30 seconds. Samples were resolved and quantified as described above.

3.7 Figures and Figure Legends

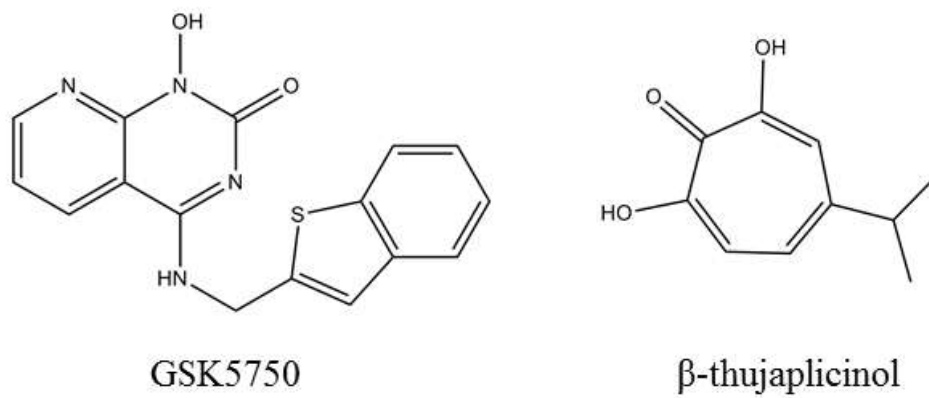


Figure 3.1: Structures of active site RNase H inhibitors GSK5750 and β -thujaplicinol.

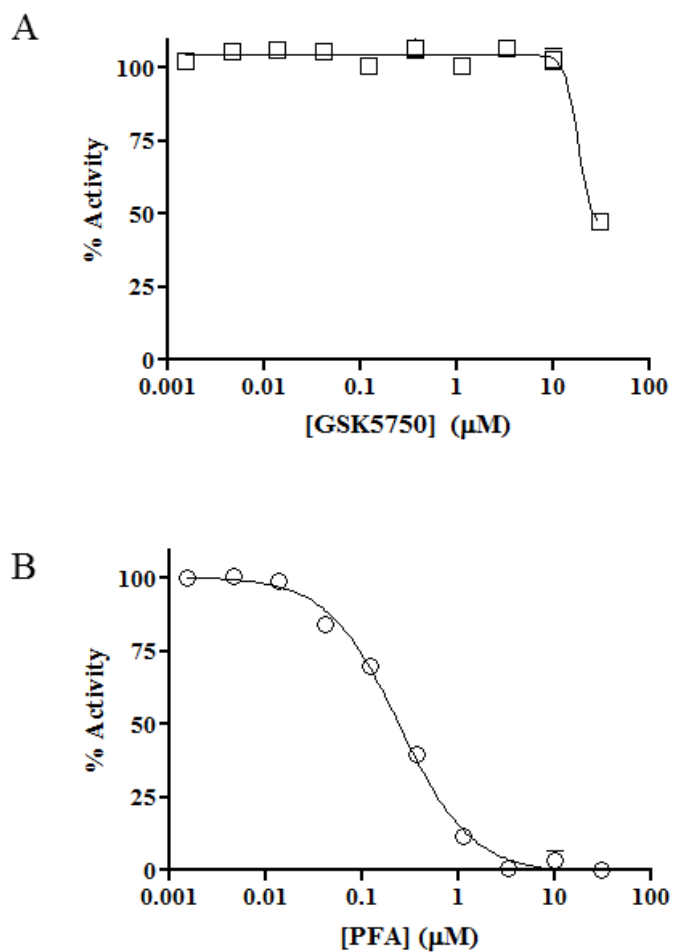


Figure 3.2: Results from the FRET assay. (A) Titration of GSK5750, with the highest concentration being 30 μM, and subsequent concentrations decreasing by a factor of 1/3. IC_{50} value is $> 30 \mu M$. (B) Same as (A) but with PFA. The IC_{50} value was found to be $0.244 \pm 0.019 \mu M$.

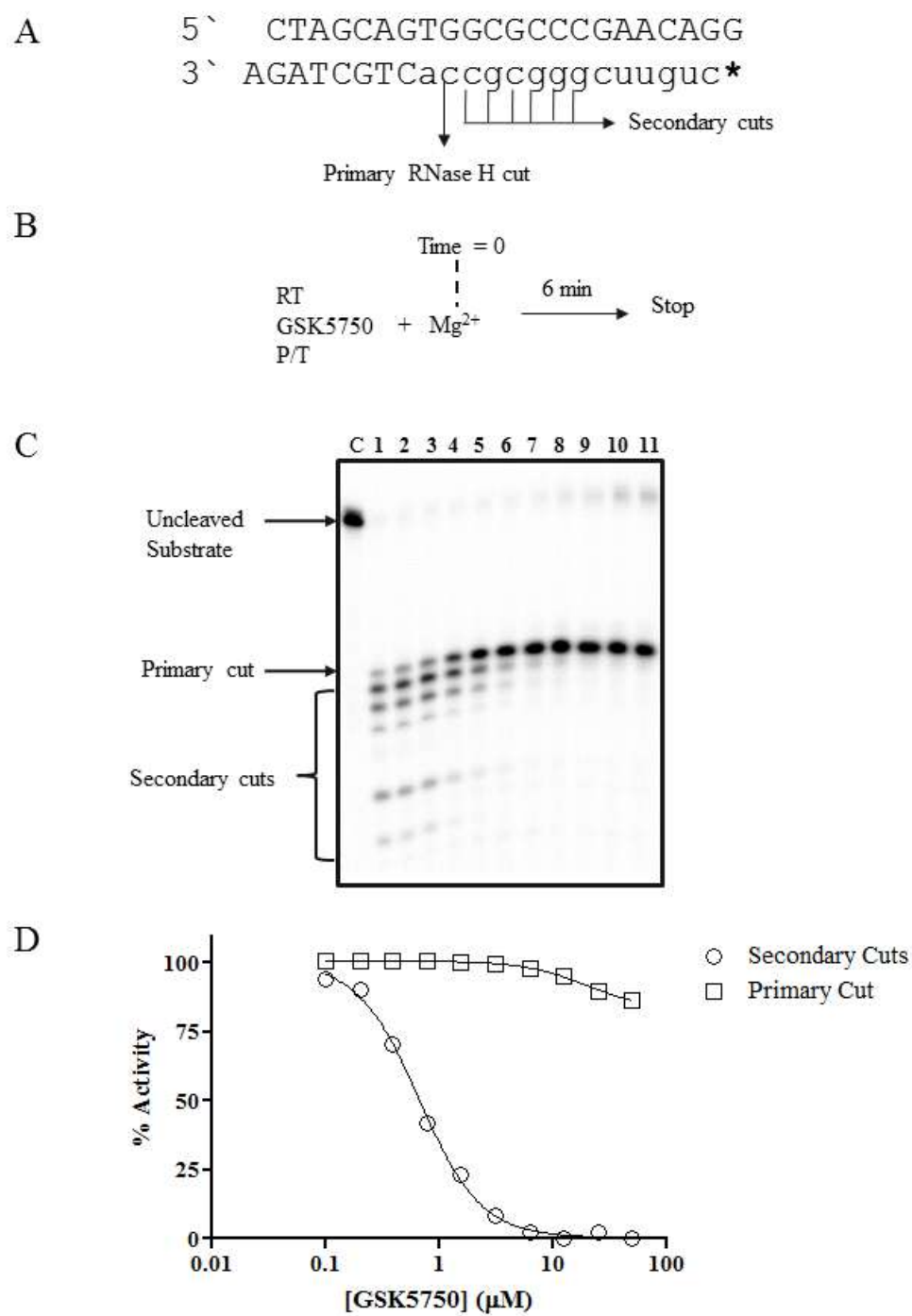


Figure 3.3.

Figure 3.3: Inhibition of RNase H activity by GSK5750 on a chimeric DNA-RNA/DNA substrate. (A) Substrate used in the assay mimics the (-)-strand primer removal reaction. The primary RNase H cleavage occurs at the DNA/RNA junction +1, as indicated. Secondary cleavages are also indicated downstream. (B) RNase H activity on the chimeric substrate shown in (A). Reaction conditions as shown in Experimental Procedures. Lane C is a control in the absence of MgCl_2 . Lanes 1-11 contain 0, 0.1, 0.2, 0.39, 0.78, 1.56, 3.12, 6.13, 12.5, 25 and 50 μM of GSK5750, respectively. (C) Results from (B) shown graphically. IC_{50} values for primary (\square) and secondary (\circ) cleavages are $>50 \mu\text{M}$ and $0.6 \mu\text{M}$, respectively.

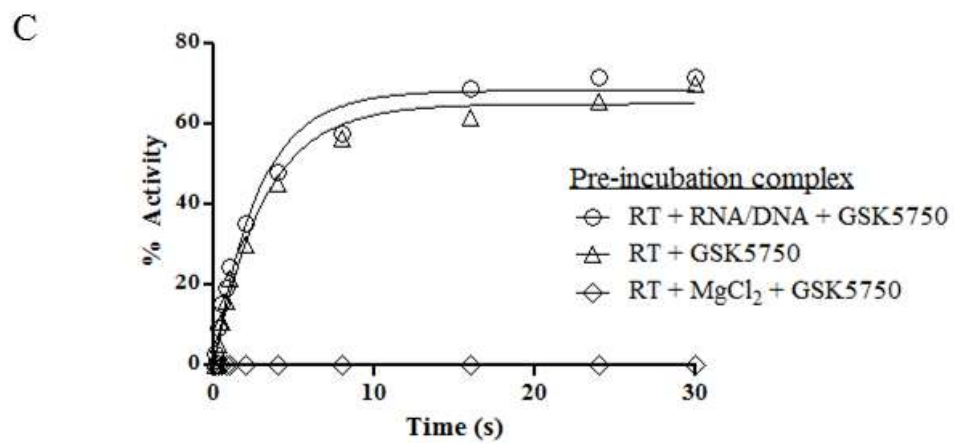
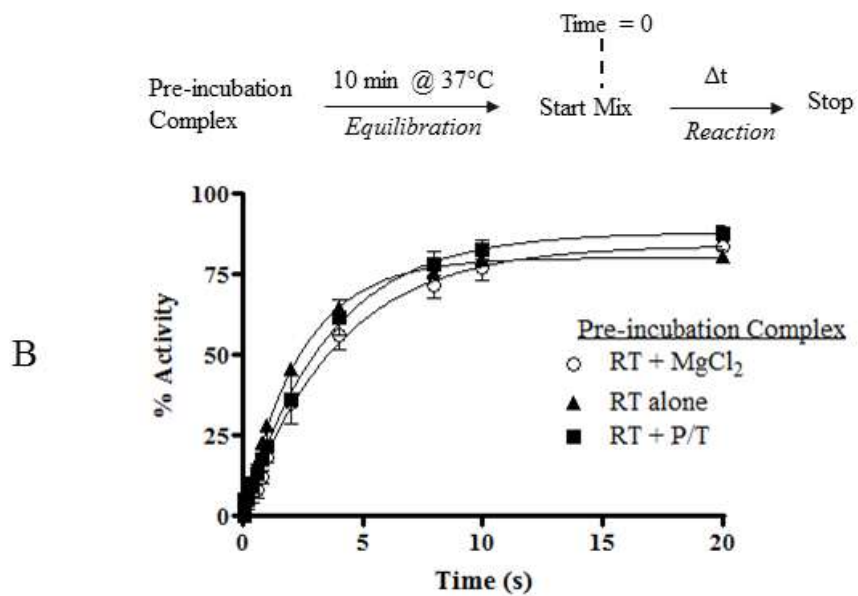
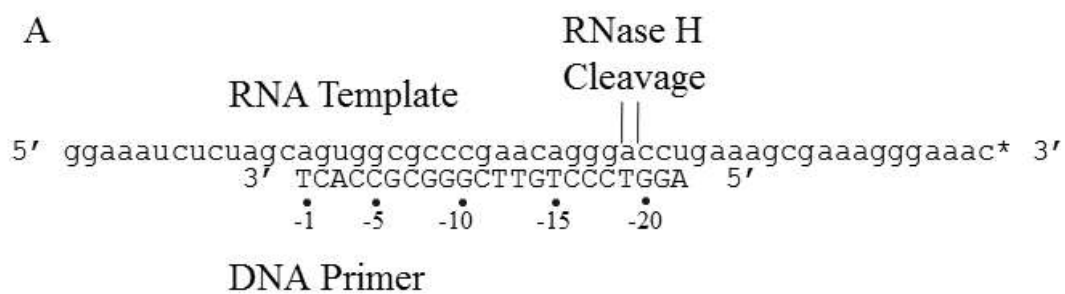


Figure 3.4.

Figure 3.4: (A) Schematic of polymerase-dependent substrate used in this assay. Primary RNase H cleavages are indicated. (B) Effects of the order-of-addition of reaction components in the absence of inhibitor. Percent activity refers to the percentage of RNA template converted into the -18 and -19 major products. From Beilhartz *et al.*, 2009 (277). (C) Effects of the order-of-addition of reaction components during inhibition by GSK5750 under pre-steady-state conditions. % activity refers to the percentage of RNA template converted into the -18 and -19 major products.

