Emergence of Compensatory Mutations Reveals the Importance of Electrostatic Interactions between HIV-1 Integrase and Genomic RNA

Christian Shema Mugisha,[a] Tung Dinh,[b] Abhishek Kumar,[a] Kasyap Tenneti,[a] Jenna E. Eschbach,[a] Keanu Davis,[a] Robert Gifford,[c] Mamuka Kvaratskhelia,[b] Sebla B. Kutluay[a]

[a]Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, Missouri, USA
[b]Division of Infectious Diseases, University of Colorado School of Medicine, Aurora, Colorado, USA
[c]MRC-University of Glasgow Centre for Virus Research, Bearsden, Glasgow, United Kingdom

ABSTRACT HIV-1 integrase (IN) has a noncatalytic function in virion maturation through its binding to the viral RNA genome (gRNA). Class II IN substitutions inhibit IN-gRNA binding and result in the formation of virions with aberrant morphologies marked by mislocalization of the gRNA between the capsid lattice and the lipid envelope. These viruses are noninfectious due to a block at an early reverse transcription stage in target cells. HIV-1 IN utilizes basic residues within its C-terminal domain (CTD) to bind to the gRNA; however, the molecular nature of how these residues mediate gRNA binding and whether other regions of IN are involved remain unknown. To address this, we have isolated compensatory substitutions in the background of a class II IN mutant virus bearing R269A/K273A substitutions within the IN-CTD. We found that the nearby D256N and D270N compensatory substitutions restored the ability of IN to bind gRNA and led to the formation of mature infectious virions. Reinstating the local positive charge of the IN-CTD through individual D256R, D256K, D278R, and D279R substitutions was sufficient to specifically restore IN-gRNA binding and reverse transcription for the IN R269A/K273A as well as the IN R262A/R263A class II mutants. Structural modeling suggested that compensatory substitutions in the D256 residue created an additional interaction interface for gRNA binding, whereas other substitutions acted locally within the unstructured C-terminal tail of IN. Taken together, our findings highlight the essential role of CTD in gRNA binding and reveal the importance of pliable electrostatic interactions between the IN-CTD and the gRNA.

IMPORTANCE In addition to its catalytic function, HIV-1 integrase (IN) binds to the viral RNA genome (gRNA) through positively charged residues (i.e., R262, R263, R269, K273) within its C-terminal domain (CTD) and regulates proper virion maturation. Mutation of these residues results in the formation of morphologically aberrant viruses blocked at an early reverse transcription stage in cells. Here we show that compensatory substitutions in nearby negatively charged aspartic acid residues (i.e., D256N, D270N) restore the ability of IN to bind gRNA for these mutant viruses and result in the formation of accurately matured infectious virions. Similarly, individual charge reversal substitutions at D256 as well as other nearby positions (i.e., D278, D279) are all sufficient to enable the respective IN mutants to bind gRNA, and subsequently restore reverse transcription and virion infectivity. Taken together, our findings reveal the importance of highly pliable electrostatic interactions in IN-gRNA binding.

KEYWORDS HIV-1, integrase, protein-RNA interactions, virion maturation, virology

A defining feature of retroviruses is the reverse transcription of the viral RNA genome (gRNA) and integration of the resultant linear viral DNA into the host chromosome, which establishes lifelong infection. The latter process is mediated by the viral integrase

1

(IN) enzyme, which catalyzes 3′ processing and DNA strand transfer reactions (1). The catalytic activity of HIV-1 IN has been successfully targeted by several integrase strand-transfer inhibitors (INSTIs) (2–7) that have become key components of frontline antiretroviral therapy regimens due to their high efficacy and tolerance profiles (8–11). In addition, HIV-1 IN has a noncatalytic function in virus replication (12–16). 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 the C-terminal domain (CTD), which adopts an SH3 fold (17, 18). Integration is facilitated by a cellular cofactor, lens epithelium-derived growth factor (LEDGF/p75), which binds tightly to a site within the CCD dimer interface (19, 20) and guides the pre-integration complex to actively transcribed regions of the host chromosome (19–24). A group of pleotropic IN mutations distributed throughout IN, collectively known as class II mutations, disrupt viral assembly (13, 16, 25–36), morphogenesis (12, 16, 27, 32–34, 37–39), and reverse transcription in target cells (12–14, 16, 29, 31, 32, 34, 36–55), often without obstructing the catalytic activity of IN *in vitro* (13, 27, 28, 31, 41, 42, 45, 47, 56–58). 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 that is commonly referred to as the eccentric morphology (12, 15, 16, 27, 32–34, 37, 38, 59–61). Although originally designed to inhibit integration through preventing the binding of IN to LEDGF/p75 (59, 62–66), allosteric integrase inhibitors (ALLINIs) potently inhibit proper particle maturation (12, 37, 62) and lead to the formation of virions that display a similar eccentric morphology observed with class II IN mutations (12).

We have previously shown that IN binds to the gRNA in mature virions to mediate proper encapsidation of the viral ribonucleoproteins (vRNPs) inside the mature CA lattice (38). IN binds with high affinity to the TAR hairpin present within the 5′ and 3′UTRs, though multiple other distinct locations on the gRNA without apparent secondary structure are also bound (38). A parameter that appears to be critical for gRNA binding is functional tetramerization of IN. Numerous class II IN mutations scattered throughout IN block IN-gRNA binding indirectly by disrupting functional oligomerization of IN (15). Similar to these class II IN mutations, ALLINIs are thought to interfere with IN-gRNA binding indirectly by inducing aberrant IN multimerization (12, 38). ALLINIs bind to a pocket within the CCD of IN (12, 61, 62, 65, 67–69) and induce the formation of open IN polymers by engaging the CTD of a nearby IN dimer (63). In contrast, mutation of basic residues within IN-CTD (i.e., R262, R263, R269, and K273) inhibits IN-gRNA binding without altering IN oligomerization in virions and *in vitro* (15, 38), suggesting that these residues directly mediate IN binding to the gRNA.

While the involvement of IN-CTD in gRNA binding has been established, how the basic residues within IN-CTD mediate recognition of 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 nonspecific or semispecific manner, depending on the folding and structure of the cognate RNA element, driven by electrostatic interactions (70–72). Alternatively, these residues can mediate specific interactions with gRNA through H-bonding and van der Waals contacts with individual nucleobases (70–72). Though the structural details of IN-RNA complexes remain unknown, a recent modeling study suggested that a combination of electrostatic interactions with the phosphate backbone and specific interactions of hydrophobic residues with the bases may mediate specific recognition of the TAR element by IN (73). Furthermore, it is possible that other domains in IN also contribute to gRNA binding. For example, we have previously shown that K34A substitution within the NTD blocks gRNA binding without impacting IN oligomerization (15), suggesting its direct interaction with the gRNA.

To gain insight into the mode of IN-gRNA interactions, the noninfectious HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus was serially passaged in T-cells until the acquisition of compensatory mutations. Two compensatory substitutions, D256N and D270N, sequentially emerged within the IN-CTD and allowed HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ to replicate with WT kinetics. Introduction of these mutations on the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ virus backbone substantially enhanced IN-gRNA binding, reverse transcription, and virion infectivity, and restored proper virion morphogenesis. As the d-to-N substitutions resulted in loss of two negative charges, possibly overcoming the loss of two positive charges with the R269A/K273A substitutions, we tested whether restoring the overall local charge of IN-CTD would also restore IN-gRNA binding and virion infectivity. Indeed, the D256R as well as other nearby charge reversal substitutions including D278R and D279R enhanced or fully restored RNA binding, reverse transcription, and virion infectivity for the IN R269A/K273A mutant. We further extended these findings to another class II IN mutant, R262A/R263A, which was similarly suppressed by the D256N/K/R substitutions. Compensatory substitutions 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. Together, our findings highlight the essential role of IN-CTD in gRNA binding and reveal a pliable electrostatic component of IN-gRNA interactions.

RESULTS

Compensatory IN D256N/D270N substitutions emerge in the background of the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus. The IN R269A/K273A class II substitutions obstruct IN-gRNA binding directly without interfering with IN multimerization and result in the formation of particles with eccentric morphology (15, 16). To gain insight into the molecular basis of how R269 and K273 residues mediate IN-gRNA binding, HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus was serially passaged in MT-4 T-cells until the emergence of compensatory mutations at the end of passage 3, which resulted in virus growth with wild-type (WT) replication kinetics (Fig. 1A). Deep sequencing of full-length gRNA isolated from virions across the three passages revealed that the IN R269A and K273A substitutions were retained while the nearby D256N and D270N mutations within IN-CTD were sequentially acquired at the end of passage 1 and passage 2, respectively (Fig. 1B). Viruses bearing both mutations dominated cultures by the end of passage 2, and no other nonsynonymous mutations were observed in IN or elsewhere on the gRNA by the end of passage 3.

The IN D256N and D270N mutations were introduced into the replication-competent pNL4-3 HIV-1 molecular clone bearing WT IN or IN R269A/K273A. These substitutions in IN had no observable effect on Gag (Pr55) and Gag-Pol expression or processing in cells, particle release and virion-associated IN levels (Fig. 1C, Fig. S1A-F). Introduction of D256N and D270N mutations either individually or together (D2N) into HIV-1$_{NL4-3}$ bearing WT IN did not affect viral titers (Fig. 1D; Fig. S1G in the supplemental material). Remarkably, while the individual D256N and D270N mutations introduced into the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus increased virus titers by approximately 4-fold, the D2N substitutions increased virus titers by 30-fold (Fig. 1D; Fig. S1G). Despite no observable differences in multicycle replication kinetics (Fig. 1A), HIV-1$_{NL4-3\ IN\ (R269A/K273A/D2N)}$ virus appeared ~4-fold less infectious than WT HIV-1 in single cycle replication assays (Fig. 1D; Fig, S1G), likely due to an impact on the catalytic activity of IN as detailed below. Overall, these results demonstrate that the combination of D256N and D270N mutations is sufficient to substantially increase the replication competency of the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus.

