Spring 5-15-2022

Molecular Characterization of Integrase-RNA Interactions and Their Role in the Replication of HIV-1 and Other Retroviruses

Christian Shema Mugisha
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Cell Biology

Dissertation Examination Committee:
Sebla B. Kutluay, Chair
Megan Baldridge
Siyuan Ding
Sergej Djuranovic
Liang Shan

Molecular Characterization of Integrase-RNA Interactions and Their Role in the Replication of HIV-1 and Other Retroviruses

by

Christian Shema Mugisha

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2022
St. Louis, Missouri

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Acknowledgements

This page is dedicated to the many brilliant and generous scientists who inspired and supported my scientific and academic growth, including Sebla Kutluay, Jenna Eschbach, Maritza Puray-Chavez, Kyle Vuong, Dana Lawson, Keanu Davis, Kasyap Tenneti, Shawn Mohammed, Nakyung Lee, Mike Wang, Tung Dinh, Mamuka Kvaratskhelia, Robert Gifford, Alan Engelman, Wen Li, Wandy Beatty, Sergej Djuranovic, Liang Shan, Megan Baldridge, Siyuan Ding, Sondra and Milton Schlesinger, Stacy Kiel, Zhongsheng You, and Roberta Faccio.

I am particularly grateful to my advisor and mentor Dr. Sebla Kutluay. By her unwavering support and encouragement, Dr. Kutluay has given the best opportunity and space to learn and grow as a scientist. Without her mentorship, I would not have made it this far.

I also wish to acknowledge my family, friends, and classmates who have believed in me and supported me during many up and downs. I am grateful for my teachers and professors who encouraged my academic aspirations. Finally, I want to thank the DBBS program and Washington university for providing some of the learning resources during my PhD training.

Christian Shema Mugisha

Washington University

May 2022

ABSTRACT OF THE DISSERTATION

Molecular Characterization of Integrase-RNA interactions and their role in the replication of HIV-1 and other retroviruses

by

Christian Shema Mugisha

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Cell Biology

Washington University in St. Louis, 2022

Professor Sebla B. Kutluay, Chair

HIV-1 integrase (IN) enzyme has an emerging non-catalytic role in particle maturation, which involves its binding to the viral genome in virions. Allosteric integrase inhibitors (ALLINIs) and class II integrase substitutions inhibit the binding of IN to the viral genome and cause formation of eccentric non-infectious HIV-1 particles. These viruses are characterized by the mislocalization of the viral ribonucleoprotein complexes between the translucent conical CA lattice and the viral lipid envelope. We have previously demonstrated that IN binding to the viral genome is mediated by basic residues within the C-terminal domain of IN.

In the first chapter, we show how basic residues of the IN CTD mediate RNA binding. We report that we have isolated secondary site suppressors of a class II IN mutant (R269A/K273A) which directly inhibits IN binding to the viral genome. Full-genome deep sequencing revealed the sequential emergence of D256N and D270N mutations within three passages. Reintroduction of these substitutions nearly fully restored the replication defect of the R269A/K273A virus, restored the ability of IN to bind RNA and led to the formation of particles

with mature morphology. Furthermore, we found that D256R and D256R/270R substitutions also increased the infectivity of R269A/K273A as well as R262A/R263A IN viruses. The nature of these suppressor mutations suggests that IN-RNA binding is partly dictated electrostatic interactions between IN CTD basic residues and RNA. Though these findings imply some level of non-specificity towards gRNA binding, CLIP-seq and in-vitro binding experiments (reported in the third chapter) revealed a striking preference of IN for binding to purine-rich sequences on the viral genome. Taken together, our findings suggest that a combination of electrostatic interactions and semi-specific binding to the viral genome underlies the non-catalytic role of IN in virion maturation. Additional preliminary findings reported in the third chapter show that IN-RNA binding is a conserved property of retroviruses, further reaffirming this characteristic would be an archetypical target for new antiretroviral agents.

In the fourth chapter, we report how we used our expertise of RNA viruses to create an assay that has been usefully in screening for antiviral agents for the SARS-Cov-2 virus during the COVID-19 pandemic.

Overall, this dissertation illustrates how the basic molecular properties of integrase and viral genomic RNA that underlie their binding. Furthermore, it shows how these properties could be potentially targeted by the next generation of antiretroviral agents.

Chapter 1: Introduction

1.1 Introduction

The human immunodeficiency virus type 1 (HIV-1), the causative agent of AIDS, has been one of the deadliest viruses of the last four decades. As of 2018, there were about 37 million living with HIV worldwide, and there were 1 million HIV-related deaths the same year (8). HIV1 is a lentivirus that contains two copies of positive sense single-stranded genomic RNA (gRNA), 9 kb in size (10). The viral gRNA forms a ribonucleoprotein (RNP) complex together with the nucleocapsid protein (NC), and reverse transcriptase (RT) and integrase (IN) enzymes in virions (11). The dense vRNP complex is encased within a conical lattice composed of ~2000 copies of capsid (CA) protein, together forming the viral "core" . A defining feature of retroviruses is the reverse transcription of the viral RNA genome (gRNA) and integration of the resultant linear viral DNA into a host chromosome, which establishes lifelong infection. The integration reaction is catalyzed by the viral integrase (IN) enzyme, which possesses 3' processing and DNA strand transfer activities . This catalytic role of integrase has been efficiently targeted by a group of antiretroviral agents known as integrase strand-transfer inhibitors (INSTIs) (16, 17). Currently, there are four FDA-approved INSTIs –raltegravir (18-20), elvitegravir , dolutegravir , and bictegravir (18, 19, 21-23); they have become essential constituents anti-retroviral therapy regimen because of their high potency and tolerance profiles . Even with the high efficacy of INSTIs, mutations that confer resistance to several INSTIs have been observed in patients (24) and treatment selects for the drug resistant viruses (24-27). These observations reiterate the need for further research and development of novel antiretroviral therapies.

The integrase enzyme also has a non-catalytic function in HIV-1 virion morphogenesis . IN binds to the viral genomic RNA (gRNA), and this binding is essential for proper positioning of the gRNA within the capsid lattice during virion maturation. This second function of IN could

be targeted by a new generation/class of antiretroviral agents, which would complement existing regimens and increase the barrier to INSTI resistance. This introduction chapter briefly reviews the HIV-1 life cycle while expanding on the roles IN's functions in integration and virion morphogenesis. A comprehensive mechanistic understanding of IN's functions in the HIV-1 life cycle would be pivotal to the development of novel and more efficient antiretroviral drugs.

1.2 Overview of the HIV-1 replication cycle

The human immunodeficiency virus type 1 (HIV-1) is a lentivirus that contains two copies of positive sense single-stranded RNA, 9 kb size (31-33). In their mature form, HIV-1 virions have a viral core consisting of a conical capsid (of about of ~2000 copies of capsid (CA) monomers) which surrounds the vRNP complex composed of viral RNA, the nucleocapsid protein (NC), integrase (IN), and reverse transcriptase (RT) (Figure 1). The viral core is surrounded by a lipid bilayer envelope, derived from the host cell plasma membrane, traversed by glycoprotein spikes (Env) that serve as receptors during viral entry (34). During viral entry, the HIV protein envelope (Env) binds to the primary cellular receptor CD4 and then to a cellular coreceptor (CXCR4/CCR5); these receptors are mostly on CD4+ T cells and monocyte/macrophage lineage cells. Following the binding of Env to the cell receptors, the viral and cellular membranes fuse and then the viral core is released into the cytoplasm (33).

Subsequent to viral entry, the viral core is then moved towards the nucleus as the RT enzymes synthesizes the viral DNA (vDNA) from the HIV-1 RNA genome . While multiple studies have previously concluded that the viral core uncoats during this stage , several new studies have shown that the capsid lattice stays almost intact until it is imported into the nucleus . After the synthesis of vDNA, the integrase enzyme, which stays associated with the RT complex, forms multimers that engage both ends of the vDNA forming the intasome complex. In the cell

nucleus, as part of the intasome complex, IN catalyzes the integration of viral DNA into the host cell chromosome .

Viral gene expression from the integrated provirus depends on viral and cellular transcription factors which lead to the transcription of single full-length viral mRNAs by the host RNA polymerase II machinery (46-48). All HIV-1 mRNAs are generated from a single primary polycistronic viral transcript that undergoes extensive alternative splicing . The completely spliced 1.8kb mRNAs encode Tat, Rev and Nef regulatory proteins, they are shuttled to the cytoplasm through the NXF1/NXT1nuclear export complex . The partially spliced 4 kb size class encode Vif, Vpr, Env, Vpu and a truncated form of Tat. The unspliced primary transcript encodes Gag and Gag-Pol proteins, and serves as genomic RNA that is packaged into progeny virions . In contrast to Gag, Gag-Pol has a pol extension which codes for the viral enzymes protease (PR), RT, and IN. Gag-Pol is produced as result of programmed ribosomal frame shifting of the Gag ORF during translation of the unspliced viral RNA (52, 53). The export of the partially spliced and unspliced viral RNA to the cytoplasm is Rev-dependent. Rev is translated in the cytoplasm from the fully spliced HIV-1 mRNA. Rev is then translocated to the nucleus, binds to the Rev-response-element present in partially and fully spliced viral RNA, and then mediates their export through the CRM1 nuclear export complex (54, 55).

The viral envelope protein (Env) is synthesized at the endoplasmic reticulum, glycosylated, and cleaved into gp120 and gp41 at the Golgi complex . Cytoplasmic assembly of virions is initiated by interaction of NC, a domain of Gag and Gag-Pol, with the packaging signal in the unspliced genomic RNA (58-60). The Gag and Gag-Pol proteins are then transported to the plasma membrane, where the membrane-bound MA mediates oligomerization and assembly of Gag and

Gag-Pol units. MA domain of Gag and Gag-Pol also associates with Env to initiate budding , the virions released from cells are initially in an immature state where Gag and Gag-Pol are radially arranged across the virion. Shortly after virion release, Gag and Gag-Pol polyproteins are cleaved and modified by the viral PR enzyme. This results in the formation of mature infectious virions that are spherical with a conical CA lattice that encloses the vRNP complex (Fig. 1), forming the viral core .

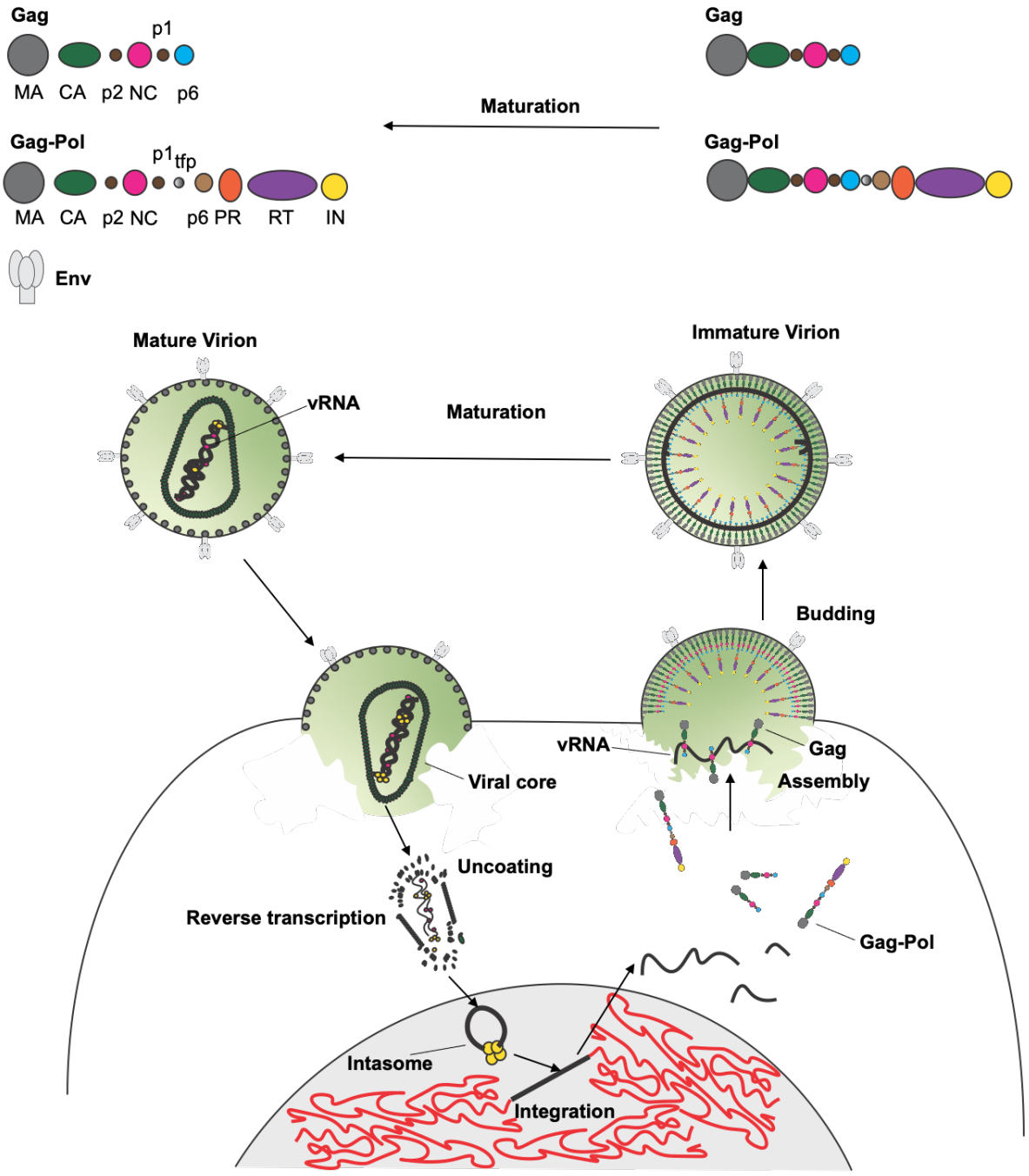

Figure 1: Overview of the HIV-1 life cycle
HIV-1 virions contain two copies of a single-stranded RNA (vRNA, in black) genome enclosed inside a conical capsid lattice. After viral entry the viral core is released inside the cytoplasm and transported towards the nucleus. The vRNA is reversed transcribed into double-stranded DNA (vDNA, in black.) which is integrated into the host DNA. Following transcription of vDNA, the produced vRNA and viral proteina assemble at the plasma membrane and bud of from the cell as immature virions. During virion maturation, Gag and Gag-Pol are cleaved leading to formation of mature virions where the capsid lattice reforms again around the viral ribonucleoproteins complex.

1.3 HIV-1 Integrase and its catalytic Role

Similarly to other retroviruses, the lentivirus HIV-1 must integrate the DNA copy of its RNA genome into the host cell chromosome; it is this step of the virus replication that helps maintain lifelong infection in patients. The integration process is mediated by the HIV-1 integrase (IN) enzyme. IN is a special recombinase that is incorporated into the virus particles as part of Gag-Pol. Integration occurs within the context of a large nucleoprotein structure known as the preintegration complex (PIC) . During integration, both ends of the linear vDNA are held together by a multimer of IN catalyzes integration . The IN enzyme has three functionally distinct domains; an N-terminal domain (NTD) that has conserved HCCH Zn2+-binding motif, an RNase H-like catalytic core domain (CCD), and a C-terminal domain (CTD) which has been shown to play a role in nucleic acid binding (67). While the NTD and CTD domains facilitate DNA binding and formation of the intasome complex, the CCD possesses a conserved D,D,-35-E motif essential for IN's catalytic activity (70). Mutations at these conserved residues block the catalytic activity of IN and thus virus replication (72), they are collectively known to as class I IN mutations.

To perform its catalytic functions, IN multimerizes to different levels forming monomers, dimers, and tetramers (75). Furthermore, there had been studies showing that small molecules could allosterically block the catalytic function of IN by binding at the interface of IN dimers within an IN tetramer, thus disrupting the proper multimerization of IN around vDNA into a fully functional nucleoprotein complex (76). Even though the enzymatic functions of IN has been studied since the 1980s, a complete understanding of its catalytic functions was only recently discovered owing to the detailed high-resolution structure of IN-DNA complexes or intasomes for the spumavirus prototype foamy virus (PFV) . Since then, intasome structures of

the α-retrovirus Rous sarcoma virus (RSV) and β-retrovirus mouse mammary tumor virus (MMTV) have been characterized . While the intasome structures of the spumavirus PFV are primarily tetrameric, RSV and MMTV have octameric intasome structures revealing an evolutionary diversity among retroviral intasomes (Figure 2) .

Resolving a detailed structure of the HIV-1 intasome has long been hindered by the tendency of HIV-1 IN to aggregate in solution. A recent study overcame this hurdle by using hyper-active HIV-1 IN mutant protein with improved solubility; using cryo-EM, it was shown that HIV-1 intasomes form a range of oligomeric configurations from tetramers to higher order dodecameric complexes for integration (84, 85) (Figure 2).

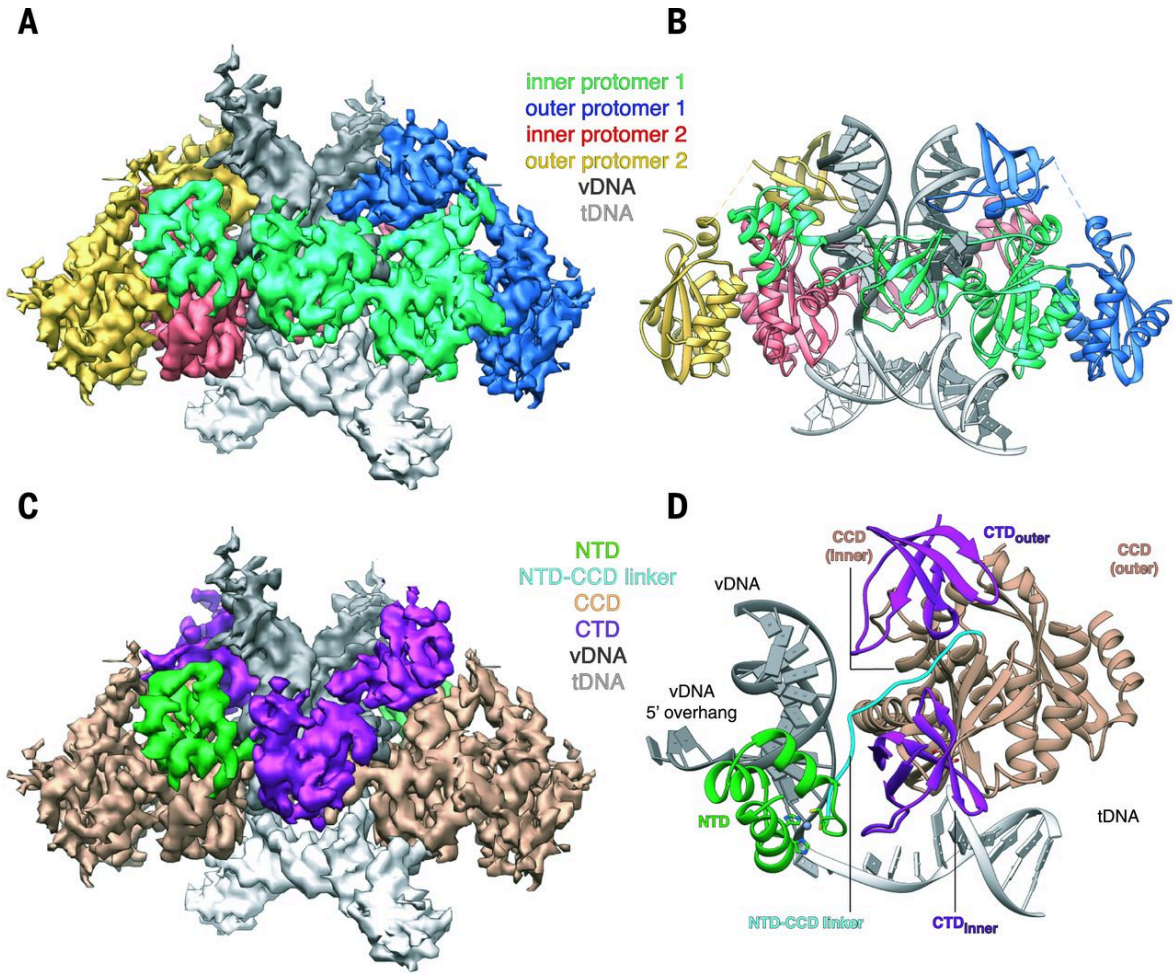

Figure *2:* Structures of HIV-1 strand transfer complex intasomes (pdb code: 5U1C)
(A) CryoEM reconstruction of the STC, segmented by IN protomer (red, green, yellow, blue) and product DNA component (dark and light grey). (B) Atomic model, similarly, colored as in A. (C) segmented cryoEM density and (D) asymmetric subunit of the atomic model colored by IN domain; NTD (green), CCD (beige), NTD-CCD linker (cyan), CTD (purple).

Following the formation of the intasome complex, IN catalyzes integration through two biochemically and temporally distinct biomolecular nucleophilic substitution reactions: 3' processing and strand transfer (Figure 3) (86, 87). During the 3' processing, IN catalyzes the hydrolysis of a phosphodiester bond from each 3' vDNA end which removes two to three nucleotides; this step leaves reactive 3' hydroxyl groups . These free 3' hydroxyl groups act as nucleophiles in the strand transfer reaction where they are covalently inserted into the major groove of the host DNA in staggered fashion . This produces an integration intermediate with

unjoined 5' vDNA overhangs (89, 90). Thereafter, the strand transfer complex (intasome) dissembles and DNA polymerase, 5' flap endonuclease, and DNA ligase enzymes respectively fill in the single-stranded regions, excise 5' vDNA overhangs, and join the vDNA 5' ends to the host DNA strands (21, 24, 91, 92). A characteristic consequence of retroviral integration and the subsequent gap repair is the generation of short target DNA duplications flanking the integrated provirus. The size of these target site duplications varies between various retroviruses. For example, while the 5 bp target site duplication is the hallmark of HIV-1, spumavirus prototype foamy virus integration results in a 4bp target site duplication .

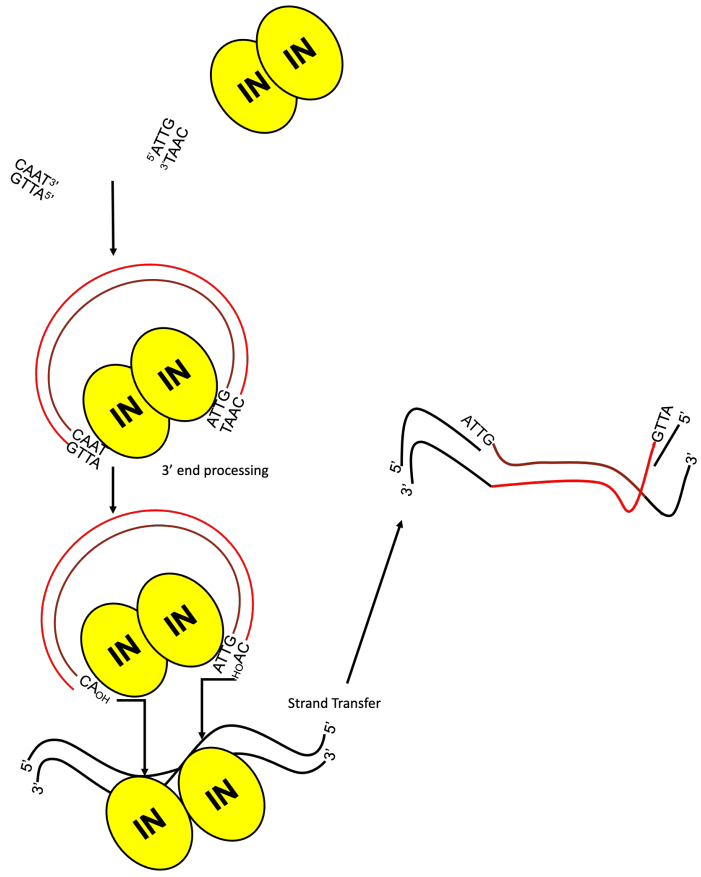

Figure 3: Mechanism of retroviral integration
The integration process has two step-3' processing and strand transfer- both catalyzed by integrase. During 3' processing, an IN dimer makes the phosphodiester bonds at the site of vDNA cleavage susceptible to nucleophilic attach by water or other nucleophiles. This leads to the removal of a dinucleotide from the 3' ends of the viral DNA (red and brown) to expose free 3' hydroxyls. During strand transfer, another IN dimer inserts the 3' ends of the viral DNA into the host DNA (black), leaving gaps in the target DNA and the loose 5' ends of viral DNA. The gaps and loose ends are subsequentially repaired by host cell machinery.

HIV-1 strongly prefers integration within transcriptionally active units in the nuclear periphery . These active transcription units are associated with regions of high G/C content, high gene density, high CpG island, short introns, high frequencies of Alu repeats, low frequencies of

LINE repeats, and characteristic epigenetic modifications . HIV takes advantage of different cell factors to promote efficient integration and optimize the integration at favored sites. Most notably, IN binds to the chromatin-associated cellular protein lens epithelium-derived growth factor (LEDGF) through its CTD (76). The PWWP motif within the N-terminus of LEDGF binds to the nucleosomes trimethylated at Lys36 of histone 3 (H3K36me3), one of the common epigenetic markers for transcriptionally active sites (99). In fact, mapping of integration sites in LEDGF/p75 knockdown cells showed that integration targeting to transcription sites was significantly reduced; on the contrary, integration targeting to transcriptionally active sites was rescued upon restoration of LEDGF/p75 expression (14, 24, 73, 75, 76). Imaging studies have showed that LEDGF/p75 binds to condensed chromosomes at mitosis, and IN also colocalizes with the chromatin in the presence of LEDGF/p75 . Altogether, the convention is a model where LEDGF/p75 guides and tethers IN to chromatin at active transcription units.

Some of the most potent available antiretroviral therapies are integrase inhibitors known as INSTIs (integrase strand transfer inhibitors). Four INSTIs are part of clinically regimes, raltegravir , elvitegravir (96), dolutegravir (97), and bictegravir (98, 101). A fifth INSTI– cabotegravir - is currently in late-stage clinical trials. INSTIs bind to the catalytic core domain (CCD) of IN; this displaces the reactive 3' end of the viral DNA thus preventing its insertion into the host DNA (102). However, treatment with INSTIs can lead to selection of drug resistance and resistance is often acquired through mutations within the CCD of IN. Although such mutations can affect the fitness of the virus, it has been observed that compensatory mutations that increase IN catalytic activity are found in patients on INSTIs.

1.4 Virion Morphogenesis

The infectious virus particles (virions) of HIV-1 is spherical with a diameter of about 120 nm with viral glycoproteins protruding from its envelope. Following the release of virions, they undergo an interior morphology transformation, named maturation. The maturation process changes the virion from an immature form where the Gag and Gag-Pol proteins are radially arranged across the viral envelope in a donut-shaped particle to a mature virion particle lined with viral matrix proteins (MA) containing a condensed conical core composed of capsid lattice that encapsidates the vRNP complex (103). The maturation process is essential for the formation infectious virions.

The assembly of HIV-1 virions is driven by the viral major sturctual protein, Gag. The precursor proteins Gag (Pr55) and Gag-Pol (Pr160) are expressed from an unspliced full-length viral RNA. The N-terminal myristoylation target Gag and Gag-Pol to the plasma membrane and the adequate accumulation of precursor proteins leads to their assembly at the host cell membrane (108). HIV-1 particle assembly is initiated when the nucleocapsid domain (NC) of Gag binds with high specificity to a single dimer of viral genomic RNA (gRNA) (109, 110). NC interacts with RNA through electrostatic binding with its positively charged residues and more specifically via its zinc finger binding pockets that interact with unpaired guanosine residues . HIV-1 RNAs are 5' capped and 3' polyadenylated, this makes HIV-1 RNA similar to the host cell's mRNA which is challenging for selective packaging of viral RNA during virion assembly. However, the HIV-1 RNA contains a packaging sequence called psi (ψ) which is specifically recognized by NC as mentioned above. The packaging sequence is formed by four stem loop structures-SL1, SL2, SL3, and SL4; these structures allow viral RNA to dimerize, and they are

part of the 5' untranslated region (5' UTR) (112). SL1 initiates the dimerization of viral RNA molecules, which is crucial for selective packaging of two copies of full-length vRNA (113). The dimer initiation site of SL1 contains a bulge of 9 bases; six of these bases form a palindrome which leads to the formation of Watson-Crick base pairs with the second RNA molecule.

The main constituents of HIV-1 virions are Gag, Gag-Pol, Env, viral membrane lipids, and the two copies of genomic RNA. In addition to these, the virion packages some viral accessory proteins which are necessary at different stages of the virus replication. Gag molecules have been shown to interact with various components of the intracellular vesicle trafficking pathways-AP-1, AP-2, and AP-3 -and the cellular motor protein KIF4 (114). The myristoylation of Gag, the basic patch on the MA domain, and the plasma membrane-specific lipid phosphatidyl inositol (4,5) biphosphate (PI(4,5)P2) facilitate the targeting of Gag to the plasma membrane (24, 104, 118). The binding between the MA domain of Gag and PI(4,5)P2 exposes the amino-terminal myristoyl group enabling the stable anchoring of Gag on the inner leaflet of the cellular membrane . Gag molecules arrive at the plasma membrane as small oligomers (dimers or monomers) before polymerizing around nucleation sites composed of Gag-RNA complexes (21, 23, 91, 94, 95, 105, 108, 110, 119-124). Env is an integral membrane protein and is co-translationally incorporated in the endoplasmic reticulum (ER) membranes. Env is delivered to the plasma membrane through vesicular secretory pathways where it is glycosylated, assembled into tetrameric complexes, and processed into the trans-membrane (gp41) and surface (gp120) subunits by the cellular protease furin .

Through a complex combination of Gag–lipid, Gag–Gag, and Gag–RNA interactions, a multimeric budding structure forms at the inner leaflet of the plasma membrane. The budding virus particle is ultimately released from the cell surface in a process that is promoted by an

interaction between the late domain in the p6 region of Gag and host proteins, most notably the endosomal sorting factor, TSG101(tumor susceptibility gene 101) . The released virions are in an immature state with 2000-4000 Gag molecules that are radially arranged along the inside of the viral envelope in a donut shape; in this form, MA remains anchored to the membrane at one end of Gag and NC still binds to vRNA . Following budding, PR enzyme cleaves Gag and Gag-Pol at multiple sites in a defined sequence to generate mature proteins matrix (MA), capsid (CA), nucleocapsid (NC), p6, two spacer peptides (SP1 and SP2) integrase (IN), reverse transcriptase (RT), and PR . The cleavage of Gag and Gag-Pol triggers a structural rearrangement called maturation, in which the immature virions become mature infectious virion characterized by an electron dense, conical core. The core is made of 1000-1500 monomers of CA assembled in a conical lattice around two single-stranded HIV-1 RNA molecules bound by the NC and associated with IN and RT . This virion maturation process is essential for the released virus particles to become infectious and initiate a new round of infection.

1.5 The role of integrase in virion morphogenesis

Integrase has long been known for its catalytic function of integrating the vDNA into the host chromosome. Several decades ago, it was noticed that IN might have other non-catalytic functions in virus replication based on observations from mutation analyses. Most notably, a group of pleiotropic mutations known as class II mutations caused defects in virion assembly , morphogenesis , and reverse transcription . Most of the class II mutations did not affect the catalytic activity of IN in vitro studies . A hallmark of class II IN mutant viruses is a phenotype known as “eccentric morphology’, characterized by the mislocalization of the vRNP outside of the conical capsid lattice (Figure 5). Deletion of IN causes similar morphological defects as class II mutations (45, 55, 57, 123, 149-151).

