Single-Molecule Fluorescence Studies of HIV-1 Reverse Transcriptase/Nucleic Acid Interactions

Ryan A. Marko

Degree of Master of Science

Department of Chemistry
Faculty of Science
McGill University
Montreal, Quebec, Canada
June 5th, 2013

A thesis submitted to McGill University in partial fulfillment of the requirements for the Master of Science degree.

© Ryan A. Marko, 2013. All rights reserved.

Dedication

This thesis is dedicated to Michael Marko, Ginny Marko, and Tiam Lie (Opa)

Abstract

The human immunodeficiency virus type-1 (HIV-1) is an epidemic affecting approximately 34 million people worldwide. The HIV-1 enzyme reverse transcriptase (RT) plays an essential role in the viral replication cycle by catalyzing the conversion of HIV-1's single-stranded RNA genome into double-stranded DNA. RT is highly dynamic and possesses both DNA polymerase and ribonuclease H activities. While the biologically relevant form of RT is a heterodimeric enzyme consisting of a 66 kDa subunit (p66) and a 51 kDa subunit (p51) (i.e., p66-p51), both p66-p66 and p51-51 homodimers are also catalytically active. Currently, over half of the clinically approved drugs to treat HIV-1 infection target RT, however, the development of drug resistance results in drug therapy failure over time. A better understanding of the complex reverse transcription process and of the dynamic nature of RT is required in order to identify new avenues for therapeutic intervention.

The recent emergence of single-molecule spectroscopy has been instrumental towards elucidating the astonishing complexity of fundamental biological processes such as genome replication. Single-molecule techniques possess the unique ability to report on molecular dynamics that would otherwise be hidden in conventional ensemble measurements and have been implemented successfully to better understand complex enzyme mechanisms.

This thesis describes the development of a single-molecule fluorescence spectroscopic assay to directly visualize the binding/unbinding of RT to a model fluorescently labeled DNA-DNA substrate. Our single-molecule assay relies on a phenomenon called protein-induced fluorescence enhancement (PIFE) and does not require chemical modification of RT. Using our single-molecule PIFE assay, we obtained equilibrium dissociation constants and association/dissociation rate constants of hetero- and homodimeric forms of RT binding to a double-stranded DNA substrate and observed discrepancies in the association/dissociation dynamics among the three RT isoforms. Furthermore, we show that the apparent binding affinity of p51-p51 is to a great extent time-dependent and is more likely determined by the dimer dissociation into constituent monomers rather than the

intrinsic binding affinity of dimeric RT. Finally, we describe the potential versatility of our assay as a tool to study RT enzymatic activity (e.g., RNase H activity) and conformational dynamics.

Overall, our single-molecule PIFE assay represents a promising platform that is primed to help us understand the intricate nature of the viral reverse transcription process. We hope to extend our single-molecule methodology to gain insight to the molecular basis of drug resistance in hopes to identifying novel therapeutic targets against HIV-1.

Résumé

Le virus de l'immunodéficience humaine de type 1 (VIH-1) est une épidémie qui affecte environ 34 millions de personnes à travers le monde. L'enzyme transcriptase inverse (TI) du VIH-1 joue un rôle essentiel dans le cycle de réplication virale en catalysant la conversion du génome de l'ARN simple brin du VIH-1 en de l'ADN double brin. L'enzyme TI, qui est associée à de procédés très dynamiques, possède à la fois des fonctions enzymatiques de polymérase d'ADN et de ribonuclease H. Bien que la forme biologiquement pertinente de TI est une enzyme hétérodimère constituée d'une sous-unité de 66 kDa (p66) et une sous-unité de 51 kDa (p51) (c.-à-d. p66-p51), les homodimères p66-p66 et p51-51 sont également catalytiquement actif. Actuellement, plus de la moitié des médicaments cliniquement approuvés pour le traitement du VIH-1 cible l'enzyme TI. Cependant, le développement de résistance aux médicaments par cet enzyme mène à un échec thérapeutique qui augmente avec le temps. Afin d'identifier de nouvelles pistes pour l'intervention thérapeutique, une meilleure compréhension du processus complexe de transcription inverse et de la nature dynamique de l'enzyme TI sont souhaitables.

L'émergence récente de la spectroscopie à molécule unique a contribué à élucider l'étonnante complexité de processus biologiques fondamentaux tels que la réplication du génome. Les techniques spectroscopiques à molécule unique possèdent la capacité unique d'observer la dynamique moléculaire qui serait autrement non-observable lors de mesures d'ensemble conventionnelles. Par exemple, celles-ci ont été appliquées afin de mieux comprendre des mécanismes enzymatiques complexes.

Cette thèse décrit le développement d'une analyse spectroscopique par fluorescence à molécule unique qui permet de visualiser l'association/dissociation de l'enzyme TI avec un substrat modèle d'ADN-ADN marqué avec un fluorophore. Notre analyse spectroscopique à molécule unique utilise le phénomène d'augmentation de la fluorescence induite par les protéines (PIFE) et ne nécessite pas de modification chimique de l'enzyme TI. Grâce à notre analyse de molécules uniques utilisant PIFE, nous avons obtenu les constantes de

dissociation et les constantes de vitesse d'association / dissociation des formes hétéro- et homodimériques de l’enzyme TI se liant à un substrat d'ADN double brin et nous avons observé les divergences dans la dynamique de l’association / dissociation entre les trois isoformes de l’enzyme TI. En plus, nous démontrons que l'affinité apparente de l’association de p51-p51 varie en fonction du temps et que cette association est plus probablement déterminée par la dissociation du dimère en monomères constitutifs plutôt que l'affinité de liaison intrinsèque du dimère TI. Enfin, nous décrivons la polyvalence potentielle de notre analyse en tant qu’outil pour étudier l'activité enzymatique de TI (par exemple, l'activité ribonucléase H) et son changement dynamique de conformations.

Dans l'ensemble, notre analyse par fluorescence de molécules uniques par l’effet PIFE représente une plate-forme prometteuse qui pourrait aider à comprendre la nature complexe du processus de transcription inverse virale. Nous espérons appliquer notre méthodologie par fluorescence de molécules uniques afin de mieux comprendre la base moléculaire de la résistance aux médicaments dans l'espoir d'identifier de nouvelles cibles thérapeutiques contre le VIH-1.

Acknowledgments

First, I would like to thank my research advisor, Professor Gonzalo Cosa, for his excellent guidance, patience, and unconditional support. I've had a fulfilling and enriching experience in the Cosa group and I wish him the best of luck in the future.

I thank past and present members of the Cosa group for all their support and for providing an incredibly pleasant working environment: Christina Calver, Dr. Mathieu Frenette, Viktorija Glembockyté, Robert Godin, Lana Greene, General Amani Hariri, Dr. Pierre Karam, Dr. Katerina Krumova, Ricky Lincoln, Dr. Hsiao-Wei Liu, and Laura Smith. Additionally, I would like to thank Chris Ablenas, Dr. Maryam Ehteshami, Suzanne McCormick, and Professor Matthias Götte for their support and collaboration on this project.

I am especially grateful to Dr. Hsiao-Wei Liu for training and guiding me over the course of this project. Your expertise and patience is greatly appreciated.

Most importantly, I want to thank my family, Mei-Lan Morton, Mark Morton, Alec Marko, Colin Marko, and Kahoko Lie (Oma) for their everlasting love and support.

Contributions of Authors

The contents of Chapter 3 have been published in the Journal of Physical Chemistry B (Marko, R. A.; Liu, H.-W.; Ablenas, C. J.; Ehteshami, M.; Gotte, M.; Cosa, G. *Journal of Physical Chemistry B* **2013**, *117*, 4560-4567.) and have been reproduced with permission. The contributions of other authors are described below:

All wild-type and mutant HIV-1 reverse transcriptase proteins were expressed and purified by Mr. Chris Ablenas, Dr. Maryam Ehteshami, and Ms. Suzanne McCormick from the laboratory of Professor Matthias Götte (McGill University, Montreal, QC).

Chapter 3: Single-molecule protein-induced fluorescence enhancement experiments conducted with 300 nM RTp66-p51 were conducted and analyzed by Dr. Hsiao-Wei Liu (Figure 3.3B and Figure 3.8).

Chapter 4: The polyacrylamide gel electrophoresis assay for reverse transcriptase RNase H activity was conducted in collaboration with Mr. Chris Ablenas (Figure 4.5).

Table of Contents

1. Introduction 1

1.1 Human Immunodeficiency Virus-1 (HIV-1) 1
  1.1.1 HIV-1 Life Cycle 1
  1.1.2 Reverse Transcription 4
  1.1.3 Reverse Transcriptase (RT) 7
    1.1.3.1 Structure 7
    1.1.3.2 Function 9
  1.1.4 RT as a Drug Target: Mechanisms of Inhibition and Drug Resistance 12
    1.1.4.1 Nucleoside RT Inhibitors (NRTIs) 13
    1.1.4.2 NRTIs and Drug Resistance 14
    1.1.4.3 Non-Nucleoside RT Inhibitors (NNRTIs) 15
    1.1.4.4 NNRTIs and Drug Resistance 16
    1.1.4.5 Connection Domain Mutations and Drug Resistance 17
1.2 Fluorescence Techniques: Applications In Single-Molecule Spectroscopy 18
  1.2.1 Fluorescence & The Jablonski Diagram 18
    1.2.1.1 Non-Radiative Decay 20
    1.2.1.2 Emission Quantum Yield 20
    1.2.1.3 Förster Resonance Energy Transfer (FRET) 21
  1.2.2 Fluorescence and Single-Molecule Detection 22
  1.2.3 Single-Molecule Versus Ensemble Measurements 22
  1.2.4 The Single-Molecule Advantage 23
  1.2.5 Towards the Era of Single-Molecule Biology 24
    1.2.5.1 Real-Time Single-Molecule Fluorescence Imaging in Aqueous Solution 26
    1.2.5.2 Single-Molecule Enzymology 27
    1.2.5.3 Single-Molecule FRET 29
    1.2.5.4 Single-Molecule Protein-Induced Fluorescence Enhancement 30
  1.2.6 Summary 32
1.3 Research Goals 32

2. General Methodology 34

2.1 Total Internal Reflection Fluorescence Microscopy 34
  2.1.1 Total Internal Reflection 34
  2.1.2 Total Internal Reflection Microscopy 35
  2.1.3 Prism-Based TIRF 38
  2.1.4 Objective-Based TIRF 38
  2.1.5 Detection 39
2.2 Our TIRF Microscope 40
2.3 Fluorophores for SMS Studies 41
2.4 Coverslip and Sample Preparation 44
  2.4.1 Cleaning and Aminosilanization 44
  2.4.2 Surface Functionalization with PEG/biotin-PEG 45
  2.4.3 Oxygen Scavenging System 47

3. Binding Kinetics and Affinities of Heterodimeric versus Homodimeric HIV-1 Reverse Transcriptase on DNA-DNA Substrates at the Single-Molecule Level 49

3.1 Preface 49
  3.1.1 Single-Molecule FRET Studies of HIV-1 RT 50
3.2 Introduction 51

3.3 Results and Discussion 54
  3.3.1 Binding of p66-p51 Heterodimers with Cy3-Primer-Template Substrates 54
    3.3.1.1 Fluorescence Enhancement of Cy3-dP-T Induced by RT Binding 55
    3.3.1.2 Steady Fluorescence of Cy3-dP-T in the Absence of RT 57
    3.3.1.3 Formation of Higher-Order RT/Cy3-dP-T Complexes at High RTp66-p51 Concentrations 58
  3.3.2 Affinity and Kinetics of RTp66-p51 at Different Concentrations 59
  3.3.3 Binding Kinetics and Affinity for RT Hetero- and Homodimers 62
  3.3.4 Binding Affinities of Homodimeric RTp66-p66 and RTp51-p51 as a Function of Time 66
3.4 Conclusions 69
3.5 Experimental 70
  3.5.1 Protein Expression and Purification 70
  3.5.2 Preparation of Primer/Template Strands 70
  3.5.3 Single-Molecule Sample Preparation and Imaging 71
  3.5.4 Electrophoretic Mobility Shift Assays (EMSA) 73
3.6 Appendix 74
  3.6.1 On/Off-time Analysis 74
  3.6.2 Determination of Apparent Dissociation Constants (K_d) by Ensemble Normalized Intensity Histograms 75
  3.6.3 Optimization RTp51 Binding to Cy3-dP-T 75

4. Extending SM-PIFE to HIV-1 RT Enzymatic Activity and Conformational Dynamics 81
4.1 SM-PIFE and RT Dynamics: Towards Characterizing the Interplay Between Polymerase and RNase H Activity 81
4.2 SM-PIFE Studies of RT RNase H Activity 82
  4.2.1 RT RNase H Cleaves RNA Close to the 3' End of 3'Cy3 RNA/DNA Substrates 86
    4.2.1.1 Gel Assay of RT RNase H Activity: Experimental 86
    4.2.1.2 Gel Assay of RT RNase H Activity: Results 87
  4.2.2 SM-PIFE Characterization of RT RNase H Activity 89
  4.2.3 Future Work 90
4.3 Probing Polymerase- and RNase H-competent Binding Orientations of RT: The E478Q Mutant 90
4.4 Multi-Step Enhancement of Cy3-rP-T Fluorescence in the Presence of RT E478Q and Nevirapine (NVP) 94
4.5 Future Avenues 97
  4.5.1 SM-PIFE Studies of the RT Polymerase Domain 97
  4.5.2 Towards Characterizing the Interplay Between RT Polymerase and RNase H Activity 101

5. References 103

List of Figures

Figure 1.1 Schematic of HIV-1 life cycle. The cycle begins with host CD4 receptor and coreceptor binding to the viral gp120 (1). Following binding, membrane fusion occurs and the contents of the virus are released into the cytoplasm of the cell (2). Uncoating of the viral capsid releases two copies of the viral RNA genome. Reverse transcriptase converts the RNA genome into double-stranded DNA (3). The newly formed DNA is transported into the nucleus and integrated into a host chromosome by the viral integrase (4). The resulting proviral DNA is transcribed (5) and translated (6) by the host cellular machinery and the viral components are assembled at the plasma membrane (7). Finally, the virion buds from the host cell and undergoes maturation, resulting in a new infectious virus (8). Figure adapted with permission from reference [6]. **2**

Figure 1.2 Schematic of the HIV-1 viral genome. Viral RNA possesses a 5' cap and a 3' poly-A tail, similar to cellular mRNA. Sequences unique to the 5' and 3' (U5 and U3, respectively) sit adjacent to identical repeat (R) sequences. The primer-binding site (PBS) is complementary in sequence to $tRNA_3^{lys}$, the primer for minus-strand DNA synthesis. The polypurine tract (PPT) sequence serves as a primer for plus-stranded DNA synthesis. **4**

Figure 1.3 Steps of reverse transcription. Viral RNA is depicted in light blue, while newly synthesized DNA is shown in dark blue. The dashed lines represent degraded RNA. (A) A host cell-derived $tRNA_3^{lys}$ binds to the primer-binding site (pbs) of the RNA genome. (B) RT catalyzes minus-strand DNA synthesis and simultaneous RNA degradation until the 5' end is reached, generating minus-strand strong-stop DNA (–sssDNA) (C) The first strand-transfer step can be inter- or intramolecular and occurs between the complementary R sequences of the –sssDNA and viral RNA. (D) Minus-strand DNA synthesis continues until the 5' end, along with RNase H-catalyzed RNA degradation. Short purine-rich sequences (ppt) are resistant to cleavage and act as primers for plus-strand DNA synthesis. (E) Plus-strand synthesis begins from a PPT primer and continues until a portion of the $tRNA_3^{lys}$ primer is transcribed before being degraded by RT RNase H. (F) Minus- and plus-strand DNA synthesis continue until a full double-stranded DNA copy of the viral RNA is obtained. Figure reproduced with permission from reference [15]. **6**

Figure 1.4 Structure of the HIV-1 RT p66-p51 heterodimer bound to a DNA/DNA duplex (grey). The p66 domain is composed of fingers (blue), palm (red), thumb (green), connection (yellow), and RNase H (orange) domains. The p51 subunit is depicted in brown. The structure shown here was generated using PyMOL molecular visualization software (www.pymol.org) and RCSB protein databank structure access code 2HMI from reference [21]. **8**

Figure 1.5 Binding of nucleic acid by RT requires a conformational change in the thumb domain in order to expose the nucleic acid binding cleft. The fingers domain also possesses conformational flexibility and closes down tightly over an RT/nucleic acid/dNTP ternary complex. Figure reproduced with permission from reference [15]. **9**

Figure 1.6 Steps involved in RT-catalyzed DNA polymerization. Polymerization begins with RT binding to primer/template (P/T) complex (1), followed by formation of a ternary complex with complementary dNTP (2). Nucleotide incorporation (3) results in the loss of pyrophosphate (PP_i) (4). RT must translocate one base pair in order to accept the next incoming dNTP. Figure adapted from reference [16]. **10**

Figure 1.7 Mechanism of RT RNase H-catalyzed RNA cleavage. The reaction occurs via a two divalent metal ion-dependent S_n2 mechanism with water as the nucleophile. The two divalent metals, A and B, are coordinated by a highly conserved DEDD motif in the active site. Figure reproduced with permission from reference [34]. **11**

Figure 1.8 FDA-approved nucleoside HIV-1 reverse transcriptase inhibitors (NRTIs). (A) Zidovudine (AZT; approved in 1987). (B) Didanosine (1991). (C) Stavudine (1994). (D) Lamivudine (1995). (E) Abacavir (1998) (F) Tenofovir (2001) (G) Emtricitabine (2003). **13**

Figure 1.9 FDA-approved non-nucleoside HIV-1 reverse transcriptase inhibitors (NNRTIs). (A) Nevirapine (approved in 1996). (B) Delavirdine (1997). (C) Efavirenz (1998). (D) Etravirine (2008). (E) Rilviripine (2011). **15**

Figure 1.10 Jablonski diagram. S_0 represents the singlet electronic ground state of the molecule and S_1, and S_2 refer to the first and second singlet electronic excited states, respectively (bold horizontal black lines). Thin horizontal black lines represent the different vibrational energy levels for each electronic state. Following excitation to a higher electronic energy state via absorption of a photon ($h\nu_a$), the electron relaxes to the lowest vibrational energy level of S_1 (internal conversion). Relaxation of the electron from S_1 to S_0 via the emission of a photon is called fluorescence ($h\nu_f$). Alternatively, the excited electron in S_1 can undergo intersystem crossing to T_1. Emission of a photon from T_1 is termed phosphorescence ($h\nu_p$). **19**

Figure 1.11 Change in FRET efficiency with increasing distance between a donor and acceptor fluorophore. Values were calculated using Equation 1.4 for the Cy3 (donor) Cy5 (acceptor) pair. The R_0 value (5.3 nm) was taken from reference [75] and is a typical value for the Cy3/Cy5 pair. **22**

Figure 1.12 Three sequential images of individual $diIC_{12}$ dyes imaged using near-field optical fluorescence microscopy at an acquisition rate of 1/30 s. The field of view for all three images is the same. Figure reproduced with permission from reference [83]. **25**

Figure 1.13 Individual association-hydrolysis-dissociation events captured using total-internal reflection fluorescence microscopy. (A) A series of images showing the appearance and disappearance of Cy3 emission over time. The emission corresponds to the position of a Cy5-labeled S-1 molecule, therefore the appearance of fluorescence represents the association between Cy3-ATP/Cy3-ADP and Cy5-S-1. (B) Cy3 emission intensity over time for the images series shown in (A). Figure reproduced with permission from reference [84]. **27**

Figure 1.14 Real-time observation of single cholesterol oxidase (COx) turnovers. (A) Reaction scheme for the COx-catalyzed oxidation of cholesterol. The reaction is monitored by observing the on/off switching between fluorescent FAD and non-fluorescent $FADH_2$. (B) Fluorescence intensity time-trajectory of a single COx molecule embedded in a polyacrylamide matrix in the presence of 0.2 mM cholesterol and 0.25 mM O_2. The upper and lower levels correspond to fluorescent FAD (on) and non-fluorescent $FADH_2$ (off). Each on/off cycle represents a single COx-catalyzed turnover. (C) Dwell-time distribution of the durations of the fluorescent on-times. The distribution does not follow first-order single-exponential kinetics, suggesting a more complex kinetic scheme. Figure adapted with permission from reference [87]. **28**

Figure 1.15 Probing intra- (A) and intermolecular (B) dynamics using single-molecule FRET. In both cases, the donor (green) is irradiated. When the acceptor dye (red) comes into close proximity of the donor, e.g., during a conformational change (A) or ligand binding (B), energy transfer occurs and emission from the acceptor increases. **29**

Figure 1.16 Cartoon representation of single-molecule protein induced fluorescence enhancement (SM-PIFE). Binding of a protein (orange circle) close to a photoisomerizable dye leads to fluorescence enhancement of the dye due to a decrease in the non-emissive relaxation rate. **30**

Figure 1.17 Single-molecule PIFE assay of T7 DNA polymerase binding to a surface-immobilized Cy3-labeled DNA duplex. (A) Intensity-time trajectories of Cy3-DNA in the absence (upper panel) and presence (lower panel) of 5 nM T7 DNA polymerase. In the lower panel, the fluorescence toggles between two intensities, corresponding to un-bound and protein-bound Cy3-DNA. (B) Intensity time-trajectories of Cy3-DNA in the presence of 10 nM T7 polymerase and increasing concentrations of dCTP. Introduction of dCTP introduces a third intensity level corresponding to a conformational change in the T7 polymerase fingers domain. Figure adapted from reference [96]. **32**

Figure 2.1 Light paths at an interface between two media with refractive indices n_1 and n_2, with $n_2 < n_1$. (A) When the incident angle (θ_1) is smaller than the critical angle (θ_c), light is refracted at the interface into the second medium at an angle θ_2. (B) When $\theta_1 = \theta_c$, light propagates parallel to the interface (i.e., $\theta_2 = 90°$). (C) When $\theta_1 > \theta_c$, total-internal reflection occurs. Figure reproduced with permission from reference [110]. **35**

Figure 2.2 Penetration depth (d) of an evanescent wave as a function of incident angle (θ_1). Values were calculated using equation 2.4 for a glass/water interface ($n_{glass} = 1.52$, $n_{water} = 1.33$) where $\lambda = 532$ nm ($\theta_c = 61°$). **36**

Figure 2.3 Changes in the relative intensity of the evanescent wave with increasing perpendicular distance (z) from the interface for different penetration depths (d). Values were calculated using Equation 2.3. .... **37**

Figure 2.4 Cartoon depicting the generation of an evanescent wave at the glass/water interface. The intensity of the evanescent wave decays exponentially with distance, producing a thin sheet of illumination. Molecules within the excitation volume (yellow stars) undergo fluorescence, while molecules outside of the excitation volume (white circles) are not excited and do not fluoresce. **37**

Figure 2.5 Prism- vs. objective-based TIR illumination. In prism-based TIR, the incident light enters a prism from above, generating an evanescent wave at the slide/water interface. In objective-type TIR, the incident light is focused onto the edge of the objective's back focal plane and leaves the objective. Figure adapted with permission from reference [113]. **39**

Figure 2.6 Schematic of two-colour TIRF microscopy setup. **41**

Figure 2.7 (A) General structure of Cy3. (B) Structure of Cy3 (green) covalently attached to a 3'-terminal deoxyribonucleotide (black). **43**

Figure 2.8 Reaction scheme depicting the functionalization of an aminosilanized glass coverslip with mPEG and biotin-mPEG. **45**

Figure 2.9 Picture of a coverslip masked with a silicone film. The center oval is the unprotected area to which the mPEG-SVA/biotin-PEG/$NaHCO_3$ solution is added. **46**

Figure 2.10 Picture of a coverslip with an assembled imaging chamber. A polycarbonate film is pressed onto a PEG-coverslip and silicone connectors are attached with double-sided tape to create inlet and outlet ports. **47**

Figure 2.11 Enzymatic oxygen scavenging system. (A) Glucose oxidase catalyzes the oxidation of β-D-glucose using molecular oxygen, producing gluconic acid and hydrogen peroxide. (B) Hydrogen peroxide is converted to water and oxygen by catalase. **48**

Figure 3.1 Cartoon depicting the SM-FRET assay developed by Zhuang and coworkers to study the orientational dynamics of HIV-1 RT. Nucleic acid substrates labeled with Cy5 (acceptor) were immobilized onto a quartz coverslip via biotin/streptavidin interactions and imaged using TIRF microscopy. RT labeled with Cy3 (donor) was flowed into the sample chamber and FRET was used to report on RT binding to the nucleic acid substrates. Figure reproduced with permission from [121]. **50**

Figure 3.2 Sequence of Cy3-dP-T employed in our SM-PIFE measurements showing the structure of the Cy3 linkage. **54**

Figure 3.3 A cartoon illustrating the single-molecule protein-induced fluorescence enhancement (SM-PIFE) assay for detecting RT binding to a DNA-DNA substrate. (A) A total internal reflection fluorescence microscopy setup combined with surface-immobilized Cy3-labeled primer-template (Cy3-dP-T) was employed for our SM-PIFE measurements. (B) A representative normalized intensity-time trajectory of Cy3-dP-T acquired while 300 nM RTp66-p51 flowed into the imaging chamber shows the binding of RT to Cy3-dP-T and the induced fluorescence enhancement. **55**

Figure 3.4 Normalized fluorescence intensity-time trajectories of Cy3-dP-T and 300 nM RTp66-p51. The red dashed line indicates the protein arrival and first intensity enhancement observed among the majority of molecules. **57**

Figure 3.5 (A) A representative intensity-time trajectory of a Cy3-dP-T molecule and (B) an ensemble normalized intensity histogram of 300 Cy3-dP-T molecules in the absence of RT fitted with a single Gaussian distribution (red curve). **58**

Figure 3.6 Binding of dimeric RT to Cy3-dP-T. EMSA experiments were performed on Cy3-dP-T in the presence of increasing concentrations of RTp66-p51 (left) and RTp66-p66 (right). **59**

Figure 3.7 Monitoring RT binding/unbinding kinetics in the presence of 50 nM (left) and 2 nM RTp66-p51 (right). Panel A depicts single-molecule intensity-time trajectories of Cy3-dP-T at two RTp66-p51 concentrations. Each trajectory was normalized to its average intensity during the initial period when RT was not present. The red arrows correspond to the first RT binding event and the shaded areas correspond to the off-states where Cy3-dP-T is unbound to RT. Panels B and C show on- and off-time histograms, respectively, constructed based on at least 60 single-molecule intensity-time trajectories. Also shown are the fittings (red) with a single exponential decay function. **59**

Figure 3.8 Normalized ensemble intensity histograms describing the fluorescence enhancement induced by two different domains of RT. Cy3-dP-T with the dye attached to (A) the 3' terminus and (B) the 5' terminus of the primer were employed to determine the extent of fluorescence enhancement induced by binding of the polymerase domain and the RNase H domain of RT to Cy3-dP-T. **62**

Figure 3.9 Binding of heterodimeric and homodimeric RT to Cy3-dP-T. (A-C) Ensemble on-time histograms for Cy3-dP-T in the presence of one of the 3 RT isoforms. Exponential fittings are also shown (red curves). (D-F) Examples of individual normalized intensity-time trajectories acquired for Cy3-dP-T in the presence of one of the 3 RT isoforms. Each trajectory was normalized to its initial average intensity when no protein was present in the imaging chamber. Red arrows indicate the onset of the first intensity enhancement event due to RT binding. (G-I) Ensemble histograms constructed with the individual normalized intensity-time trajectories of Cy3-dP-T acquired in the presence of one of the 3 RT isoforms. The ensemble histograms were fitted with a bimodal Gaussian distribution. The protein concentrations used in these experiments were 25 nM RTp66-p51 (left), 50 nM RTp66 (middle) and 1.4 μM RTp51 incubated with 70 μM efavirenz (EFZ) (right). **63**

Figure 3.10 Intensity-time trajectories (left) and ensemble normalized intensity histograms fitted with Gaussian bimodal distributions (right) for Cy3-dP-T in the presence of 1.4 μM RTp51 and 50 nM RTp66, as a function of incubation time following dilution of the RT isoforms from a concentrated (ca. 200 μM) solution. (A-C) Intensity-time trajectories and ensemble normalized intensity histograms acquired in the presence of RTp51-p51 at (A) 0 min and (B) 20 min (C) 60 min incubation periods; trajectories are recorded 2 min after protein is flowed into chambers. (D-E) Intensity-time trajectories and ensemble normalized intensity histograms acquired in the presence of RTp66 at (D) 0 hr and (E) 144 hr incubation periods. Red arrows indicate the onset of the first intensity enhancement event due to RT binding; trajectories are recorded as soon as protein is flowed. **68**

Figure 3.11 (A) A representative normalized intensity-time trajectory superimposed over a trajectory generated by hidden Markov fitting (red curve). (B) An ensemble on-time histogram with the exponential fitting curve superimposed (red). These data were acquired at 2 nM RTp66-p51. .... **74**

Figure 3.12 (A) A raw (left) and normalized (right) intensity-time trajectories of Cy3-dP-T fluctuating between an on- and an off-state. (B) A normalized intensity histogram, constructed based on at least 100 molecules, fitted with a bimodal Gaussian distribution (red curves). These data were acquired at 50 nM RTp66 after a 144 h incubation period at 4 °C. **75**

Figure 3.13 Binding of 50 nM RTp51 to Cy3-dP-T. (A) Representative intensity-time trajectory showing intensity fluctuations between two states. Red arrow indicates protein arrival in the imaging chamber. (B) Ensemble histogram constructed with the individual normalized intensity-time trajectories of Cy3-dP-T acquired in the presence of RTp51. The ensemble histogram was fitted with a bimodal Gaussian distribution. (C, D) Ensemble on- and off-time histograms for Cy3-dP-T. Exponential fittings are also shown (red curves). These results could not be reproduced. **77**

Figure 3.14 Intensity-time trajectories (left) and ensemble normalized intensity histograms fitted with Gaussian bimodal distributions (right) for Cy3-dP-T in the presence of 50 nM RTp51. (A-B) Intensity-time trajectories and ensemble normalized intensity histograms acquired in the presence of RTp51-p51 at (A) 10 min and (B) 20 min incubation periods. These results could not be reproduced. **77**

Figure 3.15 Ensemble normalized intensity histograms for Cy3-dP-T in the absence of RT (A) and in the presence of 10 μM RTp51 (B). **78**

Figure 3.16 RTp51 does not bind to Cy3-dP-T. EMSA experiments were performed on Cy3-dP-T in the presence of increasing concentrations of RTp51-p51. Two independent batches of RTp51 (1, 2) were tested. (A) Migration of Cy3-dP-T in the absence of RT. (B) Binding of 100 nM RTp66-p51 to Cy3-dP-T (positive control). **79**

Figure 4.1 Schematic of the FRET-based assay described by Parniak et al. RT RNase H-catalyzed cleavage results in the diffusion of fluorescein (green) away from its quencher (dabcyl; purple). Upon cleavage, fluorescein becomes emissive. **83**

Figure 4.2 Cartoon depicting our SM-PIFE assay for RT RNase H activity. The onset of RT binding to a surface-immobilized Cy3-labeled RNA/DNA duplex is observed as an increase in fluorescence intensity of the substrate. RNase H-catalyzed cleavage results in loss of the Cy3 signal as the Cy3-fragment rapidly diffuses out of the excitation volume. **84**

Figure 4.3 RNA/DNA substrates used in our studies. Substrates labeled as 3'Cy3 correspond to labeling at the 3' end of the primer strand. RNA primer sequences are based on the previously reported sequence A and the polypurine tract (PPT).[121] RNA and DNA strands are shown in orange and black, respectively. **85**

Figure 4.4 Predicted positioning of RT on a Cy3-labeled RNA/DNA duplex (3'Cy3-rP-T). RT should adopt an orientation where the RNase H domain is positioned close to the 3' end of the RNA primer.[121] Thus, we hypothesized that RNase H-catalyzed cleavage should occur close to the 3' end of the RNA primer (red arrows). **86**

Figure 4.5 (A) PAGE assay of RT RNase H activity. Experiments were conducted with 3'Cy3-rP-T, 5'Cy3-rP-T, or 3'Cy3-rP-T PPT substrates in the presence of WT or the N348I double-mutant RTp66-p51 (100 nM). The effects of complementary dNTPs, NVP, and foscarnet (fos) on RNase H activity were tested. Black arrow indicates full-length primer. Major cleavage products are highlighted in red. Lanes denoted 1, 2, and 3 are 3'Cy3-rP-T, 5'Cy3-rP-T, and 3'Cy3-rP-T PPT, respectively, in the absence of RT. (B) Full gel scan showing the slow migration of the unique 3'Cy3-rP PPT sequence in the denaturing polyacrylamide gel. Three major bands are observed (blue box). **89**

Figure 4.6 Using SM-PIFE to monitor the different binding orientations of RT. (A) RT is expected to bind Cy3-dP-T in polymerase-competent orientation, with the "F" fingers domain positioned close to the 3' end of the DNA (black) primer. (B) RT is expected to bind Cy3-rP-T in an RNase H-competent orientation, with the "H" RNase H domain positioned close to the 3' end of the RNA (orange) primer. **91**

Figure 4.7 Intensity-time trajectories showing synchronous fluorescence enhancement of (A) Cy3-dP-T and (B) Cy3-rP-T upon flowing 300 nM of RT E478Q into the imaging chamber. **92**

Figure 4.8 Ensemble normalized intensity histograms fitted with single Gaussian functions for (A) Cy3-dP-T and (B) Cy3-rP-T in the presence of 300 nM RT E478Q. **93**

Figure 4.9 Intensity-time trajectories showing multi-step fluorescence enhancement of Cy3-rP-T upon flowing 300 nM of RT E478Q and 100 μM NVP into the imaging chamber. Multi-step fluorescence enhancement was observed in 25% of the 112 molecules analyzed. Red lines indicate the onset of the first binding event........ **96**

Figure 4.10 Ensemble normalized intensity histograms for (A) Cy3-rP-T in the presence of 300 nM RT E478Q and (B) Cy3-rP-T in the presence of 300 nM RT E478Q and 100 μM NVP. **97**

Figure 4.11 Chemical structure of the Cy3 linkage to the 3' end of the DNA primer strand.. **100**

Figure 4.12 Alternative DNA/DNA substrate design for SM-PIFE studies of polymerase dynamics. In order to study the effects of dNTPs on RT polymerase domain dynamics, the substrate should retain an unmodified 3'OH group on the primer strand. Nucleotides labeled in red correspond to the commercially available Cy3 linkage via a thymidine base. To prevent nucleotide incorporation the 3' terminal nucleotide (blue) could be modified to a dideoxy nucleotide. **101**

Figure 4.13 Characterizing the interplay between RT polymerase domain dynamics and RNase H activity at the single-molecule level. The integration of SM-FRET and SM-PIFE approaches may be used to investigate the temporal relationship between polymerase-catalyzed nucleotide incorporation and RNase H-catalyzed RNA hydrolysis. FRET between a donor/acceptor pair such as Cy3 (green) and Cy5 (red) along with changes in PIFE of both Cy3 and Cy5 could be monitored simultaneously. **102**

List of Tables

Table 3.1 Dissociation/association rate constants and equilibrium dissociation constants for the 3 dimeric RTs and their nucleic acid substrates**... 66**

List of Abbreviations

–sssDNA	Minus-strand strong-stop DNA
+sssDNA	Plus-strand strong-stop DNA
3TC	(–)-β-L-2',3'-dideoxy-3'-thiacytidine
AIDS	Acquired immunodeficiency syndrome
ATP	Adenosine triphosphate
AZT	3'-azido-3'-deoxythymidine
bp	Base pair
CA	Viral capsid protein
CCD	Charge-coupled device
COx	Cholesterol oxidase
Cy3	Cyanine 3
Cy3-dP-T	Cy3-labeled DNA primer-DNA template
Cy3-rP-T	Cy3-labeled RNA primer-DNA template
Cy5	Cyanine 5
DLV	Delavirdine
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
ds	Double-stranded
EFZ	Efavirenz
EM-CCD	Electron multiplying charge-coupled device
EMSA	Electrophoretic mobility shift assay
ε	Absorption extinction coefficient
FAD	Flavin adenine dinucleotide (oxidized)
$FADH_2$	Flavin adenine dinucleotide (reduced)
FDA	Food and Drug Administration
Fos	Foscarnet
FRET	Förster resonance energy transfer
HIV-1	Human immunodeficiency virus type-1
IN	Viral integrase
K	Kelvin
k_a, k_d	Association and dissociation rate constant, respectively
kb	Kilobase
K_d	Equilibrium dissociation constant
kDa	Kilodalton
λ	Wavelength
MA	Viral matrix protein
MM	Heavy meromyosin
mRNA	Messenger RNA
n	Refractive index
N-site	Nucleotide-binding site
NA	Numerical aperture
NC	Nucleocapsid protein
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NNRTI-BP	Non-nucleoside reverse transcriptase inhibitor binding pocket
NRTI	Nucleoside reverse transcriptase inhibitor
NSOM	Near-field scanning optical microscopy

NVP	Nevirapine
P-site	Priming site
P-T	Primer-template
PAGE	Polyacrylamide gel electrophoresis
PBS	Primer binding site
PEG	Poly(ethylene) glycol
Φ	Emission quantum yield
PIFE	Protein-induced fluorescence enhancement
PPT	Polypurine tract
PR	Viral protease
R	Repeat sequence
RNA	Ribonucleic acid
RNase H	Ribonuclease H
RT	Reverse transcriptase
S-1	Single-headed myosin subfragment
SM-FRET	Single-molecule Förster resonance energy transfer
SM-PIFE	Single-molecule protein-induced fluorescence enhancement
ss	Single-stranded
TAM	Thymidine analogue mutation
τ_{dwell}	Duration of dwell time
τ_{on}, τ_{off}	Duration of on and off time, respectively
θ, θ_c	Angle, critical angle (degrees)
TIR(F)	Total internal reflection (fluorescence)
tRNA	Transfer RNA
vDNA	Viral deoxyribonucleic acid
WT	Wild type

1. Introduction

1.1 Human Immunodeficiency Virus-1 (HIV-1)

Retroviruses comprise a diverse group of viruses that utilize RNA as their genetic material. Common to all retroviruses is a unique replicative strategy called reverse transcription, where the single-stranded viral RNA genome is converted into linear double-stranded DNA. The human immunodeficiency virus (HIV) is a widely known retrovirus and is the causative agent of acquired immunodeficiency syndrome (AIDS). AIDS, first diagnosed in 1981, was considered a death sentence up until the early 1990s. However, with the emergence of antiretroviral therapy in the late 1980s, the life expectancy of HIV-infected individuals has increased dramatically.[1] Despite the improved prognosis, HIV/AIDS remains a global epidemic due to the lack of an effective vaccine. HIV currently affects over 34 million people worldwide.[2]

HIV is classified into two types, HIV-1 and HIV-2. HIV-1 is pandemic and the major cause of AIDS globally due to its high transmissibility.[3] In contrast, the HIV-2 strain is less pathogenic compared to HIV-1 and geographically confined to parts of west Africa.[3] The following sections will discuss HIV-1 exclusively.