IN D256N/D270N substitutions restore IN-gRNA binding and accurate virion maturation for the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ virus. We next assessed whether the IN D256N and IN D270N substitutions rendered the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ virus replication competent by restoring IN-gRNA binding and proper virion maturation. To this end, IN-gRNA complexes were immunoprecipitated from UV-cross-linked virions and visualized per CLIP protocol as described previously (38, 74, 75). IN-RNA complexes were readily visible for viruses bearing WT IN, and the D256N, D270N, and D2N substitutions
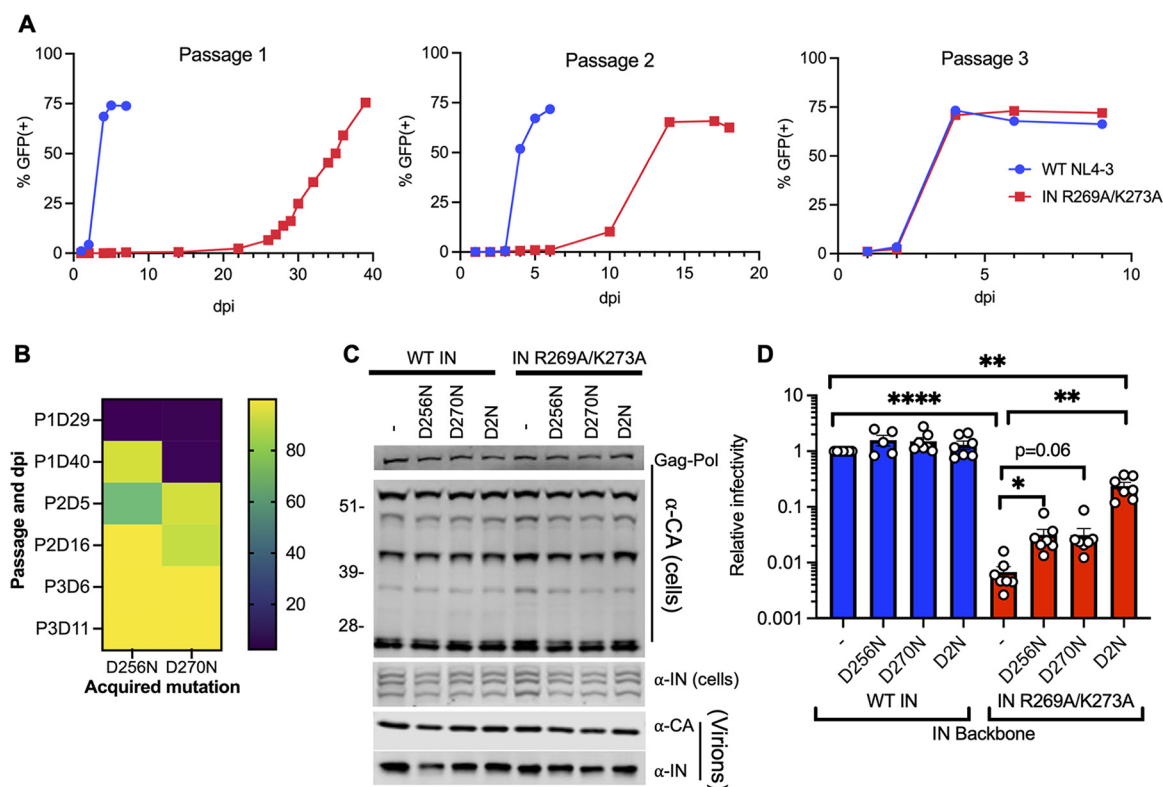
FIG 1 D256N and D270N compensatory substitutions in HIV-1 IN suppress the replication defect of the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II mutant virus. (A) MT4-LTR-GFP indicator cells were infected with HIV-1$_{NL4-3}$ at an MOI of 2 i.u./cell or an equivalent particle number of the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus (based on reverse transcriptase [RT] activity) as detailed in Materials and Methods. HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ viruses were serially passaged for three rounds until the emergence of compensatory mutations that allowed virus replication with WT kinetics. The graphs represent the percentage of GFP positive cells as assessed by flow cytometry over three passages at the indicated days postinfection (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 postinfection (i.e., D29, D40, etc.) as described in Materials and Methods. Heatmap shows the percentage of IN D256N and D270N substitutions at the indicated passages and days postinfection (dpi) as assessed by whole-genome deep sequencing. (C) HEK293T cells were transfected with full-length pNL4-3 expression plasmids carrying *pol* mutations coding for the IN D256N, D270N, and D256N/D270N (D2N) substitutions introduced on the WT IN and IN R269A/K273A backbones. Cell lysates and cell culture supernatants containing virions were collected 2 days posttransfection and analyzed by immunoblotting for CA and IN. The image is representative of five independent experiments. See Fig. S1 for quantitative analysis of immunoblots. (D) HEK293T cells were transfected as in panel C, and cell culture supernatants containing viruses were titered on TZM-bl indicator cells, whereby virus replication was limited to a single cycle by addition of dextran sulfate (50 μg/mL). The titers were normalized relative to particle numbers as assessed by an RT activity assay and are presented relative to WT (set to 1). Also see Fig. S1 for titer and RT activity values prior to normalization. The columns represent the average of 5–6 independent biological replicates, and the error bars represent standard error of the mean (SEM) (****, $P < 0.0001$; **, $P < 0.01$; *, $P < 0.05$, by one-way ANOVA multiple comparison test with Dunnett's correction).

introduced on the WT IN backbone did not impact RNA binding (Fig. 2A and Fig. S2). Expectedly (15, 38), the IN R269A/K273A class II mutant had substantially lower levels of IN-gRNA complexes isolated from virions (Fig. 2A; Fig. S2). The D256N substitution modestly enhanced the ability of IN R269A/K273A to bind RNA, whereas the D270N substitution had no observable impact (Fig. 2A; Fig. S2). In contrast, the D2N substitution substantially enhanced the ability of IN R269A/K273A to bind RNA (Fig. 2A; Fig. S2) and resulted in a significant increase in the accumulation of early and late reverse transcription (RT) products (Fig. S3A, B). 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 IN R269A/K273A class II mutant virions (~68%) had a clear eccentric morphology (Fig. 2B). Consistent with effects on virus titers and RNA-binding, the introduction of D2N substitutions largely restored the ability of the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ virus to form properly matured virions (Fig. 2B). Cumulatively, these data show that D256N and D270N substitutions restore infectivity for the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus through reestablishing gRNA binding and
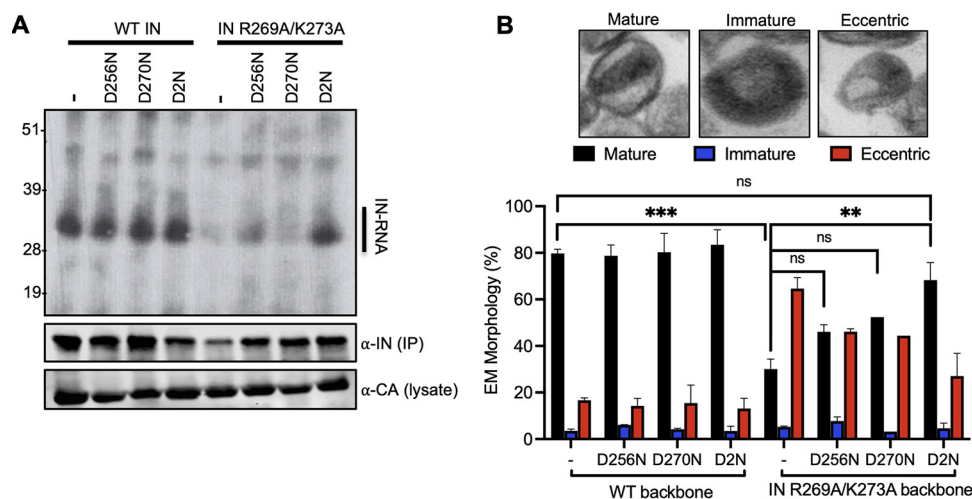
FIG 2 D256N and D270N substitutions restore IN-gRNA binding and accurate virion maturation for the HIV-1$_\text{NL4-3 IN (R269A/K273A)}$ class II mutant virus. (A) Autoradiogram of IN-RNA adducts immunoprecipitated from virions bearing the indicated substitutions in IN. Immunoblots below show immunoprecipitated (IP) IN or CA protein in lysates prior to immunoprecipitation. Data shown are representative of three independent experiments. See also Fig. S2 for quantitative analysis of autoradiographs. (B) Examination of virion maturation in WT and IN mutant viruses by thin section electron microscopy (TEM). Data show quantification of virion morphologies across 100 particles for each sample and replicate experiment. Data show the average of two independent biological replicates, and error bars represent the SEM (***, $P < 0.001$; **, $P < 0.01$; ns = not significant, by one-way ANOVA multiple comparison test with Dunnett's correction).

thereby enabling proper virion maturation and subsequently reverse transcription in target cells.

A single charge reversal substitution alone enhances IN-gRNA binding and virion infectivity for the HIV-1$_\text{NL4-3 IN (R269A/K273A)}$ virus. IN-CTD is decorated with several acidic and basic amino acids, and a long stretch of basic residues spanning 258–273 residues is notable (Fig. 3A). In effect, the D2N substitutions restored the overall local charge of this region, suggesting that electrostatic interactions 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 suppress the class II phenotype observed with HIV-1$_\text{NL4-3 IN (R269A/K273A)}$. We focused our initial analysis on D256 and D270 residues for the following reasons: (i) these amino acids were amenable to substitutions during virus passaging experiments; and (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-gRNA interactions and virion maturation. We introduced D256R, D270R, and D256R/D270R (D2R) substitutions into the HIV-1$_\text{NL4-3 IN R269A/K273A}$ backbone and transfected HEK293T cells with the resulting plasmids. Cell lysates and cell-free virions were then analyzed for Gag and Gag-Pol processing, particle release, and infectivity. Overall, d-to-R substitutions had no major effect on Gag or Gag-Pol expression and processing in cells, or virion release and virion-associated IN levels (Fig. 3B; Fig. S4A–F).

Remarkably, the D256R substitution alone increased the titers of HIV-1$_\text{NL4-3 IN (R269A/K273A)}$ by 26-fold to a level comparable with the D2N substitution (Fig. 3C; Fig. S4G). In contrast, the D270R substitution did not increase virus titers and the D2R substitution increased the infectivity of HIV-1$_\text{NL4-3 IN (R269A/K273A)}$ relatively modestly (~13-fold) compared to D256R and D2N substitutions (Fig. 3C, Fig. S4G). Both the D256R and the D2R substitutions restored IN-gRNA binding to WT levels (Fig. 3D; Fig. S4H) and increased the accumulation of early and late RT products (Fig. S3A, B). These findings demonstrate that the increase in viral titers with the D256R and D2R substitutions correlates well with enhancement of IN-gRNA binding and subsequently reverse transcription.

