STRUCTURAL BASIS FOR HIV-1 REVERSE TRANSCRIPTASE MATURATION

by

Ryan Lee Slack

Bachelor of Science, Centenary College, 2006

Submitted to the Graduate Faculty of

The School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Ryan Lee Slack

It was defended on

December 14, 2018

and approved by

Jinwoo Ahn, Associate Professor, Department of Structural Biology

Zandrea Ambrose, Associate Professor, Department of Microbiology and Molecular Genetics

Pei Tang, Professor of Anesthesiology, Pharmacology & Chemical Biology, and Computational and Systems Biology

Dissertation Director: Rieko Ishima, Associate Professor, Department of Structural Biology

STRUCTURAL BASIS FOR HIV-1 REVERSE TRANSCRIPTASE MATURATION

Ryan Lee Slack, PhD

University of Pittsburgh, 2019

Reverse transcriptase (RT) is the viral enzyme responsible for generating a double stranded DNA copy of the viral RNA genome, which is an essential step in the replication process of Human Immunodeficiency Virus-1. Mature RT consists of a 66 kDa subunit (p66) and a smaller 51 kDa subunit. The smaller subunit, p51, is formed from proteolytic cleavage between residues F440 and Y441 in the ribonuclease H domain (RNH) of p66 by HIV-1 protease (PR). Evidence suggests that the RT heterodimer is formed via processing of a p66/p66 homodimer intermediate, but a structure of this homodimer has yet to be reported. The p51-RNH processing site is located between residues 440/441 and lies in the β-sheet region of the RNH domain. According to available structures of the RNH domain, this site would be inaccessible to PR. Because the available structural data fails to explain the molecular basis for this processing event, the mechanism of RT maturation remains unclear. In an effort to clarify the mechanism of RT maturation, we initially studied structural integrity of the RNH domain by introducing mutations at the p51-RNH processing site. We proceeded to characterize the conformational symmetry and long-term stability of p66/p66 homodimer using principal component analysis of solution NMR spectra. We then studied in vitro maturation of RT, and the effect of nucleic acids on the kinetics and specificity of this process. Finally, we investigated the conformational change required for the enhancement of the RT maturation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank my mentor, Rieko Ishima, for her generosity. Without her time, patience, wisdom, and resourcefulness, none of this work would have been possible. I feel an indelible debt of gratitude for the role she has played in my professional and personal development, and will strive to give to others with the same generosity.

I would like to thank all of the members of the Ishima, Gronenborn and Ahn laboratories for their assistance and insightful discussions over the years. I would like to thank my thesis committee patience and guidance throughout this work. I also want to thank the MBSB faculty and staff. Much of their hard work goes unappreciated or unknown, but all of the students have directly benefitted from their contributions. I would like to thank my MBSB class: Abhishek, Ben, Patricia and Sean. I could not have asked for a better group of peers with whom to have shared this experience.

I would like to express my genuine gratitude for my family, whose support and belief in me will always be appreciated. Finally, I would like to thank my wife, Alexandra, for motivating our achievements; and for the love, time and support she has so generously provided.

1.0 INTRODUCTION

Human Immunodeficiency Virus Type-1 (HIV-1) is a causative agent of acquired immune deficiency syndrome (AIDS). An estimated 37 million people are presently infected with HIV and an additional 35 million individuals have died of HIV/AIDS related illness to date, making it one of the deadliest infectious diseases globally (1, 2).

Current treatment of HIV-1 infection involves drugs which act to suppress viral replication, primarily by inhibiting the activity of viral enzymes which are essential for replication. However, treatment does not eradicate infection thus necessitating chronic use of antiretroviral drugs by infected individuals (3).

This makes the development of antiretroviral drugs particularly challenging as HIV-1 is genetically diverse between and within infected individuals (4) largely due to the error prone nature of HIV-1 replication (5, 6), and the high rate of viral replication. The ability to rapidly genetically diversify in response to environmental factors allows HIV-1 to evade the host immune system and develop resistance to antiretroviral drugs in response to treatment (7); and viral evolution of resistance mutations is sufficiently rapid that treatments can only be effective by completely eliminating replication, and consequently blocking the evolution of resistance.

HIV-1 assembly, release and maturation have recently been studied to identify novel therapeutic targets (8). Although maturation of the Gag polyprotein has been characterized, a detailed understanding of viral enzyme maturation from the Gag-Pol polyprotein remains unclear presumably because their activity is indispensable to the viral life cycle. Therefore, a

detailed understanding of viral enzymes and the underlying mechanisms which regulate viral replication are of significant interest to the field and could provide a basis for the rational design of novel therapeutics targeting HIV-1 infection.

1.1 THE HIV-1 ENZYMES

HIV-1 belongs to *Retroviridea* family, characterized by a replication cycle which involves unique processes called reverse transcription and integration. In HIV-1, reverse transcription is the process by which the single stranded viral RNA (ssRNA) genome packaged by HIV is converted into a double stranded viral DNA genome (vDNA) by the viral enzyme reverse transcriptase. This vDNA is subsequently integrated into the chromosomal DNA of the host cell by another viral enzyme, integrase (8, 9). The integrated vDNA, called the provirus, is the template for the transcription of viral RNAs, which are translated into proteins that assemble into progeny viruses. These critical viral enzymes are encoded as a polyprotein, Gag-Pol, and cleaved by HIV-1 protease.

1.1.1 Protease (PR)

HIV-1 Protease processes all of the viral structural proteins and enzymes required for the formation of mature, infectious, virions. In its functional homodimeric state, catalytic aspartic acid residues from each monomer subunit meet at the dimer interface to form active site of the enzyme. Dynamic β-sheet rich sequences, called the flaps, from each monomer recognize an 8 amino acid stretch containing the substrate cleavage site and coordinate its positioning for catalysis at the active site (**Figure 1.1 A**). Because PR must cleave a number of sequentially diverse sites within viral

polyproteins, it has been hypothesized that the basis for cleavage site recognition involves a conserved structural motif rather than a conserved sequence (10).

The gene products of the HIV-1 genome are far beyond the scope of this work, and have been reviewed in great detail elsewhere (8, 9, 11). Suffice it to say that HIV-1 proteolysis is a highly ordered, multifactorial, process which requires PR to recognize and proteolytically catalyze a variety of viral substrates, including the Gag-Pol polyprotein. A simplified schematic representation of the Gag-Pol polyprotein is shown in Figure 1.2, and the sites of PR cleavage within Gag-Pol are depicted by red lines. The immature precursor of PR is known to have weak homodimer affinity(12, 13), with the monomer structure similar to that of the matured homodimer (14, 15). Of particular interest to this work is the dashed red line in Figure 1.2, corresponding to the site of PR-mediated proteolytic maturation of RT.

1.1.2 Integrase (IN)

Integrase is generated by PR-mediated cleavage of the C-terminal 228 residues of Gag-Pol. It is composed of three functionally distinct domains: the N-terminal domain which contains a zinc-finger motif and triggers IN multimerization(16), the catalytic core domain which is responsible for enzymatic activity (17-19), and the C-terminal domain which is involved in DNA binding and IN multimerization (20-22). These domains are connected to one another by flexible linker regions (**Figure 1.1 B**). IN is known to form higher ordered oligomeric complexes, and enzymatic activity is attributed to IN tetramers. Only recently, insights into structures of the functional form of integrase-DNA complexes, called intasomes, have been reported (19, 23-25).

Integration of viral genomic DNA into a host chromosome requires IN to catalyze two distinct reactions. First, IN cleaves a dinucleotide from each 3’ end of the vDNA (26), creating a

substrate with 3'-recessed ends for subsequent integration. In the second reaction, these recessed ends are covalently linked to 5' ends of target DNA in the host genome, resulting in an intermediate integration complex with unpaired 5' vDNA overhangs (27, 28). Host enzymes are then required to excise these 5' vDNA overhangs and join the vDNA 5' ends to host DNA, completing integration (29).

Non-enzymatic functions of IN related to viral packaging, reverse transcription, capsid uncoating, and nuclear import, are critical for viral replication cycle (30-38). Among them, IN has a crucial role in the viral maturation step by directly interacting with viral genomic RNA (39-42). Proper dimer formation or assembly of IN in the precursor form has been proposed to affect the viral maturation (43).

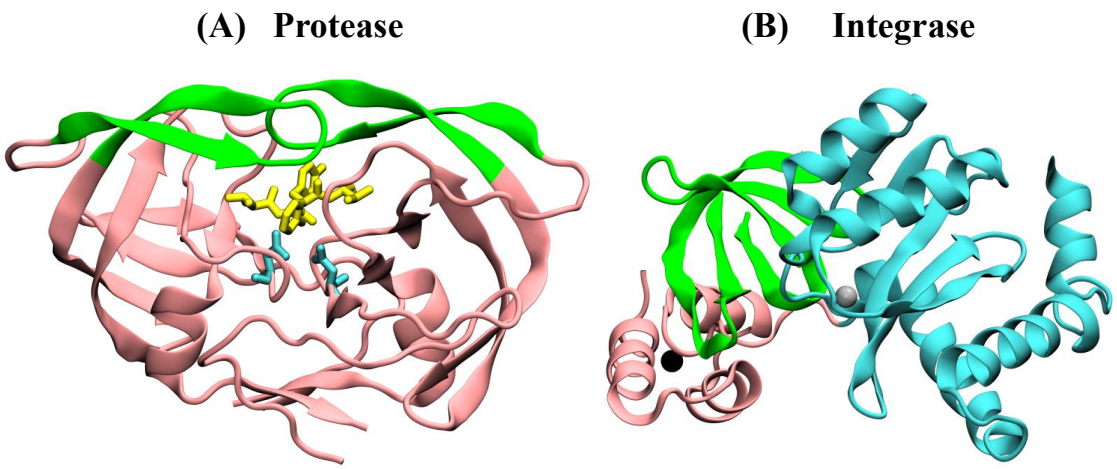

Figure 1.1: Structures of HIV-1 Protease and Integrase

(A) Cartoon representation of the PR homodimer (PDB code: 1KJG (10)) highlighting the flap regions that bind to the substrate (green), the active site residues (cyan), and a substrate peptide positioned for cleavage (yellow). (B) Cartoon representation of IN monomer (PDB code: 5U1C (25)) showing the N-terminal domain (pink), the catalytic core domain (cyan), and the C-terminal domain (green).

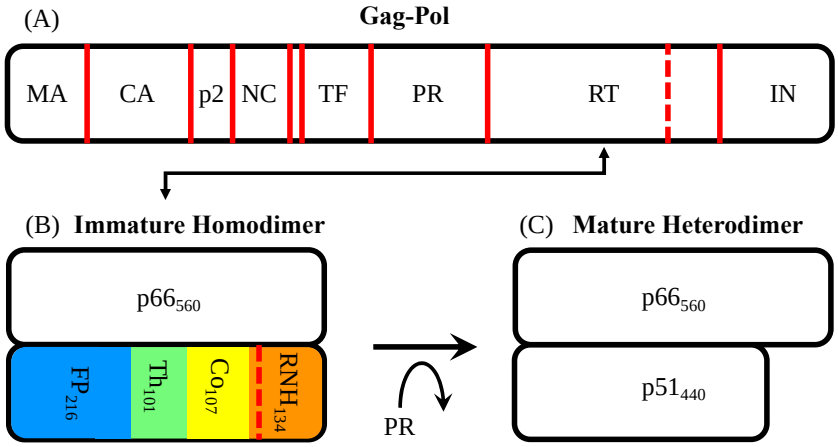

Figure 1.2: Schematic Models of HIV-1 Gag-Pol Polyprotein, Immature and Mature RT

Red lines indicate sites of cleavage by HIV-1 Protease during viral maturation. The Gag-Pol polyprotein (A) is comprised of matrix (MA), capsid (CA), spacer peptide (p2), nucleocapsid (NC), and transframe (TF) proteins as well as the C-terminal Pol region containing the HIV-1 enzymes: protease (PR), reverse transcriptase (RT) and integrase (IN). The dashed red line indicates the PR cleavage site for maturation of RT. In (B) and (C) the subunits (in white) and domains (in color) are labeled and their respective amino acid lengths indicated as subscripts.

1.1.3 Reverse Transcriptase (RT)

In order to accomplish reverse transcription, RT possesses two distinct enzymatic activities: DNA polymerase from a DNA or RNA template, and cleavage of RNA from a DNA/RNA duplex, i.e., ribonuclease H (RNH) activity. Initiation of reverse transcription also requires a cellular tRNA primer ($tRNA^{Lys3}$ in the case of HIV), which is selectively packaged during viral assembly (44). This primer tRNA is annealed to an 18-nucleotide sequence near the 5' end of the viral genome called the primer binding site, where cDNA synthesis can begin. The genomic RNA template is then copied up to its 5' end, forming an RNA/DNA hybrid. The RNase H degrades RNA within this duplex and the resulting single stranded DNA is annealed to 3' end of the viral RNA genome, in a process called first strand transfer (45-48). Following first strand transfer, RT resumes polymerization of cDNA and incomplete digestion of the genomic RNA template. A purine rich region called the polypurine tract, within the genomic RNA template, acts as the primer for synthesis of the second DNA strand from the newly formed ssDNA template. Synthesis of this second strand continues until RT recognizes a modified base in the tRNA primer, halting polymerization. RNH reaction then degrades the tRNA primer, and complementary sequences in both DNA strands anneal in a process called second strand transfer. Both DNA strands are then elongated until the entire genome is double stranded. This vDNA product serves as the substrate for subsequent integration.

Initially translated as part of Gag-Pol, RT is flanked by PR and IN at its N-terminus and C-terminus, respectively (**Figure 1.2 A**). Cleavage of PR and IN results in a 66 kDa protein, known as p66 (49). In mature viral particles, RT exists as a heterodimer consisting of p66 and a smaller 51 kDa protein, known as p51 (50). The aforementioned maturation of RT is the process by which p51 is formed, i.e., cleavage of part of the C-terminus of p66 by PR creates p51.

It has been demonstrated that the p66 subunit of the heterodimer is responsible for both polymerase and RNH activities (9), while p51 plays a role in correctly positioning the nucleic acid substrate for catalysis (51). Based upon its distinct activities, p66 is known to contain two spatially distinct domains, aptly named: polymerase and RNH. The polymerase domain has been further classified into Finger-Palm (residues 1-216), Thumb (residues 217-318) and Connection (residues 319-426) subdomains (**Figure 1.2 B**). The polymerase active site, located within the Finger-Palm region, contains three aspartic acid residues (D110, D185, D186) that coordinate two divalent metal ions which are required for catalysis (52, 53). The RNH domain (residues 427-560) comprises the C-terminus of p66, and cleavage by PR to form p51 occurs between residues F440 and Y441 within this domain (**Figure 1.2 C**) (9). The ribonuclease active site contains a highly conserved DEDD motif (resides D443, E478, D498, D549) that coordinate divalent metal ions required to hydrolyze the RNA substrate (54).

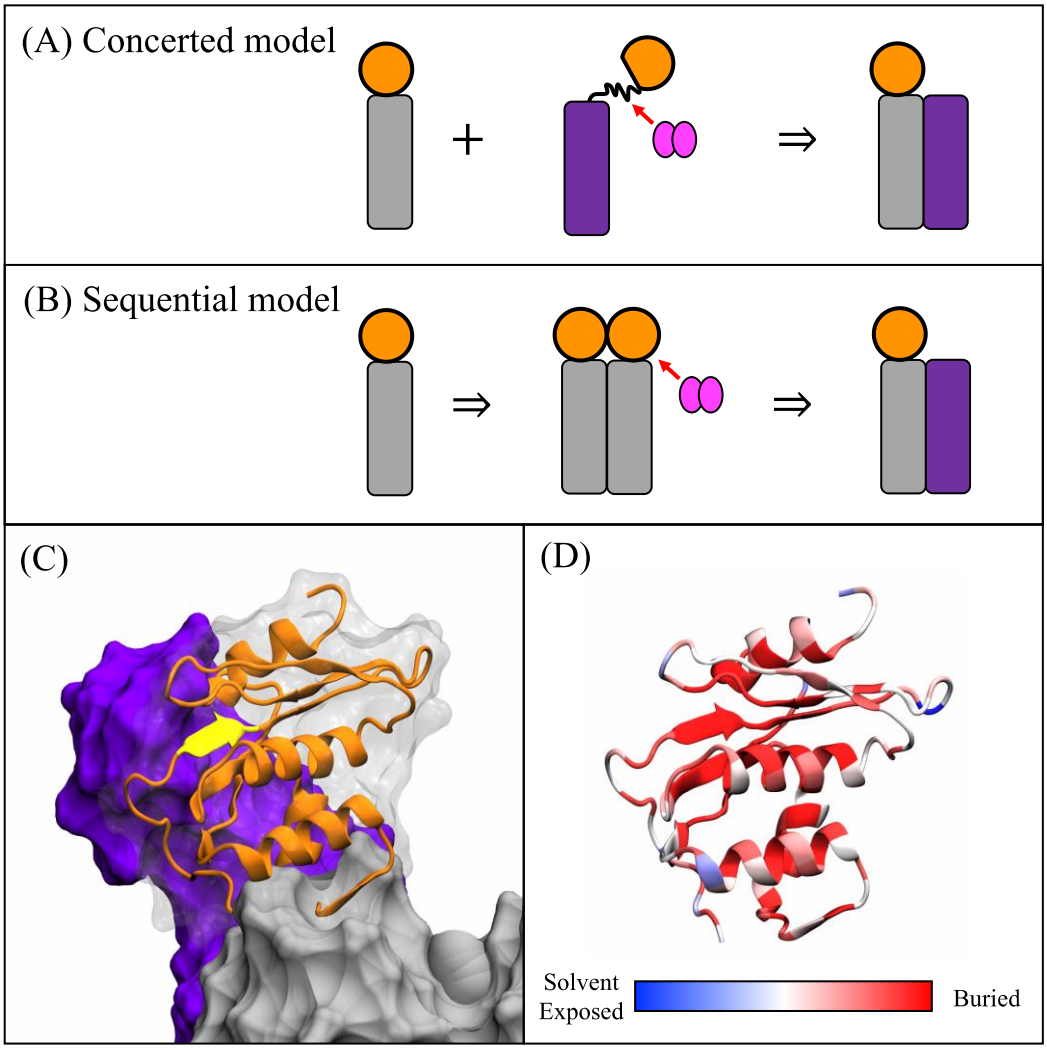

Figure 1.3: Models of RT maturation and structure of the p51-RNH cleavage site

Two models for RT maturation: (A) the concerted model and (B) sequential model, and (C) cartoon showing the p51-RNH processing site and (D) the solvent accessibility of the RNH domain that indicate complete burial of the processing site. In panels (A) and (B), polymerase and RNH domains in a p66 subunit are shown in grey and orange, respectively. The p51 subunit is shown in purple, and PR is shown in pink. In (C), the RNH domain is shown in orange, with the p51-RNH processing site indicated in yellow. In (D), solvent accessibility per residue was calculated with a probe radius of 1.4 Å, using the Solvent Accessible Surface Area module (55) in VMD(56).

1.1.4 Maturation of Reverse Transcriptase

Based on biochemical evidence and *in vitro* studies, two models of RT maturation have been proposed. In the "concerted" model, p66 and p51 are cleaved by PR from separate Gag-Pol polyproteins and subsequently assemble to form heterodimeric RT (**Figure 1.3 A**). The rationale for this model is based on comparison of the dimer dissociation constants- the p66/p51 heterodimer is approximate 10 times and 100 times lower than p66/p66 and p51/p51 homodimers, respectively (57-61). The "sequential" model proposes that RT maturation begins with processing of p66 from Gag-Pol and formation of an intermediate p66/p66 homodimer, followed by cleavage of one subunit to form the p66/p51 heterodimer (**Figure 1.3 B**). The model is supported from studies of PR-mediated processing of a pol polyprotein to form p66/p51 in an inducible bacterial expression vector, which demonstrate that synthesis of pol is followed by formation of p66, and p51 is observed only after significant accumulation of p66 (62). While this indicates that the p66 homodimer is the precursor for mature RT, no direct evidence for the sequential model has been reported. When mutations known disrupt p66/p51 dimerization were introduced in studies of pol polyprotein processing, pleiotropic effects including processing defects were observed (63, 64), suggesting that dimerization may be necessary but insufficient for RT maturation.

Over two hundred structures of the mature heterodimer a have been deposited to RCSB protein data bank (www.rcsb.org) (65) to date; many of which include RT in complex with inhibitory compounds, nucleic acid substrates, or various combinations of both. Interestingly, in all of these structures, the site which is recognized by PR and cleaved to form mature RT is buried within the hydrophobic core of the remaining RNH domain (**Figure 1.3 C and D**). Despite considerable efforts (66, 67), *no structure for the immature p66 homodimer has been reported*.

Therefore, the structural basis for p51-RNH processing site recognition and cleavage by PR remains unclear.

1.2 APPLICATION OF NMR TO STUDY REVERSE TRANSCRIPTASE

RT is a dynamic enzyme which must sample a range of conformations, orientations and translations along different substrates in order to carry out its diverse functions (68, 69). Nuclear Magnetic Resonance (NMR) spectroscopy has emerged as a powerful technique for characterizing protein structure and dynamics in solution. Resolution and sensitivity of resonances in solution-state NMR is mainly determined by the abundance of the observed nuclear spin, the rotational correlation time of the molecule that enhances transverse relaxation of the nuclear spin, and instrumental sensitivity. To resolve many NMR signals, two-dimensional (2D) and three-dimensional (3D) NMR spectra are recorded for protein studies. Typical 3D NMR experiments for signal assignments and structural determination are applicable for proteins with a molecular weight of up to 30 - 40 kDa (70). The isolated RNH domain of p66 is 15 kDa, and its structure has previously been determined both by x-ray crystallography and solution NMR.

NMR studies of large proteins in solution are technically challenging and of considerable interest in the NMR field. This is primarily due to slowing of molecular tumbling in solution as molecular mass increases, resulting in line-broadening of NMR signals (i.e., fast transverse relaxation rate). Although isotope labeling can reduce signal overlap (i.e., by observing ^{1}H, ^{13}C, and ^{15}N signals and silencing undesired hydrogens by ^{2}H labeling (71-75)), line-broadening remains an issue in practice. There are several general strategies to overcome this issue.

One approach is to utilize single-site labeling, in which only one residue is labeled with an NMR probe. This strategy has been used to study conformational changes of RT and the mutants upon inhibitor interaction (76, 77). In these studies, a ^{19}F NMR probe, trifluoromethyl-phenylalanine (tfmF), was introduced at a target site in either p66 or p51 polypeptide with the orthogonal amber tRNA/tRNA synthetase pair (78). Advantages of this single-site labeling approach include the exquisite sensitivity of chemical sifts to changes in the local chemical environment of the NMR probe, and the lack of spectral overlap due to the use of a single probe. In contrast, this approach gives only site-specific information which may not directly reflect global conformational changes within the protein.

A second strategy is selective amino acid labeling, which decreases the total number of NMR probes in the sample, thereby reducing the potential number of overlapping signals compared to uniform labeling. For example, labeling of ^{13}C methionine methyl group in a large protein significantly reduces the number of NMR signals, compared to uniformly ^{13}C labeled protein, and has been applied for the last two decades (79). Similarly, Ile $^{\delta 1}$-[$^{13}CH_3$] labeling significantly reduces the number of NMR signals (80-82). There are 6 methionine residues in each p66 and p51 subunit. There are 41 and 32 isoleucine residues in p66 and p51, respectively. Both of these labeling strategies have been previously applied to study HIV-1 RT (83-87). Signal assignments of selectively labelled proteins typically involve generating single amino acid substitutions, followed by comparison of the WT and mutant spectra. This can become fairly resource intensive because, for a protein containing 6 methionine residues, one needs at least 7 samples to identify and verify the methionine methyl signals unambiguously.

A third strategy to circumvent the issue of line-broadening is called the "divide-conquer approach" (88, 89). It involves partitioning the large protein into lower molecular weight

constituent parts, such as isolated domains, which are amenable to solution NMR studies. If there is a high degree of spectral similarity and signal assignments for the isolated domains have been obtained, they can be transferred to spectra obtained for the full-length protein for further analysis. It is worth noting that the divide-conquer approach is particularly useful when domain assignments are obtained. On the other hand, in a rigid multi-domain protein, there may be resonances that are not detected due to its fast relaxation rate. We have applied the divide-conquer approach to investigate ^{1}H-^{15}N NMR spectrum of the uniformly ^{15}N labeled per-deuterated p66 (75). Although this demonstrated that the thumb and RNH domains undergo significant domain motion, details of the domain packing of connection and finger-palm domains were not clarified. Also, since the p66 undergoes monomer-dimer equilibrium, our interpretation of the p66 NMR spectrum was performed by using dimer dissociation constant of p66 and assuming the on/off-rate of the homodimer.

As described above, there are pros-and-cons in these three strategies to investigate large proteins by NMR. Observing a small number of explicitly assigned NMR resonances simplifies spectra, but may omit information about the global features of large proteins. Observing NMR signals of uniformly labeled samples may reveal global conformational changes, at the expense of added spectral complexity. As mentioned previously, many proteins interconvert between unique conformations over a wide range of timescales, and these molecular motions are often essential for function (68, 69). Therefore, the parallel use of all of these labelling strategies can be particularly advantageous to analyze large proteins with solution NMR.

1.3 SCOPE OF THIS WORK

In this work, we seek to elucidate the structural basis for p51-RNH processing site recognition and PR-mediated maturation of RT by pursuing the following targeted aims.

1.3.1 Aim 1: Characterize the conformation of the p51-RNH processing site in the isolated RNH domain (Chapter 2)

Although the processing site is buried in all of the available x-ray crystallographic structures, there may be dynamics of the RNH domain in solution at a level that is undetectable for the wild-type RNH but detectable in mutants. A previous study found that mutations within the p51-RNH processing site result in partial or complete loss of virion associated p66 (90). A distal compensatory mutation within the RNH domain was found to rescue the aberrant processing of p66 (91). We will characterize isolated RNH domains bearing WT and mutant sequences using solution NMR, biophysical and computational methods. We hypothesize that these processing site mutations stabilize partially folded conformations of the RNH domain, consistent with the over-processing viral phenotype reported for the mutants. Concurrently, we hypothesize that the conformational equilibrium of mutants bearing the compensatory mutation resemble that of the WT, providing a mechanistic basis for the rescued p66 processing observed in vivo.

1.3.2 Aim 2: Investigate the structural integrity of the RNH domain in the immature p66 homodimer (Chapter 3)

It has been proposed that formation of asymmetric p66 homodimer in which one RNH domain is unfolded could explain the selective recognition and cleavage of one p51-RNH processing site by PR (87, 92, 93). However, unfolding of one RNH domain in p66 was not evident in our previous solution NMR studies of backbone amide (^{1}H-^{15}N) resonances in perdeuterated p66 (75). Alternative NMR studies using perdeuterated, Isoleucine side-chain ($^{\delta 1}$-[$^{13}CH_3$]) resonances in p66 reported unfolding of one RNH domain occurred in a time dependent manner (87, 92, 93). In order to clarify these incongruent results, we will monitor time dependent spectral features of p66 using backbone amide (^{1}H-^{15}N) resonances and Isoleucine side-chain ($^{\delta 1}$-[$^{13}CH_3$]) resonances. We hypothesize that the spectral features of p66 will be stable as a function of time, consistent with our previous observations. In order to more rigorously test for the possibility of differential unfolding of one RNH domain, we will evaluate assigned resonances from the thumb and RNH domains, and unassigned resonances using PCA. We hypothesize that domain-specific unfolding would be readily detectible using this numerical approach, and that experimental data will not provide evidence of time-dependent unfolding of one RNH domain.

1.3.3 Aim 3: Identify conformational changes in the immature p66 homodimer which facilitate maturation (Chapter 4, Chapter 5)

Our previous studies suggested that the p51- RNH cleavage sites in both subunits of p66/p66 were buried and thus poorly accessible to protease. Therefore, we postulated that some virion-associated factor may play a role in promoting the RT maturation process. HIV-1 virions are known to contain

substantial amounts of cellular tRNA [121-125], and RT binds these tRNAs with nanomolar affinity [126]. tRNAs are single stranded RNAs which form intramolecular base-pairs to fold into compact three-dimensional structures which are inherently asymmetric compared to canonical double stranded nucleic acids (94). We will investigate the effect of tRNA binding on the conformation of p66/p66 using solution NMR, and the effect of tRNA on the kinetics of PR-mediated RT maturation using an *in vitro* maturation assay. We hypothesize that the binding of tRNA to p66/p66 may introduce conformational asymmetry in the immature homodimer, facilitating efficient recognition and cleavage by PR to form the mature heterodimer.

2.0 STRUCTURAL INTEGRITY OF THE RIBONUCLEASE H DOMAIN IN HIV-1 REVERSE TRANSCRIPTASE

2.1 SYNOPSIS

The mature form of reverse transcriptase (RT) is a heterodimer comprising the intact 66-kDa subunit (p66) and a smaller 51-kDa subunit (p51) that is generated by removal of most of the RNase H (RNH) domain from a p66 subunit by proteolytic cleavage between residues 440 and 441. Viral infectivity is eliminated by mutations such as F440A and E438N in the proteolytic cleavage sequence, while normal processing and virus infectivity are restored by a compensatory mutation, T477A, that is located more than 10 Å away from the processing site. The molecular basis for this compensatory effect has remained unclear. We therefore investigated structural characteristics of RNH mutants using computational and experimental approaches. Our Nuclear Magnetic Resonance and Differential Scanning Fluorimetry results show that both F440A and E438N mutations disrupt RNH folding. Addition of the T477A mutation restores correct folding of the RNH domain despite the presence of the F440A or E438N mutations. Molecular dynamics simulations suggest that the T477A mutation affects the processing site by altering relative orientations of secondary structure elements. Predictions of sequence tolerance suggest that phenylalanine and tyrosine are structurally preferred at residues 440 and 441, respectively, which are the P1 and P1' substrate residues known to require bulky side chains for substrate specificity. Interestingly, our study demonstrates that the processing site residues, which are critical for protease substrate specificity and must be exposed to the solvent for efficient processing, also function to maintain proper RNH folding in the p66/p51 heterodimer.

The work presented in this chapter was reproduced with permission from: Slack RL, Spiriti J, Ahn J, Parniak MA, Zuckerman DM and Ishima, R. (2015), Structural integrity of the ribonuclease H domain in HIV-1 reverse transcriptase. Proteins, 83: 1526-1538. doi:10.1002/prot.24843 (95). This work was done in collaboration with the Daniel M. Zuckerman lab. Specifically, I mainly worked on Sections 2.3.1 - 2.3.4, and 2.3.7 in the Materials and Methods while Dr. Justin Spiriti in Zuckerman's group carried out Molecular Dynamics Simulations in Section 2.3.5 and 2.3.6. Both the Zuckerman and Ishima groups discussed about the data analysis and presentation. I was responsible for all of the figures, tables, and manuscript writing.

2.2 INTRODUCTION

The gene for HIV-1 reverse transcriptase (RT) encodes a 66 kDa protein, but mature HIV-1 reverse transcriptase is a heterodimeric protein comprised of 66 kDa and 51 kDa subunits. RT is initially translated as part of a much larger 160 kDa Gag-Pol polyprotein which is then processed by HIV-1 protease in a still poorly understood manner to yield the mature RT p66/p51 heterodimer. The smaller p51 subunit is generated by removal of most of the ribonuclease H (RNH) domain from a p66 subunit [Fig. 2.1(A)](9, 59, 96-101). Structures of both the RT heterodimer as well as the isolated RNH domain indicate that the p51-RNH cleavage site is located within the folded RNH domain, sequestered into the core b-sheet, and thus likely inaccessible to the protease [Fig. 2.1(B)] (54, 102-106).

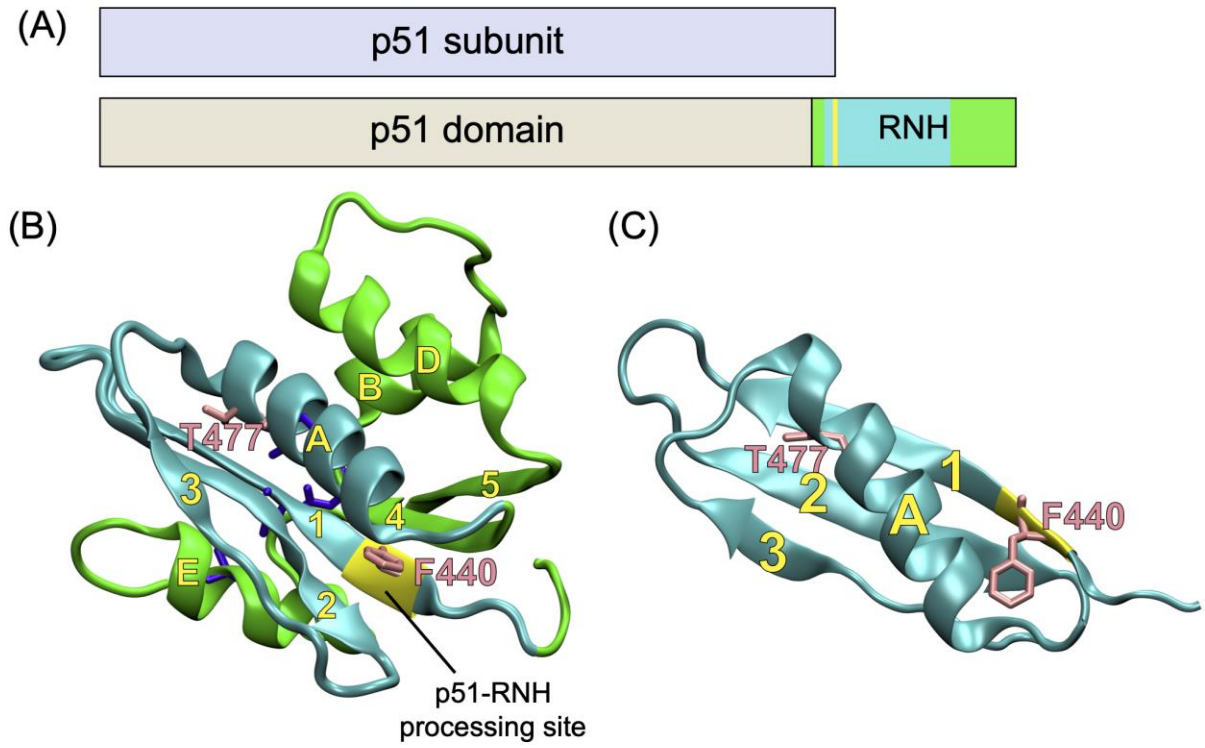

Figure 2.1 : Schematic Representation of RT and Structure of the RNH Domain

(A) Domain organization of RT, illustrating p66 (below) and p51 (above); the location of the protease processing site in p66 is indicated in yellow. Ribbon representation of the structures of (B) the RNH domain and (C) the part of the RNH domain, highlighting α-helix A and the β-sheet that includes β–strand 1, 2, and 3. In (B) and (C), the p51-RNH processing site is shown by yellow ribbon, and side chains of F440 and T477 are shown by pink color sticks.

A previous study introduced mutations within and surrounding the p51–RNH cleavage site expecting to find a relative accumulation of p66 subunits due to deceased processing efficiency of these cleavage-site mutants (90). Instead, these mutations resulted in dramatic phenotypic

alterations characterized by reduced viral infectivity, a significant reduction in virion RT p66 subunits, and a concomitant increase in the relative number of p51 subunits and fragments smaller than p51, suggesting unregulated degradation of RT during proteolytic maturation (90). Degradation of RT upon mutation of the processing site has been further demonstrated recently (107). These findings are not easily explained by changes in the side-chain volumes or charges because severe reduction of viral infectivity was observed for mutations both to hydrophobic residues, F440A and F440V, and to a hydrophilic residue, E438N. Further study revealed that an additional mutation (T477A) that arose during continued passage of the cleavage mutants rescued the p66 processing defects of these cleavage-site mutants and significantly restored viral infectivity (91). This revertant mutation site is located >10 A˚ away from the processing site: T477 is on a-helix A which interfaces with the core b-sheet containing the processing site [Fig. 2.1(C)]. The molecular mechanism for the compensatory impact of the T477A mutation was thus unclear.

In the present study the structural basis for the differences in proteolytic stability of the processing-site mutants and the revertants were studied using nuclear magnetic resonance (NMR), differential scanning fluorimetry (DSF), molecular dynamics (MD) simulations, and computational predictions of residues tolerated at the mutation sites. In particular, we characterized conformations of two isolated RNH domain mutants that contain either the F440A or E438N processing defect mutation (denoted RNH_{F440A} and RNH_{E438N}, respectively), and those with the additional T477A rescue mutation (denoted $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$, respectively). Comparison of the ^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) spectra of the mutant RNHs with wild-type (WT) shows that RNH_{F440A} and RNH_{E438N} are unfolded in solution but those with the compensatory T477A are not. The stark difference between the RNH_{F440A} and RNH_{E438N} and those with the compensatory T477A was also observed in the DSF study. Consistent with this

observation, 200 ns MD simulations exhibit wider structural variations for RNH_{F440A} and RNH_{E438N} compared with those of $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$, respectively. Structural ensembles obtained by the MD simulations for T477A mutants all exhibit a slight increase in the relative orientation of the a-helix A against the core b-sheet, compared with the WT. Predictions of sequence tolerance using RosettaBackrub (108, 109) suggest that phenylaniline and tyrosine are structurally preferred for residues 440 and 441, respectively. Our results suggest that specificity as a protease substrate is coupled to the structural requirement to maintain the RNH fold.

2.3 METHODS

2.3.1 Sample Preparation

Isolated WT RNH domain constructs were prepared by expressing the domain in Escherichia coli. In brief, the cDNA encoding RT residues 427 to 560 was inserted into the pE-SUMO vector (LifeSensors, Malvern, PA) with a six histidine tag (His_6-) at the N-terminus of the SUMO- fusion construct. As opposed to our previous WT RNH construct containing three additional amino acid residues S-E-L at the N-terminus of native RNH (110), the current construct encodes only native amino acid sequence of RNH after removal of SUMO. Mutations, E438N, E438N/T477A, F440A and F440A/T477A were introduced to the WT construct using QuickChange kits (Stratagene, La Jolla, CA) and verified by DNA sequencing. All the proteins were expressed in E. coli Rosetta 2 (DE3) cells. Cell cultures were grown at 37 °C to an OD of 1.0, induced with IPTG, and grown at 16 °C for an additional 18 h. Isotopic labeling was achieved by growing cultures in modified minimal media containing ^{15}N NH_4Cl as the sole nitrogen source using the published protocol

(111). Cells were harvested by centrifugation, resuspended in 50 mM Tris buffer, pH 7.5, and lysed by microfluidation. The His_6-SUMO-fusion RNHs were isolated from the cell lysate using a HisTrap HP columns (GE Healthcare, Piscataway, NJ) with a linear gradient of 0.02M to 0.5M Imidazole, followed by gel filtration on a Superdex75 26/60 column (GE Healthcare, Piscataway, NJ). The N-terminal His_6-SUMO fusion was then removed by digestion with histidine tagged ubiquitin-like-protein specific protease (ULP1). Finally, the RNH was separated from His-tagged proteins using a HisTrap HP column (GE Health- care, Piscataway, NJ), and polished with a Superdex75 26/60 column (GE Healthcare, Piscataway, NJ) equilibrated with a buffer containing 25 mM sodium phosphate, 100 mM NaCl, and 3 mM NaN_3, at pH 7.0. For the purification of RNH_{F440A} and RNH_{E438N}, the proteins were further purified using HiTrap Q HP column (GE Healthcare, Piscataway, NJ) to remove fragmented products. Purity of the proteins at the final step of purification were confirmed by running 20% acrylamide gels in both SDS denatured and non-denaturing (native) conditions (PhastGel system, GE Healthcare, Piscataway, NJ). Protein samples were stored at 80 °C.

2.3.2 NMR Experiments

All NMR experiments were performed at a protein concentration of ~200 μM in NMR buffer (25 mM sodium phosphate, 100 mM NaCl, pH 7.0) and supplemented with 10% D_2O. All ^{1}H-^{15}N HSQC spectra were recorded at 293 K on Bruker 600 AVANCE spectrometers, equipped with a 5-mm triple-resonance, z-axis gradient cryogenic probe. All data were processed with NMRPipe and analyzed with CCPNNMR analysis (112, 113).

2.3.3 Light Scattering Measurements

Size-exclusion multiangle light scattering (SEC-MALS) measurements were collected at room temperature using an analytical Superdex 75 HR 10/30 column (GE Healthcare, Piscataway, NJ) with in-line multiangle light scattering (HELEOS, Wyatt Technology), UV (Agilent 1100, Agilent Technology), and refractive index (Optilab rEX, Wyatt Technology) detectors. Protein samples (protein concentration: 100 μM, sample volume: 100 μL, in NMR Buffer) were injected into the column, pre-equilibrated with sterile-filtered and degassed NMR buffer. Molecular masses of the eluted proteins were analyzed using ASTRA software, version 5.3.4 (Wyatt Technologies).