1.1.1 HIV-1 Life Cycle

HIV-1 possesses two copies of a ~9 kb positive-sense single-stranded RNA genome. HIV-1 possesses an outer host-derived lipid membrane, with both host and viral membrane proteins embedded within. Lining the inner portion of the lipid membrane is a shell composed of the HIV-1 matrix (MA) protein. Within the MA shell is a conical-shaped core particle composed of HIV-1 capsid protein (CA) called the capsid core, which encapsulates both copies of the RNA genome. The HIV-1 genome encodes three structural proteins, two envelope proteins, three enzymes, and six accessory proteins.[4, 5]

Common to all viruses, HIV-1 relies entirely on the host cellular machinery to reproduce. The successful reproduction of HIV-1 requires four essential steps: 1) binding and entry into a host cell; 2) conversion of the viral single-stranded RNA genome into double-stranded DNA (vDNA); 3) integration of the vDNA into a host cell chromosome; and 4) assembly, budding, and maturation of the new HIV-1 virion. A general schematic of the HIV-1 life cycle is depicted in Figure 1.1.

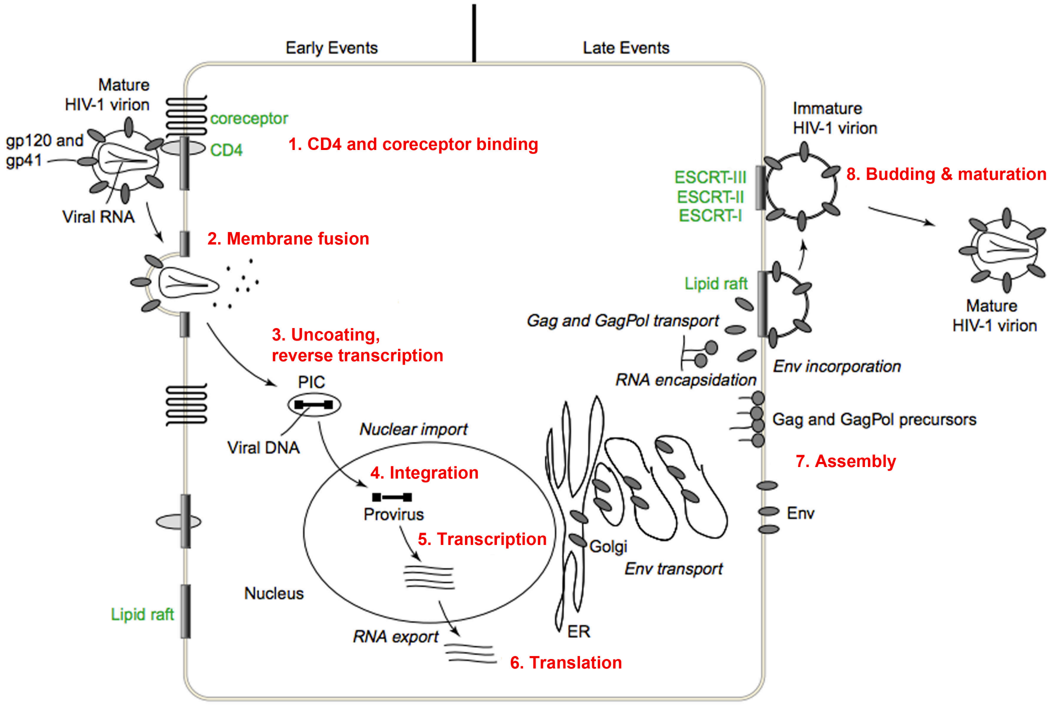

Figure 1.1 Schematic of HIV-1 life cycle. The cycle begins with host CD4 receptor and coreceptor binding to the viral gp120 (1). Following binding, membrane fusion occurs and the contents of the virus are released into the cytoplasm of the cell (2). Uncoating of the viral capsid releases two copies of the viral RNA genome. Reverse transcriptase converts the RNA genome into double-stranded DNA (3). The newly formed DNA is transported into the nucleus and integrated into a host chromosome by the viral integrase (4). The resulting proviral DNA is transcribed (5) and translated (6) by the host cellular machinery and the viral components are assembled at the plasma membrane (7). Finally, the virion buds from the host cell and undergoes maturation, resulting in a new infectious virus (8). Figure adapted with permission from reference [6].

The HIV-1 life cycle begins when a mature viral particle interacts with a host cell membrane. The outer host-derived lipid membrane of HIV-1 contains the viral envelope protein, Env, a complex composed of gp120 and gp41 glycoproteins. The initial interaction between the virus and host cell occurs via the binding of gp120 to CD4 receptors on host T-lymphocytes.[7] Upon establishing the initial virus/host-cell interaction, conformational changes in the CD4 receptor enable the recruitment of host-cell co-receptors, CXCR4 and/or CCR5.[8] Following co-receptor binding, a hydrophobic portion of gp41 inserts itself into the host cell membrane, triggering membrane fusion.

Upon membrane fusion, the contents of the HIV-1 virion are released into the cytoplasm of the host cell. Disassembly of the capsid core releases both copies of the RNA genome and ~50-100 copies of the enzyme reverse transcriptase (RT).[9] RT catalyzes reverse transcription of single-stranded viral RNA into double-stranded DNA in the cytoplasm of the host cell (see section 1.1.2).

Following reverse transcription, vDNA associates with various host cell proteins and the viral integrase (IN) to form a pre-integration complex.[5, 10] The pre-integration complex is transported into the host cell nucleus where IN catalyzes the covalent incorporation of vDNA into a host chromosome. Integrated vDNA (proviral DNA) serves as a template for the host cell's transcription machinery, which results in the production of positive-sense viral messenger RNA (mRNA) transcripts. Viral mRNA is then effectively translated into the various viral proteins by the host cell.[10]

Completion of the HIV-1 lifecycle requires assembly, budding, and maturation of the virus. The viral structural proteins MA, CA, nucleocapsid (NC), and p6 enzymes and viral enzymes RT, protease (PR), and IN are translated as polyprotein precursors called called Gag and Pol, respectively.[9] Following translation, Gag polyproteins accumulate at the host cell membrane and form spherical particles through Gag/Gag interactions.[11] Gag/Gag interactions drive viral particle assembly by recruiting the viral proteins and RNA to the membrane and promoting membrane curvature. With the additional help of host cell endosomal sorting proteins, new virions bud from the host cell and immediately undergo a process called maturation.[11] During maturation, the PR domain of the Pol polyprotein catalyzes the proteolytic cleavage of Gag and Pol into fully processed MA, CA, NC, p6, RT, IN, and PR, resulting in a fully infectious virion.[11]

1.1.2 Reverse Transcription

Reverse transcription is the process whereby RNA is converted into DNA, and is a defining feature of retroviruses such as HIV-1. In 1964, Howard Temin was the first to propose that the genome of RNA viruses could be converted into DNA.[12] Temin's proposal was controversial as it violated the central dogma of molecular biology, which states that genetic information can only flow from DNA to RNA to protein.[9] In 1970, Temin and David Baltimore independently confirmed reverse transcription when they discovered RNA-dependent DNA polymerase activity in purified virions of Rous sarcoma and Rauscher mouse leukemia viruses.[13, 14]

Reverse transcription occurs in the cytoplasm of the infected cell following viral fusion. Although viral and host cell factors facilitate reverse transcription *in vivo*, the catalytic activities of the dual-function enzyme reverse transcriptase (RT) are sufficient to convert the single-stranded viral RNA genome into double-stranded DNA. RT is both a DNA polymerase, which can use RNA or DNA as a template, as well as an RNase H endonuclease, which degrades RNA only when it is part of a RNA/DNA hybrid.

A unique feature of retroviruses such as HIV-1 is that each virion contains two copies of the single-stranded RNA genome. A cartoon depicting the structural organization of a single RNA template is shown in Figure 1.2.

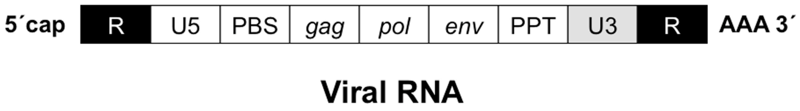

Figure 1.2 Schematic of the HIV-1 viral genome. Viral RNA possesses a 5' cap and a 3' poly-A tail, similar to cellular mRNA. Sequences unique to the 5' and 3' (U5 and U3, respectively) sit adjacent to identical repeat (R) sequences. The primer-binding site (PBS) is complementary in sequence to $tRNA_3^{lys}$, the primer for minus-strand DNA synthesis. The polypurine tract (PPT) sequence serves as a primer for plus-stranded DNA synthesis.

Viral genomic RNA is produced by the host cell RNA synthesis machinery and thus possesses features commonly found in cellular mRNA such as a 5' 7-methylguanosine cap and a 3' polyadenylated tail. Repeat sequences (R) are located at both ends of the genome next to sequences that are unique to the 5' end (U5) and 3' end (U3). The R, U5, and U3 sequences play important roles in the strand transfer steps of reverse transcription (discussed below). Minus-strand DNA synthesis begins from a cellular $\text{tRNA}_3^{\text{lys}}$ primer, which is complementary in sequence to the primer-binding site (PBS). A purine-rich RNA sequence called the polypurine tract (PPT), which is found close to the 3' end of the RNA template, serves as a primer for plus-strand DNA synthesis. Viral proteins are translated as polyprotein precursors called Gag, Pol, or Env from the *gag*, *pol*, and *env* genes, respectively (Figure 1.2). Following expression, the Gag, Pol, and Env polyproteins are cleaved by the viral protease into the various viral proteins.

The steps of reverse transcription are summarized schematically in Figure 1.3. RT uses its RNA-dependent DNA polymerase activity to initiate minus-strand DNA synthesis from a $\text{tRNA}_3^{\text{lys}}$ primer annealed to the PBS near the 5' end of the RNA template (Figure 1.3A). Minus-strand DNA synthesis continues until the 5' end of the template is reached, while the RNase H activity of RT simultaneously degrades the RNA template (Figure 1.3B). The resulting strand of newly synthesized DNA is called minus-strand strong-stop DNA (–sssDNA) and is approximately 100-150 bases long.[9]

Viral genomic RNA contains identical R sequences at each terminus (Figure 1.2), therefore –sssDNA can be transferred to the 3' end of viral RNA in a process called minus-strand transfer (Figure 1.3C). Since HIV-1 virions package two copies of viral RNA, minus-strand transfer can occur either intramolecularly (strand transfer occurs to the 3' end of the same RNA template) or intermolecularly (strand transfer occurs to the 3' end of the other RNA template copy). Following the first strand transfer event, minus-strand DNA synthesis continues until the end of the RNA template is reached (Figure 1.3D). As before, minus-strand synthesis is accompanied by RNase H-catalyzed degradation of the RNA template, however, a purine-rich RNA sequence called the polypurine tract (PPT) located near the 3' end of the RNA template is resistant to degradation. The PPT sequence serves as the primer for plus-strand DNA synthesis.

Plus-strand DNA synthesis begins at the PPT primer and continues until a portion (~18 nucleotides) of the tRNA$_3^{lys}$ primer has been transcribed (Figure 1.3E).[15] The resulting segment of newly synthesized plus-strand DNA is referred to as plus-strand strong-stop DNA (+sssDNA). The ensuing removal of the tRNA$_3^{lys}$ primer by RNase H sets the stage for the second strand transfer step, plus-strand transfer. The 3' end of the +sssDNA contains 18 nucleotides that were copied from the tRNA$_3^{lys}$ primer and are thus of the same sequence as the PBS sequence; therefore plus-strand transfer occurs between the PBS sequence at the 3' ends of +sssDNA and minus-strand DNA (Figure 1.3F). After plus-strand transfer, both minus- and plus-strand DNA synthesis continue bi-directionally, resulting in a complete double-stranded viral DNA (Figure 1.3G).

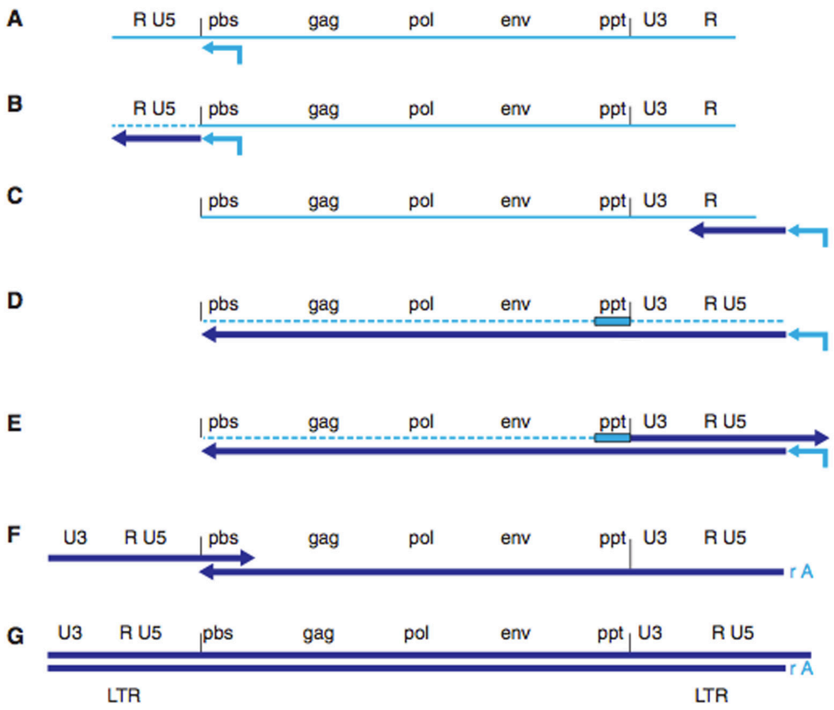

Figure 1.3 Steps of reverse transcription. Viral RNA is depicted in light blue, while newly synthesized DNA is shown in dark blue. The dashed lines represent degraded RNA. (A) A host cell-derived tRNA$_3^{lys}$ binds to the primer-binding site (pbs) of the RNA genome. (B) RT catalyzes minus-strand DNA synthesis and simultaneous RNA degradation until the 5' end is reached, generating minus-strand strong-stop DNA (–sssDNA) (C) The first strand-transfer step can be inter- or intramolecular and occurs between the complementary R sequences of the –sssDNA and viral RNA. (D) Minus-strand DNA synthesis continues until the 5' end, along with RNase H-catalyzed RNA degradation. Short purine-rich sequences (ppt) are resistant to cleavage and act as

primers for plus-strand DNA synthesis. (E) Plus-strand synthesis begins from a PPT primer and continues until a portion of the $tRNA_3^{lys}$ primer is transcribed before being degraded by RT RNase H. (F) Minus- and plus-strand DNA synthesis continue until a full double-stranded DNA copy of the viral RNA is obtained. Figure reproduced with permission from reference [15].

1.1.3 Reverse Transcriptase (RT)

Reverse transcription, the conversion of single-stranded viral RNA into double-stranded DNA, is an essential step in the HIV-1 life cycle. Although other viral proteins such as NC and IN may facilitate reverse transcription *in vivo*, HIV-1 RT contains all the necessary activities to carry out the conversion by itself.[16] RT is a multifunctional heterodimeric enzyme possessing both DNA polymerase and RNase H activities.[17, 18]

1.1.3.1 Structure

HIV-1 RT is a heterodimer comprised of a 66 kDa, 560 amino acid subunit (p66) and 51 kDa, 440 amino acid subunit (p51).[19, 20] RT is encoded by a single open reading frame of the *gag-pol* gene, however the p51 subunit of the p66-p51 heterodimer is derived via PR mediated cleavage of p66 in between residues F440 and Y441.[19] The overall three dimensional structure of RT is depicted in Figure 1.4.[21] The p66 subunit consists of two domains, the polymerase and RNase H domains, which house the DNA polymerase and RNase H active sites, respectively.[22, 23] The structure of the polymerase domain has been likened to that of a right hand and is further divided into fingers (amino acids 1-85, 118-155), palm (86-117, 156-236), and thumb (237-318) subdomains. A connection domain (319-426) bridges the N-terminal polymerase domain and the C-terminal RNase H domain.[23] The p51 subunit contains the same four subdomains as the p66, however p51 adopts a different tertiary structure that is more compact and rigid relative to its p66 counterpart.[23-25] When associated with p66, p51 is thought to provide structural support to p66 and additional contacts to bound nucleic acid substrates.[23, 26]

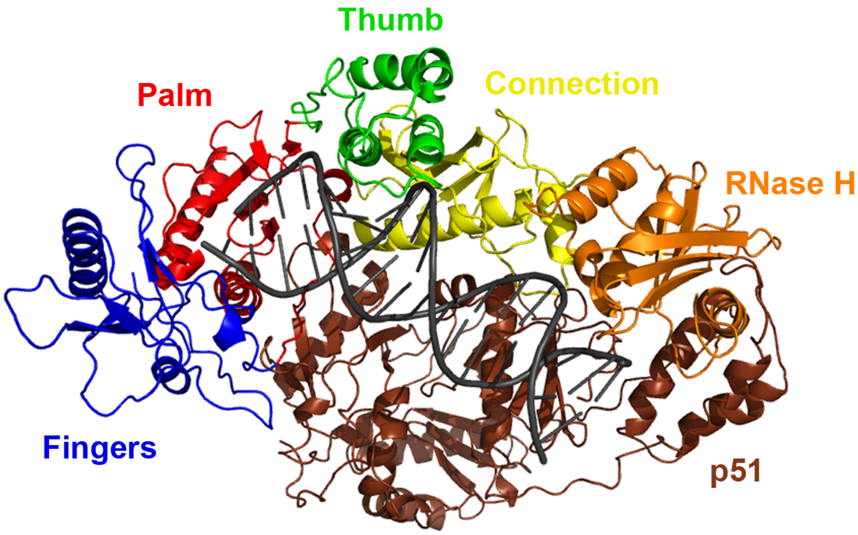

Figure 1.4 Structure of the HIV-1 RT p66-p51 heterodimer bound to a DNA/DNA duplex (grey). The p66 domain is composed of fingers (blue), palm (red), thumb (green), connection (yellow), and RNase H (orange) domains. The p51 subunit is depicted in brown. The structure shown here was generated using PyMOL molecular visualization software (www.pymol.org) and RCSB protein databank structure access code 2HMI from reference [21].

The fingers, palm, and thumb domains of the p66 subunit form the RT nucleic acid binding cleft, which can accommodate ~18-19 base pairs of double-stranded nucleic acid.[22] In unliganded RT crystal structures, the p66 thumb domain is folded inwards toward the nucleic acid binding cleft such that it almost touches the p66 fingers domain.[27] Before binding a nucleic acid substrate, the thumb domain must expose the binding cleft by undergoing a large-scale conformational change (Figure 1.5).[21] In the presence of complementary dNTPs, the fingers domain will also close down and establish additional contacts with the nucleic acid substrate. RT/nucleic acid interactions are mainly electrostatic in nature, with interactions occurring between the sugar-phosphate backbone of the nucleic acid substrate and highly conserved amino acid motifs in the p66 domain.[24]

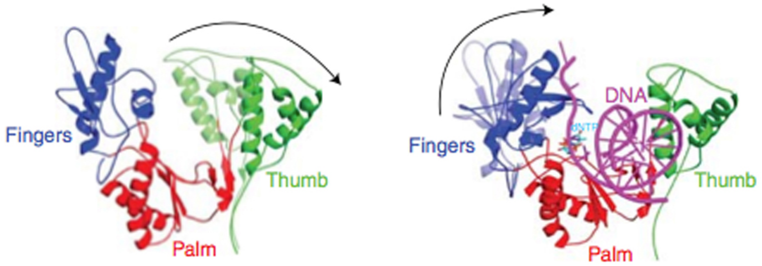

Figure 1.5 Binding of nucleic acid by RT requires a conformational change in the thumb domain in order to expose the nucleic acid binding cleft. The fingers domain also possesses conformational flexibility and closes down tightly over an RT/nucleic acid/dNTP ternary complex. Figure reproduced with permission from reference [15].

The most important motif for substrate-binding is called the primer grip, a β-hairpin structure comprised of p66 amino acids 227-235 of the palm domain. The primer grip is involved in in proper positioning of the primer's 3'-OH group during DNA polymerization.[22, 28] A similar motif located in the RNase H domain called the RNase H primer grip (p66 amino acids 473-476, 501, 505, 539) is required for the proper positioning of an RNA/DNA duplex during RNase H-catalyzed cleavage.[29]

1.1.3.2 Function

DNA Polymerase Activity

RT possesses RNA- and DNA-dependent DNA polymerase activities, which are required for minus- and plus-strand DNA synthesis, respectively. RT catalyzes the addition of free dNTPs to the free 3'-OH group of a recessed primer. Common to many known nucleic acid polymerases, RT's polymerase active site contains a catalytic triad made of three aspartic acid residues (D110, D185, and D186), which coordinate two divalent metal ions.[23]

The RT-catalyzed DNA polymerization reaction begins with binding of a primer/template nucleic acid duplex (Figure 1.6). Next, an incoming complementary dNTP binds to the nucleotide binding site (N-site) to form the ternary complex.[15, 30] The fingers domain undergoes a conformational change and closes down over the N-site, which aligns the dNTP α-phosphate in an optimal geometry for nucleophilic attack by the primer 3'-OH.[31] Incorporation of the dNTP requires two coordinated Mg^{2+} ions *in vivo* and is accompanied by the loss of pyrophosphate. In order for polymerization to continue, the enzyme needs to translocate such that the newly incorporated nucleotide moves from the N-site to the priming site (P-site), allowing the next complementary dNTP to bind.

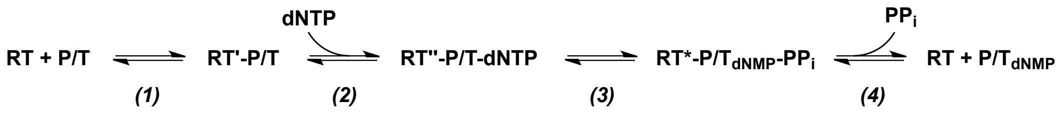

Figure 1.6 Steps involved in RT-catalyzed DNA polymerization. Polymerization begins with RT binding to primer/template (P/T) complex (1), followed by formation of a ternary complex with complementary dNTP (2). Nucleotide incorporation (3) results in the loss of pyrophosphate (PP_i) (4). RT must translocate one base pair in order to accept the next incoming dNTP. Figure adapted from reference [16].

RNase H Activity

The RNase H domain of RT is an endonuclease that hydrolyzes the phosphodiester backbone of the RNA strand in an RNA/DNA duplex. RT RNase H is structurally homologous to many other RNase H enzymes, such as RNase H from humans, *E. coli*, *B. halodurans*, and murine leukemia virus.[32] Indeed, a significant amount of mechanistic insight into HIV-1 RT RNase H function has come from studies of homologous enzymes.[33] The active site of RT RNase H contains the highly conserved DEDD catalytic motif, which facilitates a two-metal ion cleavage mechanism.[32] RNase H-mediated cleavage requires two divalent metal ions (Mg^{2+} *in vivo*) and occurs via an S_n2 mechanism, with H_2O acting as the nucleophile and attacking the phosphate group of the RNA strand (Figure 1.7).[34]

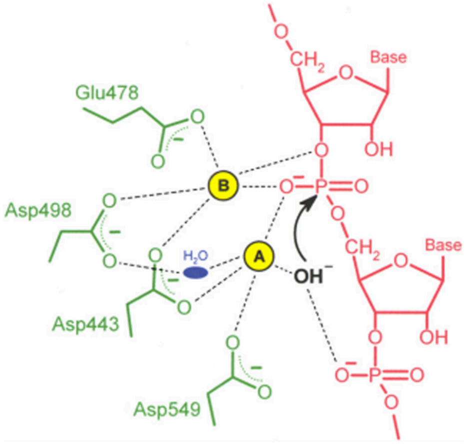

Figure 1.7 Mechanism of RT RNase H-catalyzed RNA cleavage. The reaction occurs via a two divalent metal ion-dependent S_n2 mechanism with water as the nucleophile. The two divalent metals, A and B, are coordinated by a highly conserved DEDD motif in the active site. Figure reproduced with permission from reference [34].

Residues D443 and D549 of the DEDD motif coordinate one Mg^{2+} ion, which coordinates and reduces the pK_a of one H_2O molecule. The other Mg^{2+} ion is coordinated by D443, E478, and D498 and stabilizes the build up of negative charge on the RNA phosphate in the transition state.

During reverse transcription, RT RNase H cleavage can occur in either a polymerase-dependent or polymerase-independent mode. Polymerase-dependent mode occurs during minus-strand DNA synthesis when the polymerase active site is engaged with the 3' terminus of the DNA primer.[34] As minus-strand synthesis occurs, RT RNase H produces cuts ~18 base pairs upstream of the polymerase active site.[32] For each nucleotide incorporation event, polymerase-dependent RNase H cleavage can produce two different cuts depending on the translocational status of RT.[32, 35] In the pre-translocational state, the 3' primer terminus is still in the N-site immediately after nucleotide incorporation. In the post-translocational state, RT has moved one base pair along the template and the 3' primer terminus occupies the P-site. RT can toggle between both pre- and post-translocational states, thus two populations of polymerase-dependent cleavage products are produced for each incorporated nucleotide.[35] Since the rate of polymerization is significantly faster than the rate of RNase H cleavage,[34] polymerase-dependent cleavage does not result in complete RNA template degradation.

Polymerase-independent RNase H cleavage is any cleavage that occurs when the polymerase active site is not engaged with the 3' terminus of the primer. Polymerase-independent RNase H cleavage can be classified as RNA 5'-end directed or internal. RNA 5'-end directed cleavage occurs when the polymerase domain interacts with the 5' end of an RNA strand hybridized to a longer DNA strand. Internal cleavages are non-directed, synthesis-independent cleavage events that occur on long RNA/DNA segments.

1.1.4 RT as a Drug Target: Mechanisms of Inhibition and Drug Resistance

There is no vaccine or cure for HIV-1; however, antiretroviral drug therapy has been extremely successful in improving the prognosis for infected patients. As of August 2012, 23 anti-HIV drugs have been approved by the U.S. Food and Drug Administration (FDA).[36] Of the 23 FDA-approved anti-HIV drugs, 12 target RT, emphasizing the importance of RT as a therapeutic target. Drugs targeting RT can be broadly classified into two groups: 1) nucleoside analogue RT inhibitors (NRTIs) and 2) non-nucleoside RT inhibitors (NNRTIs). Both NRTIs and NNRTIs interfere with viral replication by inhibiting RT's DNA polymerase activity. Inhibitors targeting RT RNase H function have been the focus of intense research for many years, however few inhibitors have been discovered and none have made it through clinical trials.[37]

Anti-HIV drug therapy has numerous limitations including long-term toxicity, intolerability, and the emergence of drug-resistant viral strains.[38] Drug resistance is a particularly significant problem due to HIV-1's high mutation rate of ~10^{-5}-10^{-4} mutations per nucleotide,[38] equivalent to ~1-9 mutations per replication cycle for the ~9 kb viral genome. As a result, the selective pressure introduced by HIV therapy can quickly lead to drug-resistant variants becoming dominant species. The ultimate consequence of HIV's rapid evolution is that drug resistance will develop in all patients over time.[16] In order to delay the emergence of drug-resistant variants, patients undergo highly active antiretroviral therapy

(HAART) regimens consisting of one or two NRTIs plus one NNRTI and/or one protease inhibitor.

1.1.4.1 Nucleoside RT Inhibitors (NRTIs)

NRTIs are analogues of the naturally occurring dNTP substrates lacking the 3'–OH group on the deoxyribose ring. In 1987, the NRTI zidovudine (3'-azido-3'-deoxythymidine; AZT) became the first drug approved by the FDA to treat HIV infection (Figure 1.8).[37] Since then, six more NRTIs have been approved by the FDA and have found widespread use in the clinic, including lamivudine (Figure 1.8D), which was invented at McGill University by Dr. Bernard Belleau. NRTIs are technically pro-drugs since they must be phosphorylated by host cell kinases to 5'-triphosphates before RT can utilize them as substrates.[39] Due to their high degree of structural similarity to natural dNTPs, it is not surprising that they may also be used as substrates by cellular DNA polymerases, leading to drug toxicity effects.

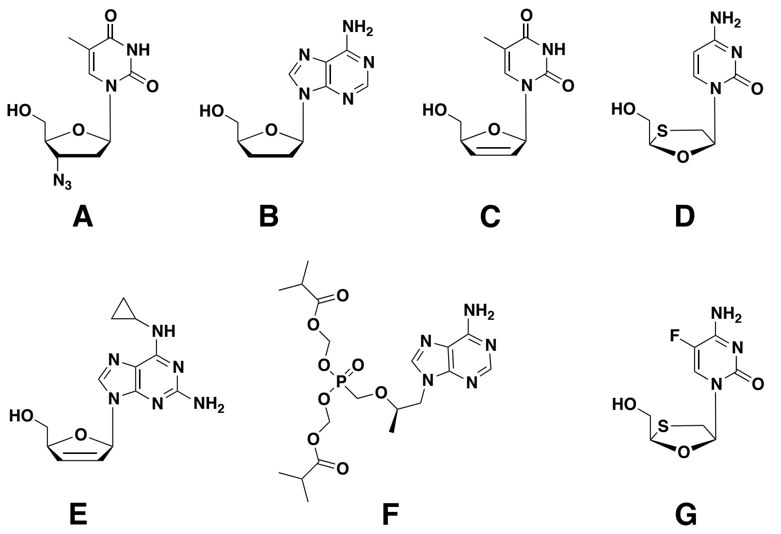

Figure 1.8 FDA-approved nucleoside HIV-1 reverse transcriptase inhibitors (NRTIs). (A) Zidovudine (AZT; approved in 1987). (B) Didanosine (1991). (C) Stavudine (1994). (D) Lamivudine (1995). (E) Abacavir (1998) (F) Tenofovir (2001) (G) Emtricitabine (2003).

Following *in vivo* phosphorylation, NRTIs inhibit RT via two mechanisms, competitive inhibition and chain-termination, the latter being the primary mechanism of action.[37] Competitive inhibition is based on the competition between NRTIs and the naturally occurring dNTPs for the binding site in the polymerase active site. The lack of 3'–OH on the NRTI's deoxyribose ring results in chain termination. Following RT-catalyzed incorporation of an NRTI into a growing DNA strand, DNA synthesis is terminated because a new phosphodiester bond cannot be formed to the next incoming dNTP.

1.1.4.2 NRTIs and Drug Resistance

RT lacks a proofreading mechanism therefore theoretically, it should only require two NRTI molecules to terminate the replication of one virus (HIV-1 possesses two copies of its RNA genome).[37] In practice, NRTIs aren't nearly as potent due to varying degrees of phosphorylation by cellular kinases and drug resistance conferring mutations in RT. As described above, HIV-1 RT is a low fidelity DNA polymerase, thus the viral genome is highly prone to mutations. Before the initiation of antiviral therapy, a population of viruses will be dominated by the wild-type phenotype. However, the population is not homogenous and low levels of drug resistant mutant strains will normally be present. Antiviral therapy (e.g., with NRTIs) introduces selective pressure leading to drug-resistant phenotypes becoming the dominant species.

Mutations in RT are associated with resistance to NRTIs.[37] In general, resistance to NRTIs occurs via two mechanisms: nucleotide discrimination and excision. The nucleotide discrimination mechanism is based on mutations that improve RT's ability to discriminate between natural dNTPs and NRTIs. The M184V/I mutation confers high-level resistance to the NRTI lamivudine ((–)-β-L-2',3'-dideoxy-3'-thiacytidine; 3TC) (Figure 1.8D). Substitution of methionine for the β-branched amino acids valine or isoleucine introduces steric bulk into the RT polymerase active site, which is thought to interfere with the binding of the 3TC oxathiolane ring, preventing catalysis.[40]

The second mechanism of NRTI resistance is the excision of an incorporated NRTI by RT.[41] Once the NRTI is excised from a new DNA strand, RT can continue DNA polymerization and viral replication can proceed. Thymidine analogue mutations (TAMs) comprise a group of common mutations associated with increased excision rates of thymidine-based NRTIs, such as AZT. TAMs facilitate the phosphorolytic excision of NRTIs using ATP or pyrophosphate as the phosphate donor. TAMs include M41L, D67N, K70R, L210W, T215F/Y, and K219E/Q, all of which are localized near the DNA polymerase active site.[38]

1.1.4.3 Non-Nucleoside RT Inhibitors (NNRTIs)

Non-nucleoside RT inhibitors (NNRTIs) comprise a structurally diverse class of molecules that are chemically distinct from dNTPs.[42] NNRTIs are allosteric inhibitors that bind to a single hydrophobic region in the palm domain of p66 called the NNRTI binding pocket (NNRTI-BP) located ~10 Å from the polymerase active site.[43] As of 2012, five NNRTIs have been approved by the FDA for clinical use, including nevirapine (NVP), delavirdine (DLV), efavirenz (EFV), etravirine (ETR), and rilpivirine (RPV) (Figure 1.9).[36]

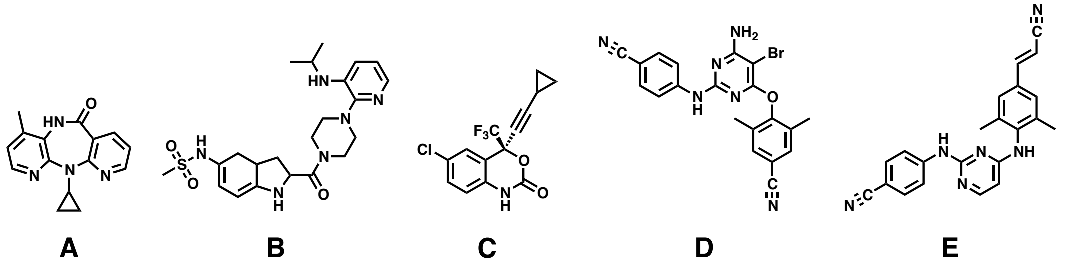

Figure 1.9 FDA-approved non-nucleoside HIV-1 reverse transcriptase inhibitors (NNRTIs). (A) Nevirapine (approved in 1996). (B) Delavirdine (1997). (C) Efavirenz (1998). (D) Etravirine (2008). (E) Rilviripine (2011).

NNRTIs function primarily by inhibiting RT-catalyzed DNA synthesis during reverse transcription. Kinetic studies have demonstrated that NNRTIs interfere with the chemical step of nucleotide incorporation during DNA synthesis, but do not inhibit the binding of dNTPs or primer/template complexes.[44-46] While the exact mechanism of inhibition is still under debate, X-ray crystallography studies have shown that the binding of NNRTIs to the RT NNRTI-BP results in conformational changes that perturb the geometry of the conserved catalytic YMDD motif in the polymerase active site.[47] Additionally, numerous crystal structures of RT complexed with NNRTIs including 1-(2,2-hydroxyethoxymethyl-6-(phenylthio)thymine (HEPT),[27, 47, 48] tetrahydro-imidazo[4,5,1-jk][1,4]-benzodiazapine-2(1H)-one compounds (TIBO),[49, 50] NVP,[23, 48, 51] and EFZ[52] display hyperextension of the p66 fingers and thumb domains. An abundance of structural data supports the "molecular arthritis" hypothesis, which states that NNRTI-induced hyperextension of the fingers and thumb domains restricts conformational flexibility, thus impairing RT's ability to form a productive ternary complex during DNA synthesis.[23]

In addition to inhibiting DNA polymerization during reverse transcription, some evidence suggests that certain NNRTIs such as EFZ can interfere with Gag-Pol polyprotein processing, which occurs during the late stages of the viral replication cycle.[53, 54]

1.1.4.4 NNRTIs and Drug Resistance

Approximately 60 different amino acid mutations in RT associated with increased resistance to NNRTIs have been identified.[55] The majority of these mutations are in or in close proximity to the NNRTI-BP and reduce the RT/drug binding affinity.[37] NNRTIs form numerous contacts within the RT NNRTI-BP, including stacking, electrostatic, hydrogen bond, and van der Waals interactions.[38] Amino acid mutations that disrupt these interactions result in high-level resistance to NNRTIs. The NNRTI-BP is very plastic and will adapt to bind different NNRTIs.[38] As a result, some mutations confer resistance to one type of NNRTI, but not others.