Interestingly, viruses produced from cells treated with allosteric integrase inhibitors (ALLINIs) display eccentric morphology like that of class II mutations . ALLINIs have, however, been designed to inhibit integration by blocking the binding of LEDGF/p75 to IN . ALLINIs inhibit integration by competing with LEDGF for IN by engaging a V-shaped binding pocket which is created by the catalytic core domains of two IN dimers in the intasome complex (152) and by causing aberrant IN multimerization in a way that forms catalytically inactive multimers that cannot form active intasomes (92, 93). The effect of ALLINIs and class II IN mutations on virion morphology suggested that IN plays a crucial role in HIV-1 particle morphogenesis. Although it had been shown that most class II mutations and ALLINIs interfered with proper IN multimerization, the exact mechanism by which IN was involved in vRNA encapsidation remained elusive in the retrovirology community for many years.

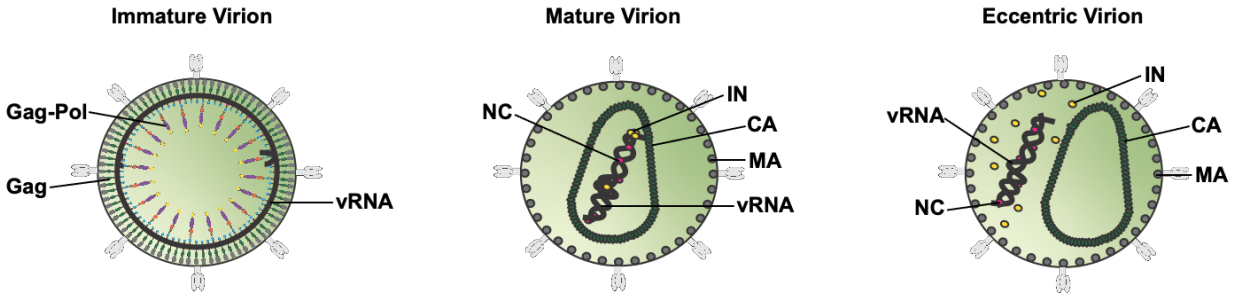

Figure 4: HIV-1 virion morphologies

Immature viral particles consist of 2000-4000 molecules of Gag and Gag-Pol radially arranged along the inner leaflet of the viral membrane, the genomic vRNA is bound to Gag's NC domain. In mature virions, the vRNP (vRNA bound by NC and condensed with RT and IN) is enclosed in the conical capsid lattice forming the viral core. In eccentric viral particles the vRNA is mislocalized outside of the capsid.

In 2016, a groundbreaking study from our lab and collaborators discovered that IN binds to vRNA in mature virions, this binding is necessary for the proper encapsidation of vRNPs inside the CA lattice after virion maturation (94) **(Fig. 5).** In these studies, the binding between IN and gRNA was discovered using a method known as crosslinking immunoprecipitation that is

coupled with next generation sequencing (CLIP-seq), a method which captures protein-RNA complexes in physiological conditions (1). CLIP-seq revealed that IN binds to gRNA at discrete sites that are different from other RNA-binding HIV-1 proteins such NC; in-vitro RNA-binding assays revealed that IN had high affinity for structured RNA elements such as TAR (2). IN not only binds RNA, but also modulates RNA structure in vitro by bridging multiple RNA molecules together (3). Several basic residues of the IN CTD- K264, K266, and K273 - were shown to directly bind to viral RNA, mutations in any of these residues blocked IN-gRNA binding (4, 5). Later studies by Elliott et al revealed that class II IN mutations impeded IN-RNA binding through three distinct means: 1) reducing packaged IN levels thus precluding formation of IN complexes with viral RNA; 2) impairing functional IN multimerization hence disrupting IN-RNA binding; 3) directly impairing IN-RNA binding without adversely affecting IN levels or functional IN multimerization . All class II mutations cause formation of morphologically eccentric particles by blocking IN-gRNA binding; following viral infection, the viral genome of such particles was degraded in target cells as a result of the mislocalization of the genomic RNA outside of the protective capsid lattice . Similarly, IN-gRNA binding is blocked and eccentric virions are formed when virions were produced from cells treated with ALLINIs .

Overall, IN-gRNA binding permits the encapsidation of the viral genomic RNA into the protective capsid lattice. Class II IN mutations and ALLINIs block IN-gRNA binding causing formation of virions with eccentric morphology, the resulting degradation of IN and gRNA in target cells, as well as spatial separation of reverse transcriptase and gRNA during the early stages of viral infection.

1.6 Conclusions

The HIV-1 integrase (IN) is a multifunctional enzyme that plays key roles during the integration and virion maturation steps of the virus replication. During integration, IN engages viral and host DNA catalyzing the incorporation of viral DNA into the host chromosome, this is the step that establishes life-long infection in HIV-1 patients. During virion morphogenesis, IN multimers bind to gRNA giving it a nucleation point and allowing it to be packaged inside the protective capsid lattice. Class II mutations and allosteric integrase inhibitors (IN) block IN-gRNA binding, thus leading to formation of non-infectious virions with an eccentric morphology where the gRNA is mislocalized outside of the capsid lattice.

IN presents itself as a great target for developing new antiretroviral therapies because both its functions in integration and virion maturation are essential for HIV-1 replication. The catalytic activity of IN has been successfully targeted by IN strand-transfer inhibitors (INSTIs) which are a key component of anti-retroviral therapy regimens today. The discovery of INSTI-resistant in clinics has raised concerns and it necessitates the development of novel IN-targeting therapies. The secondary function of IN in virion maturations is a great target for developing anti-HIV agents; however, there is still much to be understood about IN-RNA binding and it regulates virion maturation. There are several ongoing studies on ALLINIs and how they can be developed into efficient anti-retroviral therapies. Clear understanding of IN-RNA binding and virion maturation will inform the development of new and improved integrase-targeting antiretroviral therapies.

References

1. Harris ME, Hope TJ. 2000. RNA export: insights from viral models. Essays Biochem 36:115-27.

2. Cullen BR. 2003. Nuclear mRNA export: insights from virology. Trends Biochem Sci 28:419-24.

3. Wodrich H, Kräusslich HG. 2001. Nucleocytoplasmic RNA transport in retroviral replication. Results Probl Cell Differ 34:197-217.

4. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331:280-3.

5. Wilson W, Braddock M, Adams SE, Rathjen PD, Kingsman SM, Kingsman AJ. 1988. HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. Cell 55:1159-69.

6. Pollard VW, Malim MH. 1998. The HIV-1 Rev protein. Annu Rev Microbiol 52:491-532.

7. Fornerod M, Ohno M, Yoshida M, Mattaj IW. 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. Cell 90:1051-60.

8. Fukuda M, Asano S, Nakamura T, Adachi M, Yoshida M, Yanagida M, Nishida E. 1997. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. Nature 390:308-11.

9. Behrens AJ, Crispin M. 2017. Structural principles controlling HIV envelope glycosylation. Curr Opin Struct Biol 44:125-133.

10. Gorelick RJ, Nigida SM, Jr., Bess JW, Jr., Arthur LO, Henderson LE, Rein A. 1990. Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. J Virol 64:3207-11.

11. von Poblotzki A, Wagner R, Niedrig M, Wanner G, Wolf H, Modrow S. 1993. Identification of a region in the Pr55gag-polyprotein essential for HIV-1 particle formation. Virology 193:981-5.

12. Ferguson MR, Rojo DR, von Lindern JJ, O'Brien WA. 2002. HIV-1 replication cycle. Clin Lab Med 22:611-35.

13. Spearman P. 2016. HIV-1 Gag as an Antiviral Target: Development of Assembly and Maturation Inhibitors. Curr Top Med Chem 16:1154-66.

14. Sundquist WI, Krausslich HG. 2012. HIV-1 assembly, budding, and maturation. Cold Spring Harb Perspect Med 2:a006924.

15. Freed EO. 2015. HIV-1 assembly, release and maturation. Nat Rev Microbiol 13:484-96.

16. Engelman AN, Singh PK. 2018. Cellular and molecular mechanisms of HIV-1 integration targeting. Cell Mol Life Sci 75:2491-2507.

17. Bowerman B, Brown PO, Bishop JM, Varmus HE. 1989. A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev 3:469-78.

18. Lesbats P, Engelman AN, Cherepanov P. 2016. Retroviral DNA Integration. Chem Rev 116:12730-12757.

19. Engelman AN. 2019. Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. J Biol Chem 294:15137-15157.

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

21. Jenkins TM, Engelman A, Ghirlando R, Craigie R. 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J Biol Chem 271:7712-8.

22. Wang JY, Ling H, Yang W, Craigie R. 2001. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J 20:7333-43.

23. Engelman A, Craigie R. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J Virol 66:6361-9.

24. 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. J Virol 69:2729-36.

25. Engelman A. 1999. In vivo analysis of retroviral integrase structure and function. Adv Virus Res 52:411-26.

26. Feng L, Larue RC, Slaughter A, Kessl JJ, Kvaratskhelia M. 2015. HIV-1 integrase multimerization as a therapeutic target. Curr Top Microbiol Immunol 389:93-119.

27. Kessl JJ, Eidahl JO, Shkriabai N, Zhao Z, McKee CJ, Hess S, Burke TR, Kvaratskhelia M. 2009. An allosteric mechanism for inhibiting HIV-1 integrase with a small molecule. Mol Pharmacol 76:824-32.

28. Grawenhoff J, Engelman AN. 2017. Retroviral integrase protein and intasome nucleoprotein complex structures. World J Biol Chem 8:32-44.

29. Maertens GN, Hare S, Cherepanov P. 2010. The mechanism of retroviral integration from X-ray structures of its key intermediates. Nature 468:326-9.

30. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. 2010. Retroviral intasome assembly and inhibition of DNA strand transfer. Nature 464:232-6.

31. Engelman AN, Cherepanov P. 2017. Retroviral intasomes arising. Curr Opin Struct Biol 47:23-29.

32. Passos DO, Li M, Yang R, Rebensburg SV, Ghirlando R, Jeon Y, Shkriabai N, Kvaratskhelia M, Craigie R, Lyumkis D. 2017. Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. Science 355:89-92.

33. Engelman A, Mizuuchi K, Craigie R. 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67:1211-21.

34. Roth MJ, Schwartzberg PL, Goff SP. 1989. Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence on IN function and terminal DNA sequence. Cell 58:47-54.

35. Katzman M, Katz RA, Skalka AM, Leis J. 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. J Virol 63:5319-27.

36. Sherman PA, Fyfe JA. 1990. Human immunodeficiency virus integration protein expressed in Escherichia coli possesses selective DNA cleaving activity. Proc Natl Acad Sci U S A 87:5119-23.

37. Lee YM, Coffin JM. 1991. Relationship of avian retrovirus DNA synthesis to integration in vitro. Mol Cell Biol 11:1419-30.

38. Pauza CD. 1990. Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. Virology 179:886-9.

39. Brown PO, Bowerman B, Varmus HE, Bishop JM. 1989. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. Proc Natl Acad Sci U S A 86:2525-9.

40. Fujiwara T, Mizuuchi K. 1988. Retroviral DNA integration: structure of an integration intermediate. Cell 54:497-504.

41. Skalka AM, Katz RA. 2005. Retroviral DNA integration and the DNA damage response. Cell Death Differ 12 Suppl 1:971-8.

42. Neves M, Peries J, Saib A. 1998. Study of human foamy virus proviral integration in chronically infected murine cells. Res Virol 149:393-401.

43. Schweizer M, Fleps U, Jackle A, Renne R, Turek R, Neumann-Haefelin D. 1993. Simian foamy virus type 3 (SFV-3) in latently infected Vero cells: reactivation by demethylation of proviral DNA. Virology 192:663-6.

44. Vincent KA, York-Higgins D, Quiroga M, Brown PO. 1990. Host sequences flanking the HIV provirus. Nucleic Acids Res 18:6045-7.

45. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell 110:521-9.

46. Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD. 2007. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Res 17:1186-94.

47. Marini B, Kertesz-Farkas A, Ali H, Lucic B, Lisek K, Manganaro L, Pongor S, Luzzati R, Recchia A, Mavilio F, Giacca M, Lusic M. 2015. Nuclear architecture dictates HIV-1 integration site selection. Nature 521:227-31.

48. Craigie R, Bushman FD. 2012. HIV DNA integration. Cold Spring Harb Perspect Med 2:a006890.

49. Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debyser Z. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J Biol Chem 278:372-81.

50. Emiliani S, Mousnier A, Busschots K, Maroun M, Van Maele B, Tempé D, Vandekerckhove L, Moisant F, Ben-Slama L, Witvrouw M, Christ F, Rain JC, Dargemont C, Debyser Z, Benarous R. 2005. Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. J Biol Chem 280:25517-23.

51. Cherepanov P, Devroe E, Silver PA, Engelman A. 2004. Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. J Biol Chem 279:48883-92.

52. Vanegas M, Llano M, Delgado S, Thompson D, Peretz M, Poeschla E. 2005. Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. J Cell Sci 118:1733-43.

53. Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA. 2012. Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS Genet 8:e1002717.

54. Eidahl JO, Crowe BL, North JA, McKee CJ, Shkriabai N, Feng L, Plumb M, Graham RL, Gorelick RJ, Hess S, Poirier MG, Foster MP, Kvaratskhelia M. 2013. Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Res 41:3924-36.

55. Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, Ecker JR, Bushman F. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. Nat Med 11:1287-9.

56. Marshall HM, Ronen K, Berry C, Llano M, Sutherland H, Saenz D, Bickmore W, Poeschla E, Bushman FD. 2007. Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. PLoS One 2:e1340.

57. Shun MC, Raghavendra NK, Vandegraaff N, Daigle JE, Hughes S, Kellam P, Cherepanov P, Engelman A. 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. Genes Dev 21:1767-78.

58. Maertens G, Cherepanov P, Pluymers W, Busschots K, De Clercq E, Debyser Z, Engelborghs Y. 2003. LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J Biol Chem 278:33528-39.

59. Summa V, Petrocchi A, Bonelli F, Crescenzi B, Donghi M, Ferrara M, Fiore F, Gardelli C, Gonzalez Paz O, Hazuda DJ, Jones P, Kinzel O, Laufer R, Monteagudo E, Muraglia E, Nizi E, Orvieto F, Pace P, Pescatore G, Scarpelli R, Stillmock K, Witmer MV, Rowley M. 2008. Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. J Med Chem 51:5843-55.

60. Ramanathan S, Mathias AA, German P, Kearney BP. 2011. Clinical pharmacokinetic and pharmacodynamic profile of the HIV integrase inhibitor elvitegravir. Clin Pharmacokinet 50:229-44.

61. Min S, Song I, Borland J, Chen S, Lou Y, Fujiwara T, Piscitelli SC. 2010. Pharmacokinetics and safety of S/GSK1349572, a next-generation HIV integrase inhibitor, in healthy volunteers. Antimicrob Agents Chemother 54:254-8.

62. Tsiang M, Jones GS, Goldsmith J, Mulato A, Hansen D, Kan E, Tsai L, Bam RA, Stepan G, Stray KM, Niedziela-Majka A, Yant SR, Yu H, Kukolj G, Cihlar T, Lazerwith SE, White KL, Jin H. 2016. Antiviral Activity of Bictegravir (GS-9883), a Novel Potent HIV-1 Integrase Strand Transfer Inhibitor with an Improved Resistance Profile. Antimicrob Agents Chemother 60:7086-7097.

63. Smith SJ, Zhao XZ, Burke TR, Jr., Hughes SH. 2018. Efficacies of Cabotegravir and Bictegravir against drug-resistant HIV-1 integrase mutants. Retrovirology 15:37.

64. Bukrinskaya AG. 2004. HIV-1 assembly and maturation. Arch Virol 149:1067-82.

65. Frankel AD, Young JA. 1998. HIV-1: fifteen proteins and an RNA. Annu Rev Biochem 67:1-25.

66. Adamson CS, Freed EO. 2010. Novel approaches to inhibiting HIV-1 replication. Antiviral Res 85:119-41.

67. Adamson CS, Salzwedel K, Freed EO. 2009. Virus maturation as a new HIV-1 therapeutic target. Expert Opin Ther Targets 13:895-908.

68. Briggs JA, Krausslich HG. 2011. The molecular architecture of HIV. J Mol Biol 410:491-500.

69. Ganser-Pornillos BK, Yeager M, Sundquist WI. 2008. The structural biology of HIV assembly. Curr Opin Struct Biol 18:203-17.

70. Chen J, Nikolaitchik O, Singh J, Wright A, Bencsics CE, Coffin JM, Ni N, Lockett S, Pathak VK, Hu WS. 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. Proc Natl Acad Sci U S A 106:13535-40.

71. Berkowitz R, Fisher J, Goff SP. 1996. RNA packaging. Curr Top Microbiol Immunol 214:177-218.

72. Berkowitz RD, Ohagen A, Hoglund S, Goff SP. 1995. Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric Gag polyproteins during RNA packaging in vivo. J Virol 69:6445-56.

73. Moore MD, Hu WS. 2009. HIV-1 RNA dimerization: It takes two to tango. AIDS Rev 11:91-102.

74. Kuzembayeva M, Dilley K, Sardo L, Hu W-S. 2014. Life of psi: How full-length HIV-1 RNAs become packaged genomes in the viral particles. Virology:362-370.

75. Clever JL, Wong ML, Parslow TG. 1996. Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. J Virol 70:5902-8.

76. Muriaux D, Fossé P, Paoletti J. 1996. A kissing complex together with a stable dimer is involved in the HIV-1Lai RNA dimerization process in vitro. Biochemistry 35:5075-82.

77. Batonick M, Favre M, Boge M, Spearman P, Honing S, Thali M. 2005. Interaction of HIV-1 Gag with the clathrin-associated adaptor AP-2. Virology 342:190-200.

78. Camus G, Segura-Morales C, Molle D, Lopez-Verges S, Begon-Pescia C, Cazevieille C, Schu P, Bertrand E, Berlioz-Torrent C, Basyuk E. 2007. The clathrin adaptor complex AP-1 binds HIV-1 and MLV Gag and facilitates their budding. Mol Biol Cell 18:3193-203.

79. Dong X, Li H, Derdowski A, Ding L, Burnett A, Chen X, Peters TR, Dermody TS, Woodruff E, Wang JJ, Spearman P. 2005. AP-3 directs the intracellular trafficking of HIV-1 Gag and plays a key role in particle assembly. Cell 120:663-74.

80. Martinez NW, Xue X, Berro RG, Kreitzer G, Resh MD. 2008. Kinesin KIF4 regulates intracellular trafficking and stability of the human immunodeficiency virus type 1 Gag polyprotein. J Virol 82:9937-50.

81. Ono A, Ablan SD, Lockett SJ, Nagashima K, Freed EO. 2004. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. Proc Natl Acad Sci U S A 101:14889-94.

82. Saad JS, Loeliger E, Luncsford P, Liriano M, Tai J, Kim A, Miller J, Joshi A, Freed EO, Summers MF. 2007. Point mutations in the HIV-1 matrix protein turn off the myristyl switch. J Mol Biol 366:574-85.

83. Saad JS, Miller J, Tai J, Kim A, Ghanam RH, Summers MF. 2006. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. Proc Natl Acad Sci U S A 103:11364-9.

84. Jouvenet N, Simon SM, Bieniasz PD. 2009. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. Proc Natl Acad Sci U S A 106:19114-9.

85. Yu X, Yuan X, McLane MF, Lee TH, Essex M. 1993. Mutations in the cytoplasmic domain of human immunodeficiency virus type 1 transmembrane protein impair the incorporation of Env proteins into mature virions. J Virol 67:213-21.

86. Wang D, Lu W, Li F. 2015. Pharmacological intervention of HIV-1 maturation. Acta Pharm Sin B 5:493-9.

87. Briggs JA, Simon MN, Gross I, Kräusslich HG, Fuller SD, Vogt VM, Johnson MC. 2004. The stoichiometry of Gag protein in HIV-1. Nat Struct Mol Biol 11:672-5.

88. Hill M, Tachedjian G, Mak J. 2005. The packaging and maturation of the HIV-1 Pol proteins. Curr HIV Res 3:73-85.

89. Ansari-Lari MA, Donehower LA, Gibbs RA. 1995. Analysis of human immunodeficiency virus type 1 integrase mutants. Virology 213:680.

90. Bukovsky A, Gottlinger H. 1996. Lack of integrase can markedly affect human immunodeficiency virus type 1 particle production in the presence of an active viral protease. J Virol 70:6820-5.

91. Kalpana GV, Reicin A, Cheng GS, Sorin M, Paik S, Goff SP. 1999. Isolation and characterization of an oligomerization-negative mutant of HIV-1 integrase. Virology 259:274-85.

92. 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. J Virol 70:721-8.

93. Liao WH, Wang CT. 2004. Characterization of human immunodeficiency virus type 1 Pr160 gag-pol mutants with truncations downstream of the protease domain. Virology 329:180-8.

94. Lu R, Ghory HZ, Engelman A. 2005. Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. J Virol 79:10356-68.

95. Lu R, Limon A, Devroe E, Silver PA, Cherepanov P, Engelman A. 2004. Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. J Virol 78:12735-46.

96. Nakamura T, Masuda T, Goto T, Sano K, Nakai M, Harada S. 1997. Lack of infectivity of HIV-1 integrase zinc finger-like domain mutant with morphologically normal maturation. Biochem Biophys Res Commun 239:715-22.

97. 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.

98. Shin CG, Taddeo B, Haseltine WA, Farnet CM. 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. J Virol 68:1633-42.

99. Taddeo B, Haseltine WA, Farnet CM. 1994. Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. J Virol 68:8401-5.

100. 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. J Virol 73:2126-35.

101. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, Gorelick RJ, Engelman AN, Steven AC. 2015. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. J Virol 89:9765-80.

102. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, Wang W, Ballandras-Colas A, Patel PA, Fuchs JR, Kvaratskhelia M, Engelman A. 2013. Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. Proc Natl Acad Sci U S A 110:8690-5.

103. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268 e12.

104. Busschots K, Voet A, De Maeyer M, Rain JC, Emiliani S, Benarous R, Desender L, Debyser Z, Christ F. 2007. Identification of the LEDGF/p75 binding site in HIV-1 integrase. J Mol Biol 365:1480-92.

105. Engelman A, Liu Y, Chen H, Farzan M, Dyda F. 1997. Structure-based mutagenesis of the catalytic domain of human immunodeficiency virus type 1 integrase. J Virol 71:3507-14.

106. Limon A, Devroe E, Lu R, Ghory HZ, Silver PA, Engelman A. 2002. Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import. J Virol 76:10598-607.

107. Lloyd AG, Ng YS, Muesing MA, Simon V, Mulder LC. 2007. Characterization of HIV-1 integrase N-terminal mutant viruses. Virology 360:129-35.

108. Lu R, Vandegraaff N, Cherepanov P, Engelman A. 2005. Lys-34, dispensable for integrase catalysis, is required for preintegration complex function and human immunodeficiency virus type 1 replication. J Virol 79:12584-91.

109. 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. J Virol 69:6687-96.

110. Rahman S, Lu R, Vandegraaff N, Cherepanov P, Engelman A. 2007. Structure-based mutagenesis of the integrase-LEDGF/p75 interface uncouples a strict correlation between in vitro protein binding and HIV-1 fitness. Virology 357:79-90.

111. Riviere L, Darlix JL, Cimarelli A. 2010. Analysis of the viral elements required in the nuclear import of HIV-1 DNA. J Virol 84:729-39.

112. Tsurutani N, Kubo M, Maeda Y, Ohashi T, Yamamoto N, Kannagi M, Masuda T. 2000. Identification of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient proviral DNA formation at steps prior to integration in dividing and nondividing cells. J Virol 74:4795-806.

113. Wiskerchen M, Muesing MA. 1995. Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. J Virol 69:376-86.

114. 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. J Virol 78:5045-55.

115. De Houwer S, Demeulemeester J, Thys W, Rocha S, Dirix L, Gijsbers R, Christ F, Debyser Z. 2014. The HIV-1 integrase mutant R263A/K264A is 2-fold defective for TRN-SR2 binding and viral nuclear import. J Biol Chem 289:25351-61.

116. Johnson BC, Metifiot M, Ferris A, Pommier Y, Hughes SH. 2013. A homology model of HIV-1 integrase and analysis of mutations designed to test the model. J Mol Biol 425:2133-46.

117. Mohammed KD, Topper MB, Muesing MA. 2011. Sequential deletion of the integrase (Gag-Pol) carboxyl terminus reveals distinct phenotypic classes of defective HIV-1. J Virol 85:4654-66.

118. Shehu-Xhilaga M, Hill M, Marshall JA, Kappes J, Crowe SM, Mak J. 2002. The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of pol protein. J Virol 76:4331-40.

119. Lutzke RA, Plasterk RH. 1998. Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding. J Virol 72:4841-8.

120. Lutzke RA, Vink C, Plasterk RH. 1994. Characterization of the minimal DNA-binding domain of the HIV integrase protein. Nucleic Acids Res 22:4125-31.

121. Balakrishnan M, Yant SR, Tsai L, O'Sullivan C, Bam RA, Tsai A, Niedziela-Majka A, Stray KM, Sakowicz R, Cihlar T. 2013. Non-catalytic site HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in target cells. PLoS One 8:e74163.

122. Desimmie BA, Schrijvers R, Demeulemeester J, Borrenberghs D, Weydert C, Thys W, Vets S, Van Remoortel B, Hofkens J, De Rijck J, Hendrix J, Bannert N, Gijsbers R, Christ F, Debyser Z. 2013. LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the virions. Retrovirology 10:57.

123. Elliott JL, Eschbach JE, Koneru PC, Li W, Puray-Chavez M, Townsend D, Lawson DQ, Engelman AN, Kvaratskhelia M, Kutluay SB. 2020. Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion morphogenesis. Elife 9.

124. Sharma A, Slaughter A, Jena N, Feng L, Kessl JJ, Fadel HJ, Malani N, Male F, Wu L, Poeschla E, Bushman FD, Fuchs JR, Kvaratskhelia M. 2014. A new class of multimerization selective inhibitors of HIV-1 integrase. PLoS Pathog 10:e1004171.

125. Gupta K, Brady T, Dyer BM, Malani N, Hwang Y, Male F, Nolte RT, Wang L, Velthuisen E, Jeffrey J, Van Duyne GD, Bushman FD. 2014. Allosteric inhibition of human immunodeficiency virus integrase: late block during viral replication and abnormal multimerization involving specific protein domains. J Biol Chem 289:20477-88.

126. Gupta K, Turkki V, Sherrill-Mix S, Hwang Y, Eilers G, Taylor L, McDanal C, Wang P, Temelkoff D, Nolte RT, Velthuisen E, Jeffrey J, Van Duyne GD, Bushman FD. 2016. Structural Basis for Inhibitor-Induced Aggregation of HIV Integrase. PLoS Biol 14:e1002584.

127. 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. Antimicrob Agents Chemother 56:4365-74.

128. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, Bardiot D, Van der Veken NJ, Van Remoortel B, Strelkov SV, De Maeyer M, Chaltin P, Debyser Z. 2010. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. Nat Chem Biol 6:442-8.

129. Maehigashi T, Ahn S, Kim UI, Lindenberger J, Oo A, Koneru PC, Mahboubi B, Engelman AN, Kvaratskhelia M, Kim K, Kim B. 2021. A highly potent and safe pyrrolopyridine-based allosteric HIV-1 integrase inhibitor targeting host LEDGF/p75-integrase interaction site. PLoS Pathog 17:e1009671.

130. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, de Silva S, Wu L, Le Grice SF, Engelman A, Fuchs JR, Kvaratskhelia M. 2012. Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. J Biol Chem 287:16801-11.

131. Le Rouzic E, Bonnard D, Chasset S, Bruneau JM, Chevreuil F, Le Strat F, Nguyen J, Beauvoir R, Amadori C, Brias J, Vomscheid S, Eiler S, Lévy N, Delelis O, Deprez E, Saïb A, Zamborlini A, Emiliani S, Ruff M, Ledoussal B, Moreau F, Benarous R. 2013. Dual inhibition of HIV-1 replication by integrase-LEDGF allosteric inhibitors is predominant at the post-integration stage. Retrovirology 10:144.

132. Tsiang M, Jones GS, Niedziela-Majka A, Kan E, Lansdon EB, Huang W, Hung M, Samuel D, Novikov N, Xu Y, Mitchell M, Guo H, Babaoglu K, Liu X, Geleziunas R, Sakowicz R. 2012. New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. J Biol Chem 287:21189-203.

133. Shema Mugisha C, Tenneti K, Kutluay SB. 2020. Clip for studying protein-RNA interactions that regulate virus replication. Methods 183:84-92.

134. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268.e12.

135. Elliott JL, Kutluay SB. 2020. Going beyond Integration: The Emerging Role of HIV-1 Integrase in Virion Morphogenesis. Viruses 12.

136. Madison MK, Lawson DQ, Elliott J, Ozanturk AN, Koneru PC, Townsend D, Errando M, Kvaratskhelia M, Kutluay SB. 2017. Allosteric HIV-1 Integrase Inhibitors Lead to Premature Degradation of the Viral RNA Genome and Integrase in Target Cells. J Virol 91.

137. Engelman AN. 2019. Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. J Biol Chem doi:10.1074/jbc.REV119.006901.

138. Hazuda DJ, Felock P, Witmer M, Wolfe A, Stillmock K, Grobler JA, Espeseth A, Gabryelski L, Schleif W, Blau C, Miller MD. 2000. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science 287:646-50.

139. Chen J, Rahman SA, Nikolaitchik OA, Grunwald D, Sardo L, Burdick RC, Plisov S, Liang E, Tai S, Pathak VK, Hu WS. 2016. HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. Proc Natl Acad Sci U S A 113:E201-8.

140. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220:868-71.

141. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, Palker TJ, Redfield R, Oleske J, Safai B, et al. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science 224:500-3.

142. Rockstroh JK, DeJesus E, Lennox JL, Yazdanpanah Y, Saag MS, Wan H, Rodgers AJ, Walker ML, Miller M, DiNubile MJ, Nguyen BY, Teppler H, Leavitt R, Sklar P, Investigators S. 2013. Durable efficacy and safety of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naive HIV-1-infected patients: final 5-year results from STARTMRK. J Acquir Immune Defic Syndr 63:77-85.

143. Taha H, Das A, Das S. 2015. Clinical effectiveness of dolutegravir in the treatment of HIV/AIDS. Infect Drug Resist 8:339-52.

144. Smith SJ, Zhao XZ, Passos DO, Lyumkis D, Burke TR, Hughes SH. 2020. HIV-1 Integrase Inhibitors that are active against Drug-Resistant Integrase Mutants. Antimicrob Agents Chemother doi:10.1128/AAC.00611-20.

145. Brooks KM, Sherman EM, Egelund EF, Brotherton A, Durham S, Badowski ME, Cluck DB. 2019. Integrase Inhibitors: After 10 Years of Experience, Is the Best Yet to Come? Pharmacotherapy 39:576-598.

146. Rossouw TM, Hitchcock S, Botes M. 2016. The end of the line? A case of drug resistance to third-line antiretroviral therapy. South Afr J HIV Med 17:454.

147. Engelman A, Cherepanov P. 2014. Retroviral Integrase Structure and DNA Recombination Mechanism. Microbiol Spectr 2.

148. Engelman A, Cherepanov P. 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat Rev Microbiol 10:279-90.

149. Radzio-Basu J, Council O, Cong ME, Ruone S, Newton A, Wei X, Mitchell J, Ellis S, Petropoulos CJ, Huang W, Spreen W, Heneine W, García-Lerma JG. 2019. Drug resistance emergence in macaques administered cabotegravir long-acting for pre-exposure prophylaxis during acute SHIV infection. Nat Commun 10:2005.

150. Anstett K, Brenner B, Mesplede T, Wainberg MA. 2017. HIV drug resistance against strand transfer integrase inhibitors. Retrovirology 14:36.

151. Christ F, Debyser Z. 2013. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. Virology 435:102-9.

152. Darlix JL, de Rocquigny H, Mauffret O, Mely Y. 2014. Retrospective on the all-in-one retroviral nucleocapsid protein. Virus Res 193:2-15.

153. Mattei S, Schur FK, Briggs JA. 2016. Retrovirus maturation-an extraordinary structural transformation. Curr Opin Virol 18:27-35.