To determine whether the D256N/R and D270N/R substitutions affected the catalytic activity of IN, we conducted a transcomplementation assay that relies on incorporation
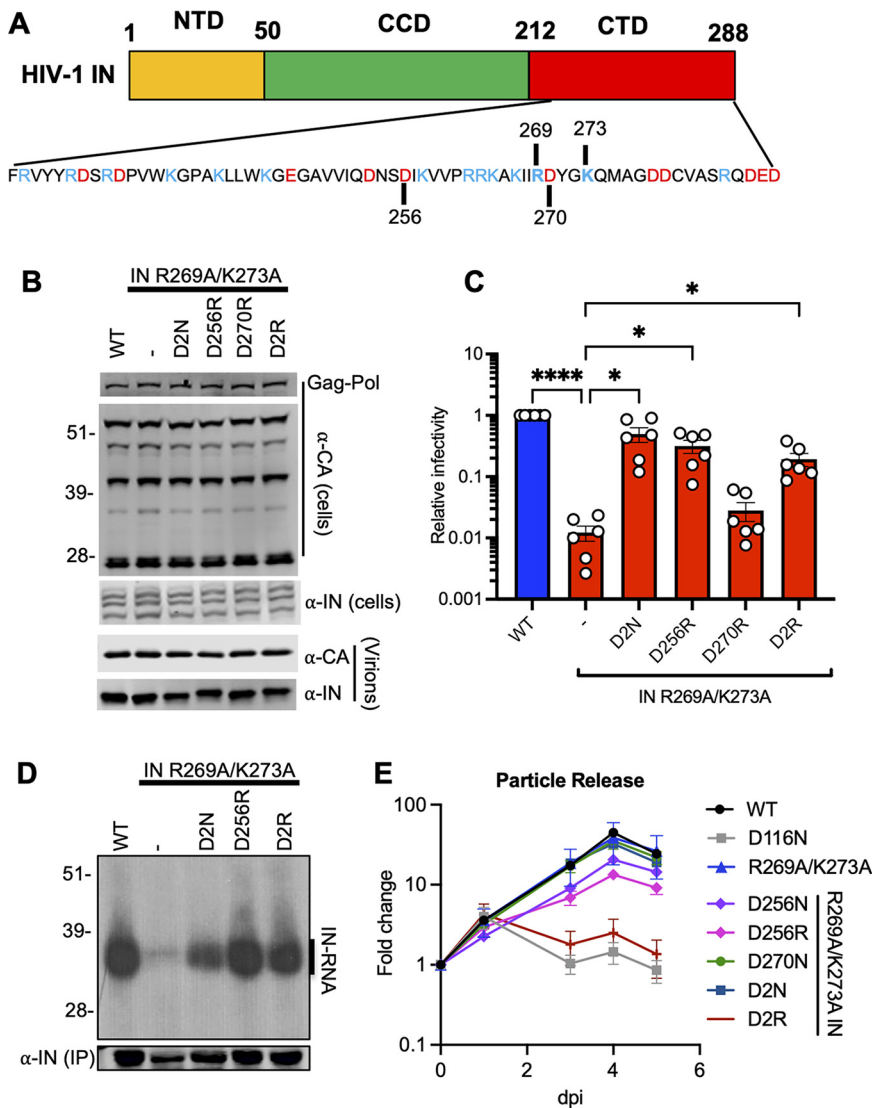
FIG 3 Restoring the local net charge of IN-CTD restores RNA binding and infectivity for the IN R269A/K273A class II mutant. (A) Schematic diagram of IN and sequence of CTD residues with basic and acidic amino acids highlighted in blue and red, respectively. (B, C) HEK293T cells were transfected with full-length proviral WT HIV-1$_{NL4-3}$ expression plasmid or its derivatives carrying *pol* mutations encoding for D2N, D256R, D270R, and D256R/D270R (D2R) IN substitutions on the backbone of HIV-1$_{NL4-3 IN (R269A/K273A)}$. (B) Cell lysates and purified virions harvested 2 days posttransfection were analyzed by immunoblotting for CA and IN. Representative image from five independent experiments is shown. See also Fig. S4A–E for quantitative analysis of immunoblots. (C) Cell culture supernatants containing viruses were titered on TZM-bl indicator cells, whereby virus replication was limited to a single cycle by addition of dextran sulfate (50 μg/mL). The titers were normalized relative to particle numbers as assessed by an RT activity assay and are presented relative to WT (set to 1). Also see Fig. S4F, G for titer and RT activity values prior to normalization. The columns represent the average of six independent experiments, and the error bars represent SEM (****, $P < 0.0001$; *, $P < 0.05$, by one-way ANOVA multiple comparison test with Dunnett's correction). (D) Autoradiogram of IN-RNA adducts immunoprecipitated from HIV-1$_{NL4-3}$ virus particles bearing WT IN or the indicated IN mutants. Immunoblot below shows immunoprecipitated (IP) IN protein. Results are representative of three independent replicates. See Fig. S4H for quantitative analysis of autoradiographs. (E) MT4 T-cells were infected with HIV-1$_{NL4-3 IN (D116N)}$ viruses that were transcomplemented with the indicated Vpr-IN mutant proteins as described in Materials and Methods. Virion release was assessed by RT activity assays over the course of 5 days postinfection. *y* axis indicates fold increase in RT activity over day 0. Data are from 3 independent replicates; error bars show the SEM.

of class II IN mutants into a class I IN mutant virus (HIV-1$_{NL4-3 IN (D116N)}$) bearing a catalytically inactive IN through fusion to the viral Vpr protein (76). In this assay, the RNA-binding defect of a class II IN mutant is complemented in *trans* by the catalytically inactive IN, whereas the class II IN mutant can complement the catalytic activity defect of the class I

mutant. This allows the specific assessment of class II and compensatory substitutions on the catalytic activity of IN, bypassing the need for successful completion of steps prior to integration. We found that the D256R substitution modestly decreased and the D2R substitutions completely abolished the ability of the R269A/K273A mutant to transcomplement the catalytically inactive IN D116N (Fig. 3E). None of the other IN substitutions appeared to have an impact. Interestingly, within the context of full-length viruses, we found that despite restoring reverse transcription to WT levels, the D2N substitution led to higher levels of 2-LTR accumulation and a 5-fold defect in integration (Fig. S3C, D). Similarly, D256R and D2R substitutions resulted in a modestly higher level of 2-LTR accumulation and a significantly lower level of integration (Fig. S3A–D). Taken together, these findings indicate that despite a decrease in the catalytic activity of IN, restoring the overall charge of IN-CTD through D2N and D256R substitutions specifically reestablishes RNA binding, overcomes the reverse transcription defect, and increases virion infectivity for the IN R269A/K273A class II mutant.

Charge reversal substitutions suppress the replication defect of a separate class II IN mutant virus through restoring gRNA binding. Mutation of other basic residues within the IN-CTD, such as R262A/R263A, also directly inhibits IN-gRNA binding without compromising oligomerization of IN (15, 38). If IN-gRNA binding is mediated by electrostatic interactions, we reasoned that substitutions of D256 and D270 residues in a way that restores the local positive charge should restore IN-gRNA binding and infectivity in the background of HIV-1$_{NL4-3 \text{ IN (R262A/R263A)}}$. D256N, D256K, D256R, D270N, and D2N substitutions introduced into HIV-1$_{NL4-3 \text{ IN (R262A/R263A)}}$ did not have a major impact on Gag and Gag-Pol expression, though a modest increase in distinct Gag processing intermediates was observed (Fig. 4A, C; Fig. S5A–D, G–J). Notwithstanding, neither virion release nor virion-associated IN levels were impacted (Fig. 4A, C; Fig. S5E, K, L). IN D256N/K/R substitutions increased virion infectivity by 10–13-fold, D2N substitutions 5-fold, whereas the D270N substitution had no impact (Fig. 4B, D; Fig. S5F, M). In line with the titer data, D256N/K/R substitutions all significantly increased IN-gRNA binding and accumulation of early and late RT products (Fig. 4E; Fig. S3A, B; Fig. S5N). D256N/K/R IN successfully transcomplemented a class I IN mutant (D116N) but displayed catalytic activity defects within the context of full-length viruses (Fig. S3C, D), suggesting that they specifically restored the second noncatalytic function of IN, namely, gRNA binding (Fig. 4F).

Charge reversal substitutions at acidic residues other than D256 can suppress the class II phenotype. Given the above findings, we next determined whether charge reversal substitutions at other nearby acidic residues, such as D278 and D279, could restore IN-gRNA binding and infectivity for the HIV-1$_{NL4-3 \text{ IN (R269A/K273A)}}$ virus. Remarkably, while the D279R substitution increased infectivity by 18-fold, the D278R substitution and the D278R/D279R (D2R') substitutions completely restored infectivity to WT levels (Fig. 5A). We also tested whether substitutions of the original D256 and D270 residues into Ile would restore virus titers at levels similar to the d-to-N substitutions. However, we found that neither D256I alone nor the D256I/D270I (D2I) substitutions increased viral titers (Fig. 5A). Of note, these substitutions did not impact Gag and Gag-Pol expression, processing, virion release, or virion-associated IN levels, though we noted the presence of an aberrantly processed IN with the D278R and D2I mutants (Fig. 5B). In line with the titer data, D279R substitution increased and the D278R and D2R' substitutions restored IN-RNA binding for the HIV-1$_{NL4-3 \text{ IN (R269A/K273A)}}$ virus (Fig. 5C). Subsequently, D278R and D2R' substitutions substantially increased the accumulation of reverse transcription products (Fig. S3A, B) without impairment of IN catalytic activity as evident in lack of 2-LTR accumulation (Fig. S3C) and near WT levels of integration (Fig. S3D). These and above results demonstrate that restoring the local positive charge of IN-CTD can efficiently restore/enhance RNA binding, reverse transcription, and virion infectivity to two distinct class II IN mutants. Cumulatively, these findings strongly suggest the importance of electrostatic interactions in mediating IN-gRNA binding.