NVP (Figure 1.9A), the first clinically-approved NNRTI, interacts with RT via stacking interactions between NVP's pyridine rings and NNRTI-BP aromatic side-chains Y181 and Y188.[23] Accordingly, mutations such as Y181C/I and Y188C/I/L/H reduce the binding affinity of NVP and confer high-level drug resistance. Y181/Y188 mutations also confer resistance to DLV, but not EFZ. Mutations that increase the steric bulk in the NNRTI-BP, such as G190/A/E/Q/T also confer resistance to NNRTIs such as NVP and EFZ.[37]

1.1.4.5 Connection Domain Mutations and Drug Resistance

Since the most common drug resistance-conferring mutations are located in the N-terminal polymerase domain, genotypic analyses of HIV-1 isolates in patients are often conducted only for the first 250 amino acids of RT.[56] However, recent studies have shown that mutations in the connection (residues 319-426) and RNase H (427-560) domains are also associated with resistance to both NRTIs and NNRTIs.[57-59]

The amino acid mutation N348I has received a considerable amount of attention as it emerges early on in antiretroviral therapy.[58] N348I is thought to confer resistance to both AZT (NRTI) and NVP (NNRTI) by decreasing the rate of RT RNase H-catalyzed RNA template degradation.[60] In the case of AZT resistance, it has been proposed that reduced RNA template degradation observed with the N348I RT mutant stabilizes the RNA/DNA duplex, thus giving RT more time to excise AZT from a chain-terminated primer.[61, 62] Götte and co-workers have also provided evidence for an additional RNase H-independent mechanism leading to increased AZT resistance, where N348I mutants lacking RNase H activity show increased processivity during DNA synthesis, thereby increasing the rate of AZT-excision.[63]

The mechanism for increased resistance to the NNRTI NVP remains elusive, however numerous hypotheses have been proposed, including decreased inhibitor binding to the N348I mutant,[64] decreased RNase H activity,[65] and improper positioning of RT on the nucleic acid substrate during plus-strand DNA synthesis initiation from PPT primers.[66]

1.2 Fluorescence Techniques: Applications In Single-Molecule Spectroscopy

Fluorescence has emerged as a central methodology in all areas of physics, chemistry, and biology. For example, fluorescence is widely used in the study of protein structure and function due to the intrinsic fluorescence of tryptophan amino acid residues.[67] Additionally, other properties of fluorescence such as anisotropy and Förster resonance energy transfer (FRET) have been exploited to develop sensitive, high-throughput screening methods for drug development purposes.[68, 69]

Over the course of the past 30 years, fluorescence has been used to achieve the ultimate level of sensitivity: single-molecule detection. Single-molecule fluorescence detection has enabled the study of complex biochemical mechanisms, such as the function of DNA polymerases,[70] protein translation by the ribosome,[71] and protein folding.[72]

In general, single-molecule techniques can be classified as either optical-based (e.g., fluorescence) or force-based (e.g., optical/magnetic tweezers, atomic force microscopy), however the following sections will focus only on the former. Specifically, I will overview the phenomenon of fluorescence, describe the advantages of single-molecule detection, and discuss the important discoveries and contributions that helped shape the field of single-molecule fluorescence spectroscopy, with a specific focus on the study of complex biochemical processes.

1.2.1 Fluorescence & The Jablonski Diagram

Fluorescence is the emission of light from the singlet excited state of a substance.[73] The processes that can occur leading up to fluorescence can be conveniently depicted in a Jablonski diagram (named after Professor Alexander Jablonski) (Figure 1.10).

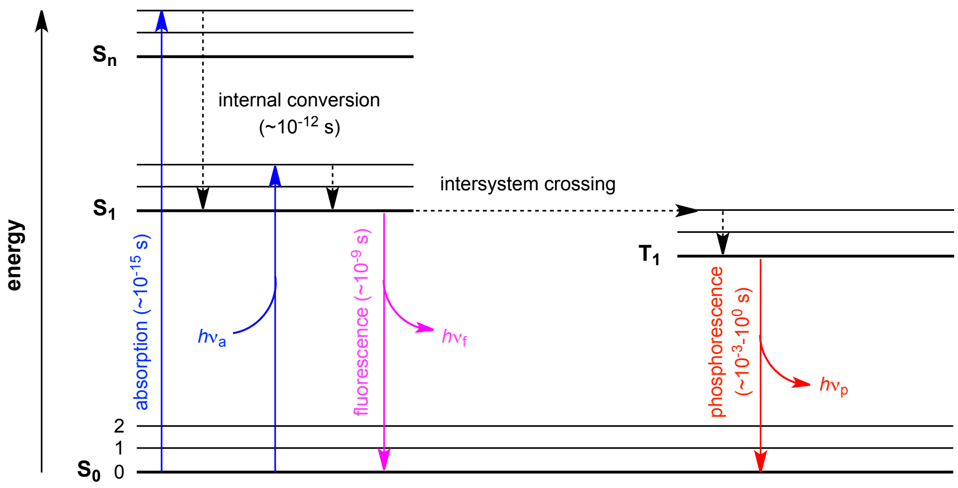

Figure 1.10 Jablonski diagram. S_0 represents the singlet electronic ground state of the molecule and S_1, and S_2 refer to the first and second singlet electronic excited states, respectively (bold horizontal black lines). Thin horizontal black lines represent the different vibrational energy levels for each electronic state. Following excitation to a higher electronic energy state via absorption of a photon ($h\nu_a$), the electron relaxes to the lowest vibrational energy level of S_1 (internal conversion). Relaxation of the electron from S_1 to S_0 via the emission of a photon is called fluorescence ($h\nu_f$). Alternatively, the excited electron in S_1 can undergo intersystem crossing to T_1. Emission of a photon from T_1 is termed phosphorescence ($h\nu_p$).

In a Jablonski diagram, S_0 represents the singlet electronic ground state of the molecule, while S_1, S_2, and S_n refer to the first, second, and nth singlet electronic excited states, respectively. Electronic energy levels are typically represented as bold horizontal black lines. The thin horizontal black lines represent the different vibrational energy levels for each electronic state.

The process of fluorescence begins upon the instantaneous (~10^{-15} s) absorption of a photon by the substance. The energy provided by the photon excites an electron from the singlet ground state of the molecule (S_0) to some vibrational energy level in a higher singlet excited state, S_n, (e.g., S_1). Following excitation to a higher energy state, the electron can relax and dissipate the excess energy through a variety of pathways. In the majority of cases, the first process to occur following excitation is called internal conversion, where the excited electron relaxes to the lowest vibrational energy level of S_1. Fluorescence occurs when the excited electron in S_1 returns to S_0 via the emission of a photon. Emission typically occurs on a femto to nanosecond timescale.

Alternatively, an excited electron in S_1 can undergo spin conversion to an excited triplet state, T_n, in a process called intersystem crossing. During intersystem crossing, the spin multiplicity of the electron changes to match the spin of its corresponding partner electron in S_0. Relaxation of the excited electron in T_n to S_0, though a quantum mechanically "forbidden" transition, can in fact occur. Emission of a photon during the T_n to S_0 transition is called phosphorescence and occurs on a much slower timescale than fluorescence (10^{-3} – 10^{0} s). A third common deactivation pathway for S_1 involves non-radiative decay concomitant with heat dissipation (i.e., internal conversion). This process typically competes with fluorescence and is characterized by rate constants in the range of $1\times10^{15} – 1\times10^{8}$ s^{-1}.

1.2.1.1 Non-Radiative Decay

Non-radiative decay occurs when molecules in the excited state follow relaxation pathways that do not result in the emission of a photon, such as internal conversion to the ground state (i.e., $S_1 \rightarrow S_0$) or intersystem crossing to the triplet excited state. In the former case, the excess energy is given off as heat. Molecules in the excited state are also prone to chemical reactions such as *cis-trans* isomerization (see section 1.2.5.4) or redox reactions, which result in non-luminescent relaxation from the excited state. Energy transfer between two molecules is another example of a non-radiative process (section 1.2.1.3).

1.2.1.2 Emission Quantum Yield

Due to the contribution of non-radiative decay pathways, the number of photons emitted by a molecule will always be equal to or smaller than the number of photons absorbed. The ratio between the number of photons emitted to the number of photons absorbed is defined as the emission quantum yield (Φ) (Equation 1.1). The emission quantum yield (Φ) can also be expressed as the ratio of decay rate constants, where k_r is the rate of radiative decay and Σk_{nr} is the sum of the rate constants for all non-radiative decay pathways.

$$\phi = \frac{\#\,photons\ emitted}{\#\,photons\ absorbed} = \frac{k_r}{k_r + \sum k_{nr}} \tag{1.1}$$

1.2.1.3 Förster Resonance Energy Transfer (FRET)

FRET is a powerful and widely used fluorescence technique used to study intra- and intermolecular dynamics of biomolecules.[74] The phenomenon is based on non-radiative energy transfer between a donor and acceptor dye via dipole-dipole interactions. Energy transfer via dipole-dipole interactions results in a decrease in donor emission intensity, and a concomitant increase in acceptor emission. The rate of energy transfer between a donor and acceptor (k_T) is given by:

$$k_T(r) = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \tag{1.2}$$

Where τ_D is the fluorescence lifetime of the donor (i.e., the average time the donor spends in the excited state before emitting a photon), r is the distance between the donor and acceptor, and R_0 is the Förster radius: the distance between the donor and acceptor that gives 50% energy transfer. R_0 is calculated as:

$$R_0 = \left(\frac{9(ln10)\phi_D \kappa^2 J(v)}{128\pi^5 N_A n^4}\right)^{\frac{1}{6}} \tag{1.3}$$

Where Φ is the emission quantum yield of the donor, κ is a factor describing the spatial orientation of the donor and acceptor, N_A is Avogadro's number, and $J(v)$ is an integral describing the spectral overlap between the donor emission and acceptor absorption spectra. The dipole-dipole nature of the donor/acceptor interactions results in an energy transfer efficiency (E_{FRET}) that is dependent on the 6th power of distance, making FRET very sensitive to small distance changes according to:

$$E_{FRET} = \frac{{R_0}^6}{{R_0}^6 + r^6} \tag{1.4}$$

The distance sensitivity of FRET is depicted in Figure 1.11 for a standard donor/acceptor pair Cy3/Cy5.

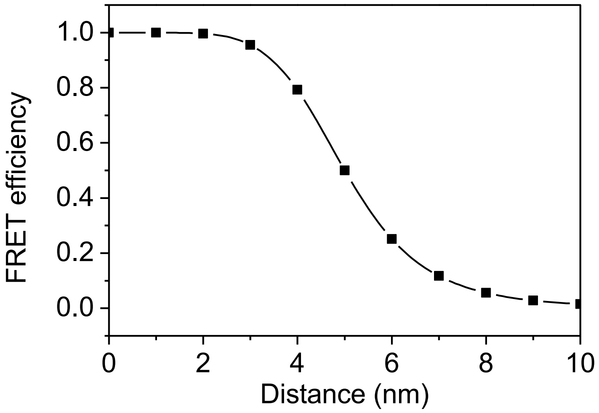

Figure 1.11 Change in FRET efficiency with increasing distance between a donor and acceptor fluorophore. Values were calculated using Equation 1.4 for the Cy3 (donor) Cy5 (acceptor) pair. The R_0 value (5.3 nm) was taken from reference [75] and is a typical value for the Cy3/Cy5 pair.

1.2.2 Fluorescence and Single-Molecule Detection

Due to its high sensitivity, fluorescence is well suited for single-molecule detection. Numerous experimental strategies are employed to detect single-molecules via fluorescence and these techniques are often based on greatly reducing the excitation volume of a sample and immobilizing molecules of interest. The following sections will describe the advantages of single-molecule detection and highlight important advances in the field of single-molecule fluorescence spectroscopy from its inception to its use in the study of complex biochemical process under physiological conditions.

1.2.3 Single-Molecule Versus Ensemble Measurements

In order to better understand the single-molecule paradigm and why single-molecule techniques are useful, it is helpful to consider the following analogy.[76] Suppose you wish to describe the traffic on a busy highway. It is possible to determine the average speed that the vehicles are traveling at. The average speed is useful as it tells you whether or not traffic is flowing, or if cars are, on average, adhering to the speed limit. The average speed of cars traveling on the highway is a single number reflecting a large population of cars, i.e., it is an ensemble measurement. What such a measurement fails to tell you is the underlying behavior of the individual cars as they move along the highway. What is the distribution of speeds among the cars on the highway? Are all cars moving at more or less the same speed, or is there a mix of drivers on the road: some with a penchant for speed, others more cautious? Are cars changing their speed over time? The answers to these questions are not apparent from the average speed measurement. It is impossible to understand and appreciate the full diversity of the "cars on the highway" scenario from a single number.

Biology is one area where single-molecule techniques have proven to be extremely powerful, as most biological processes are simply too complex to be completely understood using conventional ensemble techniques. Consider 1 mL of a 1 mg/mL enzyme solution. If the enzyme's molecular weight is 50 kDa, then there are 20 nmol of enzyme. Multiplying 20 nmol by Avogadro's number results in approximately 1.2×10^{16} enzyme molecules. An ensemble measurement on this solution reflects a property averaged over quadrillions of molecules, while in reality, the copy number of an enzyme *in vivo* may be much lower and highly sensitive to its local microenvironment.

1.2.4 The Single-Molecule Advantage

Some molecular properties are simply too difficult to study using conventional ensemble techniques. The highly coordinated functions of high-order biological assemblies often necessitate analyses at the single-molecule level. For example, the multi-subunit ATP synthase is a rotary motor protein whose orientation defines its function. Single-molecule techniques can report on the orientation of individual molecules and how orientation changes with time. Indeed, single-molecule fluorescence spectroscopy provided the first direct proof that ATP synthase was in fact a rotary motor.[77]

Another advantage of single-molecule techniques is that they enable the researcher to characterize the distribution of values of a molecular property, thus shedding light on the statistical nature of that property.[78] Statistical distributions of properties are an excellent tool for identifying molecular heterogeneity, which is an intrinsic element of any biological system.

Heterogeneity within a system can be characterized as either static or dynamic. Static heterogeneity is when different subpopulations of molecules, which do not interconvert over the course of the observation period, exist in a sample, such as active versus inactive species in a population of enzymes. For example, when determining the activity of an enzyme (k_{cat}) by ensemble methods, it is impossible to know from that k_{cat} value whether all enzyme molecules in the sample are equally active, or if a significant catalytically inactive fraction exists. Dynamic heterogeneity exists when different subpopulations of species can interconvert over the timescale of the observation, i.e., transitions of a molecule between different states, such as the stochastic fluctuations between different conformational states of a protein. Dynamic heterogeneity is often obscured in ensemble measurements due to intrinsic averaging. Single-molecule techniques are advantageous when studying dynamic heterogeneity, as they allow for the observation of transitions in real-time, without the need for synchronization.

1.2.5 Towards the Era of Single-Molecule Biology

One of the earliest accounts of single-molecule fluorescence detection came from Thomas Hirschfeld in 1976.[79] Hirschfeld's experiments demonstrated the main requirement for single-molecule detection: reduction of the excitation volume in order to extract signal from background luminescence and light scattering. In Hirschfeld's experiments, polyethyleneimine molecules were first conjugated to 80-100 molecules of fluorescein isothiocyanate and then linked to γ-globulin antibodies in a 1:1 ratio. The resulting reagent was then passed through a thin layer of high intensity illumination generated using total internal reflection (see section 2.1.1) and individual molecules were detected optically and photoelectrically as short fluorescent pulses.

Hirschfeld's method was based on the detection of emission from many fluorophores, presumably attached to a single molecule of interest. It wasn't until 1989 and 1990 that Moerner and Orrit independently reported the detection of fluorescence from individual molecules: single pentacene molecules within a *p*-terphenyl crystal at temperatures of ~1.5 K.[80, 81] In an effort to detect the fluorescence from single molecules in solution at room temperature, Shera and co-workers utilized a 70 ps pulsed laser to excite and count single Rhodamine-6G molecules as they passed through a highly focused laser beam in a thin flow cell.[82] However, even by using a microfluidic flow cell to reduce the excitation volume, the signal to noise ratio was poor and detection efficiency (~50 photons/molecule) was low.

In 1993, Betzig and Chichester pioneered a technique called near-field scanning optical microscopy (NSOM) that achieved the first repetitive imaging of a single molecule.[83] The NSOM technique allowed the authors to surpass the diffraction limit by drastically reducing the excitation volume using an optical probe with a nanometer-sized aperture placed very close (tens of nm) to the sample. By repeatedly raster scanning the sample surface relative to the optical probe, the authors obtained images and recorded the fluorescence from individual carbocyanine dyes (diIC$_{12}$) spread over a polymethylmethacrylate film at room temperature (Figure 1.12).

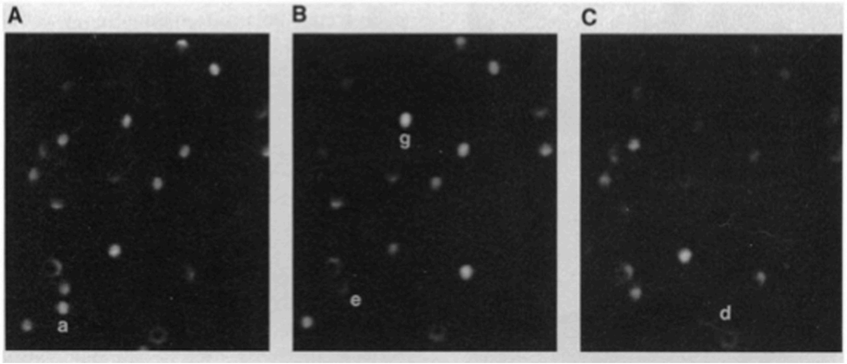

Figure 1.12 Three sequential images of individual diIC$_{12}$ dyes imaged using near-field optical fluorescence microscopy at an acquisition rate of 1/30 s. The field of view for all three images is the same. Figure reproduced with permission from reference [83].

While NSOM can be used effectively to extract the fluorescence of single molecules from background scattering and luminescence, it suffers from not being amenable to real-time imaging of protein dynamics and catalysis due to the time it takes (minutes) to scan the optical probe.

1.2.5.1 Real-Time Single-Molecule Fluorescence Imaging in Aqueous Solution

The group of Professor Yanagida made a major advance in single-molecule fluorescence imaging in 1995 by optimizing epifluorescence and total-internal reflection fluorescence (TIRF) microscopy setups to image single myosin molecules and detect single ATP hydrolysis turnovers.[84] Unlike previous work, wide-field TIRF microscopy was combined with charge-coupled device (CCD) detection, which enabled the simultaneous imaging of many surface-immobilized molecules for several seconds, in aqueous solution, and under physiological conditions. In the first experiment, the authors labeled a myosin subfragment, heavy meromyosin (MM), with a carbocyanine dye (Cy3) and spread the Cy3-MM onto a coverslip at a low density. They found that the least intense fluorescent spots (corresponding to singly-labeled MM) photobleached in one step and that the spots observed by epifluoresence corresponded to individual MM molecules imaged by electron microscopy.

Then, using TIRF microscopy, Yanagida and co-workers were able to reduce the background photon count even further and follow the intensity of many individual fluorescent spots over time at an acquisition rate of 1 frame per ~33 ms. With the capability to conduct real-time single-molecule fluorescence imaging, the authors developed a system to detect single ATP turnover events catalyzed by single-headed myosin subfragments (S-1).[84] The system involved immobilizing Cy5-labeled S-1 molecules onto a quartz coverslip at a density low enough such that single S-1 molecules could be detected and localized. Following immobilization and localization of S-1, a low concentration of Cy3-labeled ATP (10 nM) was applied to the surface. The arrival and association of Cy3-ATP with surface bound Cy5-S-1 was monitored by observing emission from Cy3 in a position corresponding to a Cy5-S-1 molecule (Figure 1.13).[84]

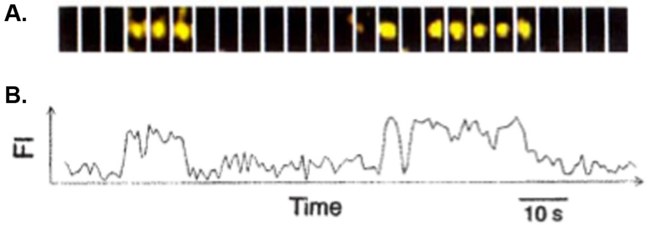

Figure 1.13 Individual association-hydrolysis-dissociation events captured using total-internal reflection fluorescence microscopy. (A) A series of images showing the appearance and disappearance of Cy3 emission over time. The emission corresponds to the position of a Cy5-labeled S-1 molecule, therefore the appearance of fluorescence represents the association between Cy3-ATP/Cy3-ADP and Cy5-S-1. (B) Cy3 emission intensity over time for the images series shown in (A). Figure reproduced with permission from reference [84].

The use of TIRF microscopy to monitor single association-hydrolysis-dissociation events demonstrated the general applicability of single-molecule fluorescence techniques towards studying biological systems under physiological conditions. The wide-field TIRF and immobilization methodologies were quickly adopted to study the sliding of kinesin and myosin over microtubules and actin filaments, respectively.[85, 86] The ability to monitor fluorescence trajectories of single molecules on a biologically relevant time scale proved that it was possible to obtain quantitative information (e.g., kinetics) about interesting biomolecular processes, such as protein/ligand interactions and enzyme catalytic activity. The studies conducted by Yanagida and co-workers were an important step away from proof-of-principle and towards novel biological insight.

1.2.5.2 Single-Molecule Enzymology

In 1998, the group of Professor Sunney Xie reported the real-time observation of single turnovers catalyzed by the enzyme cholesterol oxidase (COx) and provided new insight into a seemingly well characterized and understood system.[87] COx is a flavoenzyme that catalyzes the oxidation of cholesterol using molecular oxygen. The enzyme's active site contains a flavin adenine dinucleotide (FAD) prosthetic group that acts at the redox center. FAD is fluorescent, but loses its fluorescence upon reduction to $FADH_2$. By confining single cholesterol oxidase molecules within a 1% w/w agarose gel in the presence of excess cholesterol and oxygen, single turnover events were observed by monitoring the appearance and disappearance of FAD fluorescence. The authors discovered that single cholesterol oxidase enzymes exhibited “on-off” fluorescence behavior, i.e., fluorescence turned on and off as the flavin prosthetic group switched between oxidized (FAD) and reduced ($FADH_2$) states, respectively (Figure 1.14).

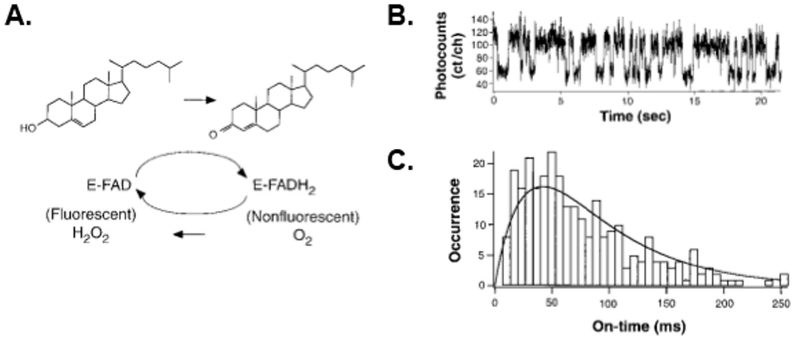

Figure 1.14 Real-time observation of single cholesterol oxidase (COx) turnovers. (A) Reaction scheme for the COx-catalyzed oxidation of cholesterol. The reaction is monitored by observing the on/off switching between fluorescent FAD and non-fluorescent $FADH_2$. (B) Fluorescence intensity time-trajectory of a single COx molecule embedded in a polyacrylamide matrix in the presence of 0.2 mM cholesterol and 0.25 mM O_2. The upper and lower levels correspond to fluorescent FAD (on) and non-fluorescent $FADH_2$ (off). Each on/off cycle represents a single COx-catalyzed turnover. (C) Dwell-time distribution of the durations of the fluorescent on-times. The distribution does not follow first-order single-exponential kinetics, suggesting a more complex kinetic scheme. Figure adapted with permission from reference [87].

Through careful analysis of the distribution of fluorescence on-times (the waiting time before FAD reduction), it was evident that the process did not follow single exponential kinetics, suggesting the presence of an intermediate enzyme/substrate complex. Unexpectedly, further statistical analyses showed that individual fluorescence trajectories displayed dynamic heterogeneity: the catalytic rate constant for a single COx molecule changed over time. The origin of the heterogeneity was attributed to interconverting conformations of COx, each with a characteristic rate constant, resulting in a “molecular memory” effect.

These discoveries had far reaching implications. At the *in vitro* ensemble level, the catalytic rate of an enzyme is a value averaged over many molecules and the presence of dynamic heterogeneity is not evident. However, in a cell, the copy number of an enzyme may be low and the dynamic heterogeneity from a single enzyme could have very important physiological consequences. Indeed, dynamic heterogeneity in an enzyme's catalytic rate constant due to conformational fluctuations has since been observed for a variety of different enzymes.[88-91] The single-molecule enzymology work by Xie and co-workers was a seminal contribution to the field as it solidified single-molecule spectroscopy as a powerful tool to unravel unique mechanistic features of an enzyme-catalyzed chemical reaction.

1.2.5.3 Single-Molecule FRET

One of the most widely used single-molecule fluorescence techniques is single-molecule Förster resonance energy transfer (SM-FRET). FRET can be used to report on inter- and intramolecular dynamics by exploiting the distance-dependent energy transfer between a donor and acceptor fluorophore (Figure 1.15). Ha et al. described the first detection of SM-FRET in 1996.[92]

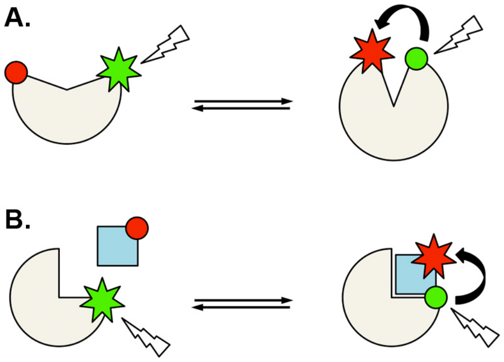

Figure 1.15 Probing intra- (A) and intermolecular (B) dynamics using single-molecule FRET. In both cases, the donor (green) is irradiated. When the acceptor dye (red) comes into close proximity of the donor, e.g., during a conformational change (A) or ligand binding (B), energy transfer occurs and emission from the acceptor increases.

The widespread adoption of SM-FRET has led to important insights into many complex biochemical processes, such as RNA and protein folding,[72, 91, 93] protein and nucleic acid conformational dynamics,[75, 94-96] DNA unwinding,[97] homologous recombination,[98] and translation.[71]

1.2.5.4 Single-Molecule Protein-Induced Fluorescence Enhancement

Recently, single-molecule protein-induced fluorescence enhancement (SM-PIFE) has emerged as a means of studying biomolecular dynamics in the distance range of 0-4 nm.[99] PIFE is a phenomenon that exploits the environmental sensitivity of photoisomerizable dyes, such as those in the cyanine family. The basic premise is that when a protein binds close to a photoisomerizable dye, the fluorescence intensity of that dye increases (Figure 1.16). Cyanine dyes such as Cy3 can undergo *cis-trans* isomerization around the C-C bonds in the polymethine chain. Isomerization of Cy3 in the excited state results in nonradiative decay, which ultimately lowers the overall emission quantum yield.[100] If *cis-trans* isomerization is constrained due to sterics and/or an increase in local viscosity (e.g., when a protein binds nearby), its contribution to the rate of nonradiative decay decreases. The end result is an increase in emission quantum yield. The PIFE effect is akin to the increase in Cy3 emission quantum yield in viscous solvents such as ethylene glycol.[101]

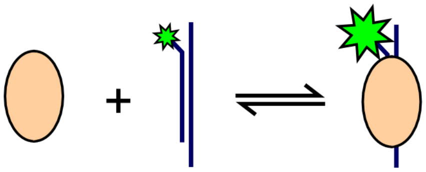

Figure 1.16 Cartoon representation of single-molecule protein induced fluorescence enhancement (SM-PIFE). Binding of a protein (orange circle) close to a photoisomerizable dye leads to fluorescence enhancement of the dye due to a decrease in the non-emissive relaxation rate.

PIFE has been demonstrated at the ensemble level,[102, 103] but has recently been adapted to report on the dynamics of single biomolecules.[96, 99, 104, 105] In one of the earliest demonstrations of SM-PIFE, Xie and co-workers utilized TIRF microscopy and a surface-immobilized Cy3-labeled double-stranded DNA substrate to study the kinetics of a conformational change in the bacteriophage T7 DNA polymerase.[96] In the absence of protein, the fluorescence from intensity-time trajectories of single Cy3-labeled DNA molecules was found to be steady (Figure 1.17A). However, upon the addition of 5 nM T7 polymerase, the intensity-time trajectories began to show fluctuations between high and low emission intensity, which the authors attribute to single T7 polymerase binding/unbinding events, respectively.

The authors extended the assay to look at a conformational change in the T7 polymerase fingers domain during formation of the polymerase/DNA/dNTP ternary complex. By adding the next complementary dNTP, a third intensity level was observed (Figure 1.17B). The duration of the third intensity level was dependent on dNTP concentration.

In contrast to SM-FRET, SM-PIFE requires only one molecular species to be fluorescently labeled. This is a major advantage because it vastly increases the repertoire of proteins available for single-molecule studies. For example, site-specific protein labeling is generally achieved using maleimide chemistry, necessitating the requirement of a unique and solvent accessible cysteine residue. Since many proteins contain multiple cysteines, site-directed mutagenesis is carried out to knock out the undesirable cysteine residues. In a biological system it is always desirable to minimize the amount of site-directed mutagenesis and covalent modifications to reduce the chances of perturbing the normal function of the system.

Importantly, distance calibration studies of SM-PIFE have been conducted and show that the effect is sensitive in the range of < 4 nm.[99] As a result, SM-PIFE may be used to detect small distance changes that fall outside the range of SM-FRET.

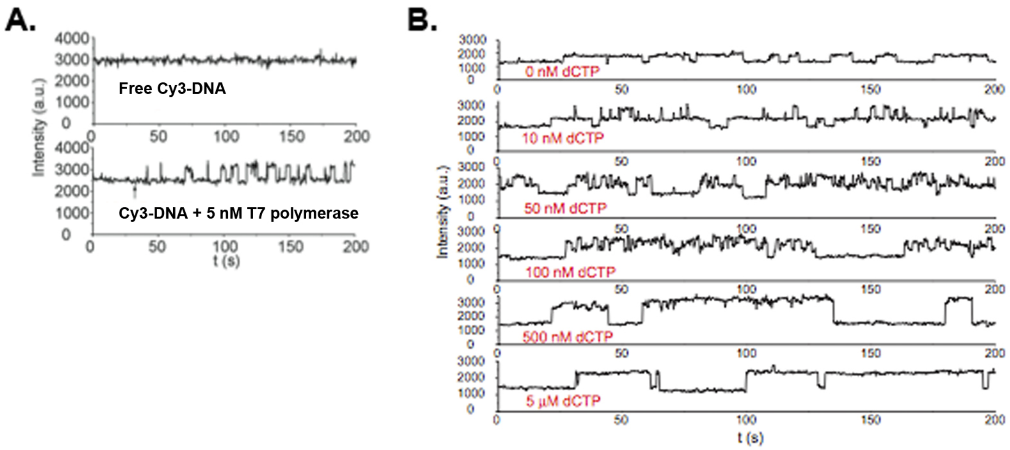

Figure 1.17 Single-molecule PIFE assay of T7 DNA polymerase binding to a surface-immobilized Cy3-labeled DNA duplex. (A) Intensity-time trajectories of Cy3-DNA in the absence (upper panel) and presence (lower panel) of 5 nM T7 DNA polymerase. In the lower panel, the fluorescence toggles between two intensities, corresponding to un-bound and protein-bound Cy3-DNA. (B) Intensity time-trajectories of Cy3-DNA in the presence of 10 nM T7 polymerase and increasing concentrations of dCTP. Introduction of dCTP introduces a third intensity level corresponding to a conformational change in the T7 polymerase fingers domain. Figure adapted from reference [96].

1.2.6 Summary

Since the introduction of single-molecule techniques in the late 1970's, the number of published papers with the words "single-molecule" in the title has grown exponentially over time.[106] Tremendous strides have been made towards being able to study complex biological processes under physiological conditions and single-molecule fluorescence techniques have provided a wealth of novel insight into processes not amenable to ensemble biochemical studies. As single-molecule techniques become more and more accessible to the general research community, the single-molecule field will continue to grow. Chapter 2 will introduce some of the general single-molecule fluorescence imaging methodologies used in our lab, specifically total-internal reflection fluorescence microscopy and sample preparation. Many single-molecule techniques exist (optical- and force-based) and are reviewed elsewhere.[78, 106-108]

1.3 Research Goals

The current status of HIV-1 as a global epidemic together with the inevitable failure of antiretroviral drug therapies has accentuated the need for new avenues of therapeutic intervention. Due to its critical role in the HIV-1 lifecycle, disrupting reverse transcription is and continues to be an attractive strategy for combating HIV-1 infection. We need to better understand the dynamic nature and intricacies of critical viral genome replication steps (i.e., reverse transcription) such that we can 1) identify new potential drug targets and 2) rationally design and develop novel inhibitors that are both potent and specific.

Crystallographic studies have provided a vast collection of atomic "snapshots" of important viral proteins at different stages of the replication process. Despite the wealth of knowledge obtained from these snapshots, the dynamic behavior of key viral enzymes is often inferred from static measurements or ensemble-averaged data. As described above, the emergence of single-molecule fluorescence techniques in recent years has made it possible to probe complex and dynamic biological processes with millisecond time- and nanometer spatial-resolution, effectively producing molecular "movies".

Our goal is to develop and implement highly sensitive single-molecule fluorescence techniques to gain a detailed mechanistic understanding of the HIV-1 replication process. Specifically, we wish to understand the dynamics of HIV-1 RT and its interactions with different replication intermediates that arise during reverse transcription. Ultimately, we wish to extend our single-molecule methodology to examine how drugs (e.g., NRTIs & NNRTIs) and drug-resistance conferring amino acid mutations modulate RT dynamics at different stages of reverse transcription. We hope that single-molecule fluorescence will allow us to correlate changes in RT dynamics to changes in RT catalytic activity, which will help us elucidate the molecular mechanisms of drug inhibition and drug resistance.

Towards achieving these goals, we have developed a single-molecule protein-induced fluorescence enhancement assay that allows us to visualize the binding and dissociation of hetero- and homodimeric RT to and from individual fluorescently labeled model DNA/DNA nucleic acid substrates.

2. General Methodology

2.1 Total Internal Reflection Fluorescence Microscopy

Total internal reflection fluorescence (TIRF) microscopy exploits the fundamental properties of reflected light at an interface to achieve a small excitation volume adjacent to that interface. In recent years, TIRF microscopy has become the go-to technique for single-molecule studies due to its excellent signal-to-noise ratio and ability to simultaneously monitor the fluorescence from many single molecules in real-time. The following sections will discuss the theory behind the technique and provide a description of the TIRF microscopy system used in our lab.

2.1.1 Total Internal Reflection

When a beam of light travelling through a medium of refractive index n_1 encounters the interface with a second medium of lower refractive index n_2, the incident light may be partially reflected at the interface and refracted into the second medium. The angle of the refracted light depends on the incident angle and the refractive indices of the two media. The relationship between the angles of incidence (θ_1) and refraction (θ_2) and the indices of refraction of the two media (n_1, n_2) is governed by Snell's law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{2.1}$$

When the refracted beam is parallel to the interface (i.e., $\theta_2 = 90°$), the incident angle is defined as the critical angle (θ_c). Rearranging Snell's law, the critical angle may be written as:

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{2.2}$$

For incident angles equal to or higher than the critical angle, all of the incident light is reflected at the interface back into the high refractive index medium in a phenomenon called total internal reflection.

2.1.2 Total Internal Reflection Microscopy

Above the critical angle, incident light is totally internally reflected at the interface. However, electrical and magnetic fields cannot become discontinuous at a boundary and an electromagnetic oscillation called an evanescent wave is generated that extends into the lower refractive index medium (Figure 2.1).[109]

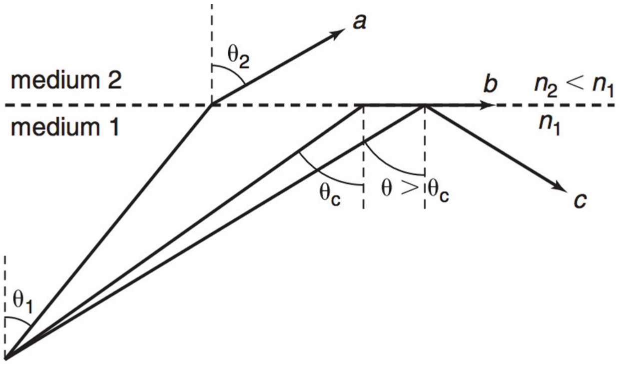

Figure 2.1 Light paths at an interface between two media with refractive indices n_1 and n_2, with $n_2 < n_1$. (A) When the incident angle (θ_1) is smaller than the critical angle (θ_c), light is refracted at the interface into the second medium at an angle θ_2. (B) When $\theta_1 = \theta_c$, light propagates parallel to the interface (i.e., $\theta_2 = 90°$). (C) When $\theta_1 > \theta_c$, total-internal reflection occurs. Figure reproduced with permission from reference [110].