Chapter 2: HIV-1 integrase binding to the genomic RNA is mediated by electrostatic interactions

Christian Shema Mugisha[1], Tung Dinh[2], Kasyap Tenneti[1], Jenna E. Eschbach[1], Keanu Davis[1], Robert Gifford[3], Mamuka Kvaratskhelia[2], Sebla B. Kutluay[1]

Under review in Journal of Virology

Affiliations:

[1] Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110, USA

[2] Division of Infectious Diseases, University of Colorado School of Medicine, Aurora, CO 80045

[3] MRC-University of Glasgow Centre for Virus Research, 464 Bearsden Rd., Bearsden, Glasgow G61 1QH, UK.

Contributions:

Christian Shema Mugisha and Sebla B. Kutluay made the hypothesis and designed the experiments. Tung Dinh perfomed the biochemical assays including SEC, integrase invitro binding assays, and structural modeling of IN. Robert Gifford did the aligned of integrase mutations found in latent CD4+ T cells.

ACKNOWLEDGEMENTS

This work was supported by NIH grants AI150497 (SBK), AI1508470 (U54 Center for HIV RNA Studies, SBK, RG), R01 AI143649 (MK) and Milton Schlesinger Student Fellowship (CSM). We thank all members of the Kutluay lab for critical suggestions and feedback.

2.1 Abstract

Independent of its catalytic activity, HIV-1 integrase (IN) enzyme regulates proper particle maturation by binding to and packaging the viral RNA genome (gRNA) inside the mature capsid lattice. Allosteric integrase inhibitors (ALLINIs) and class II IN substitutions inhibit the binding of IN to the gRNA and cause the formation of non-infectious virions characterized by mislocalization of the viral ribonucleoprotein complexes between the translucent conical capsid lattice and the viral lipid envelope. To gain insight into the molecular nature of IN-gRNA interactions, we have isolated compensatory substitutions in the background of a class II IN (R269A/K273A) variant that directly inhibits IN binding to the gRNA. We found that additional D256N and D270N substitutions in the C-terminal domain (CTD) of IN restored its ability to bind gRNA and led to the formation of infectious particles with correctly matured morphology. Furthermore, reinstating the overall positive electrostatic potential of the CTD through individual D256R or D256K substitutions was sufficient to restore IN-RNA binding and infectivity for the R269A/K273A as well as the R262A/R263A class II IN mutants. The compensatory mutations did not impact functional IN oligomerization, suggesting that they directly contributed to IN binding to the gRNA. Interestingly, HIV-1 IN R269A/K273A, but not IN R262A/R263A, bearing compensatory mutations was more sensitive to ALLINIs providing key genetic evidence that specific IN residues required for RNA binding also influence ALLINI activity. Structural modeling provided further insight into the molecular nature of IN-gRNA interactions and ALLINI mechanism of action. Taken together, our findings highlight an essential role of IN-gRNA interactions for proper virion maturation and reveal the importance of electrostatic interactions between the IN CTD and the gRNA.

2.2 Importance

HIV-1 integrase (IN) binds to the viral RNA genome through basic residues present within its C-terminal domain (CTD) and regulates proper virion maturation. Inhibition of IN binding to the HIV-1 genome through mutations of CTD basic residues results in non-infectious particles in which the viral genomes are mislocalized in virions. Here we isolated suppressor mutations that restored the ability to bind RNA to an IN-CTD mutant lacking two basic residues. We found that mutation of nearby acidic residues (i.e. D256 and D270) and restoring the overall charge of the CTD can restore RNA binding, particle maturation and virion infectivity, but sensitize the viruses to the inhibitory action of allosteric integrase inhibitors (ALLINIs) that target IN-RNA interactions. Taken together, our findings highlight the electrostatic nature of IN-RNA interactions mediated by the IN CTD and the gRNA phosphate backbone and provide key genetic evidence that the ALLINIs engage the CTD of IN.

2.3 Introduction

A defining feature of retroviruses is the reverse transcription of the viral RNA genome (gRNA) and integration of the resultant linear viral DNA into a host chromosome, which establishes lifelong infection. The latter reaction is mediated by the viral integrase (IN) enzyme, which catalyzes 3' processing and DNA strand transfer reactions . The catalytic activity of HIV-1 IN has been successfully targeted by several integrase strand-transfer inhibitors (INSTIs) that have become key components of frontline anti-retroviral therapy regimens due to their high efficacy and tolerance profiles . In addition, HIV-1 IN has an emerging non-catalytic function in virus replication (4, 17). Successful targeting of this second function can complement the existing antiviral regimens and substantially increase the barrier to INSTI resistance.

HIV-1 IN consists of three independently folded protein domains: the N-terminal domain (NTD) bears the conserved His and Cys residues (HHCC motif) that coordinate Zn^{2+} binding for 3-helix bundle formation; the catalytic core domain (CCD) adopts an RNase H fold and harbors the enzyme active site composed of an invariant DDE motif, and C-terminal domain (CTD) which adopts an SH3 fold . Integration is facilitated by a cellular co-factor, lens epithelium-derived growth factor (LEDGF/p75), which binds tightly to a site within the CCD dimer interface and guides the preintegration complex to actively transcribed regions of the host chromosome . A group of pleotropic IN substitutions distributed throughout IN, collectively known as class II mutations, disrupt viral assembly , morphogenesis and reverse transcription in target cells often without obstructing the catalytic activity of IN in vitro . A hallmark of class II IN mutant viruses is the mislocalization of the viral ribonucleoprotein complexes (vRNP) outside of the viral capsid (CA) lattice, a deformation which is often referred to as eccentric morphology . Although originally designed to inhibit integration through preventing the binding of IN to LEDGF/p75 ,

allosteric integrase inhibitors (ALLINIs) potently inhibit proper virion maturation (60) and lead to the formation of virions that display a similar eccentric morphology observed with class II IN mutations .

We have recently shown that binding of IN to the gRNA in mature virions accounts for the non-catalytic function of IN in virus replication and is required for proper encapsidation of the viral ribonucleoproteins (vRNPs) inside the mature CA lattice (63-65). Class II IN substitutions and ALLINIs block IN-gRNA binding, thus causing formation of virions with an eccentric morphology . IN binds to multiple distinct locations on the gRNA, including the TAR hairpin present within the 5' and 3'UTRs of the gRNA and constitutes a high affinity binding site (63, 64, 66). IN preferentially binds to the gRNA in a tetrameric state, and many class II IN mutations block IN-gRNA binding by disrupting the functional oligomerization of IN . Mutation of basic residues within IN-CTD (i.e. R262, R263, R269 and K273) inhibits IN-gRNA binding without altering functional IN oligomerization in virions and in vitro , suggesting that these residues are directly involved in binding to the gRNA. On the other hand, the precise mechanism of how these residues mediate recognition of specific sequence elements on the gRNA remains unknown. For example, it is possible that the positively charged Lys and Arg residues interact with the negatively charged RNA phosphate backbone in a non-specific or semi-specific manner, depending on the folding and structure of the cognate RNA element, driven by electrostatic interactions . Alternatively, these residues can mediate specific interactions with RNA targets through H-bonding and van der Waals contacts with individual nucleobases .

To gain insight into the mode of IN-gRNA interactions, the non-infectious IN R269A/K273A class II mutant virus was serially passaged in T-cells until the acquisition of compensatory mutations. We found that two compensatory mutations, D256N and D270N, within the IN-CTD

restored virion infectivity, IN-RNA interactions and accurate virion morphogenesis. As the D-to-N mutations resulted in loss of two negative charges, possibly overcoming the loss of two positive charges with the R269A/K273A class II substitutions, we tested whether restoring the overall charge of CTD through other mutations would restore IN-gRNA binding and virion infectivity. Indeed, the D256R substitution alone restored virion infectivity and RNA binding for the IN R269A/K273A mutant. We further extended these findings to another class II mutant, R262A/R263A, which was similarly suppressed by the D256R as well as D256K substitutions. Compensatory mutations did not affect the ability of IN to multimerize *in vitro* or in virions, suggesting that they restored the RNA-binding ability of IN directly. Interestingly, the IN R269A/K273A, but not the IN R262A/R263A, mutant viruses bearing the compensatory mutations had increased sensitivity to ALLINIs, providing key genetic evidence that specific residues within the CTD, which are required for RNA binding, also contribute to ALLINI mechanism of action. Together, our findings strongly suggest that IN-RNA interactions are at least in part driven by electrostatic interactions between the basic residues within IN-CTD and phosphate backbone of the gRNA and highlight that ALLINIs engage CTD residues to inhibit RNA binding.

2.4 Results

2.4.1. Compensatory IN D256N/D270N substitutions emerge in the background of the IN R269A/K273A class II mutant virus

The R269A/K273A class II IN substitutions obstruct IN-gRNA binding directly without interfering with IN multimerization and result in the formation of particles with eccentric morphology (73). To

better understand the molecular basis of how R269 and K273 residues mediate IN-gRNA binding, viruses bearing the IN R269A/K273A mutations were serially passaged in MT-4 cells until the emergence of compensatory mutations at the end of passage 3, which completely restored virion infectivity (**Fig. 1A**). Deep sequencing of full-length gRNA isolated from virions across the three passages revealed that viruses retained the IN R269A/K273A substitutions while sequentially acquiring D256N and D270N mutations in IN (**Fig. 1B**). The D256N substitution emerged at the end of passage 1 whereas the D270N mutation emerged later in passage 2 (**Fig. 1B**). Both mutations were fixed by the end of passage 2 and no other mutations were observed elsewhere on the viral genome.

The IN D256N and D270N mutations were introduced into the replication-competent pNL4-3 molecular clone bearing WT or R269A/K273A IN. Introduction of D256N and D270N substitutions in IN had no observable effect on Gag (Pr55) expression or processing in cells or particle release (**Fig. 1C**). Introduction of D256N and D270N mutations either individually or together (D2N) on the WT IN backbone did not affect viral titers (**Fig. 1D**). Remarkably, while the individual D256N and D270N mutations in the R269A/K273A IN backbone increased virus titers by 5-10-fold, the D2N substitutions increased virus titers by 100-fold (**Fig. 1D**). Overall, these results demonstrate that the combination of D256N and D270N mutations are sufficient to restore the replication competency of the R269A/K273A class II IN mutant virus.

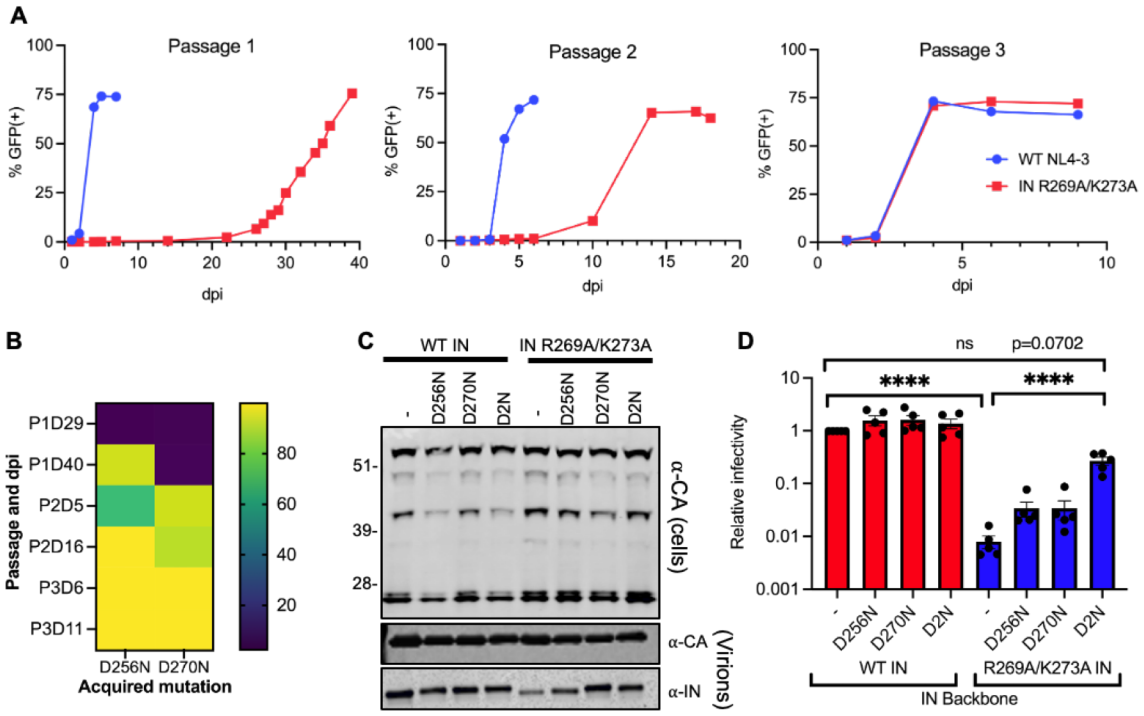

Figure 1. D256N and D270N substitutions in HIV-1 IN suppresses the replication defect of R269A/K273A class II IN mutant virus. (A) MT4-LTR-GFP cells were infected with equal particle numbers of either WT or R269A/K273A IN mutant HIV-1 (NL4-3). The R269A/K273A IN viruses were serially passaged for three times until the emergence of compensatory mutations. The graphs represent the percentage of GFP positive cells as assessed by FACS over three passages at the indicated days post-infection (dpi). **(B)** HIV-1 genomic RNA was isolated from viruses collected from cell culture supernatants over the three passages (i.e. P1, P2, P3) and at the indicated days post-infection (i.e. D29, D40, etc.). Whole-genome deep sequencing revealed the acquisition of D256N and D270N compensatory mutations. Heatmap shows the percentage of mutations at the indicated passages and days post-infection (dpi). **(C)** HEK293T cells were transfected with full-length pNL4-3 expression plasmids carrying the D256N, D270N and D256N/D270N (D2N) IN mutations introduced on the WT IN and IN R269A/K273A backbones. Cell lysates and virions were purified two days post transfection and analyzed by immunoblotting for CA and IN. The image is representative of five independent experiments. **(D)** HEK293T cells were transfected as in C and cell culture supernatants containing viruses were titered on TZM-bl indicator cells. The titers are presented relative to WT (set to 1). The columns represent the average of five independent experiments and the error bars represent SEM (****p<0.0001, by one-way ANOVA with Dunnett's multiple comparison test).

2.4.2 D256N/D270N substitutions restore IN-gRNA binding for the R269A/K273A class II IN mutant viruses and lead to formation of correctly matured virions

We next assessed whether D256N and D270N substitutions rendered the R269A/K273A class II IN mutant virus replication competent by restoring IN-gRNA binding and proper virion maturation. To this end, IN-gRNA complexes were immunoprecipitated from UV-crosslinked virions and visualized per CLIP protocol as described previously (75). Equivalent amounts of immunoprecipitated IN from WT and R269A/K273A IN viruses, or those additionally bearing the IN D256N and D270N substitutions, were analyzed for their ability to bind RNA. IN-RNA complexes were readily visible for WT viruses as well as D256N, D270N and D2N mutants (introduced on the WT backbone), but not from the R269A/K273A class II IN mutant viruses (**Fig. 2A**). Introduction of the D256N substitution on R269A/K273A IN backbone modestly enhanced the ability of IN to bind RNA, whereas the D270N substitution had no observable impact (**Fig. 2A**). In contrast, the D2N substitution substantially enhanced the ability of R269A/K273A IN to bind RNA (**Fig. 2A**). The virion morphology of WT and IN mutant viruses was assessed by transmission electron microscopy (TEM). As expected, more than 80% of WT particles had an electron-dense condensate that represents vRNPs inside the CA lattice, whereas the majority of R269A/K273A class II IN mutant virions (~68%) had a clear eccentric morphology (**Fig. 2B**). Consistent with effects on virus titers and RNA-binding, the introduction of D256N/D270N substitutions restored the ability of the R269A/K273A IN mutant to form properly mature virions (**Fig. 2B**). Cumulatively, these data show that D256N and D270N IN substitutions restore infectivity for the R269A/K273A class II IN mutant virus through reestablishing RNA binding and subsequently accurate virion maturation.

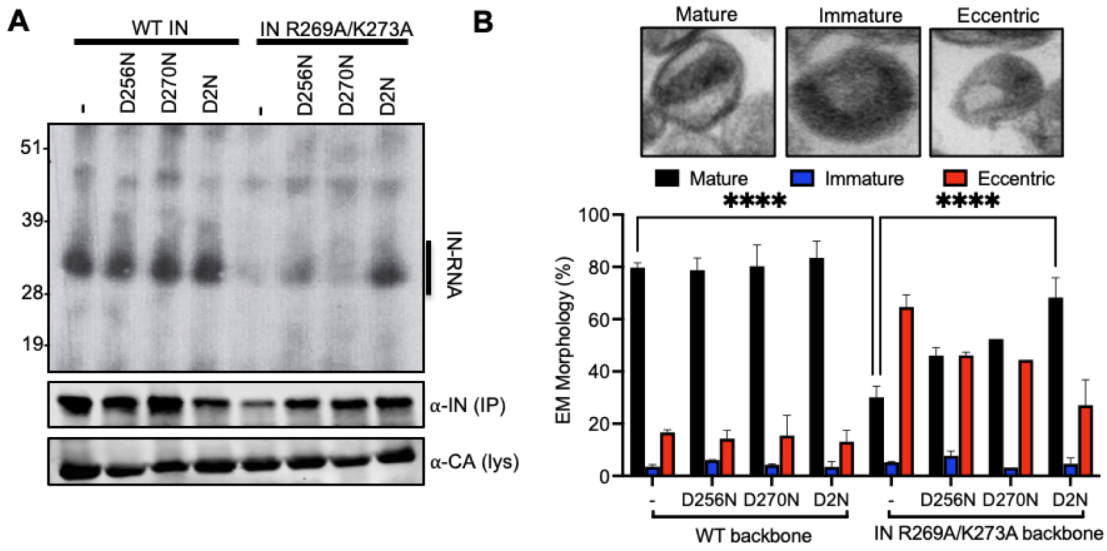

Figure 2. D256N and D270N substitutions restore IN-gRNA binding and accurate virion maturation for the R269A/K273A class II IN mutant virus. (A) Autoradiogram of IN-RNA adducts immunoprecipitated from virions bearing the indicated substitutions in IN. Immunoblots below show the amount of IP'ed IN or CA protein in lysates. Data shown are representative of at least three independent experiments. (B) Examination of virion maturation in WT and mutant IN viruses by thin section electron microscopy (TEM). The graph quantifies virion morphologies; each column is the average of two independent experiments and error bars represent the SEM (****$p<0.0001$, by repeated measures one-way ANOVA).

2.4.3 Electrostatic interactions are required for IN-gRNA binding.

IN-CTD is decorated with several acidic and basic amino acids resulting in a net charge of +3 (**Fig. 3A**). The D2N substitutions in effect restored the overall charge of the IN R269A/K273A CTD, suggesting that the net electrostatic charge of the IN-CTD may be a key parameter in gRNA binding. To test this hypothesis, we investigated whether restoring the overall charge of IN-CTD through other mutations would also restore RNA binding and infectivity for the R269A/K273A IN virus. We focused our analysis on D256 and D270 residues for the following reasons: i) these amino acids were amenable to substitutions during virus passaging experiments; ii) D256N and D270N substitutions in the context of WT HIV-1 yielded infectious virions, suggesting that these

compensatory mutations do not significantly contribute to the catalytic activity of IN thus allowing us to specifically probe their roles for IN-RNA interactions and virion maturation. To extend these studies we introduced D256R, D270R and D256R/D270R (D2R) substitutions into the HIV-$1_{NL4-3\ IN\ R269A/K273A}$ backbone and transfected HEK293T cells with the resulting plasmids. Cell lysates and cell-free virions were then analyzed for Gag processing, particle release, and infectivity. Overall, D-to-R substitutions had no major effect on Gag (Pr55) expression or processing in cells (**Fig. 3B**).

Remarkably, the D256R substitution alone increased the titers of the R269A/K273A class II IN mutant virus by 100-fold and to a level comparable to that of WT viruses (**Fig. 3C**). In contrast, D270R substitution only had a modest impact on virus titers and the D2R substitution increased virion infectivity at a modestly lower level than that of D256R and D2N (**Fig. 3C**). Importantly, the D256R and the D2R substitutions completely restored IN-gRNA binding (**Fig. 3D**), demonstrating that the increase in viral titers with the D256R and D2R substitutions correlates well with enhancement of IN-gRNA binding. Further assessment of these substitutions with Vpr-IN transcomplementation assays revealed that none of the IN substitutions impacted the catalytic activity of IN, except for the D2R substitution (**Fig. 3E**), suggesting that the inability of the D270R mutation to restore infectivity may be due to its adverse effects on the catalytic activity of IN. Taken together, these findings indicate that restoring the overall charge of IN-CTD through D256R mutation reestablishes virion infectivity and RNA binding for the IN R269A/K273A class II mutant.

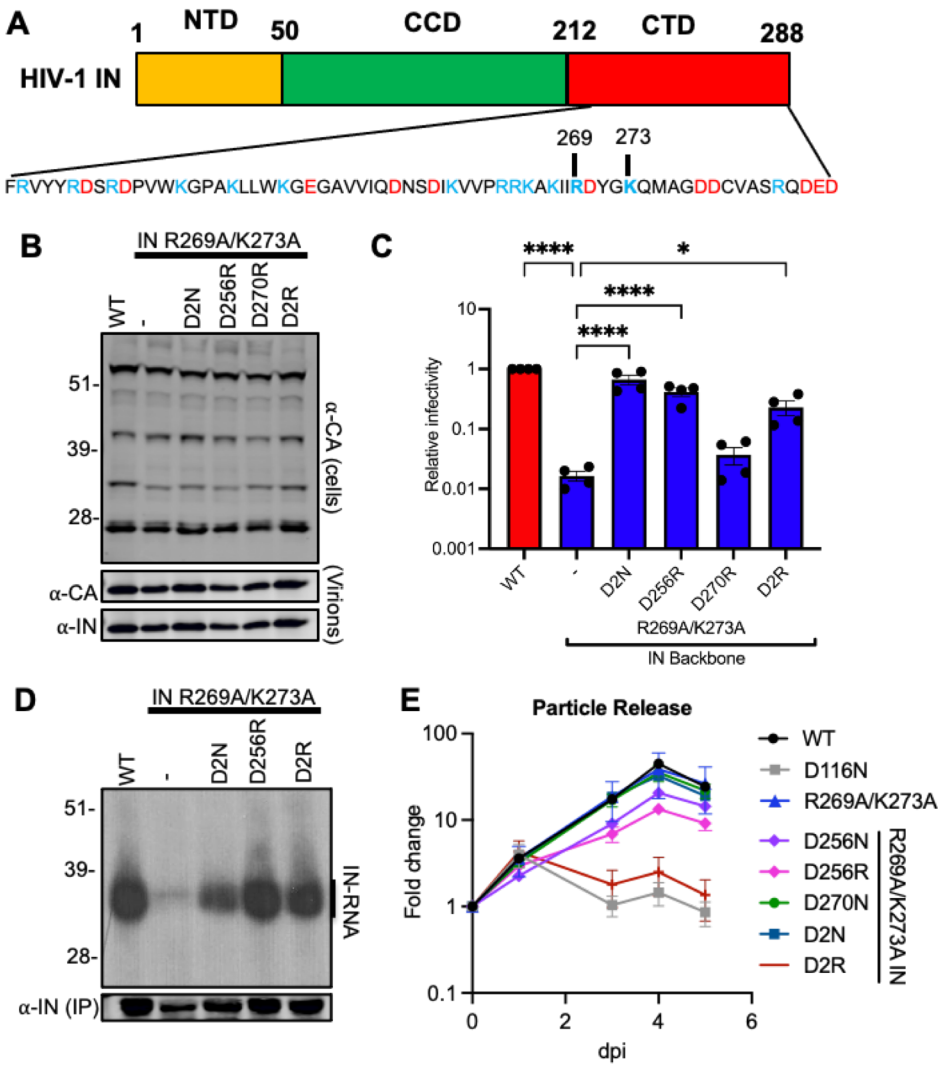

Figure 3. Restoring the net charge of IN-CTD restores RNA binding and infectivity for the R269A/K273A class II mutant. (A) Schematic diagram of IN and sequence of CTD residues with basic and acidic amino acids highlighted in blue and red, respectively. **(B, C)** HEK293T cells were transfected with full-length proviral HIV-1_{NL4-3} expression plasmids carrying *pol* mutations encoding for the indicated IN substitutions. **(B)** Cell lysates and purified virions were harvested two days post transfection and analyzed by immunoblotting for CA and IN. Representative image of one of five independent experiments is shown. **(C)** Infectious titers of WT or IN mutant HIV-1_{NL4-3} viruses in cell culture supernatants were determined on TZM-bl indicator cells and normalized relative to particle number based on RT activity. Titer values are expressed relative to WT (set to 1). The columns represent the average of four independent experiments and the error bars represent SEM (****p<0.0001, by one-way ANOVA with Dunnett's multiple comparison test). **(D)** Autoradiogram of IN-RNA adducts IP'ed from WT or IN mutant HIV-1_{NL4-3} virions. The amount of immunoprecipitated IN protein was assessed by the immunoblot below. Results are representative of four independent replicates. **(E)** A representative growth curve of HIV-$1_{NL4-3\ IN\ (D116N)}$ viruses that were trans-complemented with the indicated Vpr-IN mutant proteins in cell culture. Y-axis indicates fold increase in virion yield over day 0 as measured by RT activity in culture supernatants. Error bars show SEM from 3 independent many replicates.

2.4.4 D256R substitution fully restores virion infectivity and RNA binding for a separate class II IN (R262A/R263A) mutant

We have previously shown that mutation of other basic residues within the IN-CTD (i.e. R262A/R263A) also directly inhibit IN-gRNA binding without compromising functional oligomerization of IN (78). We next wanted to extend our observations and test how substitutions of D256 and D270 residues affect IN-gRNA binding and infectivity in the background of the IN R262A/R263A mutant virus. Introducing the D256N, D270N and D2N compensatory mutations on the IN R262A/R263A backbone did not affect Gag expression, processing, and virion release (**Fig. 4A**). While the D256N substitution increased virion infectivity by 10-fold, the D270N substitution had no impact and the D2N substitution had an intermediate phenotype (**Fig. 4B**), demonstrating the context dependency of these compensatory mutations. The D256K and D256R substitutions introduced in the R262A/R263A IN backbone also had minimal effects on Gag expression, processing, and virion release (**Fig. 4C**). In contrast, the D256K substitution increased viral titers at a greater degree than D256N and the D256R substitution completely restored virion infectivity (**Fig. 4D**). In line with the titer data, D256K significantly increased RNA binding whereas the D256R completely restored it in the context of the IN R262A/R263A mutant virus (**Fig 4E**). D256N, D256K and D256R IN successfully transcomplemented a class I IN mutant, suggesting that they did not distort the catalytic activity of IN (**Fig 4F**). Taken together, these findings indicate that restoring the overall positive charge of IN-CTD through D256R or D256K mutations is sufficient to restore virion infectivity and RNA binding to class II IN mutants that directly inhibit RNA binding.

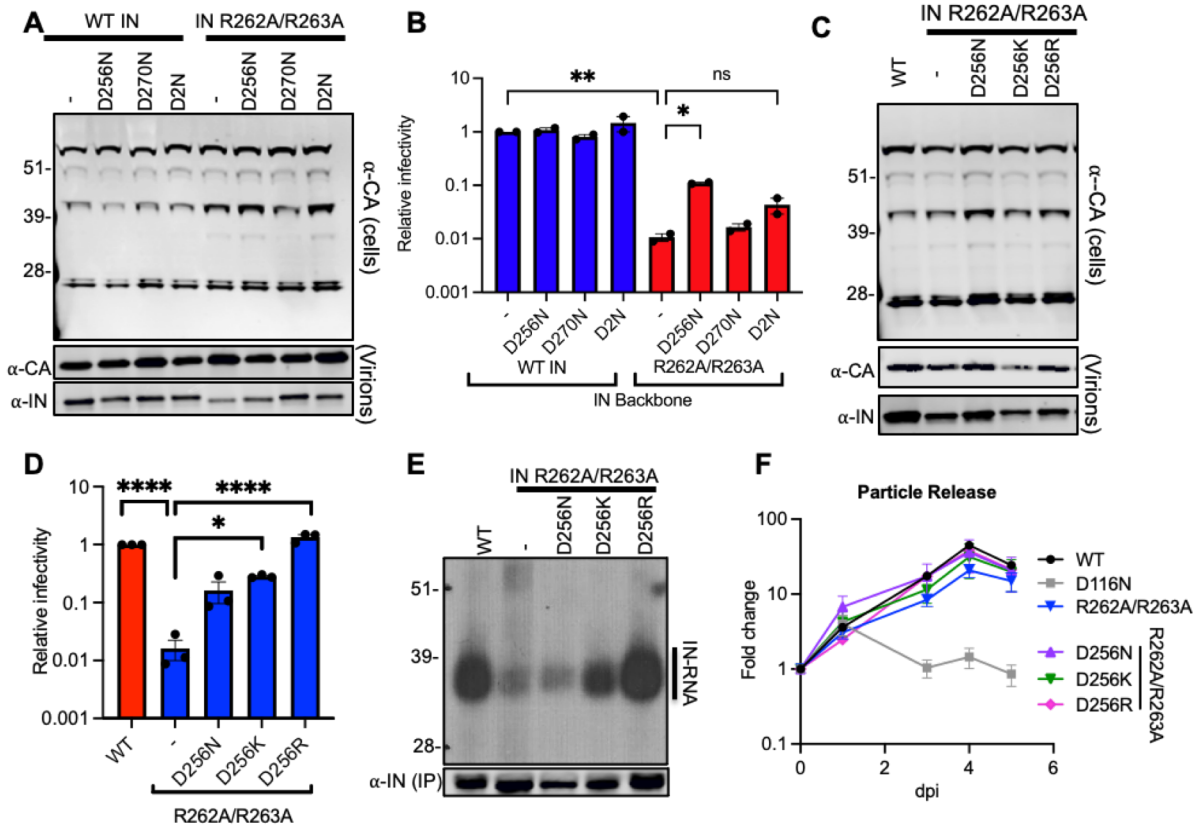

Figure 4. D256R and D256K substitutions restore IN-RNA binding and infectivity for the HIV-$1_{NL4-3\ IN(R262A/R263A)}$ virus. (A-E) HEK293T cells were transfected with proviral HIV-1_{NL4-3} expression plasmids carrying *pol* mutations for the indicated IN substitutions. **(A, C)** Cell lysates and virions were purified two days post transfection and analyzed by immunoblotting for CA and IN. **(B, D)** WT or IN mutant HIV-1_{NL4-3} viruses in cell culture supernatants were titered on TZM-bl indicator cells. The titer values are presented relative to WT (set to 1). The columns represent the average of two-three independent experiments, and the error bars represent SEM (****p<0.0001, by one-way ANOVA with Dunnett's multiple comparisons). **(E)** Autoradiogram of IN-RNA adducts immunoprecipitated from WT or IN mutant HIV-1_{NL4-3} virions. The amount of immunoprecipitated IN was assessed by the immunoblot shown below. Immunoblots and CLIP autoradiographs are representative of four independent replicates. **(F)** A representative growth curve of HIV-$1_{NL4-3\ IN\ (D116N)}$ viruses that were trans-complemented with the indicated IN mutant proteins in cell culture. Y-axis indicates fold increase in virion yield over day 0 as measured by RT activity in culture supernatants. Error bars show SEM from 3 independent many replicates.