Compensatory substitutions directly and specifically restore the ability of IN to bind vRNA. As IN oligomerization is a prerequisite for RNA binding (15), we next examined how the compensatory substitutions affected IN oligomerization. For in-virion
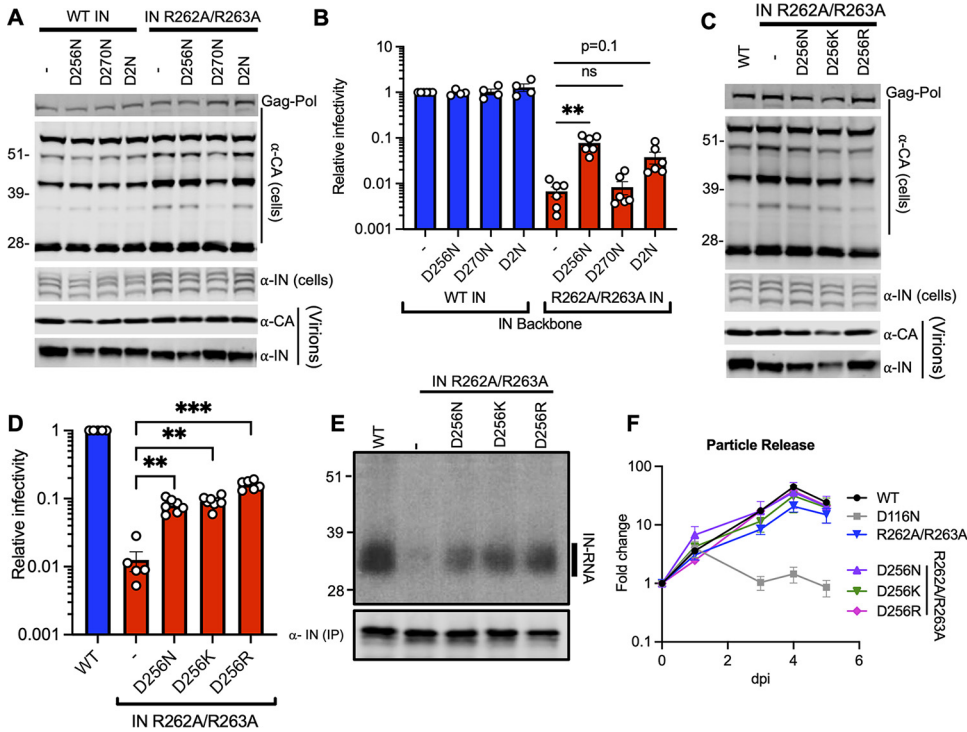
FIG 4 D256R and D256K substitutions restore IN-RNA binding and infectivity for the HIV-1$_{NL4-3\ IN\ (R262A/R263A)}$ class II mutant virus. (A–E) HEK293T cells were transfected with proviral HIV-1$_{NL4-3}$ expression plasmids carrying *pol* mutations for the indicated IN substitutions on WT IN or IN R262A/R263A backbones. (A, C) Cell lysates and purified virions collected 2 days posttransfection were analyzed by immunoblotting for CA and IN. See also Fig. S5 for quantitative analyses of immunoblots. (B, D) Cell culture supernatants containing viruses were titered on TZM-bl indicator cells, whereby virus replication was limited to a single cycle by addition of dextran sulfate (50 μg/mL). The titers were normalized relative to particle numbers as assessed by an RT activity assay and are presented relative to WT (set to 1). See also Fig. S5F, L, M for titer and RT activity values prior to normalization. The columns represent the average of four-to-six independent experiments, and the error bars represent SEM (**, $P < 0.01$; *, $P < 0.05$; ns = not significant by one-way ANOVA multiple comparison test with Dunnett's correction). (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 three independent replicates. See Fig. S5N for quantitative analysis of autoradiographs. (F) MT4 T-cells were infected with HIV-1$_{NL4-3\ IN\ (D116N)}$ viruses that were transcomplemented with the indicated Vpr-IN mutant proteins as described in Materials and Methods. Virion release was assessed by RT activity assays over the course of 5 days postinfection. *y* axis indicates fold increase in RT activity over day 0. Data are from 3 independent replicates; error bars show the SEM.

analysis, purified HIV-1$_{NL4-3}$ virions (WT and bearing the relevant IN substitutions) were treated with ethylene glycol bis (succinimidyl succinate) (EGS) to covalently cross-link 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 HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ and HIV-1$_{NL4-3\ IN\ (R262A/R263A)}$ 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 to C; reference 15).

Complementary *in vitro* experiments assessed oligomeric states of purified recombinant IN proteins by size exclusion chromatography (SEC). Results in Fig. 6D to H show that while tetramer/dimer/monomer ratios varied for different mutant proteins, the tetrameric form was a predominant species for all analyzed proteins under our experimental conditions. We also examined the ability of IN mutant proteins to bind LEDGF/p75, which preferentially interacts with and stabilizes IN tetramers (77). The results in Fig. 6I show that all tested proteins effectively interacted with LEDGF/p75. Collectively, the results in Fig. 6D–I suggest that D2N and D256R substitutions in the background of IN R269A/K273A or the IN R262A/R263A class II mutants did not grossly alter IN tetramerization.
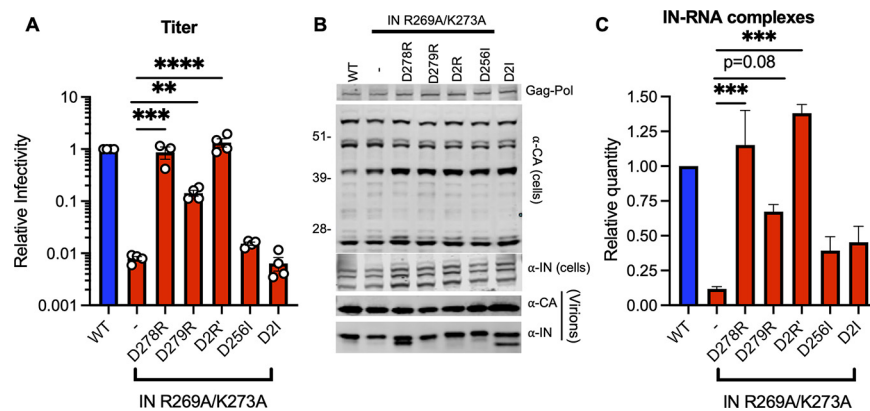
FIG 5 D279R substitution increases and D278R and D2R' substitutions restore IN-RNA binding and infectivity for the HIV-1$_{NL4-3 \text{ IN (R269A/K273A)}}$ virus. HEK293T cells were transfected with full-length pNL4-3 expression plasmids carrying mutations coding for the IN D278R, D279R, D278R/D279R (D2R'), D256I, and D256I/D270I (D2I) substitutions introduced on the IN R269A/K273A backbone. (A) Cell culture supernatants containing viruses were titered on TZM-bl indicator cells, whereby virus replication was limited to a single cycle by addition of dextran sulfate (50 μg/mL). The titers were normalized relative to particle numbers by an RT activity assay and are presented relative to WT (set to 1). The columns represent the average of 4 independent biological replicates, and the error bars represent standard error of the mean (SEM) (*, $P < 0.05$, by one-way ANOVA multiple comparison test with Dunnett's correction). (B) Cell lysates and purified virions collected 2 days posttransfection were analyzed by immunoblotting for CA and IN. The image is representative of four independent experiments. (C) IN-RNA adducts immunoprecipitated from WT or IN mutant HIV-1$_{NL4-3}$ virions per CLIP protocol were analyzed by autoradiography. Graph shows the quantification of IN-RNA adducts from three independent biological replicates; error bars show the SEM.

Next, we examined the ability of the mutant proteins to perform two critical functions *in vitro*: (i) catalyze integration (Fig. 6J), and (ii) bind and bridge synthetic RNA (Fig. 6K). Consistent with our previous findings (38), R269A/K273A substitutions only modestly (up to 2-fold) affected the catalytic function of the recombinant protein (Fig. 6J), whereas these substitutions abolished binding of the mutant protein to the synthetic TAR RNA (Fig. 6K). Strikingly, additional D2N substitutions engineered in the background of R269A/K273 IN exhibited opposite effects on integration activity versus RNA interactions. In line with infection-based experiments (Fig. S3), these additional substitutions further reduced the catalytic activity of the parental R269A/K273 mutant (Fig. 6J). IN R269A/K273A/D2N was ~5-fold less active than its WT counterpart (Fig. 6J) but exhibited WT levels of RNA binding (Fig. 6K). These biochemical results suggest that the D2N substitutions specifically restore the ability of the parental R269A/K273A IN to bind RNA. Similarly, the D256R substitution specifically enhanced the abilities of IN R269A/K273A and IN R262A/R263A to bridge between RNA molecules (Fig. 6K) without substantially affecting catalytic activities of these proteins. Taken together, the biochemical results in Fig. 6 demonstrate that the compensatory mutations directly and specifically restore the ability of IN to bind RNA.