2.3.4 Differential Scanning Fluorimetry (DSF)

Thermal stability of the WT, RNH_{F440A}, RNH_{E438N}, $RNH_{F440A/T477A}$, and $RNH_{E438N/T477A}$ were monitored by differential scanning fluorimetry. A fluorescence microplate reader (FluoDia T70, Photon Technology International, Edison, NJ) was used to measure binding of the hydrophobic dye SYPRO Orange to the unfolded fraction of the protein (Life Technologies, Carlsbad, CA) (114, 115). DSF samples were prepared at a protein concentration of 5 μM in a solution containing 25 mM sodium phosphate, 100 mM NaCl, and 5x SYPRO Orange at pH 5.0, 6.0, 7.0, 8.0, or 9.0. Sample volumes of 25 μL per well were loaded into 96-well PCR plates (Bio-Rad, Hercules, CA). Plates were heated from 25 to 75 °C in increments of 0.5 °C. Fluorescence intensity was measured using excitation/emission wavelengths of 465 and 590 nm, respectively. Fluorescence data were analyzed in MATLAB (The Mathworks Inc., Natick, MA), and melting temperatures of the proteins determined from the maximum of the first derivative of normalized fluorescence intensity signals, as described by *Niesen et al* (116). All assays were performed in triplicate.

2.3.5 Molecular Dynamics (MD) Simulations

All MD Simulations were performed by the Zuckerman Lab. Wild-type and mutant RNH systems were prepared using the CHARMM package (117, 118) and simulated with the NAMD package(119, 120). Initial structures for RNH_{WT}, RNH_{F440A}, RNH_{E438N}, RNH_{T477A}, $RNH_{E438N/T477A}$, and $RNH_{F440A/T477A}$ were generated using CHARMM based on residues 427 to 556 from the crystal structure of the WT RNH domain (PDB ID: 1DLO (121)), together with additional residues having the sequence SEF at the N-terminus and RKVL at the C-terminus in order to match the sequence of the constructs on which NMR studies had been performed. The coordinates of these residues, as well as any atoms not present in the original crystal structure (including hydrogen atoms), were generated using the internal coordinate facility in CHARMM; the additional residues were initially assumed to have extended configurations. The resulting structure was then energy-minimized using harmonic restraints together with the CHARMM36 force field (122, 123) and GBMV solvation model (124, 125). After minimization, each structure was surrounded with TIP3 water (126) in a rhombic dodecahedral box, allowing a 12 Å margin on all sides of the protein. A total of 21 Na^+ and 18 Cl^- ions were added to each system and placed using the SOLVATE program (127), bringing the salt concentration to approximately 100 mM while neutralizing the charge. The water and ions in each system were then energy-minimized while keeping the protein fixed.

The systems were then simulated using NAMD (119, 120). Each system was heated to 293 K over 1.5 ns with harmonic restraints of 1.0 kcal/(mol $Å^2$) on each nonhydrogen atom in the protein. The harmonic restraints were then gradually relaxed while equilibrating the system for an additional 1 ns.

Production simulations were carried out for 200 ns for each system with NAMD (119, 120) using the CHARMM36 force field and TIP3 water model. We employed a 2 fs time step, using

the SHAKE (128) and SETTLE (129) algorithms to constrain all bonds involving hydrogen in the protein and water respectively to their equilibrium values. Periodic boundary conditions were used; long range electrostatics was treated with the particle mesh Ewald method, and a switching function between 8 and 12 Å was applied to the van der Waals interactions. Constant temperature was maintained using Langevin dynamics with a damping coefficient of 5 ps^{-1}, and constant pressure was maintained using a Langevin piston with an oscillation time of 100 fs and a decay time constant of 50 fs. Frames were recorded every 1 ps.

2.3.6 MD Simulation Analysis

The trajectories were analyzed using CHARMM (117, 118) in order to better understand the effect of the mutations on the structure and dynamics of the RNase H domain. The backbone RMSD relative to the starting structure (involving N, C^{α}, and C atoms of residues 427–556) was calculated for each frame in the trajectory. The overall sampling quality of the simulation was evaluated by comparing the distribution of this backbone RMSD in the first half of the trajectory to that in the second.

The structure of the protein near the mutated residues was analyzed in greater detail. In order to determine the effect of mutations on the packing of nearby residues, the number of atoms within 4 Å around each residue was counted in each frame. The effect of the mutations on the local hydrogen bonding network was investigated by calculating the fraction of time individual bonds were present. In these calculations, a hydrogen bond was defined to be present if the hydrogen-acceptor distance was less than 2.4 Å and the donor-hydrogen-acceptor angle was greater than 150°.

In order to characterize the relationship of α-helix A to the first three β-sheet strands (1, 2, 3) for each frame, the helical axis of helix A was determined by applying the algorithm of Aqvist (130) to the α-carbons of residues 474 to 488. This helical axis was then represented in a coordinate system defined by the principal axes of the moment of inertia of the backbone atoms of residues 439 to 446, 453 to 459, and 467 to 469. The helical tilt angle θ was then defined to be the angle between the helical axis and the plane formed by the two principal axes with the smaller moments. The angle ϕ was defined to be the polar angle of the projection of the helical axis on this plane, relative to the principal axis with the smallest moment. For each trajectory, the average and standard deviation of all observables except the hydrogen-bonding fractions was calculated.

2.3.7 Prediction of Sequence Tolerance to Maintain Structure at the Mutation Sites

The structurally preferred amino acid type, at residues 438, 440, 441, and 447, was examined using RosettaBackrub (108, 109, 131, 132). This software uses flexible backbone modeling and a sequence tolerance protocol to predict amino acid substitutions which preserve near-native folding stability of the protein. We input the RNH domain structure and specified an ensemble size of 100 structures to score structural stability of each mutation. This procedure was performed using the RNH_{WT} and the RNH_{T477A} coordinates, which were the same as those used for the MD simulation. To check sensitivity to the choice of structure, the calculations were repeated using the RNH_{WT} and RNH_{T477A} coordinates obtained from the MD simulations at 100 ns time point.

2.4 RESULTS

2.4.1 NMR Experiments of the Processing Site and Revertant Mutants

To clarify the effect of the processing-site mutants and the revertant on the RNH structure, ^{1}H-^{15}N HSQC spectra of RNH_{F440A}, RNH_{E438N}, $RNH_{F440A/T477A}$, and $RNH_{E438N/T477A}$, as well as the WT, were recorded (Fig. 2.2). WT spectrum exhibits folded RNH signals similar to those published previously (106, 133-135). In contrast to the WT spectrum, most of the resonances in the ^{1}H-^{15}N HSQC spectra of RNH_{F440A} and RNH_{E438N} were observed at the random coil region, that is, at a narrow ^{1}H chemical shift ranges (8.0–8.5 ppm), indicating that these proteins are disordered, most likely unfolded, in solution [red spectra in Fig. 2.2(B,C)]. The observed disordered spectral feature of RNH_{F440A} and RNH_{E438N} is not due to unfolding during the expression and purification of the mutants because refolding experiments at various conditions, such as in low salt condition or by using denaturants, did not change the results at pH 7. Instead, because changing the sample condition from pH 7 to 8 increased the folded signals in the ^{1}H-^{15}N HSQC spectra of RNH_{F440A} and RNH_{E438N} (Appendix A Fig. A1), we believe that charge effects, such as salt-bridges or hydrogen bonding, may contribute to the folding of these mutants.

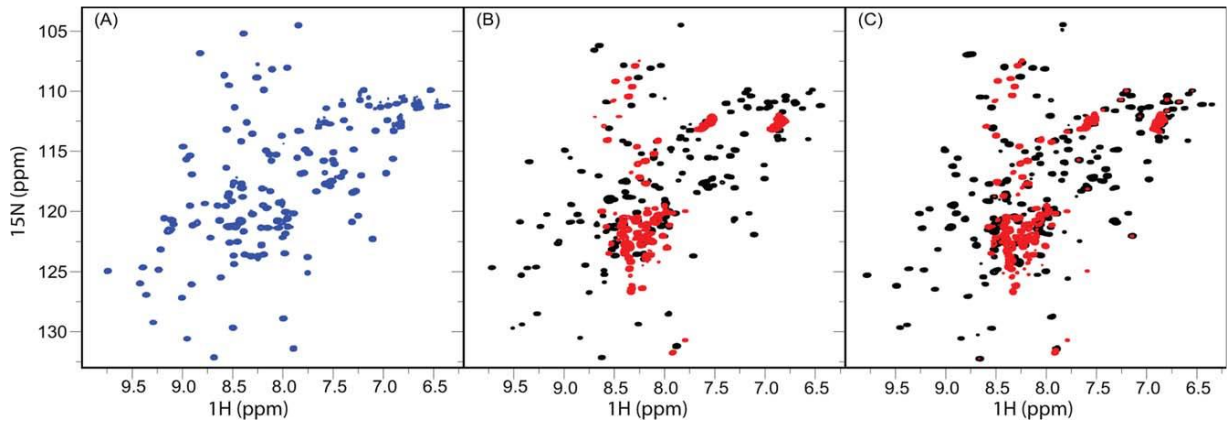

Figure 2.2 : ^{1}H-^{15}N HSQC spectra of the RNH Wild Type (WT) and mutants

(A) The RNH WT spectrum exhibits well disperse and sharp cross-peaks, characteristic of a well folded protein in solution (blue). (B) Superimposition of the RNH_{F440A} mutant (red) and the $RNH_{F440A/T477A}$ mutant (black). (C) Superimposition of the RNH_{E438N} mutant (red) and the $RNH_{E438N/T477A}$ mutant (black). All the spectra were obtained on a Bruker AVANCE 600 Spectrometer at 20 °C.

^{1}H-^{15}N HSQC spectra with the revertant, that is, $RNHF_{440A/T477A}$ and $RNH_{E438N/T477A}$, showed dispersed NMR signals that are similar to those of the WT [black spectra in Fig. 2.2(B,C)]. For example, signals with ^{1}H chemical shifts above 9 ppm are not observed in the spectra of the RNH_{F440A} and RNH_{E438N} [red color spectra in Fig. 2.2(B,C)], but are detected in the spectra of $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ similar to that of the WT [black color spectra in Fig. 2.2(B,C)]. These results indicate that restored folding was the major factor in reactivating the infectivity. However, comparison of the NMR spectra between the $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ [black color spectra in Fig. 2.2(B,C)] shows that signals from the unfolded fraction still remains in the $RNH_{F440A/T477A}$ spectrum whereas unfolding signals are less significant in the $RNH_{E438N/T477A}$ spectrum. Since viral infectivity of $RNH_{F440A/T477A}$ is approximately 20% greater than

$RNH_{E438N/T477A}$ (90, 91), the difference in the infectivity is not determined only by the folding ratio but also by a structural factor that affects the RNH activity of the mutants.

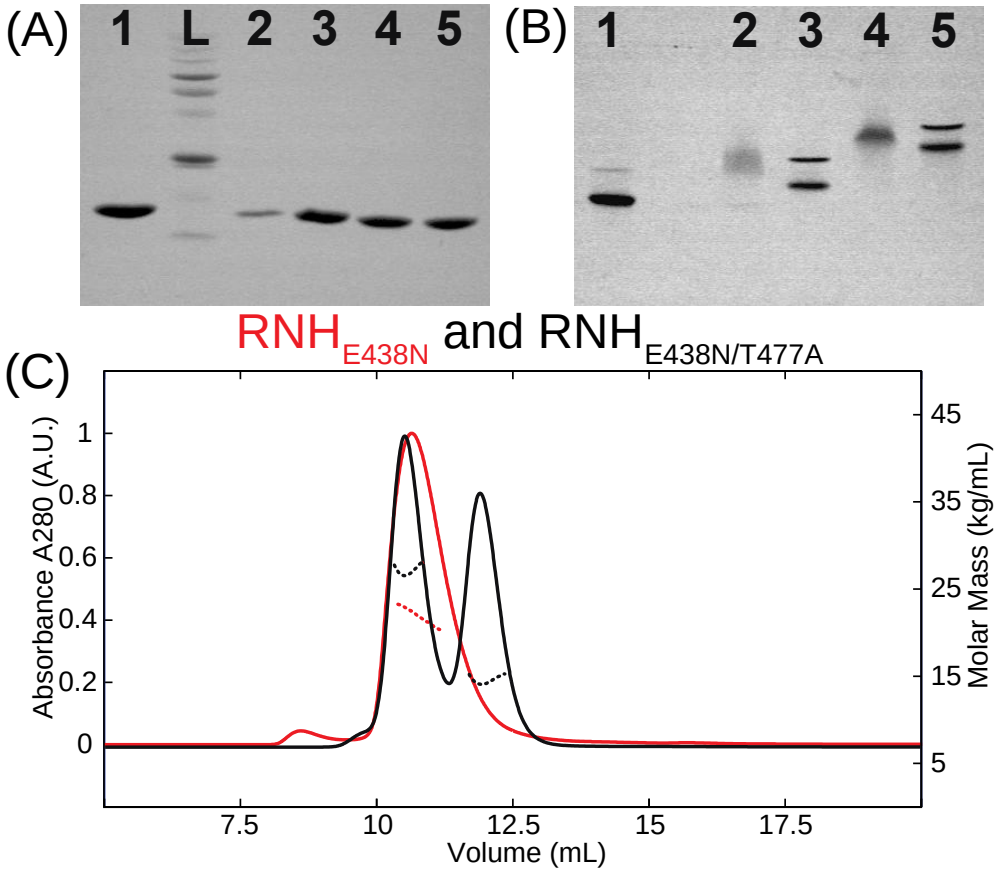

Figure 2.3 : Characterization of Mutant and Wild Type RNH Domains by Native PAGE and SEC-MALS

Gel electrophoresis profiles in (A) denatured and (B) native conditions, and (C) SEC-MALS UV profiles (solid) and molecular mass profiles (dashed) of RNHE438N (red) and RNHE438N/T477A (black). In (A), L indicates a molecular weight size marker. In (A) and (B), Lane 1: WT RNH, Lane 2: RNHF440A, Lane 3: RNHF440A/T477A, Lane 4: RNHE438N, Lane 5: RNHE438N/T477A. In (C) the average molecular mass of the eluted RNHE438N peak was determined to be 22.1 kDa, and those of the eluted RNHE438N/T477A peaks at 10.5 and 12.1 ml were determined to be 14.88 kDa and 27.25 kDa, respectively.

The profiles of SDS gel electrophoresis demonstrate single bands for all the RNH mutants, indicative of a singular molecular weight species [Fig. 2.3(A)]. This result also confirms that the observed random coil chemical shifts of RNH_{F440A} and RNH_{E438N} are not due to proteolytic fragmentation by E. coli enzymes. In contrast to the SDS gel profiles, native gel electrophoresis profiles of $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ show migration patterns with two distinct bands which stem from the monomer and dimer species, demonstrating that the double mutants have dimerization characteristics predominantly similar to that of the WT [Fig. 2.3(B)] (110). Although the band positions in the native gel electrophoresis for $RNH_{E438N/T477A}$ significantly differ from those of the WT and $RNH_{F440A/T477A}$ [Fig. 2.3(B)], monomer and dimer molecular masses were confirmed using SEC- MALS [Fig. 2.3(C)]. In both single mutants, RNH_{F440A} and RNH_{E438N}, a diffuse band was observed, probably due to the surface charge variations in the unfolded protein [Fig. 2.3(B)]. Overall, these gel profiles of the WT and the mutant RNHs as well as SEC-MALS elution profiles support the NMR observations: F440A and E438N mutations reduce the stability of the RNH folding at pH 7, while T477A rescues the folding.

2.4.2 Differential Scanning Fluorimetry of the Processing Site Mutants and the Revertant

To expand on our NMR observations, we used DSF to determine the thermal stability of RNH_{F440A}, RNH_{E438N}, $RNH_{F440A/T477A}$, $RNH_{E438N/T477A}$, and RNH_{WT} over a range of different buffer pH values (136, 137) (Table 2.1). The WT exhibited a maximum melting temperature (T_m) of 55.1 ± 2.4°C, at neutral pH, and slightly lower T_m values in alkaline buffer conditions. In all of the pH conditions used for this experiment, RNH_{F440A} and RNH_{E438N} showed high fluorescence intensity throughout the examined temperatures, and the T_m values could not be determined (116) (see Appendix A Fig. A.2). Table 2.1 shows that T_m values for $RNH_{F440A/T477A}$ were lower than that of the WT at all pH

conditions tested. T_m values for $RNH_{E438N/T477A}$ were similar to those of $RNH_{F440A/T477A}$ at neutral or alkaline pH conditions, but could not be obtained at pH 5 or 6, showing a similar profile to those of RNH_{F440A} and RNH_{E438N}. Overall, although the NMR data for $RNH_{F440A/T477A}$ and the $RNH_{E438N/T477A}$ showed a profile similar to that of WT, the thermal stability is low, which may explain the residual unfolded signals in their NMR spectra.

Table 2.1: Melting Temperature of the WT RNH and the Mutants, at Different pHs, Determined by Differential Scanning Fluorimetry*

Melting Temperature, T_m (°C)					
pH	5.0	6.0	7.0	8.0	9.0
WT	53.2 ± 0.4	54.3 ± 2.3	55.1 ± 2.4	53.8 ± 1.6	53.8 ± 0.8
F440A	NA	NA	NA	NA	NA
F440A/T477A	47.9 ± 2.4	42.7 ± 2.2	44.6 ± 1.5	43.3 ± 0.8	46.8 ± 2.9
E438N	NA	NA	NA	NA	NA
E438N/T477A	NA	NA	43.8± 0.5	43.8 ± 0.5	47.5 ± 2.1

*NA indicates that the T_m could not be determined.

2.4.3 Conformational Ensembles Obtained by MD Simulations

As seen above, $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ contain a small unfolded population and a significant dimer population, respectively (Figs. 2.2 and 2.3). Thus, it is impossible to determine unambiguous high-resolution structures of these mutants by NMR. To gain atomic level information that could help explain the experimentally observed characteristics of these RNH mutants, MD simulations for 200 ns were performed for RNH_{F440A}, RNH_{E438N}, $RNH_{F440A/T477A}$, $RNH_{E438N/T477A}$, and RNH_{T477A} as well as the WT. MD simulation data has systematic inaccuracies in force fields and limited trajectory times compared with biological timescales (138-141). In the present study, for example, the trajectories are not long enough for the unstable mutant RNH_{F440A} to unfold, as described below. On the other hand, MD simulation is advantageous for relative comparison of protein dynamics *on the timescale of the trajectories* (200 ns, in the present study), which in turn provides clues to the physical reasons for the cross-talk between the processing site and the T477 site.

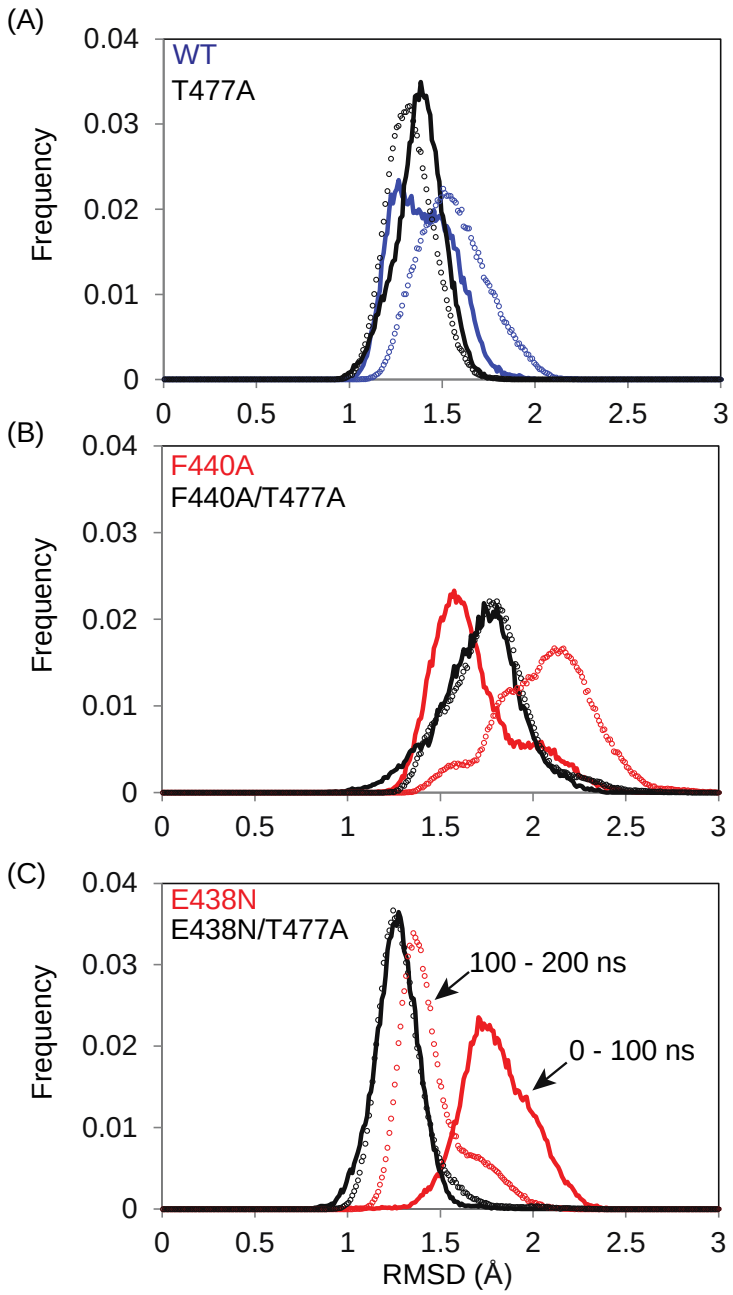

Figure 2.4 : Global assessment of sampling and structural diversity via MD simulations

Histograms of the backbone RMSD (Å) of structures obtained in the 1-100 ns (solid line) and of 100-200 ns (dashed line) MD simulations for (A) WT, (B) RNH_{F440A}, and (C) RNH_{E438N} (blue line, for WT; red lines, F440A and E438N mutants without T477A; black lines, with T477 mutation).

To obtain an overview of the conformational fluctuations observed during the simulation, the root-mean-square deviation (RMSD) of the backbone was calculated against the wild-type crystal structure for all frames. The distributions of this RMSD in the first and second half of each trajectory were compared in order to assess the quality of sampling in each simulation (Fig. 2.4, solid lines vs. symbols). Ideally, if each trajectory were of sufficient length, these distributions would be the same to within statistical error. Indeed, the RMSD was distributed from 1 to 2 Å in the WT with a 68% overlap between the 0 to 100 ns and 100 to 200 ns periods. Interestingly, the RMSD distribution was narrowed, with about 80% overlap, in RNH_{T477A}, compared with the WT [Fig. 2.4(A)].

A similar but more pronounced effect by the T477A mutation was observed in the F440A and E438N mutants: the RMSD distributions of the two time periods differ in RNH_{F440A} and RNH_{E438N} whereas those of the two time periods are almost identical to each other in their T477A mutants [Fig. 2.4(B,C)], implying a high degree of stability. Since RNH_{F440A} and RNH_{E438N} are mostly unfolded in solution (Fig. 2.2), the 200 ns simulations likely do not reflect the entire conformational ensembles of the mutants, but instead reflect conformational fluctuation around the initial folded RNH structure. Larger and longer-lived fluctuations are observed for RNH_{F440A} and RNH_{E438N} compared with the other trajectories, consistent with their structures being experimentally less stable. In addition, the RMSD comparisons demonstrate that the structures of the RNHs with the revertant mutation, T477A, are more structurally ordered on the timescale of the simulation, compared with those without the T477A mutation. Because 200 ns of MD simulation is not sufficient to characterize a protein exhibiting significant fluctuations (for example, unfolding), we also performed Weighted Ensemble (WE) simulations, which validate the observation that RNH_{F440A} is more unstable than $RNH_{F440A/F477A}$ or WT.

2.4.4 Conformational Characteristics Observed by MD Simulations

In order to investigate the changes in the conformational ensemble caused by the mutations, characteristics of the local structures around the processing site and residue 477 were compared in the various trajectories. As expected from the volume changes of the side chains, the average number of protein atoms which surround the mutation sites decreased upon mutations from 39.5 ± 2.6 for the WT compared with 31.8 ± 3.0 for the E438N mutation; from 48.3 ± 3.2 for the WT compared with 29.1 ± 1.9 for the F440A mutation; and 40.5 ± 2.3 for the WT compared with 33.3 ± 2.3 for the T477A mutation (Table 2.2).

Interestingly, the reduction in the number of residues surrounding the processing sites remains even in $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$, which exhibit stable conformations in the 200 ns simulations. These simulation results suggest that T477A mutation does not counteract the loss of side chain packing at the processing site residues due to the F440A and E438N mutations.

Table 2.2: Average and Standard Deviation of Number of Nonhydrogen Atoms in Protein Within 4 Å Around Residues 438, 440, and 477

Simulation	Res. 438	Res. 440	Res. 477
WT	39.5 ± 2.6	48.3 ± 3.2	40.5 ± 2.3
T477A	40.2 ± 2.5	48.3 ± 3.1	33.3 ± 2.3
F440A	42.0 ± 3.4	29.1 ± 1.9	40.7 ± 2.4
F440A/T477A	40.5 ± 2.6	28.7 ± 2.0	34.4 ± 2.5
E438N	31.8 ± 3.0	43.2 ± 3.2	40.0 ± 2.5
E438N/T477A	31.6 ± 2.7	42.9 ± 2.9	33.7 ± 2.3

Table 2.3: Average Number of Selected Hydrogen Bonds Around the Processing Site, Observed in each Simulation*

Simulation	E438 O^{ε}- R463 H^{ε}, H^{η}	D488 O- R463 H^{ε}, H^{η}	D460 O^{δ} - R461 H^{ε}, H^{η}	D460 O^{δ}- T439 HN	res. 438 O^{δ}/O^{ε}- T459 H^{γ}
WT	1.92	0.40	0.67	0.66	0.95
T477A	1.91	0.43	0.79	0.86	0.95
F440A	1.75	0.22	1.37	0.92	0.95
F440A/T477A	1.85	0.63	0.46	0.41	0.94
E438N	0.00	0.00	1.42	0.88	0.00
E438N/T477A	0.00	0.00	1.52	0.86	0.00

*Hydrogen bond criteria are hydrogen-acceptor distance < 2.4 Å and donor-hydrogen-acceptor angle > 150°.

More quantitative conformational changes upon mutations were investigated by monitoring the stability of the hydrogen bond network around residue E438 [Fig. 2.5(A)]. When residue 438 is a glutamate, i.e., in the WT, the O^{ϵ} atoms from the carboxyl group form hydrogen bonds with the H^{ϵ} and H^{η} atoms from the guanidinum group of R463. At the same time, the carboxyl group also forms a hydrogen bond with H^{γ} from the hydroxyl group of T459 (Table 2.3). When residue 438 is mutated to asparagine, it becomes neutral, and this hydrogen bond net- work is eliminated; instead, R463 faces outward toward the solvent. Because of this change of the R463 side chain orientation, the stability of the hydrogen bond between the D488 backbone carbonyl and the R463 side chain is also reduced upon the E438N mutation. In short, the E438N mutation induces the loss of the local hydrogen bond network.

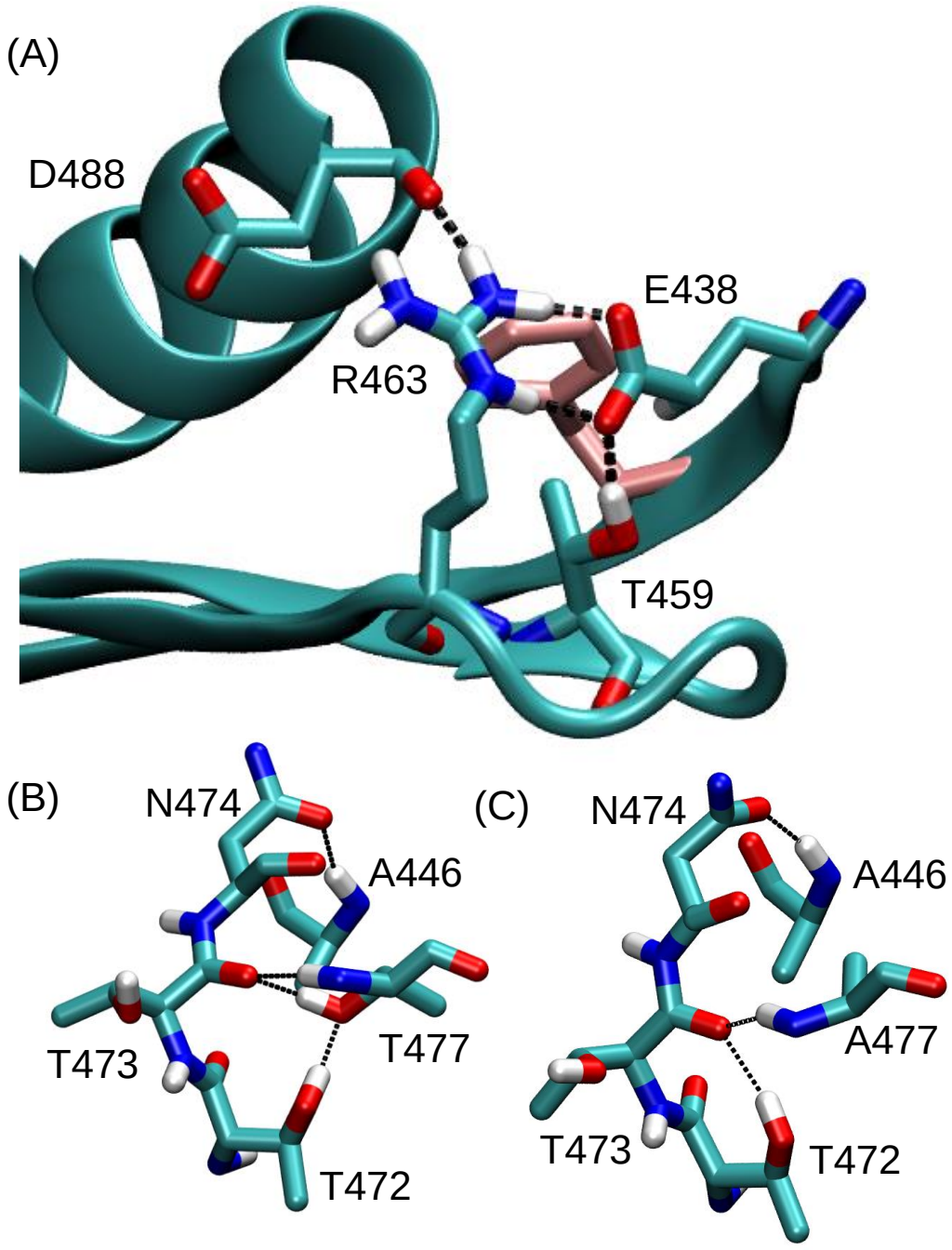

Figure 2.5 : Hydrogen bonding networks observed in MD simulations

(A) Side chain orientations in the WT RNH around (A) residue 440 (pink), and residue 477 in (B) the WT and (C) the RNH_{T477A}. In (A), dashed lines indicate salt bridge network that involves E438. In (B) and (C), black lines indicate hydrogen bonds that are frequently observed in the WT RNH and the RNH_{T477A}, respectively (see Table 2.3).

Similar to the E438N mutation, significant reduction of the average number of surrounding atoms (from 48.3 ± 3.2 for the WT compared with 29.1 ± 1.9 for the RNH_{F440A}) occurs upon the F440A mutation. The F440A mutation slightly reduces the occupancy of the hydrogen bond from the D488 backbone and increases the hydrogen bond occupancy from the D460 and R461 sidechains. Although the F440A mutation did not exhibit such drastic changes in the hydrogen bond network, the mutation results in larger backbone RMSD deviation compared with those of the WT and the E438N mutation (Fig. 2.4). Overall, reduction of the Phe to Ala side chain probably affects packing of the protein core, including hydrogen bonding networks and hydrophobic interactions.

The region around residue 477 is well folded with different hydrogen bond interactions occurring in RNH with and without the T477A mutation [Fig. 2.5(B,C)]. When residue 477 is a threonine, the hydroxyl side chain of T477 forms a hydrogen bond with the T473 backbone carbonyl oxygen or with T472 hydroxyl group, depending on the protonation of the hydroxyl side chain (Table 2.4). These hydrogen bonds, along with the hydrogen bond from N474 to A446, help to maintain the position of the N-terminal end of helix 1 relative to the first three β-strands. By contrast, the mutation of residue 477 to alanine does not allow formation of such inter helix-loop hydrogen bonds. Instead, the loop region is stabilized by forming two different hydrogen bonds, one between N477 side chain OG1 (in the loop) and A466 backbone NH (in the β-1 strand) and another between T472 hydroxyl side chain (in the loop) and T473 backbone carbonyl [Table 2.4 and Fig. 2.5(C)].

Table 2.4: Fraction of Selected Hydrogen Bonds Around Residue 477, Observed in Simulation*

Simulation	E438 O^{ϵ}- R463 H^{ϵ}, H^{η}	D488 O- R463 H^{ϵ}, H^{η}	D460 O^{δ} - R461 H^{ϵ}, H^{η}	D460 O^{δ}- T439 HN	res. 438 O^{δ}/O^{ϵ}- T459 H^{γ}
WT	1.92	0.40	0.67	0.66	0.95
T477A	1.91	0.43	0.79	0.86	0.95
F440A	1.75	0.22	1.37	0.92	0.95
F440A/T477A	1.85	0.63	0.46	0.41	0.94
E438N	0.00	0.00	1.42	0.88	0.00
E438N/T477A	0.00	0.00	1.52	0.86	0.00

*Hydrogen bond criteria are hydrogen-acceptor distance < 2.4 Å and donor-hydrogen-acceptor angle > 150°.

Table 2.5: Axis of Helix A (C^{α} Atoms) Relative to Inertia Axes of Backbone Atoms of β Strands 1-3*

Simulation	θ (°)	ϕ (°)
WT	7.7 ± 1.9	41.5 ± 1.8
T477A	10.5 ± 1.6	41.1 ± 1.8
F440A	6.0 ± 2.3	40.5 ± 2.2
F440A/T477A	10.2 ± 1.7	40.7 ± 1.6
E438N	7.2 ± 2.0	40.0 ± 1.8
E438N/T477A	10.0 ± 1.6	40.0 ± 1.9

*The indicated error bars represent the standard deviation over each trajectory. Since there is a slow conformational change, we were unable to obtain a reliable estimate of the standard error from the block averaging approach. [49]

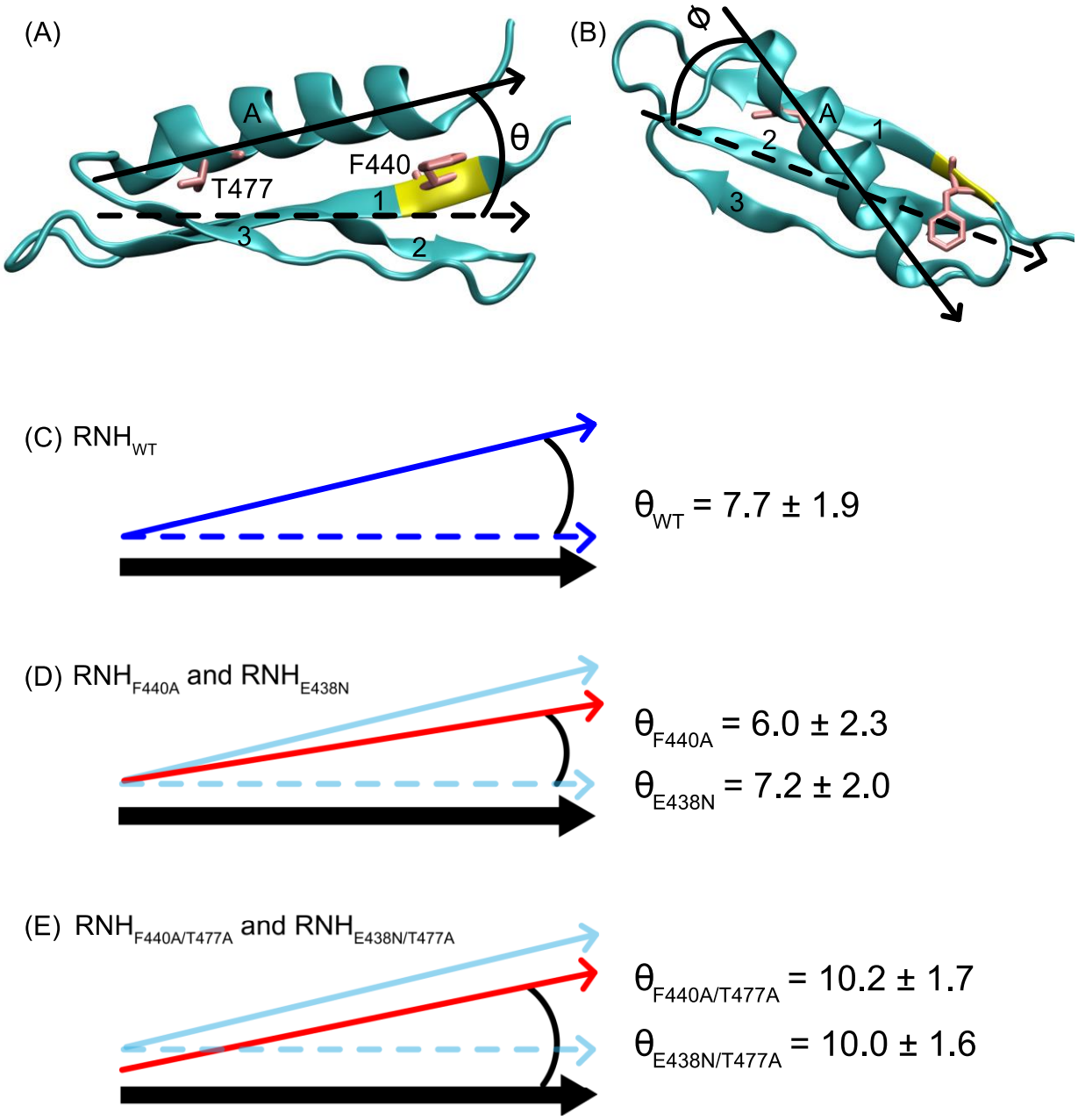

Figure 2.6 : Structural geometry for investigating crosstalk between sites 440 and 477 via MD

The axis of the α-helix A (Cα atoms) relative to inertia axes of backbone atoms of β strands 1-3 listed in Table 2.5 is defined by (A) θ and (B) ϕ. The θ value resulting from simulation of (C) RNH_{WT} was 7.7 ± 1.9°. (D) The θ decreased in the RNH_{F440A} and RNH_{E438N} mutants (red arrow). (E) $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ result in an increased θ value. In (C) – (E), schematically, a black thick arrow represents a position of the β sheet, a blue (or light blue) arrow represents a position of the α-helix A, and a blue (or light blue) dashed-arrow represents a relative position of the β sheet.

2.4.5 Crosstalk Between the Processing Site and Residue 477

The MD trajectories suggest a hypothesis for the mechanism of crosstalk between the region around the residue 477 and the somewhat distant processing site. The simulation results show a tendency for the helical tilt angle, θ, of RNH_{T477A} to exceed that of WT, in contrast to negligible observed changes in ϕ. The reduced size of A477 compared with T477 leads to the movement of the N-terminal end of the helix toward the beta strands, thereby increasing θ by ~3° (Table 2.5, Fig. 2.6). The θ angles of $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ are similar to that of RNH_{T477A}, reflecting similar movement at the N-terminal end of the helix A (Table 2.5, and Fig. 2.6). The number of residues surrounding the processing site near the C-terminus of helix A of $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ is smaller than that of WT (Table 2.2). Thus, on the whole, this simple "levering" picture appears to explain the crosstalk between the distant sites.

2.4.6 Sequence Tolerance of the Processing Site Mutations

Preference of amino acid sequences on the RNH structure was systematically evaluated for the processing-site residues using the Rosettabackrub software by randomly sampling of any of the 20 amino acids except Cys and repacking of residues within 4 Å of the newly designed residue for energy minimization (108, 109, 131, 132). First, amino acid preferences at residues 438, 440, and 441 were calculated on the structural platforms of the WT and RNH_{T477A}. For the residue 440, the calculated frequency of the preferred amino acid demonstrates that Phe and Tyr residues are strongly preferred to maintain the WT-like structure, in both WT and RNH_{T477A} [Fig. 2.7(A)]. Similarly, Glu or Asp residue is preferred for the residue 438 [Fig. 2.7(B)], indicating that these processing site residues are favored to maintain the RNH folding.