We next used reported X-ray structure and molecular modelling to visualize how class II and compensatory IN mutations affect electrostatic potential of the CTD surface. Mutation of R269 and K273 residues expectedly resulted in a substantial loss of a basic patch in IN (**Fig. 5A, B**). Both the D256R and D256N substitutions resulted in more positively charged surface distal from

the R269/K273A residues (**Fig. 5C, D**), suggesting that compensatory mutations likely created an additional interacting interface with gRNA binding. A similar outcome was observed when class II R262A/R263A and compensatory D256R changes were introduced in the CTD (**Fig. 5E-G**). While the D256R change is substantially distanced from R262/R263 and R269/K273 residues, we note the following. R262, R263, R269, and K273 are positioned within the same highly flexible C-terminal tail (aa 261-275), whereas D256 belongs to another, shorter (aa 252-257) loop (**Fig. 5H**). The highly pliable nature of the tail and the loop could be crucial for IN to optimally engage with cognate RNA as well as allow for emergence of compensatory mutations at alternative sites positioned in these IN segments.

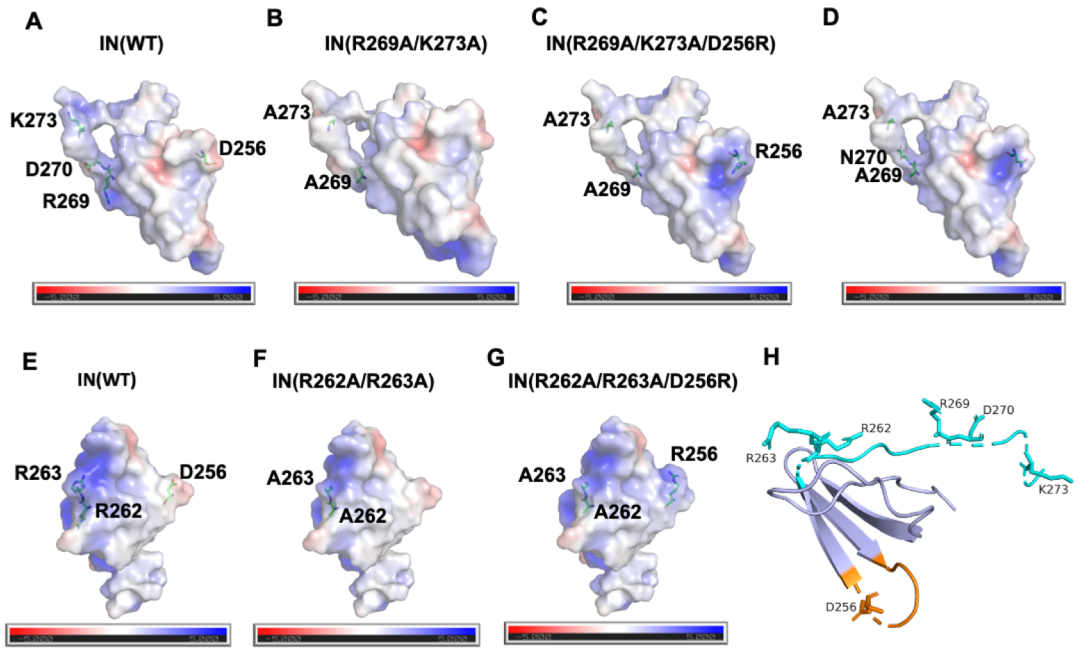

Figure 5. Electrostatic potential maps of HIV-1 IN bearing class II and compensatory mutations. **(A-G)** Electrostatic potential maps of the indicated HIV-1 IN mutants derived from the crystal structure of Gupta et al. is depicted. Calculation results are displayed as an electrostatic potential molecular surface. The low, mid, and high range values are -5, 0, and 5, respectively. (**H**) A cartoon view of the CTD structure is shown with the C-terminal tail (aa 261-275) and the loop (aa 252-257) colored in cyan and orange, respectively. Side chains of indicated residues are shown.

2.4.5 Effects of the compensatory mutations on functional oligomerization of IN and IN-RNA interactions.

As functional IN oligomerization is a prerequisite for RNA binding (13), we next examined how the compensatory substitutions affected IN oligomerization. For in virion analysis, purified HIV-1_{NL4-3} virions were treated with ethylene glycol bis (succinimidyl succinate) (EGS) to covalently crosslink IN in situ and virus lysates were analyzed by immunoblotting. As in WT virions, IN species that migrated at molecular weights consistent with those of monomers, dimers, trimers, and tetramers were readily distinguished in R269A/K273A or R262A/R263A viruses with additional compensatory mutations (D256N/D270N, D256R, D256K, and D256R/D270R) but not with the canonical class II IN mutant V165A that is unable to form functional oligomers (**Fig 6A-C**). Complementary in vitro assessment of purified recombinant IN proteins by size exclusion chromatography (SEC) revealed that D2N and D256R substitutions in the background of R269A/K273A or the R262A/R263A class II mutants did not impact functional IN tetramerization (**Fig. 6D-H**).

We have previously shown that recombinant IN binds to TAR RNA with high affinity and provides a nucleation point to bridge and condense RNA (43). We next examined the ability of class II mutant INs bearing compensatory mutations to bind and bridge TAR RNA. Consistent with findings from CLIP, the D2N and D256R substitutions also enhanced or restored the ability of R269A/K273 and R262A/R263A mutants to bridge between RNA molecules (**Fig. 6I**). Together, these data demonstrate that the compensatory mutations directly restore the ability of IN to bind RNA without altering functional IN oligomerization.

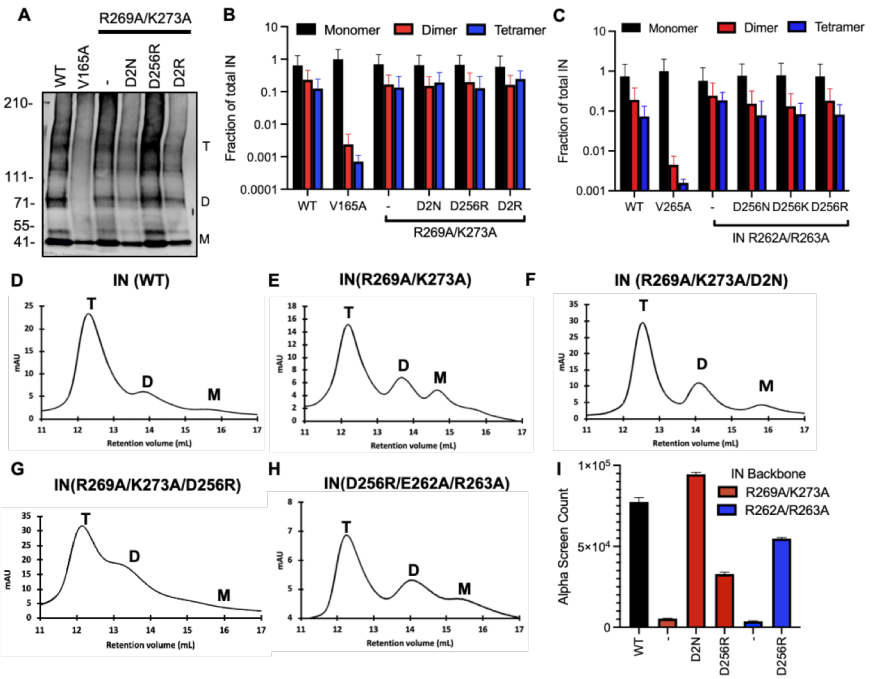

Figure 6. Assessing multimerization properties of IN in mutant viruses. (A) Purified WT or IN mutant HIV-1_{NL4-3} virions were treated with 1 mM EGS, and virus lysates analyzed by immunoblotting using antibodies against IN following separation on 6% Tris-acetate gels. The position of monomers (M), dimers (D), and tetramers (T) are indicated by arrows in a representative western blot. **(B,C)** Quantification of IN multimerization in virions from experiments conducted as in A. Error bars show the SEM from three independent experiments. **(D-I)** Biochemical analysis of IN mutants multimerization and their interactions with HIV-1 TAR RNA. **(D-H)** Representative SEC traces for indicated recombinant IN proteins. The X-axis indicates elution volume (mL) and Y-axis indicates the intensity of absorbance (mAU). Tetramers (T), dimers (D), and monomers (M) are indicated. **(I)** Summary of mutant INs bridging TAR RNA compared to WT IN. Alpha screen counts at 320 nM for each protein is shown. The graphs show average values of three independent experiments and the error bars indicate standard deviation.

2.4.6 Sensitivity to ALLINIs is determined by distinct residues within the CTD

ALLINIs potently disrupt proper virion maturation through inducing aberrant IN multimerization and consequently inhibiting IN-gRNA binding (43, 52). Recent structural studies have shown that the ALLINI, GSK-1264, can directly engage residues within the IN-CTD through its tert-butoxy and carboxylic acid moieties and induce open polymers (53). These findings are in line with previous biochemical and modeling studies that also showed the involvement of the IN-CTD, in particular residues K264 and K266, in ALLINI-induced aberrant IN multimerization (43, 79-81).

Based on these prior findings, we wanted to test how adjacently positioned IN R262A/R263A and R269A/K273A substitutions and the compensatory mutations affected ALLINI activities.

To this end, we examined effects of representative quinoline-based ALLINIs, BI-D or BI-B2, on the viruses bearing the class II and compensatory mutations. While the titers of WT viruses only decreased at ALLINI concentrations greater than 1 μM, the titers of IN R269A/K273A viruses bearing D2N, D256R and D2R substitutions were significantly reduced by ALLINIs at concentrations as low as 0.1 μM (**Fig. 7A, B**). In contrast, viruses bearing the R262A/R263A/D256R IN were less sensitive to low concentrations of ALLINIs (**Fig. 7D, E**), suggesting that the R269A and/or K273A mutations underlie the increased sensitivity to ALLINIs. The increased sensitivity to ALLINIs correlated well with inhibition of RNA binding. While 0.1 μM of the ALLINI, BI-D, did affect the level of IN-gRNA binding in WT viruses, it significantly reduced IN-gRNA binding in viruses with class II compensatory mutations on the R269A/K273A but not R262A/R263A backbone (**Fig. 7C,F**). Taken together, our findings provide key genetic and virological evidence that specific CTD residues required for RNA binding are also crucial for the ALLINI mechanism of action.

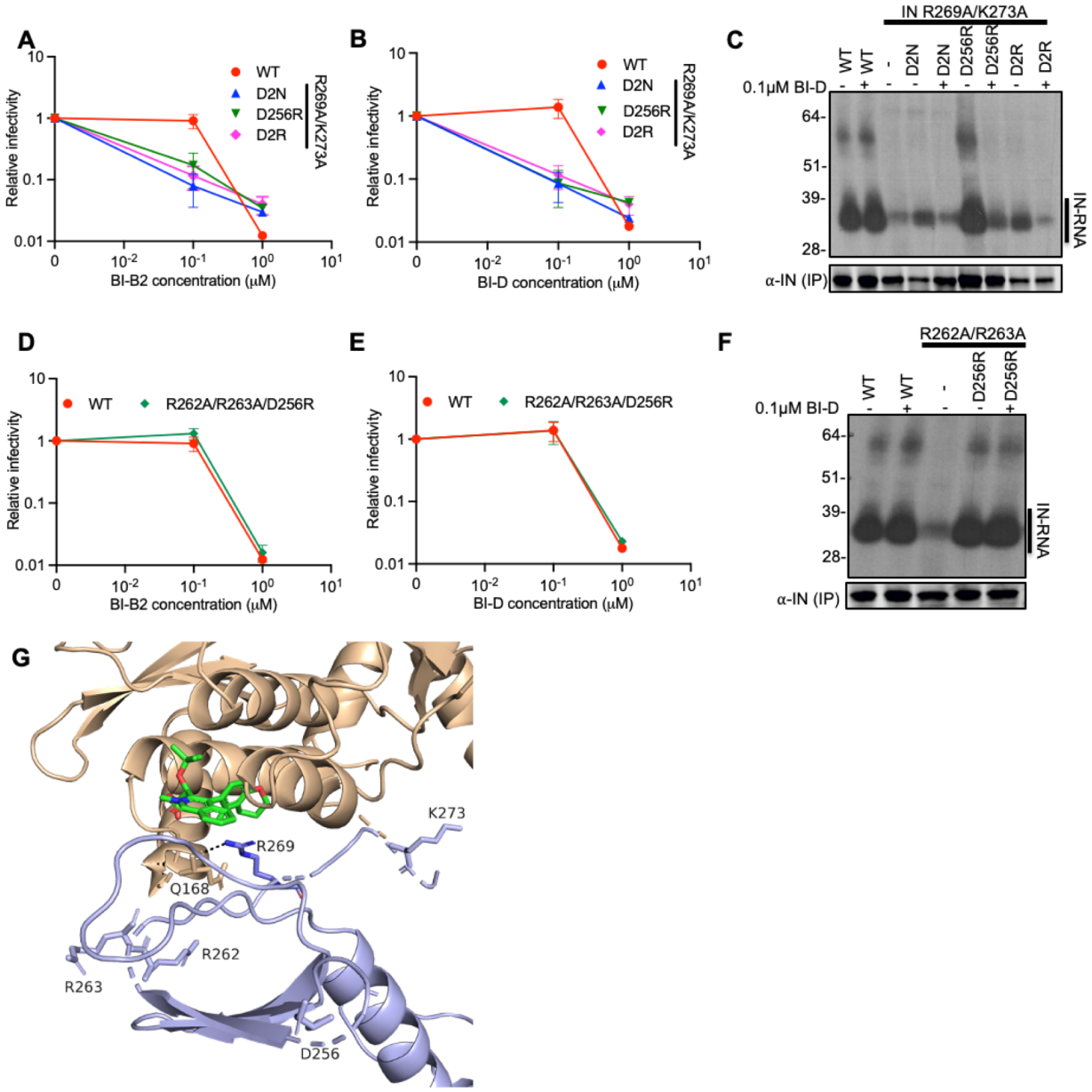

Figure 7. Secondary site suppressors of the IN R269A/K273A mutant increase susceptibility to ALLINIs. (A, B, D, E) Titers of viruses bearing the indicated substitutions in IN produced from HEK293T cells at different concentrations of ALLINIs, BI-B2 or BI-D. WT or IN mutant HIV-1_{NL4-3} in cell culture supernatants were titered on TZM-bl indicator cells. The titer values are represented relative to the mock control of each mutant (set to 1). Data are from two independent biological replicates. **(C, F)** Autoradiogram of IN-RNA adducts immunoprecipitated from WT or IN mutant HIV-1_{NL4-3} virions produced from HEK293T cells in the presence of 0.1μM of BI-D. The amount of immunoprecipitated IN protein was visualized by the immunoblot shown below. Immunoblots and CLIP autoradiographs results are a representative of three independent replicates. **(G)** The GSK-1264 binding pocket between CCD-CCD (brown) and CTD (light blue) from different IN subunits is shown (3) (PDB ID: 5HOT). GSK-1264 is in green with nitrogen and oxygen atoms colored blue and red, respectively. The side chain of R269 is shown with nitrogen atoms colored in dark blue. Interactions between CTD R269 and CCD Q168 are indicated by dashed lines.

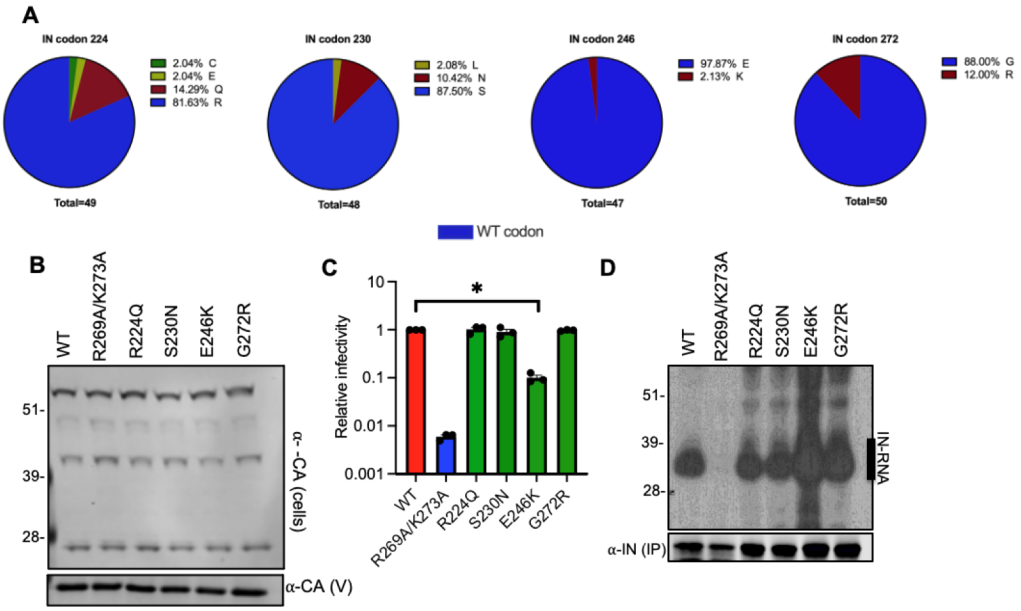

Figure 8. Characterization of IN mutations present in latently infected CD4+ T-cells. (A) Alignment from Gifford- (B) HEK293T cells were transfected with proviral HIV-1_{NL4-3} expression plasmids carrying the R224Q, S230N, E246K, and G272R IN mutations. Cell lysates and virions were purified two days post transfection and analyzed by immunoblotting for CA and IN. The image is representative of two independent experiments. **(C)** WT or IN mutant HIV-1L4-3 viruses in cell culture supernatants were titered on TZM-bl indicator cells. The titers are presented relative to WT (set to 1). The columns represent the average of three independent experiments and the error bars represent SEM (****p<0.0001, by one-way ANOVA with Dunnett's multiple comparison test). **(D)** Autoradiogram of IN-RNA adducts immunoprecipitated from virions bearing the indicated substitutions in IN. Immunoblots below show the amount of immunoprecipitated IN.

2.4.7 Characterization of IN mutations present in latently infected cells

Persistence of HIV-1 in memory CD4+ T-cells as latent proviruses constitutes a major barrier to HIV-1 cure. Although the majority of HIV-1 proviruses in these cells are defective , recent evidence suggests that defective proviruses can be transcribed into RNAs that are spliced, translated and can be recognized by HIV-1-specific cytotoxic T lymphocytes . We decided to characterize IN mutations isolated from latently infected cells, given the possibility that class II IN mutations existing in latently infected cells can result in the formation of defective particles that may subsequently modulate immune responses. Though relatively uncommon, we found the presence of R224Q, S230N, E246K and G272R substitutions in IN-CTD (**Fig. 8A**). Of note, only the R224Q substitution resulted in loss of a positive charge, whereas the E246K and G273R

substitutions resulted in gain of positive charge. These mutations were introduced into the NL4-3 proviral backbone with minimal effects on Gag expression and particle release (**Fig. 8B**). Although the E246K virus was significantly less infectious (**Fig. 8C**), we did not find any evidence for inhibition of IN-gRNA binding (**Fig. 8D**), suggesting that this mutant likely displays a class I phenotype. Thus, we conclude that the class II mutant viruses are rarely present in the latently infected cells and therefore unlikely to contribute to chronic immune activation.

2.5 Discussion

Class II IN mutations impair virion particle maturation by blocking IN-gRNA binding in virions and those within the CTD, including R269A/K273A and R262A/R263A, impede IN-gRNA binding without affecting functional oligomerization of IN . During serial passaging experiments to identify compensatory mutations of the R269A/K273A class II IN mutant, viruses acquired D256N and D270N IN mutations sequentially. We initially anticipated that mutations outside of IN, such as CA and NC, could also arise, given that the IN R269A/K273A mutant is still catalytically active and a compensatory mutation in CA or NC could presumably allow the proper packaging of the gRNA within virions. On the other hand, we did not observe any such substitutions showcasing the distinct role of IN:gRNA binding in proper virion maturation. Notably, D256N and D270N substitutions each arose through a single mutation (D256N: GAC→AAC, D270N:GAU→AAU) and thus likely provided an easier pathway for suppression than reverting back to R269 and K273, each of which would require two mutations. Though other mutations in IN could in principle restore the ability of IN to bind RNA, rise of such mutations was likely constrained in part by the necessity to maintain a catalytically active IN.

IN CTD has a net positive charge of +3. While R269A/K273A class II mutations reduced the net charge to +1, the introduction of D2N substitutions restored the overall charge back to a +3. This

observation suggested that the net electrostatic charge of the IN-CTD may be a key parameter for IN-gRNA binding. Consistent with this hypothesis, the charge reversal by the D256R substitution was sufficient to restore IN-gRNA binding for the R269A/K273A IN mutant. Though these findings point to an electrostatic component of IN-RNA interactions, we cannot exclude the possibility that the Asn and Arg residues also mediate H-bonding and van der Waals contacts with distinct nucleobases in the cognate RNA molecules. Inspection of the available X-ray structure (83) indicates that all IN residues implicated in RNA binding by the present study are positioned either in the highly flexible C-terminal tail (aa 261-275) or the 252-257 loop (**Fig. 5H**). The pliable nature of these CTD regions could be essential for allowing IN to optimally bind to cognate gRNA.

These observations extended to another class II mutant, R262A/R263A, whereby the D256R and to a lesser extent the D256K substitutions restored RNA binding. In contrast, the D2N substitutions did not enhance RNA-binding or infectivity in this setting, demonstrating a degree of context dependency. It is possible that the proximity of the D270N to R269A/K273A residues may explain why it was more effective in restoring RNA binding for this class II mutant but not for IN R262A/R263A. The role of Arg and Lys residues in RNA-binding proteins has been noted (81). Interestingly, Arg residues are overall more heavily involved in interactions with all bases, contributing with 16%–20% of all contacts while Lys residues only provide 3%–9%, which may explain why the D256R substitution restored RNA binding to a greater extent (40, 51, 85). Altogether, our findings suggest that IN-gRNA binding is mediated in part by electrostatic interactions between the basic residues in IN CTD and the negatively charged RNA backbone.

Though the electrostatic component of IN-RNA interactions imply a level of non-specificity, IN bound to distinct locations on the gRNA and displays high binding affinity to structured elements, such as the HIV-1 TAR element . Thus, it is likely that IN-gRNA interactions are mediated by

both the non-specific interactions of the basic residues with the RNA phosphate backbone and specific interactions with the cognate RNA. For example, recognition of the TAR loop by Tat and the super elongation complex is based on a complex set of interactions that results primarily in the readout of the structure as opposed to the sequence by SEC, and additional interaction of the Tat arginine rich motif (ARM) with the TAR bulge and the major groove through electrostatic interactions with the RNA phosphates, H-bonding with specific bases, as well as multiple van der Waals contacts of Cyclin T with the flipped-out G32 TAR base to strengthen the TAR complex . Structural studies of IN in complex with cognate RNA molecules are needed to tease apart the specificity determinants for IN-RNA interactions.

RNA binding proteins commonly encode modular RNA-binding domains (i.e. RRMs and KH domains), which form specific contacts with short degenerate sequences (13, 83, 86-88). Utilization of multiple RRMs/KH domains is thought to create a much larger binding interface, which allows recognition of longer sequences on and enhanced affinity and specificity for target RNAs . IN tetramerization is required for RNA binding in vitro (43, 54, 89-92) and may serve a similar purpose through generation of a larger RNA-binding surface possibly for recognition of shape, RNA-backbone and base-specific interactions as discussed above.

IN-gRNA binding is essential for HIV-1 virion morphological maturation and infectivity, thus an excellent target for novel antiretroviral compounds. Clinically used HIV-1 IN inhibitors target the catalytic activity of IN by blocking the strand transfer step of integration . Despite high barriers to resistance with the second-generation INSTIs, treatment continues to select for drug resistant HIV-1 variants (94, 95). ALLINI-mediated inhibition of the non-catalytic function of IN can complement existing INSTI-based therapies and increase the barrier to drug resistance

substantially. Although some ALLINIs displayed toxicity in animals, a highly potent and safe pyrrolopyridine-based ALLINI, STP0404, has advanced to a human trial (99, 100).

ALLINIs potently disrupt virion maturation indirectly through inducing aberrant IN multimerization resulting in inhibition of IN-gRNA binding . Recent findings have shown that the ALLINI, GSK-1264, is buried between the CTD of one IN dimer and the CCD of another dimer resulting in the formation of open, inactive IN polymers (101). Mutation of Y226, W235, K264 and K266 residues within the CTD prevents ALLINI-induced aberrant multimerization of IN (67). However, the role of these residues in ALLINI-mediated inhibition cannot be easily assessed in relevant infection settings given that their mutation alters the catalytic activity as well as RNA binding properties of IN (69). Our findings fill this gap and provide key genetic evidence that ALLINI mechanism of action indeed involves the CTD. In particular, introducing the same D256R change in the background of different class II substitutions similarly restored their infectivity and thus enabled us to compare effects of R262/R263 vs R269A/K273A IN substitutions on ALLINI activities. While HIV-$1_{IN(R262A/R263A/D256R)}$ and WT viruses were similarly susceptible to ALLINIs (**Fig. 7A,B**), HIV-$1_{IN(R269A/K273A/D256R)}$ was substantially more sensitive to these inhibitors (**Fig. 7D,E**). The X-ray structure of GSK-1264 induced IN polymers (12, 73, 77, 102, 103) reveals that R262, R263, K273A as well as D256 are distanced from the inhibitor binding pocket, whereas the side chain of R269 points toward the inhibitor bound at the CTD-CCD dimer interface (**Fig. 7G**). The guanidine group of R269 engages with Q168 of the CCD from another IN subunit and is positioned within 3 Å of GSK-1264. Yet, there is no interaction between R269 and the inhibitor as the positively charged guanidine group faces a hydrophobic part of the benzodihydropyran moiety of GSK-1264 (**Fig. 7G**). Conversely, the R269A substitution could provide more favorable

hydrophobic environment for the CTD-ALLINI-CCD interactions that may explain its increased sensitivity to ALLINIs.

Overall, our studies reveal that electrostatic interactions play an important role in mediating IN-gRNA interactions and demonstrate that CTD is a key determinant of ALLINI sensitivity. Structural characterization of IN-gRNA complexes will be crucial to determine the precise mechanism of IN-gRNA interactions. Such studies will also help better understand the ALLINI mechanism of action and aid in the development of therapeutics that directly target IN-gRNA interactions.

2.6 Materials and methods

Plasmids

IN mutations were introduced into the HIV-1_{NL4-3} full-length proviral plasmid (pNL4-3) by overlap extension PCR. Forward and reverse primers containing IN mutations were used in PCR reactions with anti-sense (with EcoRI restriction endonuclease site) and sense (with AgeI restriction site) outer primers. The resulting fragments containing the desired mutations were mixed at a 1:1 ratio and overlapped subsequently using the sense and antisense primer pairs. The overlap fragments were digested with Age-I and EcoR-I before cloning into pNL4-3 plasmids. The pLR2P-vprIN plasmid expressing a Vpr-IN fusion protein has been previously described . IN mutations were introduced in the pLR2P-VprIN plasmid using the QuikChange Site-Directed Mutagenesis kit (Agilent Technologies). The presence of the desired mutations and the absence of unwanted secondary changes was assessed by Sanger sequencing.

Cell lines

All cell lines were obtained from the American Type Culture Collection and NIH AIDS Reagents where STR profiling was performed. MT-4 cells were further STR profiled at Washington University School of Medicine Genome Engineering and iPSC center. The cell lines are consistently inspected for mycoplasma contamination using MycoAlert mycoplasma detection kit (Lonza) and checked for being free of any other contaminations. HEK293T cells (ATCC CRL-11268) and Hela-derived TZM-bl cells (NIH AIDS Reagent Program) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. MT-4 cells were cultured in in RPMI 1640 medium supplemented with 10% fetal bovine serum.

Analysis of compensatory mutations

Compensatory mutations of the R269A/K273A class-II IN mutation were isolated by serial passaging. In this experiment, MT4 T-cells that express GFP under the control of HIV-1 LTR (MT4-LTR-GFP), were infected with HIV-1pNL4-3 carrying the R269A/K273A class II IN mutation. One million cells were infected at an MOI of 2 by WT cells, input of mutant virus was normalized relative to particle numbers using a reverse transcriptase (RT) activity assay . Infections were monitored by FACS. At the end of each passage, virions were collected and wielded to infect cells in the next passage while normalizing virus input before each infection. Aliquots of infected and viral particles in the cell culture supernatants were collected over the duration of each passage. Virions collected over the three passages were concentrated and gRNA was isolated from virions using Trizol per manufacturer's instructions. Extracted RNA was prepared for deep sequencing usingthe Illumina® TruSeq® Stranded Total RNA library prep workflow kit omitting the rRNA depletion step. Resulting libraries were sequenced by an Illumina HiSeq 2000 platform.

Immunoblotting

Viral and cell lysates were resuspended in sodium dodecyl sulfate (SDS) sample buffer, separated by electrophoresis on Bolt 4-12% Bis-Tris Plus gels (Life Technologies) and transferred to nitrocellulose membranes. The membranes were then probed overnight at 4°C with a mouse monoclonal anti-HIV p24 antibody (183-H12-5C, NIH AIDS reagents) or a mouse monoclonal anti-HIV integrase antibody in Odyssey Blocking Buffer (LI-COR). Membranes were probed with fluorophore-conjugated secondary antibodies (LI-COR) and scanned using an LI-COR Odyssey system. IN and CA levels in virions were quantified using Image Studio software (LI-COR).

Vpr-IN trans-complementation experiments

Viruses bearing the class I IN mutation D116N were trans-complemented with class II mutant proteins as previously described ; two-hundred thousand HEK293T cells were co-transfected with a derivative of pNL4-3-derived plasmid bearing the IN D116N mutation, VSV-G, and derivatives of the pLR2P-VprIN plasmids bearing class II IN mutations (or the compensatory mutations thereof) at a ratio of 6:1:3. Cell-free virions were collected from cell culture two days post-transfection. MT-4 cells were infected by the virions and the integration proficiency of trans-complemented class II IN mutants was measured by the yield of progeny virions in cell culture supernatants over a 6-day period as described before (4). Briefly, MT-4 cells were incubated with virus inoculum in 96 V-bottom well plates for 4hr at 37°C before washing away the inoculum and replacing it with fresh media. Right after the addition of fresh media and over the ensuing 6 days, the number of virions in culture supernatant was quantified by measuring RT activity using a Q-PCR-based assay (15).

CLIP experiments

CLIP experiments were conducted as previously described (10). In short, cells in 15 cm cell culture plates were transfected with 30 μg full-length proviral plasmid (pNL4-3) DNA containing the WT sequence or indicated IN mutations. 4-thiouridine (4SU) was added to the cell culture media for 16hr before virus harvest. Cell culture supernatants were filtered through 0.22 μm filters and pelleted by ultracentrifugation through a 20% sucrose cushion using a Beckman SW32-Ti rotor at 28,000rpm for 1.5hr at 4°C. The virus pellets were resuspended in phosphate-buffered saline (PBS) and UV-crosslinked. Following lysis in 1X RIPA buffer, IN-RNA complexes were immunoprecipitated using a mouse monoclonal anti-IN antibody (106) . Bound RNA was end-labeled with γ-^{32}P-ATP and T4 polynucleotide kinase. The isolated protein-RNA complexes were separated by SDS-PAGE, transferred to nitrocellulose membranes, and exposed to autoradiography films to visualize IN-RNA complexes. Lysates and immunoprecipitates were also analyzed by immunoblotting using antibodies against IN.

IN multimerization in virions

In 10 cm dishes, HEK293T cells were transfected with 10 μg pNL4-3 plasmid DNA containing the WT sequence or indicated pol mutations within IN coding sequence. Two days post-transfection, cell-free virions in cell culture supernatants were pelleted through a 20% sucrose gradient using a Beckman SW41-Ti rotor at 28,000 rpm for 1.5hr at 4°C. Pelleted virions were resuspended in 1xPBS and treated with a membrane-permeable crosslinker, EGS (ThermoFisher Scientific), at a concentration of 1mM for 30 min at room temperature. Crosslinking was stopped by the addition of SDS sample buffer. The cross-linked samples were then separated on 3-8% Tris-acetate gels and analyzed by immunoblotting using a mouse monoclonal anti-IN antibody (107).

