Title
Characterizing Human Immunodeficiency Virus Type 1 Reverse Transcriptase and Integrase Interaction

Author
Tekeste, Shewit

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Characterizing Human Immunodeficiency Virus Type 1 Reverse Transcriptase and Integrase Interaction

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular and Medical Pharmacology

By

Shewit Tekeste

2014

ABSTRACT OF THE DISSERTATION

Characterizing Human Immunodeficiency Virus Type 1 Reverse Transcriptase and Integrase Interaction

By

Shewit Tekeste

Doctor of Philosophy in Molecular and Medical Pharmacology

University of California, Los Angeles, 2014

Professor Samson A. Chow, Chair

Human immunodeficiency virus type 1 (HIV-1) replication requires the reverse transcription of its RNA genome into double-stranded DNA copies within the cytoplasm before integration into the host chromosome. Reverse transcriptase (RT) and integrase (IN) are the viral enzymes responsible for catalyzing the essential steps of reverse transcription and integration, respectively. While numerous studies have led to a greater understanding of the functional roles that RT and IN individually play in HIV-1 replication, little is known about the functional role of RT-IN complex formation in vivo. We hypothesize that RT-IN interaction has functional significance in HIV-1 reverse transcription and replication kinetics. We have mapped the putative binding domain of RT on IN to nine residues on the IN C-terminal domain (CTD). We tested the significance of RT-IN interaction on reverse transcription and viral replication, and

identified the step at which viral replication of these IN mutants become defective. We observed impairment of viral cDNA synthesis in viruses harboring IN mutations at the putative RT-binding surface, supporting our hypothesis that the RT-IN interaction during the reverse transcription step is biologically relevant. We have developed a pharmacological approach to study and screen for inhibitors against the RT-IN interaction. Lastly, we have also initiated biochemical studies to determine the IN binding domain domain on RT to contribute to the full understanding of the binding mechanism.

The dissertation of Shewit Tekeste is approved by.

James Wohlschlegel

Paul Krogstad

Jing Haung

Jerome A. Zack

Samson A. Chow, Committee Chair

University of California, Los Angeles

2014

To my loving parents, thank you for the endless sacrifices and struggles you overcame to provide me with the gift of education.

Table of Contents

Acknowledgements

First and foremost, I thank God for blessing me with great opportunities and people in my life.

I thank my advisor Dr. Samson Chow for giving me the opportunity to conduct research under his supervision. He gave me freedom to explore different research areas, taught me how to face problems as challenges, and was always there whenever I needed anything. Thank you for being patient and for pushing me to grow. I am grateful for your continuous mentorship through my time as a UCLA graduate student. I thank the members of my thesis committee, Dr. James Wohlschlegel, Dr. Paul Krogstad, Dr. Jing Haung, and Dr. Jerome Zack for their expert counsel and mentorship. I thank my undergraduate mentor, Dr. Melisa Jurica, for giving me a chance to explore the world of research.

I thank the past and present members of the Chow lab members, especially those mentioned below. I thank Dr. Cora Woodward, and Dr. Marissa Briones, for their constant advice through graduate school and their patience when teaching me new experiments. Thanks to Dr. Thomas Wilkinson for the encouragement through failed experiments, endless discussion of science, life and careers. Thanks to Xiaowen Xu, Jonathan Wu, my first undergraduate student Carman Chan, for making lab an exciting place to be. I'm thankful for being a member of pharmacology department, for Dr. Michael Phelps for his constant encouragement and fostering a friendly atmosphere. To Dr. Sherly Mosessian, thank you for being a great mentor. You always knew what to say to get my spirit uplifted. You have taught me to keep pushing barriers, and find the leader in me. Most of all, thanks for the fiction books you provided for my mental vacations. Special thanks to Mr. Lloyd Gill, for encouraging me to keep my head up and try again when experiments didn't work. For making my late nights in the lab lively, ensuring I kept a positive

outlook on life and that I was always safe. You can officially retire Mr. Gill, thank you for being my UCLA adopted father.

I want to especially thank my very dear friend, Angelica Riestra, for constantly pushing me to be better and grounded. We have gone through the ups and downs of graduate school together, constantly picking each other up to face the next challenge. I am truly blessed to have strong-minded, driven friend to keep each other on our toes. I thank all my other ACCESS friends, Miguel Nava, Randall Chin, Dr. Ronik Khachatoorian, and Sacha Prashad, for always willing to help me with reagents, brainstorming sessions over coffee, and fun times away from lab. To Ms. Cortnee Wilson, I have seen you battle breast cancer and still manage to laugh. You have taught me that every day is a blessing and tomorrow is never promised. You have helped me change my perspective, to always see the good in every situation, and when life throws you lemons, instead of complaining you happily and gratefully make lemonade. Thank you for pushing me to become a better person.

To those that helped me through my dissertation, Angelica Riestra, Dr. Edwin Paz, Dr. Michael Ghebreab, Dr. Ronik Khachatoorian, Kristin Yamada, thank you for being patient with me. I truly appreciate your time, valuable suggestions, and never ending motivation. I especially want to thank Dr. Aaron Chapman for providing a writing environment, and spiritual motivation to push through the mental hurdles; I am a better Godly woman because of you.

I am mostly thankful for my family. To my father, Mr. Stifanos Tekeste, you have always been there for me. You carried me to my first day in first grade, since then you have always encouraged me to dream and aspire to be better. You have sacrificed a lot to provide us with education. Thank you so much for your love, for being my therapist, and allowing me to become an independent woman not confined in traditional ideology. I love you so much dad, and I hope I

have made you proud. To my caring mother, Teka, I aspire to be half the woman you are. Your constant prayers, unconditional love, a fighter in all areas of life, you give hope a new meaning. Your encouraging words and prayers, beautiful laughs, and our dance sessions helped me get through graduate school. Thank you mom and dad for showing me what love is and providing me with a safe loving environment. Thank you big brothers, Dejen and Denden, you two have set the standards high, and blessed to have your wise counsel and encouragement through this journey. To my younger siblings, Siye, Degol and Awate, thank you for bringing the kid out of me when I got too serious. I am truly blessed to have a loving family.

Chapter 2 is a version of a paper that is in preparation for submission: Characterizing HIV-1 reverse transcriptase and integrase interaction. Co-authors include Thomas Wilkinson, Robert Clubb, Samson Chow (Principle Investigator). This work could not be completed without the gracious gift of purified recombinant RT heterodimer (p66/p51) from Stuart Le Grice at the National Cancer Institute and expression plasmid encoding IN residues 220-270 provided by Dr. Robert Craigie from the National Institutes of Health. I am also thankful for the Dr. Dimitrios Vatakis and Dr. Jerome Zack lab for providing a PCR-clean environment to analyze reverse transcription products of the IN mutants analyzed. Chapter 3 is in collaboration with the UCLA molecular screening shared resource facility. I want to thank Dr. Robert Damoiseaux for his expert guidance through each step of the assay development. Thank you for ordering the special Alphascreen reader, for your patient mentorship, and believing in my work. Dr. Thomas Wilkinson first initiated chapter 4 co-crystallization idea, and I took over the analysis. The mass spectrometry optimization was made possible by collaborating with Dr. James Wohlschlegel, and especially Dr. Ajay Vashisht for running the mass spectrometry analysis.

Finally I would like to thank the UCLA ACCESS program, Virology and Gene Therapy Training Grant, and Molecular and Medical Pharmacology program for their financial contributions to me throughout my graduate school tenure.

Vita

Research Positions and Employment

2007-2008	Undergraduate research, Department of Molecular, Cell and Development, University of California Santa Cruz, UCSC
2009-Present	Graduate researcher, Department of Molecular and Medical Pharmacology, University of California Los Angeles, UCLA
2014-Present	Adjunct Professor, Department of Health and Human Sciences, Loyola Marymount University, Los Angeles, CA

Awards/Honors

2004-2008	Horatio Alger College Scholar California Scholar
2006-2007	Minority Biomedical Research Support, MBRS, UCSC
2007-2008	Minority Access to Research Careers, MARC, UCSC
2007-2008	Center for Biomolecular Science and Engineering Diversity Award for Undergraduate Research, UCSC
2008	Alliance for Graduate Education and the Professoriate (AGEP), UCLA
2009	FORD Predoctoral Fellowship, Honorable mention recipient
2009	NSF Predoctoral Fellowship, Honorable mention recipient
2010	Pfizer 3D Drug Discovery Workshop, First Place Team

2011	Centers of AIDS Research (CFAR), UCLA Graduate Student Award
2013	Business of Science Certificate, UCLA
2010-2013	Virology and Gene Therapy Training Grant, #5T32AI060567-08

Publications and Presentations

- Wilkinson, T.A., Januszyk, K., Phillips, M.L., **Tekeste, S.S.,** Zhang, M., Miller, J.T., Le Grice, S.F.J., Clubb, R.T. and Chow, S.A. Identifying and characterizing a functional HIV-1 reverse transcriptase-binding site on integrase. *J. Biol. Chem*. **284:** 7931-7939, 2009.

- **Tekeste, S.S.,** Wilkinson, T.A., Januszyk, K., Miller, J.T., Le Grice, S.F.J., Clubb, R.T. and Chow, S.A. Interaction between reverse transcriptase and integrase is required for reverse transcription and infectivity of HIV-1 *in vivo*. In preparation for submission, 2014.

- Briones, M., Rusmevichientong, A., **Tekeste, S.S.** and Chow, S.A. Binding of human and rhesus macaques TRIM5a restriction factors to human immunodeficiency virus type 1 viral cores and the effect on core disassembly. In preparation, 2014.

- **Tekeste, S.,** Chow S.A. (2012). HIV-1 integrase interaction with reverse transcriptase is biologically significant for viral replication. The 24th Workshop on Retroviral Pathogenesis, Philadelphia, PA.

- Wilkinson, T.A., **Tekeste, S.,** Weiner, E.M., Clubb, R.T., Miller, J.T., Le Grice, S.F.J. and Chow, S.A. (2012) Biophysical and functional studies of the role of HIV-1 integrase-reverse transcriptase interactions during viral replication. Structural Biology Related to HIV/AIDS, Bethesda, MD.

- Briones, M., **Tekeste, S.,** Rusmevichientong, A. and Chow, S.A. (2011) Human and rhesus macaque TRIM5alpha can bind HIV-1 cores and stabilize CA disassembly. AIDS Vaccine 2011, Bangkok, Thailand.

- **Tekeste, S.,** Chow S.A. (2010). Investigating the biological significance of HIV-1 integrase interactions with reverse transcriptase. Annual Biomedical Conference for Minority Students ABRCMS, Charlotte, NC.

- **Tekeste, S.,** Chow S.A. (2010). Biological significance of HIV-1 integrase interactions with reverse transcriptase. Annual Molecular and Medical Pharmacology Department retreat, Huntington Beach, CA.

- **Tekeste, S.,** Chow S.A. (2010). Characterizing HIV-1 reverse transcriptase binding domain on integrase. AIDS institute Cross-disciplinary conference at UCLA, CA.

- **Tekeste, S.,** Chow S.A. (2009). Biological significance of HIV-1 integrase interactions with reverse transcriptase. Annual Molecular and Medical Pharmacology Department research day, UCLA, CA.

- **Tekeste, S.,** Roybal, G., Jurica, M. (2008). Structural analysis of Spliceosome assembly in an AG dependent substrate using PP7 binding sequence. California Louis Stokes Alliances for Minority Participation (LSAMP) Statewide Undergraduate Research Symposium, Irvine, CA.

- **Tekeste, S.,** Roybal, G., Jurica, M. (2008). Structural analysis of Spliceosome assembly in an AG dependent substrate using PP7 binding sequence. University of California, Santa Cruz Undergraduate Research Symposium, Santa Cruz, CA.

Chapter 1: Introduction to Human Immunodeficiency Virus Type 1

3.1. Epidemiology of HIV

Thirty-three years have passed since the discovery of human immunodeficiency virus (HIV) as the etiological agent known to cause Acquired immunodeficiency diseases (AIDS). The World Health Organization reported over 35 million people worldwide are living with HIV, and 2.3 million new infections by the end of 2012 (1). HIV clearly continues to be a worldwide pandemic with more than 1.6 million AIDS related deaths in 2012. The route of transmission is through exposure to bodily fluids from infected individuals such as blood, semen, and vaginal secretions as well as in breast milk. Despite the time lapse since discovery and the large population affected, there is still no cure for HIV. Different ways to contain the spread of infection have been done by educating individuals on safer practices to prevent transmission, treating infected patients with combination of highly active anti-retroviral therapy (HAART), and new efficacy studies advocate for prophylactic use of current HAART for prevention.

3.2. Life cycle of HIV-1

HIV-1, a lentivirus and a member of the retrovirus family, has a 9.7 Kb long genome and encodes 15 proteins. The genomic material is packaged as two copies of single stranded RNA. Viral infection begins when the binding of viral glycoproteins to the host cell CD4-receptors and co-receptors. Viral glycoprotein attaches and promotes fusion of the host membrane receptors and viral envelope, allowing for the entry of the viral core into the host cell cytoplasm. The viral core is an intricate composition of capsid protein arranged in a hexamaric structure. The uncoating process has to happen in a dynamic fashion that allows proper reverse transcription process of the viral RNA genome. The reverse transcription of the RNA genome to double stranded cDNA is catalyzed by viral reverse transcriptase enzyme in the context of a reverse transcription complex. Subsequently, a pre-integration complex encompassing the new

transcribed cDNA and other viral proteins is imported through the host nuclear membrane. The viral enzyme integrase then catalyzes the integration of the viral cDNA into the host genome to establish a provirus. HIV hijacks host machinery to make copies of viral genome and translation of proteins, and packages in the host cytoplasm. Immature virions are released, and undergo the final maturation step of polyprotein processing catalyzed by the viral protease.

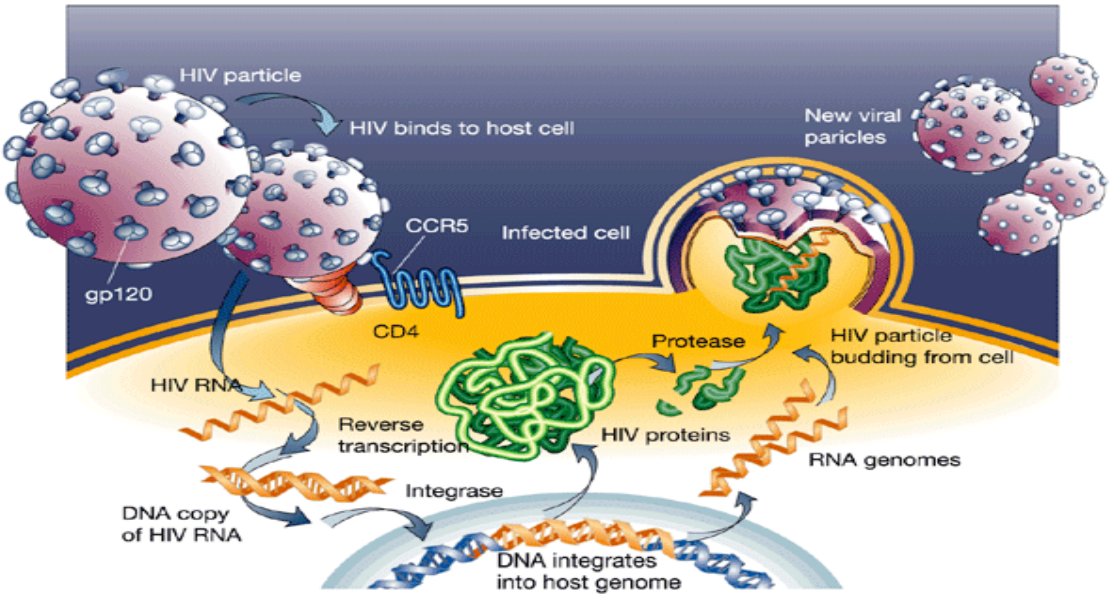


Figure 1. HIV-1 life cycle. Weiss et al. 2001. Nat Rev.

3.3. Role of Reverse transcriptase in HIV-1 Life Cycle

Reverse transcriptase (RT) is a heterodimer protein consisting of a 66 and 51 kDa molecular weight. The two polypeptides share sequence similarities except that the p66 is the active subunit and contains an RNase H domain. These distinct structure features enable RT to possess three enzymatic activities: RNA-dependent DNA polymerase, DNA-dependent DNA polymerase, and RNase H. The reverse transcription of the viral single stranded genomic RNA to double stranded DNA is initiated by the annealing of cellular $tRNA^{3}Lys$ primer to the primer binding site (PBS) on the 5'-end of RNA template. The cascade of events, outlined on Figure 2, progresses from successful initiation and elongation of early reverse transcription products (2), to the completion of a double stranded DNA species with Long Terminal Repeats (LTR) at its terminals.

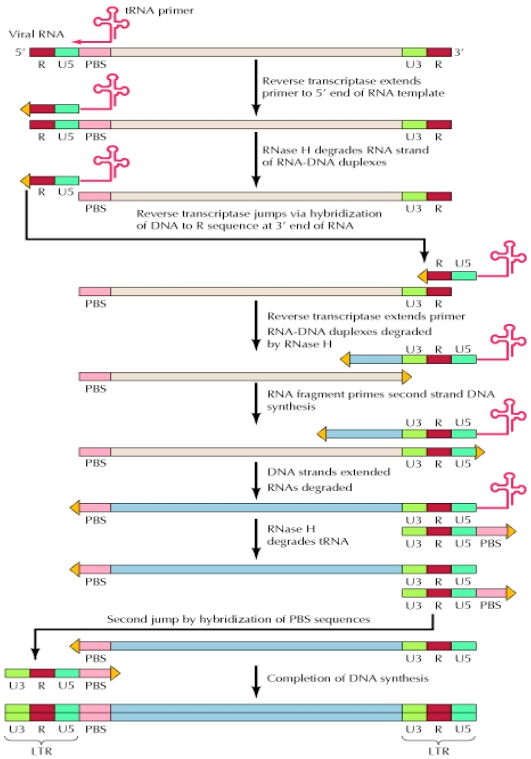


Figure 2. Reverse transcription of viral RNA to double stranded DNA.
Cooper GM.The Cell: A Molecular Approach. 2nd edition. 2000.