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 (78), recent evidence suggests that defective proviruses can be transcribed into RNAs that are spliced and translated, and can be recognized by HIV-1-specific cytotoxic T lymphocytes (79). 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. 7A). Of note, only the R224Q substitution resulted in loss of a positive charge, whereas the E246K and G273R substitutions resulted in gain of positive charges. These mutations were introduced into the NL4-3 proviral backbone with minimal effects on Gag expression and particle release (Fig. 7B).
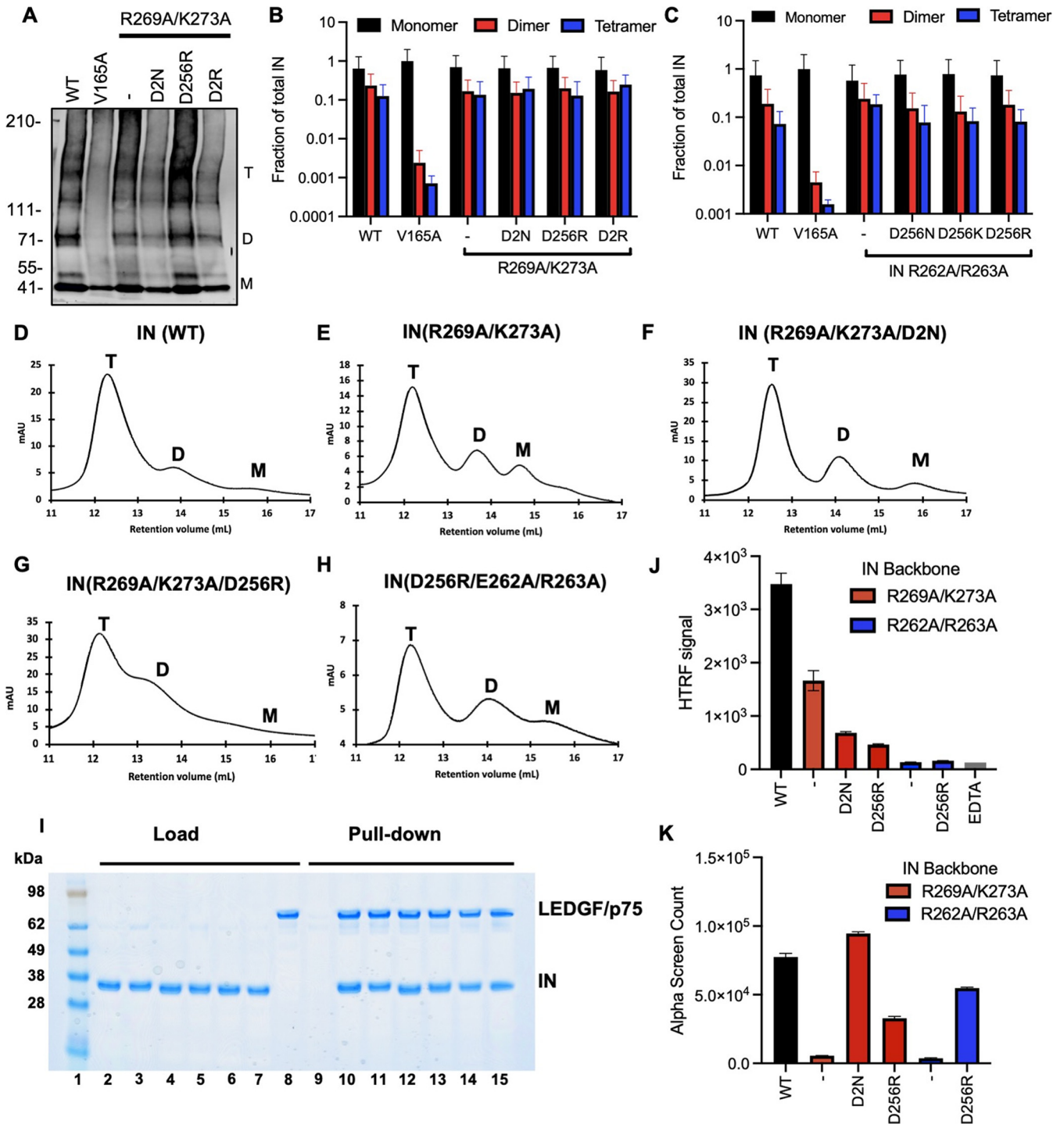
FIG 6 Assessing multimerization properties of IN in mutant viruses. (A) Purified WT or IN mutant HIV-1_{NL4-3} virions were treated with 1 mM EGS, and virus lysates analyzed by immunoblotting using antibodies against IN following separation on 3–6% Tris-acetate gels as detailed in Materials and Methods. 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–H) Representative SEC traces for indicated recombinant IN proteins at 20 μM. 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. Representative chromatograms from two independent analyses are shown. (I) Affinity-pulldown assays showing binding of WT and mutant INs to LEDGF/p75. Lane 1: molecular weight marker; Lanes 2–8: loads of 6xHis-tagged WT IN, IN(R269A/K273A), IN(R269A/K273A/D256N/ D270N), IN(R269A/K273A/D256R), IN(D256R/R262A/R263A), IN(D256R/R262A/R263A), and FLAG-tagged LEDGF/p75; Lanes 9–15: affinity pull-down using Ni beads of LEDGF/p75 with buffer only (control), 6xHis-tagged WT IN, IN(R269A/K273A), IN(R269A/K273A/D256N/D270N), IN(R269A/K273A/D256R), IN(D256R/ R262A/R263A), and IN(D256R/R262A/R263A). (J) Catalytic activities of mutant IN molecules in the presence of LEDGF/p75 monitored by HTRF based assay. The bars represent the ratio of emission signal at 665 nm over that of 615 nm, which was then multiplied by a factor of 10,000. The error bars indicate the standard error of the mean from triplicate experiments. (K) 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.
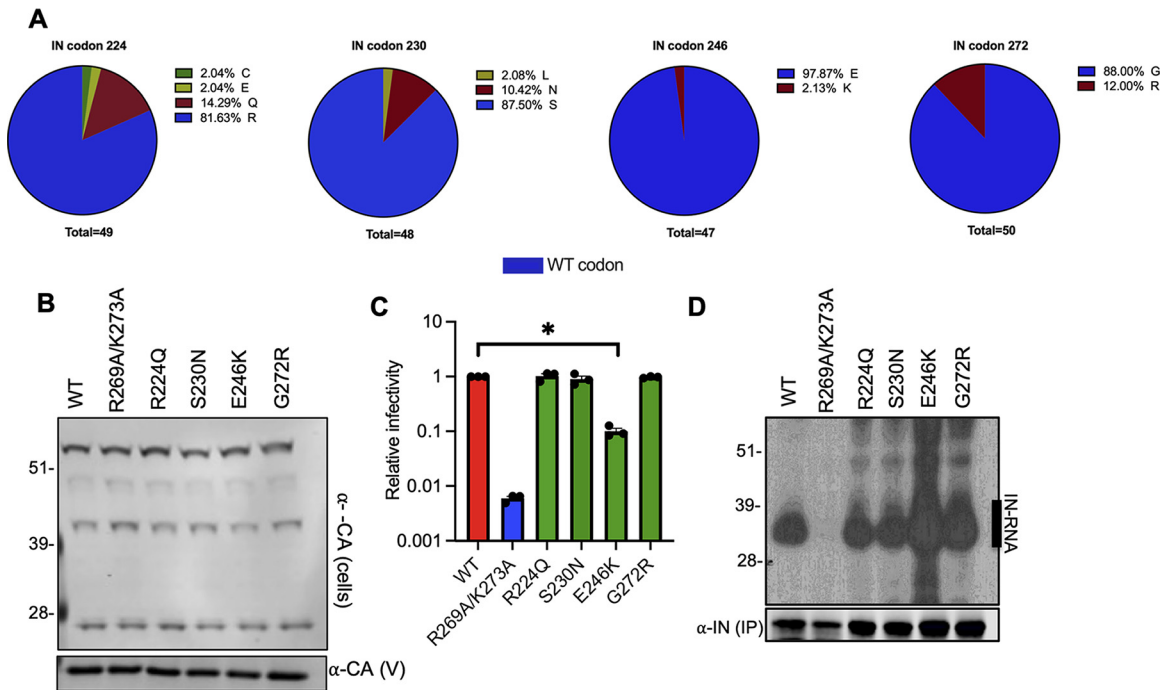
FIG 7 Characterization of IN mutations present in latently infected CD4$^+$ T cells. (A) Frequency of IN substitutions at amino acids 224, 230, 246, and 272 derived from HIV-1 sequences from latently infected CD4$^+$ T-cells is shown. (B) HEK293T cells were transfected with proviral HIV-1$_{NL4-3}$ expression plasmids carrying the R224Q, S230N, E246K, and G272R IN mutations. Cell lysates and virions were purified 2 days posttransfection 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 (*, $P < 0.05$, by one-way ANOVA multiple comparison test with Dunnett's correction). (D) Autoradiogram of IN-RNA adducts immunoprecipitated from virions bearing the indicated substitutions in IN. Immunoblots below show the amount of immunoprecipitated IN.

Although the E246K virus was significantly less infectious (Fig. 7C), we did not find any evidence for loss of IN-gRNA binding (Fig. 7D), 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.

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 substitutions, impede IN-gRNA binding without affecting oligomerization of IN (15, 38). During serial passaging of the IN R269A/K273A class II mutant virus, the D256N and D270N substitutions emerged sequentially. We also wondered whether mutations outside of IN, such as CA and NC, could also arise, given that the IN R269A/K273A mutant is still catalytically active and compensatory mutations in CA and NC could potentially provide alternative mechanisms for the proper packaging of the gRNA within virions. However, we did not observe the emergence of such substitutions showcasing the distinct role of IN-gRNA binding in proper virion maturation. Furthermore, the emergence of compensatory substitutions only in CTD, but not NTD or CCD of IN, highlights the importance of the IN-CTD in mediating direct RNA binding.

It is worth noting that D256N and D270N substitutions each arose through a single mutation (D256N: GAC→AAC, D270N: GAU→AAU), and thus possibly provided an easier pathway for suppression than reverting back to R269 and K273, each of which would require two mutations. Likewise, d-to-R and d-to-K substitutions require two and three mutations, respectively, and possibly through an unfavorable intermediate amino acid substitution, creating a higher barrier to arise in culture. While the D278R

substitution was the most effective in restoring gRNA binding and without any impact on the catalytic activity of IN, the rise of this mutation was likely additionally limited by the necessity to maintain a functional Vif given the overlap of IN-CTD and Vif ORFs. Though other mutations in IN could in principle restore the ability of IN to bind RNA, rise of such mutations was mostly likely constrained in part by the necessity to maintain a catalytically active IN.

The IN-CTD is decorated by numerous acidic and basic residues (Fig. 3A). While the IN R269A/K273A substitutions resulted in the loss of 2 positive charges, the compensatory D2N substitutions restored the overall local charge within the CTD, suggesting that electrostatic interactions contribute to IN-CTD binding to the gRNA. Consistent with this hypothesis, the D256R, D278R, and D279R substitutions that resulted in the gain of two positive charges locally were all sufficient to substantially increase or restore IN-gRNA binding, and subsequently reverse transcription and infectivity for the IN R269A/K273A mutant. Furthermore, these observations extended to another class II mutant, IN R262A/R263A, whereby the D256N/K/R substitutions specifically enhanced RNA binding and subsequently increased the accumulation of RT products and infectivity. In contrast, the D2N substitutions did not enhance RNA-binding or infectivity further for the IN R262A/R263A mutant, demonstrating a degree of context/localization dependency. In other words, it is possible that the proximity of the D270N to R269A/K273A residues may explain why it was more effective in restoring RNA binding particularly for this class II mutant but not for IN R262A/R263A. Cumulatively, our findings revealed a key electrostatic component of IN-gRNA interactions.

Inspection of the available X-ray structure (63) indicates that all IN residues implicated in RNA binding by the present study are positioned either in the highly flexible C-terminal tail (amino acids [aa] 261–275) or the 252–257 loop (Fig. S6H). R269A and K273A substitutions expectedly resulted in a substantial loss of a basic patch in IN (Fig. S6A, B). Both the D256R and D256N substitutions resulted in a more positively charged surface distal from the R269/K273A residues (Fig. S6C, D), suggesting that compensatory mutations likely created an additional interaction interface for gRNA binding. A similar outcome was observed when class II R262A/R263A and compensatory D256R changes were introduced in the CTD of IN (Fig. S6E–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. S6H). 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.