Virus production and transmission electron microscopy

HEK293T cells in a 15 cm plate were transfected with 30 μg full-length proviral plasmid (pNL4-3) DNA containing the WT sequence or indicated pol mutations within IN coding sequence. Two days post transfection, cell culture supernatants were filtered through 0.22 μm filters, and pelleted by ultracentrifugation using a Beckman SW32-Ti rotor at 28,000 rpm for 1.5 hr at 4°C. Fixative (2% paraformaldehyde/2.5% glutaraldehyde (Polysciences Inc., Warrington, PA) in 100 mM sodium cacodylate buffer, pH 7.2) was gently added to resulting pellets, and samples were incubated overnight at 4°C. Samples were washed in sodium cacodylate buffer and postfixed in 1% osmium tetroxide (Polysciences Inc.) for 1 hr. Samples were then rinsed extensively in dH_20 prior to en bloc staining with 1% aqueous uranyl acetate (Ted Pella Inc., Redding, CA) for 1 hr. After several rinses in dH_20, samples were dehydrated in a graded series of ethanol and embedded in Eponate 12 resin (Ted Pella Inc.). Sections of 95 nm were cut with a Leica Ultracut UCT ultramicrotome (Leica Microsystems Inc., Bannockburn, IL), stained with uranyl acetate and lead citrate, and viewed on a JEOL 1200 EX transmission electron microscope (JEOL USA Inc., Peabody, MA) equipped with an AMT 8 megapixel digital camera and AMT Image Capture Engine V602 software (Advanced Microscopy Techniques, Woburn, MA).

Size exclusion chromatography (SEC)

All of the indicated mutations were introduced into a plasmid backbone expressing His_6 tagged pNL4-3-derived IN by QuikChange site directed mutagenesis kit (Agilent) (1, 108). His_6 tagged recombinant pNL4-3 WT and mutant Ins were expressed in BL21 (DE3) *E. coli* cells followed by nickel and heparin column purification as described previously . Recombinant WT and mutant Ins were analyzed on Superdex 200 10/300 GL column (GE Healthcare) with running buffer

containing 20 mM HEPES (pH 7.5), 1 M NaCl, 10% glycerol and 5 mM BME at 0.3 mL/min flow rate. The proteins were diluted to 10 μM with the running buffer and incubated for 1 h at 4°C followed by centrifugation at 10,000g for 10 min. Multimeric form determination was based on the standards including bovine thyroglobulin (670,000 Da), bovine gamma-globulin (158,000 Da), chicken ovalbumin (44,000 Da), horse myoglobin (17,000 Da) and vitamin B12 (1,350 Da). Retention volumes for different oligomeric forms of IN were as follows: tetramer ~12.5 mL, dimer ~14 mL, monomer ~15-16 mL.

Analysis of IN-RNA binding in vitro

To monitor IN-RNA interactions we utilized AlphaScreen-based assay (1), which allows to monitor the ability of IN to bind and bridge between two TAR RNAs. Briefly, equal concentrations (1 nM) of two synthetic TAR RNA oligonucleotides labeled either with biotin or DIG were mixed and then streptavidin donor and anti-DIG acceptor beads at 0.02 mg/mL concentration were supplied in a buffer containing 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 1 mg/mL BSA, and 25 mM Tris (pH 7.4). After 2 hr incubation at 4°C, 320 nM IN was added to the reaction mixture and incubated further for 1.5 hr at 4°C. AlphaScreen signals were recorded with a PerkinElmer Life Sciences Enspire multimode plate reader.

Structural modeling of IN

Electrostatic potential maps of WT and mutant IN CTDs were created by Adaptive Poisson-Boltzmann Solver (APBS) program (2) with macromolecular electrostatic calculations performed in PyMOL. The published crystal structure (3) (PDB ID: 5HOT) was used as a template. The calculation results are displayed as an electrostatic potential molecular surface. The low, mid, and high range values are -5, 0, and 5, respectively.

Analysis of IN mutations from latently infected cells

Sequences identified from latently infected CD4+ T-cells (4) were downloaded from NCBI GenBank based on their accession numbers (KF526120-KF526339). Sequences were imported into the GLUE software framework (5) and aligned. Multiple sequence alignments (MSAs) containing subtype B and subtype C sequences were constructed using MUSCLE, manually inspected in AliView (6) and imported into a GLUE project database. Within GLUE, MSAs were constrained to the pNL4-3 reference to establish a standardized coordinate space for the gene being analyzed. Amino acid frequencies at each alignment position were summarized using GLUE's amino-acid frequency calculation algorithm, which accounts for contingencies such as missing data and incomplete codons.

References

1. Wilen CB, Tilton JC, Doms RW. 2012. HIV: cell binding and entry. Cold Spring Harb Perspect Med 2.

2. Herschhorn A, Hizi A. 2010. Retroviral reverse transcriptases. Cell Mol Life Sci 67:2717-47.

3. Campbell EM, Hope TJ. 2015. HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nat Rev Microbiol 13:471-83.

4. Lu R, Limon A, Devroe E, Silver PA, Cherepanov P, Engelman A. 2004. Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. J Virol 78:12735-46.

5. Nakamura T, Masuda T, Goto T, Sano K, Nakai M, Harada S. 1997. Lack of infectivity of HIV-1 integrase zinc finger-like domain mutant with morphologically normal maturation. Biochem Biophys Res Commun 239:715-22.

6. 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.

7. Shin CG, Taddeo B, Haseltine WA, Farnet CM. 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. J Virol 68:1633-42.

8. Taddeo B, Haseltine WA, Farnet CM. 1994. Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. J Virol 68:8401-5.

9. 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. J Virol 73:2126-35.

10. 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. J Virol 69:2729-36.

11. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, Gorelick RJ, Engelman AN, Steven AC. 2015. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. J Virol 89:9765-80.

12. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, Wang W, Ballandras-Colas A, Patel PA, Fuchs JR, Kvaratskhelia M, Engelman A. 2013. Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. Proc Natl Acad Sci U S A 110:8690-5.

13. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268 e12.

14. Jenkins TM, Engelman A, Ghirlando R, Craigie R. 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J Biol Chem 271:7712-8.

15. Engelman A. 1999. In vivo analysis of retroviral integrase structure and function. Adv Virus Res 52:411-26.

16. 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. J Virol 70:721-8.

17. Lu R, Ghory HZ, Engelman A. 2005. Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. J Virol 79:10356-68.

18. Ao Z, Fowke KR, Cohen EA, Yao X. 2005. Contribution of the C-terminal tri-lysine regions of human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA nuclear import. Retrovirology 2:62.

19. Busschots K, Voet A, De Maeyer M, Rain JC, Emiliani S, Benarous R, Desender L, Debyser Z, Christ F. 2007. Identification of the LEDGF/p75 binding site in HIV-1 integrase. J Mol Biol 365:1480-92.

20. Engelman A, Liu Y, Chen H, Farzan M, Dyda F. 1997. Structure-based mutagenesis of the catalytic domain of human immunodeficiency virus type 1 integrase. J Virol 71:3507-14.

21. Limon A, Devroe E, Lu R, Ghory HZ, Silver PA, Engelman A. 2002. Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import. J Virol 76:10598-607.

22. Lloyd AG, Ng YS, Muesing MA, Simon V, Mulder LC. 2007. Characterization of HIV-1 integrase N-terminal mutant viruses. Virology 360:129-35.

23. Lu R, Vandegraaff N, Cherepanov P, Engelman A. 2005. Lys-34, dispensable for integrase catalysis, is required for preintegration complex function and human immunodeficiency virus type 1 replication. J Virol 79:12584-91.

24. 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. J Virol 69:6687-96.

25. Rahman S, Lu R, Vandegraaff N, Cherepanov P, Engelman A. 2007. Structure-based mutagenesis of the integrase-LEDGF/p75 interface uncouples a strict correlation between in vitro protein binding and HIV-1 fitness. Virology 357:79-90.

26. Riviere L, Darlix JL, Cimarelli A. 2010. Analysis of the viral elements required in the nuclear import of HIV-1 DNA. J Virol 84:729-39.

27. Tsurutani N, Kubo M, Maeda Y, Ohashi T, Yamamoto N, Kannagi M, Masuda T. 2000. Identification of critical amino acid residues in human immunodeficiency virus type 1 IN required for efficient proviral DNA formation at steps prior to integration in dividing and nondividing cells. J Virol 74:4795-806.

28. Wiskerchen M, Muesing MA. 1995. Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. J Virol 69:376-86.

29. 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. J Virol 78:5045-55.

30. De Houwer S, Demeulemeester J, Thys W, Rocha S, Dirix L, Gijsbers R, Christ F, Debyser Z. 2014. The HIV-1 integrase mutant R263A/K264A is 2-fold defective for TRN-SR2 binding and viral nuclear import. J Biol Chem 289:25351-61.

31. Johnson BC, Metifiot M, Ferris A, Pommier Y, Hughes SH. 2013. A homology model of HIV-1 integrase and analysis of mutations designed to test the model. J Mol Biol 425:2133-46.

32. Mohammed KD, Topper MB, Muesing MA. 2011. Sequential deletion of the integrase (Gag-Pol) carboxyl terminus reveals distinct phenotypic classes of defective HIV-1. J Virol 85:4654-66.

33. Shehu-Xhilaga M, Hill M, Marshall JA, Kappes J, Crowe SM, Mak J. 2002. The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of pol protein. J Virol 76:4331-40.

34. Engelman A, Craigie R. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J Virol 66:6361-9.

35. Kalpana GV, Reicin A, Cheng GS, Sorin M, Paik S, Goff SP. 1999. Isolation and characterization of an oligomerization-negative mutant of HIV-1 integrase. Virology 259:274-85.

36. Lutzke RA, Plasterk RH. 1998. Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding. J Virol 72:4841-8.

37. Lutzke RA, Vink C, Plasterk RH. 1994. Characterization of the minimal DNA-binding domain of the HIV integrase protein. Nucleic Acids Res 22:4125-31.

38. Balakrishnan M, Yant SR, Tsai L, O'Sullivan C, Bam RA, Tsai A, Niedziela-Majka A, Stray KM, Sakowicz R, Cihlar T. 2013. Non-catalytic site HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in target cells. PLoS One 8:e74163.

39. Desimmie BA, Schrijvers R, Demeulemeester J, Borrenberghs D, Weydert C, Thys W, Vets S, Van Remoortel B, Hofkens J, De Rijck J, Hendrix J, Bannert N, Gijsbers R, Christ F, Debyser Z. 2013. LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the virions. Retrovirology 10:57.

40. Elliott JL, Eschbach JE, Koneru PC, Li W, Puray-Chavez M, Townsend D, Lawson DQ, Engelman AN, Kvaratskhelia M, Kutluay SB. 2020. Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion morphogenesis. Elife 9.

41. Sharma A, Slaughter A, Jena N, Feng L, Kessl JJ, Fadel HJ, Malani N, Male F, Wu L, Poeschla E, Bushman FD, Fuchs JR, Kvaratskhelia M. 2014. A new class of multimerization selective inhibitors of HIV-1 integrase. PLoS Pathog 10:e1004171.

42. Gupta K, Brady T, Dyer BM, Malani N, Hwang Y, Male F, Nolte RT, Wang L, Velthuisen E, Jeffrey J, Van Duyne GD, Bushman FD. 2014. Allosteric inhibition of human immunodeficiency virus integrase: late block during viral replication and abnormal multimerization involving specific protein domains. J Biol Chem 289:20477-88.

43. Gupta K, Turkki V, Sherrill-Mix S, Hwang Y, Eilers G, Taylor L, McDanal C, Wang P, Temelkoff D, Nolte RT, Velthuisen E, Jeffrey J, Van Duyne GD, Bushman FD. 2016. Structural Basis for Inhibitor-Induced Aggregation of HIV Integrase. PLoS Biol 14:e1002584.

44. 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. Antimicrob Agents Chemother 56:4365-74.

45. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, Bardiot D, Van der Veken NJ, Van Remoortel B, Strelkov SV, De Maeyer M, Chaltin P, Debyser Z. 2010. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. Nat Chem Biol 6:442-8.

46. Maehigashi T, Ahn S, Kim UI, Lindenberger J, Oo A, Koneru PC, Mahboubi B, Engelman AN, Kvaratskhelia M, Kim K, Kim B. 2021. A highly potent and safe pyrrolopyridine-based allosteric HIV-1 integrase inhibitor targeting host LEDGF/p75-integrase interaction site. PLoS Pathog 17:e1009671.

47. Serganov A, Patel DJ. 2008. Towards deciphering the principles underlying an mRNA recognition code. Curr Opin Struct Biol 18:120-9.

48. Bahadur RP, Zacharias M, Janin J. 2008. Dissecting protein-RNA recognition sites. Nucleic Acids Res 36:2705-16.

49. Stefl R, Skrisovska L, Allain FH. 2005. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO Rep 6:33-8.

50. Shema Mugisha C, Tenneti K, Kutluay SB. 2020. Clip for studying protein-RNA interactions that regulate virus replication. Methods 183:84-92.

51. Bieniasz PD, Kutluay SB. 2018. CLIP-related methodologies and their application to retrovirology. Retrovirology 15:35.

52. Deng N, Hoyte A, Mansour YE, Mohamed MS, Fuchs JR, Engelman AN, Kvaratskhelia M, Levy R. 2016. Allosteric HIV-1 integrase inhibitors promote aberrant protein multimerization by directly mediating inter-subunit interactions: Structural and thermodynamic modeling studies. Protein Sci 25:1911-1917.

53. Shkriabai N, Dharmarajan V, Slaughter A, Kessl JJ, Larue RC, Feng L, Fuchs JR, Griffin PR, Kvaratskhelia M. 2014. A critical role of the C-terminal segment for allosteric inhibitor-induced aberrant multimerization of HIV-1 integrase. J Biol Chem 289:26430-26440.

54. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540-51.

55. Pollack RA, Jones RB, Pertea M, Bruner KM, Martin AR, Thomas AS, Capoferri AA, Beg SA, Huang SH, Karandish S, Hao H, Halper-Stromberg E, Yong PC, Kovacs C, Benko E, Siliciano RF, Ho YC. 2017. Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. Cell Host Microbe 21:494-506 e4.

56. Bayer TS, Booth LN, Knudsen SM, Ellington AD. 2005. Arginine-rich motifs present multiple interfaces for specific binding by RNA. RNA 11:1848-57.

57. Jones S, Daley DT, Luscombe NM, Berman HM, Thornton JM. 2001. Protein-RNA interactions: a structural analysis. Nucleic Acids Res 29:943-54.

58. Jeong E, Kim H, Lee SW, Han K. 2003. Discovering the interaction propensities of amino acids and nucleotides from protein-RNA complexes. Mol Cells 16:161-7.

59. Kim OT, Yura K, Go N. 2006. Amino acid residue doublet propensity in the protein-RNA interface and its application to RNA interface prediction. Nucleic Acids Res 34:6450-60.

60. Kruger DM, Neubacher S, Grossmann TN. 2018. Protein-RNA interactions: structural characteristics and hotspot amino acids. RNA 24:1457-1465.

61. Schulze-Gahmen U, Hurley JH. 2018. Structural mechanism for HIV-1 TAR loop recognition by Tat and the super elongation complex. Proc Natl Acad Sci U S A 115:12973-12978.

62. Schulze-Gahmen U, Echeverria I, Stjepanovic G, Bai Y, Lu H, Schneidman-Duhovny D, Doudna JA, Zhou Q, Sali A, Hurley JH. 2016. Insights into HIV-1 proviral transcription from integrative structure and dynamics of the Tat:AFF4:P-TEFb:TAR complex. Elife 5.

63. Han SP, Tang YH, Smith R. 2010. Functional diversity of the hnRNPs: past, present and perspectives. Biochem J 430:379-92.

64. Jeong S. 2017. SR Proteins: Binders, Regulators, and Connectors of RNA. Mol Cells 40:1-9.

65. Lunde BM, Moore C, Varani G. 2007. RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol 8:479-90.

66. Ban T, Zhu JK, Melcher K, Xu HE. 2015. Structural mechanisms of RNA recognition: sequence-specific and non-specific RNA-binding proteins and the Cas9-RNA-DNA complex. Cell Mol Life Sci 72:1045-58.

67. Hazuda DJ, Felock P, Witmer M, Wolfe A, Stillmock K, Grobler JA, Espeseth A, Gabryelski L, Schleif W, Blau C, Miller MD. 2000. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science 287:646-50.

68. Summa V, Petrocchi A, Bonelli F, Crescenzi B, Donghi M, Ferrara M, Fiore F, Gardelli C, Gonzalez Paz O, Hazuda DJ, Jones P, Kinzel O, Laufer R, Monteagudo E, Muraglia E, Nizi E, Orvieto F, Pace P, Pescatore G, Scarpelli R, Stillmock K, Witmer MV, Rowley M. 2008. Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. J Med Chem 51:5843-55.

69. Ramanathan S, Mathias AA, German P, Kearney BP. 2011. Clinical pharmacokinetic and pharmacodynamic profile of the HIV integrase inhibitor elvitegravir. Clin Pharmacokinet 50:229-44.

70. Min S, Song I, Borland J, Chen S, Lou Y, Fujiwara T, Piscitelli SC. 2010. Pharmacokinetics and safety of S/GSK1349572, a next-generation HIV integrase inhibitor, in healthy volunteers. Antimicrob Agents Chemother 54:254-8.

71. Tsiang M, Jones GS, Goldsmith J, Mulato A, Hansen D, Kan E, Tsai L, Bam RA, Stepan G, Stray KM, Niedziela-Majka A, Yant SR, Yu H, Kukolj G, Cihlar T, Lazerwith SE, White KL, Jin H. 2016. Antiviral Activity of Bictegravir (GS-9883), a Novel Potent HIV-1 Integrase Strand Transfer Inhibitor with an Improved Resistance Profile. Antimicrob Agents Chemother 60:7086-7097.

72. Smith SJ, Zhao XZ, Burke TR, Jr., Hughes SH. 2018. Efficacies of Cabotegravir and Bictegravir against drug-resistant HIV-1 integrase mutants. Retrovirology 15:37.

73. Wijting IEA, Lungu C, Rijnders BJA, van der Ende ME, Pham HT, Mesplede T, Pas SD, Voermans JJC, Schuurman R, van de Vijver DAMC, Boers PHM, Gruters RA, Boucher CAB, van Kampen JJA. 2018. HIV-1 Resistance Dynamics in Patients With Virologic Failure to Dolutegravir Maintenance Monotherapy. J Infect Dis 218:688-697.

74. Radzio-Basu J, Council O, Cong ME, Ruone S, Newton A, Wei X, Mitchell J, Ellis S, Petropoulos CJ, Huang W, Spreen W, Heneine W, García-Lerma JG. 2019. Drug resistance emergence in macaques administered cabotegravir long-acting for pre-exposure prophylaxis during acute SHIV infection. Nat Commun 10:2005.

75. Anstett K, Brenner B, Mesplede T, Wainberg MA. 2017. HIV drug resistance against strand transfer integrase inhibitors. Retrovirology 14:36.

76. Rossouw TM, Hitchcock S, Botes M. 2016. The end of the line? A case of drug resistance to third-line antiretroviral therapy. South Afr J HIV Med 17:454.

77. Zhang WW, Cheung PK, Oliveira N, Robbins MA, Harrigan PR, Shahid A. 2018. Accumulation of Multiple Mutations In Vivo Confers Cross-Resistance to New and Existing Integrase Inhibitors. J Infect Dis 218:1773-1776.

78. Ndashimye E, Li Y, Reyes PS, Avino M, Olabode AS, Kityo CM, Kyeyune F, Nankya I, Quinones-Mateu ME, Barr SD, Arts EJ. 2021. High-level resistance to bictegravir and cabotegravir in subtype A- and D-infected HIV-1 patients failing raltegravir with multiple resistance mutations. J Antimicrob Chemother doi:10.1093/jac/dkab276.

79. Kessl JJ, Sharma A, Kvaratskhelia M. 2016. Methods for the Analyses of Inhibitor-Induced Aberrant Multimerization of HIV-1 Integrase. Methods Mol Biol 1354:149-64.

80. Madison MK, Lawson DQ, Elliott J, Ozanturk AN, Koneru PC, Townsend D, Errando M, Kvaratskhelia M, Kutluay SB. 2017. Allosteric HIV-1 Integrase Inhibitors Lead to Premature Degradation of the Viral RNA Genome and Integrase in Target Cells. J Virol 91.

81. Liu H, Wu X, Xiao H, Conway JA, Kappes JC. 1997. Incorporation of functional human immunodeficiency virus type 1 integrase into virions independent of the Gag-Pol precursor protein. J Virol 71:7704-10.

82. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. 2009. A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. J Virol Methods 156:1-7.

83. Bouyac-Bertoia M, Dvorin JD, Fouchier RA, Jenkins Y, Meyer BE, Wu LI, Emerman M, Malim MH. 2001. HIV-1 infection requires a functional integrase NLS. Mol Cell 7:1025-35.

84. Kutluay SB, Bieniasz PD. 2016. Analysis of HIV-1 Gag-RNA Interactions in Cells and Virions by CLIP-seq. Methods Mol Biol 1354:119-31.

85. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, Errando M, Bieniasz PD. 2014. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell 159:1096-1109.

86. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, de Silva S, Wu L, Le Grice SF, Engelman A, Fuchs JR, Kvaratskhelia M. 2012. Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. J Biol Chem 287:16801-11.

87. Cherepanov P. 2007. LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. Nucleic Acids Res 35:113-24.

88. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K, Chun M, Li P, Gohara DW, Dolinsky T, Konecny R, Koes DR, Nielsen JE, Head-Gordon T, Geng W, Krasny R, Wei GW, Holst MJ, McCammon JA, Baker NA. 2018. Improvements to the APBS biomolecular solvation software suite. Protein Sci 27:112-128.

89. Singer JB, Thomson EC, Hughes J, Aranday-Cortes E, McLauchlan J, da Silva Filipe A, Tong L, Manso CF, Gifford RJ, Robertson DL, Barnes E, Ansari MA, Mbisa JL, Bibby DF, Bradshaw D, Smith D. 2019. Interpreting Viral Deep Sequencing Data with GLUE. Viruses 11.

90. Singer JB, Thomson EC, McLauchlan J, Hughes J, Gifford RJ. 2018. GLUE: a flexible software system for virus sequence data. BMC Bioinformatics 19:532.

91. Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276-8.

92. Butan C, Winkler DC, Heymann JB, Craven RC, Steven AC. 2008. RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis. J Mol Biol 376:1168-81.

93. Briggs JA, Watson BE, Gowen BE, Fuller SD. 2004. Cryoelectron microscopy of mouse mammary tumor virus. J Virol 78:2606-8.

94. Forster F, Medalia O, Zauberman N, Baumeister W, Fass D. 2005. Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. Proc Natl Acad Sci U S A 102:4729-34.

95. Cao S, Maldonado JO, Grigsby IF, Mansky LM, Zhang W. 2015. Analysis of human T-cell leukemia virus type 1 particles by using cryo-electron tomography. J Virol 89:2430-5.

96. Briggs JA, Grunewald K, Glass B, Forster F, Krausslich HG, Fuller SD. 2006. The mechanism of HIV-1 core assembly: insights from three-dimensional reconstructions of authentic virions. Structure 14:15-20.

97. Hamann MV, Mullers E, Reh J, Stanke N, Effantin G, Weissenhorn W, Lindemann D. 2014. The cooperative function of arginine residues in the Prototype Foamy Virus Gag C-terminus mediates viral and cellular RNA encapsidation. Retrovirology 11:87.

98. Chen J, Nikolaitchik O, Singh J, Wright A, Bencsics CE, Coffin JM, Ni N, Lockett S, Pathak VK, Hu WS. 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. Proc Natl Acad Sci U S A 106:13535-40.

99. Darlix JL, de Rocquigny H, Mauffret O, Mely Y. 2014. Retrospective on the all-in-one retroviral nucleocapsid protein. Virus Res 193:2-15.

100. Perilla JR, Gronenborn AM. 2016. Molecular Architecture of the Retroviral Capsid. Trends Biochem Sci 41:410-420.

101. Engelman AN. 2019. Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. J Biol Chem doi:10.1074/jbc.REV119.006901.

102. Brooks KM, Sherman EM, Egelund EF, Brotherton A, Durham S, Badowski ME, Cluck DB. 2019. Integrase Inhibitors: After 10 Years of Experience, Is the Best Yet to Come? Pharmacotherapy 39:576-598.

103. Smith SJ, Zhao XZ, Passos DO, Lyumkis D, Burke TR, Hughes SH. 2020. HIV-1 Integrase Inhibitors that are active against Drug-Resistant Integrase Mutants. Antimicrob Agents Chemother doi:10.1128/AAC.00611-20.

104. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268.e12.

105. Chen J, Rahman SA, Nikolaitchik OA, Grunwald D, Sardo L, Burdick RC, Plisov S, Liang E, Tai S, Pathak VK, Hu WS. 2016. HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. Proc Natl Acad Sci U S A 113:E201-8.

106. Checkley MA, Luttge BG, Freed EO. 2011. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J Mol Biol 410:582-608.

107. Wang Q, Finzi A, Sodroski J. 2020. The Conformational States of the HIV-1 Envelope Glycoproteins. Trends Microbiol 28:655-667.

108. Chen B. 2019. Molecular Mechanism of HIV-1 Entry. Trends Microbiol 27:878-891.

Chapter 3: Studying the role of integrase and other RNA-binding proteins in the replication of retroviruses

Christian Shema Mugisha[1], Pratibha C. Koneru[2], Kasyap Tenneti[1], Wen Li[3,4], Mamuka Kvaratskhelia[2], Alan N. Engelman[3,4], Sebla B. Kutluay[1]

Part of this work has been published in Methods

Affiliations:

[1] Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110, USA

[2] Division of Infectious Diseases, University of Colorado School of Medicine, Aurora, CO 80045

[3] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02215

[4] Department of Medicine, Harvard Medical School, Boston, MA 02115

Contributions:

Christian Shema Mugisha and Sebla B. Kutluay made the hypothesis and designed the experiments. Pratibha C. Koneru perfomed the integrase invitro binding assays. Wen Li and the Engelman provided the retroviral plasmids and integrase antibodies for MLV, EIAV, and MMTV.

3.1 Introduction

For decades, the HIV-1 integrase (IN) was solely known for catalyzing the 3' end processing and strand transfer reactions which lead to integration of the newly reverse transcribed viral DNA (vDNA) into the host chromosome (5, 7-9). Recent findings from our lab have showed that integrase binds to the viral genomic RNA (gRNA) in mature virion particles (10-12). Crosslinking immunoprecipitation coupled with next generation sequencing (CLIPseq) show that IN binds to multiple distinct regions of the gRNA including TAR, a hairpin within the 5' and 3' UTRs of the HIV-1gRNA (12). Furthermore, *in vitro* binding assays showed that IN had particularly high binding affinity for TAR (12). This RNA-binding property of IN is essential for maturation of virion particles; class II IN mutations and allosteric integrase inhibitors (ALLINIs) cause a virion deformation referred to as eccentric morphology by inhibiting IN-gRNA binding (10, 12-22). The eccentric morphology of HIV-1 virions is characterized by the mislocalization of the viral ribonucleoprotein complex (vRNP) outside of the capsid lattice (10, 12-22). The gRNA and IN of eccentric virions gets degraded in the earliest stages of the HIV replication cycle (11, 23).

Our previous studies have shown that IN binds to gRNA as a tetramer and the binding is mediated by electrostatic affinity between basic residues of the IN C-terminal domain (CTD) and the negatively charged phosphodiester RNA backbone (10). This observation led to the question of whether IN only specifically binds to the viral gRNA. For instance, although the IN has high affinity for the TAR region on 5' UTR gRNA, it also binds to discrete regions all around the length of gRNA. In this chapter, we used CLIPseq (described in 3.2) and *in vitro* binding assays

to show that IN binds to gRNA regions rich in adenosine and guanosine nucleotides (A-G). Further, CLIP experiments showed that IN could bind to non-viral RNA packaged by virions. Altogether, these observations suggested that the binding of IN to RNA might be semi-specific in nature.

The RNA-binding property of IN is essential for the formation of replication competent virions, which makes it a great target for the future antiretroviral agents. The presence of ALLINIs in virion producer cells inhibits IN-gRNA binding by causing aberrant IN multimerization (10-12, 23). However, there has been no evidence that IN binds to gRNA in producer cells before virion release as gRNA is packaged and condensed by the nucleocapsid (NC) protein in virions (12). In this chapter, we describe our investigations and findings on whether IN binds to RNA before or after virions release; such information could be very useful in designing less toxic and more potent ALLINIs. The development of IN-gRNA binding inhibitors will depend on a full understanding of the IN-gRNA complex structures. Nonetheless, integrase protein has been virtually impossible to crystallize by standard structural biology techniques due to its low solubility (24-28). This chapter details our new approaches for isolating and resolving the structures of IN-gRNA complexes.

3.2 Identifying the RNA targets of RNA binding proteins with CLIP

Viruses containing RNA genomes as their genetic material (i.e. RNA viruses and retroviruses) cause a number of diseases ranging from the common cold to AIDS, and are major contributors to the global infectious disease burden. To replicate and propagate their RNA genomes, these viruses must efficiently utilize host cell machinery to facilitate a number of processes including viral RNA transcription, splicing, transport between subcellular compartments and translation.

Viral RNAs also must retain unique sequence and structural features in order for recognition by viral RNA binding proteins (RBPs) that mediate RNA polymerization and export, selective genome packaging, RNA localization in virions, and RNA stability during early stages of infection. While beneficial for virus replication, the distinct features of viral RNAs can also be recognized as foreign by several host defense proteins that mount antiviral defenses through the production of cytokines upon binding to viral RNAs. In sum, a myriad of protein-RNA interactions play crucial roles in the life cycle of viruses, and as such identifying these interactions is crucial to understanding their pathogenesis.

Development of the CLIP approach has revolutionized the study of protein-RNA interactions (29-32). Traditionally, RBP-RNA interactions have been studied by methods that required the a *priori* knowledge of the binding site, such as in vitro binding assays or purification of protein-RNA complexes from cell lysates with downstream analysis of bound RNA by Q-RT-PCR. Coupling of protein-RNA complex isolation with microarray analysis, and more recently, high throughput sequencing provided a more global picture of the bound RNAs, but suffered from high background and low specificity due to purification of non-specific RNAs or multiple RBPs in complex with their bound RNAs (33, 34). While in vitro approaches such as SELEX (35, 36), RNA-compete (37), RNA-bind-n-Seq (38), and RNA-Map (39) can provide detailed nucleotide resolution information of the target site and the biochemical properties of these interactions, they cannot determine which sequence the RBP of interest will bind to at physiologically relevant concentrations. What makes CLIP a powerful approach is that it addresses all of the shortcomings of these previous approaches by yielding nucleotide resolution information of the RNA molecules bound by the RBP of interest in physiological settings, ranging from virus particles to animal tissues, at a global scale.

The key steps of the existing CLIP methodologies are (Figure 1): (1) Protein-RNA complexes are covalently crosslinked in live cells/tissues/virions typically by UV crosslinking; (2) Cells/tissues/virions are lysed and treated with limited amounts of RNases leaving small fragments of RNA molecules (~20–50 nucleotides) protected by the protein of interest; (3) Protein-RNA complexes are immunoprecipitated, and non-specific RNAs and proteins are removed by stringent washes. Because the protein-RNA complexes are covalently crosslinked, these stringent conditions, in principle, do not affect purification of target protein-RNA adducts.