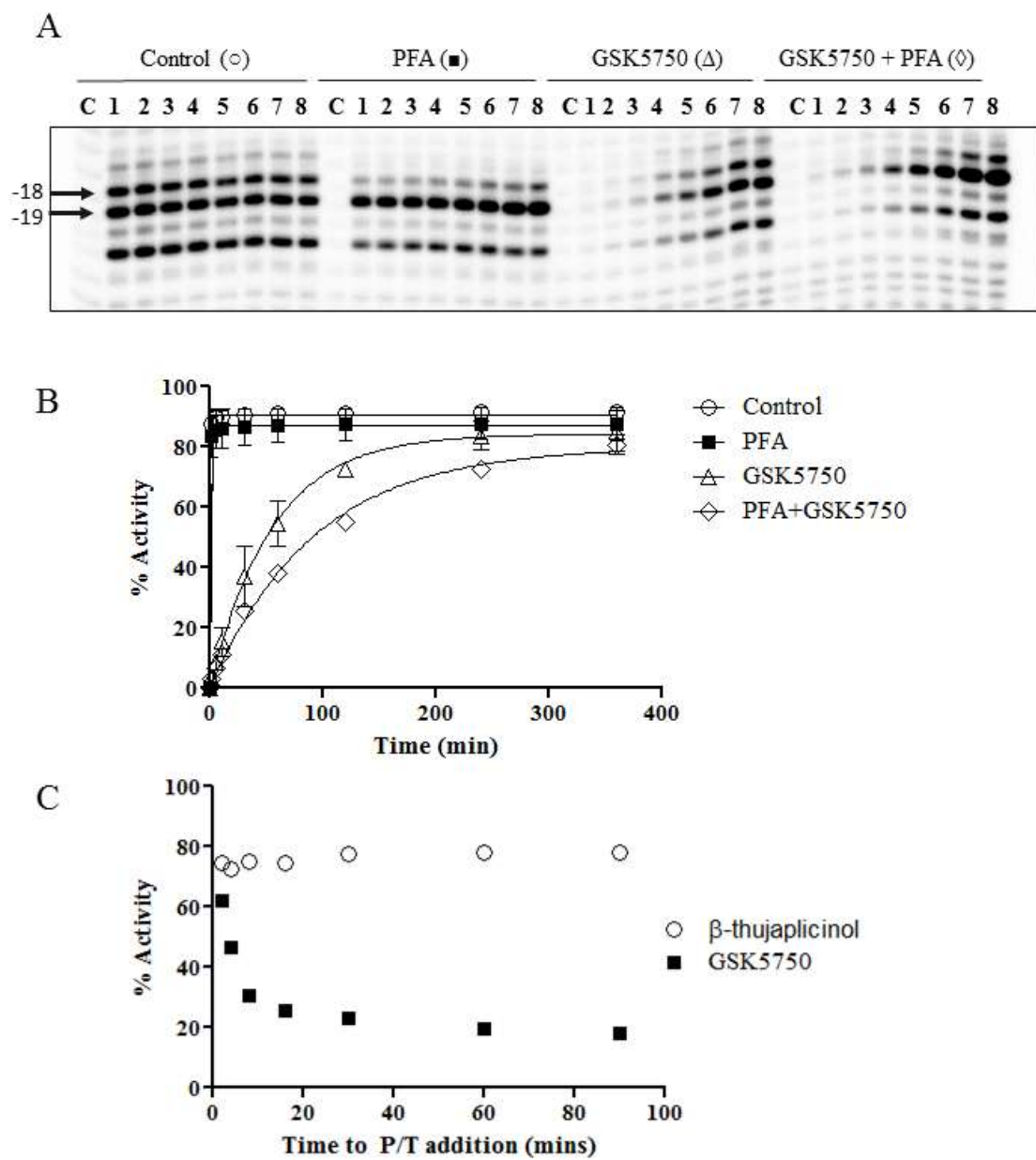


Figure 3.5.

Figure 3.5: Timecourse assay showing dissociation of the E-I-S complex. (A) In all cases, all reaction components were allowed to equilibrate at 37°C for 10 minutes and the reactions were started with primer/template. PFA and GSK5750 were added at 500 μ M and 500 nM, respectively, where indicated. Lane C is in the absence of $MgCl_2$. Lanes 1 – 8 represent time points of 1, 5, 10, 30, 60, 120, 240 and 360 minutes, respectively. (B) Results from (A) shown graphically. Timecourse in the absence of inhibitor (\circ), with PFA (\blacksquare), with GSK5750 (Δ) and with both PFA and GSK5750 (\diamond) were plotted and fit to a single exponential function. (C) Timecourse assay showing association of GSK5750 with free enzyme. X axis represents time from the creation of the E-I complex to the addition of substrate (S) for both 50 μ M β -thujaplicinol (\circ) and 0.5 μ M GSK5750 (\blacksquare).

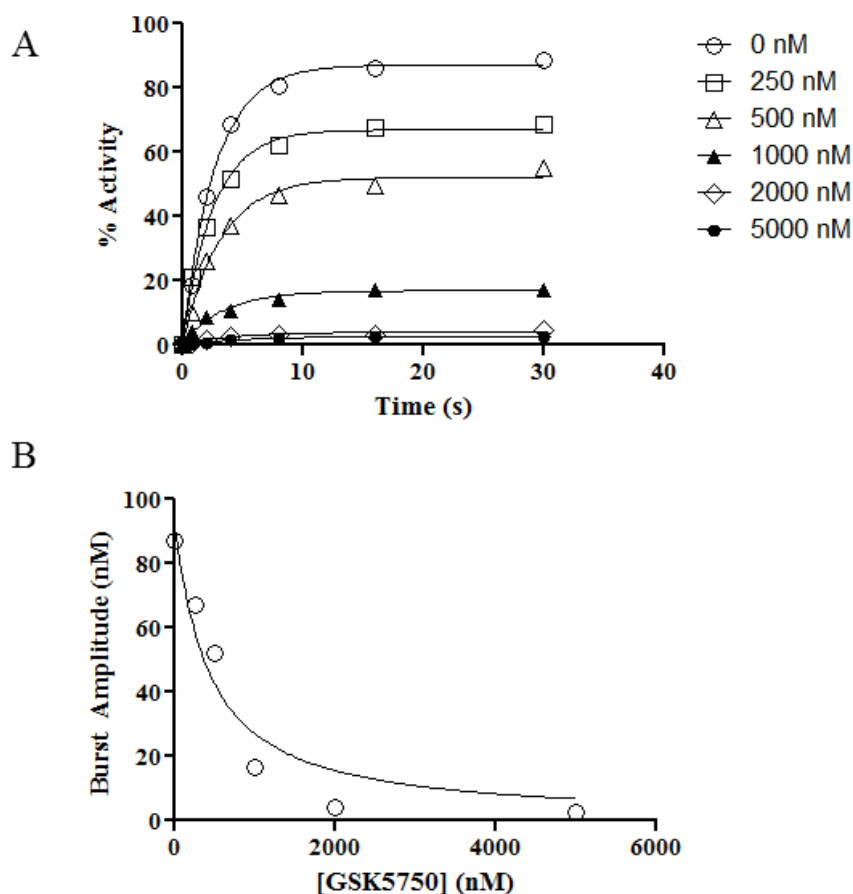


Figure 3.6: Pre-steady-state RNase H activity in the presence of GSK5750. (A) Various concentrations of GSK5750 were pre-incubated with MgCl_2 and RT. Reactions were started with DNA/RNA. Concentrations of GSK5750 are 0 nM (\bullet), 100 nM (\square), 250 nM (\blacktriangle), 500 nM (\blacktriangledown), 1250 nM (\blacklozenge) and 2500 nM (\circ). (B) Burst amplitudes from (A) (\circ) plotted against the concentration of GSK5750 and fit to a quadratic equation. A $k_D(\text{inhibitor})$ value of 386.5 ± 177.5 nM was calculated from the experiment done in triplicate.

Chapter 4

Probing the effects of the RNase H primer grip using a novel site-specific footprinting method on HIV-1 reverse transcriptase

Greg L. Beilhartz, Egor P. Tchesnokov, Matthias Götte

Dept. of Microbiology and Immunology, McGill University, Montreal, QC

4.1 Preface

Prior to my work on RNase H inhibitors, site-specific footprinting was used to determine the specific location of HIV-1 RT on a DNA/DNA hybrid to a 1 nt resolution. Both methods of footprinting (Fe^{2+} and KOONO-based) were through residues in contact with the DNA template at the RNase H and polymerase domains, respectively. We believed that if we could provide another KOONO-based footprint, then it could be used in tandem with the polymerase KOONO footprint to potentially measure changes in the trajectory of the nucleic acid. In theory this could be accomplished by retaining one footprint, and diminishing the second in the presence of some type of ligand. Since my work with β -thujaplicinol and GSK5750 suggested that the binding of these RNase H inhibitors should be accompanied by a change in the trajectory of the substrate due to steric clashes with the bound inhibitor at the RNase H active site, we wanted to develop this second footprint in order to confirm our hypothesis. If successful, it would be a fast, simple biochemical tool to measure the trajectory of nucleic acids bound to HIV-1 RT.

4.2 Abstract

Site-specific footprinting have been used to monitor the translocational equilibrium of HIV-1 RT, or to pinpoint the position of RT on its substrate to a 1 nucleotide resolution. Previous methods include KOONO or Fe^{2+} -based methods can make non-catalytic cleavages on the DNA template at the polymerase domain and RNase H active site, respectively. Here we developed another KOONO-based footprinting method based on the RNase H primer grip mutant T473C. In the presence of KOONO, this T473C cleaves the DNA primer at positions -15/-16 or -16/-17 based on whether the template is DNA or RNA. As the positions of the footprinting signals are not predicted by available crystal structures, we propose that flexibility inherent in the RT-substrate interactions in the RNase H primer grip motif can be measured by our novel footprinting technique that is not observable from rigid crystal structures. The T473C footprint can be used to measure flexibility or local perturbations in the structure of the nucleic acid substrate near the RNase H primer grip, as demonstrated by our experiments with methylphosphonate substitutions in the phosphate background of the DNA primer.

4.3 Introduction

The human immunodeficiency virus type 1 (HIV-1) is the etiological agent of acquired immune deficiency syndrome (AIDS). The HIV-1 reverse transcriptase (RT) is a heterodimeric (p66/p51) multifunctional enzyme that catalyzes the transformation of the single-stranded viral RNA genome into a double-stranded DNA provirus through two spatially and temporally distinct active sites (22). Both active sites reside on the larger,

catalytic p66 subunit. The N-terminal polymerase active site catalyzes nucleotide incorporation, while the C-terminal ribonuclease (RNase) H active site hydrolyzes the RNA strand of DNA/RNA intermediates that accumulate during reverse transcription (22). HIV-1 RT can accommodate dsDNA, dsRNA and DNA/RNA hybrids in its nucleic acid binding cleft, which stretches approximately 60 Å from the polymerase active site to the RNase H active site, which includes 18 bp of dsDNA and 19 bp of both dsRNA and DNA/RNA hybrids (57,184,188,210). During the course of nucleotide incorporation by HIV-1 RT, the polymerase active site accepts an incoming dNTP, which is then incorporated into the nascent DNA chain. At this point, the 3' primer terminus occupies the polymerase active site, an orientation referred to as pre-translocation. After addition of the dNTP, RT must translocate one bp downstream to vacate the polymerase active site to accommodate binding of the next templated dNTP (208). This orientation is referred to as post-translocation. RT exists in a thermodynamic equilibrium between pre- and post-translocational orientation. Once a dNTP binds, RT is trapped in the post-translocational position. This is referred to as the Brownian Ratchet model of polymerase translocation (208). We have previously developed site-specific footprinting technologies with HIV-1 RT that measure the translocational equilibrium of an enzyme-substrate complex using hydroxyl radicals generated from both metal and non-metal sources to a resolution of 1 nucleotide. Fe²⁺ footprinting involves the binding of two Fe²⁺ ions at the RNase H active site in a manner analogous to the catalytic Mg²⁺ ions present at the active site during canonical RNase H activity (181). Hydroxyl radicals are produced from the oxidation of the Fe²⁺ and cleave a DNA/DNA duplex on the template strand near the RNase H active

site with major cleavages at position -17 (post-translocation) and -18 (pre-translocation). KOONO footprinting produces hydroxyl radicals through its conjugate acid (ONOOH⁺) which leads to the formation of a thiyl radical on cysteine 280 (151). The thiyl radical abstracts a hydrogen from the sugar moiety of the proximal DNA template, leading to strand scission at positions -7 (post-translocation) or -8 (pre-translocation). Both of the footprinting techniques described above cleave the template DNA strand, and as such do not offer any direct information as to the behavior of the primer strand.

The RNase H primer grip is a structural motif near the RNase H domain in HIV-1 RT in contact with the primer strand (positions -10 to -15), comprising residues T473, A360, G359, H361, I505, Y501, K476, and Q475 in the p66 subunit, as well as K395 and E396 in the p51 subunit (57,303). The RNase H primer grip motif is responsible for aligning the bound DNA/RNA hybrid on the proper trajectory for RNase H cleavage at the active site. Mutations in the RNase H primer grip can cause deficiencies in RNase H activity as well as cleavage specificity, and in some cases decreases in viral titer (201,203). Since the RNase H primer grip is in close proximity to the primer strand, it is a prime candidate for site-specific footprinting. Here we present a novel site-specific footprint on the DNA primer strand at the mutant residue T473C using a metal-free radical source. This novel technology can be used to directly probe subtle interactions between the enzyme and the bound nucleic acid in the vicinity of the RNase H primer grip that was previously undetectable using other techniques.

4.4 Results

Identification of the T473C mutation – Site-specific footprinting employs Fe^{2+} or KOONO to generate hydroxyl radicals which cleave a DNA chain in close proximity to the source of radicals (Fig. 4.1B and C). Site-specific footprinting using KOONO as a source of hydroxyl radicals requires the thiol-containing side chain of a cysteine residue in close proximity to the DNA strand. As such, cysteine-scanning experiments were undertaken using the RNase H primer grip residues. The primer grip mutation T473C, is shown to be in close proximity to the DNA primer in both DNA/DNA and DNA/RNA crystal structures (Fig. 4.1A). The T473C mutation has a modest effect on substrate binding, while RNase H activity is approximately 25% of wildtype (data not shown). The effect on substrate binding is in agreement with previously published results (294).

Site-specific footprint mediated by T473C – Footprinting with T473C gave a specific cleavage on the DNA primer strand at positions -15 and -16, representing the post- and pre-translocational positions of RT, respectively (Fig. 4.2). Since the hyperreactive cleavage takes place on the DNA primer strand, we investigated the effects of the chemical nature of the template strand on the footprint. The footprint was visible with both a DNA and an RNA template, however with an RNA template the footprint is shifted upstream by one base pair, with the footprint at positions -16 and -17, representing the post- and pre-translocated positions respectively. Phosphonoformic acid (PFA) is a pyrophosphate analog that binds near the polymerase active site of RT and traps the enzyme-substrate complex in the pre-translocational orientation (228). Addition of PFA to the footprinting reaction provides a single cut at position -16 with a DNA template, and position -17 with an RNA template. Similarly, the post-translocational position can be

stabilized by first incorporating a ddNTP, then adding the next templated dNTP which will bind to RT in the post-translocated position but will not be incorporated due to the lack of a 3'-OH group on the primer terminus. Therefore the position of the footprint will represent the post-translocational position +1, and we observe a cleavage at position -14 with a DNA template and -15 with an RNA template, as expected. It is important to note, however, that the post-translocational position is actually -15 (DNA/DNA) or -16 (DNA/RNA), and only appears at -14/-15 because the primer has been extended by a single nucleotide (Fig. 4.2).

Methylphosphonate-modified primers affect DNA contacts with the RNase H primer grip

– DNA primers with the phosphate backbone selectively replaced by a methylphosphonate backbone have been shown to both modify the position of the RNase H active site on a single static complex, as well as the position of the enzyme on the nucleic acid substrate (303). Methylphosphate replacement has the effect of neutralizing the charge on the phosphate backbone, which is how the majority of RNase H primer grip residues make contact with the primer. It has been hypothesized that charge neutralization can affect a change in trajectory of the nucleic acid relative to RT (303). We employed three distinct triple methylphosphonate substitutions creating “neutral patches” that vary in position from upstream of the T473C mutation to downstream (Fig. 4.3A). This allows us to directly probe the effects of charge neutralization of the DNA backbone on the interaction with the RNase H primer grip motif (Fig. 4.3B). We employed three different footprinting methods, namely KOONO-based template footprinting (C280 cut), KOONO-based primer footprinting (T473C cut), and Fe^{2+} -based template footprinting (Fe^{2+} cut).