The evanescent wave possesses the same frequency as the incident light, however the intensity of the evanescent wave decays exponentially with distance from the interface according to the following equation:

$$I(z) = I_o e^{-\frac{z}{d}} \tag{2.3}$$

Where I(z) is the intensity of the evanescent wave at a perpendicular distance z from the interface between the two media and I_o is the intensity of the evanescent wave at the interface (i.e., z = 0). The penetration depth, d, is a characteristic distance that depends on the wavelength of the incident light (λ), the refractive indices of the media (n_1, n_2), and the angle of incidence (θ_1) according to:

$$d = \frac{\lambda}{4\pi\sqrt{n_1^2 \sin^2\theta_1 - n_2^2}} \tag{2.4}$$

The dependence of d on the incident angle is illustrated in Figure 2.2 for λ = 532 nm at the glass/water interface ($n_{glass} = 1.52$, $n_{water} = 1.33$), where $\theta_c = 61°$.

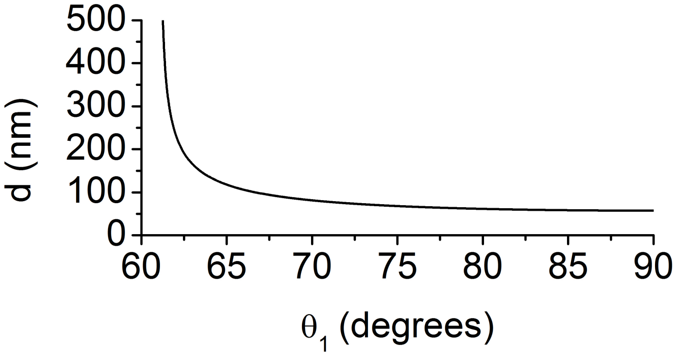

Figure 2.2 Penetration depth (d) of an evanescent wave as a function of incident angle (θ_1). Values were calculated using equation 2.4 for a glass/water interface ($n_{glass} = 1.52$, $n_{water} = 1.33$) where λ = 532 nm ($\theta_c = 61°$).

Since the intensity of the evanescent wave decays exponentially with distance (Figure 2.3), only a thin layer (ca. 500 nm) is subject to illumination in the lower refractive index medium.

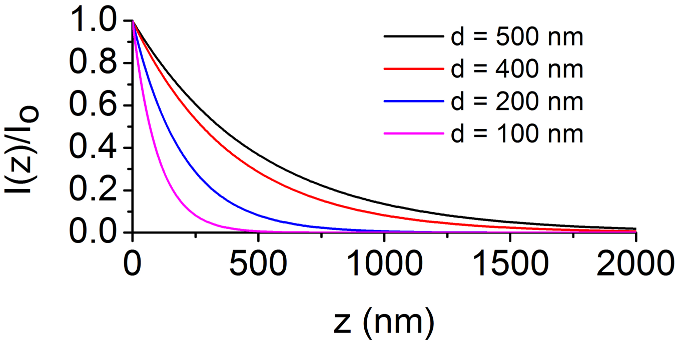

Figure 2.3 Changes in the relative intensity of the evanescent wave with increasing perpendicular distance (z) from the interface for different penetration depths (d). Values were calculated using Equation 2.3.

Total internal reflection fluorescence microscopy (TIRFM) exploits the optical properties of the evanescent wave to selectively illuminate a small volume of sample adjacent to the interface of two media, such as the glass/water interface of a microscope slide (Figure 2.4).

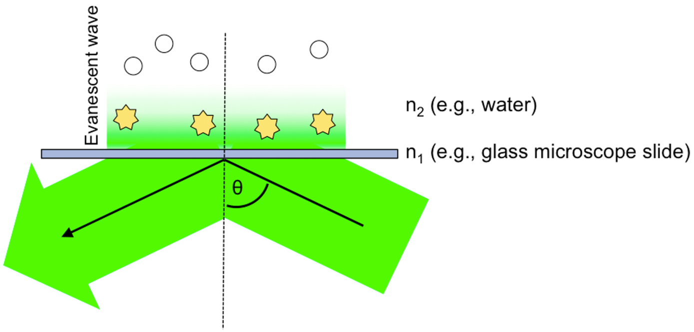

Figure 2.4 Cartoon depicting the generation of an evanescent wave at the glass/water interface. The intensity of the evanescent wave decays exponentially with distance, producing a thin sheet of illumination. Molecules within the excitation volume (yellow stars) undergo fluorescence, while molecules outside of the excitation volume (white circles) are not excited and do not fluoresce.

Studies by Ambrose in 1961 on the locomotory mechanism of chick heart fibrocytes describes one of the earliest incarnations of total internal reflection microscopy.[111] Termed "surface contact microscopy", the light from a mercury arc lamp was directed through a glass prism via a slit and reflected at the interface between the glass microscope slide and the water-mounted sample. By directing the incident light at an angle greater than the critical angle, Ambrose was able to visualize parts of the cells only in direct contact with the glass slide.

Twenty years later, Axelrod extended total internal reflection microscopy to fluorescence by introducing a laser beam as the light source.[112] By combining total internal reflection with fluorescence, Axelrod presented a strategy to greatly reduce background light scattering and autofluorescence of components within the cell.

2.1.3 Prism-Based TIRF

Prism-based TIRF utilizes a silica prism, which is placed on top of the sample on an inverted microscope. A laser beam is focused onto the prism from above at a high incident angle, passes through the prism, and is totally internally reflected at the microscope slide/water interface (Figure 2.5). A water immersion objective at the bottom of the imaging chamber collects the emission and directs it to a detector such as an electron multiplied charge coupled device (EM-CCD) camera. Prism-based TIRF microscopy is relatively cheap and easy to setup and offers low background since the excitation light does not enter the objective. However, a major disadvantage of prism-based TIRF is that it is more laborious, as the prism must be reassembled for every new sample.

2.1.4 Objective-Based TIRF

In objective-based TIRF, total internal reflection at the interface is achieved by focusing the excitation beam onto the edge of the objective's back focal plane (Figure 2.5). Fluorescence from the sample is then collected with the same objective. Objective-based TIRF has numerous advantages, including easier sample accessibility and a higher photon-collection efficiency.[113]

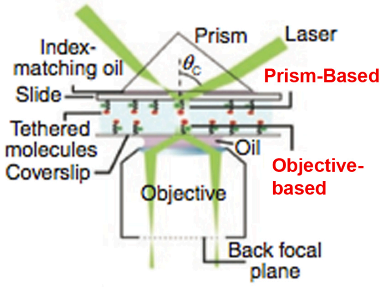

Figure 2.5 Prism- vs. objective-based TIR illumination. In prism-based TIR, the incident light enters a prism from above, generating an evanescent wave at the slide/water interface. In objective-type TIR, the incident light is focused onto the edge of the objective's back focal plane and leaves the objective. Figure adapted with permission from reference [113].

In objective-based TIRF, the incident angle is limited by the numerical aperture (NA) of the objective lens. For an objective with a given NA, the maximum achievable incident angle is given by:

$$\theta_{NA} = \sin^{-1}\left(\frac{NA}{n_m}\right) \tag{2.5}$$

Where n_m is the refractive index of the coverslip material, in this case, glass. Since TIR requires an incident angle > θ_c, then θ_{NA} must > θ_c. For the glass/water interface ($\theta_c = 61°$), this means that NA > 1.33. Objective lens with NA = 1.4-1.45 are typically used in practice.

For a typical objective-based TIRF setup, the incident angle is adjusted by moving the beam laterally away from the center of the objective (i.e., perpendicular to the direction of light propagation) towards the edge of the objective back focal plane.

2.1.5 Detection

Charge-coupled device (CCD) cameras are the main method of detection in single-molecule TIRF applications.[109] A CCD chip is a semiconducting silicone wafer that is divided into thousands of pixels. Typically, arrays of 512 × 512 (i.e., 262 144 pixels) are used. CCD detectors function by converting the energy from photons into an electrical charge. During an exposure, photons from the sample emission reach the chip surface and as photons hit the chip surface, electrons are continuously released and stored in potential wells. Each pixel has its own potential well. Following each exposure, the charge from each potential well moves down the chip, one row at a time. An analog-digital converter subsequently reads out each row one pixel at a time, converting the voltage into a digital signal.

Given the high sensitivity required in single-molecule experiments, we utilize novel electron multiplying CCD (EM-CCD) cameras with maximum gain and maximum quantum efficiency. EM-CCD cameras rely on an on-chip electron multiplication device that enables the detection of low levels of light by multiplying the photon-generated charge above background read-out noise in each pixel. Electron multiplication results in greater sensitivity and enables faster sampling rates by decreasing the required exposure time.[109]

2.2 Our TIRF Microscope

Our laboratory is equipped with two Olympus IX71 inverted microscopes configured for TIRF (Figure 2.6). The schematic below depicts the optical configuration used in this thesis and is suitable for Cy3 and/or Cy5 detection. To excite Cy3 we use a diode-pumped solid-state 532 nm cw laser (Crystal Laser, Reno, NV). A dichoric mirror (z532rdc, Chroma Technology, Rockingham, VA) directs the laser beam to a 60× PlanApo oil-immersion microscope objective (N.A. = 1.45, Olympus). The beam is focused onto the edge of the back focal plane of the objective using a TIRF illumination module (IX2-RFAEVA-2, Oympus) to generate an evanescent wave at the glass/water interface. Fluorescence from the sample is collected through the same objective and is focused by the microscope tube lens with a focal length of 150 mm at the image plane. A second lens with a focal length of 150 mm acts as a relay lens projecting the image at the image plane onto the EM-CCD chip. A dichroic mirror (640dcxr) then splits the emission: light of wavelength >640 nm passes through the mirror, while light of wavelength <640 nm is reflected. Bandpass filters (e.g., HQ 590/70 and HQ 585/80 for Cy3 and Cy5 emission, respectively) are used to remove scattered light before the photons from each emission beam are projected onto separate halves of an EM-CCD chip.

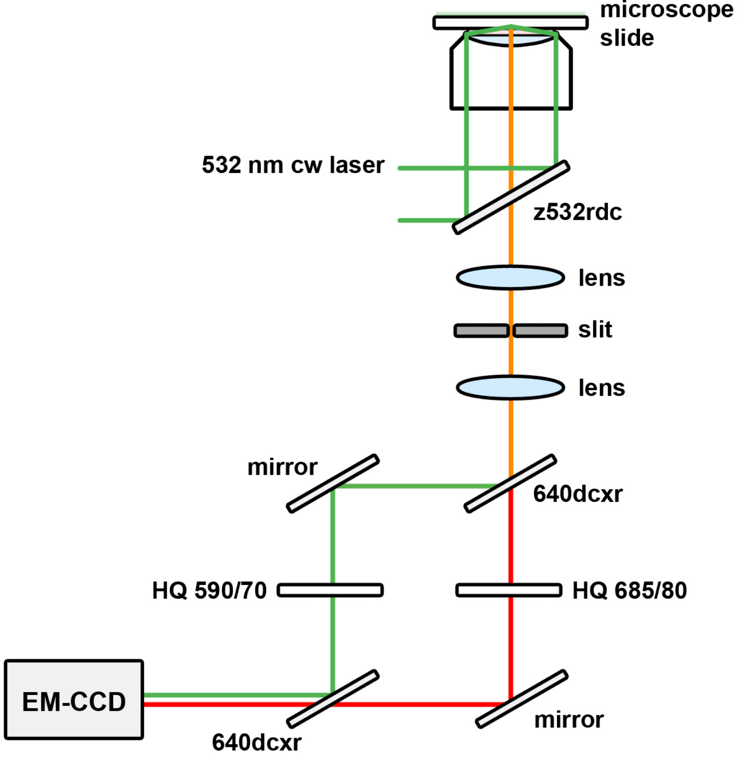

Figure 2.6 Schematic of two-colour TIRF microscopy setup.

2.3 Fluorophores for SMS Studies

Since the optical detection of single biomolecules is currently not feasible, it is necessary to utilize some sort of reporter molecule. Due to its high sensitivity, fluorescence is an ideal method for single-molecule detection. Fluorophores are excellent reporter molecules and well suited to probe biological dynamics, as numerous different fluorescence properties can be monitored:

1. *Intensity*: First, simply being able to detect a single fluorophore-tagged biomolecule is useful in itself. Changes in fluorescence intensity can give insight into changes in the local environment around a fluorophore, as seen in single-molecule protein-induced fluorescence enhancement (see

section 1.2.5.4). Additionally, fluorescence intensity can be tracked as a labeled molecule moves in space to study diffusional properties.

2. *Anisotropy*: Fluorescence polarization and anisotropy measurements can report on the orientation and rotational dynamics, such as the movement of motor proteins along helical substrates[114] and the molecular rotor proteins.[115]
3. *Energy transfer*: Single-molecule FRET (see section 1.2.5.3) exploits the distance-dependent energy transfer between a donor and acceptor dye and can report on intramolecular dynamics such as protein conformational changes,[95] and intermolecular interactions.[116]
4. *Lifetime*: Fluorescence lifetime measurements may be sensitive to changes in local environment around the fluorophore.[117, 118]

Choosing the proper fluorophore(s) is a critical step when designing single-molecule fluorescence experiments. There exists an extremely wide variety of fluorophores to choose from and each individual fluorophore will have its own unique properties. General desirable properties of fluorophores for single-molecule experiments include:[113]

1. *Brightness*: Fluorophores must be bright enough such that they can be detected. Brightness is defined as the product of the absorption extinction coefficient and the emission quantum yield (i.e., $\varepsilon \times \Phi$).
2. *Photostability*: Photobleaching and blinking are the main limitations in single-molecule fluorescence experiments and fluorescence microscopy in general. Photobleaching is the process of irreversible destruction of the fluorophore, while photoblinking (or fluorescence intermittency) is the reversible process of switching between bright and dark states of a fluorophore (e.g., due to excursions to the triplet state). Methods of enhancing photostability are discussed in section 2.4.3.
3. *Minimal aberrant photophysical effects*: Changes in fluorescence properties (e.g., intensity, energy transfer) should reflect biomolecular

dynamics and not photophysics of the fluorophore. For example, in SM-FRET experiments, fluctuations in Cy5 intensity due to photoblinking can be mistaken for true FRET fluctuations.[100]

4. *Amenable to conjugation chemistry*: Labeling of biomolecules with fluorophores requires reliable and efficient conjugation chemistries. Protein-labeling is commonly achieved by reacting free thiol or amino groups with maleimide- or *N*-hydroxysuccinimide-functionalized fluorophores, respectively.
5. *Water solubility*: For biological applications, fluorophores must be able to dissolve and function in aqueous solutions. To achieve water solubility, charged functional groups such as carboxylates and sulfonates are often added to fluorophores.

For the single-molecule studies described in this thesis, we chose to utilize indocarbocyanine dyes, specifically Cy3 (Figure 2.7). Cyanine dyes are ideal for biophysical studies due to their excellent photostability, high absorption extinction coefficient (ε), relatively high emission quantum yield (Φ) (and thus high overall brightness, i.e., ε × Φ), and amenability to a variety of conjugation chemistries to biomolecules.

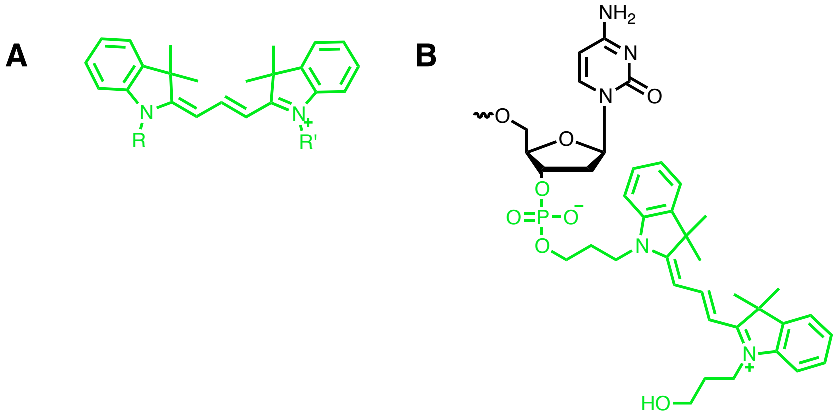

Figure 2.7 (A) General structure of Cy3. (B) Structure of Cy3 (green) covalently attached to a 3'-terminal deoxyribonucleotide (black).

2.4 Coverslip and Sample Preparation

In our SMS system we utilize a surface-immobilization strategy in order to monitor the fluorescence of single molecules over the course of minutes. Various immobilization strategies have been described in the literature including biotin-streptavidin chemistry, chelation of metal ions (e.g., Ni^{2+}, Cu^{2+}), and click chemistry.[119] We exploit biotin-steptavidin chemistry by using coverslips functionalized with poly(ethylene) glycol (PEG) and biotinylated-PEG. The following section will describe our general coverslip cleaning and surface functionalization protocols.

2.4.1 Cleaning and Aminosilanization

The first step in coverslip preparation is treatment of standard glass coverslips with a piranha solution consisting of 30% H_2O_2 and concentrated H_2SO_4 mixed in a 1:3 (v/v) ratio. Coverslips are placed in a glass microscope slide-staining jar and left in piranha solution for at least 1 h. After 1 h, the piranha solution is removed and neutralized with $NaHCO_3$ and the coverslips are rinsed three times with distilled deionized water, followed by three times with acetone (99.5% HPLC grade, ACP Chemicals, Montreal, QC, Canada).

The piranha solution is a strong oxidizing reagent that reacts with most organic material, ensuring that the coverslips are thoroughly clean, minimizing the chance of detecting fluorescence from surface impurities. Additionally, the piranha solution hydroxylates the glass surface, generating many Si–OH groups, which is required for the following aminosilanization reaction.

The resulting Si–OH groups need to be converted to free amino ($–NH_2$) groups so that the surface can be functionalized with PEG (see following section). To achieve this we use the aminosilanization reagent Vectabond™ (Vector Laboratories, Inc., Burlingame, CA), a proprietary compound of unknown structure (though surrogates such as 3-aminopropyltriethoxysilane may also be used). Aminosilanization with Vectabond™ is carried out by adding 25 mL of dry acetone to the coverslips, followed by 0.5 mL of Vectabond™. The reaction is mixed by gently shaking the jar, and then allowed to sit for 5 min. After 5 min, the reaction is quenched by the addition of water (HyClone Molecular Biology Grade, Fisher Scientific, Logan, UT). Coverslips are rinsed with water twice for 30 s and dried under a stream of N_2.

2.4.2 Surface Functionalization with PEG/biotin-PEG

The purpose of functionalizing the surface with PEG is to minimize the amount of non-specific surface interactions, particularly with proteins. The overall reaction is shown schematically in Figure 2.8. We treat the aminosilanized surface with a mixture of MW 5000 methoxy-PEG-succinimidyl valerate (mPEG-SVA) and MW 5000 biotinylated-PEG-SVA (biotin-PEG-SVA) in a ratio of 99:1 (w/w) (Laysan Bio, Inc.). The $–NH_2$ groups on the surface can displace the succinimidyl group, resulting in the formation of an amide bond between the surface and the PEG. The reaction is also carried out in 0.1 M $NaHCO_3$ (pH 8.0) to ensure that a sufficient number of surface $–NH_2$ groups remain unprotonated.

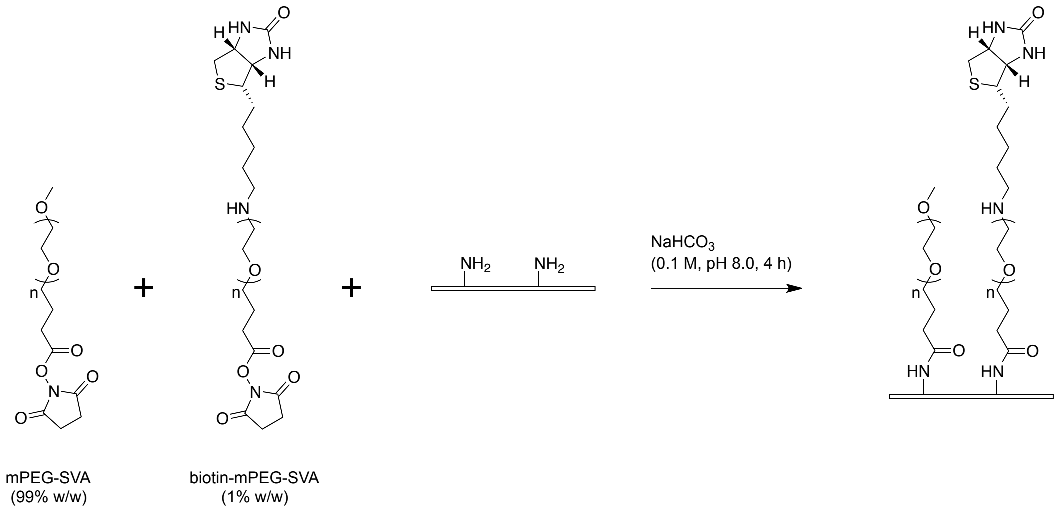

Figure 2.8 Reaction scheme depicting the functionalization of an aminosilanized glass coverslip with mPEG and biotin-mPEG.

Clean coverslips are masked with silicone films (Grace Bio-Labs, Bend, OR) and the mPEG-SVA/biotin-PEG/$NaHCO_3$ solution (45 μL) is added to the unprotected area (Figure 2.9).

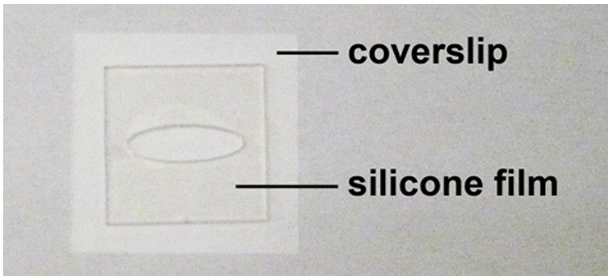

Figure 2.9 Picture of a coverslip masked with a silicone film. The center oval is the unprotected area to which the mPEG-SVA/biotin-PEG/$NaHCO_3$ solution is added.

The coverslips are left to incubate in the dark, at room temperature, for 4 h. Following the 4 h incubation, the silicone films are removed, the PEG is washed off with Molecular Biology Grade water, and the coverslips are dried under a stream of N_2. Imaging chambers are constructed by pressing a polycarbonate film with an adhesive gasket (Grace Bio-Lab, Bend, OR) onto a PEG-coated coverslip. Two silicone connectors (Grace Bio-Lab, Bend, OR) are then glued onto the predrilled holes of the film, serving as inlet and outlet ports (Figure 2.10). After assembling the imaging chambers, streptavidin (0.2 mg/mL; Sigma Aldrich, St. Louis, MO) is added, incubated for 10 min and washed. Next, a biotinylated sample in the 100 pM range is added to achieve surface immobilization of a molecule of interest with molecules sufficiently apart from each other to ensure that they can be resolved by the objective. Typically, surface densities will be smaller than 1 molecule/μm^2.

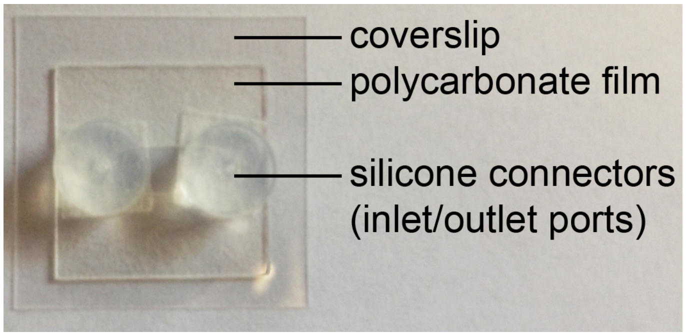

Figure 2.10 Picture of a coverslip with an assembled imaging chamber. A polycarbonate film is pressed onto a PEG-coverslip and silicone connectors are attached with double-sided tape to create inlet and outlet ports.

2.4.3 Oxygen Scavenging System

The limited photostability of organic fluorophores (e.g., due to blinking and photobleaching in the presence of singlet oxygen) is a major impediment when conducting single-molecule fluorescence experiments. Singlet oxygen (O_2), sensitized upon generation of a fluorophore's triplet excited state, readily reacts with almost all known fluorophores,[73] therefore it is often necessary to remove dissolved O_2 from single-molecule samples. One of the most popular O_2-removal methods (and the one used in all single-molecule experiments herein) is the use of an enzymatic scavenging system composed of glucose oxidase and catalase (Figure 2.11). In this system, glucose oxidase catalyzes the oxidation of β-D-glucose to D-glucono-1,5-lactone, which subsequently hydrolyzes to gluconic acid. The hydrogen peroxide produced in the reaction is converted into water and oxygen by catalase.

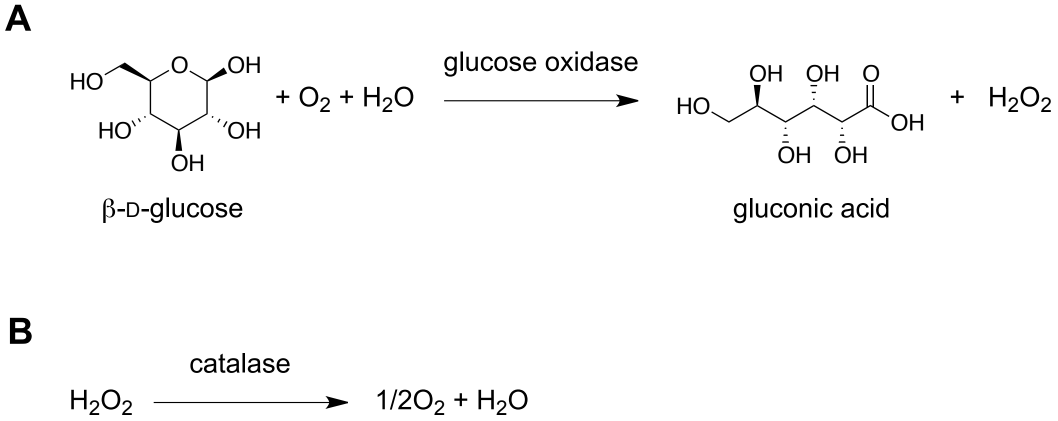

Figure 2.11 Enzymatic oxygen scavenging system. (A) Glucose oxidase catalyzes the oxidation of β-D-glucose using molecular oxygen, producing gluconic acid and hydrogen peroxide. (B) Hydrogen peroxide is converted to water and oxygen by catalase.

When employing the glucose oxidase-based oxygen scavenging system, one must exert caution due to the continuous production and accumulation of gluconic acid in solution. If the single-molecule sample is left open to air, the build-up of gluconic acid can lead to decreases in buffer pH of as high as ~2 pH units (i.e., a ~100-fold increase in H^+ concentration).[120] Such drastic changes in pH can have undesirable consequences in single-molecule biological assays considering, for example, the sensitivity of a protein's protonation state or rate of enzymatic activity to pH. In order to maintain a constant pH throughout the course of single-molecule experiments, oxygen scavenger solutions are prepared fresh and immediately transferred to syringes.

Excited-state singlet oxygen is highly reactive and leads to fluorophore photobleaching. In contrast, ground-state triplet oxygen is an efficient quencher of a fluorophore's triplet dark-state. By removing O_2 from the system, the fluorophore's residence time in the triplet dark-state increases. The oxygen scavenger solution therefore needs to be supplemented with a triplet-state quencher, such as β-mercaptoethanol or Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid). Details on their mechanism of action may be found in reference [113].

3. Binding Kinetics and Affinities of Heterodimeric versus Homodimeric HIV-1 Reverse Transcriptase on DNA-DNA Substrates at the Single-Molecule Level

3.1 Preface

Single-molecule studies of HIV-1 RT were initiated in our group by Dr. Hsiao-Wei Liu in collaboration with the group of Prof. Matthias Götte from the Department of Microbiology and Immunology at McGill University. In 2008, the group of Xiaowei Zhuang at Harvard University developed a SM-FRET assay to demonstrate that RT could bind DNA and RNA primers in either DNA polymerase- and RNase H-dependent orientations, respectively.[121] Unexpectedly, the authors also discovered that RT could flip rapidly between both orientations, without dissociating from the substrate, when bound to nucleic acid duplexes containing the unique RNA PPT primer. Furthermore, small molecules such as dNTPs and the NNRTI nevirapine could modulate the rate of enzyme flipping.

As a long-term goal, we wished to extend the SM-FRET methodology of Zhuang and co-workers to address important questions related to mechanisms of drug action and antiretroviral drug resistance in the context of reverse transcription. How do drug-resistance conferring amino acid mutations affect the molecular acrobatics of HIV-1 RT? Can we correlate changes in the binding orientation equilibrium to changes in RT polymerase or RNase H activity observed in drug resistant mutants? How do resistance-conferring mutations affect different stages of reverse transcription (i.e., minus- versus plus-strand DNA synthesis)?

The following chapter will provide a brief account of the origins of the single-molecule HIV-1 RT project in our laboratory, followed by a detailed description of an alternative single-molecule fluorescence assay for HIV-1 RT that we have developed. The new single-molecule assay is based on protein-induced fluorescence enhancement of a fluorescently labeled DNA substrate and requires no protein modification.

3.1.1 Single-Molecule FRET Studies of HIV-1 RT

The initial goal of this project was to implement in our laboratory the SM-FRET methodology described by Zhuang and coworkers (Figure 3.1).[121, 122] The project was initiated by Dr. Hsiao-Wei Liu, a postdoctoral fellow in our group. In the original assay, Zhuang and coworkers designed DNA/DNA and RNA/DNA primer/template complexes based on the unique polypurine tract (PPT) and a random sequence called "random sequence A", or simply "sequence A". The substrates consist of a biotinylated 50-nucleotide template strand labeled with Cy5 (FRET acceptor) and a 19-nucleotide primer (RT's nucleic acid binding cleft can accommodate 19 base pairs). The nucleic acid substrates were immobilized onto a quartz coverslip via biotin/streptavidin interactions and imaged using TIRF microscopy. RT labeled with Cy3 (FRET donor) was flowed into the sample chamber and FRET was used to report on RT binding. The two binding orientations of RT were observed as distinct FRET efficiency values.

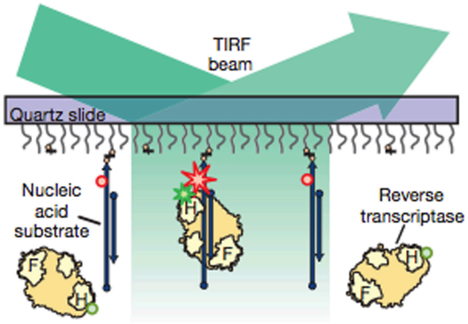

Figure 3.1 Cartoon depicting the SM-FRET assay developed by Zhuang and coworkers to study the orientational dynamics of HIV-1 RT. Nucleic acid substrates labeled with Cy5 (acceptor) were immobilized onto a quartz coverslip via biotin/streptavidin interactions and imaged using TIRF microscopy. RT labeled with Cy3 (donor) was flowed into the sample chamber and FRET was used to report on RT binding to the nucleic acid substrates. Figure reproduced with permission from [121].

In order to utilize the SM-FRET depicted in Figure 3.1, one requires that RT be fluorescently labeled. Recombinant wild-type RT and various RT mutants have been expressed and purified by our collaborators in the group of Prof. Matthias Götte (Department of Microbiology & Immunology, McGill University, Montreal, QC) and our lab has attempted many different fluorescent labeling strategies in order to site-specifically label RT. Dr. Hsiao-Wei Liu has conducted extensive studies in an attempt to implement the SM-FRET RT assay in our laboratory, however, the results of Zhuang and co-workers have not been faithfully reproduced, in part due to difficulties in the fluorescent labeling of RT.

As a result, we sought an alternative single-molecule assay based on protein-induced fluorescence enhancement (PIFE). PIFE, which does not require labeling of the protein, allows us to directly visualize the binding/unbinding of RT to a double-stranded DNA substrate. The remainder of this chapter describes the use of SM-PIFE to study the binding kinetics and affinities of hetero- and homodimeric RT on a DNA-DNA substrate. We provide values for the association and dissociation rate constants of the RT homodimers p66-p66 and p51-p51 with a double-stranded DNA substrate and compare those to values recorded for the RT heterodimer p66-p51. We also report values for the equilibrium dissociation constants for the three RT isoforms. Our data reveal great similarities in the intrinsic binding affinities of p66-p51 and p66-p66, with characteristic K_d values in the nanomolar range, much smaller (50–100-fold) than that of p51-p51). Our data also show discrepancies in the association/dissociation dynamics among the three RT isoforms. Our results further show that the apparent binding affinity of p51-p51 for its DNA substrate is to a great extent time-dependent when compared to that of p66-p66 and p66-p51, and is more likely determined by the dimer dissociation into its constituent monomers rather than the intrinsic binding affinity of dimeric RT.

The remainder of this chapter (except for section 3.6.3) has been published in The Journal of Physical Chemistry B (Marko, R. A.; Liu, H.-W.; Ablenas, C. J.; Ehteshami, M.; Götte, M.; Cosa, G. *J. Phys. Chem. B* **2013**, *117*, 4560-4567).

3.2 Introduction

During human immunodeficiency virus type 1 (HIV-1) replication, reverse transcriptase (RT) is responsible for an obligatory step involving the conversion of single-stranded genomic RNA into double-stranded proviral DNA. RT isolated from virions is a heterodimeric protein consisting of two polypeptide chains, a 66-

kDa subunit (p66) and a 51-kDa subunit (p51).[19, 20] The p66 subunit comprises a polymerase and a ribonuclease (RNase) H domain, with a connection domain linking the two catalytically active sites. The RNase H domain cleaves the RNA moiety of RNA/DNA replication intermediates, which is likewise essentially required for proviral DNA synthesis.[17] The p51 subunit, processed by proteolytic cleavage of p66 during viral maturation, corresponds to the N-terminus of the polymerase domain and lacks the RNase H domain. Although the two subunits are encoded by the same gene, distinct foldings result in an asymmetric dimeric structure.[23, 26] Consequently, the p51 subunit is catalytically incompetent; its role is limited to stabilizing the p66 subunit and it also provides contacts to the bound nucleic acid substrate. However, the nucleic acid binding channel is primarily formed by elements of the large p66 subunit.[17, 123]

Sluis-Cremer *et al.* have proposed two possible mechanisms for HIV-1 RT heterodimerization during viral maturation.[124] In the sequential model, RT initially exists as p66-p66 homodimers, after being processed from the Gag-Pol polyprotein precursor by the viral protease. Subsequent proteolytic cleavage of one subunit of the p66-p66 homodimer removes the C-terminal (p15) RNase H domain, yielding the p66-p51 heterodimer. Alternatively, in the concerted model, viral protease excises the Gag-Pol polyprotein precursor into separate p66 and p51 monomers, followed by rapid heterodimerization of p66 and p51. To test these models, Sluis-Cremer *et al.* expressed precursors containing the L234A mutation, known to impair p66-p66 dimerization.[28, 125] The formation of processed p66 accompanied by little p51 provides support for the sequential processing pathway. Here, the formation of p66-p51 occurs via a p66-p66 homodimer intermediate and yields practically equivalent amounts of p66 and p51 subunits in HIV-1 virions.[20, 126]

In addition to heterodimeric p66-p51, active homodimeric forms of HIV-1 RT p66-p66 and p51-p51 have also been characterized.[127-129] Biochemical studies have shown that p66-p66 homodimers possess polymerase and RNase H activity comparable to their heterodimeric counterparts.[20, 130, 131] In sharp contrast, several groups have reported that p51-p51 homodimers possess modest to negligible

DNA polymerase activity.[127, 130, 132-136] However, little information is available on the binding of these dimers to their nucleic acid substrates.

Single-molecule techniques are ideally positioned to unravel the biochemical mechanism and biological function of protein-nucleic acid interactions. Whereas static and dynamic protein-nucleic acid interactions have mostly been studied using single-molecule fluorescence resonance energy transfer (SM-FRET), this technique often necessitates site-specific fluorescent labeling of protein and/or nucleic acid species.[70, 94, 121, 137-140] For example, the growth of RecA filaments along single-stranded DNA (ssDNA) can be monitored solely by employing ssDNA labeled with Cy3 and Cy5 fluorescent dyes.[141] In contrast, in order to observe FRET changes induced by the association of heterodimeric HIV-1 RT with its nucleic acid substrates, both enzyme and substrate have been fluorescently labeled.[121] Introduction of protein labeling to homomultimer systems (e.g., RecA filaments or homodimeric RT) is, however, problematic due to the possibility that multiple fluorescently tagged subunits are complexed with singly-labeled DNA substrates, making it challenging to interpret SM-FRET signals.

Here we report a single-molecule protein-induced fluorescence enhancement (SM-PIFE) assay designed to directly visualize binding of RT to DNA-DNA primer-template with no protein-labeling requirement. In PIFE, protein binding to, e.g., double-stranded DNA leads to a decrease in the non-radiative decay pathway of a DNA-tethered fluorophore by reducing the weight of internal conversion in the rate of relaxation of the photoexcited dye. This reduction is concomitant with an increase in emission intensity.[96, 99] The PIFE assay has facilitated the understanding of protein-nucleic acid interactions in biological systems such as the conformational change of bacteriophage T7 DNA polymerase during DNA synthesis, the translocation of a DNA helicase, and the formation of RecA filaments along their DNA substrates.[96, 99]

Employing SM-PIFE, here we investigate both the binding affinity and kinetics of HIV-1 RTp66-p51 heterodimer versus RTp66-p66 or RTp51-p51 homodimers when exposed to their DNA-DNA substrates. We directly visualize the binding/unbinding process of RT on a DNA-DNA substrate by following RT-induced fluorescence fluctuations. The binding parameters derived from the SM-PIFE data reveal great similarities in the intrinsic binding affinities of RTp66-p51 and RTp66-p66 to the DNA substrate, with characteristic K_d values in the nanomolar range, much smaller (50-100-fold) than that measured for p51-p51. The binding parameters also show discrepancies in the association/dissociation kinetics among the three dimeric RT isoforms.