3.4. Role of Integrase in viral life cycle

Integrase (IN) is a 32-kDa protein. It is compromised of 3 distinct domains: a N-terminal domain with highly conserved zinc-binding HHCC motif, a catalytic core domain containing invariant catalytic DD(35)E motif important in coordinating divalent cation for catalytic activity, and a C-terminal domain (IN-CTD) known to bind DNA unspecificaly. Integration of newly transcribed viral double stranded DNA occurs in two steps. First, IN recognizes and catalyzes a 3'-end dinucleotide cleavage of the LTR ends, followed by a concerted cleavage-ligation reaction in which IN makes a staggered cut on the target DNA and joins the pre-processed viral cDNA (Figure 3) (3). The nicked DNA is then repaired by host machinery (4, 5). In vitro, IN can resolve the established provirus to its viral and target components, a reaction termed disintegration (6) .

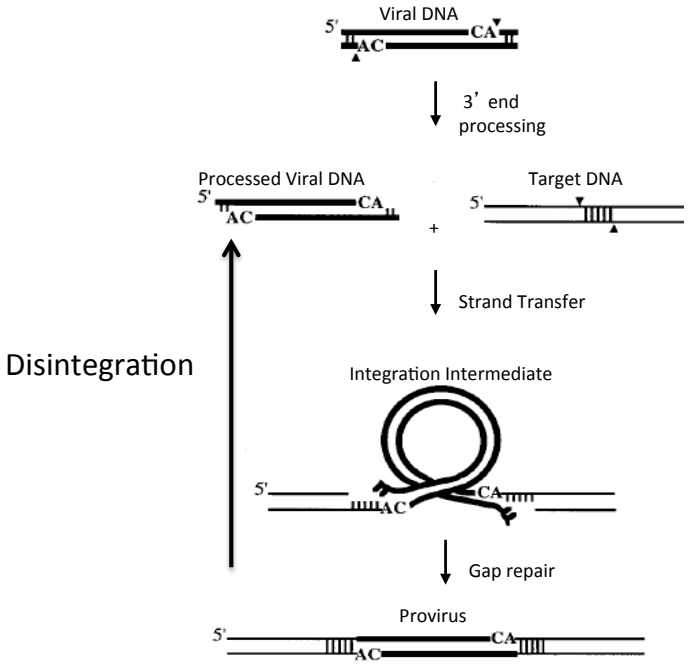


Figure 3. Schematic of HIV integration and *in vitro* disintegration activity.
Chow, S. A. 1997.

3.5. Current advancement of antiviral therapy

The discovery of HIV-1 as the causative agent of AIDS led to a search for treatment targeting the retrovirus. The first FDA approved treatment was in clinics by 1987. Zidovidine (AZT), a nucleoside analog inhibitor, competes with the incorporation of natural occurring dNTPs during reverse transcription. The high error rate of reverse transcription, caused by lack of proofreading capability of reverse transcriptase, introduces mutations that can overcome drug selection pressure. Patients soon developed resistance to mono-therapy and a new highly active antiretroviral therapy (HAART) regimen was established. Currently, over 26 antiretroviral therapies are available targeting different steps of the viral lifecycle. The six major classes of antivirals target HIV fusion, entry, nucleoside and non-nucleoside RT inhibitors, IN and protease inhibitors. The HAART regimen includes treating patients with combination of two nucleoside, one non-nucleoside RT inhibitor and one other class, either protease or integrase inhibitor.

Towards this end, the presented work illuminates the biological significance of the RT-IN interaction during retroviral replication as a potential novel target. Specifically, in Chapter 2 we characterize the effects of introducing amino acid substitutions at the RT-binding surface of HIV-1 IN on viral replication and reverse transcription in cell-based viral replication assays. The study elucidates the importance of RT-IN interactions during viral replication, thus confirming the *in vitro*-derived binding interface residues and that disrupting the RT-IN interaction impacts reverse transcription. To help fill the pharmacological need of a novel antiretroviral inhibitor chemical class, in Chapter 3 we develop and conduct a high throughput screen for small molecule inhibitors of RT-IN interaction. We screen small molecule libraries for inhibitors that disrupt the RT-IN interaction. Lastly, Chapter 4 outlines biochemical methods initiated to identify the surface on RT that interacts with IN. These findings will contribute to our

understanding of the binding mechanism of RT-IN as well as provide a rationale for using the RT-IN protein-protein interaction as a new anti-viral drug target.

Chapter 2: Reverse Transcriptase and Integrase Interaction Impact Viral Replication

3.1. Abstract

Human immunodeficiency virus type 1 (HIV-1) replication requires the reverse transcription of its RNA genome into a double-stranded DNA copy within the cytoplasm before integration into the host chromosome. Reverse transcriptase (RT) and Integrase (IN) are the viral enzymes responsible for catalyzing the essential steps of reverse transcription and integration, respectively. While numerous studies have led to a greater understanding of the functional roles that RT and IN individually play in HIV-1 replication, little is known about the functional role of RT-IN complex formation in vivo. We hypothesize that RT-IN interaction has functional significance in HIV-1 reverse transcription and replication kinetics. We have mapped the putative binding domain of RT on IN to nine residues on the IN C-terminal domain (CTD). To test the significance of RT-IN interaction on reverse transcription and viral replication, site-directed mutagenesis PCR was used to introduce single amino acid substitutions of nine different residues in the CTD to disrupt the RT-IN interaction. To identify the step at which viral replication of these IN mutants becomes defective, we carried out infection assays and analyzed viral polyprotein processing and packaging by western blot analysis. In addition, defects at reverse transcription for these nine mutant viruses were analyzed using quantitative real-time PCR to monitor early and late stages of viral cDNA synthesis in infected cells. Preliminary data show that six out of the nine mutants are replication-defective, and characterization of the viral replication defect of these mutants by quantitative real time PCR analysis show decreased copy number of early reverse transcription products. The observed impairment of viral cDNA synthesis in viruses harboring IN mutations at the putative RT-binding surface supports our hypothesis that the RT-IN interaction during the reverse transcription step is biologically relevant.

3.2. Introduction

To establish an infection after entry into the cell, human immunodeficiency virus type 1 (HIV-1) has to reverse transcribe its RNA genome to cDNA, followed by its integration into the host genome. Reverse transcriptase (RT) and Integrase (IN) are the viral enzymes responsible for catalyzing the essential steps of reverse transcription and integration, respectively. Both RT and IN are synthesized as part of the Gag-Pol polyprotein, which is later processed by the viral protease to produce active RT and IN during HIV-1 maturation (7).

RT is a heterodimeric enzyme consisting of 66 and 51 kDa subunits and catalyzes the RNA- and DNA-dependent reverse transcription of the viral RNA genome into double-stranded cDNA through a complex cascade of events (2, 8). The 32 kDa IN has three domains: an N-terminal zinc-binding domain, a catalytic core domain, and a C-terminal domain (IN-CTD) that binds DNA nonspecifically. IN catalyzes the integration of the viral cDNA into the host genome in two steps: an initial 3'-end processing step that removes two nucleotides at each 3'-end and exposes a highly-conserved CA 5'-overhang, followed by an end-joining step that inserts both processed viral DNA ends into the host cell genome (3). *In vitro,* IN can also catalyze a reverse reaction, termed disintegration, resolving a DNA mimic of the integrated provirus to products corresponding to a 3'-processed viral DNA end and a target duplex DNA (6). The dimeric form of IN catalyzes the 3'-end processing, while a tetramer has been shown to be required for strand transfer activity (9, 10).

Mutations in IN can result in pleiotropic effects throughout the viral life cycle (11), indicating that IN participates in other events beyond integration. IN mutants are grouped as either Class 1 or Class 2, depending on their effects on viral processes (12, 13). Class 1 IN mutants exhibit phenotypes that are enzymatically inactive at the 3'-end processing and end-

joining steps. Class 2 IN mutants exhibit full integrase activity *in vitro* but are nonetheless replication-defective due to pleiotropic effects at stages of replication other than the integration step. Several mutagenesis studies have described the effect of IN upon both the early stages of viral replication, such as uncoating (14), reverse transcription (15-18), nuclear import of preintegration complexes (19) and late steps post-integration, including polyprotein processing, packaging and maturation (12, 20).

Kappes and colleagues have demonstrated the impact of IN on reverse transcription through trans-complementation experiments using the HIV-1 accessory protein Vpr fused to wild-type IN, which could be packaged into virions, rescuing reverse transcription-defective viruses bearing mutant IN genes (21). The IN mutations that disrupt reverse transcription steps are scattered throughout the three domains of IN (11, 16, 17), suggesting the effect of IN upon reverse transcription may rely on multiple mechanisms that are yet to be understood.

One straightforward explanation for the influence of IN on reverse transcription is that IN binds RT and facilitates viral cDNA synthesis. Under *in vitro* conditions, HIV-1 IN physically interacts with RT and facilitates the early steps of reverse transcription (18). Deletion analyses of different IN domains and their ability to co-immunoprecipitate with recombinant RT revealed that the IN-CTD is necessary and sufficient to bind RT (15, 22). An RT-binding surface on IN has been previously determined by NMR spectroscopy using residues 220-270 of IN and RT lacking bound primer-template substrate (23). However, various experimental data suggest that IN can interact with RT in the context of bound nucleic acid substrate: *in vitro* reverse transcription assays have shown that IN enhances synthesis of early products, stimulating both the initiation and elongation phases of reverse transcription by increasing RT proccessivity and

suppressing the formation of pause products (18, 24). Other studies demonstrate that IN can interact with nucleic acid-bound RT to inhibit DNA-dependent DNA polymerization (25).

The precise RT conformation and functional form (either free or nucleic acid-bound) that IN initially targets during any particular IN-mediated reverse transcription pathway step remains largely unexamined. To explore the RT-IN interaction under conditions where RT is bound to nucleic acid, we have performed NMR chemical shift perturbation experiments to characterize a binding surface on the IN CTD in the presence of RT that had been pre-bound to a duplex DNA construct that corresponds to the primer binding site in the HIV-1 genome. These NMR experiments describe a putative binding surface on IN that is expanded relative to the previously-reported RT-binding surface on IN, which was ascertained in the absence of RT-bound DNA (23). We selectively substituted different IN residues that form the putative binding surface for both free and DNA-bound RT and characterized IN functional properties *in vitro* and in viral replication assays. We found that disruption of the putative binding surface on IN impaired viral replication due to the inability to synthesize viral cDNA, further indicating the functional significance of the RT-IN interaction during viral replication.

3.3. Results

Mapping an RT-DNA binding surface on IN by NMR spectroscopy

An RT-binding surface on IN was previously determined by NMR spectroscopy using recombinant IN 220-270 and RT lacking bound nucleic acid substrate (23). To determine whether the IN-RT interaction changes significantly when RT is bound to nucleic acid, we performed NMR chemical shift perturbation experiments to characterize the binding surface on IN 220-270 in the presence of RT that had been pre-bound to a duplex DNA segment corresponding to the primer binding site in the HIV-1 genome. X-ray crystal structures of this particular duplex DNA construct bound to RT are available (26, 27), and a dissociation constant of 38 nM ± 16 nM for RT bound to this DNA duplex was measured (28). Such dissociation constants, along with the RT-DNA complex concentrations used in the NMR experiments (14-54 μM) suggests that > 91% of RT is captured in its DNA-bound form under the conditions used in the NMR experiments.

Successive aliquots of unlabeled RT pre-bound with duplex DNA (1:1 molar ratio) were added to 150 μM ^{15}N-labeled sample of IN 220-270 to produce IN 220-270:RT-DNA ratios of 10:1, 5:1, and 2:1. No peaks showed any changes in chemical shift as the titration progressed, and most IN 220-270 signals showed relatively gradual reductions in peak intensity during the course of the titration. At a 2:1 IN 220-270:RT-DNA molar ratio, most of the peaks in the acquired spectrum became broadened to the noise level, and therefore no further titrations were performed.

Notably, amide backbone peaks associated with 16 residues showed significant decreases in intensity during the titration, as compared to other peaks in the spectra (Figure. 1A). Nine of

those residues (R231, L242, W243, G247, A248, V250, I251, Q252, and K258) had been previously mapped to an RT-binding surface on IN using analogous experiments with ^{15}N-labelled IN 220-270 and free RT (23). In the current titration experiment, seven additional residues (N222, K240, L241, K244, V249, D253, and I257) also showed significant amide peak decreases. These seven additional residues cluster with the initial set of nine residues that were identified in NMR titration using free RT (Figure. 1B). The data suggests that inclusion of the DNA substrate in these NMR experiments results in a somewhat expanded binding surface on IN. In addition, the selective broadening of the peaks for these 16 amino acids indicates that binding of IN 220-270 to RT is specific, as non-specific binding would produce uniform changes across all of the peaks in the spectra. Finally, the signal decreases observed during the titration suggest that these 16 residues experience intermediate exchange on the NMR timescale (i.e. chemical exchange occurs at a frequency that is on the order of the difference between free and bound amide group chemical shift) (29, 30).

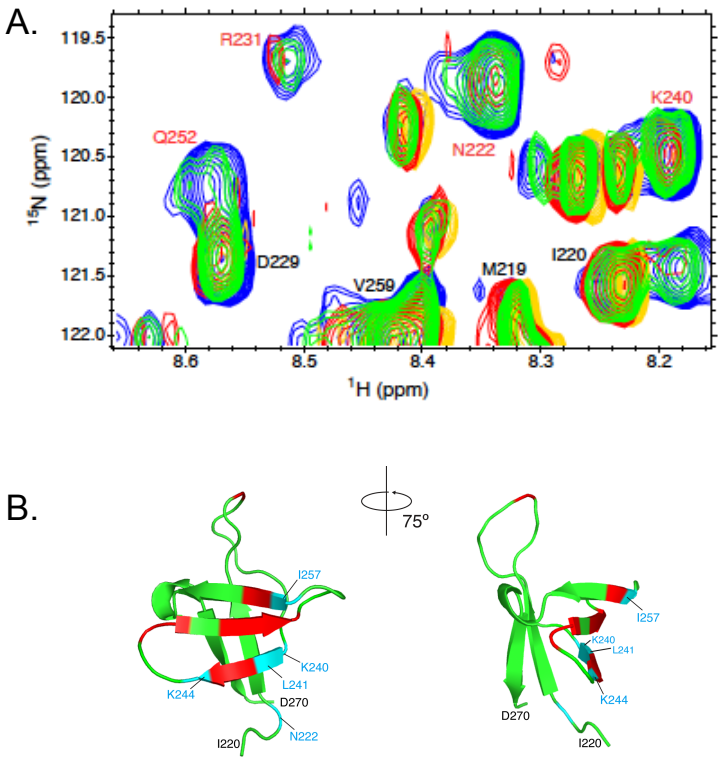


Figure 1. Mapping the amino acids with significant change in peak intensity to the IN 220-270 structure. A) Chemical shift perturbation measurements using ^{15}N-labeled IN 220-270 from the IN CTD and unlabeled RT pre-bound with DNA. The HSQC spectra show free IN 220-270 (blue) and IN 220-270 mixed with RT•DNA at molar ratios of 10:1 (green), 5:1 (red), and 2:1 (yellow). IN residues labeled in red show significant decreases in amide peak intensity upon addition of RT•DNA. B) The Apo-RT-binding surface on IN 220-270 previously identified[14] is shown in red, and the expanded set of 7 residues that are newly-identified in the current study are shown in blue. Collectively, these 16 residues form the RT•DNA-binding surface on IN 220-270.

Generation of viruses and analysis of replication kinetics

The functional relevance of this newly mapped surface on IN can be determined by evaluating the effect of perturbing the putative interacting surface upon viral replication. Since nine IN residues (R231, L242, W243, G247, A248, V250, I251, Q252 and K258) appear to

interact with RT with or without DNA, we elected to focus our mutagenesis efforts upon IN residues within this particular set. We have previously investigated the effects of introducing substitutions at two of the putative nine RT-binding residues, and found that IN substitutions W243E and V250E significantly impaired viral replication in tissue culture (23). The exact phase of the viral life cycle affected by these particular amino acid substitutions, as well as substitutions at other IN positions within this set of interest, are analyzed. We hypothesize that RT-IN interactions are necessary for reverse transcription during the viral life cycle, and that perturbing the RT-IN interaction will lead to a defect at the reverse transcription step. To investigate the functional relevance of the putative RT-binding surface on the IN C-terminal domain, we introduced single amino acid substitutions that alter surface polarity or electrostatic charge using site-directed mutagenesis.

We investigated the effects of these various substitutions upon viral replication kinetics, infecting CEM cells with equal amounts of wild-type virus or IN mutant viruses, as measured by p24 ELISA. Viral production was monitored by measuring p24 content in media every two days over ten days. Negative controls included a mock infection (media only without virus) and an infection using heat-inactivated wild-type virus. Of ten IN mutants tested (R231E, W243E, G247A, G247E, A248E, V250A, V250E, I251E, Q252A, and Q252E), six were replication-defective (Figure. 2). Cell syncytia were clearly evident after day six in cells infected with either wild-type virus or IN mutant G247A, and the time at which syncytia were observed correlated with a peak in p24 detected in culture media at day six, indicating wild-type replication kinetics for the IN G247A mutant (Figure. 2C). IN mutants V250A, Q252A, and Q252E exhibited decreased peak p24 levels relative to wild-type virus as quantified by p24 ELISA, but were above background levels (data not shown). Neither a single alanine nor a single glutamic acid at

position 252 (Q252A/E) was able to completely disrupt viral replication. We deemed these three IN mutants to be replication-competent and performed no further characterization of these variants. In contrast, IN mutants R231E, G247E, A248E, and I251E were non-viable. IN mutants W243E and V250E were also replication-defective in these experiments, as previously reported (23). We subsequently performed additional experiments to determine whether the block in replication for these six IN mutant viruses takes place during polyprotein processing and packaging, attachment and entry, or reverse transcription. We also tested whether recombinant IN proteins bearing single IN substitutions R231E, W243E, G247E, A248E, V250E, or I251E were catalytically-active in biochemical assays, and assessed the ability of these recombinant Ins to bind purified RT in pulldown assays.