Arg and Lys residues are commonly employed by numerous RNA-binding proteins, and Arg residues are generally more heavily involved interactions with all bases (71, 80–84). Though the electrostatic interactions between the Arg/Lys residues of IN-CTD and gRNA imply a level of nonspecificity, IN binds to distinct locations on the gRNA in virions and displays high binding affinity to structured elements, such as TAR (38). Thus, it is likely that IN-gRNA interactions are mediated by both the nonspecific interactions of the basic residues with the RNA phosphate backbone and specific interactions with the cognate RNA, as has also been suggested by a recent modeling study (73). Importantly, we note that while the present study elucidates crucial roles of the IN-CTD basic residues, other yet-to-be-identified IN amino acids are likely to also contribute to high affinity binding of select gRNA segments. For example, recognition of the TAR loop by Tat and the super elongation complex is based on a complex set of interactions: the super elongation complex primarily reads out the structure as opposed to the sequence, whereas the Tat arginine rich motif (ARM) interacts with the TAR bulge and the major groove through electrostatic interactions with the RNA backbone and H-bonding with specific bases (85, 86). Structural studies of IN in complex with cognate RNA oligonucleotides are crucial to tease apart the specificity determinants for IN-gRNA interactions.

RNA binding proteins commonly encode modular RNA-binding domains (i.e., RRMs and KH domains), which form specific contacts with short degenerate sequences (87–89). Utilization of multiple RRMs/KH domains is thought to create a much larger binding interface, which in turn allows for specific, high-affinity binding to target RNAs (87, 88, 90). Functionally important IN tetramerization (15) may similarly enhance binding affinity and specificity to cognate gRNA elements. In this respect, it is worth noting that CTDs of the two inner protomers in the tetrameric intasome structure contact extensively with DNA, whereas the outer CTDs are positioned in close proximity to the inner CTDs and only partially contribute DNA binding (91). It remains to be seen whether the inner CTDs could also preferentially engage gRNA, or CTD-nucleic acid interactions within IN-gDNA and IN-gRNA complexes differ substantially.

Overall, our studies reveal that pliable electrostatic interactions play an important role in mediating IN-CTD-gRNA interactions and demonstrate that CTD is a key determinant of gRNA binding. IN-gRNA binding is essential for HIV-1 virion morphological maturation and infectivity, thus an excellent target for novel antiretroviral compounds. ALLINI-mediated inhibition of the noncatalytic function of IN can complement existing INSTI-based therapies and increase the barrier to drug resistance substantially. Of note, a highly potent and safe pyrrolopyridine-based ALLINI, STP0404, has recently advanced to human trials (66). Further structural characterization of IN-gRNA and IN-ALLINI complexes will be crucial to determine the precise rules that govern IN-gRNA interactions and guide the development of novel therapeutics that target this noncatalytic function of IN.

MATERIALS AND METHODS

Plasmids and compounds. 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 PCRs with forward (with EcoRI restriction endonuclease site) and reverse (with AgeI restriction endonuclease site) outer primers. The resulting fragments containing the desired mutations were mixed at a 1:1 ratio and overlapped subsequently using the outer primer pairs. The overlap fragments were digested with AgeI and EcoRI before cloning into pNL4-3 plasmids. The pLR2P-VprIN plasmid expressing a Vpr-IN fusion protein has been previously described (76). 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.

ALLINIs, BI-D and BI-B2, were synthesized by Aris Pharmaceuticals using previously published protocols (92, 93). Compounds were dissolved in dimethyl sulfoxide (DMSO) at a final concentration of 10 mM and stored at −80 °C.

Cell lines and viruses. HeLa-derived TZM-bL cells that express CD4/CXCR4/CCR5 receptor/coreceptors and bear an HIV-1 LTR driven β-galactosidase reporter were obtained from the NIH AIDS Reagent Program. HEK293T cells (ATCC CRL-11268) and TZM-bL cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Human CD4$^+$ T-cell line, MT4 (NIH AIDS Reagents), were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum. A derivative of MT4-LTR-GFP (MT4-GFP) indicator cells bearing an HIV-1 LTR driven GFP reporter have been described before (94) and similarly maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum. All cell lines were obtained from the American Type Culture Collection and NIH AIDS Reagents, where short tandem repeat (STR) profiling was performed. MT-4 T-cells were further authenticated by STR profiling at Washington University School of Medicine Genome Engineering and iPSC center. The cell lines are regularly inspected for mycoplasma contamination using the MycoAlert mycoplasma detection kit (Lonza) and checked for being free of any other contaminations.

For generation of HIV-1 stocks, HEK293T cells grown in 10-cm dishes were transfected with the pNL4-3 and vesicular stomatitis virus glycoprotein (VSV-G) expression plasmids at a ratio of 4:1 using polyethyleneimine (PolySciences, Warrington, PA). Cell culture supernatants containing infectious virions were collected 2 days posttransfection, filtered, aliquoted, and stored at −80°C. For characterization of the effects of compensatory mutations on virus replication, HEK293T cells were transfected in 24-well plates with pNL4-3-derived plasmids similarly but without VSV-G pseudotyping. Cell lysates collected at 2 days posttransfection were analyzed by immunoblotting as detailed below. In parallel, cell culture supernatants containing virions were titered on TZM-bL cells using a β-galactosidase assay (Galactostar, Thermo Fisher) per manufacturer's instructions. Virus replication in TZM-bL cells was limited to a single cycle via addition of dextran sulfate (50 μg/mL) 6–14 h postinfection. An aliquot of the viruses was also subjected to a qPCR-based reverse transcriptase (RT) activity assay (95). Virus titers obtained from infection of TZM-bL cells was normalized to particle numbers based on values obtained from RT activity assays to determine the infectiousness (i.e., infectious units/particle) of viruses. A fraction of the cell culture supernatant collected from transfected HEK293T cells was also concentrated by centrifugation (16,100 × g, 4°C, 90 min) on a 20% sucrose cushion prepared in 1 × SDS-PAGE sample buffer for analysis of virion-associated proteins by immunoblotting as described below.

Analysis of compensatory mutations. Compensatory substitutions that arose in the background of HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ class II IN mutant virus were isolated by serial passaging as follows. In brief, one million MT4-GFP cells were infected at a multiplicity of infection (MOI) 2 for the WT virus. An equivalent particle number (as determined by a qPCR-based reverse transcriptase (RT) activity assay [95]) of the HIV-1$_{NL4-3\ IN\ (R269A/K273A)}$ virus was used as inoculum. Infected cells were split at a ratio of 1:5 every 3–4 days and infections were monitored by microscopy as well as flow cytometry of GFP-positive cells at the indicated intervals shown in Fig. 1A. Cell culture supernatants containing virions were also collected at the same time intervals for subsequent sequencing analysis. At the end of each passage (i.e., when infections plateaued at >75% GFP+ cells), cell culture supernatants containing infectious virions were collected. The resultant viruses were first titered on MT4-GFP cells and used to infect MT4-GFP cells in the next passage at an MOI of 2 infectious units (i.u.)/cell for WT and an equivalent particle number for the class II IN mutant virus (or its derivative bearing compensatory mutations). Aliquots of infected cells and viral particles in the cell culture supernatants were collected over the duration of each passage as above.

For sequencing analysis of viruses bearing compensatory mutations, virions collected over the three passages were concentrated on 20% sucrose cushions (prepared in 1× phosphate-buffered saline [PBS]) by ultracentrifugation (28,000 rpm, 4°C, 90 min, Beckman SW41 rotor). Genomic RNA was isolated from pelleted virions using TRIzol per manufacturer's instructions. Extracted RNA was prepared for deep sequencing using the Illumina TruSeq Stranded Total RNA library kit following the manufacturer's instructions but omitting the rRNA depletion step. Resulting libraries were sequenced on an Illumina HiSeq 2000 platform at the Genome Technology Access Center at Washington University School of Medicine. Sequencing reads were mapped to the pNL4-3 reference viral genome allowing for 2 mismatches using the Bowtie algorithm (i.e., -v 2, -m 10 parameters) (96), and the frequency of compensatory mutations was determined thereafter.

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 (Hybond ECL, Amersham). 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 (97) in Odyssey Blocking Buffer (LI-COR). Membranes were probed with fluorophore-conjugated secondary antibodies (LI-COR) and scanned using a LI-COR Odyssey system. IN and CA levels in virions were quantified using Image Studio software (LI-COR).

Vpr-IN trans-complementation experiments. HIV-1$_{NL4-3}$ bearing the class I IN mutation D116N was transcomplemented with class II mutant proteins as previously described (76). In brief, HEK293T cells plated in 24-well dishes were co-transfected with full-length proviral plasmids expressing HIV-1$_{NL4-3\ IN\ (D116N)}$, 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 using polyethyleneimine as described above. Cell-free virions were collected from cell culture 2 days posttransfection. MT-4 T-cells were infected by the resultant virus stocks, and the integration capability of transcomplemented class II IN mutants was measured by the yield of progeny virions in cell culture supernatants over a 6-day period using the aforementioned reverse transcriptase (RT) activity assay as described before (76). Briefly, MT-4 T-cells were incubated with virus inoculum in 96 V-bottom well plates for 4 h 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, an aliquot of the media was collected and the number of virions in culture supernatant was quantified by measuring a qPCR-based RT activity using assay (95), as also described above.

CLIP experiments. CLIP experiments were conducted as previously described (15, 74, 75, 98, 99). In short, cells in 15-cm cell culture plates were transfected with 30 μg full-length proviral plasmid (pNL4-3) DNA or derivatives carrying the IN mutations. 4-thiouridine (4SU) was added to the cell culture media for 16 h before virus harvest. Cell culture supernatants containing virions were filtered through 0.22-μm filters and pelleted by ultracentrifugation through a 20% sucrose cushion prepared in 1× phosphate-buffered saline (PBS) using a Beckman SW32-Ti rotor at 28,000 rpm for 1.5 h at 4°C. The virus pellets were resuspended in 1× PBS and UV-cross-linked for two consecutive times at an energy setting of 500 mJ in a Boekel UV-cross-linking chamber equipped with UV 368-nm bulbs. Following lysis in 1× RIPA buffer, IN-RNA complexes were immunoprecipitated using a mouse monoclonal anti-IN antibody (97). 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 (Hybond ECL, Amersham), and exposed to autoradiography films to visualize IN-RNA complexes. If needed, following immunoblotting and quantitation of immunoprecipitated IN, equivalent amounts of immunoprecipitated IN were rerun to accurately assess the ability of IN variants to bind RNA. Autoradiographs were quantitated by the ImageStudio software (LI-COR). Lysates and immunoprecipitates were also analyzed by immunoblotting using the aforementioned mouse monoclonal antibodies against IN.