This approach allows a high-resolution readout of enzyme-substrate interactions throughout the nucleic acid binding channel. The nomenclature for neutral patches denotes the methylphosphonate link just upstream of the base number on the DNA primer. For example, the neutral patch labeled -14/-13/-12 denotes methylphosphonate links between nucleosides -12 and -13, -13 and -14, and -14 and -15. The C280 footprint of the control substrate (no methylphosphonate substitutions) shows a stabilization of the pre-translocated position in the presence of PFA (Fig. 4.3B). Neutral patches at -14/-13/-12, -15/-14/-13 and -17/-16/-15 all show a strong footprint in the pre-translocational position, although the complex forms at higher PFA concentrations for the -17/-16/-15 primer. In contrast, the T473C footprint showed a distinct footprint in the control primer, while the signal was lost as the neutral patch was moved through the RNase H primer grip toward the RNase H active site. The footprint of the -14/-13/-12 primer was weaker than the control primer, and the footprint of the -15/-14/-13 was weaker still. There was no footprinting signal with the -17/-16/-15 primer (Fig. 4.3B). The Fe^{2+} footprint showed a mostly consistent signal irrespective of the location of the neutral patch. However, when the neutral patch was positioned at -15/-14/-13 the lane without PFA shows multiple bands that gradually fade with increasing concentrations of PFA. The same effect is visible with the neutral patch over positions -17/-16/-15 to a slightly lesser extent. In all cases, at high PFA concentrations, there is a strong footprinting signal with Fe^{2+} .

4.5 Discussion

We have previously used site-specific footprinting to determine the precise positioning of HIV-1 RT on its nucleic acid substrate, especially in terms of the translocational equilibrium (209,228). Both KOONO-based and Fe^{2+} -based footprinting have been used in this regard, which probe the distance between RT and the DNA template near the polymerase and RNase H active sites, respectively. These methods have been satisfactory for determining the translocational equilibrium and enzyme position, however they lack the ability to measure the nature of the enzyme-substrate relationship in between the active sites. To this end, we developed a novel footprinting method. As with the KOONO-based C280 footprint, the new footprinting method would require a cysteine residue to be in close proximity to the nucleic acid. Residue threonine 473 in the RNase H primer grip motif, when mutated to a cysteine produced a single cleavage on the DNA primer at position -16 (pre-translocation) and -15 (post-translocation). Since the cleavage reaction took place on the DNA primer, we also tested whether the new footprint would accommodate an RNA template, something the earlier footprinting methods did not allow. On a DNA/RNA primer/template we observed the same footprinting signal but shifted by one nucleotide, at -17 and -16 (pre- and post-translocation, respectively) (Fig. 4.2B). The T473C footprint also provided multiple cleavages beyond the expected pre- and post-translocational cleavages, from -15 to -18 (Fig. 4.2B). These were visible even in the presence of ligands used to stabilize the complex, and prevent translocation of RT. Since C280 footprinting confirms that PFA and ddNTPs stabilize pre- and post-translocated complexes respectively, we require another explanation for multiple footprinting cleavages.

When we compare our footprinting data to the available crystal structures, we notice that in both DNA/DNA and DNA/RNA structures, T473 is in proximity to the sugar of primer base -15. This is in agreement with our T473C footprint on DNA/DNA, but less conclusive with DNA/RNA substrates. Since there are multiple cuts, it is difficult to determine where the primary cleavage is. We considered the structural differences in the substrates, namely, there is one extra base pair in the DNA/RNA structure in between the RT active sites than there is with DNA/DNA (57). This agrees with the observation that the primary cleavages appear to be -16/-17 with DNA/RNA, but does not explain the discrepancy with the crystal structure.

A third explanation for the appearance of multiple footprinting signals in a stabilized complex is flexibility in the protein-substrate interactions, something that would not be obvious in a rigid crystallographic structure. According to the Brownian Ratchet model of polymerase translocation, thermodynamic forces are sufficient to break the contacts between enzyme and substrate to allow RT to shift downstream by a single base pair. This suggests that a certain amount of flexibility may exist in the enzyme-substrate interactions, especially in the RNase H primer grip, which has generally fewer contacts than in the polymerase domain (189).

In order to measure internal flexibility, we then replaced the phosphate backbone with methylphosphate linkers (Fig. 4.3). This has the effect of neutralizing the charge on the backbone at the location of the substitution. We employed triple methylphosphonate substitutions to assess their effect on primer grip contacts. The C280 footprint showed the formation of a pre-translocated ternary complex upon addition of PFA to the control

(unmodified) primer. There was a slight loss of signal as the neutral patch created by the methylphosphonate substitutions was moved from positions -14/-13/-12, to -15/-14/-13, to -17/-16/-15. However, at high concentrations of PFA, the signal was still quite strong even at position -17/-16/-15. The signal also shows that PFA is bound to RT in the pre-translocational position, as expected. This demands that the substrate is in contact with RT at the polymerase active center. The novel T473C footprint shows a gradual loss of the footprinting signal as the neutral patch moves from -14/-13/-12 to -17/-16/-15 to the point where it completely disappears at -17/-16/-15. This is the point where the neutral patch is positioned directly where T473C makes contact with the primer. The loss of the footprint signal suggests that RT is not in contact with the substrate. This is likely due to the loss of the charge on the primer backbone. Interestingly, the Fe^{2+} footprint shows no loss of signal, and a strong signal at position -17/-16/-15. Although the Fe^{2+} footprint cleaves the template only 2 base pairs from the T473C footprint, it retains a strong signal, where the T473C signal is lost. This suggests that RT is in contact with the substrate at the RNase H active site. Also, visible on the Fe^{2+} footprint where the neutral patch is at position -15/-14/-13 in the absence of PFA are multiple bands beyond the canonical -17/-18 pre/post-translocational bands. This suggests that the enzyme is sliding on its substrate due to reduced contacts with the RNase H primer grip. This is in agreement with Dash *et al.* who describe a similar effect with Fe^{2+} and KOONO-based footprinting, as well as a change in RNase H cleavage specificity with methylphosphonate-modified primers.

Collectively, these data suggest that local flexibility in enzyme-substrate contacts in the vicinity of the RNase H primer grip best explains the T473C footprinting pattern. The

C280 and Fe²⁺ footprints show strong signals in the presence of methylphosphonate-modified primers when stabilized with PFA, but the T473C footprint shows a signal loss. This suggests that while RT can make contact with the substrate at both active sites simultaneously, it can lose contacts with the substrate in between the active sites. The T473C footprint is, to our knowledge, the first technique that can measure the trajectory of the nucleic acid substrate at both active sites, as well as in between active sites relative to HIV-1 RT. As a substrate trajectory change is believed to be the inhibitory mechanism behind several putative RNase H inhibitors (e.g. vinyllogous ureas), a method to probe the nucleic acid trajectory could be useful in determining novel mechanisms of inhibition and resistance to RNase H inhibitors.

4.6 Materials and Methods

Expression and Purification of HIV-1 RT variants – Heterodimeric p66/p51 HIV-1 RT was expressed and purified as previously described (344). The T473C RT mutation was created using the QuikChange site-directed mutagenesis kit according to the manufacturer's protocol. Sequencing of mutated plasmids was done by Genome Quebec.

Nucleic Acids – DNA primer PBS-30 (5'-TTCCCTTTCGCTTTCAGGTCCCTGTTCGGG-3') as well as DNA template PBS-50 (5'-AAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAAC-3') and RNA template PBS-52r (5'-

AAAUCUCUAGCAGUGGCGCCCGAACAGGGACCUGAAAGCGAAAGGGAAAC -
3') were purchased from IDT DNA Technologies. Control primer for figure 4.3 (5'-
TGGAGTCTTATTGCCATATCGAGTAGCTAG-3') was also purchased from IDT.
Methylphosphonate-modified primers were kindly provided by Dr. Stuart LeGrice, NCI,
Frederick.

KOONO site-specific footprinting – 50 nM of 5'-radiolabeled DNA was heat-annealed to
a 3-fold molar excess of either DNA or RNA template. The hybrid was then added to 750
nM HIV-1 RT in a buffer containing 120 mM cacodylate buffer (pH 7), 0.3 mM DTT, 10
mM MgCl₂, and 20 mM NaCl. PFA was added when described in the manuscript.
Reactions were incubated for 10 minutes at 37°C prior to addition of KOONO or Fe²⁺.
Treatment with KOONO and/or Fe²⁺ was performed essentially as described (181,209).

4.7 Figures and Figure Legends

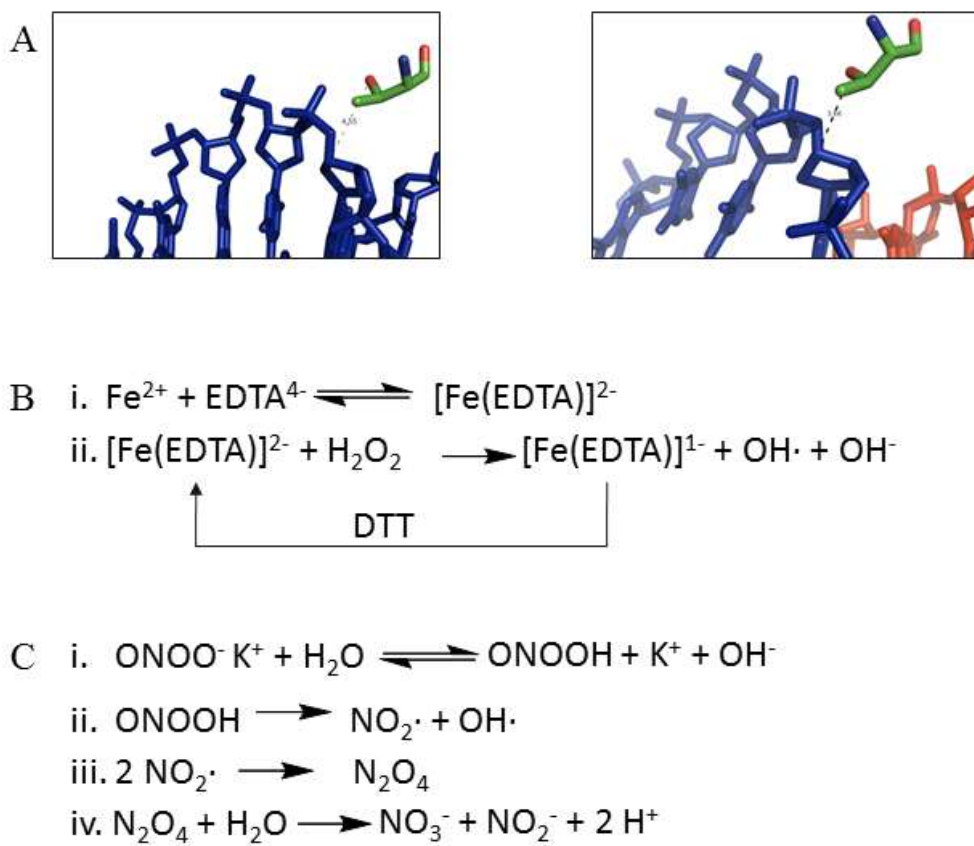


Figure 4.1

Figure 4.1: (A) (Left Panel) Position of T473 in the HIV-1 RT structure relative to the DNA primer on a DNA/DNA substrate (PDB code: 1rtd). Distance shown is from the sidechain of T473 to the sugar of nucleotide -15. (Right Panel) Same as left panel except with RT crystallized with a DNA/RNA primer/template. Distance shown is again from the sidechain of T473 to the sugar of primer nucleotide -15. (B) Reaction scheme for creating OH^\cdot radicals from Fe^{2+} . The modified fenton reaction using DTT to reduce $[\text{Fe}(\text{EDTA})]^{1-}$ to regenerate $[\text{Fe}(\text{EDTA})]^{2-}$ and maintain production of OH^\cdot radicals. (C) Potassium peroxyxynitrite reaction. OH^\cdot radicals in this case were generated by the degradation of the conjugate acid at neutral pH. Used in both C280 and T473C footprinting.

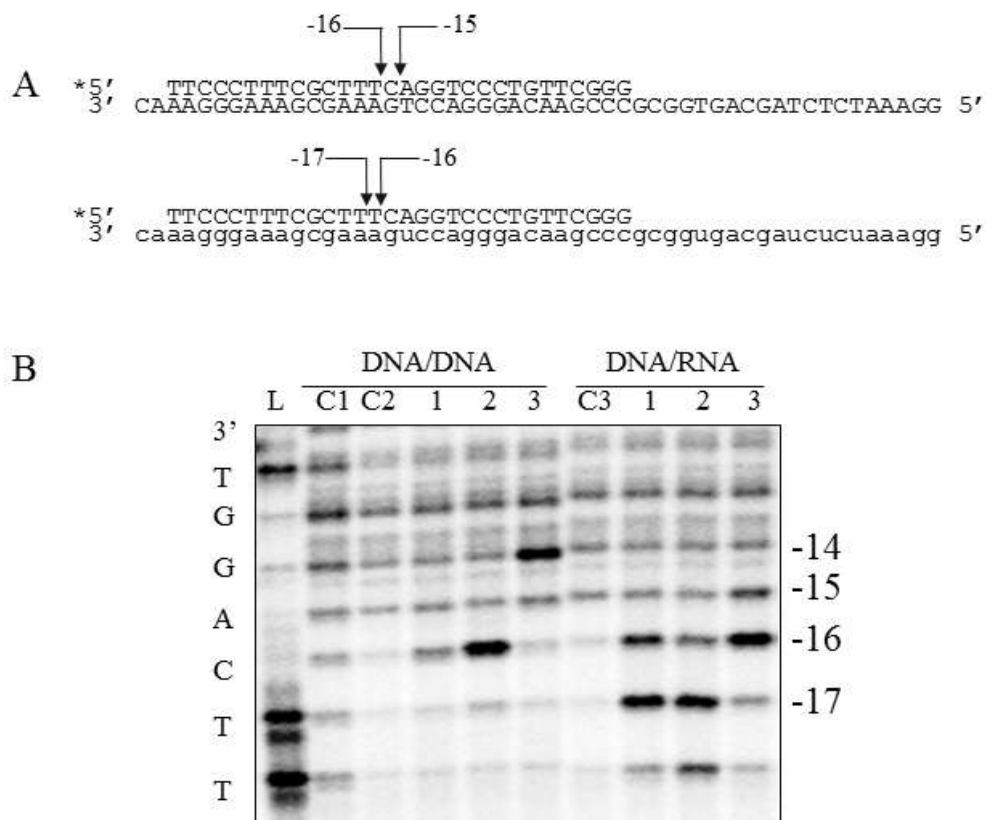


Figure 4.2: (A) Sequences of the primer (PBS-30) and templates (PBS-DNA and –RNA) used in (B). Arrows indicate the position of the KOONO cleavages. (B) (Left-panel) Lane 1 - RT carrying the T473C mutation is treated with KOONO in the presence of the DNA/DNA substrate in (A). Faint cleavages are visible at positions -15 and -16 in the presence of KOONO. Lane 2 - In the presence of 200 μ M PFA, the stabilized ternary complex shows a strong cleavage at position -16, indicative of a pre-translocated complex. Lane 3 - In the presence of 50 μ M of the chain-terminator ddCTP and 500 μ M of the next templated nucleotide (dGTP), a dead-end complex is formed in the post-translocated position. A cleavage representing this complex is seen at position -14

(instead of -15, because the primer is extended by one nucleotide following the incorporation of ddCTP). (Right panel) Same as the left panel, except the template is RNA, instead of DNA and RT carries the E478Q mutation, as well as T473C to prevent degradation of the RNA template. All of the cleavages shown in the left panel are shifted by one base pair upstream, e.g. the pre-translocated cleavage is at -17 instead of -16.