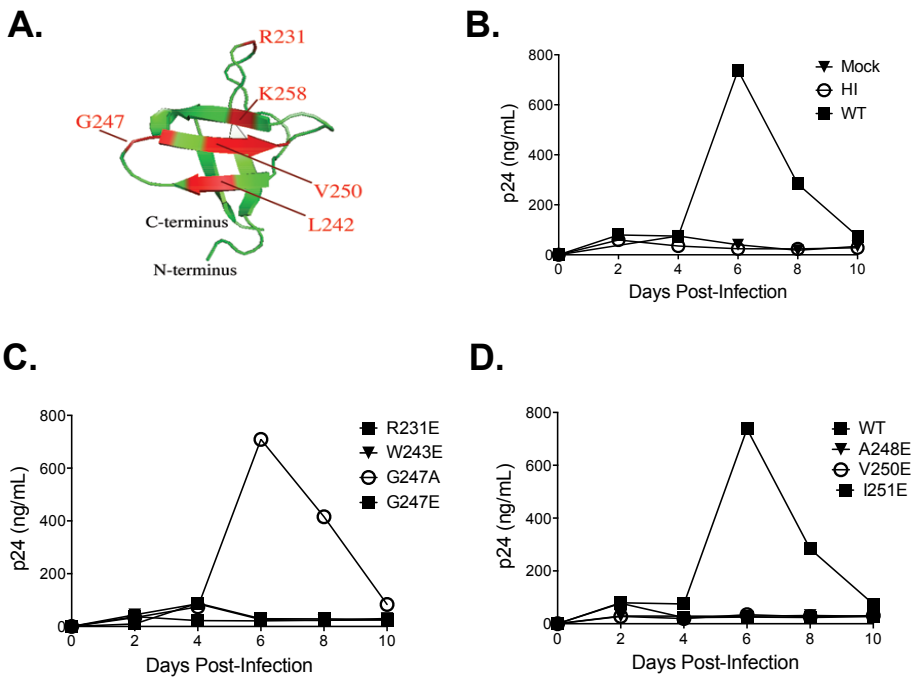


Figure 2. Replication kinetics of wild-type and IN mutant HIV-1 viral clones. A) Location of some of the IN residues targets in the IN 220-270. B-D) CEM cells were infected with 100 ng of p24-equivalent of the wild-type or respective IN mutant viruses and the culture media monitored for p24 production at the indicated time points post-infection. Compared to wild-type, six of the ten mutants exhibited complete defects in replication: R231E, W243E, G247E, A248E, V250E, I251E. IN mutants V250A, Q252A, Q252E exhibited decreased peak p24 levels but were still viable (data not shown). Data representative of three replicates is shown.

Analysis of polyprotein processing and packaging

Several IN mutants have been previously reported to be replication-defective due to defects at the processing step (31, 32). We examined the possible effects of the six IN mutations described above upon precursor processing and viral protein packaging by western blot analysis. One hundred nanograms of p24-equivalent mutant IN or wild-type virus were lysed with SDS-containing gel-loading buffer, separated on a 12% denaturing polyacrylamide gel, and transferred to nitrocellulose membranes. Blots were probed with either human anti-HIV serum to

detect all packaged viral proteins, or with antibodies specific for RT or IN. All IN mutants exhibited proper Gag-Pol processing and packaging of viral proteins comparable to wild-type control (Figure. 3A). When blots were probed to specifically detect either RT or IN packaged within virions, mutants W243E, V250E, and I251E showed a notable decrease in packaged IN (Figure. 3C). This decrease in packaged IN for these IN mutants may contribute in part to the replication-defective phenotype, but nonetheless may not fully explain the block in virus production.

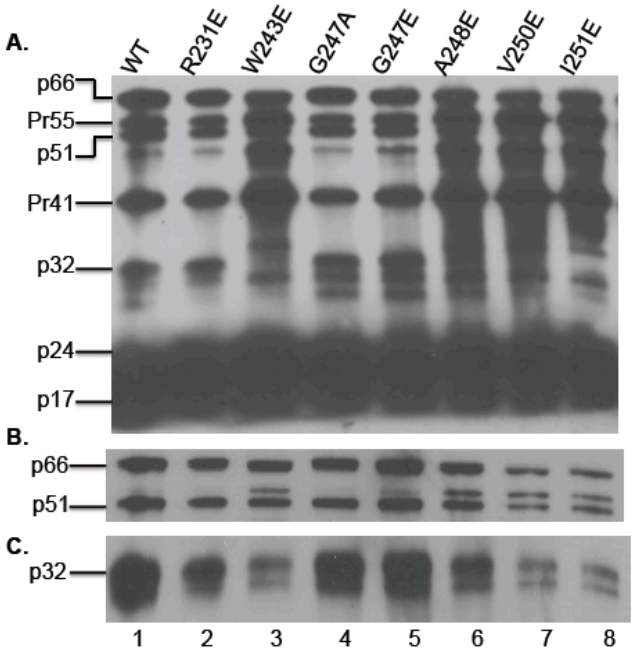


Figure 3. Western blot analysis of HIV-1 proteins packaged in wild-type and IN mutant viral particles and resolved on a 12% denaturing polyacrylamide gel. Each lane contains 100 ng of p24-equivalent of viral lysate. Blots were probed with A) human anti-HIV-1 serum, B) anti-RT and, C) anti-IN. Labels on the left correspond to positions of viral proteins with their molecular masses in kilodaltons; p66 and p51 correspond to the RT 66 and 51 subunits; Pr55, Gag-Pol precursor; p32, IN; p24, capsid; p17, matrix.

Replication-defective IN mutants are competent in viral attachment and entry

We determined the ability of these six replication-defective mutant viruses to attach to and enter cells. CEM cells were exposed to equal p24-equivalent amounts of either wild-type virus or each of the six IN mutant viruses. Following 4 h incubation with virus, cells were either treated with trypsin to facilitate subsequent detection of intracellular p24 levels in cell lysates, or simply washed and lysed without trypsin treatment to quantify total amounts of both cell-associated and intracellular p24. Negative controls included both mock-infected cells and cells exposed to heat-inactivated wild-type virus. We also included the replication-competent IN G247A mutant as a positive control. All of the IN mutants showed comparable amounts of cell-associated and intracellular p24 as wild-type virus in these assays (Figure. 4). These results show that all six replication-defective IN mutants have no defects in the attachment and entry steps of viral replications.

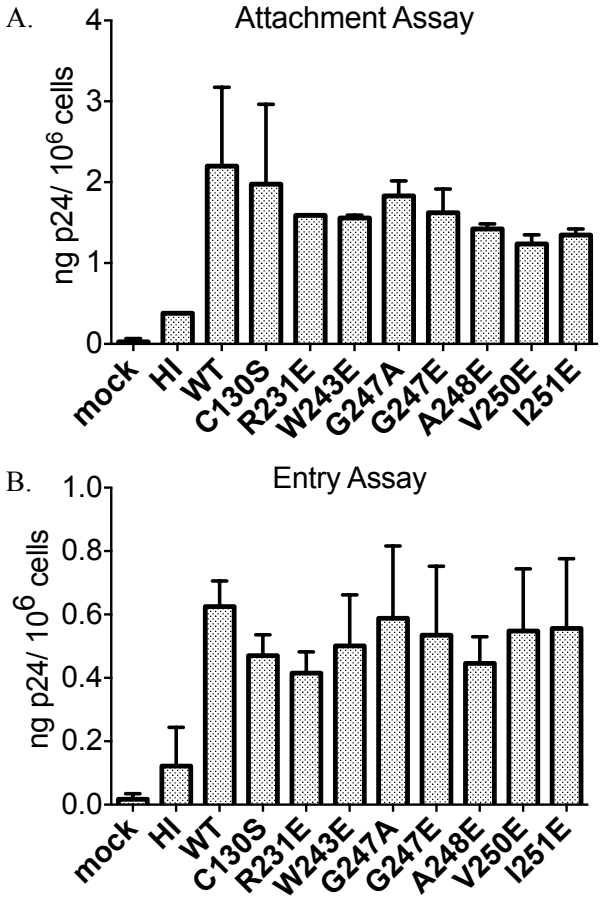


Figure 4. Attachment and entry cell assay analysis of replication-defective IN mutants. CEM cells were infected with 100 ng p24-equivalent of wild-type or IN mutant viruses and analyzed for A) viral attachment and B) entry as described in Materials and Methods. Treatment with heat-inactivated wild-type virus and mock infections using culture medium only were used as negative controls.

Replication-defective IN mutants are impaired in early viral cDNA synthesis

We explored the ability of these six replication-defective IN mutants to synthesize viral cDNA. Experiments using mock-infected cells and cells infected with heat-inactivated wild-type virus served as negative controls. Viruses were treated with DNase prior to infection assays to eliminate contamination from any residual NL4-3 transfection plasmid. Total DNA was

extracted from target cells 16 h post-infection and analyzed for early reverse transcription products using qPCR, as previously described (33). DNA input from each sample was normalized to levels of the human β-globin gene, which was amplified under identical conditions. All six replication-defective IN mutants were significantly impaired for early reverse transcription product synthesis, with decreases ranging from 75-100% of wild-type levels (Figure. 5A).

To account for the possibility of a defective RT resulting in the lack of reverse transcription products, we examined the endogenous reverse transcription activity in each of these IN mutant viruses. Equal amounts of detergent-lysed IN mutant virions were supplied with an exogenous template and radiolabelled nucleotides as substrates, and the amount of newly transcribed radioactive transcripts measured. All IN mutants packaged RT at comparable levels to that of wild-type virus (Fig. 3B). We observed no significant difference in RT activity in any of the mutant IN viruses relative to that found in wild-type virus, ruling out the possibility of a defective RT enzyme leading to the observed lack of reverse transcription products (Fig. 5B). These results indicate that IN substitutions at the RT-binding surface both adversely impact reverse transcription and ultimately lead to defects in viral replication without affecting endogenous reverse transcriptase activity.

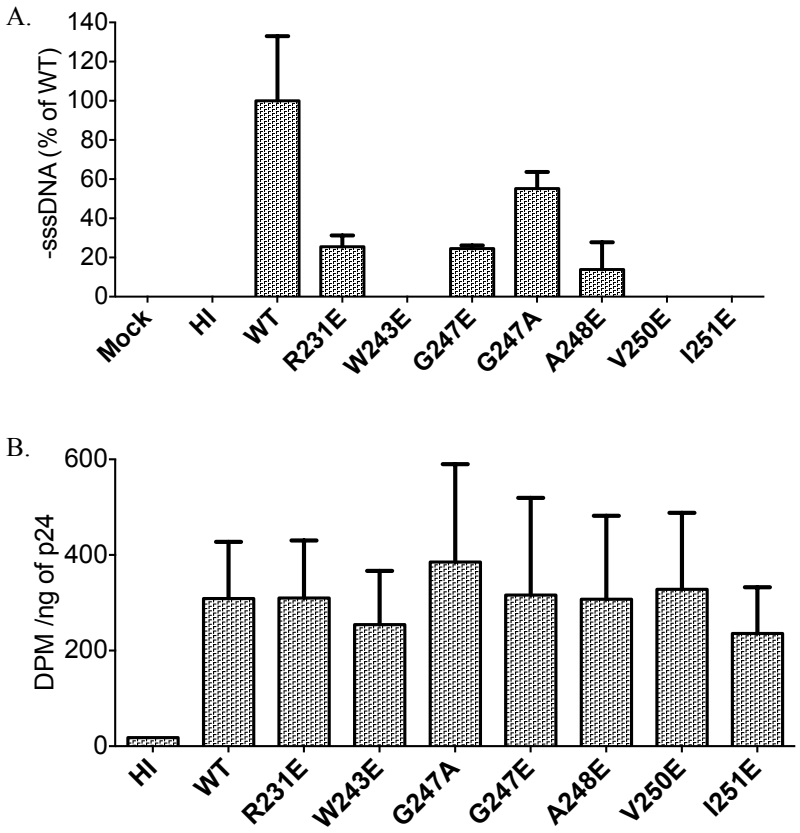


Figure 5. qPCR analysis of viral cDNA synthesis in cells infected with IN mutant viral clones. Graphs represent the mean of two replicates. A) Amount of early reverse-transcribed viral DNA present in infected cells was measured 16 h after infection. DNA input of each sample was normalized to the human β-globin gene amplified under identical conditions. B) Virus associated RT activity assay of wild-type and IN mutant viral lysate.

Recombinant mutant IN proteins retain *in vitro* catalytic activity

To investigate whether the introduced substitutions affect proper IN folding and activity, we assessed recombinant IN catalytic ability using IN activity assays for 3'-end processing and disintegration. The ability of mutant INs to catalyze the 3'-end processing and disintegration reactions was monitored by comparing amounts of DNA products produced in mutant IN reactions to that seen in wild-type IN reactions. All IN replication-defective mutant proteins

maintained *in vitro* catalytic activity similar to that of wild-type enzyme (Figure. 6). Since the substitutions at the IN CTD did not adversely affect catalysis, we infer that these substitutions do not perturb protein folding and IN dimerization required for proper 3'-end processing and disintegration.

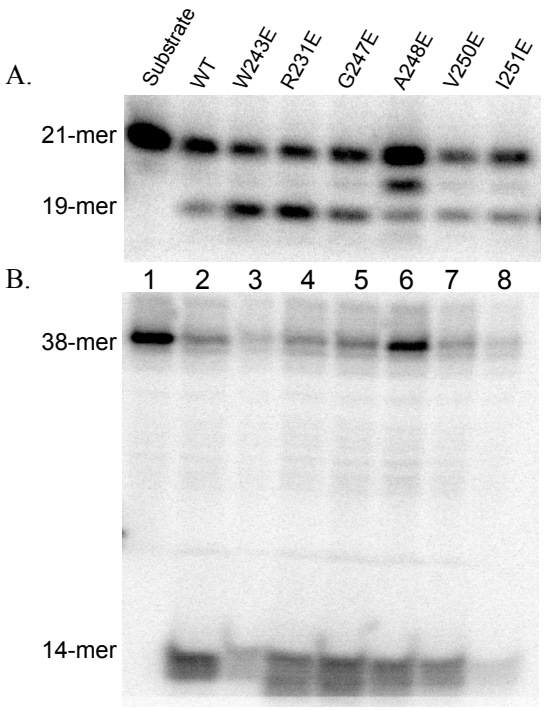


Figure 6. *In vitro* IN activity assays for wild-type and IN mutants A) 3'-end processing. 5'-end radiolabeled 21-mer substrate incubated with 40 pmol IN wild-type and mutant proteins to assess dinucleotide cleavage activity, resulting in a 19-mer product. B) Disintegration activity of 5' end labeled substrate catalyzed by wild-type and IN mutants. Lane 1: 38-mer substrate only control. IN wild-type and mutants were analyzed for their ability to resolve the viral sequence from the target DNA to form a 14-nucleotide labeled hairpin products and a 24-nucleotide circular molecule.

IN mutants show decreased binding affinity to RT

Targeted substitutions on the RT-binding surface of IN results in replication-defective viruses and a disruption of the viral cDNA synthesis step. To determine whether these amino acid substitutions can disrupt RT-IN complex formation, we applied two independent methods to characterize the ability of the recombinant mutant IN to bind RT. First, we employed a pulldown assay where we conjugated RT to a solid support consisting of magnetic beads (Dynabeads®) via amine coupling chemistry, and analyzed the ability of the mutated IN to bind immobilized RT. A previously characterized IN K258A substitution in the context of IN 220-270 shows a 8-fold decrease in RT binding compared to wild-type IN, as measured by SPR (23). The solid support without any conjugated RT served as a negative control to assess any nonspecific binding of IN to the beads, and no IN bound to this solid support under the conditions tested. The IN 220-270 constructs bearing each of the six substitutions showed relatively poor binding to RT. However, RT-binding defects associated with some of the substitutions in the context of 220-270 construct, such as W243E, G247E and A248E, were partly alleviated in the context of full-length IN (Figure. 7). These observations are consistent with previously published data showing that the C-terminal domain of IN is essential in binding to RT, but that other domains of IN may also play a role in the interaction (18).

To get a more quantitative RT-IN binding kinetics of the IN mutants, we utilized SPR analysis with RT protein immobilized on the chip gold-surface and IN protein flowed over the RT. Similar studies have been done by our group and others indicating a high affinity of wild-type IN binding to RT in the lower nanomolar range (23, 34). We used concentrations ranging from 0 nM to 500 nM of full length and 220-270 IN constructs, flowed over the RT-Immobilized surface and recorded the real-time association and dissociation. As a control, we also tested BSA

binding to the surface. We further analyzed the RT-binding residue G247 because depending on the substitution introduced, G247A and G247E, the virus was either replication competent or defective, respectively. Our SPR studies of G247E IN mutant exhibited a ten-fold decrease in binding affinity compared to wild-type (data not shown). When IN 220-270 constructs bearing the substitutions were used for the SPR, the inability of IN mutants to bind to RT was drastic, limiting further kinetics calculations.

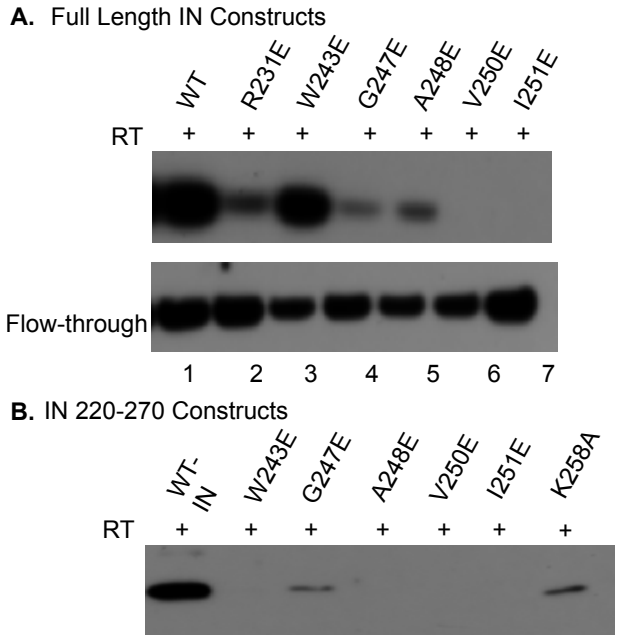


Figure 7. *In vitro* pulldown assay to determine the ability of IN wild-type and mutants to bind to RT. RT-coated magnetic beads (Invitrogen) were incubated with A) full length or, B) IN 220-270 wild-type and mutants, washed and eluted proteins were analyzed on a SDS-13% polyacrylamide gel probed with anti-IN. One control included IN 220-270/K258A, which has 10-fold lower affinity for RT as compared to wild-type.