3.3 Results and Discussion

3.3.1 Binding of p66-p51 Heterodimers with Cy3-Primer-Template Substrates

We first investigated whether PIFE would be applicable to monitor the binding of HIV-1 RTp66-p51 to surface-immobilized Cy3-primer-template, hereafter noted as Cy3-dP-T (Figure 3.2).

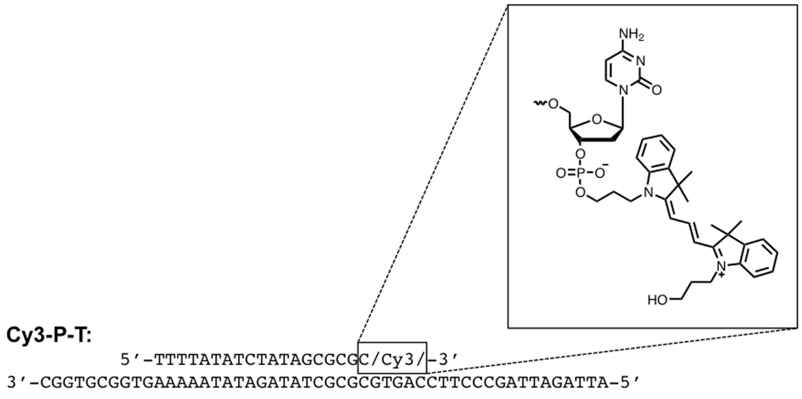

Figure 3.2 Sequence of Cy3-dP-T employed in our SM-PIFE measurements showing the structure of the Cy3 linkage.

To ensure saturation binding of RTp66-p51 to Cy3-dP-T, large enzyme concentrations, ca. 300 nM, were employed. A diagram of the assay used for SM-PIFE measurements is depicted in Figure 3.3.

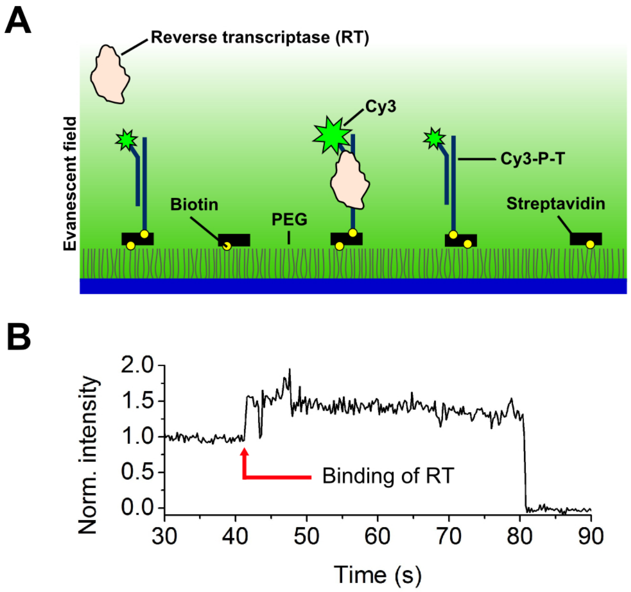

Figure 3.3 A cartoon illustrating the single-molecule protein-induced fluorescence enhancement (SM-PIFE) assay for detecting RT binding to a DNA-DNA substrate. (A) A total internal reflection fluorescence microscopy setup combined with surface-immobilized Cy3-labeled primer-template (Cy3-dP-T) was employed for our SM-PIFE measurements. (B) A representative normalized intensity-time trajectory of Cy3-dP-T acquired while 300 nM RTp66-p51 flowed into the imaging chamber shows the binding of RT to Cy3-dP-T and the induced fluorescence enhancement.

We employed a flow system allowing for the rapid delivery of RT into the chamber where the buffer solution was readily switched from no RT to 300 nM RTp66-p51 (see experimental section). During the buffer exchange, images were acquired over time wherefrom the fluorescence intensity of hundreds of individual Cy3-dP-T molecules were simultaneously recorded in order to extract their intensity-time trajectories. As shown in Figure 3.3B, we observed a ca. 30% fluorescence enhancement for individual Cy3 due to RT binding.

3.3.1.1 Fluorescence Enhancement of Cy3-dP-T Induced by RT Binding

For RT concentrations ≥ 50 nM, fluorescence enhancement occurred synchronously among the majority of molecules recorded, supporting the notion that the observed intensity increases are due to RT binding (Figure 3.4). The immobilized Cy3-dP-T was initially incubated in a buffer containing no RT. We initiated image acquisition at a repetition rate of 5 Hz immediately after we began flowing RT into the imaging chamber at a rate of 10 μL per min. Due to the dead volume of the flow tubing, up to 1 min elapsed before RT reached the imaging chamber and the binding of RT was observed. Each intensity-time trajectory thus contains an initial segment (recorded in the absence of protein), whose average intensity was utilized to normalize the overall trajectory. Figure 3.4 shows intensity-time trajectories normalized to the average intensity of the initial off-state segment. These trajectories were randomly chosen from one experiment at 300 nM RTp66-p51. The set of trajectories demonstrates that fluorescence enhancement occurs synchronously among all the intensity-time trajectories within an error of ~5 s, likely corresponding to the time required for RT to diffuse throughout the imaging chamber and to bind to Cy3-dP-T productively.

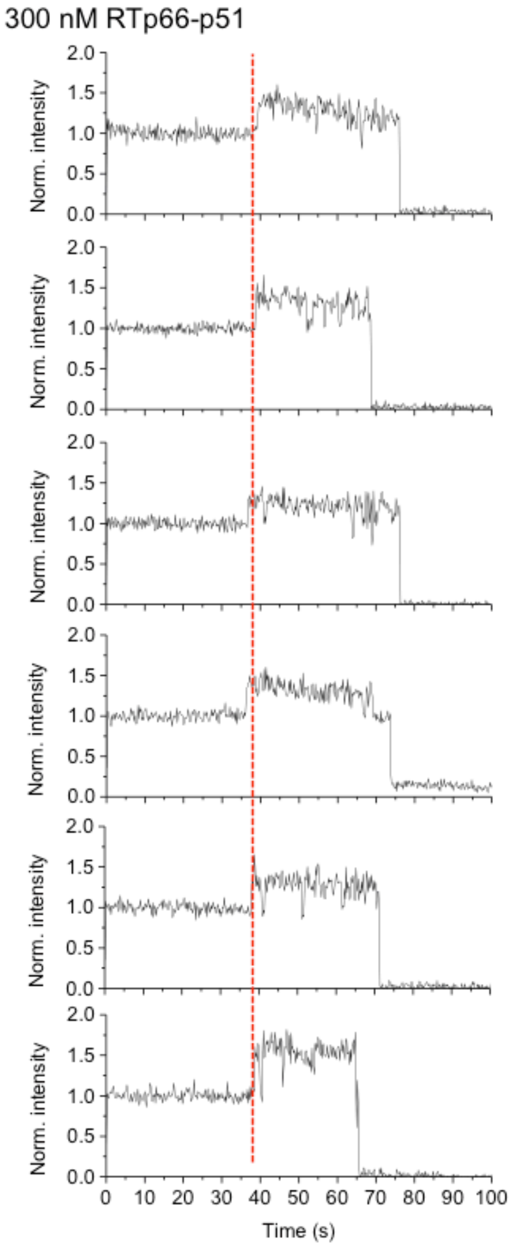

Figure 3.4 Normalized fluorescence intensity-time trajectories of Cy3-dP-T and 300 nM RTp66-p51. The red dashed line indicates the protein arrival and first intensity enhancement observed among the majority of molecules.

3.3.1.2 Steady Fluorescence of Cy3-dP-T in the Absence of RT

As a control, we monitored the fluorescence of Cy3-dP-T in the absence of RT, but under otherwise identical conditions (Figure 3.5). As shown in the representative normalized intensity-time trajectory of Cy3-dP-T (Figure 3.5A), Cy3-dP-T exhibited steady fluorescence intensity when RT was not present in the imaging chamber. Furthermore, the ensemble normalized intensity histogram constructed based on 300 normalized intensity-time trajectories of Cy3-dP-T molecules (Figure 3.5B) exhibited a single distribution centered at 1. We therefore assign the baseline intensity and the elevated intensity in Figure 3.3B to an RT-unbound state (off-state) and an RT-bound state (on-state), respectively.

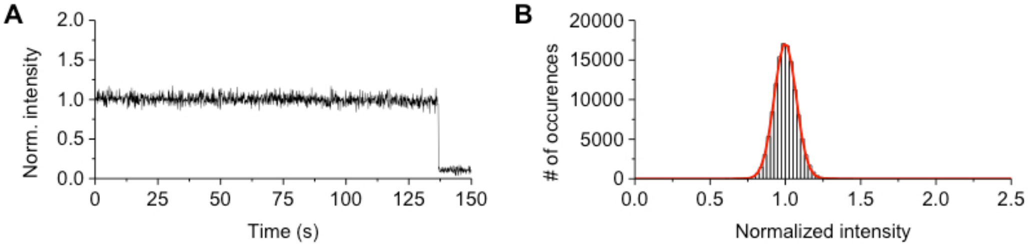

Figure 3.5 (A) A representative intensity-time trajectory of a Cy3-dP-T molecule and (B) an ensemble normalized intensity histogram of 300 Cy3-dP-T molecules in the absence of RT fitted with a single Gaussian distribution (red curve).

3.3.1.3 Formation of Higher-Order RT/Cy3-dP-T Complexes at High RTp66-p51 Concentrations

Since the nucleic acid binding cleft of RTp66-p51 can accommodate a nucleic acid duplex of 18-22 base-pairs between both active sites,[21] we expected to observe a 1:1 complex composed of RTp66-p51 and Cy3-dP-T (the duplex region in our single molecule substrate consists of 19 base pairs). However an electrophoretic mobility shift assay (EMSA) also revealed the formation of higher-order complexes for RTp66-p51 concentrations ≥ 300 nM and RTp66-p51/Cy3-dP-T molar ratios ≥ 4 (Figure 3.6). At concentrations > 300 nM of both RTp66-p51 and RTp66-p66, we observed an additional RT/Cy3-dP-T complex that migrated more slowly within the gel, suggesting the formation of higher order assemblies of RT/Cy3-dP-T.[48] It is plausible that the aggregated RT/Cy3-dP-T complex involves RT oligomers as previously observed in a chemical cross linking assay.[142] To avoid PIFE effects due to aggregated RT, the remaining single-molecule experiments were acquired with HIV-1 RT concentrations ≤ 50 nM.

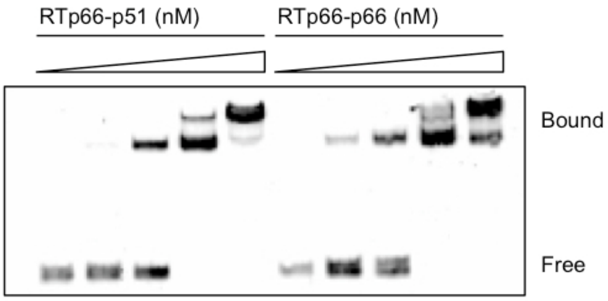

Figure 3.6 Binding of dimeric RT to Cy3-dP-T. EMSA experiments were performed on Cy3-dP-T in the presence of increasing concentrations of RTp66-p51 (left) and RTp66-p66 (right).

3.3.2 Affinity and Kinetics of RTp66-p51 at Different Concentrations

To gain insight into the affinity and kinetics of RTp66-p51 binding to Cy3-dP-T, we conducted SM-PIFE studies at different RTp66-p51 concentrations. Figure 3.7 shows two intensity-time trajectories acquired with 50 and 2 nM RT.

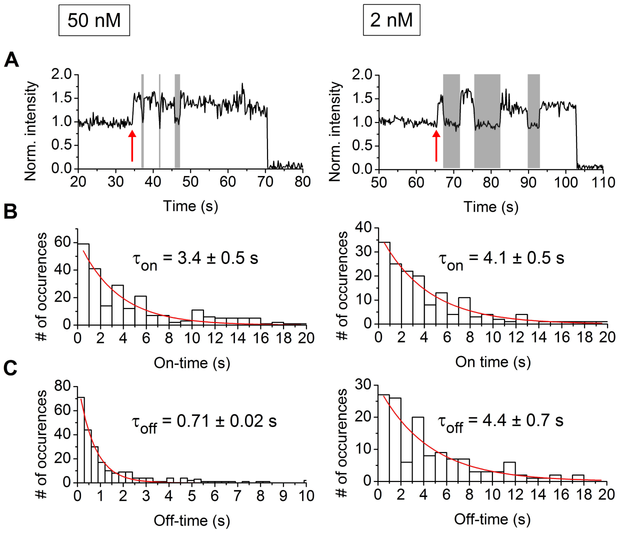

Figure 3.7 Monitoring RT binding/unbinding kinetics in the presence of 50 nM (left) and 2 nM RTp66-p51 (right). Panel A depicts single-molecule intensity-time trajectories of Cy3-dP-T at two RTp66-p51 concentrations. Each trajectory was normalized to its average intensity during the initial period when RT was not present. The red arrows correspond to the first RT binding event

and the shaded areas correspond to the off-states where Cy3-dP-T is unbound to RT. Panels B and C show on- and off-time histograms, respectively, constructed based on at least 60 single-molecule intensity-time trajectories. Also shown are the fittings (red) with a single exponential decay function.

After the introduction of RTp66-p51 into the imaging chamber, the intensity-time trajectories fluctuated between the elevated (on) and baseline (off) levels, indicating that P-T switched between the RT-bound and RT-unbound states, respectively. The duration of the on-state was independent of RTp66-p51 concentration (consistent with a protein dissociation rate which is first order in the concentration of RT/Cy3-dP-T complex). In contrast, the duration of the off-state was shorter with increasing RTp66-p51 concentration; see shaded areas in Figure 3.7A (this result is consistent with a second order process, first order in both [RT] and [Cy3-dP-T]). The association and dissociation dynamics follow the elementary reaction steps shown in Equation 3.1:

$$RT + Cy3\text{-}P\text{-}T \underset{k_d}{\overset{k_a}{\rightleftharpoons}} RT/Cy3\text{-}P\text{-}T \tag{3.1}$$

In order to gain a quantitative insight on the rate constants involved in the association (k_a) and dissociation (k_d) processes, we conducted further analysis of the intensity-time trajectories using a hidden Markov model based on a three-state system[119] (see section 3.6.1) wherefrom the off and on interval durations could be extracted.[143] We next constructed histograms of the durations of Cy3-dP-T in the on-states and off-states (hereafter referred to as on-time and off-time, respectively) based on at least 60 molecules for each RTp66-p51 concentration. Single exponential fittings of the histograms yielded the dissociation rate constant, k_d ($1/\tau_{on}$) and the pseudo-first-order association rate constant ($1/\tau_{off}$), where [RTp66-p51] >> [Cy3-dP-T], under our experimental conditions (see Figures 3.7B-C and Table 3.1, also see section 3.6.1). We obtained values of $k_d = 0.28 \pm 0.03$ s^{-1} and of $k_a = 1.1\times10^8$ $M^{-1}s^{-1}$, the latter being calculated from measurements conducted with 2 nM RT by dividing the pseudo-first-order rate constant obtained by [RTp66-p51]. The value of k_a we found is comparable to that of 2×10^8 $M^{-1}s^{-1}$ obtained from stopped-flow experiments reported by Kruøhfter et al.[144] We note that although the τ_{off} values derived from the SM-PIFE data at 25 and 50 nM RTp66-p51 were shorter than the τ_{off} obtained with 2 nM RTp66-p51, the values did not decrease in a concentration-dependent manner. The overestimated τ_{off} at higher RT concentration is due to the time resolution in our measurements, which is unable to resolve the shorter unbound events with off-times < 280 ms.

Hwang *et al.* have demonstrated that the extent of PIFE is markedly dependent on the distance between the fluorescent dye and the protein for separations below 4 nm[99] and may be intrinsically related to the steric hindrance the protein provides.[96, 100] RTp66-p51 may bind primer-template substrates in two distinct orientations with either the polymerase or RNase H domain positioned at the 3' terminus of the primer.[66, 145] As a result, one might expect each orientation to induce a different degree of fluorescence enhancement. Abbondanzieri et al. have demonstrated however that, for the DNA duplex employed in our studies, the polymerase domain of RTp66-p51 preferentially binds the 3' termini of primers ~90% of the time.[121] Based on the results by Abbondanzieri et al. we are predominantly probing interactions between the polymerase active site of RTp66-p51 and Cy3 on the Cy3-dP-T construct employed in our studies.

Control SM-PIFE studies performed on a Cy3-dP-T with a different Cy3 labeling site (attached to the 5' terminus of the primer rather than the 3' terminus) yielded a similar fluorescence enhancement upon binding of RTp66-p51 (Figure 3.8).

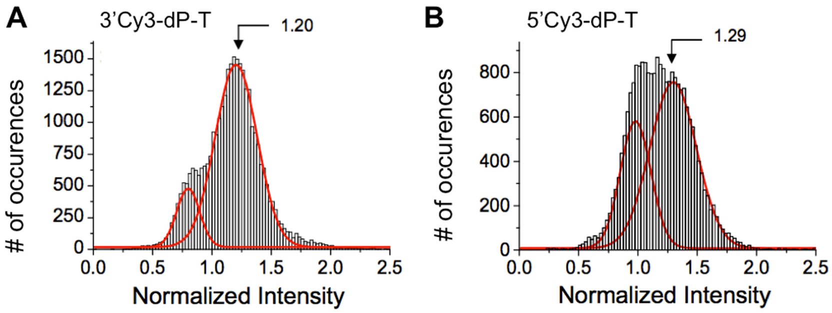

Figure 3.8 Normalized ensemble intensity histograms describing the fluorescence enhancement induced by two different domains of RT. Cy3-dP-T with the dye attached to (A) the 3' terminus and (B) the 5' terminus of the primer were employed to determine the extent of fluorescence enhancement induced by binding of the polymerase domain and the RNase H domain of RT to Cy3-dP-T.

The dominant binding mode in the latter case positions the RNase domain close to Cy3. Thus SM-PIFE may not enable us to distinguish polymerase vs RNase domain binding since they both gave rise to a similar enhancement. Albeit no difference may be appreciated between the binding of polymerase and RNase domains based on SM-PIFE, the two-state rather than multi-state intensity fluctuations observed in the Cy3-dP-T trajectories in the presence of RTp66-p51 (Figure 3.7A and 3.9G), together with results from Abbondanzieri et al. support the notion that we are probing a specific binding orientation of RT. Two-state intensity fluctuations for Cy3-dP-T are observed in the presence of all three RT dimers (see Figure 3.9 in the following section).

3.3.3 Binding Kinetics and Affinity for RT Hetero- and Homodimers

In order to determine whether or not significant differences in binding affinity to Cy3-dP-T exist between heterodimeric and homodimeric forms of RT, we next conducted SM-PIFE measurements with homodimeric constructs RTp66-p66 and RTp51-p51. The parameters related to the binding kinetics of the three RT dimers with Cy3-dP-T are summarized in Figure 3.9 and Table 3.1.

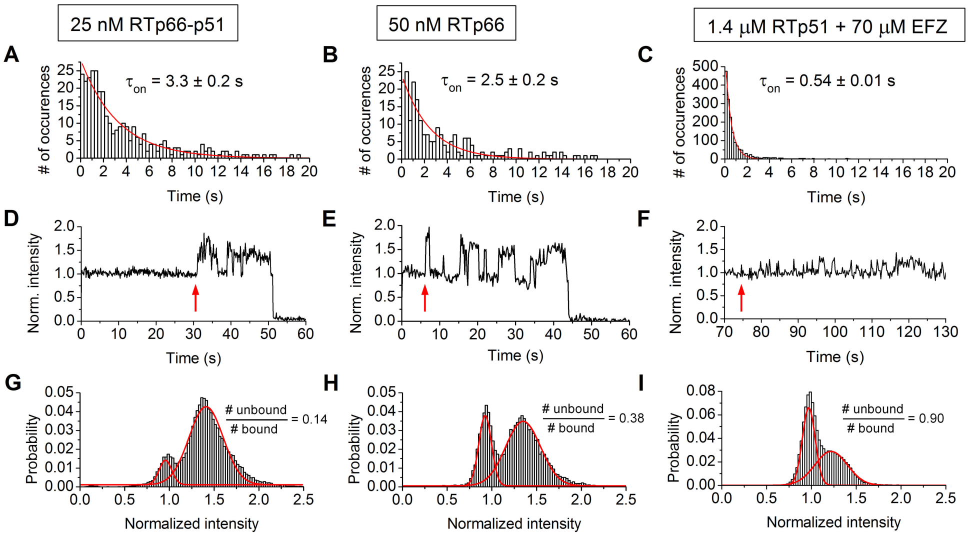

Figure 3.9 Binding of heterodimeric and homodimeric RT to Cy3-dP-T. (A-C) Ensemble on-time histograms for Cy3-dP-T in the presence of one of the 3 RT isoforms. Exponential fittings are also shown (red curves). (D-F) Examples of individual normalized intensity-time trajectories acquired for Cy3-dP-T in the presence of one of the 3 RT isoforms. Each trajectory was normalized to its initial average intensity when no protein was present in the imaging chamber. Red arrows indicate the onset of the first intensity enhancement event due to RT binding. (G-I) Ensemble histograms constructed with the individual normalized intensity-time trajectories of Cy3-dP-T acquired in the presence of one of the 3 RT isoforms. The ensemble histograms were fitted with a bimodal Gaussian distribution. The protein concentrations used in these experiments were 25 nM RTp66-p51 (left), 50 nM RTp66 (middle) and 1.4 μM RTp51 incubated with 70 μM efavirenz (EFZ) (right).

The average time required for RT to dissociate from the complexes, τ_{on} ($1/k_d$, derived from the on-time histograms in Figure 3.9A-C), is indicative of the stability of RT/Cy3-dP-T complexes. The k_d value we obtained for RTp51-p51/Cy3-dP-T (in the presence of efavirenz (EFZ), see below) was the largest among the three RT/Cy3-dP-T complexes (see Table 3.1). A closer look at the normalized intensity-time trajectories of individual Cy3-dP-T molecules in Figure 3D-F reveals faster dissociation dynamics for RTp51-p51. While the RT unbound/bound ratios of Cy3-dP-T for p66-p51 and p66-p66 were within the same order of magnitude (Figure 3.9G-H), binding of RTp51-p51 to Cy3-dP-T was scarce under the same monomer concentration (50 nM). As we increased the RTp51 monomer concentration to 1.4 μM, the majority of Cy3-dP-T remained in the unbound state, with an unbound/bound ratio of ~5; the intensity-time trajectories under this condition did not exhibit sufficient binding/unbinding events for obtaining reliable τ_{on} and τ_{off}. Only upon incubating RTp51 with 50-fold EFZ (a non-nucleoside RT inhibitor (NNRTI) known to enhance the dimer formation of RT), did the unbound/bound ratio drastically decrease to 0.9. The resulting ensemble normalized intensity histogram and on-time histogram (Figures 3.9C and F) indicated that binding of RTp51-p51 to Cy3-dP-T is highly dynamic, characterized by 5-7-fold faster dissociation rates when compared to those of RTp66-p51 and RTp66-p66. (see Table 3.1).

The analysis of the intensity-time trajectories not only provides dynamic information but also may yield information on the affinity of RT, either homo- or heterodimer, for its P-T substrate. The ensemble normalized intensity histograms may provide a quantitative value for the RT unbound/bound ratio at a given RT concentration. Next, it is possible to estimate the apparent dissociation constant of RT/Cy3-dP-T (K_d) according to Equation 3.2:

$$K_d = \frac{\#\, of\ events{,}unbound}{\#\, of\ events{,}bound}[RT] \tag{3.2}$$

Table 3.1 lists the K_d values for RTp66-p51 and the homodimers RTp66-p66 and RTp51-p51 obtained based on Equation 3.2. Values for RTp66-p66 and RTp51-p51 (in the presence of efavirenz) are shown to be 2.5- and 112-fold larger, respectively, than those acquired for RTp66-p51 (K_d = 3.9 nM).

Critical to the analysis of the above results and those presented in the ensuing section is the consideration of the dynamic equilibrium between the RT dimer and its constituent monomers. The documented equilibrium dissociation constant for the dimer-monomer equilibrium of RTp66-p51 ($K_{d,p66\text{-}p51}$) has been reported to be 310 nM.[146] Such a large value for the dissociation constant of the heterodimer would imply that only ~0.6% (or 13 pM) RTp66-p51 actually exists in its dimeric state under our experimental conditions (initial 2 nM RTp66-p51). In turn, such a low effective RT dimer concentration would require revising Equation 2 by inserting the real dimer concentration, leading to a K_d of ~8 pM for the RTp66-p51/Cy3-dP-T complex. We note however that in our SM-PIFE assay, care was taken to rapidly dilute RTp66-p51 from a concentrated stock solution (~100 μM). The dissociation kinetics of the RT heterodimer have been shown to be significantly slow, with a dimer dissociation half-life ($t_{1/2}$) of ~48 h at 5 °C,[67] much longer than the course of our SM-PIFE measurements (~5 min). Thus, under our conditions, all the RTp66-p51 should be fully dimerized, leading to a K_d of 3.9 nM. This K_d value, derived based on 100% heterodimeric RT, is consistent with K_d values of a few nanomolar previously reported by two independent groups using longer DNA-DNA primer-template constructs (see Table 3.1).[147, 148]

Likewise, to calculate the K_d of RTp51-51/Cy3-dP-T with EFZ, it is necessary to estimate the amount of dimeric species in the RTp51 solution. Several analytical methods have been applied to characterize the homodimerization of RTp51, with reported $K_{d,p51\text{-}p51}$ values ranging from 230 to 670 μM. Therefore, it is not surprising that only ~20% Cy3-dP-T were bound with RTp51-p51 (at 1.4 μM monomer concentration) in our SM-PIFE measurements.[134, 146] Venezia et al. further suggested that incubation of RTp51 with saturating levels of EFZ enhanced the p51 dimerization and decreased the $K_{d,p51\text{-}p51}$ to 0.37 μM.[146] Presumably, the monomer-dimer equilibrium should have established after 16 hrs of incubation with EFZ. Based on this $K_{d,p51\text{-}p51}$ value, we obtained 0.49 μM of dimeric RTp51-p51 and a K_d of 440 nM for RTp51-51/Cy3-dP-T with EFZ (Table 3.1).

Table 3.1 Dissociation/association rate constants and equilibrium dissociation constants for the 3 dimeric RTs and their nucleic acid substrates.

Constant	RTp66-p51	RTp66-p66	RTp51-p51	RTp51-p51 + EFZ[a]
k_a (s^{-1} M^{-1})	1.1×10^{8} [b] $(2 \pm 1) \times 10^{8}$ [c]	1.2×10^{7} [b]	N/A	1.6×10^{6} [b]
k_d (s^{-1})	0.28 ± 0.03 [d] 0.13 ± 0.01 [e] 0.2 [f]	0.40 [d]	N/A	1.9 [d]
K_d (nM)	3.9 ± 2.9 [d] 9.33 ± 0.09 [g] 4.7 ± 0.4 [f]	9.8 [d]	N/A	440 [d]

[a] Data acquired after 16 h of incubation of RTp51 with 50-fold EFZ (70 μM).

[b] Data were measured by SM-PIFE under pseudo-first-order conditions, where [RT] was significantly greater than [Cy3-dP-T]. The apparent association rate constants were estimated by dividing the pseudo-first-order association rate constants by the RT concentration. The value reported for RTp66-p51 was calculated using [RTp66-p51] = 2 nM (see main text). The values for RTp66-p66 and RTp51-p51 were calculated using [RTp66-p66] = 25 nM based on the assumption that the enzyme was fully dimerized and [RTp51-p51] = 0.49 μM based on the $K_{d,p51-p51}$ value in the presence of efavirenz (EFZ) reported by Venezia et al.[146]

[c] Data on a 19/53-mer DNA-RNA primer-template reported by KruØfter et al. at 6 mM Mg^{2+}.[144]

[d] Data were measured by SM-PIFE under pseudo-first-order conditions, where [RT] was significantly greater than [Cy3-dP-T]. The value reported for RTp66-p51 represents the average ± standard deviation of three independent trials conducted at 2, 25, and 50 nM. The values reported for RTp66-p66 and RTp51-p51 (with EFZ) were obtained at the same dimer concentrations as listed in entry b above, respectively. Experimental errors are expected to be similar to those reported for RTp66-p51.

[e] Data on a 22/44-mer DNA-DNA primer-template reported by Reardon at 5 mM Mg^{2+}.[149]

[f] Data on a 25/45-mer DNA-DNA primer-template reported by Kati et al. at 10 mM Mg^{2+}.[147]

[g] Data on a 25/30-mer DNA-DNA primer-template reported by DeStefano et al. at 6 mM Mg^{2+}.[148]

3.3.4 Binding Affinities of Homodimeric RTp66-p66 and RTp51-p51 as a Function of Time

Nuclear magnetic resonance (NMR) spectroscopy studies have revealed that RTp51-p51 possesses a nucleic acid binding cleft where one p51 subunit of the homodimer adopts an open and extended structure resembling the p66 subunit of RTp66-p51.[150] Based on such structural similarity between RTp66-p51 and RTp51-p51 (to our knowledge there is no structural data available for RTp66-

p66), the binding affinities of homodimeric and heterodimeric RT to nucleic acid substrates are arguably comparable to one another. It is thus surprising that the K_d value of RTp51-p51 in complexation with Cy3-dP-T, calculated from data in Figure 3.9I and listed in Table 3.1, is ~50-100-folder bigger than that of RTp66-p51 and RTp66-p66. It must be emphasized that all SM-PIFE measurements were conducted rapidly to minimize the effect of RT dimer dissociation into its constituent monomers. Thus RTp66-p66 or RTp51-p51 were diluted from a concentrated stock solution (~200 μM), and immediately flowed into imaging chambers containing the immobilized Cy3-dP-T, in an analogous manner to the aforementioned preparation of RTp66-p51 imaging studies (see also the experimental section for further details).

Restle *et al.* have reported $t_{1/2}$ values of 19 h for RTp66-p66 and 3 h for RTp51-p51 at 0 °C.[134] Therefore, it is highly plausible that the population of RT-bound states displayed in Figure 3.9I is due to residual RTp51-p51 dimers existing in the solution, where the RTp51 monomer/dimer equilibrium has not yet been established following rapid dilution of the concentrated stock solution. To test the above hypothesis, we incubated diluted RTp66 or RTp51 solutions, prepared from the 200 μM stock, in imaging buffer for increasing time intervals before introducing the protein solutions to Cy3-dP-T in the imaging chamber. Considering that the reported dimer dissociation constants for RTp66-p66 and RTp51-p51 ($K_{d,p66\text{-}p66}$ and $K_{d,p51\text{-}p51}$) are a few micromolar and hundreds of micromolar, respectively,[134, 146] we anticipated that the diluted RT homodimers would readily dissociate into monomers over time. Since monomeric p66 or p51 are incapable of binding to DNA primer-template substrates,[151] dimer dissociation of RT would shift the equilibrium of Cy3-dP-T from the bound towards unbound states.

Following an incubation period (up to 1 hr) of a diluted RTp51 (1.4 μM) at room temperature in imaging buffer prior to its inclusion in the imaging chamber, we observed a dramatic reduction in the binding affinity of RTp51-p51 for its DNA substrate, consistent with the dissociation of the homodimer into its constituent monomers within this short time window (Figures 3.10A and 3.10B).

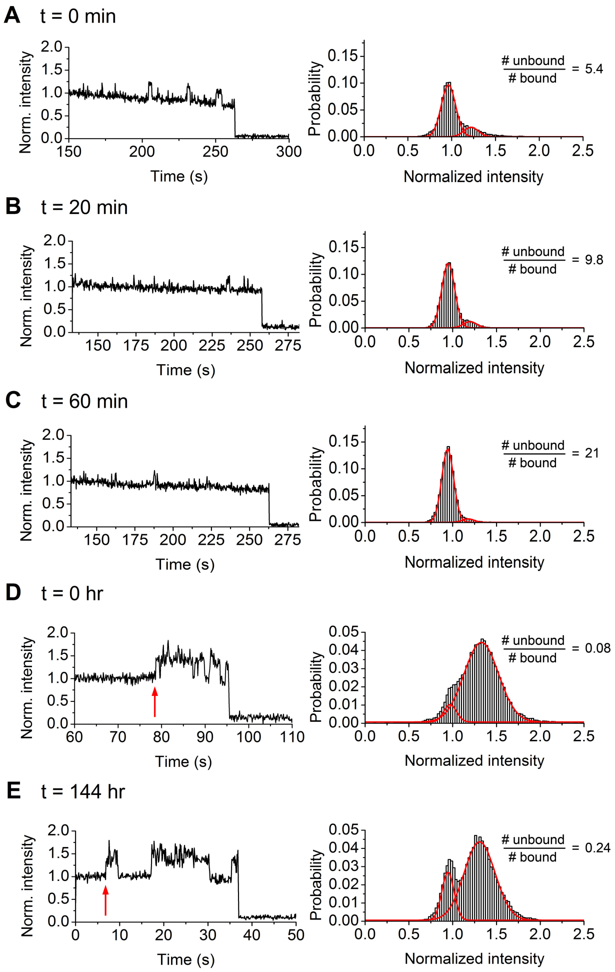

Figure 3.10 Intensity-time trajectories (left) and ensemble normalized intensity histograms fitted with Gaussian bimodal distributions (right) for Cy3-dP-T in the presence of 1.4 μM RTp51 and 50 nM RTp66, as a function of incubation time following dilution of the RT isoforms from a concentrated (ca. 200 μM) solution. (A-C) Intensity-time trajectories and ensemble normalized intensity histograms acquired in the presence of RTp51-p51 at (A) 0 min and (B) 20 min (C) 60 min incubation periods; trajectories are recorded 2 min after protein is flowed into chambers. (D-E) Intensity-time trajectories and ensemble normalized intensity histograms acquired in the presence of RTp66 at (D) 0 hr and (E) 144 hr incubation periods. Red arrows indicate the onset of the first intensity enhancement event due to RT binding; trajectories are recorded as soon as protein is flowed.

Attempts to perform this set of experiments with 50 nM RTp51 (whether it was diluted directly from the 200 μM stock solution or from the RTp51 (1.4 μM) with EFZ (70 μM) after 16 hour incubation) failed due to the scarce binding events caused by the rapid dissociation of RTp51. In contrast, an appreciable decrease in the RT-bound fraction of the Cy3-dP-T population appeared only upon incubating diluted RTp66-p66 (50 nM) for over 144 h at 4 °C (see Figures 3.10C and 3.10D), despite the fact that the RTp66-p66 concentration was in the nanomolar range, far below the dimerization $K_{\mathrm{d,p66\text{-}p66}}$ value of ~4 μM. Similar results to those of RTp66-p66 were obtained for RTp66-p51 (not shown).[146] The slow reduction in binding affinity for RTp66-p66 suggests that the rate of dissociation of RTp66-p66 into its constituent monomers is significantly smaller than the $t_{1/2}$ = 19 h reported by Restle *et al*.[134] In their work, Restle *et al*. induced the dissociation of dimers by diluting RT solutions with 20% acetonitrile. The effect of the organic solvent on promoting dimer dissociation (as suggested by Venezia *et al*.[67]) may account for the discrepancy in the dissociation rate values reported here and those listed in the previously published work.[134]

3.4 Conclusions

By using a flow system combined with the SM-PIFE approach, we have been able to investigate the binding of HIV-1 RTp66-p51 heterodimer and the two recombinant homodimers RTp66-p66 and RTp51-p51 to a DNA-DNA substrate. Binding kinetics of RT on nucleic acid substrates have typically been estimated indirectly via the amount of DNA synthesized or RNA degradation catalyzed by RT.[148, 152] In our SM-PIFE studies, insights into the parameters pertaining to the binding kinetics and affinities of the three HIV-1 RT dimers were obtained by directly visualizing the binding/unbinding process. The SM-PIFE results demonstrate that the intrinsic binding affinities of RTp66-p51 and RTp66-p66 with their DNA-DNA substrates are 50-100-fold larger than that of RTp51-p51. The bound RTp51-p51/Cy3-dP-T complex appears to be less stable (5-7 times bigger k_{d}) and highly dynamic, exhibiting faster binding/unbinding rates compared to RTp66-p51/Cy3-dP-T or RTp66-p66/Cy3-dP-T complexes.

It is well known that the stability and the enzymatic activity of the three RT dimers exhibit a hierarchical order with RTp66-p51 > RTp66-p66 > RTp51-p51.[20, 130, 131, 134] The SM-PIFE data presented here reveal that the initial binding of the three RT dimers to a DNA-DNA substrate also follows the same trend where heterodimeric RTp66-p51 has a slightly higher affinity than homodimeric RTp66-

p66 and much greater affinity than RTp51-p51. There exists a time-dependent decrease in the binding affinity of RT for its DNA substrate. While the binding events of either RTp66-p51 or RTp66-p66 with Cy3-dP-T exhibited an appreciable but small reduction after 6 days, those of RTp51-p51 diminished within ca. 20 min. These data strongly support the notion that the dissociation of the RTp66-p51 heterodimer and the RTp66-p66 homodimer into their constituent monomers is significantly slower compared to that of RTp51-p51. The binding of both hetero and homodimeric HIV-1 RT with their nucleic acid substrates is likely dominated by the protein dimerization process rather than by the intrinsic binding affinities of the functionally active dimeric forms.