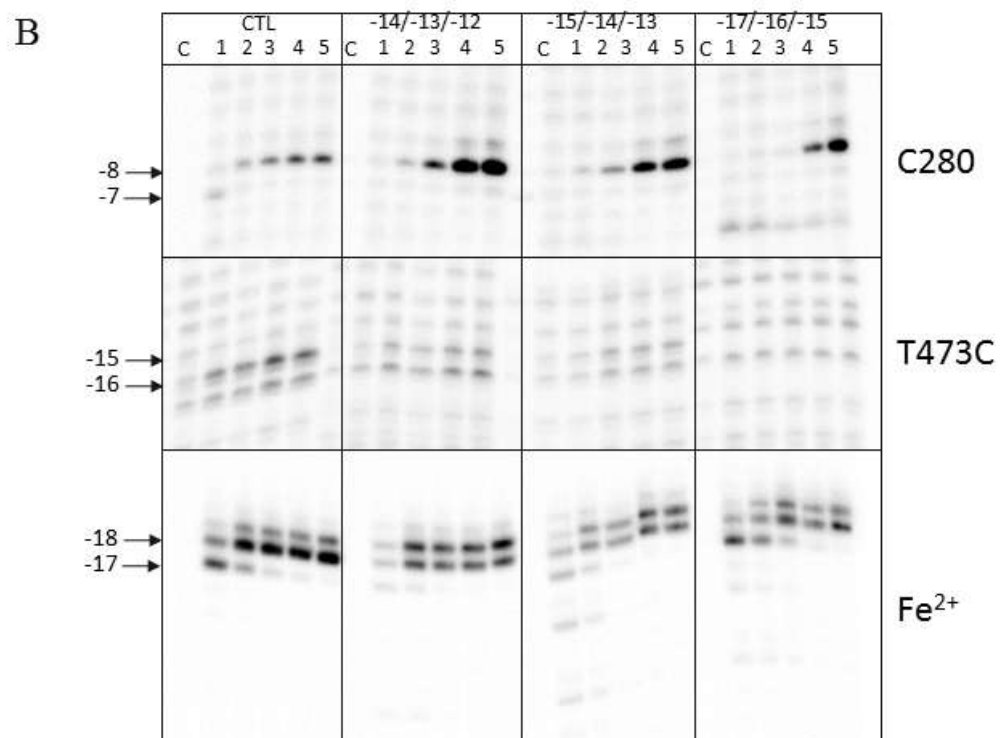
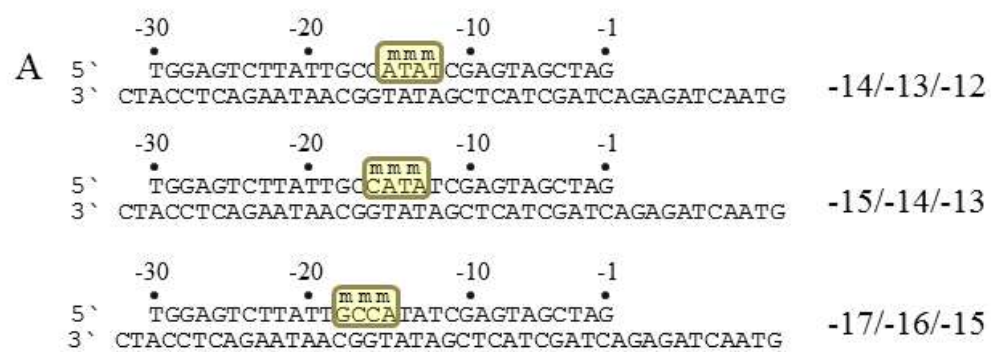


Figure 4.3.

Figure 4.3: (A) Sequences used in this experiment. Neutral patches are surrounded by the brown box, and single methylphosphonate substitutions are marked by an “m”. (B) Site-specific footprints of methylphosphonate-modified primers. Lane C represents samples without addition of either KOONO or Fe²⁺. Lanes 1-5 represent PFA concentrations of 0, 1, 10, 50, and 500 μ M. C280 and Fe²⁺ footprints are 5'-radiolabeled on the template strand, while T473C footprints are 5'-radiolabeled on the primer strand. Pre- and post-translocated cleavages are indicated with arrows, as well as their position from the 3' primer terminus.

Chapter 5

Discussion

At the onset of my PhD work, several different chemicals had been discovered that inhibit the RNase H activity of HIV-1 RT. Some had multiple binding sites on the RT enzyme, or were suspected to. Others involved the p51 thumb subdomain in inhibitor binding. The vast majority of the anti-RNase H compounds belonged to a class called active site RNase H inhibitors, so named due to their binding at the RNase H active site. Specifically, active site inhibitors shared a 3-oxygen pharmacophore that were correctly spaced so as to chelate the two divalent metal ions that are coordinated by the conserved DEDD motif in the RNase H active site. Research had shown that chelation of these critical metal ions prohibited hydrolysis of the RNA template. However, despite the success of active site inhibitors *in vitro*, only one had ever been shown to be active in cell-based assays, and none had proceeded to clinical trials. The reason for this discrepancy was unknown, and represented an important obstacle to the development of clinically-relevant drugs against HIV-1 RNase H.

To my knowledge, we were the first to characterize active site RNase H inhibitors under pre-steady-state conditions, something that had already been done for polymerase inhibitors. Pre-steady-state conditions differ from steady-state in several ways, but most importantly they differ in relative enzyme-substrate concentrations. Pre-steady-state analyses are concerned with a single turnover of the enzyme, as opposed to the repeated activity, dissociation, re-association and activity on the enzyme on multiple substrates (steady-state). Using the active site inhibitor β -thujaplicinol as a test compound, we found that we could not observe inhibition by β -thujaplicinol under a single turnover, unless we pre-formed the E-I complex and started the reaction with substrate, and even under these

circumstances, inhibition was minor and β -thujaplicinol dissociated in the time required for one enzyme turnover event. Therefore, the conditions of the experiment were critical to determining the potency of β -thujaplicinol. *In vivo*, HIV RT is in excess of substrate, although estimates vary of the amount. Thus, the pre-steady-state conditions of our experiment more closely resemble the situation *in vivo* than steady-state experiments that are traditionally used to characterize these compounds.

We then performed experiments using a hybrid substrate, with a one DNA strand, and the other strand a DNA-RNA hybrid. This represents the (-)-strand primer removal reaction during reverse transcription. In this reaction, there are two types of RNase H cleavages: the primary cleavage, which takes place first, and at the DNA-RNA junction +1, and secondary cleavages, which are kinetically slower and take place further downstream on the RNA portion of the template. We observed that β -thujaplicinol inhibits these cleavages at different rates. The primary cleavage was not inhibited at all, while the secondary cleavages were strongly inhibited. These data, combined with the pre-steady-state data described above, resulted in our hypothesis that β -thujaplicinol competed for space in the RNase H active site with the nucleic acid substrate. Firstly, β -thujaplicinol only inhibited RNase H activity during a single enzyme turnover if it was allowed to first form an E-I complex, then the substrate is added later. Even in this situation, the inhibition is transient. Secondly, β -thujaplicinol was unable to inhibit the primary RNase H cleavage, when the RNase H active site was occupied by the substrate, but easily inhibited secondary cleavages after the primary cleavage had been made, and the inhibitor had gained access to the active site. We hypothesized that the nucleic acid substrate was

outcompeting the inhibitor in the active site. The only time inhibition is observed with β -thujaplicinol was during multiple turnover events, where the transient inhibition we observed during a single turnover would be repeated many, many times, amplifying the effect and giving the illusion of potent inhibition. If this hypothesis could be applied to other active site inhibitors, then we had successfully identified a reason why active site RNase H inhibitors were inactive in cell-based assays.

Another example of the success of pre-steady-state analysis in determining true RNase H inhibitors came from our experiments with the PPi analog foscarnet. Foscarnet had been described as a potent RNase H inhibitor in steady-state assays, despite being a well-known and clinically-relevant polymerase inhibitor. We demonstrated that foscarnet did not directly inhibit the RNase H activity of RT, but instead stabilized an E-I-S complex, which decreased enzyme turnover, resulting in less polymerase and RNase H activity, which is in agreement with previously published work showing that all foscarnet resistance mutations appear near the polymerase active site.

During the course of this work, a study was published that contained a crystal structure of human RNase H1 complexed with an RNA substrate in contact with the active site. The authors aligned this structure with the structure of RT complexed with a PPT-based DNA/RNA substrate, where the substrate was in contact with the polymerase active site, and not the RNase H active site. In this model, it was impossible to “connect” the RNA strands, leading the authors to conclude that the substrate must toggle between the polymerase and RNase H active sites of HIV-1 RT, and cannot occupy both sites simultaneously. This would have implications for our hypothesis of active site inhibitor

binding, since the toggling of the substrate between active sites would leave space for active site RNase H inhibitors to bind while the RT enzyme was in “polymerase mode”. Our data was not in agreement with this conclusion.

It has been shown that foscarnet binds only to pre-translocated E-S complexes and stabilizes them. We can also show that by adding a chain-terminating nucleotide in concert with the next templated nucleotide, we can trap the post-translocated complex. We can monitor the translocational status of RT by looking at RNase H activity on polyacrylamide gels. In the absence of ligand, two primary cleavages are observed, representing the pre- and post-translocated complexes of RT. We have shown that under a single enzyme turnover, the addition of foscarnet will result in a single RNase H cleavage corresponding to the pre-translocated complex, while the addition of the next templated nucleotide will also result in a single RNase H cleavage corresponding to the post-translocated complex. In both situations, the RNA template is completely cleaved, and both foscarnet and nucleotides bind at the polymerase active site, while the experimental readout is RNase H activity. In summary, because these ligands bind at the polymerase active site, produce a single RNase H cut, and do not inhibit RNase H activity, the substrate must engage both active sites simultaneously. However, this does not require that the substrate must always engage both active sites at the same time, nor does it require that polymerase and RNase H activities are temporally coordinated. For example, it has been shown that polymerase activity is faster than RNase H activity. Also, we have measured equilibrium binding constants (K_d) of the substrate for both polymerase and RNase H active sites. We observed that K_d values at the RNase H active site are

approximately 5-fold higher than at the polymerase active site. This is consistent with the observation that most RT-substrate contacts are located in the polymerase domain.

Collectively with the data discussed above, we proposed a model of RNase H activity based on the “breathing” of the substrate in the vicinity of the RNase H active site, while the substrate remains bound at the polymerase active site. This is consistent with the observation that RT cleaves its substrate predominantly during RT pausing events, but does not support the idea that substrate binding at the active sites of HIV RT are mutually exclusive. Therefore, although RT can engage its substrate at both active sites simultaneously, it is not strictly required to do so.

In 2009, Himmel *et al.* published the crystal structure of β -thujaplicinol bound at the RNase H active site through a metal ion-mediated mechanism as predicted. Himmel *et al.* reported a K_d of the inhibitor of 0.73 μM using intrinsic protein fluorescence and 6 μM using surface plasmon resonance. They reported a K_i of 0.14 μM determined kinetically, using a high-turnover system. They also report non-competitive kinetics, in agreement with previously published results. Himmel *et al.* show a slight decrease in K_d (inhibitor) in the presence of an RNA-RNA substrate, thus showing the presence of an E-S-I complex, and therefore non-competitive inhibition by strict definition. Our model of local “breathing” by the substrate at the RNase H active site does not contradict a non-competitive model of inhibition. However, I believe it should be noted that although β -thujaplicinol binding does not interfere with the global binding of the substrate, i.e. enzyme-substrate contacts throughout the nucleic acid binding channel, it does appear to compete directly with the scissile phosphate for access to the RNase H active site.

Mutually exclusive binding (to the RNase H active site specifically) of inhibitor and substrate suggests competitive inhibition. Thus, although β -thujaplicinol displays classical non-competitive kinetics, it should not be compared to an NNRTI-style of non-competitive inhibition. As this may be the case with many active site inhibitors of RNase H activity, an effort should be made to address this potentially confusing contradiction, i.e. a competitive inhibitor displaying non-competitive kinetics, and both being true.

It should be noted that we published a structural model including an artificial amino acid at position 501, benzoyl-L-phenylalanine, which in biochemical assays conferred resistance to β -thujaplicinol. We suggested that the sidechain of the artificial residue rotated away from β -thujaplicinol to form other interactions, depriving β -thujaplicinol of necessary π -interactions. The crystal structure published by Himmel *et al.* does not agree with this model however, by showing that β -thujaplicinol is not available for interactions with Y501, and that the resistance we observed was likely due to steric clashes between β -thujaplicinol and the artificial amino acid, which were not a problem with the natural substrate (347). This, however, is separate from our “breathing” model and does not affect other conclusions in chapter 2.

There are at least two methods to avoid the obstacle of substrate competition in developing novel RNase H inhibitors. First, develop inhibitors that bind at another site that does not overlap with the substrate such as the vinylogous ureas, or secondly, develop inhibitors that bind tightly enough to the active site to out-compete the substrate. The active site inhibitor GSK5750 is a good example of the latter. Like β -thujaplicinol,

GSK5750 is ineffective at inhibiting primary RNase H cuts when the inhibitor is introduced to a pre-formed E-S complex, but is highly potent against secondary cuts. However, under single turnover conditions, GSK5750 could inhibit the primary RNase H cleavage if an E-I complex was pre-formed. Recall that under these conditions, only minor and transient inhibition was observed with β -thujaplicinol. GSK5750 bound tightly enough that we were able to pre-form an E-I complex, and then add substrate to monitor the dissociation of GSK5750 over time. Whereas β -thujaplicinol dissociated from the complex within a single turnover event, GSK5750 dissociated from the enzyme in a period of approximately 6 hours. When foscarnet was added along with the substrate, only pre-translocated complexes were observed (inferred by a single RNase H cleavage instead of the usual two). This suggests that the substrate can bind simultaneously with the bound inhibitor, forming an E-I-S ternary complex. The observation that GSK5750 does not affect the polymerase activity of RT further supports this conclusion. Evidence that β -thujaplicinol can also form an E-I-S complex is provided by Himmel *et al.*. We also monitored the time between inhibitor binding and substrate addition and found that GSK5750 is a slow-binding inhibitor, taking about 8 minutes for maximum inhibition to occur. From this experiment, we cannot tell whether GSK5750 has one slow binding event, or an initial fast-binding event, followed by a slow conformational change. We can speculate that it is the latter, since we know that GSK5750 cannot bind to an E-S complex, and therefore a slow-binding inhibitor would have difficulty competing with a fast-binding substrate. However, if it is a two-stage binding, the nature of the fast-binding complex must allow for RNase H cleavage to occur. This is difficult to reconcile with the

observation that GSK5750 binding is Mg^{2+} -dependent, and presumably chelates metal ions in a similar mechanism to β -thujaplicinol. The nature of GSK5750 binding to RT is therefore not fully understood, and more research is required to accurately describe the characteristics of the presumed fast-binding complex.