2.4. Discussion

In this study, we determined the binding surface on IN of RT bound to DNA and investigated the biological significance of the complex formation in the context of viral replication. Building on our previous findings of RT-IN interaction (23), the NMR binding experiments with RT-DNA complex revealed seven additional residues encompassing the RT putative binding surface on IN. This findings further support published data that showed nucleic acid is not needed to bridge the RT-IN interaction (15), but rather IN acts to increase RT binding to vDNA (18). We introduced substitutions on the binding surface of IN and investigated the effect of these mutations on viral replication. Six of the nine IN mutants exhibited defects in replication compared to WT. Moreover, all six IN mutants exhibited significant decrease in copy number of viral cDNA compared to WT. Though the precise mechanism is still under investigation, we speculate that IN substitution disrupted the RT-DNA-IN complex formation required in the early steps of reverse transcription, resulting in short abortive transcripts rendering the virus replication incompetent. Our data confirms the *in vitro*-derived binding surface residues on IN and the biological significance of RT-IN interaction during viral replication, specifically at the reverse transcription step.

The structure of the free form of RT that is targeted by IN in this particular study can be inferred from past crystallographic data: crystal structures of free and nucleic acid-bound RT complexes show that the p66 subunit adopts an open conformation when bound to nucleic acid substrate, with the thumb subdomain rotating away from the fingers subdomain to expose a large cleft, while the apo-RT form assumes a closed conformation with the thumb domain filling this cleft (26, 35-37). RT lacking a bound nucleic acid substrate can impact IN activity, and has been previously shown (under various conditions) to either stimulate (22) or inhibit (34, 38) IN strand

transfer activity in biochemical assays, as well as inhibit IN 3'-processing activity (22, 34). In addition, other studies demonstrate that IN can interact with nucleic acid-bound RT to either promote RNA-dependent DNA polymerization (18, 24) or inhibit DNA-dependent DNA polymerization (38). Results from inhibition studies of DNA-dependent DNA polymerization by Hübscher and colleagues suggest that IN has greater affinity for RT that is bound to a dsDNA template-primer pair than for RT alone (38), implying that IN targets RT in the open conformation in this particular setting. We show data that supports a specific RT-binding surface on IN that is expandable when RT is prebound to vDNA.

Deletion analysis indicated that IN CTD is both necessary and sufficient for binding RT (15, 22). We focused our analysis on the nine amino acid residues that appeared in both titrations of free RT and vDNA prebound RT complex. Site directed point mutations of the nine putative binding residues on IN resulted to six out of nine, replication defective virus. Furthermore, these six replication-defective IN mutants were impaired at the early reverse transcription step. Purified recombinant IN mutants also had decreased RT binding affinity in the pull down assays performed. Our data suggest that a direct RT-IN interaction is required for efficient reverse transcription to occur.

IN has been shown to have pleiotropic tendency throughout the viral cycle (11, 39). Genetic studies have expanded the list of IN amino acid substitutions that impact viral cDNA synthesis, with the substitutions located at all three IN domains. IN exerts its effect on reverse transcription through multiple mechanisms. We investigated the major steps of viral replication and compared our results to published data. RT and IN are processed as a gag-pol precursor; some IN mutants have been reported to be defective at reverse transcription due to defects at maturation. Analysis of viral protein processing of packaged virions that were used for

subsequent infection, revealed decreased signal of packaged IN. Nevertheless, the decrease is not significant to explain the replication defective phenotype observed for W243E, V250E, and I251E, compared to previously reported polyprotein processing defective IN mutant (32). In contrast to the studied processing defect E246K IN mutant, we were able to pick up high levels of processed gag products such as capsid. Another well studied IN mutant, C130S, exhibited accelerated capsid disassembly resulting in unstable core phenotype. As a result, no reverse transcription products were detected (14). We evaluated core stability of two mutants, replication competent G247A and replication defective G247E, compared to WT. There was no discernable difference in the uncoating kinetics of both mutants compared to WT, narrowing the window of the observed reverse transcription defect past uncoating. Lastly, IN has been reported to interact with other viral proteins as well as host proteins to facilitate viral reverse transcription. One such example is the interaction of IN with host protein Gemin2/SIP1 (24). SIP1 is part of a complex that mediates the assembly of snRNPs. Masuda group showed that SIP1 knockdown abrogates reverse transcription, and that IN-SIP1 is a prerequisite for efficient viral infection. IN double mutant LL241, 242AA, lose their ability to bind to SIP1, and thus exhibit defects at reverse transcription. Though we are selectively targeting the putative mapped RT-binding surface on IN, we have not ruled out the possibility that our IN mutants might indirectly inhibit IN-SIP1 interaction.

The preferential multermeric state of IN required to bind to RT is unknown. IN forms a dimer to catalyze the 3'-end processing and requires a functional tetramer to catalyze the strand transfer step. Several studies investigating host protein LEDGF and IN interaction have shed light to the importance of IN ability to form higher order oligomers for proper core formation (40). *In vitro* IN activity assay carried out with recombinant IN replication-defective mutants

retained enzymatic activity similar to WT, indicating no significant structural defects to the IN protein attributed to the introduced mutation. These results eliminate secondary defects that might be anticipated with IN, collaborating on the specificity of the mapped RT binding surface on IN.

In summary, substitutions at the RT-binding surface on IN significantly impaired viral replication in tissue culture, specifically at the reverse transcription step. These results support a functional significance for RT-IN interaction during viral replication. In addition, considering the increasing viral resistance observed against many clinically relevant anti-HIV drugs that target the enzymatic active sites of both RT and IN, the insight gained here may help identify potential allosteric targets that may prove beneficial in antiviral drug design.

2.5. Materials and Methods

Construction of IN mutant viral clones and protein expression plasmids. Mutations within the IN coding sequence were introduced into a pBluescript cloning vector bearing the AgeI-EcoRI fragment from the infectious HIV-1 NL4-3 molecular clone (nucleotide positions 3486-5744) (41) using mutagenic primers and PCR-based site-directed mutagenesis (42). The IN gene containing the desired mutation was then excised from this cloning vector with AgeI and EcoRI and ligated into AgeI/EcoRI-digested NL4-3. Chemically competent Top10 cells (Invitrogen) were transformed by NL4-3 plasmids encoding the mutant IN sequences, and the resulting colonies were screened for desired constructs by DNA sequencing (Laragen Inc.). PCR-based site-directed mutagenesis (42) was also used to introduce mutations directly into either the pT7-7 (His) H-IN expression plasmid containing the full-length IN gene (43) or an expression plasmid encoding IN residues 220-270 (provided by Dr. Robert Craigie from the National Institutes of Health).

Protein Expression and Purification. CodonPlus *E. coli* cells (Agilent) were transformed by expression constructs encoding either full-length IN or IN 220-270 amino acid sequences having an N-terminal histidine tag for facilitating protein purification. All full-length IN and IN 220-270 constructs were purified under nondenaturing conditions as previously described, with few modifications (15). Briefly, transformed cells were grown in LB medium at 32°C until the optical density at 600 nm was between 0.8 and 1.0. Protein expression was then induced by adding 0.4 mM isopropyl-1-thio-β-D-galactopyranoside and left to grow for an additional 4 h. Pelleted cells were suspended in a lysis buffer containing 20 mM HEPES (pH 7.5), 5 mM 2-

mercaptoethanol, 1 M NaCl, 0.2 mM ethylenediaminetetraacetic acid (EDTA), 10% glycerol, 0.5% IGEPAL® CA-630 (Sigma-Aldrich), and EDTA-free protease inhibitor tablets (Roche; 1 tablet/10 mL lysis buffer), and further disrupted by sonication. Lysates were then clarified by centrifugation at 100,000 × *g* for 1 h at 4°C, followed by dialysis in buffer C (20 mM HEPES pH 7.5, 1 M NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, and 0.1% IGEPAL® CA-630). Purified IN constructs were then obtained from clarified lysates at > 95% homogeneity using Ni^{2+}-nitrilotriacetic acid immobilized metal affinity chromatography and cation exchange chromatography, as determined by SDS-PAGE analysis. For NMR studies, the IN 220-270 construct was expressed in CodonPlus *E. coli* cells at 37°C in minimal M9 media containing ^{15}N-labelled ammonium chloride. Purified recombinant RT heterodimer (p66/p51) was a kind gift from Stuart Le Grice (National Cancer Institute). Purified proteins were quantitated using extinction coefficients that were calculated based upon amino acid sequence.

Generation of viruses. Virus stocks were generated by transient transfection of 293T cells with NL4-3 constructs encoding the desired IN mutation. Cells were supplemented with 10% fetal bovine serum in RPMI 1640 with 10,000 IU/mL penicillin and 10,000 µg/mL streptomycin, and were transfected at 80-90% confluence in T-75 flasks using eight micrograms of each viral DNA construct according to a PolyFect®-based protocol (Qiagen). The medium-containing virus was collected 36-48 h post-transfection and gravity-filtered through 0.45 µm cellulose acetate filters. Harvested virus particles were concentrated through a 20% (w/v) sucrose cushion by ultracentrifugation at 100,000 × *g* for 3 h at 4°C, suspended in cellgro® sterile phosphate-buffered saline (PBS; Corning), and stored at -80°C. Viral titers were quantified by levels of HIV-1 p24 antigen, as determined by enzyme-linked immunoabsorbent assays (ELISA; PerkinElmer).

Western blot analysis of viruses. One hundred nanograms of p24-equivalent viruses were electrophoresed on 12% SDS-PAGE gels and transferred onto nitrocellulose membranes (Thermo Fisher). Membranes were blocked with 5% nonfat milk in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.05% Tween-20 (TBS-T) for 1 h, and then probed for 1 h with either a 1:500 dilution of anti-HIV human serum (Scripps Laboratory, Inc.), a 1:1000 dilution of a polyclonal anti-RT antibody, or a 1:500 dilution of a polyclonal anti-IN antibody in TBS-T containing 1% nonfat milk. Anti-RT and anti-IN antibodies were obtained from the AIDS Reagent Program at NIH. Blots were subsequently washed three times for 10 min each in TBS-T, and then incubated for 1 h with a 1:10,000 dilution of anti-human or anti-rabbit secondary antibodies conjugated with horseradish peroxidase (Santa Cruz Biotechnology, Inc.) in TBS-T with 1% nonfat milk. Detection of signal was accomplished using a chemiluminescence substrate kit (SuperSignal® West Pico; Pierce) according to manufacturer instructions.

Replication kinetics assays. Infection assays were carried out as previously described (23). Briefly, 100 ng of p24-equivalent viruses were used to infect 2×10^6 CEM cells, and virus-containing medium was carefully collected at 48 h intervals without disturbing the sedimented cells for 10 days post-infection. Harvested viruses were quantified by p24 antigen ELISA. Following virus collection at each time point, the infected cells were split 1:3. Negative controls included mock-infected cells (no virus added) as well as cells treated with heat-inactivated (HI) virus, produced by heating wild-type virions at 95°C for 10 min.

Viral attachment and entry assays. One hundred nanograms of p24-equivalent viruses were added to 2×10^6 CEM cells, which were then placed at 4°C for 30 min to synchronize attachment of the virus to the cell surface. Virus-exposed cells were then incubated at 37 °C for 4

h to allow infection to proceed, pelleted, and washed 3 times in 1 mL ice-cold sterile PBS to remove any loosely-bound virus from the cell surface. To quantify the extent of viral attachment, infected cells were disrupted in a lysis buffer consisting of PBS containing 1% Triton X-100 and 0.1 mM phenylmethylsulfonyl fluoride, and HIV-1 p24 levels in the resulting lysate were subsequently measured by ELISA. To quantify viral entry, cells were initially treated similar to the viral attachment assay, except that virus-exposed cells were washed only once in 1 mL of ice-cold sterile PBS following the 37°C incubation step to remove loosely-bound virus. The infected cells were then incubated at 37°C for 5 min in 0.5 mL of 0.25% trypsin (Invitrogen) to remove any stably attached virus. The cells were then washed twice more in PBS and then resuspended in lysis buffer prior to quantification of levels of viral entry by p24 ELISA detection. Data is represented as means of two replicates.

Real-time PCR analysis of reverse transcription products. Wild-type or mutant IN viruses collected from 293T producer cells were treated with 2 U/mL DNase I (New England BioLabs) in the presence of 10 mM $MgCl_2$ for 2 h at room temperature, and 2×10^6 CEM cells were infected with 100 ng of p24-equivalent DNase-treated virus. The infected cells were pelleted by centrifugation 16 h post-infection, washed twice in 5 mL PBS, and then lysed in a buffer containing 6.7 mM Tris-HCl (pH 8.0), 0.23 M NaCl, 0.67 mM EDTA, 4.7 M urea, and 1.3% (w/v) SDS. Total DNA from the lysed cells was obtained by phenol/chloroform extraction and ethanol precipitation, and subsequently used in Taqman quantitative real-time PCR (q-PCR) assays (Applied Biosystem) as previously described (33). The primer pair M667 (5'-GCTAACTAGGGAACCCACTG-3') and AA55 (5'-CTGCTAGAGATTTTCCACACTGAC-3') (44) that amplifies R-U5 region along with the probe ZXF (5'-FAM-TGTGACTCTGGTAACTAGAGATCCCTCAGACCC-TAMRA-3') (15) were used to measure

early reverse transcriptase product levels. A primer/probe pair BGF1 (5'-CAACCTCAAACAGACACCATG-3'), and BGR1 (5'-TCCACGTTCACCTTGCCC-3') that amplifies the β-globin gene, along with the probe BGX1 (5'-FAM-CTCCTGAGGAGAAGTCTGCCGTTACTGCC-TAMRA-3') were used to develop an internal standard for each DNA sample (15). All amplifications were performed in parallel with a set of linearized HIV-1 DNA standards of known concentrations. Viral DNA copy numbers were extrapolated from the standard curves and normalized to the internal control.

NMR sample preparation, data collection, and analysis.

Synthetic 18-mer (5'-GTCCCTGTTCGGGCGCCA-3') and 19-mer (5'-ATGGCGCCCGAACAGGGAC-3') DNA oligonucleotides with sequences corresponding to the 18 3'-terminal nucleotides of the $tRNA^{Lys3}$ primer and the viral RNA primer-binding site, respectively, were purchased from Integrated DNA Technologies. Concentrations of these oligodeoxynucleotides were determined using extinction coefficients that were provided by the manufacturer (158,700 and 185,600 $M^{-1}cm^{-1}$ for the 18-mer and 19-mer oligonucleotides, respectively). Equimolar amounts of the 19-mer and 18-mer were annealed to one another by heating to 90°C and slow-cooling in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 20 μM ethylenediaminetetrraacetic acid (EDTA) to form a DNA duplex with a single A-nucleotide overhang. Purified RT heterodimer and the 19mer/18mer DNA duplex were each dialyzed at 4°C in NMR buffer (50 mM sodium phosphate buffer, pH 6.5, 100 mM NaCl, and 0.5 mM EDTA), concentrated using centrifugal filter units (10,000 and 3,500 MWCO, respectively; Millipore), and combined in equimolar amounts to give a final RT-duplex DNA complex concentration of 200 μM.

The chemical shift perturbation experiment was performed by adding aliquots of this unlabeled RT-DNA complex to 150 μM ^{15}N-labeled IN 220-270 (concentration calculated assuming monomeric IN 220-270) in NMR buffer with 93% H_2O/7% D_2O. To facilitate IN 220-270 peak assignments, the same buffer and temperature conditions used in the NMR structure determination of IN 220-270 were also employed in our NMR experiments (45). An initial two-dimensional ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) spectrum (23) of free ^{15}N-labeled IN 220-270 was recorded at 25°C on a Bruker 800 MHz Avance spectrometer for the purpose of confirming peak assignments, which were made by comparison of peak positions with the NMR data archived at the Protein Data Bank (pdb 1IHV). Successive aliquots of RT-DNA complex were then added to IN 220-270 to produce IN 220-270:RT-19mer/18mer molar ratios of 10:1 and 5:1, and 2:1, with another ^{15}N-^{1}H HSQC spectrum being recorded after each addition. Molar ratios were calculated assuming monomeric amounts of IN 220-270. Spectra were examined and peak intensities were quantified using the program Sparky (46). Structural representations for IN 220-270 were produced using the program MacPyMOL (47).

Chapter 3: Alphascreen Assay Development for High Throughput Screen for Inhibitors of Reverse Transcriptase and Integrase Interaction

3.1. Abstract

Despite the arsenal of drugs available to treat HIV-1, there is emergence of drug resistant viral strains. This necessitates the need to identify novel targets to better aid in the development of new therapeutics. A promising approach to decrease the selection pressure for drug-resistant strains is to design inhibitors that target key protein-protein interactions in the HIV-1 life cycle. Reverse transcriptase and integrase are the viral enzymes responsible for catalyzing the essential steps of reverse transcription and integration, respectively. Our current research validates the functional importance of reverse transcriptase and integrase interaction during viral replication. We have developed a pharmacological approach to screen small molecule inhibitors against this interaction as a potential novel therapeutic target. An amplified luminescent proximity homogenous assay (ALPHA) screening assay was miniaturized to 10 µl final well volumes. Smart library compounds were screened at final concentrations of 10µM compounds. An initial pilot screen of 5440 compounds resulted in 34 validated hits that share distinct structure-activity relationships. Overall, we have developed the first high throughput screen for inhibitors of the reverse transcription and integrase interaction.