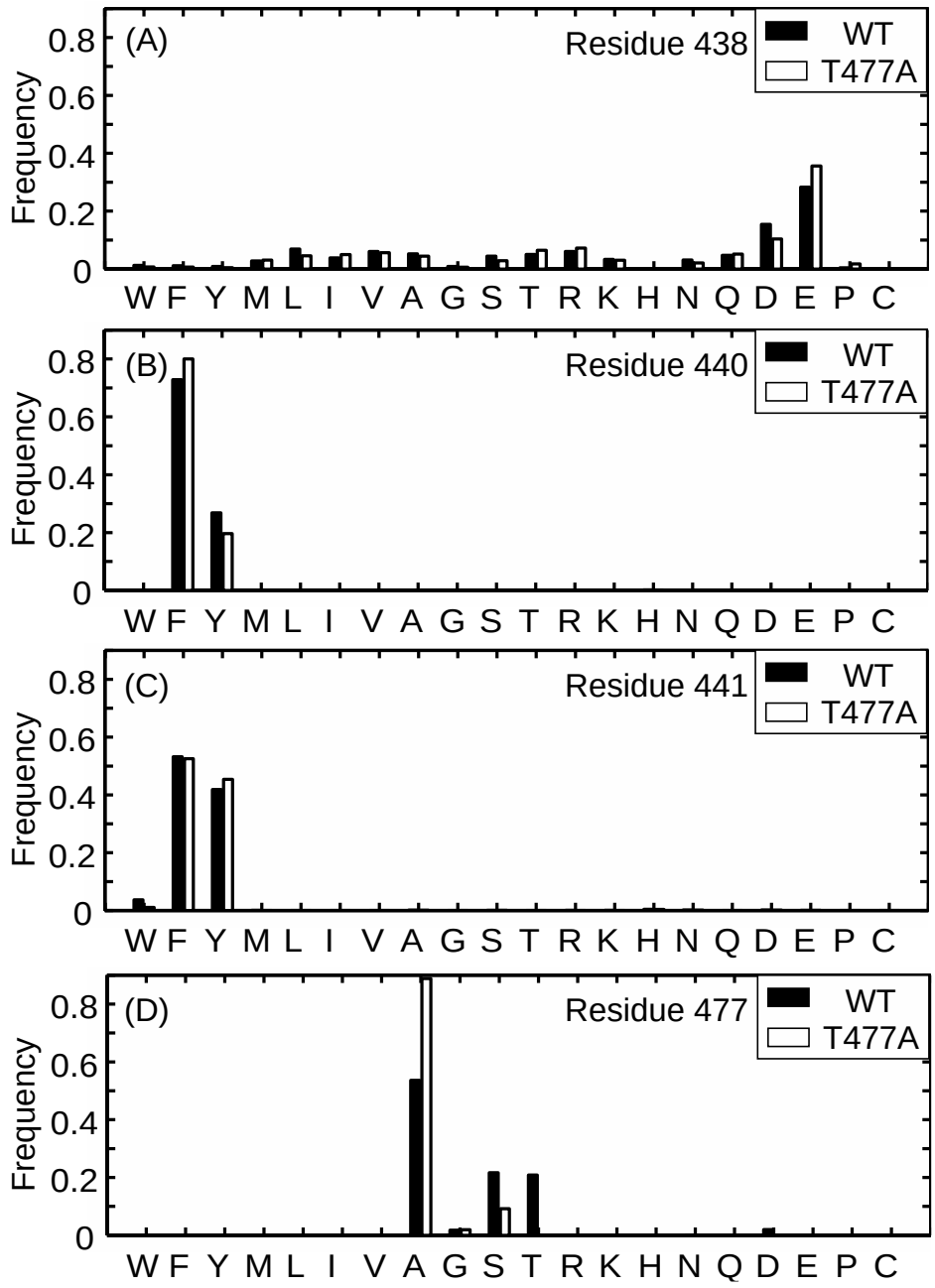

Figure 2.7 : Prediction of sequence tolerance to preserve folding stability as a function of residue position

Histograms of preferred amino acids at residues (A) 438, (B) 440, (C) 441 and (D) 477 for RNH WT (filled bar) and RNHT477A (open bar) coordinates (see Methods 2.3.7).

Since residue 440 is the P1 site as a substrate for the HIV-1 protease, for comparison, sequence tolerance was also tested for residue 441, which corresponds to the P1' site as a substrate for the HIV-1 protease. The calculation indicates that Tyr and Phe are structurally preferred for the P1' site [Fig. 2.7(C)]. Since the hydroxyl group of Tyr side chain of the residue 441 forms a hydrogen bond with the backbone carbonyl of Lys 287 in the p51 subunit, Tyr must be preferred in the actual RNH domain in RT rather than Phe. Our observation suggests that although P1 and P1' sites are typically exposed to the solvent for protease cleavage, those of the p51-RNH processing site, F440 and Y441, contribute to structural stability in order to maintain the protein core (discussed later).

Interestingly, Ala is preferred for the residue 477 compared with Thr [Fig. 2.7(D)], which is consistent to the narrower RMSD distribution in the RNH_{T477A} compared with the WT [Fig. 2.4(A)], and consistent with the above observation that the processing site mutations cause folding defects in the WT but less in the T477A mutants (Figs. 2.2 and 2.4). Nevertheless, the sequence tolerance calculations using the RNH_{T477A} provided essentially the same results (>92%) as those using the WT for residues 438, 440, and 441 (Fig. 2.7). Since the results were similar even when the RosettaBackrub calculation was done using coordinates from the MD simulations at the 100 ns time point (Appendix Figure A.4), the observed tendencies of the residue preferences appear to be insensitive to fine details of the structure used.

2.5 DISCUSSION

The structural behavior of the p51-RNH processing site in the RNH domain is not well understood. Although the RNH domain is rigidly folded in the known crystal structures of the isolated RNH

fragment and the p51/p66 RT heterodimer (54, 102-106, 142, 143), p51-RNH processing would be very inefficient if the processing site were located within a structured domain (144). Indeed, other protease processing sites in HIV-encoded polyproteins, such as N and C-terminus of MA, CA, PR, RT (the N-terminus and the C-terminus of the p66 and p51 subunits), and IN, are exposed to solution at least in the monomer forms (105, 145-148). Thus, it has been hypothesized that the RNH domain may be unfolded or in another conformation in the pre-matured form in RT, that is, p66 monomer or p66 homodimer (54, 92, 97, 102, 149). Alternatively, even though the isolated RNH domain is stably folded in solution (106, 134, 135), the structure may have a potential plasticity to allow conformational change for the protease processing at the p51-RNH site.

To further investigate structural characteristics of the p51-RNH processing site, our study employed a combination of experimental and computational methods to enhance understanding of previously observed phenotypic changes arising from processing site mutations including the impact of the revertant mutation (90, 91). Since the isolated HIV-1 RNH domain itself is not enzymatically active, comparison of the activity cannot be examined (150). Our experimental data show that the processing site mutations, RNH_{F440A} and RNH_{E438N}, result in substantial unfolding of the protein while the revertant, that is, $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$, are significantly folded. Simulations show that the processing site mutations cause changes in side-chain packing and the hydrogen bond network (Table 2.2). Further, simulations of $RNH_{F440A/T477A}$ and $RNH_{E438N/T477A}$ indicate that the T477A mutation shifts the position of helix A relative to the first three β-sheets.

In HIV-1, the T477A variant is extremely common in subtypes F and G (60% and 43%, respectively), in contrast to its rarity in subtype B (1.9%) (151). According to the RNH sequence database, eight residues in the RNH domain of subtypes F and G are highly different from those

in B: T477A, R463K, V466I, D471E, H/Y483Q, L491S, K/Q512R, and A534S. These mutation sites are not located between α-helix A and the β-sheet that includes β-strand 1, 2, and 3, except for R463K in the β3-strand. Indeed, residue R463 is a hydrogen bonding partner of E438. Thus, in subtypes F and G, disruption of the hydrogen bond by R463K, which likely causes similar unfolding of the protein as in the case of E438N, may be rescued by the additional T477A mutation. This evaluation of the sequence variations in the subtypes F and G provides a consistent cooperative mutation effect to the observations reported above.

Together with the NMR, DSF, and MD simulations, the sequence tolerance calculation demonstrates that the processing site residues are important to maintain the RNH structure, that is, buried in the protein core. How- ever, these residues are also preferred as a protease substrate. It is known the P1 and P1' sites of the HIV-1 protease are mostly occupied by bulky side chains: the P1 site amino acids in the HIV-1 substrate are Phe, Leu, Asn, Met, Tyr, and the P1' amino acids are Phe, Pro, Leu, Tyr, Ala, Met (with most frequent residues listed first) (152). The P1 and P1' sites for the p51-RNH processing are occupied by F440 and Y441, respectively. Compared with other protease cleavage sites in HIV-1 polyproteins, polymorphism at the p51-RNH site is small with high conservation of the processing site residues from E438 to V442 (132, 152). Based on our analyses, this is because these residues are needed for the structural stability. Taken together, our results demonstrate that having F440 and Y441 as P1 and P1', respectively, is important for the substrate specificity as well as the structural stability of the protein core.

The structural behavior of RNH in the context of the RT dimers remains to be clarified. As a substrate cleaved by the viral protease, the RNH processing site has to be accessible to the solvent. On the other hand, as revealed above, the p51-RNH processing residues are well arranged in the protein core to maintain RNH folding. Such a coupling of opposite characteristics, substrate

specificity and the structural stability, is puzzling. One possible explanation may be, as hypothesized and proposed previously, the RNH domain is unfolded or in an extended conformation in the RT precursor(54, 92, 97, 102, 149). However, if so, since the secondary and tertiary processing sites within the RNH are not protected(90, 96, 144, 153, 154), the p51 subunit must have variation of the amino acid lengths. Based on the changes in relative fraction of the hydrogen bond network at the processing site (Tables 2.3 and 2.4) and our recent observation of the RNH fold in the p66/p66 homodimer(75), we rather postulate a model in which the hydrogen-bond network may be weakened in the p66 homodimer, possibly due to fluctuation in the domain linker orientation or by the protease interaction to the linker region, increasing the population of the minor open conformation. Indeed, the hydrogen bond network observed in the simulation was not static but exhibited significant fluctuations (Table 2.3).

2.6 CONCLUSION

Our combined NMR and computational results fill in missing pieces of the HIV RT structural story. Our data explain why p66 did not accumulate when the p51-RNH processing site was mutated, and how the revertant mutation, T477A, was able to restore RNH folding, leading to normal proteolytic processing to the p66/p51 heterodimer despite the continued presence of the p51-RNH processing site mutations. A plausible “levering” mechanism for the crosstalk between the region around T477 and the processing site has been proposed based on a total of >1 μs of all-atom MD simulation studies. Sequence tolerance calculations, as well as MD simulations and NMR experiments, indicate that the P1 residue, F440, that is critical for substrate specificity is

also important for the RNH folding; this observation is consistent with the fact that the p51-RNH processing site in the matured RT is protected within the protein core.

2.7 ACKNOWLEDGMENTS

The authors thank Maria DeLucia for technical assistance in cloning, protein expression, and purification, Martin Christen for support of the NMR data analysis, and Troy Krzysiak for guidance and assistance of the Differential Scanning Fluorimetry experiment.

3.0 PRINCIPAL COMPONENT ANALYSIS OF NMR SIGNAL INTENSITY CHANGES TO ASSESS STRUCTURAL INTEGRITY OF A LARGE MULTI-DOMAIN PROTEIN

3.1 SYNOPSIS

Principal component analysis (PCA) permits the assessment of structural integrity of proteins even in the absence of full signal assignments. We applied PCA to investigate time-dependence of spectral characteristics of a ^{15}N-labeled HIV-1 reverse transcriptase precursor, p66, a 130 kDa protein in its homodimeric form. Although complete backbone assignments of p66 are unavailable, spectra of the two isolated domains have been assigned. With incorporation of these previous assignments of the isolated domains, PCA demonstrated that spectral features were stable, with no sign of differential unfolding between the domains. Observation was validated using the ^{13}Cδ1-Isoleucine labeled p66.

3.2 INTRODUCTION

Principal component analysis (PCA) evaluates correlations among the observed variables of a given function. PCA has been applied to a variety of NMR data, including metabolomics data, assessing spectral similarity, and in ligand-titration (155-164). PCA can be also used to assess high-order structure of large proteins, for which signal assignment is not achieved (165). However, PCA is not likely applied to routine inspection of spectral reproducibility and integrity. Given the

effectiveness of PCA, a simple ^{1}H-^{15}N or ^{1}H-^{13}C correlation spectroscopy of a protonated sample in combination with PCA may be a powerful method to assess spectroscopic characteristics of proteins even when amino-acid type labeling in a deuterated background is technically applicable.

HIV-1 reverse transcriptase (RT) precursor, p66, forms a 130 kDa homodimer with a dissociation constant of ~4 μM in solution (75) Although complete assignments of all backbone resonances of the p66 homodimer are not available, spectra of several isolated domains, such as the Ribonuclease H (RNH) and thumb domains, have been assigned and it has been established that the domain structures in the isolated domains is essentially identical that that in mature RT (75, 106, 166, 167). For the p66/p66 homodimer a slow conformational change, involving selective unfolding of the RNH domain in one of the p66 subunits has been proposed. This process was reported to occur over 5-40 hours and purportedly results in an asymmetric homodimer (86, 93). We previously observed NMR spectra of the homodimer but did not find evidence to support RNH unfolding (75).

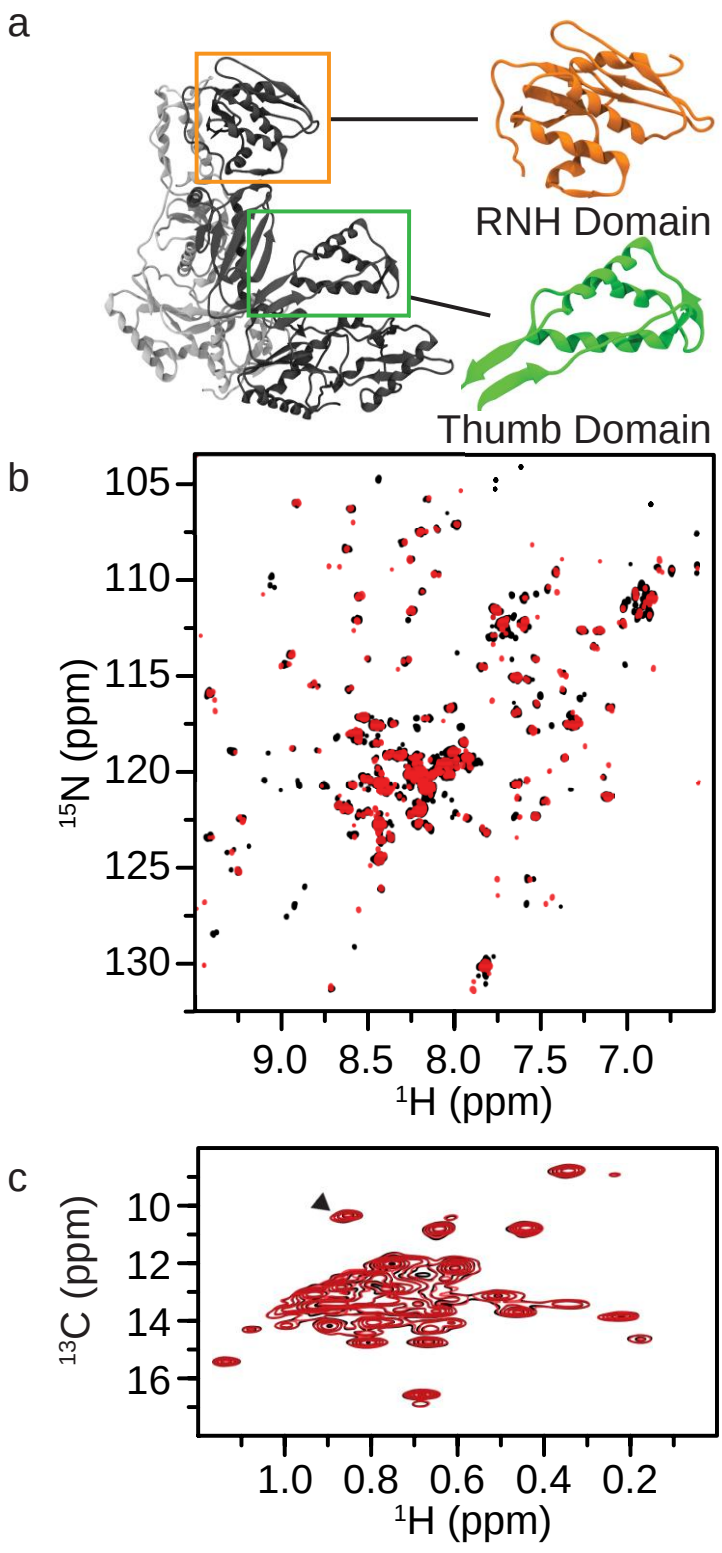

Figure 3.1: Structure of Mature RT and Solution NMR Spectra of Immature p66

(a) Structure of the matured, p66/p51 RT, indicating the location of thumb and RNH domains and (b) ^{1}H-^{15}N TROSY-HSQC spectrum of [U-^{15}N] labeled p66 and (c) ^{1}H-^{13}C SOFAST-HMQC spectrum of [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] labeled p66. In (a) the ribbon structure was generated using pdb: 1DLO (168). In (b), spectra were recorded for 170 µM [U-^{15}N]-labeled p66 in 25 mM Tris buffer, containing 100 mM KCl and 0.02% NaN_3 at pH 7.1. In (c), spectra were recorded for 70 µM [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] labeled p66 in a deuterated 25 mM Bis-Tris buffer, containing 100 mM KCl, 0.02% NaN_3, and 5% v/v Glycerol-d8, at pH 7.1. All the experiments were performed at 35 °C. The initial spectrum (black) and those at 64.3 hours and 56.95 hours (red) are shown in panel (b) and (c), respectively.

Here, we examine the spectral characteristic of the p66/p66 homodimer over an extended time period following the ^{1}H-^{15}N TROSY-HSQC spectra of a [U-^{15}N] labeled p66 over 64 hours to elucidate the structural integrity. We analyze this data using PCA and compare the results with simulated data, considering two scenarios: in Model (I) NMR signals from the RNH domain exhibit similar time-dependent characteristics as those of the thumb domain and unassigned signals, exhibiting changes only due to sample or instrumental instability; in Model (II) NMR signals of the RNH domain change significantly due to unfolding of one of the RNH domains, with concomitant appearance of sharp signals in the unfolded region of the spectra. We verify these observations extracted from ^{1}H-^{15}N spectra by analyzing a time course of ^{1}H-^{13}C SOFAST-HMQC spectra of the [U-^{2}H] and Ile $^{\delta 1}$-[$^{13}CH_3$] labeled protein.

3.3 METHODS

We used the same coding sequence of the RT p66 as was used previously (75), except (i) a V559I polymorphism mutation was included to increase the number of Ile NMR resonances in ^{1}H-^{13}C NMR spectra and (ii) an N-terminal His_6-fusion tag containing a TEV-protease cleavage site was added. [U-^{15}N] labeled p66 was expressed using a minimum media described previously (75). [U-^{2}H] and Ile $^{\delta 1}$-[$^{13}CH_3$] labeled p66 was expressed using the published protocol (74). ^{1}H-^{15}N TROSY-HSQC spectra were recorded using Bruker sequence, trosyetfpf3gpsi, on a Bruker Avance 800 spectrometer equipped with a cryogenic probe head, at time points 0, 12.86, 25.72, 38.58, 51.44 and 64.3 hours. The time point indicates the starting time of each HSQC spectrum that takes 12.86 hours. ^{1}H-^{13}C SOFAST-HMQC spectra were recorded using Bruker sequence, sf_metrosygpph, at time points 0, 4.85, 15.29, 20.12, 24.97, 29.80, 37.27, 42.42, 47.27, 52.10 and 56.95 hours, on a Bruker Avance 900 spectrometer, equipped with a cryogenic probehead. Time required to record each spectrum was 4.85 hours. NMR spectra were processed using nmrPipe and the signal intensities were detected using nmrDraw (169).

To analyze the time-course NMR data of the [U-^{15}N]-labeled p66 protein sample, we generated a total of 200 sets of synthetic time course intensity curves, using the following equation:

$$I(t) = e^{-(t/\tau)} + \varepsilon,$$

Where *I(t)* is intensity at time *t*, and τ is a decay constant, and ε is an additional error term. For Model (I), in which NMR signals of the RNH domain decrease at a rate similar to those of the thumb domain, one hundred intensity curves were generated using τ values which were uniformly varied from 50 hours to 70 hours. For Model (II), in which signals in the two domains undergo

signal intensity changes at different decay rates, fifty intensity curves were generated using τ values which were uniformly varied from 60 to 80 hours, while another fifty intensity curves were generated using τ values which were uniformly varied from 90 to 110 hours. All synthetic intensity curves decayed from an initial value, *I(0)*, of 1 consistent with the normalization procedure used for experimental data. In all subsequent time points ε, was given as a normally distributed random error with variance of ±15% of *I(0)*. For each set of experimental or synthesized time courses, a matrix of n × m was generated. Here, n is the number of residues (or number of synthetic data sets), and m is the number of time course data points. PCA were performed using Matlab software (Mathworks Inc., Natick, MA). Because the peak intensities were normalized as a function of their initial intensity, no additional scaling was applied for the calculation of the principal components. Plotted scores of the first two principal components were used to evaluate the differential signal intensity changes.

3.4 RESULTS AND DISCUSSION

Despite the limited signal-to-noise of the non-deuterated sample, we observed ^{1}H-^{15}N TROSY-HSQC spectra of the [U-^{15}N] labeled p66 with features similar to those previously published (75) (**Figure 3.1 b**). A total 145 peaks were detected and selected for analysis. We used a spectral region from 6.2 to 9.5 ppm in the ^{1}H dimension and 99.0 to 134.9 ppm in the ^{15}N dimension, excluding the side-chain region of 6.7 to 7.8 ppm in ^{1}H dimension and 109.0 to 113.5 ppm in the ^{15}N dimension. A 48% median reduction in the intensity of all signals was observed over the duration of the ^{1}H-^{15}N TROSY-HSQC time-course experiment (**Figure 3.2 a**), presumably due to

sample solubility and stability in the experimental conditions, at a 170 μM protein concentration. Moreover, we did not observe the generation of intense peaks in the unfolded region of the spectrum over the duration of the time course experiment (**Figure 3.1**). If selective unfolding of one domain occurred, it would result in the observation of intense peaks in the unfolded region of the spectrum over the duration of the time course. The appearance of such peaks would be even more pronounced in spectra of the non-deuterated protein.

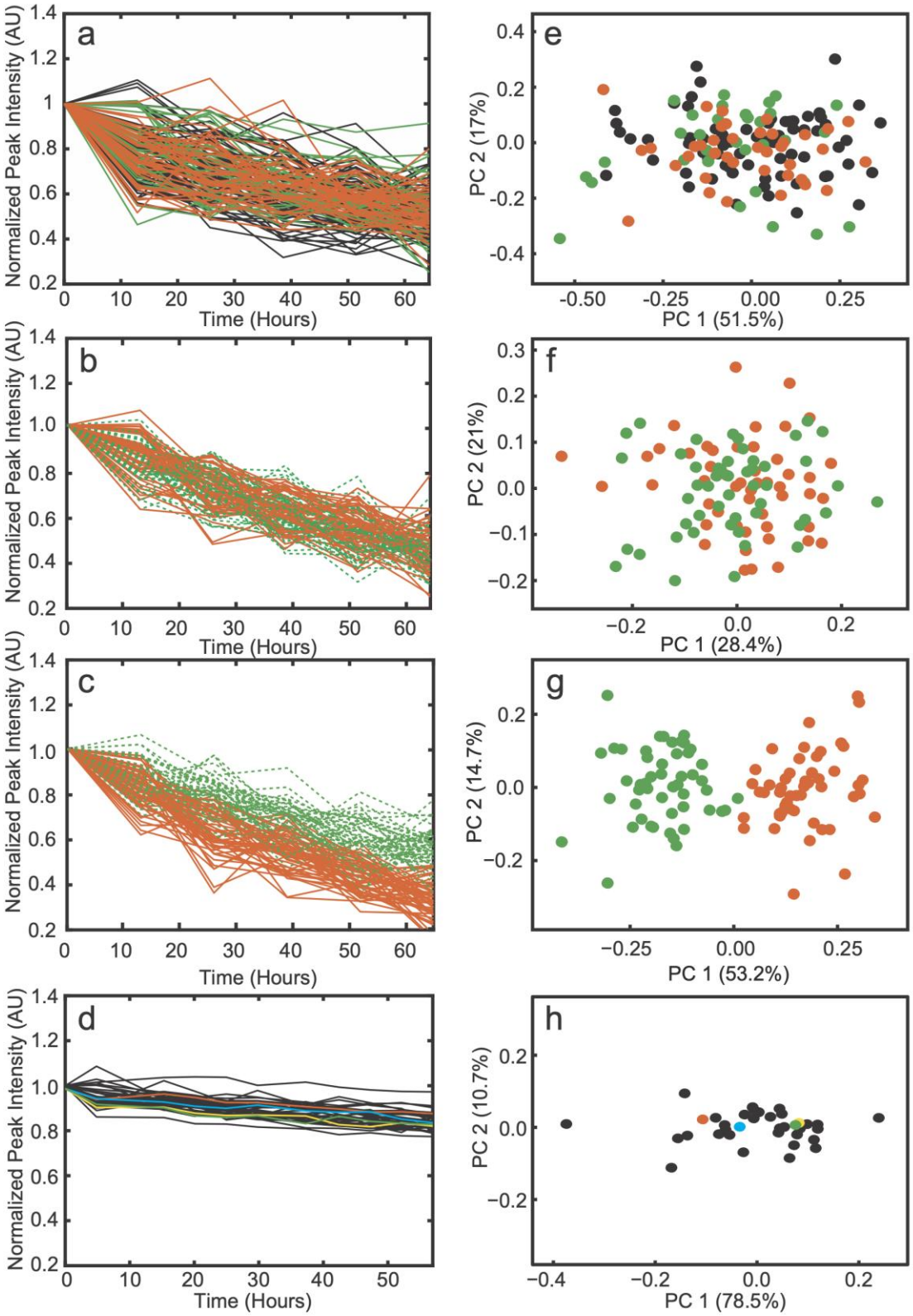

Figure 3.2 : Principal component analysis of experimental and simulated data

(a – d) Time course of the NMR signal intensity changes and (e-h) the PCA score plot: (a, e) the time course and the analysis of ^{1}H-^{15}N TROSY-HSQC spectra; (b, f) simulated time course and the analysis assuming model (I); (c, g) simulated time course and the analysis assuming model (II); (d, h) the time course and the analysis of ^{1}H-^{13}C HMQC spectra.

PCA was performed using the experimental data obtained for peak intensities as a function of time (**Figure 3.2 e-h**). To identify characteristics of the principal components from RNH domain signals, peaks in the p66 spectra were tentatively assigned, based on their chemical shift proximity, within 0.075 ppm of the normalized shift to ^{1}H, to those of isolated thumb and RNH domains (75, 106). Using this criterion, we tentatively identified that 32 and 40 peaks in the p66 spectrum stem from the thumb and RNH domains, respectively. A PCA score plot obtained from the ^{1}H-^{15}N TROSY-HSQC time-course show fairly uniform distributions (**Figure 3.2 e**): no clear trend differentiates the changes observed in the RNH domain from those in the thumb, or unassigned signals. To understand the characteristics seen in the PCA score plot, PCA was also performed for two sets of the synthetic data generated assuming Models (I) and (II) (**Figure 3.2 b** and **3.2 c**, respectively). Synthetic time course data for model (II) show that even an on-average 35% decrease in the decay constant is sufficient to completely resolve independent components in the PC plot (**Figure 3.2 g**), compared to Model (I) (**Figure 3.2 f**). Overall, PCA of ^{1}H-^{15}N spectra did not support differential RNH unfolding (i.e., generation of random-coil resonances) nor differential decay.

To verify the observation, we conducted similar experiments by recording ^{1}H-^{13}C SOFAST-HMQC spectra of [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$]-labeled p66 at low protein concentration, 70

μM. Our ^{1}H-^{13}C SOFAST-HMQC spectra was similar to the initial time-course spectrum published by Zheng, et al (92), with an additional peak resulting from the V559I polymorphism, indicated with a black triangle (**Figure 3.1 c**). Similar to the above ^{1}H-^{15}N TROSY-HSQC time course, a gain of signal intensity as a function of time, indicative of domain unfolding, was not detected (**Figure 3.1 c**). In the initial ^{1}H-^{13}C HMQC spectrum, 36 total peaks were detected from -0.4 to 1.8 ppm in the ^{1}H dimension and 2.0 to 25.5 ppm in the ^{13}C dimension. The median reduction in the intensity of all signals by 16% was observed over the duration of the time course experiment (**Figure 3.2 d**). A PCA score plot of the ^{1}H-^{13}C spectrum time course showed one cluster, with two "outliers" representing the peaks which decayed the most and least compared to their initial intensities (**Figure 3.2 h**). Since the sample size of the Ile $^{\delta 1}$-data is small, these data points seemed to be outliers even though the entire data set, including these two points, was correctly explained with a normal distribution. Taken together, the ^{1}H-^{13}C HMQC time course experiment exhibited results consistent to those of the ^{1}H-^{15}N spectrum time course.

3.5 CONCLUSION

This study demonstrates the application of PCA for the assessment of structural integrity of a large protein as a function of time. Data recorded using a [U-^{15}N]-labeled p66 had lower sensitivity, but an advantage of a larger sampling size, than the [U-^{2}H] and Ile $^{\delta 1}$-[$^{13}CH_3$]-labeled p66, demonstrating complimentary information to each other. Such a simple analysis using a [U-^{15}N]-labeled protein is useful to provide insight into the structural integrity of a protein during a long-term NMR experiment. We did not observe slow unfolding of either RNH domain within the p66 homodimer nor did we observe the presence of discrete populations as a result of differential

domain unfolding. HIV-1 RT, a heterodimer of p66 and p51 subunits, is generated by proteolytic cleavage of one RNH domain from the p66/p66 homodimer (99, 170, 171). Given that the exact mechanism by which only one RNH domain in the immature p66/p66 homodimer is unclear, the results presented here may help to provide a mechanistic basis of the maturation.

3.6 ACKNOWLEDGEMENTS

We thank Mike Delk for NMR support, Michel Guerrero for the support of the protein purification, Arthur Okura for helpful discussion, and Teresa Brosenitsch, Angela M. Gronenborn, and Andrew P. Hinck for critical reading of the manuscript. The project is supported by NIH NIGMS and NIAID (R01GM105401 and P50GM082251).

4.0 EFFECT OF TRNA ON THE MATURATION OF HIV-1 REVERSE TRANSCRIPTASE

4.1 SYNOPSIS

The mature HIV-1 reverse transcriptase is a heterodimer that comprises 66 kDa (p66) and 51 kDa (p51) subunits. The latter is formed by HIV-1 protease-catalyzed removal of a C-terminal ribonuclease H domain from a p66 subunit. This proteolytic processing is a critical step in virus maturation and essential for viral infectivity. Here, we report that tRNA significantly enhances in vitro processing even at a substoichiometric tRNA:p66/p66 ratio. Other double-stranded RNAs have considerably less pronounced effect. Our data support a model where interaction of p66/p66 with tRNA introduces conformational asymmetry in the two subunits, permitting specific proteolytic processing of one p66 to provide the mature RT p66/p51 heterodimer.

The work presented in this chapter was reproduced with permission from: Ilina TV, Slack RL, Elder JH, Sarafianos SG, Parniak MA, Ishima R. (2018), Effect of tRNA on the Maturation of HIV-1 Reverse Transcriptase. Journal of Molecular Biology, 430(13):1891-900. doi: 10.1016/j.jmb. In the work, in-vitro proteolytic processing experiments and fluorescence-based determination of binding constants were performed by Dr. Tatiana Ilinia.

4.2 INTRODUCTION

HIV-1 reverse transcriptase (RT) is a multifunctional enzyme with both DNA polymerase and ribonuclease H (RNH) activities and is essential for HIV-1 replication(9). While the gene for RT encodes a 66-kDa protein, RT is initially translated as part of the 160-kDa Gag-Pol precursor polyprotein, which is proteolytically processed by HIV-1 protease (PR) during virion assembly and maturation. The mature RT is a heterodimer comprising 66-kDa (p66) and 51-kDa (p51) subunits; the p51 subunit is generated upon proteolytic cleavage of the p66 subunit between residues 440 and 441, thereby removing most of the RNH domain from this p66 subunit (9, 59, 96, 97, 172). This cleavage event is essential for viral infectivity (90, 91).

Formation of the mature p66/p51 heterodimer is generally thought to proceed via a p66/p66 homodimer intermediate (Sequential model in Fig. 4.1 a), rather than generation of p51 followed by p66/p51 heterodimer formation (Concerted model in Fig. 4.1 a) (99, 170, 171). However, data in support of the Sequential model are primarily based on studies that introduce mutations leading to dissociation of subunits in p66/p66 in vitro or at a viral level (99, 170, 171). Although lack of p66/p51 production using the mutants may support the requirement of the homodimer formation for the RT maturation (i.e., Sequential model), the mutants were originally known to affect p66/p51 heterodimer formation (173-176). Therefore, such experiments do not consider whether the mutations themselves diminish maturation of p66/p66 or whether they cause p66/p51 dissociation which in turn results in decreased RT detection. Furthermore, structures of both p66/p51 (Fig. 4.1 b) and the isolated RNH domain indicate that the p51-RNH cleavage site is buried within the core β-sheet of the folded RNH domain (Fig. 4.1 C) and likely inaccessible to the protease (54, 102-106). Despite efforts to determine the structure of the RNH domain in the immature RT p66/p66 homodimer, a detailed structure of p66/p66 is not available (75, 86, 87, 93).

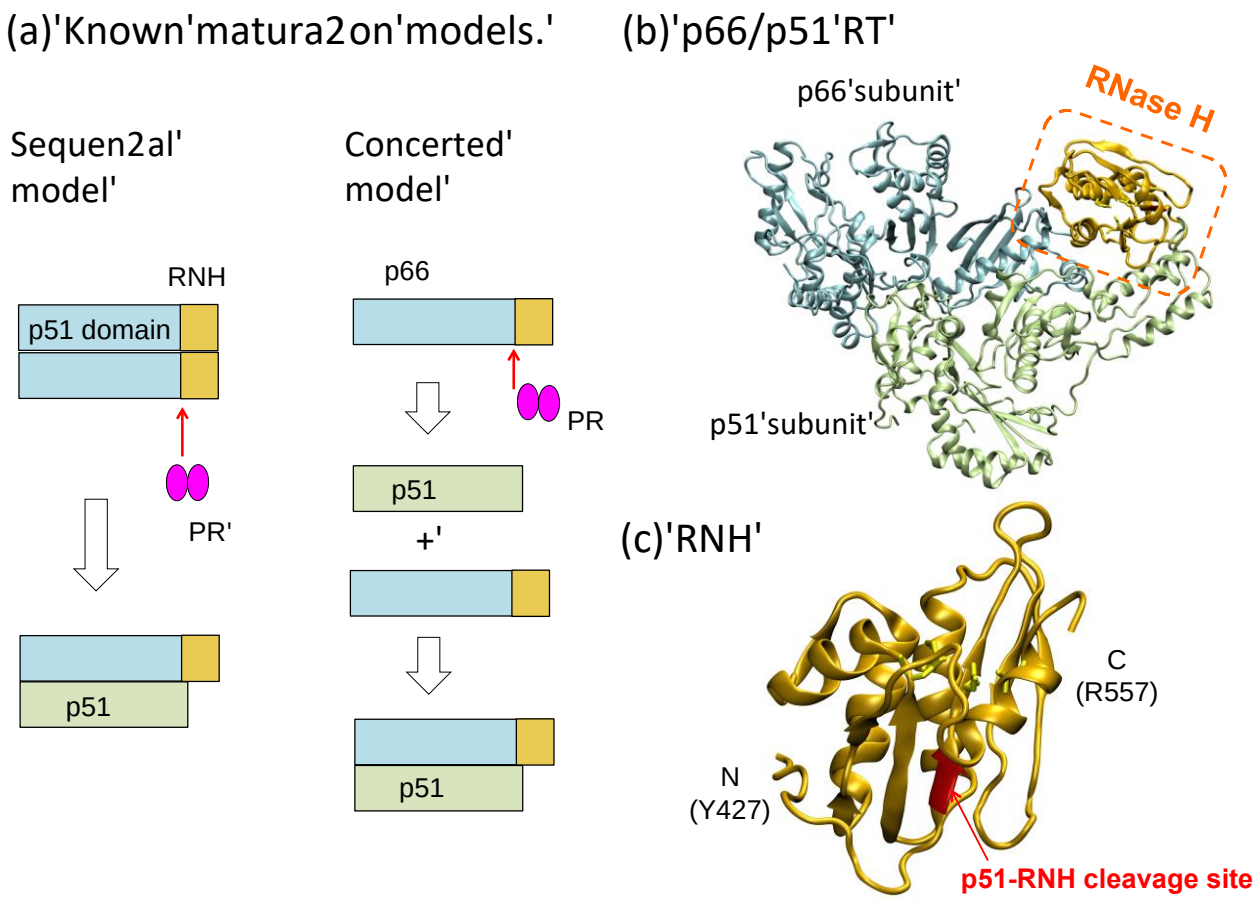

Figure 4.1: Proposed Models of RT Maturation and Structure of Mature RT

(a) Concerted and Sequential Models for RT maturation (59, 99, 170, 171). (b) Location of the RNH domain in the known p66/p51 RT structure. (c) An expanded view of the RNH domain. In panels a–c, the p51 and the RNH domains in the p66 subunit are shown in cyan and orange, respectively, while the p51 subunit is shown in lime color. In panels b and c, the p51-RNH cleavage site (F440-Y441) is shown as a red ribbon. The structures were generated using PDB code 1DLO (121).

Our previous NMR studies suggested that the p51- RNH cleavage sites in both subunits of p66/p66 were buried and thus poorly accessible to protease (75). We therefore postulated that some virion-associated factor may play a role in promoting the RT maturation process. HIV-1 virions

are known to contain substantial amounts of cellular tRNA (44, 177-180), and RT can bind such tRNAs with nanomolar affinity (181). Here, we describe biochemical experiments to assess the impact of tRNA on the in vitro HIV-1 PR-catalyzed cleavage of p66/p66 to mature RT p66/p51. We show that processing of p66/p66 is slow and results in numerous aberrant products in the absence of tRNA. Surprisingly, in the presence of even a substoichiometric tRNA:p66/p66 ratio, processing of p66/ p66 by HIV-1 PR is greatly accelerated and with few or no aberrant cuts. Our data, obtained using two different p66 concentrations, show that the mature p66/p51 arises from proteolytic processing of the p66/p66, and not from processing of p66 monomers, in agreement with previously published data (99, 170, 182, 183). Although a certain processing enhancement was observed in the presence of ds-RNA, based on the fact that HIV-1 virions are known to contain substantial amounts of cellular tRNA (44, 177-180), we propose a model where virion-encapsidated tRNA facilitates RT maturation to p66/p51 that is essential for HIV-1 replication.

4.3 MATERIALS AND METHODS

4.3.1 Protein Expression and Purification

p66/p51 HIV-1 RT was prepared using the p6HRT-PROT plasmid (184) as previously described (185, 186). The p66 sequence from p6HRT-PROT used for p66/ p51 expression (184) was cloned into the pPSG-IBA3 vector using the StarGate cloning system (IBA Solutions for Life Sciences, Göttingen, Germany). p66 protein was expressed in BL21 (DE3) Escherichia coli cells and purified using a Strep-Tactin gravity flow column (IBA Solutions). Protein concentration (calculated as p66 monomer) was determined by measuring absorbance at 280 nm with an extinction coefficient

of 137,405 M^{-1} cm^{-1}. Purified p66 and p66/p51 were stored in 25 mM sodium phosphate (pH 7.0), 250 mM NaCl, and 50% v/v glycerol at −80 °C. HIV-1 PR was expressed and purified as previously described (187, 188).

4.3.2 HIV-1 PR-catalyzed processing of p66/p66

Proteolytic processing of p66 protein was evaluated using kinetic (time-course) experiments that determined the rate of processing and fixed-time experiments that assessed the impact of differences in tRNA:protein ratio on extent of processing. Since a monomer:dimer ratio of p66 depends on protein concentration and tRNA concentration, all p66 concentrations are reported as monomer protein concentration for the Materials and Methods purpose. Processing experiments were carried out in 20 mM sodium acetate buffer (pH 5.2) at 37 °C, unless otherwise noted.