The final manuscript presented in this thesis is a slight departure from the first two, in that it does not directly address the problem of efficient inhibition of the RNase H activity of HIV-1 RT. Site-specific footprinting is a technique used to quantitatively measure the percentage of RT bound to the substrate in a specific position, to a 1 nucleotide resolution. It is frequently used to determine the translocation equilibrium of various substrates, as well as the effects of various ligands on RT translocation. As described in chapter 4, there are two distinct types of site-specific footprinting. A metal-dependent method, which employs Fe^{2+} ions in the RNase H active site as a source of OH^- radicals used to cleave the DNA template at position -17, and a metal-independent method which employs KOONO which activates the thiol sidechain of cysteine 280 used to cleave the DNA template at position -7. It was our goal to use site-specific footprinting not to discern where RT was located on the substrate, but to measure a change in trajectory of the template due to the presence of a bound inhibitor at the RNase H active site which relates to our earlier work from chapters 2 and 3. Previous footprinting experiments with β -thujaplicinol had shown no effect with KOONO footprinting, and a concentration-dependent decrease in the footprinting signal with Fe^{2+} . However, these data were inconclusive, as Fe^{2+} ions bind to the RNase H active site in a similar fashion to Mg^{2+} , and β -thujaplicinol could have merely been chelating Fe^{2+} in the active site, instead of

measuring a trajectory change in the template. This led to the development of a second KOONO footprinting site, based on site-directed mutagenesis of residues that were in close proximity to the substrate. Cysteine scanning experiments determined that the RNase H primer grip mutant T473C provided a KOONO-dependent cleavage at position -15 on the DNA primer. A benefit of this T473C footprint is that we can cleave the DNA primer paired to both DNA and RNA templates. In fact, with an RNA template, the T473C footprint occurs at position -16, as opposed to -15 for a DNA template.

Footprinting experiments have already shown that 17 bps of DNA/DNA can fit in between the active sites of RT, while 18 bps of DNA/RNA can fit in the same distance. This is due to the more compact shape of DNA/RNA hybrids. Further, while foscarnet and ddNTPs have been used to trap the pre- and post-translocated complexes on DNA/DNA substrates, multiple footprinting signals (KOONO-dependent cleavages) are observed on DNA/RNA substrates, even when the complex is stabilized by ligands. This suggests a certain degree of inherent flexibility in the RT-DNA/RNA complex.

In order to monitor a change in trajectory of the substrate in the presence of an active site RNase H inhibitor, we expected to see a concentration-dependent decrease in the T473C footprint, while observing no change in the C280 footprint. Experiments with both β -thujaplicinol and GSK5750 have shown no effect on the T473C footprint, or the C280 footprint. This suggests that either a bound RNase H inhibitor does not alter the trajectory of the nucleic acid substrate, or that the T473C footprint is not sensitive enough to measure it. Based on the conclusions stated in the previous chapters, it seems the latter is more likely. We then used modified primers with methylphosphonate substitutions, which

have the effect of negating the charge on the primer backbone, through which the majority of primer grip residues contact the primer. Using methylphosphonate substitutions to create “neutral patches”, we used 3 different primers that placed the neutral patch in different areas of the primer grip. As these neutral patches moved toward the RNase H active site, the T473C footprint decreased in intensity, as well as changed position. However, strong footprinting signals were still observed at the RNase H active site (Fe^{2+} footprint) and near the polymerase active site (C280 footprint). We concluded that our T473C footprint can measure a local perturbation in the primer grip – substrate interaction.

With the above data, we can hypothesize that there is significant flexibility between the RNase H primer grip and the nucleic acid substrate. As mentioned above, RT can accommodate an “extra” base of DNA/RNA in between its active sites than DNA/DNA. We questioned “where” that extra base occurs. It is made up gradually throughout the entire enzyme length, or does it occur rapidly in a specific place? Structural alignments have shown differing results based on whether the template or primer is observed. This is not trivial, since all biochemical studies showing the number of bases between active sites have focussed on template bases, and it was assumed that primer bases would match based on Watson-Crick complementarity. However, only one crystal structure has been published with RT complexed with a DNA/RNA substrate (57). This substrate is the HIV-1 PPT sequence, which is resistant to RNase H cleavage. In this structure, an unusual departure from Watson-Crick base pairing occurs at the RNase H primer grip. Template base 15 is unpaired, followed by two mispaired bases, followed by a unpaired

primer base (bp 17) bringing it back into register immediately before the RNase H active site. So in this substrate, the “extra” base is made up in different places on the primer and on the template. This “unzipping” of the PPT, as it is referred to by the authors, requires a certain amount of flexibility. Although the substrate used in our experiments were not PPT-based, they do show a certain amount of flexibility in RT-substrate contacts. Perhaps the “unzipping” observed in the PPT crystal structure is not strictly a function of PPT DNA/RNA hybrids, but a property of all DNA/RNA hybrids bound to RT, and the RNase H primer grip is responsible for this deformation of the nucleic acid. The RT-DNA/RNA structure predicts that T473 would contact the sugar of primer base 15, identical to the RT-DNA/DNA structure (57,188). This does not agree with our footprinting data. However, crystal structures by their nature show static complexes, and flexibility is difficult to deduce from these complexes. Our footprinting data could represent an alternate configuration, which for a number of reasons was not the configuration observed in the crystal structure.

Collectively, T473C footprinting provides a relatively simple biochemical method of probing enzyme-nucleic acid contacts at the RNase H primer grip. We have shown that despite the proximity to the RNase H active site, T473C footprinting responds differently than C280 footprinting and Fe^{2+} footprinting. As a technique, it has shown insight into the flexible motion of the RT-substrate interaction that is unattainable by crystallographic methods.

References

1. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dautuet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983) *Science* **220**, 868-871
2. Coffin, S. R., Hollis, T., and Perrino, F. W. (2011) *J Biol Chem* **286**, 16984-16991
3. Pulliero, A., Fazzi, E., Cartiglia, C., Orcesi, S., Balottin, U., Uggetti, C., La Piana, R., Olivieri, I., Galli, J., and Izzotti, A. (2011) *Mutation research* **717**, 99-108
4. Reeves, J. D., and Doms, R. W. (2002) *J Gen Virol* **83**, 1253-1265
5. Douglas, N. W., Munro, G. H., and Daniels, R. S. (1997) *J Mol Biol* **273**, 122-149
6. Popovic, M., Read-Connole, E., and Gallo, R. C. (1984) *Lancet* **2**, 1472-1473
7. Ono, A. (2010) *Biol Cell* **102**, 335-350
8. Chertova, E., Bess, J. W., Jr., Crise, B. J., Sowder, I. R., Schaden, T. M., Hilburn, J. M., Hoxie, J. A., Benveniste, R. E., Lifson, J. D., Henderson, L. E., and Arthur, L. O. (2002) *J Virol* **76**, 5315-5325
9. Ott, D. E. (2008) *Rev Med Virol* **18**, 159-175
10. Zhu, P., Chertova, E., Bess, J., Jr., Lifson, J. D., Arthur, L. O., Liu, J., Taylor, K. A., and Roux, K. H. (2003) *Proc Natl Acad Sci U S A* **100**, 15812-15817
11. Adamson, C. S., and Freed, E. O. (2007) *Adv Pharmacol* **55**, 347-387
12. Dimitrov, D. S., Xiao, X., Chabot, D. J., and Broder, C. C. (1998) *J Membr Biol* **166**, 75-90
13. Gorry, P. R., and Ancuta, P. (2011) *Curr HIV/AIDS Rep* **8**, 45-53
14. Dalgleish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) *Nature* **312**, 763-767
15. Maddon, P. J., Dalgleish, A. G., McDougal, J. S., Clapham, P. R., Weiss, R. A., and Axel, R. (1986) *Cell* **47**, 333-348
16. Doms, R. W., and Peiper, S. C. (1997) *Virology* **235**, 179-190
17. Moore, J. P., Trkola, A., and Dragic, T. (1997) *Curr Opin Immunol* **9**, 551-562
18. Berger, E. A., Doms, R. W., Fenyo, E. M., Korber, B. T., Littman, D. R., Moore, J. P., Sattentau, Q. J., Schuitemaker, H., Sodroski, J., and Weiss, R. A. (1998) *Nature* **391**, 240
19. Caffrey, M. (2011) *Trends Microbiol* **19**, 191-197
20. Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G. B. (2009) *Cell* **137**, 433-444
21. Basu, V. P., Song, M., Gao, L., Rigby, S. T., Hanson, M. N., and Bambara, R. A. (2008) *Virus Res* **134**, 19-38
22. Telesnitsky, A., and Goff, S. (eds). (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)
23. Ellison, V., Abrams, H., Roe, T., Lifson, J., and Brown, P. (1990) *J Virol* **64**, 2711-2715
24. Farnet, C. M., and Haseltine, W. A. (1990) *Proc Natl Acad Sci U S A* **87**, 4164-4168
25. Arhel, N. (2010) *Retrovirology* **7**, 96

26. Bukrinsky, M. (2004) *Mol Med* **10**, 1-5
27. Dvorin, J. D., and Malim, M. H. (2003) *Curr Top Microbiol Immunol* **281**, 179-208
28. Lehmann-Che, J., and Saib, A. (2004) *AIDS Rev* **6**, 199-207
29. Suzuki, Y., and Craigie, R. (2007) *Nat Rev Microbiol* **5**, 187-196
30. Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G., and Stevenson, M. (1993) *Proc Natl Acad Sci U S A* **90**, 6125-6129
31. Fassati, A., and Goff, S. P. (2001) *J Virol* **75**, 3626-3635
32. Heinzinger, N. K., Bukinsky, M. I., Haggerty, S. A., Ragland, A. M., Kewalramani, V., Lee, M. A., Gendelman, H. E., Ratner, L., Stevenson, M., and Emerman, M. (1994) *Proc Natl Acad Sci U S A* **91**, 7311-7315
33. Iordanskiy, S., Berro, R., Altieri, M., Kashanchi, F., and Bukrinsky, M. (2006) *Retrovirology* **3**, 4
34. Karageorgos, L., Li, P., and Burrell, C. (1993) *AIDS Res Hum Retroviruses* **9**, 817-823
35. McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M., and Hope, T. J. (2002) *J Cell Biol* **159**, 441-452
36. Nermut, M. V., and Fassati, A. (2003) *J Virol* **77**, 8196-8206
37. Warrilow, D., Tachedjian, G., and Harrich, D. (2009) *Rev Med Virol* **19**, 324-337
38. Forshey, B. M., von Schwedler, U., Sundquist, W. I., and Aiken, C. (2002) *J Virol* **76**, 5667-5677
39. Klarmann, G. J., Schaubert, C. A., and Preston, B. D. (1993) *J Biol Chem* **268**, 9793-9802
40. Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000) *Nature* **407**, 409-413
41. Arhel, N. J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prevost, M. C., Allen, T. D., and Charneau, P. (2007) *EMBO J* **26**, 3025-3037
42. Dismuke, D. J., and Aiken, C. (2006) *J Virol* **80**, 3712-3720
43. Yamashita, M., Perez, O., Hope, T. J., and Emerman, M. (2007) *PLoS Pathog* **3**, 1502-1510
44. Mak, J., Jiang, M., Wainberg, M. A., Hammariskjold, M. L., Rekosh, D., and Kleiman, L. (1994) *J Virol* **68**, 2065-2072
45. Arts, E. J., Ghosh, M., Jacques, P. S., Ehresmann, B., and Le Grice, S. F. (1996) *J Biol Chem* **271**, 9054-9061
46. Isel, C., Keith, G., Ehresmann, B., Ehresmann, C., and Marquet, R. (1998) *Nucleic Acids Res* **26**, 1198-1204
47. Wakefield, J. K., Kang, S. M., and Morrow, C. D. (1996) *J Virol* **70**, 966-975
48. Isel, C., Westhof, E., Massire, C., Le Grice, S. F., Ehresmann, B., Ehresmann, C., and Marquet, R. (1999) *EMBO J* **18**, 1038-1048
49. Yusupova, G., Lanchy, J. M., Yusupov, M., Keith, G., Le Grice, S. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (1996) *J Mol Biol* **261**, 315-321
50. Liu, S., Harada, B. T., Miller, J. T., Le Grice, S. F., and Zhuang, X. (2010) *Nature structural & molecular biology* **17**, 1453-1460
51. Gotte, M. (2011) *Viruses* **3**, 331-335

52. Lanchy, J. M., Ehresmann, C., Le Grice, S. F., Ehresmann, B., and Marquet, R. (1996) *EMBO J* **15**, 7178-7187
53. van Wamel, J. L., and Berkhout, B. (1998) *Virology* **244**, 245-251
54. Palaniappan, C., Kim, J. K., Wisniewski, M., Fay, P. J., and Bambara, R. A. (1998) *J Biol Chem* **273**, 3808-3816
55. Peliska, J. A., Balasubramanian, S., Giedroc, D. P., and Benkovic, S. J. (1994) *Biochemistry* **33**, 13817-13823
56. Powell, M. D., and Levin, J. G. (1996) *J Virol* **70**, 5288-5296
57. Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D., Jr., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) *EMBO J* **20**, 1449-1461
58. Gotte, M., Maier, G., Onori, A. M., Cellai, L., Wainberg, M. A., and Heumann, H. (1999) *J Biol Chem* **274**, 11159-11169
59. Pullen, K. A., Ishimoto, L. K., and Champoux, J. J. (1992) *J Virol* **66**, 367-373
60. Ooms, M., Abbink, T. E., Pham, C., and Berkhout, B. (2007) *Nucleic Acids Res* **35**, 5253-5261
61. Charneau, P., Mirambeau, G., Roux, P., Paulous, S., Buc, H., and Clavel, F. (1994) *J Mol Biol* **241**, 651-662
62. Lavigne, M., Roux, P., Buc, H., and Schaeffer, F. (1997) *J Mol Biol* **266**, 507-524
63. Lavigne, M., and Buc, H. (1999) *J Mol Biol* **285**, 977-995
64. Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., and Charneau, P. (2000) *Cell* **101**, 173-185
65. Lewis, P. F., and Emerman, M. (1994) *J Virol* **68**, 510-516
66. Sherman, M. P., and Greene, W. C. (2002) *Microbes and infection / Institut Pasteur* **4**, 67-73
67. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002) *Cell* **110**, 521-529
68. Ji, H., Moore, D. P., Blomberg, M. A., Braiterman, L. T., Voytas, D. F., Natsoulis, G., and Boeke, J. D. (1993) *Cell* **73**, 1007-1018
69. Kirchner, J., Connolly, C. M., and Sandmeyer, S. B. (1995) *Science* **267**, 1488-1491
70. Jordan, A., Defechereux, P., and Verdin, E. (2001) *EMBO J* **20**, 1726-1738
71. Jouvenet, N., Simon, S. M., and Bieniasz, P. D. (2009) *Proc Natl Acad Sci U S A* **106**, 19114-19119
72. Gartner, S., Markovits, P., Markovitz, D. M., Kaplan, M. H., Gallo, R. C., and Popovic, M. (1986) *Science* **233**, 215-219
73. Pelchen-Matthews, A., Kramer, B., and Marsh, M. (2003) *J Cell Biol* **162**, 443-455
74. Cimorelli, A., Sandin, S., Hoglund, S., and Luban, J. (2000) *J Virol* **74**, 3046-3057
75. Muriaux, D., Mirro, J., Nagashima, K., Harvin, D., and Rein, A. (2002) *J Virol* **76**, 11405-11413
76. Weiss, E. R., and Gottlinger, H. (2011) *J Mol Biol* **410**, 525-533
77. Neil, S. J., Zang, T., and Bieniasz, P. D. (2008) *Nature* **451**, 425-430
78. Van Damme, N., Goff, D., Katsura, C., Jorgenson, R. L., Mitchell, R., Johnson, M. C., Stephens, E. B., and Guatelli, J. (2008) *Cell Host Microbe* **3**, 245-252