3.2. Introduction

HIV-1 infections still remain a global pandemic (1) despite its initial discovery in 1984 implicating HIV-1 as the etiology for AIDS. There are 26 FDA approved treatment targeting major steps of viral replication cycle (48) . Yet, still no cure is in sight due to the emerging resistant viral strains in treatment-exposed patients. The high turnover rate of the virus (49), low fidelity during reverse transcription step, and the low patient adherence to treatments, and viral latency are some compounding examples that are driving the high drug resistance viral strains. The alarming rate of resistance substantiates the importance of returning to the study of basic virology and discovering new therapeutic targets to supplement the current available treatments. Outstanding progress has been made towards gene therapy treatments; however, such types of therapy are still not well established and hence not readily available to patients (50).

Biochemical studies showing a functional relevance of the RT-IN interaction in Rous sarcoma virus, avian leukosis virus and human T-lymphotropic leukemia virus-type 1 have been reported, in which IN increased the activity of RT and facilitated viral nuclear localization (51, 52). In chapter two, we studied the biological significance of the HIV-1 RT-IN interaction in viral replication. We selectively substituted different IN residues that form the putative binding surface for both free and DNA-bound RT and characterized IN functional properties *in vitro* and in viral replication assays. We found that disruption of the putative binding surface on IN impaired viral replication due to the inability to synthesize viral cDNA, further indicating the functional significance of the RT-IN interaction during viral replication. The binding kinetics and affinity constant of the RT-IN interaction has been determined by surface plasmon resonance (SPR), which revealed a tight binding affinity (K_D = 61.2 nM) (23). *In vitro* transcription assays have shown that IN stimulates both the initiation and elongation modes of the RT-catalyzed

reverse transcription by increasing the proccessivity of RT and suppressing the formation of paused products (18). Collectively, we have accumulated proof of concept data that prompts inhibitor targeting RT-IN interaction as a novel potential anti-viral target.

A pharmacological approach to examine the biological significance and functional role of RT-IN interaction is necessary. An amplified luminescent proximity homogenous assay (ALPHA) was utilized to screen small molecule libraries for specific inhibitors against the RT-IN interaction. This method detects energy transfer between binding partners through a singlet oxygen that can diffuse up to 200 nm. In this ALPHA assay, 6X-Histidine-RT p66/55 binds to the Ni-chelate acceptor beads and Flag-IN binds to the anti-Flag donor beads. When RT and IN interact, the two beads are brought in close proximity and excitation of the donor beads results in generation of a signal (Figure 1). The goal is to screen large diverse libraries for compounds that will inhibit RT and IN interaction, indicative of decreased signal output. We have performed a pilot screen of over 5000 compounds and validated the hits obtained. The initial screen was done in singlets and all positive leads were subjected to quality control and activity verification in triplicate.

3.3. Results

Alphascreen assay

Lack of a full length crystal structure of IN limits our structural understanding of the RT-IN interaction complex. Therefore, we chose to perform an alphascreen assay over other fluorescents-based screening assays. Commonly used methods to study protein-protein interactions such as fluorescence resonance energy transfer (FRET), fluorescence polarization (FP), and ELISA have limitations when analyzing multiple-component protein complexes. ALPHA technology is less restrictive than FRET (200 nm versus 7 nm), has higher sensitivity requiring less (nanomolar) protein as compared to FP (micromolar), and complex formation is formed in solution whereas in ELISA one component is immobilized to a surface.

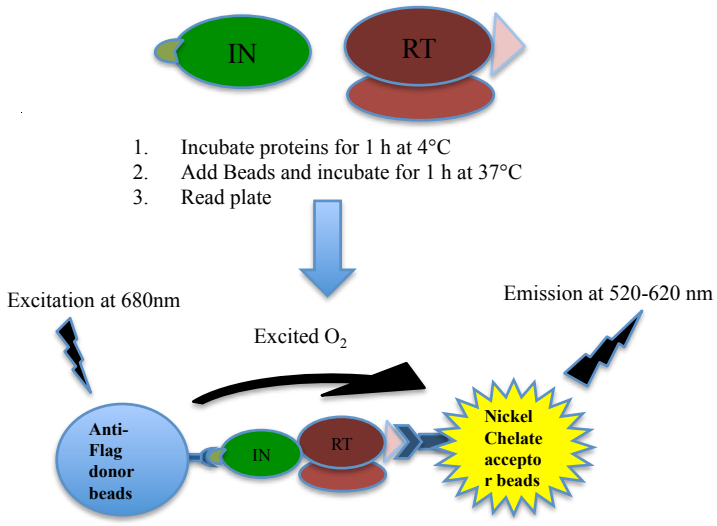


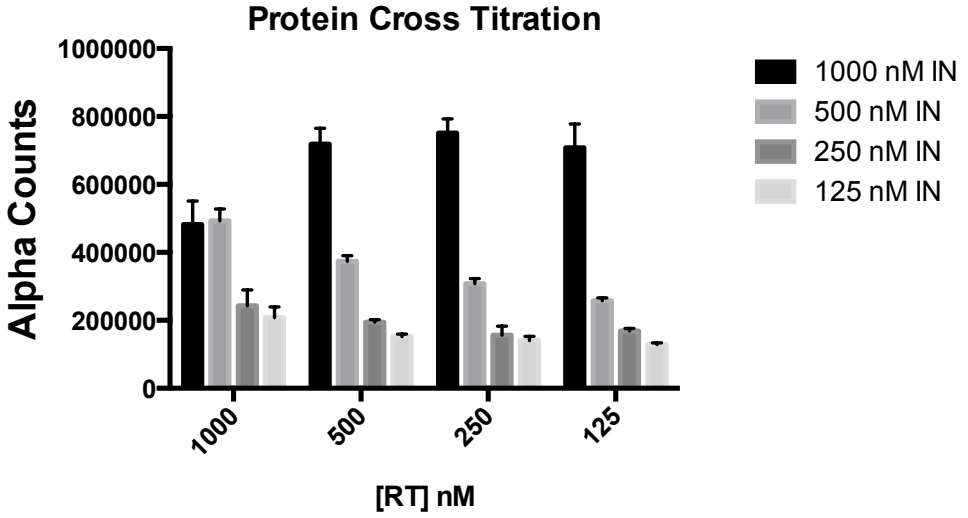
Figure 1. Initial alphascreen assay workflow.

Determining molar combinations of reverse transcriptase and integrase

Cross-titration of His-RT and IN-Flag concentrations needed to drive the equilibrium to the RT-IN complex was first determined. Initial cross-titrations were done in 25 µl reaction volumes ranging from 1000 nM to 125 nM (Figure. 2). All molar ratios, 1:1, 1:2, 1:4, 1:8 of RT to IN tested were optimal, generating signal to background ratios above 10. The quality of each experimental condition was assessed by calculating the Z' factor and choosing those within the optimal ranges of 0.5-1. The Z' factor is a statistical measure that takes into account the mean and standard deviations of both positive and negative controls, and gives a more accurate and reproducible estimate of the inhibitor screening window compared to looking at signal to background ratio. The quality of the assays in wells with the highest concentrations of proteins that showed higher signal to background ratios, did not correlate with the calculated Z' values. For example, 1000 nM RT: 250 nM IN had a signal/background levels of approximately 38 with the calculated Z' value of 0.36. Since we aim to screen a large library of compounds, we choose the lowest amount of protein combination that was in the acceptable screening range. We proceeded with concentrations of RT:IN of 125 nm with optimal signal/background ratio of 10 and Z' value of 0.8.

$$Z` = 1 - \left(\frac{3\sigma_{+c} + 3\sigma_{-c}}{| \mu_{-c} - \mu_{+c} |} \right)$$

Zhang, J-H, Chung, TDY and Oldenburg, KR, J. Biomol Screen (1999) 4: 67-73



	S/B, Z'- Values			
	1000 nM IN	**500 nM IN**	**250 nM IN**	**125 nM IN**
1000 nM RT	37, 0.5	38, 0.9	18, 0.4	16, 0.5
500 nM RT	55, 0.8	28, 0.8	15, 0.8	11, 0.8
250 nM RT	57, 0.8	23,0.7	12, 0.4	10, 0.6
125 nM RT	54, 0.7	19, 0.9	13, 0.8	10, 0.8

Figure 2: Determining the optimal molar combination of reverse transcriptase and integrase. A) Different molar combinations of both proteins were first incubated at room temperature for 1 h before beads added. The reaction was left to proceed for 1 h before plates were read. Result is represented as a mean of triplicate wells. B) Table summarizing the respective signal to background and Z' factor calculations for each combination tested.

Validation of screening assay

To validate that the signal output was due to specific RT and IN interaction, we performed a displacement assay using GST-IN. Titrations of GST-IN will compete with Flag-IN for binding to RT, therefore reducing the number of Flag-IN bound to His-RT. When Flag-acceptor and Nickel-donor beads are added to the complex, signal output is expected to decrease due to less His-RT:Flag-IN complex formed. Concentrations of His-RT and Flag-IN were kept constant at 125 nM, and concentrations of GST-IN ranging from 0 nm to 250 nM were titrated

into each well. The proteins were incubated together for one hour before beads were added and left to incubate for another hour before reading the plates. The alpha counts decreased with increase in GST-IN concentrations (Figure. 3). The result validates the specificity of the assay and confirms that the signal generated is specifically from RT-IN complex formation.

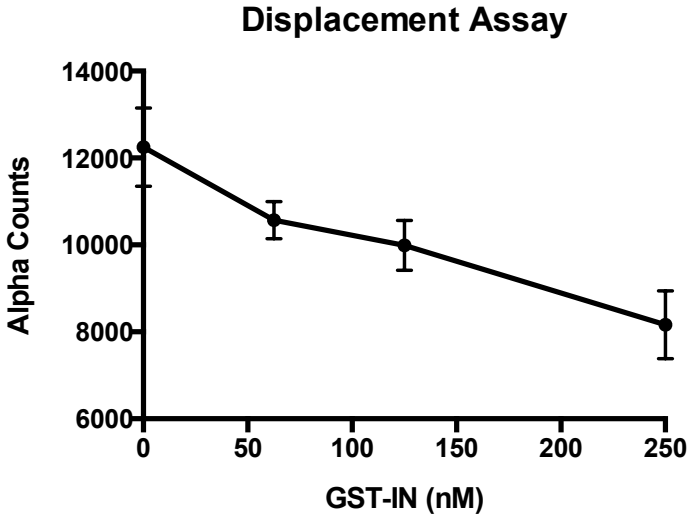


Figure 3. Alphascreen assay validation. Increasing concentrations of GST-tagged IN was titrated in to the optimal 125 nM combination RT:IN. Reactions were incubated for 1 h before beads were added, then followed by 1 h incubation at 37°C. Results are a representative of triplicate wells.

Order of addition

The initial assay workflow entailed a 1 h incubation of the proteins at 4°C, mainly due to IN tendency to precipitate. The second incubation of the protein complex with beads was at 37°C, as per manufacturer protocol. However, incubations of low volume reactions at 37°C can lead to evaporation thus increasing error per well. We examined different order of additions and

incubation temperatures (data not shown). A one step assay where all the components (RT, IN, bead combinations) were added at the same time, resulted in poor reproducibility of values. A two step assay where proteins were first incubated and allowed to form complexes and then beads were added resulted in optimal Z'-values. Incubation of the two-step assay was all done at room temperature. We concluded that a two-step assay with room temperature incubations would simplify automation of the screen and allow for miniaturization.

Assay Miniaturization

The cost of the ALPHA screen reagents in 25 µl reactions was $ 1.16/well. A pilot screen of at least 5000 compounds under these conditions would cost $ 5,810 for just the beads. Miniaturization of the assay will not only be cost effective but also lower the amount of proteins required. With the concentrations of RT and IN held constant (125 nM each), all components were tested at 25 µl or scaled down to 10 µl and 5 µl final reaction volumes. In addition, different concentrations of beads were tested in each reaction volume (Figure 4). We were able to scale down the assay to 5 µl reaction volumes and with lower bead concentrations. Though the Z' values obtained in 5 µl reactions (with fewer beads) was in the optimal range (0.5-1), we anticipated some upstream complications with evaporation when the screen is automated for HTS. We proceeded to further optimize the 10 µl reaction volumes with 25 ng/µl final bead concentrations, costing $ 0.29/well.

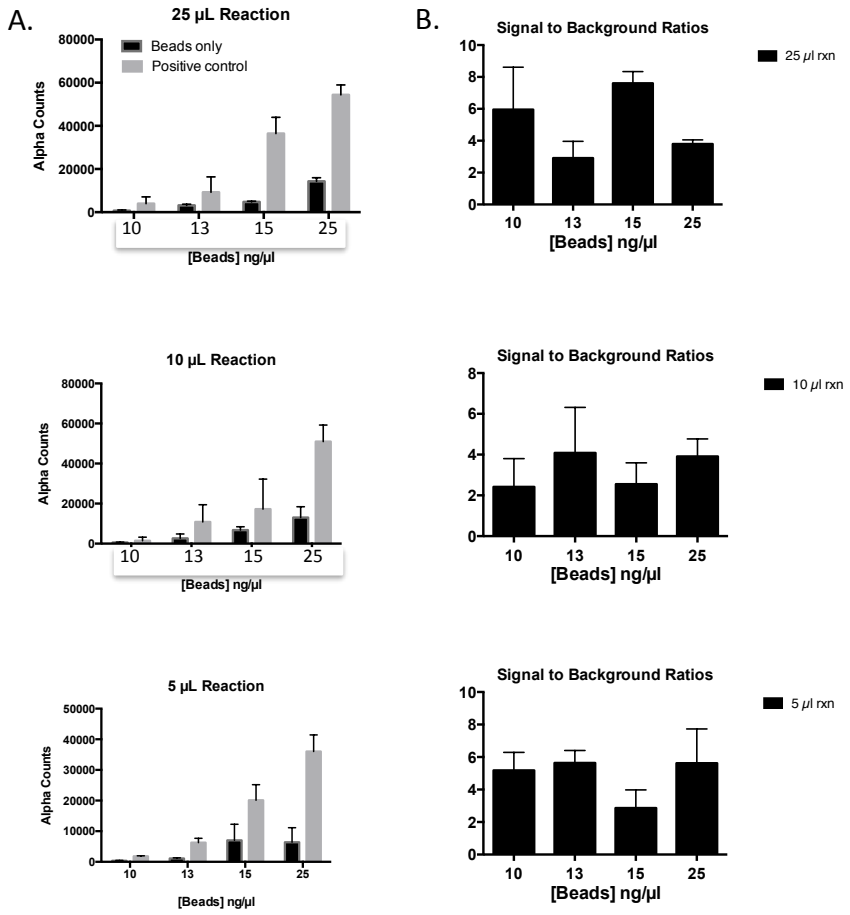


Figure 4. Alphascreen miniaturization and beads titration. A) Concentrations of RT and IN were kept constant at 125 nM each, and components scaled down proportionally to lower volumes. Different concentrations of 1:1 mix of donor and acceptor beads were titrated. Values are a representative of 5 replicate wells. B) Bar graph representation of signal to background values of each condition tested.

Screening buffer optimization

The low values of signal to background ratio in the 10 µl reactions prompted further optimization. So far, the screening buffer (25 mM Tris pH 7.4, 150 mM NaCl, 1mM $MgCl_2$, 1 mM DTT, 0.1% Tween 20 and 0.1% BSA) was adapted from a previously reported alphascreen assay for screening inhibitors of IN dimerization (53). Bovine serum albumin is commonly added in high throughput screening buffers to facilitate stability of proteins and reduce

adsorption of the reagents. However, the unspecific binding property of BSA has been shown to have negative effects due to compound masking (54). In our case, the background signals were very low as expected, but the positive controls were also low. We determined if addition of BSA in the screening buffer was masking the RT:IN complex formation and resulting in lower signal counts. Omitting only BSA from the screening buffer resulted in a 10-fold increase in the positive control signal output (Figure 5A).

The reverse transcription process has been extensively studied both biochemically and in the context of cell infection. The optimum salt concentration that facilitates the best enzymatic activity has been reported to be in the range of 50 to 150 mM NaCl. We intended to mimic the optimum enzymatic buffer condition for reverse transcription. We performed titrations of 25 mM to 100 mM NaCl in the screening buffer without BSA. Positive control wells were compared to beads-only negative control wells. There was a notable effect with increasing salt concentrations in the buffer (Figure 5B). Sodium chloride concentrations of 75 mM gave the best signal to background ratio of 300 and high Z' values > 0.8.

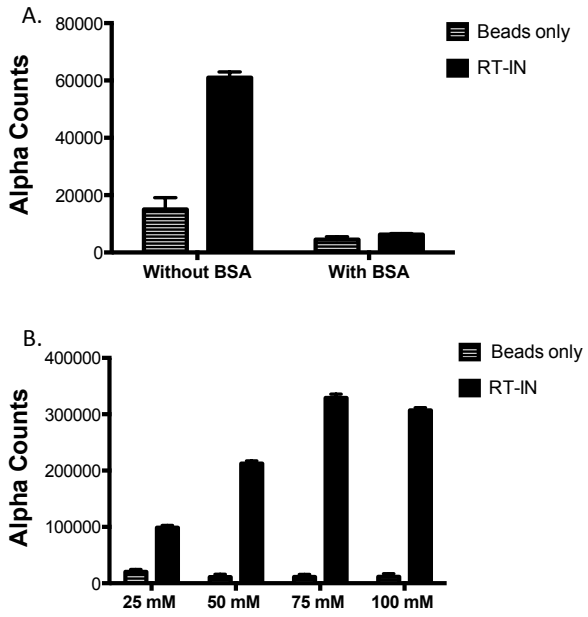


Figure 5. Screening buffer optimization. A) Initial screening buffer without BSA or supplemented with 0.1% BSA. B) Titration of NaCl in screening buffer without BSA.