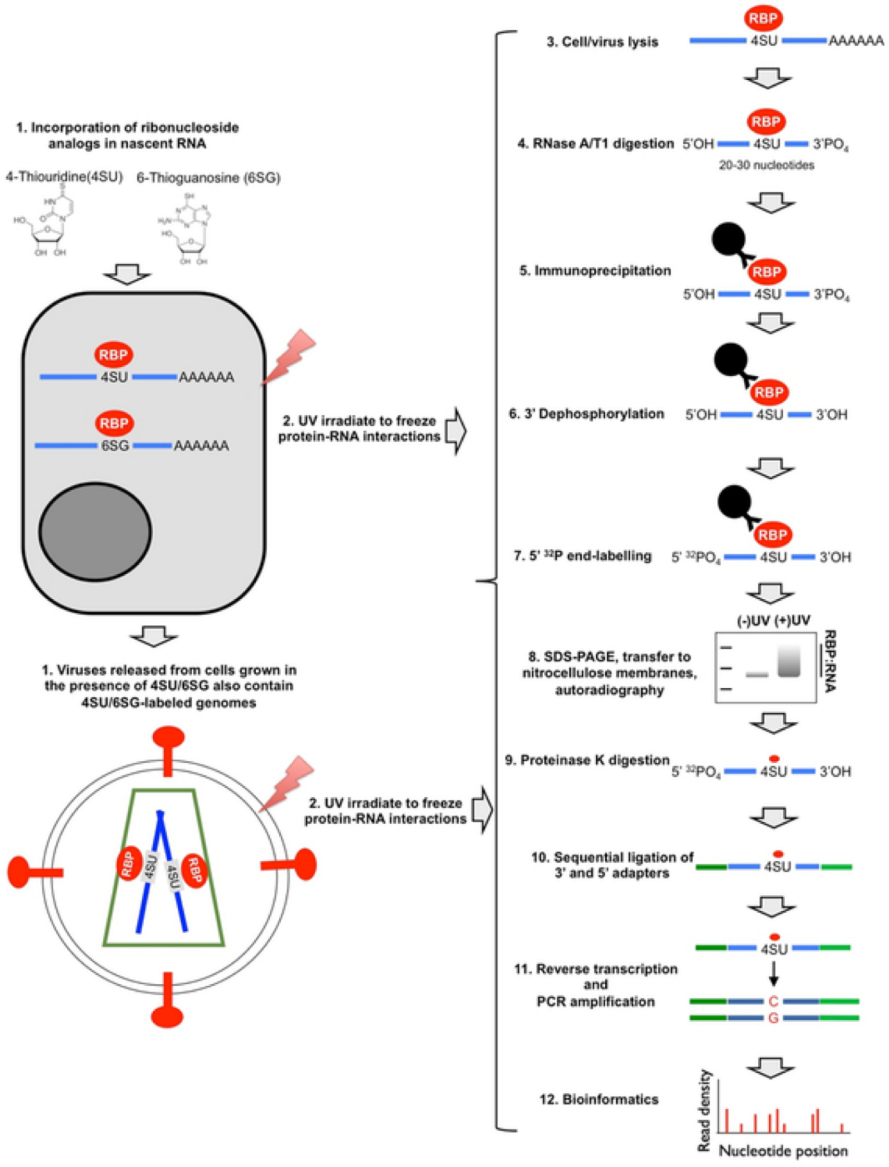

Figure 1. Schematic diagram of steps involved in PAR-CLIP experiments.

(4) The purified protein-RNA complexes are radioactively labeled and separated by SDS-PAGE. (5) Bound RNA is isolated by Proteinase K treatment. (5) Eluted RNA is ligated to adapters, reverse transcribed, and the resulting cDNA is PCR amplified and subjected to sequencing. (6) Sequencing reads are processed and mapped to reference genomes. Depending on the method used, the resulting library contains nucleotide substitutions or deletions (40) at the site of crosslinking, which allows mapping of the site of protein-RNA interactions at nucleotide resolution.

Various versions of CLIP have been developed and the details of these alternative approaches have been reviewed elsewhere (33, 41). The approach we commonly use depends on efficient UV crosslinking mediated by ribonucleoside analogs, including 4-thiouridine (4SU) and 6-thioguanosine (6SG), as in the original PAR-CLIP protocol (29). In PAR-CLIP experiments, cells are grown in the presence of ribonucleoside analogs for up to 16 hours and UV-crosslinked at a longer wavelength (365 nm). With the majority of the viral and cellular RNA-binding proteins we studied in the past few years, this method significantly enhanced the amount of protein-RNA adducts obtained (12, 42-44). In addition, PAR-CLIP allows accurate nucleotide resolution mapping of target RNA sites due to mutations introduced by the reverse transcriptase enzyme (T-to-C for 4SU and G-to-A for 6SG) precisely at the site of crosslinking during cDNA synthesis. On the other hand, use of ribonucleoside analogs may inadvertently enrich RNA elements with distinct nucleotide composition or alter RNA structure (45), which may subsequently affect protein binding. Thus, we routinely validate our findings with the use of both ribonucleoside analogs and HITS CLIP, which utilizes conventional UV-crosslinking.

For library generation, we follow a combination of PAR-CLIP and HITS-CLIP protocols (30-32). As reviewed elsewhere (33, 34, 41), this is the stage where most variant CLIP protocols differ. While the HITS-CLIP and many other protocols call for ligation of adapters while the protein-RNA complexes are on beads, the solution-phase PAR-CLIP library generation protocol was significantly more efficient in our experience with 3' and 5' adapter ligations routinely functioning at >90% efficiency. Another common alternative is the utilization of a two-part cleavable adapter introduced into cDNA during reverse transcription in the iCLIP approach (46). This is followed by circularization and restriction enzyme digestion, which allows the recovery of a larger fraction of truncated cDNAs a result of reverse transcriptase stalling at crosslinking sites. Due to this enrichment, iCLIP can yield higher complexity libraries and has been proposed to perform better than previous approaches in identification of the precise site of crosslinking (46-48).

All CLIP approaches are technically challenging with numerous labor-intensive steps. In addition, the loss of the starting material at several inefficient steps is a major drawback. This problem is further exacerbated if the initial protein-RNA complexes are not abundant due to low levels of expression, low crosslinking or immunoprecipitation efficiencies. These problems can often lead to a final library with insufficient complexity and over-enrichment of environmental contaminating sequences.

3.3 Integrase preferably binds to purine-rich sequences on viral RNA

Although CLIPseq and *in vitro* binding assays have shown that IN has high affinity for TAR, a hairpin structure within the 5' UTR of gRNA (12), CLIPseq also shows that IN binds to gRNA sequences without distinguishable secondary structure. Altogether, the nature of IN-gRNA

binding is poorly understood, and it is not certain whether the binding is specific or IN binds to distinct characteristic sequences of viral gRNA. The version of CLIP (PAR-CLIP) that has previously been used to study IN-gRNA binding relies on the incorporation of photoreactive analogs such as 4-thiouridine (4-SU) into nascent gRNAs in producer cells (12, 49); this enhances the efficiency of crosslinking, the amount IN-gRNA adducts that are immunoprecipitated, and helps identify IN-gRNA binding sites at nearly nucleotide resolution. Although the incorporation of 4-SU is very helpful, it can bias binding site identification towards U-rich elements. We used a guanosine analog, 6-thioguanosine (6-SG), in CLIP experiments to get a distinct picture of gRNA elements that are bound IN. The 6-SG based CLIPseq identified more distinctive binding sites with high frequency (Figure 2A). A deeper analysis of the motifs within the binding sites showed that IN binds to A/G rich (Figure 2B) sequences most of which contained the 5'-GGAAAGGA-3' sequence.

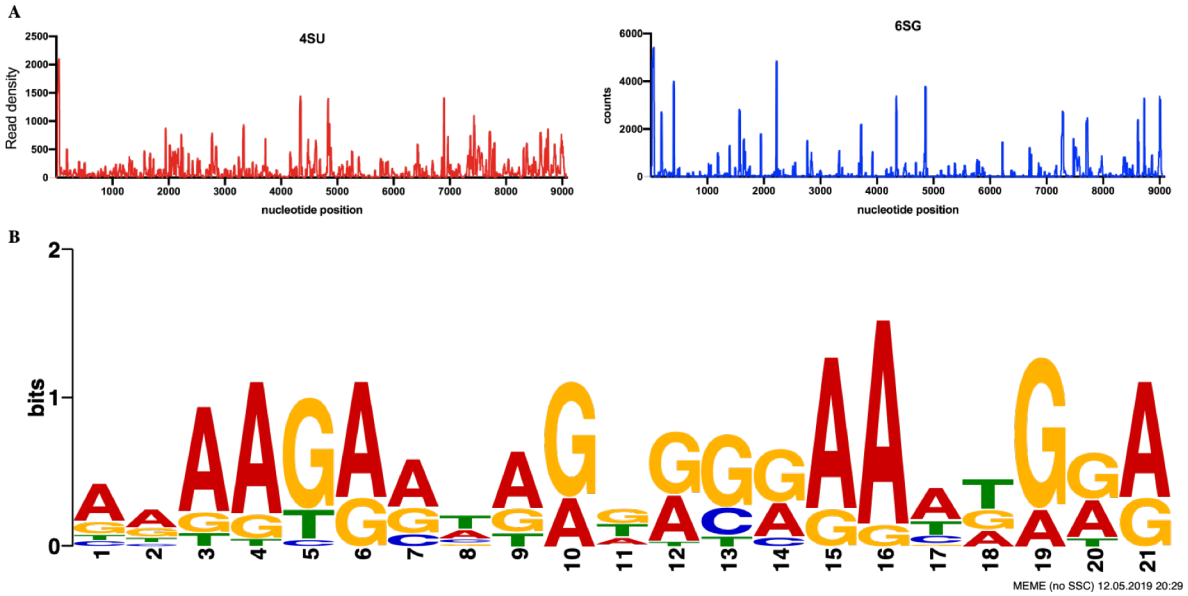

Figure 2. Purine-rich sequences are preferentially bound by IN. A. CLIP was performed on virions generated in the presence of 4-SU and 6-SG. Footprints of IN on the 9kb viral genome is shown. **B.** Meme plot of the most prevalent nucleotides in integrase bound regions.

This observation was validated through in-vitro biochemical binding assays in collaboration with Dr. Mamuka Kvaratskhelia's lab at the University of Colorado. IN showed high propensity to binding the 5'-GGAAAGGA-3' consensus sequence, but not its reverse complement or the double stranded form (Figure. 3A). Furthermore, the addition of purine nucleotides to the 5' and 3' of the sequence increased the binding affinity to IN (Figure 3B).

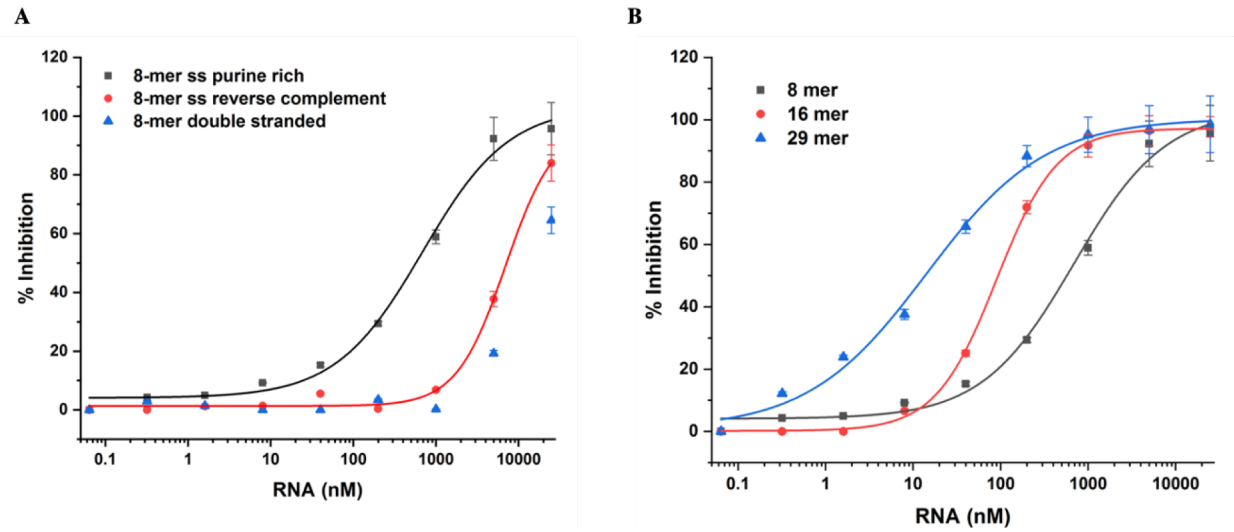

Figure 3. **Purine-rich sequences are preferentially bound by IN**. **A.** In vitro assessment of IN binding affinity to the 5'-GGAAAGGA-3' RNA sequences (black), its reverse compliment (red), or its double stranded version (blue). **B.** Similar to A but with addition of several AG nucleotides on both ends of the 5'-GGAAAGGA-3' sequence.

The observed high binding frequency and affinity of IN to purine-rich sequences might be a result of the fact that the HIV-1 genome has a particularly high A-nucleotide content (50, 51). In the absence of HIV-1 gRNA, virions can package cellular RNAs (52). To test whether IN can bind to non-viral packaged RNA, we generated virus like particles (VLPs) devoid of the viral genome but package cellular RNA. These VLPs were made using plasmids bearing a CMV promoter driving the expression of a synthetic codon-optimized Gag-Pol (SYNGP) gene. CLIP autoradiographs show that IN in VLPs could bind to RNA (Figure 4A). Moreover, the ALLINI

BI-D and class II mutation R269A/K273A blocked RNA binding (Figure 4B). Although this observation suggests that IN can bind to cellular RNAs, it remains unclear whether IN binds to purine rich sequences on these RNAs as of now. CLIP to identify IN-bound sequences from the VLPs will be instrumental to distinguish whether the observed RNA-binding specificity of IN is a result of AG-rich genome content or specific interaction of IN with purine rich sequence elements.

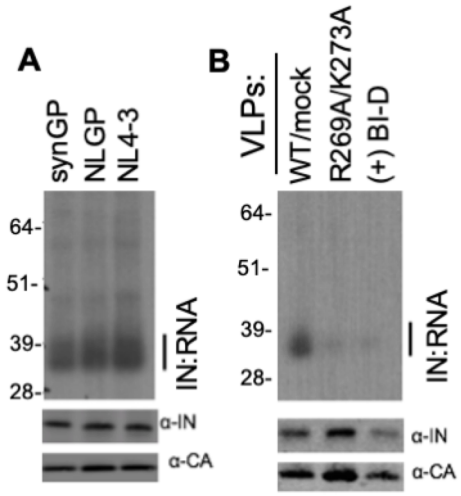

Figure 4. HIV-1 IN can bind to cellular RNAs in the absence of the genome. A, synGP and NLGP-derived VLPs were subjected to CLIP analysis alongside with full-length NL4-3. **B,** synGP-derived VLPs bearing the R269A/K273A substitution and grown in the presence of 10 uM BI-D were subjected to CLIP. Autorads displaying IN-RNA complexes are shown on top. W.blots of IN and CA are shown on the bottom.

3.4 Conservation of integrase-RNA binding in retroviruses

IN-gRNA binding is essential for maturation of HIV-1 virions (10-12, 23). Following the release of immature virions from cells, the protease domain of Gag-Pol cleaves Gag and Gag-Pol into their constituent domains which triggers the reorganization of these proteins into a mature HIV-1 virion. In mature HIV-1virions, the capsid lattice surrounds the viral ribonucleoprotein complex (53, 54). While proteolytic cleavage of Gag and Gag-Pol polyproteins during particle maturation

is common to all retroviruses, virion morphologies vary widely. Looking at different retroviral species, the capsids of mature virions adopt a multitude of shapes and sizes (Figure 5)(55). While HIV-1 and lentivirus capsid lattices are mostly cone-shaped (56-59), retroviruses such as RSV and MLV display polyhedral or nearly spherical capsids (58-62). It has been shown that HTLV-1 has a poorly defined polyhedral capsid, with angular polygon-like regions and at least one curved region in each capsid (63).

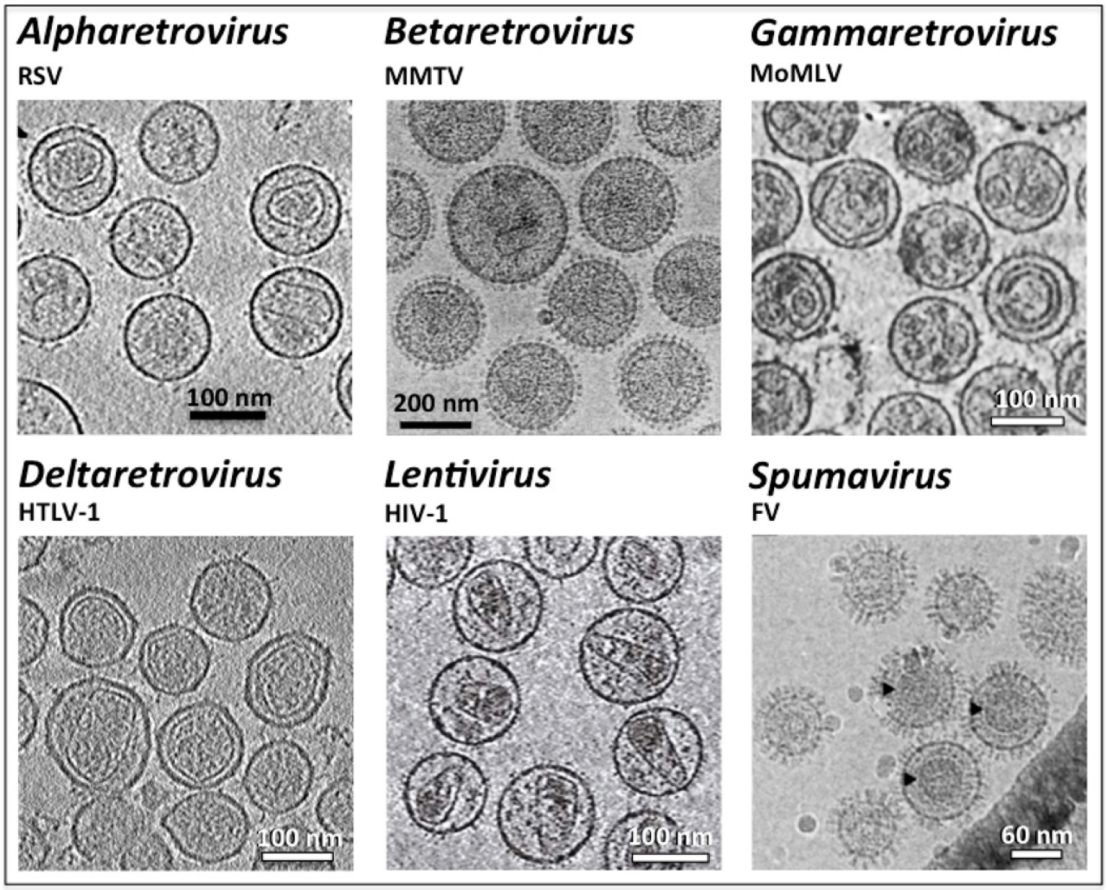

Figure 5. Morphology of mature retrovirus virions. RSV: Rous sarcoma virus (1, 2), MMTV: mouse mammary tumor virus (1-3), MoMLV: Moloney murine leukemia virus , HTLV-1: human T-cell leukemia virus type 1 (4), HIV-1: human immunodeficiency virus type 1(5), FV: Foamy virus (6, 7). The FV Gag protein is not processed into the classical orthoretroviral MA, CA, and NC subunits during particle morphogenesis. The black arrowheads show regular Gag assemblies in the wild type virus .

It is currently unknown whether IN-gRNA binding is essential for the encapsidation of the vRNP of all retroviruses. Using CLIP, we have investigated whether IN binds to the genomic RNA in other retroviruses: lentiviruses (HIV-1, MVV, and EIAV), beta-retrovirus (MMTV), delta-retrovirus (HTLV-1), and gamma-retrovirus (MLV). IN-RNA adducts on our CLIP autoradiographs show that IN binds to the genomic RNA of such viruses (Figure 6). Furthermore, some of the mutations that were predicted as class II mutations based on alignment with HIV-1 IN sequence also inhibited IN-gRNA (Figure 6). The introduction of such mutations in some viruses, however, resulted in loss of IN in virions and the mutations were thus not interpretable (Figure 6). These initial findings show that IN-gRNA binding is conserved across retroviruses, and they suggest that so might the role of IN in virion morphogenesis. Our future experiments will focus on more detailed characterization of class II phenotype with other retroviruses using a combination of thin section electron microscopy and CLIP to investigate how inhibition of IN-gRNA binding affects retroviral morphology.

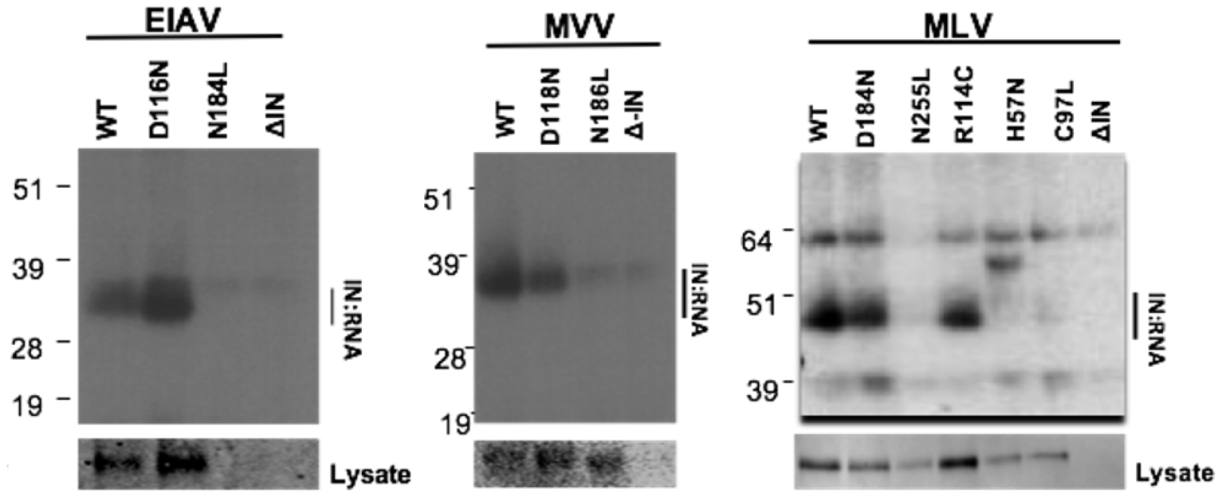

Figure 6. Autoradiographs (top) showing IN-RNA adducts and Western blots of IN (bottom) in MLV, EIAV, MVV, and MMTV

3.5 Discussion

This chapter provides tools and data that can serve as the basis for further studies into how IN-RNA binding regulates the morphogenesis of HIV-1 and other retroviruses. CLIP not only provides a powerful tool to identify whether a protein of interest binds to RNA in physiologically relevant settings, but it also permits to identify protein-bound sequences at nearly nucleotide resolution. CLIPseq has revealed that HIV-1 IN preferentially binds to AG-rich sequences on the genomic RNA packaged inside virions. Furthermore, *in vitro* binding assays have shown that IN has higher binding affinity for AG-rich sequences. Although IN binds to specific sequences of the viral genomic RNA with high affinity, we observed that IN can also bind to packaged cellular RNAs in the absence of the viral genome. This observation suggests that the binding of IN-gRNA is semi specific. The nature of IN-gRNA will not be completely understood until the structure of IN-gRNA complexes is available. Over the years, studying the full structure of IN has been nearly impossible with traditional protein isolation methods due to the low solubility of integrase. We are building immunoprecipitation-based tools to purify IN-RNA complexes from virions, and initial findings show that full length IN is isolated with relatively high efficiency. The structure of IN complexes purified in such ways can be resolved using cryo-EM through our existing collaborations. Over and above that, we show that IN-RNA binding is conserved in other retroviruses. Characterizing the structure and properties of IN-gRNA binding and the molecular mechanism through they regulate virion morphology will be crucial in designing a new class of integrase-targeting antiretroviral agents.

References

1. Chen B. 2019. Molecular Mechanism of HIV-1 Entry. Trends Microbiol 27:878-891.

2. Hu WS, Hughes SH. 2012. HIV-1 reverse transcription. Cold Spring Harb Perspect Med 2.

3. Dharan A, Bachmann N, Talley S, Zwikelmaier V, Campbell EM. 2020. Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nat Microbiol doi:10.1038/s41564-020-0735-8.

4. Burdick RC, Li C, Munshi M, Rawson JMO, Nagashima K, Hu WS, Pathak VK. 2020. HIV-1 uncoats in the nucleus near sites of integration. Proc Natl Acad Sci U S A 117:5486-5493.

5. Lesbats P, Engelman AN, Cherepanov P. 2016. Retroviral DNA Integration. Chem Rev 116:12730-12757.

6. Karn J, Stoltzfus CM. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med 2:a006916.

7. Craigie R, Bushman FD. 2012. HIV DNA integration. Cold Spring Harb Perspect Med 2:a006890.

8. Engelman A, Mizuuchi K, Craigie R. 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67:1211-21.

9. Engelman AN, Singh PK. 2018. Cellular and molecular mechanisms of HIV-1 integration targeting. Cell Mol Life Sci 75:2491-2507.

10. Elliott JL, Eschbach JE, Koneru PC, Li W, Puray-Chavez M, Townsend D, Lawson DQ, Engelman AN, Kvaratskhelia M, Kutluay SB. 2020. Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion morphogenesis. Elife 9.

11. Elliott JL, Kutluay SB. 2020. Going beyond Integration: The Emerging Role of HIV-1 Integrase in Virion Morphogenesis. Viruses 12.

12. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268 e12.

13. 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. J Virol 69:2729-36.

14. Nakamura T, Masuda T, Goto T, Sano K, Nakai M, Harada S. 1997. Lack of infectivity of HIV-1 integrase zinc finger-like domain mutant with morphologically normal maturation. Biochem Biophys Res Commun 239:715-22.

15. 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.

16. Shin CG, Taddeo B, Haseltine WA, Farnet CM. 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. J Virol 68:1633-42.

17. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, Gorelick RJ, Engelman AN, Steven AC. 2015. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. J Virol 89:9765-80.

18. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, Wang W, Ballandras-Colas A, Patel PA, Fuchs JR, Kvaratskhelia M, Engelman A. 2013. Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. Proc Natl Acad Sci U S A 110:8690-5.

19. Jenkins TM, Engelman A, Ghirlando R, Craigie R. 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J Biol Chem 271:7712-8.

20. Balakrishnan M, Yant SR, Tsai L, O'Sullivan C, Bam RA, Tsai A, Niedziela-Majka A, Stray KM, Sakowicz R, Cihlar T. 2013. Non-catalytic site HIV-1 integrase inhibitors disrupt core maturation and induce a reverse transcription block in target cells. PLoS One 8:e74163.

21. Desimmie BA, Schrijvers R, Demeulemeester J, Borrenberghs D, Weydert C, Thys W, Vets S, Van Remoortel B, Hofkens J, De Rijck J, Hendrix J, Bannert N, Gijsbers R, Christ F, Debyser Z. 2013. LEDGINs inhibit late stage HIV-1 replication by modulating integrase multimerization in the virions. Retrovirology 10:57.

22. Sharma A, Slaughter A, Jena N, Feng L, Kessl JJ, Fadel HJ, Malani N, Male F, Wu L, Poeschla E, Bushman FD, Fuchs JR, Kvaratskhelia M. 2014. A new class of multimerization selective inhibitors of HIV-1 integrase. PLoS Pathog 10:e1004171.

23. Madison MK, Lawson DQ, Elliott J, Ozanturk AN, Koneru PC, Townsend D, Errando M, Kvaratskhelia M, Kutluay SB. 2017. Allosteric HIV-1 Integrase Inhibitors Lead to Premature Degradation of the Viral RNA Genome and Integrase in Target Cells. J Virol 91.

24. 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. Proc Natl Acad Sci U S A 97:8233-8.

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

26. Li M, Jurado KA, Lin S, Engelman A, Craigie R. 2014. Engineered hyperactive integrase for concerted HIV-1 DNA integration. PLoS One 9:e105078.

27. Li X, Krishnan L, Cherepanov P, Engelman A. 2011. Structural biology of retroviral DNA integration. Virology 411:194-205.

28. Passos DO, Li M, Yang R, Rebensburg SV, Ghirlando R, Jeon Y, Shkriabai N, Kvaratskhelia M, Craigie R, Lyumkis D. 2017. Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. Science 355:89-92.

29. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141:129-41.

30. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456:464-9.

31. Ule J, Jensen K, Mele A, Darnell RB. 2005. CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods 37:376-86.

32. Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. 2003. CLIP identifies Nova-regulated RNA networks in the brain. Science 302:1212-5.

33. Lee FCY, Ule J. 2018. Advances in CLIP Technologies for Studies of Protein-RNA Interactions. Mol Cell 69:354-369.

34. Ule J, Hwang HW, Darnell RB. 2018. The Future of Cross-Linking and Immunoprecipitation (CLIP). Cold Spring Harb Perspect Biol 10.

35. Ellington AD, Szostak JW. 1990. In vitro selection of RNA molecules that bind specific ligands. Nature 346:818-22.

36. Tuerk C, Gold L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505-10.

37. Ray D, Kazan H, Chan ET, Pena Castillo L, Chaudhry S, Talukder S, Blencowe BJ, Morris Q, Hughes TR. 2009. Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. Nat Biotechnol 27:667-70.

38. Lambert N, Robertson A, Jangi M, McGeary S, Sharp PA, Burge CB. 2014. RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. Mol Cell 54:887-900.

39. Buenrostro JD, Araya CL, Chircus LM, Layton CJ, Chang HY, Snyder MP, Greenleaf WJ. 2014. Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. Nat Biotechnol 32:562-8.

40. Zhang C, Darnell RB. 2011. Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol 29:607-14.

41. Bieniasz PD, Kutluay SB. 2018. CLIP-related methodologies and their application to retrovirology. Retrovirology 15:35.

42. Kutluay SB, Emery A, Penumutchu SR, Townsend D, Tenneti K, Madison MK, Stukenbroeker AM, Powell C, Jannain D, Tolbert BS, Swanstrom RI, Bieniasz PD. 2019. Genome-Wide Analysis of Heterogeneous Nuclear Ribonucleoprotein (hnRNP) Binding to HIV-1 RNA Reveals a Key Role for hnRNP H1 in Alternative Viral mRNA Splicing. J Virol 93.

43. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, Errando M, Bieniasz PD. 2014. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell 159:1096-1109.

44. York A, Kutluay SB, Errando M, Bieniasz PD. 2016. The RNA Binding Specificity of Human APOBEC3 Proteins Resembles That of HIV-1 Nucleocapsid. PLoS Pathog 12:e1005833.

45. Testa SM, Disney MD, Turner DH, Kierzek R. 1999. Thermodynamics of RNA-RNA duplexes with 2- or 4-thiouridines: implications for antisense design and targeting a group I intron. Biochemistry 38:16655-62.

46. Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 17:909-15.

47. Haberman N, Huppertz I, Attig J, Konig J, Wang Z, Hauer C, Hentze MW, Kulozik AE, Le Hir H, Curk T, Sibley CR, Zarnack K, Ule J. 2017. Insights into the design and interpretation of iCLIP experiments. Genome Biol 18:7.

48. Sugimoto Y, Konig J, Hussain S, Zupan B, Curk T, Frye M, Ule J. 2012. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. Genome Biol 13:R67.

49. Shema Mugisha C, Tenneti K, Kutluay SB. 2020. Clip for studying protein-RNA interactions that regulate virus replication. Methods 183:84-92.

50. Kypr J, Mrazek J. 1987. Unusual codon usage of HIV. Nature 327:20.

51. van der Kuyl AC, Berkhout B. 2012. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. Retrovirology 9:92.

52. Rulli SJ, Jr., Hibbert CS, Mirro J, Pederson T, Biswal S, Rein A. 2007. Selective and nonselective packaging of cellular RNAs in retrovirus particles. J Virol 81:6623-31.

53. Sundquist WI, Krausslich HG. 2012. HIV-1 assembly, budding, and maturation. Cold Spring Harb Perspect Med 2:a006924.

54. Freed EO. 2015. HIV-1 assembly, release and maturation. Nat Rev Microbiol 13:484-96.

55. Zhang W, Cao S, Martin JL, Mueller JD, Mansky LM. 2015. Morphology and ultrastructure of retrovirus particles. AIMS Biophys 2:343-369.