79. Wills, J. W., and Craven, R. C. (1991) *AIDS* **5**, 639-654
80. Swanstrom, R., and Wills, J. W. (1997) Synthesis, assembly and processing of viral proteins. in *Retroviruses* (Coffin, J., Hughes, S. H., and Varmus, H. eds.), Cold Spring Harbor Laboratory Press. pp 263-334
81. Parent, L. J., Bennett, R. P., Craven, R. C., Nelle, T. D., Krishna, N. K., Bowzard, J. B., Wilson, C. B., Puffer, B. A., Montelaro, R. C., and Wills, J. W. (1995) *J Virol* **69**, 5455-5460
82. Huang, Y., Khorchid, A., Wang, J., Parniak, M. A., Darlix, J. L., Wainberg, M. A., and Kleiman, L. (1997) *J Virol* **71**, 4378-4384
83. Dorfman, T., Mammano, F., Haseltine, W. A., and Gottlinger, H. G. (1994) *J Virol* **68**, 1689-1696
84. Freed, E. O., and Martin, M. A. (1995) *J Virol* **69**, 1984-1989
85. Ono, A., Huang, M., and Freed, E. O. (1997) *J Virol* **71**, 4409-4418
86. Yu, X., Yuan, X., Matsuda, Z., Lee, T. H., and Essex, M. (1992) *J Virol* **66**, 4966-4971
87. Huang, M., and Martin, M. A. (1997) *J Virol* **71**, 4472-4478
88. Smith, A. J., Srinivasakumar, N., Hammarskjold, M. L., and Rekosh, D. (1993) *J Virol* **67**, 2266-2275
89. Franke, E. K., Yuan, H. E., and Luban, J. (1994) *Nature* **372**, 359-362
90. Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C. T., Sodroski, J., and Gottlinger, H. G. (1994) *Nature* **372**, 363-365
91. Stremlau, M., Owens, C. M., Perron, M. J., Kiessling, M., Autissier, P., and Sodroski, J. (2004) *Nature* **427**, 848-853
92. Maurer, B., Bannert, H., Darai, G., and Flugel, R. M. (1988) *J Virol* **62**, 1590-1597
93. Meric, C., Darlix, J. L., and Spahr, P. F. (1984) *J Mol Biol* **173**, 531-538
94. Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., and Barre-Sinoussi, F. (1990) *J Mol Biol* **216**, 689-699
95. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) *Science* **279**, 384-388
96. Bennett, R. P., Nelle, T. D., and Wills, J. W. (1993) *J Virol* **67**, 6487-6498
97. Franke, E. K., Yuan, H. E., Bossolt, K. L., Goff, S. P., and Luban, J. (1994) *J Virol* **68**, 5300-5305
98. Sandefur, S., Varthakavi, V., and Spearman, P. (1998) *J Virol* **72**, 2723-2732
99. Li, X., Quan, Y., Arts, E. J., Li, Z., Preston, B. D., de Rocquigny, H., Roques, B. P., Darlix, J. L., Kleiman, L., Parniak, M. A., and Wainberg, M. A. (1996) *J Virol* **70**, 4996-5004
100. Berthoux, L., Pechoux, C., Ottmann, M., Morel, G., and Darlix, J. L. (1997) *J Virol* **71**, 6973-6981
101. Strack, B., Calistri, A., Craig, S., Popova, E., and Gottlinger, H. G. (2003) *Cell* **114**, 689-699
102. Huang, M., Orenstein, J. M., Martin, M. A., and Freed, E. O. (1995) *J Virol* **69**, 6810-6818

103. Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988) *Nature* **331**, 280-283
104. Parkin, N. T., Chamorro, M., and Varmus, H. E. (1992) *J Virol* **66**, 5147-5151
105. Ashorn, P., McQuade, T. J., Thaisrivongs, S., Tomasselli, A. G., Tarpley, W. G., and Moss, B. (1990) *Proc Natl Acad Sci U S A* **87**, 7472-7476
106. Wlodawer, A., and Erickson, J. W. (1993) *Annu Rev Biochem* **62**, 543-585
107. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2000) *J Mol Biol* **301**, 1207-1220
108. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) *Structure* **10**, 369-381
109. Acosta, E. P. (2002) *J Acquir Immune Defic Syndr* **29 Suppl 1**, S11-18
110. Rathbun, R. C., and Rossi, D. R. (2002) *Ann Pharmacother* **36**, 702-706
111. Bushman, F. D., Fujiwara, T., and Craigie, R. (1990) *Science* **249**, 1555-1558
112. LaFemina, R. L., Callahan, P. L., and Cordingley, M. G. (1991) *J Virol* **65**, 5624-5630
113. Sherman, P. A., and Fyfe, J. A. (1990) *Proc Natl Acad Sci U S A* **87**, 5119-5123
114. Bukrinsky, M., Sharova, N., and Stevenson, M. (1993) *J Virol* **67**, 6863-6865
115. Gao, K., Gorelick, R. J., Johnson, D. G., and Bushman, F. (2003) *J Virol* **77**, 1598-1603
116. Li, L., Yoder, K., Hansen, M. S., Olvera, J., Miller, M. D., and Bushman, F. D. (2000) *J Virol* **74**, 10965-10974
117. Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003) *J Biol Chem* **278**, 33528-33539
118. Miller, M. D., Farnet, C. M., and Bushman, F. D. (1997) *J Virol* **71**, 5382-5390
119. Cherepanov, P., Ambrosio, A. L., Rahman, S., Ellenberger, T., and Engelman, A. (2005) *Proc Natl Acad Sci U S A* **102**, 17308-17313
120. Hare, S., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2010) *Nature* **464**, 232-236
121. Maertens, G. N., Hare, S., and Cherepanov, P. (2010) *Nature* **468**, 326-329
122. Leavitt, A. D., Rose, R. B., and Varmus, H. E. (1992) *J Virol* **66**, 2359-2368
123. Pommier, Y., Johnson, A. A., and Marchand, C. (2005) *Nat Rev Drug Discov* **4**, 236-248
124. Yoder, K. E., and Bushman, F. D. (2000) *J Virol* **74**, 11191-11200
125. Billamboz, M., Bailly, F., Barreca, M. L., De Luca, L., Mouscadet, J. F., Calmels, C., Andreola, M. L., Witvrouw, M., Christ, F., Debyser, Z., and Cotellet, P. (2008) *J Med Chem* **51**, 7717-7730
126. Allan, J. S., Coligan, J. E., Barin, F., McLane, M. F., Sodroski, J. G., Rosen, C. A., Haseltine, W. A., Lee, T. H., and Essex, M. (1985) *Science* **228**, 1091-1094
127. Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N., and Gregory, T. J. (1990) *J Biol Chem* **265**, 10373-10382
128. Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H. D., and Garten, W. (1992) *Nature* **360**, 358-361
129. Pinter, A., Honnen, W. J., Tilley, S. A., Bona, C., Zaghouani, H., Gorny, M. K., and Zolla-Pazner, S. (1989) *J Virol* **63**, 2674-2679

130. Egan, M. A., Carruth, L. M., Rowell, J. F., Yu, X., and Siliciano, R. F. (1996) *J Virol* **70**, 6547-6556
131. Rowell, J. F., Ruff, A. L., Guarnieri, F. G., Staveley-O'Carroll, K., Lin, X., Tang, J., August, J. T., and Siliciano, R. F. (1995) *J Immunol* **155**, 1818-1828
132. Dimitrov, D. S., Willey, R. L., Sato, H., Chang, L. J., Blumenthal, R., and Martin, M. A. (1993) *J Virol* **67**, 2182-2190
133. Jolly, C., Kashefi, K., Hollinshead, M., and Sattentau, Q. J. (2004) *J Exp Med* **199**, 283-293
134. Waki, K., and Freed, E. O. (2010) *Viruses* **2**, 1603-1620
135. Karn, J. (1999) *J Mol Biol* **293**, 235-254
136. Harrich, D., Hooker, C. W., and Parry, E. (2000) *J Virol* **74**, 5639-5646
137. Wu-Baer, F., Sigman, D., and Gaynor, R. B. (1995) *Proc Natl Acad Sci U S A* **92**, 7153-7157
138. Nekhai, S., Zhou, M., Fernandez, A., Lane, W. S., Lamb, N. J., Brady, J., and Kumar, A. (2002) *Biochem J* **364**, 649-657
139. Chiu, Y. L., Coronel, E., Ho, C. K., Shuman, S., and Rana, T. M. (2001) *J Biol Chem* **276**, 12959-12966
140. Brigati, C., Giacca, M., Noonan, D. M., and Albin, A. (2003) *FEMS Microbiol Lett* **220**, 57-65
141. Chen, D., Wang, M., Zhou, S., and Zhou, Q. (2002) *EMBO J* **21**, 6801-6810
142. Kekow, J., Wachsmann, W., McCutchan, J. A., Cronin, M., Carson, D. A., and Lotz, M. (1990) *Proc Natl Acad Sci U S A* **87**, 8321-8325
143. Sastry, K. J., Marin, M. C., Nehete, P. N., McConnell, K., el-Naggar, A. K., and McDonnell, T. J. (1996) *Oncogene* **13**, 487-493
144. Yang, Y., Dong, B., Mittelstadt, P. R., Xiao, H., and Ashwell, J. D. (2002) *J Biol Chem* **277**, 19482-19487
145. Yang, Y., Tikhonov, I., Ruckwardt, T. J., Djavani, M., Zapata, J. C., Pauza, C. D., and Salvato, M. S. (2003) *J Virol* **77**, 6700-6708
146. Zhang, M., Li, X., Pang, X., Ding, L., Wood, O., Clouse, K., Hewlett, I., and Dayton, A. I. (2001) *J Biomed Sci* **8**, 290-296
147. Bonifaci, N., Sitia, R., and Rubartelli, A. (1995) *AIDS* **9**, 995-1000
148. Huang, L., Bosch, I., Hofmann, W., Sodroski, J., and Pardee, A. B. (1998) *J Virol* **72**, 8952-8960
149. Olsen, H. S., Cochrane, A. W., Dillon, P. J., Nalin, C. M., and Rosen, C. A. (1990) *Genes Dev* **4**, 1357-1364
150. Fischer, U., Pollard, V. W., Luhrmann, R., Teufel, M., Michael, M. W., Dreyfuss, G., and Malim, M. H. (1999) *Nucleic Acids Res* **27**, 4128-4134
151. Qu, J., Yang, Z., Zhang, Q., Liu, W., Li, Y., Ding, Q., Liu, F., Liu, Y., Pan, Z., He, B., Zhu, Y., and Wu, J. (2011) *FEBS letters* **585**, 4002-4009
152. Roeth, J. F., and Collins, K. L. (2006) *Microbiol Mol Biol Rev* **70**, 548-563
153. Cohen, G. B., Gandhi, R. T., Davis, D. M., Mandelboim, O., Chen, B. K., Strominger, J. L., and Baltimore, D. (1999) *Immunity* **10**, 661-671
154. Fenard, D., Yonemoto, W., de Noronha, C., Cavrois, M., Williams, S. A., and Greene, W. C. (2005) *J Immunol* **175**, 6050-6057

155. Fortin, J. F., Barat, C., Beausejour, Y., Barbeau, B., and Tremblay, M. J. (2004) *J Biol Chem* **279**, 39520-39531
156. Swinger, S., Brichacek, B., Jacque, J. M., Ulich, C., Zhou, J., and Stevenson, M. (2003) *Nature* **424**, 213-219
157. Swinger, S., Mann, A., Jacque, J., Brichacek, B., Sasseville, V. G., Williams, K., Lackner, A. A., Janoff, E. N., Wang, R., Fisher, D., and Stevenson, M. (1999) *Nat Med* **5**, 997-103
158. James, C. O., Huang, M. B., Khan, M., Garcia-Barrio, M., Powell, M. D., and Bond, V. C. (2004) *J Virol* **78**, 3099-3109
159. Lenassi, M., Cagney, G., Liao, M., Vaupotic, T., Bartholomeeusen, K., Cheng, Y., Krogan, N. J., Plemenitas, A., and Peterlin, B. M. (2010) *Traffic* **11**, 110-122
160. Bour, S., Schubert, U., and Strebel, K. (1995) *J Virol* **69**, 1510-1520
161. Willey, R. L., Maldarelli, F., Martin, M. A., and Strebel, K. (1992) *J Virol* **66**, 7193-7200
162. Douglas, J. L., Viswanathan, K., McCarroll, M. N., Gustin, J. K., Fruh, K., and Moses, A. V. (2009) *J Virol* **83**, 7931-7947
163. Akari, H., Bour, S., Kao, S., Adachi, A., and Strebel, K. (2001) *J Exp Med* **194**, 1299-1311
164. Nishizawa, M., Myojin, T., Nishino, Y., Nakai, Y., Kamata, M., and Aida, Y. (1999) *Virology* **263**, 313-322
165. Sherman, M. P., Schubert, U., Williams, S. A., de Noronha, C. M., Kreisberg, J. F., Henklein, P., and Greene, W. C. (2002) *Virology* **302**, 95-105
166. Feldherr, C. M., and Feldherr, A. B. (1960) *Nature* **185**, 250-251
167. Kogan, M., and Rappaport, J. (2011) *Retrovirology* **8**, 25
168. Connor, R. I., Chen, B. K., Choe, S., and Landau, N. R. (1995) *Virology* **206**, 935-944
169. Tristem, M., Marshall, C., Karpas, A., and Hill, F. (1992) *EMBO J* **11**, 3405-3412
170. Fletcher, T. M., 3rd, Brichacek, B., Sharova, N., Newman, M. A., Stivahtis, G., Sharp, P. M., Emerman, M., Hahn, B. H., and Stevenson, M. (1996) *EMBO J* **15**, 6155-6165
171. Laguette, N., Sobhian, B., Casartelli, N., Ringear, M., Chable-Bessia, C., Segeral, E., Yatim, A., Emiliani, S., Schwartz, O., and Benkirane, M. (2011) *Nature* **474**, 654-657
172. Sheehy, A. M., Gaddis, N. C., Choi, J. D., and Malim, M. H. (2002) *Nature* **418**, 646-650
173. Bogerd, H. P., and Cullen, B. R. (2008) *RNA* **14**, 1228-1236
174. Harris, R. S., and Liddament, M. T. (2004) *Nat Rev Immunol* **4**, 868-877
175. Holmes, R. K., Malim, M. H., and Bishop, K. N. (2007) *Trends Biochem Sci* **32**, 118-128
176. Mehle, A., Goncalves, J., Santa-Marta, M., McPike, M., and Gabuzda, D. (2004) *Genes Dev* **18**, 2861-2866
177. Russell, R. A., and Pathak, V. K. (2007) *J Virol* **81**, 8201-8210
178. Baltimore, D. (1970) *Nature* **226**, 1209-1211
179. Temin, H. M., and Mizutani, S. (1970) *Nature* **226**, 1211-1213