3.5 Experimental

3.5.1 Protein Expression and Purification

Recombinant HIV-1 RTp66-p51 heterodimers were expressed and purified as previously described.[128] M15 competent cells were transformed with the pRT6Hpr plasmid, which allows for the expression of RTp66 in the absence of HIV-1 protease. The lack of protease allows for the expression of RTp66 without a subsequent cleavage to generate RTp51. The remainder of the purification method was identical to that of RTp66-p51 heterodimers. The RTp51 harboring a C-terminal glutathione S-transferase (GST) tag was generated separately by transforming BL21DE3 *E. coli* cells with the pGEXp51 vector. Induced RTp51 was purified through a GST column, with increasing concentrations of glutathione. Purified RT (p66-p51, p66, and p51) was then stored at high concentrations (100-300 μM) in RT buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl) with 50% glycerol at -20 °C or -80 °C.

3.5.2 Preparation of Primer/Template Strands

The DNA sequences used in this work (primer: 5'-TTTTATATCTATAGCGCGC-3'; template: 5'-ATTAGATTAGCCCTTCCAGTGCGCGCTATAGATATAAAAGTGGCGTGGC-3') were acquired from Integrated DNA Technologies (Coralville, IA). The strands were purified by HPLC. The 3'-terminus of the primer strand was conjugated to a Cy3 dye via phosphoramidite chemistry. In order to enable surface immobilization of the DNA duplex, the 3'-terminus of the template strand was attached to a biotin moiety via a C6 linker, which was in turn coupled to the 3'OH of the deoxyribose. The Cy3-biotinylated primer-template duplexes were prepared by mixing the primer strands and the template strands in annealing buffer (50 mM Tris-HCl and 40 mM NaCl, pH 8.0) at a ratio of 2:1. The mixture was then incubated in a thermal cycler (Eppendorf, Mississauga, Canada) at 95 °C for 2 min followed by a gradual cooling step of 2 °C/min to 25 °C. Duplex formation of the primer and template strands was confirmed using 10% native polyacrylamide gel electrophoresis.

3.5.3 Single-Molecule Sample Preparation and Imaging

In order to prevent non-specific adsorption of biomolecules onto the glass surface, coverslips were functionalized prior to use with a mixture of poly(ethylene glycol) succinimidyl valerate, MW 5000 (mPEG-SVA) and biotin-PEG-SVA (Laysan Bio, Inc.) at a ratio of 99/1 (w/w), following a previously described protocol.[153] The surface was incubated with 10 μL of 0.2 mg/mL streptavidin solution (Sigma Aldrich, St. Louis, MO) for 10 min followed by ~20 μL of 100 pM biotinylated Cy3-primer-template duplexes to achieve a good surface density (~300 fluorescent spots per 30 μm × 30 μm region). Unbound duplexes were then flushed out with imaging buffer (50 mM Tris-HCl pH 8.0, 40 mM NaCl, 6 mM $MgCl_2$). RT solutions ranging from 2-300 nM were prepared in the otherwise identical imaging buffer. Experiments with RTp51 were conducted at much higher concentration to promote homodimer formation. The non-nucleoside RT inhibitor (NNRTI) efavirenz (EFZ), known to enhance the dimer formation of RT, was also employed in studies with RTp51. RTp51 (1.46 μM) was initially incubated with efavirenz (72 μM) in imaging buffer for 16 h at 4 °C. Immediately prior to image acquisition, RTp51 and efavirenz were diluted to final concentrations of 1.4 μM and 70 μM, respectively, upon the addition of oxygen scavenging and triplet quenching components (see below). For monitoring the time-dependent dissociation of homodimeric RT, we diluted the protein from concentrated stock solutions (~200-300 μM) to 58 nM (RTp66) or 1.46 μM (RTp51) in imaging buffer and incubated the samples for various time intervals, at 4 °C. The protein was further diluted to a final concentration of 50 nM (RTp66) or 1.4 μM (RTp51) upon the addition of oxygen scavenging and triplet-state quenching components immediately before being introduced to the imaging chamber.

Imaging chambers (~8 μL) were constructed by pressing a polycarbonate film with an adhesive gasket (Grace Bio-Lab, Bend, OR) onto a PEG-coated coverslip. Two silicone connectors (Grace Bio-Lab, Bend, OR) were glued onto the predrilled holes of the film and served as inlet and outlet ports.[154] Prior to image acquisition, tubing was inserted into the inlet port, connecting the chamber to a syringe containing the RT solution that was then placed on a syringe pump. Image acquisition began simultaneously as the pump started to flow the RT solution at 10 μL per min into the chamber to replace the pre-existing imaging buffer. Due to the dead volume of the flow tubing, the RT-induced fluorescence enhancement of Cy3 on primer-template duplexes was observed ~1 min after initiating RT flow.[155] To delay the photobleaching of Cy3, all the solutions employed in single-molecule imaging experiments contained an oxygen scavenging component containing 0.1 mg/mL glucose oxidase (Sigma Aldrich, St. Louis, MO), 0.02 mg/mL catalase (Sigma Aldrich, St. Louis, MO), and 3% (w/w) β-D-glucose (Sigma Aldrich, St. Louis, MO) combined with a triplet-state quenching component of 143 mM β-mercaptoethanol (Sigma Aldrich, St. Louis, MO).

The fluorescence imaging of surface-anchored primer-template duplexes was conducted using an objective-type total internal reflection fluorescence (TIRF) microscope (Olympus IX-71 inverted microscope) in conjunction with a back-illuminated electron multiplying charge coupled device (EM-CCD) detector (Cascade II:512; Roper Scientific, Tucson, AZ). Excitation was carried out with a diode-pumped solid-state green laser (532-nm cw from Crystal Laser, Reno, NV) with a power output of ca. 10 mW measured from the objective in the wide field mode. The excitation was directed by a dichroic mirror (Z532rdc, Chroma Technology, Rockingham, VA) to an oil immersion microscope objective (60X PlanApo N; N.A. = 1.45; Olympus). The laser beam was focused onto the back focal plane of the objective using a TIRF illumination module (IX2-RFAEVA-2, Olympus) to yield an evanescent wave excitation at the glass-water interface. The fluorescence of the primer-template duplexes was collected by the objective and imaged onto the EM-CCD detector. A bandpass filter (HQ 590/70, Chroma Technology, Rockingham, VA) was placed in front of the EM-CCD detector to remove the scattered light. An additional magnification of 2X was introduced by a relay lens located between the camera and the microscope resulting in pixels of ca. 135 nm. Images consisted of ca. 70 μm x 70 μm regions. The camera was controlled using Image-Pro Plus 5.1 (Media Cybernetics), capturing 8-bit 512 x 512 pixel images with a conversion gain of 3, and multiplication gain of 4095. The exposure time per frame ranged from 100 to 280 ms and a total of 1000 to 1500 frames were acquired in each measurement. Fluorescence intensity-time trajectories of individual molecules were extracted from the videos using a self-written algorithm in IDL and Matlab.

3.5.4 Electrophoretic Mobility Shift Assays (EMSA)

For all electrophoretic mobility shift gel assay described herein, various concentrations of RTp66-p51 and RTp66-p66 (0, 50, 100, 300, 600 nM) and RTp51-p51 (0, 100, 300, 600, 1500 nM) were incubated with 75 nM Cy3-dP-T in a total volume of 10 μL containing 50 mM Tris-HCl pH 8.0, 40 mM NaCl, and 6 mM $MgCl_2$. After a 30-minute incubation period at room temperature, 2 μL of gel loading dye (50% w/v sucrose, 0.5% w/v bromophenol blue) were added and the RT/Cy3-dP-T complexes were resolved in 10% non-denaturing polyacrylamide gels. The bands were identified by Cy3 fluorescence.

3.6 Appendix

3.6.1 On/Off-time Analysis

On/off-time analysis of RT binding to Cy3-dP-T was conducted by fitting normalized intensity-time trajectories to a 3-state hidden Markov model (HMM) corresponding to RT-bound (on), unbound (off), and dark states.[143] For the on/off-time analysis, only partial intensity-time trajectories starting from the onset of the first binding event until photobleaching were taken into account. The last event before the molecule reached the dark state in each trajectory was discarded since the dwell time at that level is unknown. On/off-time histograms were each fitted with a single exponential decay function to yield the average on/off-time, τ. Figure 3.11 shows the ensemble on-time histogram acquired with 2 nM RTp66-p51. As shown by Equation 3.3, the average on-time, τ_{on}, provided by the fit (red curve in Figure 3.11) is inversely proportional to the dissociation rate constant, k_d:

$$k_d = \frac{1}{\tau_{on}} \tag{3.3}$$

To determine the second-order association rate constant, k_a, from the exponential fitting of the ensemble off-time histogram, an additional term, [RT], must be taken into account, as shown in Equation 3.4.

$$k_a = \frac{1}{\tau_{off}[RT]} \tag{3.4}$$

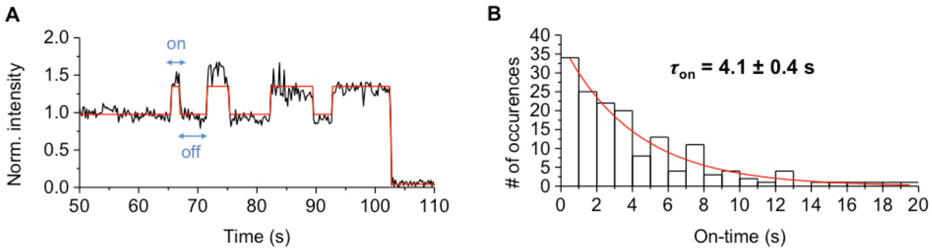

Figure 3.11 (A) A representative normalized intensity-time trajectory superimposed over a trajectory generated by hidden Markov fitting (red curve). (B) An ensemble on-time histogram with the exponential fitting curve superimposed (red). These data were acquired at 2 nM RTp66-p51.

3.6.2 Determination of Apparent Dissociation Constants (K_d) by Ensemble Normalized Intensity Histograms

Single-molecule intensity-time trajectories depicting binding/unbinding events as evidenced by two distinct states were selected for analysis shown in Figure 3.12. The trajectories were normalized to the average intensity of the initial off-state prior to the first binding event. The normalized intensity-time trajectories from at least 100 molecules were used to construct an ensemble normalized intensity histogram. The histogram was further divided by the total number of occurrences to yield the probability distribution (Figure 3.12B). Histograms were then fitted with a bimodal Gaussian function, with the first distribution centered at ~1 and the second distribution centered around 1.3 corresponding to the off- and on- states, respectively. Occasionally, the peak of the off-state distribution deviated from 1 due to the gradual decrease in the intensity-time trajectories, which might result from the background decrease of the sample or the microscope stage drift. The ratio of the areas for the off- and on-state distributions was calculated and multiplied by the concentration of RT used in the experiment to derive K_d, see also Equation 2 in the main document.

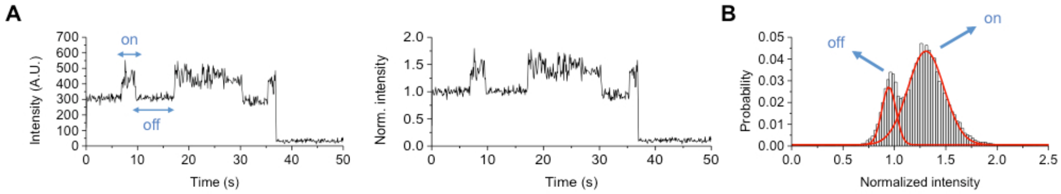

Figure 3.12 (A) A raw (left) and normalized (right) intensity-time trajectories of Cy3-dP-T fluctuating between an on- and an off-state. (B) A normalized intensity histogram, constructed based on at least 100 molecules, fitted with a bimodal Gaussian distribution (red curves). These data were acquired at 50 nM RTp66 after a 144 h incubation period at 4 °C.

3.6.3 Optimization RTp51 Binding to Cy3-dP-T

The homodimerization of RTp51 has been described in the literature as being "hypersensitive" to experimental conditions and numerous $K_{d,p51\text{-}p51}$ values ranging from 4-230 μM have been independently reported.[146, 150] As described in section 3.3.3, appreciable binding of RTp51-p51 to Cy3-dP-T could only be observed after first incubating 1.4 μM RTp51 with a 50-fold molar excess of EFZ overnight at 4 °C. At RTp51 concentrations below 1.4 μM and in the absence of EFZ, we could not observe any appreciable binding of RTp51-p51 to Cy3-dP-T at the single-molecule level. The lack of observable binding by RTp51 in the absence of EFZ was not consistent with previous SM-PIFE results that I obtained with 50 nM RTp51.

In September 2011, I conducted initial SM-PIFE experiments using 50 nM RTp51, without EFZ. RTp51 was diluted from a concentrated stock solution into the oxygen scavenger solution and immediately flowed into the imaging chamber containing Cy3-dP-T. The resulting intensity-time trajectories showed that RTp51 bound Cy3-dP-T and that the binding was highly dynamic, characterized by relatively fast dissociation rate compared to RTp66-p51 and RTp66-p66 (Figure 3.13 and Table 3.1).

The normalized intensity histogram for 50 nM RTp51 gave an unbound/bound ratio of 0.75 (Figure 3.13B). Using the unbound/bound ratio of 0.75, we originally calculated the K_d for RTp51-p51/Cy3-dP-T to be 19 nM, assuming 100% dimer formation, i.e., [RTp51-p51] = 25 nM when [RTp51] = 50 nM. Since the true proportion of RTp51 monomer to RTp51-51 dimer was unknown, we concluded that the K_d value of 19 nM was an upper limit for RTp51-p51 binding to and dissociating from Cy3-dP-T, at the salt concentration used in our experiments. We utilized hidden Markov modeling to analyze many intensity-trajectories to obtain dwell-time distributions of the RTp51-p51-bound (on) and unbound (off) states (Figure 3.13C & D). We obtained k_d and k_a values of 1.4 s^{-1} and 3.1×10^{-7} s^{-1} M^{-1}, respectively, assuming 100% RTp51-p51 dimerization (i.e., [RTp51-p51] = 25 nM).

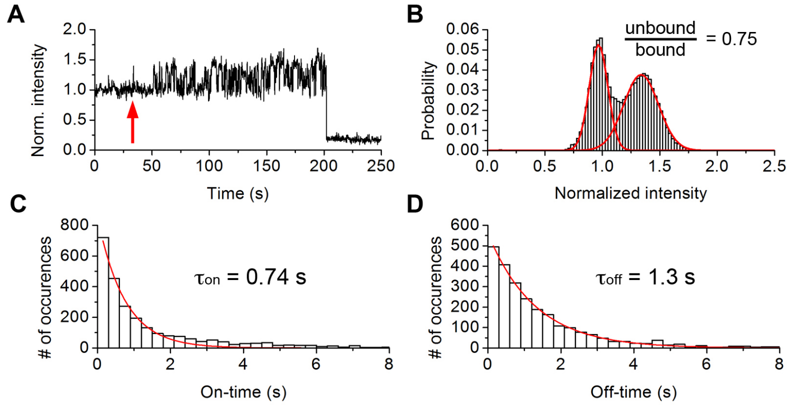

Figure 3.13 Binding of 50 nM RTp51 to Cy3-dP-T. (A) Representative intensity-time trajectory showing intensity fluctuations between two states. Red arrow indicates protein arrival in the imaging chamber. (B) Ensemble histogram constructed with the individual normalized intensity-time trajectories of Cy3-dP-T acquired in the presence of RTp51. The ensemble histogram was fitted with a bimodal Gaussian distribution. (C, D) Ensemble on- and off-time histograms for Cy3-dP-T. Exponential fittings are also shown (red curves). These results could not be reproduced.

Additionally, we observed a time-dependent decrease in the amount of binding of RTp51-p51 to Cy3-dP-T (Figure 3.14). The data shown in Figure 3.14 were acquired after 10 and 20 min incubation periods of RTp51 in the imaging chamber, at room temperature. Since literature values for $K_{d,p51\text{-}p51}$ in the hundreds of micromolar have been reported,[146] we hypothesized that the reduction in the RTp51-p51 binding affinity was due to the dissociation of the homodimer over time following the rapid dilution of RTp51-p51 from a concentrated stock solution.

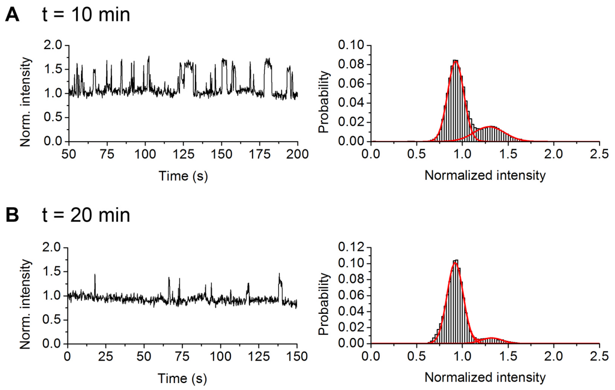

Figure 3.14 Intensity-time trajectories (left) and ensemble normalized intensity histograms fitted with Gaussian bimodal distributions (right) for Cy3-dP-T in the presence of 50 nM RTp51. (A-B)

Intensity-time trajectories and ensemble normalized intensity histograms acquired in the presence of RTp51-p51 at (A) 10 min and (B) 20 min incubation periods. These results could not be reproduced.

Despite these observations and resulting analyses, attempts to reproduce the results shown in Figures 3.13 and 3.14 were unsuccessful. Subsequent SM-PIFE experiments with RTp51 yielded intensity-time trajectories displaying no fluorescence enhancement. We first suspected that the original RTp51-p51 stock had lost its binding capability after prolonged storage at -20 °C, therefore we expressed and purified new RTp51 and evaluated its binding to Cy3-dP-T using SM-PIFE within 1 day of purification. After testing newly purified RTp51 up to concentrations of 10 μM, we observed no RTp51 binding using SM-PIFE (Figure 3.15).

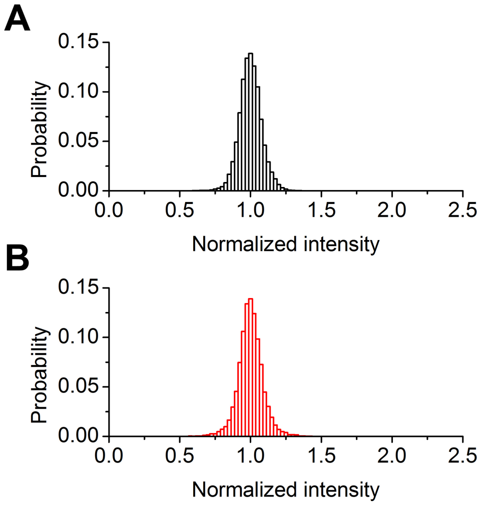

Figure 3.15 Ensemble normalized intensity histograms for Cy3-dP-T in the absence of RT (A) and in the presence of 10 μM RTp51 (B).

Since all purified RT stock solutions are stored in RT buffer containing 50% v/v glycerol (see experimental section 3.5.1), concentrations of RTp51 higher than 10 μM could not be tested as the high viscosity of concentrated RT solutions are not compatible with our flow pump system (section 3.5.3).

In order to rule out the possibility that RTp51 was binding but simply not inducing fluorescence enhancement, we conducted an ensemble EMSA (Figure

3.16). For two different batches of RTp51 at concentrations up to 30 μM, we observed no binding. The results from the EMSA are consistent with both RTp51 existing predominantly in its monomeric form, since monomeric RT does not bind to double-stranded nucleic acid substrates.[151] It is also possible that both batches of RTp51 were inactive due to improper protein folding or denaturation.

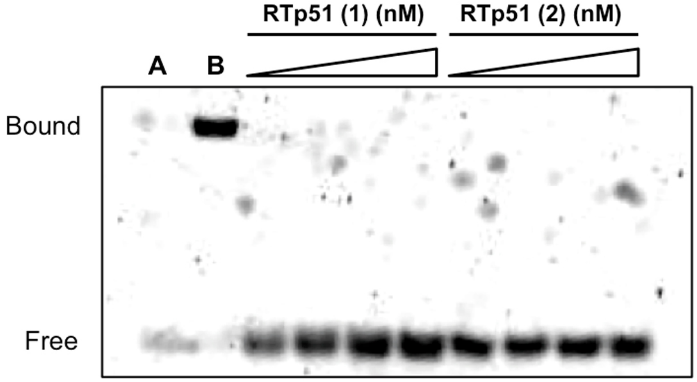

Figure 3.16 RTp51 does not bind to Cy3-dP-T. EMSA experiments were performed on Cy3-dP-T in the presence of increasing concentrations of RTp51-p51. Two independent batches of RTp51 (1, 2) were tested. (A) Migration of Cy3-dP-T in the absence of RT. (B) Binding of 100 nM RTp66-p51 to Cy3-dP-T (positive control).

We hypothesized that the dissociation kinetics of the RTp51-p51 homodimer into its constituent monomers were much faster than we originally observed in Figure 3.14. Since it typically takes ~5 min to prepare the RTp51 sample (with the oxygen scavenging components; see section 3.5.3), assemble the flow system on the microscope, and flow the solution into the imaging chamber, it is conceivable that the RTp51-p51 homodimer could dissociate into monomers before image acquisition.

Upon expressing and purifying a new RTp51 batch in December 2012, we observed reproducible, albeit scarce, binding to Cy3-dP-T at [RTp51] = 1.4 μM which decreased dramatically over the course of 1 h (Figure 3.10). To determine if we could shift the RTp51 monomer/dimer equilibrium towards the dimeric species, we pre-incubated RTp51 with EFZ, a NNRTI known to promote the dimerization of all forms of RT. Upon overnight incubation of RTp51 with EFZ, we observed a marked increase in the degree of binding to Cy3-dP-T (Figure 3.9F & I).

It is possible that even slight differences in experimental conditions (e.g., salt concentration, pH, temperature) could have contributed to our one-time observation of high RTp51 binding affinity shown in Figure 3.13, although the precise reason is still unclear. Indeed, discrepancies in literature $K_{\mathrm{d,p51\text{-}p51}}$ values have been attributed to a "hypersensitivity" of RTp51 dimerization on experimental conditions.[150] Using small-angle X-ray scattering (SAXS), Zheng et al. have reported a $K_{\mathrm{d,p51\text{-}p51}}$ value of ~4 μM at 25 °C with 200 mM KCl and 4 mM Mg^{2+}. In contrast, Venezia et al. used sedimentation equilibrium analysis to obtain a $K_{\mathrm{d,p51\text{-}p51}}$ value of 230 μM in the absence of Mg^{2+} at substantially lower temperatures (5 °C) and 25 mM NaCl.[146] Additionally, differences in pH have been shown to strongly influence the dimerization of RTp66-p66, thus it is reasonable to hypothesize that pH may have a similar effect on RTp51 dimerization.[156] It would be necessary to conduct an in depth characterization of RTp51 binding to Cy3-dP-T over a range of salt concentrations and pH values in order to determine the sensitivity of RTp51 dimer formation to experimental conditions for our SM-PIFE assay.

4. Extending SM-PIFE to HIV-1 RT Enzymatic Activity and Conformational Dynamics

4.1 SM-PIFE and RT Dynamics: Towards Characterizing the Interplay Between Polymerase and RNase H Activity

HIV-1 RT is a highly dynamic enzyme capable of undergoing a series of “molecular acrobatics” to coordinate and catalyze the multiple steps of reverse transcription.[157] In our laboratory it is of great interest to develop single-molecule methods to probe the dynamics of RT such that we can begin to examine the molecular mechanisms of how small-molecule inhibitors and drug resistance-conferring mutations impart their effects.

Zhuang and coworkers have designed an elegant SM-FRET assay capable of probing the binding and orientational dynamics of RT, however, a method based on SM-PIFE offers numerous advantages, the most prominent being that SM-PIFE has no protein-labeling requirement. RT contains 4 Cys residues, with 2 in each subunit (Cys38 and Cys280), therefore site-specific fluorescent labeling of RT requires site-directed mutagenesis of 3 of the 4 Cys residues. The C280S double mutation is widely used to fluorescently label RT via Cys38 of the p66 subunit, since it does not significantly alter RT polymerase or RNase H activity.[158] However, despite the widespread use of these mutations, one study has shown that RT with the C280S/C38S mutations in both RT subunits possess impaired DNA polymerase activity and enhanced RNase H activity, relative to the WT enzyme.[158] Nevertheless, the use RT variants containing Cys mutations has proved useful in providing insight into RT function.[67, 121, 122, 146, 159, 160]

If one wishes to study naturally occurring mutations, e.g., drug resistance mutations, the ability to compare the mutant to a purely wild-type background is advantageous. By not having to fluorescently label RT, SM-PIFE could prove to be an ideal method to investigate dynamic RT structure-function relationships. The following sections will describe ongoing work and future directions related to exploiting SM-PIFE to probe other aspects of RT structure/function, such as RNase H activity, binding orientation, and polymerase domain dynamics.

4.2 SM-PIFE Studies of RT RNase H Activity

An ongoing project initiated by a post-doctoral fellow in our group, Dr. Hsiao-Wei Liu, has been the utilization of SM-PIFE to monitor the RNase H activity of HIV-1 RT. In the following section I will summarize and place the results of Dr. Liu into context and describe my ongoing contribution as well as future directions for this project.

As described in section 1.1.3.2, RT catalyzes the endolytic hydrolysis of RNA of RNA/DNA hybrids. The RNase H function of RT is an absolutely essential component of the HIV-1 lifecycle and is thus an attractive target for the development of novel anti-HIV therapeutics. Over half of the 23 FDA-approved anti-HIV drugs target RT, all of which interfere with RT's DNA polymerase activity (section 1.1.4). Although the rate of discovery of potent RNase H inhibitors has increased in recent years, there are currently no RNase H inhibitors approved for clinical use.[32]

A major reason for the lack of clinically approved inhibitors targeting RT RNase H function has been due to the scarcity of high-throughput screening methods to measure RT RNase H activity. Widely used gel-based RNase H assays are time consuming and not amenable to high-throughput screening of compound libraries. Parniak et al. addressed this issue in 2003 by developing an automatic well-plate format FRET assay that employed an RNA/DNA substrate composed of a 3' fluorescein-labeled RNA annealed to a 5' dabcyl-labeled DNA. The principle of the assay relies on the quenching of fluorescein by dabcyl in the intact RNA/DNA substrate (Figure 4.1). The sequence was designed such that upon RNase H-catalyzed cleavage, the 3' fluorescein fragment would diffuse away from the dabcyl quencher, immediately generating a fluorescence signal.

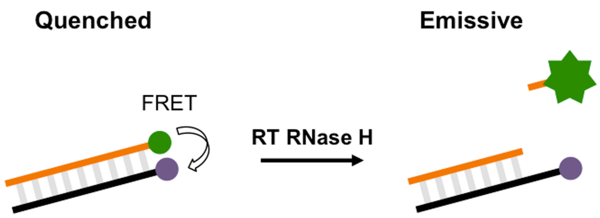

Figure 4.1 Schematic of the FRET-based assay described by Parniak et al. RT RNase H-catalyzed cleavage results in the diffusion of fluorescein (green) away from its quencher (dabcyl; purple). Upon cleavage, fluorescein becomes emissive.

In 2006, Nakayama et al. introduced a similar well-plate format RNase H assay that used fluorescence polarization as a read-out instead of FRET.[69] The main advantage of using polarization over FRET was that only one strand of the RNA/DNA substrate required fluorescent labeling, thus minimizing the amount of perturbation to the system. Indeed, Wahba et al. have demonstrated that the fluorescein/dabcyl labeling scheme conceived by Parniak et al. lowers the catalytic efficiency of RT RNase H.[161]

Here we describe an extension of our SM-PIFE assay to monitor the RNase H activity of HIV-1 RT. Compared to conventional ensemble fluorescence assays, our SM-PIFE approach has the advantage of providing rich mechanistic insight into RNase H function while possessing the potential for improved high-throughput capabilities.

A real-time single-molecule fluorescence approach such as SM-PIFE allows one to not only monitor RNase H-catalyzed reaction kinetics, but also to examine reaction dynamics, such as the precise motions (e.g., conformational rearrangements, orientational dynamics) of RT as it engages and positions itself to cleave an RNA strand. Ideally, the single-molecule approach may be used to 1) evaluate the inhibitory capacity of a compound and 2) understand how, mechanistically, it achieves its effect. Ultimately, the integration of single-molecule fluorescence RNase H assay with a high-throughput screening platform will help guide the rational design of novel therapeutics against RT RNase H function. For example, the combination of single-molecule fluorescence with an automated microfluidics-based measurement system should enable rapid, precise, and high-throughput evaluation of enzyme activity and/or conformational dynamics all while minimizing the amount of sample consumption. Such a platform will require the development of rapid analysis software, as well as robust micro- or nanofluidic devices. Indeed, recent strides have been made in the development of high-throughput single-molecule optofluidic methodologies to examine the activity of the *E. coli* RNA polymerase under a plethora of experimental conditions.[162]

A schematic depicting the principle of our assay is shown in Figure 4.2. As described in the previous chapter, SM-PIFE can be used to follow individual binding events of RT to a Cy3-labeled DNA/DNA substrate. When the substrate is an RNA/DNA hybrid, an additional process can take place where RT RNase H activity cleaves the RNA strand. Ideally, the substrate would be designed such that RNase H-catalyzed RNA cleavage results in the immediate disappearance of the Cy3-signal as the small (a few bases) cleaved-fragment rapidly diffuses out of the TIRF excitation volume, thus providing a direct readout of enzymatic activity.

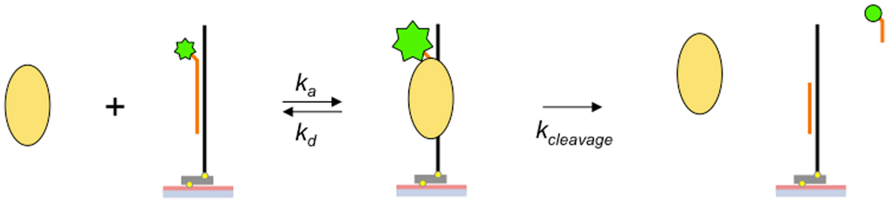

Figure 4.2 Cartoon depicting our SM-PIFE assay for RT RNase H activity. The onset of RT binding to a surface-immobilized Cy3-labeled RNA/DNA duplex is observed as an increase in fluorescence intensity of the substrate. RNase H-catalyzed cleavage results in loss of the Cy3 signal as the Cy3-fragment rapidly diffuses out of the excitation volume.

The substrates used in our SM-PIFE assay were designed such that RNase H cleavage occurs close to the Cy3 fluorophore, ensuring the immediate dissociation of the cleaved fragment. The sequences of our substrates are depicted in Figure 4.3. RNA primer sequences labeled at the 3' end (3'Cy3-rP) are based on either random sequence A or the unique PPT sequence. The 3'Cy3-rP sequences are then annealed to a 50 nucleotide DNA template (T). The resulting duplexes are hereafter referred to as 3'Cy3-rP-T (sequence A) and 3'Cy3-rP-T PPT (polypurine tract). The double-stranded regions are 19 bp (3'Cy3-rP-T) and 20 bp (3'Cy3-rP-T PPT) and correspond to the size of RT's nucleic acid binding cleft.[121] In the case of 3'Cy3-rP-T PPT, the RNA primer was extended to 20 nucleotides with an additional adenosine ribonucleotide to mitigate the potential fluorescence quenching effects of the proximal guanosine bases.

3'Cy3-rP-T

```
                     5'-UUU UAU AUC UAU AGC GCG C/3Cy3Sp/-3'
3'-/3Bio/CG GTG CGG TGA AAA ATA TAG ATA TCG CGC GTG ACC TTC CCG ATT AGA TTA-5'
```

3'Cy3-rP-T PPT

```
                     5'-UUU UAA AAG AAA AGG GGG GA (C6-NH)(Cy3)-3'
3'-/3Bio/CG GTG CGG TGA AAA ATT TTC TTT TCC CCC CTG ACC TTC CCG ATT AGA TTA-5'
```

Figure 4.3 RNA/DNA substrates used in our studies. Substrates labeled as 3'Cy3 correspond to labeling at the 3' end of the primer strand. RNA primer sequences are based on the previously reported sequence A and the polypurine tract (PPT).[121] RNA and DNA strands are shown in orange and black, respectively.

Based on previous SM-FRET results,[121] we predicted that RT should adopt the RNase H-competent binding orientation on our 3'Cy3-rP-T construct (Figure 4.4). The double stranded region of the substrates depicted in Figure 4.3 correspond to 19-20 base pairs, approximately equal to the size of the RT nucleic acid binding cleft.[22] Thus, the RT RNase H domain should be positioned close to the Cy3-labeled 3' terminus of the RNA strands. Since RT RNase H produces cuts ~18 base pairs upstream of the polymerase active site,[32] we hypothesized that the initial RNase H-catalyzed cleavage of 3'Cy3-rP-T should result in small fragments (~1-3 bases).

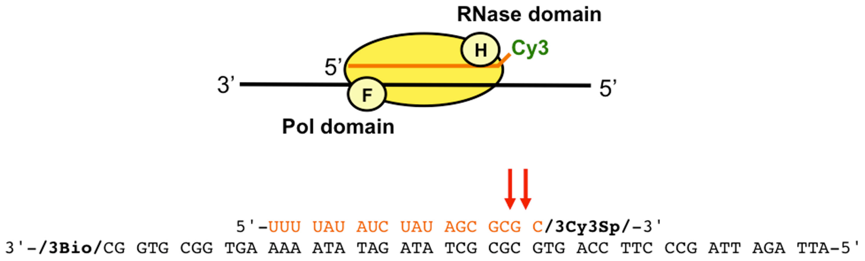

Figure 4.4 Predicted positioning of RT on a Cy3-labeled RNA/DNA duplex (3'Cy3-rP-T). RT should adopt an orientation where the RNase H domain is positioned close to the 3' end of the RNA primer.[121] Thus, we hypothesized that RNase H-catalyzed cleavage should occur close to the 3' end of the RNA primer (red arrows).

4.2.1 RT RNase H Cleaves RNA Close to the 3' End of 3'Cy3 RNA/DNA Substrates

To validate our substrate design and to identify the major RNase H cleavage products, we conducted a standard polyacrylamide gel electrophoresis-based RNase H activity assay[161] using our 3'Cy3-rP-T and 3'Cy3-rP-T PPT substrates, under a variety of conditions (Figure 4.5). As a control, we also tested a substrate of identical sequence to 3'Cy3-rP-T where the Cy3 fluorophore was attached to the 5' end of the RNA strand (i.e., 5'Cy3-rP-T). Cleavage of 5'Cy3-rP-T should result in larger Cy3-labeled fragments compared to 3'Cy3-rP-T. Both 3'Cy3-rP-T and 5'Cy3-rP-T should give rise to the same cleavage patterns, however the labeled fragments may be different in each case. We employed both WT RTp66-p51 and the N348I RTp66-p51 double mutant, the latter of which has been shown to exhibit reduced RNase H activity.[63] Additionally, we tested the effects of complementary dNTPs, NVP, and foscarnet (fos), all of which have been shown to modulate RNase H activity.[43, 163]

4.2.1.1 Gel Assay of RT RNase H Activity: Experimental

Polyacrylamide Gel Electrophoresis (PAGE) RNase H Assay

RTp66-p51 (WT or N348I double-mutant; 100 nM) was pre-incubated with or without dNTP (1 mM), foscarnet (Fos; 0.1 mM), or NVP (0.1 mM) in reaction buffer (50 mM Tris-HCl, 40 mM NaCl, 6 mM Mg^{2+}) for 10 min at 37 ºC. Following pre-incubation, reactions were initiated by the addition of RNA/DNA primer/template substrate to a final concentration of 50 nM. Reactions were allowed to proceed for 30 min at 37 ºC and then quenched by the addition of an equal volume of stop buffer (98% deionized formamide, 250 mM EDTA) followed by heat inactivation (5 min at 95 ºC). Cleavage products were resolved in a 12% denaturing polyacrylamide gel containing 7 M urea and visualized by fluorescence.

Preparation of Primer/Template Strands

The Cy3-biotinylated RNA/DNA primer-template duplexes were prepared by mixing the primer and template strands in annealing buffer (50 mM Tris-HCl and 40 mM NaCl, pH 8.0) at a ratio of 2:1. The mixture was incubated in a thermal cycler (Eppendorf, Mississauga, Canada) at 95 °C for 2 min followed by a gradual cooling step of 2 °C/min to 25 °C. Duplex formation of the primer and template strands was confirmed using 10% native polyacrylamide gel electrophoresis.

4.2.1.2 Gel Assay of RT RNase H Activity: Results

Both WT and N348I RTp66-p51 catalyzed the cleavage of our three fluorescently labeled RNA/DNA constructs, as indicated by the intensities of the major fluorescent products (red boxes, Figure 4.5). In the absence of protein (lanes 1-3, Figure 4.5A), multiple bands are observed for both 3'Cy3-rP-T and 5'Cy3-rP-T substrates, which is indicative of non-specific RNA degradation. It is possible that the specific sequences we have employed are sensitive to the conditions used to terminate our reactions (see section 4.2.1.1).

Experimental determination of the major fluorescent product sizes is difficult due to the lack of an appropriate ladder for short oligos. However, based on the positions of the major fluorescent cleavage products relative to the position of the full-length primer and the observed non-specific RNA hydrolysis pattern, we can obtain a rough estimate of the product sizes. Assuming that the size difference between sequential bands in each lane is 1 base, the major fluorescent products for 3'Cy3-rP-T and 3'Cy3-rP-T PPT are likely ~1-3 bases in size. In contrast, the major fluorescent product observed for 5'Cy3-rP-T is much larger (~14-15 bases).