Kinetic experiments were conducted at two p66 concentrations: “high p66 concentration” (8 μM), where RT is predominantly in the p66/p66 homodimer form and processed by 1 μM PR in the presence of 0, 0.5, or 8 μM synthetized $tRNA^{Lys3}$ of human sequence (TriLink BioTechnologies LLC, San Diego, CA), and “low p66 concentration” (1 μM), where RT is predominantly in the p66 monomer form and processed using 0.25 μM PR in the presence of 0, 0.1, or 2 μM tRNA. Additional kinetic experiments at high p66 (8 μM) were conducted in the sodium acetate buffer containing 100 mM NaCl at 20 °C. Fixed-time experiments were conducted using high p66 (8 μM) conditions only, at different concentrations of tRNA or NaCl or heparin (Sigma-Aldrich. St. Louis, MO), which were added to p66 protein prior to starting the reaction, incubated at room temperature for 5 min, and allowed to react at 37 °C for 10– 20 min. The fixed-time experiments at different tRNA concentrations were performed in sodium acetate buffer with/without 100 mM NaCl.

The kinetic (time-course) modes of p66 proteolytic processing were performed at 8 μM p66 and 1 μM PR, in the presence of 0.5 μM ds-RNA (40/22 nt), ss-RNA (40 nt), and ss-RNA (20 nt), of which 40- and 20-nt sequences are 5′-AGGUGAGUGAGAUGAUAACAA AUUUGCGAGCCCCAGAUG and 5′-GCAUCUGGG GCUCGCAAAUUUG, respectively (TriLink BioTechnologies LLC).

In all the processing experiments, reactions were stopped by addition of Tricine sample loading buffer (Bio-Rad Laboratories, Berkeley, CA) and denaturation at 95 °C for 5 min. Samples were loaded onto precast 4–15% Tris-glycine gels (Bio-Rad) and stained with Bio-safe Coomassie stain (Bio-Rad). Band intensities were quantified using an Odyssey CLX gel imaging system by Image Studio software (Li-Cor Biosciences, Lincoln, NE) or an Amersham Imager 600 (GE Healthcare Life Sciences). For quantification, these gel experiments were repeated at least three times. Production of RT p66/p51 heterodimer was determined based on the ratio of p66 to p51 band intensities in the following way: the ratio of p66 to p51 band intensities of a reference heterodimer was first quantified in the same gel, and production of the heterodimer against the initial p66 intensity was determined using the p51 band intensity normalized by the reference p66/p51 intensity ratio. An average of the three quantified p66/p51 production was plotted with the standard deviation as an error bar. Trends of intensity changes were shown by fit curves using Igor (Wavemetrics, Inc., Lake Oswego, OR).

4.3.3 Analytical Size Exclusion Chromatography to Monitor p66/p66-tRNA Interaction

All size exclusion chromatography (SEC) experiments used a 24-ml analytical Superdex 200 Increase 10/300 GL column (GE Healthcare), mobile phase of 25 mM Bis–Tris buffer, pH 7.0, containing 100 mM NaCl with 0.02% sodium azide at a flow rate of 0.5 ml/min. Injection volume

was 50 μl, and protein elution was monitored by UV absorbance at 254 and 280 nm. Elution profiles of 40 μM p66 RT were evaluated in the absence of tRNA, or following preincubation with 5 μM tRNA containing tracer tRNA 3′-end labeled with pCp-Cy3 (Jena Bioscience, Jena, Germany). As a control, the labeled tRNA was injected without mixing with p66. With the SEC experiments that contain labeled tRNA, in addition to UV, the fluorescence emission at 560 nm (excitation 485 at nm) was also measured using an in-line Shimadzu RF-10AXL Fluorescence Detector.

4.3.4 Fluorescence Spectroscopic Analysis of p66/p66-tRNA Interaction

The interaction of tRNA with RT p66 was evaluated in 25 mM Bis–Tris (pH 7.0) and 100 mM NaCl. RT p66 protein in various concentrations was incubated for 30 min with 1 μM total tRNA containing tracer Cy3-labeled tRNA. Emission spectra were collected using a FluoroMax-4 (Horiba Scientific, Edison, NJ) with excitation at 485 nm. All experiments were carried out at least three separate times to determine an average and a standard deviation of the data. The change in fluorescence at 560 nm was plotted at different protein concentrations 0–20 μM, with error bars representing one standard deviation. The tRNA dissociation constant, K_D, was determined assuming two models, one with a single K_D and the other with two independent K_Ds, using $\chi 2$-minimization routine in Matlab (MathWorks, Natick, MA), and evaluated using the F test.

4.4 RESULTS AND DISCUSSION

4.4.1 p66/p66 Homodimer Formation is Essential for Efficient RT Proteolytic Processing

We evaluated the kinetics of in vitro proteolytic processing by HIV-1 PR at different concentrations of RT p66 in the absence and in the presence of tRNA (Fig. 4.2).

With 8 μM p66, over 50% of RT is predicted to be the p66/p66 homodimer, based on a p66/p66 dissociation constant of ~4 μM (60, 61, 189). In the absence of tRNA, minimal formation (<10%) of p51 is noted after 1 h (Fig. 4.2 a and b). In contrast, with stoichiometric amounts of tRNA, essentially complete processing to equivalent amounts of p66 and p51 (the RT p66/p51 heterodimer) is seen within 30–40 min. Under the same conditions, enhancement of processing is noted even with substoichiometric levels of tRNA (0.5 μM). Surprisingly, the extent of processing under these sub-stoichiometric conditions, after 20 min, was more than 50% of that observed for the 8 μM tRNA samples. This suggests that the tRNA may act as a catalytic factor in the processing event.

In contrast, at low p66 concentrations (1 μM) where p66/p66 homodimer formation is minimal (<40%) (60, 61, 189), very little p51 was formed even after 1-h incubation, either in the absence of tRNA or in the presence of excess tRNA (2 μM) (Fig. 4.2 c and d). Significant aberrant processing was also evident with low p66 and excess tRNA. Because of such p66 degradation, p66/p51 production did not necessarily saturate as a function of time (Fig. 4.2). These results suggest that p66/p66 formation is important for proper maturation of RT to the p66/p51 heterodimer.

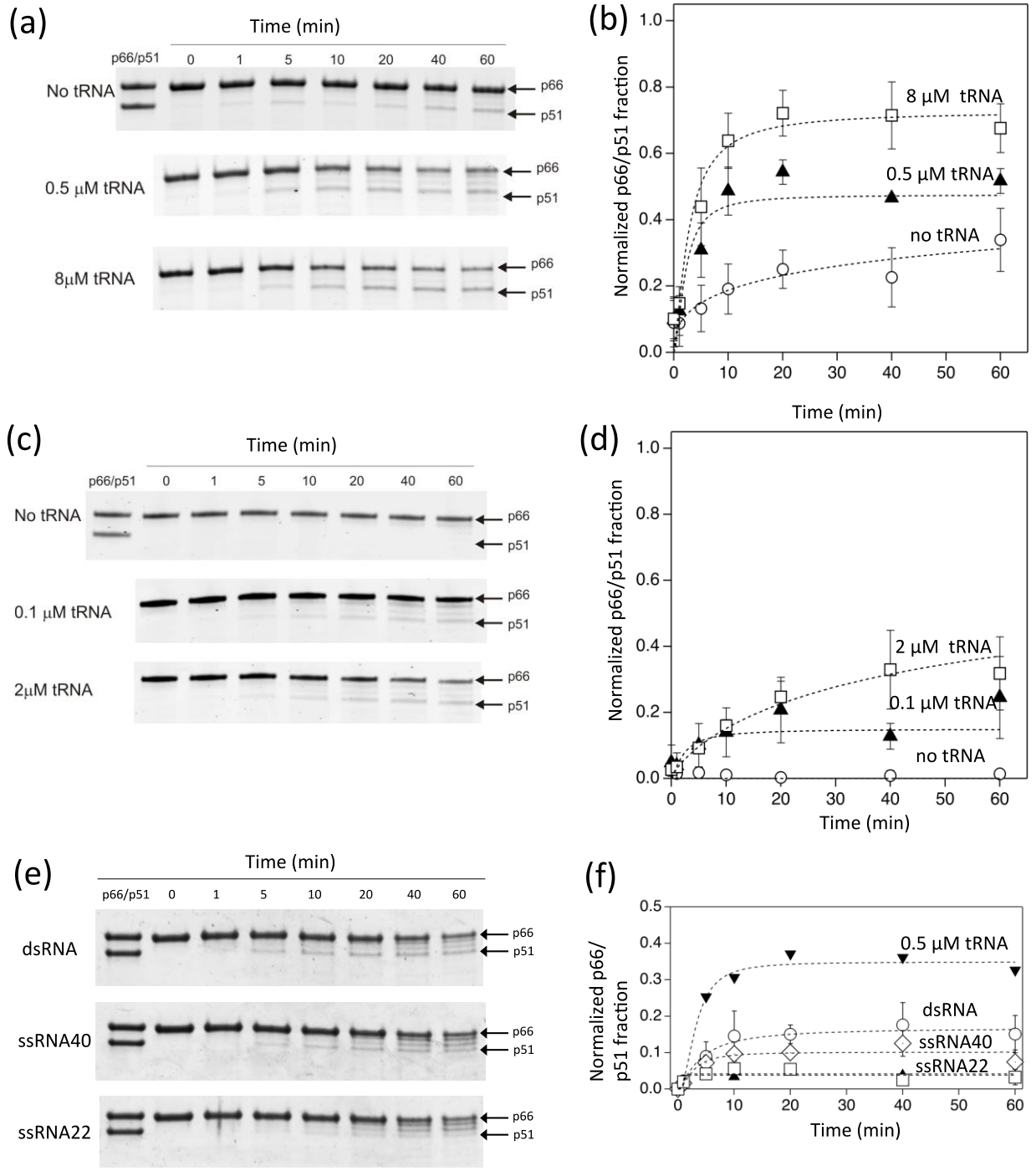

Figure 4.2 : Time dependence of p66 processing by HIV-1 PR monitored by SDS-PAGE

All Reactions performed in 20 mM sodium acetate buffer at pH 5.2 and 37 °C, (a and b) at a high concentration of p66 (8 µM as a p66 monomer, 4 µM as a dimer) proteolytically processed by 1 µM HIV-1 PR in the presence of 0, 0.5, 8 µM tRNA concentrations, (c and d) at a low concentration of p66 (1 µM p66 monomer concentration, 0.5 µM

as a dimer) processed by 0.25 μM HIV-1 PR in the presence of 0, 0.1, and 2 μM tRNA concentrations, and (e and f) at a high concentration of p66 (8 μM as a p66 monomer, 4 μM as a dimer) proteolytically processed by 1 μM HIV-1 PR in the presence of 0.5 μM ds-RNA (40/22 nt), ss-RNA (40 nt), and ss-RNA (22 nt). In panels b, d, and f, p66/p51 fractions, determined from the gel images in panels a, c, and e, are shown, respectively. Average data points were used to fit a curve with one standard deviation as an uncertainty of each data point. Since p66/p51 production occurs in parallel with p66 degradation by PR, the build-up curves do not necessarily reach a plateau. Because of these multiple factors, normalized χ^2, χ^2_n, values for the curve fits were between 0.2 and 10.7.

To determine whether maturation enhancement occurs in the presence of other RNAs, a similar set of experiments was performed using some other small RNA molecules: ds-RNA (40/22 nt), ss-RNA (40 nt), and ss-RNA (22 nt) (Fig. 4.2 e and f). We chose the substoichiometric nucleic acid concentration, 0.5 μM, to see the effect of RNA on the RT maturation but not the effect of stabilizing the matured RT. The positive control, 0.5 μM tRNA, showed similar, although slightly lower, activity compared to that obtained in our first set of experiments (see Fig. 4.2 a and b), presumably due to slightly lower activity of PR used in the experiments in Fig. 4.2 e. The efficiency of RT maturation in the presence of ds-RNA was about half, ~ 47%, of that obtained in the presence of tRNA (Fig. 4.2 e and f). In the presence of the ss-RNAs, the efficiency was even lower, 29%. The results demonstrate that ds-RNA, in addition to tRNA, also enhances RT maturation, although to a lesser degree for those tested in our study.

4.4.2 Other Factors Which Potentially Impact tRNA-mediated p66/p66 Processing

The data presented in Fig. 4.2 show that tRNA facilitates the formation of RT p66/p51 heterodimers. Confirmation was obtained using dose-dependency experiments, in which p66/p51 production by PR was assessed at different tRNA concentrations (Fig. 4.3 a). Since both p66/p66 dimer formation and PR activity may be influenced by ionic strength (190, 191), and tRNA may act as a polyanion (192-194), we investigated the impact of different salt concentrations on p66 processing. In the absence of tRNA, at physiological NaCl concentrations (50–150 mM), aberrant non-specific cleavage products (Fig. 4.3 b, arrows), with molecular sizes between those of p51 and p66, were seen; significant p66/p51 formation became evident only above 400 mM NaCl (Fig. 4.3 b). In contrast, even at low levels of tRNA, substantial RT p66/p51 formation was noted in the presence of 100 mM NaCl (Fig. 4.3 c). Heparin did not enhance p66/p66 maturation (Fig. 4.3 d), and PR activity in the assay with fluorescent HIV protease substrate 1 (Sigma) was actually less in the presence of 0.5, 1.0, or 5.0 μM tRNA (concentrations at which we observed enhanced p66 processing efficiency) compared to its activity without tRNA or at 10 μM tRNA (195), indicating that a simple PR activity enhancement by polyanions does not explain our results (194). These data show that while higher ionic strength may enhance proteolytic activity, as well as dimer formation, it does not impact the total production of p66/p51 (Fig. 4.3 f).

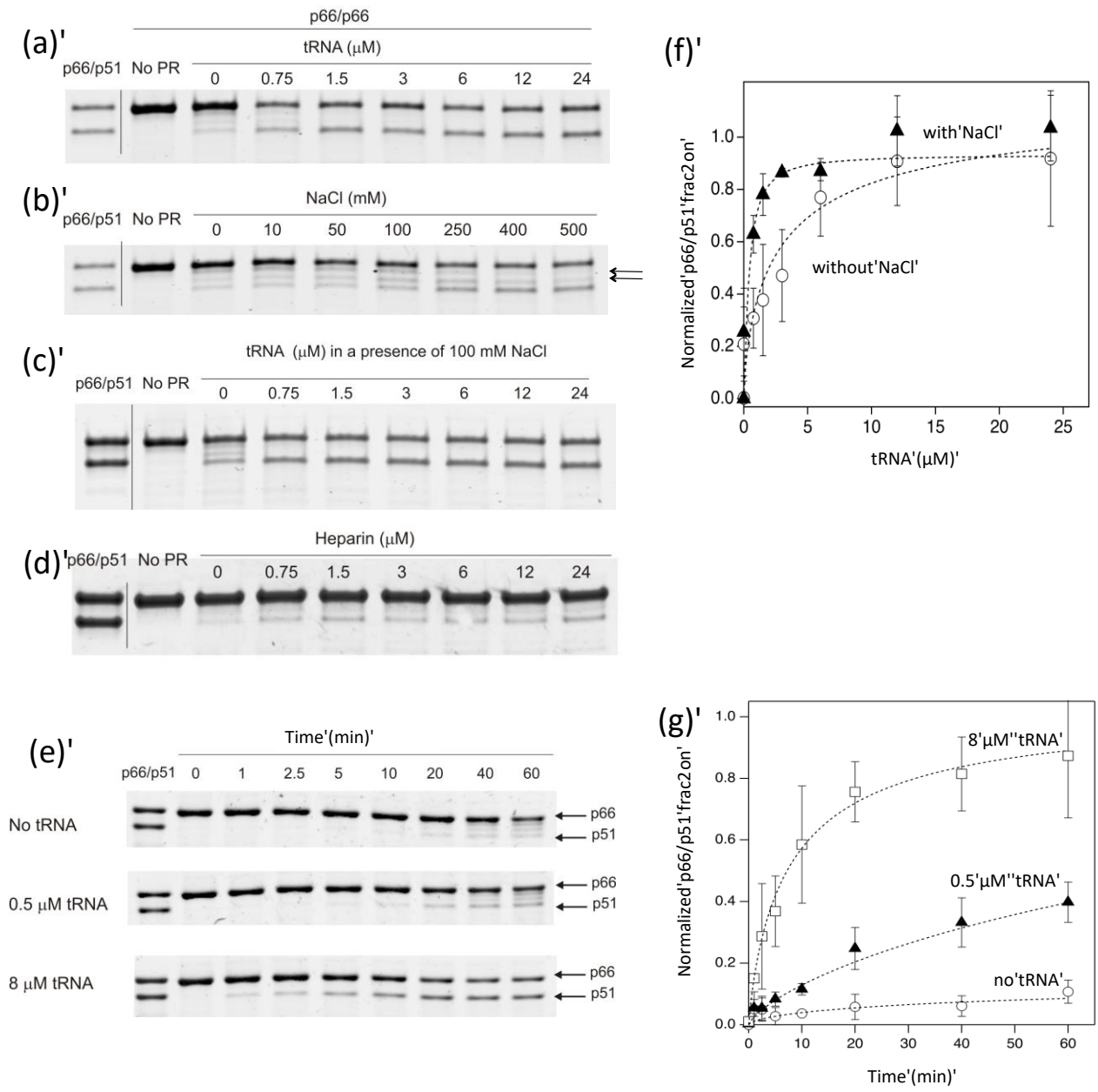

Figure 4.3 : Concentration dependent effects of tRNA, ionic strength, and heparin on HIV-1 PR mediated maturation of p66 monitored by SDS-PAGE.

Varying amounts of (a) tRNA, (b) NaCl, (c) tRNA in the presence of 100 mM NaCl, and (d) heparin on p66 processing by HIV-1 PR for 20 min at 37 °C, and (e) the time dependence of p66 processing by HIV-1 PR in the presence of 100 mM NaCl at 20 °C. In all the experiments, reactions were initiated by the addition of 1 μM HIV-1 PR to p66 at a high concentration of p66 (8 μM as a monomer, 4 μM as a dimer) and monitored by SDS-PAGE, and the buffer contains 20 mM sodium acetate at pH 5.2. In panel e, experiments were carried out in the presence of 0, 0.5, and 8 μM tRNA.

(f) Plots of intensity changes and the fit curves, χ^2_n values 2.7 and 4.5, are shown for panels a and c, respectively. (g) Plots of intensity changes and the fit curves, χ^2_n values from 0.88 to 2.2, are shown for panel f. In panel b), aberrant non-specific cleavage products are shown by arrows.

Kinetic experiments of p66 processing in a buffer that contained 100 mM NaCl (Fig. 4.3 e and g) also showed a pattern of the p66/p51 production similar to that observed in a buffer lacking 100 mM NaCl (Fig. 4.2 a), that is, tRNA impacts p66/p51 production. Of note, the reaction was performed at 20 °C instead of 37 °C because the PR catalytic rate increases in the presence of 100 mM NaCl, compared to the absence of NaCl (37 °C). Overall, our data, obtained with different NaCl conditions (Figs. 4.2 and 4.3), consistently show that tRNA influences the selectivity of cleavage at the processing site. Although there are reports that tRNA may act as a polyanion (192-194), tRNA clearly increases selectivity of the processing site for p66/p51 production in the presented RT maturation experiments.

4.4.3 p66/p66 Interaction with tRNA

Our biochemical in vitro proteolytic processing studies strongly suggest that RT p66/p66 homodimer is the substrate for HIV-1 PR processing to mature RT p66/p51 heterodimer. To further evaluate this, we carried out biophysical analysis of the RT species using size-exclusion chromatography (Fig. 4.4).

In the absence of tRNA, both monomer and dimer peaks of p66 were observed, consistent with previous data (Fig. 4.4 a) (75). tRNA alone eluted at a slightly greater volume than p66

monomer alone, based on both UV and the fluorescence detection (Fig. 4.4 b). In the mixture of p66 and tRNA, two additional peak signals, compared to p66 alone, were noted (Fig. 4.4 c); these stem from p66/p66–tRNA and p66–tRNA complex species, as confirmed by the fluorescence emission of Cy3-labeled tRNA. The loaded p66 and tRNA concentrations, 40 and 5 μM, respectively, are empirically estimated to be 4 μM p66 and 0.5 μM tRNA on a column, which are similar to the conditions used in the processing experiments (Fig. 4.2 a).

The tRNA interaction with p66 protein was quantified by recording changes in fluorescence emission of Cy3-labeled tRNA at varying p66 concentrations in the fluorescence spectroscopic analysis (Fig. 4.4 d). The emission changes at 560 nm could not be described by a single-site binding model (Fig. 4.4 e, dashed line) and were better explained with a two-binding mode model (Fig. 4.4 e, solid line) that gave two K_D values: 65.8 ± 26.9 nM and 2.38 ± 0.69 μM. When considered in the context of the known concentrations of p66, we conclude that these dissociation constants reflect tRNA dissociation from p66/p66 (dimer) and from p66 (monomer), respectively. Competitive gel mobility shift assay of a solution containing tRNA, p66, and PR suggests that tRNA more strongly interacts with p66 than PR (Cite Ilina, 2018). Taken together, we confirmed the p66/p66–tRNA interaction. In addition, observation of p66–tRNA species explains why p66 processing at low p66 concentration showed tRNA dependence (Fig. 4.2 c and d).

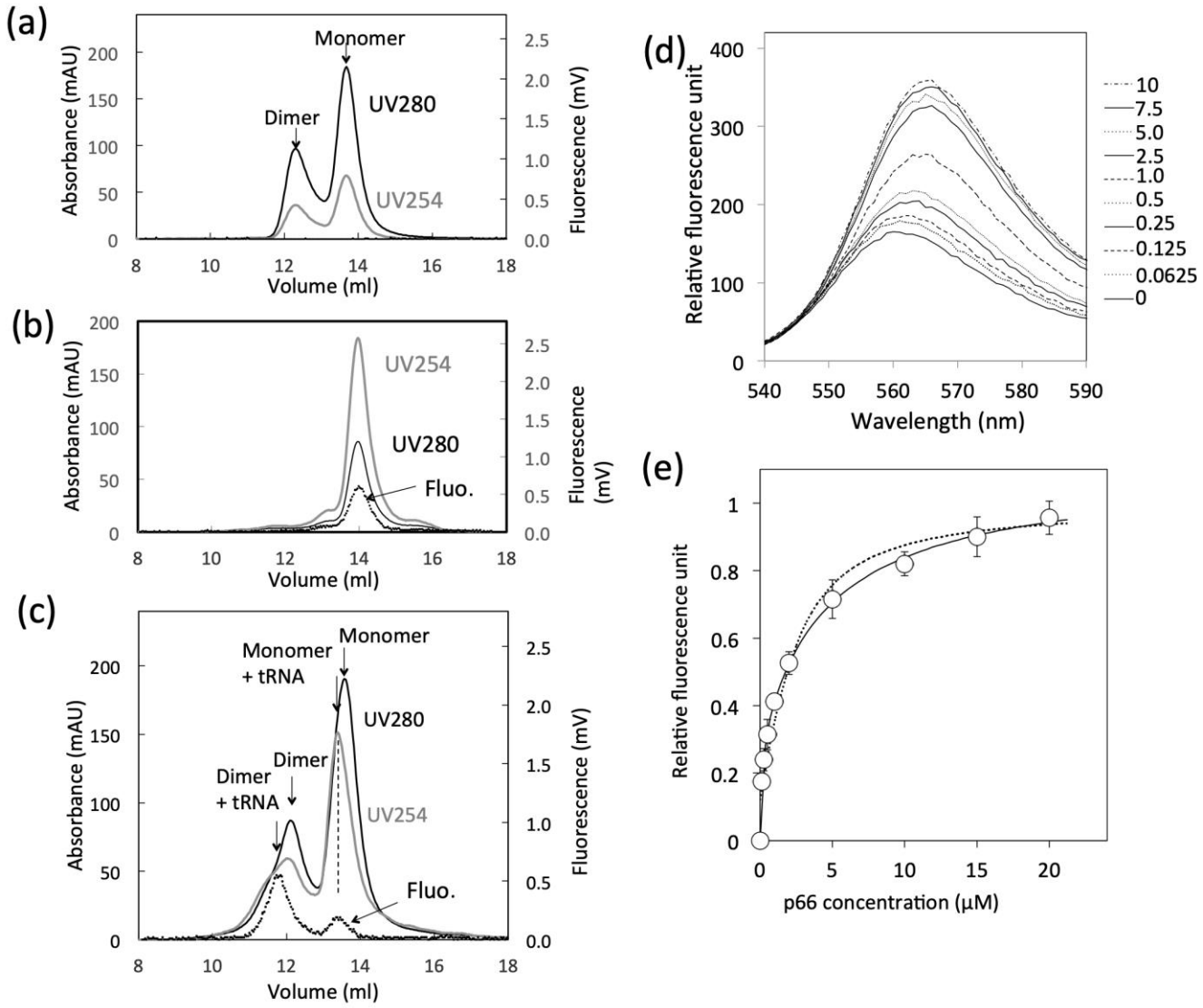

Figure 4.4 : Interaction of tRNA with p66 monitored by SEC and fluorescence spectroscopy.

SEC elution profiles of (a) free p66 protein, (b) free tRNA, and (c) p66 + tRNA, and (d, e) change of fluorescence emission of Cy3-labeled tRNA at varying p66 concentrations obtained by spectrofluorometry. In panels a–c, UV absorbance at 254 nm (gray line) and 280 nm (black line), and fluorescence emission at 560 nm (dashed line) are shown. In panel d, changes in fluorescence of total 1 μM tRNA, containing 40 nM Cy3-labeled tRNA, as a function of p66 concentration were recorded. In panel e, fluorescence intensity changes at 560 nm (d) are plotted. The dash line is a fit curve calculated with a single binding mode model (χ^2_n = 6.5), and the solid line indicates a fit-curve calculated with a two- binding mode model (χ^2_n = 0.8). The null hypothesis was rejected with p = 0.00036.

4.4.4 A Proposed Model for RT Maturation

Here, we demonstrate that tRNA interacts with p66/ p66 homodimer and facilitates selective cleavage at the p51-RNH site by HIV-1 PR (Fig. 4.2). We also demonstrate that the effect of tRNA on the selectivity of p66 processing is distinct from the effect of ionic strength (Fig. 4.3 and C.2). Consistent with the processing experiments, both tRNA–p66/p66 and tRNA–p66 forms were observed. Although tRNA can bind p66/ p51 tightly (K_D = 3–50 nM (59, 181, 196)), tRNA interaction with p66/p66 is the same as, or weaker than, the p66/ p51 binding (Fig. 4.4 e). Because of this moderate binding and observation of the significant tRNA effect at a substoichiometric concentration, it is possible that transient interaction of tRNA with p66/p66 mediates the selective processing at the p51-RNH site, with tRNA likely being released before heterodimer formation is complete (59, 61, 191). Such gain of RT maturation at the substoichiometric concentration suggests that the tRNA may serve as a catalytic molecule in RT maturation, rather than just stabilizing the matured RT in solution.

Our *in vitro* data suggest that ds-RNA, in addition to tRNA, also enhances RT maturation, although to a lesser degree for those tested in our study. Considering this *in vitro* data, we cannot conclude which RNA mediates the RT maturation *in virio*. However, these data are consistent with previously observed changes in enzymatic activities of p66 upon tRNA interaction (197, 198). The proposed model also explains why p66/p51 formation in cells occurs efficiently within 1 h (199, 200), whereas in vitro, in the absence of nucleic acid, RT heterodimer formation takes significantly longer and with lower yield. HIV-1 is known to contain numerous copies of multiple tRNA species in addition to the essential primer tRNA (44, 177-180); thus, these other tRNA species may play a role in directing appropriate proteolytic maturation of HIV polyproteins, especially the conversion of the RT p66/p66 homodimer to the mature RT p66/p51 heterodimer.

4.5 CONCLUSIONS

Our data show that the RT p66/p66 can interact with tRNA and that this interaction facilitates selective cleavage at the p51-RNH site by HIV PR. Importantly, we also show that this selective cleavage is independent of ionic strength, but dependent on the concentration of RT p66, a factor directly related to the RT p66 subunit dissociation strength. Facilitation of the selective cleavage at the p51-RNH site by HIV-1 PR is significant even at a substoichiometric tRNA concentration. We propose a model in which interaction of the p66/p66 homodimer with tRNA introduces conformational asymmetry in the two subunits, permitting specific proteolytic processing of one of the p66 subunits leading to formation of the mature p66/p51 RT.

4.6 ACKNOWLEDGMENTS

This study was supported by grants from the National Institutes of Health (R01GM105401 to R.I. and R.L.S.; R01GM109767 to R.I.; R01AI00890 to T. I., S.G.S., and M.A.P.; P50GM103368 to J.H.E., S. G.S., and M.A.P.). We thank Michel Guerrero for his technical assistance, Michael Tsang at the University of Pittsburgh for use of Odyssey CLX gel imaging system, and Teresa Brosenitsch for reading the manuscript.

5.0 CONFORMATIONAL CHANGES IN HIV-1 REVERSE TRANSCRIPTASE WHICH FACILITATE ITS MATURATION

5.1 SYNOPSIS

HIV-1 reverse transcriptase (RT) is translated as part of the Gag-Pol polyprotein that is proteolytically processed by HIV-1 protease (PR) to finally become a mature heterodimer, composed of a p66 and a p66-derived 51-kDa subunit, p51. Our previous work suggested that $tRNA^{Lys3}$ binding to p66/p66 introduces conformational changes in the ribonuclease (RNH) domain of RT that facilitate efficient cleavage of p66 to p51 by PR. In this study, we characterized the conformational changes in the RNH domain of p66/p66 imparted by $tRNA^{Lys3}$ using NMR. Moreover, the importance of $tRNA^{Lys3}$ in RT maturation was confirmed *in cellulo* by modulating the levels of Lys-tRNA synthetase, which affects recruitment of $tRNA^{Lys3}$ to the virus. We also employed nonnucleoside RT inhibitors, to modulate the p66 dimer-monomer equilibrium and monitor the resulting structural changes. Taken together, our data provide unique insights into the conformational changes in p66/p66 that drive PR mediated cleavage.

The work presented in this chapter was reproduced with permission from: Slack RL, Ilina TV, Xi Z, Giacobbi NS, Kawai G, Parniak MA, Sarafianos SG, Sluis-Cremer N, Ishima R. (2019), Conformational Changes in HIV-1 Reverse Transcriptase that Facilitate Its Maturation. Structure, 27(10):1581-1593.e3. doi: 10.1016/j.str.2019.08.004. In this work, in-vitro proteolytic processing experiments were performed by Dr. Tatiana Ilinia, ^{1}H-^{15}N TROSY-HSQC NMR spectra we acquired by Dr. Zhaoyong Xi, and Lys-tRNA synthetase knockdown experiments were conducted by Nicholas S. Giacobbi.

5.2 INTRODUCTION

Efficient maturation of HIV-1 proteins is critical for virus replication. HIV-1 reverse transcriptase (RT) is expressed as part of the viral Gag-Pol polyprotein, which is cleaved by HIV-1 protease (PR) to finally form a mature RT heterodimer composed of 66-kDa (p66) and 51-kDa (p51) subunits (p66/p51) (**Figure 5.1A**) (9, 172). The p51 subunit is generated upon removal of most of the ribonuclease H (RNH) domain from p66 (59, 201, 202). Two models of RT maturation have been proposed: a concerted model, in which the p66 and p51 subunits are cleaved independently from Gag-Pol, and a sequential model, in which PR first cleaves p66 from the polyprotein and, following p66 dimerization, the p66/p51 RT heterodimer is formed (62, 63, 86, 87, 154, 160, 199, 200, 203, 204). Regarding these models, previous biochemical data, including ours, demonstrated that p66/p66 homodimer formation is absolutely necessary for efficient in vitro RT maturation, thus supporting the sequential model (Figure 1C) (62, 90, 91). Paradoxically, the p66/p66 homodimer adopts a symmetrical conformation in solution in which both RNH domains are folded and the p51-RNH cleavage sites are inaccessible to PR (75). Interestingly, in all structures of the mature p66/p51 heterodimer, the p51-RNH cleavage site is sequestered in a β sheet within the RNH domain and is inaccessible to PR (**Figure 5.1B**) (Davies et al., 1991, Jacobo-Molina and Arnold, 1991, Jacobo-Molina et al., 1993, Kohlstaedt et al., 1992). Consequently, the pathways involved in p66/p51 RT maturation have not been defined. However, characteristic differences between the immature p66/p66 homodimer and the mature p66/p51 heterodimer, such as a ∼10-fold decrease in the dimer dissociation constant (Sharaf et al., 2014, Sluis-Cremer et al., 2000, Venezia et al., 2006), have led to the hypothesis that significant structural differences exist between these RT proteins.

In regard to these models, prior biochemical data, including ours, demonstrated that p66/p66 homodimer formation is absolutely necessary for efficient *in vitro* RT maturation, thus supporting the sequential model (**Figure 5.1C**) (90, 91, 99). Paradoxically, the p66/p66 homodimer adopts a symmetrical conformation in solution in which both RNH domains are folded and the p51-RNH cleavage sites are inaccessible to PR (75). Interestingly, in all structures of the mature p66/p51 heterodimer, the p51-RNH cleavage site is sequestered in a β-sheet within the RNH domain and is inaccessible to PR (**Figure 5.1B**) (54, 102, 104, 105). Consequently, the pathways involved in p66/p51 RT maturation have not been defined. However, characteristic differences between the immature p66/p66 homodimer and the mature p66/p51 heterodimer, such as a ~ 10-fold decrease in the dimer dissociation constant (60, 75, 189), have led to the hypothesis that significant structural differences exist between these RT proteins *in vitro*.

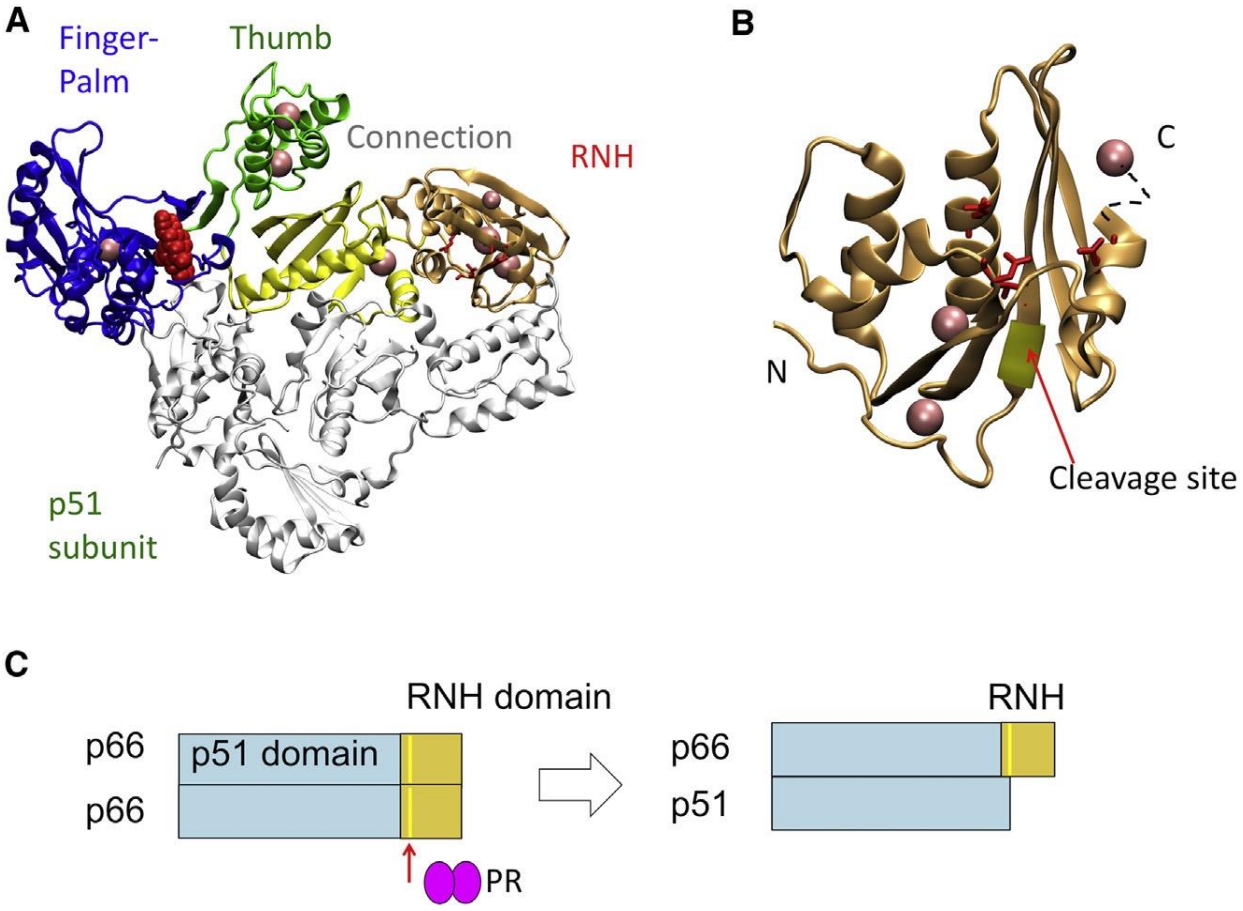

Figure 5.1 : Structure of p66/p51 HIV-1 RT

(A) Overall structure of the p66/p51 heterodimer. The fingers-palm, thumb, connection, and RNH domains in the p66 subunit are purple, green, yellow, and orange, respectively. The p51 subunit is white. (B) Structure of the RNH domain highlighting that the p51-RNH cleavage site (F440-Y441, yellow ribbon) is sequestered in the protein core. The RNH active site residues are shown by red sticks. (C) Schematic highlighting how p66/p51 is generated from p66/p66 by HIV-1 PR-mediated cleavage. In (A and B), graphics were generated using the structure of PDB: 3MEE (Lansdon et al., 2010); the location of RPV is shown by red spheres in (A); locations of the Ile-δ1 methyl groups that were uniquely observed in the NMR data are shown by pink spheres. These are residues 202 in the fingers-palm domain, 254 and 259 in the thumb domain, 393 in the connection domain, and 434, 495, and 559 in the RNH domain. Note, because crystallographic coordinates are not available for residue 559, the position of residue 559 is approximated.

Recently, we developed an *in vitro* RT maturation assay that evaluates processing of p66 by active HIV-1 PR to yield p66/p51 heterodimer, and we proposed that interaction of $tRNA^{Lys3}$ with the p66/p66 homodimer enhances specific cleavage by PR at the p51-RNH cleavage site (195). Although this study identified key factors in RT maturation including: (i) the fundamental importance of homodimer formation; (ii) an interaction between $tRNA^{Lys3}$ and p66/p66; and (iii) enhancement of p66/p51 production in the presence of $tRNA^{Lys3}$, the *conformational changes that the* p66/p66 homodimer undergoes during maturation are unknown. Another detail of the sequential model that remained unclear was whether $tRNA^{Lys3}$ enhanced p66/p51 production due to its ability to increase p66/p66 homodimer formation, or if a specific p66/p66 conformation induced by $tRNA^{Lys3}$ was required for the RT maturation. Although tRNA, especially $tRNA^{Lys3}$, is abundantly present in the virus (44, 177, 178, 180, 205), it is also unclear whether $tRNA^{Lys3}$ affects RT maturation in the virus to impact viral replication.

Herein, we present an analysis of the conformational changes of p66/p66 homodimer upon tRNALys3 interaction in solution, as well as changes in PR-mediated production of p66/p51, using NMR spectroscopy. Since our previous data suggested that p66 undergoes fast monomer-dimer equilibrium (75), and that tRNALys3 interacts with p66 monomer as well as the p66/p66 homodimer(195), we designed experiments to distinguish the homodimer interaction with tRNALys3 from that of the monomer. We achieved this by using non-nucleoside reverse transcriptase inhibitors (NNRTIs) known to interact with the p66/p66 homodimer at a 1:1 stoichiometry, to enhance p66/p66 homodimer formation, and to change the environment of the NNRTI binding pocket in p66/p66 similar to that of the p66/p51 (77, 176, 206). Importantly, using in vitro RT maturation experiments and by employing size exclusion chromatography (SEC) of the protein in various conditions, we show that the application of NNRTIs alone does not alter our

underlying premise that tRNALys3 binding to p66/p66 generates a conformational change in the homodimer that facilitates RT maturation. Notably, in HIV-1, the primer tRNA, tRNALys3, is required for reverse transcription initiation complex formation, and is recruited to the virus by interacting with lysine-tRNA synthetase (KARS) (44, 179, 207-211). Therefore, we also assessed the impact of tRNALys3 on RT maturation by KARS knockdown in HIV-1 producing cells.