56. Briggs JA, Wilk T, Welker R, Krausslich HG, Fuller SD. 2003. Structural organization of authentic, mature HIV-1 virions and cores. EMBO J 22:1707-15.

57. Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI. 1999. Assembly and analysis of conical models for the HIV-1 core. Science 283:80-3.

58. Kingston RL, Olson NH, Vogt VM. 2001. The organization of mature Rous sarcoma virus as studied by cryoelectron microscopy. J Struct Biol 136:67-80.

59. Yeager M, Wilson-Kubalek EM, Weiner SG, Brown PO, Rein A. 1998. Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. Proc Natl Acad Sci U S A 95:7299-304.

60. Butan C, Winkler DC, Heymann JB, Craven RC, Steven AC. 2008. RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis. J Mol Biol 376:1168-81.

61. Ganser-Pornillos BK, Yeager M, Pornillos O. 2012. Assembly and architecture of HIV. Adv Exp Med Biol 726:441-65.

62. Heymann JB, Butan C, Winkler DC, Craven RC, Steven AC. 2008. Irregular and Semi-Regular Polyhedral Models for Rous Sarcoma Virus Cores. Comput Math Methods Med 9:197-210.

63. Cao S, Maldonado JO, Grigsby IF, Mansky LM, Zhang W. 2015. Analysis of human T-cell leukemia virus type 1 particles by using cryo-electron tomography. J Virol 89:2430-5.

Chapter 4: A simplified quantitative real-time PCR assay for monitoring SARS-CoV-2 growth in cell culture

Christian Shema Mugisha [1,*], Hung R. Vuong [1,*], Maritza Puray-Chavez [1], Adam L. Bailey[2], Julie M. Fox[3], Rita E. Chen[2,3], Alex W. Wessel[2,3], Jason M. Scott[3], Houda H. Harastani[3], Adrianus C. M. Boon[1,2,3], Haina Shin[3], Sebla B. Kutluay[1]

This work has been published in mSphere journal

Affiliations

[1] Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110, USA

[2] Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA

[3] Department of Medicine, Washington University School of Medicine, Saint Louis, MO 63110, USA

Contributions:

Christian Shema Mugisha, Maritza Puray-Chavez, and Sebla B. Kutluay made the hypothesis and designed the experiments. Hung R. Vuong carried the infection of cells and the BSL3-part of the experiments. The rest of the authors provided PCR reagents or virus inoculum.

***** Christian Shema Mugisha and Hung R. Vuong contributed equally to this work. Author order was determined both alphabetically and in order of increasing seniority.

4.1 Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected millions within just a few months causing severe respiratory disease and mortality. Assays to monitor SARS-CoV-2 growth *in vitro* depend on time-consuming and costly RNA extraction steps, hampering progress in basic research and drug development efforts. Here we developed a simplified quantitative real-time PCR (Q-RT-PCR) assay that bypasses viral RNA extraction steps and can monitor SARS-CoV-2 growth from a small amount of cell culture supernatants. In addition, we show that this approach is easily adaptable to numerous other RNA and DNA viruses. Using this assay, we screened the activities of a number of compounds that were predicted to alter SARS-CoV-2 entry and replication as well as HIV-1-specific drugs in a proof of concept study. We found that E64D (inhibitor of endosomal proteases cathepsin B & L) and apilimod (endosomal trafficking inhibitor) potently decreased the amount of SARS-CoV-2 RNA in cell culture supernatants with minimal cytotoxicity.,Surprisingly, we found that the macropinocytosis inhibitor EIPA similarly decreased SARS-CoV-2 RNA levels in supernatants suggesting that entry may additionally be mediated by an alternative pathway. HIV-1-specific inhibitors nevirapine (an NNRTI), amprenavir (a protease inhibitor), and ALLINI-2 (an allosteric integrase inhibitor) modestly inhibited SARS-CoV-2 replication, albeit the IC_{50} values were much higher than that required for HIV-1. Taken together, this simplified assay will expedite basic SARS-CoV-2 research, be amenable to mid-throughput screening assays (i.e. drugs, CRISPR, siRNA, etc.), and be applicable to a broad number of RNA and DNA viruses.

4.2 Importance

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiological agent of the COVID-19 pandemic is continuing to cause immense respiratory disease and social and economic disruptions. Conventional assays that monitor SARS-CoV-2 growth in cell culture rely on costly and time-consuming RNA extraction procedures, hampering progress in basic SARS-CoV-2 research and development of effective therapeutics. Here we developed a simple quantitative real-time-PCR assay to monitor SARS-CoV-2 growth in cell culture supernatants that does not necessitate RNA extraction, and is as accurate and sensitive as existing methods. In a proof-of-concept screen, we found that E64D, apilimod, EIPA and remdesivir can substantially impede SARS-Cov-2 replication providing novel insight into viral entry and replication mechanisms. In addition, we show that this approach is easily adaptable to numerous other RNA and DNA viruses. This simplified assay will undoubtedly expedite basic SARS-CoV-2 and virology research, and be amenable to drug screening platforms to identify therapeutics against SARS-CoV-2.

4.3 Observation

Severe acute respiratory syndrome coronavirus, SARS-CoV-2 is continuing to cause substantial morbidity and mortality around the globe (1, 2). Lack of a simple assay to monitor virus growth is slowing progress basic SARS-CoV-2 research as well as drug discovery. Current methods to quantify SARS-CoV-2 growth in cell culture supernatants rely on time-consuming and costly RNA extraction protocols followed by quantitative real-time PCR (Q-RT-PCR) (3). In this study, we developed a simplified Q-RT-PCR assay that bypasses the RNA extraction steps, can detect viral RNA from as little as 5 µL of cell culture supernatants and works equally well with TaqMan and SYBR-Green-based detection methods.

A widely used assay to measure virus growth in the retrovirology field relies on determining the activity of virion-associated reverse transcriptase enzyme collected from a small amount of infected cell culture supernatants (4). We reasoned that we could adapt this approach to monitor SARS-CoV-2 growth. First, we tested whether the more stringent lysis conditions used to inactivate SARS-CoV-2 would interfere with the subsequent Q-RT-PCR step. To do so, 5 µL of serially diluted RNA standards prepared by in vitro transcription from a plasmid containing the entire SARS-CoV-2 nucleoprotein (N) gene were mixed with 5 µL of 2x RNA lysis buffer (2% Triton X-100, 50 mM KCl, 100 mM TrisHCl pH7.4, 40% glycerol, 0.4 U/µL of SuperaseIN (Life Technologies)), followed by addition of 90 µL of 1X core buffer (5 mM $(NH_4)_2SO_4$, 20 mM KCl and 20 mM Tris–HCl pH 8.3). 8.5 µL of the diluted samples were added to 11.5 µL of a reaction mix consisting of 10 µL of a 2x TaqMan RT-PCR mix, 0.5 µL of a 40x Taqman RT enzyme mix (containing ArrayScript™ UP Reverse Transcriptase, RNase Inhibitor), and 1 µL of a mixture containing 10 pmoles of forward and reverse primers as well as 2 pmoles of Taqman Probe Table S1) resulting in a final reaction volume of 20 µL. The reactions were run on the

ViiA™ 7 Real-Time PCR system (Applied Biosystems) using the following cycling parameters: 48 ºC for 15 min, 95ºC for 10 min, 50 cycles of 95ºC for 15 sec and 60ºC for 1 min of signal acquisition. We found that the modified sample preparations did not impact the sensitivity, efficiency or the dynamic range of the Q-RT-PCR assay as evident in the virtually identical cycle threshold (Ct) values obtained for a given RNA concentration and the similar slopes of linear regression curves (Fig. 1A).

To determine whether this approach would work equally well for virus preparations, 100 µL of virus stock ($1.4x10^5$ pfu) was lysed via the addition of an equal volume of buffer containing 40 mM TrisHCl, 300 mM NaCl, 10 mM MgCl2, 2% Triton X-100, 2 mM DTT, 0.4 U/μL SuperaseIN RNase Inhibitor, 0.2% NP-40. RNA was then extracted using the Zymo RNA clean and concentrator^{TM}-5 kit and was serially diluted afterwards. In parallel, 5 µL of virus stock and its serial dilutions prepared in cell culture media were lysed in 2X RNA lysis buffer and processed as above. Samples were analyzed by Q-RT-PCR alongside with RNA standards. A standard curve was constructed by plotting the cycle threshold (Ct) value against the corresponding $\log_2$(copy number) of the RNA standards, which was subsequently used to determine copy numbers in samples. We then calculated the number of copies per milliliter of the original virus stock, assuming 100% recovery for samples subjected to RNA extraction. We found that the modified assay performed equivalently well, if not better, with a similarly broad dynamic range (Fig. 1B).

We next used this assay to monitor virus growth on infected Vero cells. Cell culture supernatants containing virus collected at various times post infection were either used to extract viral RNA or subjected to Q-RT-PCR directly (non-extracted) as above. The modified assay with non-extracted samples yielded virtually identical number of copies/mL of SARS-CoV-2 RNA in cell

culture supernatants even at low concentrations of viral RNAs (Fig. 1C). Collectively, these results suggest that RNA extraction from cell culture supernatants can be bypassed without any compromise on the sensitivity or the dynamic range of Q-RT-PCR detection.

Next, we wanted to test whether this assay could work equally well with SYBR-Green-based detection methods. In addition to the N primer pair used in the above TaqMan-based assays, we utilized the N2 primer set designed by CDC and targeting the N region of the SARS-CoV-2 genome (Table S1). Serially diluted RNA standards were processed in RNA lysis and core buffers, and 7.5 uL of each dilution was used in a 20 uL SYBR-Green Q-RT-PCR reaction containing 10μL of a 2X POWERUP SYBR Green mix (Life Technologies ref: A25742), 1.25units/ μL of MultiScribe Reverse Transcriptase (Applied Biosystems), 1X random primers and 25 pmoles of forward and reverse primers. Both primer pairs yielded reasonably broad dynamic ranges, but were modestly less sensitive than TaqMan-based assays with a detection limit of ~3500 RNA copies/mL (Fig. 1D).

We next tested whether this simplified Q-RT-PCR assay can be adapted to other RNA and DNA viruses. Dilutions of stocks of influenza A virus (IAV/PR8), herpes simplex virus type-2 (HSV-2), alphaviruses (Ross River virus (RRV), Chikungunya virus (CHIKV), Mayorovirus (MAYV)) and flaviviruses (Dengue virus (DENV-4), West Nile virus (WNV NY99) and Zikavirus (ZIKV-Dakar)) collected from cell culture supernatants were either subjected to RNA/DNA extraction or the simplified lysis protocol as above followed by SYBR-Green or Taqman-based Q-RT-PCR with the indicated primers (Table S1). For HSV-2, the reaction mixture did not include the reverse transcription enzyme and the initial reverse transcription step was skipped. We found that for IAV (Fig. 1E), HSV-2 (Fig. 1F) and RRV (Fig. 1G), the non-extracted samples worked equally well, and for CHIKV non-extracted samples gave lower Ct values across various virus

dilutions (Fig. 1H). For MAYV the dynamic range obtained from non-extracted samples was low compared to extracted samples (Fig. 1I), likely due to the incompatibility between lysis and PCR conditions. Although the Ct values were generally higher for non-extracted samples of ZIKV (Fig. 1J), WNV (Fig. 1K) and DENV (Fig. 1L), the dynamic range was still broad with similar PCR efficiencies between extracted and non-extracted samples. Taken together, these results demonstrate that the simplified Q-RT-PCR developed here can in principle be easily adapted to a large number of viruses provided that the lysis conditions are appropriate and working primer sets are present.

One immediate application of this simplified assay is in mid-throughput drug screening platforms (i.e. compound, CRISPR, siRNA screens) given the ease of quantitatively assessing viral growth from small quantities of cell culture media containing virions. To demonstrate this, we next conducted a proof-of-concept drug screen to validate the antiviral activities of various compounds that have been reported to inhibit SARS-CoV-2 and HIV-1 replication as well as non-specific entry inhibitors (Table S2). Vero E6 cells plated in 96-well plates were infected in the presence of varying concentrations of the indicated compounds. Viral RNA in cell culture supernatants was quantified by the SYBR-Green-based Q-RT-PCR assay as above at 6, 24 and 48 hpi. Compound cytotoxicity was assessed in parallel by the RealTime-Glo™ MT Cell Viability Assay (Promega). While viral RNA was at background levels at 6 hpi (data not shown), we found that, at 24hpi, remdesivir (inhibitor of RNA-dependent RNA polymerase, (5)), E64D (inhibitor of the endosomal protease cathepsin B, K and L), and apilimod (PIKfyve inhibitor resulting in endosomal trafficking defects, (6, 7)) substantially decreased SARS-CoV-2 viral RNAs in supernatants (Fig. 2). IC_{50} values of these compounds (2.8 μg/mL (remdesivir), 3.3 μM (E64D) and 12nM (apilimod)) were within the same range of published IC_{50} values of these

compounds (6-8) (Fig. 2). Similar results were obtained at 48 hpi, albeit E64D and apilimod appeared to be less potent at this time point either due to virus overgrowth or compound turnover (data not shown). We found that EIPA, which inhibits Na^+/H^+ exchanger and macropinocytosis, substantially decreased viral RNA in supernatants at sub-cytotoxic levels (Fig. 2D), suggesting that macropinocytosis may contribute to viral entry and/or subsequent steps in virus replication. HIV-1 specific inhibitors nevirapine, amprenavir and ALLINI-2 modestly inhibited SARS-CoV-2 replication without apparent cytotoxicity at high concentrations, albeit the concentrations required for this inhibition were much higher than those that inhibit HIV-1 (Fig. S1). Overall, these findings demonstrate that this simplified assay can be adapted for screening platforms and support previous reports which demonstrated that SARS-CoV-2 entry is dependent on processing of the Spike protein by cellular proteases and requires endosomal fusion (7, 9, 10).

In conclusion, we have developed a simple Q-RT-PCR assay to monitor the growth of SARS-CoV-2 as well as other viruses from cell culture supernatants, bypassing the time consuming and costly RNA extraction procedures. This simplified assay will undoubtedly expedite basic SARS-CoV-2 research, might be amenable to mid-throughput screens to identify chemical inhibitors of SARS-CoV-2 and can be applicable to the study of numerous other RNA and DNA viruses.

4.4 ACKNOWLEDGEMENTS

We thank members of the Whelan, Diamond and Amarasinghe labs for their generosity in providing reagents and support. This study was supported by Washington University startup funds to SBK, NIH grant U54AI150470 (the Center for HIV RNA Studies) to SBK, 1U01AI151810 - 01 to ACMB, NSF grant DGE-1745038 to HRV and Stephen I. Morse

fellowship to MPC. The Boon laboratory has scientific research agreements with AI therapeutics, Greenlight Biosciences and Nano Targeting & Therapy Biopharma Inc.

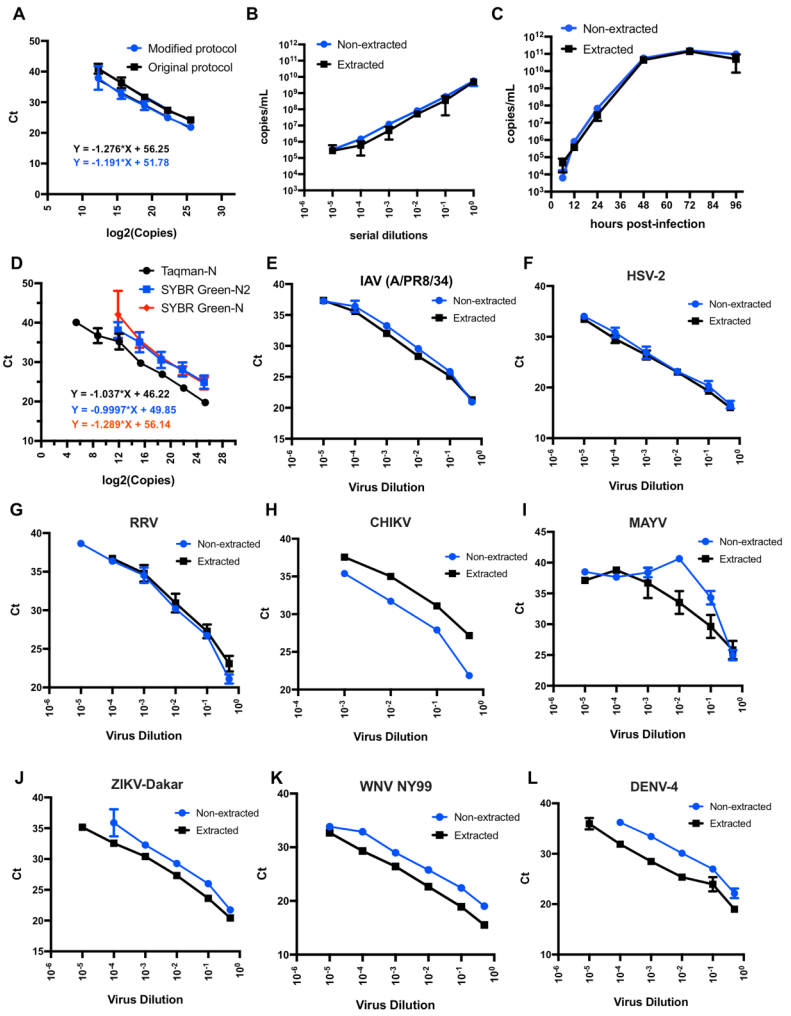

Figure 1. Development of a simplified Q-RT-PCR assay for SARS-CoV-2 viral RNA detection in cell culture supernatants. A. Serially diluted RNA standards were either directly subjected to Q-RT-PCR or processed as in the modified protocol detailed in the text prior to Q-RT-PCR. Log_2 (copies) are plotted against the cycle threshold (Ct) values. Linear regression analysis was done to obtain the equations. Data show the average of three independent biological replicates. Error bars show the SEM. **B.** Comparison of the efficiency and detection ranges for quantifying SARS-CoV-2 RNA using purified RNA or lysed supernatants from virus stocks. Data are derived from three independent replicates. Error bars show the SEM. **C.** Vero E6 cells were infected at an MOI of 0.01 and cell culture supernatants were analyzed for SARS-CoV-2 RNA following the conventional RNA extraction protocol vs. the modified protocol developed herein at various times post infection. Cell-associated viral RNA was analyzed in parallel following RNA extraction for reference. Data are from three independent biological replicates. Error bars show the SEM. **D.** Illustration of the efficiency and detection ranges of Taqman-based and SYBR-Green-based Q-RT-PCR quantifying known amounts of SARS-CoV-2 RNA. Data is from 2-3 independent replicates. Error bars show the SEM. **E-L.** Indicated viruses were subjected to RNA or DNA extraction (extracted) and diluted 10-fold, or used directly following dilution (non-extracted) in the SYBR-Green **E-I**. or Taqman-based (J-L) Q-RT-PCR assay as above. Samples were normalized such that equivalent amount of the original virus stock was added to PCR reactions for extracted and non-extracted samples. Plots show the corresponding cycle threshold (Ct, y-axis) per virus dilution (x-axis). Data are from two independent replicates with error bars showing the SEM.

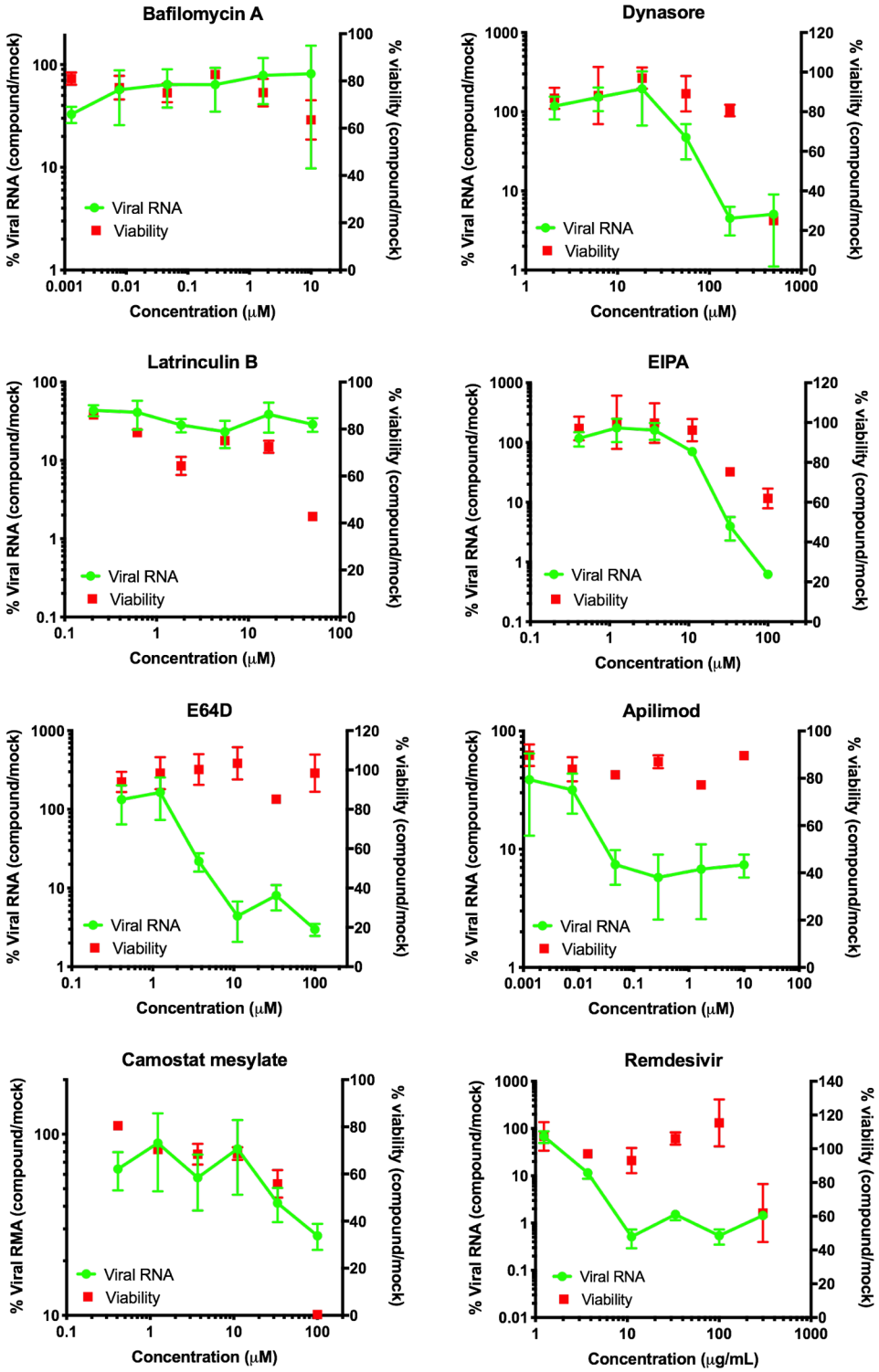

Fig 2. A compound screen to validate SARS-CoV-2-specific inhibitors and entry pathways. Vero E6 cells were infected with SARS-CoV-2 at an MOI of 0.01 and inhibitors were added concomitantly at concentrations shown in the figures following virus adsorption. Supernatants from infected cells were lysed and used in a SYBR-Green based Q-RT PCR to quantify the viral RNA in cell culture supernatants. Compound cytotoxicity was monitored by RealTime-Glo™ MT Cell Viability Assay Kit (Promega) in parallel plates. Data show the cumulative data from 2-5 independent biological replicates. Error bars show the SEM.

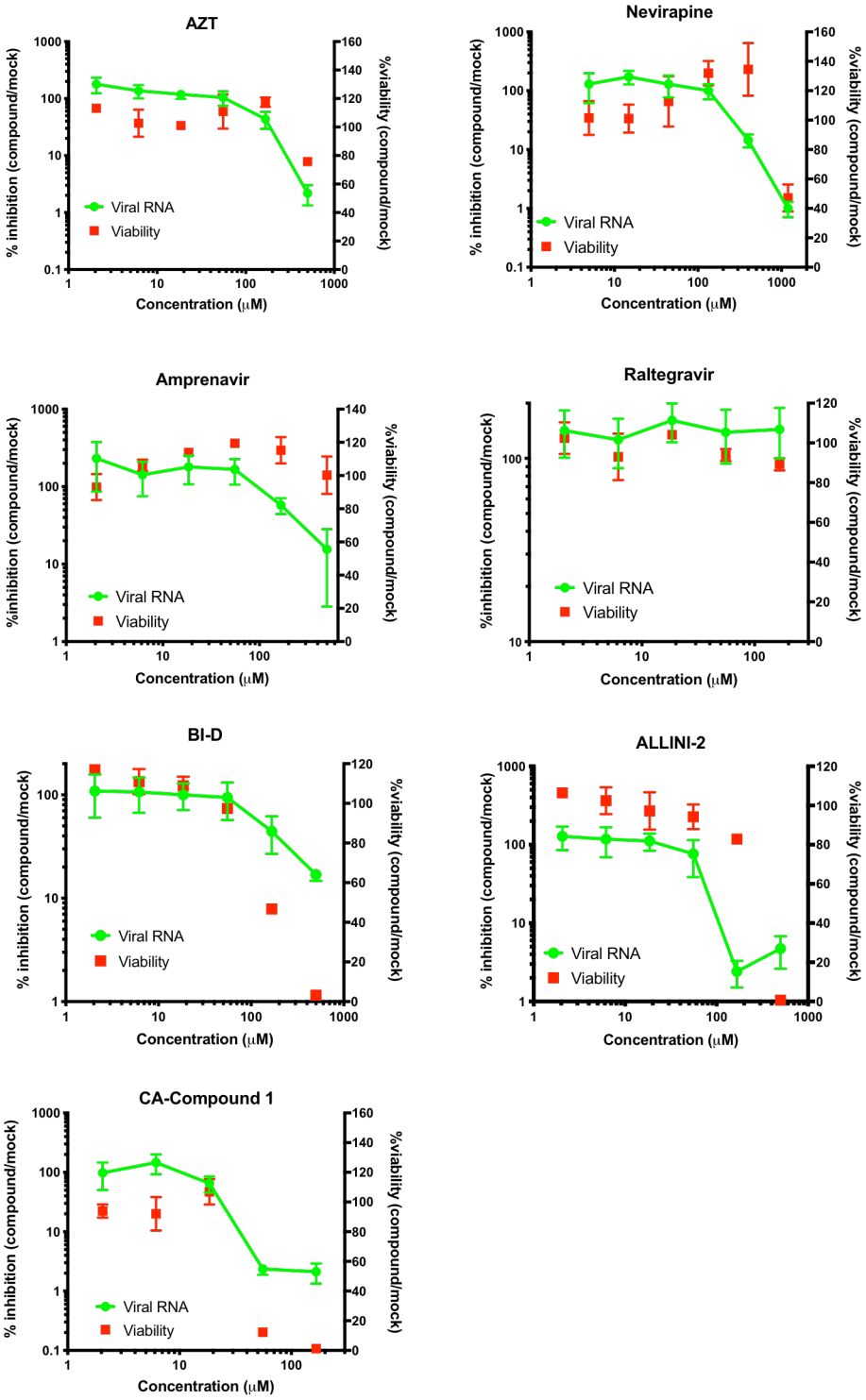

Fig S1. A screen to test the antiviral activities of various HIV-1-specific inhibitors. Vero E6 cells were infected with SARS-CoV-2 at an MOI of 0.01 and inhibitors were added concomitantly at concentrations shown in the figures following virus adsorption. Supernatants from infected cells were lysed and used in a SYBR-Green based Q-RT PCR to quantify the viral RNA in cell culture supernatants. Compound cytotoxicity was monitored by RealTime-Glo™ MT Cell Viability Assay Kit (Promega) in parallel plates. Data show the cumulative data from 2-3 independent biological replicates. Error bars show the SEM.

Table S1. Sequences of the primers and probes used in this study

Virus (Gene)	Primer/Probe Sequence (Forward/Reverse)
SARS-CoV-2 (N)	5'-ATGCTGCAATCGTGCTACAA-3' (F)
SARS-CoV-2 (N)	5'-GACTGCCGCCTCTGCTC-3' (R)
SARS-Cov-2 (N)	5'-FAM/TCAAGGAAC/ZEN/AACATTGCCAA/3IABkFQ/
SARS-CoV-2 (N2)	5'-TTACAAACATTGGCCGCAAA-3' (F)
SARS-CoV-2 (N2)	5'-GCGCGACATTCCGAAGAA-3' (R)
IAV A/PR8/34 (M)	5'-AAGACCAATCCTGTCACCTCTGA-3' (F)
IAV A/PR8/34 (M)	5'-CAAAGCGTCTACGCTGCAGTCC-3' (R)
HSV-2 (ICP27)	5'-TGT CGG AGA TCG ACT ACA CG-3' (F)
HSV-2 (ICP27)	5'-CGGTGCGTGTCCAGTATTTC-3' (R)
CHIKV (E1)	5'-TCGACGCGCCCTCTTTAA-3' (F)
CHIKV (E1)	5'-ATCGAATGCACCGCACAC T-3' (R)
RRV (nsP3)	5'-GTGTTCTCCGGAGGTAAAGATAG-3' (F)
RRV (nsP3)	5'-TCGCGGCAATAGATGACTAC-3' (R)
MAYV (nsp1-3)	5'-AAGCTCTTCCTCTGCATTGC-3' (F)
MAYV (nsp1-3)	5'-TGCTGGAAACGCTCTCTGTA-3' (R)
ZIKV-Dakar (pp)	5'-TTCGGACAGCCGTTGTCCAACACAAG-3' (F)
ZIKV-Dakar (pp)	5'-CCACCAATGTTCTCTTGCAGACATATTG-3' (R)
ZIKV-Dakar (pp)	5'-FAM/AGCCTACCT/ZEN/TGACAAGCAGTC/3IABkFQ/
WNV-NY99 (pp)	5'-TCAGCGATCTCTCCACCAAAG-3' (F)
WNV-NY99 (pp)	5'-GGGTCAGCACGTTTGTCATTG-3' (R)
WNV-NY99 (pp)	5'-FAM/TGCCCGACC/ZEN/ATGGGAGAAGCTC/3IABkFQ/

Table S2. Sources and properties of the compounds used in this study.

Compound name	Source	Function	Reference
Camostat mesylate	Tocris Bioscience	TMPRSS2 inhibitor	(9), (10)
Bafilomycin A	Sigma Aldrich	Inhibits receptor mediated endocytosis	(7)
Apilimod	A gift of Sean Whelan	Inhibits PIKfyve kinase and endosomal trafficking	(6, 7)
E64D	Sigma Aldrich	Inhibits cathepsins B and L	(7)
EIPA	Sigma Aldrich	Inhibits macropinocytosis	(11)
Dynasore	Sigma Aldrich	inhibits dymanin and clathrin-mediated endocytosis	(11)
Latrunculin B	Sigma Aldrich	Inhibits actin polymerization	
Remdesivir	A gift of Gaya Amarasinghe	RNA-dependent RNA polymerase	(5)
Compound 1	Custom synthesis	Inhibits HIV-1 capsid stability	(12, 13)
Nevirapine	NIH AIDS Reagents	HIV-1 non-nucleoside reverse transcriptase inhibitor (NNRTI)	(14)
Azidothymidine (AZT)	NIH AIDS Reagents	HIV-1 nucleoside reverse transcriptase inhibitor (NRTI)	(15, 16)
Raltegravir	NIH AIDS Reagents	Integrase strand transfer inhibitor	(17)
ALLINI-2	Custom synthesis	Allosteric integrase inhibitor	(18)
BI-D	Custom synthesis	Allosteric integrase inhibitor	(19)
Amprenavir	NIH AIDS Reagents	HIV-1 protease inhibitor	(20)

REFERENCES

1. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. 2020. A new coronavirus associated with human respiratory disease in China. Nature 579:265-269.

2. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579:270-273.