180. Gotte, M., Fackler, S., Hermann, T., Perola, E., Cellai, L., Gross, H. J., Le Grice, S. F., and Heumann, H. (1995) *EMBO J* **14**, 833-841
181. Gotte, M., Maier, G., Gross, H. J., and Heumann, H. (1998) *J Biol Chem* **273**, 10139-10146
182. Cameron, C. E., Ghosh, M., Le Grice, S. F., and Benkovic, S. J. (1997) *Proc Natl Acad Sci U S A* **94**, 6700-6705
183. Arts, E. J., Stetor, S. R., Li, X., Rausch, J. W., Howard, K. J., Ehresmann, B., North, T. W., Wohrl, B. M., Goody, R. S., Wainberg, M. A., and Grice, S. F. (1996) *Proc Natl Acad Sci U S A* **93**, 10063-10068
184. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., and et al. (1993) *Proc Natl Acad Sci U S A* **90**, 6320-6324
185. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) *Science* **256**, 1783-1790
186. Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark, A. D., Jr., Jacobo-Molina, A., Tantillo, C., Hughes, S. H., and Arnold, E. (1998) *J Mol Biol* **284**, 1095-1111
187. Ding, J., Hughes, S. H., and Arnold, E. (1997) *Biopolymers* **44**, 125-138
188. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) *Science* **282**, 1669-1675
189. Bohlayer, W. P., and DeStefano, J. J. (2006) *Biochemistry* **45**, 7628-7638
190. Huber, H. E., McCoy, J. M., Seehra, J. S., and Richardson, C. C. (1989) *J Biol Chem* **264**, 4669-4678
191. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) *J Biol Chem* **267**, 25988-25997
192. Reardon, J. E. (1992) *Biochemistry* **31**, 4473-4479
193. Reardon, J. E., Furfine, E. S., and Cheng, N. (1991) *J Biol Chem* **266**, 14128-14134
194. Ehteshami, M., Beilhartz, G. L., Scarth, B. J., Tchesnokov, E. P., McCormick, S., Wynhoven, B., Harrigan, P. R., and Gotte, M. (2008) *J Biol Chem* **283**, 22222-22232
195. Arnott, S., Chandrasekaran, R., Millane, R. P., and Park, H. S. (1986) *J Mol Biol* **188**, 631-640
196. Fedoroff, O., Salazar, M., and Reid, B. R. (1993) *J Mol Biol* **233**, 509-523
197. Lane, A. N., Ebel, S., and Brown, T. (1993) *Eur J Biochem* **215**, 297-306
198. Cote, M. L., and Roth, M. J. (2008) *Virus Res* **134**, 186-202
199. Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T., and Morikawa, K. (1990) *Nature* **347**, 306-309
200. Gutierrez-Rivas, M., and Menendez-Arias, L. (2001) *Nucleic Acids Res* **29**, 4963-4972
201. Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2002) *Proc Natl Acad Sci U S A* **99**, 9515-9520
202. Rausch, J. W., Lener, D., Miller, J. T., Julias, J. G., Hughes, S. H., and Le Grice, S. F. (2002) *Biochemistry* **41**, 4856-4865

203. Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Alvord, W. G., Arnold, E., and Hughes, S. H. (2003) *J Virol* **77**, 8548-8554
204. Larder, B. A., Purifoy, D. J., Powell, K. L., and Darby, G. (1987) *Nature* **327**, 716-717
205. Poch, O., Sauvaget, I., Delarue, M., and Tordo, N. (1989) *EMBO J* **8**, 3867-3874
206. Martin-Hernandez, A. M., Domingo, E., and Menendez-Arias, L. (1996) *EMBO J* **15**, 4434-4442
207. Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H., and Arnold, E. (2009) *J Mol Biol* **385**, 693-713
208. Gotte, M. (2006) *Curr Pharm Des* **12**, 1867-1877
209. Marchand, B., and Gotte, M. (2003) *J Biol Chem* **278**, 35362-35372
210. Davies, J. F., 2nd, Hostomska, Z., Hostomsky, Z., Jordan, S. R., and Matthews, D. A. (1991) *Science* **252**, 88-95
211. Lim, D., Gregorio, G. G., Bingman, C., Martinez-Hackert, E., Hendrickson, W. A., and Goff, S. P. (2006) *J Virol* **80**, 8379-8389
212. Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990) *Science* **249**, 1398-1405
213. Nowotny, M., Gaidamakov, S. A., Crouch, R. J., and Yang, W. (2005) *Cell* **121**, 1005-1016
214. Nowotny, M., Gaidamakov, S. A., Ghirlando, R., Cerritelli, S. M., Crouch, R. J., and Yang, W. (2007) *Mol Cell* **28**, 264-276
215. Keck, J. L., and Marqusee, S. (1996) *J Biol Chem* **271**, 19883-19887
216. Keck, J. L., and Marqusee, S. (1995) *Proc Natl Acad Sci U S A* **92**, 2740-2744
217. Smith, J. S., Gritsman, K., and Roth, M. J. (1994) *J Virol* **68**, 5721-5729
218. Klumpp, K., Hang, J. Q., Rajendran, S., Yang, Y., Derosier, A., Wong Kai In, P., Overton, H., Parkes, K. E., Cammack, N., and Martin, J. A. (2003) *Nucleic Acids Res* **31**, 6852-6859
219. Yang, W., Lee, J. Y., and Nowotny, M. (2006) *Mol Cell* **22**, 5-13
220. Nowotny, M., and Yang, W. (2006) *Embo J* **25**, 1924-1933
221. Bochner, R., Duvshani, A., Adir, N., and Hizi, A. (2008) *FEBS Lett* **582**, 2799-2805
222. Sevilya, Z., Loya, S., Adir, N., and Hizi, A. (2003) *Nucleic Acids Res* **31**, 1481-1487
223. DeStefano, J. J., Mallaber, L. M., Fay, P. J., and Bambara, R. A. (1993) *Nucleic Acids Res* **21**, 4330-4338
224. Furfine, E. S., and Reardon, J. E. (1991) *J Biol Chem* **266**, 406-412
225. Gopalakrishnan, V., Peliska, J. A., and Benkovic, S. J. (1992) *Proc Natl Acad Sci U S A* **89**, 10763-10767
226. Wohrl, B. M., Volkmann, S., and Moelling, K. (1991) *J Mol Biol* **220**, 801-818
227. Beilhartz, G. L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and Gotte, M. (2009) *J Mol Biol*
228. Marchand, B., Tchesnokov, E. P., and Gotte, M. (2007) *J Biol Chem* **282**, 3337-3346

229. Tong, W., Lu, C. D., Sharma, S. K., Matsuura, S., So, A. G., and Scott, W. A. (1997) *Biochemistry* **36**, 5749-5757
230. Sarafianos, S. G., Clark, A. D., Jr., Das, K., Tuske, S., Birktoft, J. J., Ilankumaran, P., Ramesha, A. R., Sayer, J. M., Jerina, D. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2002) *EMBO J* **21**, 6614-6624
231. Champoux, J. J., and Schultz, S. J. (2009) *Febs J* **276**, 1506-1516
232. Schultz, S. J., Zhang, M., and Champoux, J. J. (2004) *J Mol Biol* **344**, 635-652
233. Schultz, S. J., and Champoux, J. J. (2008) *Virus Res* **134**, 86-103
234. Schultz, S. J., Zhang, M., and Champoux, J. J. (2006) *J Biol Chem* **281**, 1943-1955
235. Schultz, S. J., Zhang, M., and Champoux, J. J. (2009) *J Biol Chem* **284**, 32225-32238
236. Abbondanzieri, E. A., Bokinsky, G., Rausch, J. W., Zhang, J. X., Le Grice, S. F., and Zhuang, X. (2008) *Nature* **453**, 184-189
237. Fuentes, G. M., Rodriguez-Rodriguez, L., Fay, P. J., and Bambara, R. A. (1995) *J Biol Chem* **270**, 28169-28176
238. Huber, H. E., and Richardson, C. C. (1990) *J Biol Chem* **265**, 10565-10573
239. Palaniappan, C., Fay, P. J., and Bambara, R. A. (1995) *J Biol Chem* **270**, 4861-4869
240. Pullen, K. A., and Champoux, J. J. (1990) *J Virol* **64**, 6274-6277
241. Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Alvord, W. G., Arnold, E., and Hughes, S. H. (2004) *J Virol* **78**, 13315-13324
242. Grobler, J. A., Dornadula, G., Rice, M. R., Simcoe, A. L., Hazuda, D. J., and Miller, M. D. (2007) *J Biol Chem* **282**, 8005-8010
243. Gao, L., Balakrishnan, M., Roques, B. P., and Bambara, R. A. (2007) *J Biol Chem* **282**, 6222-6231
244. Chen, Y., Balakrishnan, M., Roques, B. P., and Bambara, R. A. (2005) *J Biol Chem* **280**, 14443-14452
245. Hwang, C. K., Svarovskaia, E. S., and Pathak, V. K. (2001) *Proc Natl Acad Sci U S A* **98**, 12209-12214
246. Telesnitsky, A., and Goff, S. P. (1993) *Embo J* **12**, 4433-4438
247. Smith, J. S., and Roth, M. J. (1992) *J Biol Chem* **267**, 15071-15079
248. Bushman, F. D., and Craigie, R. (1991) *Proc Natl Acad Sci U S A* **88**, 1339-1343
249. Oh, J., McWilliams, M. J., Julias, J. G., and Hughes, S. H. (2008) *J Virol* **82**, 719-727
250. Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D., and et al. (1995) *Nat Struct Biol* **2**, 293-302
251. Archer, R. H., Dykes, C., Gerondelis, P., Lloyd, A., Fay, P., Reichman, R. C., Bambara, R. A., and Demeter, L. M. (2000) *J Virol* **74**, 8390-8401
252. Brehm, J. H., Koontz, D., Meteer, J. D., Pathak, V., Sluis-Cremer, N., and Mellors, J. W. (2007) *J Virol* **81**, 7852-7859
253. Cane, P. A., Green, H., Fearnhill, E., and Dunn, D. (2007) *Aids* **21**, 447-455
254. Gotte, M. (2007) *PLoS Med* **4**, e346

255. Hachiya, A., Kodama, E. N., Sarafianos, S. G., Schuckmann, M. M., Sakagami, Y., Matsuoka, M., Takiguchi, M., Gatanaga, H., and Oka, S. (2008) *J Virol* **82**, 3261-3270
256. Ntemgwa, M., Wainberg, M. A., Oliveira, M., Moisi, D., Lalonde, R., Micheli, V., and Brenner, B. G. (2007) *Antimicrob Agents Chemother* **51**, 3861-3869
257. Santos, A. F., Lengruher, R. B., Soares, E. A., Jere, A., Sprinz, E., Martinez, A. M., Silveira, J., Sion, F. S., Pathak, V. K., and Soares, M. A. (2008) *PLoS One* **3**, e1781
258. Yap, S. H., Sheen, C. W., Fahey, J., Zanin, M., Tyssen, D., Lima, V. D., Wynhoven, B., Kuiper, M., Sluis-Cremer, N., Harrigan, P. R., and Tachedjian, G. (2007) *PLoS Med* **4**, e335
259. Delviks-Frankenberry, K. A., Nikolenko, G. N., Barr, R., and Pathak, V. K. (2007) *J Virol* **81**, 6837-6845
260. Delviks-Frankenberry, K. A., Nikolenko, G. N., Boyer, P. L., Hughes, S. H., Coffin, J. M., Jere, A., and Pathak, V. K. (2008) *Proc Natl Acad Sci U S A* **105**, 10943-10948
261. Nikolenko, G. N., Delviks-Frankenberry, K. A., Palmer, S., Maldarelli, F., Fivash, M. J., Jr., Coffin, J. M., and Pathak, V. K. (2007) *Proc Natl Acad Sci U S A* **104**, 317-322
262. Nikolenko, G. N., Palmer, S., Maldarelli, F., Mellors, J. W., Coffin, J. M., and Pathak, V. K. (2005) *Proc Natl Acad Sci U S A* **102**, 2093-2098
263. Radzio, J., and Sluis-Cremer, N. (2008) *Mol Pharmacol* **73**, 601-606
264. Ehteshami, M., Beilhartz, G. L., Scarth, B. J., Tchesnokov, E. P., McCormick, S., Wynhoven, B., Harrigan, P. R., and Gotte, M. (2008) *J Biol Chem*
265. Brehm, J. H., Mellors, J. W., and Sluis-Cremer, N. (2008) *Biochemistry* **47**, 14020-14027
266. Hachiya, A., Shimane, K., Sarafianos, S. G., Kodama, E. N., Sakagami, Y., Negishi, F., Koizumi, H., Gatanaga, H., Matsuoka, M., Takiguchi, M., and Oka, S. (2009) *Antiviral Res* **82**, 115-121
267. Ehteshami, M., and Gotte, M. (2008) *AIDS Rev* **10**, 224-235
268. Mocroft, A., Vella, S., Benfield, T. L., Chiesi, A., Miller, V., Gargalianos, P., d'Arminio Monforte, A., Yust, I., Bruun, J. N., Phillips, A. N., and Lundgren, J. D. (1998) *Lancet* **352**, 1725-1730
269. De Clercq, E. (2010) *Current opinion in pharmacology* **10**, 507-515
270. Colucci, P., Pottage, J. C., Robison, H., Turgeon, J., and Ducharme, M. P. (2009) *Antimicrob Agents Chemother* **53**, 646-650
271. Hurwitz, S. J., Otto, M. J., and Schinazi, R. F. (2005) *Antiviral chemistry & chemotherapy* **16**, 117-127
272. Cahn, P., and Wainberg, M. A. (2010) *The Journal of antimicrobial chemotherapy* **65**, 213-217
273. Madruga, J. V., Cahn, P., Grinsztejn, B., Haubrich, R., Lalezari, J., Mills, A., Pialoux, G., Wilkin, T., Peeters, M., Vingerhoets, J., de Smedt, G., Leopold, L., Trefiglio, R., and Woodfall, B. (2007) *Lancet* **370**, 29-38