5.3 MATERIAL AND METHODS

5.3.1 *In vitro* RT Maturation Experiments

RT proteins, p66/p51 and p66 alone were prepared using the p6HRT-PROT plasmid (184), as described previously (195). p66 protein was expressed in BL21 (DE3) E. coli cells and purified using a Strep-Trap HP column (GE Healthcare Lifesciences, Piscataway, NJ) and gel filtration on a Superdex 200 column (GE Healthcare, Piscataway, NJ). Purified proteins were stored in 25 mM sodium phosphate, pH 7.0, 250 mM NaCl and 50% v/v glycerol at -80°C. HIV-1 PR, clone purchased from ATUM (Newark, CA), was expressed and purified as described previously (212). Rilpivirine (RPV) and efavirenz (EFV) (Selleckchem, Houston, TX, and NIH AIDS reagent program) stock solutions, at 50 mM concentration, were prepared in 100% DMSO (76).

Proteolytic processing of p66 protein by HIV-1 PR was carried out in 20 mM sodium acetate buffer, pH 5.2, at 37°C. RPV was added to the reaction with a final DMSO concentration of 2%. All reactions that did not contain RPV included 2% DMSO for control. 4 μM of p66/p66,

calculated as a dimer, was incubated for 5 min at room temperature in four conditions: p66 alone; with 4 μM tRNALys3; with 4 μM RPV; with 4 μM tRNALys3 and 4 μM RPV. Processing was initiated by addition of HIV-1 PR to a final concentration 1 μM and incubated at 37°C. Aliquots were collected following different time intervals and quenched by the addition of Tricine sample loading buffer (Bio-Rad Laboratories, Berkeley, CA) and denatured at 95°C for 5 min. Samples were loaded onto precast 4-15% Tris-glycine gels (Bio-Rad), stained with Bio-safe Coomassie stain (Bio-Rad) and analyzed with Amersham Imager 600 (GE Healthcare Life Sciences).

5.3.2 *In Vitro* tRNA Transcription

tRNALys3 was prepared using a DNA template for tRNA transcription, which was PCR amplified using the following oligonucleotides as primers (213):

Coding strand: 5'-
GCCCGGATAGCTCAGTCGGTAGAGCATCAGACTTTTAATCTGAGGGTCCA
GGGTTCAAGTCCCTGTTCGGGCGCCA
Reverse primer: 5'-mUmGGCGCCCGAACAGGGACTTG
Forward Primer: 5'-
AATTCCTGCAGTAATACGACTCACTATAGCCCGGATAGCTCAGTCG

PCR products were purified with phenol/chloroform extraction and used in transcription reactions. In vitro transcription of tRNALys3 was performed using NTPs and T7 RNA polymerase (Thermo Fisher Scientific, USA), as described by Sherlin et al. (214), and purified by anion exchange using Hi Trap Q HP column (GE Healthcare, Piscataway, NJ), desalted with PD-10 columns (GE

Healthcare, Piscataway, NJ) and aliquoted for further use. Prior to each experiment tRNA was reannealed by heating at 95°C for 5 min followed by slow cooling to room temperature.

5.3.3 Analytical Size Exclusion Chromatography to Monitor p66/p66-tRNA Interaction

Size Exclusion Chromatography (SEC) experiments were performed using a 24-ml analytical Superdex 200 Increase 10/300 GL column (GE Healthcare), at room temperature at a flow rate of 0.5 mL/min. Either p66 protein at 20 – 46 μM, with 0 - 25 μM tRNALys3 or the tRNALys3 alone that contains tracer tRNA 3'-end labeled with pCp-Cy3 (Jena Bioscience, Jena, Germany) was prepared in 25 mM Bis-tris buffer, pH 7.0, containing 100 mM NaCl, 1% DMSO, and 0.02% sodium azide. RPV was added at [p66:p66]:[RPV] = 1: 1.3 or 1.5 ratio. Each injection volume was 50 μL. Elution profiles were monitored by UV absorbance at 254 and 280 nm and, for the samples with tRNALys3, additionally by in-line Shimadzu RF-10AXL Fluorescence Detector with fluorescence excitation at 485 nm and the emission at 560 nm.

5.3.4 Sample Preparation

We used the same coding sequence of the RT p66 as described previously (75), except (i) a V559I polymorphism mutation was included to increase the number of Ile NMR resonances in ^{1}H-^{13}C NMR spectra and (ii) an N-terminal His_6-fusion tag containing a TEV-protease cleavage site was added. [U-^{2}H] and Ile $^{\delta 1}$-[$^{13}CH_3$] labelled p66 and [U-^{2}H,^{15}N] labelled p66 was expressed using a published protocol (74). In brief, isotopes were purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA) or MilliporeSigma (St. Louis, MO). Proteins were expressed in Rosetta 2(DE3) cells and were purified using HisTrap HP columns (GE Healthcare, Piscataway, NJ) and

gel filtration on a Superdex 200 column (GE Healthcare, Piscataway, NJ). The N-terminal fusion tag was digested with His_6-TEV-protease. The p66 was separated from the remaining digestion products using a HisTrap HP column (GE Healthcare, Piscataway, NJ), followed by a final purification step on a Superdex 200 column (GE Healthcare, Piscataway, NJ). Purified proteins were exchanged to a buffer containing 50 mM Tris, 250 mM NaCl, 0.02% NaN_3 and 50% v/v Glycerol, pH 8.0 and stored at -80°C.

5.3.5 NMR Experiments

All the NMR experiments were recorded on a Bruker Avance 900 spectrometer. Prior to NMR, the buffers of all proteins were exchanged to a deuterated and non-deuterated, the latter containing 5% D_2O, 25 mM Bis-Tris buffer pH 7.1 containing 100 mM KCl, 0.02% NaN_3, and 5% v/v Glycerol-d8 (named NMR buffer hereafter) for [U-^{2}H] and Ile $^{\delta1}$-[$^{13}CH_3$] labelled p66/p66, and for [U-^{2}H,^{15}N] labelled p66/p66, respectively. NMR experiments in the presence of RPV were performed by adding 0.5-1.0% d6-DMSO in the NMR buffer.

The time-dependent spectral changes of 35 μM, [U-^{2}H] and Ile $^{\delta1}$-[$^{13}CH_3$] labelled p66/p66 were monitored by recording ^{1}H-^{13}C SOFAST-HMQC spectra using Bruker sequence, sf_metrosygpph, at time points 0, 4.85, 15.29, 20.12, 24.97, 29.80, 37.27, 42.42, 47.27, 52.10 and 56.95 hours at 35°C. The time point indicates the starting time of each HSQC spectrum which took 4.85 hours to complete. Because over 90% of p66 is expected to form a homodimer (60, 75), we simplify the expression as p66/p66, to complement the p66/p51 heterodimer description. The NMR spectra of the Ile $^{\delta1}$-labelled p66/p66 were compared to those of (1) p66/p51; and (2) of partially matured RT *in vitro*. For (1), p66/p51 protein was prepared by using [U-^{2}H], Ile $^{\delta1}$-[$^{13}CH_3$]-labelled p66 and unlabelled p51. Both p66/p66 and p66/p51 spectra were recorded using

the NMR buffer, but without 100 mM KCl, at 35°C. For (2), partial maturation was achieved by adding 3 μM PR and ~20 μM $tRNA^{Lys3}$ to the 25 μM [U-^{2}H], Ile $^{\delta1}$-[$^{13}CH_3$]-labelled p66/p66 solution in the NMR buffer, incubating at 35°C for 3 hours, followed by addition of the PR inhibitor, darunavir (obtained from NIH AIDS reagent program), to stop the reaction.

Spectral changes of p66/p66 upon interaction with tRNALys3 or RPV and with both $tRNA^{Lys3}$ and RPV were monitored by recording ^{1}H-^{13}C SOFAST-HMQC spectra of a ~25 μM [U-^{2}H], Ile $^{\delta1}$-[$^{13}CH_3$]-labelled p66/p66 in the deuterated NMR buffer at the anticipated ratios of $tRNA^{Lys3}$/[p66/p66] = 1.4, RPV/[p66/p66] = 2.0, and $tRNA^{Lys3}$/[p66/p66-RPV] = 1.0, at 35°C. Titrations of p66/p66 with RPV and with $tRNA^{Lys3}$ in the presence of RPV were monitored by recording ^{1}H-^{13}C SOFAST-HMQC spectra of the Ile $^{\delta1}$-labelled p66/p66, at 35°C, at relative concentrations of [RPV]:[p66/p66] of 0:1, 0.5:1, 1:1, and 2:1 and at [tRNA]:[p66/p66-RPV] of 0:1,1:1, 2:1. These p66/p66 spectra, in the presence of $tRNA^{Lys3}$, RPV, or both, were also recorded at 20°C.

Spectral changes of p66/p66 upon interaction with these molecules were also monitored by recording ^{1}H-^{15}N TROSY-HSQC spectra of 75 μM [U-^{2}H, ^{15}N]-labelled p66/p66 in the protonated NMR buffer at anticipated ratios of $tRNA^{Lys3}$/[p66/p66] = 1.0, RPV/[p66/p66] = 1.3, and $tRNA^{Lys3}$/[p66/p66-RPV] = 1.0, at 20°C. All the NMR spectra were processed using nmrPipe and analyzed using nmrDraw, nmrView or ccpNMR (112, 113, 215).

5.3.6 KARS Knockdown Experiments

Small interfering RNAs (siRNAs) targeting KARS, as well as a control scrambled sequence control siRNA, were purchased from Sigma (St Louis, MO, USA). 293T cells were transfected with 80-nmol/L siRNA using the Neon Transfection System (Thermo Fisher Scientific, USA),

according to the manufacturer's protocol. The efficiency of gene knockdown was assessed by western blot analyses of protein expression. Anti-KARS antibodies were also purchased from Sigma. HEK 293T cells (ATCC® CRL-3216™) were transfected with HIV-1^{LAI} (216) and RT and p24 antigen expression levels were measured by Western Blot (90). Viral infectivity was assessed using TZM-bl cells (217).

5.4 RESULTS

5.4.1 NMR Spectra of the p66/p66 Homodimer

To gain insight into the conformational changes in p66/p66 that facilitate in vitro RT maturation, we monitored the ^{1}H-^{13}C HMQC spectral features of [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] p66 protein over time in solution at 35°C (Figures 5.2A and 5.2B). As described in the Methods, we used four columns to purify p66 for NMR, to avoid contamination by E. coli proteases that can process p66 to p51 (Bavand et al., 1993, Clark et al., 1995, Lowe et al., 1988)(66, 218, 219). These experiments were performed using a concentration of 35 μM (as p66/p66), a concentration at which we would expect 80% of the protein to exist as a homodimer. The overall HMQC spectral features of the protein did not change over 50 h (Figure S1), consistent with our previous observation using ^{1}H-^{15}N NMR of [U-^{2}H,^{15}N] p66 (75). A small reduction in the NMR signal intensity (~15%) was observed, presumably due to instability of the magnet or protein solution (Figure 5.2C); a gain in signal intensity would be expected if protein unfolding occurred. The spectrum shows a set of protein resonances that are indicative of a symmetric homodimeric form of p66/p66, consistent with

previously published p66/p66 spectra, albeit lacking evidence of previously described slow conformational changes that had occurred over 40 h (86, 87, 220). In addition, as previously observed, several resonances overlapped in the central region of the spectrum (dashed rectangle in Figure 5.2A) even though the Ile $^{\delta 1}$-methyl-labeled residues are distributed across the p66 domains (highlighted by pink spheres in Figure 5.1A) (86, 87, 220).

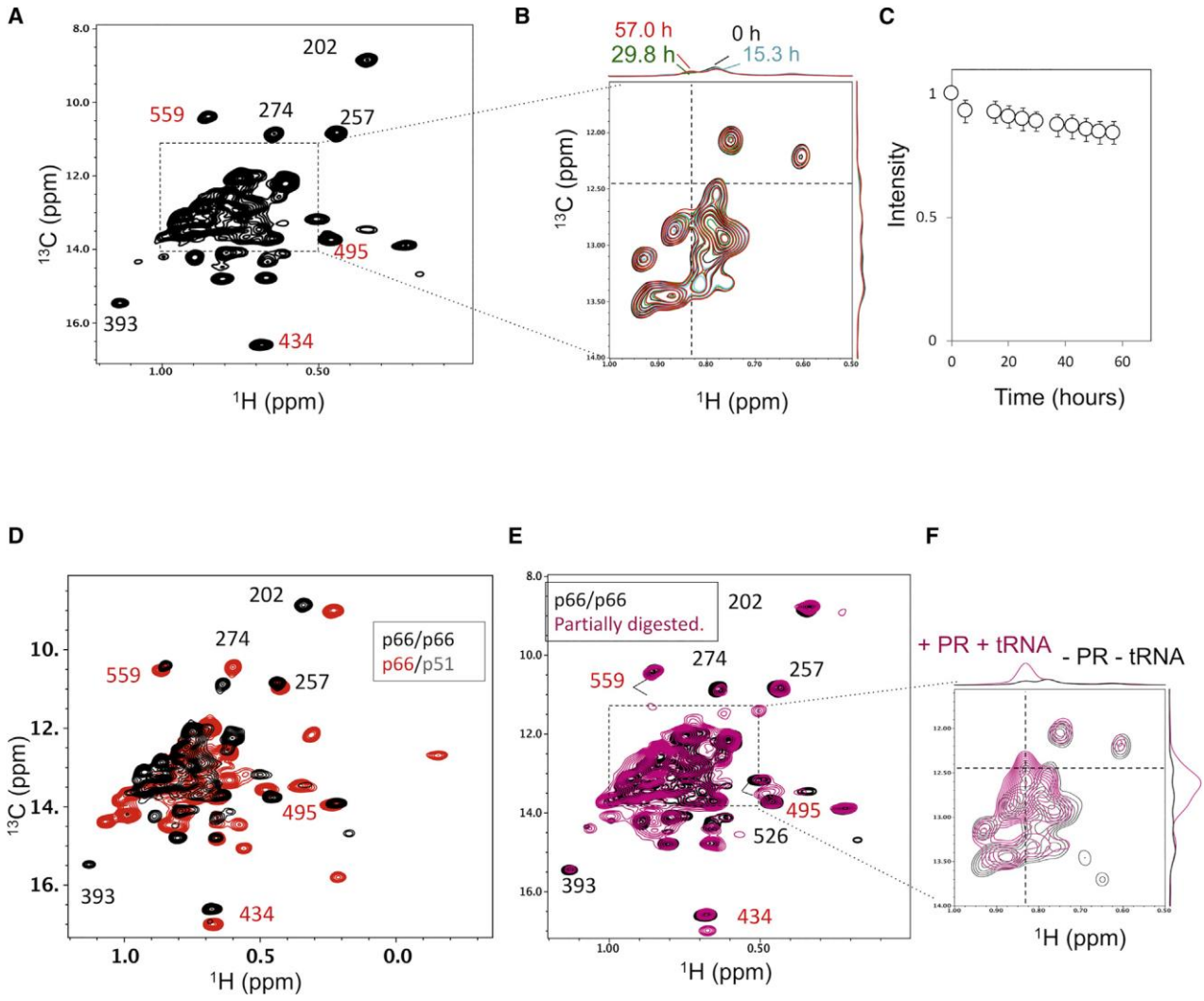

Figure 5.2 ^{1}H-^{13}C SO-FAST TROSY-HMQC NMR Spectra of [U-^{2}H], Ile $^{\delta1}$-[$^{13}CH_3$]-labelled p66/p66 at 35°C (A) p66/p66 spectrum immediately following NMR sample preparation; (B) the random coil region of the spectrum shown in (A) is overlaid with NMR spectra obtained for the same sample at 2.9, 25, and 57 h; (C) a plot of the average intensity decay of 42 resonances collected from p66/p66 over 57 h; (D) comparison of the p66/p66 spectrum (black) with that of a p66/p51 sample (red) in which only the p66 subunit was labeled with [U-^{2}H], Ile $^{\delta1}$-[$^{13}CH_3$]; (E) comparison of the p66/p66 spectrum (black) with that after 3 h digestion of the NMR sample by active PR in the presence of unlabelled tRNA^{Lys3} (pink); (F) the indicated area of the spectrum in (E) is shown at high threshold level. Residue numbers (those in the RNH domain are colored red), except for I559, were based on previous literature (86). In (B and F), slices taken at the dashed lines, in which cross-section, ^{1}H 0.83 ppm and ^{13}C 12.6 ppm, is nearly at the random coil position of the Ile-$^{\delta1}$ methyl group (221), are plotted along the outer edge of the spectrum. Note, since

approximately 80% of the p66 forms a homodimer at the p66 concentration used in this study, we use the notation of p66/p66, to compare with the p66/p51. See also Figure C1.

The observed spectral features of p66/p66 homodimer were compared with those of the p66/p51 heterodimer (Figure 2D) and with a partially matured sample in which p66/p66 was incubated with HIV-1 PR in the presence of $tRNA^{Lys3}$ (Figure 5.2E). Consistent with previously published data (86), the p66/p66 NMR spectrum was distinct compared with that of p66/p51. When PR was added to p66/p66 in the presence of $tRNA^{Lys3}$, we observed an increase in the signal intensity at the random coil position (Figure 5.2F) and a spectral pattern similar to that of p66/p51 (Figure 5.2E). Because the cleavage of p66/p66 to p66/p51 was incomplete, spectral patterns for both p66/p66 and p66/p51 can be observed in Figure 5.2E. Overall, the observed spectral features of p66/p66 are distinct compared with those of p66/p51 and are suggestive of a symmetric homodimer conformation.

5.4.2 NMR Spectra of p66 in the Presence of $tRNA^{Lys3}$

We hypothesized that $tRNA^{Lys3}$ interaction with p66/p66 introduces conformational changes in a single RNH domain that facilitate efficient cleavage of p66 to p51 by HIV-1 PR (195). To address the hypothesis, we monitored changes in the ^{1}H-^{13}C HMQC spectrum of [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] p66 protein in the absence and presence of unlabeled $tRNA^{Lys3}$ at a [$tRNA^{Lys3}$]:[p66/p66] = 1.4:1 molar ratio. We observed that upon addition of $tRNA^{Lys3}$, the p66 spectrum, which shows only p66 signals and not $tRNA^{Lys3}$, exhibited a slight increase in the signal intensity, nearly at the random

coil position of the Ile $^{\delta 1}$-methyl group at ^{1}H 0.83 ppm and ^{13}C 12.6 ppm (221) (cross-section of the dashed lines in Figure 5.3A), suggesting partial unfolding of the protein, although many of the resonance positions did not change (Figure 5.3A, discussed below).

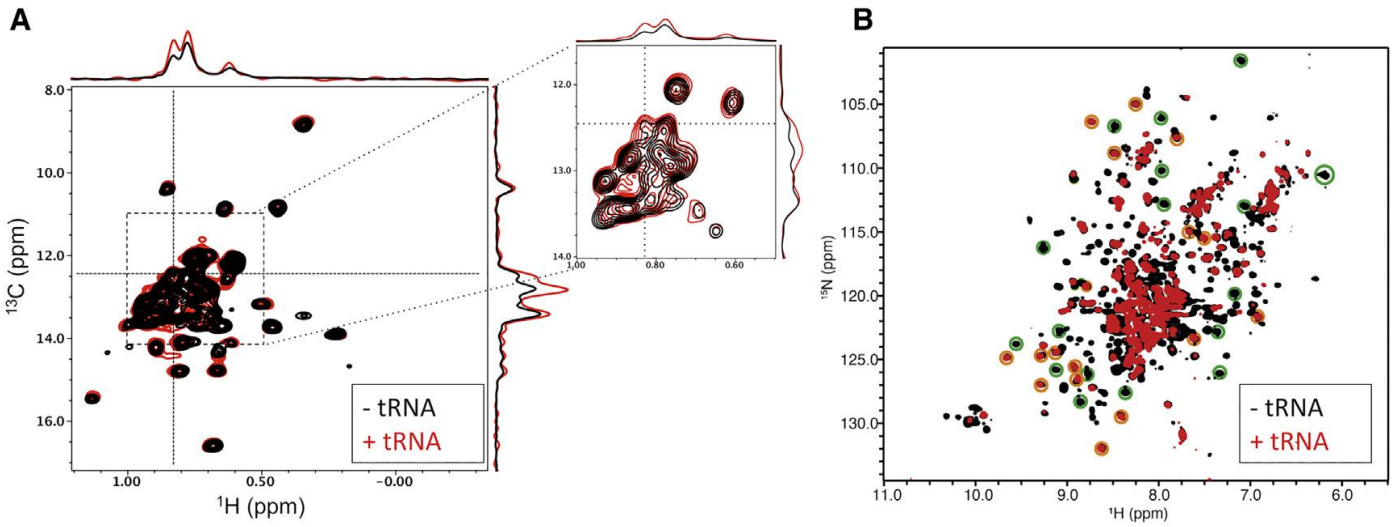

Figure 5.3 NMR Spectra of p66 in the Presence of tRNA^{Lys3}

(A) ^{1}H-^{13}C SOFAST-HMQC NMR spectra of [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] p66/p66 recorded at 35°C and (B) ^{1}H-^{15}N TROSY-HSQC NMR spectra of [U-^{2}H, ^{15}N] p66/p66, in the absence (black) or presence (red) of unlabeled tRNA^{Lys3}, recorded at 20°C. In (A), a selected region (box) is shown at high threshold level with slices taken at the dashed lines, the cross-section of which is nearly at the random coil position. In (B), resonances that were previously found to overlap with the isolated thumb and RNH domains are circled by green and orange colors, respectively (75). Note, a lower temperature was used for the ^{1}H-^{15}N experiments compared with the temperature used for the ^{1}H-^{13}C experiments, presuming greater protein stability at a lower temperature.

Because methyl signal intensities, as peak height, are mainly determined by a fast methyl three-site jump (222), Ile $^{\delta 1}$-methyl ^{1}H-^{13}C spectra are less sensitive to reductions in domain motion compared with backbone ^{1}H-^{15}N amide NMR spectra, which are more sensitive to changes in molecular tumbling, domain motion, and internal motion. To further elucidate the conformational changes of p66/p66 imparted by tRNA^{Lys3} binding, we also recorded ^{1}H-^{15}N TROSY-HSQC spectra of [U-^{2}H,^{15}N] p66 in the absence and presence of tRNA^{Lys3} at a [tRNA^{Lys3}]:[p66/p66] = 1:1 molar ratio (Figure 3B). At 75 μM, p66/p66 contains ~90% dimer, with fast exchange occurring between monomer and dimer. The ^{1}H-^{15}N TROSY-HSQC spectrum exhibits a single set of clearly identifiable signals stemming from the RNH and thumb domains (circled orange and green, respectively, in Figure 5.3B) based on previous assignments, suggesting symmetrical p66/p66 conformation in solution (75). As reported previously (75), the resonance positions of the thumb and RNH domains within p66 spectra were highly similar to those of the isolated domains, and also to those in p51. Reported dissociation constants indicate that the p66 homodimer has a 10-fold higher affinity than that of the p51 homodimer at equilibrium (60, 75). Based on these observations, and the assumption that resonances observed for the p66 dimer/monomer equilibrium were in the fast exchange regime, we derived a model in which the thumb and RNH domains undergo domain motion, allowing for the observed resonance similarity with respect to spectra of the isolated domains (75). Indeed, the observation of one set of resonances in the Ile $^{\delta 1}$-methyl ^{1}H-^{13}C spectra is consistent with the ^{1}H-^{15}N data (Figure 5.3).

Upon tRNA^{Lys3} addition, many resonances in the ^{1}H-^{15}N NMR spectrum of p66/p66 exhibited a significant reduction in intensity. In particular, signals from the thumb domain significantly decreased, undergoing line broadening or disappearance upon interaction with tRNA^{Lys3}, while resonances from the RNH domain remained mostly unchanged. This reduction in

the signal intensity of the thumb domain resonances is reasonable in the sense that nucleic acid binding would reduce the domain motion of the thumb, resulting in a decrease in signal intensities. It is also consistent with existing structural data, which clearly show that the canonical nucleic acid binding site in RT involves extensive contacts with the p66 thumb and fingers-palm subdomains (105, 223). On the other hand, if the RNH domain were in an equilibrium between either rigid and mobile domain states or folded and unfolded states, then the signals would be expected to broaden as a result of the exchange equilibrium. Thus, our observation of the RNH resonances may signify that one RNH domain remains mobile in the $tRNA^{Lys3}$-bound form of p66/p66 (discussed below).

5.4.3 NNRTI Minimizes p66 Monomer Interaction with $tRNA^{Lys3}$

As previously mentioned, the p66/p66 sample contains both monomer and dimer species in equilibrium. Because NMR resonance intensities are inversely proportional to the hydrodynamic radius of macromolecules in solution, even small amounts of monomer-bound $tRNA^{Lys3}$ could potentially complicate our interpretation of NMR data. NNRTIs have been shown to promote homodimerization of the polyprotein Pol and p66 in cells and in vitro(77, 176, 206). We therefore hypothesized that the inclusion of an NNRTI could be used to reduce the monomeric p66 species within our NMR sample. To confirm that NNRTIs are useful to reduce the monomer component of our p66/p66 samples, we first performed analytical SEC experiments with p66 in the absence or presence of NNRTI and/or $tRNA^{Lys3}$. As previously reported, the SEC elution profile of p66 protein alone showed both monomer and dimer elution peaks with a UV254/UV280 ratio of ~0.5, while the SEC profile of $tRNA^{Lys3}$, also monitored by fluorescence, showed a single elution peak with a UV254/UV280 ratio of ~2 (Figures 5.4A and 5.4B). In the presence of a small amount of $tRNA^{Lys3}$ ([p66/p66]:[tRNALys3] = 1:0.22 molar ratio), the elution peak of the dimer shifted to a

larger molecular mass, presumably a tRNA-bound form (Figure 5.4C). However, with excess $tRNA^{Lys3}$ ([p66/p66]:[tRNALys3] = 1:1.25), a new elution peak, located between the monomer and dimer peaks and presumably tRNA-bound p66 monomer, appeared (Figure 5.4D). Indeed, as described previously, even in the presence of a small amount of $tRNA^{Lys3}$ (Figure 5.4C), the maximum position of the p66 monomer at UV254 slightly shifted from that of UV280, suggesting the existence of a $tRNA^{Lys3}$-p66 monomer-bound form, the retention of which may depend on the rate of exchange between the monomer and dimer fractions, or of the complexes.

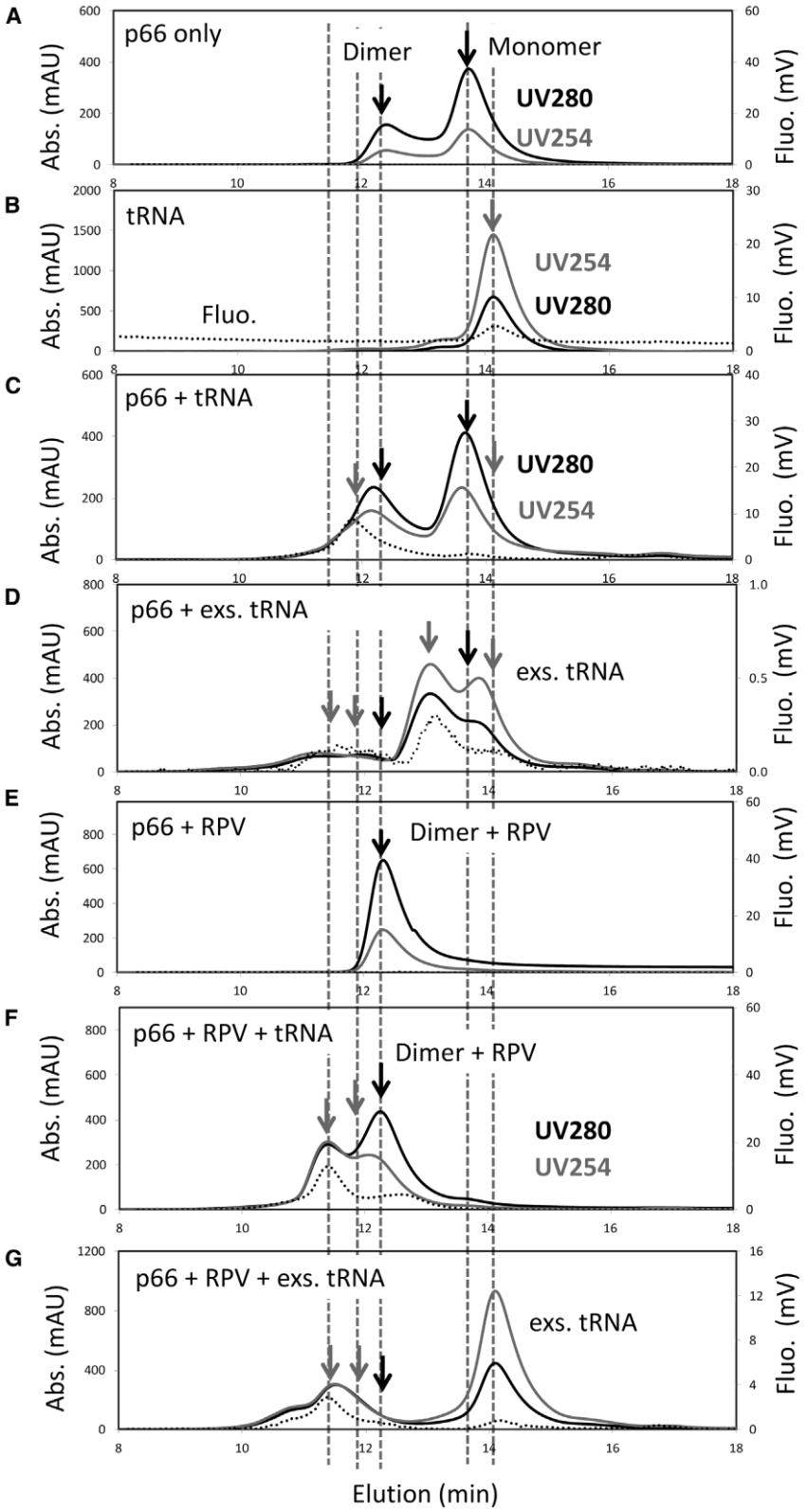

Figure 5.4 SEC Elution Profiles of p66

(A) p66 only, (B) $tRNA^{Lys3}$ only, (C) p66 and $tRNA^{Lys3}$ at [p66/p66]:[$tRNA^{Lys3}$] = 1:0.22, (D) p66 and $tRNA^{Lys3}$ at [p66/p66]:[$tRNA^{Lys3}$] = 1:1.25, (E) p66 and RPV at [p66:p66]:[RPV] = 1:1.3 ratio, (F) p66, RPV, and $tRNA^{Lys3}$ at [p66/p66]:[RPV]:[$tRNA^{Lys3}$] = 1:1.3:0.22, and (G) p66, RPV, and $tRNA^{Lys3}$ at [p66/p66]:[RPV]:[$tRNA^{Lys3}$] = 1:1.5:2.0. Elution profiles were monitored by UV absorbance at 280 nm (black line) and 254 nm (gray line), and by fluorescence detection for the labeled $tRNA^{Lys3}$ (dotted line). Black and gray arrows indicate protein alone elution peaks and those containing $tRNA^{Lys3}$, respectively. Note, the molar extinction coefficient of $tRNA^{Lys3}$ at 254 and 280 nm are 10.2 and 1.8 times those of p66, respectively, and in (D) the elution peaks of free $tRNA^{Lys3}$ and monomer p66 partially overlap with that of the $tRNA^{Lys3}$-bound monomer p66, which we estimate to be 20%–40% of total p66.

The NNRTI rilpivirine (RPV) is known to enhance p66 homodimer formation with an apparent RPV-p66/p66 dissociation constant of 0.86 ± 0.064 μM (60, 77, 176, 206, 224, 225). Thus, not surprisingly, in the presence of RPV, the SEC profile of p66/p66 exhibited only a homodimer fraction, and a monomer elution peak was not detected (Figure 5.4E). RPV is known to bind to p66/p66 with a 1:1 ratio (76, 77). Incubation of RPV-bound p66/p66 with a small amount of $tRNA^{Lys3}$ did not produce the $tRNA^{Lys3}$-p66 monomer peak that was seen in the absence of RPV, but, instead, produced a single elution peak, earlier than the p66/p66 homodimer, which is presumably $tRNA^{Lys3}$-bound p66/p66 and may include its oligomer in an exchange equilibrium (Figure 5.4F). Even with 2-fold excess tRNALys3, the $tRNA^{Lys3}$-bound p66 monomer form was not observed in the presence of RPV (Figure 5.4G).

5.4.4 Effect of NNRTIs on the Maturation of HIV-1 RT In Vitro

Our SEC data clearly indicate that an NNRTI can suppress the amount of p66 monomers in our p66/p66 samples. To further characterize how NNRTI-mediated reduction in the p66 monomer component modulates RT maturation, we conducted in vitro RT maturation assays. In these experiments, purified p66 is incubated with HIV-1 PR and the cleavage of p66 to p51 is monitored by SDS-PAGE (195), and generation of equivalent amounts of p66 and p51 is indicative of p66/p51 heterodimer production. Incubation of p66 alone with PR does not result in significant p66/p51 heterodimer formation (Figure 5.5A), whereas addition of $tRNA^{Lys3}$ does (Figure 5.5B) (195). Of note, in the absence of NNRTI, p66 exists in monomer-dimer equilibrium with a 4–10 μM dissociation constant for the homodimer (60, 75, 189), similar to the concentration of p66 used in these experiments (4 μM as p66/p66). Previously, we have also shown that heterodimer production is more efficient at higher protein concentration, i.e., homodimer formation is necessary (195).

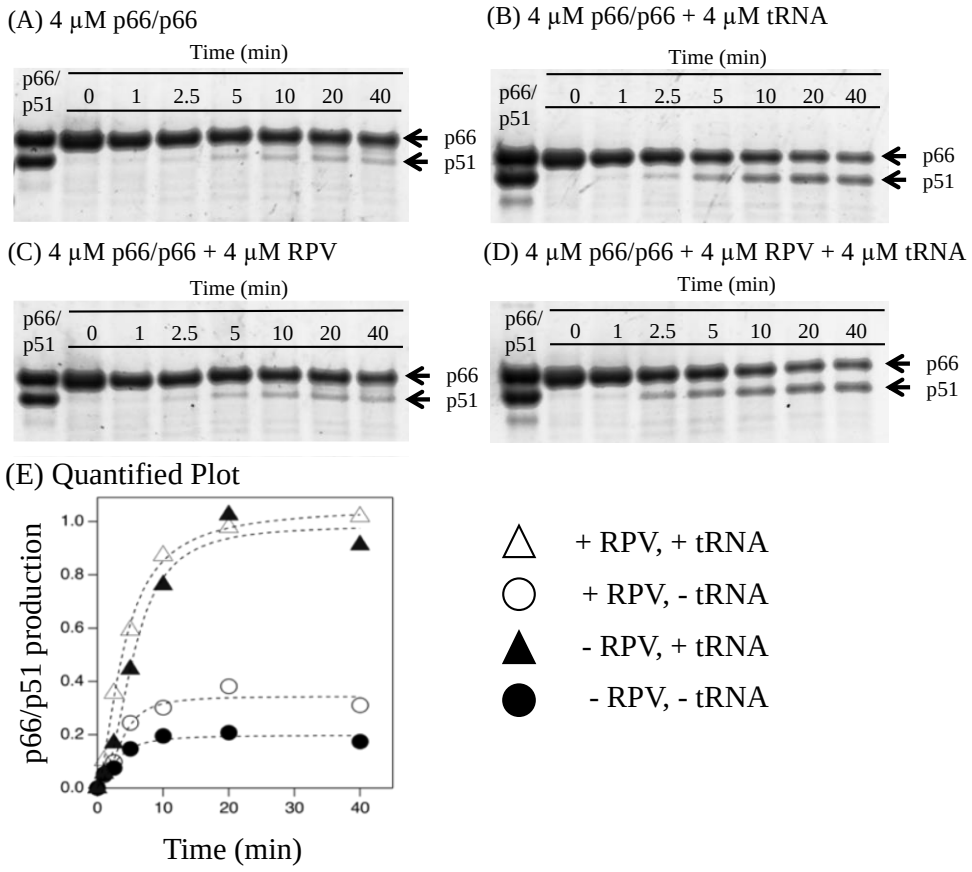

Figure 5.5 : Time-dependent Proteolytic Cleavage of p66 by HIV-1 PR Monitored by SDS-PAGE Cleavage experiments were conducted in the absence (A and B) or presence (C and D) of RPV, and presence (B and D) or absence (A and C) of tRNALys3. In (E), p51 band intensities shown in (A–D) were quantified and plotted. Concentrations of p66 and PR were 4 μM, as p66/p66 homodimer, and 1 μM, respectively. Both $tRNA^{Lys3}$ and RPV concentrations were 4 μM.

When the monomer component of the p66 sample was suppressed by addition of 4 μM RPV (a similar molar ratio was used for the SEC experiments), we found that PR-mediated processing of p66 was mostly unchanged (Figures 5.5C and C2). Neither varying the RPV

concentration (Figure C2), nor using efavirenz, which belongs to a different NNRTI class (Figure C3), altered the p66 processing kinetics. In contrast, addition of $tRNA^{Lys3}$ to p66/p66 in the presence of RPV promoted efficient RT maturation, which was similar to that observed in the presence of $tRNA^{Lys3}$ alone (Figures 5.5D and 5.5E). Collectively, these data show that NNRTIs, which induce p66/p66 homodimer formation, have minimal impact on p66 processing, suggesting that p66/p66 homodimer formation alone is not sufficient to drive proteolytic processing of p66 to p51, and the presence of $tRNA^{Lys3}$ introduces a change that allows efficient processing.

5.4.5 Probing RPV-Bound p66/p66 Conformation in the Absence and Presence of $tRNA^{Lys3}$

To further probe the RPV-induced conformational changes in p66/p66, we first titrated [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] p66/p66 with increasing concentrations of RPV and monitored chemical shifts of the protein by ^{1}H-^{13}C SOFAST-HMQC experiments. We found that some signal intensities in the RNH domain decreased and new signals appeared (Figure 5.6A). The most salient example is residue I434. In the apo form, residue 434 appeared as a single isolated resonance near 16 ppm 13C chemical shift (bottom of the spectrum); the intensity of this resonance (labeled A) decreased while a new resonance, B, appeared with increasing RPV concentration (Figure 5.6A). An analysis of the intensity change indicates a binding ratio [RPV]:[p66/p66] of 1:1 (Figure 5.6A, graph). This observation of two sets of signals in RPV-bound p66/p66 suggests introduction of conformational asymmetry in the RNH domain of p66/p66 (or rigorously speaking, chemical shift), with two folded RNH domains that have different resonance positions relative to each other. Despite this asymmetry, introduced by inhibitor binding, in vitro RT maturation of RPV-p66/p66 was similar

to p66/p66 in the absence of RPV, with low efficiency (Figure 5.5). This suggests that conformational asymmetry alone is not sufficient to facilitate efficient RT maturation.