IN multimerization in virions. HEK293T cells grown in 10-cm dishes were transfected with 10 μg pNL4-3 plasmid DNA bearing WT IN or the indicated *pol* mutations within IN coding sequence as above. Two days posttransfection, cell-free virions in cell culture supernatants were pelleted through a 20% sucrose cushion prepared in 1× PBS using a Beckman SW41-Ti rotor at 28,000 rpm for 1.5 h at 4°C. Pelleted virions were resuspended in 1× PBS and treated with a membrane-permeable cross-linker, EGS (ThermoFisher Scientific), at a concentration of 1 mM for 30 min at room temperature. Crosslinking was stopped by the addition of SDS-PAGE 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 (97).

Virus production and transmission electron microscopy. HEK293T cells grown in 15-cm plates were transfected with 30 μg full-length proviral plasmid (pNL4-3) DNA containing WT IN or the indicated

pol mutations within IN coding sequence as above. Two days posttransfection, 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 h 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 post-fixed in 1% osmium tetroxide (Polysciences, Inc.) for 1 h. Samples were then rinsed extensively in dH$_2$0 prior to *en bloc* staining with 1% aqueous uranyl acetate (Ted Pella Inc., Redding, CA) for 1 h. After several rinses in dH$_2$0, 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 mega-pixel digital camera and AMT Image Capture Engine V602 software (Advanced Microscopy Techniques, Woburn, MA). One hundred virions from each sample were classified for virion morphology.

Quantitation of early and late RT products, 2-LTR circles, and integrated proviral DNA. We infected 1×10^6 MT4-LTR-GFP indicator cells (per time point/per condition) with equivalent particle numbers of WT and class II IN mutant viruses bearing compensatory mutations. Cells were collected at 4, 8, and 24 hours postinfection (hpi) and pelleted, and total DNA extracted by lysis of cells in Proteinase K (0.1 mg/mL) containing buffer (100 mM NaCl, 10 mM Tris pH 8.0, 25 mM EDTA, 0.5% SDS) for 2 h at 50°C on a thermal mixer set to 1,800 rpm. DNA was extracted by phenol:chloroform:isoamyl alcohol and precipitated with ammonium acetate and ethanol. The resulting DNA pellet was resuspended in 100 μL nuclease-free water, and 5 μL of the sample was subjected to SYBR-green-based qPCR for quantitation of early RT (Forward: 5'-GCTAACTAGGGAACCCACTGCTT; Reverse: 5'-CAACAGACGGGCACACACTACT), late RT (Forward: 5'-TGGGCAAGCAGGGAGCTA; Reverse: 5'-TCCTGTCTGAAGGGATGGTTGT), and 2-LTR (Forward: 5'-CTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTA; Reverse: 5'-TGACCCCTGGCCCTGGTGTGTAG) products as described previously (54). For quantitation of integrated proviral DNA by Alu-PCRs, 4 μL of the DNA samples were subjected to an initial PCR (Forward: 5'-TCCCAGCTACTCGGGAGGCTGAGG; Reverse: 5'-AGGCAAGCTTTATTGAGGCTTAAGC) followed by SYBR green-based qPCR quantification (Forward: 5'-GAAGGGCTAATTCACTCCCA; Reverse: 5'-CTTGAAGTACTCCGGATGCAG) as detailed previously (54).

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) (67). His$_6$ tagged recombinant pNL4-3 WT and mutant IN molecules were expressed in BL21 (DE3) *E. coli* cells followed by nickel and heparin column purification as described previously (67, 100). Recombinant WT and mutant IN molecules 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, 7.5 mM CHAPS, and 5 mM BME at 0.2 mL/min flow rate. The proteins were diluted to 20 μM with the running buffer and incubated for 1 h at 4°C followed by centrifugation at 10,000 $\times$ *g* 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 B$_{12}$ (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.

LEDGF/p75 pulldown assay. The method was modified from that of McKee et al., 2008 (77) in which LEDGF/p75 (2 μM) was incubated with 2 μM His-tagged IN (WT or mutant) in binding buffer (50 mM HEPES (pH 7.5), 200 mM NaCl, 2 mM MgCl$_2$, 15 mM imidazole, 0.1% (vol/vol) Nonidet P40) for 60 min at room temperature. Samples were then briefly centrifuged for 2 min at 1,000 $\times$ *g*, and supernatants were pulled down by Ni-NTA resin for 30 min in the presence of bovine serum albumin (0.1 mg/mL). The resin was then washed three times with the same buffer, and the bound proteins were separated by SDS-PAGE. LEDGF and INs were separated by SDS–PAGE electrophoresis and visualized by staining with Coomassie-Blue-like AcquaStain (Bulldog-Bio).

Homogeneous time resolved fluorescence (HTRF). Catalytic activities of WT and mutant INs were analyzed in the presence of LEDGF/p75 using homogeneous time-resolved fluorescence (HTRF) according to Koneru et al., 2019 (101). Briefly, 100 nM each IN was incubated with 100 nM LEDGF/p75, 50 nM Cy-5 labeled donor DNA, and 10 nM biotinylated target acceptor DNA in 20 mM HEPES (pH 7.5), 1 mM DTT, 10 mM MgCl$_2$, 10% glycerol, 0.05% Brij-35, and 0.1 mg/mL BSA. End detection was based on europium-streptavidin that binds to the biotinylated DNA and brings donor europium cryptate closer to acceptor Cy5 fluorophore in integrated DNA. This proximity results in energy transfer to yield a fluorescent signal that was recorded by a PerkinElmer Life Sciences Enspire multimode plate reader. The data were plotted as Ratio, being the ratio of emission signal at 665 nm over that of 615 nm, which was then multiplied by a factor of 10,000.

Analysis of IN-RNA binding *in vitro*. To monitor IN-RNA interactions, we utilized AlphaScreen-based assay (38), 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 with either biotin or digoxigenin (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 h incubation at 4°C, 320 nM IN was added to the reaction mixture and incubated further for 1.5 h 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 the Adaptive Poisson-Boltzmann Solver (APBS) program (102) with macromolecular electrostatic calculations performed in PyMOL. The published crystal structure (63) (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 (78) were downloaded from NCBI GenBank based on their accession numbers (KF526120– KF526339). Sequences were imported into the GLUE software framework (103, 104) and aligned. Multiple sequence alignments (MSAs) containing subtype B sequences were constructed using MUSCLE, manually inspected in AliView (105), 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.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 2.6 MB.
FIG S2, TIF file, 2.7 MB.
FIG S3, TIF file, 1.4 MB.
FIG S4, TIF file, 1.8 MB.
FIG S5, JPG file, 1 MB.
FIG S6, TIF file, 1.8 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Engelman AN. 2019. Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. J Biol Chem 294:15137–15157. https://doi.org/10.1074/jbc.REV119.006901.

2. 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 287:646–650. https://doi.org/10.1126/science.287.5453.646.

3. 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 51:5843–5855. https://doi.org/10.1021/jm800245z.

4. Ramanathan S, Mathias AA, German P, Kearney BP. 2011. Clinical pharmacokinetic and pharmacodynamic profile of the HIV integrase inhibitor elvitegravir. Clin Pharmacokinet 50:229–244. https://doi.org/10.2165/11584570-000000000-00000.

5. 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 54:254–258. https://doi.org/10.1128/AAC.00842-09.

6. 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. https://doi.org/10.1128/AAC.01474-16.

7. 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. https://doi.org/10.1186/s12977-018-0420-7.

8. 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. https://doi.org/10.1097/QAI.0b013e31828ace69.

9. Taha H, Das A, Das S. 2015. Clinical effectiveness of dolutegravir in the treatment of HIV/AIDS. Infect Drug Resist 8:339–352. https://doi.org/10.2147/IDR.S68396.

10. 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 64:e00611-20. https://doi.org/10.1128/AAC.00611-20.

11. 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. https://doi.org/10.1002/phar.2246.

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 110:8690–8695. https://doi.org/10.1073/pnas.1300703110.

13. 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 78:12735–12746. https://doi.org/10.1128/JVI.78.23.12735-12746.2004.

14. Engelman A. 1999. In vivo analysis of retroviral integrase structure and function. Adv Virus Res 52:411–426. https://doi.org/10.1016/s0065-3527(08)60309-7.

15. 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:e54311. https://doi.org/10.7554/eLife.54311.

16. 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 69:2729–2736. https://doi.org/10.1128/JVI.69.5.2729-2736.1995.

17. Engelman A, Cherepanov P. 2014. Retroviral integrase structure and DNA recombination mechanism. Microbiol Spectr 2:1–22. https://doi.org/10.1128/microbiolspec.MDNA3-0024-2014.

18. Engelman AN, Cherepanov P. 2017. Retroviral intasomes arising. Curr Opin Struct Biol 47:23–29. https://doi.org/10.1016/j.sbi.2017.04.005.

19. Craigie R, Bushman FD. 2012. HIV DNA integration. Cold Spring Harb Perspect Med 2:a006890. https://doi.org/10.1101/cshperspect.a006890.

20. Engelman A, Cherepanov P. 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat Rev Microbiol 10:279–290. https://doi.org/10.1038/nrmicro2747.

21. 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 11:1287–1289. https://doi.org/10.1038/nm1329.

22. 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 21:1767–1778. https://doi.org/10.1101/gad.1565107.

23. 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 110:521–529. https://doi.org/10.1016/s0092-8674(02)00864-4.

24. Christ F, Debyser Z. 2013. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. Virology 435:102–109. https://doi.org/10.1016/j.virol.2012.09.033.

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

26. 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 70:6820–6825. https://doi.org/10.1128/JVI.70.10.6820-6825.1996.

27. 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 271:7712–7718. https://doi.org/10.1074/jbc.271.13.7712.

28. 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 259:274–285. https://doi.org/10.1006/viro.1999.9767.

29. 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 70:721–728. https://doi.org/10.1128/JVI.70.2.721-728.1996.

30. Liao WH, Wang CT. 2004. Characterization of human immunodeficiency virus type 1 Pr160 gag-pol mutants with truncations downstream of the protease domain. Virology 329:180–188. https://doi.org/10.1016/j.virol.2004.08.010.