Miniaturized protein concentrations determination

Moving forward with the final optimized buffer, we needed to determine the optimal molar combination of RT and IN. Cross titrating different concentrations of RT and IN was performed as previously done in 25 µl reactions, where we sought to obtain a combination that gave high signal to background ratio, Z' values > 0.6, and lower protein concentrations. We titrated concentrations ranging from 0 nM to 500 nM of RT and IN (Figure 6A). Extrapolating from these results, we were able to determine that the best combination was using 8 nM and 16

nM of RT:IN concentrations, respectively. This combination gives an optimal Z' value of 0.91, a wide signal/background window of 65, and is in the linear range (Figure 6B).

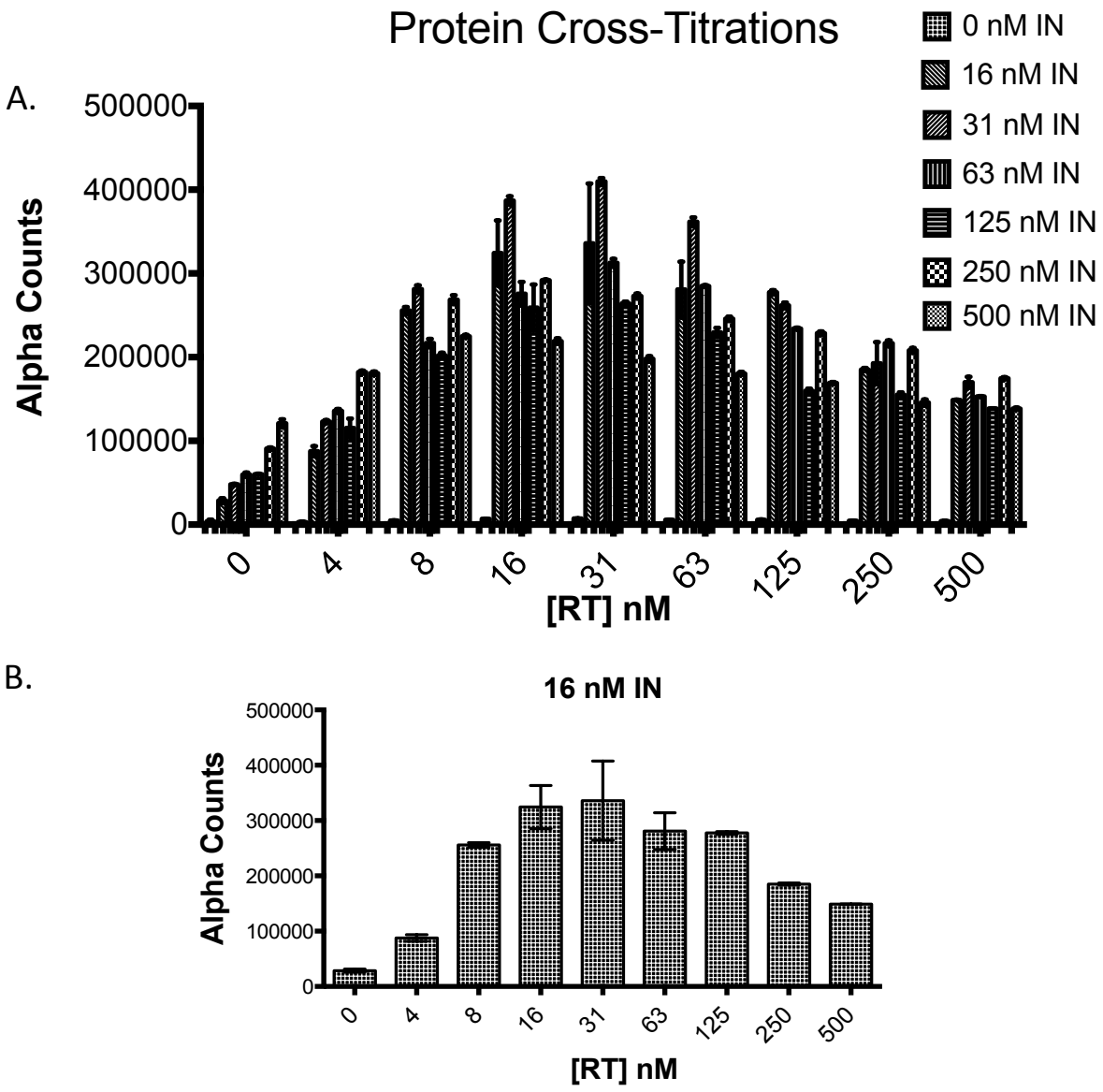


Figure 6. Final miniaturized optimal protein combination.

Time course of protein and bead incubation

The current assay flow entailed two 1 h incubation break points. We next assessed a time course of protein incubations before addition of beads. We did not observe a profound variation in signal output over the five time points evaluated (Figure 7A). Incubation times between 2 h and 3 h displayed the least signal change, hence be more suitable in limiting signal variations when handling a large number of plates. The second incubation time point of protein with beads was evaluated. The goal was to identify a screening time window that will have less signal change. Proteins were incubated for 2 h at room temperature, and then beads were added. Plates were read between 1h to 4.5 h post beads addition. Signal output increased with increasing incubation time (Figure 7B), with less signal change observed between 3.5 h to 4.5 h incubation. The 1 h window allows for up to 8 plates to be read in the alphascreen plate reader.

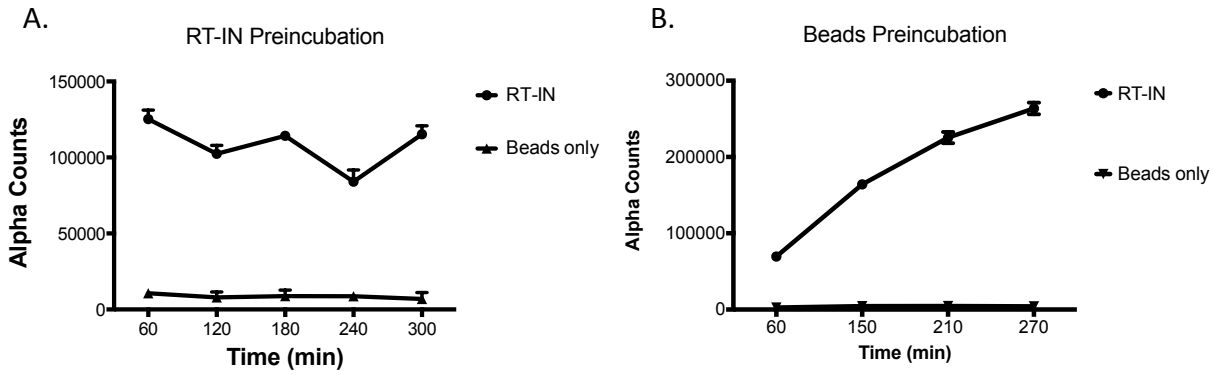


Figure 7. Assay workflow optimization. A) Optimized concentrations of RT:IN combination (8 nM: 16 nM) were incubated at room temperature at 1 h increments. Beads were added after each time point and incubated for 2 h before plates were read. B) Proteins were incubated for 2 h and beads added. Plates were read at the indicated time points after beads were added. All data are represented as average values of 5 replicate wells.

A schematic representative of the final workflow is outlined in Figure 8. We also performed a displacement assay, as described earlier, following the optimized workflow. Concentrations of His-RT and Flag-IN were kept constant at 8 nM and 16 nM, respectively, and concentrations of GST-IN ranging from 0 nM to 2 nM were titrated into each well. The proteins were incubated together for 2 h before beads were added and left to incubate for another 4 h before reading the plates. The alpha counts decreased with increase in GST-IN concentrations (Figure. 8B), validating that the miniaturization of the assay did not compromise assay integrity.

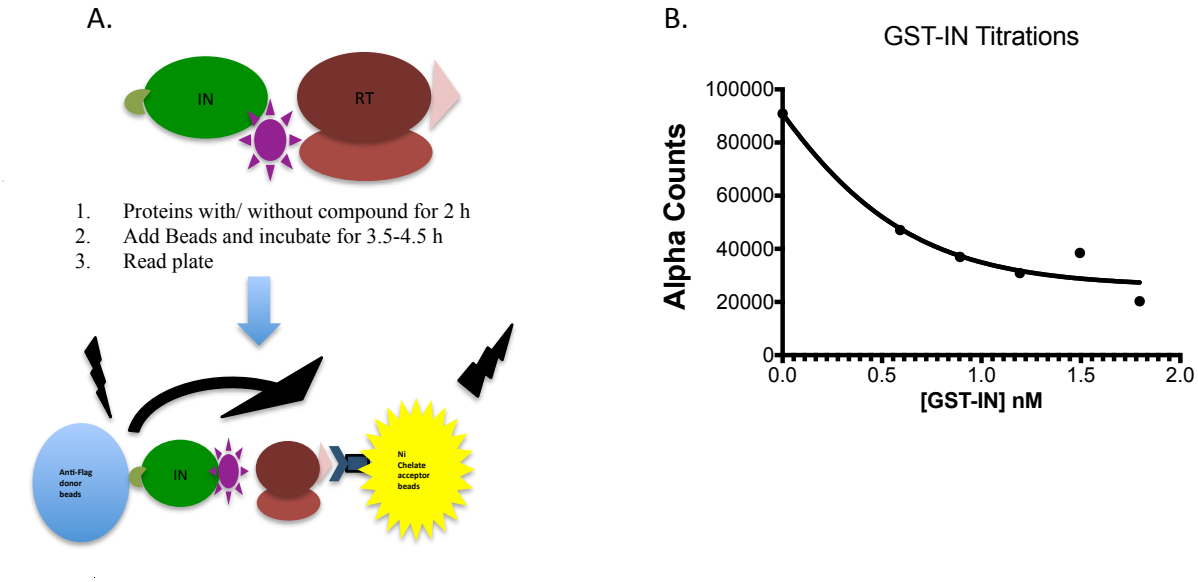


Figure 8. Final optimized screening assay and B) GST-IN displacement assay.

Counter-screening assay

Screening of compounds in assays that use tagged proteins have a tendency of accumulating false hits that target the binding of the tagged protein to their respective beads. This type of off target effects by compounds will also result in decreased signal output.

Additional false hits can arise with compounds that quench the fluorescence emitted thus also exhibiting decreased signal. We have constructed a double tagged IN protein that has a N-terminal Flag-tag and a C-terminal His tag (F-IN-H) for counter screening such off-target and false positive hits (Figure 9A). Compounds that are only specific for the RT-IN interaction should not exhibit any decrease in signal when incubated with the double tagged IN. A decrease of signal in this counter screen will indicate that the compound was an off-target hit (Figure 9B).

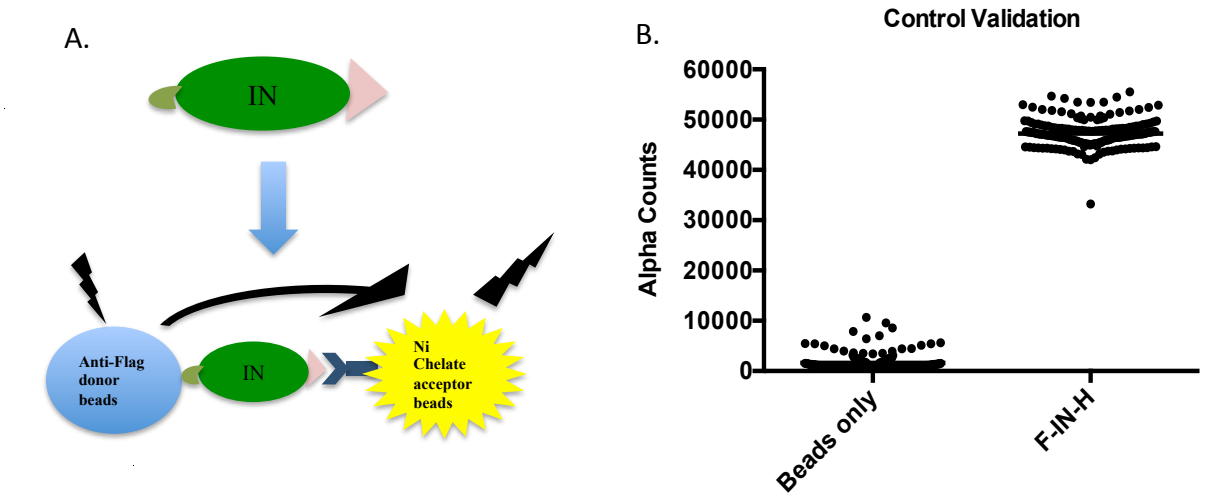


Figure 9. Counter screening assay. Double tagged Flag-Integrase-His protein is used to counter screen off target hits or quenchers. True hits are compounds that will not decrease signal output in this counter screen assay.

Pilot screening assay

Compounds were sampled from seven diverse libraries and screened at 10 µM concentrations. A total of 5540 structurally diverse compounds were incubated with proteins for 2 h then beads were added. Plates were read 3.5 h after addition of beads. A potential hit

compound was defined as any compound that reduced the Z-score (mean average signal output/plate) by 3 standard deviations. Figure 10 represents all the compounds screened and ± 3 STDEV cut off lines. Out of the 5440 compounds screened, 89 compounds met the first cut-off resulting in a 1.64% hit rate, which is acceptable for protein-protein inhibitors screens.

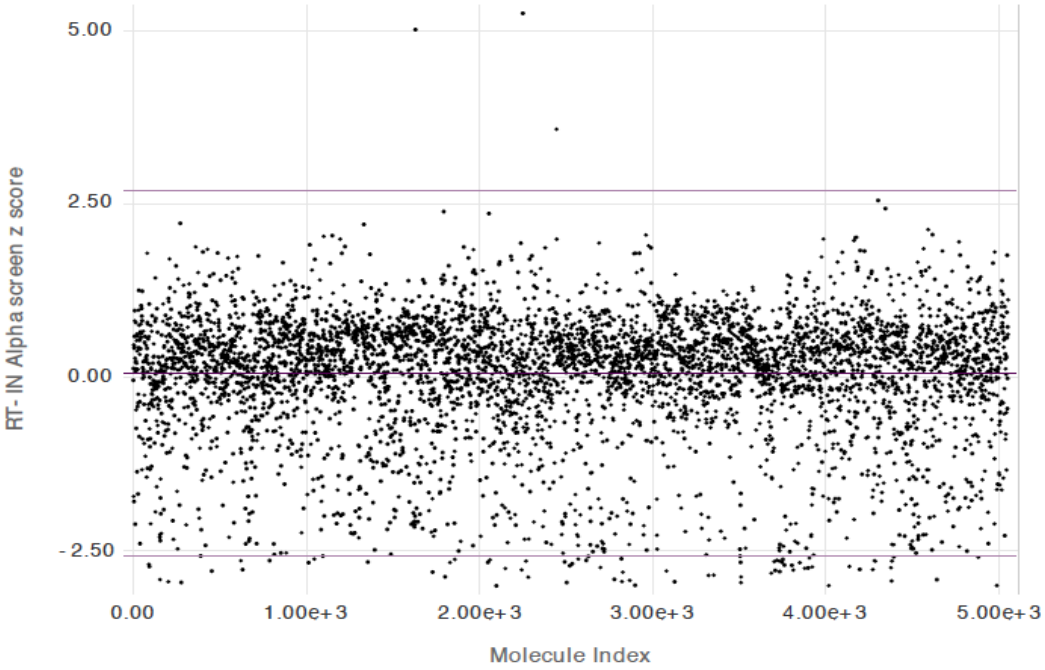


Figure 10. Pilot compound screen. Seven different libraries were sampled and screened in singlet. Potential hits were defined as compounds that exhibited 3 STDEV below mean Z score value of the plate. 89 compounds out of 5544 were selected for secondary screening, 1.6% hit rate.

We proceeded with selecting the best hits out of the 89, considering the errors introduced in certain well locations on the plate, commonly seen in the outer most wells. Secondary validation entailed repeating the same assay workflow but analyzing the hits in triplicates. Out of the 89 hits, we have validated 62 (Figure 11A). Most of the hits could be validated, indicating

the high reproducibility of this developed assay for screening compounds. Moreover, we counter screened for any off-target compounds and quenchers using our double tagged IN. Figure 11B summarizes the results from the counter screen. Overall, this screening assay had 54% reconfirmation rate, resulting in 34 lead compounds that passed the secondary validation and counter screening checkpoints.

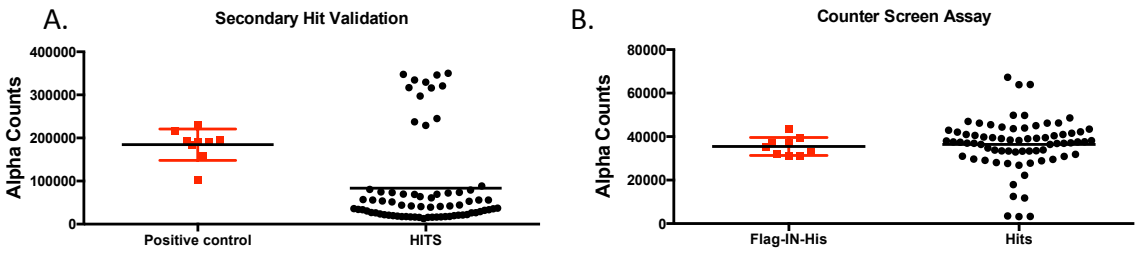


Figure 11. Hit validation. A) Sixty four potential hit compounds from the initial screen, B) and were also subjected to a counter screening assay again Flag-IN-His analyzed in triplicates. Positive control refers to wells with only RT and IN proteins in the presence of 1% DMSO. Data is represented as a scatter plot with each dot representing individual compound/well analyzed.

Structure-Activity Relationship

The 34 validated hit compounds were assessed for common structural-activity relationships (SAR). Commonly, screens result in random hits that do not share chemical functional groups. This is usually indicative of a poor screen set up or library selection. The SAR information acquired from our 34 compounds reveal a selection of distinct chemical moieties, indicating the selectivity of the assay. Interestingly, some major functional groups observed are also present in some FDA approved non-nucleoside RT inhibitors. We will use the hits to perform *in silico* screening of more than 250,000 compound libraries available at the MSSR core.