Interestingly, the unique 3'Cy3-rP-T PPT sequence showed only one hydrolysis product (red box, Figure 4.5). Furthermore, unlike the 3'Cy3-rP-T and 5'Cy3-rP-T sequences, a major apparent high-molecular weight species was observed for 3'Cy3-rP-T PPT (Figure 4.5B). The slow migration of the 3'Cy3-rP PPT primer sequence suggests that this sequence can adopt significant secondary structure or form aggregates.

Under the conditions employed in our gel assay (see section 4.2.1.1), we did not observe drastic modulation of RT RNase H activity in the presence of any of the three compounds (dNTP, NVP, fos). Further optimization of experimental conditions is required. We note that the substrates used in Figure 4.5 were annealed at a primer:template ratio of 2:1. The 2:1 annealing ratio is typically used for single-molecule experiments, where excess Cy3-labeled primer is washed away and does not interfere with experimental measurements. In the case of the gel shown in Figure 4.5, the presence of excess primer decreases the sensitivity of the assay, making it difficult to assess 1) whether or not the reaction has gone to completion and 2) the effects of various compounds on RNase H activity. We originally used the substrates prepared at a 2:1 primer:template annealing ratio instead of preparing new substrates annealed at a 1:1 ratio due to lack of available material.

We have since repeated the assay in Figure 4.5 using substrates annealed at a 1:1 primer:template ratio and shown that under the aforementioned experimental conditions, the RNase H catalyzed reaction goes essentially to completion (not shown). It will be necessary to conduct time-course experiments to determine the appropriate reaction incubation time in order to better evaluate the effects of inhibitors, nucleic acid sequence, and RT amino acid mutations on RNase H activity.

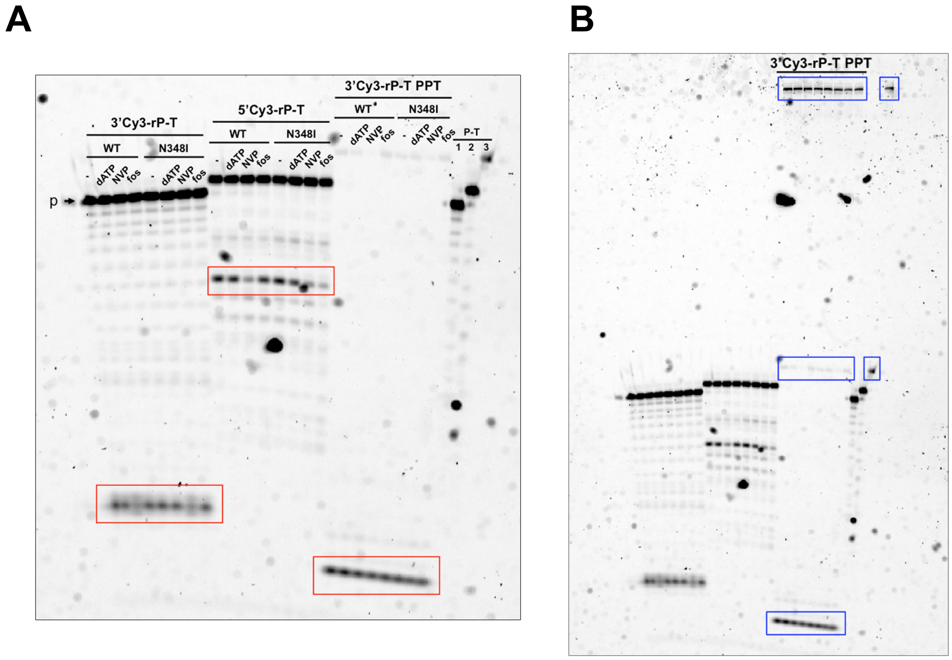

Figure 4.5 (A) PAGE assay of RT RNase H activity. Experiments were conducted with 3'Cy3-rP-T, 5'Cy3-rP-T, or 3'Cy3-rP-T PPT substrates in the presence of WT or the N348I double-mutant RTp66-p51 (100 nM). The effects of complementary dNTPs, NVP, and foscarnet (fos) on RNase H activity were tested. Black arrow indicates full-length primer. Major cleavage products are highlighted in red. Lanes denoted 1, 2, and 3 are 3'Cy3-rP-T, 5'Cy3-rP-T, and 3'Cy3-rP-T PPT, respectively, in the absence of RT. (B) Full gel scan showing the slow migration of the unique 3'Cy3-rP PPT sequence in the denaturing polyacrylamide gel. Three major bands are observed (blue box).

4.2.2 SM-PIFE Characterization of RT RNase H Activity

SM-PIFE characterization of RT RNase H activity has been conducted by Dr. Hsiao-Wei Liu. Using the substrates in Figure 4.3, Dr. Liu has established an analytical method for the detection of RT RNase H-catalyzed cleavage events by observing the initial binding of RT via PIFE, followed by a transition to a dark-state due to RNase H cleavage of the primer strand. Briefly, intensity-time trajectories of individual 3'Cy3-rP-T molecules show an increase in fluorescence intensity corresponding to the binding of RT, and for each enhancement event, the duration of the enhancement until the transition to a dark-state (τ_{dwell}), is determined. Histograms composed of τ_{dwell} values for individual 3'Cy3-rP-T molecules can be constructed and the resulting distribution can be used to obtain insight into the kinetics of the reaction. The results are in preparation for publication and will not be discussed further.

4.2.3 Future Work

We have demonstrated the potential of SM-PIFE as a versatile and sensitive method to monitor the RNase H activity of RT. Future work will include RT concentration titrations to gain insight into RNase H processivity. Does RT RNase H catalyze one cleavage event for every association event with its substrate, or must RT undergo multiple association/dissociation events prior to RNase H-catalyzed cleavage? SM-PIFE experiments may be suitable to address these questions.

It will also be interesting to examine the relationship between RNA secondary structure and the rate of RNA cleavage. For example, Liu et al. have demonstrated using SM-FRET that RT exhibits extensive orientational dynamics on substrates modeled after the reverse transcription initiation complex and that these dynamics are influenced by RNA sequence and secondary structure.[159]

4.3 Probing Polymerase- and RNase H-competent Binding Orientations of RT: The E478Q Mutant

SM-FRET studies have shown that RT can engage its nucleic acid substrates in two different orientations, depending on the composition of the primer.[121] On DNA/DNA substrates, RT engages in a DNA polymerase-competent binding orientation, while on RNA/DNA substrates, RT engages in a DNA RNase H-competent binding orientation. In order to see if the two different binding orientations of RT could also be observed using SM-PIFE, we utilized the Cy3-labeled 19/50-nucleotide primer/template DNA duplex shown in Figure 3.2 (i.e., Cy3-dP-T) as well as an identical substrate where the primer was RNA (Cy3-rP-T). Based on the SM-FRET results of Zhuang and co-workers, RT should bind to RNA/DNA substrates exclusively in an RNase H-competent orientation and DNA/DNA substrates in a flipped, polymerase-competent orientation (Figure 4.6).

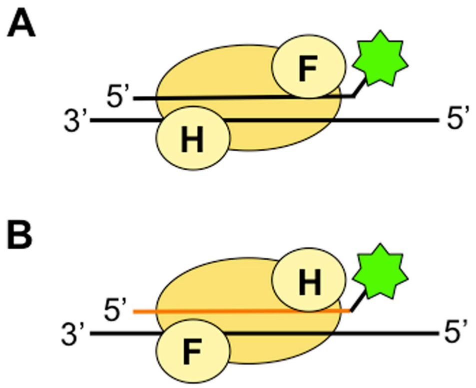

Figure 4.6 Using SM-PIFE to monitor the different binding orientations of RT. (A) RT is expected to bind Cy3-dP-T in polymerase-competent orientation, with the “F” fingers domain positioned close to the 3' end of the DNA (black) primer. (B) RT is expected to bind Cy3-rP-T in an RNase H-competent orientation, with the “H” RNase H domain positioned close to the 3' end of the RNA (orange) primer.

We hypothesized that each binding orientation of RT may induce a different degree of enhancement for our Cy3-rP-T versus Cy3-dP-T substrates. To test this hypothesis, we employed the RT E478Q mutant that lacks RNase H activity in order to monitor the enhancements arising from both orientations.[164] Upon flowing 300 nM of RT E478Q into imaging chambers containing surface-immobilized Cy3-rP-T or Cy3-dP-T, we observed the synchronous fluorescence enhancement of the majority of molecules recorded (Figure 4.7).

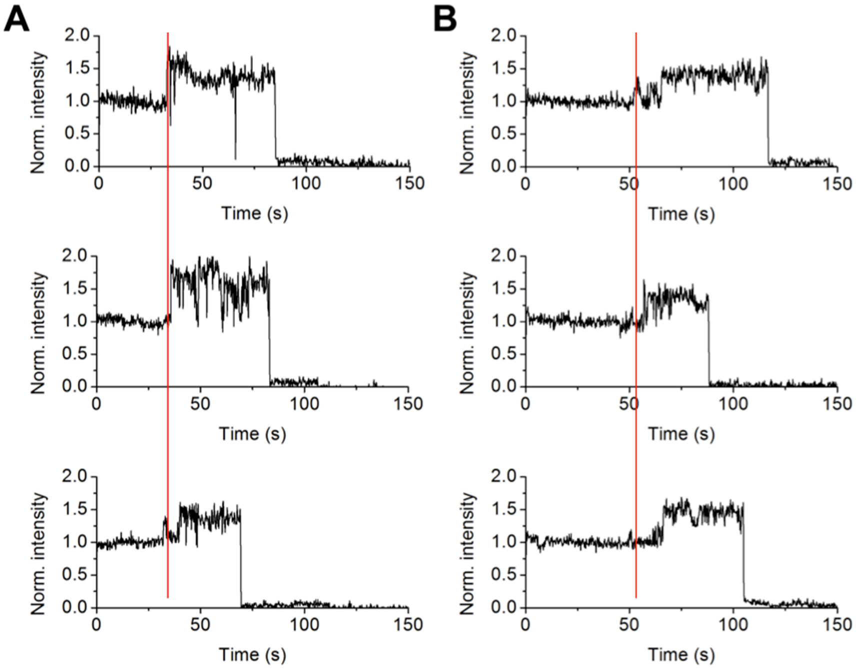

Figure 4.7 Intensity-time trajectories showing synchronous fluorescence enhancement of (A) Cy3-dP-T and (B) Cy3-rP-T upon flowing 300 nM of RT E478Q into the imaging chamber.

Ensemble normalized intensity histograms were constructed for Cy3-rP-T and Cy3-dP-T and fitted with Gaussian functions (Figure 4.8). At the high concentrations of RT E478Q used in these experiments (300 nM), we rarely observed intensity-time trajectories containing long off-states. As a result there is no appreciable peak corresponding to the off-state in the ensemble normalized intensity histograms. Gaussian fitting of the ensemble normalized intensity histograms resulted in peak center values of 1.20 and 1.26 and peak width (i.e., standard deviation) values of 0.38 and 0.52 for Cy3-rP-T (Figure 4.8A) and Cy3-dP-T (Figure 4.8B), respectively. Since SM-PIFE has been shown to be sensitive in the 0-4 nm range,[99] we anticipated that RT E478Q might induce fluorescence enhancement of different magnitudes between Cy3-dP-T and Cy3-rP-T substrates. Considering the fact that the double-stranded region of our substrates is 19 bp and the RT binding cleft can accommodate ~18-19 bp,[22] it is not completely surprising that the average enhancements are so similar between Cy3-dP-T (ca. 26%) and Cy3-rP-T (ca. 20%).

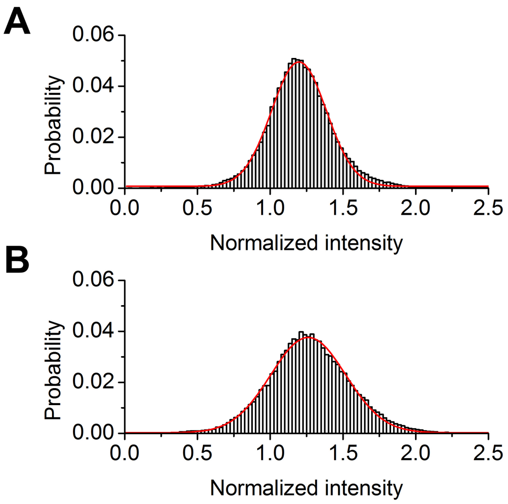

Figure 4.8 Ensemble normalized intensity histograms fitted with single Gaussian functions for (A) Cy3-dP-T and (B) Cy3-rP-T in the presence of 300 nM RT E478Q.

The larger standard deviation of the normalized intensity histogram for Cy3-dP-T (0.52; Figure 4.8A) compared to the corresponding value for Cy3-rP-T (0.38; Figure 4.8B) suggests that RT E478Q may be more dynamic around the 3' end of Cy3-dP-T compared to Cy3-rP-T, though the difference in distribution widths is modest. If RT E478Q binds to Cy3-dP-T in a polymerase-competent binding orientation, then the polymerase domain should be positioned close to the 3' end of the DNA primer, putting the fingers domain in close proximity to the Cy3 dye (Figure 4.6). Structural and computational studies show that the fingers domain of RT exhibits a high degree of conformational flexibility,[16, 25] thus we might expect to see a broader distribution of intensities when RT E478Q engages Cy3-dP-T in a polymerase-competent binding mode. Compared to the RT p66 polymerase domain, the conformational flexibility of the RNase H domain was originally thought to be quite rigid based on structural and biochemical studies of homologous human and bacterial RNase H enzymes. However, NMR and hydrogen exchange mass spectrometry of the isolated HIV-1 RNase H domain and heterodimeric RT, respectively, demonstrate that the RNase H domain is indeed flexible, especially near the C-terminus.[165, 166] Since both the polymerase and RNase H domains of RT may exhibit considerable dynamics, it is not surprising that the normalized intensity histograms in Figure 4.8 are so similar.

Overall, the results shown in Figures 4.7 and 4.8 indicate that the two different RT binding orientations cannot be conclusively distinguished from one another.

It is also important to note that these experiments were conducted using 300 nM RT E478Q. We have since shown using an electrophoretic mobility shift assay that at RTp66-p51 concentrations ≥ 300 nM, multiple RT/Cy3-dP-T complexes are observed suggesting the formation of higher order assemblies of RT/Cy3-dP-T (Figure 3.6). Thus we cannot be sure to what extent potential RT aggregation effects contribute to the intensity-time trajectories and ensemble normalized intensity histograms seen in Figures 4.7 and 4.8. Further optimization of experimental conditions is required.

To investigate the use of SM-PIFE as an assay to monitor the orientational dynamics of RT in more depth, future experiments should include probing the RT E478Q-induced fluorescence enhancement of Cy3-rP-T and Cy3-dP-T at lower RT concentrations. Additionally, RT E478Q-induced fluorescence enhancement should be investigated using the unique RNA PPT sequence since RT will bind to PPT primers and flip between both orientations without dissociating from the substrate.[121, 122] The PPT primer may help emphasize differences in the intensity-time trajectories and ensemble normalized intensity distributions between the polymerase- and RNase H-competent binding modes. By employing the PPT sequence, it would be possible to bias RT towards a polymerase-competent binding orientation using complementary dNTPs and the RNase H-competent binding orientation using an NNRTI such as NVP.[121] Furthermore, chimeric PPT primers containing 3' terminal deoxyribonucleotides promote the formation of polymerase-competent complexes.[121]

4.4 Multi-Step Enhancement of Cy3-rP-T Fluorescence in the Presence of RT E478Q and Nevirapine (NVP)

Small-molecule modulators of RT activity such as NNRTIs are known to induce large-scale conformational changes in RT[23, 47] or changes in RT orientational dynamics on RNA/DNA substrates.[121] Since SM-PIFE has been shown to be sensitive in the ~0-4 nm distance range,[99] we hypothesized that our

SM-PIFE assay may be sensitive enough to monitor inhibitor-induced RT conformational changes or binding dynamics, in real-time.

To test this hypothesis, we examined the effects of the NNRTI NVP on RT E478Q binding to the RNA/DNA substrate Cy3-rP-T. NVP (Figure 1.9A) is an NNRTI that binds allosterically to a hydrophobic pocket in the p66 domain ~10 Å from the polymerase active site.[43] X-ray crystal structures show that NNRTI binding to RT induces large-scale conformational changes, particularly in the fingers and thumb domain of the p66 subunit.[23] NNRTIs exert their inhibitory effect by perturbing the geometry of important catalytic residues in the polymerase active site.[47] Kohlstaedt et al. have proposed a "molecular arthritis" model stating that NNRTIs restrict the conformational flexibility of the polymerase domain via the hyperextension of the fingers and thumb.[23] In addition to X-ray crystallography data, molecular dynamics simulations have supported this hypothesis.[167] Recent single-molecule studies using the model PPT sequences also show that the NNRTIs NVP and EFZ increase the rate of RT flipping from the polymerase- to the RNase H-competent binding orientation, suggesting that NNRTI-induced hyperextension of the polymerase fingers and thumb domains ultimately loosens RT's grip on the substrate.[121] The conformation of RT where the fingers and thumb domain are hyperextended is referred to as the "open" conformation.

We have examined the effects of NVP on the binding of RT E478Q to our Cy3-rP-T substrate using SM-PIFE. Interestingly, a subset of the intensity-time trajectories (~25% of the 112 molecules analyzed) showed two- or multi-step enhancements, with the first enhancement occurring synchronously at ca. 40 s, corresponding to the arrival of RT E478Q in the imaging chamber (Figure 4.9). An ensemble normalized intensity histogram was constructed and compared to the histogram obtained for RT E478Q binding to Cy3-rP-T in the absence of NVP (Figure 4.10).

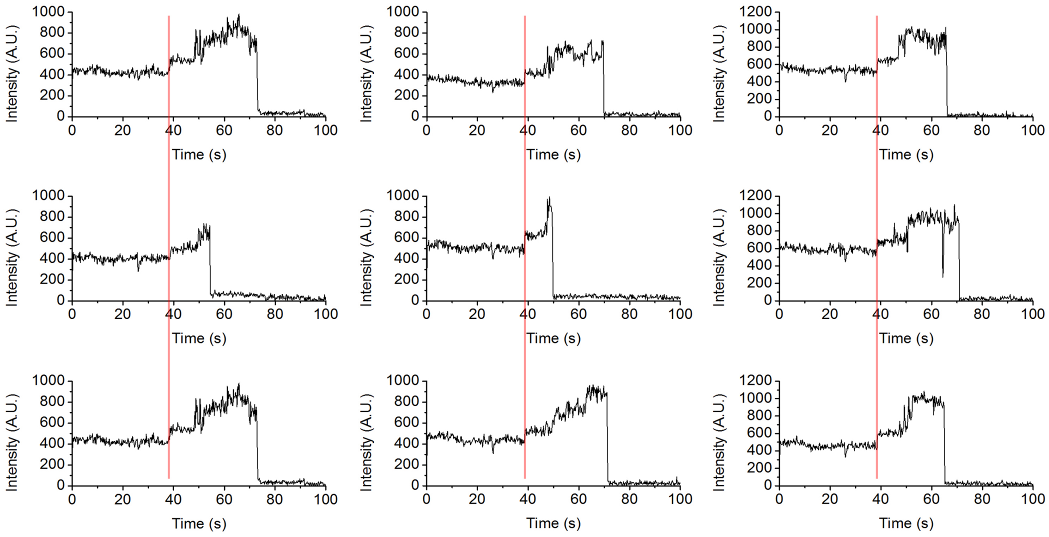

Figure 4.9 Intensity-time trajectories showing multi-step fluorescence enhancement of Cy3-rP-T upon flowing 300 nM of RT E478Q and 100 μM NVP into the imaging chamber. Multi-step fluorescence enhancement was observed in 25% of the 112 molecules analyzed. Red lines indicate the onset of the first binding event.

The normalized intensity histogram for RT E478Q binding to Cy3-rP-T in the absence of NVP shows one clear peak centered around ~1.20 (Figure 4.10A). In contrast, the distribution obtained in the presence of NVP is broader and less well defined (Figure 4.10B), with a significantly higher number of events occurring at normalized intensity values > 1.5.

Experiments with NVP were conducted using [NVP] = 100 μM ([NVP]:[RT E478Q] = 333:1), far above the K_d value of 0.08 μM for NVP binding to RT.[64] Thus we assume that the binding of NVP to RT E478Q is saturated prior to flowing the protein solution into the imaging chamber. The broad distribution of normalized intensities observed in the presence of NVP (Figure 4.10B) suggests that RT E478Q/NVP complexes may bind to Cy3-rP-T in an alternate conformation compared to RT E478Q in the absence of NVP. Indeed, crystal structures show that the open conformation of NNRTI-bound RT is fully capable of binding DNA/DNA[17,48] and RNA/DNA[168] hybrids. In crystal structures of RT/NNRTI/nucleic acid ternary complexes, the open structure of RT makes significantly fewer contacts with the nucleic acid compared to RT/nucleic acid binary complexes, especially involving residues at the tip of the p66 fingers domain. While the open conformation of RT/NNRTI complexes bound/unbound to double-stranded nucleic acids is evident from static crystallographic snapshots, little is known about the equilibria between different conformational states. It is possible that the step-wise fluorescence enhancements observed in Figure 4.9 are the result of conformational changes or a change in positioning of the RT on the substrate.

The preliminary results described above demonstrate that SM-PIFE may be used as a tool to investigate changes in the conformational landscape of RT in response to allosteric modulators such as NNRTIs. Crystallographic data do not provide a complete description of inhibition mechanisms, as crystal-packing effects may influence the relative occupancy of different RT conformations. SM-PIFE offers the opportunity to identify and characterize the thermodynamics and kinetics of important RT conformational changes under physiological solution conditions.

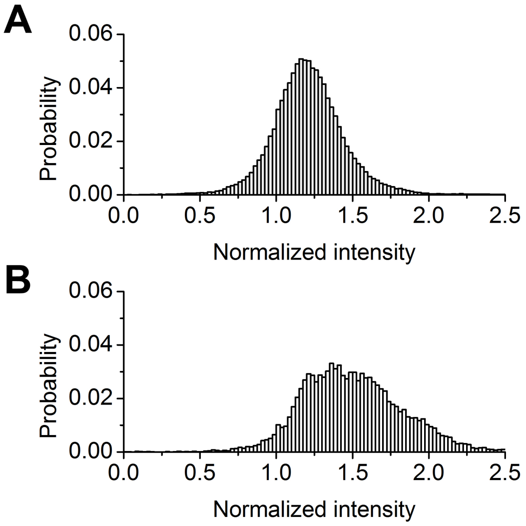

Figure 4.10 Ensemble normalized intensity histograms for (A) Cy3-rP-T in the presence of 300 nM RT E478Q and (B) Cy3-rP-T in the presence of 300 nM RT E478Q and 100 μM NVP.

4.5 Future Avenues

4.5.1 SM-PIFE Studies of the RT Polymerase Domain

All currently FDA-approved drugs targeting HIV-1 RT interfere with the enzyme's polymerase activity. While therapeutic intervention with RT inhibitors has drastically extended the life expectancy of those infected with HIV-1, the inevitable emergence of drug-resistant strains has necessitated the need for novel drug targets. In addition to the problem of drug resistance, RT inhibitors suffer from a lack of selectivity leading to drug toxicity. Often, NRTIs can be incorporated by host-cell DNA polymerases and are thus particularly notorious for inducing undesired side effects in patients undergoing antiretroviral therapy.

One of the defining features of retroviruses such as HIV-1 RT is a relatively low fidelity compared to other DNA polymerases. In higher organisms, high DNA synthesis fidelity is absolutely required as even single nucleotide mutations can have disastrous consequences on organism viability. Many eukaryotic DNA polymerases have evolved highly efficient proofreading mechanisms and are able to achieve extraordinary accuracy with base-substitution error rates in the range of one error per 1×10^7 - 1×10^8 nucleotides incorporated.[169] In contrast, low fidelity DNA synthesis is advantageous from an evolutionary standpoint and is often found in viruses and microbes where rapid evolution is required in order to adapt to environmental changes or evade an immune response.

During DNA synthesis, a DNA polymerase must discriminate the next complementary dNTP from a pool of 4 dNTPs and 4 rNTPs. Biochemical and kinetic studies have revealed that the origin of base selectivity arises primarily from differences in catalytic efficiency (i.e., k_{cat}/K_m) between the incorporation of complementary versus non-complementary nucleotides, with up to 10^7-fold differences.[170, 171] However, the precise molecular mechanism of nucleotide discrimination remains elusive.

Recent single-molecule fluorescence studies have shed light on the mechanism of how the high-fidelity T7 bacteriophage DNA polymerase and bacterial DNA polymerase I select for complementary nucleotides.[96, 104] Luo et al. utilized SM-PIFE and ensemble stop-flow fluorescence measurements to determine the rate of the T7 polymerase conformational change induced by the binding of a complementary dNTP. The rate of the conformational change (the closing of the fingers domain) was found to be significantly reduced for non-complementary dNTPs.[96] Similarly, SM-PIFE and SM-FRET studies of *E. coli* DNA polymerase I by Markiewicz et al. suggest that the DNA/polymerase complex is destabilized by the binding of a non-complementary dNTP or rNTP and that nucleotide rejection occurs prior to the closing of the fingers domain.[104] Both the aforementioned single-molecule studies highlight the important role of polymerase conformational dynamics in the mechanism of nucleotide selection.

What is the molecular basis for the low fidelity of HIV-1 RT? As described in section 1.1.4, RT-catalyzed DNA synthesis is highly error-prone leading to the rapid emergence of drug-resistant viral strains. Additionally, recent studies have shown that RT can efficiently incorporate rNTPs into growing DNA strands in macrophages.[172] A better understanding of the HIV-1 RT nucleotide selection mechanism and how it compares to homologous high fidelity polymerases may aid in the design of novel and more selective inhibitors against RT.

Single-molecule techniques are ideally positioned to probe mechanistic features related to RT polymerase function and nucleotide selection such as fingers domain dynamics during the formation of the polymerase/nucleic acid/dNTP ternary complex. SM-PIFE is particularly well suited in this regard, since no protein modification is required. The ability to monitor conformational changes of individual RT molecules in real-time may reveal unique kinetic checkpoints that are absent in non-viral polymerases, leading to the identification of new potential drug targets. Ideally, mechanistic insight into HIV-1 RT would also be broadly applicable to other low-fidelity viral polymerases.

Crystal structures of ternary RT complexes show that the fingers domain of RT moves ~5 Å towards the palm domain upon dNTP binding.[31] Since the PIFE effect is sensitive in the ~0-4 nm range,[99] we hypothesize that our SM-PIFE system may be sensitive to conformational changes in the RT polymerase domain. To test this hypothesis, the first logical experiment would be to examine the effects of RT in the presence of the next complementary dNTP, which in the case of Cy3-dP-T would be dATP (Figure 3.2). However our current Cy3-dP-T construct is not ideal for experiments involving dNTP due to the chemistry of the Cy3 linkage (Figure 4.11).

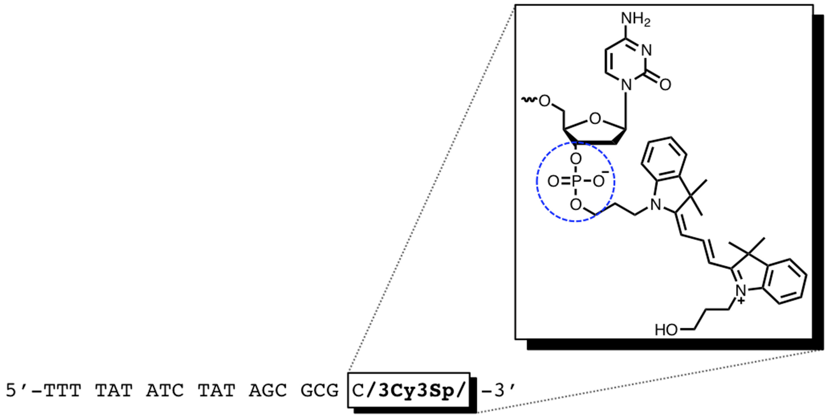

Figure 4.11 Chemical structure of the Cy3 linkage to the 3' end of the DNA primer strand.

The Cy3 dye is covalently linked to the terminal 3'OH group of the primer via a phosphate group. Thus in the RT/Cy3-dP-T complex, the Cy3 dye would likely interfere with the binding of the incoming dNTP. A potential alternative substrate based on the Cy3-dP-T sequence is proposed in Figure 4.12. The two important considerations in the substrate design are 1) a labeling site that does not interfere with dNTP binding and 2) 3' dideoxy-terminated nucleotide to prevent nucleotide incorporation during the experiment. Possible alternative labeling sites via thymidine bases are highlighted in red in Figure 4.12 and do not rely on conjugation to the 3'OH of the DNA primer strand. Thymidine was chosen as the base for fluorophore-conjugation based on commercial availability. The primer 3' terminal nucleotide (cytidine) is highlighted in blue and indicates the position of the dideoxy nucleotide modification to prevent nucleotide incorporation.

```
              5'-TTT TAT ATC TAT AGC GCG C-3'
3'-/3Bio/CG GTG CGG TGA AAA ATA TAG ATA TCG CGC GTG ACC TTC CCG ATT AGA TTA-5'
```

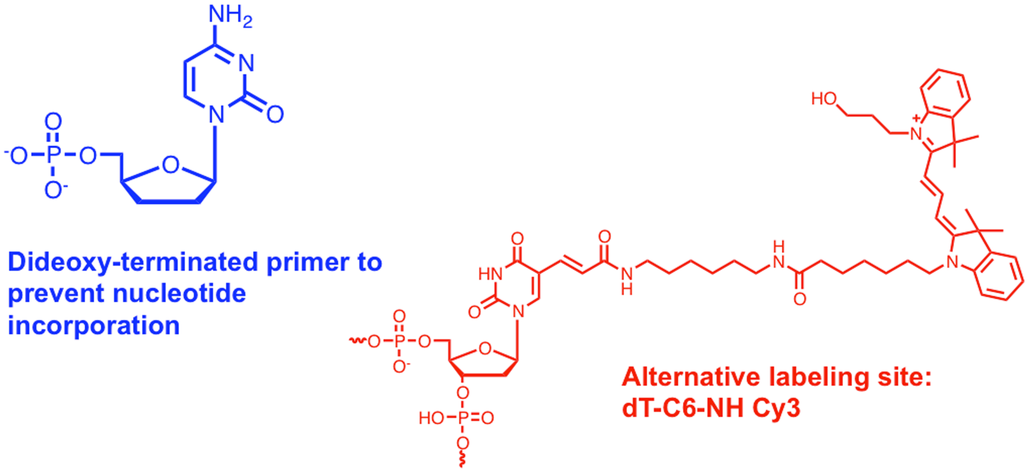

Figure 4.12 Alternative DNA/DNA substrate design for SM-PIFE studies of polymerase dynamics. In order to study the effects of dNTPs on RT polymerase domain dynamics, the substrate should retain an unmodified 3'OH group on the primer strand. Nucleotides labeled in red correspond to the commercially available Cy3 linkage via a thymidine base. To prevent nucleotide incorporation the 3' terminal nucleotide (blue) could be modified to a dideoxy nucleotide.

4.5.2 Towards Characterizing the Interplay Between RT Polymerase and RNase H Activity

We have shown that the use of SM-PIFE is a promising approach towards understanding the catalytic function and dynamic conformational landscape of RT. One interesting aspect of RT that would be difficult to study at the ensemble level would be the spatial and temporal regulation between the RT polymerase and RNase H active sites during different stages of reverse transcription. For example, during minus-strand DNA synthesis, RT engages in polymerase-dependent RNase H activity, where RNase H activity degrades the RNA template strand as DNA polymerase-catalyzed nucleotide incorporation occurs (section 1.1.3). While polymerase-dependent RNase H cleavage is spatially coordinated, the rate of RNase H cleavage is significantly slower (~7-fold) than the rate of nucleotide incorporation.[147] A system that enables the direct and simultaneous observation of both polymerase domain dynamics (e.g., fingers domain closing during nucleotide incorporation) and RNase H cleavage would help shed light on the mechanism of temporal regulation between the two RT active sites and how this regulation differs in response to different nucleic acid length/sequence/base composition, inhibitors, and drug resistance mutations.

An example of a potential construct to interrogate both RT active sites simultaneously integrates both SM-FRET and SM-PIFE (Figure 4.13). Using a FRET donor/acceptor pair such as Cy3/Cy5, RNase H activity could be monitored by following PIFE of Cy5[99] and the disappearance of the FRET signal. Motion in the fingers of the polymerase domain could be probed simultaneously via PIFE at the other end of the nucleic acid duplex.

Single-molecule fluorescence represents a powerful strategy that is primed to help us understand the fundamental relationship between enzyme conformational dynamics and catalytic activity and how this relationship is affected in response to small-molecule inhibitors and drug resistance-conferring mutations.

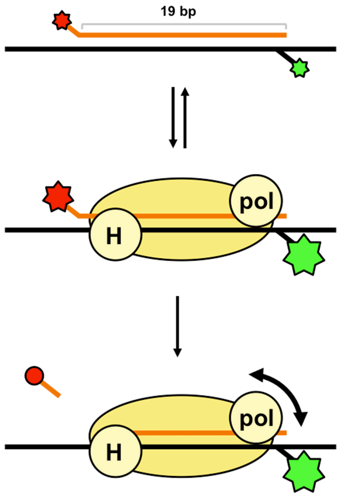

Figure 4.13 Characterizing the interplay between RT polymerase domain dynamics and RNase H activity at the single-molecule level. The integration of SM-FRET and SM-PIFE approaches may be used to investigate the temporal relationship between polymerase-catalyzed nucleotide incorporation and RNase H-catalyzed RNA hydrolysis. FRET between a donor/acceptor pair such as Cy3 (green) and Cy5 (red) along with changes in PIFE of both Cy3 and Cy5 could be monitored simultaneously.

5. References

1. Folkers, G. K.; Fauci, A. S., Controlling and ultimately ending the HIV/AIDS pandemic: a feasible goal. *Journal of the American Medical Association* **2010,** *304* (3), 350-351.

2. 2009 AIDS Epidemic Update. http://www.unaids.org/en/dataanalysis/knowyourepidemic/epidemiologypublications/2009aidsepidemicupdate/ (accessed April 4, 2013).

3. Reeves, J. D.; Doms, R. W., Human immunodeficiency virus type 2. *Journal of General Virology* **2002,** *83*, 1253-1265.

4. Frankel, A. D.; Young, J. A. T., HIV-1: Fifteen proteins and an RNA. *Annual Review of Biochemistry* **1998,** *67*, 1-25.

5. Turner, B. G.; Summers, M. F., Structural biology of HIV. *Journal of Molecular Biology* **1999,** *285* (1), 1-32.

6. Freed, E. O., HIV-1 and the host cell: an intimate association. *Trends in Microbiology* **2004,** *12* (4), 170-177.

7. Wilen, C. B.; Tilton, J. C.; Doms, R. W., HIV: Cell binding and entry. *Cold Spring Harbor Perspectives in Medicine* **2012,** *2* (8).

8. Friedrich, B. M.; Dziuba, N.; Li, G.; Endsley, M. A.; Murray, J. L.; Ferguson, M. R., Host factors mediating HIV-1 replication. *Virus Research* **2011,** *161* (2), 101-114.

9. Retroviruses. Coffin, J. M.; Hughes, S. H.; Varmus, H. E., Eds. Cold Spring Harbor Laboratory Press: New York, 1997. http://www.ncbi.nlm.nih.gov/books/NBK19376/.

10. Craigie, R.; Bushman, F. D., HIV DNA integration. *Cold Spring Harbor Perspectives in Medicine* **2012,** *2* (7).

11. Sundquist, W. I.; Kräusslich, H.-G., HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspectives in Medicine* **2012,** *2* (7).

12. Temin, H. M., Homology between RNA from Rous sarcoma virus + DNA from Rous sarcoma virus-infected cells. *Proceedings of the National Academy of Sciences of the United States of America* **1964,** *52* (2), 323-329.

13. Baltimore, D., Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **1970,** *226* (5252), 1209-1211.

14. Temin, H. M.; Mizutani, S., Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **1970,** *226* (5252), 1211-1213.

15. Hu, W.-S.; Hughes, S. H., HIV-1 reverse transcription. *Cold Spring Harbor Perspectives in Medicine* **2012,** *2* (10).

16. Sarafianos, S. G.; Marchand, B.; Das, K.; Himmel, D. M.; Parniak, M. A.; Hughes, S. H.; Arnold, E., Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *Journal of Molecular Biology* **2009,** *385* (3), 693-713.

17. Le Grice, S. F. J.; Naas, T.; Wohlgensinger, B.; Schatz, O., Subunit-selective mutagenesis indicates minimal polymerase-activity in heterodimer-associated p51 HIV-1 reverse transcriptase. *EMBO Journal* **1991,** *10* (12), 3905-3911.

18. Schatz, O.; Mous, J.; Legrice, S. F. J., HIV-1 RT-associated ribonuclease H displays both endonuclease and 3'-5' exonuclease activity *EMBO Journal* **1990,** *9* (4), 1171-1176.

19. di Marzo Veronese, F.; Copeland, T. D.; DeVico, A. L.; Rahman, R.; Oroszlan, S.; Gallo, R. C.; Sarngadharan, M. G., Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. *Science* **1986,** *231* (4743), 1289-1291.