31. 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 79:10356–10368. https://doi.org/10.1128/JVI.79.16.10356-10368.2005.

32. 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 239:715–722. https://doi.org/10.1006/bbrc.1997.7541.

33. 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. https://doi.org/10.1006/viro.1996.0219.

34. Shin CG, Taddeo B, Haseltine WA, Farnet CM. 1994. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. J Virol 68:1633–1642. https://doi.org/10.1128/JVI.68.3.1633-1642.1994.

35. Taddeo B, Haseltine WA, Farnet CM. 1994. Integrase mutants of human immunodeficiency virus type 1 with a specific defect in integration. J Virol 68:8401–8405. https://doi.org/10.1128/JVI.68.12.8401-8405.1994.

36. 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 73:2126–2135. https://doi.org/10.1128/JVI.73.3.2126-2135.1999.

37. 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 89:9765–9780. https://doi.org/10.1128/JVI.01522-15.

38. 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. https://doi.org/10.1016/j.cell.2016.07.044.

39. Mathew S, Nguyen M, Wu X, Pal A, Shah VB, Prasad VR, Aiken C, Kalpana GV. 2013. INI1/hSNF5-interaction defective HIV-1 IN mutants exhibit impaired particle morphology, reverse transcription and integration in vivo. Retrovirology 10:66. https://doi.org/10.1186/1742-4690-10-66.

40. 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. https://doi.org/10.1186/1742-4690-2-62.

41. 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 365:1480–1492. https://doi.org/10.1016/j.jmb.2006.10.094.

42. 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 71:3507–3514. https://doi.org/10.1128/JVI.71.5.3507-3514.1997.

43. 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 76:10598–10607. https://doi.org/10.1128/jvi.76.21.10598-10607.2002.

44. Lloyd AG, Ng YS, Muesing MA, Simon V, Mulder LC. 2007. Characterization of HIV-1 integrase N-terminal mutant viruses. Virology 360:129–135. https://doi.org/10.1016/j.virol.2006.10.007.

45. 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 79:12584–12591. https://doi.org/10.1128/JVI.79.19.12584-12591.2005.

46. 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 69:6687–6696. https://doi.org/10.1128/JVI.69.11.6687-6696.1995.

47. 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. https://doi.org/10.1016/j.virol.2006.08.011.

48. Riviere L, Darlix JL, Cimarelli A. 2010. Analysis of the viral elements required in the nuclear import of HIV-1 DNA. J Virol 84:729–739. https://doi.org/10.1128/JVI.01952-09.

49. 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 74:4795–4806. https://doi.org/10.1128/jvi.74.10.4795-4806.2000.

50. 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 69:376–386. https://doi.org/10.1128/JVI.69.1.376-386.1995.

51. 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 78:5045–5055. https://doi.org/10.1128/jvi.78.10.5045-5055.2004.

52. 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 289:25351–25361. https://doi.org/10.1074/jbc.M113.533281.

53. 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 425:2133–2146. https://doi.org/10.1016/j.jmb.2013.03.027.

54. 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 85:4654–4666. https://doi.org/10.1128/JVI.02374-10.

55. 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 76:4331–4340. https://doi.org/10.1128/jvi.76.9.4331-4340.2002.

56. Engelman A, Craigie R. 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J Virol 66:6361–6369. https://doi.org/10.1128/JVI.66.11.6361-6369.1992.

57. 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 72:4841–4848. https://doi.org/10.1128/JVI.72.6.4841-4848.1998.

58. Lutzke RA, Vink C, Plasterk RH. 1994. Characterization of the minimal DNA-binding domain of the HIV integrase protein. Nucleic Acids Res 22: 4125–4131. https://doi.org/10.1093/nar/22.20.4125.

59. 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. https://doi.org/10.1371/journal.pone.0074163.

60. 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. https://doi.org/10.1186/1742-4690-10-57.

61. 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. https://doi.org/10.1371/journal.ppat.1004171.

62. 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 289:20477–20488. https://doi.org/10.1074/jbc.M114.551119.

63. 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. https://doi.org/10.1371/journal.pbio.1002584.

64. 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: 4365–4374. https://doi.org/10.1128/AAC.00717-12.

65. 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 6:442–448. https://doi.org/10.1038/nchembio.370.

66. 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. https://doi.org/10.1371/journal.ppat.1009671.

67. 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. Multi-mode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. J Biol Chem 287:16801–16811. https://doi.org/10.1074/jbc.M112.354373.

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

69. 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 287:21189–21203. https://doi.org/10.1074/jbc.M112.347534.

70. Serganov A, Patel DJ. 2008. Towards deciphering the principles underlying an mRNA recognition code. Curr Opin Struct Biol 18:120–129. https://doi.org/10.1016/j.sbi.2007.12.006.

71. Bahadur RP, Zacharias M, Janin J. 2008. Dissecting protein-RNA recognition sites. Nucleic Acids Res 36:2705–2716. https://doi.org/10.1093/nar/gkn102.

72. Stefl R, Skrisovska L, Allain FH. 2005. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO Rep 6:33–38. https://doi.org/10.1038/sj.embor.7400325.

73. Dixit U, Bhutoria S, Wu X, Qiu L, Spira M, Mathew S, Harris R, Adams LJ, Cahill S, Pathak R, Rajesh Kumar P, Nguyen M, Acharya SA, Brenowitz M, Almo SC, Zou X, Steven AC, Cowburn D, Girvin M, Kalpana GV. 2021. INI1/SMARCB1 Rpt1 domain mimics TAR RNA in binding to integrase to

74. facilitate HIV-1 replication. Nat Commun 12:2743. https://doi.org/10.1038/s41467-021-22733-9.

74. Shema Mugisha C, Tenneti K, Kutluay SB. 2020. Clip for studying protein-RNA interactions that regulate virus replication. Methods 183:84–92. https://doi.org/10.1016/j.ymeth.2019.11.011.

75. Bieniasz PD, Kutluay SB. 2018. CLIP-related methodologies and their application to retrovirology. Retrovirology 15:35. https://doi.org/10.1186/s12977-018-0417-2.

76. 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 71:7704–7710. https://doi.org/10.1128/JVI.71.10.7704-7710.1997.

77. McKee CJ, Kessl JJ, Shkriabai N, Dar MJ, Engelman A, Kvaratskhelia M. 2008. Dynamic modulation of HIV-1 integrase structure and function by cellular lens epithelium-derived growth factor (LEDGF) protein. J Biol Chem 283:31802–31812. https://doi.org/10.1074/jbc.M805843200.

78. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent non-induced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551. https://doi.org/10.1016/j.cell.2013.09.020.

79. 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. https://doi.org/10.1016/j.chom.2017.03.008.

80. Bayer TS, Booth LN, Knudsen SM, Ellington AD. 2005. Arginine-rich motifs present multiple interfaces for specific binding by RNA. RNA 11:1848–1857. https://doi.org/10.1261/rna.2167605.

81. Jones S, Daley DT, Luscombe NM, Berman HM, Thornton JM. 2001. Protein–RNA interactions: a structural analysis. Nucleic Acids Res 29:943–954. https://doi.org/10.1093/nar/29.4.943.

82. Jeong E, Kim H, Lee SW, Han K. 2003. Discovering the interaction propensities of amino acids and nucleotides from protein-RNA complexes. Mol Cells 16:161–167.

83. 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 34:6450–6460. https://doi.org/10.1093/nar/gkl819.

84. Kruger DM, Neubacher S, Grossmann TN. 2018. Protein–RNA interactions: structural characteristics and hotspot amino acids. RNA 24:1457–1465. https://doi.org/10.1261/rna.066464.118.

85. 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. https://doi.org/10.1073/pnas.1806438115.

86. 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:e15910. https://doi.org/10.7554/eLife.15910.

87. Han SP, Tang YH, Smith R. 2010. Functional diversity of the hnRNPs: past, present and perspectives. Biochem J 430:379–392. https://doi.org/10.1042/BJ20100396.

88. Jeong S. 2017. SR proteins: binders, regulators, and connectors of RNA. Mol Cells 40:1–9. https://doi.org/10.14348/molcells.2017.2319.

89. Lunde BM, Moore C, Varani G. 2007. RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol 8:479–490. https://doi.org/10.1038/nrm2178.

90. 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 72:1045–1058. https://doi.org/10.1007/s00018-014-1779-9.

91. 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. https://doi.org/10.1126/science.aah5163.

92. 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 40:11518–11530. https://doi.org/10.1093/nar/gks913.

93. 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 288:15813–15820. https://doi.org/10.1074/jbc.M112.443390.

94. Kane M, Zang TM, Rihn SJ, Zhang F, Kueck T, Alim M, Schoggins J, Rice CM, Wilson SJ, Bieniasz PD. 2016. Identification of interferon-stimulated genes with antiretroviral activity. Cell Host Microbe 20:392–405. https://doi.org/10.1016/j.chom.2016.08.005.

95. 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. https://doi.org/10.1016/j.jviromet.2008.10.012.

96. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25.

97. 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 7:1025–1035. https://doi.org/10.1016/s1097-2765(01)00240-4.

98. Kutluay SB, Bieniasz PD. 2016. Analysis of HIV-1 Gag-RNA interactions in cells and virions by CLIP-seq. Methods Mol Biol 1354:119–131. https://doi.org/10.1007/978-1-4939-3046-3_8.

99. 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. https://doi.org/10.1016/j.cell.2014.09.057.

100. Cherepanov P. 2007. LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. Nucleic Acids Res 35:113–124. https://doi.org/10.1093/nar/gkl885.

101. Koneru PC, Francis AC, Deng N, Rebensburg SV, Hoyte AC, Lindenberger J, Adu-Ampratwum D, Larue RC, Wempe MF, Engelman AN, Lyumkis D, Fuchs JR, Levy RM, Melikyan GB, Kvaratskhelia M. 2019. HIV-1 integrase tetramers are the antiviral target of pyridine-based allosteric integrase inhibitors. Elife 8:e46344. https://doi.org/10.7554/eLife.46344.

102. 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. https://doi.org/10.1002/pro.3280.

103. 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 11:323. https://doi.org/10.3390/v11040323.

104. Singer JB, Thomson EC, McLauchlan J, Hughes J, Gifford RJ. 2018. GLUE: a flexible software system for virus sequence data. BMC Bioinformatics 19:532. https://doi.org/10.1186/s12859-018-2459-9.

105. Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276–3278. https://doi.org/10.1093/bioinformatics/btu531.