3.4. Discussion

We have developed a HTS assay for small molecule inhibitors of RT-IN interaction. Extensive proof of concept experiments has been reported showing that RT-IN directly interacts. We reported the biological significance of RT-IN interaction in viral replication is required for reverse transcription, a step prior to integration. The innovation of the presented study lies in the protein-protein target. Taking advantage of high sensitivity and low background noise properties of the alphascreen technology, we were able to miniaturize the final volume to 10 µl while maintain excellent assay reproducibility, Z' values of > 0.7 and high signal to background ratios. Additionally, we implemented a counter screening assay using double tagged IN, His-IN-Flag, enabling a high throughput false-positives detections platform of hits from the initial screen.

Alphascreen assay has been used to screen for inhibitor against other IN interacting proteins. Alphascreen technology is based on luminescent readout, where upon excitation of the donor bead ambient oxygen molecules are excited and can travel distances across 200 nm to a neighboring acceptor beads. Interaction of two proteins in solution brings the acceptor beads in close proximity to the donor beads, allowing for energy transfer and emission. In contrast to other florescent screening technology, alphascreen affords a wider distance range between protein-protein interactions increasing feasibility of targeting multicomplex protein interactions (55). The stoichiometry of multi-subunit RT and IN interaction is unknown. This limited structural information as hindered the pharmacological advancement of developing screening assay targeting RT-IN interaction. Another well established application of alphascreen drug discovery is targeting host protein LEDGF and IN interaction. Small molecule compounds hits from initial hits were refined and currently in clinical trials as a new class of allosteric IN inhibitor (56). Moreover, dimerization of IN as an essential oligomerization state for enzymatic

catalysis has also been targeted (53). However, though alphascreen technology was commonly applied, the experimental conditions are not universal. To our best knowledge, we are reporting the first developed screening assay that is specific to RT-IN interaction.

RT and IN are very attractive antiviral targets that have been extensively studied. The crystal structure of RT has attributed to the tremendous discovery of therapeutics currently in treatment. Two classes of RT inhibitors exists; nucleoside RT inhibitors (NRTI) mimic nucleic acid and bind the RT active site, while non-nucleoside RT inhibitor (NNRTI) indirectly inhibits RT activity by binding to and changing the structural conformation of the RT active site. NNRTI share similar pharmacophore and the binding pocket determined by crystallography structure of RT bound to the inhibitor (57). On the other hand, there are only three FDA approved IN inhibitors. Lack of full length structure of IN has stalled the advancement of therapeutics targeting integration steps. IN inhibitors have a carboxyl moieties that form hydrophilic bonds with IN and inhibit catalysis. A new class of allosteric IN inhibitors, ALLINIs, targeting IN dimerization, and host protein-IN interactions, are currently being studied.

Recent studies are advancing towards *in silico* designing of a dual inhibitor targeting both RT and IN. The idea is to add IN pharmacophore, the carboxyl group, to a NNRTI pharmocophore. Several potent components compounds have been reported (58). Our pilot screen of 5440, resulted in 31 validated hit compounds that directly targeted RT-In interaction. Structural relationship studies revealed a selection of amine-containing rings similar to NNRTI. We are currently assessing the anti-viral activity in tissue culture and EC_{50} values. All the compounds obey the rule of 5 (59), increasing their ability to enter the cell and have a higher bio-availability upstream of the drug development.

The evolution and improved efficacy of antiretroviral drugs has been instrumental in controlling HIV-1 infection. However, the high rate of mutations observed in the HIV genome render these drugs ineffectual after repeated clinical use due to the inevitable selection of drug-resistant viral strains. These findings will contribute to our understanding of the binding mechanism of RT-IN as well as provide a potential new platform for the RT-IN protein-protein interaction as a new anti-viral drug target.

3.5. Materials and Methods

Protein Expression and Purification. CodonPlus *E. coli* cells (Agilent) were transformed by expression constructs encoding full-length IN amino acid sequences having an N-terminal Flag-tag. The full length IN were purified under non-denaturing conditions as previously described, with few modifications (15). Briefly, transformed cells were grown in LB medium at 37 °C until the optical density at 600 nm was between 0.8 and 1.0. Protein expression was induced by adding 0.25 mM isopropyl-1-thio-β-D-galactopyranoside and growing the culture for an additional 4 h. Cells were pelleted and resuspended in a lysis buffer 1 (*20 mM Tris-HCL [pH 8.0], 100 mM NaCl, 7.5 mM CHAPS, 2 mM BME*) and further disrupted by sonication. Lysates were then clarified twice by centrifugation at 19,000 RPM for 45 minutes at 4 °C. The pellet was suspended in lysis buffer 2 (*20 mM Tris-HCL [pH 8.0], 1M NaCl, 7.5 mM CHAPS, 2 mM BME*) and protein pre-cleared by adding 0.8 M Ammonium sulfate to the supernatant containing solubilized IN. IN was then obtained from clarified lysates at > 95% homogeneity using Phenyl Sepharose column chromatography and a Heparin column for cation exchange chromatography. Protein purity was determined by SDS-PAGE analysis. Purified proteins were quantitated using extinction coefficients that were calculated based upon amino acid sequence. Purified recombinant RT heterodimer (p66/p51) was provided by Stuart Le Grice (National Cancer Institute).

Screening buffers and reagents. Alphascreen Ni^{2+}donor and Anti-Flag acceptor beads were obtained from PerkinElmer. Initial screening buffer was adapted from alphascreen assay designed for IN dimerization inhibitors screen and includes 25 mM Tris pH 7.4, 150 mM NaCl, 1mM $MgCl_2$, 1 mM DTT, 0.1% Tween 20 and 0.1% BSA (7). Final optimized screening buffer contains 20mM HEPES pH 7.5, 75mM NaCl, 2mM $MgCl_2$, 2mM DTT and 0.05% TWEEN-20.

All optimizations and screenings were done in 384-well white low-volume plates with lids (Corning®). BioTek EL 406 liquid dispenser with 1μL cassette was used to automate dispensing of proteins and beads suspension into each well. Biomek fx liquid handler was used to add 100 nL of compounds, in which the final compound concentrations were 10 μM in 1% DMSO. Biomex Fx tips were cleaned in after each use and in-between libraries. The standard cleaning flow of washes was water, DMSO, ethanol, methanol, and a final 60 seconds fan drying of the tip.

Library selection. Compound libraries were obtained from the UCLA Molecular Screening Shared Resource (MSSR) core. Seven different libraries were tested in the pilot screen reported here. Compounds from all seven libraries are considered smart libraries that result from only selecting drug-like compounds from over 250K compounds available at the core. Specifically, the compound libraries cover a wide chemical space and are selectively designed to have higher rate of cell entry, increasing chances of obtaining lead compounds.

Chapter 4: Biochemical Approach to Map the Integrase Binding Surface on Reverse Transcriptase

4.1. Abstract

Determining the IN binding region on RT and identification of the amino acid residues that facilitate the interaction will contribute to the full understanding of the binding mechanism. Moreover, screening for small molecules inhibitors of the RT-IN interaction resulted in several hit compounds that are currently under study for antiviral activity in tissue culture. Acquiring the knowledge of the binding surface on both proteins will aid in rational design of small molecule inhibitors and open a new window of *in silico* drug screening. To map the IN-interacting domain on RT, we are employing two different approaches: 1) co-crystallization of RT in complex with the IN-CTD, 2) high-resolution mass-spectrometry (MS) foot-printing technique that compares chemically-modified surface-exposed lysine residues of unbound RT to that of the RT-IN complex.

4.2. Introduction

Retroviruses require the reverse transcription of their RNA genome into viral cDNA before integration into a host chromosome to establish a productive infection. Reverse transcriptase (RT) and integrase (IN), present in all retroviruses, are the viral enzymes responsible for catalyzing the essential steps of reverse transcription and integration, respectively. RT-IN interaction has been observed in Rous sarcoma virus, avian leukosis virus and human T-lymphotropic leukemia virus type 1 where they are part of a single polypeptide forming a heterodimer (1,2). In HIV-1, both RT and IN are synthesized as part of the Gag-Pol polyprotein that is processed by the viral protease to produce active RT and IN during viral maturation.

Biochemical evidence supports a direct RT-IN interaction using recombinant proteins. Others and we have shown that HIV-1 IN physically interacts with RT and helps facilitate the early steps of reverse transcription, initiation and elongation, *in vitro* (3). In addition, the binding kinetics and rate constants, as well as nine amino acid residues in IN that form the RT-binding surface, have been determined by nuclear magnetic resonance (NMR) spectroscopy and surface plasmon resonance (SPR)-based optical biosensor (4). In chapter two, we applied site-directed mutagenesis to characterize the NMR mapped RT-binding surface on IN. We found that substitutions on the RT-binding surface on IN perturbed the RT-IN interaction and resulted in replication-defective virus. The results strongly support the biological significance of RT-IN interaction in the context of viral replication. Analysis of the replication steps revealed that mutation of the RT-binding surface of IN mainly caused defects in the early reverse transcription stage.

A glutathione S-transferase (GST) based pull-down approach was used to map the binding domain of IN on individual subunits of RT. The results showed a bipartite binding region, which

included residues 1-242 in the finger-palm domain of both the p66 and the p51 subunits of RT, as well as residues 387-422 in the connection domain (14, 22). However, since the monomeric p51 subunit folds similarly to the monomeric p66 subunit (60), these results might not reflect the true binding parameters as found in the native heterodimeric RT. In addition, the particular amino acids involved in the RT-IN interaction are still unknown.

4.3. Results and Discussion

Co-crystallization of RT-IN 220-270 Complex Buffer Screen

Crystallization is a two-step process. First the protein solute is slowly concentrated from its buffer and creates stable nuclei, a process called nucleation. The stable nucleus then serves as a building block for crystal growth. Nucleation and the stability of the nuclei are highly depended on the operating conditions such as the precipitating agent, pH, temperature and the purity of the protein. Several groups have solved diffraction quality crystals of both RT bound to various non-nucleoside reverse transcription inhibitors (NNRTI) (61-63), and unliganded RT p66/p55 with detailed crystallography methodology outlined (2, 3). However, full length IN has been difficult to crystallize due to its poor solubility at high concentration. A co-domain crystal structure of IN Core-C terminal domain has been resolved to a 2.8 A (64), only after several point mutations were introduced to reduce IN hydrophobicity and oxidation. Recently, a retrovirus full length IN belonging to Proto foamy virus (PFV) was crystallized (65). PFV IN shares 15% sequence homology and the conserved three-domain structure present in HIV-1 IN. The advancements in PFV IN crystallography studies provide new insights on the possible structure that HIV-1 IN may adopt.

We performed three high throughput sparse matrix screens in a 96 well format with only two overlapping conditions. We screened the protein complex suite (Qiagen) based on a previous report (4) to optimize the crystallization conditions optimal for protein-protein complexes. This screen employs different molecular weight polyethylene glycols as the main precipitant with varying pH, salt concentrations, as well as inorganic solvents such as isopropanol and ethanol. In contrast, we screened the ammonium sulfate AmSO4 suite (Qiagen), which tests different concentrations of AmSO4 as the main precipitant with various salts, buffers and pH conditions to

promote nucleation and crystal growth. In addition, the AmSO4 suite screen had several solutions with co-precipitants, salts and buffers conditions derived Biological Macromolecule Crystallization Database that proved to be successful for other proteins. Lastly, we chose a JCSG-plus screen devised at the joint center for structural genomics to broaden crystallization conditions. The JCSG-plus screen also includes a range of neutralized organic acids as the precipitants and a wider pH range.

Most of the Protein Complex suite conditions resulted in under-saturated precipitation; the AmSO4 screening conditions resulted in supersaturated precipitation whereas the JCSG-plus gave a mixture. We observed some crystal-like growth in the AmSO4 screen but we need to distinguish salts crystals from complex crystals by UV-rays, which are absorbed by tryptophan residues in the complex. Nonetheless, the observations from the three screens can be optimized further using the phase diagram for crystallization to decrease super saturation to a metastable or liable condition where nucleation can occur spontaneously in order to facilitate crystal growth.

Temperature is a critical variable condition that affects crystal growth. Since all these screens were done at 4 °C, due to the tendency of IN to precipitate at room temperature, we plan on varying the temperature in our next screen. In addition, we will investigate the appropriate molar ratios of RT:IN220-270 that forms stable complex at equilibrium using non-denaturing PAGE gels.

High-resolution MS foot-printing method to map the IN-binding surface on RT

Although NMR was used to map the putative binding residues on IN, the same approach cannot be done to map the IN-binding surface on RT due to the high molecular weight (117 kDa) of RT prohibiting a similar analysis by NMR spectra. We performed a high-resolution mass

spectrometry foot-printing study to map the IN binding surface on RT. This approach has been successfully applied to map the IN binding domain of the cellular cofactor lens epithelium derived growth factor (LEDGF)/p75 (16), as well as to identify the specific RT contacts with the viral RNA: tRNA complex (11). This method compares accessibility of surface exposed residues in the free protein versus that of the IN- bound form using a selective small chemical reagent specific for modifying lysine side-chains (Figure 1). Since lysine is the most abundant residue in RT, solvent-accessible lysine residues in free RT and in the RT-IN complex will be modified using the amine-specific reagent N-hydroxysuccinimidobiotin (NHS-biotin). The NHS-biotin reagent is a reversible crosslinker that adds a biotin tag to solvent-accessible lysine. The use of a reversible crosslinker will enable separation of individual components of the RT-IN complex by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for a more detailed MS analysis.

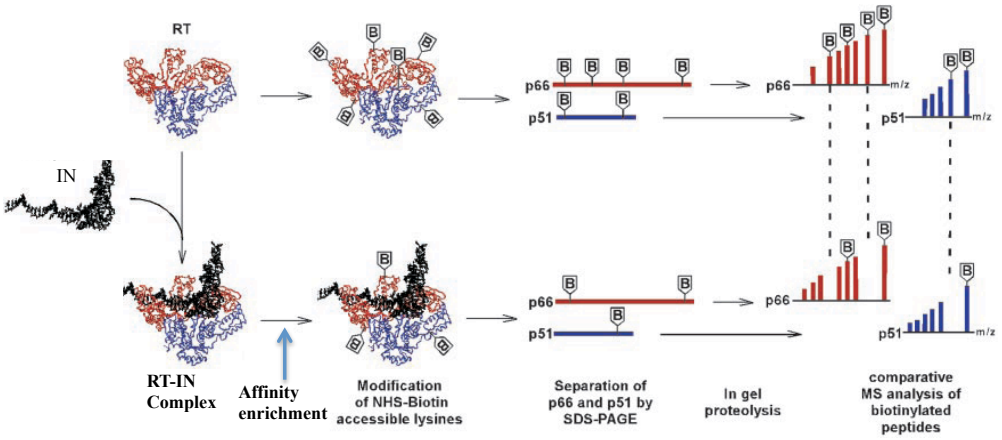


Figure 1. Process for mapping the IN-binding surface on RT.
Schematic adapted from Kvaratskhelia, M., *et al.* 2002. PNAS.

One major limitation of using mass spectrometry to detect residue modifications is that the modified residues are usually of low abundance and might be masked by background noise. To circumvent this problem, we modified a published protocol and included an enrichment step. We performed a pilot optimization experiment to ensure the enrichment of the RT-IN complex prior to any residue modifications. Affinity enrichment of RT p66/p51 by GST-IN pull downs was evident by western blot analysis. Silver stained gels showed three bands representing RT p66, p51 and GST-IN p55. These bands were excised and subjected to in-gel trypsin proteolysis prior to MS analysis. The sequence coverage was obtained was low and might be attributed to low levels of protein pulldown.

Table 1. Mass spectrometry analysis of RT from GST-IN pulldown experiment.

RT peptide	Sequence coverage	Sequence count	Spectral count
p66	26.30%	23	37
p51	36.60%	48	296

4.4. Materials and Methods

Protein Expression and Purification. CodonPlus *E. coli* cells (Agilent) were transformed by expression constructs encoding either full-length IN or IN 220-270 amino acid sequences having an N-terminal histidine tag for facilitating protein purification. All full-length IN and IN 220-270 constructs were purified under nondenaturing conditions as previously described, with few modifications (15). Briefly, transformed cells were grown in LB medium at 32 °C until the optical density at 600 nm was between 0.8 and 1.0. Protein expression was then induced by adding 0.4 mM isopropyl-1-thio-β-D-galactopyranoside and left to grow for an additional 4 h. Pelleted cells were resuspended in a lysis buffer containing 20 mM HEPES (pH 7.5), 5 mM 2-mercaptoethanol, 1 M NaCl, 0.2 mM ethylenediaminetetraacetic acid (EDTA), 10% glycerol, 0.5% IGEPAL® CA-630 (Sigma-Aldrich), and EDTA-free protease inhibitor tablets (Roche; 1 tablet/10 mL lysis buffer), and further disrupted by sonication. Lysates were then clarified by centrifugation at 100,000 × *g* for 1 h at 4°C, followed by dialysis in buffer C (20 mM HEPES pH 7.5, 1 M NaCl, 10% glycerol, 5 mM 2-mercaptoethanol, and 0.1% IGEPAL® CA-630). Purified IN constructs were then obtained from clarified lysates at > 95% homogeneity using Ni^{2+}-nitrilotriacetic acid immobilized metal affinity chromatography and cation exchange chromatography, as determined by SDS-PAGE analysis. For NMR studies, the IN 220-270 construct was expressed in CodonPlus *E. coli* cells at 37 °C in minimal M9 media containing ^{15}N-labelled ammonium chloride. Purified recombinant RT heterodimer (p66/p51) was a kind gift from Stuart Le Grice (National Cancer Institute). Purified proteins were quantitated using extinction coefficients that were calculated based upon amino acid sequence.

Co-crystal screening. The Protein Complex suite (Qiagen), the ammonium sulfate $AmSO_4$ suite (Qiagen), and the JCSG-plus. Co-crystals of equal molar ratio of RT:IN 220-270 are grown by

batch crystallization vapor diffusion method in hanging drops at 4°C. For each individual well, we employed a three-drop pattern (1:1, 1:2, 2:1) that varies the concentration of protein solution to crystallization reagent to increase the chances of getting crystals.