20. Starnes, M. C.; Gao, W. Y.; Ting, R. Y.; Cheng, Y. C., Enzyme activity gel analysis of human immunodeficiency virus reverse transcriptase. *Journal of Biological Chemistry* **1988,** *263* (11), 5132-5134.

21. Ding, J.; Das, K.; Hsiou, Y.; Sarafianos, S. G.; Clark Jr, A. D.; Jacobo-Molina, A.; Tantillo, C.; Hughes, S. H.; Arnold, E., Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody fab fragment at 2.8 Å resolution. *Journal of Molecular Biology* **1998,** *284* (4), 1095-1111.

22. Jacobo-Molina, A.; Ding, J. P.; Nanni, R. G.; Clark, A. D.; Lu, X. D.; Tantillo, C.; Williams, R. L.; Kamer, G.; Ferris, A. L.; Clark, P.; Hizi, A.; Hughes, S. H.; Arnold, E., Crystal structure of human immunodeficiency virus type-1 reverse transcriptase complexed with double-stranded DNA at 3.0 angstrom resolution shows bent DNA. *Proceedings of the National Academy of Sciences of the United States of America* **1993,** *90* (13), 6320-6324.

23. Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A., Crystal structure at 3.5 angstrom resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **1992,** *256* (5065), 1783-1790.

24. Le Grice, S. F. J., Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise. *Journal of Biological Chemistry* **2012,** *287* (49), 40850-40857.

25. Bahar, I.; Erman, B.; Jernigan, R. L.; Atilgan, A. R.; Covell, D. G., Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function. *Journal of Molecular Biology* **1999,** *285* (3), 1023-1037.

26. Wang, J.; Smerdon, S. J.; Jager, J.; Kohlstaedt, L. A.; Rice, P. A.; Friedman, J. M.; Steitz, T. A., Structural basis of asymmetry in the human immunodeficiency virus type-1 reverse transcriptase heterodimer. *Proceedings of the National Academy of Sciences of the United States of America* **1994,** *91* (15), 7242-7246.

27. Hsiou, Y.; Ding, J.; Das, K.; Clark, A. D.; Hughes, S. H.; Arnold, E., Structure of unliganded HIV-1 reverse transcriptase at 2.7 angstrom resolution: Implications of conformational changes for polymerization and inhibition mechanisms. *Structure* **1996,** *4* (7), 853-860.

28. Ghosh, M.; Jacques, P. S.; Rodgers, D. W.; Ottman, M.; Darlix, J. L.; LeGrice, S. F. J., Alterations to the primer grip of p66 HIV-1 reverse transcriptase and their consequences for template-primer utilization. *Biochemistry* **1996,** *35* (26), 8553-8562.

29. Rausch, J. W.; Lener, D.; Miller, J. T.; Julias, J. G.; Hughes, S. H.; Le Grice, S. F. J., Altering the RNase H primer grip of human immunodeficiency virus reverse transcriptase modifies cleavage specificity. *Biochemistry* **2002,** *41* (15), 4856-4865.

30. Hsieh, J. C.; Zinnen, S.; Modrich, P., Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *Journal of Biological Chemistry* **1993,** *268* (33), 24607-24613.

31. Huang, H. F.; Chopra, R.; Verdine, G. L.; Harrison, S. C., Structure of a covalently trapped catalytic complex of HIV-I reverse transcriptase: Implications for drug resistance. *Science* **1998,** *282* (5394), 1669-1675.

32. Beilhartz, G. L.; Götte, M., HIV-1 Ribonuclease H: Structure, catalytic mechanism, and inhibitors. *Viruses* **2010,** *2* (4), 900-926.

33. Nowotny, M.; Gaidamakov, S. A.; Ghirlando, R.; Cerritelli, S. M.; Crouch, R. J.; Yang, W., Structure of human RNase H1 complexed with an RNA/DNA hybrid: Insight into HIV reverse transcription. *Molecular Cell* **2007,** *28* (2), 264-276.

34. Schultz, S. J.; Champoux, J. J., RNase H activity: Structure, specificity, and function in reverse transcription. *Virus Research* **2008,** *134* (1-2), 86-103.

35. Marchand, B.; Tchesnokov, E. P.; Götte, M., The pyrophosphate analogue foscarnet traps the pre-translocational state of HIV-1 reverse transcriptase in a Brownian ratchet model of polymerase translocation. *Journal of Biological Chemistry* **2007,** *282* (5), 3337-3346.

36. AIDSinfo FDA-Approved Anti-HIV Medications. http://aidsinfo.nih.gov/contentfiles/ApprovedMedstoTreatHIV_FS_en.pdf (accessed March 18, 2013).

37. Ilina, T.; Parniak, M. A., Inhibitors of HIV-1 reverse transcriptase. *Advances in Pharmacology* **2008,** *56*, 121-167.

38. Menendez-Arias, L., Molecular basis of human immunodeficiency virus drug resistance: An update. *Antiviral Research* **2010,** *85* (1), 210-231.

39. Mitsuya, H.; Weinhold, K. J.; Furman, P. A.; Stclair, M. H.; Lehrman, S. N.; Gallo, R. C.; Bolognesi, D.; Barry, D. W.; Broder, S., 3'-Azido-3'-deoxythymidine (BW A509U) - an antiviral agent that inhibits the infectivity and cytopathic effect of human lymphotropic-T virus type-III lymphadenopathy-associated virus in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **1985,** *82* (20), 7096-7100.

40. Sarafianos, S. G.; Das, K.; Clark, A. D.; Ding, J. P.; Boyer, P. L.; Hughes, S. H.; Arnold, E., Lamivudine (3TC) resistance in HIV-1 reverse transcriptase involves steric hindrance with beta-branched amino acids. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** *96* (18), 10027-10032.

41. Meyer, P. R.; Matsuura, S. E.; Mian, A. M.; So, A. G.; Scott, W. A., A mechanism of AZT resistance: An increase in nucleotide-dependent primer unblocking by mutant HIV-1 reverse transcriptase. *Molecular Cell* **1999,** *4* (1), 35-43.

42. de Bethune, M. P., Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: A review of the last 20 years (1989-2009). *Antiviral Research* **2010,** *85* (1), 75-90.

43. Sluis-Cremer, N.; Tachedjian, G., Mechanisms of inhibition of HIV replication by non-nucleoside reverse transcriptase inhibitors. *Virus Research* **2008,** *134* (1-2), 147-156.

44. De Clercq, E., The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. *Antiviral Research* **1998,** *38* (3), 153-179.

45. Spence, R. A.; Kati, W. M.; Anderson, K. S.; Johnson, K. A., Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Science* **1995,** *267* (5200), 988-993.

46. Rittinger, K.; Divita, G.; Goody, R. S., Human immunodeficiency virus reverse transcriptase substrate-induced conformational changes and the mechanism of inhibition by nonnucleoside inhibitors. *Proceedings of the National Academy of Sciences* **1995,** *92* (17), 8046-8049.

47. Esnouf, R.; Ren, J. S.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D., Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Nature Structural Biology* **1995,** *2* (4), 303-308.

48. Ren, J. S.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; Stammers, D., High-resolution structures of HIV-1 RT from 4 RT-inhibitor complexes. *Nature Structural Biology* **1995,** *2* (4), 293-302.

49. Das, K.; Ding, J. P.; Hsiou, Y.; Clark, A. D.; Moereels, H.; Koymans, L.; Andries, K.; Pauwels, R.; Janssen, P. A. J.; Boyer, P. L.; Clark, P.; Smith, R. H.; Smith, M. B. K.; Michejda, C. J.; Hughes, S. H.; Arnold, E., Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant. *Journal of Molecular Biology* **1996,** *264* (5), 1085-1100.

50. Ding, J. P.; Das, K.; Moereels, H.; Koymans, L.; Andries, K.; Janssen, P. A. J.; Hughes, S. H.; Arnold, E., Structure of HIV-1 RT/TIBO R-86183 complex reveals similarity in the binding of diverse nonnucleoside inhibitors. *Nature Structural Biology* **1995,** *2* (5), 407-415.

51. Smerdon, S. J.; Jager, J.; Wang, J.; Kohlstaedt, L. A.; Chirino, A. J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A., Structure of the binding site for nonnucleoside inhibitors of the reverse transcriptase of human immunodeficiency virus type-1. *Proceedings of the National Academy of Sciences of the United States of America* **1994,** *91* (9), 3911-3915.

52. Ren, J.; Milton, J.; Weaver, K. L.; Short, S. A.; Stuart, D. I.; Stammers, D. K., Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Structure* **2000,** *8* (10), 1089-1094.

53. Tachedjian, G.; Moore, K. L.; Goff, S. P.; Sluis-Cremer, N., Efavirenz enhances the proteolytic processing of an HIV-1 pol polyprotein precursor

and reverse transcriptase homodimer formation. *FEBS Letters* **2005,** *579* (2), 379-384.

54. Figueiredo, A.; Moore, K. L.; Mak, J.; Sluis-Cremer, N.; de Bethune, M. P.; Tachedjian, G., Potent nonnucleoside reverse transcriptase inhibitors target HIV-1 Gag-Pol. *PLoS Pathogens* **2006,** *2* (11), 1051-1059.

55. Johnson, V. A.; Calvez, V.; Gunthard, H. F.; Paredes, R.; Pillay, D.; Shafer, R.; Wensing, A. M.; Richman, D. D., 2011 Update of the drug resistance mutations in HIV-1. *Topics in Antiviral Medicine* **2011,** *19* (4), 156-164.

56. Ren, J.; Stammers, D. K., HIV reverse transcriptase structures: Designing new inhibitors and understanding mechanisms of drug resistance. *Trends in Pharmacological Sciences* **2005,** *26* (1), 4-7.

57. Brehm, J. H.; Koontz, D.; Meteer, J. D.; Pathak, V.; Sluis-Cremer, N.; Mellors, J. W., Selection of mutations in the connection and RNase H domains of human immunodeficiency virus type 1 reverse transcriptase that increase resistance to 3'-azido-3'-dideoxythymidine. *Journal of Virology* **2007,** *81* (15), 7852-7859.

58. Yap, S. H.; Sheen, C. W.; Fahey, J.; Zanin, M.; Tyssen, D.; Lima, V. D.; Wynhoven, B.; Kuiper, M.; Sluis-Cremer, N.; Harrigan, P. R.; Tachedjian, G., N348I in the connection domain of HIV-1 reverse transcriptase confers zidovudine and nevirapine resistance. *PLoS Medicine* **2007,** *4* (12), 1887-1900.

59. Hachiya, A.; Shimane, K.; Sarafianos, S. G.; Kodama, E. N.; Sakagami, Y.; Negishi, F.; Koizumi, H.; Gatanaga, H.; Matsuoka, M.; Takiguchi, M.; Oka, S., Clinical relevance of substitutions in the connection subdomain and RNase H domain of HIV-1 reverse transcriptase from a cohort of antiretroviral treatment-naive patients. *Antiviral Research* **2009,** *82* (3), 115-121.

60. Menendez-Arias, L.; Betancor, G.; Matamoros, T., HIV-1 reverse transcriptase connection subdomain mutations involved in resistance to approved non-nucleoside inhibitors. *Antiviral Research* **2011,** *92* (2), 139-149.

61. Delviks-Frankenberry, K. A.; Nikolenko, G. N.; Boyer, P. L.; Hughes, S. H.; Coffin, J. M.; Jere, A.; Pathak, V. K., HIV-1 reverse transcriptase connection subdomain mutations reduce template RNA degradation and enhance AZT excision. *Proceedings of the National Academy of Sciences* **2008,** *105* (31), 10943-10948.

62. Nikolenko, G. N.; Palmer, S.; Maldarelli, F.; Mellors, J. W.; Coffin, J. M.; Pathak, V. K., Mechanism for nucleoside analog-mediated abrogation of

HIV-1 replication: Balance between RNase H activity and nucleotide excision. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (6), 2093-2098.

63. Ehteshami, M.; Beilhartz, G. L.; Scarth, B. J.; Tchesnokov, E. P.; McCormick, S.; Wynhoven, B.; Harrigan, P. R.; Götte, M., Connection domain mutations N348I and A360V in HIV-1 reverse transcriptase enhance resistance to 3'-azido-3'-deoxythymidine through both RNase H-dependent and -independent mechanisms. *Journal of Biological Chemistry* **2008,** *283* (32), 22222-22232.

64. Schuckmann, M. M.; Marchand, B.; Hachiya, A.; Kodama, E. N.; Kirby, K. A.; Singh, K.; Sarafianos, S. G., The N348I mutation at the connection subdomain of HIV-1 reverse transcriptase decreases binding to nevirapine. *Journal of Biological Chemistry* **2010,** *285* (49), 38700-38709.

65. Nikolenko, G. N.; Delviks-Frankenberry, K. A.; Pathak, V. K., A novel molecular mechanism of dual resistance to nucleoside and nonnucleoside reverse transcriptase inhibitors. *Journal of Virology* **2010,** *84* (10), 5238-5249.

66. Biondi, M. J.; Beilhartz, G. L.; McCormick, S.; Götte, M., N348I in HIV-1 reverse transcriptase can counteract the nevirapine-mediated bias toward RNase H cleavage during plus-strand initiation. *Journal of Biological Chemistry* **2010,** *285* (35), 26966-26975.

67. Venezia, C. F.; Meany, B. J.; Braz, V. A.; Barkley, M. D., Kinetics of association and dissociation of HIV-1 reverse transcriptase subunits. *Biochemistry* **2009,** *48* (38), 9084-9093.

68. Parniak, M. A.; Min, K.-L.; Budihas, S. R.; Le Grice, S. F. J.; Beutler, J. A., A fluorescence-based high-throughput screening assay for inhibitors of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H activity. *Analytical Biochemistry* **2003,** *322* (1), 33-39.

69. Nakayama, G. R.; Bingham, P.; Tan, D.; Maegley, K. A., A fluorescence polarization assay for screening inhibitors against the ribonuclease H activity of HIV-1 reverse transcriptase. *Analytical Biochemistry* **2006,** *351* (2), 260-265.

70. Berezhna, S. Y.; Gill, J.; Lamichhane, R.; Millar, D. P., Single-molecule FRET reveals an innate fidelity checkpoint in DNA polymerase I. *Journal of the American Chemical Society* **2012,** *134* (27), 11261-11268.

71. Blanchard, S. C.; Kim, H. D.; Gonzalez, R. L.; Puglisi, J. D.; Chu, S., tRNA dynamics on the ribosome during translation. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (35), 12893-12898.

72. Deniz, A. A.; Laurence, T. A.; Beligere, G. S.; Dahan, M.; Martin, A. B.; Chemla, D. S.; Dawson, P. E.; Schultz, P. G.; Weiss, S., Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** *97* (10), 5179-5184.

73. Lakowicz, J., *Principles of Fluorescence Spectroscopy*. 3 ed.; Springer: 2006.

74. Selvin, P. R., The renaissance of fluorescence resonance energy transfer. *Nature Structural Biology* **2000,** *7* (9), 730-734.

75. Ha, T.; Zhuang, X.; Kim, H. D.; Orr, J. W.; Williamson, J. R.; Chu, S., Ligand-induced conformational changes observed in single RNA molecules. *Proceedings of the National Academy of Sciences* **1999,** *96* (16), 9077-9082.

76. Knight, A., *Single molecule biology*. 1 ed.; Elsevier: 2009.

77. Noji, H.; Yasuda, R.; Yoshida, M.; Kinosita, K., Direct observation of the rotation of F1-ATPase. *Nature* **1997,** *386* (6622), 299-302.

78. Kapanidis, A. N.; Strick, T., Biology, one molecule at a time. *Trends in Biochemical Sciences* **2009,** *34* (5), 234-243.

79. Hirschfeld, T., Optical microscopic observation of single small molecules. *Applied Optics* **1976,** *15* (12), 2965-2966.

80. Moerner, W. E.; Kador, L., Optical detection and spectroscopy of single molecules in a solid. *Physical Review Letters* **1989,** *62* (21), 2535-2538.

81. Orrit, M.; Bernard, J., Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal. *Physical Review Letters* **1990,** *65* (21), 2716-2719.

82. Shera, E. B.; Seitzinger, N. K.; Davis, L. M.; Keller, R. A.; Soper, S. A., Detection of single fluorescent molecules. *Chemical Physics Letters* **1990,** *174* (6), 553-557.

83. Betzig, E.; Chichester, R. J., Single molecules observed by near-field scanning optical microscopy. *Science* **1993,** *262* (5138), 1422-1425.

84. Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K.; Yanagida, T., Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature* **1995,** *374* (6522), 555-559.

85. Vale, R. D.; Funatsu, T.; Pierce, D. W.; Romberg, L.; Harada, Y.; Yanagida, T., Direct observation of single kinesin molecules moving along microtubules. *Nature* **1996,** *380* (6573), 451-453.

86. Sase, I.; Miyata, H.; Corrie, J. E. T.; Craik, J. S.; Kinosita, K., Real-time imaging of single fluorophores on moving actin with an epifluorescence microscope. *Biophysical Journal* **1995,** *69* (2), 323-328.

87. Lu, H. P.; Xun, L.; Xie, X. S., Single-molecule enzymatic dynamics. *Science* **1998,** *282* (5395), 1877-1882.

88. English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G. B.; Sun, H. Y.; Cherayil, B. J.; Kou, S. C.; Xie, X. S., Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nature Chemical Biology* **2006,** *2* (2), 87-94.

89. van Oijen, A. M.; Blainey, P. C.; Crampton, D. J.; Richardson, C. C.; Ellenberger, T.; Xie, X. S., Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* **2003,** *301* (5637), 1235-1238.

90. Velonia, K.; Flomenbom, O.; Loos, D.; Masuo, S.; Cotlet, M.; Engelborghs, Y.; Hofkens, J.; Rowan, A. E.; Klafter, J.; Nolte, R. J. M.; de Schryver, F. C., Single-enzyme kinetics of CALB-catalyzed hydrolysis. *Angewandte Chemie International Edition* **2005,** *44* (4), 560-564.

91. Zhuang, X. W.; Kim, H.; Pereira, M. J. B.; Babcock, H. P.; Walter, N. G.; Chu, S., Correlating structural dynamics and function in single ribozyme molecules. *Science* **2002,** *296* (5572), 1473-1476.

92. Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S., Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences* **1996,** *93* (13), 6264-6268.

93. Zhuang, X.; Bartley, L. E.; Babcock, H. P.; Russell, R.; Ha, T.; Herschlag, D.; Chu, S., A single-molecule study of RNA catalysis and folding. *Science* **2000,** *288* (5473), 2048-2051.

94. Cosa, G.; Harbron, E. J.; Zeng, Y.; Liu, H.-W.; O Connor, D. B.; Eta-Hosokawa, C.; Musier-Forsyth, K.; Barbara, P. F., Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophysical Journal* **2004,** *87* (4), 2759-2767.

95. Santoso, Y.; Joyce, C. M.; Potapova, O.; Le Reste, L.; Hohlbein, J.; Torella, J. P.; Grindley, N. D. F.; Kapanidis, A. N., Conformational

transitions in DNA polymerase I revealed by single-molecule FRET. *Proceedings of the National Academy of Sciences* **2010,** *107* (2), 715-720.

96. Luo, G.; Wang, M.; Konigsberg, W. H.; Xie, X. S., Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proceedings of the National Academy of Sciences* **2007,** *104* (31), 12610-12615.

97. Ha, T.; Rasnik, I.; Cheng, W.; Babcock, H. P.; Gauss, G. H.; Lohman, T. M.; Chu, S., Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* **2002,** *419* (6907), 638-641.

98. McKinney, S. A.; Declais, A. C.; Lilley, D. M. J.; Ha, T., Structural dynamics of individual Holliday junctions. *Nature Structural Biology* **2003,** *10* (2), 93-97.

99. Hwang, H.; Kim, H.; Myong, S., Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences* **2011,** *108* (18), 7414-7418.

100. Levitus, M.; Ranjit, S., Cyanine dyes in biophysical research: The photophysics of polymethine fluorescent dyes in biomolecular environments. *Quarterly Reviews of Biophysics* **2011,** *44* (1), 123-151.

101. Aramendia, P. F.; Negri, R. M.; Sanroman, E., Temperature-dependence of fluorescence and photoisomerization in symmetrical carbocyanines: Influence of medium viscosity and molecular structure. *Journal of Physical Chemistry* **1994,** *98* (12), 3165-3173.

102. Divita, G.; Mueller, B.; Immendoerfer, U.; Gautel, M.; Rittinger, K.; Restle, T.; Goody, R. S., Kinetics of interaction of HIV reverse transcriptase with primer/template. *Biochemistry* **1993,** *32* (31), 7966-7971.

103. Fischer, C. J.; Maluf, N. K.; Lohman, T. M., Mechanism of ATP-dependent translocation of E. coli UvrD monomers along single-stranded DNA. *Journal of Molecular Biology* **2004,** *344* (5), 1287-1309.

104. Markiewicz, R. P.; Vrtis, K. B.; Rueda, D.; Romano, L. J., Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucleic Acids Research* **2012,** *40* (16), 7975-7984.

105. Myong, S.; Cui, S.; Cornish, P. V.; Kirchhofer, A.; Gack, M. U.; Jung, J. U.; Hopfner, K.-P.; Ha, T., Cytosolic viral sensor RIG-I Is a 5'-triphosphate-dependent translocase on double-stranded RNA. *Science* **2009,** *323* (5917), 1070-1074.

106. Cornish, P. V.; Ha, T., A survey of single-molecule techniques in chemical biology. *ACS Chemical Biology* **2007,** *2* (1), 53-61.

107. Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T., Advances in single-molecule fluorescence methods for molecular biology. In *Annual Review of Biochemistry*, Annual Reviews: Palo Alto, 2008; Vol. 77, pp 51-76.

108. Weiss, S., Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature Structural Biology* **2000,** *7* (9), 724-729.

109. Millis, B. A., Evanescent-wave field imaging: an introduction to total internal reflection fluorescence microscopy. *Methods in Molecular Biology* **2012,** *823*, 295-309.

110. Trache, A.; Meininger, G. A., Total internal reflection fluorescence (TIRF) microscopy. *Current Protocols in Microbiology* **2008,** *Chapter 2*, 2A.2.1-2A.2.22.

111. Ambrose, E. J., The movements of fibrocytes. *Experimental Cell Research* **1961,** *8*, 54-73.

112. Axelrod, D., Cell-substrate contacts illuminated by total internal reflection fluorescence. *Journal of Cell Biology* **1981,** *89* (1), 141-145.

113. Roy, R.; Hohng, S.; Ha, T., A practical guide to single-molecule FRET. *Nature Methods* **2008,** *5* (6), 507-516.

114. Forkey, J. N.; Quinlan, M. E.; Goldman, Y. E., Measurement of single macromolecule orientation by total internal reflection fluorescence polarization microscopy. *Biophysical Journal* **2005,** *89* (2), 1261-1271.

115. Adachi, K.; Yasuda, R.; Noji, H.; Itoh, H.; Harada, Y.; Yoshida, M.; Kinosita, K., Stepping rotation of F-1-ATPase visualized through angle-resolved single-fluorophore imaging. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** *97* (13), 7243-7247.

116. Karymov, M. A.; Krasnoslobodtsev, A. V.; Lyubchenko, Y. L., Dynamics of synaptic sfil-DNA complex: Single-molecule fluorescence analysis. *Biophysical Journal* **2007,** *92* (9), 3241-3250.

117. Yang, H.; Luo, G. B.; Karnchanaphanurach, P.; Louie, T. M.; Rech, I.; Cova, S.; Xun, L. Y.; Xie, X. S., Protein conformational dynamics probed by single-molecule electron transfer. *Science* **2003,** *302* (5643), 262-266.

118. Sorokina, M.; Koh, H. R.; Patel, S. S.; Ha, T., Fluorescent lifetime trajectories of a single fluorophore reveal reaction intermediates during transcription initiation. *Journal of the American Chemical Society* **2009,** *131* (28), 9630-9631.

119. Lamichhane, R.; Solem, A.; Black, W.; Rueda, D., Single-molecule FRET of protein-nucleic acid and protein-protein complexes: Surface passivation and immobilization. *Methods* **2010,** *52* (2), 192-200.

120. Shi, X. H.; Lim, J.; Ha, T., Acidification of the oxygen scavenging system in single-molecule fluorescence studies: In situ sensing with a ratiometric dual-emission probe. *Analytical Chemistry* **2010,** *82* (14), 6132-6138.

121. Abbondanzieri, E. A.; Bokinsky, G.; Rausch, J. W.; Zhang, J. X.; Le Grice, S. F. J.; Zhuang, X., Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature* **2008,** *453* (7192), 184-189.

122. Liu, S. X.; Abbondanzieri, E. A.; Rausch, J. W.; Le Grice, S. F. J.; Zhuang, X. W., Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science* **2008,** *322* (5904), 1092-1097.

123. Hostomsky, Z.; Hostomska, Z.; Fu, T. B.; Taylor, J., Reverse transcriptase of human immunodeficiency virus type 1: functionality of subunits of the heterodimer in DNA synthesis. *Journal of Virology* **1992,** *66* (5), 3179-3182.

124. Sluis-Cremer, N.; Arion, D.; Abram, M. E.; Parniak, M. A., Proteolytic processing of an HIV-1 pol polyprotein precursor: Insights into the mechanism of reverse transcriptase p66/p51 heterodimer formation. *The International Journal of Biochemistry & Cell Biology* **2004,** *36* (9), 1836-1847.

125. Tachedjian, G.; Aronson, H.-E. G.; Goff, S. P., Analysis of mutations and suppressors affecting interactions between the subunits of the HIV type 1 reverse transcriptase. *Proceedings of the National Academy of Sciences* **2000,** *97* (12), 6334-6339.

126. Wu, J.; Amandoron, E.; Li, X.; Wainberg, M. A.; Parniak, M. A., Monoclonal antibody-mediated inhibition of HIV-1 reverse transcriptase polymerase activity: Interaction with a possible deoxynucleoside triphosphate binding domain. *Journal of Biological Chemistry* **1993,** *268* (14), 9980-9985.

127. Müller, B.; Restle, T.; Weiss, S.; Gautel, M.; Sczakiel, G.; Goody, R. S., Co-expression of the subunits of the heterodimer of HIV-1 reverse transcriptase in Escherichia coli. *Journal of Biological Chemistry* **1989,** *264* (24), 13975-13978.

128. Le Grice, S. F. J.; Grüninger-Leitch, F., Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography. *European Journal of Biochemistry* **1990,** *187* (2), 307-314.

129. Maier, G.; Dietrich, U.; Panhans, B.; Schröder, B.; Rübsamen-Waigmann, H.; Cellai, L.; Hermann, T.; Heumann, H., Mixed reconstitution of mutated subunits of HIV-1 reverse transcriptase coexpressed in Escherichia coli – two tags tie it up. *European Journal of Biochemistry* **1999,** *261* (1), 10-18.

130. Bavand, M. R.; Wagner, R.; Richmond, T. J., HIV-1 reverse transcriptase: Polymerization properties of the p51 homodimer compared to the p66/p51 heterodimer. *Biochemistry* **1993,** *32* (40), 10543-10552.

131. Fletcher, R. S.; Holleschak, G.; Nagy, E.; Arion, D.; Borkow, G.; Gu, Z.; Wainberg, M. A.; Parniak, M. A., Single-step purification of recombinant wild-type and mutant HIV-1 reverse transcriptase. *Protein Expression and Purification* **1996,** *7* (1), 27-32.

132. El Dirani-Diab, R.; Andreola, M. L.; Nevinsky, G.; Tharaud, D.; Barr, P. J.; Litvak, S.; Tarrago-Litvak, L., Biochemical characterization of the p51 sub-unit of human immunodeficiency virus reverse transcriptase in homo- and heterodimeric recombinant forms of the enzyme. *FEBS Letters* **1992,** *301* (1), 23-28.

133. Hizi, A.; McGill, C.; Hughes, S. H., Expression of soluble, enzymatically active, human immunodeficiency virus reverse transcriptase in Escherichia coli and analysis of mutants. *Proceedings of the National Academy of Sciences of the United States of America* **1988,** *85* (4), 1218-1222.

134. Restle, T.; Müller, B.; Goody, R. S., Dimerization of human immunodeficiency virus type 1 reverse transcriptase: A target for chemotherapeutic intervention. *Journal of Biological Chemistry* **1990,** *265* (16), 8986-8988.

135. Tisdale, M.; Ertl, P.; Larder, B. A.; Purifoy, D. J.; Darby, G.; Powell, K. L., Characterization of human immunodeficiency virus type 1 reverse transcriptase by using monoclonal antibodies: Role of the C terminus in antibody reactivity and enzyme function. *Journal of Virology* **1988,** *62* (10), 3662-3667.

136. Hansen, J.; Schulze, T.; Mellert, W.; Moelling, K., Identification and characterization of HIV-specific RNase H by monoclonal-antibody. *EMBO Journal* **1988,** *7* (1), 239-243.

137. Liu, H.-W.; Zeng, Y.; Landes, C. F.; Kim, Y. J.; Zhu, Y.; Ma, X.; Vo, M.-N.; Musier-Forsyth, K.; Barbara, P. F., Insights on the role of nucleic

acid/protein interactions in chaperoned nucleic acid rearrangements of HIV-1 reverse transcription. *Proceedings of the National Academy of Sciences* **2007,** *104* (13), 5261-5267.

138. Christian, T. D.; Romano, L. J.; Rueda, D., Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proceedings of the National Academy of Sciences* **2009,** *106* (50), 21109-21114.

139. Myong, S.; Rasnik, I.; Joo, C.; Lohman, T. M.; Ha, T., Repetitive shuttling of a motor protein on DNA. *Nature* **2005,** *437* (7063), 1321-1325.

140. Rasnik, I.; Myong, S.; Cheng, W.; Lohman, T. M.; Ha, T., DNA-binding Orientation and Domain Conformation of the E. coli Rep Helicase Monomer Bound to a Partial Duplex Junction: Single-molecule Studies of Fluorescently Labeled Enzymes. *Journal of Molecular Biology* **2004,** *336* (2), 395-408.

141. Joo, C.; McKinney, S. A.; Nakamura, M.; Rasnik, I.; Myong, S.; Ha, T., Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* **2006,** *126* (3), 515-527.

142. Debyser, Z.; De Clercq, E., Chemical crosslinking of the subunits of HIV-1 reverse transcriptase. *Protein Science* **1996,** *5* (2), 278-286.

143. McKinney, S. A.; Joo, C.; Ha, T., Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical Journal* **2006,** *91* (5), 1941-1951.

144. KruhØfter, M.; Urbanke, C.; Grosse, F., Two step binding of HIV-1 reverse transcriptase to nucleic acid substrates. *Nucleic Acids Research* **1993,** *21* (17), 3943-3949.

145. Götte, M.; Maier, G.; Onori, A. M.; Cellai, L.; Wainberg, M. A.; Heumann, H., Temporal coordination between initiation of HIV (+)-strand DNA synthesis and primer removal. *Journal of Biological Chemistry* **1999,** *274* (16), 11159-11169.

146. Venezia, C. F.; Howard, K. J.; Ignatov, M. E.; Holladay, L. A.; Barkley, M. D., Effects of efavirenz binding on the subunit equilibria of HIV-1 reverse transcriptase. *Biochemistry* **2006,** *45* (9), 2779-2789.

147. Kati, W. M.; Johnson, K. A.; Jerva, L. F.; Anderson, K. S., Mechanism and fidelity of HIV reverse transcriptase. *Journal of Biological Chemistry* **1992,** *267* (36), 25988-25997.

148. DeStefano, J. J.; Bambara, R. A.; Fay, P. J., Parameters that influence the binding of human immunodeficiency virus reverse transcriptase to nucleic acid structures. *Biochemistry* **1993,** *32* (27), 6908-6915.

149. Reardon, J. E., Human immunodeficiency virus reverse transcriptase: A kinetic analysis of RNA-dependent and DNA-dependent DNA polymerization. *Journal of Biological Chemistry* **1993,** *268* (12), 8743-8751.

150. Zheng, X.; Mueller, G. A.; Cuneo, M. J.; DeRose, E. F.; London, R. E., Homodimerization of the p51 subunit of HIV-1 reverse transcriptase. *Biochemistry* **2010,** *49* (13), 2821-2833.

151. Ignatov, M. E.; Berdis, A. J.; Le Grice, S. F. J.; Barkley, M. D., Attenuation of DNA replication by HIV-1 reverse transcriptase near the central termination sequence. *Biochemistry* **2005,** *44* (14), 5346-5356.

152. Bohlayer, W. P.; DeStefano, J. J., Tighter binding of HIV reverse transcriptase to RNA−DNA versus DNA−DNA results mostly from interactions in the polymerase domain and requires just a small stretch of RNA−DNA. *Biochemistry* **2006,** *45* (24), 7628-7638.

153. Ngo, A. T.; Karam, P.; Fuller, E.; Burger, M.; Cosa, G., Liposome encapsulation of conjugated polyelectrolytes: Toward a liposome beacon. *Journal of the American Chemical Society* **2007,** *130* (2), 457-459.

154. Karam, P.; Ngo, A. T.; Rouiller, I.; Cosa, G., Unraveling electronic energy transfer in single conjugated polyelectrolytes encapsulated in lipid vesicles. *Proceedings of the National Academy of Sciences* **2010,** *107* (41), 17480-17485.

155. Liu, H.-W.; Ngo, A. T.; Cosa, G., Enhancing the emissive properties of poly(p-phenylenevinylene)-conjugated polyelectrolyte-coated SiO2 nanoparticles. *Journal of the American Chemical Society* **2011,** *134* (3), 1648-1652.

156. Cabodevilla, J. F.; Odriozola, L.; Santiago, E.; Martínez-Irujo, J. J., Factors affecting the dimerization of the p66 form of HIV-1 reverse transcriptase. *European Journal of Biochemistry* **2001,** *268* (5), 1163-1172.

157. Götte, M.; Rausch, J. W.; Marchand, B.; Sarafianos, S.; Le Grice, S. F. J., Reverse transcriptase in motion: Conformational dynamics of enzyme-substrate interactions. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2010,** *1804* (5), 1202-1212.

158. Loya, S.; Tal, R.; Hughes, S. H.; Hizi, A., The effects of cysteine mutations on the catalytic activities of the reverse transcriptase of human immunodeficiency virus type-1. *Journal of Biological Chemistry* **1992,** *267* (20), 13879-13883.

159. Liu, S.; Harada, B. T.; Miller, J. T.; Le Grice, S. F. J.; Zhuang, X., Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nature Structural & Molecular Biology* **2010,** *17* (12), 1453-1460.

160. Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wohrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. M., Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. *Proceedings of the National Academy of Sciences* **2003,** *100* (4), 1655-1660.

161. Wahba, A. S.; Esmaeili, A.; Damha, M. J.; Hudson, R. H. E., A single-label phenylpyrrolocytidine provides a molecular beacon-like response reporting HIV-1 RT RNase H activity. *Nucleic Acids Research* **2010,** *38* (3), 1048-1056.

162. Kim, S.; Streets, A. M.; Lin, R. R.; Quake, S. R.; Weiss, S.; Majumdar, D. S., High-throughput single-molecule optofluidic analysis. *Nature Methods* **2011,** *8* (3), 242-245.

163. Shaw-Reid, C. A.; Munshi, V.; Graham, P.; Wolfe, A.; Witmer, M.; Danzeisen, R.; Olsen, D. B.; Carroll, S. S.; Embrey, M.; Wai, J. S.; Miller, M. D.; Cole, J. L.; Hazuda, D. J., Inhibition of HIV-1 ribonuclease H by a novel diketo acid, 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid. *Journal of Biological Chemistry* **2003,** *278* (5), 2777-2780.

164. Schatz, O.; Cromme, F. V.; Gruningerleitch, F.; Legrice, S. F. J., Point mutations in conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function. *FEBS Letters* **1989,** *257* (2), 311-314.

165. Seckler, J. M.; Howard, K. J.; Barkley, M. D.; Wintrode, P. L., Solution structural dynamics of HIV-1 reverse transcriptase heterodimer. *Biochemistry* **2009,** *48* (32), 7646-7655.

166. Mueller, G. A.; Pari, K.; DeRose, E. F.; Kirby, T. W.; London, R. E., Backbone dynamics of the RNase H domain of HIV-1 reverse transcriptase. *Biochemistry* **2004,** *43* (29), 9332-9342.

167. Zhou, Z. G.; Madrid, M.; Evanseck, J. D.; Madura, J. D., Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. *Journal of the American Chemical Society* **2005,** *127* (49), 17253-17260.

168. Lapkouski, M.; Tian, L.; Miller, J. T.; Le Grice, S. F. J.; Yang, W., Complexes of HIV-1 RT, NNRTI and RNA/DNA hybrid reveal a structure compatible with RNA degradation. *Nature Structural & Molecular Biology* **2013,** *20* (2), 230-236.

169. Kunkel, T. A., DNA replication fidelity. *Journal of Biological Chemistry* **2004,** *279* (17), 16895-16898.

170. Beard, W. A.; Shock, D. D.; Vande Berg, B. J.; Wilson, S. H., Efficiency of correct nucleotide insertion governs DNA polymerase fidelity. *Journal of Biological Chemistry* **2002,** *277* (49), 47393-47398.

171. Beard, W. A.; Wilson, S. H., Structural insights into the origins of DNA polymerase fidelity. *Structure* **2003,** *11* (5), 489-496.

172. Kennedy, E. M.; Amie, S. M.; Bambara, R. A.; Kim, B., Frequent incorporation of ribonucleotides during HIV-1 reverse transcription and their attenuated repair in macrophages. *Journal of Biological Chemistry* **2012,** *287* (17), 14280-14288.