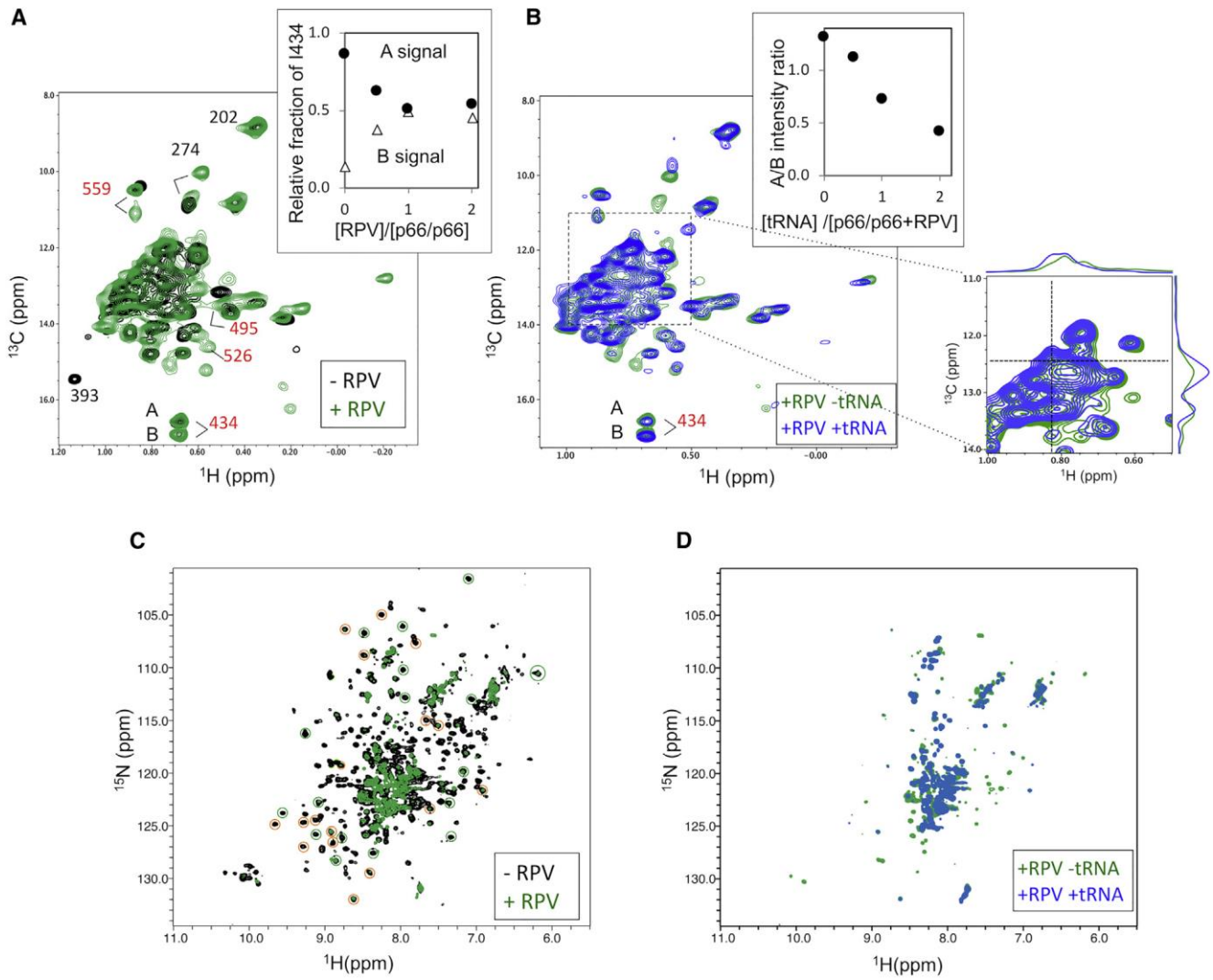

Figure 5.6 ^{1}H-^{13}C SOFAST-HMQC NMR Spectra of [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] p66/p66 HMQC spectra (A) in the absence (black) or presence (green) of RPV, and (B) in the presence of RPV (green) or RPV plus tRNA^{Lys3} (blue), recorded at 35°C, and ^{1}H-^{15}N TROSY-HSQC NMR spectra of [U-^{2}H, ^{15}N] p66/p66 (C) in the absence (black) or presence (green) of RPV, and (D) in the presence of RPV (green) or RPV plus tRNA^{Lys3} (blue), recorded at 20°C. In (A and B), inset graphs show relative intensity changes of residue 434 resonances (A and B) at different [RPV]:[p66/p66] or [tRNA]:[p66/p66-RPV] ratios. In (B), a selected region (box) is shown at high threshold level with slices taken at the dashed lines, the cross-section of which is nearly at a random coil position. In (C), resonances that overlap with previously assigned resonances of the isolated thumb and RNH domains are circled by green and orange colors, respectively (75).

Next, we titrated p66/p66-RPV with $tRNA^{Lys3}$, and monitored the p66 signals by ^{1}H-^{13}C SOFAST-HMQC experiments, to gain further insight into the conformational changes that facilitate RT maturation (Figure 5.6B). We found that resonance A of residue 434 decreased in signal intensity while resonance B remained stable, indicating that $tRNA^{Lys3}$ influences the A RNH domain in p66/p66 homodimer more than the B RNH domain (Figure 5.6B, described below). The observed change in the chemical shift of resonance A suggests either conformational change of a region that includes I434 or changes of the chemical environment surrounding I434, presumably by domain orientation changes or by interaction with $tRNA^{Lys3}$. Consistent with $tRNA^{Lys3}$ titration into p66/p66 alone (Figure 5.3A), $tRNA^{Lys3}$titration into p66/p66-RPV produced a slight increase in the resonances located in the random coiled region (Figure 5.6B, side panel)

In contrast to the ^{1}H-^{13}C SOFAST-HMQC experiments of Ile $^{\delta 1}$-methyl groups, ^{1}H-^{15}N TROSY-HSQC did not show clear changes in signal positions for the amide backbone signals of p66/p66 upon RPV interaction; instead, only a reduction of amide backbone signal intensity was observed, including some of the thumb and RNH domain resonances (Figure 5.6C). This is consistent with the fact that the NNRTI-bound p66/p66 conformation is similar to that of the NNRTI-bound heterodimer (77), including possibly reduced RNH and thumb domain mobility, and that NNRTI rigidifies the thumb conformation (69, 226, 227). $tRNA^{Lys3}$ interaction with p66/p66-RPV further reduced molecular tumbling (Figure 5.6D), presumably due to dimer-oligomer equilibrium as seen in the SEC (Figures 5.4F and 5.4G).

In Figure 5.7, we summarize the observed NMR spectral changes, recorded at 35°C, for three different residues in the RNH domain of RT: I434, I495, and I559. As described above, I434 exhibits a single signal in the spectrum of p66/p66 (resonance A; Figure 5.7A, top), under experimental conditions in which ~80% of the p66 exists as a homodimer (60, 61, 75, 189). Upon

RPV binding, a second I434 signal is observed (resonance B; Figure 5.7B, top). Addition of $tRNA^{Lys3}$ eliminates resonance A and slightly changes the position of resonance B (resonance B′; Figure 5.7C, top). This resonance B′ is similar to that observed in the partially processed RT, i.e., containing both p66/p66 and p66/p51 (Figure 5.7D, top). Although we do not know where the A resonance moved, we note that the resonance at the random coil position was increased in the presence of $tRNA^{Lys3}$. Thus, it is possible that the region including I434 in subunit A (light yellow RNH domain in the cartoon of Figure 5.7C) is unfolded. Importantly, resonance A is absent in the p66/p51 spectrum, while resonance B′ is clearly present (Figure 5.7E, top). Similar spectral changes were also observed at RNH domain residues I495 and I559 (Figure 5.7, the second and third rows). However, the chemical resonance at I495 in the p66/p51 heterodimer does not line up with the same chemical shift as the partially digested p66 (Figures 5.7D and 5.7E, the second row), suggesting that this residue may undergo internal dynamics. Data recorded at 20°C indicate similar tendencies but exhibit more conformers, presumably due to slower rates of exchange (Figure C4).

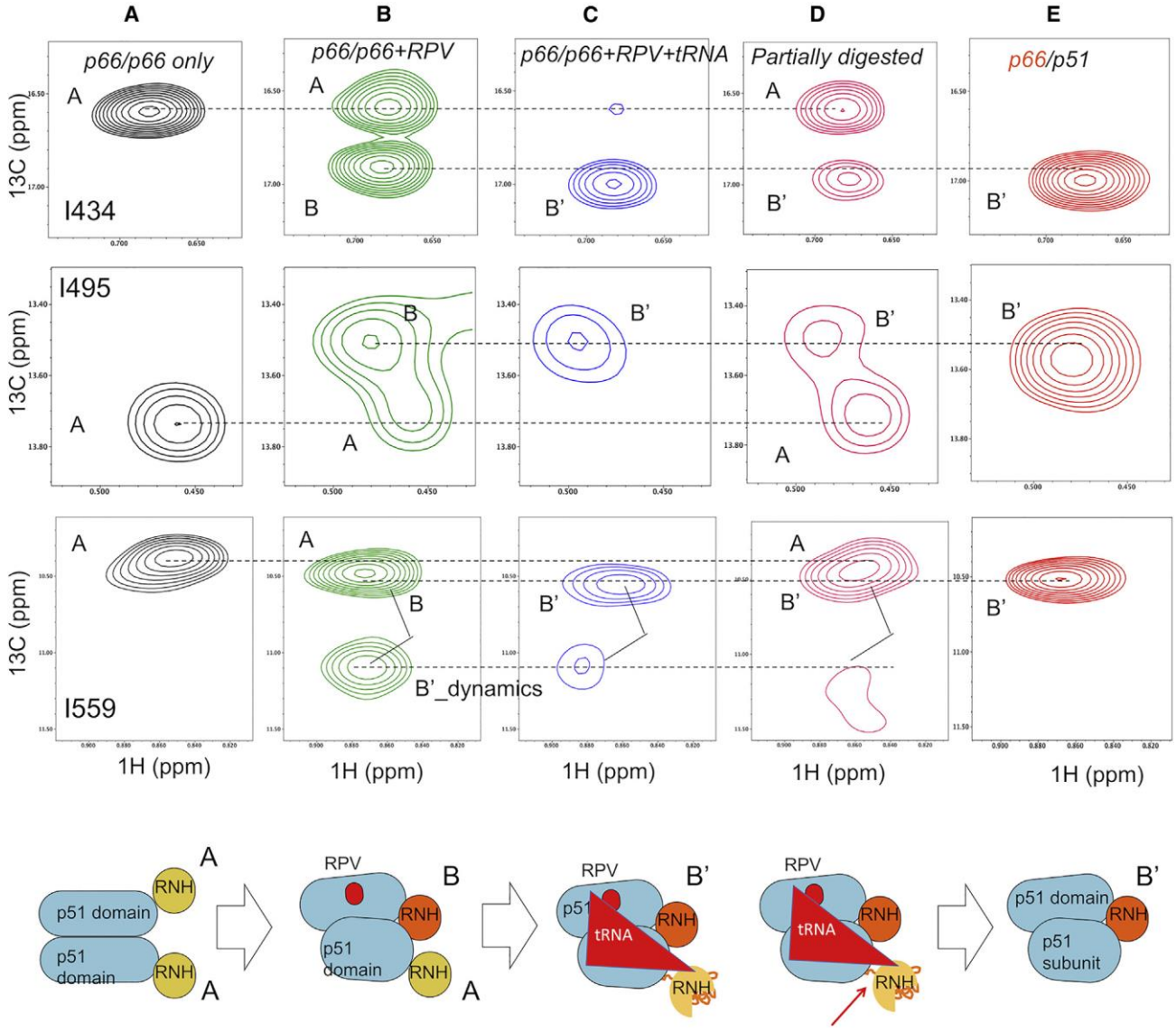

Figure 5.7 Overview of the Observed Signal Patterns of p66/p66 RNH Domain Residues I434, I495, and I559 (A) p66/p66 only, (B) p66/p66 + RPV, (C) p66/p66 + RPV + tRNA, (D) partially digested p66/p66 sample, and (E) p66/p51 in which only the p66 subunit is [U-^{2}H] and Ile δ1-[$^{13}CH_3$] labeled. All the spectra were recorded at 35°C. Cartoon at the bottom indicates conformational changes deduced from the observed spectra in each condition. The designation A in the spectra indicates NMR resonance positions stemming from the p66/p66 homodimer, while B and B′ indicate newly generated resonance positions upon RPV interaction and partial digestion of p66/p66, respectively.

Based on these spectral changes and previous ^{19}F NMR that monitored residue 181 located at the NNRTI binding pocket of p66/p51 and p66/p66 (76, 77), the following scenario for p66/p66 conformational changes is derived (cartoons at the bottom of Figures 5.7A–5.7E). (1) p66/p66 is in equilibrium with p66 monomer and exhibits one set of stable signals, with the two RNH domains in the p66/p66 homodimer folded and symmetrical. (2) RPV binding induces some asymmetry, or conformational change, in one RNH domain to create an environment similar to that of the RNH domain in p66/p51, such that the p51-RNH site is protected (orange in the cartoon of Figure 5.7C). However, this conformational change is not sufficient to drive proteolytic processing. (3) tRNA interaction affects the A peak within the RNH domain (yellow in the cartoon of Figure 5.7C), which is the subunit that is cleaved by PR (note the reduced intensity of peak A in Figures 5.7D compared with 5.7A). Overall, $tRNA^{Lys3}$ generates partial unfolding of the protein, presumably of the RNH domain region, in the presence and absence of RPV (Figure 5.3).

5.4.6 Knockdown of KARS in 293T Cells Affects Intracellular RT Processing and Reduces Virus Particle Production

Collectively, our data underscore that tRNALys3 binding to the p66/p66 RT homodimer triggers the necessary conformational changes that facilitate PR access to the cleavage site. During the HIV-1 life cycle, tRNALys3 is essential as a primer for reverse transcription reaction and is recruited into the budding virus through its interaction with KARS and Gag-Pol (44, 179, 207-211). However, it is unknown whether $tRNA^{Lys3}$ affects RT maturation during viral replication. Thus, to investigate the role of $tRNA^{Lys3}$ in RT maturation, we knocked down KARS expression in 293T cells by small interfering RNA (siRNA), and then transfected these cells with a full-length replication competent molecular clone of HIV-1 ($HIV\text{-}1^{LAI}$). We anticipated that this knockdown

would impact virus replication, Gag-Pol polyprotein processing in the host cell and/or virus, and possibly the formation of the reverse transcription initiation complex. It has been previously shown that KARS knockdown does not alter general protein translation (228). Forty-eight hours after HIV-1^{LAI} transfection, we evaluated RT expression in both the 293T cells and in purified viral particles.

The siRNA knockdown of KARS expression in the 293T cells was stable for the duration of the experiment (Figure 5.8A). We found intracellular accumulation of RT that was not observed in the control cells (Figures 5.8B and 5.8C). Interestingly, the KARS knockdown cells accumulated p66, which did not appear to be efficiently processed to p51 by HIV-1 PR (Figures 5.8B and 5.8C), indicating a possible effect of $tRNA^{Lys3}$ on RT maturation within the cellular environment. Compared with control cells, KARS knockdown resulted in a significant reduction in viral particle production, as assessed by quantification of p24 (Figure 5.8D). However, in the viral particles that were produced from KARS knockdown cells, the p66:p51 ratio was 1:1 (Figures 5.8E and 5.8F), with no difference in relative infectivity (as measured in TZM-bL cells) between the viruses generated from the control and KARS knockdown cell lines (Figure 5.8G). These observations may suggest a role for $tRNA^{Lys3}$ in viral assembly, which is beyond the scope of the current study. We could not measure $tRNA^{Lys3}$ levels in the viral particles produced by KARS knockdown cell lines due to low virus production of these cells.

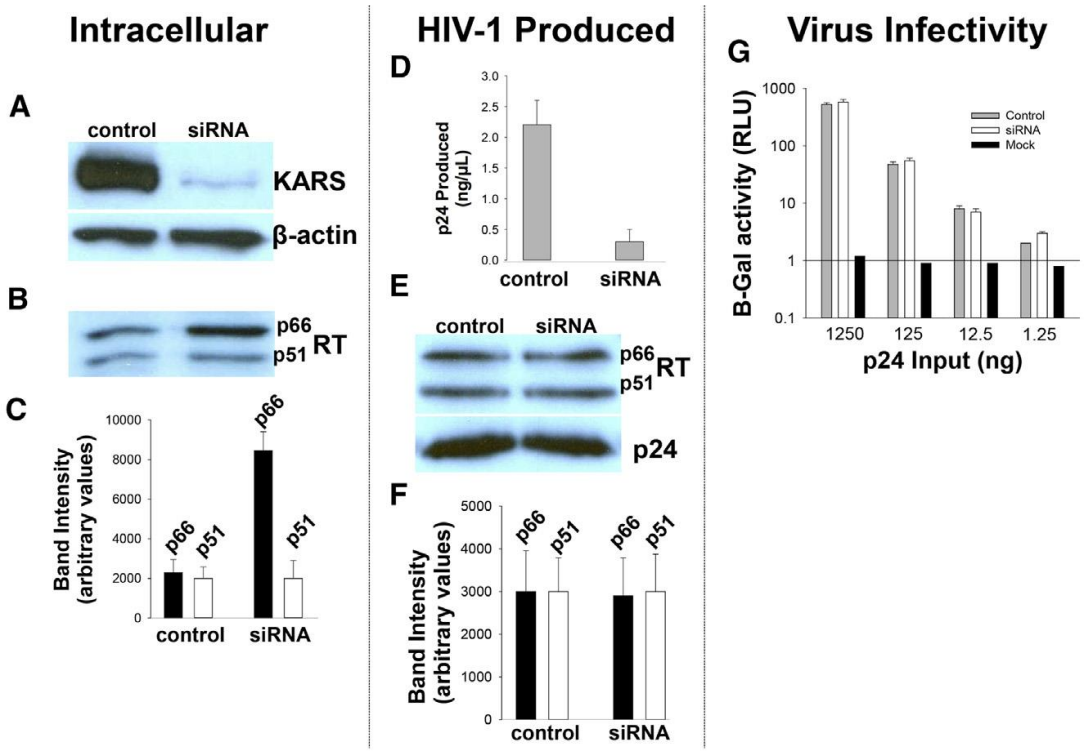

Figure 5.8 siRNA-Mediated Knockdown of KARS in 293T Cells

(A) Western blot analysis of KARS and B-actin expression in 293T cells 48 h after HIV-1 transfection; (B) western blot of intracellular RT expression; (C) densiometric analysis of (B); amount of HIV-1 produced, as assessed by p24, from the KARS knockdown and control 293T cells; (E) virion-associated RT and p24; (F) densiometric analysis of (E); single-cycle infectivity of HIV-1 generated from KARS knockdown and control 293T cells as assessed in TZM-bL cells. The data in (C, D, F, and G) are reported as the mean ± SE from three independent experiments.

5.5 DISCUSSION

In all structures of the mature p66/p51 heterodimer, the p51-RNH cleavage site is sequestered in a β sheet within the RNH domain and is inaccessible to PR (Figure 5.1B). Thus, the pathways involved in maturation of the asymmetric p66/p51 heterodimer are unknown. Using an in vitro RT maturation assay, we previously demonstrated that interaction of $tRNA^{Lys3}$ with the p66/p66

homodimer enhances specific cleavage by PR at the p51-RNH cleavage site, resulting in p66/p51 formation (195). The mechanisms by which $tRNA^{Lys3}$ enhances p66/p51 heterodimer production, however, are unclear and could involve its ability to increase p66/p66 homodimer formation and/or introduce conformational changes, particularly in the RNH domain of RT, which facilitate PR-mediated processing of p66 to p51. Thus, we assessed the impact of $tRNA^{Lys3}$ interaction on p66/p66 conformation in solution using ^{1}H-^{13}C HMQC NMR of ^{13}C $^{\delta 1}$-Ile signals of p66, to assess specific signals with high sensitivity, and using ^{1}H-^{15}N TROSY-HSQC NMR of [U-^{2}H,^{15}N] p66, to gain insight into overall conformational changes in the protein.

The results indicate a partial unfolding of p66/p66, reduction of thumb domain motion, and reduction in the mobility of at least one RNH domain upon $tRNA^{Lys3}$ interaction (Figure 5.3). In addition, a slight increase in the signal intensity proximal to the random coil position of the Ile-$^{\delta 1}$ methyl group was observed (Figure 5.3). In large protein NMR, even a small number of fragments can give significant signals, because of the small rotational correlation time of fragments compared with that of the large protein. Such fragments could be introduced to a sample upon cleavage by contaminated E. coli enzyme; however, in this study, the increases in the unfolded signals detected in the ^{1}H-^{13}C and ^{1}H-^{15}N data are not due to generation of fragmented protein products, as shown in the SDS gels (Figures C6 and C7).

To reduce the p66 monomer fraction in our samples, we utilized NNRTIs that are potent chemical enhancers of p66/p66 homodimer formation (224, 225). Using an in vitro maturation assay (Figure 5.5), we unequivocally show that NNRTIs do not facilitate proteolytic processing of p66 to p51, whereas addition of $tRNA^{Lys3}$ to the NNRTI-bound p66/p66 resulted in efficient cleavage. Our NMR experiments show that NNRTI binding induced conformational changes in p66/p66 homodimer that extended to the RNH domains (Figure 5.6A). However, these

conformational changes were not sufficient to induce efficient PR processing at the p51-RNH site (Figure 5.5C). Indeed, the addition of $tRNA^{Lys3}$ induced additional changes particularly in the RNH domain (Figure 5.6B), with an increase in the unfolded resonance similar to that observed in the experiments without RPV. Collectively, these data show that specific conformational changes in the p66/p66 homodimer, enhanced by nucleic acid, are needed for efficient RT maturation.

In the Ile-$^{\delta 1}$ methyl ^{1}H-^{13}C HMQC spectrum of p66/p66-RPV in the presence of $tRNA^{Lys3}$, we observed resonance changes for residues in the finger-palm and thumb domains, 202 and 274, as well as those in the RNH domain. Thus, $tRNA^{Lys3}$ is expected to bind at the canonical nucleic acid binding site that spans the entire p51 domain in p66 (229). This notion is consistent with the observed reduction in thumb domain signals upon $tRNA^{Lys3}$ binding to p66, monitored by ^{1}H-^{15}N TROSY-HSQC NMR (Figure 5.3B). Previous ^{19}F NMR studies have suggested that the NNRTI binding pocket of p66/p66 is similar to that of p66/p51 (76, 77). If a p66/p51-like structure is present in the p66/p66-RPV form, our observation that the p66/p66-RPV-bound conformation is asymmetric but not ideal for p66/p51 production in the absence of $tRNA^{Lys3}$ suggests a steric effect of $tRNA^{Lys3}$ on one of the RNH domains. Since we were unable to identify the specific region that undergoes partial unfolding, the RNH domain signal observed in the $tRNA^{Lys3}$-bound p66/p66 in the ^{1}H-^{15}N TROSY-HSQC NMR spectrum (Figure 5.3B) could correspond to domains A or B in the Ile-$^{\delta 1}$ methyl ^{1}H-^{13}C HMQC of Figure 5.6B or to the $tRNA^{Lys3}$-p66 monomer form. In either case, our data do not support the model that p66/p66 alone, in the absence of nucleic acid or PR, slowly changes conformation in solution (86, 87, 93), as we did not observe such a conformational change. Even if there is a minor conformer, it may be separated by a high energy barrier from the major population in p66/p66 alone.

During the HIV-1 life cycle, $tRNA^{Lys3}$ is recruited into the budding virus through its interaction with KARS and Gag-Pol (44, 179, 207-211). To investigate the role of $tRNA^{Lys3}$ in RT maturation, we knocked down KARS expression in 293T cells by siRNA, and then transfected these cells with a full-length replication competent molecular clone of HIV-1 ($HIV\text{-}1^{LAI}$). Interestingly, we found intracellular accumulation of inefficiently processed RT in cells with reduced KARS. Since KARS knockdown significantly reduced the amount of virus production, we think that the accumulated Gag-Pol and the products in the cell showed such difference in RT maturation in the intracellular environment. Thus, the result is not a direct evidence of the RT maturation in virus, but suggests a possible role of KARS in the intracellular maturation of RT and supports application of $tRNA^{Lys3}$ in our in vitro data.

The reduction of the amount of virus production is consistent with the notion that KARS is important for viral packaging of $tRNA^{Lys3}$ and Gag-Pol (230, 231). The virus that was produced from the KARS knockdown cells, however, contained p66/p51 RT and exhibited similar infectivity to the control virus. While we observed robust knockdown of KARS in the 293T cells, there was residual protein expression, which may have been sufficient to facilitate some virus production. However, how $tRNA^{Lys3}$ affects viral packaging is beyond the scope of the current study. Similarly, there are studies that have investigated the order of PR cleavage sites in Gag-Pol using different systems (91, 154, 204, 232). In this regard, our study does not address the entire RT maturation pathway from Gag-Pol processing to p66/p51 production, but illuminates the RT conformational characteristics in relation to functional heterodimer maturation.

5.6 ACKNOWLEDGMENTS

We thank Michel Guerrero for technical support and Teresa Brosenitsch for reading the manuscript. This study was supported by grants from the NIH (R01GM105401 to R.I., U54AI150472 and R01GM118012 to S.G.S, R01GM068406 to N.S.C, and P50AI150481 to R.I. and N.S.C.), and the MEXT-Supported Program for the Strategic Research Foundation at Private Universities in Japan (the term, 2011–2015, to G.K.).

6.0 SUMMARY AND FUTURE DIRECTIONS

The underlying focus of this work has been an attempt to understand the mechanism by which a single RNH domain in the immature p66/p66 homodimer is specifically recognized and cleaved by PR, resulting in the formation of mature heterodimeric (p66/p51) RT. In Chapter 2, building upon observations that mutations to the p51-RNH cleavage site result in aberrant proteolytic processing of p66/p66 by HIV-1 PR *in vivo* (90), we found that these cleavage site mutations destabilize the hydrophobic core of the RNH domain (95). The addition of a distal compensatory mutation, shown to ameliorate the aberrant processing of cleavage site mutants *in vivo* (91), restored folding of the isolated RNH domain *in vitro* (95). Altogether, these findings suggest that the folding stability of the RNH domain is coupled with the specificity of the p51-RNH cleavage site in the proteolytic maturation of RT by PR. This work also suggests that the maturation of RT is more complex than models which propose selective unfolding of one RNH domain upon homodimerization of p66.

It had been previously been reported that one of the RNH domains in the p66/p66 homodimer exhibited slow, time-dependent unfolding, as monitored by solution NMR of Ile δ1-[$^{13}CH_3$] probes (86, 93). We previously observed solution NMR spectra of p66, but did not find evidence for RNH domain unfolding, as monitored by backbone probes (75). Therefore, in Chapter 3, we examined the time-dependent stability of p66 solution NMR spectra using backbone ([^{1}H-^{15}N]) and side-chain (Ile $^{\delta 1}$-[$^{13}CH_3$]) probes. We found that our samples exhibited stable spectral features, with no sign of differential domain unfolding, over the duration of both 60+ hour experiments. We therefore concluded that homodimerization is not sufficient to induce conformational asymmetry in one of the two RNH domains in p66/p66.

In addition to these observations, we demonstrated in Chapter 4 that the interaction of p66/p66 with tRNA significantly enhances the efficiency and specificity of PR mediated maturation *in vitro*. Interpreted within the context of our findings from Chapter 3, we proposed a model whereby binding of tRNA to the immature p66/p66 homodimer induces a conformational change which facilitates specific cleavage by HIV-1 PR to form the mature RT. In Chapter 5 we showed that conformational asymmetry was induced upon ligand binding by the p66/p66 homodimer. Furthermore, addition of $tRNA^{Lys3}$ to this homodimer induced a conformational change which is sufficient to enhance RT maturation, supporting the proposed model.

While this model helps to explain the structural basis for p51-RNH processing site recognition and cleavage by PR, it also raises several interesting questions. Although the precise timing and location of RT maturation are unknown, evidence from several independent studies suggest that PR mediated cleavage of RT from Gag-Pol is highly regulated during assembly, budding and maturation of the viral particle (90, 137, 203). The structures of full length Gag-Pol and of biologically relevant Gag-Pol processing intermediates preceding mature RT could shed light on the regulation of RT maturation *in vivo* (90, 154, 204, 232).

Endogenous cellular factors play essential roles in HIV-1 replication (233, 234), therefore a more detailed understanding of host-virus interactions may lead to the discovery of novel therapeutic targets. Lysyl-tRNA synthetase (LysRS), the cellular factor which binds and aminoacylates the 3'-end of $tRNA^{Lys}$, is selectively packaged into viral particles via a direct interaction with the Capsid (CA) region of Gag (207, 235). It was therefore suggested that LysRS was responsible for viral incorporation of $tRNA^{Lys}$. However, it was subsequently demonstrated that the thumb domain of RT within unprocessed Gag-Pol is required for viral packaging of $tRNA^{Lys3}$ (44, 207, 230, 236), and that $tRNA^{Lys3}$ annealed to the PBS must be in a

nonaminoacylated or “uncharged” state in order to serve as the primer for reverse transcription (237). It has been recently reported that upon HIV-1 infection, a substantial fraction of LysRS is phosphorylated at Ser207, resulting in a loss of its aminoacylation activity and its relocalization to the nucleus (238). Intriguingly, tRNAs lacking 3’ CCA ends affected the efficiency of nuclear localization of Reverse Transcription complexes in an *in vitro* transport assay designed to identify cellular factors involved in nuclear import (239). Additional studies may delineate the roles of LysRS and tRNA in viral trafficking, assembly and maturation.

APPENDIX A

A.1 CHARACTERIZATION OF WILD TYPE AND MUTANT RNH DOMAINS

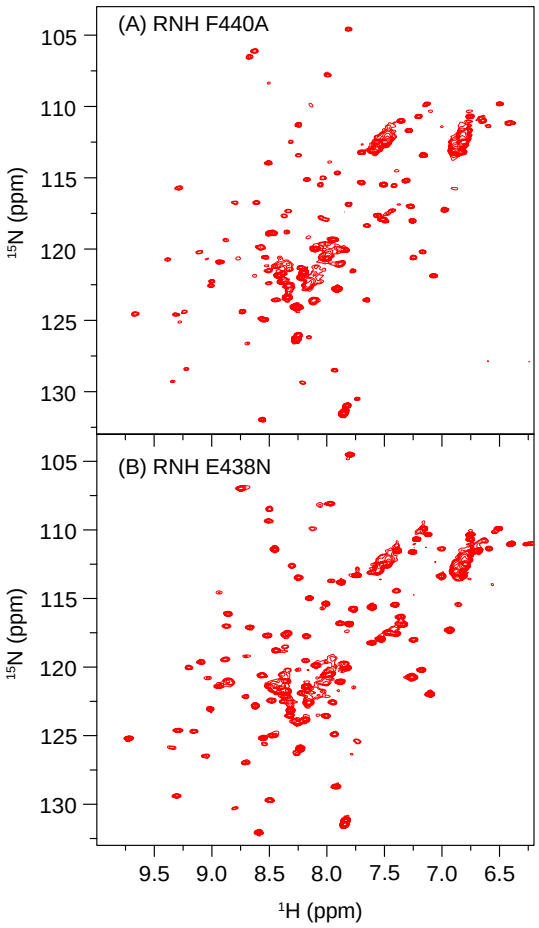

Figure A.1 : 1H-15N HSQC spectra of cleavage site mutants at pH 8.0

(A) the RNH_{F440A} mutant and (B) the RNH_{E438N} mutant.

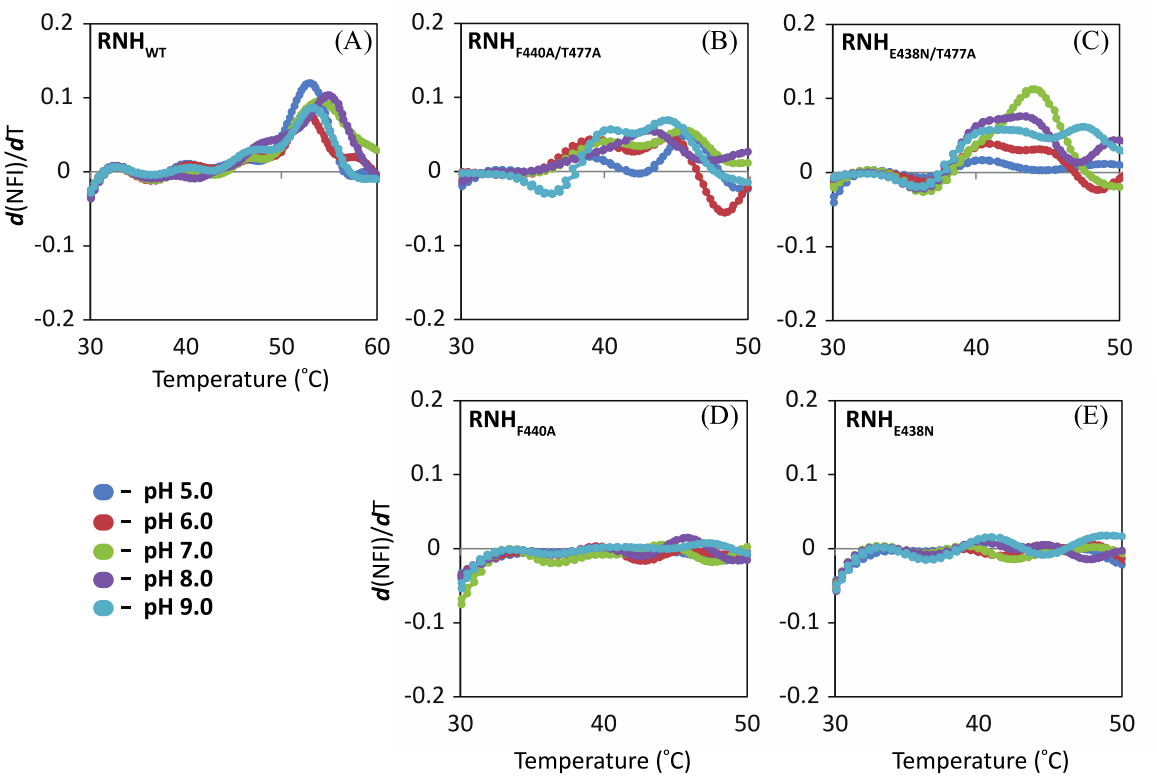

Figure A.2 : Melting Curves for WT and mutant RNH domains obtained by Differential Scanning Fluorimetry.

The first derivative of normalized fluorescence intensity signals, d(NFI)/dT, as a function of temperature for (A) WT RNH, (B) $RNH_{F440A/T477A}$, (C) $RNH_{E438N/T477A}$, (D) RNH_{F440A}, and (E) RNH_{E438N} mutants at pH 5.0, 6.0, 7.0, 8.0, and 9.0. The peak of the d(NFI)/dT was clearly observed for WT RNH, $RNH_{F440A/T477A}$, and $RNH_{E438N/T477A}$. However, no change of the derivative was observed for RNH_{F440A} and RNH_{E438N}.

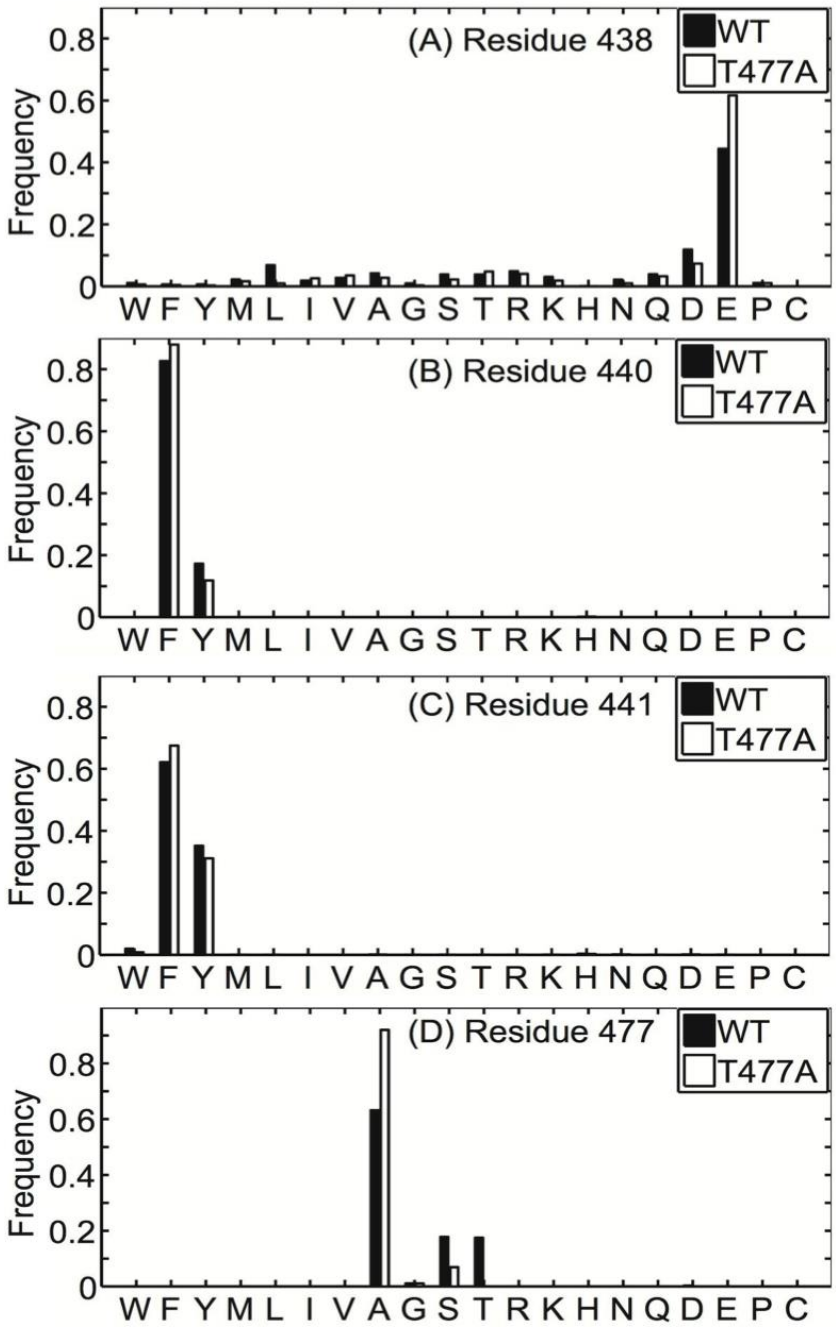

Figure A.3: Prediction of sequence tolerance to preserve folding stability as a function of residue position.

(A) 438, (B) 440, (C) 441 and (D) 477 for RNH WT (filled bar) and RNH_{T477A} (open bar), calculated using the structural coordinates at a 100 ns simulation point.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR PRINCIPAL COMPONENT ANALYSIS OF NMR SIGNAL INTENSITY CHANGES TO ASSESS STRUCTURAL INTEGRITY OF A LARGE MULTI-DOMAIN PROTEIN

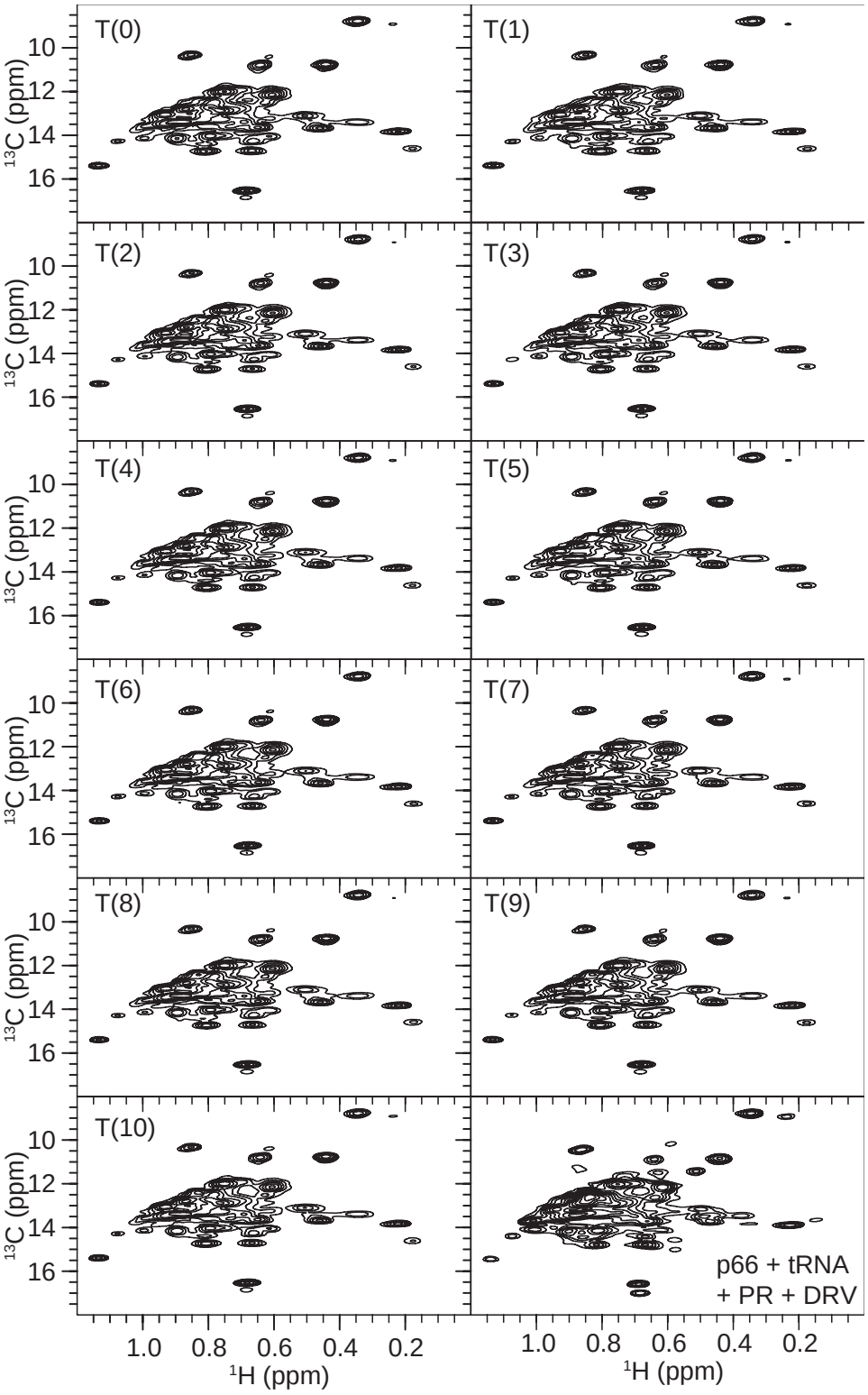

Figure B.1 : Individual 1H-13C SOFAST-HMQC time course Spectra of p66.