274. Azijn, H., Tirry, I., Vingerhoets, J., de Bethune, M. P., Kraus, G., Boven, K., Jochmans, D., Van Craenenbroeck, E., Picchio, G., and Rimsky, L. T. (2010) *Antimicrob Agents Chemother* **54**, 718-727
275. Hostetler, K. Y., Aldern, K. A., Wan, W. B., Ciesla, S. L., and Beadle, J. R. (2006) *Antimicrob Agents Chemother* **50**, 2857-2859
276. Budihias, S. R., Gorshkova, I., Gaidamakov, S., Wamiru, A., Bona, M. K., Parniak, M. A., Crouch, R. J., McMahon, J. B., Beutler, J. A., and Le Grice, S. F. (2005) *Nucleic Acids Res* **33**, 1249-1256
277. Beilhartz, G. L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and Gotte, M. (2009) *J Mol Biol* **388**, 462-474
278. De Clercq, E. (2009) *Int J Antimicrob Agents* **33**, 307-320
279. Parkes, K. E., Ermert, P., Fassler, J., Ives, J., Martin, J. A., Merrett, J. H., Obrecht, D., Williams, G., and Klumpp, K. (2003) *J Med Chem* **46**, 1153-1164
280. Hang, J. Q., Rajendran, S., Yang, Y., Li, Y., In, P. W., Overton, H., Parkes, K. E., Cammack, N., Martin, J. A., and Klumpp, K. (2004) *Biochem Biophys Res Commun* **317**, 321-329
281. Klumpp, K., and Mirzadegan, T. (2006) *Curr Pharm Des* **12**, 1909-1922
282. Grobler, J. A., Stillmock, K., Hu, B., Witmer, M., Felock, P., Espeseth, A. S., Wolfe, A., Egbertson, M., Bourgeois, M., Melamed, J., Wai, J. S., Young, S., Vacca, J., and Hazuda, D. J. (2002) *Proc Natl Acad Sci U S A* **99**, 6661-6666
283. Hazuda, D. J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J. A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C., and Miller, M. D. (2000) *Science* **287**, 646-650
284. Shaw-Reid, C. A., Munshi, V., Graham, P., Wolfe, A., Witmer, M., Danzeisen, R., Olsen, D. B., Carroll, S. S., Embrey, M., Wai, J. S., Miller, M. D., Cole, J. L., and Hazuda, D. J. (2003) *J Biol Chem* **278**, 2777-2780
285. Tramontano, E., Esposito, F., Badas, R., Di Santo, R., Costi, R., and La Colla, P. (2005) *Antiviral Res* **65**, 117-124
286. Shaw-Reid, C. A., Feuston, B., Munshi, V., Getty, K., Krueger, J., Hazuda, D. J., Parniak, M. A., Miller, M. D., and Lewis, D. (2005) *Biochemistry* **44**, 1595-1606
287. Kirschberg, T. A., Balakrishnan, M., Squires, N. H., Barnes, T., Brendza, K. M., Chen, X., Eisenberg, E. J., Jin, W., Kutty, N., Leavitt, S., Licican, A., Liu, Q., Liu, X., Mak, J., Perry, J. K., Wang, M., Watkins, W. J., and Lansdon, E. B. (2009) *J Med Chem* **52**, 5781-5784
288. Borkow, G., Fletcher, R. S., Barnard, J., Arion, D., Motakis, D., Dmitrienko, G. I., and Parniak, M. A. (1997) *Biochemistry* **36**, 3179-3185
289. Keck, J. L., Goedken, E. R., and Marqusee, S. (1998) *J Biol Chem* **273**, 34128-34133
290. Tsunaka, Y., Takano, K., Matsumura, H., Yamagata, Y., and Kanaya, S. (2005) *J Mol Biol* **345**, 1171-1183
291. Sluis-Cremer, N., Arion, D., and Parniak, M. A. (2002) *Mol Pharmacol* **62**, 398-405
292. Arion, D., Sluis-Cremer, N., Min, K. L., Abram, M. E., Fletcher, R. S., and Parniak, M. A. (2002) *J Biol Chem* **277**, 1370-1374

293. Himmel, D. M., Sarafianos, S. G., Dharmasena, S., Hossain, M. M., McCoy-Simandle, K., Ilina, T., Clark, A. D., Jr., Knight, J. L., Julias, J. G., Clark, P. K., Krogh-Jespersen, K., Levy, R. M., Hughes, S. H., Parniak, M. A., and Arnold, E. (2006) *ACS Chem Biol* **1**, 702-712
294. Wendeler, M., Lee, H. F., Bermingham, A., Miller, J. T., Chertov, O., Bona, M. K., Baichoo, N. S., Ehteshami, M., Beutler, J., O'Keefe, B. R., Gotte, M., Kvaratskhelia, M., and Le Grice, S. (2008) *ACS Chem Biol* **3**, 635-644
295. Gao, H. Q., Boyer, P. L., Arnold, E., and Hughes, S. H. (1998) *J Mol Biol* **277**, 559-572
296. Jacques, P. S., Wohrl, B. M., Howard, K. J., and Le Grice, S. F. (1994) *J Biol Chem* **269**, 1388-1393
297. Loya, S., Tal, R., Kashman, Y., and Hizi, A. (1990) *Antimicrob Agents Chemother* **34**, 2009-2012
298. Min, B. S., Miyashiro, H., and Hattori, M. (2002) *Phytother Res* **16 Suppl 1**, S57-62
299. Hannoush, R. N., Carriero, S., Min, K. L., and Damha, M. J. (2004) *Chembiochem* **5**, 527-533
300. James, W. (2007) *J Gen Virol* **88**, 351-364
301. Kissel, J. D., Held, D. M., Hardy, R. W., and Burke, D. H. (2007) *AIDS Res Hum Retroviruses* **23**, 699-708
302. Ji, J. P., and Loeb, L. A. (1992) *Biochemistry* **31**, 954-958
303. Hill, A. L., Rosenbloom, D. I., and Nowak, M. A. (2012) *J Mol Med (Berl)*
304. Goldschmidt, V., and Marquet, R. (2004) *The international journal of biochemistry & cell biology* **36**, 1687-1705
305. Gotte, M. (2004) *Expert Rev Anti Infect Ther* **2**, 707-716
306. Menendez-Arias, L. (2008) *Virus Res* **134**, 124-146
307. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2001) *J Virol* **75**, 4832-4842
308. Tu, X., Das, K., Han, Q., Bauman, J. D., Clark, A. D., Jr., Hou, X., Frenkel, Y. V., Gaffney, B. L., Jones, R. A., Boyer, P. L., Hughes, S. H., Sarafianos, S. G., and Arnold, E. (2010) *Nature structural & molecular biology* **17**, 1202-1209
309. Lafeuillade, A., and Tardy, J. C. (2003) *AIDS Rev* **5**, 80-86
310. Margot, N. A., Isaacson, E., McGowan, I., Cheng, A. K., Schooley, R. T., and Miller, M. D. (2002) *AIDS* **16**, 1227-1235
311. Catucci, M., Venturi, G., Romano, L., Riccio, M. L., De Milito, A., Valensin, P. E., and Zazzi, M. (1999) *J Acquir Immune Defic Syndr* **21**, 203-208
312. Larder, B. A., Kemp, S. D., and Harrigan, P. R. (1995) *Science* **269**, 696-699
313. Gao, H. Q., Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2000) *J Mol Biol* **300**, 403-418
314. Matsumi, S., Kosalaraksa, P., Tsang, H., Kavlick, M. F., Harada, S., and Mitsuya, H. (2003) *AIDS* **17**, 1127-1137
315. Deval, J., Selmi, B., Boretto, J., Egloff, M. P., Guerreiro, C., Sarfati, S., and Canard, B. (2002) *J Biol Chem* **277**, 42097-42104

316. Valer, L., Martin-Carbonero, L., de Mendoza, C., Corral, A., and Soriano, V. (2004) *AIDS* **18**, 2094-2096
317. Delaugerre, C., Flandre, P., Marcelin, A. G., Descamps, D., Tamalet, C., Cottalorda, J., Schneider, V., Yerly, S., LeGoff, J., Morand-Joubert, L., Chaix, M. L., Costagliola, D., and Calvez, V. (2008) *Journal of medical virology* **80**, 762-765
318. Sluis-Cremer, N., Sheen, C. W., Zelina, S., Torres, P. S., Parikh, U. M., and Mellors, J. W. (2007) *Antimicrob Agents Chemother* **51**, 48-53
319. McColl, D. J., Chappey, C., Parkin, N. T., and Miller, M. D. (2008) *Antiviral therapy* **13**, 189-197
320. Miranda, L. R., Gotte, M., Liang, F., and Kuritzkes, D. R. (2005) *Antimicrob Agents Chemother* **49**, 2648-2656
321. Lacey, S. F., and Larder, B. A. (1994) *Antimicrob Agents Chemother* **38**, 1428-1432
322. Ren, J., and Stammers, D. K. (2008) *Virus Res* **134**, 157-170
323. Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D., and Stuart, D. (1995) *Nature structural biology* **2**, 303-308
324. De Clercq, E. (1994) *Biochemical pharmacology* **47**, 155-169
325. Ren, J., Nichols, C. E., Chamberlain, P. P., Weaver, K. L., Short, S. A., Chan, J. H., Kleim, J. P., and Stammers, D. K. (2007) *J Med Chem* **50**, 2301-2309
326. Fujiwara, T., Sato, A., el-Farrash, M., Miki, S., Abe, K., Isaka, Y., Kodama, M., Wu, Y., Chen, L. B., Harada, H., Sugimoto, H., Hatanaka, M., and Hinuma, Y. (1998) *Antimicrob Agents Chemother* **42**, 1340-1345
327. Smerdon, S. J., Jager, J., Wang, J., Kohlstaedt, L. A., Chirino, A. J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1994) *Proc Natl Acad Sci U S A* **91**, 3911-3915
328. Himmel, D. M., Das, K., Clark, A. D., Jr., Hughes, S. H., Benjahad, A., Oumouch, S., Guillemont, J., Coupa, S., Poncelet, A., Csoka, I., Meyer, C., Andries, K., Nguyen, C. H., Grierson, D. S., and Arnold, E. (2005) *J Med Chem* **48**, 7582-7591
329. Das, K., Bauman, J. D., Clark, A. D., Jr., Frenkel, Y. V., Lewi, P. J., Shatkin, A. J., Hughes, S. H., and Arnold, E. (2008) *Proc Natl Acad Sci U S A* **105**, 1466-1471
330. Miller, C. D., Crain, J., Tran, B., and Patel, N. (2011) *Drugs Today (Barc)* **47**, 5-15
331. De Clercq, E. (2004) *Nat Rev Microbiol* **2**, 704-720
332. Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Jr., Swanstrom, R., Burch, C. L., and Weeks, K. M. (2009) *Nature* **460**, 711-716
333. Telesnitsky, A., and Goff, S. (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA
334. De Clercq, E. (2007) *Nat Rev Drug Discov* **6**, 1001-1018
335. Hang, J. Q., Li, Y., Yang, Y., Cammack, N., Mirzadegan, T., and Klumpp, K. (2007) *Biochem Biophys Res Commun* **352**, 341-350
336. Tchesnokov, E. P., Obikhod, A., Schinazi, R. F., and Gotte, M. (2008) *J Biol Chem* **283**, 34218-34228

337. Didierjean, J., Isel, C., Querre, F., Mouscadet, J. F., Aubertin, A. M., Valnot, J. Y., Piettre, S. R., and Marquet, R. (2005) *Antimicrob Agents Chemother* **49**, 4884-4894
338. Furfine, E. S., and Reardon, J. E. (1991) *Biochemistry* **30**, 7041-7046
339. Parniak, M. A., Min, K. L., Budihias, S. R., Le Grice, S. F., and Beutler, J. A. (2003) *Anal Biochem* **322**, 33-39
340. Guajardo, R., and Sousa, R. (1997) *J Mol Biol* **265**, 8-19
341. Meyer, P. R., Rutvisuttinunt, W., Matsuura, S. E., So, A. G., and Scott, W. A. (2007) *J Mol Biol* **369**, 41-54
342. Purohit, V., Roques, B. P., Kim, B., and Bambara, R. A. (2007) *J Biol Chem* **282**, 12598-12609
343. Jochmans, D., Deval, J., Kesteleyn, B., Van Marck, H., Bettens, E., De Baere, I., Dehertogh, P., Ivens, T., Van Ginderen, M., Van Schoubroeck, B., Ehteshami, M., Wigerinck, P., Gotte, M., and Hertogs, K. (2006) *J Virol* **80**, 12283-12292
344. Le Grice, S. F., and Gruninger-Leitch, F. (1990) *Eur J Biochem* **187**, 307-314
345. Ryu, Y., and Schultz, P. G. (2006) *Nat Methods* **3**, 263-265
346. Klarmann, G. J., Eisenhauer, B. M., Zhang, Y., Sitaraman, K., Chatterjee, D. K., Hecht, S. M., and Le Grice, S. F. (2004) *Protein Expr Purif* **38**, 37-44
347. Himmel, D. M., Maegley, K. A., Pauly, T. A., Bauman, J. D., Das, K., Dharia, C., Clark, A. D., Jr., Ryan, K., Hickey, M. J., Love, R. A., Hughes, S. H., Bergqvist, S., and Arnold, E. (2009) *Structure* **17**, 1625-1635
348. Su, H. P., Yan, Y., Prasad, G. S., Smith, R. F., Daniels, C. L., Abeywickrema, P. D., Reid, J. C., Loughran, H. M., Kornienko, M., Sharma, S., Grobler, J. A., Xu, B., Sardana, V., Allison, T. J., Williams, P. D., Darke, P. L., Hazuda, D. J., and Munshi, S. (2010) *J Virol* **84**, 7625-7633
349. Chung, S., Wendeler, M., Rausch, J. W., Beilhartz, G., Gotte, M., O'Keefe, B. R., Bermingham, A., Beutler, J. A., Liu, S., Zhuang, X., and Le Grice, S. F. (2010) *Antimicrob Agents Chemother* **54**, 3913-3921
350. Auger, A., Beilhartz, G. L., Zhu, S., Cauchon, E., Falguyret, J. P., Grobler, J. A., Ehteshami, M., Gotte, M., and Melnyk, R. A. (2011) *J Biol Chem* **286**, 29575-29583
351. Burney, T., and Dusheiko, G. (2011) *Expert Rev Anti Infect Ther* **9**, 151-160
352. Foote, B. C., Spooner, L. M., and Belliveau, P. P. (2011) *Ann Pharmacother*
353. Lin, C. (2006)
354. Cauchon, E., Falguyret, J. P., Auger, A., and Melnyk, R. A. (2011) *J Biomol Screen* **16**, 518-524