3. Chan JF, Yip CC, To KK, Tang TH, Wong SC, Leung KH, Fung AY, Ng AC, Zou Z, Tsoi HW, Choi GK, Tam AR, Cheng VC, Chan KH, Tsang OT, Yuen KY. 2020. Improved Molecular Diagnosis of COVID-19 by the Novel, Highly Sensitive and Specific COVID-19-RdRp/Hel Real-Time Reverse Transcription-PCR Assay Validated In Vitro and with Clinical Specimens. J Clin Microbiol 58.

4. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. 2009. A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. J Virol Methods 156:1-7.

5. Sheahan TP, Sims AC, Leist SR, Schafer A, Won J, Brown AJ, Montgomery SA, Hogg A, Babusis D, Clarke MO, Spahn JE, Bauer L, Sellers S, Porter D, Feng JY, Cihlar T, Jordan R, Denison MR, Baric RS. 2020. Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. Nat Commun 11:222.

6. Kang Y-L, Chou Y-Y, Rothlauf PW, Liu Z, Soh TK, Cureton D, Case JB, Chen RE, Diamond MS, Whelan SPJ, Kirchhausen T. 2020. Inhibition of PIKfyve kinase prevents infection by EBOV and SARS-CoV-2. bioRxiv doi:10.1101/2020.04.21.053058:2020.04.21.053058.

7. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, Guo L, Guo R, Chen T, Hu J, Xiang Z, Mu Z, Chen X, Chen J, Hu K, Jin Q, Wang J, Qian Z. 2020. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. Nat Commun 11:1620.

8. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, Shi Z, Hu Z, Zhong W, Xiao G. 2020. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res 30:269-271.

9. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, Li F. 2020. Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci U S A 117:11727-11734.

10. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Muller MA, Drosten C, Pohlmann S. 2020. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181:271-280 e8.

11. Burkard C, Verheije MH, Wicht O, van Kasteren SI, van Kuppeveld FJ, Haagmans BL, Pelkmans L, Rottier PJ, Bosch BJ, de Haan CA. 2014. Coronavirus cell entry occurs through the endo-/lysosomal pathway in a proteolysis-dependent manner. PLoS Pathog 10:e1004502.

12. Goudreau N, Lemke CT, Faucher AM, Grand-Maitre C, Goulet S, Lacoste JE, Rancourt J, Malenfant E, Mercier JF, Titolo S, Mason SW. 2013. Novel inhibitor binding site discovery on HIV-1 capsid N-terminal domain by NMR and X-ray crystallography. ACS Chem Biol 8:1074-82.

13. Lemke CT, Titolo S, Goudreau N, Faucher AM, Mason SW, Bonneau P. 2013. A novel inhibitor-binding site on the HIV-1 capsid N-terminal domain leads to improved crystallization via compound-mediated dimerization. Acta Crystallogr D Biol Crystallogr 69:1115-23.

14. Merluzzi VJ, Hargrave KD, Labadia M, Grozinger K, Skoog M, Wu JC, Shih CK, Eckner K, Hattox S, Adams J, et al. 1990. Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. Science 250:1411-3.

15. Mitsuya H, Weinhold KJ, Furman PA, St Clair MH, Lehrman SN, Gallo RC, Bolognesi D, Barry DW, Broder S. 1985. 3'-Azido-3'-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. Proc Natl Acad Sci U S A 82:7096-100.

16. Yarchoan R, Klecker RW, Weinhold KJ, Markham PD, Lyerly HK, Durack DT, Gelmann E, Lehrman SN, Blum RM, Barry DW, et al. 1986. Administration of 3'-azido-3'-deoxythymidine, an inhibitor of HTLV-III/LAV replication, to patients with AIDS or AIDS-related complex. Lancet 1:575-80.

17. Cahn P, Sued O. 2007. Raltegravir: a new antiretroviral class for salvage therapy. Lancet 369:1235-1236.

18. Feng L, Sharma A, Slaughter A, Jena N, Koh Y, Shkriabai N, Larue RC, Patel PA, Mitsuya H, Kessl JJ, Engelman A, Fuchs JR, Kvaratskhelia M. 2013. The A128T resistance mutation reveals aberrant protein multimerization as the primary mechanism of action of allosteric HIV-1 integrase inhibitors. J Biol Chem 288:15813-20.

19. Wang H, Jurado KA, Wu X, Shun MC, Li X, Ferris AL, Smith SJ, Patel PA, Fuchs JR, Cherepanov P, Kvaratskhelia M, Hughes SH, Engelman A. 2012. HRP2 determines the

efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. Nucleic Acids Res 40:11518-30.

20. Anderson J, Schiffer C, Lee SK, Swanstrom R. 2009. Viral protease inhibitors. Handb Exp Pharmacol doi:10.1007/978-3-540-79086-0_4:85-110.

Chapter 5: Summary and Future Studies

5.1 Summary

This thesis dissertation presents the findings of several studies on the molecular mechanisms that regulate the binding between integrase and the viral genomic RNA in HIV-1 and other retroviruses. The second chapter of the dissertation elaborates the findings that integrase binding to RNA is mediated by electrostatic interactions. The third chapter provides evidence from CLIP showing that integrase has higher binding affinity to purine-rich sequences of RNA and the binding is conserved in different retrovirus genera. The fourth chapter demonstrates an assay that we have developed at the dawn of the COVID-19 pandemic and its use in screening new antiviral agents against the SARS-Cov-2 virus. Altogether, these studies illustrate the molecular properties underlying integrase-RNA binding and its role in HIV-1 virion morphogenesis. Furthermore, we illustrate how our methods can be used to study different viruses and applied in fighting future pandemics.

HIV-1 integrase binding to the genomic RNA is mediated by electrostatic interactions

The human immunodeficiency virus type 1 (HIV-1) is a retrovirus that has been responsible for millions of deaths over the course of the AIDS pandemic (1, 2). HIV-1 contains two copies of single stranded genomic RNA (gRNA) that surrounded conical capsid lattice as part of the viral ribonucleoprotein complex (vRNP), altogether they form the viral core (1, 3, 4). Similarly to other retroviruses, the gRNA of HIV-1 is reverse transcribed by the viral reverse transcriptase enzyme in infected cells (5). The viral DNA is then integrated into the host chromosome in a reaction catalyzed by the viral enzyme integrase (5). Some of the most efficient anti-HIV-1 therapies target this catalytic role of integrase (6, 7); these IN-targeting compounds, known as

integrase strand-transfer inhibitors (INSTIs), have been essential components of anti-retroviral regimens (7-14).

Over the years, there has been multiple lines of evidence which suggested that IN plays a role in the maturation of viral particle maturation. A group of pleotropic mutations, known as class II IN mutations, causes defects in virion assembly (15-28) and morphogenesis (17, 18, 24-26, 29-31) without affecting the catalytic activity of IN in vitro studies. Furthermore, virions produced from cells treated with allosteric integrase inhibitors (ALLINIs) showed a similar characteristic eccentric morphological defect where the vRNP is mislocalized outside the capsid lattice (29, 30, 32). In a groundbreaking study done by our lab and collaborators, it was shown that IN binds to gRNA in mature virions, blocking IN-gRNA binding by either ALLINIs or class II IN mutations caused formation of eccentric virion particles (31, 33, 34). ALLINIs and most class II mutations blocked IN-gRNA binding by disrupting proper IN multimerization (33, 35, 36).

The R262, R263, R269 and the K273 are directly involved in the binding of IN to gRNA, their mutations (R269A/K273A and R262A/R263A) inhibit IN-gRNA binding without affecting IN multimerization (33). To understand how these residues mediate IN-gRNA, we isolated secondary site suppressor mutations of the R269A/K273A mutation in an evolution-style serial passaging experiment. By acquisition of additional D256N and D270N mutations, the R269A/K273A class II mutant virus regained the ability of IN to bind gRNA and led to formation of mature virion particles. Moreover, other mutations, such as D256R or D256K, that restored the overall positive electrostatic charge of the IN CTD could restore IN-gRNA binding and infectivity for the R269A/K273A and the R262A/R263A class II IN mutants. None of these compensatory mutations affected the functional multimerization of IN, which suggests that they

directly contribute to the binding of IN to gRNA. HIV-1 IN R269A/K273A viruses bearing the compensatory mutations were more sensitive to ALLINIs in comparison to WT viruses, providing the first key genetic evidence that specific IN residues required for RNA binding also influence ALLINI activity. Supplementary structural modeling provides more insights into the molecular dynamics underlying IN-gRNA binding and the ALLINI mechanism of action. Altogether, this project elucidated the role of electrostatic interactions in IN-gRNA binding and how such a property affects virus susceptivity to ALLINIs.

Molecular characterization of IN-RNA binding in HIV-1 and its conservation in retroviruses

Prior studies have used CLIPseq to show that IN binds to genomic RNA in mature virions, and that the binding is essential for the replication of HIV-1 viruses in target cells (31, 33, 37). These studies showed that IN has high binding affinity for structured RNA elements of the HIV-1 genome such TAR (31). Using a modified CLIP approach and *in vitro* studies, we showed that IN also has high affinity for purine-rich sequences, a characteristic feature of the HIV-1 genome. In addition, we showed that IN could bind to packaged cellular RNAs in the absence of the viral genome.

Although, virion morphologies are different from one retrovirus to another, the maturation process is generally conserved (38-42). We have examined whether the IN-RNA binding is conserved in other retroviruses and whether its role in virion maturation is conserved. Our preliminary findings demonstrates that IN-gRNA binding is conserved in retroviruses such as MLV, EIAV, MMTV, and MVV. Moreover, this binding can also be blocked by mutations that are similar to HIV-1 class II mutations. In our future studies, we will examine how IN-

gRNA binding or the lack of it affects the maturation and morphologies of different retroviruses by using electron microscopy, CLIP and other biochemical approaches as previously described (33). Furthermore, we will investigate what gRNA sequences are bound by IN and whether ALLINIs can block IN-gRNA binding in such retroviruses.

Altogether, our studies have improved our understanding of the molecular aspects of IN-gRNA binding and how they can be targeted by the new generation of IN-targeting antiretroviral agents. In addition, we propose a series of experiments that would reveal key details of IN-gRNA complex structure. Finally, we have shown how our tools could be modified and deployed to study and screen new inhibitors of the SARS-CoV-2 virus during the COVID-19 pandemic.

4.2 Future Studies

Temporal assessment of IN-RNA interaction during virion morphogenesis

After transcription and translation of viral RNAs in the host cells, viral structural proteins and enzymes are part of two polyproteins, Gag and Gag-Pol (43-47). Gag and Gag-Pol are assembled and released from the cell as part of immature virions, which later mature into virions where the vRNP is enclosed a capsid lattice (48, 49). Understanding when IN binds to gRNA during this process of virion morphogenesis would help understand how the encapsidation of the vRNP is coordinated by IN and other gRNA-binding proteins such as NC.

To analyze whether IN-gRNA binding is initiated in producer cells or immature virions, we will use an in-vitro PR cleavage assay to separate IN from the rest of Gag-Pol polyprotein. We have previously developed a similar in vitro assay to study the cleavage of the host CARD8 protein by HIV-1 PR (50). Briefly, HEK293T cells will be transfected with proviral plasmids encoding a catalytically inactive protease (PR-). Two days post transfection, cells and released immature virions

will be UV-crosslinked as detailed in the CLIP protocol (51), then lysed and treated with recombinant HIV-1 PR (available from NIH AIDS reagents) to cleave IN from Gag-Pol. The released IN protein will then be immunoprecipitated and the γ-P^{32}-ATP labelled RNA complexes will be analyzed for binding to IN by autoradiography. Further sequencing could reveal the sequences that are bound by IN prior to virion maturation.

Isolation and structural characterization of integrase-RNA complexes

HIV IN has three domains: N-terminal dimerization domain (NTD) that has a conserved HCCH Zn+ binding motif, an RNase H-like catalytic core domain (CCD), and a C-terminal domain that aids in DNA binding (18, 52-55). The IN protein has been virtually impossible to crystallize by standard structural biology techniques due to its low solubility (54, 56-59). Nonetheless, each of the IN domains has been purified, crystallized and characterized individually, in complex with other proteins, or as double domain partial structure; but never in complex with RNA (18, 54, 55, 58). However, crystallization of the full-length HIV-1 IN structure has been elusive and none of the HIV-1 double-domain partial structures has been crystallized together with DNA or RNA (60). The absence of the integrase-viral RNA structure is not just an obstacle for understanding the key role of integrase at virion morphogenesis, but it is also a challenge in the development of new integrase-inhibitors.

We are developing a biochemical method that could isolate full-domain integrase with bound RNA. This assay is based on immunoprecipitation of IN-gRNA complexes; full length IN and bound viral RNA have previously been immunoprecipitated in CLIP assays (31, 33, 37). The production and isolation IN-gRNA complexes is done in three steps: 1) Virion production, 2)

Separation of the viral core on sucrose density gradients, and 3) isolation of IN-gRNA complexes by immunoprecipitation.

1. Virion Production: HIV-1 virions are produced by transfecting 7 million HEK293T with 30μg of HIV-1 pNL4-3 plasmids. Two days after transfection, virions are concentrated by ultracentrifugation at 28000rpm for 90 minutes on an SW32Ti rotor. The virions can then be resuspended and lysed in 500uL of 1XPBS containing 0.5% Triton-X.
2. Separation of virion components by sucrose density sedimentation: In mature HIV-1 virions, IN-gRNA complexes are surrounded by the capsid lattice as part of the vRNP and viral core (61, 62). Being surrounded by the capsid lattice, components of the viral core (RT, MA, CA, and IN) move together to denser fractions during density sedimentation (33, 63, 64). Some capsid mutations (K203A and P38A) destabilize the capsid and lead to the mislocalization of IN-gRNA complexes and other viral core outside of the capsid lattice (63-65). During density sedimentation, the CA, MA, and RT of such mutant viruses migrated separately from IN-gRNA complexes (Figure 1A,B).
3. Isolation of IN-gRNA complexes by immunoprecipitation: Following density sedimentation, the fractions containing IN-RNA complexes (7-10) are combined (Figure 1C) and Triton is added to a final concentration of 0.5% to aid in immunoprecipitation. Furthermore, RNase A can be added to fragment RNA that is attached to the integrase. IN complexes are then immunoprecipitated using 40μL of Protein G Dynabeads® (Life Technologies) conjugated with a monoclonal anti-integrase mouse antibody. Following immunoprecipitation, the beads are washed three times with 1XSTE (100 mM NaCl, 10 mM Tris-Cl (pH 8.0) to remove any contaminating proteins in complex with IN.

Immunoblotting (Figure 1C) shows that a vast majority of integrase was immunoprecipitated using the above approach, proving the efficiency of immunoprecipitation technique.

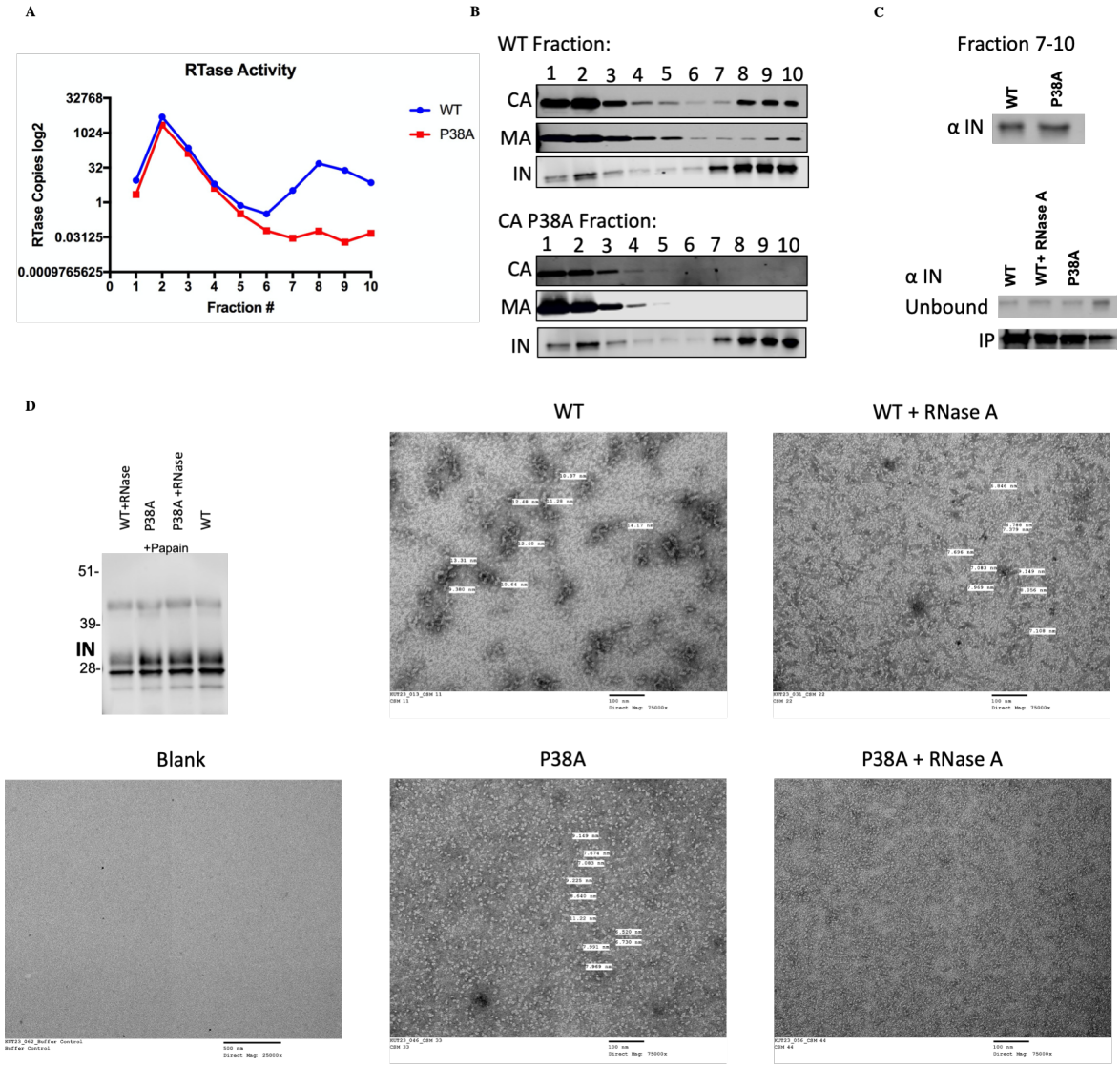

Figure 1. Isolation of IN-RNA complexes. Components of envelope-stripped virions were separated by density segmentation on a linear sucrose gradient. **A, B.** Migration of Capsid, Matrix, and RT as seen in the fraction1-10 (top-bottom). **C.** Collective IN in fraction 7-10 (top), and immunoprecipitated integrase from fractions 7-10 compared to integrase left in the IP solution (Unbound)(bottom). **D.** Immunoprecipitated integrase was separated from the IP antibody using immobilized Papain (top left) and Negative staining images of integrase complexes eluted from papain beads.

To serve as specimen for cryo-EM, the majority of the immunoprecipitated integrase had to be removed from the immunoprecipitated beads. Immobilized papain© (Thermo Fisher Scientific Inc 20341) fragments apart the Fc and Fab regions of most IgG antibodies and it was used to release IN from the IP antibody. Papain was activated in a reaction buffer containing 20mM sodium phosphate, 10mM EDTA, 20mM cysteine.HCl; pH 7.0. Cleavage by papain was done in a thermomixer at 37^0C 800rpm overnight. Papain cleaved the majority of integrase away from the antibody (Figure 1D). Diagnostic negative staining was done for initial inquiries on whether the purified complexes were the size of integrase or met the quantity. Negative staining showed that there was a good yield of IN complexes (Figure 1D). RNAseA did not seem to alter the size of the complexes (Fig. 1D); however, P38A particles generated slightly smaller and more uniform complexes, which are more suitable for mass spectrometry and Cryo EM (Figure 1D). Future experiments will make a definitive confirmation on whether the complexes observed by negative staining were integrase complexes by using mass spectrometry. The structure of these complexes will then be studied by small-particle cryo-EM using facilities at the Washington University School of Medicine.

References

1. Chen J, Nikolaitchik O, Singh J, Wright A, Bencsics CE, Coffin JM, Ni N, Lockett S, Pathak VK, Hu WS. 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. Proc Natl Acad Sci U S A 106:13535-40.

2. Chen J, Rahman SA, Nikolaitchik OA, Grunwald D, Sardo L, Burdick RC, Plisov S, Liang E, Tai S, Pathak VK, Hu WS. 2016. HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. Proc Natl Acad Sci U S A 113:E201-8.

3. Darlix JL, de Rocquigny H, Mauffret O, Mely Y. 2014. Retrospective on the all-in-one retroviral nucleocapsid protein. Virus Res 193:2-15.

4. Perilla JR, Gronenborn AM. 2016. Molecular Architecture of the Retroviral Capsid. Trends Biochem Sci 41:410-420.

5. Engelman AN. 2019. Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. J Biol Chem doi:10.1074/jbc.REV119.006901.

6. Hazuda DJ, Felock P, Witmer M, Wolfe A, Stillmock K, Grobler JA, Espeseth A, Gabryelski L, Schleif W, Blau C, Miller MD. 2000. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science 287:646-50.

7. Brooks KM, Sherman EM, Egelund EF, Brotherton A, Durham S, Badowski ME, Cluck DB. 2019. Integrase Inhibitors: After 10 Years of Experience, Is the Best Yet to Come? Pharmacotherapy 39:576-598.

8. Ramanathan S, Mathias AA, German P, Kearney BP. 2011. Clinical pharmacokinetic and pharmacodynamic profile of the HIV integrase inhibitor elvitegravir. Clin Pharmacokinet 50:229-44.

9. Summa V, Petrocchi A, Bonelli F, Crescenzi B, Donghi M, Ferrara M, Fiore F, Gardelli C, Gonzalez Paz O, Hazuda DJ, Jones P, Kinzel O, Laufer R, Monteagudo E, Muraglia E, Nizi E, Orvieto F, Pace P, Pescatore G, Scarpelli R, Stillmock K, Witmer MV, Rowley M. 2008. Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. J Med Chem 51:5843-55.

10. Tsiang M, Jones GS, Goldsmith J, Mulato A, Hansen D, Kan E, Tsai L, Bam RA, Stepan G, Stray KM, Niedziela-Majka A, Yant SR, Yu H, Kukolj G, Cihlar T, Lazerwith SE, White KL, Jin H. 2016. Antiviral Activity of Bictegravir (GS-9883), a Novel Potent HIV-1 Integrase Strand Transfer Inhibitor with an Improved Resistance Profile. Antimicrob Agents Chemother 60:7086-7097.

11. Min S, Song I, Borland J, Chen S, Lou Y, Fujiwara T, Piscitelli SC. 2010. Pharmacokinetics and safety of S/GSK1349572, a next-generation HIV integrase inhibitor, in healthy volunteers. Antimicrob Agents Chemother 54:254-8.

12. Rockstroh JK, DeJesus E, Lennox JL, Yazdanpanah Y, Saag MS, Wan H, Rodgers AJ, Walker ML, Miller M, DiNubile MJ, Nguyen BY, Teppler H, Leavitt R, Sklar P, Investigators S. 2013. Durable efficacy and safety of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naive HIV-1-infected patients: final 5-year results from STARTMRK. J Acquir Immune Defic Syndr 63:77-85.

13. Taha H, Das A, Das S. 2015. Clinical effectiveness of dolutegravir in the treatment of HIV/AIDS. Infect Drug Resist 8:339-52.

14. Smith SJ, Zhao XZ, Passos DO, Lyumkis D, Burke TR, Hughes SH. 2020. HIV-1 Integrase Inhibitors that are active against Drug-Resistant Integrase Mutants. Antimicrob Agents Chemother doi:10.1128/AAC.00611-20.

15. Ansari-Lari MA, Donehower LA, Gibbs RA. 1995. Analysis of human immunodeficiency virus type 1 integrase mutants. Virology 213:680.

16. Bukovsky A, Gottlinger H. 1996. Lack of integrase can markedly affect human immunodeficiency virus type 1 particle production in the presence of an active viral protease. J Virol 70:6820-5.

17. 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. J Virol 69:2729-36.

18. Jenkins TM, Engelman A, Ghirlando R, Craigie R. 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J Biol Chem 271:7712-8.

19. Kalpana GV, Reicin A, Cheng GS, Sorin M, Paik S, Goff SP. 1999. Isolation and characterization of an oligomerization-negative mutant of HIV-1 integrase. Virology 259:274-85.

20. 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. J Virol 70:721-8.

21. Liao WH, Wang CT. 2004. Characterization of human immunodeficiency virus type 1 Pr160 gag-pol mutants with truncations downstream of the protease domain. Virology 329:180-8.

22. Lu R, Ghory HZ, Engelman A. 2005. Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. J Virol 79:10356-68.

23. Lu R, Limon A, Devroe E, Silver PA, Cherepanov P, Engelman A. 2004. Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. J Virol 78:12735-46.

24. Nakamura T, Masuda T, Goto T, Sano K, Nakai M, Harada S. 1997. Lack of infectivity of HIV-1 integrase zinc finger-like domain mutant with morphologically normal maturation. Biochem Biophys Res Commun 239:715-22.

25. 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.

26. Shin CG, Taddeo B, Haseltine WA, Farnet CM. 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. J Virol 68:1633-42.

27. Taddeo B, Haseltine WA, Farnet CM. 1994. Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. J Virol 68:8401-5.

28. 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. J Virol 73:2126-35.

29. Fontana J, Jurado KA, Cheng N, Ly NL, Fuchs JR, Gorelick RJ, Engelman AN, Steven AC. 2015. Distribution and Redistribution of HIV-1 Nucleocapsid Protein in Immature, Mature, and Integrase-Inhibited Virions: a Role for Integrase in Maturation. J Virol 89:9765-80.

30. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, Wang W, Ballandras-Colas A, Patel PA, Fuchs JR, Kvaratskhelia M, Engelman A. 2013. Allosteric integrase inhibitor potency is determined through the inhibition of HIV-1 particle maturation. Proc Natl Acad Sci U S A 110:8690-5.

31. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268 e12.

32. Gupta K, Brady T, Dyer BM, Malani N, Hwang Y, Male F, Nolte RT, Wang L, Velthuisen E, Jeffrey J, Van Duyne GD, Bushman FD. 2014. Allosteric inhibition of human immunodeficiency virus integrase: late block during viral replication and abnormal multimerization involving specific protein domains. J Biol Chem 289:20477-88.

33. Elliott JL, Eschbach JE, Koneru PC, Li W, Puray-Chavez M, Townsend D, Lawson DQ, Engelman AN, Kvaratskhelia M, Kutluay SB. 2020. Integrase-RNA interactions underscore the critical role of integrase in HIV-1 virion morphogenesis. Elife 9.

34. Elliott JL, Kutluay SB. 2020. Going beyond Integration: The Emerging Role of HIV-1 Integrase in Virion Morphogenesis. Viruses 12.

35. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, de Silva S, Wu L, Le Grice SF, Engelman A, Fuchs JR, Kvaratskhelia M. 2012. Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. J Biol Chem 287:16801-11.

36. Tsiang M, Jones GS, Niedziela-Majka A, Kan E, Lansdon EB, Huang W, Hung M, Samuel D, Novikov N, Xu Y, Mitchell M, Guo H, Babaoglu K, Liu X, Geleziunas R, Sakowicz R. 2012. New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. J Biol Chem 287:21189-203.

37. Madison MK, Lawson DQ, Elliott J, Ozanturk AN, Koneru PC, Townsend D, Errando M, Kvaratskhelia M, Kutluay SB. 2017. Allosteric HIV-1 Integrase Inhibitors Lead to Premature Degradation of the Viral RNA Genome and Integrase in Target Cells. J Virol 91.

38. Zhang W, Cao S, Martin JL, Mueller JD, Mansky LM. 2015. Morphology and ultrastructure of retrovirus particles. AIMS Biophys 2:343-369.

39. Briggs JA, Wilk T, Welker R, Krausslich HG, Fuller SD. 2003. Structural organization of authentic, mature HIV-1 virions and cores. EMBO J 22:1707-15.

40. Ganser BK, Li S, Klishko VY, Finch JT, Sundquist WI. 1999. Assembly and analysis of conical models for the HIV-1 core. Science 283:80-3.

41. Kingston RL, Olson NH, Vogt VM. 2001. The organization of mature Rous sarcoma virus as studied by cryoelectron microscopy. J Struct Biol 136:67-80.

42. Yeager M, Wilson-Kubalek EM, Weiner SG, Brown PO, Rein A. 1998. Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. Proc Natl Acad Sci U S A 95:7299-304.

43. Harris ME, Hope TJ. 2000. RNA export: insights from viral models. Essays Biochem 36:115-27.

44. Cullen BR. 2003. Nuclear mRNA export: insights from virology. Trends Biochem Sci 28:419-24.

45. Wodrich H, Kräusslich HG. 2001. Nucleocytoplasmic RNA transport in retroviral replication. Results Probl Cell Differ 34:197-217.

46. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331:280-3.

47. Wilson W, Braddock M, Adams SE, Rathjen PD, Kingsman SM, Kingsman AJ. 1988. HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. Cell 55:1159-69.

48. Sundquist WI, Krausslich HG. 2012. HIV-1 assembly, budding, and maturation. Cold Spring Harb Perspect Med 2:a006924.

49. Freed EO. 2015. HIV-1 assembly, release and maturation. Nat Rev Microbiol 13:484-96.

50. Wang Q, Gao H, Clark KM, Mugisha CS, Davis K, Tang JP, Harlan GH, DeSelm CJ, Presti RM, Kutluay SB, Shan L. 2021. CARD8 is an inflammasome sensor for HIV-1 protease activity. Science 371.

51. Shema Mugisha C, Tenneti K, Kutluay SB. 2020. Clip for studying protein-RNA interactions that regulate virus replication. Methods 183:84-92.

52. Lesbats P, Engelman AN, Cherepanov P. 2016. Retroviral DNA Integration. Chem Rev 116:12730-12757.

53. Engelman AN. 2019. Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. J Biol Chem 294:15137-15157.

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

55. Wang JY, Ling H, Yang W, Craigie R. 2001. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J 20:7333-43.

56. 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. Proc Natl Acad Sci U S A 97:8233-8.

57. Li M, Jurado KA, Lin S, Engelman A, Craigie R. 2014. Engineered hyperactive integrase for concerted HIV-1 DNA integration. PLoS One 9:e105078.

58. Li X, Krishnan L, Cherepanov P, Engelman A. 2011. Structural biology of retroviral DNA integration. Virology 411:194-205.

59. Passos DO, Li M, Yang R, Rebensburg SV, Ghirlando R, Jeon Y, Shkriabai N, Kvaratskhelia M, Craigie R, Lyumkis D. 2017. Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome. Science 355:89-92.

60. Quashie PK, Han YS, Hassounah S, Mesplede T, Wainberg MA. 2015. Structural Studies of the HIV-1 Integrase Protein: Compound Screening and Characterization of a DNA-Binding Inhibitor. PLoS One 10:e0128310.

61. Mattei S, Schur FK, Briggs JA. 2016. Retrovirus maturation-an extraordinary structural transformation. Curr Opin Virol 18:27-35.

62. Kessl JJ, Kutluay SB, Townsend D, Rebensburg S, Slaughter A, Larue RC, Shkriabai N, Bakouche N, Fuchs JR, Bieniasz PD, Kvaratskhelia M. 2016. HIV-1 Integrase Binds the Viral RNA Genome and Is Essential during Virion Morphogenesis. Cell 166:1257-1268.e12.

63. Eschbach JE, Elliott JL, Li W, Zadrozny KK, Davis K, Mohammed SJ, Lawson DQ, Pornillos O, Engelman AN, Kutluay SB. 2020. Capsid Lattice Destabilization Leads to Premature Loss of the Viral Genome and Integrase Enzyme during HIV-1 Infection. J Virol 95.

64. Forshey BM, von Schwedler U, Sundquist WI, Aiken C. 2002. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. J Virol 76:5667-77.

65. von Schwedler UK, Stray KM, Garrus JE, Sundquist WI. 2003. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J Virol 77:5439-50.