Spectra were recorded for 70 µM [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] labeled p66 in a deuterated 25 mM Bis-Tris buffer, containing 100 mM KCl, 0.02% NaN_3, and 5% v/v Glycerol-d8, at pH 7.1. All experiments were performed at 35 °C. spectra

were recorded for 70 µM [U-^{2}H], Ile $^{\delta 1}$-[$^{13}CH_3$] labeled p66 in a deuterated 25 mM Bis-Tris buffer, containing 100 mM KCl, 0.02% NaN_3, and 5% v/v Glycerol-d8, at pH 7.1. All the experiments were performed at 35 °C. Spectra were recorded sequentially as indicated by the number in the top left corner, starting at T(0) = 0 hours after the start of acquisition and finishing at T(10) = 56.95 hours after the start of acquisition. The lowest panel on the right is a spectrum recorded for after p66 was incubated with HIV-1 PR in the presence of tRNA for 2 hours, then quenched by the addition of the PR inhibitor Darunavir (DRV).

APPENDIX C

SUPPLEMENTARY MATERIAL FOR CONFORMATIONAL CHANGES IN HIV-1 REVERSE TRANSCRIPTASE WHICH FACILITATE ITS MATURATION

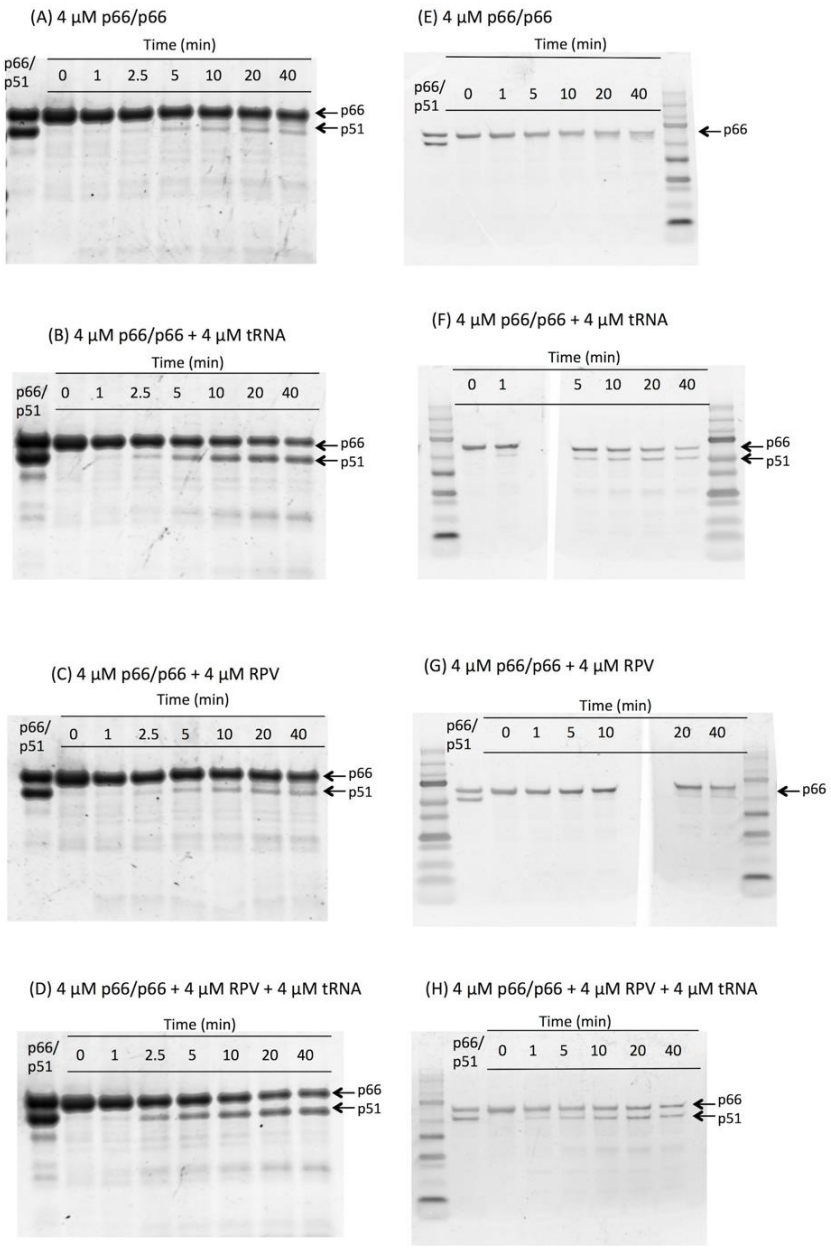

Figure C.1: Entire gels from Figure 5.5, with additional gels to show data reproducibility.

(A-D) Entire gels used to generate Figure 5.5, showing the time dependence of p66 processing by HIV-1 PR in 20 mM sodium acetate buffer at pH 5.2, at 37 °C: in the (A, B) absence or (C, D) presence of RPV, and (B, D) presence or (A, C) absence of tRNALys3. (E-H) the data repeated, from an independent experiment. In the latter, since three gel cassettes were used to run the four sets of data, each gel cassette contains p66/p51 reference and a molecular weight ladder.

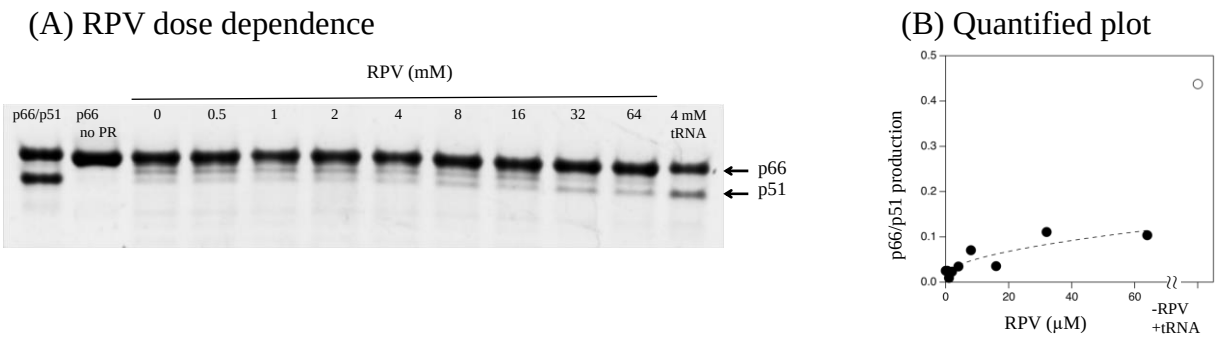

Figure C.2: Dose Dependence and Quantification of the Effect of RPV on *in vitro* maturation of RT

(A) RPV-dose dependence of p66/p66 (4 μM) processing by HIV-1 PR in 20 mM sodium acetate buffer at pH 5.2 and 37 °C and (B) the quantification.

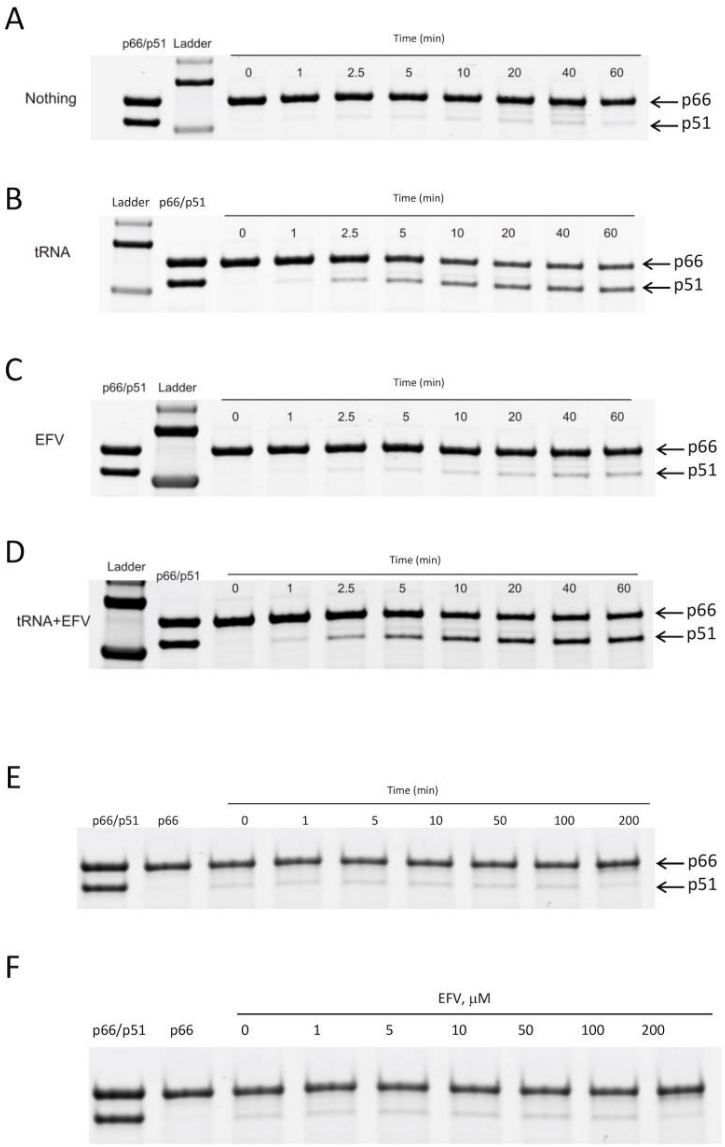

Figure C.3: Time dependence of p66 processing by PR in the absence and presence of EFV

(A – D) Time dependence of p66 processing by 1 μM HIV-1 PR in 20 mM sodium acetate buffer, containing 1% DMSO, at pH 5.2, at 37 °C: in the (A, B) absence or (C, D) presence of EFV, and (B, D) presence or (A, C) absence of tRNALys3. Concentrations of p66/p66, tRNA and EFV are 5 μM, 20 μM, and 20μM, respectively. (E) repeat of A. (F) the EFV dose-dependence. Protein, tRNA, and EFV concentrations are indicated in the Figure.

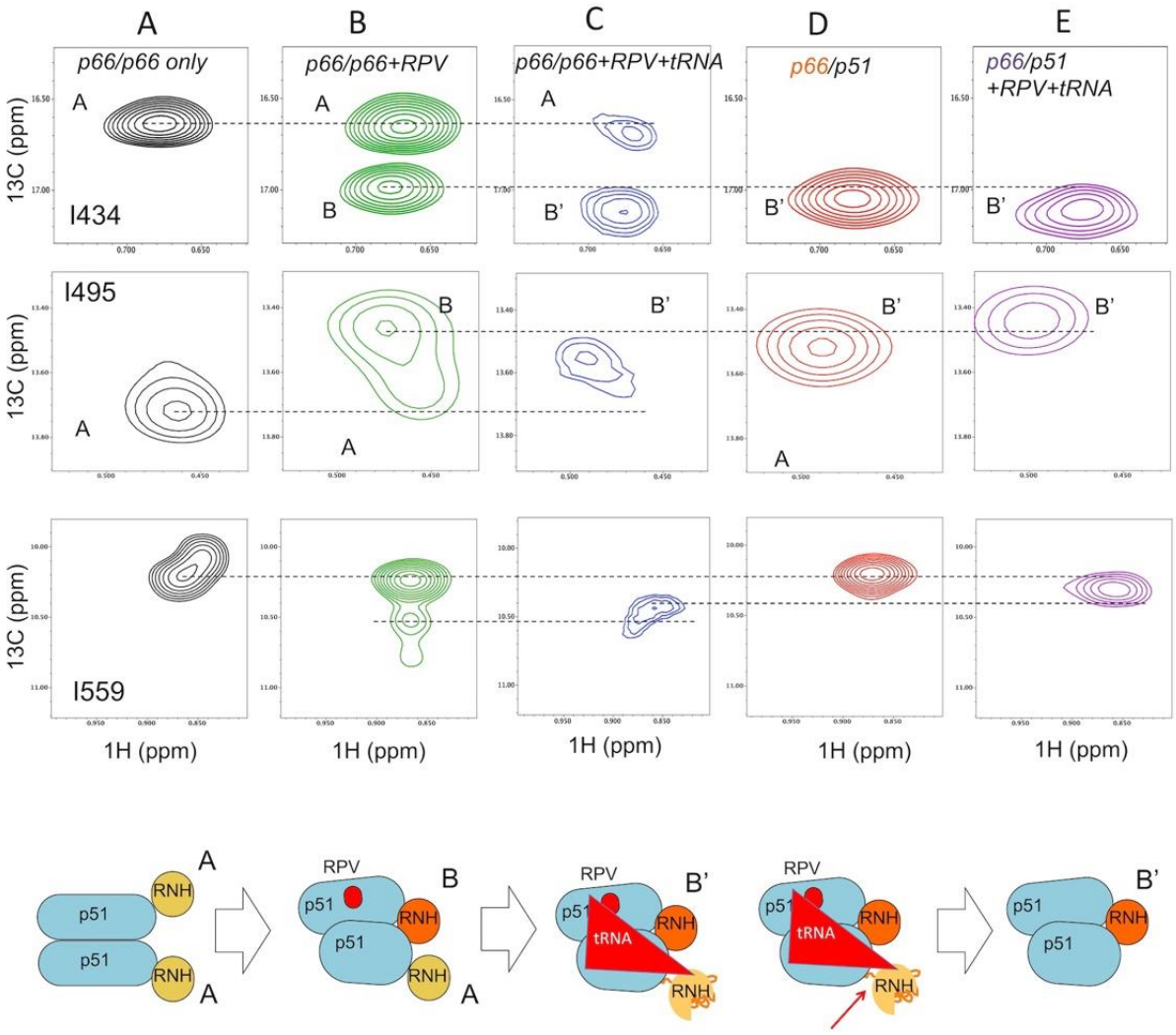

Figure C.4: Overview of the observed signal patterns of p66 RNH domain residues I434, I495 and I559 recorded at 20 °C

(A) p66/p66 only, (B) p66/p66 + RPV, (C) p66/p66 + RPV + tRNA, (D) p66/p51 in which only the p66 subunit is [U-^{2}H]- and Ile $^{\delta 1}$-[$^{13}CH_3$] labelled, and (E) is the same sample as D but in the presence of RPV. Cartoon at the bottom indicates conformational changes deduced from the observed spectra in each condition. The "A" designation in the spectra indicates NMR resonance positions stemming from the p66/p66 homodimer while B and B' indicate newly generated resonance positions upon RPV interaction and partial digestion of p66/p66, respectively. Note, in (B), the p66/p66 was not saturated, presumably due to lower RPV solubility in aqueous solution at 20 °C, compared to 35 °C. In (C), signal broadening was more severe at 20 °C, compared to 35 °C (see Figure 5.7).

BIBLIOGRAPHY

1. Selik RM, Mokotoff ED, Branson B, Owen SM, Whitmore S, Hall HI. Revised Surveillance Case Definition for HIV Infection - United States, 2014. Mmwr Recomm Rep. 2014;63(3). PubMed PMID: WOS:000334352300001.

2. Dailey AF, Hoots BE, Hall HI, Song RG, Hayes D, Fulton P, Prejean J, Hernandez AL, Koenig LJ, Valleroy LA. Vital Signs: Human Immunodeficiency Virus Testing and Diagnosis Delays - United States. Mmwr-Morbid Mortal W. 2017;66(47):1300-6. PubMed PMID: WOS:000417329200004.

3. Iyidogan P, Anderson KS. Current perspectives on HIV-1 antiretroviral drug resistance. Viruses. 2014;6(10):4095-139. Epub 2014/10/25. doi: 10.3390/v6104095. PubMed PMID: 25341668; PMCID: PMC4213579.

4. Butler IF, Pandrea I, Marx PA, Apetrei C. HIV genetic diversity: biological and public health consequences. Curr HIV Res. 2007;5(1):23-45. Epub 2007/02/03. PubMed PMID: 17266555.

5. Coffin JM, Hughes SH, Varmus HE. The Interactions of Retroviruses and their Hosts. In: Coffin JM, Hughes SH, Varmus HE, editors. Retroviruses. Cold Spring Harbor (NY)1997.

6. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH. Nature, Position, and Frequency of Mutations Made in a Single Cycle of HIV-1 Replication. Journal of Virology. 2010;84(19):9864-78. doi: 10.1128/Jvi.00915-10. PubMed PMID: WOS:000282641800020.

7. Coffin JM. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science. 1995;267(5197):483-9. Epub 1995/01/27. PubMed PMID: 7824947.

8. Freed EO. HIV-1 assembly, release and maturation. Nat Rev Microbiol. 2015;13(8):484-96. doi: 10.1038/nrmicro3490. PubMed PMID: WOS:000358000700007.

9. Coffin JM, Hughes SH, Varmus HE. Retroviruses. Plainview, NY: Cold Spring Harbor Laboratory Press; 1997.

10. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. Structure. 2002;10(3):369-81. Epub 2002/05/15. PubMed PMID: 12005435.

11. Li GD, De Clercq E. HIV Genome-Wide Protein Associations: a Review of 30 Years of Research. Microbiol Mol Biol R. 2016;80(3):679-731. doi: 10.1128/Mmbr.00065-15. PubMed PMID: WOS:000382247500009.

12. Wondrak EM, Nashed NT, Haber MT, Jerina DM, Louis JM. A transient precursor of the HIV-1 protease. Isolation, characterization, and kinetics of maturation. J Biol Chem. 1996;271(8):4477-81. Epub 1996/02/23. doi: 10.1074/jbc.271.8.4477. PubMed PMID: 8626801.

13. Wondrak EM, Louis JM. Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease. Biochemistry. 1996;35(39):12957-62. Epub 1996/10/01. doi: 10.1021/bi960984y. PubMed PMID: 8841142.

14. Ishima R, Ghirlando R, Tozser J, Gronenborn AM, Torchia DA, Louis JM. Folded monomer of HIV-1 protease. J Biol Chem. 2001;276(52):49110-6. doi: 10.1074/jbc.M108136200. PubMed PMID: 11598128.

15. Ishima R, Torchia DA, Lynch SM, Gronenborn AM, Louis JM. Solution structure of the mature HIV-1 protease monomer - Insight into the tertiary fold and stability of a precursor. Journal of Biological Chemistry. 2003;278(44):43311-9. doi: 10.1074/jbc.M307549200. PubMed PMID: WOS:000186157000074.

16. Zheng R, Jenkins TM, Craigie R. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. Proc Natl Acad Sci U S A. 1996;93(24):13659-64. Epub 1996/11/26. PubMed PMID: 8942990; PMCID: PMC19383.

17. Drelich M, Wilhelm R, Mous J. Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. Virology. 1992;188(2):459-68. Epub 1992/06/01. PubMed PMID: 1585629.

18. Leavitt AD, Shiue L, Varmus HE. Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. J Biol Chem. 1993;268(3):2113-9. Epub 1993/01/25. PubMed PMID: 8420982.

19. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. Retroviral intasome assembly and inhibition of DNA strand transfer. Nature. 2010;464(7286):232-6. Epub 2010/02/02. doi: 10.1038/nature08784. PubMed PMID: 20118915; PMCID: PMC2837123.

20. Engelman A, Hickman AB, Craigie R. The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. J Virol. 1994;68(9):5911-7. Epub 1994/09/01. PubMed PMID: 8057470; PMCID: PMC236996.

21. Lodi PJ, Ernst JA, Kuszewski J, Hickman AB, Engelman A, Craigie R, Clore GM, Gronenborn AM. Solution structure of the DNA binding domain of HIV-1 integrase. Biochemistry. 1995;34(31):9826-33. Epub 1995/08/08. PubMed PMID: 7632683.

22. Jenkins TM, Engelman A, Ghirlando R, Craigie R. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J Biol Chem. 1996;271(13):7712-8. Epub 1996/03/29. PubMed PMID: 8631811.

23. Yin Z, Shi K, Banerjee S, Pandey KK, Bera S, Grandgenett DP, Aihara H. Crystal structure of the Rous sarcoma virus intasome. Nature. 2016;530(7590):362-6. Epub 2016/02/19. doi: 10.1038/nature16950. PubMed PMID: 26887497; PMCID: PMC4881392.

24. Yin Z, Lapkouski M, Yang W, Craigie R. Assembly of prototype foamy virus strand transfer complexes on product DNA bypassing catalysis of integration. Protein science : a publication of the Protein Society. 2012;21(12):1849-57. Epub 2012/09/27. doi: 10.1002/pro.2166. PubMed PMID: 23011895; PMCID: PMC3575915.

25. Passos DO, Li M, Yang R, Rebensburg SV, Ghirlando R, Jeon Y, Shkriabai N, Kvaratskhelia M, Craigie R, Lyumkis D. Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. Science. 2017;355(6320):89-92. Epub 2017/01/07. doi: 10.1126/science.aah5163. PubMed PMID: 28059769; PMCID: PMC5508583.

26. Pauza CD. Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. Virology. 1990;179(2):886-9. Epub 1990/12/01. PubMed PMID: 2238479.

27. Fujiwara T, Mizuuchi K. Retroviral DNA integration: structure of an integration intermediate. Cell. 1988;54(4):497-504. Epub 1988/08/12. PubMed PMID: 3401925.

28. Brown PO, Bowerman B, Varmus HE, Bishop JM. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. Proc Natl Acad Sci U S A. 1989;86(8):2525-9. Epub 1989/04/01. PubMed PMID: 2539592; PMCID: PMC286949.

29. Yoder KE, Bushman FD. Repair of gaps in retroviral DNA integration intermediates. J Virol. 2000;74(23):11191-200. Epub 2000/11/09. PubMed PMID: 11070016; PMCID: PMC113210.

30. Engelman A, Englund G, Orenstein JM, Martin MA, Craigie R. Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. J Virol. 1995;69(5):2729-36. Epub 1995/05/01. PubMed PMID: 7535863; PMCID: PMC188965.

31. Zhu K, Dobard C, Chow SA. Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. J Virol. 2004;78(10):5045-55. Epub 2004/04/29. doi: 10.1128/jvi.78.10.5045-5055.2004. PubMed PMID: 15113886; PMCID: PMC400327.

32. Dobard CW, Briones MS, Chow SA. Molecular mechanisms by which human immunodeficiency virus type 1 integrase stimulates the early steps of reverse transcription. J Virol. 2007;81(18):10037-46. Epub 2007/07/13. doi: 10.1128/JVI.00519-07. PubMed PMID: 17626089; PMCID: PMC2045400.

33. Masuda T. Non-Enzymatic Functions of Retroviral Integrase: The Next Target for Novel Anti-HIV Drug Development. Front Microbiol. 2011;2:210. Epub 2011/10/22. doi: 10.3389/fmicb.2011.00210. PubMed PMID: 22016749; PMCID: PMC3192317.

34. Feng L, Larue RC, Slaughter A, Kessl JJ, Kvaratskhelia M. HIV-1 integrase multimerization as a therapeutic target. Curr Top Microbiol Immunol. 2015;389:93-119. Epub 2015/03/18. doi: 10.1007/82_2015_439. PubMed PMID: 25778682; PMCID: PMC4791179.

35. Briones MS, Dobard CW, Chow SA. Role of human immunodeficiency virus type 1 integrase in uncoating of the viral core. J Virol. 2010;84(10):5181-90. Epub 2010/03/12. doi: 10.1128/JVI.02382-09. PubMed PMID: 20219923; PMCID: PMC2863833.

36. Ikeda T, Nishitsuji H, Zhou X, Nara N, Ohashi T, Kannagi M, Masuda T. Evaluation of the functional involvement of human immunodeficiency virus type 1 integrase in nuclear import of viral cDNA during acute infection. J Virol. 2004;78(21):11563-73. Epub 2004/10/14. doi: 10.1128/JVI.78.21.11563-11573.2004. PubMed PMID: 15479797; PMCID: PMC523288.

37. Wilkinson TA, Januszyk K, Phillips ML, Tekeste SS, Zhang M, Miller JT, Le Grice SF, Clubb RT, Chow SA. Identifying and characterizing a functional HIV-1 reverse transcriptase-binding site on integrase. J Biol Chem. 2009;284(12):7931-9. Epub 2009/01/20. doi: 10.1074/jbc.M806241200. PubMed PMID: 19150986; PMCID: PMC2658086.

38. Ao Z, Huang G, Yao H, Xu Z, Labine M, Cochrane AW, Yao X. Interaction of human immunodeficiency virus type 1 integrase with cellular nuclear import receptor importin 7 and its impact on viral replication. J Biol Chem. 2007;282(18):13456-67. Epub 2007/03/16. doi: 10.1074/jbc.M610546200. PubMed PMID: 17360709.

39. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell. 2016;166(5):1257-68 e12. Epub 2016/08/28. doi: 10.1016/j.cell.2016.07.044. PubMed PMID: 27565348; PMCID: PMC5003418.

40. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, Bardiot D, Van der Veken NJ, Van Remoortel B, Strelkov SV, De Maeyer M, Chaltin P, Debyser Z. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. Nat Chem Biol. 2010;6(6):442-8. Epub 2010/05/18. doi: 10.1038/nchembio.370. PubMed PMID: 20473303.

41. Le Rouzic E, Bonnard D, Chasset S, Bruneau JM, Chevreuil F, Le Strat F, Nguyen J, Beauvoir R, Amadori C, Brias J, Vomscheid S, Eiler S, Levy N, Delelis O, Deprez E, Saib A, Zamborlini A, Emiliani S, Ruff M, Ledoussal B, Moreau F, Benarous R. Dual inhibition of HIV-1 replication by integrase-LEDGF allosteric inhibitors is predominant at the post-integration stage. Retrovirology. 2013;10:144. Epub 2013/11/23. doi: 10.1186/1742-4690-10-144. PubMed PMID: 24261564; PMCID: PMC4222603.

42. Fader LD, Malenfant E, Parisien M, Carson R, Bilodeau F, Landry S, Pesant M, Brochu C, Morin S, Chabot C, Halmos T, Bousquet Y, Bailey MD, Kawai SH, Coulombe R, LaPlante S, Jakalian A, Bhardwaj PK, Wernic D, Schroeder P, Amad M, Edwards P, Garneau M, Duan J, Cordingley M, Bethell R, Mason SW, Bos M, Bonneau P, Poupart MA, Faucher AM, Simoneau B, Fenwick C, Yoakim C, Tsantrizos Y. Discovery of BI 224436, a Noncatalytic Site Integrase Inhibitor (NCINI) of HIV-1. ACS Med Chem Lett. 2014;5(4):422-7. Epub 2014/06/06. doi: 10.1021/ml500002n. PubMed PMID: 24900852; PMCID: PMC4027581.

43. Takahata T, Takeda E, Tobiume M, Tokunaga K, Yokoyama M, Huang YL, Hasegawa A, Shioda T, Sato H, Kannagi M, Masuda T. Critical Contribution of Tyr15 in the HIV-1 Integrase (IN) in Facilitating IN Assembly and Nonenzymatic Function through the IN Precursor Form with Reverse Transcriptase. J Virol. 2017;91(1). Epub 2016/11/01. doi: 10.1128/JVI.02003-16. PubMed PMID: 27795445; PMCID: PMC5165222.

44. Mak J, Jiang M, Wainberg MA, Hammarskjold ML, Rekosh D, Kleiman L. Role of Pr160gag-pol in mediating the selective incorporation of tRNA(Lys) into human immunodeficiency virus type 1 particles. J Virol. 1994;68(4):2065-72. Epub 1994/04/01. PubMed PMID: 7511167; PMCID: PMC236680.

45. Tanese N, Telesnitsky A, Goff SP. Abortive reverse transcription by mutants of Moloney murine leukemia virus deficient in the reverse transcriptase-associated RNase H function. J Virol. 1991;65(8):4387-97. Epub 1991/08/01. PubMed PMID: 1712862; PMCID: PMC248878.

46. Telesnitsky A, Blain SW, Goff SP. Defects in Moloney murine leukemia virus replication caused by a reverse transcriptase mutation modeled on the structure of Escherichia coli RNase H. J Virol. 1992;66(2):615-22. Epub 1992/02/01. PubMed PMID: 1370551; PMCID: PMC240759.

47. DeStefano JJ, Mallaber LM, Rodriguez-Rodriguez L, Fay PJ, Bambara RA. Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. J Virol. 1992;66(11):6370-8. Epub 1992/11/01. PubMed PMID: 1383563; PMCID: PMC240129.

48. DeStefano JJ, Bambara RA, Fay PJ. The mechanism of human immunodeficiency virus reverse transcriptase-catalyzed strand transfer from internal regions of heteropolymeric RNA templates. J Biol Chem. 1994;269(1):161-8. Epub 1994/01/07. PubMed PMID: 7506252.

49. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature. 1988;331(6153):280-3. Epub 1988/01/21. doi: 10.1038/331280a0. PubMed PMID: 2447506.

50. di Marzo Veronese F, Copeland TD, DeVico AL, Rahman R, Oroszlan S, Gallo RC, Sarngadharan MG. Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. Science. 1986;231(4743):1289-91. Epub 1986/03/14. PubMed PMID: 2418504.

51. Cameron CE, Ghosh M, Le Grice SF, Benkovic SJ. Mutations in HIV reverse transcriptase which alter RNase H activity and decrease strand transfer efficiency are suppressed by HIV nucleocapsid protein. Proc Natl Acad Sci U S A. 1997;94(13):6700-5. Epub 1997/06/24. PubMed PMID: 9192628; PMCID: PMC21221.

52. Larder BA, Purifoy DJ, Powell KL, Darby G. Site-specific mutagenesis of AIDS virus reverse transcriptase. Nature. 1987;327(6124):716-7. Epub 1987/06/01. doi: 10.1038/327716a0. PubMed PMID: 2439916.

53. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science. 1998;282(5394):1669-75. Epub 1998/11/30. PubMed PMID: 9831551.

54. Davies JFn, Hostomska Z, Hostomsky Z, Jordan SR, Matthews DA. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. Science. 1991;252(5002):88-95.

55. Sanner MF, Olson AJ, Spehner JC. Reduced surface: an efficient way to compute molecular surfaces. Biopolymers. 1996;38(3):305-20. Epub 1996/03/01. doi: 10.1002/(SICI)1097-0282(199603)38:3%3C305::AID-BIP4%3E3.0.CO;2-Y. PubMed PMID: 8906967.

56. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996;14(1):33-8, 27-8. Epub 1996/02/01. PubMed PMID: 8744570.

57. Restle T, Muller B, Goody RS. Dimerization of human immunodeficiency virus type 1 reverse transcriptase. A target for chemotherapeutic intervention. J Biol Chem. 1990;265(16):8986-8. Epub 1990/06/05. PubMed PMID: 1693146.

58. Divita G, Restle T, Goody RS. Characterization of the dimerization process of HIV-1 reverse transcriptase heterodimer using intrinsic protein fluorescence. FEBS Lett. 1993;324(2):153-8. Epub 1993/06/14. PubMed PMID: 7685295.

59. Divita G, Rittinger K, Geourjon C, Deleage G, Goody RS. Dimerization kinetics of HIV-1 and HIV-2 reverse transcriptase: A two step process J Mol Biol. 1995;245:508-21.

60. Venezia CF, Howard KJ, Ignatov ME, Holladay LA, Barkley MD. Effects of efavirenz binding on the subunit equilibria of HIV-1 reverse transcriptase. Biochemistry. 2006;45(9):2779-89. Epub 2006/03/01. doi: 10.1021/bi051915z. PubMed PMID: 16503633.

61. Venezia CF, Meany BJ, Braz VA, Barkley MD. Kinetics of association and dissociation of HIV-1 reverse transcriptase subunits. Biochemistry. 2009;48(38):9084-93. Epub 2009/09/01. doi: 10.1021/bi9010495. PubMed PMID: 19715314; PMCID: PMC2770954.

62. Sluis-Cremer N, Arion D, Abram ME, Parniak MA. Proteolytic processing of an HIV-1 pol polyprotein precursor: insights into the mechanism of reverse transcriptase p66/p51 heterodimer formation. Int J Biochem Cell B. 2004;36(9):1836-47. doi: 10.1016/j.biocel.2004.02.020. PubMed PMID: WOS:000222255900014.

63. Wapling J, Moore KL, Sonza S, Mak J, Tachedjian G. Mutations that abrogate human immunodeficiency virus type 1 reverse transcriptase dimerization affect maturation of the reverse transcriptase heterodimer. J Virol. 2005;79(16):10247-57. Epub 2005/07/30. doi: 10.1128/JVI.79.16.10247-10257.2005. PubMed PMID: 16051818; PMCID: PMC1182633.

64. Nishitsuji H, Yokoyama M, Sato H, Yamauchi S, Takaku H. Identification of amino acid residues in HIV-1 reverse transcriptase that are critical for the proteolytic processing of Gag-Pol precursors. FEBS Lett. 2011;585(21):3372-7. Epub 2011/10/19. doi: 10.1016/j.febslet.2011.09.034. PubMed PMID: 22004763.

65. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic acids research. 2000;28(1):235-42. Epub 1999/12/11. PubMed PMID: 10592235; PMCID: PMC102472.

66. Lowe DM, Aitken A, Bradley C, Darby GK, Larder BA, Powell KL, Purifoy DJ, Tisdale M, Stammers DK. HIV-1 reverse transcriptase: crystallization and analysis of domain structure by limited proteolysis. Biochemistry. 1988;27(25):8884-9. Epub 1988/12/13. PubMed PMID: 2466481.

67. Deibel MR, Jr., McQuade TJ, Brunner DP, Tarpley WG. Denaturation/refolding of purified recombinant HIV reverse transcriptase yields monomeric enzyme with high enzymatic activity. AIDS Res Hum Retroviruses. 1990;6(3):329-40. Epub 1990/03/11. doi: 10.1089/aid.1990.6.329. PubMed PMID: 1692723.

68. Abbondanzieri EA, Bokinsky G, Rausch JW, Zhang JX, Le Grice SF, Zhuang X. Dynamic binding orientations direct activity of HIV reverse transcriptase. Nature. 2008;453(7192):184-9. Epub 2008/05/10. doi: 10.1038/nature06941. PubMed PMID: 18464735; PMCID: PMC2655135.

69. Schauer GD, Huber KD, Leuba SH, Sluis-Cremer N. Mechanism of allosteric inhibition of HIV-1 reverse transcriptase revealed by single-molecule and ensemble fluorescence. Nucleic acids research. 2014;42(18):11687-96. Epub 2014/09/19. doi: 10.1093/nar/gku819. PubMed PMID: 25232099; PMCID: PMC4191400.

70. Huang C, Kalodimos CG. Structures of Large Protein Complexes Determined by Nuclear Magnetic Resonance Spectroscopy. Annu Rev Biophys. 2017;46:317-36. doi: 10.1146/annurev-biophys-070816-033701. PubMed PMID: 28375736.

71. Tugarinov V, Kay LE. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. J Am Chem Soc. 2003;125(45):13868-78. Epub 2003/11/06. doi: 10.1021/ja030345s. PubMed PMID: 14599227.

72. Godoy-Ruiz R, Guo C, Tugarinov V. Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins. J Am Chem Soc. 2010;132(51):18340-50. Epub 2010/12/09. doi: 10.1021/ja1083656. PubMed PMID: 21138300.

73. Sheppard D, Sprangers R, Tugarinov V. Experimental approaches for NMR studies of side-chain dynamics in high-molecular-weight proteins. Prog Nucl Magn Reson Spectrosc. 2010;56(1):1-45. Epub 2010/07/17. doi: 10.1016/j.pnmrs.2009.07.004. PubMed PMID: 20633347.

74. Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc. 2006;1(2):749-54. doi: 10.1038/nprot.2006.101. PubMed PMID: 17406304.

75. Sharaf NG, Poliner E, Slack RL, Christen MT, Byeon IJ, Parniak MA, Gronenborn AM, Ishima R. The p66 immature precursor of HIV-1 reverse transcriptase. Proteins. 2014;82(10):2343-52. Epub 2014/04/29. doi: 10.1002/prot.24594. PubMed PMID: 24771554; PMCID: PMC4441793.

76. Sharaf NG, Ishima R, Gronenborn AM. Conformational plasticity of the NNRTI-binding pocket in HIV-1 reverse transcriptase - A fluorine NMR study. Biochemistry. 2016;55(28):3864-73. doi: 10.1021/acs.biochem.6b00113. PubMed PMID: 27163463.

77. Sharaf NG, Xi Z, Ishima R, Gronenborn AM. The HIV-1 p66 homodimeric RT exhibits different conformations in the binding-competent and -incompetent NNRTI site. Proteins. 2017;85(12):2191-7. doi: 10.1002/prot.25383. PubMed PMID: 28905420.

78. Sharaf NG, Gronenborn AM. (19)F-modified proteins and (19)F-containing ligands as tools in solution NMR studies of protein interactions. Methods Enzymol. 2015;565:67-95.

79. Siivari K, Zhang M, Palmer AG, 3rd, Vogel HJ. NMR studies of the methionine methyl groups in calmodulin. FEBS Lett. 1995;366(2-3):104-8. Epub 1995/06/12. PubMed PMID: 7789524.

80. Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE. Selective methyl group protonation of perdeuterated proteins. J Mol Biol. 1996;263(5):627-36. Epub 1996/11/15. doi: 10.1006/jmbi.1996.0603. PubMed PMID: 8947563.

81. Gardner KH, Kay LE. Production and incorporation of N-15, C-13, H-2 (H-1-delta 1 methyl) isoleucine into proteins for multidimensional NMR studies. J Am Chem Soc. 1997;119(32):7599-600. doi: DOI 10.1021/ja9706514. PubMed PMID: WOS:A1997XR11800032.

82. Ruschak AM, Kay LE. Methyl groups as probes of supra-molecular structure, dynamics and function. J Biomol NMR. 2010;46(1):75-87. Epub 2009/09/29. doi: 10.1007/s10858-009-9376-1. PubMed PMID: 19784810.

83. Zheng X, Mueller GA, DeRose EF, London RE. Solution characterization of [methyl-(13)C]methionine HIV-1 reverse transcriptase by NMR spectroscopy. Antiviral research. 2009;84(3):205-14. doi: 10.1016/j.antiviral.2009.07.021. PubMed PMID: 19665484; PMCID: 3807244.

84. Zheng X, Mueller GA, Cuneo MJ, Derose EF, London RE. Homodimerization of the p51 subunit of HIV-1 reverse transcriptase. Biochemistry. 2010;49(13):2821-33. doi: 10.1021/bi902116z. PubMed PMID: 20180596.

85. Zheng X, Mueller GA, DeRose EF, London RE. Metal and ligand binding to the HIV-RNase H active site are remotely monitored by Ile556. Nucleic acids research. 2012;40(20):10543-53. doi: 10.1093/nar/gks791. PubMed PMID: 22941642; PMCID: 3488238.

86. Zheng X, Pedersen LC, Gabel SA, Mueller GA, Cuneo MJ, Derose EF, Krahn JM, London RE. Selective unfolding of one Ribonuclease H domain of HIV reverse transcriptase is linked to homodimer formation. Nucleic Acids Res. 2014;42(8):5361-77.

87. Zheng X, Perera L, Mueller GA, DeRose EF, London RE. Asymmetric conformational maturation of HIV-1 reverse transcriptase. eLife. 2015;4:e06359. doi: 10.7554/eLife.06359. PubMed PMID: 26037594; PMCID: 4452869.

88. Xiao Y, Warner LR, Latham MP, Ahn NG, Pardi A. Structure-Based Assignment of Ile, Leu, and Val Methyl Groups in the Active and Inactive Forms of the Mitogen-Activated Protein Kinase Extracellular Signal-Regulated Kinase 2. Biochemistry. 2015;54(28):4307-19. Epub 2015/07/02. doi: 10.1021/acs.biochem.5b00506. PubMed PMID: 26132046; PMCID: PMC4943881.

89. Wiesner S, Sprangers R. Methyl groups as NMR probes for biomolecular interactions. Current opinion in structural biology. 2015;35:60-7. Epub 2015/09/26. doi: 10.1016/j.sbi.2015.08.010. PubMed PMID: 26407236.

90. Abram ME, Parniak MA. Virion instability of human immunodeficiency virus type 1 reverse transcriptase (RT) mutated in the protease cleavage site between RT p51 and the RT RNase H domain. J Virol. 2005;79(18):11952-61.

91. Abram ME, Sarafianos SG, Parniak MA. The mutation T477A in HIV-1 reverse transcriptase (RT) restores normal proteolytic processing of RT in virus with Gag-Pol mutated in the p51-RNH cleavage site. Retrovirology. 2010;7:6.