Affinity pulldown for RT and IN complex. Equal molar ratios of His-RT p66/p51 and GST-IN recombinant proteins were incubate in 200 µl of binding buffer (50 mM HEPES pH 7.4, 400 mM NaCl, 2 mM MgCl2, and 0.1% NP-40) on ice for 1 h. Pre-cleared 20 µl of GST-beads was then added and the reaction incubated end-over-end for 2 h at room temperature. The beads were then spun down at 5000 X g for 30 sec, washed three times with binding buffer before elution with 15 µl 0.1 M Glycine-HCl pH 3.5. The elution was neutralized with 1 M Tris pH 9 and subjected to 10% SDS-PAGE analysis. Bands corresponding to p66 and p51 were excised for mass spectrometry analysis.

Mass spectrometry foot printing of RT-IN complex. The excised gel slices were digested with trypsin, fractionated online using a C18 reversed phase column, and analyzed by MS/MS on a Thermofisher LTQ- Orbitrap XL as previously described (1,2). MS/MS spectra were subsequently analyzed using the ProLuCID and DTASelect algorithms (3,4).

Chapter 5: Summary and Conclusions

The goal of this dissertation was to characterize the biological significance of HIV-1 reverse transcriptase (RT) and integrase (IN) interaction during viral replication, and to assess if the RT-IN interaction can be a target for small molecule screening. RT-IN interaction has been observed in Rous sarcoma virus, avian leukosis virus and human T-lymphotropic leukemia virus type 1 where they are part of a single polypeptide forming a heterodimer (1,2). In HIV-1, both RT and IN are synthesized as part of the Gag-Pol polyprotein that is processed by the viral protease to produce active RT and IN during viral maturation. The binding kinetics and affinity constant of the RT-IN interaction has been determined by surface plasmon resonance (SPR), which revealed a tight binding affinity (K_D = 61.2 nM) (23). *In vitro* transcription assays have shown that IN stimulates both the initiation and elongation modes of the RT-catalyzed reverse transcription by increasing the proccessivity of RT and suppressing the formation of paused products (18).

In chapter two, we determined the binding surface on IN of RT bound to DNA and investigated the biological significance of the complex formation in the context of viral replication. Building on our previous findings of RT-IN interaction, the NMR binding experiments with RT-DNA complex revealed seven additional residues encompassing the RT putative binding surface on IN. We introduced substitutions on the binding surface of IN and investigated the effect of these mutations on viral replication. Six of the nine IN mutants exhibited defects in replication compared to WT. Moreover, all six IN mutants exhibited a significant decrease in copy number of viral cDNA compared to WT. We focused our analysis on the nine amino acid residues that appeared in both titrations of free RT and vDNA prebound to RT complex. Mutating six of the nine putative binding residues on IN resulted in replication defects. Furthermore, these six replication-defective IN mutants were impaired at the early reverse transcription step (minus strand strong stop DNA). The preferential multermeric state of

IN required to bind to RT is unknown. IN forms a dimer to catalyze the 3'-end processing and requires a functional tetramer to catalyze the strand transfer step. *In vitro* IN activity assayes carried out with recombinant IN replication-defective mutants retained enzymatic activity similar to WT, indicating no significant structural defects to the IN protein attributable to the introduced mutation. These results eliminate secondary defects that might be anticipated with IN, collaborating the specificity of the mapped RT binding surface on IN.

In chapter 3, we developed a pharmacological approach to examine the functional role of HIV-1 reverse transcriptase and integrase interaction during viral replication. An amplified luminescent proximity homogenous assay (ALPHA) was utilized to screen small molecule libraries for specific inhibitors against the RT-IN interaction. This method detects energy transfer between binding partners through singlet oxygen that can diffuse up to 200 nm. To our knowledge, we are reporting the first screening assay developed that is specific for the RT-IN interaction. Taking advantage of high sensitivity and low background noise properties of the alphascreen technology, we were able to miniaturize the final volume to 10 µl while maintain excellent assay reproducibility, Z' values of > 0.7 and high signal to background ratios. Additionally, we implemented a counter screening assay using double tagged IN, His-IN-Flag, enabling a high throughput false-positives detections platform of hits from the initial screen. Our pilot screen of 5440 compounds, resulted in 31 validated hits that directly targeted RT-IN interaction. Structural relationship studies revealed a selection of amine containing ring structures similar to NNRTI. We are currently assessing their anti-viral activity in tissue culture and half maximal effective concetration (EC_{50}) values. Potential lead compounds will be tested for their ability to inhibit viral replication and reverse transcription in cell culture. This analysis will verify the specificity of inhibition and corroborate the functional role of RT-IN interaction in

viral replication. Lastly, we embarked on mapping the IN-binding surface on RT. We performed crystallization buffer screening of three different matrixes and a high-resolution mass spectrometry foot-printing study. Acquiring the knowledge of the binding surface on both proteins will aid in rational design of small molecule inhibitors and open a new window of *in silico* drug screening.

Taken together, our findings will greatly augment our understanding of the biological significance and functional role of the RT-IN interaction in HIV-1 replication. Characterization of the RT-IN interaction and determination of its biological significance may reveal new functional roles for IN. In addition, considering the increasing viral resistance observed against many clinically relevant anti-HIV drugs that target the enzymatic active sites of both RT and IN, the insight gained here may identify new potential allosteric targets that prove beneficial for antiviral drug design. Our overall goal is to reveal new therapeutic targets for combating HIV-1 infection.

References

1. **UNAIDS.** 2013. UNAIDS report on the global AIDS epidemic 2013.
2. **Lanchy JM, Isel C, Ehresmann C, Marquet R, Ehresmann B.** 1996. Structural and functional evidence that initiation and elongation of HIV-1 reverse transcription are distinct processes. Biochimie **78:**1087-1096.
3. **Engelman A, Mizuuchi K, Craigie R.** 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell **67:**1211-1221.
4. **Brin E, Yi J, Skalka AM, Leis J.** 2000. Modeling the late steps in HIV-1 retroviral integrase-catalyzed DNA integration. J Biol Chem **275:**39287-39295.
5. **Daniel R, Katz RA, Skalka AM.** 1999. A role for DNA-PK in retroviral DNA integration. Science **284:**644-647.
6. **Chow SA, Vincent KA, Ellison V, Brown PO.** 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. Science **255:**723-726.
7. **Telesnitsky A, Goff SP.** 1997. Reverse Transcriptase and the Generation of Retroviral DNA.
8. **Jacobo-Molina A, Arnold E.** 1991. HIV reverse transcriptase structure-function relationships. Biochemistry-Us **30:**6351-6356.
9. **Asante-Appiah E, Skalka AM.** 1999. HIV-1 integrase: structural organization, conformational changes, and catalysis. Adv Virus Res **52:**351-369.
10. **Hare S, Di Nunzio F, Labeja A, Wang J, Engelman A, Cherepanov P.** 2009. Structural basis for functional tetramerization of lentiviral integrase. PLoS pathogens **5:**e1000515.

11. **Engelman A, Englund G, Orenstein JM, Martin MA, Craigie R.** 1995. Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. Journal of virology **69:**2729-2736.

12. **Shin CG, Taddeo B, Haseltine WA, Farnet CM.** 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. Journal of virology **68:**1633-1642.

13. **Engelman A.** 1999. In vivo analysis of retroviral integrase structure and function. Adv Virus Res **52:**411-426.

14. **Briones MS, Dobard CW, Chow SA.** Role of human immunodeficiency virus type 1 integrase in uncoating of the viral core. Journal of virology **84:**5181-5190.

15. **Zhu K, Dobard C, Chow SA.** 2004. Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. Journal of virology **78:**5045-5055.

16. **Masuda T, Planelles V, Krogstad P, Chen IS.** 1995. Genetic analysis of human immunodeficiency virus type 1 integrase and the U3 att site: unusual phenotype of mutants in the zinc finger-like domain. Journal of virology **69:**6687-6696.

17. **Leavitt AD, Robles G, Alesandro N, Varmus HE.** 1996. Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. Journal of virology **70:**721-728.

18. **Dobard CW, Briones MS, Chow SA.** 2007. Molecular mechanisms by which human immunodeficiency virus type 1 integrase stimulates the early steps of reverse transcription. Journal of virology **81:**10037-10046.

19. **Woodward CL, Prakobwanakit S, Mosessian S, Chow SA.** 2009. Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. Journal of virology **83:**6522-6533.

20. **Quillent C, Borman AM, Paulous S, Dauguet C, Clavel F.** 1996. Extensive regions of pol are required for efficient human immunodeficiency virus polyprotein processing and particle maturation. Virology **219:**29-36.

21. **Wu X, Liu H, Xiao H, Conway JA, Hehl E, Kalpana GV, Prasad V, Kappes JC.** 1999. Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex. Journal of virology **73:**2126-2135.

22. **Hehl EA, Joshi P, Kalpana GV, Prasad VR.** 2004. Interaction between human immunodeficiency virus type 1 reverse transcriptase and integrase proteins. Journal of virology **78:**5056-5067.

23. **Wilkinson TA, Januszyk K, Phillips ML, Tekeste SS, Zhang M, Miller JT, Le Grice SF, Clubb RT, Chow SA.** 2009. Identifying and characterizing a functional HIV-1 reverse transcriptase-binding site on integrase. J Biol Chem **284:**7931-7939.

24. **Nishitsuji H, Hayashi T, Takahashi T, Miyano M, Kannagi M, Masuda T.** 2009. Augmentation of reverse transcription by integrase through an interaction with host factor, SIP1/Gemin2 Is critical for HIV-1 infection. PloS one **4:**e7825.

25. **Tasara T, Maga G, Hottiger MO, Hubscher U.** 2001. HIV-1 reverse transcriptase and integrase enzymes physically interact and inhibit each other. Febs Lett **507:**39-44.

26. **Ding J, Das K, Hsiou Y, Sarafianos SG, Clark AD, Jr., Jacobo-Molina A, Tantillo C, Hughes SH, Arnold E.** 1998. 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 A resolution. J Mol Biol **284:**1095-1111.

27. **Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Jr., Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, et al.** 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proc Natl Acad Sci U S A **90:**6320-6324.

28. **Patel PH, Jacobo-Molina A, Ding J, Tantillo C, Clark AD, Jr., Raag R, Nanni RG, Hughes SH, Arnold E.** 1995. Insights into DNA polymerization mechanisms from structure and function analysis of HIV-1 reverse transcriptase. Biochemistry-Us **34:**5351-5363.

29. **Palmer AG, 3rd, Kroenke CD, Loria JP.** 2001. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods Enzymol **339:**204-238.

30. **Lian L-Y, Roberts GCK.** 1993. Effects of chemical exchange on NMR spectra, p. 153-182. *In* Roberts GCK (ed.), NMR of Macromolecules: A Practical Approach. Oxford University Press Inc., New York.

31. **Mandal D, Feng Z, Stoltzfus CM.** 2008. Gag-processing defect of human immunodeficiency virus type 1 integrase E246 and G247 mutants is caused by activation of an overlapping 5' splice site. Journal of virology **82:**1600-1604.

32. **Lu R, Ghory HZ, Engelman A.** 2005. Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. Journal of virology **79:**10356-10368.

33. **Korin YD, Zack JA.** 1998. Progression to the G1b phase of the cell cycle is required for completion of human immunodeficiency virus type 1 reverse transcription in T cells. Journal of virology **72:**3161-3168.

34. **Oz I, Avidan O, Hizi A.** 2002. Inhibition of the integrases of human immunodeficiency viruses type 1 and type 2 by reverse transcriptases. The Biochemical journal **361:**557-566.

35. **Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC.** 1995. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc Natl Acad Sci U S A **92:**1222-1226.

36. **Hsiou Y, Ding J, Das K, Clark AD, Jr., Hughes SH, Arnold E.** 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure **4:**853-860.

37. **Sarafianos SG, Das K, Tantillo C, Clark AD, Jr., Ding J, Whitcomb JM, Boyer PL, Hughes SH, Arnold E.** 2001. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. The EMBO journal **20:**1449-1461.

38. **Tasara T, Maga G, Hottiger MO, Hubscher U.** 2000. HIV-1 reverse transcriptase and integrase enzymes physically interact and inhibit each other. Febs Lett **507:**39-44.

39. **Cannon PM, Wilson W, Byles E, Kingsman SM, Kingsman AJ.** 1994. Human immunodeficiency virus type 1 integrase: effect on viral replication of mutations at highly conserved residues. Journal of virology **68:**4768-4775.

40. **Christ F, Shaw S, Demeulemeester J, Desimmie BA, Marchand A, Butler S, Smets W, Chaltin P, Westby M, Debyser Z, Pickford C.** 2012. Small-molecule inhibitors of

the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. Antimicrobial agents and chemotherapy **56:**4365-4374.

41. **Adachi A, Ono N, Sakai H, Ogawa K, Shibata R, Kiyomasu T, Masuike H, Ueda S.** 1991. Generation and characterization of the human immunodeficiency virus type 1 mutants. Archives of virology **117:**45-58.

42. **Fisher CL, Pei GK.** 1997. Modification of a PCR-based site-directed mutagenesis method. Biotechniques **23:**570-&.

43. **Shibagaki Y, Holmes ML, Appa RS, Chow SA.** 1997. Characterization of feline immunodeficiency virus integrase and analysis of functional domains. Virology **230:**1-10.

44. **Khan M, Garcia-Barrio M, Powell MD.** 2001. Restoration of wild-type infectivity to human immunodeficiency virus type 1 strains lacking nef by intravirion reverse transcription. Journal of virology **75:**12081-12087.

45. **Lodi PJ, Ernst JA, Kuszewski J, Hickman AB, Engelman A, Craigie R, Clore GM, Gronenborn AM.** 1995. Solution structure of the DNA binding domain of HIV-1 integrase. Biochemistry-Us **34:**9826-9833.

46. **Goddard TD, Kneller DG.** Sparky 3. University of California, San Francisco.

47. **Schrodinger, LLC.** 2010. The PyMOL Molecular Graphics System, Version 1.3r1.

48. **De Clercq E.** 2009. Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. International journal of antimicrobial agents **33:**307-320.

49. **Arora P, Dixit NM.** 2009. Timing the emergence of resistance to anti-HIV drugs with large genetic barriers. PLoS computational biology **5:**e1000305.

50. **Peterson CW, Younan P, Jerome KR, Kiem HP.** 2013. Combinatorial anti-HIV gene therapy: using a multipronged approach to reach beyond HAART. Gene therapy **20:**695-702.

51. **Werner S, Hindmarsh P, Napirei M, Vogel-Bachmayr K, Wohrl BM.** 2002. Subcellular localization and integration activities of rous sarcoma virus reverse transcriptase. Journal of virology **76:**6205-6212.

52. **Trentin B, Rebeyrotte N, Mamoun RZ.** 1998. Human T-cell leukemia virus type 1 reverse transcriptase (RT) originates from the pro and pol open reading frames and requires the presence of RT-RNase H (RH) and RT-RH-integrase proteins for its activity. Journal of virology **72:**6504-6510.

53. **Demeulemeester J, Tintori C, Botta M, Debyser Z, Christ F.** 2012. Development of an AlphaScreen-based HIV-1 integrase dimerization assay for discovery of novel allosteric inhibitors. Journal of biomolecular screening **17:**618-628.

54. **Pena I, Dominguez JM.** 2010. Thermally denatured BSA, a surrogate additive to replace BSA in buffers for high-throughput screening. Journal of biomolecular screening **15:**1281-1286.

55. **Wagstaff KM, Jans DA.** 2006. Intramolecular masking of nuclear localization signals: analysis of importin binding using a novel AlphaScreen-based method. Analytical biochemistry **348:**49-56.

56. **Hou Y, McGuinness DE, Prongay AJ, Feld B, Ingravallo P, Ogert RA, Lunn CA, Howe JA.** 2008. Screening for antiviral inhibitors of the HIV integrase-LEDGF/p75 interaction using the AlphaScreen luminescent proximity assay. Journal of biomolecular screening **13:**406-414.

57. **Zhan P, Liu X, Li Z.** 2009. Recent advances in the discovery and development of novel HIV-1 NNRTI platforms: 2006-2008 update. Current medicinal chemistry **16:**2876-2889.

58. **Wang Z, Tang J, Salomon CE, Dreis CD, Vince R.** 2010. Pharmacophore and structure-activity relationships of integrase inhibition within a dual inhibitor scaffold of HIV reverse transcriptase and integrase. Bioorganic & medicinal chemistry **18:**4202-4211.

59. **Pollastri MP.** 2010. Overview on the Rule of Five. Current protocols in pharmacology / editorial board, S.J. Enna **Chapter 9:**Unit 9 12.

60. **Zheng X, Mueller GA, Cuneo MJ, Derose EF, London RE.** Homodimerization of the p51 subunit of HIV-1 reverse transcriptase. Biochemistry-Us **49:**2821-2833.

61. **Ding J, Das K, Tantillo C, Zhang W, Clark AD, Jr., Jessen S, Lu X, Hsiou Y, Jacobo-Molina A, Andries K, et al.** 1995. Structure of HIV-1 reverse transcriptase in a complex with the non-nucleoside inhibitor alpha-APA R 95845 at 2.8 A resolution. Structure **3:**365-379.

62. **Wang J, Smerdon SJ, Jager J, Kohlstaedt LA, Rice PA, Friedman JM, Steitz TA.** 1994. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. Proc Natl Acad Sci U S A **91:**7242-7246.

63. **Smerdon SJ, Jager J, Wang J, Kohlstaedt LA, Chirino AJ, Friedman JM, Rice PA, Steitz TA.** 1994. Structure of the binding site for nonnucleoside inhibitors of the reverse transcriptase of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A **91:**3911-3915.

64. **Chen JC, Krucinski J, Miercke LJ, Finer-Moore JS, Tang AH, Leavitt AD, Stroud RM.** 2000. Crystal structure of the HIV-1 integrase catalytic core and C-terminal

domains: a model for viral DNA binding. Proceedings of the National Academy of Sciences of the United States of America **97:**8233-8238.

65. **Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P.** 2010. Retroviral intasome assembly and inhibition of DNA strand transfer. Nature **464:**232-236.