92. Zheng XH, Pedersen LC, Gabel SA, Mueller GA, Cuneo MJ, DeRose EF, Krahn JM, London RE. Selective unfolding of one Ribonuclease H domain of HIV reverse transcriptase is linked to homodimer formation. Nucleic Acids Res. 2014;42(8):5361-77. doi: 10.1093/nar/gku143. PubMed PMID: WOS:000336092300059.

93. Zheng XH, Pedersen LC, Gabel SA, Mueller GA, DeRose EF, London RE. Unfolding the HIV-1 reverse transcriptase RNase H domain - how to lose a molecular tug-of-war. Nucleic Acids Res. 2016;44(4):1776-88. doi: 10.1093/nar/gkv1538. PubMed PMID: WOS:000371519700034.

94. Giege R, Juhling F, Putz J, Stadler P, Sauter C, Florentz C. Structure of transfer RNAs: similarity and variability. Wiley Interdiscip Rev RNA. 2012;3(1):37-61. Epub 2011/10/01. doi: 10.1002/wrna.103. PubMed PMID: 21957054.

95. Slack RL, Spiriti J, Ahn J, Parniak MA, Zuckerman DM, Ishima R. Structural integrity of the ribonuclease H domain in HIV-1 reverse transcriptase. Proteins. 2015;83(8):1526-38. Epub 2015/06/11. doi: 10.1002/prot.24843. PubMed PMID: 26061827; PMCID: PMC4509971.

96. Chattopadhyay D, Evans DB, Deibel MR, Jr., Vosters AF, Eckenrode FM, Einspahr HM, Hui JO, Tomasselli AG, Zurcher-Neely HA, Heinrikson RL, et al. Purification and characterization of heterodimeric human immunodeficiency virus type 1 (HIV-1) reverse transcriptase produced by in vitro processing of p66 with recombinant HIV-1 protease. J Biol Chem. 1992;267(20):14227-32. Epub 1992/07/25. PubMed PMID: 1378437.

97. Sharma SK, Fan N, Evans DB. Human immunodeficiency virus type 1 (HIV-1) recombinant reverse transcriptase. Asymmetry in p66 subunits of the p66/p66 homodimer. FEBS Lett. 1994;343(2):125-30.

98. Katz RA, Skalka AM. The retroviral enzymes. Ann Rev Biochem. 1994;63:133-73.

99. Sluis-Cremer N, Arion D, Abram ME, Parniak MA. Proteolytic processing of an HIV-1 pol polyprotein precursor: insights into the mechanism of reverse transcriptase p66/p51 heterodimer formation. Int J Biochem Cell Biol. 2004;36(9):1836-47.

100. Herschhorn A, Hizi A. Retroviral reverse transcriptases. Cell Mol Life Sci 2010;67(16):2717-47.

101. Hizi A, Herschhorn A. Retroviral reverse transcriptases (other than those of HIV-1 and murine leukemia virus): a comparison of their molecular and biochemical properties. Virus Res. 2008;134(1-2):203-20.

102. Jacobo-Molina A, Arnold E. HIV reverse transcriptase structure-function relationships. . Biochemistry. 1991;30:6351-6.

103. Hostomska Z, Matthews DA, Davies JFn, Nodes BR, Hostomsky Z. Proteolytic release and crystallization of the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase. J Biol Chem. 1991;266(22):14697-702.

104. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science. 1992;256(5065):1783-90.

105. Jacobo-Molina A, Ding J, Nanni RG, Clark ADJ, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, Hizi A, Hughes SH, Arnold E. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proc Natl Acad Sci U S A. 1993;90(13):6320-4.

106. Pari K, Mueller GA, DeRose EF, Kirby TW, London RE. Solution structure of the RNase H domain of the HIV-1 reverse transcriptase in the presence of magnesium. Biochemistry. 2003;42(3):639-50.

107. Boso G, Orvell C, Somia NV. The Nature of the N-Terminal Amino Acid Residue of HIV-1 RNase H Is Critical for the Stability of Reverse Transcriptase in Viral Particles. Journal of Virology. 2015;89(2):1286-97. doi: 10.1128/Jvi.02312-14. PubMed PMID: WOS:000347178900034.

108. Smith CA, Kortemme T. Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. J Mol Biol. 2008;380(4):742-56.

109. Smith CA, Kortemme T. Predicting the tolerated sequences for proteins and protein interfaces using RosettaBackrub flexible backbone design. PLoS One. 2011;6(7):e20451.

110. Christen MT, Menon L, Myshakina NA, Ahn J, Parniak MA, Ishima R. Structural Basis of the Allosteric Inhibitor Interaction on the HIV-1 Reverse Transcriptase RNase H domain. Chem Biol Drug Des. 2012;80(5):706-16.

111. Ahn J, Byeon IJ, Dharmasena S, Huber K, Concel J, Gronenborn AM, Sluis-Cremer N. The RNA binding protein HuR does not interact directly with HIV-1 reverse transcriptase and does not affect reverse transcription in vitro. Retrovirology. 2010;7:40.

112. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. J Biomol NMR. 1995;6(3):277-93. PubMed PMID: ISI:A1995TH72500006.

113. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinás M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. Proteins: Structure, Function, and Bioinformatics. 2005;59:687-96.

114. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, Carver T, Asel E, Springer BA, Lane P, Salemme FR. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Biomol Screen. 2001;6(6):429-40. doi: Doi 10.1177/108705710100600609. PubMed PMID: WOS:000173150900007.

115. Lavinder JJ, Hari SB, Sullivan BJ, Magliery TJ. High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. J Am Chem Soc. 2009;131(11):3794-+. doi: 10.1021/ja8049063. PubMed PMID: WOS:000264792700002.

116. Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc. 2007;2(9):2212-21. doi: 10.1038/nprot.2007.321. PubMed PMID: WOS:000253139600019.

117. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. Journal of Computational Chemistry. 1983;4(2):187-217.

118. Brooks BR, Brooks CL, III, MacKerell AD, Jr., Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. CHARMM: The biomolecular simulation program. Journal of Computational Chemistry. 2009;30(10):1545-614.

119. Kalé LV, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K. NAMD2: Greater scalability for parallel molecular dynamics. Journal of Computational Physics. 1999;151(1):283-312.

120. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry. 2005;26:1781-802.

121. Hsiou Y, Ding J, Das K, Clark ADJ, Hughes SH, Arnold E. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure. 1996;4(7):853-60.

122. Best RB, Mittal J, Feig M, MacKerell AD, Jr. Inclusion of Many-Body Effects in the Additive CHARMM Protein CMAP Potential Results in Enhanced Cooperativity of alpha-Helix and beta-Hairpin Formation. Biophysical Journal. 2012;103(5):1045-51.

123. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD, Jr. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. Journal of Chemical Theory and Computation. 2012;8(9):3257-73.

124. Lee MS, Feig M, Salsbury FR, Brooks III CL. New analytical approximation to the standard molecular volume definition and its application to generalized Born calculations. Journal of Computational Chemistry. 2003;25:265-84.

125. Lee MS, Salsbury FR, Brooks III CL. Novel generalized Born methods. J Chem Phys. 2002;116:10606-14.

126. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics. 1983;79(2):926-35.

127. Grubmüller H, Heymann B, Tavan P. Ligand Binding: Molecular Mechanics Calculation of the Streptavidin-Biotin Rupture Force. Science. 1996;271:997-9.

128. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. Journal of Computational Physics. 1977;23(3):327-41. PubMed PMID: 18.

129. Miyamoto S, Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. Journal of Computational Chemistry. 1992;13(8):952-62.

130. Aqvist J. A SIMPLE WAY TO CALCULATE THE AXIS OF AN ALPHA-HELIX. Computers & Chemistry. 1986;10(2):97-9.

131. Lauck F, Smith CA, Friedland GF, Humphris EL, Kortemme T. RosettaBackrub--a web server for flexible backbone protein structure modeling and design. Nucleic Acids Res. 2010;38:W569-75.

132. Humphris-Narayanan E, Akiva E, Varela R, Ó Conchúir S, Kortemme T. Prediction of mutational tolerance in HIV-1 protease and reverse transcriptase using flexible backbone protein design. PLoS Comput Biol. 2012;8(8):e1002639.

133. Christen MT, Menon L, Myshakina NS, Ahn J, Parniak MA, Ishima R. Structural basis of the allosteric inhibitor interaction on the HIV-1 reverse transcriptase RNase H domain. Chem Biol Drug Des. 2012;80(5):706-16. doi: 10.1111/cbdd.12010. PubMed PMID: 22846652; PMCID: 3465473.

134. Powers R, Clore GM, Stahl SJ, Wingfield PT, Gronenborn AM. Analysis of the backbone dynamics of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using nitrogen-15 relaxation measurements. Biochemistry. 1992;31(38):9150-7.

135. Yan J, Wu H, Tom T, Brodsky O, Maegley K. Targeting Divalent Metal Ions at the Active Site of the HIV-1 RNase H Domain: NMR Studies on the Interactions of Divalent Metal Ions with RNase H and Its Inhibitors. Am J Anal Chem. 2011;2:639-49.

136. Chung S, Miller JT, Lapkouski M, Tian L, Yang W, Le Grice SFJ. Examining the Role of the HIV-1 Reverse Transcriptase p51 Subunit in Positioning and Hydrolysis of RNA/DNA Hybrids. Journal of Biological Chemistry. 2013;288(22):16177-84. doi: 10.1074/jbc.M113.465641. PubMed PMID: WOS:000319822300068.

137. Dunn LL, Boyer PL, Clark PK, Hughes SH. Mutations in HIV-1 reverse transcriptase cause misfolding and miscleavage by the viral protease. Virology. 2013;444(1-2):241-9. doi: 10.1016/j.virol.2013.06.017. PubMed PMID: WOS:000323873600028.

138. Freddolino PL, Harrison CB, Liu Y, Schulten K. Challenges in protein-folding simulations. Nat Phys. 2010;6(10):751-8.

139. Grossfield A, Zuckerman DM. Quantifying uncertainty and sampling quality in biomolecular simulations. Annual Reports in Computational Chemistry. 2009;5:23-48.

140. Zuckerman DM. Equilibrium Sampling in Biomolecular Simulation. Annual Review of Biophysics. 2011;40:41-62.

141. van der Vaart A. Simulation of conformational transitions. Theor Chem Acc. 2006;116:183-93.

142. Jäger J, Smerdon S, Wang J, Boisvert DC, Steitz TA. Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion. Structure. 1994;2(9):869-76.

143. Unge T, Knight S, Bhikhabhai R, Lövgren S, Dauter Z, Wilson K, Strandberg B. 2.2 A resolution structure of the amino-terminal half of HIV-1 reverse transcriptase (fingers and palm subdomains). Structure. 1994;2(10):953-61.

144. Tomasselli AG, Sarcich JL, Barrett LJ, Reardon IM, Howe WJ, Evans DB, Sharma SK, Heinrikson RL. Human immunodeficiency virus type-1 reverse transcriptase and ribonuclease H as substrates of the viral protease. Protein Sci. 1993;2(12):2167-76.

145. Wlodawer A, Erickson JW. Structure-based inhibitors of HIV-1 protease. Annu Rev Biochem. 1993;62:543-85.

146. Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science. 1994;266(5193):1981-6.

147. Hill CP, Worthylake D, Bancroft DP, Christensen AM, Sundquist WI. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. Proc Natl Acad Sci U S A. 1996;93(7):3099-104.

148. Momany C, Kovari LC, Prongay AJ, Keller W, Gitti RK, Lee BM, Gorbalenya AE, Tong L, McClure J, Ehrlich LS, Summers MF, Carter C, Rossmann MG. Crystal structure of dimeric HIV-1 capsid protein. Nat Struct Biol. 1996;3(9):763-70.

149. Wang J, Smerdon SJ, Jager J, Kohlstaedt LA, Rice PA, Friedman JM, Steitz TA. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. Proc Natl Acad Sci U S A. 1994;91:7242-6.

150. Keck JL, Marqusee S. Substitution of a Highly Basic Helix Loop Sequence into the Rnase-H Domain of Human-Immunodeficiency-Virus Reverse-Transcriptase Restores Its Mn2+-Dependent Rnase-H Activity. P Natl Acad Sci USA. 1995;92(7):2740-4. doi: DOI 10.1073/pnas.92.7.2740. PubMed PMID: WOS:A1995QP88900068.

151. Rhee SY, Kantor R, Katzenstein DA, Camacho R, Morris L, Sirivichayakul S, Jorgensen L, Brigido LF, Schapiro JM, Shafer RW, Group ftiNSBH-W. HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. AIDS. 2006;20(5):643-51.

152. Côté HC, Brumme ZL, Harrigan PR. Human immunodeficiency virus type 1 protease cleavage site mutations associated with protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir. J Virol. 2001;75(2):589-94.

153. Schulze T, Nawrath M, Moelling K. Cleavage of the HIV-1 p66 reverse transcriptase/RNase H by the p9 protease in vitro generates active p15 RNase H. Arch Virol. 1991;118(3-4):179-88.

154. Pettit SC, Lindquist JN, Kaplan AH, Swanstrom R. Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. Retrovirology. 2005;2:66.

155. Edlund U, Grahn H. Multivariate data analysis of NMR data. Journal of Pharmaceutical and Biomedical Analysis. 1991;9(8):655-8. doi: 10.1016/0731-7085(91)80191-b.

156. Stoyanova R, Brown TR. NMR spectral quantitation by principal component analysis. NMR in biomedicine. 2001;14(4):271-7. doi: 10.1002/nbm.700. PubMed PMID: 11410945.

157. Stoyanova R, Brown TR. NMR spectral quantitation by principal component analysis. III. A generalized procedure for determination of lineshape variations. Journal of magnetic resonance. 2002;154(2):163-75. doi: 10.1006/jmre.2001.2486. PubMed PMID: 11846573.

158. Stoyanova R, Kuesel AC, Brown TR. Application of Principal-Component Analysis for NMR Spectral Quantitation. Journal of Magnetic Resonance, Series A. 1995;115(2):265-9. doi: 10.1006/jmra.1995.1177.

159. Hiltunen Y, Heiniemi E, Alakorpela M. Lipoprotein-Lipid Quantification by Neural-Network Analysis of 1H-NMR Data from Human Blood Plasma. Journal of Magnetic Resonance, Series B. 1995;106(2):191-4. doi: 10.1006/jmrb.1995.1032.

160. Fedrigo M, Fogolari F, Viglino P, Esposito G. Improved Processing of Selective NMR Spectra of Biopolymers by Separation of Noise and Signal Subspaces through Singular-Value Decomposition. Journal of Magnetic Resonance, Series B. 1996;113(2):160-6. doi: 10.1006/jmrb.1996.0169.

161. Ochs MF, Stoyanova RS, Arias-Mendoza F, Brown TR. A new method for spectral decomposition using a bilinear Bayesian approach. Journal of magnetic resonance. 1999;137(1):161-76. doi: 10.1006/jmre.1998.1639. PubMed PMID: 10053145.

162. Lindon JC, Holmes E, Nicholson JK. Metabonomics techniques and applications to pharmaceutical research & development. Pharmaceutical research. 2006;23(6):1075-88. doi: 10.1007/s11095-006-0025-z. PubMed PMID: 16715371.

163. Sakurai K, Goto Y. Principal component analysis of the pH-dependent conformational transitions of bovine beta-lactoglobulin monitored by heteronuclear NMR. Proc Natl Acad Sci U S A. 2007;104(39):15346-51.

164. Chen K, Long DS, Lute SC, Levy MJ, Brorson KA, Keire DA. Simple NMR methods for evaluating higher order structures of monoclonal antibody therapeutics with quinary structure. Journal of pharmaceutical and biomedical analysis. 2016;128:398-407. doi: 10.1016/j.jpba.2016.06.007. PubMed PMID: 27344629.

165. Arbogast LW, Delaglio F, Schiel JE, Marino JP. Multivariate Analysis of Two-Dimensional (1)H, (13)C Methyl NMR Spectra of Monoclonal Antibody Therapeutics To Facilitate Assessment of Higher Order Structure. Analytical chemistry. 2017;89(21):11839-45. doi: 10.1021/acs.analchem.7b03571. PubMed PMID: 28937210; PMCID: 5886730.

166. Mueller GA, Pari K, DeRose EF, Kirby TW, London RE. Backbone dynamics of the RNase H domain of HIV-1 reverse transcriptase. Biochemistry. 2004;43(29):9332-42. doi: 10.1021/bi049555n. PubMed PMID: 15260476.

167. Sharaf NG, Brereton AE, Byeon IL, Andrew Karplus P, Gronenborn AM. NMR structure of the HIV-1 reverse transcriptase thumb subdomain. J Biomol NMR. 2016. doi: 10.1007/s10858-016-0077-2. PubMed PMID: 27858311.

168. Bhatt D, Zuckerman DM. Heterogeneous Path Ensembles for Conformational Transitions in Semiatomistic Models of Adenylate Kinase. Journal of Chemical Theory and Computation. 2010;6(11):3527-39. doi: 10.1021/ct100406t. PubMed PMID: WOS:000283884300023.

169. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995;6(3):277-93. PubMed PMID: 8520220.

170. Wapling J, Moore KL, Sonza S, Mak J, Tachedjian G. Mutations that abrogate human immunodeficiency virus type 1 reverse transcriptase dimerization affect maturation of the reverse transcriptase heterodimer. Journal of Virology. 2005;79(16):10247-57. doi: 10.1128/Jvi.79.16.10247-10257.2005. PubMed PMID: WOS:000230884700016.

171. Figueiredo A, Moore KL, Mak J, Sluis-Cremer N, de Bethune MP, Tachedjian G. Potent nonnucleoside reverse transcriptase inhibitors target HIV-1 Gag-Pol. PLoS pathogens. 2006;2(11):e119. doi: 10.1371/journal.ppat.0020119. PubMed PMID: 17096588; PMCID: 1635531.

172. Katz RA, Skalka AM. The Retroviral Enzymes. Annual Review of Biochemistry. 1994;63:133-73. doi: DOI 10.1146/annurev.bi.63.070194.001025. PubMed PMID: WOS:A1994NV05900005.

173. Ghosh M, Jacques PS, Rodgers DW, Ottman M, Darlix JL, Le Grice SF. Alterations to the primer grip of p66 HIV-1 reverse transcriptase and their consequences for template-primer utilization. Biochemistry. 1996;35(26):8553-62. Epub 1996/07/02. doi: 10.1021/bi952773j. PubMed PMID: 8679616.

174. Tachedjian G, Aronson HE, de los Santos M, Seehra J, McCoy JM, Goff SP. Role of residues in the tryptophan repeat motif for HIV-1 reverse transcriptase dimerization. J Mol Biol. 2003;326(2):381-96. Epub 2003/02/01. PubMed PMID: 12559908.

175. Tachedjian G, Aronson HE, Goff SP. Analysis of mutations and suppressors affecting interactions between the subunits of the HIV type 1 reverse transcriptase. Proc Natl Acad Sci U S A. 2000;97(12):6334-9. Epub 2000/06/07. PubMed PMID: 10841542; PMCID: PMC18603.

176. Tachedjian G, Moore KL, Goff SP, Sluis-Cremer N. Efavirenz enhances the proteolytic processing of an HIV-1 pol polyprotein precursor and reverse transcriptase homodimer formation. FEBS Lett. 2005;579(2):379-84. Epub 2005/01/12. doi: 10.1016/j.febslet.2004.11.099. PubMed PMID: 15642347.

177. Kleiman L, Caudry S, Boulerice F, Wainberg MA, Parniak MA. Incorporation of tRNA into normal and mutant HIV-1. Biochem Biophys Res Commun. 1991;174(3):1272-80. Epub 1991/02/14. PubMed PMID: 1705120.

178. Jiang M, Mak J, Ladha A, Cohen E, Klein M, Rovinski B, Kleiman L. Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1. J Virol. 1993;67(6):3246-53. Epub 1993/06/01. PubMed PMID: 8497049; PMCID: PMC237665.

179. Kleiman L, Cen S. The tRNALys packaging complex in HIV-1. Int J Biochem Cell Biol. 2004;36(9):1776-86. Epub 2004/06/09. doi: 10.1016/j.biocel.2004.02.022. PubMed PMID: 15183344.

180. Pavon-Eternod M, Wei M, Pan T, Kleiman L. Profiling non-lysyl tRNAs in HIV-1. RNA. 2010;16(2):267-73. Epub 2009/12/17. doi: 10.1261/rna.1928110. PubMed PMID: 20007329; PMCID: PMC2811655.

181. Arion D, Harada R, Li X, Wainberg MA, Parniak MA. HIV-1 reverse transcriptase shows no specificity for the binding of primer tRNA(Lys3). Biochem Biophys Res Commun. 1996;225(3):839-43. Epub 1996/08/23. doi: 10.1006/bbrc.1996.1260. PubMed PMID: 8780699.

182. Chiang CC, Tseng YT, Huang KJ, Pan YY, Wang CT. Mutations in the HIV-1 reverse transcriptase tryptophan repeat motif affect virion maturation and Gag-Pol packaging. Virology. 2012;422(2):278-87. Epub 2011/11/23. doi: 10.1016/j.virol.2011.11.001. PubMed PMID: 22104208.

183. Chiang CC, Wang SM, Pan YY, Huang KJ, Wang CT. A single amino acid substitution in HIV-1 reverse transcriptase significantly reduces virion release. J Virol. 2010;84(2):976-82. Epub 2009/11/06. doi: 10.1128/JVI.01532-09. PubMed PMID: 19889767; PMCID: PMC2798345.

184. Le Grice SF, Gruninger-Leitch F. Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography. Eur J Biochem. 1990;187(2):307-14. Epub 1990/01/26. PubMed PMID: 1688798.

185. Fletcher RS, Holleschak G, Nagy E, Arion D, Borkow G, Gu Z, Wainberg MA, Parniak MA. Single-step purification of recombinant wild-type and mutant HIV-1 reverse transcriptase. Protein Expr Purif. 1996;7(1):27-32. Epub 1996/02/01. doi: 10.1006/prep.1996.0004. PubMed PMID: 9172779.

186. Gong Q, Menon L, Ilina T, Miller LG, Ahn J, Parniak MA, Ishima R. Interaction of HIV-1 reverse transcriptase ribonuclease H with an acylhydrazone inhibitor. Chem Biol Drug Des. 2011;77(1):39-47. doi: 10.1111/j.1747-0285.2010.01052.x. PubMed PMID: 21114787; PMCID: 3320797.

187. Lee T, Laco GS, Torbett BE, Fox HS, Lerner DL, Elder JH, Wong CH. Analysis of the S3 and S3' subsite specificities of feline immunodeficiency virus (FIV) protease: development of a broad-based protease inhibitor efficacious against FIV, SIV, and HIV in vitro and ex vivo. Proc Natl Acad Sci U S A. 1998;95(3):939-44. Epub 1998/03/14. PubMed PMID: 9448264; PMCID: PMC18632.

188. Heaslet H, Kutilek V, Morris GM, Lin YC, Elder JH, Torbett BE, Stout CD. Structural insights into the mechanisms of drug resistance in HIV-1 protease NL4-3. J Mol Biol. 2006;356(4):967-81. Epub 2006/01/13. doi: 10.1016/j.jmb.2005.11.094. PubMed PMID: 16403521.

189. Sluis-Cremer N, Dmitrienko GI, Balzarini J, Camarasa MJ, Parniak MA. Human immunodeficiency virus type 1 reverse transcriptase dimer destabilization by 1-[Spiro[4"-amino-2",2" -dioxo-1",2" -oxathiole-5",3'-[2', 5'-bis-O-(tert-butyldimethylsilyl)-beta-D-ribofuranosyl]]]-3-ethylthy mine. Biochemistry. 2000;39(6):1427-33. Epub 2000/02/24. PubMed PMID: 10684624.

190. Nashed NT, Louis JM, Sayer JM, Wondrak EM, Mora PT, Oroszlan S, Jerina DM. Continuous spectrophotometric assay for retroviral proteases of HIV-1 and AMV. Biochem Biophys Res Commun. 1989;163(2):1079-85. Epub 1989/09/15. PubMed PMID: 2551268.

191. Cabodevilla JF, Odriozola L, Santiago E, Martinez-Irujo JJ. Factors affecting the dimerization of the p66 form of HIV-1 reverse transcriptase. Eur J Biochem. 2001;268(5):1163-72. Epub 2001/03/07. PubMed PMID: 11231267.

192. Porter DJ, Hanlon MH, Carter LH, 3rd, Danger DP, Furfine ES. Effectors of HIV-1 protease peptidolytic activity. Biochemistry. 2001;40(37):11131-9. Epub 2001/09/12. PubMed PMID: 11551211.

193. Lipfert J, Doniach S, Das R, Herschlag D. Understanding nucleic acid-ion interactions. Annu Rev Biochem. 2014;83:813-41. Epub 2014/03/13. doi: 10.1146/annurev-biochem-060409-092720. PubMed PMID: 24606136; PMCID: PMC4384882.

194. Potempa M, Nalivaika E, Ragland D, Lee SK, Schiffer CA, Swanstrom R. A Direct Interaction with RNA Dramatically Enhances the Catalytic Activity of the HIV-1 Protease In Vitro. J Mol Biol. 2015;427(14):2360-78. Epub 2015/05/20. doi: 10.1016/j.jmb.2015.05.007. PubMed PMID: 25986307; PMCID: PMC4465046.

195. Ilina TV, Slack RL, Elder JH, Sarafianos SG, Parniak MA, Ishima R. Effect of tRNA on the Maturation of HIV-1 Reverse Transcriptase. Journal of Molecular Biology. 2018;430(13):1891-900. doi: 10.1016/j.jmb.2018.02.027. PubMed PMID: WOS:000436224800006.

196. Lanchy JM, Ehresmann C, Le Grice SF, Ehresmann B, Marquet R. Binding and kinetic properties of HIV-1 reverse transcriptase markedly differ during initiation and elongation of reverse transcription. The EMBO journal. 1996;15(24):7178-87. Epub 1996/12/16. PubMed PMID: 9003793; PMCID: PMC452545.

197. Andreola ML, Nevinsky GA, Barr PJ, Sarih-Cottin L, Bordier B, Fournier M, Litvak S, Tarrago-Litvak L. Interaction of tRNALys with the p66/p66 form of HIV-1 reverse transcriptase stimulates DNA polymerase and ribonuclease H activities. J Biol Chem. 1992;267(27):19356-62. Epub 1992/09/25. PubMed PMID: 1382072.

198. Robert D, Sallafranque-Andreola ML, Bordier B, Sarih-Cottin L, Tarrago-Litvak L, Graves PV, Barr PJ, Fournier M, Litvak S. Interactions with tRNA(Lys) induce important structural changes in human immunodeficiency virus reverse transcriptase. FEBS Lett. 1990;277(1-2):239-42. Epub 1990/12/17. PubMed PMID: 1702735.

199. Lindhofer H, von der Helm K, Nitschko H. In vivo processing of Pr160gag-pol from human immunodeficiency virus type 1 (HIV) in acutely infected, cultured human T-lymphocytes. Virology. 1995;214(2):624-7. Epub 1995/12/20. doi: 10.1006/viro.1995.0074. PubMed PMID: 8553565.

200. Speck RR, Flexner C, Tian CJ, Fu XF. Comparison of human immunodeficiency virus type 1 Pr55(Gag) and Pr160(Gag-Pol) processing intermediates that accumulate in primary and transformed cells treated with peptidic and nonpeptidic protease inhibitors. Antimicrobial agents and chemotherapy. 2000;44(5):1397-403. doi: Doi 10.1128/Aac.44.5.1397-1403.2000. PubMed PMID: WOS:000086625400052.

201. Chattopadhyay D, Evans DB, Deibel MR, Vosters AF, Eckenrode FM, Einspahr HM, Hui JO, Tomasselli AG, Zurcher-Neely HA, Heinrikson RL, Sharma SK. Purification and characterization of heterodimeric human immunodeficiency virus type 1 (HIV-1) reverse transcriptase produced by in vitro processing of p66 with recombinant HIV-1 protease. J Biol Chem. 1992;267:14227-32.

202. Sharma SK, Fan N, Evans DB. Human immunodeficiency virus type 1 (HIV-1) recombinant reverse transcriptase. Asymmetry in p66 subunits of the p66/p66 homodimer. FEBS Lett. 1994;343(2):125-30. Epub 1994/04/25. PubMed PMID: 7513287.

203. Mattei S, Anders M, Konvalinka J, Krausslich HG, Briggs JA, Muller B. Induced maturation of human immunodeficiency virus. J Virol. 2014;88(23):13722-31. Epub 2014/09/19. doi: 10.1128/JVI.02271-14. PubMed PMID: 25231305; PMCID: PMC4248963.

204. Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH. Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. J Virol. 2004;78(16):8477-85. Epub 2004/07/29. doi: 10.1128/JVI.78.16.8477-8485.2004. PubMed PMID: 15280456; PMCID: PMC479095.

205. Jiang M, Mak J, Wainberg MA, Parniak MA, Cohen E, Kleiman L. Variable tRNA content in HIV-1IIIB. Biochem Biophys Res Commun. 1992;185(3):1005-15. Epub 1992/06/30. PubMed PMID: 1627125.

206. Braz VA, Holladay LA, Barkley MD. Efavirenz binding to HIV-1 reverse transcriptase monomers and dimers. Biochemistry. 2010;49(3):601-10. Epub 2009/12/31. doi: 10.1021/bi901579y. PubMed PMID: 20039714; PMCID: PMC2896556.

207. Cen S, Khorchid A, Javanbakht H, Gabor J, Stello T, Shiba K, Musier-Forsyth K, Kleiman L. Incorporation of lysyl-tRNA synthetase into human immunodeficiency virus type 1. J Virol. 2001;75(11):5043-8. Epub 2001/05/03. doi: 10.1128/JVI.75.11.5043-5048.2001. PubMed PMID: 11333884; PMCID: PMC114908.

208. Cen S, Javanbakht H, Kim S, Shiba K, Craven R, Rein A, Ewalt K, Schimmel P, Musier-Forsyth K, Kleiman L. Retrovirus-specific packaging of aminoacyl-tRNA synthetases with cognate primer tRNAs. J Virol. 2002;76(24):13111-5. Epub 2002/11/20. doi: 10.1128/jvi.76.24.13111-13115.2002. PubMed PMID: 12438642; PMCID: PMC136713.

209. Khorchid A, Javanbakht H, Wise S, Halwani R, Parniak MA, Wainberg MA, Kleiman L. Sequences within Pr160gag-pol affecting the selective packaging of primer tRNA(Lys3) into HIV-1. J Mol Biol. 2000;299(1):17-26. Epub 2000/06/22. doi: 10.1006/jmbi.2000.3709. PubMed PMID: 10860720.

210. Kleiman L, Jones CP, Musier-Forsyth K. Formation of the tRNALys packaging complex in HIV-1. FEBS Lett. 2010;584(2):359-65. Epub 2009/11/17. doi: 10.1016/j.febslet.2009.11.038. PubMed PMID: 19914238; PMCID: PMC2820394.

211. Mak J, Khorchid A, Cao Q, Huang Y, Lowy I, Parniak MA, Prasad VR, Wainberg MA, Kleiman L. Effects of mutations in Pr160gag-pol upon tRNA(Lys3) and Pr160gag-plo incorporation into HIV-1. J Mol Biol. 1997;265(4):419-31. Epub 1997/01/31. doi: 10.1006/jmbi.1996.0742. PubMed PMID: 9034361.

212. Khan SN, Persons JD, Paulsen JL, Guerrero M, Schiffer CA, Kurt-Yilmaz N, Ishima R. Probing Structural Changes among Analogous Inhibitor-Bound Forms of HIV-1 Protease and a Drug-Resistant Mutant in Solution by Nuclear Magnetic Resonance. Biochemistry. 2018;57(10):1652-62. Epub 2018/02/20. doi: 10.1021/acs.biochem.7b01238. PubMed PMID: 29457713; PMCID: PMC5850901.

213. Miller JT, Khvorova A, Scaringe SA, Le Grice SF. Synthetic tRNALys,3 as the replication primer for the HIV-1HXB2 and HIV-1Mal genomes. Nucleic acids research. 2004;32(15):4687-95. Epub 2004/09/03. doi: 10.1093/nar/gkh813. PubMed PMID: 15342789; PMCID: PMC516074.

214. Sherlin LD, Bullock TL, Nissan TA, Perona JJ, Lariviere FJ, Uhlenbeck OC, Scaringe SA. Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. RNA. 2001;7(11):1671-8. Epub 2001/11/27. PubMed PMID: 11720294; PMCID: PMC1370207.

215. Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol. 2004;278:313-52. Epub 2004/08/20. doi: 10.1385/1-59259-809-9:313. PubMed PMID: 15318002.

216. Shi C, Mellors JW. A recombinant retroviral system for rapid in vivo analysis of human immunodeficiency virus type 1 susceptibility to reverse transcriptase inhibitors. Antimicrobial agents and chemotherapy. 1997;41(12):2781-5. Epub 1998/01/07. PubMed PMID: 9420060; PMCID: PMC164210.

217. Giacobbi NS, Sluis-Cremer N. In Vitro Cross-Resistance Profiles of Rilpivirine, Dapivirine, and MIV-150, Nonnucleoside Reverse Transcriptase Inhibitor Microbicides in Clinical Development for the Prevention of HIV-1 Infection. Antimicrobial agents and chemotherapy. 2017;61(7). Epub 2017/05/17. doi: 10.1128/AAC.00277-17. PubMed PMID: 28507107; PMCID: PMC5487620.

218. Bavand MR, Wagner R, Richmond TJ. HIV-1 reverse transcriptase: polymerization properties of the p51 homodimer compared to the p66/p51 heterodimer. Biochemistry. 1993;32(40):10543-52. Epub 1993/10/12. PubMed PMID: 7691176.

219. Clark AD, Jr., Jacobo-Molina A, Clark P, Hughes SH, Arnold E. Crystallization of human immunodeficiency virus type 1 reverse transcriptase with and without nucleic acid substrates, inhibitors, and an antibody Fab fragment. Methods Enzymol. 1995;262:171-85. Epub 1995/01/01. PubMed PMID: 8594346.

220. Zheng X, Mueller GA, Kim K, Perera L, DeRose EF, London RE. Identification of drivers for the metamorphic transition of HIV-1 reverse transcriptase. The Biochemical journal. 2017;474(19):3321-38. doi: 10.1042/BCJ20170480. PubMed PMID: 28811321; PMCID: 5614878.

221. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. (1)H, (13)C and (15)N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. J Biomol NMR. 1995;5(3):332. Epub 1995/04/01. doi: 10.1007/BF00211764. PubMed PMID: 22911508.

222. Nicholson LK, Kay LE, Baldisseri DM, Arango J, Young PE, Bax A, Torchia DA. Dynamics of methyl groups in proteins as studied by proton-detected 13C NMR spectroscopy. Application to the leucine residues of staphylococcal nuclease. Biochemistry. 1992;31(23):5253-63. Epub 1992/06/16. doi: 10.1021/bi00138a003. PubMed PMID: 1606149.

223. Bakhanashvili M, Hizi A. Interaction of the reverse transcriptase of human immunodeficiency virus type 1 with DNA. Biochemistry. 1994;33(40):12222-8. Epub 1994/10/11. doi: 10.1021/bi00206a027. PubMed PMID: 7522556.

224. Hughes SH. Molecular matchmaking: NNRTIs can enhance the dimerization of HIV type 1 reverse transcriptase. Proc Natl Acad Sci U S A. 2001;98(13):6991-2. Epub 2001/06/21. doi: 10.1073/pnas.141222698. PubMed PMID: 11416176; PMCID: PMC34610.

225. Tachedjian G, Orlova M, Sarafianos SG, Arnold E, Goff SP. Nonnucleoside reverse transcriptase inhibitors are chemical enhancers of dimerization of the HIV type 1 reverse transcriptase. Proc Natl Acad Sci U S A. 2001;98(13):7188-93. Epub 2001/06/21. doi: 10.1073/pnas.121055998. PubMed PMID: 11416202; PMCID: PMC34644.

226. Ivetac A, McCammon JA. Molecular recognition in the case of flexible targets. Curr Pharm Des. 2011;17(17):1663-71. Epub 2011/05/31. doi: 10.2174/138161211796355056. PubMed PMID: 21619526; PMCID: PMC3271729.

227. Temiz NA, Bahar I. Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. Proteins. 2002;49(1):61-70. Epub 2002/09/05. doi: 10.1002/prot.10183. PubMed PMID: 12211016.

228. Nam SH, Kim D, Lee MS, Lee D, Kwak TK, Kang M, Ryu J, Kim HJ, Song HE, Choi J, Lee GH, Kim SY, Park SH, Kim DG, Kwon NH, Kim TY, Thiery JP, Kim S, Lee JW. Noncanonical roles of membranous lysyl-tRNA synthetase in transducing cell-substrate signaling for invasive dissemination of colon cancer spheroids in 3D collagen I gels. Oncotarget. 2015;6(25):21655-74. Epub 2015/06/20. doi: 10.18632/oncotarget.4130. PubMed PMID: 26091349; PMCID: PMC4673294.

229. Sarafianos SG, Das K, Tantillo C, Clark AD, Jr., Ding J, Whitcomb JM, Boyer PL, Hughes SH, Arnold E. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. The EMBO journal. 2001;20(6):1449-61. Epub 2001/03/17. doi: 10.1093/emboj/20.6.1449. PubMed PMID: 11250910; PMCID: PMC145536.

230. Cen S, Niu M, Saadatmand J, Guo F, Huang Y, Nabel GJ, Kleiman L. Incorporation of pol into human immunodeficiency virus type 1 Gag virus-like particles occurs independently of the upstream Gag domain in Gag-pol. J Virol. 2004;78(2):1042-9. Epub 2003/12/25. PubMed PMID: 14694138; PMCID: PMC368740.

231. Guo F, Cen S, Niu M, Javanbakht H, Kleiman L. Specific inhibition of the synthesis of human lysyl-tRNA synthetase results in decreases in tRNA(Lys) incorporation, tRNA(3)(Lys) annealing to viral RNA, and viral infectivity in human immunodeficiency virus type 1. J Virol. 2003;77(18):9817-22. Epub 2003/08/28. doi: 10.1128/jvi.77.18.9817-9822.2003. PubMed PMID: 12941890; PMCID: PMC224598.

232. Pettit SC, Clemente JC, Jeung JA, Dunn BM, Kaplan AH. Ordered processing of the human immunodeficiency virus type 1 GagPol precursor is influenced by the context of the embedded viral protease. J Virol. 2005;79(16):10601-7. Epub 2005/07/30. doi: 10.1128/JVI.79.16.10601-10607.2005. PubMed PMID: 16051852; PMCID: PMC1182631.

233. Friedrich BM, Dziuba N, Li G, Endsley MA, Murray JL, Ferguson MR. Host factors mediating HIV-1 replication. Virus Res. 2011;161(2):101-14. Epub 2011/08/30. doi: 10.1016/j.virusres.2011.08.001. PubMed PMID: 21871504.

234. Goff SP. Host factors exploited by retroviruses. Nat Rev Microbiol. 2007;5(4):253-63. Epub 2007/02/28. doi: 10.1038/nrmicro1541. PubMed PMID: 17325726.

235. Javanbakht H, Halwani R, Cen S, Saadatmand J, Musier-Forsyth K, Gottlinger H, Kleiman L. The interaction between HIV-1 Gag and human lysyl-tRNA synthetase during viral assembly. J Biol Chem. 2003;278(30):27644-51. Epub 2003/05/21. doi: 10.1074/jbc.M301840200. PubMed PMID: 12756246.

236. Huang Y, Mak J, Cao Q, Li Z, Wainberg MA, Kleiman L. Incorporation of excess wild-type and mutant tRNA(3Lys) into human immunodeficiency virus type 1. J Virol. 1994;68(12):7676-83. Epub 1994/12/01. PubMed PMID: 7966556; PMCID: PMC237227.

237. Rigourd M, Bec G, Benas P, Le Grice SF, Ehresmann B, Ehresmann C, Marquet R. Effects of tRNA 3 Lys aminoacylation on the initiation of HIV-1 reverse transcription. Biochimie. 2003;85(5):521-5. Epub 2003/05/24. PubMed PMID: 12763311.

238. Duchon AA, St Gelais C, Titkemeier N, Hatterschide J, Wu L, Musier-Forsyth K. HIV-1 Exploits a Dynamic Multi-aminoacyl-tRNA Synthetase Complex To Enhance Viral Replication. J Virol. 2017;91(21). Epub 2017/08/18. doi: 10.1128/JVI.01240-17. PubMed PMID: 28814526; PMCID: PMC5640844.

239. Zaitseva L, Myers R, Fassati A. tRNAs promote nuclear import of HIV-1 intracellular reverse transcription complexes. PLoS Biol. 2006;4(10):e332. Epub 2006/10/06. doi: 10.1371/journal.pbio.0040332. PubMed PMID: 17020411; PMCID: PMC1584419.